

**Análise do impacto da pandemia no perfil
e na afluência de utentes de um Hospital
do Norte**

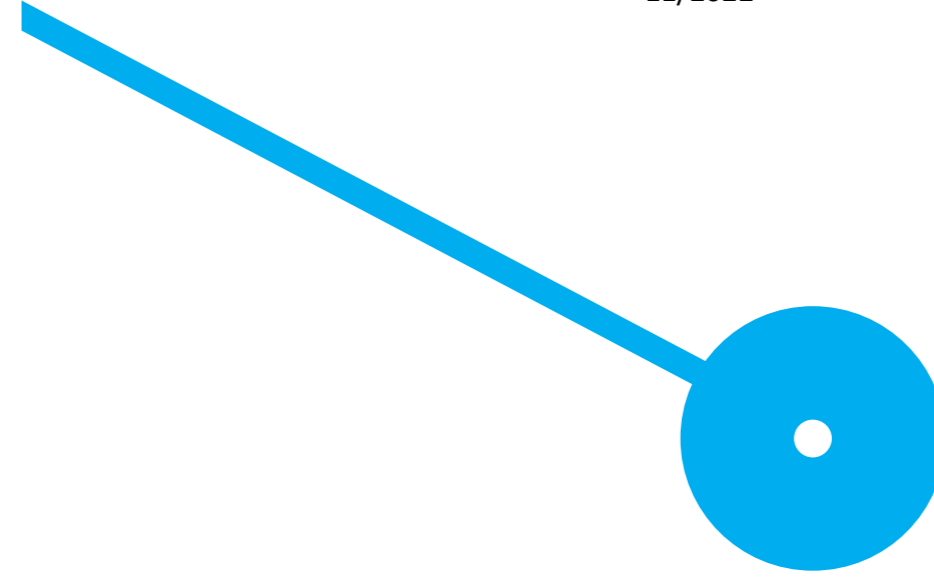
Miguel Silva Maia

11/2022

Miguel Silva Maia. Análise do impacto da pandemia no perfil e na afluência de utentes de um Hospital do Norte

**Análise do impacto da pandemia
no perfil e na afluência de utentes
de um Hospital do Norte**
Miguel Silva Maia

11/2022





Análise do impacto da pandemia no perfil e na afluência de utentes de um Hospital do Norte

Miguel Silva Maia

Orientador(es): Ana Borges e Mariana Carvalho

Dedicatória

Aos meus pais, Fátima e Francisco,
O meu Muito Obrigado.

“A persistência é o caminho mais curto para o êxito”
Charles Chaplin

Agradecimentos

O meu sincero agradecimento às minhas orientadoras, Professora Doutora Ana Borges e Professora Doutora Mariana Carvalho, que me ajudaram a fazer com que este trabalho se tornasse possível de se concretizar, tanto pela orientação científica, como também pelas palavras de encorajamento, incentivo e pela confiança que depositaram em mim ao acreditarem nas minhas capacidades.

Um agradecimento muito especial à minha família, mais concretamente os meus pais, Fátima Maia e Francisco Maia por me terem ajudado a concretizar este sonho e agradeço de forma sincera a oportunidade que me deram. Agradeço ainda mais, pelo facto de me terem acompanhado todos os meus passos e acreditando nas minhas capacidades e competências.

Por último, um agradecimento importantíssimo aos meus amigos que me ajudaram a superar todas as minhas dificuldades nas horas mais difíceis e que me acompanharam nos momentos de alegria e tristeza. Levar-vos-ei comigo para a vida. Quero também agradecer à Escola Superior de Tecnologia e Gestão do Politécnico do Porto, pela oportunidade de ter conseguido estudar nesta instituição ao longo destes cinco anos e por me ter dado os melhores amigos, porque sem eles, este caminho seria muito mais difícil de concluir.

A todos, o meu mais sincero obrigado!

Resumo

Atualmente, um dos principais problemas encarados pelas instituições de saúde, a nível mundial, é o absentismo dos pacientes perante as consultas marcadas. O número de consultas e/ou exames marcados e não realizados, chegam a atingir níveis elevados, sendo este fenómeno visível em vários países e em diferentes tipos de especialidades e atendimentos. Este comportamento gera às instituições: perda de receita, desperdícios de recursos, desorganização da oferta de serviços e limitação dos utentes na garantia do atendimento para a sua assistência. Adicionalmente, gera o uso inadequado da oferta, intensifica os tempos de espera e as respetivas filas, além de acarretar um prejuízo financeiro. Por isso, torna-se importante prever o não comparecimento dos pacientes, assim como perceber o perfil destes pacientes ausentes para que seja possível criar mecanismos que o mitiguem e descobrir as razões que os levam a não comparecer às consultas marcadas. Adicionalmente, a pandemia originada pela COVID-19 apresentou-se como um desafio à gestão hospitalar que rapidamente se teve de adaptar à nova realidade. Como tal, este estudo apresenta dois objetivos principais: (i) entender os fatores associados ao risco de absentismo de pacientes e (ii) entender se houve alterações nos perfis os pacientes perante o risco de absentismo após o início da pandemia.

Assim, este trabalho apresenta a aplicação do modelo logístico binário, identificado na literatura como um dos modelos mais utilizados e eficazes, para ajudar a prever se o paciente comparecerá ou não à consulta agendada. Permitindo entender quais os fatores relacionados com o risco de faltar e comparando os perfis dos pacientes em dois horizontes temporais distintos: pré-pandemia e durante a pandemia, e expondo sugestões de melhoria para a diminuição das taxas de absentismo.

Com base em estudos anteriores, foram incluídos todos os fatores de risco significativos ao absentismo de utentes, tais como: idade, género, estado civil, distância do hospital até à residência do paciente, número de consultas anteriores, primeira consulta/acompanhamento, mês, dia da semana da consulta, precipitação, velocidade do vento, tempo de espera, intervalo entre consultas, número de não comparências anteriores e estação do ano;

Por último, o modelo de regressão logístico permitiu identificar diferentes perfis de pacientes por diferentes especialidades, assim o hospital deverá considerar um modelo logístico distinto por cada tipo de especialidade. O Covid-19 provocou mudanças no comportamento dos pacientes resultando num absentismo destes às consultas. Os fatores de risco comuns identificados nos modelos antes e durante a pandemia, foram o tempo de espera, o número de não comparências anteriores e o número de consultas anteriores.

Relativamente aos fatores de risco característicos do período anterior à pandemia, destacam-se o estado civil e a idade, enquanto durante a pandemia realça-se o mês da consulta.

Palavras-Chave: absentismo; regressão logística; profiling; agendamentos médicos; horizontes temporais

Abstract

Nowadays, one of the main problems faced by health institutions worldwide is healthcare patients not attending to their scheduled appointments. The number of appointments and/or exams scheduled and not performed can reach high levels, and this phenomenon is visible in several countries and in different types of specialties and treatments. This behavior generates loss of revenue, waste of resources, disorganizes the services, and limits the assurance of healthcare to the patients. All this negative behavior generates inadequate use of services, intensifies waiting times and queues, and causes financial loss. Therefore, it is important to predict the non-attendance of patients, as well as to understand the profile of these absent patients so that it is possible to create mechanisms to mitigate it and find out the reasons that lead them not to attend the scheduled appointments in the first place. Additionally, the pandemic originated by the COVID-19 virus also presented a challenge to hospital management, that quickly had to adapt to a new reality. Therefore, this study has two main goals: (i) to understand the factors associated with the risk of patient non-attendance and (ii) to understand if there were changes in the patients' profiles regarding the risk of non-attendance after the beginning of the pandemic.

For this, this study will present the application of the binary logistic model - identified in the literature as one of the most effective and used model- to help predict whether or not the patient will attend the scheduled appointment. This will allow us to understand which factors are related to the risk of non-attendance and compare patient profiles in two distinct time horizons: pre-pandemic and during the pandemic, as well as setting out suggestions for improving the decrease of non-attendance rates.

Based on previous studies, all significant risk factors for patient absenteeism were included, such as: age, gender, marital status, distance from the hospital to the patient's residence, number of previous visits, first visit/attendance, month, day of the week of the visit, rainfall, wind speed, waiting time, interval between visits, number of previous no-shows and season of the year.

Finally, the logistic regression model allowed us to identify different patient profiles by different specialties, so the hospital should consider a distinct logistic model by each type of specialty. Covid-19 caused changes in patient behavior resulting in patient

absenteeism from appointments. The common risk factors identified in the models before and during the pandemic were waiting time, number of prior no-shows, and number of prior visits.

Regarding the risk factors characteristic of the period before the pandemic, marital status and age stand out, while during the pandemic the month of the visit stands out.

Keywords: patient no-show; logistic regression; profiling; scheduled appointment; time horizons.

Índice

Dedicatória	III
Agradecimentos.....	IV
Resumo	V
Abstract.....	VII
Índice de figuras	X
Índice de gráficos.....	X
Capítulo 1: Introdução	11
Capítulo 2: Revisão de literatura	15
2.1 – Absentismo hospitalar: definições e impactos	15
2.2 – Métodos de previsão de absentismo de pacientes.....	16
2.3 – Fatores relacionados com o absentismo de pacientes.....	19
Capítulo 3 – Metodologia.....	30
3.1 – Recolha de dados	31
3.2 – Tarefas de Pré-processamento dos dados	31
3.3 – Modelo de regressão logística.....	36
3.3.1– Teste de Omnibus.....	37
3.3.2– Teste de Hosmer-Lemeshow.....	38
3.4 – Curva de ROC	39
Capítulo 4: Resultados	41
4.1 – Análise Descritiva e Exploratória da Base de Dados	41
4.1.1 – Caracterização dos utentes	41
4.1.2 – Caracterização das consultas.....	43
4.1.2.1 - Antes da pandemia	43
4.1.2.2 - Durante a pandemia.....	45
4.1.2.3 – Comparação pré-pandemia com pandemia.....	47
4.2 – Resultados por especialidade.....	49
4.3 – Diagnóstico dos modelos	58
4.3.1 – Teste de Omnibus.....	58
4.3.2 – Teste de Hosmer e Lemeshow	58
4.3.3 – Curva de ROC.....	59
Discussão e Conclusão	62
Bibliografia	65

Índice de figuras

Figura 1 - Metodologia utilizada neste estudo.....	30
Figura 2 - Exemplo da curva de ROC.....	39

Índice de tabelas

Tabela 1 - Métodos utilizados na literatura para a previsão do absentismo de pacientes.....	19
Tabela 2 - Fatores demográficos dos pacientes do Hospital.....	22
Tabela 3 - Histórico médico dos pacientes do Hospital	22
Tabela 4 - Detalhes do agendamento das consultas dos pacientes do Hospital	23
Tabela 5 - Comportamento dos pacientes do Hospital.....	24
Tabela 6 - Principais conclusões dos estudos analisados face aos fatores que afetam o absentismo identificado	29
Tabela 7 - Variáveis incluídas no conjunto de dados originais.....	31
Tabela 8 - Horizontes temporais considerados	35
Tabela 9 - Variáveis independentes e dependente analisados no estudo.....	36
Tabela 10 - Descrição do modelo segundo a área que se encontra abaixo da curva de ROC	40
Tabela 11 - Estatísticas descritivas dos pacientes	42
Tabela 12 - Características descritivas das consultas (antes da pandemia).....	45
Tabela 13 - Características descritivas das consultas (durante a pandemia).....	47
Tabela 14 - Modelo logístico multivariado de ortopedia.....	51
Tabela 15 - Modelo logístico multivariado de oftalmologia.	53
Tabela 16 - Modelo logístico multivariado de cardiologia.	54
Tabela 17 - Modelo logístico multivariado de medicina dentária.....	56
Tabela 18 - Modelo logístico multivariado de medicina física e reabilitação.	58
Tabela 19 - Áreas abaixo das curvas de ROC das especialidades analisadas.	61

Índice de gráficos

Gráfico 1- antes e durante a pandemia - ortopedia.....	59
Gráfico 2 - antes e durante a pandemia - oftalmologia	60
Gráfico 3 - antes e durante a pandemia - cardiologia	60
Gráfico 4 - antes e durante a pandemia - medicina dentária	60
Gráfico 5 - antes e durante a pandemia - medicina física e reabilitação	61

Capítulo 1: Introdução

O absentismo de pacientes tornou-se uma situação alarmante, não só a nível nacional, mas também mundial, tornando-se um problema crónico (Salazar *et al.*, 2020). Para clarificar, o absentismo (ou a ausência) de um paciente face a uma consulta é quando um paciente não consegue comparecer à consulta marcada sem informar o respetivo hospital antes do horário agendado (Perez *et al.*, 2014).

Um grande número de pacientes falta às suas consultas por não conseguir cancelar as mesmas a tempo. Como um dos objetivos primordiais dos serviços de saúde é o atendimento aos pacientes para a resolução de problemas de saúde individuais e coletivos, quando esse atendimento não é realizado há a perda de oportunidade de conseguir oferecer ajuda a outro paciente que necessite de cuidados, como também há uma perda mercadológica. O não comparecimento de pacientes causa um enorme desperdício de recursos de saúde, e a falta de qualquer tipo de comunicação dificulta o reagendamento ou a inserção de outros pacientes no horário de funcionamento para melhorar a eficiência da instituição (Bittar, *et al.*, 2016) (Bush, *et al.*, 2016). O atendimento dos pacientes pode resultar no desperdício de biliões de dólares em inatividade, horas extras e tempo de espera de que envolvem hospitais e pacientes (Batool, *et al.*, 2021).

Assim, surge a necessidade de arranjar uma solução para abordar as taxas de não comparecimento dos pacientes e tentar entender a razão pelas quais estas ocorrem.

1.1 Doença Coronavírus

Nos finais do ano de 2019, alastrou-se uma doença denominada de coronavírus, também conhecida por COVID-19. Os primeiros casos da doença foram detetados na China, mais precisamente em Whuan. A dificuldade para controlar a sua propagação fez com que se espalhasse não só por toda a China como também pelos restantes países do mundo, tornando-se uma preocupação mundial (Gong *et al.*, 2020). Em dezembro do mesmo ano, estudos revelaram existir uma ligação entre o surto ocorrido em Whuan e o Huanan Seafood Wholesale Market, que vende diversos produtos como aves vivas,

animais aquáticos e selvagens. Devido à seriedade deste surto, o centro Chinês de Controlo e Prevenção de doenças e as autoridades de saúde da China comunicaram a existência de um novo coronavírus. Mais tarde, quando a doença atingiu proporções mundiais, a Organização Mundial de Saúde (OMS) declarou o coronavírus como doença e a Comissão Internacional de Taxonomia do Vírus nomeou o vírus responsável como SARS-CoV-2 (Deng & Peng, 2020).

A partir do dia 18 de fevereiro de 2020, a situação complicou-se tendo o novo vírus alastrado para cerca de 34 províncias na China e para outros 25 países, resultando assim, num total de 75.199 casos confirmados, e num total de 2.009 mortes (Deng & Peng, 2020) & (World Health Organization, 2020). Durante os meses seguintes, o número de casos foi aumentando consideravelmente em todo o mundo, o que acabou por se tornar numa grande ameaça para a saúde pública mundial. A vasta e rápida propagação da pandemia do COVID-19 originou respostas médicas e de saúde pública bastante violentas, implicando testes, triagem, rastreamento de pessoas, distanciamento social, limitações de viagem e obrigação em ficar casa após o teste positivo à COVID-19 (Berger *et al.*, 2020).

Adicionalmente, a pandemia COVID-19 dificultou o controlo dos médicos no que diz respeito ao atendimento aos pacientes, originando diversos constrangimentos na qualidade de atendimento, levando a um aumento substancial do stress sentido pelos profissionais de saúde e conseqüentemente a uma maior probabilidade de os mesmos sofrerem de esgotamento físico, mental, emocional e de verem o seu bem-estar afetado a médio/longo prazo. Assim, é fundamental que todos os líderes e médicos que estão na linha da frente no combate à doença, se encontrem mentalmente e fisicamente bem para que se possa evitar deficiências na qualidade de atendimento ao paciente. Para isso acontecer, devemos-nos empenhar e apoiar o bem-estar dos envolvidos no atendimento ao utente. (Dewey *et al.*, 2020).

1.2 Motivação e objetivos

A motivação deste trabalho é a redução das taxas de absentismo dos pacientes que frequentam o Hospital do Norte, uma vez que as taxas de absentismo acarretam inúmeras conseqüências prejudiciais ao Hospital, por exemplo, perda de milhares de

euros, aumento das filas de espera, perda de oportunidade de oferta para ajudar outro paciente, desperdícios de recursos de saúde.

O objetivo primordial deste trabalho é determinar os fatores de risco que influenciam os pacientes a não comparecer às consultas hospitalares. Trata-se de um estudo, recorrendo ao modelo logístico, com o objetivo de perfilar os pacientes, ajudando o Hospital a entender os fatores associados de forma a tomar medidas para mitigar as taxas de absentismo. Outro objetivo principal é entender se ocorreram alterações nos perfis dos pacientes perante o risco de absentismo após o início da pandemia, fazendo assim uma comparação antes e durante a pandemia.

1.3 Principais contribuições

Como principais contribuições deste trabalho apontam-se: a) a melhoria da eficiência do Hospital ajudando a entender o perfil dos seus pacientes e ajudando a entender os fatores que os levam a faltar às consultas, b) a implementação do modelo logístico que auxiliará o Hospital a entender o problema identificado, e c) contribuição para o entendimento sobre a extensão dos problemas/alterações nos comportamentos dos pacientes perante o absentismo hospitalar originadas pela pandemia.

1.4 Estrutura da Dissertação

Este estudo encontra-se organizado da seguinte forma: o capítulo 2 apresenta uma revisão de literatura, onde estão descritos a definição e os impactos relativos ao absentismo hospitalar, os métodos utilizados pelos autores a fim de mitigarem o absentismo de utentes às consultas hospitalares e, também, onde são apresentados os fatores relacionados com a comparência ou não dos utentes às consultas; no capítulo 3 é feita uma análise descritiva dos dados recolhidos. Este capítulo encontra-se dividido em 3 subcapítulos, sendo que no primeiro é realizada uma caracterização dos pacientes deste estudo, no segundo é feita uma caracterização das consultas - nos diferentes períodos (antes e durante a pandemia) - e no terceiro é realizada uma comparação entre o período de pré-pandemia e o período de pandemia; no capítulo 4 é descrita a metodologia deste estudo, iniciando-se pelo pré-processamento, posteriormente pela explicação o modelo utilizado (modelo de regressão logístico) assim como os respetivos

testes de diagnóstico do modelo: teste de Omnibus e o teste de Hosmer-Lemeshow e as curvas de ROC; O capítulo 5 expressa os resultados deste estudo explicando os resultados por especialidades assim como os diagnósticos dos modelos; Por fim, o capítulo 6 oferece a conclusão e a discussão final do trabalho.

Capítulo 2: Revisão de literatura

Neste capítulo serão apresentados os principais conceitos de absentismo hospitalar, os seus impactos encontrados na literatura, os métodos de previsão do absentismo de pacientes e os fatores relacionados com o mesmo.

2.1 – Absentismo hospitalar: definições e impactos

O absentismo ambulatorio refere-se à impossibilidade de comparecimento do paciente na consulta marcada sem informar o respetivo hospital antes do horário agendado (Perez *et al.* 2014). Esta situação tem-se tornado bastante alarmante não só a nível nacional como mundial, chegando a tornar-se um problema crónico ou, por outras palavras, um transtorno duradouro (Salazar *et al.*, 2020).

A nível mundial, uma elevada quantidade de pacientes perde as suas consultas por não conseguirem cancelar a consulta atempadamente. O número total de exames e consultas agendadas e não efetuadas, resultantes do não aparecimento dos pacientes, é bastante elevado em todo o mundo nas diferentes especialidades médicas e nos diferentes tipos de atendimento (Salazar *et al.* 2020). Este problema gera perdas de milhões de euros anualmente em consequência do tempo disponível, do tempo de espera e das horas extras que os hospitais e os pacientes têm de se submeter (Batoool *et al.*, 2021).

Um dos objetivos principais dos serviços de saúde é desenvolver o melhor atendimento possível ao cidadão de forma a resolver problemas individuais e coletivos de saúde. Quando esse atendimento não é concretizado, perde-se a oportunidade de oferecer auxílio a outro paciente, resultando num prejuízo financeiro e mercadológico, isto é, afeta a “imagem” da organização diante da sociedade. Para além disso, o absentismo dos pacientes causa um enorme desperdício de recursos de saúde e a falta de comunicação entre os órgãos hospitalares obstrui o reagendamento ou a inserção de outros pacientes no horário de atendimento (Bittar *et al.* 2016) e (Bush *et al.* 2016).

Atualmente, é comum encontrar utentes descontentes com a forma de atendimento que é disponibilizada ao público. Na maioria das vezes, a contestação relaciona-se com

a demora do agendamento das consultas médicas. De facto, o absentismo prejudica a capacidade de agendamento das consultas, dificultando que outros pacientes tenham acesso ao sistema de saúde. Este problema acaba por intensificar o tempo de espera estipulado para a consulta gerando gastos financeiros visto que este atendimento é pago pelo Estado mesmo quando o paciente decide não comparecer (Salazar *et al.*, 2020). Em suma, o conceito de absentismo em consultas hospitalares deve ser realçado pois conduz de forma direta ao desperdício de recursos estruturais e financeiros (Bittar *et al.* 2016).

O absentismo de pacientes em Hospitais tem sido cada vez maior com o passar dos anos, pelo que é fundamental investigar as principais causas que levam os utentes ao absentismo das consultas médicas é o principal motivo da atenção da saúde pública. As consultas médicas perdidas representam efeitos prejudiciais em várias vertentes, nomeadamente, no paciente, na clínica, no sistema de saúde e, por último e não menos importante, na saúde pública (Bittar *et al.* 2016) e (Kalb *et al.* (2012). Para que haja uma melhoria do acesso ao sistema público de saúde, é necessário reduzir o absentismo para que possa haver um controlo das filas de espera levando a uma melhoria na qualidade de vida da população (Salazar *et al.*, 2020).

O absentismo hospitalar provoca efeitos negativos na produtividade dos funcionários, no equipamento de consultório e nos custos, que aumentam devido à ociosidade dos recursos disponíveis. Para além disso, existem grandes perdas de lucro, que se estimam por volta dos 150 milhões de dólares anuais apenas nos EUA. Estes valores alertam uma vez mais para a necessidade da redução das taxas de absentismo. Para perceber de que forma se podem reduzir essas taxas, é necessário traçar o perfil dos utentes e a partir disso aprender a prever a probabilidade de um paciente aparecer ou não nas consultas (Batool *et al.* 2021).

2.2 – Métodos de previsão de absentismo de pacientes

Como forma de reduzir o absentismo hospital é necessário identificar os métodos de previsão associados ao absentismo de utentes.

Em 2020, Carreras-García *et al.* (2020) realizaram uma revisão de literatura onde analisaram cerca de cinquenta artigos. Os modelos, por eles encontrados, aplicados na

literatura foram os seguintes: modelos de regressão, modelos em árvore, redes neurais, modelos de Markov, modelos Bayesianos (ver tabela 1). Chegaram à conclusão de que o modelo de regressão é sem dúvida alguma o modelo que é mais utilizado para se prever as faltas de não comparecimento, apresentando elevados níveis de precisão para a maioria dos autores que estudaram este tema (Carreras-Carreras-García *et al.*, 2020).

Wahid *et al.* (2019), Salazar *et al.* (2020) e Batool *et al.* (2021) utilizaram o modelo das árvores de decisão, que é visto como o modelo mais utilizado após os modelos de regressão. Este modelo foi também bastante utilizado na literatura a partir dos anos 80 e daí em diante (Carreras-García *et al.*, 2020). Mohammadi *et al.* (2018) usaram redes neurais, que são uma das técnicas de classificação binária que apresenta melhor cuidado no ramo da inteligência artificial. Este modelo apresenta valores relativamente semelhantes aos do modelo anterior, só que é aplicado com menor frequência.

Em 2011, Alaeddini *et al.* (2011) estudou as taxas de absentismo em centros clínicos (centros de saúde comunitários e centros comunitários), através da produção de um modelo de regressão logística e Naïve Bayes, obtendo um nível de precisão de cerca de 80%.

Levy (2013) identificou diversos problemas que tem vindo a preocupar as clínicas do estado de Nova Iorque, dado que os valores de não comparecimento de pacientes às consultas tem vindo a aumentar de dia para dia. Com o intuito de entender o absentismo, o autor decidiu analisar o perfil dos pacientes numa clínica de Gastroenterologia, em Albany (Nova Iorque). O autor selecionou uma série de consultas, na qual constavam uma lista de características demográficas, clínicas e administrativas. Para analisar as variáveis em causa utilizou o modelo Bayesiano – modelo vantajoso, dado que são conhecidos pela transparência com que lidam com a incerteza. Este modelo apresentou níveis de precisão que variam entre os 60% e os 65%.

Com o intuito de perfilar dos pacientes de alguns estados da América do Norte, Bush *et al.* (2016) utilizaram um sistema como parte prática através da rotina diária de várias instituições (Hospital Pediátrico de Colorado, o Hospital da Universidade de Virgínia e o Hospital Pediátrico de Rady), na qual foram analisadas todas as consultas ambulatoriais de urologia. Para análise da probabilidade de não comparecimento foram realizados testes através de uma regressão logística multivariada (Bush, *et al.* 2016).

Mohammadi *et al.* (2018) estudaram dados que foram recolhidos de um grande centro de saúde urbano, em Indianápolis, na qual apresentavam informação sobre vários serviços de saúde que incluem cuidados primários pediatria, medicina interna, clínica familiar, ginecologia/obstetrícia, cuidados dentários/visão, saúde comportamental e cuidados preventivos. Os dados recolhidos incluíam informações clínicas (dados operacionais e financeiros) e do paciente (dados demográficos e características clínicas do paciente). Para desenvolver e comparar os modelos de previsão do não comparecimento de pacientes às consultas, os autores recorreram a três metodologias: regressão logística, rede neuronal artificial e o classificador de Naïve Bayes, que apresentaram em termos de precisão, 80%, 79% e 92%, respetivamente.

No ano anterior, Goffman *et al.* (2017) propuseram o modelo Markov como alternativa ao modelo de regressão logística de forma a justificar a dependência temporal. Estes registaram um melhor desempenho e uma melhor caracterização das faltas comparativamente ao modelo de regressão.

Para analisar os modelos propostos, Hu *et al.* (2020) propuseram três modelos para uma análise de regressão hierárquica na qual utilizaram o modelo de regressão logística.

Mais recentemente, Batool *et al.* (2021) desenvolveu o sistema de agendamento de Consultas e Gerenciamento Intuitivo (SACG), usado para prever o não comparecimento do paciente baseado no perfil de utente. Neste sistema são aplicados algoritmos de machine learning mais especificamente árvores de decisão, através de uma base de dados. Este modelo assinalou que um paciente corre o risco de faltar a uma consulta com uma precisão de 95% (Batool *et al.*, 2021).

Autores	Método
Neal <i>et al.</i> (2001); Junod-Perron <i>et al.</i> (2010); Kalb <i>et al.</i> (2012); Bush <i>et al.</i> (2016); Mohammadi <i>et al.</i> (2018); Salazar <i>et al.</i> (2020); Hu <i>et al.</i> (2020);	Regressão logística
Mohammadi <i>et al.</i> (2018);	Redes neurais
Mohammadi <i>et al.</i> (2018);	Classificador Naïve Bayes
Wahid <i>et al.</i> (2019); Salazar <i>et al.</i> (2020); Batool <i>et al.</i> (2021)	Árvore de decisão
Alaeddini, Yang, & Reddy (2011); Levy (2013);	Bayesiano
Wahid <i>et al.</i> (2019); Salazar <i>et al.</i> (2020);	Random Forest

Tabela 1 - Métodos utilizados na literatura para a previsão do absentismo de pacientes.

2.3 – Fatores relacionados com o absentismo de pacientes

Os cancelamentos das consultas dos pacientes em serviços de clínica geral (em todas as faixas etárias) são explicados por diversos fatores que advêm do não comparecimento dos pacientes (Neal *et al.* 2001).

Muitos desses fatores foram estudados em detalhe por diversos autores, nomeadamente Neal *et al.* (2001) e Junod-Perron *et al.* (2010). Estes autores, em especial, identificam inúmeros fatores clínicos como, por exemplo, a dificuldade de agendamento de uma consulta devido à difícil comunicação que se enfrenta entre a clínica e os pacientes, que resulta numa relação instável entre ambos, e os demorados tempos de espera entre a disponibilidade da consulta e o respetivo agendamento. Relativamente a fatores demográficos, estes autores apontam outras causas como: ser-se indivíduo do sexo feminino; residir numa zona socioeconómica mais baixa; e ter menos habilitações literárias.

A taxa de absentismo em centros clínicos, nomeadamente centros de saúde comunitários e centros universitários, pode variar entre os 3% e os 80% (Alaeddini *et al.* 2011). Estas taxas dependem de diversas condições, nomeadamente, o tipo de centro clínico e das informações demográficas dos pacientes. Alaeddini *et al.* (2011) concluíram também que existe um forte absentismo em dias próximos de feriados nacionais.

Levy (2013) ao analisar as clínicas do estado de Nova Iorque, em destaque uma clínica de gastroenterologia, selecionou uma série de consultas, das quais consta uma lista de características demográficas, clínicas e administrativas, sendo que as características que mais influenciaram o não comparecimento foram o tempo de espera, hora do dia, comportamento de compra do médico, o sexo e a idade. Este autor conseguiu prever o não comparecimento dos pacientes às consultas marcadas através de bases de dados que continham diversas variáveis independentes, particularmente, o tipo de clínica, o tempo de espera, a idade, o sexo, o estado civil, a raça, se detém ou não telemóvel, se possui seguro de saúde e se é fumador.

Mohammadi *et al.* (2018) percebeu que o tempo de espera (diferença de tempo (dias) entre a data em que o paciente marcou com a clínica a consulta e a data da respetiva consulta), o comportamento anterior do não comparecimento do paciente, a posse de telemóvel, o consumo de tabaco e o número de dias desde a última consulta do paciente são fatores significativos que influenciaram a respetiva adesão dos pacientes às consultas.

Recentemente, Carreras-García *et al.* (2020) analisaram o não comparecimento dos pacientes às consultas no Reino Unido. Os autores conseguiram no seu trabalho exaustivo, reunir diversos fatores que podem levar ao não comparecimento dos pacientes, fatores estes que se dividem em: fatores demográficos dos pacientes (tabela 2); historial do médico (tabela 3); detalhes de agendamento/consulta (tabela 4); e, comportamento dos pacientes (tabela 5).

Hu *et al.* (2020) investigaram o comportamento de não comparecimento dos pacientes às consultas. Este estudo foi realizado com base em dados fornecidos por um hospital público de atendimento terciário na cidade de Hangzhou, na China. Estes autores conseguiram concluir algumas variáveis fulcrais para ajudar a entender o não comparecimento dos pacientes às consultas marcadas. Usaram as taxas de não comparecimento obtidas através do número de registos do não comparecimento dividido pelo total de consultas, para se calcular a frequência com que os pacientes faltavam às consultas. Estes autores concluíram que o sexo e a idade não têm qualquer tipo de efeito significativo sobre o absentismo e um fator bastante importante na previsão do mesmo é o tempo de espera (Hu *et al.*, 2020).

Autores	Variável	
Neal <i>et al.</i> (2001); Junod-Perron <i>et al.</i> (2010); Kalb <i>et al.</i> (2012); Levy (2013); Huang & Hanauer, 2016; Bittar <i>et al.</i> (2016); Bush <i>et al.</i> (2016); Harvey <i>et al.</i> , 2017; Lu <i>et al.</i> , 2017; Ding, <i>et al.</i> , 2018; Topuz <i>et al.</i> (2017); Mohammadi <i>et al.</i> (2018); Praveena, Krupa, & SaiPreethi, 2019; Ahmadi, <i>et al.</i> , 2019; Dantas <i>et al.</i> , 2018; Aladeemy <i>et al.</i> (2020); Hu <i>et al.</i> (2020); Salazar <i>et al.</i> (2020); Batool <i>et al.</i> (2021)	Idade	Fatores demográficos dos pacientes
Neal <i>et al.</i> (2001); Junod-Perron <i>et al.</i> (2010); Kalb <i>et al.</i> (2012); Levy (2013); Huang & Hanauer, 2016; Bittar <i>et al.</i> (2016); Lu <i>et al.</i> , 2017; Bush <i>et al.</i> (2016); Harvey <i>et al.</i> , 2017; Ding, <i>et al.</i> , 2018; Topuz <i>et al.</i> (2017); Mohammadi <i>et al.</i> (2018); Praveena, Krupa, & SaiPreethi, 2019; Ahmadi, <i>et al.</i> , 2019; Dantas <i>et al.</i> , (2018); Aladeemy <i>et al.</i> (2020); Hu <i>et al.</i> (2020); Salazar <i>et al.</i> (2020); Batool <i>et al.</i> (2021)	Género	
Kalb <i>et al.</i> (2012); Huang & Hanauer, 2016; Bush <i>et al.</i> (2016); Harvey <i>et al.</i> , 2017; Lu <i>et al.</i> , 2017; Ding, <i>et al.</i> , 2018; Mohammadi <i>et al.</i> (2018)	Raça/Etnia	
Huang & Hanauer, 2016; Harvey <i>et al.</i> , 2017; Lu <i>et al.</i> , 2017	Língua	
Harvey <i>et al.</i> , 2017; Ding, <i>et al.</i> , 2018; Ahmadi, <i>et al.</i> , 2019	Situação profissional	
Bush <i>et al.</i> (2016); Harvey <i>et al.</i> , 2017; Lu <i>et al.</i> , 2017; Mohammadi <i>et al.</i> (2018); Ding, <i>et al.</i> , 2018; Ahmadi, <i>et al.</i> , 2019; Aladeemy <i>et al.</i> (2020)	Estado Civil	
Harvey <i>et al.</i> , 2017; Lu <i>et al.</i> , 2017; Ding, <i>et al.</i> , 2018; Topuz <i>et al.</i> (2017); Aladeemy <i>et al.</i> (2020)	Estatuto económico	
Harvey <i>et al.</i> , 2017; Praveena, Krupa, & SaiPreethi, 2019	Habilitações literárias	
Junod-Perron, <i>et al.</i> (2010); Kalb <i>et al.</i> (2012); Huang & Hanauer, 2016; Bush <i>et al.</i> (2016); Lu <i>et al.</i> , 2017; Harvey <i>et al.</i> , 2017; Mohammadi <i>et al.</i> (2018); Dantas <i>et al.</i> , (2018); Ahmadi, <i>et al.</i> , 2019; Salazar <i>et al.</i> (2020); Batool <i>et al.</i> (2021);	Seguros/Pagamentos	
Neal <i>et al.</i> (2001); Lu <i>et al.</i> , 2017; Ahmadi, <i>et al.</i> , 2019; Salazar <i>et al.</i> (2020); Hu <i>et al.</i> (2020); Batool <i>et al.</i> (2021)	Código Postal	

Kalb <i>et al.</i> (2012); Huang & Hanauer, 2016); Harvey <i>et al.</i> , 2017; Lu <i>et al.</i> , 2017; Topuz <i>et al.</i> (2017); Praveena, Krupa, & SaiPreethi, 2019; Dantas <i>et al.</i> , (2018); Ahmadi, <i>et al.</i> , 2019; (Aladeemy <i>et al.</i> 2020);	Distância/transporte
Huang & Hanauer, 2016; Lu <i>et al.</i> , 2017;	Religião
Bush <i>et al.</i> (2016); Lu <i>et al.</i> , 2017; Mohammadi <i>et al.</i> (2018); Salazar <i>et al.</i> (2020); Aladeemy <i>et al.</i> (2020);	Acesso telemóvel

Tabela 2 - Fatores demográficos dos pacientes do Hospital

Autores	Variável	
Bush <i>et al.</i> (2016); Lu <i>et al.</i> , 2017; Mohammadi <i>et al.</i> (2018); Hu <i>et al.</i> (2020);	Tipo de clínica	Histórico Médico
Neal <i>et al.</i> (2001); Junod-Perron <i>et al.</i> (2010); Kalb <i>et al.</i> (2012); Levy (2013); Bittar <i>et al.</i> (2016); Bush <i>et al.</i> (2016); Lu <i>et al.</i> , 2017; Mohammadi <i>et al.</i> (2018); Dantas <i>et al.</i> , (2018); Ahmadi, <i>et al.</i> , 2019; Hu <i>et al.</i> (2020); Salazar <i>et al.</i> (2020); Batool <i>et al.</i> (2021);	Especialidade	
Harvey <i>et al.</i> , 2017; Lu <i>et al.</i> , 2017; Topuz <i>et al.</i> (2017); Dantas <i>et al.</i> , (2018); Ahmadi, <i>et al.</i> , 2019; Hu <i>et al.</i> (2020); Batool <i>et al.</i> (2021)	Consultas anteriores	
Junod-Perron <i>et al.</i> (2010); Kalb <i>et al.</i> (2012); Lu <i>et al.</i> , 2017; Ahmadi, <i>et al.</i> , 2019;	Prestador de serviços	
Levy (2013); Lu <i>et al.</i> , 2017; Ding, <i>et al.</i> , 2018; Praveena, Krupa, & SaiPreethi, 2019	Diagnóstico	
Lu <i>et al.</i> , 2017; Ahmadi, <i>et al.</i> , 2019	Duração diagnóstico	
Huang & Hanauer, 2016; Lu <i>et al.</i> , 2017; Dantas <i>et al.</i> , (2018); Hu <i>et al.</i> (2020); Aladeemy <i>et al.</i> (2020)	Primeira consulta/consulta acompanhamento	

Tabela 3 - Histórico médico dos pacientes do Hospital

Autores	Variável	
Junod-Perron <i>et al.</i> (2010); Levy (2013); Huang & Hanauer, 2016; Lu <i>et al.</i> , 2017; Topuz <i>et al.</i> (2017); Dantas <i>et al.</i> , (2018); Ahmadi, <i>et al.</i> , 2019; Hu <i>et al.</i> (2020)	Mês	Detalhes do agendamento das consultas
Junod-Perron <i>et al.</i> (2010); Levy (2013); Huang & Hanauer, 2016; Lu <i>et al.</i> , 2017; Harvey <i>et al.</i> , 2017; Topuz <i>et al.</i> (2017); Ding, <i>et al.</i> , 2018; Praveena, Krupa, & SaiPreethi, 2019; Dantas <i>et al.</i> , (2018); Ahmadi, <i>et al.</i> , 2019; Hu <i>et al.</i> (2020); Aladeemy <i>et al.</i> (2020)	Dia da semana	
Kalb <i>et al.</i> (2012); Levy (2013); Huang & Hanauer, 2016; Lu <i>et al.</i> , 2017; Harvey <i>et al.</i> , 2017; Topuz <i>et al.</i> (2017); Ding, <i>et al.</i> , 2018; Dantas <i>et al.</i> , (2018); Hu <i>et al.</i> (2020); Aladeemy <i>et al.</i> (2020)	Horário da consulta	
Lu <i>et al.</i> , 2017; Topuz <i>et al.</i> (2017)	Férias	
Lu <i>et al.</i> , 2017; Ahmadi, <i>et al.</i> , 2019	Consulta de um dia	
Aladeemy <i>et al.</i> (2020)	Meteorologia	
Kalb, <i>et al.</i> (2012); Ahmadi, <i>et al.</i> , 2019;	Estação do ano	
Lu <i>et al.</i> , 2017; Ahmadi, <i>et al.</i> , 2019; Aladeemy <i>et al.</i> (2020);	Intervalo entre consultas	
Harvey <i>et al.</i> , 2017; Huang & Hanauer, 2016; Lu <i>et al.</i> , 2017; Mohammadi <i>et al.</i> (2018); Topuz <i>et al.</i> (2017); Ding, <i>et al.</i> , 2018; Ahmadi, <i>et al.</i> , 2019; Hu <i>et al.</i> (2020); Aladeemy <i>et al.</i> (2020);	Diferença entre a data da consulta e a data da criação da consulta (Lead time)	
Kalb <i>et al.</i> (2012); Bush <i>et al.</i> (2016); Mohammadi <i>et al.</i> (2018); Batool <i>et al.</i> (2021);	Tempo de espera	
Lu <i>et al.</i> , 2017; Ding, <i>et al.</i> , 2018; Batool <i>et al.</i> (2021);	Modo de agendamento	

Tabela 4 - Detalhes do agendamento das consultas dos pacientes do Hospital

Autores	Variável	
Neal <i>et al.</i> (2001); Kalb <i>et al.</i> (2012); Huang & Hanauer, 2016); Lu <i>et al.</i> , 2017; Harvey <i>et al.</i> , 2017; Ding, <i>et al.</i> , 2018; Ahmadi, <i>et al.</i> , 2019; Hu <i>et al.</i> (2020); Aladeemy <i>et al.</i> (2020); Salazar <i>et al.</i> (2020);	Não comparecências anteriores	Comportamento do paciente
Kalb <i>et al.</i> (2012); Lu <i>et al.</i> , 2017; Ahmadi, <i>et al.</i> , 2019;	Cancelamentos anteriores	
Hu <i>et al.</i> (2020); Lu <i>et al.</i> , 2017;	Estado da última consulta	

Tabela 5 - Comportamento dos pacientes do Hospital

Assim sendo, na literatura foram identificados alguns fatores independentes que explicam a razão de os pacientes não comparecerem às consultas marcadas pelas respetivas instituições de saúde. Na tabela 6 estão resumidas algumas das principais conclusões dos estudos analisados.

A idade é um dos fatores que mais apresenta resultados divergentes no não comparecimento dos pacientes. Pacientes mais jovens tendem a faltar mais do que pacientes de maior idade (Noris *et al.* 2014) (Perez *et al.* 2014) e (Harvey *et al.* 2017). Como Kalb *et al.* (2012) estudou apenas crianças entre os 4 meses e os 21 anos, este autor conclui que o não comparecimento das crianças vai aumentando gradualmente com o aumento das idades. Para Junod-Perron *et al.* (2010) e Hu *et al.* (2020), a faixa etária que maior taxa de absentismo apresenta é a dos 40 anos, o contrário acontece com Mohammadi *et al.* (2010), que comprovou que a faixa etária dos 20 anos é que menos comparece. Para Dantas *et al.* (2018) e Neal *et al.* (2001), concluíram que os jovens adultos são os que mais faltam, ou seja, entre os 30 e os 44 anos e os 18 e os 34 anos respetivamente.

Outro fator, não menos importante, é o género. É de notar que os homens são os que mais faltam às consultas (Kalb *et al.*, 2012; Perez *et al.*, 2014; Huang *et al.*, 2016; Bush *et al.*, 2016; Harvey *et al.*, 2017 e Junod-Perron *et al.*, 2010), mas pelo contrário, outros autores como Neal *et al.* (2001), Dantas *et al.* (2018) e Mohammadi *et al.* (2018)

corroboram os outros autores, afirmando que pacientes do género feminino tendem a faltar mais do que os pacientes do género masculino.

Relativamente ao estado civil, Mohammadi *et al.* (2018) conclui que pacientes que se encontram solteiros, são os que mais faltam às consultas marcadas pelas instituições de saúde, em contrapartida, pacientes enviuvados e pacientes que se encontram em união de facto tendem a comparecer com maior frequência às consultas.

Outro fator significativo que influencia diretamente o não comparecimentos dos pacientes é a distância entre o concelho de residência de cada paciente e a instituição de saúde/hospital. Para esta variável foram obtidas várias conclusões. Huang *et al.* (2016) concluiu que pacientes que estejam a menos de 8km da clínica/hospital, tem maiores probabilidade de faltar às consultas. Para Kalb *et al.* (2012), a média ronda os 25km, enquanto para Dantas *et al.* (2018), pacientes que se encontram entre os 20 e os 50km, não aparecem às consultas agendadas. Em contraposto, Perez *et al.* (2014), conclui que pacientes que estejam a menos de 25km da clínica tendem a faltar menos e pacientes que estejam a mais de 80km tem maior probabilidade de não comparecer.

Outro fator importante é o tipo de especialidade. Este é um fator que varia conforme o público-alvo que cada autor decidiu analisar. Segundo Harvey *et al.* (2017), as especialidades com maior taxa de absentismo por parte dos pacientes, são a Mamografia e Tomografia Computadorizada (CT), o contrário acontece na Tomografia por emissão de positrões (TEC) e a Ressonância Magnética que apresentam taxas baixas de não comparecimento.

Para Perez *et al.* (2014), a dermatologia e a neurologia, destacam-se pela negativa no que diz respeito ao absentismo, aumentando cada vez mais de ano para ano. Outras especialidades identificadas negativamente por Ding *et al.* (2018), foram a Pulmonar e Alergia, em contrapartida da Uroginecologia. A vascular e a Ginecologia foram também identificadas por Bittar *et al.* (2016) como especialidades com maiores taxas de absentismo. Mohammadi *et al.* (2016) identificou a pediatria como a especialidade que apresenta maior taxa de não comparecimento, por outro lado, identificou a podologia como especialidade com menor taxa de absentismo. Por fim, identificou-se também a cirurgia bariátrica (Dantas *et al.* 2018) e a terapia da fala (Kalb *et al.* 2012) como especialidades com elevado absentismo.

O facto de ser ou não a primeira consulta, também tem uma relação com o facto de os pacientes virem a faltar ou não às consultas. Assim, Dantas *et al.* (2018) concluiu que pacientes que tenham tido pelo menos 1 a 2 consultas antes naquela mesma clínica, tem maior taxa de não virem a comparecer à próxima consulta marcada, o contrário acontece com pacientes que tenham tido mais de 9 consultas anteriores naquela clínica/hospital tendem a comparecer com maior probabilidade. Vários autores como Mohammadi *et al.* (2018), Junod-Perron *et al.* (2020), Kalb *et al.* (2012) e Bush *et al.* (2016) concluíram todos o mesmo, que os pacientes que se encontram em consultas de acompanhamento estão mais propícios a faltar do que pacientes pela primeira vez na clínica.

Pacientes com historiais anteriores de não comparecimento também têm maiores indícios de virem a faltar com maior frequência. Por isso, Dantas *et al.* (2018) conclui que pacientes que tenham faltado pelo menos 50% das vezes às consultas terão maior probabilidade de faltar das próximas vezes.

Relativamente aos meses de maior absentismo, Dantas *et al.* (2018) deduziu que as maiores taxas de absentismo variam entre setembro e novembro. Huang *et al.* (2016) apresentam valores próximos a Dantas *et al.* (2018), isto é, de outubro a dezembro. Perez *et al.* (2014), obtiveram valores minimamente distantes de Dantas *et al.* (2018) e Huang *et al.* (2016), ou seja, concluíram que os pacientes tendem a faltar mais na época de junho a agosto.

As estações do ano que apresentam uma maior taxa de absentismo são a primavera (Bush *et al.* 2016; Mohammadi *et al.* 2016 e Salazar *et al.* 2020) e o outono (Kalb *et al.* 2012; Dantas *et al.* 2018 e Huang *et al.* 2016), sucedendo o verão (Kalb *et al.* 2012 e Perez *et al.* 2014) e por fim, o inverno (Mohammadi *et al.* 2016). No entanto, para Kalb *et al.* (2012) e Perez *et al.* (2014), a primavera é a estação do ano que apresenta maior taxa de comparecimento por parte dos pacientes da clínica.

Quanto aos dias da semana, segunda-feira (Noris *et al.* 2014; Huang *et al.* 2016; Mohammadi *et al.* 2018 e Dantas *et al.* (2018) e quarta-feira (Levy *et al.* 2013; Huang *et al.* 2016 e Mohammadi *et al.* 2018) são os dias da semana que obtiveram maiores taxas de absentismo. De seguida, encontra-se a quinta-feira (Harvey *et al.* 2017 e Dantas *et al.* 2018) e por último, a terça-feira (Noris *et al.* 2014) e a sexta-feira (Levy *et al.* 2013).

Segundo Huang *et al.* (2016), sábado é o dia da semana que apresenta maior taxa de comparecimento dos pacientes às consultas marcadas.

A meteorologia é outro fator a ter em consideração no absentismo dos pacientes. Noris *et al.* (2014) estudaram qual o estado do tempo no dia em que os pacientes se deslocavam à clínica. Estes autores conseguiram concluir que dias de neve e temperaturas baixas ajudam no não comparecimento dos pacientes às consultas, e que nos dias de chuva não há qualquer tipo de influência.

Outro fator bastante relevante é o intervalo entre consultas, isto é, o tempo entre a última consulta e a próxima consulta. Segundo Mohammadi *et al.* (2018), os pacientes faltam mais quando a distância entre a última consulta e a próxima é menor que um ano, para Kalb *et al.* (2012) ronda os 60 dias, 1 semana para Hu *et al.* (2020) e por último, mais de um ano para Mohammadi *et al.* (2018).

Uma das variáveis independentes e fulcrais neste estudo é o tempo de espera, dado pela diferença de dias entre a chamada de marcação da consulta e o dia da própria consulta. Diversos autores chegaram a conclusões diversas. Dantas *et al.* (2018) concluíram que o tempo de espera com maior taxa de absentismo é entre os 16 e os 29 dias, sucedendo entre os 8 e os 15 dias. Drewek *et al.* (2017) conferiu que acima dos 31 dias, a taxa de não comparecimento é maior, e que dos 0 a 30 dias, a percentagem de não comparecimento é muito menor. Perez *et al.* (2014) e Bush *et al.* (2016) constataram que o tempo de espera com maior taxa de absentismo é superior a 28 dias e a 29 dias, respetivamente. Para Mohammadi *et al.* (2018), o tempo de maior absentismo é com um tempo de espera de duas semanas, sendo a marcação no “dia seguinte” o intervalo com menor taxa. Por último, Noris *et al.* (2014) constatou de uma forma geral, que quanto maior o tempo de espera, maior a probabilidade de os pacientes não comparecerem nas consultas marcadas.

Variáveis	Autores	Efeito
Idade	Dantas <i>et al.</i> (2018)	A faixa etária que menos aparece às consultas é entre os 30 e os 44 anos
	Neal <i>et al.</i> (2001);	Entre os 18 e os 34 anos, adultos tem maior probabilidade de faltar
	Mohammadi <i>et al.</i> (2018)	Faixa etária dos 20 anos
	Junod-Perron <i>et al.</i> (2010); Hu <i>et al.</i> (2020)	Ronda a faixa etária dos 40 anos
	Kalb <i>et al.</i> (2012)	Estudou crianças entre os 4 meses e os 21 anos, a taxa de não comparecimento vai aumentando com o aumento das idades
	Noris <i>et al.</i> (2014); Perez <i>et al.</i> (2014); Harvey <i>et al.</i> (2017);	Pacientes mais jovens faltam mais do que pacientes mais velhos
Género	Neal <i>et al.</i> (2001); Dantas <i>et al.</i> (2018); Mohammadi <i>et al.</i> (2018);	As mulheres faltam muito mais vezes do que os homens
	Kalb <i>et al.</i> (2012); Perez <i>et al.</i> (2014); Huang <i>et al.</i> (2016); Bush <i>et al.</i> (2016); Harvey <i>et al.</i> (2017); Junod-Perron <i>et al.</i> (2010);	Os homens são mais propícios a não comparecer nas consultas
Estado civil	Mohammadi <i>et al.</i> (2018)	Os solteiros são os que mais faltam às consultas, o contrário acontece com os enviduados e união de facto;
Distância (Concelho - Hospital)	Dantas <i>et al.</i> (2018)	Pacientes que se encontrem entre os 20 e os 50 km, tendem a não comparecer às consultas agendadas
	Kalb <i>et al.</i> (2012)	A média de não comparecimentos, ronda os 25km
	Perez <i>et al.</i> (2014)	Mais de 80 km os pacientes tendem a faltar mais comparativamente aos pacientes que têm de percorrer menos que 25 Km.
	Huang <i>et al.</i> (2016)	Pacientes que estejam a menos de 8km da clínica, tem maiores probabilidades de faltar às consultas
Número de consultas anteriores	Dantas <i>et al.</i> (2018)	Pacientes que já tenham tido 1 a 2 consultas antes naquela clínica, tem maior taxa de não comparecimento, o contrário acontece com que tenha tido mais de 9 consultas anteriores
Mês	Dantas <i>et al.</i> (2018)	De Setembro a Novembro, é a época que apresenta maior taxa de comparecimento dos pacientes às consultas
	Perez <i>et al.</i> (2014)	Entre junho e agosto os pacientes têm maior probabilidade de faltar às consultas
	Huang <i>et al.</i> (2016)	Os meses que apresentam maior taxa de não comparecimento são de outubro a dezembro
Dia da semana	Dantas <i>et al.</i> (2018)	O dia da semana com maior taxa de não comparecimento é a segunda-feira, sucedendo a quinta-feira
	Huang <i>et al.</i> (2016)	Segunda e quarta-feira, são os dias da semana que apresentam maiores percentagens de não comparecimento, sendo o sábado o dia da semana com maior taxa de comparecimento
	Mohammadi <i>et al.</i> (2018)	Segunda e terça-feira são os dias da semana que mais pacientes tendem a faltar às consultas
	Levy <i>et al.</i> (2013)	Sexta e quarta-feira são dos dias da semana que mais apresentam taxas de absentismo às consultas
	Noris <i>et al.</i> (2014)	O dia da semana que existe mais cancelamentos é à segunda-feira, sucedendo a terça-feira.

	Harvey <i>et al.</i> (2017)	O dia da semana com maior taxa de não comparecimento é a quinta-feira
Meteorologia	Noris <i>et al.</i> (2014)	Neve e a temperatura são os estados do tempo que ajudam no não comparecimento dos pacientes, ao contrário da chuva, que não tem qualquer tipo de influência
Tempo de espera (Intervalo de tempo entre o dia da marcação da consulta e o próprio dia da consulta)	Dantas <i>et al.</i> (2018)	O tempo de espera que contém maior taxa de não comparecimentos dos pacientes é entre os 16 e os 29 dias, sucedendo entre os 8 e os 15 dias
	Drewek <i>et al.</i> (2017)	Acima de 31 dias, a taxa de não comparecimento é maior. De 0 a 30 dias a percentagem de não comparecimento é bem menor.
	Mohammadi <i>et al.</i> (2018)	Dentro de duas semanas é quando há maior taxa de absentismo, sendo o "dia seguinte" o com menor taxa de absentismo
	Bush <i>et al.</i> (2016)	Há maior taxa de não comparecimento acima dos 29 dias
	Noris <i>et al.</i> (2014)	Quanto maior o tempo de espera, maior o não comparecimentos dos pacientes às consultas
	Perez <i>et al.</i> (2014)	O tempo de espera com maior taxa de não comparecimentos é superior a 28 dias, ou seja, 4 semanas.
Não comparências anteriores	Dantas <i>et al.</i> (2018)	Pacientes que tenham faltado pelo menos 50% às consultas anteriores, têm maior probabilidade de faltar às consultas
Estação do ano	Kalb <i>et al.</i> (2012); Perez <i>et al.</i> (2014)	Pacientes faltam mais às consultas nos meses de verão, o contrário acontece na primavera
	Bush <i>et al.</i> (2016); Salazar <i>et al.</i> (2020)	A primavera é a estação do ano em que apresenta maiores taxas de não comparecência dos pacientes às consultas
	Mohammadi <i>et al.</i> (2018)	No inverno e na primavera há mais taxas de absentismo às consultas
	Kalb <i>et al.</i> (2012); Dantas <i>et al.</i> (2018); Huang <i>et al.</i> (2016);	No outono os pacientes faltam mais às consultas marcadas

Tabela 6 - Principais conclusões dos estudos analisados face aos fatores que afetam o absentismo identificado

Capítulo 3 – Metodologia

Neste capítulo serão explicados os métodos utilizados para a recolha e tratamento de dados e os métodos aplicados para analisar os fatores de risco associados ao absentismo dos pacientes. Serão considerados dois horizontes temporais para comparar o comportamento dos pacientes em período de pandemia com o período pré pandemia.

Para além disso, como estudos anteriores referem que existem diferenças significativas em termos de especialidades no campo da análise de absentismo de pacientes, decidiu-se analisar individualmente as cinco especialidades com maior afluência do hospital em análise, nomeadamente: medicina física e reabilitação, ortopedia, oftalmologia, cardiologia e medicina dentária.

Na figura 1 esquematizou-se a metodologia levada a cabo no presente estudo:

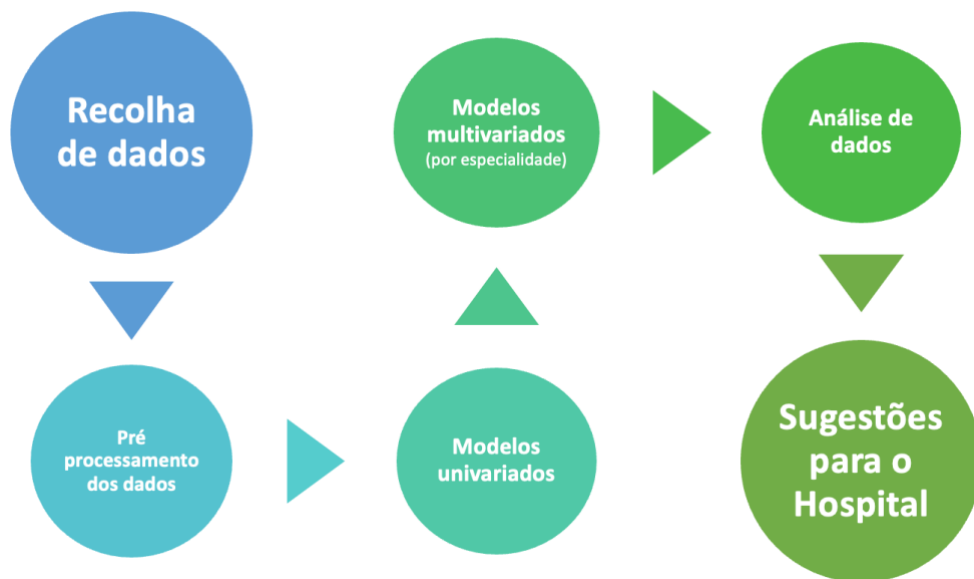


Figura 1 - Metodologia utilizada neste estudo.

3.1 – Recolha de dados

Para a presente análise foram recolhidos dados de um hospital privado localizado no norte de Portugal. Os dados coletados inicialmente (o conjunto de dados originais) são compostos por 18 variáveis e 739.401 registos, como mostra na tabela 7.

Variável	Descrição
ID_Consulta	Identificador da consulta
ID_Paciente	Identificador do paciente
Idade	Idade do paciente em anos
Género	Género do paciente ("M" se masculino, "F" se feminino)
Estado_Civil	Estado civil do paciente (categorias: viúvo, casado, divorciado, solteiro, união de facto, outro)
Profissão	Profissão do paciente
Distrito	Distrito do endereço do paciente
Município	Município de endereço do paciente
ID_especialidade	Indicador da especialidade disponível no hospital
Especialidade	Especialidade disponível no hospital (50 no total)
Tipo_consulta	Tipo de consulta
Primeira_Consulta_Flag	Indicador da primeira consulta: 1 se for a primeira consulta do paciente, 0 caso contrário
Emergência_Flag	Indicador de emergência: 1 se for uma consulta de emergência, 0 caso contrário
Tipo_consulta_flag	Tipo de consulta: "C" se for consulta, "E" se for emergência
Data_consulta	Data da consulta (formato: dd/mm/aaaa)
Estado_consulta_flag	Estado da consulta: "OK" se a consulta foi marcada, "C" se a consulta foi cancelada, "F" se o paciente faltou, "MF" se o médico faltou à consulta, "R" se a consulta foi reagendada
Estado_consulta	Estado da consulta (categorias: ocorrida, cancelada, faltou, medico_faltou, reagendamento)
Data_agendamento_consulta	Data da marcação da consulta (formato: dd/mm/aaaa)

Tabela 7 - Variáveis incluídas no conjunto de dados originais

3.2 – Tarefas de Pré-processamento dos dados

Na fase de pré-processamento dos dados, foram realizadas diversas tarefas.

Iniciando com a correção de erros de digitalização e erros no conjunto de dados. Como os dados contidos no conjunto de dados são derivados da recolha de informações

através de um software, há sempre o problema de alguns dos dados inseridos não serem validados, o que promove a existência dos erros. Mais especificamente:

- foram apagados os registos que continham a “Data_agenda_da_consulta” (data de agendamento da consulta) posterior à “Data_da_consulta”;
- foram apagados os registos que continham pacientes com idade igual a 120 anos tendo sido isso indicado como erro;
- todos os valores de *strings* vazios, ou seja, “ ”, foram declarados como indisponíveis, isto é, NA (not available);

De seguida, procedeu-se à categorização da variável idade como forma de lidar com os *outliers*. Esta variável passou a estar categorizada em 7 categorias, sendo estas: crianças (com idades compreendidas entre os 0 e os 12 anos), adolescentes (com idades compreendidas entre os 13 e os 18 anos), jovens adultos (com idades compreendidas entre os 19 e os 33 anos); adultos (com idades compreendidas entre os 34 e 48 anos); meia-idade: (com idades compreendidas entre os 44 e os 64 anos); idosos (com idades compreendidas entre os 65 e os 70 anos), e por fim, todos os pacientes que tenham mais de 80 anos.

Em relação à criação de variáveis derivadas e exclusão de variáveis foram realizadas as seguintes tarefas:

- A variável “Profissão” foi excluída, pois todos os registos estavam indisponíveis (*missing*). Uma vez que ao prever quais eram as variáveis que poderiam influenciar a não comparência dos pacientes e apenas a considerar as consultas ambulatoriais, todos os registos relacionados à sala de emergências foram eliminados.
- Novas variáveis derivadas foram criadas, tendo em conta as duas variáveis do tipo Data (Data_agendamento_consulta e Data_consulta, refletindo a decomposição dos diferentes membros da data: “Data de Agendamento da consulta ano”, “data de agendamento da consulta mês”, “data de agendamento da consulta dia”, “data de agendamento da consulta hora”, “data de agendamento da consulta minuto”, “data de agendamento da consulta dia da semana”, “dia da consulta ano”, “dia da consulta mês”, “dia da consulta dia” e “dia da consulta dia da semana”).

Considerando a revisão de literatura, existe um conjunto de variáveis importantes (que não existem na base de dados original) que nos permitem prosseguir com a análise pretendida e que podem ter impacto na não comparência do paciente.

As variáveis mais importantes usadas em estudos semelhantes são:

- Distância (distância em km da residência do paciente ao hospital);
- Número marcações anteriores (número de consultas anteriores) e a sua percentagem considerando todas as marcações anteriores;
- Não comparecimentos anteriores (número de não comparecimentos anteriores) e a sua percentagem considerando todas as marcações anteriores;
- Tempo de espera (intervalo entre o agendamento do dia da consulta e o dia da consulta, em horas);
- Comparecimento (caso o paciente compareça à consulta);
- Estação do ano (estação do ano na data da consulta: inverno, outono, primavera ou verão);
- Intervalo entre uma consulta e a primeira consulta atendida por um determinado paciente;
- Intervalo entre uma consulta e a última consulta atendida por um determinado paciente.

Todas as variáveis podem ser derivadas de variáveis já existentes na nossa base de dados, conforme descrito abaixo:

- A variável “distância” foi criada utilizando a distância do município de endereço do paciente até ao hospital em km;
- As variáveis “consultas anteriores” e “não comparecimento” foram criadas calculando-se os comparecimentos anteriores e os não comparecimentos em consultas anteriores e as respetivas percentagens também foram criadas considerando o número de consultas anteriores;
- A variável “tempo de espera” foi criada calculando a diferença entre duas variáveis da base de dados (“Data da consulta” – “Marcação da consulta”) e convertida em horas;
- A variável “Atendimento” foi criada com base no “sinalizador do estado de compromisso” (se “sinalizador do estado de compromisso” = Ok, então “sim”, senão, “não”);

- A variável “Estação do ano” foi criada a partir da data do agendamento;
- A variável “dif_primeira_consulta” foi calculada pela diferença entre a data da consulta atual e a data da primeira consulta atendida por um determinado paciente;
- A variável “dif_última_consulta” foi calculada pela diferença entre a data da consulta atual e a data da última consulta atendida por um determinado paciente.

Como forma de enriquecimento dos dados, também foram adicionados ao conjunto de dados, dados meteorológicos (informações sobre a situação do clima no dia da consulta considerando a localização do hospital). As variáveis adicionadas são: precipitação e velocidade do vento. Todos estes dados foram extraídos usando a API Global Weather (<https://www.visualcrossing.com/weather-api>). Além disso, foi adicionado um conjunto de variáveis cíclicas relacionadas aos atributos do tipo de dados (Mahajan, Singh, & Bruns, 2021).

No final, criaram-se 2 conjuntos de dados, um com os dados da pré-pandemia (intervalo entre 3 de novembro de 2018 e 15 de março de 2020) e, por fim, o conjunto de dados contendo apenas dados durante a pandemia (intervalo entre 3 de abril de 2020 e 1 de dezembro de 2020). Os conjuntos de dados finais que foram utilizados para a modelagem que será descrita na próxima secção são descritos a seguir:

- Um conjunto de dados contendo apenas informações sobre as consultas realizadas durante a pandemia com 147.300 registos e 72 variáveis, o que corresponde a 37.844 pacientes e 145.821 consultas.
- Um conjunto de dados contendo apenas informações sobre as consultas realizadas anteriormente à pandemia com 313.006 registos e 72 variáveis, o que corresponde a informações sobre 53.753 pacientes e 311.563 consultas.

Após o pré-processamento de dados reuniu-se informação de cerca de 28.709 pacientes, correspondente a um total de 465.970 consultas, num período entre 2018 e 2020.

O horizonte temporal da análise foi dividido em três momentos (Tabela 8):

- i) O primeiro momento que decorre durante um período de pré-pandemia, ou seja, entre 3 de novembro de 2018 e 15 de março de 2020.
- ii) O segundo momento - o momento de fecho – é o momento em que o hospital, por decisão própria, esteve fechado para consultas durante o começo da pandemia, estando apenas abertas as urgências - este período ocorreu entre o dia 16 de março de 2020 e o dia 2 de abril de 2020. Assim sendo, este momento não entrou para a análise, uma vez que foi um caso excepcional e não ocorreram consultas, apenas ocorreram urgências.
- iii) Por último, o terceiro momento considerado decorre durante a pandemia - entre 3 de abril de 2020 e 1 de dezembro de 2020.

Momento	Período
Pré-pandemia	3 novembro 2018 a 15 março 2020
Período de fecho (apenas urgências)	16 março 2020 a 2 abril 2020
Pandemia	3 abril 2020 a 1 dezembro 2020

Tabela 8 - Horizontes temporais considerados

Para a presente análise foram reunidas diversas informações à cerca dos pacientes nas quais as variáveis independentes e a variável dependente (Comparecimento) encontram-se mencionadas na tabela 9:

Variável	Descrição
Comparecimento	Variável dummy: 0 se o paciente faltou à consulta, 1 se o paciente compareceu à consulta
Idade	Idade dos pacientes (em anos)
Género	Feminino ou masculino
Estado Civil	Categorias: Viúvo, Casado, Divorciado, Solteiro, União de facto, Outro.
Distância	Distância em km da residência até ao hospital
Consultas anteriores	Número de consultas que os pacientes compareceram anteriormente
Primeira consulta	1 se é a primeira consulta do paciente, 0 caso contrário
Mês	Mês em que a consulta ocorreu
Dia da semana	Dia da semana em que a consulta ocorreu
Precipitação	Nível médio de precipitação no dia da consulta
Velocidade do vento	Velocidade do vento no dia da consulta
Tempo de espera	Tempo, em horas, intervalo de tempo entre o agendamento da consulta e a consulta
Intervalo entre consultas	diferença (em dias) entre a última consulta efetivamente realizada e consulta em questão
Não comparências	Número de não comparências anteriores
Estação do ano	Estação do ano: Primavera, Outono, Verão e Inverno.

Tabela 9 - Variáveis independentes e dependente analisados no estudo

Todas as variáveis independentes que constam na base de dados não apresentam valores omissos.

3.3 – Modelo de regressão logística

Um modelo de regressão logística binomial é um caso particular dos modelos lineares generalizados (GLM) que prevê a probabilidade de ocorrência de um evento (neste caso a presença do paciente na consulta), ajustando covariáveis numéricas ou categóricas a uma função Logit (Hilbe, 2009):

$$\text{Logit}(p) = \log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 \cdot X_1 + \dots + \beta_k \cdot X_k \quad (1)$$

Onde p e $1-p$ são as probabilidades de comparecer e não comparecer à consulta, respetivamente, dado um conjunto de variáveis explicativas (X_i) e coeficientes de regressão desconhecidos (β_i) ($0 < j < k$), a serem estimados por métodos de máxima verossimilhança.

A função logística pode ser escrita como:

$$p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)}} \quad (2)$$

Ou pode-se escrever o modelo em termos de probabilidade como:

$$\frac{p}{1-p} = e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k} \quad (3)$$

Nota-se que os resultados da regressão logística relatam diretamente a razão de chance (*Odds Ratio=OR*) e não as probabilidades. Quanto maior o OR, maior a chance de ocorrer o desfecho (comparecer à consulta) num determinado grupo em relação a outro e, conseqüentemente, maior o risco.

Para os modelos multivariados (para cada uma das cinco especialidades), foi utilizado o *Conditional Backward Elimination*, implementado no SPSS, para a seleção das variáveis que entram no modelo. Este procedimento remove as variáveis não significativas por um teste baseado na probabilidade da estatística razão de verossimilhança baseada em estimativas de parâmetros condicionais.

3.3.1– Teste de Omnibus

Os testes de Omnibus são testes estatísticos que provam se a variância explicada num conjunto de dados é significativamente maior do que a variância inexplicada, numa forma geral.

Este teste implementa um pressuposto geral que tenta encontrar uma significância geral entra a variância dos parâmetros enquanto investiga parâmetros do mesmo género (Stringfixer, 2022).

Assim, temos que a hipótese a testar é:

$$H_0 = \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1 = \text{pelo menos um } \beta_j \neq 0$$

3.3.2– Teste de Hosmer-Lemeshow

Para testar a qualidade do ajuste do modelo, recorreu-se ao teste de Hosmer-Lemeshow (Hosmer, Lemeshow, & Sturdivant, Applied Logistic Regression, 2013) com a hipótese nula de que as proporções observadas e esperadas são as mesmas em todas as categorias. A rejeição da hipótese nula, indica que o modelo não se ajusta adequadamente aos dados.

O teste de Hosmer-Lemeshow avalia as distâncias entre as probabilidades ajustadas e as probabilidades observadas para o modelo ajustado.

Assim, temos que a hipótese a testar é:

H_0 : *O modelo ajusta-se bem aos dados*

vs

H_1 : *O modelo não se ajusta bem aos dados*

O teste calcula a seguinte estatística de teste:

$$\chi_{HL}^2 = \sum_{k=1}^g \frac{(o_k - n'_k \bar{\pi}_k)^2}{n'_k \bar{\pi}_k (1 - \bar{\pi}_k)} \bigcap_{sob H_0} \chi_{g-2}^2$$

Em que:

n'_k - é o número de indivíduos no k-ésimo grupo

c'_k - é o número de covariáveis padrão no k-ésimo grupo

$o_k = \sum_{i=1}^{c'_k} y_i$ - é o número de respostas ao longo das X classes de variáveis

$$\bar{\pi}_k = \sum_{i=1}^{c'_k} \frac{\bar{\pi}_i m_i}{n'_i}$$

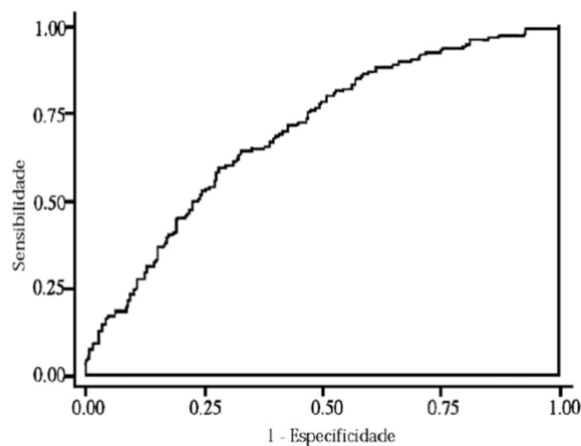
Este teste segue uma distribuição qui-quadrado (g-2 g.1.) quando o modelo se encontra especificado corretamente.

3.4 – Curva de ROC

Para diagnóstico do valor preditivo do modelo recorreremos a uma análise da curva ROC (Receiver Operating Characteristic) pois esta é uma ferramenta que nos permite analisar o desempenho de uma modelo de regressão logística binário, isto é, a variável resposta é do tipo 0-1. O intuito desta análise é reconhecer qual a qualidade do ajustamento dos modelos logísticos aplicados neste trabalho.

Apresenta-se por um gráfico simples e robusto que reporta a variação da sensibilidade e especificidade para diferentes pontos de quebra (pontes de corte) (Gonçalves, 2013).

Na figura 2 é apresentado um exemplo da Curva de ROC.



Fonte: (Nakamura, 2013)

Figura 2 - Exemplo da curva de ROC.

Devemos reportar um ponto de corte C e equiparar cada probabilidade estimada com esse mesmo valor de C. Sendo o valor mais usado para C, o de 0,5 (Hosmer & Lemeshow, 2000).

O ideal será que os valores de ROC apresentem valores próximos de 1, o que nos permite concluir que, quanto mais próxima estiver a curva relativamente ao canto superior esquerdo, mais verdadeiros positivos e menos falsos negativos iremos obter.

A área que se encontra abaixo da curva de ROC, apresenta uma medida de discriminação, que dá a possibilidade de um paciente vir a faltar a uma consulta hospitalar ter uma probabilidade estimada associada mais elevada do que um paciente que não falte às consultas.

Considerando R como o valor que corresponde à área que se encontra abaixo da curva de ROC, de forma geral seguiu-se as seguintes orientações (Gonçalves, 2013), (ver tabela 10):

Área abaixo da curva ROC	Discriminação do modelo
0,5	Não existe discriminação
$0,6 < R < 0,7$	Limitada
$0,7 < R < 0,8$	Aceitável
$0,8 < R < 0,9$	Muito bom
$R \geq 0,9$	Excelente

Tabela 10 - Descrição do modelo segundo a área que se encontra abaixo da curva de ROC

Capítulo 4: Resultados

Neste capítulo estão descritas as análises descritivas e exploratórias da base de dados, assim como a caracterização dos utentes nos períodos de pré e durante a pandemia. É ainda feita, uma comparação entre estes dois períodos. Adicionalmente, apresentam-se os resultados obtidos com o ajuste do modelo logístico por especialidade, pré e durante pandemia, e o diagnóstico do mesmo.

4.1 – Análise Descritiva e Exploratória da Base de Dados

4.1.1 – Caracterização dos utentes

Este estudo é constituído por 28.709 pacientes de um hospital do Norte de Portugal, sendo que a idade dos pacientes varia entre 0 e 113 anos, com uma média de 41,66 anos. Cerca de metade (57,9%) são pacientes do sexo feminino, o que nos dá um total de 16.620 mulheres, sendo os restantes do sexo masculino (42,1% - 12.084 homens). Para os pacientes com informação sobre o estado civil, a maioria é solteiro ou casado (1.223 e 930 pacientes, respetivamente), os restantes são viúvos, divorciados, estão em união de facto e apresentam outro estado (55, 50, 15 e 7, respetivamente).

Nesta análise existem pacientes vindos de todo o território português, isto é, tanto de Portugal Continental (18 distritos), como das regiões autónomas dos Açores (9 ilhas) e da Madeira (2 ilhas). A maioria dos pacientes que frequentaram este hospital são do distrito de Braga (83,1%). Os distritos seguintes com maior taxa de pacientes são os distritos do Porto (9,7%) e de Vila Real (1,5%). Pelo contrário, apenas 37 pacientes pertencem à região autónoma dos Açores e 16 pacientes são da região autónoma da Madeira.

Relativamente aos concelhos do distrito de Braga, são predominantes os concelhos de Guimarães, Fafe e Vizela, com cerca de 15.621, 3.205 e 1.918 pacientes sucessivamente. Quanto ao distrito do Porto, os concelhos com maior número de pacientes, são o concelho de Felgueiras e o de Santo Tirso, com um total de 1.475 e 617 pacientes, respetivamente.

No que concerne à distância entre o hospital e a zona onde moram os pacientes, o paciente que vive mais perto do hospital tem de percorrer uma distância de 2,57 km, enquanto o paciente mais longe, tem de percorrer 1946,94 km. A média da distância entre o hospital e a zona de residência dos utentes é de 84,22 km, aproximadamente.

A média de consultas realizadas neste Hospital é de 9,08 consultas por paciente, sendo o mínimo de consultas realizadas por pacientes, 1 consulta e o máximo de 639 consultas realizadas. Pelo contrário, a média de consultas não realizadas é de 3,68 consultas, apresentando um mínimo de 1 consulta não realizada e um máximo de 782 consultas não realizadas (ver tabela 11).

	Média	SD	Min	Max
Idade	41,66	21,618	0	113
Distância	84,22	339,83	2,57	1946,94
Total de consultas realizadas	9,08	14,476	1	639
Total de consultas não realizadas	3,68	8,119	1	782
	N	%		
Sexo				
Masculino	12,084	42,1		
Feminino	16,620	57,9		
Estado civil				
Casado	930	3,2		
Divorciado	50	0,2		
Outro	7	0		
Solteiro	1,223	4,3		
União de facto	15	0,1		
Viúvo	55	0,2		
Distrito*				
Braga	23,849	83,1		
Porto	2,784	9,7		
Vila Real	427	1,5		
Açores	37	0		
Madeira	16	0		
Concelho*				
Guimarães	15,621	54,4		
Fafe	3,205	11,2		
Vizela	1,918	6,7		
Felgueiras	1,475	5,1		
Santo Tirso	617	2,1		

Tabela 11 - Estatísticas descritivas dos pacientes

4.1.2 – Caracterização das consultas

4.1.2.1 - Antes da pandemia

Todas as variáveis independentes que estão incluídas para a caracterização de consultas não apresentam valores omissos.

Uma vez que se trata do período antes da pandemia, apenas foram utilizados os dados entre os dias 3 de novembro de 2018 e 15 de março de 2020, obtendo no total, 313,006 consultas.

Antes da pandemia, taxa de não comparecimento dos pacientes às consultas é de 22,9%, o que 77,1% dos pacientes compareceu à consulta marcada. Deste número total de consultas, a maioria das consultas marcadas são consultas subsequentes (19,3%), primeira consulta (18,7%), consulta de fisioterapia (16,1%), consulta de rotina (7,5%) e consulta de enfermagem (4%). As especialidades com maior taxa de recorrência foram a Medicina Física e Reabilitação (19,7%), Medicina Dentária (9,3%), Cardiologia (7,5%), Ortopedia (7,2%), e por último, Oftalmologia (7%).

Neste período de pré-pandemia, marcaram-se consultas maioritariamente nos meses de janeiro (13,2%) e novembro (12,1%) e minoritariamente em Abril (5,6%) e Junho (5,3%). Durante a semana, a segunda-feira (20,1%) e terça-feira (20%) foram os dias com maior taxa de marcação, contrariamente aos fins-de-semana as taxas de marcação são bastante reduzidas, ou seja, Domingo (0,2%) e Sábado (7,4%).

Registou-se um maior número de consultas agendadas durante o Inverno (35,7%) e a Primavera (27,8%), de forma controversa, o Verão (17,9%) e Outono (18,5%) foram as estações com menor taxa de agendamento.

O número médio de consultas realizadas por paciente é 10,75 consultas, apresentando um máximo de 412 consultas. Pelo contrário, o número médio de consultas não comparecidas é de 4,12 consultas, apresenta um máximo de 406 consultas por paciente não comparecidas. Para a grande maioria dos pacientes não se trata da primeira consulta (97,9%) naquela organização, tratando-se da primeira consulta no hospital apenas para 2,1%.

A média do tempo de espera de um paciente para obter uma determinada consulta é de 406,19 horas, o que corresponde aproximadamente 17 dias. O mínimo tempo de espera é de 12 horas enquanto o máximo é de 35100,78 horas, o que dá um

total de 4 anos de espera. Quanto à diferença, em horas, entre a primeira consulta e a respetiva consulta, a média de horas atinge 127,24 horas, o que representa 5 dias, atingido um máximo de 498 horas (20 dias). Em relação à diferença entre a última consulta administrada e a consulta em questão, a média de espera é de 26,12 horas (aproximadamente 1 dia) e um máximo de 497 horas (20 dias).

Durante estes dias de consultas, a percentagem de precipitação rondou os 6,35%, com um mínimo de 0% e um máximo de 97,36%. No que diz respeito ao vento, a média durante esse período foi de 19,93 km/h, podendo atingir um máximo de 55,8 km/h e um mínimo de 8,2 km/h (ver tabela 12).

	Média	SD	Min	Max
Número consultas realizadas	10,75	22,32	0	412
Número consultas não comparecimentos	4,12	15,02	0	406
Precipitação	6,35	12,95	0	97,36
Velocidade do vento	19,93	7	8,2	55,8
Tempo de espera (horas)	406,19	633,16	12,39	35100,78
Diferença entre a última consulta e a consulta (horas)	26,12	55,43	0	497
Diferença entre a primeira consulta e a consulta (horas)	127,24	137,08	0	498
	N	%		
Comparências				
Faltou	71,818	22,9		
Compareceu	241,188	77,1		
Especialidade*				
Medicina Física e Reabilitação	61,522	19,7		
Medicina Dentária	29,030	9,3		
Cardiologia	23,363	7,5		
Ortopedia	22,585	7,2		
Oftalmologia	21,992	7,0		
Tipo de consulta*				
Consulta Subsequente	60,352	19,3		
Primeira consulta	58,441	18,7		
Fisioterapia	50,545	16,1		
Consulta	23,178	7,4		
Enfermagem	12,470	4		
Estação do ano				
Outono	57,869	18,5		
Primavera	87,163	27,8		
Verão	56,155	17,9		
Inverno	111,819	35,7		
Dia da semana				
Domingo	538	0,2		
Segunda-feira	63,009	20,1		
Terça-feira	62,461	20		

Quarta-feira	53,697	17,2
Quinta-feira	58,462	18,7
Sexta-feira	51,829	16,6
Sábado	23,010	7,4
Mês da consulta		
Janeiro	41,471	13,2
Fevereiro	37,106	11,9
Março	23,151	7,4
Abril	17,637	5,6
Mai	19,707	6,3
Junho	16,629	5,3
Julho	20,575	6,6
Agosto	17,735	5,7
Setembro	20,821	6,7
Outubro	28,116	9
Novembro	37,740	12,1
Dezembro	32,318	10,3
Primeira vez		
Sim	6,680	2,1
Não	306,326	97,9

Tabela 12 - Características descritivas das consultas (antes da pandemia)

4.1.2.2 - Durante a pandemia

Neste subcapítulo será realizada uma análise descritiva das consultas durante o período de pandemia, isto é, de 3 de abril de 2020 a 1 de dezembro de 2020. Durante este período, marcaram-se neste hospital cerca de 147,300 consultas, nas quais a média destas consultas realizadas é de 329,80 consultas, havendo pacientes com um máximo de 634 consultas realizadas. No que diz respeito a consultas não comparecidas, a média é de 7,08 consultas não comparecidas, tendo um máximo de 781 consultas.

Durante a pandemia, a taxa de absentismo dos pacientes às consultas marcadas é de 23,6%, o que corresponde a um total de 34,820 consultas não comparecidas. Aproximadamente 4,9% (7,275 consultas) foi a 1ª consulta naquela instituição de saúde e 95,1% não foram a primeira vez.

A maioria destas consultas são consultas subsequentes (18,1%), primeiras consultas (15,7%), consultas de fisioterapia (11,5%), consultas de rotina (7,5%) e por último, consultas de enfermagem (4,7%). Quanto às especialidades, as com maior taxa de recorrência são: a Medicina Física e de Reabilitação (14,4%), Cardiologia (7,9%), Ortopedia (7,2%), Oftalmologia (6,7%) e Medicina dentária (6,4%).

Relativamente a fatores meteorológicos, a percentagem de precipitação rondou os 3,15%, com um mínimo de 0% e um máximo de 62,66%. No que diz respeito ao vento,

a média durante esse período de tempo foi de 19,64 km/h, podendo atingir um máximo de 36,6 km/h e um mínimo de 9,6 km/h.

Neste período de pandemia, marcaram-se consultas com maior frequência nos meses de novembro (15,7%) e outubro (15,1%) e minoritariamente em abril (4,9%) e dezembro (0,6%). Durante a semana, a quarta-feira (18,8%) e sábado (18,3%) foram os dias com maior taxa de marcação, contrariamente as taxas de marcação são bastante reduzidas à segunda-feira (0,1%) e ao domingo (10,5%).

Os pacientes obtiveram mais consultas agendadas durante o Verão (36,7%) e a Outono (42,2%), de forma controversa, o Inverno (16,3%) e Primavera (4,9%) foram as estações com menor taxa de agendamento.

A média do tempo de espera de um paciente para obter uma determinada consulta é de 404,75 horas, o que corresponde aproximadamente 17 dias. O mínimo tempo de espera é de 24 horas enquanto o máximo é de 17959,9 horas, o que dá um total de 2 anos de espera. Quanto à diferença, em horas, entre a primeira consulta e a respetiva consulta, a média de horas atinge 329,8 horas, o que representa 13 dias, atingido um máximo de 759 horas (31 dias). No que concerne à diferença entre a última consulta administrada e a consulta em questão, a média de espera é de 56,68 horas (aproximadamente 3 dias) e um máximo de 753 horas (31 dias) (ver tabela 13)

	Média	SD	Min	Max
Número consultas realizadas	329,80	260,68	0	634
Número consultas não comparecimentos	7,08	33,57	0	781
Precipitação	3,15	8,29	0	62,66
Velocidade do vento	19,64	5,52	9,6	36,6
Tempo de espera (Horas)	404,75	626,36	24,35	17959,9
Diferença entre a última consulta e a consulta (horas)	56,68	113,06	0	753
Diferença entre a primeira consulta e a consulta (horas)	329,8	260,68	0	759
	N	%		
Comparências				
Faltou	34,820	23,6		
Compareceu	112,480	76,4		
Especialidade*				
Medicina Física e Reabilitação	21,140	14,4		
Medicina Dentária	9,389	6,4		
Cardiologia	11,621	7,9		
Ortopedia	10,614	7,2		

Oftalmologia	9,915	6,7
Tipo de consulta*		
Consulta Subsequente	26,628	18,1
Primeira consulta	23,089	15,7
Fisioterapia	16,872	11,5
Consulta	11,121	7,5
Enfermagem	6,915	4,7
Estação do ano		
Outono	62,088	42,2
Primavera	7,150	4,9
Verão	54,022	36,7
Inverno	24,040	16,3
Dia da semana		
Domingo	15,433	10,5
Segunda-feira	201	0,1
Terça-feira	25,275	17,2
Quarta-feira	27,723	18,8
Quinta-feira	26,057	17,7
Sexta-feira	25,706	17,5
Sábado	26,905	18,3
Mês da consulta		
Abril	7,150	4,9
Maio	14,775	10
Junho	17,663	12
Julho	21,584	14,7
Agosto	18,143	12,3
Setembro	21,671	14,7
Outubro	22,274	15,1
Novembro	23,094	15,7
Dezembro	946	0,6
Primeira vez		
Sim	7,275	4,9
Não	140,025	95,1

Tabela 13 - Características descritivas das consultas (durante a pandemia)

4.1.2.3 – Comparação pré-pandemia com pandemia

A comparação entre estes períodos não pode ser feita de forma equivalente, visto que o período de pré-pandemia apresenta maior número de observações, isto é, totaliza 498 dias (1 ano, 4 meses e 2 dias) de consultas enquanto o período de pandemia representa quase metade, apresentando apenas 242 dias (7 meses e 28 dias) de marcação de consultas.

De forma geral, observa-se que a média do número de consultas realizadas é bastante superior no período de pandemia comparativamente ao período de pré-pandemia, sendo a média 10,75 consultas na pré-pandemia e 329,80 consultas na

pandemia. O mesmo acontece com o número de consultas não realizadas, ou seja, houve mais consultas não realizadas durante a pandemia.

O tempo médio de espera de um paciente para uma consulta é bastante semelhante nas duas épocas, havendo apenas uma grande diferença no valor máximo, sendo 4 anos no período de pré-pandemia e 2 anos no período de pandemia.

As especialidades com maior taxa de ocorrência são as mesmas nos dois períodos, o mesmo acontece com o tipo de consulta.

A taxa de não comparecimento do período de pandemia é muito semelhante com o período de pré-pandemia, sendo que no período de pandemia, 34,820 (23,6%) pacientes faltaram às consultas enquanto no período de pré-pandemia, cerca de 71,818 (22,9%) não compareceram às consultas agendadas.

Não esquecendo que o horizonte temporal não é idêntico nas duas épocas, sendo na pré-pandemia o quase o dobro dos dias durante a pandemia, de uma forma oposta, na pré-pandemia os dias de semana com maior taxa de marcação de consultas são às segundas-feiras e terças-feiras enquanto no período de pandemia são aos sábados e às quartas-feiras. Pelo mesmo seguimento, os meses também não são do mesmo consentimento nas duas épocas, sendo janeiro e novembro os meses com maior percentagem de consultas marcadas na pré-pandemia e outubro e novembro durante a pandemia. Quanto às estações do ano, também não se encontram em concordância, uma vez que na pré-pandemia as estações mais ocorridas foram na primavera e no inverno; e outono e verão na pandemia.

Relativamente a dados meteorológicos, antes da pandemia os valores de precipitação rondam os 6,35%, enquanto o valor durante a pandemia é muito semelhante, ou seja, 3,15%. No que diz respeito à velocidade do vento, o valor durante estas épocas é também bastante idêntico, 19,93 km/h e 19,64km/h na pré-pandemia e na pandemia, respetivamente.

Na pré-pandemia o tempo máximo de espera entre a última consulta e a consulta em causa é de 26,12 horas, sendo que na pandemia este valor é bastante superior, 56,68 horas. Quanto ao tempo máximo de espera entre a primeira consulta do paciente no respetivo hospital e a consulta em questão é de 127,24 horas durante o período de pré-pandemia e 329,8 horas durante a pandemia.

Antes da pandemia, cerca de 97,9% dos utentes não recorreu a este hospital pela primeira vez, sendo que durante a pandemia, os pacientes não se alteraram muito, uma vez que 95,1% dos pacientes durante a pandemia não recorreram também a este hospital pela primeira vez.

4.2 – Resultados por especialidade

Os modelos logísticos binários univariados do conjunto de variáveis testadas, escolhidas através da literatura, permitem entender individualmente que variáveis ao nível do paciente, como idade, sexo, estado civil e distância de casa ao hospital, apresentaram efeito estatisticamente significativo na probabilidade de comparecer a uma consulta. O número de consultas anteriores e se é ou não a primeira consulta, estão relacionados a um maior risco de comparecimento. Em relação às variáveis exógenas, verificou-se um efeito significativo do mês, dia da semana, nível médio de precipitação, nível médio da velocidade do vento e da estação do ano na probabilidade de comparecer à consulta. Além disso, variáveis relacionadas com o comportamento do paciente frente ao risco, como o tempo de espera, o intervalo entre consultas e o número de não comparecimentos anteriores, também se apresentam como fatores de risco para a probabilidade de comparecimento. Valores mais altos do tempo de espera e não comparecimentos anteriores diminuem as chances de comparecer à consulta.

Conforme mencionado anteriormente, os modelos logísticos multivariados foram ajustados para as cinco especialidades com maior número de consultas marcadas, considerando todas as variáveis significativas obtidas nos modelos univariados (Salazar, *et al.*, 2020). No entanto, o procedimento *Conditional Backward Elimination* descartou algumas variáveis, resultando em cinco modelos com variáveis significativas como idade do paciente, sexo, estado civil, distância da residência até ao hospital, número de consultas anteriores, primeira consulta no hospital, mês da consulta, dia da semana da consulta, estação do ano, níveis de precipitação, níveis da velocidade do vento, tempo de espera, intervalo entre consultas e número de não comparecimentos anteriores.

Observando os dados reportados pelo modelo logístico multivariado binário para a especialidade de ortopedia antes e durante a pandemia, é de referir que antes da

pandemia nesta especialidade, as variáveis: estado civil, idade, não comparências anteriores e tempo de espera são estatisticamente significativas enquanto durante a pandemia apenas foram estatisticamente significativas as variáveis: mês do ano, não comparências anteriores e o tempo de espera.

A partir dos resultados, como tentativa de perfilar os pacientes que recorreram à especialidade de ortopedia no período de pré-pandemia é de notar que pacientes divorciados ($e^\beta = 0,456$) tem maior chance de não comparecer do que os casados, pacientes idosos ($e^\beta = 4,686$) tem maior chance de comparecer relativamente às crianças; o tempo de espera, ou seja, o tempo entre o agendamento da consulta e a data da consulta, está significativamente relacionado à chance de não comparecer à consulta, pois quanto maior o tempo entre o agendamento e a data da consulta, maior o risco de não comparência ($e^\beta = 0,998$) e o número de não comparências anteriores, quanto maior o número de consultas não comparecidas anteriormente, maior o risco de não comparência ($e^\beta = 0,977$).

Relativamente aos pacientes que recorreram à especialidade de ortopedia durante a pandemia (ver tabela 14), é de concluir que pacientes que tiveram consultas marcadas nos meses de maio ($e^\beta = 1,43$), junho ($e^\beta = 1,425$), agosto ($e^\beta = 1,733$) e outubro ($e^\beta = 1,365$) tem maior probabilidade de comparecer às consultas do que pacientes que tiveram consultas marcadas no mês de abril; quanto ao número de não comparências anteriores ($e^\beta = 0,979$), maior o risco dos pacientes faltarem às consultas, o mesmo acontece com o tempo de espera ($e^\beta = 0,999$), isto é, quanto maior o distanciamento entre a marcação da consulta e a data da consulta, maior a chance de não comparecer às consultas de ortopedia.

Ortopedia (pré-pandemia)				Ortopedia (durante a pandemia)			
Variável	e^β	IC95%	p-valor	Variável	e^β	IC95%	p-value
Estado civil (referência:casado)				Mês (referência:abril)			
Divorciado	0,456	[0,226;0,920]	0,028	Maio	1,43	[1,110;1,843]	0,006
Solteiro	1,708	[0,957;3,047]	0,07	Junho	1,425	[1,114;1,822]	0,005
União de facto	341058534	[0;.]	1	Julho	1,223	[0,971;1,542]	0,087
Viúvo	0,531	[0,261;1,082]	0,081	Agosto	1,733	[1,350;2,225]	<,001
Idade (referência:criança)				Setembro	1,245	[0,989;1,569]	0,063
Adolescente	1,076	[0,607;1,905]	0,802	Outubro	1,365	[1,083;1,720]	0,008
Jovem adulto	0,888	[0,496;1,590]	0,689	Novembro	0,98	[0,781;1,229]	0,862
Adulto	1,983	[0,941;4,180]	0,072	Dezembro	0,715	[0,341;1,499]	0,375
Meia-idade	1,872	[0,873;4,013]	0,107	Não comparências anteriores	0,979	[0,974;0,984]	<,001

Idoso	4,686	[2,085;10,532]	<,001	Tempo de espera	0,999	[0,998;0,999]	<,001
Mais de 80	2,407	[0,928;6,242]	0,071	Teste de Omnibus	Qui-quadrado	gl	p-valor
Não comparências anteriores	0,977	[0,961;0,992]	0,003		477,195	10	<,001
Tempo de espera	0,998	[0,998;0,999]	<,001	R2	Verossimilhança de log -2	R quadrado Cox & Snell	R quadrado Nagelkerke
Teste de Omnibus	Qui-quadrado	gl	p-valor		9055,061	0,044	0,074
	101,274	12	<,001	Teste de Hosmer e Lemeshow	Qui-quadrado	gl	p-valor
R2	Verossimilhança de log -2	R quadrado Cox & Snell	R quadrado Nagelkerke		18,463	8	0,018
	1342,616	0,069	0,108				
Teste de Hosmer e Lemeshow	Qui-quadrado	gl	p-valor				
	16,205	8	0,04				

Tabela 14 - Modelo logístico multivariado de ortopedia.

Na especialidade de oftalmologia (ver tabela 15), antes da pandemia, o modelo multivariado logístico reportou que as variáveis, estado civil, idade, não comparências anteriores, dia da semana e tempo de espera são estatisticamente significativas no que diz respeito à comparência dos pacientes às consultas de oftalmologia antes da pandemia. Durante a pandemia, as variáveis estatisticamente significativas são o género, idade, não comparências anteriores, número de consultas anteriores, mês da consulta e tempo de espera.

Os utentes que recorreram ao serviço da oftalmologia antes da pandemia, é de notar que pacientes que se encontram em situação de união de facto ($e^{\beta} = 0,078$) tem menor chance comparecer às consultas do que utentes que sejam casados; pacientes idosos e que tenham mais de 80 anos ($e^{\beta} = 3,386$) e ($e^{\beta} = 2,334$) respetivamente, tem maior chance de comparecer às consultas face a pacientes que sejam crianças; utentes que tenham tido consultas às quartas-feiras ($e^{\beta} = 0,54$) tem maior chance de não comparecer do que aos domingos; não comparências anteriores ($e^{\beta} = 0,976$) e o tempo de espera ($e^{\beta} = 0,9996$) são fatores que influenciam numa maior chance de os utentes faltarem às consultas, ou seja, quanto maior o número de não comparências anteriores e maior tempo de espera, maior o risco de não comparência dos utentes às consultas de oftalmologia.

Em comparação com os pacientes que recorreram ao serviço de oftalmologia durante a pandemia, consta-se que pacientes do sexo masculino ($e^{\beta} = 1,215$) tem maior chance de comparecer do que pacientes do sexo feminino; jovens adultos ($e^{\beta} = 1,573$), adultos ($e^{\beta} = 1,207$), meia-idade ($e^{\beta} = 1,671$), envelhecidos ($e^{\beta} = 1,725$) e com mais de 80 anos ($e^{\beta} = 1,592$) tem maior probabilidade de comparecer a consultas agendas do que pacientes que sejam crianças; no que diz respeito ao tempo em horas ($e^{\beta} = 0,999$) e o número de não comparências anteriores ($e^{\beta} = 0,95$) têm maior risco de não comparência nas consultas de oftalmologia durante a pandemia; o número de consultas anteriores apresenta um efeito marginal ao nível da significância de 5%, ou seja, quanto maior o número de consultas anteriores ($e^{\beta} = 1,023$) maior a chance de comparência dos pacientes nas consultas de oftalmologia; pacientes que recorrem ao hospital para consultas de oftalmologia durante os meses de maio ($e^{\beta} = 1,894$), junho ($e^{\beta} = 1,881$), julho ($e^{\beta} = 2,148$), agosto ($e^{\beta} = 2,815$), setembro ($e^{\beta} = 1,787$), outubro ($e^{\beta} = 2,505$), novembro ($e^{\beta} = 2,166$) e dezembro ($e^{\beta} = 2,493$) tem maior chance de comparecer a consultas face ao mês de abril.

Oftalmologia (pré-pandemia)				Oftalmologia (durante a pandemia)			
Variável	e^{β}	IC95%	p-value	Variável	e^{β}	IC95%	p-value
Estado civil (referência: casado)				Género (referência: feminino)			
Divorciado	1,43	[0,594;3,444]	0,425	Masculino	1,215	[1,094;1,350]	<,001
Outro	719077095	[0;.]	1	Idade (referência: criança)			
Solteiro	1,536	[0,894;2,637]	0,12	Adolescente	1,086	[0,852;1,384]	0,504
União de facto	0,078	[0,009;0,698]	0,022	Jovem adulto	1,573	[1,277;1,936]	<,001
Viúvo	1,381	[0,623;3,065]	0,427	Adulto	1,207	[1,012;1,438]	0,036
Idade (referência: criança)				Meia-idade	1,671	[1,406;1,987]	<,001
Adolescente	0,946	[0,634;1,41]	0,784	Idoso	1,725	[1,432;2,079]	<,001
Jovem adulto	0,871	[0,569;1,333]	0,525	Mais de 80	1,592	[1,217;2,082]	<,001
Adulto	1,028	[0,575;1,840]	0,925	Não comparências anteriores	0,95	[0,940;0,960]	<,001
Meia-idade	1,212	[0,648;2,269]	0,547	Nº consultas anteriores	1,023	[1,017;1,030]	<,001
Idoso	3,386	[1,765;6,496]	<,001	Mês (referência: abril)			
Mais de 80	2,334	[1,010;5,395]	0,047	Maio	1,894	[1,451;2,473]	<,001
Não comparências anteriores	0,976	[0,952;1]	0,049	Junho	1,881	[1,467;2,411]	<,001
Dia da semana (referência: domingo)				Julho	2,148	[1,676;2,754]	<,001
Segunda	0,324	[0,60;1,748]	0,17	Agosto	2,815	[2,169;3,653]	<,001
Terça	0,917	[0,573;1,466]	0,716	Setembro	1,787	[1,406;2,271]	<,001
Quarta	0,54	[0,356;0,817]	0,004	Outubro	2,505	[1,952;3,216]	<,001
Quinta	0,74	[0,475;1,152]	0,182	Novembro	2,166	[1,683;2,788]	<,001
Sexta	0,671	[0,443;1,015]	0,059	Dezembro	2,493	[1,109;5,605]	0,027

Sábado	0,965	[0,583;1,597]	0,89	Tempo de espera	0,999	[0,998;0,999]	<,001
Tempo de espera	0,9996	[0,9995;0,9998]	<,001	Teste de Omnibus	Qui-quadrado	gl	p-valor
Teste de Omnibus	Qui-quadrado	gl	p-valor		947,347	18	<,001
	101,34	19	<,001	R2	Verossimilhança de log -2	R quadrado Cox & Snell	R quadrado Nagelkerke
R2	Verossimilhança de log -2	R quadrado Cox & Snell	R quadrado Nagelkerke		9142,433	0,091	0,143
	1918,25	0,046	0,076	Teste de Hosmer e Lemeshow	Qui-quadrado	gl	p-valor
Teste de Hosmer e Lemeshow	Qui-quadrado	gl	p-valor		67,726	8	<,001
	7,467	8	0,487				

Tabela 15 - Modelo logístico multivariado de oftalmologia.

Nas consultas de cardiologia (ver tabela 16), antes da pandemia, o resultado que se obteve no modelo logístico multivariado foi que as variáveis estatisticamente significativas foram: o estado civil, o número de consultas não comparecidas anteriormente, o número de consultas obtidas anteriormente no hospital, e por último o tempo de espera. Durante a pandemia, obteve outras variáveis significativas como: a velocidade do vento, o intervalo entre consultas, o tempo de espera e o número de consultas não comparecidas anteriormente.

No período de pré-pandemia, no serviço de cardiologia, pacientes que estejam divorciados ($e^{\beta} = 0,379$) tem maior chance de não comparecer às consultas de cardiologia do que os pacientes que sejam casados; o tempo de espera ($e^{\beta} = 0,9997$) e o número de não comparências anteriores ($e^{\beta} = 0,892$) estão diretamente relacionados com a chance de não comparências às consultas de cardiologia; Pelo contrário, o número de consultas anteriores ($e^{\beta} = 1,045$) está diretamente relacionado com a chance de comparência dos utentes face a consultas de cardiologia.

De outro modo, durante a pandemia o número de não comparências anteriores ($e^{\beta} = 0,983$), a velocidade do vento ($e^{\beta} = 0,986$) tem maior risco de os pacientes não comparecerem às consultas de cardiologia; o intervalo entre consultas ($e^{\beta} = 0,998$), tempo entre a última consulta obtida e data da consulta, também está associada à chance de não comparência dos utentes, isto é, quanto maior o intervalo entre consultas, maior o risco de absentismo dos pacientes às consultas de cardiologia durante o período de pandemia; Do mesmo modo, quanto maior o tempo de espera

($e^\beta = 0,9996$), maior o risco dos utentes faltarem às consultas de cardiologia marcadas.

Cardiologia (pré-pandemia)				Cardiologia (durante a pandemia)			
Variável	e^β	IC95%	p-value	Variável	e^β	IC95%	p-value
Estado civil (referência: casado)				Não comparências anteriores	0,983	[0,977;0,988]	<,001
Divorciado	0,379	[0,229;0,625]	<,001	Velocidade do vento	0,986	[0,978;0,994]	<,001
Outro	1,186	[0,142;9,914]	0,875	Intervalo entre consultas	0,998	[0,998;0,999]	<,001
Solteiro	1,26	[0,891;1,782]	0,191	Tempo de espera	0,9996	[0,9996;0,9997]	<,001
União de facto	6,07E+09	[0;.]	0,998	Teste de Omnibus	Qui-quadrado	gl	p-valor
Viúvo	0,717	[0,447;1,149]	0,167		319,612	4	<,001
Não comparências anteriores	0,892	[0,857;0,928]	<,001	R2	Verossimilhança de log -2	R quadrado Cox & Snell	R quadrado Nagelkerke
Nº consultas anteriores	1,045	[1,024;1,066]	<,001		11401,105	0,027	0,043
Tempo de espera	0,9997	[0,9995;0,9998]	<,001	Teste de Hosmer e Lemeshow	Qui-quadrado	gl	p-valor
Teste de Omnibus	Qui-quadrado	gl	p-valor		163,196	8	<,001
	88,506	8	<,001				
R2	Verossimilhança de log -2	R quadrado Cox & Snell	R quadrado Nagelkerke				
	1566,785	0,049	0,081				
Teste de Hosmer e Lemeshow	Qui-quadrado	gl	p-valor				
	34,136	8	<,001				

Tabela 16 - Modelo logístico multivariado de cardiologia.

Relativamente aos pacientes que recorreram ao serviço de medicina dentária (ver tabela 17) antes da pandemia, denota-se que os homens ($e^\beta = 1,236$) tem maior chance de comparecer às consultas de medicina dentária do que as mulheres; utentes sem estado civil não descrito (outro) ($e^\beta = 0,075$) tem menor probabilidade de comparecer à consulta face a utentes que se encontram casados; jovens adultos ($e^\beta = 0,753$) tem maior probabilidade de faltar às consultas do que as crianças; Em relação ao comportamento do paciente em relação ao risco, o tempo de espera ($e^\beta = 0,9996$) está significativamente relacionado à chance de não comparecer às consultas, assim

como o número de não comparências anteriores ($e^{\beta} = 0,931$); Pelo contrário, o número de consultas anteriores ($e^{\beta} = 1,038$) está positivamente relacionado com a chance de comparecer às consultas, ou seja, quanto maior o número de consultas realizadas anteriormente maior a chance de o paciente comparecer à consulta.;

Neste serviço, durante a pandemia, os homens ($e^{\beta} = 1,469$) continuam a ser os pacientes com maior chance de comparecer do que as mulheres; o tempo de espera ($e^{\beta} = 0,999$) e o número não de comparências anteriores ($e^{\beta} = 0,935$) tem um efeito significativo na chance de não comparecer às consultas de medicina dentária; Contrariamente, quanto maior o número de consultas anteriores ($e^{\beta} = 1,469$) maior a chance de os pacientes comparecerem às consultas marcadas de medicina dentária.

Medicina Dentária (pré-pandemia)				Medicina Dentária (durante a pandemia)			
Variável	e^{β}	IC95%	p-value	Variável	e^{β}	IC95%	p-value
Género (referência:feminino)				Género (referência:feminino)			
Masculino	1,236	[1,041;1,466]	0,015	Masculino	1,469	[1,099;1,965]	0,009
Estado civil (referência: casado)				Não comparências anteriores	0,935	[0,905;0,965]	<,001
Divorciado	0,62	[0,378;1,016]	0,058	Nº consultas anteriores	1,031	[1,014;1,050]	<,001
Outro	0,075	[0,006;0,887]	0,04	Tempo de espera	0,9990	[0,998;0,999]	<,001
Solteiro	1,044	[0,748;1,456]	0,801	Teste de Omnibus	Qui-quadrado	gl	p-valor
União de facto	0,447	[0,184;1,085]	0,075		128,479	4	<,001
Viúvo	1,553	[0,439;5,497]	0,495	R2	Verossimilhança de log -2	R quadrado Cox & Snell	R quadrado Nagelkerke
Idade (referência: criança)					1155,161	0,115	0,163
Adolescente	1,113	[0,843;1,469]	0,451	Teste de Hosmer e Lemeshow	Qui-quadrado	gl	p-valor
Jovem adulto	0,753	[0,578;0,981]	0,035		8,912	8	0,35
Adulto	0,707	[0,487;1,026]	0,068				
Meia-idade	1,125	[0,716;1,766]	0,61				
Idoso	0,882	[0,534;1,456]	0,624				
Mais de 80	1,168	[0,575;2,374]	0,667				
Não comparências anteriores	0,931	[0,905;0,958]	<,001				
Nº consultas anteriores	1,038	[1,019;1,057]	<,001				
Tempo de espera	0,9996	[0,9995;0,9997]	<,001				
Teste de Omnibus	Qui-quadrado	gl	p-valor				
	121,798	15	<,001				
R2	Verossimilhança de log -2	R quadrado Cox & Snell	R quadrado Nagelkerke				
	3497,76	0,037	0,055				
Teste de Hosmer e Lemeshow	Qui-quadrado	gl	p-valor				
	6,961	8	0,541				

Tabela 17 - Modelo logístico multivariado de medicina dentária.

Para a especialidade de Medicina Física e Reabilitação (ver tabela 18), no período de pré-pandemia, os homens ($e^{\beta} = 1,239$) tem maior chance de comparecer às consultas do que as mulheres; pacientes viúvos ($e^{\beta} = 0,577$) tem maior chance de não comparecer às consultas do que pacientes que estejam casados; pacientes adultos ($e^{\beta} = 0,67$) tem menor chance de comparecer às consultas do que as crianças, pelo contrário, pacientes que tenham mais de 80 anos ($e^{\beta} = 1,713$) tem maior chance de comparecer do que os crianças; utentes que tenham recorrido a consultas de medicina física e reabilitação nos meses de abril ($e^{\beta} = 0,655$), julho ($e^{\beta} = 0,535$), agosto ($e^{\beta} = 0,42$), setembro ($e^{\beta} = 0,634$) e novembro ($e^{\beta} = 0,718$) tem menor chance de não comparecer às consultas do que pacientes que recorrem este serviço no mês de janeiro; o tempo de espera ($e^{\beta} = 0,999$) e o número não de comparências anteriores ($e^{\beta} = 0,979$) tem um efeito significativo na chance de não comparecer às consultas antes da pandemia; De outro modo, quanto maior o número de consultas anteriores ($e^{\beta} = 1,008$) maior a chance de os pacientes comparecerem às consultas marcadas de medicina física e reabilitação.

Durante a pandemia de covid-19, pacientes que se encontram divorciados ($e^{\beta} = 0,253$) tem maior chance de não comparecer às consultas de medicina física e reabilitação do que pacientes casados; relativamente à idade, jovens adultos ($e^{\beta} = 2,84$) e de meia-idade ($e^{\beta} = 3,041$) tem maior chance de comparecer às consultas durante a pandemia do que utentes que sejam crianças; já o número de não comparências anteriores ($e^{\beta} = 0,991$) e o tempo de espera ($e^{\beta} = 0,9989$), estão significativamente associados ao não comparecimento dos pacientes às consultas de medicina física e reabilitação durante período de covid-19, isto significa que quanto maior o número de não comparências anteriores e o tempo de espera, menor a chance dos utentes não faltarem às consultas marcadas; pelo contrário, quanto maior a distância ($e^{\beta} = 1,123$) (que os pacientes tem de percorrer desde as suas habitações até ao hospital) e o número de consultas anteriores ($e^{\beta} = 1,012$) menor a chance de os utentes faltarem às consultas de medicina física e reabilitação; No que diz respeito a fatores de calendarização, pacientes que tenham tido consultas marcadas nos meses de

maio ($e^\beta = 3,871$), junho ($e^\beta = 3,85$), agosto ($e^\beta = 2,012$), outubro ($e^\beta = 2,34$), e novembro ($e^\beta = 2,058$) tem maior possibilidade de não faltar às consultas do que os pacientes que tiveram consultas marcadas no mês de abril; quanto à meteorologia, a velocidade do vento influencia na chance de os pacientes faltarem ou não às consultas, ou seja, quanto maior a velocidade do vento ($e^\beta = 1,022$) maior a chance dos pacientes comparecerem às consultas marcadas de medicina física e reabilitação durante a pandemia.

Medicina Física e Reabilitação (pré-pandemia)				Medicina Física e Reabilitação (durante a pandemia)			
Variável	e^β	IC95%	p-value	Variável	e^β	IC95%	p-value
Gênero (referência: feminino)				Estado civil (referência: casado)			
Masculino	1,239	[1,079;1,424]	0,002	Divorciado	0,253	[0,164;0,392]	<,001
Estado civil (referência: casado)				Solteiro	0,731	[0,343;1,555]	0,415
Divorciado	0,665	[0,395;1,117]	0,123	União de facto	6,20E+08	[0;.]	1
Solteiro	0,878	[0,622;1,239]	0,459	Viúvo	2,20E+08	[0;.]	0,999
União de facto	0,749	[0,511;1,096]	0,137	Idade (referência: criança)			
Viúvo	0,577	[0,396;0,841]	0,004	Adolescente	1,762	[0,927;3,348]	0,084
Idade (referência: criança)				Jovem adulto	2,84	[1,405;5,743]	0,004
Adolescente	0,91	[0,674;1,228]	0,538	Adulto	1,357	[0,525;3,508]	0,528
Jovem adulto	0,927	[0,713;1,205]	0,572	Meia-idade	3,041	[1,107;8,351]	0,031
Adulto	0,67	[0,46;0,977]	0,037	Idoso	2,379	[0,874;6,48]	0,09
Meia-idade	0,851	[0,564;1,285]	0,444	Mais de 80	3,109	[0,88;10,989]	0,078
Idoso	1,183	[0,779;1,796]	0,431	Não comparências anteriores	0,991	[0,987;0,995]	<,001
Mais de 80	1,713	[1,077;2,724]	0,023	Distância	1,123	[1,06;1,189]	<,001
Não comparências anteriores	0,979	[0,975;0,983]	<,001	Nº consultas anteriores	1,012	[1,006;1,017]	<,001
Nº consultas anteriores	1,008	[1,004;1,012]	<,001	Mês (referência: abril)			
Mês (referência: janeiro)				Maio	3,871	[1,886;7,948]	<,001
Fevereiro	0,85	[0,647;1,117]	0,244	Junho	3,85	[1,907;7,775]	<,001
Março	0,924	[0,701;1,219]	0,576	Julho	1,881	[0,966;3,663]	0,063
Abril	0,655	[0,477;0,9]	0,009	Agosto	2,012	[1,024;3,956]	0,043
Maio	0,78	[0,563;1,082]	0,137	Setembro	1,928	[0,987;3,763]	0,055
Junho	1,008	[0,734;1,384]	0,961	Outubro	2,34	[1,188;4,61]	0,014
Julho	0,535	[0,41;0,699]	<,001	Novembro	2,058	[1,02;4,15]	0,044
Agosto	0,42	[0,314;0,561]	<,001	Dezembro	0,781	[0,217;2,808]	0,705
Setembro	0,634	[0,475;0,845]	0,002	Velocidade do vento	1,022	[1,003;1,042]	0,022
Outubro	0,927	[0,695;1,237]	0,607	Tempo de espera	0,9989	[0,9985;0,9993]	<,001
Novembro	0,718	[0,558;0,922]	0,009	Teste de Omnibus	Qui-quadrado	gl	p-valor
Dezembro	0,87	[0,668;1,134]	0,303		272,517	23	<,001
Tempo de espera	0,999	[0,998;0,999]	<,001	R2	Verossimilhança de log -2	R quadrado Cox & Snell	R quadrado Nagelkerke
Teste de Omnibus	Qui-quadrado	gl	p-valor		2219,143	0,135	0,184
	625,46	25	<,001	Teste de Hosmer e Lemeshow	Qui-quadrado	gl	p-valor

R2	Verossimilhança de log -2	R quadrado Cox & Snell	R quadrado Nagelkerke	4,627	8	0,797
	6905,156	0,101	0,14			
Teste de Hosmer e Lemeshow	Qui-quadrado	gl	p-valor			
	11,462	8	0,177			

Tabela 18 - Modelo logístico multivariado de medicina física e reabilitação.

4.3 – Diagnóstico dos modelos

4.3.1 – Teste de Omnibus

Observando os dados reportados pelo teste de Omnibus, é de concluir que para todos modelos de regressão logística de todas as especialidades os testes de Omnibus apresentam uma significância de $< 0,01$, rejeitando-se a hipótese nula de que os efeitos estimados dos fatores considerados sejam nulos

Posto isto, para todos os modelos de regressão logística para as especialidades de ortopedia, oftalmologia, cardiologia, medicina dentária e medicina física e reabilitação as variáveis consideradas permitem explicar o risco de absentismo dos pacientes face às consultas hospitalares (antes e durante a pandemia) no hospital do Norte de Portugal.

4.3.2 – Teste de Hosmer e Lemeshow

Nas tabelas dos modelos logísticos mencionados acima, estão apresentados os resultados obtidos nos testes de Hosmer e Lemeshow.

Dos resultados obtidos conclui-se então que não existe evidência para se puder rejeitar a hipótese nula nos modelos de medicina dentária (antes e durante a pandemia), medicina física e reabilitação (antes e durante a pandemia) e no modelo de oftalmologia antes da pandemia, uma vez que os valores de prova são superiores ao nível de significância de 0,05, logo não se rejeita a hipótese nula, estando assim, os modelos bem ajustados aos dados.

Pelo contrário, para os modelos de cardiologia (antes e durante a pandemia), oftalmologia (durante a pandemia) e ortopedia (antes e durante pandemia), os valores

de prova $< 0,05$, questionando a validade do modelo. No entanto, como alguns autores apontam quando o teste de ajuste de Hosmer-Lemeshow é realizado com vários padrões de covariáveis inferiores ao número de sujeitos (como neste caso), o seu resultado pode ser impreciso (Bertolini *et al.* 2000). Como Kramer e Zimmerman (2007) discutem, o teste de Hosmer-Lemeshow é sensível ao tamanho da amostra, e um teste de Hosmer-Lemeshow significativo não significa necessariamente que um modelo preditivo não seja útil ou suspeito.

4.3.3 – Curva de ROC

Analisando os gráficos das curvas de ROC entre as cinco especialidades durante o período de pandemia (ver gráficos 1, 2, 3, 4 e 5), percebe-se que os valores das especialidades ortopedia e cardiologia são muito semelhantes aos valores de todas as especialidades na pré-pandemia, ou seja, rondam entre os 0,6 e os 0,7, o que nos permite concluir que o modelo logístico binário tem uma boa discriminação apesar de limitada. Pelo contrário, observa-se que as especialidades oftalmologia, medicina dentária e medicina física e reabilitação apresentam valores de ROC entre 0,7 e 0,8, o que é perceptível que estes modelos são considerados modelos com uma discriminação aceitável (ver tabela 19).

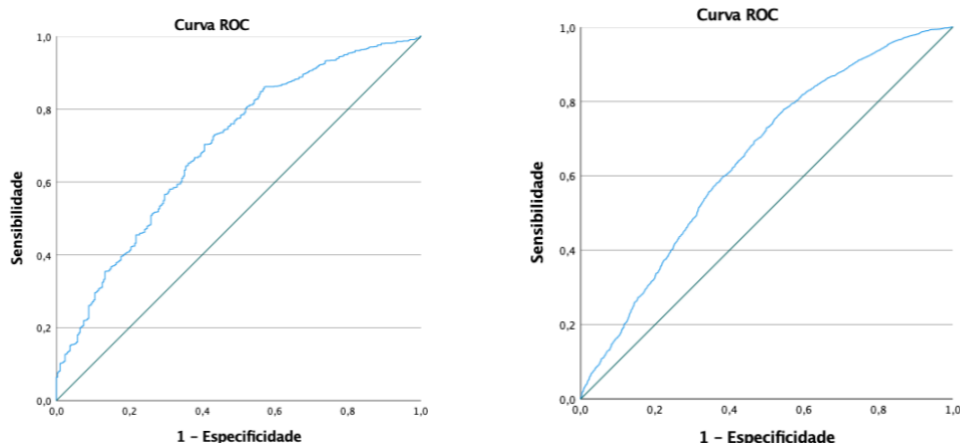


Gráfico 1- antes e durante a pandemia – ortopedia

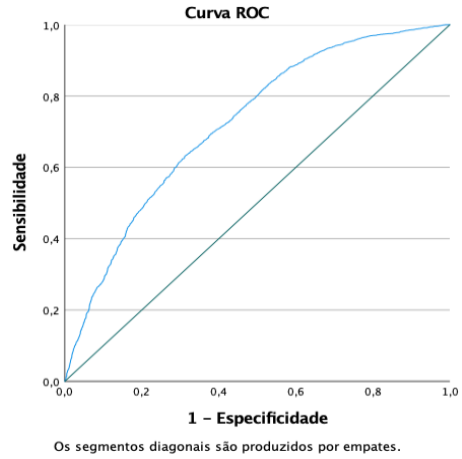
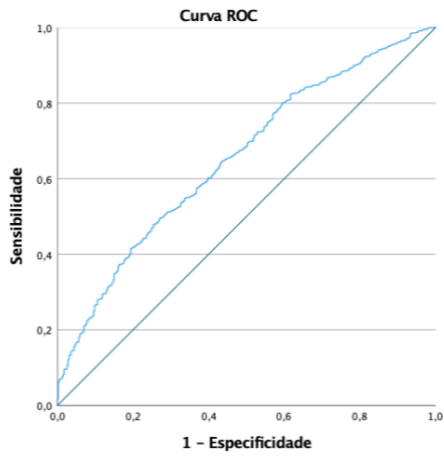


Gráfico 2 - antes e durante a pandemia – oftalmologia

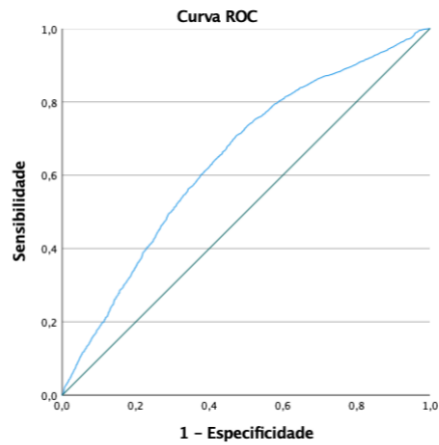
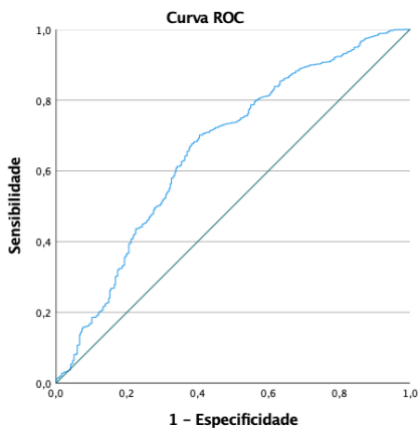


Gráfico 3 - antes e durante a pandemia – cardiologia

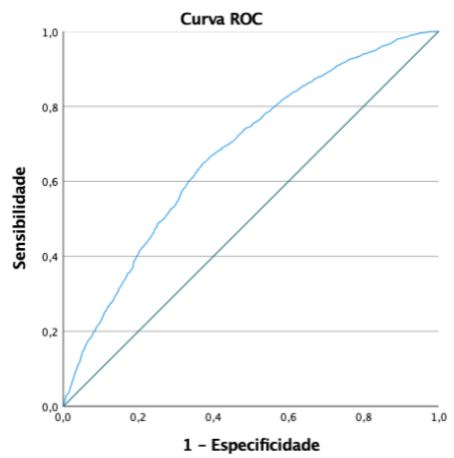
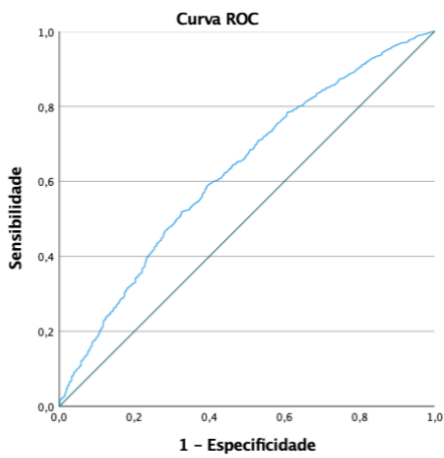


Gráfico 4 - antes e durante a pandemia - medicina dentária

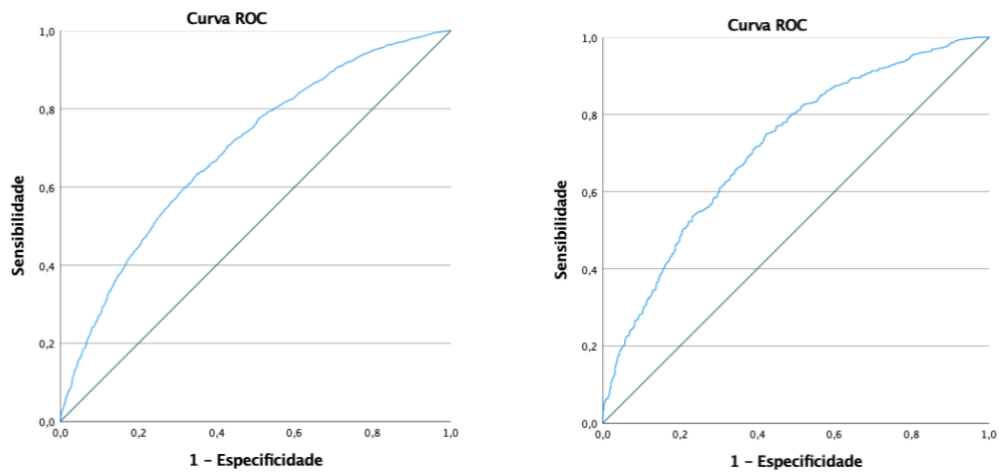


Gráfico 5 - antes e durante a pandemia - medicina física e reabilitação

Especialidade	Momento	Área	Erro	Sig assintótico	IC95%
Ortopedia	Pré-Pandemia	0,695	0,017	0,000	[0,661;0,729]
	Pandemia	0,648	0,008	0,000	[0,633;0,662]
Oftalmologia	Pré-Pandemia	0,654	0,015	0,000	[0,625;0,683]
	Pandemia	0,722	0,006	0,000	[0,709;0,734]
Cardiologia	Pré-Pandemia	0,661	0,018	0,000	[0,626;0,696]
	Pandemia	0,644	0,006	0,000	[0,631;0,656]
Medicina Dentária	Pré-Pandemia	0,626	0,011	0,000	[0,604;0,648]
	Pandemia	0,676	0,006	0,000	[0,664;0,689]
Medicina Física e Reabilitação	Pré-Pandemia	0,693	0,007	0,000	[0,679;0,707]
	Pandemia	0,715	0,012	0,000	[0,691;0,739]

Tabela 19 - Áreas abaixo das curvas de ROC das especialidades analisadas.

Discussão e Conclusão

O estudo teve como objetivo primordial determinar os fatores de risco que influenciam os pacientes a não comparecer às consultas hospitalares, utilizando o modelo de regressão logístico, o modelo mais eficiente e eficaz de acordo com a revisão de literatura. Adicionalmente, após a identificação dos fatores de riscos e quais os modelos a utilizar, fez-se uma comparação antes e durante o tempo de pandemia COVID-19.

Foram vários os fatores de risco identificados na literatura associados ao absentismo de pacientes face a consultas hospitalares, estando estes fatores divididos em 4 grupos, tais como: (i) fatores demográficos; (ii) histórico médico; (iii) detalhes de agendamento; (iv) comportamento do paciente. Neste estudo apenas foram considerados os seguintes: idade, género, estado civil, distância do hospital até à residência do paciente, número de consultas anteriores, primeira consulta ou consulta de acompanhamento, mês da consulta, dia da semana da consulta, precipitação, velocidade do vento, tempo de espera, intervalo entre consultas, número de não comparências anteriores e por último, a estação do ano.

No que diz respeito aos modelos referenciados na literatura, foram identificados os seguintes: modelo de regressão logístico, redes neurais, algoritmo de classificação de Naïve Bayes, árvores de decisão, Bayesiano e o Random Forest. Uma vez que o modelo de regressão logístico foi o modelo que está referenciado na literatura como o mais eficiente e eficaz, este foi o modelo que foi utilizado neste estudo.

No que diz respeito à análise das especialidades analisadas neste estudo (ortopedia, oftalmologia, medicina física e reabilitação, medicina dentária e cardiologia), antes da pandemia os fatores de risco comuns a todas as especialidades foram: o tempo de espera e o número de não comparências anteriores. No entanto, o estado civil, a idade e o número de consultas anteriores também foram variáveis que se destacaram como importantes fatores de risco durante o período de pré-pandemia. Analogamente, durante o período pandémico, os fatores de risco comuns a todas as especialidades foram o tempo de espera e o número de não comparências anteriores. Contrariamente ao que se registou no período de pré-pandemia, o mês da consulta e o número de

consultas anteriormente administradas tiveram um forte impacto no risco durante o período de pandemia COVID-19.

Através dos fatores de risco identificados e utilizando o modelo de regressão logístico, é possível identificar que existem perfis de pacientes diferentes por especialidade. Assim, entende-se que o hospital deverá considerar um modelo distinto por cada tipo de especialidade.

As causas que levam ao absentismo hospitalar deverão ser estudadas minuciosamente e de forma individual em cada especialidade, isto porque em cada especialidade podem existir fatores específicos, como por exemplo, a gravidade da doença, a disponibilidade dos profissionais de saúde e de equipamentos.

Foram identificados diferentes perfis dos pacientes que faltam às consultas antes e durante a pandemia, o que se poderá concluir que o vírus COVID-19 teve um impacto no que diz respeito ao absentismo dos pacientes às consultas. O que poderá querer dizer que houve uma mudança de comportamento por parte dos pacientes.

Tendo em conta os resultados obtidos, o hospital poderá tomar diversas medidas para mitigar as taxas de absentismo. Na literatura, vários autores chegaram a diversas soluções como por exemplo, a criação de um sistema de agendamento de consultas que permita realizar o agendamento de consultas de forma rápida e sem qualquer tipo de complicações para os pacientes (Batoool *et al.* (2021) e Bush *et al.*, (2016); contactar com o paciente algum tempo antes da consulta, seja ele avisado através de um funcionário da clínica ou está mesmo por um sistema automatizado (Junod-Perron, *et al.*, 2010) e (Levy, 2013); criação de um *overbooking* que permita preencher os horários de agendamento e manter de forma gradual os tempos de espera mais curtos (Alaeddini *et al.*, 2011); elaboração de sistema integrados que consistam na comunicação entre as diversas unidades de saúde, de maneira a evitar que os pacientes agendem uma consulta para mais do que um serviço - a não ser que os objetivos sejam diferentes - evitando a colisão entre horários e datas (Junod-Perron *et al.*, 2010); diminuição das listas de espera respetivas às primeiras consultas, de maneira a determinar a importância e rapidez na marcação das consultas seguintes (Batoool *et al.* 2021).

Concluindo, a diminuição das taxas de absentismo é crucial para uma gestão eficiente das unidades hospitalares. Este desafio torna-se ainda mais importante num contexto pandémico, onde é necessário alocar recursos de forma rápida e eficiente. Este

trabalho visou abordar este problema propondo um modelo preditivo dos fatores de risco de absentismo dos pacientes a consultas pré-agendadas, e fornecer informações sobre como a pandemia COVID-19 afetou o comportamento dos doentes no que diz respeito à não-comparência.

Como trabalho futuro, pretende-se testar outras técnicas tais como *clustering* e árvores de decisão e incorporar um modelo de explicabilidade. Adicionalmente, também se propõe testar possíveis fatores de risco, tal como, por exemplo, o grau de alfabetização dos pacientes, de modo que hospital possa, a título de exemplo, criar técnicas de comunicação adequadas a cada tipo de público-alvo.

Bibliografia

- (2022). Obtido em 29 de maio de 2022, de Stringfixer:
https://stringfixer.com/pt/Omnibus_test
- Ahmadi, E., Garcia-Arce, A., Masel, D., Reich, E., Puckey, J., & Maff, R. (2019). A metaheuristic-based stacking model for predicting the risk of patient no-show and late cancellation for neurology appointments. *IJSE Transactions on Health care Systems Engineering*, 9 (3), 1-31.
- Aladeemy, M., Adwan, L., Booth, A., Khasawneh, M., & Poranki, S. (2020). New feature selection methods based on opposition-based learning and self-adaptive cohort intelligence for predicting patient no-shows. *Applied Soft Computing Journal*, 86.
- Alaeddini, A., Yang, K., & Reddy, C. (June de 2011). A probabilistic model for predicting the probability of no-show in hospital appointments. *Health Care Management Science*, 14 (2), 146-157.
- Batool, T., Abuelnoor, M., Aloul, F., Sagahyroon, A., & El Boutari, O. (2021). Predicting Hospital No-Shows Using Machine Learning. *2020 IEEE Internaciona Conference on Internet of Things and Intelligent System (IoTAIS)*.
- Berger, Z., Evans, N., Phelan, A., & Silverman, R. (2020). *Covid-19: control measures must be equitable and inclusive*. Institute of Bioethics, School of Medicine.
- Bertolini, G., D'Amico, R., Tinazzi, A., & Apolone, G. (2000). One model, several results: the paradox of the Hosmer-Lemeshow goodness-of-fit test for the logistic regression model. *5(4)*, 251-253.
- Bittar, O., Magalhães, A., Martines, C., Felizola, N., & Falcão, L. (2016). Absenteeism in outpatient care specialties in São Paulo. *Secretaria de Estado da Saúde, São Paulo*, 13 (152), 19-32.
- Bush, R., Vemulakonda, V., Corbett, S., & Chiang, G. (05 de December de 2016). Can we predict a national profile of non-attendance pediatric urology patients: a multi-institucional electronic health record study. *HHS Public Access*, 21(3), 132-138.
- Carreras-García, D., Delgado-Gomez, D., & Arribas-Gil, A. (2020). Patient No-Show Prediction: A Systematic Literature Review. *Entropy*, 22(6).
- Dantas, L., Hamacher, S., Cyrino, F., Barbosa, S., & Viegas, F. (2018). Predicting Patient No-Show Behavior: a study in a bariatric clinic. *Obesity Surgery*.
- Deng, S., & Peng, H. (2020). *Characteristics of and Public Health Responses to the Coronavirus Disease 2019 Outbreak in China*. Journal of Clinical Medicine, Escola de Saúde Pública, Departamento de Biologia Patológica.
- Dewey, C., Hingle, S., Goolez, E., & Linzer, M. (2020). Supporting Clinicians During the Covid-19 Pandemic. *Annals of Internal Medicine*, 25(8), 924-930.
- Ding, X., Gellad, Z., Mather, C., Barth, P., Poon, E., Newman, M., & Goldstein, B. (2018). Designing risk prediction models for ambulatory no-shows across different specialties and clinics. *Journal of the American Medical Informatics Association*, 1-8.
- Drewek, R., Mirea, L., & Adelson, D. (2017). Lead time to appointment and no-show rates for new and follow-up patients in an ambulatory clinic. *Wolters Kluwer Health*, 36(1), 4-9.

- Goffman, R., Harris, S., May, J., Milicevic, A., Monte, R., Myaskovsky, L., . . . Vargas, D. (2017). Modeling patient no-show history and predicting future outpatient appointment behavior in the veterans health administration. *Military Medicine*, 182(5), e1708-e1714.
- Gonçalves, A. (2013). *Regressão Logística Aplicada à Pesquisa de Preditores de Morte*. Dissertação de Mestrado, Faculdade de Ciências e Tecnologia: Universidade de Coimbra , Departamento de Matemática.
- Gong, K., Xu, Z., Cai, Z., Chen, Y., & Wang, Z. (2020). *Internet Hospitals Help Prevent and Control the Epidemic of Covid-19 in China: Multicenter User Profiling Study*. *Journal of Medical Internet Research* , Xia'men University, China.
- Harvey, H., Liu, C., Ai, J., Jaworsky, C., Guerrier, C., Flores, E., & Pianykh, O. (2017). Predicting No-Shows in Radiology using regression modeling of data available in the electronic medical record. *Journal of the American College of Radiology*, 14 (10), 1303-1309.
- Hilbe, J. (2009). *Logistic regression models. Chapman and Hall/CRC*.
- Hosmer, D., & Lemeshow, S. (2000). *Applied Logistic Regression*. Wiley.
- Hosmer, D., Lemeshow, S., & Sturdivant, R. (2013). *Applied Logistic Regression* (Vol. Vol 3rd).
- Hu, M., Xu, X., Li, X., & Che, T. (2020). Managing patients' no-show behaviour to improve the sustainability of hospital appointment systems: Exploring the conscious and unconscious determinants of no-show behaviour. *Journal of cleaner production*, 269.
- Huang, Y., & Hanauer, D. (2016). Time dependent patient no-show predictive modelling development. *International Journal of Health Care Quality Assurance*, 29 (4), 475-488.
- Junod Perron, N., Dominice Dao, M., Kossovsky, M., Miserez , V., Chuard, C., Calmy, A., & Gaspoz, J. (2010). Reduction of missed appointments at an urban primary care clinic: a randomised controlled study. *BMC Family Practise*, 11(1).
- Kalb, L., Freedman, B., Foster, C., Menon, D., Landa, R., Kishfy, L., & Law, P. (2012). Determinants of Appointment Absenteeism at an Outpatient Pediatric Autism Clinic. *Journal of Development and Behavioral Pediatrics*, 33 (9), 685/697.
- Kramer, A., & Zimmerman, J. (2007). Assessing the calibration of mortality benchmarks in critical care: The Hosmer-Lemeshow test revisited. *Journal of Critical care medicine*, 35(9), 2052-2056.
- Lee, G., Wang, S., Dipuro, F., Hou, J., Grover, P., & Low, L. (2017). Leveraging on Predictive Analytics to Manage Clinic No-Show and Improve Accessibility of Care. *International Conference on Data Science and Advanced Analytics*.
- Levy, V. (October de 2013). A predictive tool for nonattendance at a specialty clinic: An application of multivariate probabilistic big data analytics. *10th Internaciona Conference and Expo on Emerging Technologies for a Smarter World (CEWIT)*.
- Mahajan, T., Singh, G., & Bruns, G. (2021). *An experimental Assessment of Treatment for Cyclical Data* .
- Mohammadi, I., Wu, H., Turkcan, A., Toscos, T., & Doebbeling, B. (2018). Data Analytics and Modeling for Appointment No-show in community health centers. *Journal of Primary Care and Community Health*, 9.
- Nakamura, K. (2013). *Multicolinearidade em Modelos de Regressão Logística*. Dissertação de Mestrado, Universidade de São Paulo , Instituto De Matemática e Estatística, São Paulo, Brasil .

- Neal, R., Neal, R., Lawlor, D., Allgar, V., Colledge, M., Ali, S., . . . Wilson, A. (October de 2001). Missed appointments in general practice: retrospective data analysis from four practices. *British Journal of General Practice*.
- Norris, J., Kumar, C., Chand, S., Moskowitz, H., Shade, S., & Willis, D. (2014). An empirical investigation into factors affecting patient cancellations and no-shows at outpatient clinic. *Decision Support Systems* , 57, 428-443.
- Perez, F., Xie, J., Sin, A., Tsai, R., Sanders, L., Cox, K., . . . Haberland, K. (2014). Characteristics and Direct Costs of Academic Pediatric Subspecialty Outpatient No-Show Events. *Journal for Healthcare Quality*, 36, 32-42.
- Praveena, M., Krupa, J., & SaiPreethi, S. (2019). Statistical Analysis of Medical Appointments using decision tree. *Fifthy Internacional conference on Science Technology Engineering and Mathematics*.
- Salazar, L., Fernandes, A., Dazzi, R., Garcia, N., & Leithardt, V. (2020). Using Different Models of Machine Learning to Predict Attendance at Medical Appointments. *Journal of Information Systems Engineering and Management*, 5 (4).
- Simsek, S., Dag, A., Tiaht, T., & Oztekin, A. (2020). A Bayesian Belief Network-based probabilistic mechanism to determine patient no-show risk categories. *Omega*.
- Topuz, K., Uner, H., Oztekin, A., & Yildirim, M. (2017). Predicting pediatric clinic no-shows: a decision analytic framework using elastic net and Bayesian belief network. *Annals of Operations Research*, 263 (1-2), 479-499.
- Vaz, F. (2020). *Modelação Linear e Extensões: Aplicação da Regressão Logística no Estudo de Câncer da Mama*. Mestrado em Estatística, Matemática e Computação , Universidade Aberta, Estatística Computacional .
- Wahid, Z., Satter, A., Al Imran, A., & Bhuiyan, T. (2019). Predicting Absenteeism at Work Using Tree-Based Learners. *Proceedings of the 3rd Internacional Conference on Machine Learning and Soft Computing*.
- World Health Organization. (2020). Obtido em Abril de 2022, de Coronavirus disease 2019 (COVID-19) Situation Reports: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>