



ANÁLISE DE DADOS E BIG DATA - PROJETO BOSTON SCIENTIFIC

O composto Topiramato dentro do cenário de venda de medicamentos do Brasil

Gabriel Gama de Mouras
Gustavo Fernando Pupo de Moraes
Lucas Augustho da Silva Candido
Matheus Costa Amorim
Vinicios Silva Santos
Gabriel Gomes Apolinário
Vinicius Scaramello Rodrigues Neres
Email para contato(gabrielmouras.5044@aluno.saojudas.br)

Professor Orientador- Evandro Ferraz

RESUMO

A elaboração do projeto em questão surge através da necessidade de fazer uma análise crítica acerca de dados fornecidos pela ANVISA a respeito da venda de medicamentos controlados no Brasil, fato que flui diretamente com o que está sendo estudado ao longo do semestre acerca da análise de dados. Sendo assim, visando tornar a análise dos dados mais assertiva, optamos pela utilização de uma divisão da base de dados em questão que traz informações da venda do composto conhecido como Topiramato, no período de 2020 até o final de 2021. Nesse viés, utilizamos a metodologia conhecida como análise quantitativa para que fosse possível traduzir os nossos dados brutos em informações que sejam mais fáceis de serem visualizadas. Outrossim, ao iniciar as pesquisas relacionados à temática, foi estabelecido pelo grupo que a teoria de hipótese era a de que o consumo do Topiramato seria maior para o público masculino durante a infância e adolescência já que essa faixa etária condiz com o período da vida em que mais indivíduos sofrem com convulsões e crises de epilepsia. Finalmente, visando obter uma forma de mapear as vendas do composto em questão e de retorno de clientes, estabelecemos como proposta de rastreabilidade uma quantidade farmacotécnica média do remédio através de cálculos na linguagem de programação R, por meio da qual atribuímos outra coluna dentro do dataset e que iria comparar o valor médio com a quantidade que foi comprada e caso esse valor seja maior que o estabelecido, teríamos a ciência de que aquele consumidor provavelmente irá demorar um período maior para retornar a consumir e será necessário captar outros clientes para suprir as demandas de venda.

Palavras-chave: Topiramato. Medicamentos. Big Data.

1 INTRODUÇÃO

Inicialmente, a temática a ser explorada ao longo do projeto é a análise de dados brutos acerca da venda do composto medicamentoso conhecido como Topiramato (LIN,2011), durante o período dos anos de 2020 até o final de 2021. Ainda deve-se dizer que todas as informações que serão utilizadas ao longo da realização das análises concordam com a base de dados fornecida pela ANVISA acerca da venda de medicamentos controlados no país, que conta com diversos compostos ao longo de um período um pouco maior do que o que foi estabelecido. Sendo assim, o grupo optou por fazer uma análise quantitativa de um recorte de dados da base em questão, que corresponde somente às vendas do Topiramato, para que seja possível tornar a investigação mais concreta e confiável.

Além disso, é importante considerar que, no início das pesquisas realizadas, o grupo não detinha nenhum conhecimento sobre as vendas do medicamento no país. No entanto, o estudo e a análise de dados realizados neste trabalho possibilitaram a obtenção de informações relevantes que podem ser utilizadas como base pelos próximos pesquisadores que vierem a estudar o assunto Análise de Dados e Big Data (TAURION,2013). Do mesmo modo, o grupo se motivou a buscar mais conhecimento sobre o tema desenvolvido e se inteirar de maneira mais profunda para possíveis discussões que possam surgir.

Finalmente, para tornar o trabalho mais eficiente, foi estabelecido como principal objetivo do projeto uma proposta de rastreabilidade que tenha como função, mapear todas as vendas do Topiramato e definir através de cálculos matemáticos a média das quantidades que estão sendo compradas com o intuito de definir se a compra em questão se refere a um possível cliente que irá retornar em breve para comprar, ou seja, caso a compra seja maior que o valor estipulado como média, é bastante provável que o cliente não retorne a comprar em um período curto de tempo, sendo assim será necessário que a empresa realize a captação de outros possíveis consumidores (FREITAS,2002).

2 METODOLOGIA

Primeiramente, ao descrever os métodos utilizados para a obtenção dos dados e das fontes de nossa pesquisa, torna-se essencial destacar que no início do projeto foram sugeridas algumas bases de dados que melhor se encaixam na temática do projeto que será alinhado com a empresa Boston Scientific, sendo assim por se tratar de uma empresa que tem como ideal a transformação da vida de indivíduos através da inovação em soluções da medicina que melhoram a qualidade de vida de seus pacientes, o grupo optou por utilizar uma base de dados fornecida pelo órgão regulamentador do Brasil acerca da distribuição de medicamentos no país.

A pesquisa em questão e tratamento dos dados corresponde a metodologia conhecida como Análise Quantitativa (LAY,2005), através de cálculos matemáticos e estabelecimento de métricas, o grupo pretende obter estatísticas, informações numéricas e percentuais acerca da venda de Topiramato no país. Do mesmo modo, o principal objetivo que temos com a execução do projeto é tornar possível rastrear quais são as vendas do Topiramato que podem proporcionar a fidelização dos indivíduos que consomem o medicamento, no entanto, a análise de tal possibilidade será traduzida através de cálculos matemáticos na linguagem de programação R.

Seguidamente, a base de dados do Topiramato conta com informações como: princípio ativo, ano de venda, sexo do paciente, idade, estado e município da venda, quantidade farmacotécnica (doses e quantidades compradas). Essas informações possibilitam aumentar a rastreabilidade de cada medicamento vendido, tornando possível determinar o consumo do medicamento a partir de um parâmetro específico, como uma faixa etária, por exemplo.

Sendo assim, é importante destacar que a base de dados que o grupo optou por utilizar conta com uma amostra de cerca de 70 mil vendas relacionadas ao composto descrito acima em um recorte de tempo de dois anos completos (2020-2021) e que conta com as vendas em todo o território nacional, possibilitando

a distribuição das estatísticas de forma equivalente e o mais fiel que conseguimos chegar de forma amadora, já que estamos utilizando uma base de dados disponível na 'internet' que seja de cunho confiável, no entanto, não fizemos essa pesquisa de campo e não conseguimos assegurar que este seja o número exato.

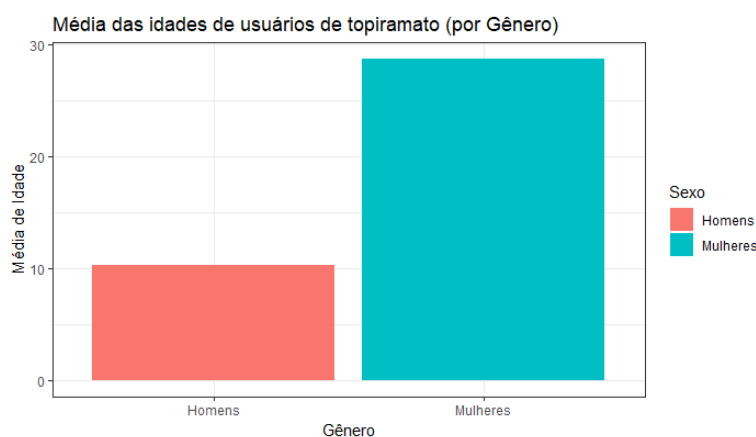
Finalmente, ao tomar ciência que o composto medicamentoso em questão é utilizado geralmente para o tratamento de crises de epilepsia (CORREDEIRA,2003), convulsão e na profilaxia de enxaqueca (LAINETTI, 2019), o grupo estabeleceu como hipótese que o consumo de tal medicamento é maior para o sexo masculino durante a infância e adolescência considerando questões de desenvolvimento genético e hormonal que são influenciados no corpo masculino já que durante as nossas pesquisas foi constatado que tais faixas etárias costumam ter maior ocorrência de crises de epilepsia e convulsão. Já uma segunda hipótese era de que a média de idade dos homens que utilizam o medicamento seria menor que a idade média das mulheres, sendo assim, nossa análise de dados terá como objetivo testar estas hipóteses.

3 RESULTADOS

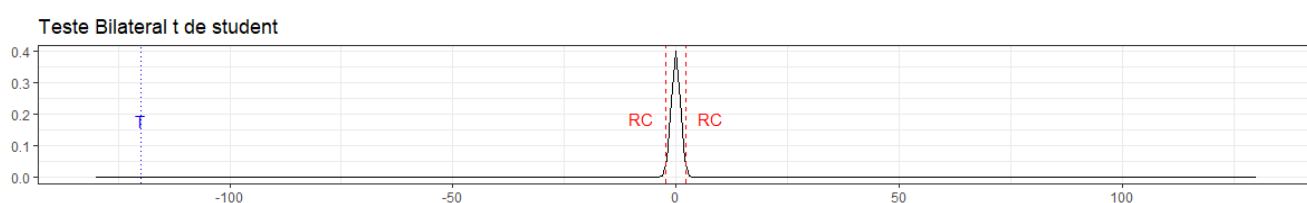
Inicialmente, deve-se dizer que o processo de testes de hipótese e de utilização dos cálculos e fórmulas em R, necessitou de grande aprendizado por parte de todos os membros da equipe, já que ao longo do semestre todos os casos e alternativas relacionadas com os testes que foram desenvolvidos foram ensinados durante as aulas.

Ainda, é fundamental salientar que durante a criação das nossas hipóteses o grupo quis possibilitar uma forma de utilizar todos os recursos que já estavam disponíveis dentro do nosso dataset e investigar os mesmos dados de maneira mais aprofundada, sendo assim as hipóteses que foram testadas culminaram em respostas positivas e negativas. Já que a suposição acerca do consumo de tal composto medicamentoso seria maior durante a infância e adolescência para o público masculino, e a devolutiva que tivemos das análises confirma que essa afirmação de hipótese que o grupo sustentou era de fato verdadeira.

Do mesmo modo, a outra hipótese elaborada pelo grupo que diz que a idade média dos homens que utilizam o medicamento seria menor que a média feminina se comprova como verdade também, já que segundo os cálculos e fórmulas que utilizamos, a média de idade do sexo masculino que utilizam o medicamento de acordo com os dados do dataset é de 10 anos e 4 meses, e a idade média das mulheres que utilizam o medicamento conhecido como Topiramato é de cerca de 28 anos de idade.



Finalmente, é importante ter como base que levando em consideração toda a base de dados utilizada , cerca de 45 mil amostras correspondem ao sexo feminino e o restante que corresponde à cerca de 32 mil amostras correspondem ao público masculino. Sendo assim, tamanha quantidade de linhas de informação permite que as nossas pesquisas e conclusões sejam o mais verídicas possíveis. De acordo com os cálculos realizados, a confiabilidade dos resultados é de 95%, podendo apresentar um desvio padrão dos dados, como por exemplo a idade média dos gêneros.



4 CONCLUSÃO

Ademais, deve-se dizer que o trabalho possui alguns objetivos principais, sendo estes: utilizar a base de dados fornecida para propor um modelo de estudos para a Análise de dados e Big Data, o segundo objetivo corresponde a uma forma de auxiliar a empresa parceira Boston Scientific (que atua na área de inovações do ramo da medicina) a analisar possibilidades de estudo acerca da utilização do Topiramato para o combate de doenças que atacam o sistema nervoso e motor dos indivíduos da sociedade brasileira, como, por exemplo as crises de convulsão e epilepsia no país, afetando em sua grande maioria o público infantil e adolescente.

Além disso, as proposições de soluções de rastreabilidade criadas pelo grupo visam auxiliar no controle mais rígido das informações de venda do composto Topiramato, levando em consideração que a proposta tem como objetivo estabelecer uma quantidade de venda média e detectar compras que sejam consideradas superiores ao valor em questão, fato que provavelmente traduz uma próxima compra muito mais demoradas por parte de tal cliente, uma vez que, as doses farmacotécnicas médias liberadas pela Anvisa correspondem a 25 mg ao dia do medicamento, sendo assim ao realizar uma compra com grande quantidade farmacotécnica possivelmente vai durar por mais alguns meses de tratamento e será necessário captar outros clientes para suprir as necessidades de venda.

Finalmente, as hipóteses que foram testadas nos trouxeram como devolutiva, afirmações verdadeiras e que possibilitaram os testes e estudos que eram necessários para comprovar as nossas proposições iniciais, ainda devemos ter como base que a nossa margem de estudos não condiz exatamente com toda a sociedade brasileira levando em consideração que a base de dados compreende apenas cerca de 70 mil indivíduos, no entanto, é de suma importância compreender que este recorte possibilita um estudo com devido embasamento para a realização dos testes.

Referência Bibliográfica

CORREDERA-GARCÍA, E. ET AL. TRATAMIENTO CON TOPIRAMATO EN PACIENTES CON EPILEPSIA REFRACTARIA. **REV NEUROL**, V. 37, P. 401-4, 2003

FREITAS, HENRIQUE; MOSCAROLA, JEAN. DA OBSERVAÇÃO À DECISÃO: MÉTODOS DE PESQUISA E DE ANÁLISE QUANTITATIVA E QUALITATIVA DE DADOS. **RAE ELETRÔNICA**, V.1, P. 1-30, 2002

LIN, KATIA. TOPIRAMATO: UMA MOLÉCULA MULTIFACETADA. **REVISTA NEUROCIÊNCIAS**, V. 19, N. 1, P. 8-9, 2011.

LAY, MARIA CRISTINA DIAS; REIS, ANTONIO TARCISIO DA LUZ. ANÁLISE QUANTITATIVA NA ÁREA DE ESTUDOS AMBIENTE-COMPORTAMENTO. **AMBIENTE CONSTRUÍDO: REVISTA DA ASSOCIAÇÃO NACIONAL DE TECNOLOGIA DO AMBIENTE CONSTRUÍDO**. VOL. 5, N. 2 (ABR./JUN. 2005), P. 21-36, 2005

”TAURION, CEZAR. **BIG DATA**. BRASPORT, 2013.”

VITOR LAINETTI; TÂNIA CARMEM PEÑARANDA GOVATO; GEHRKE, FLÁVIA; ET AL. USO DE TOPIRAMATO NA PROFILAXIA DA ENXAQUECA: REVISÃO DA LITERATURA. UNILUS ENSINO E PESQUISA, V.16, N. 43, P. 136-142, 2019. DISPONÍVEL EM:
<<http://revista.lusiada.br/index.php/ruep/article/view/1136/u2019v16n43e1136>>. Acesso em: 17 abr. 2023.