



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE MEDICINA
CURSO DE NUTRIÇÃO

GABRIEL GUARIGLIA GHIGNATTI

**O PROBLEMA OCULTO NA SAÚDE: OS PROFISSIONAIS SABEM LER? UMA
REVISÃO SISTEMÁTICA**

Porto Alegre

2024

GABRIEL GUARIGLIA GHIGNATTI

O PROBLEMA OCULTO NA SAÚDE: OS PROFISSIONAIS SABEM LER? UMA
REVISÃO SISTEMÁTICA

Trabalho de Conclusão de Curso apresentado ao
Curso de Nutrição da Faculdade de Medicina da
Universidade Federal do Rio Grande do Sul como
requisito parcial para a obtenção do título de
Bacharel em Nutrição.

Orientador(a): Vivian Cristine Luft

Porto Alegre

2024

CIP - Catalogação na Publicação

Ghignatti, Gabriel

O problema oculto na saúde: Os profissionais sabem ler? Uma revisão sistemática / Gabriel Ghignatti. -- 2024.

52 f.

Orientadora: Vivian Luft.

Trabalho de conclusão de curso (Graduação) -- Universidade Federal do Rio Grande do Sul, Faculdade de Medicina, Curso de Nutrição, Porto Alegre, BR-RS, 2024.

1. Bioestatística. 2. Epidemiologia. 3. Metodologia. 4. Lacuna de evidências. I. Luft, Vivian, orient. II. Título.

GABRIEL GUARIGLIA GHIGNATTI

O PROBLEMA OCULTO NA SAÚDE: OS PROFISSIONAIS SABEM LER? UMA
REVISÃO SISTEMÁTICA

Trabalho de conclusão de curso de graduação apresentado à Faculdade de Medicina da
Universidade Federal do Rio Grande do Sul como requisito parcial para a obtenção do título
de Bacharel em Nutrição.

Aprovado em: 19/02/2024

BANCA EXAMINADORA

Nut. Me. Isadora Jardim de Almeida
Departamento de Epidemiologia - UFRGS

Nut. Me. Tássia Rolim Camargo
Departamento de Epidemiologia - UFRGS

Prof^a. Dra. Nut. Vivian Cristine Luft

AGRADECIMENTOS

Em primeiro lugar, quero expressar minha profunda gratidão aos amigos que, com paciência, compreenderam minhas ausências, às vezes longas, durante estes intensos cinco anos. A paciência, entendimento e apoio incondicionais foram pilares para mim.

À minha amada Maria Eduarda, minha parceira leal e constante fonte de apoio, você foi a rocha sobre a qual pude me apoiar nos momentos mais turbulentos. Sua presença foi um porto seguro nos dias tempestuosos, sempre pronta a ouvir minhas reclamações e angústias com um coração aberto e uma mente compreensiva. Sua disposição inabalável em estar ao meu lado fortaleceu minha resiliência e me motivou a enfrentar cada desafio.

Um agradecimento especial ao meu avô João, cujo generoso apoio permitiu que eu me dedicasse integralmente à minha graduação sem as distrações e preocupações do mundo externo. Foi o alicerce sobre o qual construí meu sucesso acadêmico. Sua contribuição foi indispensável, e minha gratidão é imensurável.

Às minhas incansáveis companheiras de jornada, Camilla e Giovanna, minha gratidão é imensa. Com vocês, cada momento dentro e fora da sala de aula foi enriquecedor. As horas gastas em discussões pelo Discord, os trabalhos em grupo, e cada risada compartilhada foram fundamentais para minha formação. Sem a amizade e o apoio de vocês, minha experiência universitária teria sido insuportável.

Ana Teresa, minha cupinxa, sua amizade e suporte foram pilares de minha jornada. A partilha de ideias e o encorajamento mútuo foram essenciais para o nosso crescimento e desenvolvimento. Sua paixão pelo aprendizado e evolução são fontes de inspiração constantes para mim.

Camila, minha eterna dupla de estágio e enciclopédia viva de farmacologia e conhecimentos caninos. Tua sabedoria e apoio foram faróis em minha jornada acadêmica. As conversas, por mais desprovidas de sentido que parecessem, e o apoio nos momentos de trabalho

não-remunerado foram fundamentais para minha formação. Teu intelecto e conhecimento são inspirações constantes para meu progresso como profissional.

À minha orientadora, Vivian Luft, minha gratidão é profunda. Sua disposição em ajudar, debater ideias e fornecer orientação foi essencial para o sucesso do meu trabalho. Sou imensamente grato pelo conhecimento compartilhado e pelo apoio contínuo em cada etapa deste caminho.

Por último, mas não menos importante, estendo meus agradecimentos aos professores Carolina Guerini e Luís Portela, cuja dedicação e orientação enriqueceram minha experiência acadêmica e pessoal. A proximidade e o apoio que recebi de vocês foram cruciais para o meu desenvolvimento como profissional e como pessoa. Minha jornada universitária foi imensuravelmente enriquecida pela presença de cada um de vocês em minha vida.

RESUMO

Introdução: Com os avanços tecnológicos, os métodos estatísticos têm se tornado mais complexos e sofisticados, acompanhando o aumento exponencial do volume de publicações científicas. Portanto, é imperativo que os profissionais da saúde se atualizem continuamente em conceitos e técnicas estatísticas para uma interpretação acurada da literatura e evitar a aceitação de resultados falaciosos oriundos de estudos metodologicamente frágeis. Neste contexto, ferramentas como o *Biostatistical Knowledge Test Survey Instrument* (BKTSI) são fundamentais para avaliar a competência estatística, essencial na prática baseada em evidências. **Objetivo:** Avaliar o nível de conhecimento em bioestatística e metodologia de pesquisa entre diferentes grupos de profissionais de saúde, incluindo estudantes, clínicos e docentes. **Métodos:** Realizou-se uma revisão sistemática conforme as diretrizes PRISMA. Foram pesquisados artigos em bases de dados como MEDLINE, Web of Science e EMBASE, focando em estudos que avaliavam o conhecimento objetivo em bioestatística e metodologia. **Resultados:** Foram encontrados 1228 artigos, dos quais 29 foram selecionados para análise. As amostras variaram de 11 a 898 indivíduos, incluindo profissionais de enfermagem, farmácia, medicina, odontologia e saúde pública. Observou-se uma ampla variação nas ferramentas de avaliação. As maiores médias de conhecimento em bioestatística não ultrapassaram 69,8%, com a maioria dos estudos apresentando médias entre 25-60%. **Conclusão:** Os resultados deste estudo apontam para uma lacuna significativa no conhecimento em bioestatística e metodologia entre os profissionais da saúde, ressaltando a urgência de aprimorar a formação nestas áreas. A heterogeneidade das ferramentas de avaliação sugere a necessidade de um padrão mais unificado na avaliação deste conhecimento.

Palavras-chave: bioestatística, epidemiologia, metodologia, prática baseada em evidência

ABSTRACT

Introduction: With technological progress, statistical methods have evolved and become more sophisticated, coinciding with the exponential increase in scientific publications. It's crucial for healthcare professionals to continuously update their understanding of new statistical concepts and techniques for accurate interpretation of the literature and to avoid false conclusions from poorly designed studies. Tools like the Biostatistical Knowledge Test Survey Instrument (BKTSI) play a crucial role in assessing statistical competence, integral skill to evidence-based practices. **Objective:** This study aimed to evaluate biostatistical and research methodology knowledge among diverse healthcare professionals, including students, clinicians, and faculty. **Methods:** A systematic review was conducted following PRISMA guidelines, searching databases like MEDLINE, Web of Science, and EMBASE for studies on biostatistical and methodological knowledge. **Results:** Out of 1228 articles found, 29 were selected for final analysis. Samples ranged from 11 to 898 individuals across nursing, pharmacy, medicine, dentistry, and public health. A wide variation in assessment tools was observed. The highest average knowledge in biostatistics did not exceed 69.8%, with most studies showing mean scores between 25-60%. **Conclusion:** The study reveals a significant gap in biostatistical and methodological knowledge among healthcare professionals, underscoring the need for enhanced educational approaches. The heterogeneity in assessment tools underscores the necessity for a more standardized evaluation of these critical skills.

Keywords: biostatistics, epidemiology, methodology, evidence based practice

LISTA DE ABREVIATURAS E SIGLAS

ANOVA	Análise de Variância
BE	Bioestatística
BKTSI	Biostatistical Knowledge Test Survey Instrument
DP	Desvio Padrão
ELG	Escala de Literacia Gráfica
ENO	Escala de Numeração Objetiva
ENS	Escala de Numeração Subjetiva
EP	Epidemiologia
MBE	Medicina Baseada em Evidência
OLG	Obstetrician-Gynecologist Statistical Literacy Questionnaire
PBE	Prática Baseada em Evidência
VPP	Valor Preditivo Positivo
VPN	Valor Preditivo Negativo

SUMÁRIO

1 INTRODUÇÃO	9
2 JUSTIFICATIVA	12
3 OBJETIVO	12
4 MÉTODOS	12
5 RESULTADOS	15
6 DISCUSSÃO	32
7 CONSIDERAÇÕES FINAIS	36
8 REFERÊNCIAS	37

1 INTRODUÇÃO

Com o avanço tecnológico e a crescente disponibilidade de recursos computacionais avançados, o volume de publicações científicas têm apresentado um crescimento exponencial nas últimas décadas (Bornmann et al., 2015). Este crescimento não é apenas quantitativo, mas também se reflete na complexidade dos métodos e técnicas empregados. Paralelamente, o desenvolvimento da tecnologia tem impulsionado uma evolução nas técnicas de análise estatística, tornando-as mais sofisticadas (Horton et al., 2005). Esta evolução gera uma necessidade permanente de atualização constante dos profissionais da área da saúde para se manterem aptos na interpretação de novas publicações científicas (Hellems. et al., 2007; Baldi et al. 2015). A habilidade de interpretar adequadamente estes dados complexos e utilizar as informações de forma eficaz na prática clínica se tornou um componente essencial da medicina baseada em evidências (MBE) (Mayer, 2006).

O desperdício de recursos de pesquisa frequentemente decorre da má estruturação do desenho do estudo, resultando em uma carência de descobertas clinicamente ou cientificamente relevantes e significativas (Glasziou et al., 2009). Esse cenário, muitas vezes, é atribuído à insuficiência na revisão bibliográfica, a métodos de pesquisa inadequados, a falhas na seleção da amostra, entre outros fatores. Tais falhas metodológicas não apenas comprometem a eficácia da pesquisa, mas também levam ao desperdício de esforços, tempo e investimentos consideráveis (Valcaster. et al., 2021). Em um ambiente onde a eficiência e a precisão são cada vez mais valorizadas, a capacidade de desenhar e interpretar estudos de forma correta se torna uma habilidade indispensável (Thornton, 2009).

Em 1994, Douglas Altman já destacava a problemática da inconsistência metodológica e da falta de conhecimento estatístico entre os profissionais e pesquisadores da saúde. Altman defendia que, embora não seja necessário ser um especialista em estatística, um conhecimento mínimo é essencial para se ser capaz de desenhar um estudo de forma apropriada, evitando assim resultados enviesados e o conseqüente desperdício na pesquisa (Altman et al., 1994). Esta observação enfatiza a importância de um domínio fundamental em estatística para a condução de pesquisas confiáveis e éticas na área da saúde.

Apesar da importância desta temática ter sido realçada por Altman em 1994, ela continua sendo extremamente relevante nos dias de hoje, como demonstrado pelos experimentos realizados durante a pandemia da SARS-CoV-2 (Glasziou et al., 2020). Um exemplo notório foi o caso dos estudos sobre a eficácia da Hidroxicloroquina na prevenção e tratamento da COVID-19, os quais, apesar de sugerirem que o medicamento reduzia o risco,

duração e severidade da infecção, eram baseados em pesquisas com falhas metodológicas significativas (Alexander et al., 2020).

Como previamente descrito, a pesquisa na área da saúde enfrenta diversos desafios em relação à qualidade metodológica de suas publicações, seja devido à pressão pelo viés de publicação, que frequentemente leva ao *p-hacking*, ou pela falta de conhecimento estatístico e metodológico dos pesquisadores (Stefan et al, 2023). Contudo, há uma população específica que necessita ter um domínio básico de ambas as habilidades, em epidemiologia e estatística, e que muitas vezes não é focada quando se discutem problemas metodológicos: os profissionais da saúde que não estão envolvidos diretamente com pesquisa. A capacitação destes profissionais em conceitos básicos de bioestatística (BE) e metodologia de pesquisa é crucial para a interpretação adequada de estudos e para a aplicação efetiva de suas conclusões na prática clínica.

Em 1980, Weiss e Samet (1980) foram pioneiros ao questionar a importância do entendimento em metodologia por profissionais clínicos, tendo em vista os diversos problemas descritos. Eles avaliaram o conhecimento de 141 médicos praticantes (incluindo acadêmicos da medicina) com 10 questões simples abordando risco relativo, p-valor, sensibilidade e especificidade. Entretanto, foi somente após a publicação do questionário validado por Windish e colaboradores em 2007 que outros estudos semelhantes foram realizados. Estes estudos utilizaram questionários simplificados e, por serem participantes voluntários, apresentam um viés inerente (Wulff et al., 1987; Scheutz et al., 1988; Ferrill et al., 2000; McNally e Loftus, 2005). Esta limitação é importante, pois há uma alta probabilidade de que apenas aqueles que acreditam ter um desempenho superior respondam aos questionários, um fenômeno apontado por Weiss em seu trabalho (Weiss et al., 1980).

Windish e colaboradores (2007), ao desenvolverem seu questionário, revisaram 239 artigos originais publicados entre janeiro e março de 2005 em seis renomados periódicos médicos (*American Journal of Medicine*, *Annals of Internal Medicine*, *BMJ*, *JAMA*, *Lancet* e *New England Journal of Medicine*). Eles identificaram os métodos estatísticos mais frequentes e, com base nesses achados, desenvolveram questões sobre análises simples (χ^2 , testes t, análise de variância) e análises multivariáveis (ex., regressão proporcional de Cox, regressão logística múltipla). O questionário, denominado *Biostatistical Knowledge Test Survey Instrument* (BKTSI), demonstrou boa validade interna (Cronbach α - 0,81), validade discriminativa e de conteúdo. Ele é composto por quatro blocos: 1) Onze questões sobre aspectos demográficos, treinamento em bioestatística e hábitos de leitura de artigos; 2) Cinco questões sobre a atitude dos participantes em relação à estatística; 3) Quatro questões sobre a

confiança na interpretação de dados e avaliação de conceitos estatísticos; 4) Vinte questões de múltipla escolha sobre bioestatística, sem exigir cálculos, para avaliar o entendimento de métodos estatísticos, desenho de estudo e interpretação de resultados. Esta abordagem metodológica reforça a relevância prática para um questionário de avaliação do conhecimento em bioestatística (Windish et al., 2007).

Apesar de a literatura enfatizar a necessidade de conhecimento em bioestatística e epidemiologia (EP) por parte de profissionais clínicos, bem como o desenvolvimento de instrumentos para a mensuração desse conhecimento, há uma lacuna no que diz respeito à quantificação do conhecimento mínimo necessário. Frequentemente, os estudos descrevem o conhecimento dos participantes como insuficiente, pobre ou inadequado, mas sem estabelecer um limiar claro de adequação. No entanto, é viável utilizar como referência para esta métrica a amostra usada na validação do BKTSI, que consiste em médicos atuantes em medicina interna com treinamento avançado em EP. Esta escolha se justifica devido ao alto rigor metodológico do instrumento e por ser ele um precursor relevante dos estudos analisados (Windish et al., 2007).

Considerando que a equipe com treinamento avançado provavelmente possui uma capacidade interpretativa aprimorada da literatura científica, a média geral de acertos obtida por eles, 94% (variando entre 86,37% e 94,55%), pode ser usada como um padrão-ouro. Este parâmetro oferece uma base quantitativa para avaliar o conhecimento em BE e EP entre os profissionais de saúde, fornecendo um ponto de referência para futuras análises no trabalho. Ademais, é importante ressaltar que o questionário foi cuidadosamente projetado para abordar os conhecimentos fundamentais em bioestatística necessários para acompanhar e interpretar adequadamente a literatura científica contemporânea (Windish et al., 2007).

Para aprofundar nossa compreensão sobre as raízes do problema no ensino de BE e EP, seria essencial examinar e comparar as grades curriculares de diferentes cursos de saúde, bem como entre diversas universidades. Esta análise permitiria uma melhor compreensão das diferenças e lacunas existentes na formação desses profissionais. Embora o artigo de Jayakumar e Umscheid (2007) se foque na expansão dos currículos de bioestatística e epidemiologia nos Estados Unidos, a importância dada a estas disciplinas sugere a existência de desafios semelhantes em contextos educacionais diversos, ressaltando a necessidade de pesquisas mais abrangentes para explorar as variações curriculares em uma perspectiva global.

2 JUSTIFICATIVA

Com o aumento acelerado do número de publicações na área biomédica, assim como a complexidade dos testes estatísticos, é importante que os profissionais saibam interpretar os métodos e identificar possíveis fragilidades de um artigo científico. Ademais, a qualidade das publicações e conclusões podem estar distorcidas por má condução e vieses metodológicos (Carlisle et al., 2017; Ioannidis et al., 2005). A necessidade de maior compreensão de conceitos metodológicos e estatísticos por parte de profissionais da saúde ainda parece ser pouco explorada na literatura, não sendo encontrada nenhuma revisão sistemática sobre o assunto até o presente momento.

3 OBJETIVO

Este trabalho tem como objetivo avaliar conhecimento de bioestatística, metodologia e habilidade interpretativa da literatura biomédica de profissionais da área da saúde.

4 MÉTODOS

Critérios de inclusão:

Foram incluídos estudos que avaliaram de forma transversal o conhecimento em estatística e metodologia de pesquisa de profissionais da saúde, independente do desenho do estudo. Estudos de todas as línguas foram englobados

Critérios de exclusão

Foram considerados inapropriados para análise artigos que não possuíssem uma avaliação objetiva de habilidades dos participantes, estudos que relataram somente autoavaliação de conhecimento e resumos de congresso.

População

Foram considerados indivíduos com ou em formação em qualquer área da saúde como Medicina, Enfermagem, Farmácia, Nutrição, Psicologia, Odontologia e Saúde Pública que estivessem praticando a profissão de formação.

Momento da avaliação transversal

Pela natureza do objetivo da revisão, estudos com intervenção foram incluídos apenas se apresentassem dados pré-intervenção que avaliassem, de maneira objetiva, o

conhecimento dos participantes. A escolha de não utilizar dados pós-intervenção se deve pela inerência da melhora imediata após uma intervenção educativa, distorcendo os resultados desejados pela nossa revisão.

Desfechos

O desfecho analisado é o grau de conhecimento estatístico em relação aos testes aplicados nos estudos incluídos.

Método de pesquisa para identificação de estudos

A estratégia de pesquisa foi elaborada com colaboração de um bibliotecário experiente da Faculdade de Medicina da Universidade Federal do Rio Grande do Sul para a base de dados MEDLINE e posteriormente adaptada para EMBASE e Web of science. Foram incluídos todos os estudos que atendessem os critérios de inclusão, publicados até 29 de outubro de 2023. A estratégia de busca completa se encontra no quadro 1.

Seleção de estudos

Um pesquisador (G.G.G) analisou os resultados da estratégia de busca em duas fases: análise de títulos e resumos e, posteriormente, análise completa dos estudos relevantes. Na primeira fase foi utilizado o Software Rayyan para inclusão e exclusão dos artigos, assim como a remoção de duplicatas.

Quadro 1 - Estratégia de busca da literatura.

MEDLINE via Pubmed: 857 publicações
<p>(Statistics as Topic[mh:noexp] OR Biostatistics[mh] OR Data Interpretation, Statistical[mh] OR Statistic*[ti] OR Biostatistic*[ti] OR Data Analys*[ti] OR Data Interpretation*[ti]) AND (Health Personnel[mh] OR Internship and Residency[mh] OR Health Personnel[tiab] OR Healthcare Provider*[tiab] OR Health care Provider*[tiab] OR Healthcare Worker*[tiab] OR Health Care Professional*[tiab] OR Health care Worker*[tiab] OR Healthcare Professional*[tiab] OR Physician*[tiab] OR Doctor[tiab] OR Doctors[tiab] OR Nurse*[tiab] OR Medical staff*[tiab] OR Hospital staff*[tiab] OR Nursing Staff*[tiab] OR Interns*[tiab] OR Resident*[tiab] OR Residency[tiab] OR Medical student*[tiab] OR Nursing Student*[tiab]) AND (Comprehension[mh] OR Knowledge[mh] OR Professional Competence[mh] OR Health Knowledge, Attitudes, Practice[mh] OR Comprehension[ti] OR Understanding[ti] OR Knowledge[ti])</p>
Embase: 292 publicações
<p>('statistics'/de OR 'biostatistics'/de OR 'statistical analysis'/de OR (Statistic* OR Biostatistic* OR 'Data Analys*' OR 'Data Interpretation*'):ti) AND ('health care personnel'/exp OR 'residency education'/exp OR ('Health Personnel' OR 'Healthcare Provider*' OR 'Health care Provider*' OR 'Healthcare Worker*' OR 'Health Care Professional*' OR 'Health care Worker*' OR 'Healthcare Professional*' OR 'Physician*' OR 'Doctor' OR 'Doctors' OR 'Nurse*' OR 'Medical staff*' OR 'Hospital staff*' OR 'Nursing Staff*' OR 'Interns*' OR 'Resident*' OR 'Residency' OR 'Medical student*' OR 'Nursing Student*'):ti,ab,kw) AND ('comprehension'/exp OR 'knowledge'/exp OR 'professional competence'/exp OR (Comprehension OR Understanding OR Knowledge):ti) AND [embase]/lim NOT ([embase]/lim AND [medline]/lim) AND ('article'/it OR 'article in press'/it OR 'editorial'/it OR 'letter'/it OR 'note'/it OR 'review'/it OR 'short survey'/it)</p>
Web of Science: 135 publicações
<p>TI=(Statistic* OR Biostatistic* OR 'Data Analys*' OR 'Data Interpretation*') AND TS=('Health Personnel' OR 'Healthcare Provider*' OR 'Health care Provider*' OR 'Healthcare Worker*' OR 'Health Care Professional*' OR 'Health care Worker*' OR 'Healthcare Professional*' OR 'Physician*' OR 'Doctor' OR 'Doctors' OR 'Nurse*' OR 'Medical staff*' OR 'Hospital staff*' OR 'Nursing Staff*' OR 'Interns*' OR 'Resident*' OR 'Residency' OR 'Medical student*' OR 'Nursing Student*' OR Pharmacist*) AND TI=(Comprehension OR Understanding OR Knowledge)</p>

5 RESULTADOS

Um total de 1284 estudos foram identificados com as estratégias de busca descritas. Destes, 56 foram removidos como duplicatas, restando 1228 títulos e resumos para a triagem de elegibilidade, dos quais 30 foram elegíveis para a leitura completa. Houve apenas uma exclusão após a leitura na íntegra: um artigo idêntico a outro apresentando apenas autores diferentes. No total, foram incluídos 29 artigos na revisão. O fluxograma PRISMA (2020) de revisão sistemática para busca, triagem e inclusão de artigos está disponível na Figura 1.

Os estudos foram conduzidos nos Estados Unidos da América (N=14), Canadá (N=3), Dinamarca (N=2), Arábia Saudita (N=2), França (N=1), África do Sul (N=1), Grécia (N=1), Irlanda (N=1), Nova Zelândia e Austrália (N=1), Paquistão (N=1), Peru (N=1) e em um conjunto de países europeus (Bulgária, Finlândia, Alemanha, Grécia, Itália, Holanda, Noruega, Suíça, Suécia e Inglaterra) (N=1).

O conjunto dos estudos incluiu um total de 6220 participantes, sendo que as amostras variaram de 11 a 898 indivíduos. As populações dos estudos selecionados foram caracterizadas da seguinte forma:

- Profissionais de Medicina: juntos envolviam 3583 indivíduos, sendo Estudantes (N=641), Residentes de medicina (N=1767) e Médicos Clínicos (N=1175).
- Farmacêuticos: somaram 1540 indivíduos, sendo Residentes de Farmácia (N=943) e Farmacêuticos (N= 597).
- Odontologistas: um total de 476 indivíduos, sendo Residente de Odontologia (N=238) e Odontologistas (N=238).
- Enfermeiros: englobaram 442 indivíduos, os quais eram Estudantes de Enfermagem (N=165) e Enfermeiros Clínicos (N=277) .
- Saúde Pública: 179 Profissionais da Saúde Pública de um único estudo.

Quatro estudos utilizaram o questionário BKTSI (Windish et al., 2007; Alzharani et al., 2015; Sheeran et al., 2016; Araoye et al., 2020), desenvolvido por Windish e colaboradores em 2007. Por compartilharem diversas semelhanças, é possível também considerar nesse conjunto o trabalho de Polychronopoulou. e colaboradores (2011), de Bookstaver e colaboradores (2012), Best e Laskin (2013), Susarla e Redett (2014), Barreto e colaboradores (2020) e Couture e colaboradores (2020). Devido à padronização das questões, estes estudos são os únicos que permitem comparações diretas entre si. O questionário abordava tipo de variáveis, tipo de estudo, teste estatístico básico, p-valor, risco relativo e

razão de chance, teste estatístico complexo, medidas de dispersão, com exceção dos trabalhos de Polychronopoulou e colaboradores (2011), Bookstaver e colaboradores (2012), Best e Laskin (2013), Susarla e Redett (2014), Bougie e colaboradores (2015) e Couture e colaboradores (2020) que fizeram remoções de múltiplas questões. Barreto e Colaboradores (2020) fizeram adição de oito questões ao seu questionário.

O estudo de Windish e colaboradores (2007) envolveu 277 residentes médicos (52% homens) de 11 diferentes programas de residência nos Estados Unidos, incluindo 7 programas tradicionais de medicina interna, 2 de medicina de cuidados primários, 1 de medicina/pediatria e 1 de medicina/medicina preventiva. A taxa média de resposta variou entre os programas, com uma média de 54,05% (variando de 28,1% a 80%). Dos participantes, 68,8% haviam realizado algum curso de BE ou EP, e 85,1% não possuíam pós-graduação. Em média, os participantes atingiram 41,1% (10-90%) nas pontuações, destacando-se no reconhecimento do propósito de estudos duplo cegos (87,4%) e na interpretação de risco relativo (81,6%). Contudo, os menores escores foram na interpretação de análises de Kaplan-Meier (10,5%), intervalos de confiança de 95% e significância estatística (11,9%), além de regressão de risco proporcional de Cox (13%) e na determinação da força da evidência para fatores de risco (17%). Notavelmente, apenas 58,8% da amostra interpretou corretamente o significado de um p-valor maior que 0,05, e 50,2% foram capazes de interpretar o desvio padrão. Entre os participantes, aqueles que haviam concluído pós-graduação obtiveram um aumento absoluto de 9,2% (4,2 a 14%) nas questões corretas, enquanto anos sucessivos após a graduação estiveram associados a uma diminuição de 12,3% (-22,2 a -3,3%) nos escores. Aqueles que fizeram um curso de BE ou PE tiveram um aumento de 4,5% (0,8 a 8,2%) nas pontuações.

Polychronopoulou e colaboradores (2011) avaliaram conhecimentos similares em 127 alunos, com predominância feminina (62,2%), de pós-graduação em odontologia da Europa. A pesquisa abrangeu diversos países, incluindo Bulgária (1,6%), Finlândia (11%), Alemanha (27,6%), Grécia (8,7%), Itália (14,9%), Holanda (9,5%), Noruega (7%), Suíça (5,5%), Suécia (6,3%) e Inglaterra (7,9%). Os autores distribuíram os questionários para 61 diretores de instituições de residência, dos quais apenas 21 aceitaram (34,4%) participar. Ademais, o questionário apesar de ser avaliar tópicos similares ao BKTSI, possui questões focadas em odontologia e 7 questões a menos no bloco de conhecimentos (análises de Kaplan-Meier, determinação da força da evidência para fatores de risco, poder e tamanho da amostra), assim como não apresentou seções relacionadas a atitudes e confiança. Os autores na seção de métodos do trabalho não detalham a metodologia do questionário final, nem

referências que possam ter contribuído. Entre os dentistas, 44,8% possuíam mestrado ou doutorado, e 34,7% já haviam feito algum treinamento em bioestatística ou epidemiologia. De maneira similar ao trabalho de Windish e colaboradores (2007), a média de acertos foi de 43,8% (40,2-47,3%), sendo o melhor desempenho em reconhecimento da razão do cegamento duplo (77,9%), reconhecer a definição de viés (70,8%), interpretar a hipótese nula (70%), reconhecer uma meta-análise (70%). As piores taxas de acerto foram quanto ao reconhecimento de estudo de caso-controle (3,2%), identificar teste do qui-quadrado (11,8%), assim como regressão de risco proporcional de Cox (11,8%). Tal qual o trabalho de Windish e colaboradores, a capacidade de interpretação do p-valor foi baixa (33%). A única associação significativa encontrada foi treinamento prévio em EP ou BE, que foi correlacionada a uma maior taxa de acerto absoluta de 11,8% (4,3-19,4%), mesmo após ajuste para idade, gênero, tempo desde a graduação, pós-graduação e nível de treinamento.

Alzharani e Aba Al Khail (2015) aplicaram o questionário BKTSI em um programa de residência em um hospital de Jidá, Arábia Saudita, complementando-o com quatro questões sobre o envolvimento dos residentes em pesquisa e publicações. Dentre 180 residentes convidados, 162 responderam o questionário (90%). Segundo a descrição dos autores: “Homens compunham mais de dois terços da amostra dos participantes” (76,5%, de acordo com a Tabela 1 do artigo), “que na maioria dos casos tinham 30 anos ou menos” (80,2%, com base na tabela), sendo que 40% haviam concluído algum curso de bioestatística ou epidemiologia. Outras informações demográficas importantes não foram descritas no trabalho. Ademais, há uma discrepância entre a descrição textual dos autores e as informações apresentadas na tabela, na qual é possível extrair dados e calcular as proporções ao: “Menos de um décimo dos médicos liam periódicos” enquanto na tabela foi incluído 154 (95,06%) de participantes que possuíam esse hábito. Tal inconsistência sugere um possível erro de tabulação, especialmente porque durante a discussão, os autores descrevem novamente a baixa taxa de leitura pelos participantes. “Além do mais, médicos não eram muito versados em bioestatística. provavelmente devido a falta de interesse em ler periódicos, como refletido na alta proporção de residentes nesse estudo que não os lêem”.

Os autores não disponibilizam os resultados detalhados por questão, nem uma média geral. Eles afirmam que a maioria dos participantes conseguiu identificar conceitos básicos como P-valor, poder do estudo e estudo de caso-controle. Em contraponto, mais de dois terços dos residentes desconheciam termos mais “sofisticados”, sendo que a caracterização de “sofisticados” foi definida pelos autores, como ANOVA, tamanho de amostra e razão de chance. A análise baseada na idade (>30 ou ≤ 30 anos) não mostrou diferença significativa,

com 24,06% e 23,75%, respectivamente. As características associadas a diferenças significativas de conhecimento incluíam gênero masculino (Teste $t = 2,77$, p -valor= 0,006) com associação positiva, enquanto leitura de periódicos foi associada a um menor conhecimento (Teste $t = -3,34$, p -valor= 0,001). Contrariamente a outros, o estudo encontrou que os homens apresentaram um melhor desempenho no questionário (Teste $t = 2,77$, $p = 0,006$). Possíveis explicações para essas associações são a amostra ter um número pequeno de mulheres (76,5% eram homens) e ser composta por maioria de residentes que não leem periódicos (<10% da amostra estavam expostos a literatura como hábito). Foi conduzido um modelo de regressão linear múltipla para covariáveis preditoras da média dos escores. Dessas, as que tiveram significância estatística foram gênero (feminino vs. masculino: $\beta = -4,75$, $t = -2,70$, p -valor= 0,008), não ter feito curso de bioestatística ($\beta = 7,04$, $t = 2,28$, p -valor= 0,03), não ter feito curso de MBE ($\beta = -13,94$, $t = -4,65$, p -valor < 0,001). O modelo foi capaz de explicar 24% ($r^2 = 0,24$) das variações nos escores. Os pesquisadores descrevem que a relação negativa com ter feito curso de bioestatística se deve pela baixa probabilidade que cursos independentes de BE sejam capazes de retificar as habilidades de profissionais de saúde na área (Alzharani, S. et al, 2015).

Sheeran e colaboradores (2016) investigaram os conhecimentos de residentes de farmácia, aplicando o BKTSI em 501 participantes de um programa de residência ao longo de 4 anos (2009-2013), denominados P1 a P4, respectivamente. A pesquisa ocorreu em 4 momentos: Dezembro de 2009, após a conclusão do curso de estatística no primeiro ano de residência; Dezembro de 2010, após a conclusão da conclusão do curso de informação sobre medicamentos e avaliação literatura no segundo ano de residência; Abril de 2012, seguido da conclusão da sequência de laboratório de terapêutica integrativa no terceiro ano; Abril de 2013, no encerramento da Experiência Prática Avançada em Farmácia. Foi utilizado como incentivo, um ponto extra durante os três primeiros anos e 5 dólares em vale-presente de café no ano final. Os questionários ficaram disponíveis por 3 semanas e um lembrete via email foi enviado. Cada ano foi contabilizado como uma amostra independente, já que diversos alunos não completaram anualmente os questionários. A taxa de resposta >70% em todos os anos, mas apenas 63 residentes responderam a pesquisa em seus quatro momentos. A amostra de P1 foi de 131 alunos, destes 64% do sexo feminino, com idade média de 21,5 anos (19-39), dos quais 94% não haviam feito curso de BE, 85% não haviam feito curso de EP e 98% não haviam feito curso de MBE; O segundo ano (P2) de avaliação contou com 146 residentes, dos quais 68% eram mulheres, mantendo a proporção da ausência de cursos de EP (95%) e MBE (99%), enquanto BE 74% não haviam feito; P3 teve uma amostra de 115 alunos, sendo 70%

mulheres, com uma proporção similar de ausência de cursos prévios (MBE - 96%; BE - 81%; EP - 95%); No último ano de residência (P4) foram avaliados 109 residentes, dos quais 71% eram mulheres, com uma proporção similar quanto aos cursos de EP (94%), BE (83%) e MBE (97%). O Alfa de Cronbach, que mede a validade interna de questionários, para os 4 anos foi, respectivamente, de 0,23, 0,42, 0,46 e 0,57. Valores abaixo do adequado segundo os próprios autores. Os alunos de todos os anos demonstraram saber discernir as diferentes categorias de variáveis (90,8%; 93,2%; 84,4%; 83,5%), bem como identificar estudos duplo-cego (89,3%; 89%; 87,8; 87,2%). Alguns conceitos, no entanto, tiveram um baixo percentual de acerto independente do ano, incluindo interpretar Kaplan-Meier (6,1%; 8,9%; 11,3%; 4,6%), Determinar força da evidência para fatores de risco (9,9%; 10,3%; 14,8%; 9,2%), Interpretar significância estatística e intervalo de confiança (9,2%; 16,4%; 12,2%; 17,4%). As médias gerais foram de 44,5% (15 a 75%), 49,5% (15 a 80%), 45% (10 a 75%) e 47,1 (15 a 75%), para cada um dos quatro anos, respectivamente. Os 63 alunos que responderam aos questionários nos 5 anos apresentaram resultados semelhantes da coorte total.

Araoye e colaboradores (2020) conduziram um estudo similar ao de Windish, focando em residentes de ortopedia de 10 programas de residência dos Estados Unidos. Eles adaptaram o BKTSI para a ortopedia, alterando as questões demográficas e preferências por periódicos específicos, mantendo o bloco de 20 questões de conhecimento inalterados. A pesquisa obteve uma taxa de resposta de 64%, com 178 participantes, sendo 83% homens com idade entre 26 e 30 anos (59%). Uma quantidade considerável de residentes tinha concluído algum curso de BE, EP ou MBE (53%; 44%; 83%), entretanto menos de 5% após a conclusão da graduação. A média dos escores foi de 49,4% (10 a 90%). Os tópicos de maior domínio pelos residentes foram interpretar risco relativo (87,64%), identificar estudos duplo-cego (79,66%), identificar teste t (78,53%), identificar ANOVA (69,66%) e interpretar p-valor (69,1%). Quanto aos tópicos de menor taxa de acerto incluíam a interpretação do Kaplan-Meier (10,17%), interpretar intervalos de confiança (11,8%), interpretação da regressão de risco proporcional de Cox (16,85%), determinação da força da evidência para fatores de risco (17,42%). Entre os fatores demográficos que impactaram o desempenho na ferramenta, ter feito treinamento de BE prévio, mas não MBE ou EP, foi associado com melhor desempenho ($p=0,029$). Entretanto, os autores não realizaram uma análise do tamanho desse efeito. Ademais, atividades como ler publicações de pesquisa, analisar dados para a própria pesquisa e participar de conferências profissionais para desenvolvimento foram ligadas a um melhor desempenho ($p=0,0183$; $p=0,0314$; $p=0,0134$, respectivamente). Os

autores sugerem que é crucial para os programas de residência oferecerem uma introdução à bioestatística e à metodologia de pesquisa, assim como incluir questões relacionadas a esses conhecimentos em exames durante a residência.

Bookstaver e colaboradores (2012) realizaram um estudo com residentes de farmácia dos Estados Unidos. Os autores enviaram e-mails para diretores de programas de residência acreditados pela *American Society of Health-System Pharmacists*, solicitando que distribuíssem para os residentes de primeiro ano (PGY1) a pesquisa online para avaliar o conhecimento em bioestatística. O instrumento de pesquisa foi baseado no BKTSI de Windish e colaboradores (2007), mas diferia ao incluir 12 questões em vez de 20. Foi estimado que 1400 PGY1 tenham recebido o questionário, dos quais 166 (taxa de resposta de 11,8%) responderam ao menos uma pergunta do bloco de conhecimentos. A amostra foi predominantemente feminina (73,5%) e a maioria na faixa etária de 21 e 30 anos (82,2%). Dessas, 86,8% relataram ter tido treinamento prévio em bioestatística. O teste de consistência interna resultou em um alfa de Cronbach de 0,42. Não foi observada diferença significativa entre aqueles com ou sem treinamento em BE. A média geral foi de 47,3%, com maiores taxas de acerto no reconhecimento do propósito de um estudo duplo cego (92,6%), interpretação de risco relativo (75,8%) e identificação de variável nominal (69,4%). Quanto às menores taxas de acerto, destacam-se a identificação de regressão de risco proporcional de Cox (24,8%) e a definição de viés (23,9%). Menos da metade dos participantes foi capaz de interpretar desvio padrão (44,9%). Os autores não identificaram associações significativas com características demográficas.

Best e Laskin (2013) analisaram o conhecimento bioestatístico de residentes de cirurgia oral e maxilofacial dos Estados Unidos. Utilizaram os três primeiros blocos do BKTSI de forma íntegra e selecionaram 6 questões do bloco de conhecimento, abrangendo os diferentes tipos de variáveis, reconhecimento de caso controle, propósito de um estudo duplo-cego e interpretação do p-valor. Para o recrutar participantes, contataram, via e-mail, os diretores de todos os 106 programas de residência da especialidade, solicitando que encaminhassem o questionário aos seus residentes. Dos 990 residentes em todos os programas, 112 responderam à ferramenta. A amostra foi majoritariamente masculina (82%), com a maioria dos participantes na faixa etária entre 26 e 30 anos, 53% haviam concluído algum curso em epidemiologia, 49% em bioestatística e 65% em odontologia baseada em evidência. A média geral foi de 38%, sendo que a única questão que a maioria acertou foi o propósito de um estudo duplo cego (78%).

Susarla e Colaboradores (2014) conduziram uma pesquisa com residentes da cirurgia plástica na *John Hopkins University*. Para avaliar os conhecimentos dos residentes, utilizaram um questionário com 6 itens: Desenho de estudo, p-valor, intervalo de confiança, testes estatísticos e motivo de cegamento. De 24 residentes registrados, 22 completaram o questionário (91,7% de taxa de resposta). Desses, 45,5% haviam feito alguma disciplina de bioestatística, 59,1% epidemiologia e 45,5% de MBE. Cerca de 91% dos residentes relataram ler pelo menos 1 periódico de cirurgia plástica por mês, com uma média geral de leitura de 2,1 ($\pm 1,5$). A média de acertos dos residentes foi de 53% ($\pm 24,5$), com as maiores taxas de acerto em razão do cegamento (95,5%) e interpretação do p-valor (77,3%). As piores taxas foram na identificar a função da regressão de risco proporcional de Cox (9,1%), na interpretação do intervalo de confiança, função da ANOVA e desenho de estudo (45,5%). Educação formal em bioestatística foi significativamente associada a uma maior taxa de acerto ($p = 0,044$). A quantidade de periódicos lidos mensalmente teve uma correlação moderada com a taxa de acerto ($r = 0,61$, $p = 0,003$). Os autores concluem que é importante que os profissionais não apenas aprendam bioestatística por meio da educação formal, mas também se empenhem ativamente em manter uma análise crítica da literatura. Ademais, eles sugerem a inclusão de bioestatística como um componente padrão dos exames de serviço, exames de certificação de conselho e exames de manutenção de certificação.

Bougie e colaboradores (2015) convidaram todos os residentes canadenses de obstetria e ginecologia (PGY-1 a PGY-5) para completar um questionário anônimo durante o exame nacional de treinamento. O objetivo era avaliar as habilidades de avaliação crítica dos residentes. Eles utilizaram um questionário composto por 15 questões que abordaram p-valor, sensibilidade e especificidade, tipos de variáveis, desenho de estudo, viés, métodos estatísticos simples e análises multivariadas. Dos 355 que realizaram o exame nacional, 301 consentiram em participar da pesquisa (84,8% de taxa de resposta). Entre eles, 39% já haviam publicado em periódico clínico e 51% já haviam feito algum curso em epidemiologia ou bioestatística. A média dos residentes foi de 69,8% (± 15), e as maiores taxas de acerto foram em razão de chances (92,7%), p-valor (88,7%) e curvas de Kaplan-Meier (92,7%), o que difere dos resultados de outros estudos. Em contraponto, os residentes tiveram maior dificuldade em identificar o teste ANOVA (36,9%) e determinação da força da evidência para fatores de risco, poder e tamanho da amostra (42%).

Barreto e Colaboradores (2020) conduziram um estudo multicêntrico em quatro centros da Mayo Clinic Midwest durante 12 meses (julho de 2016 a junho de 2017) com residentes. Teve como objetivo administrar um curso online para os residentes a fim de avaliar

a melhora do conhecimento antes e depois da intervenção. O instrumento de avaliação era similar ao BKTSI, porém contando com 5 questões a mais e outras modificadas, totalizando 28 questões. De 15 residentes elegíveis, 8 completaram o questionário (taxa de resposta de 53,33%). Dados demográficos não foram disponibilizados no trabalho. A média dos participantes pré-intervenção foi de 15 ($\pm 2,5$) de no máximo 28.

Couture e colaboradores (2020) realizaram um estudo para avaliar o conhecimento de residentes de urologia no Canadá. O questionário continha 10 questões com conhecimento de conceitos estatísticos. O questionário foi enviado para 13 programas de urologia do país, com um total de 201 residentes. Dos 201, 68 (taxa de resposta de 33,83%) responderam o questionário completo. A mediana de idade foi de 28 anos, com 61,8% dos residentes tendo completado algum curso de epidemiologia, bioestatística ou MBE. 94,1% já publicaram ao menos uma vez e 66,2% já fizeram análise de dados. A média geral foi de 65% ($\pm 20,7\%$), sendo as maiores taxas de acerto em identificar variáveis contínuas (89,7%), propósito de cegamento (85,3%) e definir erro tipo I (70,6%). As menores taxas de acerto foram em interpretar razão de chances (50%), identificar variáveis ordinais (51,5%) e discretas (52,9%). Os autores não encontraram nenhuma associação com dados demográficos e escore.

Weiss e Samet (1980) foram um dos primeiros a investigar o conhecimento bioestatístico de profissionais da saúde. Eles convidaram uma equipe de medicina interna de um hospital-escola dos Estados Unidos. Dos 229 profissionais convidados, 141 responderam de forma satisfatória (61,57%), incluindo acadêmicos de tempo-integral (N=62), médicos em exercício (N=117) e Residentes (N=50). Para a elaborar o questionário, os autores analisaram os periódicos *American Journal of Medicine*, *Annals of Internal Medicine*, *British Medical Journal*, *Journal of the American Medical Association*, *Lancet* e *New England Journal of Medicine* para encontrar a frequência dos desenhos de estudos e técnicas estatísticas. A busca foi feita em 1977 durante um período de seis meses. O questionário possuía 10 perguntas avaliando conhecimentos sobre teste de diagnóstico, p-valor, desvio padrão, teste de hipótese, vieses, risco relativo e desenho de estudo, além de perguntas demográficas. Após testar a ferramenta com professores e alunos de pós-graduação em epidemiologia, que obtiveram média de 9,1 ($\pm 1,3$), a média geral dos participantes foi 7,4 ($\pm 1,6$). Os autores observaram que treinamento prévio em bioestatística (36%), epidemiologia (24%) ou ambos (16%) estava associado a pontuações maiores. Ressaltaram a possibilidade de viés nas médias devido à não resposta de alguns participantes, principalmente por receberem feedback de não entenderem as questões e receio de terem um desempenho ruim. Os autores concluem que médicos podem não estar adequadamente preparados para avaliar a literatura biomédica (Weiss et al., 1980).

Wulff e colaboradores (1987) selecionaram 250 médicos de maneira aleatória do registro de médicos dinamarqueses e enviaram por carta um questionário de 9 questões. Os autores não detalharam como desenvolveram a ferramenta. Dos 250 médicos, 148 (59,2%) responderam o questionário, entre eles tinham clínicos gerais (24%), especialistas (18%), residentes (50%) e outros (8%). Além destes, 97 participantes de dois cursos introdutórios de métodos de pesquisa clínica e bioestatística completaram o questionário. A mediana dos médicos dinamarqueses foi de 2,4, sendo aqueles formados há mais de 15 anos com uma média de 2,1 ($p=0,025$). Dentre os participantes dos cursos, a mediana foi de 4. Ambos os grupos tiveram dificuldades nos seguintes tópicos, entre parênteses a taxa de acerto dos médicos e dos participantes: Desvio Padrão (8% e 29%), Erro Padrão (38% e 27%), P-valor (13% e 39%), comparação de p-valor entre estudos (20% e 14%) e correlação (16% e 36%).

De forma similar, Wulff e colaboradores (1988) avaliaram o conhecimento de odontologistas dinamarqueses, utilizando a mesma metodologia de seu estudo anterior. Eles selecionaram aleatoriamente 250 profissionais registrados e enviaram um questionário de 9 questões. Dos 125 dentistas que responderam (Taxa de respostas= 50%), 64% eram clínicos gerais, 32% odontopediatras, 2% membros hospitalares e 2% acadêmicos de odontologia. O teste foi entregue para 38 alunos de odontologia, sem aviso prévio, durante uma disciplina de estatística, tendo uma taxa de participação de 71% ($N=21$). A mediana para os profissionais foi de 2,2, enquanto os estudantes tiveram uma mediana superior de 3,4 ($p=0,0004$). Ambos os grupos apresentaram dificuldades semelhantes nas seguintes áreas (taxa de acerto dos profissionais e dos estudantes, respectivamente): desvio padrão (7% e 4%), erro padrão (26% e 41%), p-valor (8% e 33%), interpretação de p-valores (26% e 22%) e correlação (6% e 26%). Essas dificuldades são consistentes com as encontradas no estudo com médicos, indicando a possibilidade de um problema estrutural no ensino de estatística na área da saúde.

Ferril e Colaboradores (2000) elaboraram um estudo de caso de uma pesquisa clínica com 5 questões de múltipla escolha a respeito de sua metodologia (objetivo do estudo, incidência de efeito adverso, tamanho amostral, falha metodológica e comparação entre grupos) para farmacêuticos. Eles enviaram o questionário e instruções, por meio de duas cartas com 4 semanas de intervalo, para 2500 farmacêuticos aleatoriamente selecionados através de uma lista de inscrição do periódico *Hospital Pharmacy*. Dessas cartas, 81 não foram entregues, e 545 questionários válidos foram retornados (Taxa de resposta de 22,5%). Dos 545 participantes, 55,2% eram bacharéis, 29% doutores e 16% outros. Em relação à atuação primária deles, 33,9% eram farmacêuticos de equipe, 23,5% clínico/consultores, 22,6% diretores de farmácia e 20% outros. A média dos participantes foi de 2,2 ($\pm 1,1$), com 93,4%

respondendo corretamente ao objetivo do estudo, enquanto apenas 18,3% identificaram a falha metodológica do caso clínico. Para identificar fatores associados a um melhor desempenho, foi conduzida uma regressão *stepwise* retrógrada, com um $r^2=0,106$, no modelo final. Dentre as variáveis analisadas, apenas duas tiveram permaneceram no modelo final: atuação primária e tempo desde o último diploma. Consultor/Farmacêutico Clínico obtiveram uma média de 2,66 enquanto a média de outras profissões foi de 2,06 ($F(1,497) = 12,05$, $p<0,01$). Quanto ao tempo desde a obtenção do último diploma foram estabelecidas 5 categorias (média): 0-5 anos (2,77), 6-10 anos (2,48), 11-15 anos (2,30), 16-20 anos (2,17), 21 ou mais anos (1,80). Independente das associações, os autores classificaram o conhecimento dos farmacêuticos como insuficientes para a interpretação adequada da literatura.

McNally e Loftus (2005) distribuíram para todos os especialistas pediátricos registados e consultores da Irlanda um questionário para avaliar os conhecimentos em estatística. O questionário foi dividido em dois blocos: um com informações demográficas e outro com 10 questões sobre estatística básica, aplicação de testes estatísticos e regressão. Os autores não apresentam o questionário na publicação, mencionando apenas que se basearam em livros-texto básicos em estatística médica. Dos 137 questionários enviados, 62 foram completados (taxa de resposta de 45,6%). A média geral foi de 5,1, porém ao analisar grupos distintos, os especialistas tiveram uma média de 5, pediatras gerais de 4,4 e sub-especialistas pediátricos de 6,6, embora a origem dos sub-especialistas não tenha sido explicada. As questões de estatística básica tiveram 67% de acerto, enquanto as aplicações de testes tiveram 38%. Nas três questões de regressão, apenas 25,8% das respostas estavam corretas. Não foi realizada uma avaliação das associações entre características demográficas e conhecimento estatístico.

Baghi e colaboradores (2013) conduziram uma pesquisa com enfermeiras em um curso de mestrado quanto aos conhecimentos estatísticos aplicados a práticas baseadas em evidência, antes e após 10 semanas de um curso introdutório de estatística. Os participantes foram 165 estudantes de mestrado, todos enfermeiros. Os autores não disponibilizaram dados demográficos. O teste de proficiência continha 30 questões, dividido em 5 áreas: estatística descritiva, análise de correlação, teste de hipótese, análise de regressão e testes estatísticos para diferenças entre grupos. Os autores disponibilizaram as médias por áreas, entretanto a soma das médias representa melhor a pontuação total no questionário, cujo valor pré-intervenção foi de 12,1 (40,30%) e pós-intervenção de 22,59 (62,13%).

Anderson e colaboradores (2014) convidaram 200 obstetras-ginecologistas (ob-gin) para participar de sua pesquisa sobre alfabetização estatística. Eles utilizaram uma

amostragem quasi-aleatória de membros do *Collaborative Ambulatory Reserve* de ob-gin membros do *American College of Obstetricians and Gynecologists*, enviando a eles o questionário junto com as instruções. Dos convidados, 94 ob-gin responderam os questionários (47% de taxa de resposta), com distribuição similar de gênero (apenas descrita sem dados apresentados) e uma idade média de 52 anos (± 10). A ferramenta consistia em 3 blocos: Perguntas demográficas, escalas de habilidade numérica de Schwartz e Lipkus e o *Obstetrician-Gynecologist Statistical Literacy Questionnaire* (OGSLQ). A estatística descritiva da amostra e as questões demográficas não foram apresentadas no estudo. As escalas de habilidade numérica avaliaram a capacidade de converter percentual em prevalência, e vice-versa, além de probabilidade básica. Já o OGSLQ, com 18 questões, avaliou 4 categorias distintas de conhecimento: 9 questões sobre fatos numéricos (questões de fatos) associados a medicina (e.g, incidência de câncer de mama nos EUA), 4 questões avaliando conhecimentos estatísticos (questões de conceitos) como incidência, 5 questões avaliando relações numéricas (questões de relações) como risco relativo e absoluto. As questões de fatos obtiveram um alfa de Cronbach de 0,01, enquanto as questões de conceitos e relações tiveram valores de 0,55 e 0,53, respectivamente. Devido à baixa consistência interna das questões de fatos, os autores combinaram as questões de conceitos e relações (conceito-relação), resultando em um alfa de cronbach de 0,61. Nenhuma associação demográfica foi encontrada com questões de fatos. As questões de conceito, por sua vez, apresentaram uma correlação negativa com idade ($r = -0,25$, $IC_{95\%} = -0,46$ a $-0,06$), enquanto questões de relações não mostraram associações significativas. Os autores não forneceram a média ou a mediana dos escores diretamente, entretanto a partir do gráfico de distribuição de acertos, foi possível calcular a média. Para isso foi utilizada a seguinte fórmula: $\Sigma(\% \text{ de respondentes} * \text{escore do \% de respondentes}) / n$
 $(0*0+0*1+0*2+3*1+4*1+5*1+6*1+7*3+8*7+9*12+17*10+19*11+19*12+12*13+14*5+15) / 94 = 11,18$ (em uma escala de 0 a 18 acertos). Os autores concluíram que são necessários mais estudos sobre a alfabetização estatística de profissionais da saúde, mas que esforços para melhorar esse conhecimento seriam benéficos.

Msaouel e colaboradores (2014) realizaram seu estudo em oito hospitais públicos da Grécia, recrutando 200 residentes de medicina, 25 de cada hospital, de forma aleatória através de um software. Dos 200, 153 residentes (taxa de resposta 76,5%) participaram, predominantemente homens (73,2%), com uma idade mediana de 32 anos, dentre os quais 81,7% haviam feito algum curso de bioestatística anteriormente. O questionário incluía 7 questões demográficas e 7 questões de conhecimento, abrangendo conceitos como desvio

padrão, erro padrão, p-valor, risco relativo, intervalo de confiança, coeficientes de correlação, sensibilidade e valor preditivo positivo (VPP). Adicionalmente, foi desenvolvido um teste de vieses cognitivos, também com 7 questões. O alfa de Cronbach foi de 0,648 para as questões de estatística, considerado como adequado pelos autores. A média dos escores foi de 2,1 ($\pm 1,1$). As questões com as menores taxas de acerto foram sobre erro padrão (22,9%) e p-valor (26,8%). Na análise bivariada, associações significativas foram encontradas entre o conhecimento bioestatístico e idade ($r=-0,166$, $p=0,041$), bem como ano de residência ($r=-0,205$, $p=0,011$). Os autores concluem que, para desenvolver e manter as habilidades adequadamente e raciocínio estatístico, os educadores precisam reavaliar e sistematizar o ensino da bioestatística. Eles enfatizam a necessidade de melhorar a abordagem educacional na área, realçando a importância de uma estratégia de ensino mais eficaz diante das dificuldades identificadas no estudo.

Chima e colaboradores (2015) avaliaram o impacto de um curso rápido de bioestatística no conhecimento de pesquisadores de uma universidade sul-africana. Participaram do curso 40 pessoas, das quais 34 (85%) responderam o questionário de conhecimento basal. Este questionário, contendo 7 questões sobre, foi aplicado antes do curso e novamente dois meses após sua conclusão. Dos 34 participantes, 65% eram mulheres e 55% eram pesquisadores qualificados, incluindo supervisores de pesquisa. A mediana do escore inicial foi de 0% (53% dos respondentes) e o percentil 75 de 28,6%. O questionário de seguimento foi concluído por apenas 6 participantes, os quais tiveram uma mediana de 28,5% e um percentil 75 de 85,7%. Os autores consideraram que o curso teve uma tendência de melhoria no entendimento bioestatístico. Apesar de não ser o objetivo desta revisão avaliar intervenções, é importante ressaltar que existe uma alta probabilidade de viés nos resultados da intervenção, por apenas participantes que viram melhora responderem o questionário de seguimento.

Hazelton e colaboradores (2016) conduziram uma pesquisa sobre os conhecimentos de termos absolutos e relativos de docentes. Enviaram um convite por e-mail, contendo um link para o questionário, a 1608 docentes de medicina da Universidade de Dalhousie, no Canadá. O questionário foi dividido em duas partes: uma de autoavaliação de conhecimentos e outra de teste de conhecimentos, incluindo 4 questões sobre número necessário para tratar (NNT), redução de risco absoluto e relativo, razão de chances e razões de risco. Dos 1608, 221 responderam (14% de taxa de resposta), entretanto apenas 137 completaram (71%) as duas etapas do questionário, e foram incluídos no estudo. A amostra foi majoritariamente masculina (62%), com uma média de 16,8 anos ($\pm 10,4$) de experiência em docência. Cerca de

34% dos participantes haviam completado algum curso em MBE nos últimos 10 anos. Os autores não forneceram a média ou a mediana das pontuações, mas relataram que 26% acertaram as quatro questões e 18% não acertaram nenhuma. Eles concluem que alguns docentes se beneficiariam de suporte adicional e recurso para aprimorar o conhecimento de estatística, assim como mentoramente em MBE, com enfoque em fortalecer o pensamento crítico na educação médica.

Shafi e colaboradores (2018) realizaram um trabalho no Centro Regional de Islamabad, na Faculdade de Médicos e Cirurgiões do Paquistão, para avaliar o conhecimento de bioestatística de estudantes de pós-graduação, antes e depois de uma oficina de metodologia. Foram registrados 270 alunos de pós-graduação durante uma oficina sobre metodologia de pesquisa e bioestatística. A ferramenta continha 21 questões relativas ao entendimento dos participantes em aplicar conhecimento estatístico no desenho de pesquisa e nas aplicações de métodos adequados. Os autores não disponibilizaram o questionário completo, apenas as perguntas e suas respectivas médias de acerto. Todos os 270 participaram, sendo eles majoritariamente mulheres (70%) com uma idade média de 28 anos. A média de acertos pré-intervenção foi de 62,9%. Os tópicos com as maiores taxas de acerto foram medidas de dispersão (86,7%), variáveis (91,9%) e exemplos de dados qualitativos (92,2%). As menores taxas de acerto foram relacionadas à identificação de quais testes de significância devem ser aplicados em diferentes circunstâncias (27%, 30,7% e 31,9%). A média após intervenção aumentou para 83,6%.

Mas e colaboradores (2018) investigaram habilidades numéricas e gráficas de estudantes de uma universidade privada de medicina do Peru. Os autores utilizaram a escala de Lipkus, Escala de Numeração Objetiva (ENO), a Escala de Numeração Subjetiva (ENS) e a Escala de Literacia Gráfica (ELG), contendo 9, 7 e 13 questões, respectivamente. De 304 alunos e residentes selecionados por amostragem de conveniência, 169 completaram a pesquisa, sendo 52% alunos do sexto ano, 18,34% do sétimo ano e 29,66% residentes. A amostra era de ambos os sexos (53% homens), com a maioria possuindo algum treinamento em metodologia (58%) e domínio da língua inglesa (65,1%). A média na ENO foi de 7,34 (\pm 1,51) de um máximo de 9, enquanto a ENS teve uma média de 34,14 (\pm 4.87) de um máximo de 42 e, por fim, a ELG teve uma média de 10,35 (\pm 1,93), de um máximo de 13. A única escala que os autores consideraram ter uma boa consistência interna foi o ENS com um alfa de Cronbach de 0,81 enquanto a ENO e a ELG tiveram 0,61 e 0,66, respectivamente. Eles realizaram uma regressão linear múltipla com as variáveis independentes sendo idade, gênero e treinamento em epidemiologia, e as dependentes sendo as escalas. Nesta análise, a ENS e

ENO mostraram associação significativa com gênero e treinamento prévio, enquanto ELG mostrou associação significativa com gênero e idade. Entretanto, os coeficientes beta não foram relevantes, variando de -0,25 a 0,38.

Tam e colaboradores (2018) conduziram uma pesquisa em médicos da Nova Zelândia e Austrália para entender como os mesmos conceituam o p-valor. Foram formuladas 2 perguntas: uma aberta, para explicarem o conceito de $p=0,05$ em um determinado cenário, e outra questão objetiva dicotômica quanto a interpretação incorreta do p-valor. Os autores disponibilizaram o link para o questionário em um grupo do Facebook para médicos atuantes na Austrália e Nova Zelândia. O grupo é constituído por aproximadamente 4000 membros, dos quais 272 iniciaram o questionário e 247 finalizaram (taxa de conclusão de 91%). A amostra foi composta principalmente por mulheres (66%), com uma idade média 39,6 anos e tempo médio de atuação de 9,7 anos. As respostas descritivas foram classificadas em seis categorias: probabilidade de acontecer no mundo real, raciocínio de limiar, significância estatística, conceituação dentro do contexto do teste estatístico de hipótese nula, sem interpretação relevante e qualidade de estudo. Dessas categorias, os autores consideraram a conceituação em relação a hipótese nula, a qual apenas 4% responderam, como a mais próxima da realidade. A interpretação mais comum foi a probabilidade de acontecer no mundo real (54%). A questão objetiva afirmava: “A partir de $p=0,05$, um leitor faz a seguinte interpretação: Isso significa que há uma probabilidade de 5% de que este resultado seja devido apenas ao acaso, ou, há uma probabilidade de 95% de que a conclusão seja verdadeira?”. Pedia-se então que o leitor indicasse se a interpretação estava majoritariamente errada ou certa. A resposta considerada correta pelos autores era a afirmativa falsa, a qual 24% responderam. Nenhuma característica demográfica teve associação significativa com as respostas.

Hayat e colaboradores (2021) realizaram uma pesquisa avaliando os conhecimentos bioestáticos de docentes universitários em cinco áreas da saúde. A amostra foi obtida por meio de amostragem probabilística, e instituições acreditadas foram selecionadas por amostragem estratificada aleatória, estratificando por curso entre 537 faculdades. A amostra consistiu de 708 docentes dos cursos de Saúde Pública (179 participantes), Medicina (117 participantes), Odontologia (109 participantes), Farmácia (139 participantes) e Enfermagem (164 participantes) de 102 universidades. A taxa de resposta variou entre os cursos, sendo de 8,6% na saúde pública, 4,2% em odontologia, 5,5% em medicina, 7,7% em enfermagem e 6,8% em farmácia, resultando em uma taxa geral de 6,5% (708/10931 docentes). A composição da amostra foi predominantemente feminina em enfermagem (90,2%) e farmácia

(64%), enquanto medicina (61,5%) e odontologia (63,3%) foram masculinos. Saúde Pública apresentava proporção similar entre homens e mulheres (48% mulheres). A maioria dos docentes de saúde pública (75,4%) e enfermagem (70,7%) reportou ter concluído três ou mais cursos de bioestatística. A maior parcela dos participantes de todas as disciplinas, com exceção de saúde pública (49,2%), referem não ter feito um curso de epidemiologia ou ter feito apenas um. O instrumento incluía itens demográficos e oito questões quanto ao conhecimento estatístico abrangendo randomização, estudos observacionais, poder estatístico, intervalos de confiança, testes múltiplos, erro padrão, resultados de regressão e razões de chances. A média geral de acertos foi de 66,2%. Os docentes de saúde pública tiveram o melhor desempenho com 80,7%, seguida por farmácia com 63,8%, comparativamente com enfermagem com 63,6% e medicina com 62,7%, e odontologia com a menor pontuação de 53,3%. Os menores escores foram em torno de 50% para randomização, intervalo de confiança e razão de chances.

Os autores realizaram um modelo geral linear para modelar o número de acertos em função da disciplina e características dos docentes. O modelo é responsável por explicar 40,8% da variabilidade de acertos ($r^2=0,408$). Em odontologia, observou-se 1,07 (-0,86 a -1,83) respostas corretas a menos em relação à saúde pública que foi a profissão de referência para a comparação. Professores que lecionavam estatística estavam ligados a uma melhora no escore em 0,91 (0,60 a 1,21) questões certas a mais. Por outro lado, o tempo de profissão (formação) foi inversamente relacionado a conhecimento estatístico; a cada 10 anos adicionais de experiência, o número médio de respostas diminuiu em 0,30 (-0,44 a -0,16). Em contraponto, o tempo de docência melhorou o escore em 0,27 (0,12 a 0,42) a cada 10 anos lecionando. Ter completado três ou mais cursos de bioestatística foi associado a uma melhora em 1,21 (0,93 a 1,50) questões corretas em relação a ter menos de três cursos. Em epidemiologia, a melhora foi mais modesta com 0,54 (0,20 a 0,89) questões (Hayat et al., 2021).

Os pesquisadores reconheceram a possibilidade viés de seleção, potencialmente superestimando os resultados devido à probabilidade de não-respondentes serem menos envolvidos com o tópico. Profissionais da saúde pública, provavelmente, tiveram um desempenho superior devido a *Association of Schools and Programs of Public Health* fornecer orientações para escolas e programas com um conjunto recomendado de diretrizes de competência em bioestatística. Dois estudos dentro do escopo desta pesquisa fazem parte desse estudo maior de Hayat e Colaboradores (2021), publicados individualmente. Hayat e colaboradores (2021) avaliaram, com o mesmo questionário, conhecimento de enfermeiras

com doutorados e Jiroutek e colaboradores (2019) investigaram os conhecimentos de farmacêuticos. Os resultados de ambos estão inclusos nesse trabalho maior.

Aloitabi e colaboradores (2022) realizaram um estudo com residentes num programa de medicina de família na Arábia Saudita para avaliar os conhecimentos estatísticos. O questionário era composto por 15 questões com opções de “verdadeiro”, “falso” ou “não sei”. Os autores categorizam o escore de 7 ou menos como conhecimento inadequado de bioestatística. Ao todo, participaram do estudo 113 residentes, com 54,9% de homens e uma idade média de 28,1 anos. A média geral foi de 34%, os participantes tiveram as menores taxas de acerto em identificar Kaplan-Meier (15,9%), regressão múltipla para análise ajustada (17,7%), reconhecer necessidade de uma regressão simples em análise não ajustada (22,1%) e identificar a interpretação de p-valor (26,5%). Apesar de não comparável, os resultados foram similares aos de Alzharani e Al-Khail (2015).

Lakhlifi e colaboradores (2023) conduziram uma pesquisa sobre metacognição em médicos, abrangendo clínicos, residentes e estudantes. Segundo os autores, a metacognição (definida como a habilidade de estimar, refletir e monitorar o próprio pensamento e processo decisório) pode atuar na mitigação de erros de raciocínio estatístico. A avaliação de metacognição comumente se dá por meio da análise de confiança, definida como o grau de certeza que um indivíduo atribui às suas decisões. Neste contexto, a 'calibração' relaciona-se à correspondência entre a confiança do indivíduo e a precisão efetiva de suas decisões. Uma calibração adequada significa que a confiança do clínico reflete com precisão a correteza de suas decisões; por outro lado, uma má calibração, como o excesso de confiança, indica uma discrepância entre a confiança e a precisão real, levando potencialmente a decisões clínicas menos informadas ou subótimas. O questionário consistiu em duas partes: na primeira, os participantes responderam a três exercícios com seis afirmações cada, abordando temas como eficácia de vacinas, p-valor e interpretação de testes (VPP, sensibilidade, especificidade e prevalência), utilizando uma escala de -100 a 100 para expressar sua confiança nas respostas, assim como questões teóricas quanto especificidade e sensibilidade, além de cálculo a partir de frequência natural ou probabilidade condicional. A segunda parte visava avaliar o impacto do uso de probabilidade condicional ou frequências naturais na primeira fase. Dos 898 médicos que participaram, constatou-se um aumento proporcional da confiança com o número de respostas corretas.

De maneira específica, a confiança média ultrapassou 50/100 para aqueles que acertaram pelo menos três das doze questões. Quando os médicos acertaram sete ou mais questões, a confiança média atingiu 80/100, indicando uma descalibração caracterizada por

um excesso de confiança à medida que o desempenho melhorava. Em relação às questões específicas, a taxa de acertos sobre a eficácia da vacina variaram entre 37,7% e 91% corretas, enquanto as respostas sobre o p-valor variaram de 28,1% a 71,4% corretas. Observou-se que a mediana de confiança para as respostas corretas foi de 78,5%, em contraste com 66,7% para as respostas incorretas, indicando um excesso de confiança. No exercício de interpretação de resultados, a maioria dos participantes identificou corretamente as definições de sensibilidade (80,4%), especificidade (66,8%), VPP (75,7%) e VPN (73,9%). A confiança dos participantes nas suas respostas corretas foi alta, com uma mediana de 100, mas apresentou maior variação quando nas respostas incorretas. Na tarefa prática de cálculo de VPP, apenas 15% dos participantes chegaram ao resultado correto, mantendo uma alta confiança em suas respostas com uma mediana de 80. Estes achados ressaltam uma tendência à descalibração, inclinada ao excesso de confiança, e destacam a influência da forma como o problema é apresentado na precisão das respostas, com taxas de acerto superiores quando as probabilidades foram apresentadas como frequências naturais em vez de condicionais.

6 DISCUSSÃO

Esta revisão lança luz sobre as deficiências no conhecimento em metodologia e bioestatística de profissionais da saúde, independente de sua área de atuação ou nível de experiência, de estudantes a docentes (Mas et al., 2018 e Hayat et al., 2021). Quando avaliamos os trabalhos que usaram questionários similares (BKTSI), apresentaram médias de acerto variando de 23,75% a 65%, com a maior parte abaixo de 50%, o que reforça o que vêm sendo demonstrado na literatura há décadas, com exceção de residentes canadenses que parecem performar acima da média (Couture et al., 2020 e Bougie et al., 2015). Weiss e Samet, em 1980, já apontavam para essa lacuna de conhecimento, anterior aos grandes avanços computacionais e testes estatísticos mais complexos. Curiosamente, o aumento na sofisticação das análises estatísticas, como as multivariadas, não resultou em um melhor entendimento das técnicas básicas (Aloitabi et al., 2022; Araoye et al., 2020; Hayat et al., 2021; Hayat et al., 2021; Jiroutek et al., 2019; Msaouel et al., 2014; Polychronopoulou et al., 2011; Sheeran et al., 2016; Susarla et al., 2014; Weiss et al., 1980; Windish et al., 2007; Wulff et al., 1988; Wulff et al., 1988).

Quando olhamos os tópicos com os menores escores, frequentemente encontrados nos trabalhos analisados, incluem intervalo de confiança, p-valor, desvio padrão, razão de chances, e erro padrão, todos tópicos fundamentais para interpretação de dados (Aloitabi et al., 2022; Araoye et al., 2020; Hayat et al., 2021; Hayat et al., 2021; Jiroutek et al., 2019; Msaouel et al., 2014; Polychronopoulou et al., 2011; Sheeran et al., 2016; Susarla et al., 2014; Weiss et al., 1980; Windish et al., 2007; Wulff et al., 1987; Wulff et al., 1988). Ademais, análises mais complexas, como na análise de sobrevivência de Kaplan-Meier e regressão de risco proporcional de Cox, também apresentaram altas frequências de baixo desempenho (Aloitabi et al., 2022; Araoye et al., 2020; Sheeran et al., 2016; Windish et al., 2007). Em contraponto, os tópicos que frequentemente tiveram maior taxa de acerto foram risco relativo, propósito do cegamento, identificar algum tipo de variável e identificar estudos observacionais (Aloitabi et al., 2022; Araoye et al., 2020; Best e Lakin, 2013; Bookstaver et al., 2012; Bougie et al., 2015; Couture et al., 2020; Hayat et al., 2021; Hayat et al., 2021; Jiroutek et al., 2019; Polychronopoulou et al., 2011; Shafi et al., 2018; Sheeran et al., 2016; Weiss et al., 1980; Windish et al., 2007). Interessantemente, o p-valor esteve entre os conceitos com maior taxa de acerto em alguns estudos. Entretanto, com exceção do trabalho de Bougie e colaboradores (2015), que os residentes tiveram um desempenho excepcional, essa taxa não passou de 59,6%, apesar de serem as mais elevadas dos trabalhos (Best e Lakin,

2013; Bookstaver et al., 2012; Bougie et al., 2015; Windish et al., 2007). Tam e colaboradores (2018) avaliaram qualitativamente o entendimento de médicos a respeito do p-valor e apenas 4% apresentaram respostas alinhadas com a definição correta, reiterando o desconhecimento da base conceitual estatística. Alguns pesquisadores argumentam que as dificuldades em compreender o p-valor não são devidas à falta de literacia estatística, mas sim à natureza abstrata do conceito, sugerindo abandoná-lo como métrica de significância em trabalhos acadêmicos (García-Berthou e Alcaraz, 2004; Goodman, 2008; Tam et al., 2018).

Dentre as características demográficas associadas ao desempenho em testes de conhecimento, observou-se consistentemente que a conclusão de cursos ou formação formal em bioestatística, epidemiologia ou práticas baseadas em evidência aumenta consideravelmente as habilidades interpretativas. Contudo, isso contrasta com os achados do estudo de Alzharani e Al-Khail (2015) que encontrou resultados opostos. (Alzharani e Al-Khail, 2015; Araoye et al., 2020; Hayat et al., 2021; Hayat et al., 2021; Polychronopoulou et al., 2011; Weiss et al., 1980; Windish et al., 2007; Wulff et al., 1987; Wulff et al., 1988). Essas associações podem ser interpretadas como um indicativo de que o interesse pelos tópicos em questão motiva a busca por conhecimento, o que por sua vez fomenta o desenvolvimento da habilidade crítica. No entanto, mesmo aqueles que avaliaram o impacto dessa formação e reportaram um aumento médio absoluto em relação à média geral, indicam que a diferença não é suficiente para atender as demandas estabelecidas pela literatura (Hayat et al., 2021; Mas et al., 2018; Polychronopoulou et al., 2011; Windish et al., 2007). West e Ficalora (2007) sintetizam bem ao afirmar que cursos independentes em bioestatística dificilmente corrigirão as deficiências estatísticas e capacidade crítica de profissionais da saúde.

Quanto aos fatores negativamente associados ao conhecimento estatístico e epidemiológico, estudos frequentemente encontraram que idade, bem como o tempo decorrido desde a graduação e o período de residência estão relacionados a um menor escore nos questionários (Aloitabi et al., 2022; Hayat et al., 2021; Hayat et al., 2021; Mas et al., 2018; Msaouel et al., 2014; Windish et al., 2007; Wulff et al., 1987; Wulff et al., 1988). Porém, os estudos de Alzharani e Al-Khail (2015) e Bougie e colaboradores (2015) encontraram efeitos contrários. Uma possível explicação para a diferença de conhecimento entre profissionais formados há mais tempo e os mais recentemente graduados pode ser atribuída às mudanças curriculares, que passaram a enfatizar o ensino de epidemiologia e bioestatística. Todavia, o fato dos tópicos básicos se manterem entre os menos compreendidos, de 1980 a 2023, sugere que o problema pode ser mais profundo. Outra

possível justificativa para a perda de conhecimentos pode ser pela falta de engajamento contínuo com a estatística e estudos científicos, ausentando o reforço necessário para manter as habilidades.

Apesar da variedade dos questionários utilizados nos estudos e isso impedir comparações diretas entre profissões, através do estudo de Hayat (2021) que comparou cinco grupos de profissionais, os trabalhos que usaram o BKTSI em diferentes populações e os resultados gerais dos outros estudos, é possível inferir tendências gerais. Observa-se que, na análise crítica da literatura biomédica, há uma tendência de desempenho insatisfatório entre as várias profissões. Notavelmente, um dos estudos incluídos nesta revisão avaliou múltiplas profissões - enfermagem, farmácia, medicina, odontologia e saúde pública - com a mesma ferramenta. Nesse estudo, os profissionais da saúde pública demonstraram um desempenho substancialmente melhor do que os demais, e profissionais da odontologia performando pior (Hayat et al., 2021). Ainda que uma extrapolação, é seguro afirmar que, embora a mais estudada, a medicina não é a única que sofre com a falta de qualificação em estatística e metodologia. Essa é uma problemática ampla na área de saúde como um todo.

A falta de formação de habilidades de pesquisa durante a graduação por parte das universidades, seja estatística, seja metodológica, gera impactos que vão além da iliteracia numérica dos profissionais clínicos. Desonestidade científica a parte, há muitas publicações que apresentam falhas críticas, tanto em seu desenho quanto em suas análises estatísticas mal executadas, resultando em descobertas falaciosas. Esses falsos resultados, advindos da falta de conhecimento dos investigadores, poderiam ser prevenidos com educação adequada. (Glasziou et al., 2009; Thornton. 2009; Valcaster et al., 2021).

Erros de pesquisa podem disseminar informações danosas para a sociedade, exemplificados pelo caso do médico Andrew Wakefield e sua falsa associação entre a vacina MMR e o autismo, cujas consequências são sentidas até hoje (Motta e Stecula, 2021). Além deste, é possível citar um evento mais recente, que causou estragos pela falta de criticidade de médicos ao analisarem os trabalhos de Raoult Didier quanto ao uso de hidroxiclороquina para o tratamento e prevenção de COVID-19 (Bertin et al., 2020; Pathak et al., 2020; Glasziou, 2020). Estes são apenas exemplos mais visíveis dos perigos inerentes à falta de proficiência em estatística, contudo, erros metodológicos mais sutis, como a utilização incorreta de testes estatísticos ou a extrapolação indevida dos resultados, também podem alterar significativamente a validade de um trabalho científico (Altman, 1994; Omar et al., 2006; Ioannides, 2005; Chalmers e Glasziou, 2009; Smith, 2014).

A necessidade de aprimoramento no ensino de epidemiologia e estatística é evidente, e

estas disciplinas deveriam ser tratadas como prioridades na formação acadêmica, a par com as disciplinas básicas da saúde. Disciplinas clínicas desprovidas de um contexto estatístico-metodológico são incompletas e limitadas em utilidade. Afinal, como um profissional de saúde pode avaliar a melhor evidência e prática para um paciente se não compreende as nuances dos resultados de um estudo? Os conteúdos atualmente ensinados a graduandos e residentes mostram-se insuficientes, como evidenciado pelos resultados desta revisão, onde a maioria falha em entender conceitos básicos como p-valor, desvio padrão e intervalo de confiança. Ainda mais preocupante é o estudo de Hayat e colaboradores (2021), que avaliou o conhecimento dos docentes e encontrou que apenas aqueles da área de saúde pública atingiram escores aceitáveis. Isso sugere que até mesmo educadores podem não ter domínio do assunto. O estudo levanta a hipótese de que os padrões de acreditação para os cursos de saúde pública, particularmente rigorosos em relação à bioestatística, podem ser responsáveis pelo desempenho superior desses profissionais (Hayat et al., 2021).

A revisão atual apresenta várias limitações importantes que merecem atenção. Inicialmente, destaca-se a ausência de um segundo revisor, o que poderia ter contribuído para uma avaliação mais robusta dos estudos incluídos, especialmente no que tange à análise do risco de viés. Adicionalmente, percebe-se que o revisor possuía uma familiaridade inicial limitada com o corpo literário relevante ao tema de pesquisa. Essa limitação na compreensão abrangente e detalhada do contexto e do processo inerente a uma revisão sistemática restringiu a profundidade das inferências realizadas.

Além disso, observa-se uma significativa heterogeneidade entre os estudos analisados, tanto no que se refere às populações estudadas quanto às ferramentas de avaliação empregadas. Isso é notável mesmo entre os trabalhos que aplicaram o BKTSI, visto que várias modificações foram feitas por diversos autores, incluindo principalmente a remoção de conteúdos, reduzindo o escopo de avaliação da ferramenta. Essas variações introduzem desafios adicionais na tentativa de generalizar os resultados, comprometendo a capacidade de derivar conclusões consistentes a partir da revisão. Outro ponto importante a se destacar, a probabilidade de um viés de seleção nos estudos incluídos, uma vez que indivíduos mais familiarizados com os tópicos tendem a participar, o que pode resultar em uma média de conhecimento mais elevada, como discutido por Weiss e colaboradores (1980). No entanto, até o momento da concepção deste trabalho, nenhuma outra revisão havia agregado achados a respeito do conhecimento estatístico e metodológico de profissionais da saúde, sendo esse o primeiro a ser feito.

7 CONSIDERAÇÕES FINAIS

Este estudo realizou uma análise abrangente sobre o conhecimento estatístico e metodológico dos profissionais da saúde, preenchendo uma lacuna crítica na literatura disponível. Revelou-se uma deficiência generalizada em conceitos fundamentais de bioestatística, que transcende áreas de especialização e níveis de experiência, desde estudantes a docentes. Este trabalho enfatiza a importância vital de aprimorar e aprofundar o ensino de epidemiologia e estatística na formação acadêmica dos profissionais de saúde, dada a sua relevância indiscutível na interpretação de evidências e na implementação de práticas clínicas baseadas em evidências.

8 REFERÊNCIAS

ALEXANDER, P. E. et al. COVID-19 coronavirus research has overall low methodological quality thus far: case in point for chloroquine/hydroxychloroquine. **Journal of Clinical Epidemiology**, v. 123, p. 120-126, 2020.
DOI: 10.1016/j.jclinepi.2020.04.016

ALOTAIBI, W. A. et al. Family medicine residents' knowledge and attitudes towards biostatistics, Taif, Kingdom of Saudi Arabia. **J Family Med Prim Care**, v. 11, n. 11, p. 7015-7023.
DOI:10.4103/jfmprc.jfmprc_1146_22

ALTMAN, D. G. The scandal of poor medical research. **BMJ**, v. 308, p. 283, 1994. DOI: 10.1136/bmj.308.6924.283.

AL-ZAHRANI, Sami H.; aba al-khail, Bahaa A. Resident physician's knowledge and attitudes toward biostatistics and research methods concepts. **Saudi Med J**, [S.l.], v. 36, n. 10, p. 1236-1240, out. 2015.
DOI: 10.15537/smj.2015.10.11842.

ANDERSON, Britta L.; GIGERENZER, Gerd; PARKER, Scott; SCHULKIN, Jay. Statistical Literacy in Obstetricians and Gynecologists. **Journal For Healthcare Quality**, v. 36, n. 1, p. 5-17, jan. 2014.
DOI: 10.1111/j.1945-1474.2011.00194.x.

ARAOYE, I. et al. A National Survey of Orthopaedic Residents Identifies Deficiencies in the Understanding of Medical Statistics. **The Journal of Bone and Joint Surgery**, v. 102, n. 5, p. e19, mar. 2020.
DOI: 10.2106/JBJS.19.01095.

BAGHI, Heibatollah; KORNIDES, Melanie L. Current and future health care professionals attitudes toward and knowledge of statistics: How confidence influences learning. **J Nurs Educ Pract**, v. 3, n. 7, p. 24-29, 2013.
DOI: 10.5430/jnep.v3n7p24.

BALDI, B.; UTTS, J. What your future doctor should know about statistics: must-include topics for introductory undergraduate biostatistics. **The American Statistician**, v. 69, n. 3, p. 231-240, 2015.
DOI: 10.1080/00031305.2015.1048903.

BARRETTO, J. et al. Effect of a multimodal multidisciplinary training program on pharmacy residents' knowledge and confidence toward research and biostatistics. **Currents in Pharmacy Teaching and Learning**, v. 12, n. 1, jan. 2020, p. 20-26.
DOI: 10.1016/j.cptl.2019.10.002.

BERTIN, P., NERA, K., & DELOUVÉE, S. Conspiracy Beliefs, Rejection of Vaccination, and Support for hydroxychloroquine: A Conceptual Replication-Extension in the COVID-19 Pandemic Context. **Front. Psychol.**, v. 18, 2020.
DOI: 10.3389/fpsyg.2020.565128

BEST, Al M.; LASKIN, Daniel M. Oral and Maxillofacial Surgery Residents Have Poor Understanding of Biostatistics. **Journal of Oral and Maxillofacial Surgery**, v. 71, n. 1, jan. 2013, p. 227-234.

DOI: 10.1016/j.joms.2012.03.010.

BOOKSTAVER, P. Brandon et al. Assessing Pharmacy Residents' Knowledge of Biostatistics and Research Study Design. **Annals of Pharmacotherapy**, v. 46, n. 7-8, p. 991-999, 2012.

DOI: 10.1345/aph.1Q772.

BORNMAN, L.; MUTZ, R. Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references. **Journal of the Association for Information Science and Technology**, v. 66, n. 11, p. 2215-2222, 2015.

DOI: 10.1002/asi.23329.

BOUGIE, OLGA et al. Critical Appraisal Skills Among Canadian Obstetrics and Gynaecology Residents: How Do They Fare?. **Journal of Obstetrics and Gynaecology Canada**, v. 37, n. 7, jul. 2015, p. 639-647.

DOI: 10.1016/S1701-2163(15)30203-6.

CARLISLE, J. B. Data fabrication and other reasons for non-random sampling in 5087 randomised, controlled trials in anaesthetic and general medical journals. **Anaesthesia**, v. 72, p. 944-952, 2017.

DOI: 10.1111/anae.13938.

CHALMERS, I.; GLASZIOU, P. Avoidable waste in the production and reporting of research evidence. **The Lancet**, v. 373, n. 9683, p. 86-89, 2009.

DOI: 10.1016/S0140-6736(09)60329-9.

CHIMA, S. C.; MKWANYANA, N. M.; ESTERHUIZEN, T. M. Impact of a short biostatistics course on knowledge and performance of postgraduate scholars: Implications for training of African doctors and biomedical researchers. **Niger J Clin Pract**, v. 18, suppl. 1 p. 62-70, dez. 2015.

DOI: 10.4103/1119-3077.170818.

COUTURE, F.; NGUYEN, D.-D.; BHOJANI, N.; LEE, J. Y.; RICHARD, P. O. Knowledge and confidence level of Canadian urology residents toward biostatistics: A national survey.

Canadian Urological Association Journal, v. 14, n. 10, e514-519, May 2020.

DOI: 10.5489/cuaj.6495

FERRILL, M. J.; BLALOCK, S. J.; LINDA. A Survey of the Literature Evaluation Skills of Health-System Pharmacists. **Hospital Pharmacy**, v. 35, n. 7, p. 721-727, 2000.

DOI: doi/10.1177/001857870003500712

GARCÍA-BERTHOU, E.; ALCARAZ, C. Incongruence between test statistics and P values in medical papers. **BMC Med Res Methodol**. v. 4, n. 13, 2004.

DOI: 10.1186/1471-2288-4-13.

GLASZIOU, P. P.; SANDERS, S.; HOFFMANN, T. Waste in covid-19 research. **BMJ**, v. 369, 2020.

DOI: 10.1136/bmj.m1847

GOODMAN, Steven. A dirty dozen: twelve p-value misconceptions. **Semin Hematol**, v. 45, n. 3, p.135-140, jul. 2008.

DOI: 10.1053/j.seminhematol.2008.04.003.

HAYAT, M. J.; KIM, M.; SCHWARTZ, T. A.; JIROUTEK, M. R. A study of statistics knowledge among nurse faculty in schools with research doctorate programs. **Nursing Outlook**, v. 69, n. 2, p. 228-233, Mar. 2021.

DOI: 10.1016/j.outlook.2020.09.004

HAYAT, M. J.; SCHWARTZ, T. A.; KIM, M.; ALI, S. Z.; JIROUTEK, M. R. A comparative cross-sectional assessment of statistical knowledge of faculty across five health science disciplines. **Journal of Clinical and Translational Science**, v. 5, n. 1, p. e153, 2021.

DOI: 10.1017/cts.2021.820

HAZELTON, L.; ALLEN, M.; MACLEOD, T.; BOUDREAU, M.; LEBLANC, C. Assessing Clinical Faculty Understanding of Statistical Terms Used to Measure Treatment Effects and Their Application to Teaching. **Journal of Continuing Education in the Health Professions**, v. 36, n. 4, p. 278-283, Fall 2016.

DOI: 10.1097/CEH.0000000000000121

HELLEMS, M. A.; GURKA, M. J.; HAYDEN, G. F. Statistical Literacy for Readers of Pediatrics: A Moving Target. **Pediatrics**, v. 119, n. 6, p. 1083-1088, 2007.

DOI: 10.1542/peds.2006-2330.

HORTON, N. J.; SWITZER, S. S. Statistical methods in the Journal. **New England Journal of Medicine**, v. 353, n. 18, p. 1977-1979, Nov. 2005.

DOI: 10.1056/NEJM200511033531823

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Normas de apresentação tabular**. Rio de Janeiro, 1993. Disponível em:

<http://biblioteca.ibge.gov.br/visualizacao/livros/liv23907.pdf>. Acesso em: 24 jan. 2023.

IOANNIDIS, J. P. A. Why Most Published Research Findings Are False. **PLOS Medicine**, v. 2, n. 8, p. e124, 2005.

DOI: 10.1371/journal.pmed.0020124.

JAYAKUMAR, K. L.; UMSCHIED, C. A. Expanding epidemiology and biostatistics curricula in undergraduate medical education to promote evidence-based practice. **BMJ Evidence-Based Medicine**, v. 23, n. 2, p. 60-61, 2018.

DOI: 10.1136/bmjebm-2017-110875.

JIROUTEK, M. R.; KIM, M.; HAYAT, M. J.; TAYLOR, M. N.; SCHWARTZ, T. A. A cross-sectional assessment of statistical knowledge among pharmacy faculty. **Currents in Pharmacy Teaching and Learning**, v. 11, n. 8, p. 793-801, Aug. 2019.

DOI: 10.1016/j.cptl.2019.04.007

LAKHLIFI, C.; LEJEUNE, F.-X.; ROUAULT, M.; KHAMASSI, M.; ROHAUT, B. Illusion of knowledge in statistics among clinicians: evaluating the alignment between objective accuracy and subjective confidence, an online survey. **Cognitive Research: Principles and Implications**, v. 8, n. 23, p. 2023.

DOI: 10.1186/s41235-023-00474-1

MAS, G.; TELLO, T.; ORTIZ, P.; PETROVA, D.; GARCIA-RETAMERO, R. Graphical and numerical skills in pre- and postgraduate medical students from a private university. **Gac Med Mex**. v. 154, n. 2, p. 163-169, 2018.

DOI: 10.24875/GMM.M18000132.

MAYER, D. Evidence-based medicine. **Epilepsia**, v. 47, supl. 1, p. 3-5, 2006.

DOI: 10.1111/j.1528-1167.2006.00651.x.

MCNALLY, P.; LOFTUS, B. G. Knowledge of statistical methods and their implications for clinical practice: A survey of paediatricians. **Irish Medical Journal**, 2005. v. 98, n. 10.

Disponível em: <https://archive.imj.ie/ViewArticleDetails.aspx?ArticleID=694>

Acessado em: 16/09/2023 às 17:30.

MOTTA M, STECULA D. Quantifying the effect of Wakefield et al. (1998) on skepticism about MMR vaccine safety in the U.S. **PLoS One**. v. 16, n. 8, p. e0256395, 2021.

DOI: 10.1371/journal.pone.0256395

MSAOUEL, P.; KAPPOS, T.; TASOULIS, A.; APOSTOLOPOULOS, A. P.; LEKKAS, I.; TRIPODAKI, E.-S.; et al. Assessment of cognitive biases and biostatistics knowledge of medical residents: a multicenter, cross-sectional questionnaire study. *Medical Education Online*. v. 19, n. 1, 2013.

DOI: 10.3402/meo.v19.23646.

OMAR, R. Z., MCNALLY, N., AMBLER, G., & POLLOCK, A. M.. Quality research in healthcare: are researchers getting enough statistical support? *BMC Health Services Research*, v. 6, n. 1, 2006

DOI: 10.1186/1472-6963-6-2

PATHAK, S. K., SALUNKE, A. A., ET AL. (2020). No benefit of hydroxychloroquine in COVID-19: Results of Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Diabetes Metab Syndr*, 14(6), 1673-1680. 10.1016/j.dsx.2020.08.033

POLYCHRONOPOULOU, A.; ELIADES, T.; TAOUFIK, K.; PAPADOPOULOS, M. A.; ATHANASIOU, A. E. Knowledge of European orthodontic postgraduate students on biostatistics. **European Journal of Orthodontics**. v. 33, n. 4, p. 434-440. 2011

DOI:10.1093/ejo/cjq098.

SHAFI, M. S.; FAISAL, T.; NASEEM, S.; JAVED, S. Postgraduate Medical Trainees Understanding of Biostatistics: A Pre- and Post-Research Methodology Workshop Experience. **J Coll Physicians Surg Pak**. v. 28, n. 3, p. 196-200, 2018.

DOI: 10.29271/jcpsp.2018.03.196.

SHEEHAN, A. H.; WAGNER, E. L.; SWOSINSKI, K. M.; et al. Student pharmacists' knowledge of biostatistical and literature evaluation concepts. **Currents in Pharmacy**

Teaching and Learning. v. 8, n. 5, p. 622-628, 2016
DOI: 10.1016/j.cptl.2016.06.013.

SMITH, RICHARD. Medical research—still a scandal. **The BMJ Opinion**, 31 jan. 2014.
Disponível em:
<https://blogs.bmj.com/bmj/2014/01/31/richard-smith-medical-research-still-a-scandal>
Acesso em: 30/11/2023

STEFAN, A. M.; SCHÖNBRODT, F. D. Big little lies: a compendium and simulation of p-hacking strategies. *Royal Society Open Science*, v. 10, n. 2, fev. 2023.
DOI: 10.1098/rsos.220346%20Review%20history

SUSARLA, S. M.; REDETT, R. J. Plastic surgery residents' attitudes and understanding of biostatistics: a pilot study. *J Surg Educ*. v. 71, n. 4, p. 574-579, 2014.
DOI: 10.1016/j.jsurg.2013.12.003

TAM, C. W. M.; KHAN, A. H.; KNIGHT, A.; RHEE, J.; PRICE, K.; MCLEAN, K. How doctors conceptualise P values: A mixed methods study. *Aust J Gen Pract*. v.47, n. 10, p. 705-710, 2018.
DOI: 10.31128/AJGP-02-18-4502.

THORNTON, H. Statistical illiteracy is damaging our health: Doctors and patients need to understand numbers if meaningful dialogues are to occur. **International Journal of Surgery**, v. 7, n. 4, p. 279-284, 2009.
DOI: 10.1016/j.ijssu.2009.06.008

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL. Faculdade de Medicina. Biblioteca da Faculdade de Medicina. **Normas**. Porto Alegre, c2023. Disponível em: <https://www.ufrgs.br/bibmed/ferramentas/normas/>. Acesso em: 24 jan. 2023.

VAN CALSTER, B. et al. Methodology over metrics: current scientific standards are a disservice to patients and society. *Journal of Clinical Epidemiology*, v. 138, p. 219-226, 2021.
DOI: 10.1016/j.jclinepi.2021.02.007.

WEISS, S. T.; SAMET, J. M. An assessment of physician knowledge of epidemiology and biostatistics. **Academic Medicine**, v. 55, n. 8, p. 692-697, 1980.
DOI: 10.1097/00001888-198008000-00007.

WEST, C. P.; FICALORA, R. D. Clinician Attitudes Toward Biostatistics. *Mayo Clinic Proceedings*, v. 82, n. 8, p. 939-943, ago. 2007.
DOI: 10.4065/82.8.939

WINDISH, D. M.; HUOT, S. J.; GREEN, M. L. Medicine residents' understanding of the biostatistics and results in the medical literature. **JAMA**, v. 298, n. 9, p. 1010-1022, 2007.
DOI: 10.1001/jama.298.9.1010

WULFF, H. R.; ANDERSEN, B.; BRANDENHOFF, P.; GUTTNER, F. What do doctors know about statistics? *Statistics in Medicine*, v. 6, n. 1, p. 3-10, Jan./Feb. 1987.
DOI: 10.1002/sim.4780060103

WULFF, H. R.; ANDERSEN, B.; SCHEUTZ, F. What do dentists know about statistics?
European Journal of Oral Sciences. v. 96, n. 4, p. 281–287, 1988
DOI: 10.1111/j.1600-0722.1988.tb01557.x.

Figura 1 - PRISMA 2020: Fluxograma para revisões sistemáticas que incluem buscas em bases de dados outras fontes

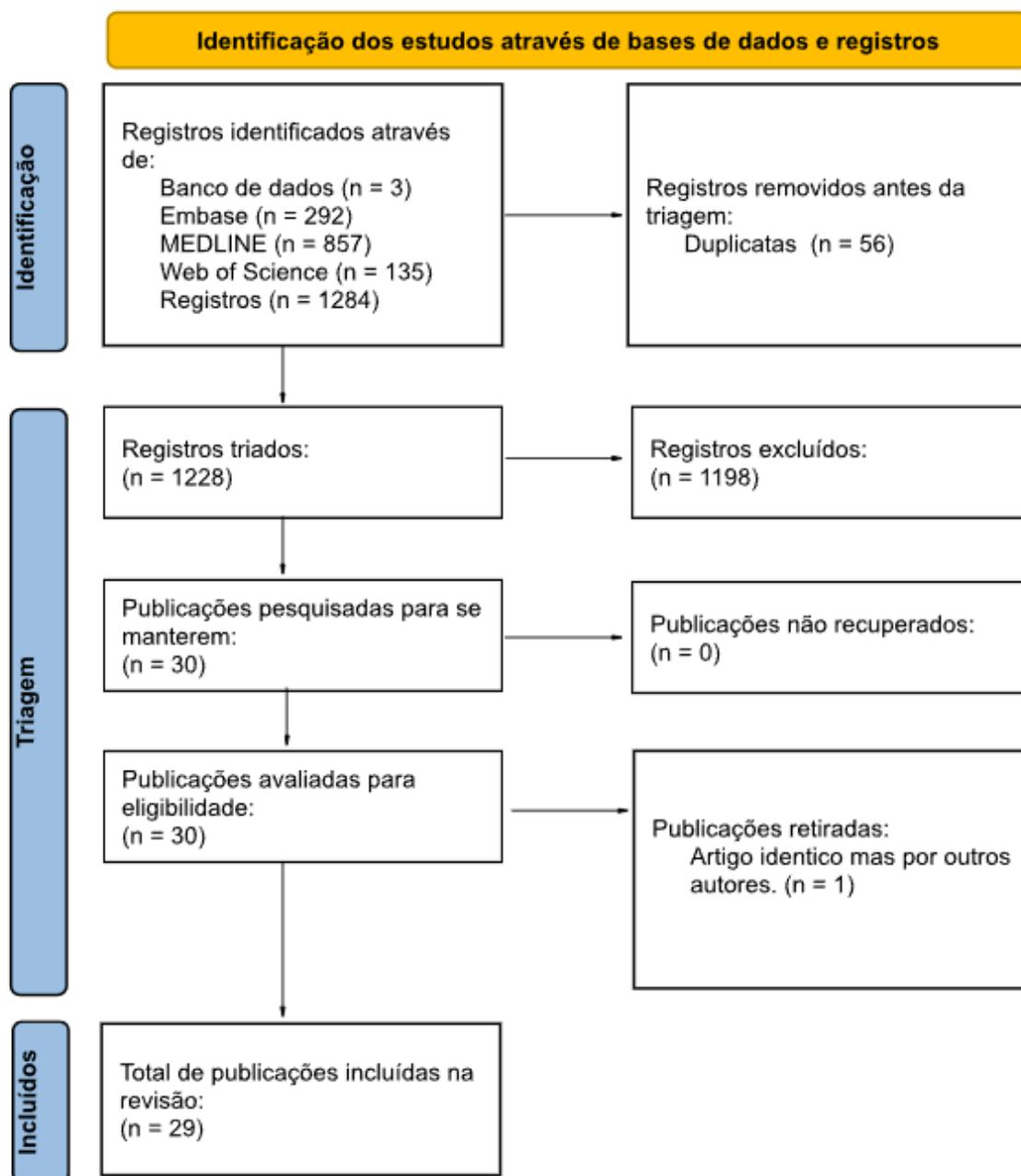


Tabela 1 - Sumário das características dos estudos

Referência, Ano e Localidade	Título	Tipo de Estudo	Questionário	População	Resultados
Alotaibi et al, 2022 Arábia Saudita	Family medicine residents' knowledge and attitudes towards biostatistics, Taif, Kingdom of Saudi Arabia	Transversal	15 Questões	Resid. de Med. n=113	Média 34%
Alzharani et al, 2015 Arábia Saudita	Resident physician's knowledge and attitudes toward biostatistics and research methods concepts	Transversal	20 Questões (BKTSI)	Resid. de Med. n=162	Média <=30 anos: 23,75% >30 anos: 24,06%
Anderson et al, 2014 Estados Unidos	Statistical literacy in obstetricians and gynecologists	Transversal	18 Questões	Obstetras e Gineco. n=94	Média Geral OGSLQ 62,67%
Araoye et al, 2020 Estados Unidos	A National Survey of Orthopaedic Residents Identifies Deficiencies in the Understanding of Medical Statistics	Transversal	20 Questões (BKTSI)	Resid. de Med. n=178	Média Geral (Range) 49,4% (10-90%)
Baghi et al, 2013 Estados Unidos	Current and future health care professionals attitudes toward and knowledge of statistics: How confidence influences learning	Pré-pós intervenção	30 Questões	Estud. de Grad. Enf n=165	Soma das médias das proficiências (%) 12,1 de 30 (40,33%)
Barreto et al, 2020 Estados Unidos	Effect of a multimodal multidisciplinary training program on pharmacy residents' knowledge and confidence toward research and biostatistics	Pré-pós intervenção	28 Questões (BKTSI)	Resid. de Farm. n=8	Média Geral 53,57%
Best e Laskin, 2013 Estados Unidos	Oral and maxillofacial surgery residents have poor understanding of biostatistics	Transversal	6 Questões (BKTSI)	Resid. de Odonto. n=111	Média Geral (Range) 38% (34,1 - 42,4)
Bookstaver et al, 2012 Estados Unidos	Assessing pharmacy residents' knowledge of biostatistics and research study design	Transversal	12 Questões (BKTSI)	Resid. de Farm. n=214	Média Geral (Range) 47,3% (44,45 - 50,12)
Bougie et al, 2015 Canadá	Critical Appraisal Skills Among Canadian Obstetrics and Gynaecology Residents: How Do They Fare?	Transversal	15 Questões	Resid. de Obst. e Gineco n=301	Média Geral (Range) 69,8% (68,2 - 71,4)

Chima et al, 2015 África do Sul	Impact of a short biostatistics course on knowledge and performance of postgraduate scholars Implications for training of African doctors and biomedical researchers	Pré-pós intervenção	7 Questões	Médicos e Biomédicos	Mediana: 0 Percentil 75: 28,6%
Couture et al, 2020 Canadá	Knowledge and confidence level of Canadian urology residents toward biostatistics: A national survey	Transversal	10 Questões (BKTSI)	Resid. de Uro. n=68	Média Geral (DP) 65% (20.7) Participantes com nota >=80% :30,9% P com nota >=60%: 72,1%
Ferril et al, 2000 Estados Unidos	A Survey of the Literature Evaluation Skills of Health-System Pharmacists	Transversal	5 Questões	Farmacêuticos n=539	Média Geral (DP) 2,2 (1,1) de 5 44%
Hayat et al, 2021 Estados Unidos	A study of statistics knowledge among nurse faculty in schools with research doctorate programs	Transversal	8 Questões	Enferm. com doc. n=113	Média Geral 63.6% Acertos Brutos (DP) 5.1 (1.6)
Hayat et al, 2021 Estados Unidos	A comparative cross-sectional assessment of statistical knowledge of faculty across five health science disciplines	Transversal	8 Questões	Docentes de Odonto, Med., Saúde Pública, Farm. e Enf. n=708	Média Independente Dent= 53.3%, Med= 62.7%, Enf=63.6%, Farm= 63.8%, SP=80.7%. Média Geral 66,2%
Hazelton et al, 2016 Canadá	Assessing Clinical Faculty Understanding of Statistical Terms Used to Measure Treatment Effects and Their Application to Teaching	Transversal	4 Questões	Médicos	26% responderam corretamente as 4 questões
Jiroutek et al, 2019 Estados Unidos	A cross-sectional assessment of statistical knowledge among pharmacy faculty	Transversal	8 Questões	Docentes da Farm. n=139	Média (DP) 5.1 (1.65)/8Q 63,75%

Lakhfili et al, 2023 França	Illusion of knowledge in statistics among clinicians: evaluating the alignment between objective accuracy and subjective confidence, an online survey	Transversal	3 Questões	Estud., Resid., Clinicos de Med. n=898	Varição dos resultados P-valor: 28,1 a 71,4% Eficácia da Vacina: 37,7 a 91% Confiança: Questões corretas: Mediana: 78,5% Questões incorretas: 66,7% Questões Teóricas (% de acerto) Definição de Sensibilidade: 80,4% Definição de Especificidade: 66,8% Definição de VPP: 75,7% Definição de VPN: 73,9% Cálculo de VPP: 15%
Mas et al, 2018 Peru	Graphical and numerical skills in pre- and postgraduate medical students from a private university	Transversal	22 Questões	Estud de Grad. Med (6 e 7 ano) e Resid. n=169	Média Geral 80,45%
McNally et al, 2005 Irlanda	Knowledge of Statistical Methods and their Implications for Clinical Practice: A Survey of Paediatrician	Transversal	10 Questões	Méd. Pedi. n=48	Média Geral 51%
Msaouel et al, 2014 Grécia	Assessment of cognitive biases and biostatistics knowledge of medical residents: a multicenter, cross-sectional questionnaire study	Transversal	7 Questões	Resid. de Med. n=153	Média Geral (Range) 30% (0 a 100%)
Polychronopoulou et al, 2011 Europa (Bulgária, Finlândia, Alemanha, Grécia, Itália, Holanda, Noruega, Suíça, Suécia e Inglaterra)	Knowledge of European orthodontic postgraduate students on biostatistics	Transversal	13 Questões (BKTSI)	Pós-grad de odonto. n=127	Média Geral (Range) 43,8% (40,2-47,3)
Weiss et al, 1988 Dinamarca	What do dentists know about statistics?	Transversal	9 Questões	Dentistas	Mediana de Acertos (%) 2.2 (24,44%)

Shafi et al, 2018 Paquistão	Postgraduate Medical Trainees Understanding of Biostatistics: A Pre- and Post-Research Methodology Workshop Experience	Pré-pós intervenção	21 Questões	Pós-grad de Med. n=270	Média Geral 62,9%
Sheeran et al, 2016 Estados Unidos	Student pharmacists' knowledge of biostatistical and literature evaluation concepts	Coorte	20 Questões (BKTSI)	Resid. de Farm. n=501	Média (DP) P1= 44,5% (10,6) P2= 49,5% (12,4) P3= 45% (12,9) P4= 47,1% (14,2)
Susarla et al, 2014 Estados Unidos	Plastic surgery residents' attitudes and understanding of biostatistics: a pilot study	Transversal	6 Questões (BKTSI)	Resid. de Cir. Plást. n=22	Média Geral (DP) 53% (24,5%)
Tam et al, 2018 Austrália e Nova Zelândia	How doctors conceptualise P values: A mixed methods study	Transversal	3 Questões	Médicos n=247	4% descreveram uma descrição aceitável de p-valor e 24% acertaram a questão objetiva
Weiss et al, 1980 Estados Unidos	An assessment of physician knowledge of epidemiology and biostatistics	Transversal	10 Questões	Médicos n=141	Média Geral (Range) 74% (33% a 100%)
Windish et al, 2007 Estados Unidos	Medicine Residents' Understanding of the Biostatistics and Results in the Medical Literature	Transversal	20 Questões (BKTSI)	Resid. de Med. n=277	Média Geral (DP) 41,1% (15,2%)
Wulff et al, 1987 Dinamarca	WHAT DO DOCTORS KNOW ABOUT STATISTICS?	Transversal	9 Questões	Médicos n=148	Mediana de Acertos (%) 2.4 (26,67%)

Tabela 2 - Sumário dos resultados específicos dos estudos

Referência, Ano e Localidade	Questionário	Menores taxas de acerto (% de acerto)	Maiores taxas de acerto (% de acerto)
Alotaibi et al, 2022	15 Questões	Análise do Kaplan-Meier (15,9%) Regressão múltipla para análise ajustada (17,7%) Regressão simples em análise não ajustada (22,1%) P-valor (26,5%)	Variável contínua (57,5%)
Alzharani et al, 2015	20 Questões (BKTSI)	Não descrito	Não descrito
Anderson et al, 2014	18 Questões	Dados de prevalência (7%; 19%; 15%)	Cálculo de percentual (89%; 77%) Incidência (74%) Prevalência (68%)
Araoye et al, 2020	20 Questões (BKTSI)	Análise do Kaplan-Meier (17,4%) Regressão proporcional de Cox (16,9%) Intervalo de confiança (11,8%) Força de associação (10,2%)	Risco relativo (87,64%) Propósito do cegamento (79,7%) Teste t (78,53%) ANOVA (69,7%)
Baghi et al, 2013	30 Questões	Não descrito	Não descrito
Barreto et al, 2020	28 Questões (BKTSI)	Não descrito	Não descrito
Best e Laskin, 2013	6 Questões (BKTSI)	Variável contínua (24%) Variável nominal (26%) Variável ordinal (28%) Reconhecer caso-controle (21%)	Propósito do cegamento (79%) P-valor (46%)
Bookstaver et al, 2012	12 Questões (BKTSI)	Regressão proporcional de Cox (24,8%) Definição de viés (23,9%) Variável ordinal (41,9%)	Propósito do cegamento (92,6%) Risco relativo (75,8%) Variável nominal (69,4%) P-valor (59,6%)
Bougie et al, 2015	15 Questões	Regressão proporcional de Cox (9,9%) ANOVA (36,9%) Relação entre poder, amostra e diferença detectável significativa (42%)	Razão de chances (92,7%) P-valor (88,7%) Análise do Kaplan-Meier (88,2%)
Chima et al, 2015	7 Questões	Todas as questões tiveram taxa de acerto baixa	Todas as questões tiveram taxa de acerto baixa
Couture et al, 2020	10 Questões (BKTSI)	Razão de chances ajustadas (50%) Variável ordinal (51,5%) Variável discreta (52,9%)	Variável contínua (89,7%) Propósito do cegamento (85,3%) Teste t (64,7%)
Ferril et al, 2000	5 Questões	Identificar erro do estudo (18,3%) Incidência (25,7%) Comparação entre intervenções (32,8%)	Objetivo do estudo (93,4%) Tamanho amostral (52,1%)
Hayat et al, 2021	8 Questões	Propósito da randomização (34,9%) Razão de chances (33,5%) Intervalo de confiança (42,7%)	Entender o potencial de múltiplos testes (82,9%) Estudo observacional (82,3%) Relação entre amostra e erro padrão (82,3%)

Hayat et al, 2021	8 Questões	Propósito da randomização (51,7%) Intervalo de confiança (49,7%) Razão de chances (47,7%)	Entender o potencial de múltiplos testes (80,9%) Estudo observacional (82,5%) Relação entre amostra e erro padrão (81,6%)
Hazelton et al, 2016	4 Questões	Não descrito	Não descrito
Jiroutek et al, 2019	8 Questões	Intervalo de confiança (42,5%) Propósito da randomização (50,4%) Razão de chances (51,8%)	Entender o potencial de múltiplos testes (80,6%) Estudo observacional (79,9%) Relação entre amostra e erro padrão (74,1%)
Lakhfili et al, 2023	3 Questões	Não descrito	Não descrito
Mas et al, 2018	22 Questões	Não descrito	Não descrito
McNally et al, 2005	10 Questões	Não descrito	Não descrito
Msaouel et al, 2014	7 Questões	Erro padrão (22,9%) P-valor (26,8%)	Sensibilidade (51,6%) Desvio padrão (32%)
Polychronopoulou et al, 2011	13 Questões (BKTSI)	Reconhecer caso controle (3,2%) Qui-quadrado (11,8%) Regressão proporcional de Cox (14,1%) P-valor (33%)	Propósito de cegamento (77,9%) Definição de viés (70,8%) Reconhecer meta-análise (70%)
Weiss et al, 1988	9 Questões	Erro padrão (24%) Desvio padrão (34%) P-valor (8%)	Não descrito
Shafi et al, 2018	21 Questões	Teste t independente (31,9%) Qui-quadrado (27%) Teste t dependente (30,7%)	Definição de variáveis (91,9%) Tipos de variáveis (92,2%) Exemplo de dado quantitativo (87,4%)
Sheeran et al, 2016 Estados Unidos	20 Questões (BKTSI)	P2 (maior média geral) Análise de Kaplan-Meier (8,9%) Força da evidência para fatores de risco (10,3%) Intervalo de confiança (16,4%)	P2 (maior média geral) Variável contínua (93,2%) Propósito de cegamento (89,0%) Risco relativo (80,8%)
Susarla et al, 2014 Estados Unidos	6 Questões (BKTSI)	Regressão proporcional de Cox (9,1%) Intervalo de confiança (45,5%) Desenho de estudo (45,5%) ANOVA (45,5%)	Não descrito
Tam et al, 2018 Austrália e Nova Zelândia	3 Questões	Não descrito	Não descrito

Weiss et al, 1980 Estados Unidos	10 Questões	Teste de Hipótese (não especificado) P-valor (40%) Hipótese nula (50%) Desenho de estudo (não especificado)	Risco relativo (97%) Desvio padrão (91%)
Windish et al, 2007 Estados Unidos	20 Questões (BKTSI)	Análise do Kaplan-Meier (10,5%) Regressão proporcional de Cox (13,0%) Intervalo de confiança (11,9%) Força da evidência para fatores de risco (17,0%)	Propósito do cegamento (87,4%) Risco relativo (81,6%) P-valor (58,8%) Teste t (58,1%)
Wulff et al, 1987 Dinamarca	9 Questões	Erro padrão (39%) Desvio padrão (30%)	Não descrito