



CURSO DE MEDICINA

NAIELI MACHADO DE ANDRADE

**A FREQUÊNCIA DE DESENHOS NÃO EXPERIMENTAIS EM ESTUDOS DE
EFICÁCIA NA ÁREA DE CARDIOLOGIA: UM ESTUDO META-EPIDEMIOLÓGICO**

Salvador

2022

Naieli Machado de Andrade

**A FREQUÊNCIA DE DESENHOS NÃO EXPERIMENTAIS EM ESTUDOS DE
EFICÁCIA NA ÁREA DE CARDIOLOGIA: UM ESTUDO META-EPIDEMIOLÓGICO**

Trabalho de conclusão de curso apresentado como requisito parcial para aprovação no 4º ano do curso de medicina da Escola Bahiana de Medicina e Saúde Pública.

Orientador: Prof. Dr. Luís Cláudio L Correia

Salvador

2022

AGRADECIMENTOS

Consoante Antoine de Saint-Exupéry, em “O Pequeno Príncipe”: aqueles que passam por nós não vão sós, deixam um pouco de si e levam um pouco de nós. Dessa forma, a finalização desse trabalho é dedicada a todos que passaram por mim durante a minha trajetória. Primeiramente, agradeço a Deus por me dar sabedoria, força e me guiar em todos os aspectos da minha vida. Gostaria de agradecer, especialmente, a minha mãe, Nailza, que com todo esforço e sacrifício, me apoiou em todos os meus sonhos e me deu todo o suporte necessário para a realização deles, bem como introduziu desde cedo a importância de batalhar com perseverança para conseguir tudo aquilo que eu anseio. Meu padrasto, Mauto, que hoje cuida de mim lá de cima, e que assim como minha mãe, me ensinou valores importantes, como o amor e a benevolência, atributos que quero deixar sempre à frente no cuidado com meus pacientes. Além disso, não poderia deixar de agradecer-lo por me inspirar sobre a importância da realização da pesquisa científica com seriedade, transparência e empenho. Às minhas avós, Maria José e Severina, que estão me assistindo e cuidando de mim do plano espiritual. Aos meus tios, especialmente à minha tia Nádia, tia Eliana, Mário e tio Wellington, que cuidam de mim com muito amor. Aos meus primos, como Geize, que incentivou e despertou meu interesse para a medicina, à Leo, que estava comigo desde criança em todos os momentos e às minhas primas Camila e Thaís, que têm um espaço reservado em meu coração. Agradeço também a Henrique, meu namorado, pela paciência, ajuda, aconselhamento, por vibrar e ficar feliz com minha felicidade e minhas conquistas. A todos os meus amigos, por tornar essa trajetória mais leve, especialmente Miriam e Gabriel, os quais quero levar para o resto da vida, independente da distância física. Sou grata à minha psicóloga Luana Cavalcanti, que cuida da minha saúde mental e me dá suporte desde 2018, assim como a Dra. Patrícia Lemos, que me forneceu um tratamento humanizado e se tornou a minha maior fonte de inspiração médica. À minha professora de metodologia Profa. Alessandra Caldas, serei eternamente grata pela orientação, paciência e pelo estímulo com seu entusiasmo pelo meu projeto. Por fim, o mais genuíno agradecimento ao meu orientador Prof. Dr. Luís Cláudio Lemos Correia, sou muito feliz de ser sua orientanda e ter a oportunidade de aprender com você um ritual científico que requer humildade, dúvida e contém a leveza da imprevisibilidade.

“Some things benefit from shocks; they thrive and grow when exposed to volatility, randomness, disorder, and stressors and love adventure, risk, and uncertainty. Yet, in spite of the ubiquity of the phenomenon, there is no word for the exact opposite of fragile. Let us call it antifragile. Antifragility is beyond resilience or robustness. The resilient resists shocks and stays the same; the antifragile gets better.”

Nassim Nicholas Taleb

RESUMO

Introdução: Estudos não experimentais podem ajudar a identificar associações entre fatores de risco ou marcadores prognósticos de doenças cardiovasculares e gerar hipóteses. No entanto, esses estudos não são suficientes para confirmar associações causais, pois a natureza e a força das associações podem ser afetadas por fatores de confusão. Apesar disso, estudos observacionais são, indevidamente, utilizados para inferir causalidade. **Objetivo:** O objetivo desse trabalho é descrever a frequência de estudos não experimentais testando eficácia de conduta intervencionista ou farmacológica na área de cardiologia. **Métodos:** Nesse estudo de caráter meta-epidemiológico, foram identificados todos os artigos publicados, de 2019 a 2020, nas cinco revistas de medicina geral e cardiologia: New England Journal of Medicine (NEJM), Journal of the American Association (JAMA), Circulation, Journal of the American College of Cardiology (JACC) e European Heart Journal (EHJ). Essas revistas foram selecionadas, propositalmente, por possuírem os maiores fatores de impacto do mundo – o que fornece uma amostra de estudos com maior probabilidade de ter um rigor metodológico elevado. Dentre essas publicações, foram incluídos os estudos cujo objetivo principal é testar a eficácia de condutas intervencionistas ou farmacológicas cardiovasculares. Posteriormente, dois investigadores independentes extraíram desses trabalhos a variável: o estudo é não experimental e, caso seja, qual o tipo metodológico de estudo – coorte retrospectiva, coorte prospectiva, estudo transversal, estudo caso-controle ou estudo ecológico. Além disso, foi coletada a variável: existe uma afirmação clara, na conclusão do estudo observacional, de que ele é apenas um estudo gerador de hipótese, ou seja, apenas um estudo exploratório. **Resultados:** Dos 2.004 artigos pré-selecionados, 408 estudos testando eficácia de tratamento foram identificados e incluídos em nossa amostra, com frequência de 24% (IC 95%, 20% - 28%) de estudos não experimentais. Entre esses estudos não experimentais, 30% (IC 95%, 21% - 39%) afirmam em sua conclusão, de maneira clara, que são apenas geradores de hipóteses. **Conclusão:** Os desenhos não experimentais não correspondem à maioria dos estudos que se propõem a avaliar benefício de tratamento nas revistas cardiológicas de alto impacto. No entanto, a frequência de 24% de estudos não experimentais testando eficácia de tratamento, associado à baixa frequência de uma afirmação clara, em suas conclusões, de que eles são apenas geradores de hipóteses, pode gerar consequências não intencionais ao ecossistema científico.

Palavras-chave: Estudos observacionais; Tratamento; Eficácia; Causalidade; Medicina baseada em evidências; Metaciência; Cardiologia.

ABSTRACT

Introduction: Non-experimental studies can help to identify associations between risk factors or prognostic markers of cardiovascular disease and generate hypotheses. However, these studies are not sufficient to confirm causal associations, as the nature and strength of associations can be affected by confounding factors. Despite this, observational studies are often misused to infer causality. **Objective:** The objective of this paper is to describe the frequency of non-experimental studies testing efficacy of interventional or pharmacological conduct in the field of cardiology. **Methods:** In a meta-epidemiological design, all papers published from 2019 to 2020 in the five journals of general medicine or cardiology were identified: New England Journal of Medicine (NEJM), Journal of the American Association (JAMA), Circulation, Journal of the American College of Cardiology (JACC) and European Heart Journal (EHJ). This sample was purposely selected because it has the highest impact factors in the world – which provides a sample of studies most likely to have high methodological rigor. Among these publications, studies whose the primary objective is to test the efficacy of interventional or pharmacological conduct in cardiology were included. Thereafter, two independent investigators extracted the variable of these studies: is the study non-experimental and, if so, what is the methodological type of study – retrospective cohort, prospective cohort, cross-sectional study, case-control study or ecological study. In addition, the variable was collected: there is a clear statement, in the conclusion of the observational study, that it is just a hypothesis generator study, that is, just an exploratory study. **Results:** Of the 2004 pre-selected articles, 408 studies suggesting treatment were identified and included in our sample, with a prevalence of 24% (95% CI, 20% - 28%) of non-experimental studies. Among these observational studies, 30% (95% CI, 21% - 39%) declare to be just a hypothesis generator in their conclusion. **Conclusion:** The non-experimental designs do not correspond to the majority of studies that aim to assess the benefit of treatment in high-impact cardiology journals. However, the frequency of 24% of non-experimental studies testing treatment, associated with the low frequency of a clear affirmation in its conclusions, that they are only hypothesis generators, can generate unintended consequences for the scientific ecosystem.

Keywords: Observational studies; Treatment; Efficacy; Causality; Evidence-based medicine; Metascience; Cardiology.

LISTA DE ILUSTRAÇÕES

Figura 1: Fluxograma de identificação e inclusão dos artigos utilizados na amostra deste estudo	21
Figura 2: Frequência de estudos não experimentais que testam eficácia de tratamento	22
Figura 3: Frequência de estudos que afirmam ser apenas geradores de hipótese entre os estudos não experimentais.....	23
Figura 4: Frequência de positividade do desfecho primário entre os estudos não experimentais.....	24
Figura 4: Frequência de positividade do desfecho primário e do reconhecimento de ser um estudo gerador de hipótese entre os trabalhos não experimentais.....	24

LISTA DE TABELAS

Tabela 1: Frequência de estudos não experimentais testando eficácia de tratamento de acordo com a revista de publicação	23
---	----

SUMÁRIO

1. INTRODUÇÃO	10
2. OBJETIVOS	11
2.1. Objetivo primário	11
2.2. Objetivo secundário	11
3. REVISÃO DE LITERATURA	12
3.1. Modelo de estudos observacionais	12
3.2. Risco de vieses em estudos observacionais	13
3.3. Modelo de estudo ideal para testar benefício de tratamento	13
4. METODOLOGIA	15
4.1. Desenho de estudo	15
4.2. Amostra estudada	15
4.3. Estratégia de busca	15
4.3.1. Critérios de inclusão	16
4.3.2. Critérios de exclusão	16
4.4. Extração de dados	16
4.4.1. Variáveis características da amostra	17
4.4.2. Variáveis de desfecho	19
4.5. Tamanho amostral	19
4.6. Plano de análise de dados	19
4.7. Considerações éticas	20
5. RESULTADOS	21
5.1. Triagem e seleção de estudos	21
5.2. Análise da frequência de estudos não experimentais que testam eficácia de tratamento	22
5.3. Característica da amostra de estudos não experimentais	22
5.4. Análise de frequência de estudos que declaram ser geradores de hipótese	23
5.5. Análise da positividade do desfecho primário	24
6. DISCUSSÃO	25
7. CONCLUSÃO	28
REFERÊNCIAS	29

1. INTRODUÇÃO

Estudos observacionais podem contribuir consideravelmente para identificar associações entre fatores de risco ou marcadores prognósticos de doença cardiovascular, avaliar as tendências geográficas e temporais nos cuidados cardiovasculares, avaliar dano de condutas e gerar hipóteses para estudos futuros^{1,2,3}. No entanto, esses estudos não são suficientes para confirmar associações causais para a maioria das terapias, uma vez que a natureza e a força das associações podem ser afetadas por fatores de confusão^{1,2}.

Sendo assim, estudos observacionais não devem ser utilizados para promover uma mudança no tratamento de doenças. O delineamento ideal para estudos que testam a eficácia de condutas medicamentosas, procedimentos cirúrgicos ou outras intervenções é o ensaio clínico randomizado, já que as diferenças observadas nas taxas de resultados do estudo podem ser atribuídas, de modo causal, aos efeitos da intervenção^{1,4,5}.

No que tange à prática clínica, um tratamento deve ser substituído por outro melhor, mas ao utilizar estudos não experimentais como evidência de benefício de tratamento, muitas vezes, observa-se o contrário. Em um estudo observacional, publicado por Go et al., em 2006, foi concluído que o uso de estatinas estava associado a uma mortalidade de 30% a 40% menor em pacientes com insuficiência cardíaca⁶. Entretanto, dois grandes ensaios clínicos randomizados realizados posteriormente não demonstraram efeito deste tratamento na mortalidade nos casos de insuficiência cardíaca^{6,7,8}. Esse é um dos exemplos que evidenciam que grandes ensaios clínicos randomizados bem delineados são imprescindíveis para estimar magnitude de efeito de tratamento^{9,10,11}.

Tendo em vista os impactos clínicos trazidos pela utilização de condutas intervencionistas e farmacológicas baseadas em evidências não experimentais, nas quais os efeitos dos tratamentos podem ser mais incertos e substancialmente menores do que o estimado^{9,12}, é relevante analisar a frequência de estudos observacionais que testam benefício de intervenções médicas. As evidências dos impactos trazidos por essa prática serão motivo de busca desse atual trabalho, visando como desfecho a mudança de percepção e comportamentos na produção de ciência e no uso dela para tomada de decisão médica.

2. OBJETIVOS

2.1. Objetivo primário

Descrever a frequência de desenhos de estudos não experimentais entre os trabalhos que testam eficácia de condutas farmacológicas ou intervencionistas, na área cardiológica, publicados nas cinco revistas de cardiologia de maior impacto.

2.2. Objetivo secundário

- Descrever a frequência de estudos que afirmam, de maneira clara, em sua conclusão, que são apenas geradores de hipótese entre os estudos não experimentais publicados nas cinco revistas de cardiologia de maior impacto.
- Descrever a frequência de positividade do desfecho primário entre os estudos não experimentais que avaliam eficácia de tratamento publicados nas cinco revistas de cardiologia de maior impacto.

3. REVISÃO DE LITERATURA

3.1. Modelos de estudos observacionais

Estudos observacionais são estudos clínicos em que o pesquisador observa um determinado comportamento ambiental e não desempenha um papel ativo no que acontece. Esse modelo de estudos é subdividido em três principais tipos: os estudos de coorte, caso-controle e transversais. No entanto, alguns autores utilizam o termo “estudo de acompanhamento” e “estudo longitudinal” como sinônimos para estudo de coorte ou “estudo de prevalência” como sinônimo para estudo transversal¹³. Os estudos observacionais de coorte são utilizados para descrever grupos de pessoas com alguma característica em comum ao serem reunidas, as quais são observadas por um determinado período, com o intuito de coletarmos informações sobre o que acontece a elas. Nos estudos de caso-controle são selecionados pacientes que desenvolveram a doença em questão e indivíduos semelhantes que não desenvolveram essa doença e os pesquisadores olham para o passado desses pacientes e medem a frequência de exposição a um determinado fator de risco nos dois grupos. Já nos estudos transversais, os investigadores acessam todos os indivíduos em uma amostra simultaneamente, e avaliam a prevalência de exposições, fatores de risco ou de determinada doença⁴.

A escolha adequada do desenho de estudo é um dos fatores essenciais para que uma pesquisa seja executada de maneira bem-sucedida e, posteriormente, utilizada na prática clínica. Sendo assim, estudos observacionais podem ser úteis, dentre outras coisas, para avaliar a ligação entre a exposição ambiental e as doenças, identificar fatores de risco ou marcadores prognósticos de doenças, identificar novos alvos potenciais de drogas e para avaliar a carga de doenças e planejar recursos^{1,2,3}. Além disso, as análises de dados observacionais que relatam associações entre exposições e desfechos podem gerar hipóteses para testes prospectivos em um ECR¹. Entretanto, frequentemente, estudos observacionais são utilizados indevidamente para inferir associações causais para terapias, ainda que a natureza e a força dessas associações possam ser afetadas por fatores de confusão^{1,2}.

3.2. Risco de vieses em estudos observacionais

Os estudos observacionais sofrem, constantemente, o risco de vieses, ou seja, um processo, em qualquer estágio da inferência, com tendência de produzir resultados que se afastem sistematicamente dos valores verdadeiros^{1,3,4}. Embora existam diversos tipos de vieses descritos, a maioria se enquadra nas categorias de viés de seleção, viés de aferição ou viés de confusão. O viés de seleção, por sua vez, ocorre quando dois grupos de pacientes são comparados e eles diferem em outros determinantes do desfecho além do que está sendo estudado, por exemplo, pacientes alocados para o grupo tratamento são mais saudáveis do que os indivíduos alocados para o grupo controle, favorecendo, assim, o grupo tratamento. Já o viés de aferição ocorre quando o método de aferição é diferente entre os grupos de pacientes, levando a resultados incorretos na pesquisa. Por fim, o viés de confusão, ocorre quando um ou mais fatores estão associados e o efeito de um se confunde com o efeito de outro ou é distorcido pelo efeito de outro, por exemplo, a relação entre a ingestão de um medicamento e o risco para o desenvolvimento de determinadas doenças é potencialmente confundido pelas características e comportamento dos pacientes^{1,4}. Sendo assim, uma vez que os estudos observacionais correm o risco de vieses não controlados, as comparações randomizadas são mais adequadas para o fornecimento de evidências de alta qualidade sobre a eficácia e segurança das intervenções¹.

3.3. Modelo de estudo ideal para testar benefício de tratamento

Os ensaios clínicos randomizados garantem distribuição equitativa de fatores de confusão conhecidos e desconhecidos entre os grupos de tratamento e controle, com uma alocação de pacientes entre esses grupos diferindo apenas pelo acaso, evitando, assim, o viés de seleção. Desse modo, as diferenças observadas em resultados podem ser atribuídas à aplicação ou não da intervenção do estudo^{1,3}. Além disso, a verificação de resultados entre os grupos sem distinção minimiza o viés na aferição dos efeitos do tratamento, o que pode ser ainda mais otimizado através do cegamento. Portanto, ensaios clínicos randomizados correspondem ao único modelo de estudo que pode ser considerado confirmatório, já que é capaz de eliminar efeitos de confusão^{1,3}. Apesar disso, há estudos observacionais, indevidamente, na tentativa de demonstrar efeito de tratamento entre as revistas cardiológicas de alto impacto, sendo

fundamental calcular a frequência desses estudos nesses periódicos, partindo do princípio que essa é uma conduta inadequada, a qual não gera contribuição científica e, conseqüentemente, impacta negativamente em condutas clínicas.

4. MATERIAIS E MÉTODO

4.1. Desenho do estudo

Trata-se de um estudo descritivo, que possui artigos científicos como unidade de análise e, portanto, de caráter meta-epidemiológico. O projeto aqui descrito faz parte do segmento Metaciência, do grupo de pesquisa Registro de Síndromes Coronarianas Agudas (RESCA), denominado “*Uma Perspectiva Geral do Overuse Clínico e Científico em Cardiologia*”. O protocolo, estabelecido a priori, está disponível no OSF Registries, em: <https://osf.io/rs9k6>.

4.2. Amostra estudada

Com o propósito de estudar uma amostra mais recente, esse projeto teve como objeto de estudo os artigos publicados, durante os anos de 2019 e 2020, nas cinco revistas de maior fator de impacto de cardiologia e de medicina geral, sendo elas: New England Journal of Medicine (NEJM), Journal of the American Association (JAMA), Circulation, Journal of the American College of Cardiology (JACC) and European Heart Journal (EHJ). Esses periódicos foram escolhidos, propositalmente, por possuírem maior quantidade de artigos científicos da área de cardiologia e, ao mesmo tempo, por possuírem os maiores fatores de impacto do mundo – o que fornece uma amostra de estudos com maior probabilidade de ter um rigor metodológico elevado.

4.3. Estratégia de busca

A coleta de dados foi dividida em duas etapas: a primeira foi realizada visando a identificar os estudos publicados, nos anos de 2019 e 2020, nas cinco revistas de medicina geral e de cardiologia de maior impacto. Essa etapa foi realizada diretamente nos websites de cada respectiva revista por dois avaliadores de forma independente. Inicialmente foram identificados todos os artigos originais publicados em cada revista de janeiro de 2019 a dezembro de 2020. Desses, foram selecionados para a amostra aqueles estudos que abordam tratamentos cardiológicos, independentemente do respectivo desenho de estudo. Essa seleção foi realizada mediante leitura de título e abstract de cada estudo publicado no período pré-definido.

Foram excluídos os estudos de outras áreas da medicina e aqueles sobre cardiologia, mas que o objetivo primário não estava relacionado a terapêutica. Os dados foram coletados por dois avaliadores independentes e armazenados no programa Microsoft Excel 2016. Ao final da coleta, esses avaliadores se reuniram e compararam suas planilhas. Nos casos de divergências entre os avaliadores, um terceiro avaliador foi consultado.

4.3.1. Critérios de inclusão

Foram selecionados para a amostra aqueles estudos que abordavam tratamentos cardiológicos, intervencionistas ou farmacológicos, os quais deveriam passar por prescrição médica, independentemente do respectivo desenho de estudo.

4.3.2. Critérios de exclusão

Foram excluídos os estudos que avaliam efetividade ou dano de condutas intervencionistas ou farmacológicas como objetivo primário. Além disso, foram excluídos da análise estudos que abordavam intervenções não prescritas por médicos, por exemplo, mudança de hábitos de vida ou dietas.

4.4. Extração dos dados

Uma vez já tendo sido selecionados os artigos que compuseram a amostra, foram coletadas as variáveis relacionadas às características desses artigos, relacionadas à transparência na análise, à divulgação dos dados e as variáveis de desfecho. Essas variáveis foram extraídas com a leitura dos estudos e nos protocolos dos estudos publicados no site <http://clinicaltrials.gov> ou em outra plataforma de registros que esteja indicada no corpo do artigo. Analogamente, essa coleta foi realizada por pares de avaliadores independentes que se reuniram ao final da coleta, compararam as suas planilhas e, em casos de divergências, um terceiro avaliador foi consultado.

4.4.1. Variáveis características da amostra

- Título do artigo;
- Ano de publicação;
- Periódico de publicação;
- País de origem;
- Número de objetivos primários;
- Número de objetivos secundários;
- Número de desfechos primários;
- Número de desfechos secundários;
- Natureza numérica do resultado primário (0 – simples, 1 – combinado), sendo que um resultado é classificado como resultado combinado se for uma combinação de 2 ou mais dados;
- Natureza clínica do desfecho primário (0 – clínico, 1 – substituto): um desfecho substituto foi definido como um desfecho que não é perceptível clinicamente pelo paciente como, por exemplo, uma variação de um parâmetro de um teste laboratorial, o qual é diferente de um desfecho clínico como alívio da dor e morte;
- Natureza constitutiva do desfecho primário (0 – duro; 1 – intermediário; 2 – mole): considerou-se como desfecho mole aquele que possui uma dimensão subjetiva como intensidade da angina, capacidade funcional relatada e dispneia, já os desfechos intermediários são morte por doença cardiovascular, IAMSSST, hospitalização por causa cardiovascular e necessidade de revascularização, por fim, um desfecho duro é mais objetivo, como STEMI, acidente vascular cerebral e morte por todas as causas¹²;
- Estruturação da coleta de dados (0 – retrospectivo; 1 – prospectivo): se o estudo foi planejado antes da coleta de dados, foi definido como um estudo prospectivo e se os procedimentos de coleta de dados ocorreram antes do planejamento do estudo, foi considerado um estudo retrospectivo;
- Percentual de perda de seguimento: calculado a partir da amostra inicial do estudo e da amostra final analisada;
- Centro (0 – estudo unicêntrico; 1 – estudo multicêntrico);

- Presença de patrocínio da indústria farmacêutica (0 – não; 1 – sim, estudo apoiado por esta indústria; 2 – sim, estudo realizado por esta indústria);
- Positividade do estudo quanto ao desfecho secundário (0 – não 1 – sim): quando a hipótese nula foi rejeitada por pelo menos um desfecho secundário;
- Presença de cálculo do tamanho amostral necessário (0 – não, 1 – sim);
- Tamanho da amostra planejada: número de participantes que seriam necessários para mensuração do desfecho primário declarado antes da realização do estudo;
- Tamanho real da amostra: número de participantes incluídos e analisados no estudo;
- Atingimento do tamanho da amostra planejado: percentual do tamanho da amostra planejado que foi alcançado pelo estudo;
- Poder estimado do estudo feito a priori: poder do estudo que foi estimado pelo autor a priori no estudo, com base no tamanho da amostra planejada;
- Poder estimado de estudo a posteriori: foi calculado o poder de estudo a posteriori, dado o tamanho final da amostra obtido e as premissas previamente estabelecidas;
- Presença de comparações múltiplas (0 – não, 1 – sim): definido como a presença de duas ou mais comparações em relação ao mesmo desfecho^{13,14};
- Presença de correções para comparações múltiplas (0 – não, 1 – sim): se houve alguma medida para corrigir as comparações múltiplas, por exemplo, a correção de Bonferroni;
- Presença de spin (0 – não, 1 – sim): definido como uma atenuação do resultado primário negativo pela menção e destaque de uma análise secundária positiva na conclusão do estudo ou no final da discussão ou resumo^{15,16};
- Existe algum ajuste para variáveis de confusão se uma análise multivariada foi realizada? (0 – não, 1 – sim, 9999 – não foi realizada análise multivariada);
- Protocolo do estudo postado a priori (0 – não, 1 – sim): verificou se havia um protocolo de estudo publicado antes do início do estudo em algum site para registro de protocolo, por exemplo, o ClinicalTrials.gov.

4.4.2. Variáveis de desfecho

- O estudo é não experimental e, caso seja, qual o tipo metodológico de estudo não experimental (9999 – não é um estudo observacional, 0 – coorte retrospectiva, 1 – coorte prospectiva, 2 – estudo transversal, 3 – estudo caso-controle, 4 – estudo ecológico);
- Caso o artigo seja não experimental, existe uma afirmação clara em sua conclusão de que ele é apenas um estudo gerador de hipótese, ou seja, um estudo exploratório? (0 – não; 1 – sim);
- Positividade do estudo quanto ao desfecho primário (0 – não; 1 – sim): verificou se a hipótese nula foi rejeitada por pelo menos um desfecho primário.

4.5. Tamanho amostral

São necessários 385 artigos avaliando benefício de tratamento, para calcular a frequência de estudos não experimentais, com mais ou menos 5% de precisão e um alfa de 5%, em se considerando uma prevalência estimada de 50%, baseada numa estimativa feita pelo artigo publicado, em 2002, por Ioannidis et. *al*⁹.

4.6. Plano de análise dos dados

A análise primária constituiu-se em calcular a frequência dos estudos não experimentais que testam eficácia de tratamento na amostra selecionada e realizar o cálculo do intervalo de confiança (IC), determinado pelo 95% exato de Fisher. Após isto, foi calculada a frequência dos estudos que declaram ser geradores de hipótese e a frequência de positividade do desfecho primário entre os estudos não experimentais, e realizado o cálculo do intervalo de confiança (IC), determinado pelo 95% exato de Fisher.

A concordância entre os avaliadores independentes em relação às variáveis de desfecho foi calculada através do coeficiente de Kappa, de modo que, valores de Kappa menores que 0.4 foram considerados concordância pobre, valores entre 0.4 e 0.75, moderada e maiores que 0.75, excelente¹⁷.

As variáveis independentes categóricas foram expressas em proporções. Já as variáveis numéricas foram expressas utilizando médias e desvio padrão (DP), caso apresentassem uma distribuição normal e, em caso de distribuição anormal, foram apresentadas utilizando medianas e intervalos interquartis. A normalidade das variáveis foi testada pelo teste de Kolmogorov-Smirnov.

Os dados foram compilados em banco de dados com o programa Microsoft Excel 2016, alimentado pelas variáveis descritas na seção de extração de dados e coletadas por dois avaliadores independentes, com a consulta de um terceiro avaliador em casos de divergências. O programa utilizado para análise estatística foi o IBM SPSS (Versão 25, SPSS Inc.) e o cálculo amostral foi realizado por meio do programa WinPepi versão 11.65.

4.7. Considerações éticas

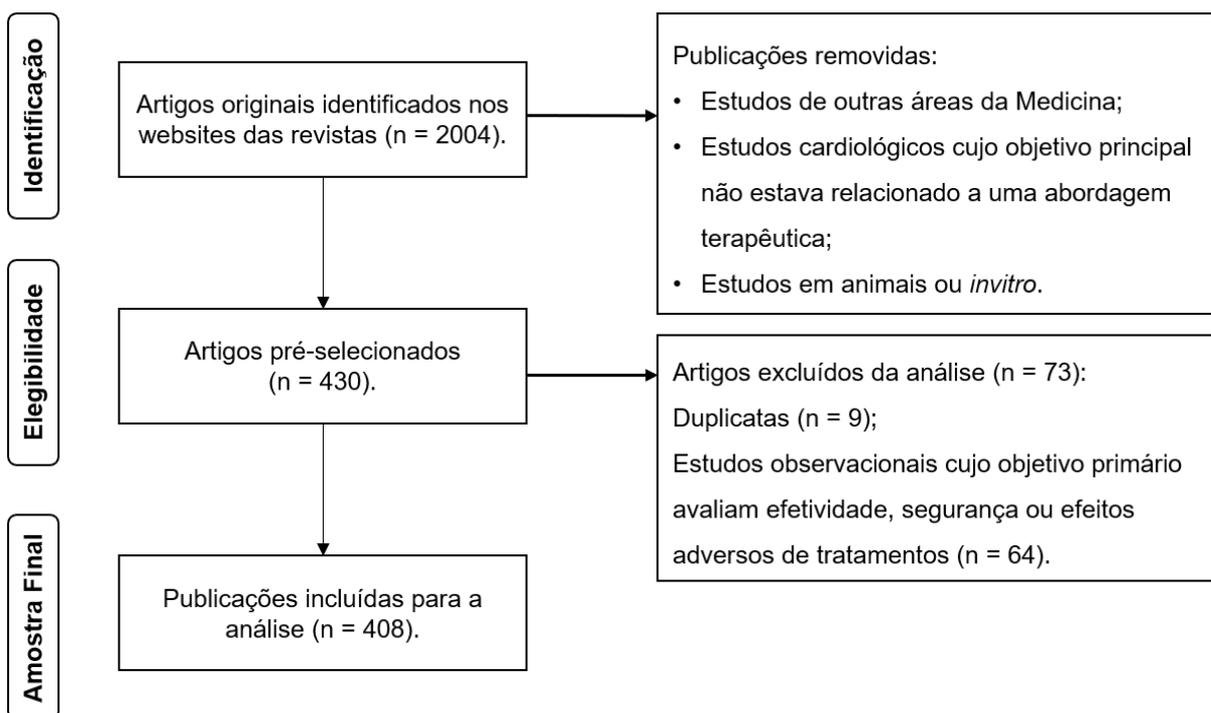
Pela resolução nº 466/12, de 12 de dezembro de 2012 do Conselho Nacional de Saúde, pesquisas que utilizam dados de domínio público não necessitam ser registradas ou avaliadas pelo Comitê Nacional de Ética em Pesquisa (CEP) ou o Comitê de Ética em Pesquisa. Assim, não foi necessária a submissão do presente estudo ao CEP.

5. RESULTADOS

5.1. Triagem e seleção dos estudos

Um total de 2004 artigos foram identificados, utilizando a estratégia de seleção descrita nos métodos, nos websites das cinco revistas de maior impacto de cardiologia e de medicina geral, das edições entre janeiro de 2019 e dezembro de 2020. Dos 2004 artigos, 481 foram pré-selecionados por se tratarem de estudos que abordavam tratamentos cardiológicos. 73 estudos foram retirados da análise, dentre eles, 9 foram considerados duplicatas e 64 tinham seu objetivo primário avaliando efetividade ou segurança de condutas intervencionistas ou farmacológicas. Por fim, 408 artigos foram classificados como estudos avaliando benefício de tratamento na área cardiológica e incluídos na análise (Figura 1).

Figura 1. Fluxograma de identificação e inclusão dos artigos utilizados na amostra deste estudo.

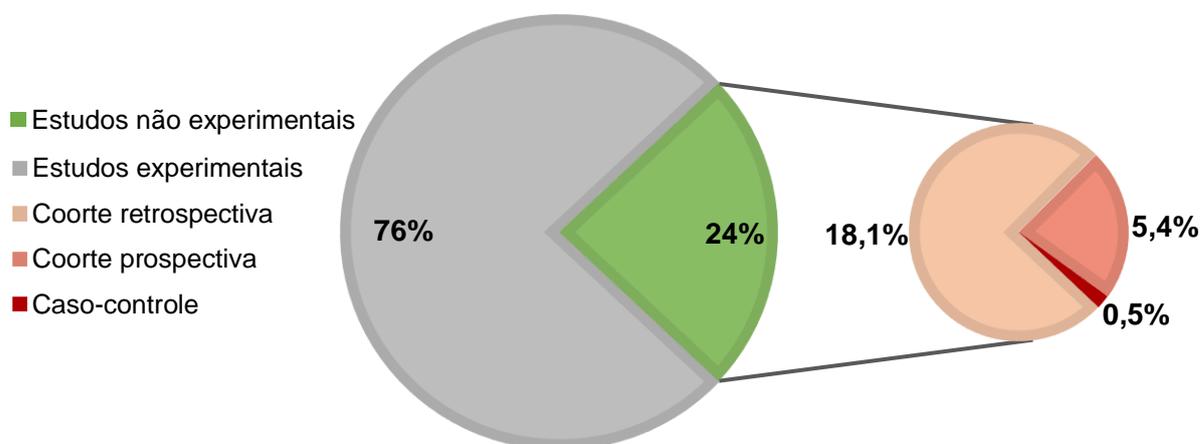


Fonte: Dados do próprio autor.

5.2. Análise da frequência de estudos não experimentais que testam eficácia de tratamento

Entre os 408 estudos que testam benefício de tratamento, 98 são trabalhos não experimentais, ou seja, 24% [95% IC, 20%-28%]. O total de coortes retrospectivas testando eficácia de tratamento foi 74 (18,1%), 22 (5,4%) coortes prospectivas e 2 (0,5%) estudos de caso-controle. A figura 2 demonstra esses resultados (Figura 2).

Figura 2. Frequência de estudos não experimentais que testam eficácia de tratamento.



Fonte: Dados do próprio autor.

5.3. Características da amostra de estudos não experimentais

A tabela 1 apresenta a caracterização da amostra de artigos não experimentais coletados em relação a revista de publicação (Tabela 1). As revistas de publicações com maior frequência desses estudos foram, de maior para menor, respectivamente, a JACC (51%), a EHJ (27,6%), a Circulation (17,3%) e a JAMA (4,1%). Não foi identificado nenhum estudo cardiológico não experimental testando eficácia de condutas intervencionistas ou farmacológicas na NEJM.

Tabela 1. Frequência de estudos não experimentais testando eficácia de tratamento de acordo com a revista de publicação.

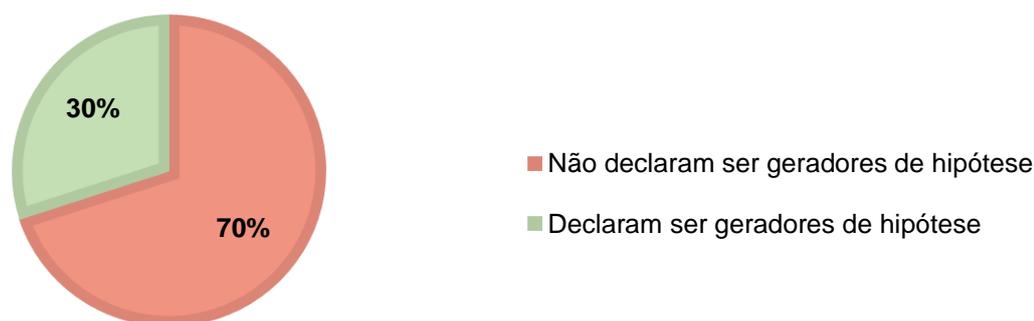
Amostra	N (%)
Local de origem	
Revista de Publicação	
JACC	51%
EHJ	27,6%
Circulation	17,3%
JAMA	4,1%
NEJM	0%

Fonte: Dados do próprio autor.

5.4. Análise da frequência de estudos que declaram ser geradores de hipótese

Entre os estudos não experimentais testando eficácia de tratamento, 30% [95% IC, 21%-39%] reconhecem a limitação de ser apenas um estudo gerador de hipótese em sua conclusão. A figura 3 demonstra essa frequência (Figura 3).

Figura 3. Frequência de estudos que afirmam ser apenas geradores de hipótese entre os estudos não experimentais.



Fonte: Dados do próprio autor.

5.5. Análise da positividade do desfecho primário

Entre os estudos não experimentais testando eficácia de condutas médicas, a frequência de trabalhos com desfecho primário positivo foi de 82% [95% IC, 74%-90%]. A figura 4 demonstra esse resultado (Figura 4).

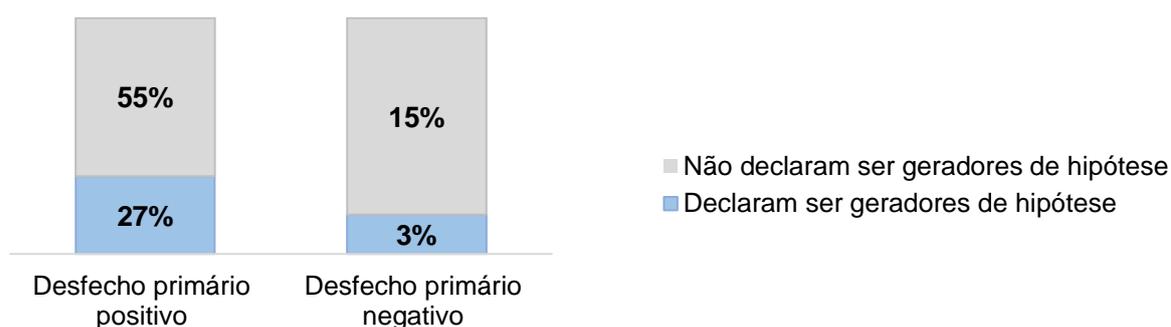
Figura 4. Frequência de positividade do desfecho primário entre os estudos não experimentais.



Fonte: Dados do próprio autor.

Entre os 98 estudos não experimentais, 27% possuem desfecho primário positivo e, simultaneamente, reconhecem a limitação de ser apenas gerador de hipótese. Ainda entre os estudos não experimentais, 55% possuem desfecho primário positivo e, concomitantemente, não declaram ser geradores de hipótese. A figura 5 demonstra esses resultados (Figura 5).

Figura 5. Frequência de positividade do desfecho primário e do reconhecimento de ser um estudo gerador de hipótese entre os trabalhos não experimentais.



Fonte: Dados do próprio autor.

6. DISCUSSÃO

Os estudos com desenhos não experimentais não são frequentes entre os artigos que estimam efeito de tratamento nas revistas de maior impacto em cardiologia. Entretanto, entre esses trabalhos, é baixa a frequência dos que afirmam, de maneira clara, que são apenas geradores de hipótese. Além disso, é elevada a frequência de estudos com desfecho primário positivo entre os trabalhos não experimentais que testam eficácia de condutas médicas.

Cabe lembrar que a frequência de 24% [95% IC, 20%-28%] de artigos não experimentais encontrados entre os trabalhos que testam eficácia de tratamento no presente estudo, comparado com os 41% encontrados por Ioannidis *et al.*⁹ no seu artigo publicado em 2001, possivelmente, está associada ao fato de o atual trabalho ter como unidade de análise apenas as revistas de maior impacto na área de cardiologia e medicina geral, aproximando os resultados à realidade de revistas com maior qualidade científica. Sendo assim, é esperado que em revistas científicas de menor impacto, a quantidade de estudos não experimentais sugerindo tratamentos cardiovasculares seja maior, já que o rigor metodológico e conceitual desses periódicos não é tão elevado, comparado ao rigor das revistas de maior fator de impacto¹⁸.

No que diz respeito aos tipos metodológicos encontrados, o mais prevalente foi de coorte retrospectiva (18,1%). Entretanto, segundo Deeks, J J *et al.*¹⁹, especialmente nos trabalhos que utilizam registros históricos, ao incluir participantes retrospectivamente nos estudos, vieses adicionais são potencialmente introduzidos¹⁹. Isso resulta do fato de que os pesquisadores podem não utilizar critérios de inclusão adequados e podem não garantir que os dados sejam completos e registrados de modo consistente, bem como não é assegurado de que informações sobre possíveis confundidores estão sendo coletadas^{5,19}. Não obstante, Ioannidis *et al.*⁹, descreve maiores taxas de discrepâncias entre os resultados de estudos de coorte retrospectiva e os resultados de estudos controlados randomizados sobre o mesmo tópico, um fator ainda mais preocupante no que tange a confiabilidade dessas evidências⁹.

Apesar de os estudos não experimentais cujo objetivo é testar a eficácia de condutas médicas não serem considerados confirmatórios, eles podem produzir resultados promissores que podem levar ensaios subsequentes a validar os benefícios de tratamentos, ou seja, gerar hipóteses para estudos futuros^{1,3,9}. Sendo assim, estes

artigos sobre tratamento devem afirmar em sua conclusão e, não em outras seções do estudo, que são necessários ensaios clínicos randomizados para confirmar a hipótese primária testada.

Contudo, como exposto nos resultados, foi encontrada uma baixa frequência de estudos que explanam de forma transparente o seu caráter exploratório, o que correspondeu a apenas 30% [95% IC, 21%-39%] da amostra de trabalhos não experimentais. Vale ressaltar que, em muitos estudos, a afirmação de que ele é apenas exploratório pode ser encontrada na sessão de limitações do artigo, tal como foi observado durante a coleta dos dados. No entanto, foi considerado que o autor reconhecia a limitação de ser apenas um estudo gerador de hipótese somente quando ele declarava isso na conclusão, já que estamos avaliando a inferência de uma associação causal entre o tratamento e um determinado desfecho e levando em consideração que a seção de limitações do artigo não é frequentemente consultada durante a análise de artigos^{20,21}.

Uma alta frequência de desfechos primários positivos (82% [95% IC, 74%-90%]) foi encontrada neste trabalho e, possivelmente, está associada ao que Ewald *et al.*²², em 2019, aponta em seu artigo – que os estudos não randomizados demonstram um efeito de tratamento maior para uma conduta experimental do que os estudos randomizados²². Esse fenômeno está ligado à ausência de uma atribuição aleatória para receber intervenções, o que gera efeitos de confusão não controlados por conta de covariáveis desconhecidas, não medidas ou medidas inadequadamente²³⁻²⁶.

Diante disso, a alta frequência de artigos não experimentais testando eficácia de tratamento com desfecho primário positivo e que não afirmam, de maneira clara, que são apenas estudos exploratórios é um fator preocupante para o cenário de pesquisa e da prática médica. Na medida em que, muitos desses trabalhos podem corresponder às “fake news científicas”, ou seja, evidências que sugerem credibilidade, porém são estudos falso positivos²⁷.

Estas evidências acabam sendo incorporadas no tratamento de alguns pacientes e podem levar anos para serem descartadas através da publicação de algum ensaio clínico randomizado que reverta essa conduta médica, provando que ela é ineficaz ou até mesmo danosa^{28,29}. Tal fenômeno, descrito como “reversão médica”, por Prasad *et al.*³⁰, tem como exemplo as indicações para intervenções coronarianas percutâneas (ICP) de rotina, as quais foram utilizadas por muitos anos e revertidas por grandes

ensaios clínicos que concluíram que o uso do ICP não altera as taxas de infarto do miocárdio, morte cardíaca ou morte por todas as causas^{31,32}.

Sendo assim, faz-se necessário a familiarização com os princípios básicos de uma boa pesquisa e adoção da dúvida e do questionamento antes de tirar conclusões prematuras com base em estudos não experimentais ou em estudos com falhas metodológicas³². Dessa forma, poderão ser evitados os danos à saúde de pacientes bem como os impactos da mudança de recomendações, como a perda de confiança no sistema médico e o movimento anticidência.

Devemos reconhecer algumas limitações do presente trabalho: por se tratar de um estudo meta-epidemiológico, as variáveis foram coletadas sob o julgamento dos avaliadores, fazendo com que alguns dados fossem avaliados de maneira mais subjetiva, o que afeta a confiabilidade do trabalho. Entretanto, para isso, realizamos a análise de Kappa para quantificar a concordância entre os avaliadores, como sugerido pelo guideline para pesquisas meta-epidemiológicas³³. Além disso, quando havia discordância entre os avaliadores, um terceiro avaliador era consultado para resolução da divergência. Outrossim, algumas variáveis, mesmo com a definição a priori no protocolo, não tiveram definições suficientemente objetivas para a extração de dados, uma vez que estas não apresentam definições padronizadas como os desfechos clínicos na literatura, o que prejudica a reprodutibilidade do estudo.

7. CONCLUSÃO

Os desenhos não experimentais não correspondem à maioria dos estudos que se propõem a avaliar benefício de tratamento nas revistas cardiológicas de alto impacto. No entanto, a frequência de 24% de estudos não experimentais testando eficácia de tratamento, associado à baixa frequência de uma afirmação clara, em sua conclusão, de que eles são apenas geradores de hipóteses, pode gerar consequências não intencionais ao ecossistema científico.

REFERÊNCIAS

1. Bowman L, Baras A, Bombien R, Califf RM, et al. Understanding the use of observational and randomized data in cardiovascular medicine. *Eur Heart J*. 2020 Jul 14;41(27):2571-2578.
2. McMurray JJV. Only Trials Tell the Truth About Treatment Effects. *J Am Coll Cardiol*. 2018 Jun 12;71(23):2640-2642.
3. Correia, L. C. L. Estudos Observacionais: quando confiar, desconfiar ou descartar? Disponível em: <http://medicinabaseadaemevidencias.blogspot.com/2018/11/estudos-observacionais-quando-confiar.html>. Acesso em: 20 de março de 2022.
4. Greenhalgh, Trisha. Como Ler Artigos Científicos: Fundamentos da Medicina Baseada em Evidências. Tradução: Ananyr Porto Fajardo. 5ª. ed. atual. [S. l.]: Artmed, 2015. 282 p. ISBN 8582712227.
5. Fletcher, Robert H. Epidemiologia Clínica: Elementos Essenciais. 5ª. ed. atual. [S. l.]: Artmed, 2014. 296 p. ISBN 978-8582710678.
6. Go AS, Lee WY, Yang J, Lo JC, Gurwitz JH. Statin therapy and risks for death and hospitalization in chronic heart failure. *JAMA*. 2006 Nov 1;296(17):2105-11.
7. Kjekshus J, Apetrei E, Barrios V, Böhm M, Cleland JGF, Cornel JH, et al. Rosuvastatin in Older Patients with Systolic Heart Failure. *New England Journal of Medicine*. 2007 Nov 29;357(22):2248–61.
8. GISSI-HF investigators. Effect of rosuvastatin in patients with chronic heart failure (the GISSI-HF trial): a randomised, double-blind, placebo-controlled trial. *The Lancet*. 2008 Aug 31; 372: 1231–1239.

9. Ioannidis JP, Haidich AB, Pappa M, et al. Comparison of evidence of treatment effects in randomized and nonrandomized studies. *JAMA*. 2001 Aug 15; 286(7):821-30.
10. MacMahon S, Collins R. Reliable assessment of the effects of treatment on mortality and major morbidity, II: observational studies. *Lancet*. 2001 Feb 10;357(9254):455-62.
11. Hemkens LG, Contopoulos-Ioannidis DG, Ioannidis JP. Agreement of treatment effects for mortality from routinely collected data and subsequent randomized trials: meta-epidemiological survey. *BMJ*. 2016 Feb 8; 352:i493.
12. Correia L. C. L. Estudo PROTECT: Medida de BNP protege indivíduos com insuficiência cardíaca? Disponível em: <http://medicinabaseadaemevidencias.blogspot.com/2011/10/>. Acesso em: 14 de dezembro de 2022.
13. Vandembroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration. *Epidemiology*. 2007;18(6):805–35.
14. Gerstein HC, McMurray J, Holman RR. Real-world studies no substitute for RCTs in establishing efficacy. *The Lancet* [Internet]. 2019;393(10168):210–1. Available from: [http://dx.doi.org/10.1016/S0140-6736\(18\)32840-X](http://dx.doi.org/10.1016/S0140-6736(18)32840-X).
15. Benjamini Y. Simultaneous and selective inference: Current successes and future challenges. *Biom J*. 2010 Dec; 52(6):708-21.
16. Miller, R.G. *Simultaneous Statistical Inference* 2nd Ed. Springer Verlag New York, 1981.
17. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther*. 2005 Mar;85(3):257-68.

18. Mahtta D, Altibi A, Gad MM, Samara A, Barakat AF, Bagur R, et al. Methodological rigor and temporal trends of cardiovascular medicine meta-analyses in highest-impact journals. *J Am Heart Assoc.* 2021; 10(18).
19. Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovich C, Song F, et al. Evaluating non-randomised intervention studies. *Health Technol Assess.* 2003;7(27):iii-x, 1-173.
20. Boutron I, Dutton S, Ravaud P, Altman DG. Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. *JAMA.* 2010 May 26;303(20):2058-64.
21. Lazarus C, Haneef R, Ravaud P, Boutron I. Classification and prevalence of spin in abstracts of non-randomized studies evaluating an intervention. *BMC Med Res Methodol.* 2015 Oct 13;15:85.
22. Ewald H, Ioannidis JPA, Ladanie A, Mc Cord K, Bucher HC, Hemkens LG. Nonrandomized studies using causal-modeling may give different answers than RCTs: a meta-epidemiological study. *J Clin Epidemiol.* 2020 Feb;118:29-41.
23. Correia L. C. L. A Provável e Inconveniente Causalidade entre Tiazídico e Câncer de Pele. Disponível em: <https://medicinabaseadaevidencias.blogspot.com/2018/12/a-provavel-e-inconveniente-causalidade.html>. Acesso em: 29 de setembro de 2022.
24. Hemkens LG, Contopoulos-Ioannidis DG, Ioannidis JP. Agreement of treatment effects for mortality from routinely collected data and subsequent randomized trials: meta-epidemiological survey. *BMJ.* 2016 Feb 8; 352:i493.
25. Klungel OH, Martens EP, Psaty BM, et al. Methods to assess intended effects of drug treatment in observational studies are reviewed. *J Clin Epidemiol.* 2004 Dec; 57(12):1223-31.

26. Groenwold RH, Hak E, Hoes AW. Quantitative assessment of unobserved confounding is mandatory in nonrandomized intervention studies. *J Clin Epidemiol*. 2009 Jan;62(1):22-8.
27. Correia L. C. L. Qual o significado de fake news científica? Disponível em: <http://medicinabaseadaemevidencias.blogspot.com/2018/08/qual-o-significado-de-fake-news.html>. Acesso em: 04 de outubro de 2022.
28. Prasad V, Cifu A. Medical reversal: why we must raise the bar before adopting new technologies. *Yale J Biol Med*. 2011 Dec; 84(4):471-8.
29. Coccheri S. Error, contradiction and reversal in science and medicine. *Eur J Intern Med*. 2017 Jun; 41:28-29.
30. Prasad V. K., Cifu A. S. *Ending Medical Reversal: Improving Outcomes, Saving Lives*. Johns Hopkins University Press. 2015. ISBN: 9781421417721.
31. Cifu AS, Prasad VK. Medical Debates and Medical Reversal. *J Gen Intern Med*. 2015 Dec;30(12):1729-30.
32. Prasad V, Cifu A. The reversal of cardiology practices: interventions that were tried in vain. *Cardiovasc Diagn Ther*. 2013 Dec;3(4):228-35.
33. Murad MH, Wang Z. Guidelines for reporting meta-epidemiological methodology research. *Evid Based Med*. 2017 Aug;22(4):139-142.