

**UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE ODONTOLOGIA
PROGRAMA DE PÓS GRADUAÇÃO EM ODONTOLOGIA**

Larissa Costa Freitas

**Natureza da Inovação em Prótese Dentária: Uma Análise Bibliométrica do
Índice de Disrupção**

Juiz de Fora

2026

Larissa Costa Freitas

**Natureza da Inovação em Prótese Dentária: Uma Análise Bibliométrica do
Índice de Disrupção**

Tese apresentada ao Programa de Pós graduação em Odontologia da Faculdade de Odontologia da Universidade Federal de Juiz de Fora, como requisito parcial para obtenção do título de Doutora em Odontologia. Área de concentração: Odontologia.

Orientadora: Prof^a. Dr^a. Fabíola Pessoa Pereira Leite

Juiz de Fora

2026

Ficha catalográfica elaborada através do programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Freitas, Larissa.

Natureza da Inovação em Prótese Dentária : Uma Análise Bibliométrica do Índice de Disrupção / Larissa Freitas. -- 2026. 64 f.

Orientador: Fabíola Pessoa Pereira Leite
Tese (doutorado) - Universidade Federal de Juiz de Fora, Faculdade de Odontologia. Programa de Pós-Graduação em Odontologia, 2026.

1. Prótese Dentária. 2. Bibliometria. 3. Inovação Científica. I. Leite, Fabíola Pessoa Pereira , orient. II. Título.

Larissa Costa Freitas

Natureza da Inovação em Prótese Dentária: Uma Análise Bibliométrica do Índice de Disrupção

Tese apresentada ao Programa de Pós-graduação em Odontologia da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Doutora em Odontologia. Área de concentração: Clínica Odontológica.

Aprovada em 13 de maio de 2026.

BANCA EXAMINADORA

Prof.^a Dr.^a Fabíola Pessôa Pereira Leite - Orientadora e Presidente

Universidade Federal de Juiz de Fora - UFJF

Prof. Dr. Ronaldo Luís Almeida de Carvalho - Membro titular interno

Universidade Federal de Juiz de Fora - UFJF

Prof. Dr. Leonardo Custódio de Lima - Membro titular interno

Universidade Federal de Juiz de Fora - UFJF - *Campus Governador Valadares*

Prof. Dr. Felipe Eduardo de Oliveira - Membro titular externo

Centro Universitário Braz Cubas

Prof.^a Dr.^a Graziela Ribeiro Batista - Membro titular externo

Indiana University School of Dentistry

Juiz de Fora, 29/04/2026.



Documento assinado eletronicamente por **Fabiola Pessoa Pereira Leite, Professor(a)**, em 13/05/2026, às 15:50, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Ronaldo Luis Almeida de Carvalho, Professor(a)**, em 13/05/2026, às 16:16, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Leonardo Custódio de Lima, Professor(a)**, em 13/05/2026, às 16:17, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Felipe Eduardo de Oliveira, Usuário Externo**, em 13/05/2026, às 18:14, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Graziela Ribeiro Batista, Usuário Externo**, em 13/05/2026, às 20:11, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no Portal do SEI-Ufjf (www2.ufjf.br/SEI) através do ícone Conferência de Documentos, informando o código verificador **2965437** e o código CRC **75BC905C**.

Dizem que se morre duas vezes: a primeira, de fato quando se enterra o corpo, a segunda, quando seu nome é mencionado pela última vez. Dedico, portanto, ao meu avô, Aníbal Salustiano da Costa, a imortalidade da escrita.

Agradecimentos

Chegar até aqui é a soma de muitas mãos estendidas, muitos gestos silenciosos e muitas presenças essenciais. Nada disso teria sido possível sem as pessoas que caminharam comigo ao longo desta trajetória.

Minha gratidão começa em Deus, fonte de força, direção e esperança. Foi Ele quem sustentou cada passo, renovou minhas energias nos dias difíceis e iluminou minhas decisões.

Realizar este sonho me trouxe um sentimento de realização, de força. Mas sei que ele foi repleto de abdições. Eu estive ausente em aniversários, datas comemorativas, momentos de comemoração, deixei de oferecer meu abraço em momentos difíceis. E ainda assim, as pessoas aqui mencionadas, me ofereceram o mais puro afeto e incentivo. Por mais que eu me expresse, jamais serei capaz de demonstrar o quanto o amor deles me manteve de pé.

À minha família, especialmente aos meus pais, José de Freitas Guimarães Júnior e Vanda Meire Costa Freitas, expresso minha eterna gratidão por todo amor, incentivo e dedicação. Em especial a minha mãe, que foi fonte de motivação constante para que eu acreditasse em mim e seguisse em frente. Ao meu irmão e meu sobrinho, agradeço por sempre me lembrarem quem eu sou.

Aos meus familiares, que sempre estiveram presentes por meio de palavras de encorajamento, pensamentos positivos e orações, fortalecendo-me mesmo à distância. Não podendo deixar de mencionar, às orações da minha avó, Alda Maria da Costa e ao encorajamento de meu avô, Aníbal Salustiano da Costa (in memoriam).

Aos amigos, que souberam respeitar meus momentos de ausência e, ainda assim, permaneceram ao meu lado, oferecendo apoio, paciência e afeto. Desejo a todos uma amizade como a da Analina, que abraçou comigo a jornada como uma irmã.

À Profa. Dra. Fabíola, por acreditar no meu potencial quando eu duvidei e me orientar com tanto desejo que eu alcançasse meus objetivos. Ao Prof. Dr. Ronaldo, a sua orientação e amizade inspiraram este trabalho.

A todos que, de alguma forma, contribuíram para que este sonho se tornasse realidade, deixo aqui meu sincero agradecimento.

À Universidade Federal de Juiz de Fora, que permanecerá sempre sendo a minha casa.

Agradeço a todos aqueles que de forma direta ou indireta fizeram parte da construção deste sonho.

“Eu sou a continuação de um sonho
Da minha mãe, do meu pai, todos que vieram antes de mim
Eu sou a continuação de um sonho
Da minha vó, do meu vô, quem sangrou pra gente poder sorrir”

Bk- Continuação de um sonho

RESUMO

A inovação científica na área de Prótese Dentária tem sido impulsionada pelo crescimento da produção científica e pelo avanço de tecnologias e materiais, tornando relevante a aplicação de métricas capazes de diferenciar contribuições incrementais de avanços disruptivos. Nesse contexto, o presente estudo teve como objetivo categorizar a natureza da inovação científica na literatura de Prótese Dentária ao longo do tempo, por meio da aplicação do Índice de Disrupção (ID), em publicações indexadas na base PubMed entre 1900 e 2025. A busca foi realizada por meio de MeSH *terms* e termos livres relacionados à prótese dentária. O Índice de Disrupção foi calculado a partir de redes de citação, considerando uma janela temporal de cinco anos. Foram conduzidas análises descritivas, temporais e correlacionais, incluindo variáveis como número de autores, referências, citações, países, financiamento, periódicos e descritores. Dos 141.829 artigos identificados, 2.980 apresentaram dados suficientes para o cálculo do DI. Entre 2003 e 2020, 73,1% das publicações apresentaram caráter disruptivo, com predominância de valores moderados ($0 < DI < 0,5$). O DI médio manteve-se positivo e estável ao longo do tempo. Não foi observada correlação entre DI e número de citações ($r = -0,001$), enquanto o número de referências apresentou correlação fraca e negativa ($r = -0,135$). Artigos com menor número de autores apresentaram maiores valores de DI. Países com maior produção científica apresentaram valores moderados de DI, enquanto países como Índia, Brasil e Japão demonstraram maior potencial disruptivo relativo. Os artigos mais disruptivos não coincidiram com os mais citados. A inovação científica na Prótese Dentária apresenta predominância de caráter disruptivo moderado, indicando um padrão de avanço incremental. O Índice de Disrupção mostrou-se útil para identificar diferentes perfis de contribuição científica, evidenciando que a disrupção não está diretamente associada ao impacto por citações ou ao volume de produção.

Palavras-chave: Prótese Dentária; Bibliometria; Inovação Científica

ABSTRACT

Scientific innovation in the field of Prosthodontics has been driven by the growth of scientific production and advances in technologies and materials, highlighting the relevance of applying metrics capable of distinguishing incremental contributions from disruptive advances. In this context, the present study aimed to categorize the nature of scientific innovation in Prosthodontics literature over time through the application of the Disruption Index (DI) in publications indexed in the PubMed database between 1900 and 2025. The search strategy included MeSH descriptors and free terms related to prosthodontics. The Disruption Index was calculated based on citation networks, considering a five-year citation window. Descriptive, temporal, and correlational analyses were performed, including variables such as number of authors, references, citations, countries, funding, journals, and descriptors. Of the 141,829 articles identified, 2,980 had sufficient data for DI calculation. Between 2003 and 2020, 73.1% of the publications exhibited a disruptive character, predominantly with moderate values ($0 < DI < 0.5$). The mean DI remained positive and stable over time. No correlation was observed between DI and the number of citations ($r = -0.001$), while the number of references showed a weak negative correlation ($r = -0.135$). Articles with fewer authors presented higher DI values. Countries with higher scientific output showed moderate DI values, whereas countries such as India, Brazil, and Japan demonstrated higher relative disruptive potential. The most disruptive articles did not coincide with the most cited ones. Scientific innovation in Prosthodontics shows a predominance of moderate disruptive character, indicating a pattern of incremental advancement. The Disruption Index proved to be a useful tool for identifying different profiles of scientific contribution, demonstrating that disruption is not directly associated with citation impact or publication volume.

Keywords: Prosthodontics; Bibliometrics; Scientific Innovation

LISTA DE TABELAS

Tabela 1	Artigos mais citados na amostra de prótese dentária, com respectivos valores de Índice de Disrupção (DI)	32
Tabela 2	Artigos com maiores valores de Índice de Disrupção (DI) na amostra, com respectivos números de citações	33
Tabela 3	Periódicos com maior Índice de Disrupção médio (DI médio) e respectivos modelos de acesso	34
Tabela 4	Distribuição dos países com maior número de publicações e respectivos valores de Índice de Disrupção médio e mediano	36
Tabela 5	Principais agências de fomento associadas às publicações e respectivos valores de Índice de Disrupção médio e mediano	
Tabela 6	Frequência dos termos utilizados na estratégia de busca e respectivos valores de Índice de Disrupção médio.	

LISTA DE FIGURAS

Figura 1	Diagrama esquemático das relações de citação entre um artigo focal e seus trabalhos predecessores	19
Figura 2	Evolução temporal do número de artigos em prótese dentária com e sem cálculo do Índice de Disrupção na base PubMed	28
Figura 3	Distribuição anual das categorias do Índice de Disrupção dos artigos publicados entre 2003 e 2020	29
Figura 4	Variação temporal do Índice de Disrupção médio e do intervalo entre valores mínimos e máximos observados na amostra.....	29
Figura 5	Gráfico de dispersão do número de referências X DI	30
Figura 6	Análise de sensibilidade do Índice de Disrupção médio em função do número mínimo de referências considerado	36
Figura 7	Relação entre o número total de citações recebidas pelos artigos e o Índice de Disrupção	36
Figura 8	Relação entre o número de autores por artigo e o Índice de Disrupção médio	37

LISTA DE ABREVIATURAS E SIGLAS

ID	Índice de Disrupção
API	<i>Application Programming Interface</i>
PMID	PubMed Identifier (Identificador do PubMed)
DOI	Digital Object Identifier
ORCID	Open Researcher and Contributor ID
NIH	National institute of health
MeShTerms	Medical Subject Headings Medical
WoS	Web Of Science
EUA	Estados Unidos da América

LISTA DE SÍMBOLOS

- + Representação gráfica que indica adição ou soma
- Representação gráfica que indica subtração.
- < Representação gráfica que indica que um valor é menor que outro.
- > Representação gráfica que indica que um valor é maior que outro.
- = Representação gráfica que indica igualdade entre dois valores.
- ≈ Representação gráfica que indica aproximadamente
- ρ Coeficiente de correlação de Spearman
- r Coeficiente de correlação de Pearson

SUMÁRIO

1	INTRODUÇÃO	13
2	REVISÃO DE LITERATURA	15
2.1	História da bibliometria	15
2.2	Bancos de dados na bibliometria	15
2.3	Análises bibliométricas	16
2.4	Relevância na ciência	13
2.5	Conceito do índice de disrupção	14
2.6	Limitações do índice de disrupção	16
2.7	Aplicabilidade do índice de disrupção	16
2.8	Estudos bibliométricos na odontologia.....	17
3	PROPOSIÇÃO	19
4	METODOLOGIA	20
4.1	CARACTERIZAÇÃO BIBLIOMÉTRICA	
4.1.1	Seleção de dados	
4.1.2	Busca dos dados	
4.1.3	Cálculo do DI	
4.2	ANÁLISES CORRELACIONAIS	
5	RESULTADOS	24
6	DISCUSSÃO	25
7	CONCLUSÃO	26
	REFERÊNCIAS	30
	GLOSSÁRIO	
	ANEXOS	

1 INTRODUÇÃO

A perda de dentes por cárie, doença periodontal ou trauma ainda é comum e reduz a capacidade mastigatória, a fala e a estética, comprometendo a qualidade de vida relacionada à saúde bucal (Emami *et al.*, 2013). A prótese dentária é uma área da odontologia que substitui dentes ausentes por peças artificiais e, ao restaurar estruturas orais e faciais, contribui para a saúde geral do paciente (Watted *et al.*, 2016; Emami, E. *et al.*, 2013). Mesmo com avanços em prevenção, o edentulismo continua sendo um problema de saúde pública, e a perda total de dentes varia entre países, sendo mais prevalente em idosos e influenciada por fatores socioeconômicos e de gênero (Watted *et al.*, 2016). Tais impactos motivaram um aumento nas pesquisas sobre técnicas, materiais e tecnologias aplicadas às próteses e tornaram o campo fértil para análises quantitativas (Hsu; Chang, 2025).

Com o crescimento exponencial da literatura científica, os estudos de avaliação de produção científica tornaram-se mais complexos e para lidar com grandes volumes de publicações, Pritchard propôs em 1969 o termo bibliometria para designar o uso de métodos estatísticos e matemáticos na quantificação de livros e artigos (Cheng *et al.*, 2024). Essa metodologia possibilita examinar tendências, identificar temas emergentes e avaliar colaborações ou impactos de pesquisas. Ferramentas computacionais modernas e bases de dados eletrônicas como PubMed, Scopus e *Web of Science* possibilitaram análises em larga escala, sendo essenciais para a chamada *ciência da ciência*, campo que usa redes de citações e de colaboração para compreender a dinâmica científica (Reia *et al.*, 2025).

Nos últimos anos, a bibliometria tem sido utilizada para mapear áreas odontológicas. Pesquisas evidenciam que o número de publicações relacionadas a próteses dentárias, especialmente em tecnologias digitais como CAD/CAM e impressão 3D, cresce continuamente (Hsu; Chang, 2025). A partir da aplicação de métricas específicas, como o índice de disrupção (ID), torna-se possível quantificar essa dinâmica e identificar que, embora o volume da produção científica global continue a crescer de forma acelerada, a proporção de trabalhos verdadeiramente disruptivos têm se mantido relativamente estável ou até mesmo diminuído em

diversas áreas (Park; Leahey; Funk, 2023). Este fenômeno pode ser atribuído a fatores complexos, como as prioridades de agências de fomento, interesses políticos e comerciais, que exige mais tempo e esforço para que os cientistas dominem o conhecimento acumulado antes de poderem avançar as fronteiras da ciência (Jones, 2009; Bloom *et al.*, 2020).

A escolha de uma estratégia de busca ampliada, baseada nos descritores *Prosthodontics* e *Dental Prosthesis*, bem como em termos livres relacionados às principais modalidades protéticas, fundamenta-se na relevância clínica da Prótese Dentária e em seu impacto direto na saúde pública global. Nesse contexto, a adoção de um conjunto abrangente de termos permitiu captar a diversidade temática da literatura, contemplando desde próteses fixas e removíveis até reabilitações implantossuportadas. Tal abordagem possibilita compreender de forma mais precisa o panorama científico da área, identificar tendências de pesquisa, materiais e tecnologias mais investigadas, bem como lacunas de conhecimento, fornecendo subsídios para o desenvolvimento de inovações mais eficazes, acessíveis e duráveis.

2 REVISÃO DE LITERATURA

2.1 História da bibliometria

A bibliometria surgiu como resposta à necessidade de mensurar e organizar o crescente volume de informações científicas e seus antecedentes remontam ao final do século XIX, quando Paul Otlet e Henri La Fontaine desenvolveram sistemas de classificação e indexação mundial (Garcia *et al.*, 2023). Na década de 1920, Gross e Gross utilizaram contagens de citações para selecionar periódicos importantes para bibliotecas, antecipando o conceito de fator de impacto (Ganti; Yu; Jose, 2025). Algumas leis estatísticas, como a lei de Zipf (distribuição de frequências de palavras) e a lei de Bradford (distribuição de artigos em periódicos), ajudaram a entender a dispersão da informação científica (Koshevoy; Miton; Morin, 2023). Pritchard formalizou o termo bibliometria em 1969 e a definiu como a aplicação de métodos estatísticos à literatura (Cheng *et al.*, 2024). Eugene Garfield introduziu o “*Science Citation Index*” em 1955, permitindo rastrear citações e desenvolver o fator de impacto das revistas e Derek de Solla Price estudou a “lei do crescimento exponencial” da ciência e Merton descreveu o “efeito Mateus” (citações tendem a se acumular em trabalhos já citados) (Ganti; Yu; Jose, 2025). Nos anos 1990 e 2000, o avanço da informática e das redes digitais permitiu análises de co-citação, acoplamento bibliográfico e visualização de redes, consolidando a bibliometria como ferramenta robusta para investigar produção científica e políticas de pesquisa (Cheng *et al.*, 2024).

2.2 Bancos de dados na bibliometria

O acesso a dados confiáveis é fundamental para análises bibliométricas. PubMed, Scopus e Web of Science (WoS) são os bancos mais utilizados nas ciências da saúde e biomédicas (Moraes; Kafure, 2020). PubMed, mantido pela Biblioteca Nacional de Medicina dos Estados Unidos da América (EUA), é direcionado às ciências da vida e utiliza termos controlados (MeSH terms) para otimizar buscas; sua cobertura é mais restrita, mas inclui periódicos de alta qualidade, assim como Scopus, operado pela Elsevier, é multidisciplinar, abrange maior número de periódicos e oferece indicadores como índice h de autores e

análise de colaborações e WoS, da Clarivate, também é multidisciplinar e permite acompanhar citações ao longo do tempo, facilitando a análise de impacto de artigos e a construção de redes de co-citação (AlryalatL; Momani, 2019).

2.3 Análises bibliométricas

Os métodos bibliométricos incluem medidas de desempenho, como contagem de publicações, citações e índices de impacto (e.g., índice h), análise de co-citação, que examina com que frequência dois artigos são citados juntos, acoplamento bibliográfico, que conecta artigos que compartilham referências; análise de coocorrência de palavras-chave, e análise de autoria e colaboração (Ganti, Yu e Jose, 2025). Softwares como VOSviewer, CiteSpace e Bibliometrix permitem construir mapas de ciência e redes de colaboração, para tal é necessário planejar adequadamente a estratégia de busca, selecionar a janela temporal e escolher métricas apropriadas são fundamentais para garantir resultados confiáveis (Agarwal) *et al.*, 2015). A literatura também alerta que indicadores de citação devem ser usados com cautela, pois os autores citam por diversos motivos (apoio, crítica, contexto) e as métricas podem ser influenciadas por fatores como idioma e reputação (Belter, 2015).

2.4 Relevância na ciência

No contexto da ciência, o campo "ciência das ciências" utiliza redes de citações e colaboração para investigar como a ciência avança; essa abordagem ajuda a melhorar a eficiência da pesquisa, compreender trajetórias de carreira e promover inovação (Reia *et al.*, 2025). O uso de métodos de redes possibilita identificar padrões de colaboração, prever mudanças de tendência e orientar políticas científicas para fomentar inovações (Agarwal *et al.*, 2015).

2.5 Conceito do índice de disrupção

O índice de disrupção foi proposto por Funk e Owen-Smith em 2016 para quantificar o quanto uma publicação ou patente substitui o conhecimento precedente em sua área. O índice mede a proporção de citações que se referem apenas ao

artigo focal em comparação às citações que também incluem as referências desse artigo (Jiang; Liu, 2023). Wu et al. (2019) adaptaram o índice para publicações científicas, designando-o como D-index:

O cálculo considera três conjuntos:

- (i) artigos que citam somente o trabalho focal;
- (ii) artigos que citam o trabalho focal e pelo menos uma referência dele;
- (iii) artigos que citam apenas as referências sem citar o trabalho focal.

Ao definir uma produção como objeto de pesquisa a ter o índice de disrupção calculado, esta será identificada como artigo focal. Os criadores do índice de disrupção, Funk e Owen-Smith (2016) desenvolveram a seguinte equação (Equação 1) para que o cálculo da natureza do artigo focal pudesse ser feito:

$$CD_t = \frac{1}{n} \sum_{n=1}^n -2f_{it}b_{it} + f_{it} \quad (1)$$

Sendo:

f_{it} = 1 se a citação for ao artigo focal e 0 se não;

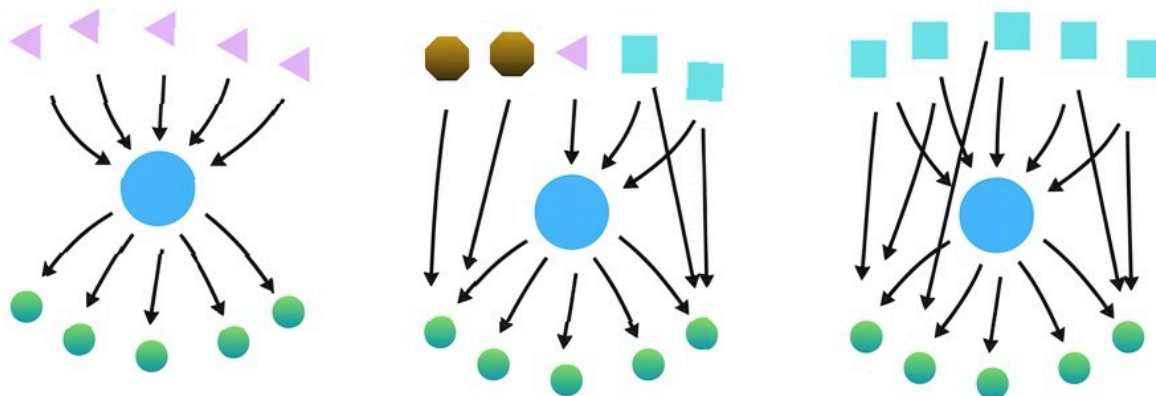
b_{it} = 1 se a citação citar o predecessor do artigo focal e 0 se não;

n_t = número total de citações ao artigo focal e/ou ao predecessor do artigo focal, no tempo t (soma de $i1$, $i2$ e $i3$).

Para isso, é importante entender que o DI é baseado em redes de citações, onde em cada rede pode existir três tipos de citações: $i1$ – as citações que citam apenas o artigo focal e nenhum de seus predecessores (triângulo rosa – Figura 1); $i2$ – as citações que citam o artigo focal e pelo menos um de seus predecessores (quadrado azul – Figura 1); $i3$ – as citações que citam apenas os predecessores do artigo focal e não o citam (hexágono marrom – Figura 1).

Figura 1- Diagrama esquemático das relações de citação entre um artigo focal e seus trabalhos predecessores. O artigo focal é representado pelo círculo azul e os artigos predecessores pelos círculos verdes. Triângulos rosa correspondem a citações apenas ao artigo focal ($i1$); quadrados azul indicam citações ao artigo focal e a pelo

menos um artigo predecessor (i_2); enquanto hexágonos marrom representam citações exclusivamente aos artigos predecessores (i_3).



Wu, Wang e Evans (2019) foram os primeiros a aplicar o índice de disrupção para dados de publicações científicas. Os autores se basearam nos princípios apresentados por Funk e Owen-Smith (2016), e apresentaram a seguinte fórmula simplificada (Equação 2):

$$DI = \frac{i_1 - i_2}{i_1 + i_2 + i_3}$$

(2)

Quando muitos artigos posteriores citam apenas o trabalho focal, diz-se que ele exerce papel disruptivo ao romper com a tradição. Essa abordagem utiliza a estrutura local da rede de citações para avaliar se um artigo substitui os trabalhos que citou e, portanto, introduz uma linha de pesquisa independente (Petersen; Arroyave; Pammolli, 2024). O ID varia de -1 (consolidante) a $+1$ (disruptivo), em que valores negativos indicam que os trabalhos futuros continuam citando as referências originais, enquanto valores positivos sugerem substituição de conhecimento e potencial inovação (Funk; Owen-Smith, 2016).

2.6 Limitações do índice de disrupção

Apesar de seu potencial para identificar contribuições inovadoras, o ID apresenta limitações inerentes (Poder, 2022). Como toda métrica baseada em citações, ele é sensível ao tempo: o número de citações varia entre campos, idiomas

e períodos, e bases de dados têm cobertura imperfeita, especialmente para publicações antigas (Leibel; Bornmann, 2024). Isso pode inflar artificialmente os valores de disrupção em estudos históricos ou reduzir a comparabilidade entre áreas. Variações no tipo de documento e na disponibilidade de palavras-chave (por exemplo, MeSH terms) também afetam o cálculo, pois índices baseados em MeSH não podem ser aplicados a publicações sem esses termos (Funk; Owen-Smith, 2016; Leibel; Bornmann, 2024). Outra crítica refere-se à convergência do índice para zero à medida que listas de referências se tornam mais longas: o aumento contínuo do número de referências e das taxas de autocitação gera “inflação de citações” e altera a densidade das redes de citações, tornando o ID temporalmente inadequado para análises comparativas ao longo de décadas (Petersen; Arroyave; Pammolli, 2024). Além disso, estudos apontam que o ID depende da janela de tempo escolhida e pode sofrer inconsistências quando o número de referências citadas pelo artigo focal é pequeno, ou seja, recomenda-se utilizar janelas maiores (por exemplo, três anos) e excluir artigos com poucas citações e referências para obter resultados mais confiáveis (Leibel; Bornmann, 2024). A própria validação do ID enfrenta desafios: julgamentos de especialistas podem estar enviesados, pois eles têm acesso a contagens de citações e podem associar novidade à disrupção, confundindo conceitos distintos (Funk; Owen-Smith, 2016).

2.7 Aplicação do índice de disrupção

Desde sua criação, o ID foi aplicado em diversos campos para identificar artigos transformadores. Funk e Owen-Smith (2016) analisaram patentes tecnológicas, enquanto Wu et al. (2019) utilizaram a métrica em artigos científicos e observaram que equipes pequenas tendem a produzir pesquisas mais disruptivas do que equipes grandes, que geralmente desenvolvem trabalhos de consolidação. Estudos aplicaram o ID para determinar trabalhos disruptivos em cirurgias craniofaciais, pediatria, biologia sintética, radiologia e outras áreas (Leibel; Bornmann, 2024). Leydesdorff, Tekles e Bornmann (2021) investigaram artigos disruptivos na revista *Scientometrics*, destacando a importância de ajustar a janela de citações para melhorar a validade do índice. O ID também foi expandido para nível de revistas, o *Journal Disruption Index*, com proposta de classificar periódicos

de acordo com a capacidade de publicar pesquisas inovadoras, revelando correlação moderada com fatores de impacto tradicionais (Jiang; Liu, 2023).

2.8 Estudos bibliométricos na odontologia

A bibliometria vem sendo aplicada a diversas especialidades odontológicas para identificar tendências e contribuições relevantes. Nath *et al.* (2024) analisaram 3 428 artigos publicados na revista *Community Dentistry and Oral Epidemiology* usando dados da Scopus e observaram aumento progressivo no número de artigos e citações, bem como países e autores mais produtivos (Estados Unidos, Reino Unido, Brasil). Khan *et al.* (2023) examinaram os 20 artigos mais citados sobre complicações protéticas de implantes dentários utilizando Web of Science e Google Scholar e a contagem de citações variou de 6 391 a 315, e os estudos mais citados eram revisões ou estudos prospectivos, sem ensaios clínicos randomizados. Rashid e Dudley (2024) conduziram uma análise bibliométrica sobre técnicas de harmonização de cor (*shade-matching*) em prótese, identificando correlação entre contagem de citações e o fator de impacto das revistas (a revista *Journal of Prosthetic Dentistry* e autores dos Estados Unidos foram os mais proeminentes). Hu (2025) realizou um mapeamento do conhecimento sobre inteligência artificial (IA) em prótese dentária com 214 artigos do Web of Science (Estados Unidos, China e Itália foram os países com mais publicações), e palavras-chave como “*accuracy*”, “*placement*” e “*dental implants*” foram frequentes, indicando tendência de uso de IA para diagnóstico e fabricação de próteses.

Especificamente na área de prótese dentária, algumas análises bibliométricas foram realizadas para mapear tendências tecnológicas e impactos sobre qualidade de vida. Uma pesquisa de mapeamento bibliométrico identificou que os estudos sobre qualidade de vida e próteses dentárias são escassos, mas concentrados em periódicos de alto impacto; Brasil e Alemanha foram os países com maior produção, e *Journal of Prosthetic Dentistry* publicou o maior número de artigos (Espinoza-Carhuancho *et al.*, 2023). Hsu e Chang (2025) examinaram duas décadas de literatura sobre tecnologias digitais para próteses removíveis usando a Web of Science e o estudo encontrou 457 documentos entre 2004 e 2023, com crescimento significativo de publicações e citações, predominância dos temas CAD/CAM e impressão 3D e liderança da *Journal of Prosthetic Dentistry*. Del Rio Silva *et al.*

(2025) avaliaram 25 anos de pesquisas sobre próteses unitárias fabricadas por CAD/CAM e relataram tendência ascendente de publicações, especialmente com manufatura aditiva e materiais estéticos; próteses cimentadas foram as mais comuns e a digitalização das impressões cresceu ao longo do tempo. Além disso, estudos bibliométricos sobre próteses parciais removíveis ressaltaram que temas relacionados à qualidade de vida, função mastigatória e sobrevivência de próteses concentram-se principalmente em países com alta produção científica e indicam lacunas em pesquisas clínicas controladas (Al-Rafee, 2020; Techapiroontong *et al.*, 2022; Venkatesan *et al.*, 2022).

3 PROPOSIÇÃO

3.1 Objetivo Geral

Categorizar a natureza da inovação científica na literatura de Prótese Dentária ao longo do tempo, utilizando o índice de disrupção para diferenciar entre publicações de caráter consolidador e disruptivo.

3.2 Objetivos Específicos

1. Calcular o índice de disrupção dos artigos relacionados à Prótese Dentária indexados na base de dados PubMed no período de 1900 a 2025;
2. Analisar a evolução temporal da disrupção científica na literatura de Prótese Dentária por meio da distribuição anual do índice de disrupção;
3. Caracterizar a produção científica quanto à distribuição geográfica, identificando os países com maior número de publicações e seus respectivos valores médios de ID;
4. Investigar a relação entre características bibliométricas dos artigos e o ID, incluindo número de autores, número de referências e número total de citações;
5. Identificar os artigos e periódicos, associados ao seu modelo de acesso a comunidade científica, com maior potencial disruptivo na literatura de Prótese Dentária;
6. Avaliar a relação entre financiamento da pesquisa e o ID, considerando as principais agências de fomento associadas às publicações;

7. Analisar os MeSH terms e termos da estratégia de busca associados às publicações, identificando sua frequência e o ID médio relacionado a cada termo.

4 METODOLOGIA

4.1 CARACTERIZAÇÃO BIBLIOMÉTRICA

4.1.1 Seleção de dados

O presente trabalho configura-se como um estudo observacional de caráter retrospectivo. A busca inicial foi composta por todos os artigos indexados na base de dados PubMed que correspondiam a palavras-chave relacionadas à prótese dentária, publicados entre 1900 e 2025.

4.1.2 Busca dos dados

Os identificadores dos artigos (PMIDs) foram obtidos por meio do serviço Application Programming Interface (API)¹, E-utilities da API do PubMed (NCBI, 2024), que permite recuperar informações diretamente do banco de dados de forma automatizada. Para isso, foi utilizado o código do Repositório SciDisruptor-PubMed (Apolinário et al., 2025) (Anexo A), modificado com auxílio do ChatGPT (OpenAI, 2026) para programação, para buscar todos os artigos recuperados da base de dados PubMed que correspondiam às palavras-chave ("Prosthodontics"[MeSH] OR "*Dental Prosthesis*"[MeSH]) OR (*prosthodontics*[Title/Abstract] OR "*fixed prosthodontics*"[Title/Abstract] OR "*removable prosthodontics*"[Title/Abstract] OR "*implant prosthodontics*"[Title/Abstract] OR "*complete denture*"[Title/Abstract] OR "*partial denture*"[Title/Abstract] OR "*overdenture*"[Title/Abstract] OR "*implant-supported prosthesis*"[Title/Abstract] OR "*fixed dental prosthesis*"[Title/Abstract] OR "*removable dental prosthesis*"[Title/Abstract] OR "*prosthetic rehabilitation*"[Title/Abstract]) com data de publicação compreendida entre 1900 e 2025. O código foi executado no Notebook Macbook (Apple, Los Altos, Califórnia, EUA, 2024), pelo aplicativo Visual Studio Code (Microsoft, Albuquerque, Novo México, EUA) ambiente

¹ Interface de programação que facilita o acesso às bases de dados da NCBI.

que organiza o código em células e facilita a execução e inspeção dos resultados. Os PMIDs recuperados foram separados por ano e salvos em subpastas no formato .csv².

A partir dos PMIDs, efetuou-se o download dos metadados completos de cada artigo pela mesma API. Esse processo, implementado em outro código (Anexo B), também foi executado no Notebook Macbook (Apple, Los Altos, Califórnia, EUA, 2024) e gerou arquivos .xml armazenados nas subpastas correspondentes ao ano de publicação. Cada arquivo XML³ contém informações detalhadas: dados gerais da publicação (identificador, título, resumo e datas), dados do periódico (nome, ISSN, volume, edição e paginação), dados dos autores (nomes, afiliações e ORCID), dados de financiamento (identificador, agência e país), descritores (MeSH terms), referências citadas e citações recebidas, histórico editorial (datas de submissão, aceitação e publicação), tipo de publicação, idioma e, quando disponível, informações sobre a revisão por pares.

Com base na estrutura de dados fornecida pelo Repositório SciDisruptor-PubMed (Apolinário et al., 2025), que foi concebido para padronizar o acesso a diversos bancos bibliométricos e disponibilizar um conjunto de métricas e algoritmos para mapeamento de redes de citações, desenvolveu-se um código (Anexo C) para tratar os arquivos XML obtidos. Esse processamento extraiu, entre outros, os elementos necessários ao cálculo do ID, como identificadores dos artigos, títulos, ano de publicação e listas de referências. O código C exportou os dados em formato .csv. no Excel (Microsoft Corporation, Redmond, WA, Estados Unidos). Todo o procedimento foi executado no ambiente Notebook Macbook (Apple, Los Altos, Califórnia, EUA, 2024), que permite organizar e executar códigos Python em células.

O código apresentado no Anexo D foi utilizado para a construção da rede de citações empregada neste estudo. Para sua elaboração, adotou-se uma janela de citação de cinco anos, considerando que períodos muito curtos podem comprometer a confiabilidade do índice de disrupção (Leibel; Bornmann, 2024). A partir da rede de citações gerada, foi possível identificar quais artigos receberam citações, bem como quantificar o número de citações de cada tipo (A, B e C).

² Tipo de arquivo de texto usado para armazenar dados de maneira estruturada, organizados em forma de linhas e colunas.

³ eXtensible Markup Language é uma linguagem de marcação usada para armazenar e transportar informações, especialmente na troca de dados entre sistemas diferentes

4.1.3 Cálculo de DI

O Anexo E foi utilizado para calcular o Índice de Disrupção, reproduzindo a lógica introduzida por Wu, Wang e Evans (2019), na qual a disruptividade de um artigo se baseia nas suas referências e nas citações a ele destinadas. Essa lógica foi implementada de acordo com a Equação (2), permitindo calcular o índice de disrupção para cada publicação analisada.

$$DI = A - B/A + B + C \quad (2)$$

Onde:

DI = Pontuação de disrupção do artigo focal;

A = número de artigos subsequentes que citam o artigo focal sem que as referências do artigo focal sejam citadas simultaneamente;

B = número de artigos subsequentes que citam o artigo focal e pelo menos uma das referências do artigo focal;

C = número de artigos subsequentes que citam uma das referências do artigo focal, mas não o artigo focal em si.

O resultado varia de -1 a +1, onde os artigos com pontuação negativa (<0) são considerados de natureza consolidante e os de pontuação positiva (>0) são considerados de natureza disruptiva.

O código para o cálculo do índice (Anexo E) foi aplicado junto ao conjunto de dados obtidos pelos Anexos C e a rede de citação produzida no anexo D, incluindo as especificações dos artigos focais, a janela de citação de 5 anos. Como resultado, foi obtida uma planilha em formato .csv, contendo todos os PMIDs dos artigos que tiveram ID calculado. Além do resultado do índice desses artigos, foi apresentado o número de total de citações, e número de citações para A, B e C separadamente dentro da janela de citação de 5 anos.

A amostra final do estudo foi composta por artigos com índice de disrupção calculado, que foram publicados até o ano de 2020 e apresentavam, necessariamente, pelo menos uma citação e referências indexadas no PubMed.

4. 2 ANÁLISES CORRELACIONAIS

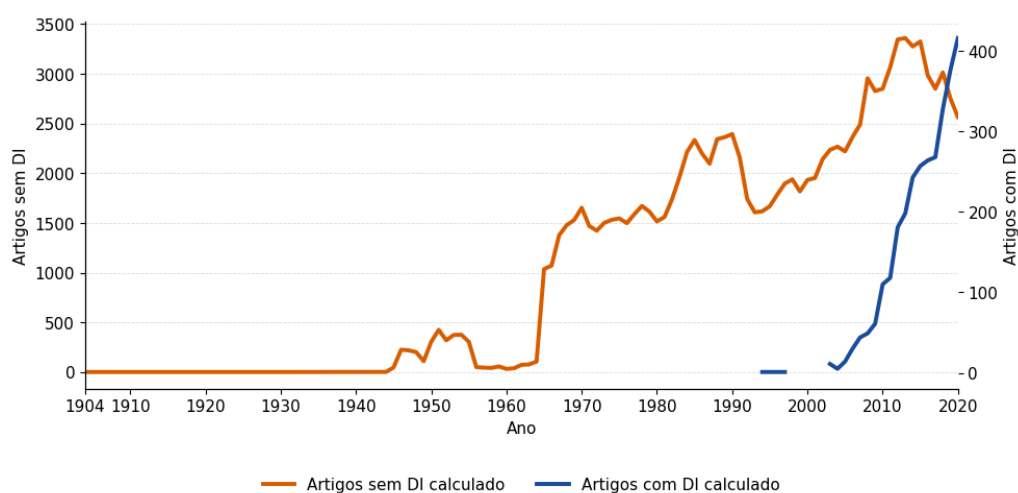
Realizou-se a caracterização da produção científica segundo país de origem, número de autores por artigo, presença de agências de financiamento. Essas informações foram extraídas diretamente dos metadados da base, com padronização de grafias e consolidação de categorias quando necessário. Os resultados foram sintetizados em tabelas, gráficos e representações visuais, incluindo contagens absolutas, proporções, médias, correlações e séries temporais por ano, década ou períodos agregados. Esse conjunto de análises permitiu mapear a estrutura da produção científica, identificar concentrações geográficas, padrões de autoria, tendências temáticas e subáreas com maior potencial disruptivo, oferecendo subsídios interpretativos para a compreensão do comportamento do índice de disrupção ao longo do tempo. O número de autores de cada artigo da amostra foi extraído dos dados. O teste de correlação de Spearman, usando a variável de Número de Autores dos Artigos versus ID, foi feito para analisar se há relação estatisticamente significativa; O número de artigos na amostra por ano, a média de ID anual, o número total de citações e de referências dos artigos da amostra. O teste de Correlação de Pearson (r) foi utilizado para avaliar a correlação entre: Total de Citações X Índice de disrupção; Total de Referências X ID; Número de artigos publicados no ano X ID.

As informações relacionadas ao número total de citações dos artigos incluídos na amostra foram obtidas por meio da plataforma iCite (<https://icite.od.nih.gov/analysis>). Para essa etapa, foi utilizada a lista de identificadores PubMed (PMIDs) correspondente aos artigos selecionados, apresentada no Anexo A, que foi inserida no sistema para a extração dos dados bibliométricos. Foram identificados os dez periódicos e os dez artigos mais disruptivos. A partir dessas duas listas, procedeu-se à verificação manual do modelo de acesso dos periódicos correspondentes. Para essa etapa, cada periódico foi classificado quanto ao tipo de acesso como acesso aberto (*open access*), híbrido ou acesso fechado (*subscription-based*). Essa classificação foi realizada por meio da consulta direta aos sites oficiais dos periódicos e às informações disponibilizadas pelas editoras.

5 RESULTADOS

Do total de 141.829 artigos, encontrados no dia 16 de dezembro de 2025, 390.473 citações foram analisadas. A partir delas, o Índice de Disrupção pôde ser calculado para 2.980 publicações (2,1%), enquanto 138.849 artigos (97,9%) não atenderam aos requisitos necessários para esse cálculo. Embora tenham sido identificados artigos a partir de 1904, houve aumento acentuado a partir da década de 1960. Enquanto os artigos com Índice de Disrupção calculado aparecem apenas nos anos mais recentes da série, com aumento gradual principalmente a partir de 2003 (Figura 2).

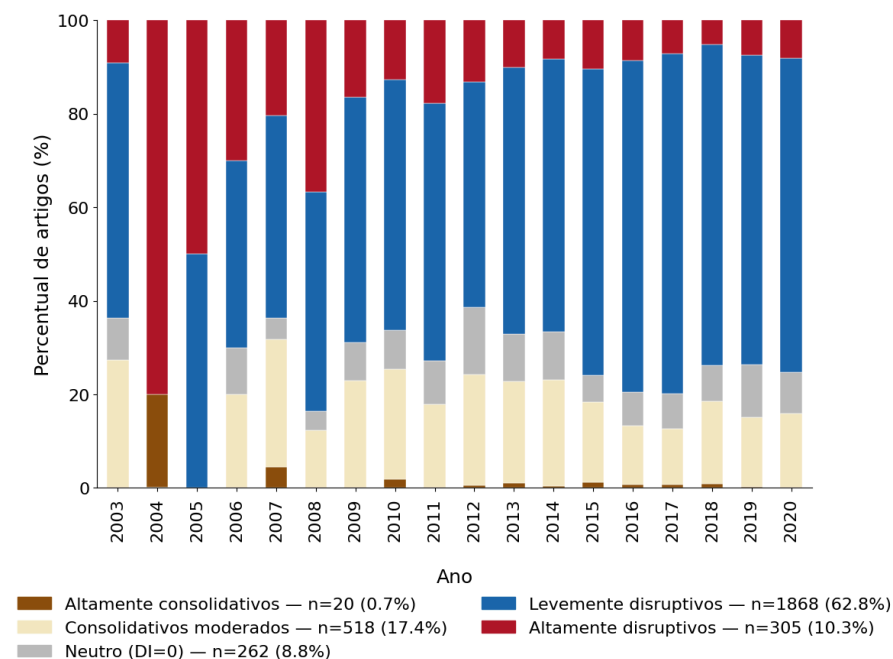
Figura 2 - Evolução temporal do número de artigos em prótese dentária com e sem cálculo do Índice de Disrupção na base PubMed.



Fonte: Autora (2026)

Entre 2003 e 2020 (Figura 3), 73,1% dos artigos apresentam caráter disruptivo, indicando predominância de contribuições que introduzem novas trajetórias de pesquisa. Desses, predominam artigos levemente disruptivos ($0 < ID < 0,5$), que representam 62,8% da amostra ($n = 1868$). Em seguida, aparecem os consolidativos moderados (17,4%; $n = 518$), os altamente disruptivos (10,3%; $n = 305$), os neutros (8,8%; $n = 262$) e, em menor proporção, os altamente consolidativos (0,7%; $n = 20$). Além disso, os artigos altamente disruptivos estão presentes em todos os anos da série.

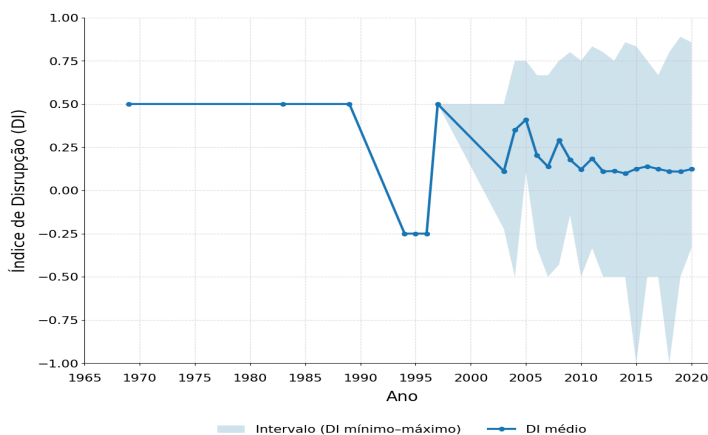
Figura 3 - Distribuição anual das categorias do Índice de Disrupção dos artigos publicados entre 2003 e 2020.



Fonte: Autora (2026)

Durante todo o período, o valor do ID médio permaneceu positivo, permanecendo constante entre os valores de 0 a +0,25 a partir de 2009. Observa-se também variação anual entre os valores mínimo e máximo de ID, onde os valores máximos aproximaram-se de 1, evidenciando a presença de artigos altamente disruptivos, enquanto valores mínimos negativos indicam a ocorrência de artigos consolidativos (Figura 4).

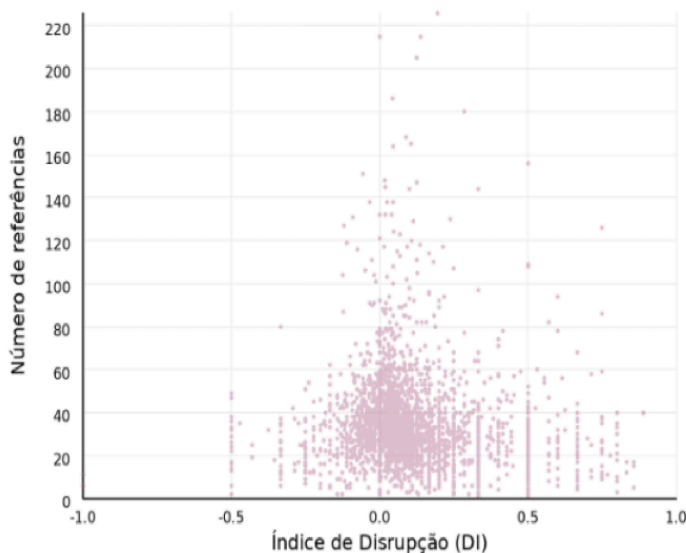
Figura 4 - Variação temporal do Índice de Disrupção médio e do intervalo entre valores mínimos e máximos observados na amostra.



Fonte: Autora (2026)

A correlação entre o Índice de Disrupção e o número de referências foi fraca e negativa ($r = -0,135$), indicando que artigos com maior número de referências tendem levemente a apresentar valores menores de disrupção (Figura 5).

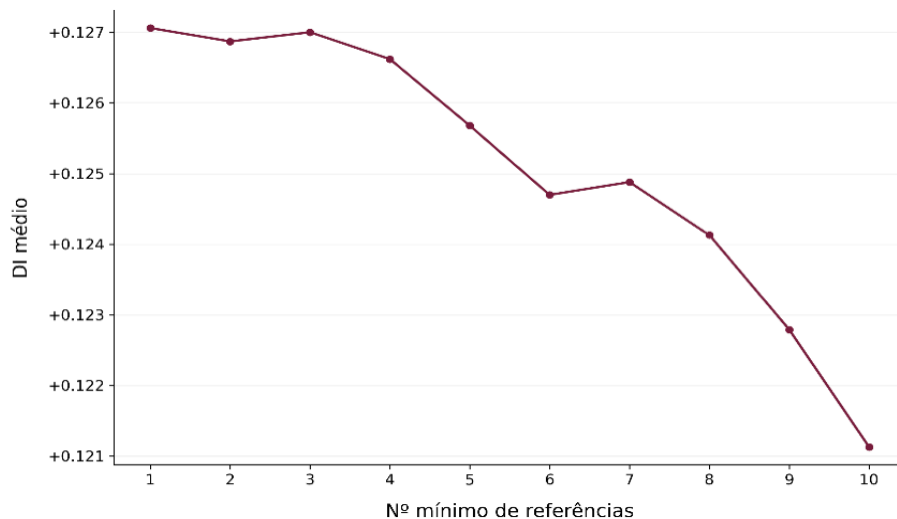
Figura 5 – Gráfico de dispersão do número de referências X ID



Fonte: Autora (2026)

Na análise de sensibilidade, o aumento do número mínimo de referências reduziu o número de artigos incluídos de 2.980 (≥ 1 referência) para 2.833 (≥ 10 referências). Nesse mesmo intervalo, o ID médio diminuiu discretamente de 0,12706 para 0,12113, a mediana variou minimamente de 0,077 para 0,071, e o desvio-padrão de 0,209 para 0,201, mostrando que o centro de distribuição não é afetado e a forma geral de distribuição é preservada. Na análise estratificada por ano, todos os períodos permaneceram representados, sem diferenças estatisticamente significativas nos valores de ID entre os cortes avaliados. Esses achados indicam que a adoção de limiares mínimos para o número de referências teve impacto pequeno sobre os valores de ID, sem introduzir distorções temporais relevantes (Figura 6).

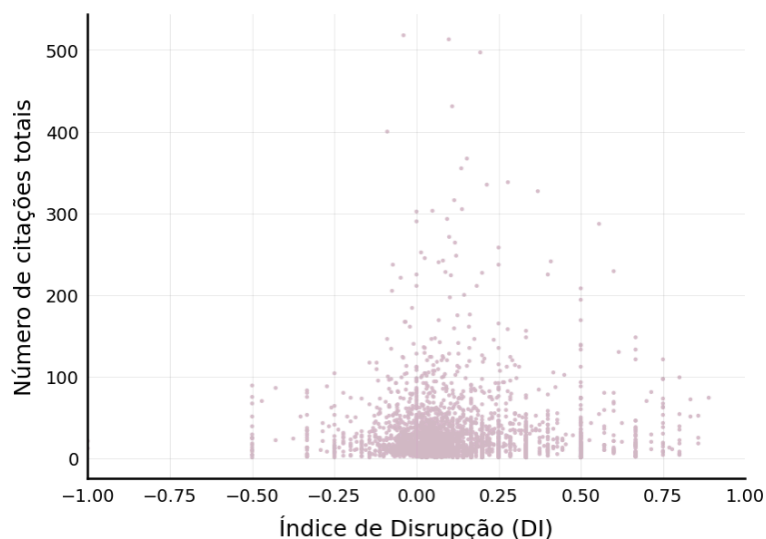
Figura 6 - Análise de sensibilidade do Índice de Disrupção médio em função do número mínimo de referências considerado.



Fonte: Autora (2026)

Por outro lado, a correlação entre o Índice de Disrupção e o número total de citações recebidas foi praticamente nula ($r = -0,001$), sugerindo ausência de associação entre o nível de disrupção e o impacto medido pelo total de citações (Figura 7).

Figura 7 - Relação entre o número total de citações recebidas pelos artigos e o Índice de Disrupção.



Fonte: Autora (2026)

Entre os artigos mais citados da amostra (Tabela 1), observam-se valores tanto positivos quanto negativos de Índice de Disrupção, indicando a presença de contribuições com caráter disruptivo e consolidativo. Esses artigos concentram-se

majoritariamente entre 2008 e 2018, refletindo maior tempo de exposição na literatura . Por sua vez, os artigos com maiores valores de ID (Tabela 2) apresentam predominantemente valores elevados e positivos ($\geq +0,80$) e distribuem-se entre 2009 e 2020 . Esse padrão indica que contribuições altamente disruptivas não estão restritas a um recorte temporal específico, ocorrendo de forma contínua ao longo do tempo, e que a disrupção e o acúmulo de citações podem seguir dinâmicas distintas.

Tabela 1 - Artigos mais citados na amostra de prótese dentária, com respectivos valores de Índice de Disrupção.

Posição	Artigo	Ano	Citações	DI
1	Longevity of posterior composite restorations: a systematic review	2014	1190	-0,039
2	3D printing in dentistry	2016	1047	+0,099
3	Guided bone regeneration: materials and biological mechanisms	2017	987	+0,194
4	A review on the wettability of dental implant surfaces	2014	932	+0,109
5	Intraoral scanners in dentistry: a review of the current literature	2017	905	-0,089
6	Degradation, fatigue, and failure of resin dental composite materials	2008	873	+0,154
7	Impact of Dental Implant Surface Modifications on Osseointegration	2016	851	+0,137
8	Trends in Dental Implant Use in the U.S., 1999–2016	2018	812	+0,278
9	Recent advances and developments in composite dental restorative materials	2010	789	+0,214
10	3D printed versus conventionally cured provisional crowns	2017	742	+0,370

Fonte: Autora (2026)

Tabela 2 - Artigos com maiores valores de Índice de Disrupção na amostra, com respectivos números de citações.

Posição	Artigo	Ano	Citações	DI
1	The Use of Titanium Mesh in Guided Bone Regeneration: A Systematic Review.	2019	74	+0,889
2	Prosthetic rehabilitation after orbital exenteration: a case series.	2014	18	+0,857
3	Immediate dental implants in fibula free flaps to reconstruct the mandible: A pilot study of the short-term effects on radiotherapy for patients with head and neck cancer.	2020	25	+0,857
4	Effects of disinfection and sterilization on the dimensional changes and mechanical properties of 3D printed surgical guides for implant therapy - pilot study.	2020	52	+0,857
5	A comparative study of the accuracy and reliability of multidetector computed tomography and cone beam computed tomography in the assessment of dental implant site dimensions.	2011	72	+0,833
6	Rehabilitation after mandibular resection and bilateral nerve grafting using implant-supported prosthesis: a case report.	2015	51	+0,833
7	Forty-year history of the dental implant.	2009	99	+0,800
8	In vitro evaluation of the accuracy of surgical guide dimensions in square and tapered implants.	2010	35	+0,800
9	A retrospective multicenter evaluation of 1,022 maxillary anterior single-unit implant restorations	2011	54	+0,800

10	Association between missing posterior teeth and body mass index in adults.	2018	22	+0,800
----	--	------	----	--------

Fonte: Autora (2026)

Os dez periódicos com maior Índice de Disrupção médio apresentaram valores variando entre +0,75 e +0,62 (Tabela 3), indicando elevada capacidade de redirecionamento das redes de citação entre os artigos publicados nesses veículos. Observa-se que, entre os periódicos identificados, apenas o *Journal of Dental Education* está diretamente relacionado ao campo da odontologia, apresentando ID médio de +0,75. Os demais periódicos pertencem a áreas adjacentes das ciências da saúde e biomédicas, evidenciando a presença de artigos potencialmente disruptivos em periódicos de escopo interdisciplinar. Entre os dez periódicos com maior Índice de Disrupção médio, observou-se distribuição equilibrada entre modelos de publicação open access e híbrido, sem predominância de periódicos de acesso fechado. Esse padrão sugere que a produção científica potencialmente mais disruptiva relacionada à prótese dentária encontra-se disseminada em diferentes modelos de comunicação científica, indicando que o potencial de disrupção não está restrito a um único regime de acesso editorial.

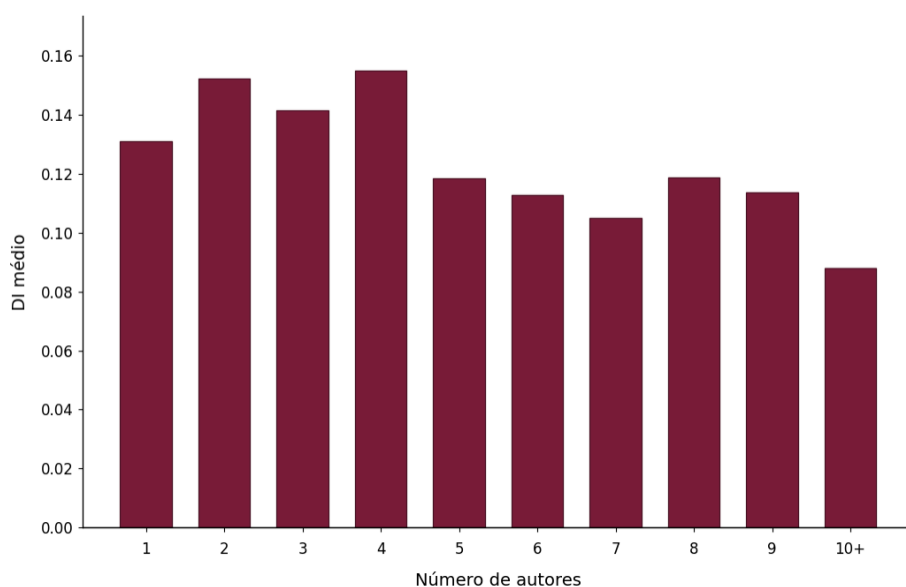
Tabela 3 - Periódicos com maior Índice de Disrupção médio e respectivos modelos de acesso.

Posição	Periódico	DI médio	Tipo de acesso
1	Molecular Therapy: The Journal of the American Society of Gene Therapy	+0,750	Acesso híbrido
2	Journal of the Royal Society of Medicine	+0,750	Acesso híbrido
3	Journal of Dental Education	+0,750	Acesso híbrido
4	Orphanet Journal of Rare Diseases	+0,667	Open Access
5	Journal of Otolaryngology – Head & Neck Surgery	+0,667	Open Access
6	Journal of Veterinary Dentistry	+0,667	Acesso híbrido
7	American Journal of Public Health	+0,667	Acesso híbrido
8	Indian Journal of Plastic Surgery	+0,667	Open Access
9	Archives of Rehabilitation Research and Clinical Translation	+0,667	Open Access
10	The Pan African Medical Journal	+0,625	Open Access

Fonte: Autora (2026)

A média do índice de disrupção foi maior entre os artigos com 2 autores (ID = +0,152) e 4 autores (ID = +0,154), seguido pelos grupos com 3 autores (ID = +0,141) e 1 autor (ID = +0,131) (Figura 8). A partir de 5 autores, os valores médios foram menores, variando de +0,104 a +0,118, e o menor ID médio foi observado nos artigos com 10 ou mais autores (ID = +0,087). De forma consistente, a análise de correlação mostrou associação negativa muito fraca entre o número de autores e o ID (Pearson: $r = -0,072$; $p < 0,001$; Spearman: $\rho = -0,072$; $p < 0,001$), indicando tendência discreta de redução do índice com o aumento do número de autores.

Figura 8 - Relação entre o número de autores por artigo e o Índice de Disrupção médio.



Fonte: Autora (2026)

A análise dos países com maior número de publicações revelou concentração da produção científica em nações europeias e nos Estados Unidos. Entretanto, a magnitude do Índice de Disrupção variou entre os países (Tabela 4). Enquanto Inglaterra, Estados Unidos e Alemanha apresentaram elevados volumes de produção, seus valores médios de ID permaneceram moderados ($\approx 0,10-0,12$). Por outro lado, países como Índia, Brasil, Japão e Coreia do Sul demonstraram maiores valores médios de disrupção, sugerindo maior potencial inovador relativo em comparação aos países tradicionalmente mais produtivos.

Tabela 4 - Distribuição dos países com maior número de publicações e respectivos valores de Índice de Disrupção médio e mediano.

Posição	País	Nº de publicações	DI médio	Mediana DI
1	Inglaterra	699	+0,111	+0,067
2	Estados Unidos	624	+0,125	+0,080
3	Alemanha	594	+0,106	+0,056
4	Suíça	184	+0,079	+0,033
5	Índia	183	+0,242	+0,200
6	Brasil	149	+0,170	+0,143
7	Espanha	109	+0,109	+0,100
8	China	58	+0,138	+0,087
9	Coreia do Sul	58	+0,157	+0,083
10	Japão	42	+0,171	+0,127

Fonte: Autora (2026)

A análise das agências de fomento (Tabela 5) revelou predominância de instituições vinculadas ao *National Institutes of Health* (NIH), especialmente do *National Institute of Dental and Craniofacial Research* (NIDCR), responsável pelo maior número de publicações financiadas na área de prótese dentária. Esse padrão evidencia a forte centralização do financiamento científico em organismos norte-americanos e reforça o papel dessas instituições na sustentação da produção científica do campo. Todavia, apesar do elevado volume de estudos financiados por essas agências, os valores médios do Índice de Disrupção observados foram relativamente moderados. Destaca-se, entretanto, a presença da *Fundação de Amparo à Pesquisa do Estado de São Paulo* (FAPESP) como a única agência não vinculada ao sistema norte-americano.

Tabela 5 - Principais agências de fomento associadas às publicações e respectivos valores de Índice de Disrupção médio e mediano.

Posição	Agência de fomento	Nº de artigos	DI médio	Mediana DI
---------	--------------------	---------------	----------	------------

1	NIDCR – National Institute of Dental and Craniofacial Research (NIH*)	654	+0,079	+0,033
2	NCATS – National Center for Advancing Translational Sciences (NIH*)	33	+0,104	+0,036
3	NCRR – National Center for Research Resources (NIH*)	33	+0,302	+0,143
4	NIAMS – National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIH*)	30	+0,206	+0,107
5	NIA – National Institute on Aging (NIH*)	19	+0,197	+0,136
6	FAPESP – Fundação de Amparo à Pesquisa do Estado de São Paulo	14	+0,062	+0,053
7	NCI – National Cancer Institute (NIH*)	13	+0,413	+0,500
8	NIEHS – National Institute of Environmental Health Sciences (NIH*)	12	-0,223	-0,238
9	NIAID – National Institute of Allergy and Infectious Diseases (NIH*)	11	+0,188	+0,200
10	NHLBI – National Heart, Lung, and Blood Institute (NIH*)	10	+0,076	+0,042

* *National Institutes of Health* (NIH)

Fonte: Autora (2026)

A Tabela 6 apresenta a frequência de ocorrência dos principais termos utilizados na estratégia de busca, bem como o índice médio de interrupção associado a cada um deles. Observa-se que o termo “Dental Prosthesis” (ID= 0,095) foi o mais prevalente entre os artigos analisados, refletindo seu caráter mais amplo e abrangente dentro da literatura da prótese dentária. Em contraste, termos mais específicos, como os relacionados a prótese removível (“*removable prosthodontics*” ID= +0,500, “*removable dental prosthesis*” ID=+0,370), próteses parciais (“*partial denture*” DI=+0,202). Embora menos frequentes, apresentaram valores de índice de interrupção relativamente mais elevados, sugerindo que pesquisas nesses subcampos podem estar associadas a contribuições científicas mais inovadoras ou capazes de alterar padrões de citação existentes.

Tabela 6- Frequência dos termos utilizados na estratégia de busca e respectivos valores de Índice de Disrupção médio.

Termo da busca	Ocorrências	DI médio
Prosthodontics	41	+0,155
Dental Prosthesis	559	+0,095
Fixed prosthodontics	9	+0,123
Removable prosthodontics	1	+0,500
Implant prosthodontics	3	+0,030
Complete denture	100	+0,196
Partial denture	94	+0,202
Overdenture	120	+0,095
Implant-supported prosthesis	11	+0,179
Fixed dental prosthesis	18	+0,154
Removable dental prosthesis	5	+0,370
Prosthetic rehabilitation	34	+0,244

Fonte: Autora (2026)

6 DISCUSSÃO

O presente estudo analisou a natureza da produção científica na área de prótese dentária por meio da aplicação do índice de disrupção, uma métrica bibliométrica baseada nas referências e citações de um estudo redes, que permite avaliar se ele tende a redirecionar ou consolidar o conhecimento existente em determinado campo científico (Funk; Owen-Smith, 2016; Wu; Wang; Evans, 2019). Diferentemente de indicadores tradicionais baseados apenas em contagem de citações, o ID considera a estrutura das citações e busca identificar se um trabalho é citado independentemente das referências que utilizou ou se continua vinculado ao corpo de conhecimento previamente estabelecido (Jiang; Liu, 2023).

Os resultados indicaram que, embora a literatura em prótese dentária tenha apresentado crescimento expressivo ao longo do tempo, especialmente a partir da década de 2010, o valor médio do índice de disrupção permaneceu relativamente estável e levemente positivo ao longo dos anos. Esse comportamento é consistente com achados observados em outros campos científicos, nos quais o aumento do volume de publicações não necessariamente se traduz em maior disrupção científica (Park, Leahey; Funk, 2023). Estudos na área da 'ciência das ciências' sugerem que a expansão do conhecimento pode tornar mais complexa a produção de avanços que redefinem paradigmas existentes, uma vez que pesquisadores precisam dominar um corpo crescente de literatura antes de propor novas abordagens (Jones, 2009; Bloom *et al.*, 2020).

No presente estudo, a maioria dos artigos foi classificada como levemente disruptiva, enquanto artigos altamente disruptivos representaram uma proporção menor da amostra. Outro achado relevante foi a ausência de correlação entre o número total de citações e o índice de disrupção, indicando que essas métricas capturam dimensões distintas da produção científica. Na presente amostra, observou-se que artigos altamente citados apresentaram valores variados de ID, incluindo valores negativos e positivos, enquanto os artigos com maiores valores de disrupção concentraram baixo número de citações. Esse padrão é consistente com a proposta teórica do índice de disrupção, que não mede o volume de reconhecimento de um artigo mas sim o quanto ele foi capaz de romper conceitos

pré-estabelecidos na comunidade científica (Funk; Owen-Smith, 2016; Wu, Wang e Evans, 2019).

A análise da relação entre o número de referências e o índice de disrupção revelou correlação negativa baixa, indicando que artigos com listas de referências mais extensas tendem a apresentar valores levemente menores de ID. Esse padrão pode refletir, em parte, o fenômeno conhecido como inflação de citações, no qual o aumento progressivo do número de referências nas publicações científicas altera a densidade das redes de citação e pode influenciar métricas baseadas nessas estruturas (Petersen; Arroyave; Pammolli, 2024). Estudos metodológicos apontam que o crescimento contínuo do número de referências pode favorecer valores mais próximos de zero no ID, uma vez que aumenta a probabilidade de citações simultâneas ao artigo focal e aos seus predecessores (Leibel; Bornmann, 2024; Petersen, Arroyave; Pammolli, 2024).

A relação entre o número de autores e o índice de disrupção também apresentou associação negativa fraca, indicando tendência de redução do ID à medida que aumenta o tamanho das equipes de pesquisa. Esse resultado está alinhado com o trabalho de Wu, Wang e Evans (2019) que sugerem que equipes menores tendem a produzir trabalhos potencialmente disruptivos, enquanto equipes maiores frequentemente se concentram em pesquisas de caráter mais consolidativo. Entretanto, essa correlação ser fraca, sugere que o tamanho da equipe exerce influência limitada sobre a natureza disruptiva da produção científica na área de prótese dentária.

A análise da distribuição geográfica da produção científica revelou concentração de publicações em países tradicionalmente líderes em pesquisa, como Estados Unidos, Inglaterra e Alemanha. No entanto, nesses países, os valores médios de ID foram apenas moderados, enquanto nações com menor volume absoluto de produção, como Índia, Brasil, Japão e Coreia do Sul, apresentaram médias mais elevadas. Esse achado indica que a maior produtividade científica de um país não implica, necessariamente, maior potencial disruptivo de sua produção. Tal interpretação está em consonância com Leibel e Bornmann (2024), que destacam que a disrupção científica não deve ser compreendida como reflexo direto da quantidade de publicações, mas como uma característica estrutural das redes de citação, relacionada à maneira como novos estudos deslocam ou redefinem a centralidade dos trabalhos anteriores.

No que se refere aos periódicos com maior índice de disrupção médio, observou-se que apenas um deles pertence diretamente ao campo da odontologia, enquanto os demais estão associados a áreas biomédicas ou de saúde pública. Esse achado pode refletir a natureza interdisciplinar de parte das pesquisas relacionadas à prótese dentária, que frequentemente incorporam conhecimentos provenientes de áreas como engenharia biomédica, ciência dos materiais, reabilitação e tecnologias digitais. O que é demonstrado em estudos como o de Hsu e Chang (2025) e Hu (2025) indicam que o desenvolvimento de tecnologias como CAD/CAM, manufatura aditiva e inteligência artificial tem ampliado a interface entre a prótese dentária e outros campos científicos.

No que se refere ao modelo de acesso dos periódicos mais disruptivos, observou-se distribuição equilibrada entre periódicos de acesso aberto e híbrido, sem predominância de um único modelo. Esse resultado indica que contribuições associadas a maiores valores de índice de disrupção circulam em diferentes sistemas de publicação científica, sem associação aparente com o modelo de acesso. Embora o acesso aberto seja frequentemente relacionado ao aumento da visibilidade e do número de citações, conforme apontado por Belter (2015), esse efeito diz respeito ao volume de citações e não à forma como elas se organizam nas redes de citação.

A análise das agências de fomento evidenciou predominância de instituições vinculadas ao *National Institutes of Health* (NIH), particularmente do *National Institute of Dental and Craniofacial Research* (NIDCR), sendo a *Fundação de Amparo à Pesquisa do Estado de São Paulo* (FAPESP) a única agência não vinculada aos Estados Unidos entre as mais frequentes. Embora agências americanas tenham financiado grande volume de estudos na área, os valores médios de ID associados a esses financiamentos foram relativamente moderados. Esse resultado pode refletir a tendência das políticas de financiamento científico, que frequentemente priorizam pesquisas com maior previsibilidade de resultados ou continuidade de linhas de investigação previamente estabelecidas, contribuindo para o desenvolvimento cumulativo do conhecimento (Bloom *et al.*, 2020).

Por fim, a análise dos descritores relacionados à estratégia de busca evidenciou variações nos valores médios do Índice de Disrupção entre os termos avaliados, destacando-se *removable prosthodontics* como aquele associado ao maior valor médio de ID (+0,500), ainda que com baixa frequência de ocorrência na

amostra. Esse resultado sugere que a disrupção, enquanto métrica baseada na estrutura das redes de citação, não se distribui necessariamente de forma proporcional à prevalência temática, podendo também emergir em subáreas menos representadas. A organização do conhecimento em subcampos distintos influencia os padrões de citação e a dinâmica das redes científicas, resultando em comportamentos heterogêneos entre diferentes áreas de pesquisa (Cheng *et al.*, 2024; Reia *et al.*, 2025).

Apesar das contribuições do presente estudo, algumas limitações devem ser consideradas. A análise foi restrita a uma única base de dados, o que pode ter influenciado a representatividade da produção científica avaliada. Além disso, a dependência do ID em relação à estrutura das redes de citação torna seus resultados sensíveis à janela temporal adotada e à disponibilidade de metadados completos. Como perspectivas futuras, destaca-se a necessidade de integrar o índice de disrupção a outras abordagens bibliométricas e qualitativas, permitindo uma análise mais abrangente da inovação científica. Além disso, a aplicação do índice em níveis mais específicos, como subáreas da prótese dentária ou tipos de tecnologia, pode aprofundar a compreensão sobre os padrões de inovação no campo.

7 CONCLUSÃO

A partir deste estudo, é possível concluir:

- O Índice de Disrupção evidenciou que a literatura em Prótese Dentária apresenta predominância de publicações com caráter disruptivo, especialmente a partir dos anos 2000;
- A evolução temporal do DI demonstrou estabilidade dos valores médios ao longo do período analisado;
- A produção científica concentrou-se em países tradicionalmente produtivos, como Estados Unidos e países europeus, enquanto nações como Brasil, Índia e Japão apresentaram maior potencial disruptivo relativo;
- Observou-se ausência de correlação significativa entre o DI e o número de citações, indicando que impacto científico não está diretamente associado à disrupção;
- Houve correlação fraca e negativa entre o número de referências e o DI, sugerindo tendência de trabalhos mais fundamentados atuarem de forma consolidativa;
- Artigos com menor número de autores apresentaram, em média, maiores valores de DI, indicando maior potencial disruptivo em equipes menores;
- Os artigos e periódicos com maior disrupção não se restringem a um único modelo de acesso, estando distribuídos entre acesso aberto e híbrido;
- O financiamento científico concentrou-se em agências norte-americanas, com valores de DI moderados, sem associação clara entre financiamento e maior disrupção;
- A análise dos descritores permitiu identificar termos com elevados valores de DI, demonstrando que a disrupção está associada a tópicos específicos dentro da literatura, e não necessariamente aos mais frequentes.

REFERÊNCIAS

- AGARWAL, A. et al. Bibliometria: rastreando o impacto da pesquisa selecionando as métricas apropriadas. **Jornal Asiático de Andrologia**, v. 18, n. 2, p. 296, 2015. DOI: 10.4103/1008-682x.171582.
- AL-RAFEE, M. Epidemiology of edentulism: literature review. **The Saudi Dental Journal**, v. 32, n. 3, p. 107-114, 2020.
- ALRYALAT, S. A.; MOMANI, W. Comparing bibliometric analysis using PubMed, Scopus, and Web of Science databases. **Journal of Informetrics**, v. 13, n. 3, p. 1168-1177, 2019.
- Apolinário, B.A. et al., “SciDisruptor-PubMed”, Mendeley Data, V1, doi: 10.17632/hyg4vfgjtn.1 (2025).
- BELTER, C. Bibliometric indicators: opportunities and limits. **The Editorial Office**, v. 4, n. 2, p. 1-8, 2015.
- BLOOM, N. et al. Are ideas getting harder to find? **American Economic Review**, v. 110, n. 4, p. 1104-1144, 2020.
- CHENG, C. et al. Scientific knowledge graphs: a survey of construction techniques and applications. **Information Processing & Management**, v. 61, n. 4, p. 102306, 2024.
- DEL RIO SILVA, L. et al. CAD/CAM single prosthesis: a 25-year bibliometric assessment. **The Journal of Prosthetic Dentistry**, v. 125, n. 3, p. 303-311, 2025.
- EMAMI, E. et al. The impact of edentulism on oral and general health. *International Journal of Dentistry*, v. 2013, p. 498305, 2013.
- ESPINOZA-CARHUANCHO, R. et al. Exploration and bibliometric mapping of scientific production on quality of life in patients with dental prostheses. **Journal of Clinical and Diagnostic Research**, v. 17, n. 11, p. 1-7, 2023.
- FUNK, R. J.; OWEN-SMITH, J. A dynamic network measure of technological change. **Management Science**, v. 63, p. 791-817, 2016.
- GANTI, S.; YU, P.; JOSE, D. Bibliometric analysis methods for the medical literature: history, methodologies, and tools. **Academic Medicine & Surgery**, v. 3, n. 1, p. e23, 2025.
- GARCIA, B. R. L. et al. Análise bibliométrica dos sistemas cyber-físicos na área da saúde (2016-2020): um olhar na base scopus (elsevier). **Cuadernos de Educación y Desarrollo**, v. 15, n. 8, p. 7696-7716, 2023.
- HSU, M.; CHANG, Y. The literature on digital technologies for removable dental prostheses: a two-decade bibliometric analysis. **Journal of Dental Sciences**, v. 20, n. 1, p. 170-177, 2025.

- HU, Y. Knowledge mapping of artificial intelligence in prosthodontics during 1995-2024: a bibliometric analysis. **International Journal of Prosthodontics**, v. 38, n. 1, p. 21-31, 2025.
- JIANG, P.; LIU, Z. A construction and empirical research of the journal disruption index based on open citation data. **Scientometrics**, v. 128, n. 7, p. 3935-3958, 2023.
- JONES, B. F. The burden of knowledge and the 'death of the renaissance man': Is innovation getting harder? **The Review of Economic Studies**, v. 76, n. 1, p. 283-317, 2009.
- KHAN, S. et al. Bibliometric analysis of the top-cited articles on prosthetic complications in implant dentistry. **The Journal of Prosthetic Dentistry**, v. 132, n. 1, p. 17-25, 2023.
- KOSHEVOY, A.; MITON, H.; MORIN, O. Zipf's Law of Abbreviation holds for individual characters across a broad range of writing systems. **Cognition**, v. 238, p. 105527, 2023.
- LEIBEL, C.; BORNMANN, L. What do we know about the disruption index in scientometrics? An overview of the literature. **Scientometrics**, v. 129, n. 3, p. 601-639, 2024.
- LEYDESDORFF, L.; TEKLES, A.; BORNMANN, L. A proposal to revise the disruption index. **Professional Information**, v. 30, p. e300121, 2021.
- MORAES, L. L.; KAFURE, I. Bibliometria e ciência de dados um exemplo de busca e análise de dados da Web of Science (WoS). **RDBCI**, v. 18, p. e020016, 2020.
- NATH, P. et al. A bibliometric analysis of the journal Community Dentistry and Oral Epidemiology using Scopus data. **Community Dentistry and Oral Epidemiology**, v. 52, n. 1, p. 84-92, 2024.
- PARK, M.; LEAHEY, E.; FUNK, R. J. Papers and patents are becoming less disruptive over time. **Nature**, v. 613, n. 7942, p. 138-144, 2023.
- PETERSEN, A. M.; ARROYAVE, F.; PAMMOLLI, F. The disruption index is biased by citation inflation. **Quantitative Science Studies**, v. 5, n. 4, p. 936-948, 2024.
- PODER, E. What is wrong with the current evaluative bibliometrics? **Frontiers in Research Metrics and Analytics**, v. 6, art. 824518, 2022.
- RASHID, A.; DUDLEY, J. Shade-matching in dentistry: a bibliometric analysis. **Journal of Esthetic and Restorative Dentistry**, v. 36, n. 4, p. 505-516, 2024.
- REIA, S. et al. Science of science: a complex network perspective. **Frontiers in Research Metrics and Analytics**, v. 4, p. 148-169, 2025.
- TECHAPIROONTONG, T. et al. The benefits of dental prostheses on masticatory function and oral health-related quality of life: a cross-sectional study. **BMC Oral Health**, v. 22, n. 1, p. 1-12, 2022.
- VENKATESAN, K. et al. Evidence-based prosthodontics—simplifying the search and selection of scientific evidence. **Journal of Indian Prosthodontic Society**, v. 22, n. 1, p. 5-10, 2022.

WATTED, N. et al. Combined orthodontic and prosthetic therapy; special considerations. **Journal of Dental and Medical Sciences**, v. 15, n. 12, p. 62-73, 2016.

WU, L.; WANG, D.; EVANS, J. A. Large teams develop and small teams disrupt science and technology. **Nature**, v. 566, p. 378-382, 2019.

GLOSSÁRIO

Algoritmo: Conjunto de instruções lógicas e sequenciais utilizadas para resolver um problema ou executar uma tarefa computacional, como o processamento de dados bibliométricos.

API (Application Programming Interface): Interface que permite a comunicação entre diferentes sistemas computacionais, possibilitando a extração automatizada de dados de bases como o PubMed.

Arquivo CSV (Comma-Separated Values): Formato de arquivo de texto utilizado para armazenar dados estruturados em forma de linhas e colunas, amplamente empregado para análise de dados.

Arquivo XML (eXtensible Markup Language): Linguagem de marcação utilizada para armazenar e transportar dados estruturados, permitindo a organização de informações complexas em bases científicas.

Co-citação: Técnica bibliométrica que analisa a frequência com que dois documentos são citados conjuntamente por outros trabalhos, indicando relação temática entre eles.

Código (Script): Conjunto de comandos escritos em linguagem de programação para automatizar tarefas específicas.

E-utilities (PubMed): Conjunto de ferramentas disponibilizadas pela *National Center for Biotechnology Information* que permite o acesso programático à base de dados PubMed por meio de requisições via API, possibilitando a busca, recuperação e extração automatizada de informações bibliográficas.

iCite: O iCite é uma ferramenta online mantida pelo National Institutes of Health (NIH) que disponibiliza indicadores de desempenho científico para publicações indexadas no PubMed desde 1980.

Metadados: Informações estruturadas que descrevem os dados principais de uma publicação científica, como título, autores, ano, resumo e referências.

Mesh terms: Descritores padronizados desenvolvidos pela National Library of Medicine (NLM) e utilizados no PubMed para classificar artigos científicos por assunto. Permitem buscas bibliográficas mais precisas e abrangentes, agrupando diferentes termos relacionados a um mesmo conceito.

Python: Linguagem de programação de alto nível amplamente utilizada para análise de dados, automação e desenvolvimento de algoritmos científicos.

Repositório SciDisruptor-PubMed: Conjunto de códigos e ferramentas computacionais desenvolvido para extração, organização e análise de dados bibliométricos provenientes da base PubMed, permitindo a construção de redes de citações e o cálculo do Índice de Disrupção em publicações científicas.

Visual Studio Code: Editor de código-fonte utilizado para desenvolvimento e execução de códigos em diversas linguagens de programação.

Anexo A

```

import os
import requests
import pandas as pd
import time
from tqdm import tqdm
# Pasta de saída
output_folder = r"C:\Users\apoli\OneDrive\Desktop\Larissa\16-12"
os.makedirs(output_folder, exist_ok=True)
# Chave da API PubMed
API_KEY = "32b027833855dd01d0bb2d7891c832d97308"
# Estratégia de busca
BASE_QUERY = (
    '(
    "Prosthodontics"[MeSH] OR '
    "Dental Prosthesis"[MeSH]'
    )'
    'OR '
    '(
    'prosthodontics[Title/Abstract] OR '
    "fixed prosthodontics"[Title/Abstract] OR '
    "removable prosthodontics"[Title/Abstract] OR '
    "implant prosthodontics"[Title/Abstract] OR '
    "complete denture*" [Title/Abstract] OR '
    "partial denture*" [Title/Abstract] OR '
    "overdenture*" [Title/Abstract] OR '
    "implant-supported prosthesis"[Title/Abstract] OR '
    "fixed dental prosthesis"[Title/Abstract] OR '
    "removable dental prosthesis"[Title/Abstract] OR '
    "prosthetic rehabilitation"[Title/Abstract]'
    )'
)
def fetch_pmids_by_year(start_year: int, end_year: int):
    base_url = "https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esearch.fcgi"
    retmax = 10000
    session = requests.Session()
    # LISTA GLOBAL (todos os anos)
    all_pmids_global = []
    for year in range(start_year, end_year + 1):
        print(f"\n===== Buscando PMIDs para o ano: {year} =====")
        start = 0
        all_pmids_year = []
        year_folder = os.path.join(output_folder, str(year))
        os.makedirs(year_folder, exist_ok=True)

```

```

search_query = f"({BASE_QUERY}) AND ({year}[PDAT])"
# Buscar contagem
params_count = {
    "db": "pubmed",
    "term": search_query,
    "retstart": 0,
    "retmax": 0,
    "retmode": "json",
    "api_key": API_KEY
}
r0 = session.get(base_url, params=params_count, timeout=60)
total_records = int(r0.json()["esearchresult"]["count"])
print(f"Total de registros para {year}: {total_records}")
if total_records == 0:
    pd.DataFrame([], columns=["PMID"]).to_csv(
        os.path.join(year_folder, f"pmids_{year}.csv"),
        index=False
    )
    continue
total_batches = (total_records + retmax - 1) // retmax
with tqdm(total=total_batches, desc=f"{year}", unit="lote") as pbar:
    while start < total_records:
        params = {
            "db": "pubmed",
            "term": search_query,
            "retstart": start,
            "retmax": retmax,
            "retmode": "json",
            "api_key": API_KEY
        }
        r = session.get(base_url, params=params, timeout=60)
        data = r.json()
        pmids = data["esearchresult"]["idlist"]
        all_pmids_year.extend(pmids)
        all_pmids_global.extend(pmids)
        start += retmax
        pbar.update(1)
        time.sleep(0.34)
# Remover duplicatas do ano
pmids_year_unique = list(dict.fromkeys(all_pmids_year))
pd.DataFrame(pmids_year_unique, columns=["PMID"]).to_csv(
    os.path.join(year_folder, f"pmids_{year}.csv"),
    index=False
)

```

```
    print(f"PMIDs salvos para {year}: {len(pmids_year_unique)}")
# ===== CSV GLOBAL =====
pmids_global_unique = list(dict.fromkeys(all_pmids_global))
global_file = os.path.join(output_folder, "PMIDs_TODOS_OS_ANOS.csv")
pd.DataFrame(pmids_global_unique, columns=["PMID"]).to_csv(global_file,
index=False)
print(f"\n✅ CSV GLOBAL salvo em: {global_file}")
print(f"Total de PMIDs únicos (1900–2025): {len(pmids_global_unique)}")
if __name__ == "__main__":
    fetch_pmids_by_year(1900, 2025)
    print("\nBusca finalizada!")
```

ANEXO B

```

import os
import time
import requests
import pandas as pd
from tqdm import tqdm
# ===== 1) Configuração de pastas
=====
# Pasta RAIZ onde estão as subpastas com os arquivos de PMIDs (CSVs)
input_folder = r"C:\Users\apoli\OneDrive\Desktop\Larissa\16-12\PMIDS"
# Pasta onde os XML serão salvos (espelhando a estrutura das subpastas)
output_folder = r"C:\Users\apoli\OneDrive\Desktop\Larissa\16-12\XML"
os.makedirs(output_folder, exist_ok=True)
# Chave de API da PubMed
API_KEY = "32b027833855dd01d0bb2d7891c832d97308"
def request_with_retries(url, params, max_retries=3, timeout=60):
    """Requisição com tentativas (retries) e backoff simples."""
    for attempt in range(1, max_retries + 1):
        try:
            r = requests.get(url, params=params, timeout=timeout)
            return r
        except Exception:
            if attempt == max_retries:
                return None
            time.sleep(1.5 * attempt) # backoff
def download_xml_from_pmids():
    """
    Faz o download dos dados XML da PubMed para TODOS os arquivos CSV
    encontrados em input_folder e em TODAS as suas subpastas.
    Mantém a estrutura de subpastas na saída.
    """
    base_url = "https://eutils.ncbi.nlm.nih.gov/entrez/eutils/efetch.fcgi"
    batch_size = 200 # limite clássico do efetch com id list
    print("===== INÍCIO DO DOWNLOAD =====")
    print(f"Pasta de entrada: {input_folder}")
    print(f"Pasta de saída: {output_folder}\n")
    # 1) Listar todos os CSVs recursivamente (para ter progresso geral)
    all_csv_paths = []
    for root_dir, _, files in os.walk(input_folder):
        for f in files:
            if f.lower().endswith(".csv"):
                all_csv_paths.append(os.path.join(root_dir, f))

```

```

if not all_csv_paths:
    print(" ⚠ Nenhum arquivo .csv encontrado na pasta (ou subpastas).")
    return
# 2) Progresso geral por arquivo CSV
for pmid_file in tqdm(all_csv_paths, desc="Arquivos CSV", unit="arquivo"):
    root_dir = os.path.dirname(pmid_file)
    csv_file = os.path.basename(pmid_file)
    # Caminho relativo dessa pasta em relação à pasta raiz
    rel_path = os.path.relpath(root_dir, input_folder)
    # Pasta de saída espelhando a estrutura da entrada
    dest_folder = os.path.join(output_folder, rel_path)
    os.makedirs(dest_folder, exist_ok=True)
    # Ler os PMIDs do arquivo
    try:
        df_pmids = pd.read_csv(pmid_file)
    except Exception as e:
        print(f"\n ⚠ Erro ao ler {pmid_file}: {e}")
        continue
    # Encontrar coluna PMID
    pmid_col = None
    for col in df_pmids.columns:
        if col.strip().lower() == "pmid":
            pmid_col = col
            break
    if pmid_col is None:
        print(f"\n ⚠ Nenhuma coluna 'PMID' encontrada em {pmid_file}.
Pulando.")
        continue
    pmids = df_pmids[pmid_col].dropna().astype(str).str.strip().tolist()
    pmids = [p for p in pmids if p.isdigit()] # higiene básica
    if not pmids:
        print(f"\n ⚠ Nenhum PMID válido encontrado em {pmid_file}. Pulando.")
        continue
    total_batches = (len(pmids) + batch_size - 1) // batch_size
    # 3) Barra de progresso por lote (dentro de cada CSV)
    lotes = range(0, len(pmids), batch_size)
    for idx in tqdm(lotes, desc=f"Lotes ({csv_file})", total=total_batches,
unit="lote", leave=False):
        pmid_batch = pmids[idx:idx + batch_size]
        batch_number = idx // batch_size + 1
        base_name = os.path.splitext(csv_file)[0]

```

```

        xml_file = os.path.join(dest_folder, f"{base_name}
_batch_{batch_number}.xml")
        # Se já existe, pula (evita retrabalho se cair no meio)
        if os.path.exists(xml_file) and os.path.getsize(xml_file) > 0:
            continue
        params = {
            "db": "pubmed",
            "id": ",".join(pmid_batch),
            "retmode": "xml",
            "api_key": API_KEY
        }
        response = request_with_retries(base_url, params=params,
max_retries=3, timeout=60)
        if response is None:
            print(f"\n ⚠ Falha de conexão após retries no lote {batch_number}
({csv_file}).")
            time.sleep(1)
            continue
        if response.status_code == 200:
            with open(xml_file, "wb") as f:
                f.write(response.content)
        else:
            # Log curto para não poluir a tela (tqdm + prints)
            print(f"\n ⚠ Erro HTTP {response.status_code} no lote
{batch_number} ({csv_file}).")
            print(" Resposta (início):", response.text[:250].replace("\n", " "), "...")
            time.sleep(1)
            # Pausa para respeitar rate limit
            time.sleep(0.34)
        print("\n===== DOWNLOAD DOS XMLs FINALIZADO!
=====")
        print(f"📁 XMLs salvos em: {output_folder}")
if __name__ == "__main__":
    download_xml_from_pmids()

```

ANEXO C

```

import os

import pandas as pd
from xml.etree import ElementTree as ET
from tqdm import tqdm
# ===== CONFIGURAÇÃO DE PASTAS
=====
input_folder = r"C:\Users\apoli\OneDrive\Desktop\Larissa\16-12\XML"
output_folder = r"C:\Users\apoli\OneDrive\Desktop\Larissa\16-12"
os.makedirs(output_folder, exist_ok=True)
# ===== FUNÇÕES AUXILIARES =====
def safe_text(elem):
    return elem.text.strip() if (elem is not None and elem.text) else None
def get_first_text(element, paths):
    """Tenta múltiplos XPath e retorna o primeiro texto encontrado."""
    for path in paths:
        found = element.find(path)
        txt = safe_text(found)
        if txt:
            return txt
    return None
def extract_total_citations(article):
    """
    Tenta achar 'citações totais' caso esse XML tenha esse tipo de dado.
    OBS: Em XMLs do efetch padrão, normalmente isso NÃO vem; mas se vier,
    costuma aparecer em alguns nós/atributos específicos dependendo da fonte/
pipe.
    Aqui tentamos algumas possibilidades comuns.
    """
    # 1) Alguns pipelines incluem contagens como tags no ArticleIdList ou em
    extensões
    candidates = [
        ".//CitedByCount",
        ".//CitedBy",
        ".//Citations",
        ".//CitationCount",

```

```

    "../PmcRefCount",
    "../ReferenceCount",
    "../RefCount",
]
for path in candidates:
    val = get_first_text(article, [path])
    if val and val.isdigit():
        return int(val)
# 2) Tenta atributos do tipo Count="123"
for node_name in ["CitedByCount", "Citations", "CitationCount", "CitedBy"]:
    for el in article.findall(f"://{node_name}"):
        count_attr = el.attrib.get("Count") or el.attrib.get("count")
        if count_attr and str(count_attr).isdigit():
            return int(count_attr)
# 3) Tenta achar algo do tipo <ArticleId IdType="citedby">123</ArticleId>
for aid in article.findall("../ArticleId"):
    idtype = (aid.attrib.get("IdType") or "").lower()
    if idtype in {"citedby", "cited_by", "citations", "citationcount"}:
        txt = safe_text(aid)
        if txt and txt.isdigit():
            return int(txt)
return None
# ===== FUNÇÃO PARA EXTRAIR DADOS DE UM ARTIGO
=====
def extract_article_data(article):
    # PMID
    pmid = get_first_text(article, ["../PMID"])
    # Ano
    year = get_first_text(article, ["../PubDate/Year"])
    if year is None:
        medline_date = get_first_text(article, ["../PubDate/MedlineDate"])
        year = medline_date[:4] if medline_date else None
    # Título (pode ter tags internas; fazemos join do itertext)
    title_el = article.find("../ArticleTitle")
    title = " ".join(title_el.itertext()).strip() if title_el is not None else None
    # ----- INFORMAÇÕES DO PERIÓDICO -----
    journal = get_first_text(article, ["../Journal/Title"])
    iso_abbrev = get_first_text(article, ["../Journal/ISOAbbreviation"])
    # ISSN (principal / print / electronic)
    issn_principal = None
    issn_electronic = None
    issn_print = None
    for jid in article.findall("../Journal/ISSN"):
        txt = safe_text(jid)

```

```

if not txt:
    continue
issn_type = (jid.attrib.get("IssnType") or "").lower()
if "electronic" in issn_type:
    issn_electronic = txt
elif "print" in issn_type:
    issn_print = txt
if issn_principal is None:
    issn_principal = txt
volume = get_first_text(article, [".//JournalIssue/Volume"])
issue = get_first_text(article, [".//JournalIssue/Issue"])
pages = get_first_text(article, [".//Page/Text/Text"])
# ----- AUTORES E ORCID -----
authors = []
for author in article.findall(".//Author"):
    last_name = get_first_text(author, ["LastName"])
    fore_name = get_first_text(author, ["ForeName"])
    orcid = get_first_text(author, ["Identifier[@Source='ORCID']"])
    if last_name and fore_name:
        full_name = f"{fore_name} {last_name}"
        if orcid:
            full_name += f" (ORCID: {orcid})"
        authors.append(full_name)
authors_text = "; ".join(authors) if authors else None
# ----- MeSH -----
mesh_terms = []
for mesh in article.findall(".//MeshHeading/DescriptorName"):
    txt = safe_text(mesh)
    if txt:
        mesh_terms.append(txt)
mesh_text = "; ".join(mesh_terms) if mesh_terms else None
# ----- Tipos de publicação -----
pub_types = []
for pub_type in article.findall(".//PublicationType"):
    txt = safe_text(pub_type)
    if txt:
        pub_types.append(txt)
pub_type_text = "; ".join(pub_types) if pub_types else None
# ----- Agências financiadoras -----
grants = []
for grant in article.findall(".//GrantList/Grant"):
    agency = get_first_text(grant, ["Agency"])
    if agency:
        grants.append(agency)

```

```

grants_text = "; ".join(grants) if grants else None
# Idioma
language = get_first_text(article, ["//Language"])
# País
country = get_first_text(article, ["//MedlineJournalInfo/Country"])
# ----- Referências (PMIDs citados) -----
references = []
for ref in article.findall("//Reference/ArticleIdList/
ArticleId[@IdType='pubmed']"):
    txt = safe_text(ref)
    if txt:
        references.append(txt)
references_text = "; ".join(references) if references else None
# ----- Citações totais (se existirem nesse XML) -----
total_citations = extract_total_citations(article)
return {
    "PMID": pmid,
    "Ano": year,
    "Titulo": title,
    # Dados do periódico
    "Journal": journal,
    "ISOAbbreviation": iso_abbrev,
    "ISSN Principal": issn_principal,
    "ISSN Eletrônico": issn_electronic,
    "ISSN Impresso": issn_print,
    "Volume": volume,
    "Issue": issue,
    "Paginas": pages,
    # Outros metadados
    "Autores (ORCID)": authors_text,
    "MeSH Terms": mesh_text,
    "Tipo de Publicacao": pub_type_text,
    "Agencia Financiadora": grants_text,
    "Idioma": language,
    "Pais": country,
    "Referencias (PMIDs citados)": references_text,
    # Citações (se disponível)
    "Citacoes Totais (se disponível no XML)": total_citations,
}
# ===== FUNÇÃO PRINCIPAL =====
def process_all_xmls():
    # 1) Lista todos os XMLs primeiro (pra barra de progresso ficar certa)
    xml_files = []
    for root_dir, _, files in os.walk(input_folder):

```

```

for f in files:
    if f.lower().endswith(".xml"):
        xml_files.append(os.path.join(root_dir, f))
if not xml_files:
    print(f"Nenhum XML encontrado em: {input_folder}")
    return
articles_data = []
erros = 0
# 2) Processa com barra de progresso
for file_path in tqdm(xml_files, desc="Lendo XMLs", unit="arquivo"):
    try:
        tree = ET.parse(file_path)
        root = tree.getroot()
        # vários PubmedArticle por arquivo
        for article in root.findall("./PubmedArticle"):
            data = extract_article_data(article)
            # adiciona info de origem (ajuda debug)
            data["Arquivo_Origem"] = os.path.basename(file_path)
            data["Pasta_Origem"] = os.path.relpath(os.path.dirname(file_path),
input_folder)
            articles_data.append(data)
    except Exception as e:
        erros += 1
        # Não interrompe o processo
        # print(f"Erro ao processar {file_path}: {e}")
df = pd.DataFrame(articles_data)
out_file = os.path.join(output_folder, "dados_extraidos_Larissa_16-12.xlsx")
df.to_excel(out_file, index=False)
print("\n===== RESUMO =====")
print(f"XMLs encontrados: {len(xml_files)}")
print(f"Arquivos com erro: {erros}")
print(f"Artigos extraídos (linhas): {len(df)}")
print(f"PMIDs únicos: {df['PMID'].nunique(dropna=True) if 'PMID' in
df.columns else 'N/A'}")
print(f"Planilha salva em: {out_file}")
print("=====\n")
# ===== EXECUTAR =====
if __name__ == "__main__":
    process_all_xmls()

```

ANEXO D

```

# =====
# ANEXO D (ROBUSTO)
# - Cria rede de citações a partir de Dados_oficiais_sem_duplicados.xlsx
# - Inclui anos do citante e do citado
# - Barra de progresso no Jupyter
# - Salva: rede_citacoes_com_anos.csv
# =====
import pandas as pd
import os
import re
from tqdm.auto import tqdm
# ===== CAMINHOS =====
input_file = r"C:
\Users\apoli\OneDrive\Desktop\Larissa\16-12\Dados_oficiais_sem_duplicados.xlsx"
output_folder = r"C:\Users\apoli\OneDrive\Desktop\Larissa\16-12"
output_file = os.path.join(output_folder, "rede_citacoes_com_anos.csv")
os.makedirs(output_folder, exist_ok=True)
# ===== LER DADOS =====
df = pd.read_excel(input_file)
# ===== NORMALIZAR NOMES DAS COLUNAS (remove
espaços invisíveis) =====
df.columns = [str(c).strip() for c in df.columns]
# ===== DETECTAR COLUNA DE REFERÊNCIAS
=====
pmid_col = "PMID"
year_col = "Ano"
possible_ref_cols = ["Referencias", "Referencias (PMIDs citados)"]
ref_col = next((c for c in possible_ref_cols if c in df.columns), None)
missing = [c for c in [pmid_col, year_col] if c not in df.columns]
if missing:

```

```

    raise ValueError(f"Faltam colunas obrigatórias: {missing}\nColunas
encontradas: {list(df.columns)}")
    if ref_col is None:
        raise ValueError(
            "Não encontrei a coluna de referências. Esperado um destes nomes:\n"
            f"{possible_ref_cols}\nColunas encontradas: {list(df.columns)}"
        )
    print(f"✅ Coluna de referências detectada: {ref_col}")
    # ===== RENOMEAR PARA NOMES 'SEGUROS' (sem
parênteses/acentos) =====
    df = df.rename(columns={pmid_col: "PMID", year_col: "Ano", ref_col:
"Referencias"})
    # ===== NORMALIZAR TIPOS =====
    # PMID às vezes vem como 18074164.0 -> vamos extrair apenas os dígitos
df["PMID"] = df["PMID"].astype(str).str.extract(r"(\d+)")[0]
df["PMID"] = df["PMID"].astype(str).str.strip()
df["Ano"] = pd.to_numeric(df["Ano"], errors="coerce")
pmid_to_year = dict(zip(df["PMID"], df["Ano"]))
    # ===== FUNÇÃO PARA EXTRAIR PMIDs DA CÉLULA
=====
    def extract_pmids(cell):
        if pd.isna(cell):
            return []
        # pega sequências numéricas (funciona para "18074164; 14531614; ...")
        return re.findall(r"\b\d+\b", str(cell))
    # ===== CRIAR REDE (COM PROGRESSO)
=====
    citations = []
    n_with_refs = 0
    total_refs_found = 0
    for row in tqdm(df.itertuples(index=False), total=len(df), desc="🔗 Construindo
rede de citações", unit="artigo", dynamic_ncols=True):
        citing_pmid = str(row.PMID).strip()
        citing_year = pmid_to_year.get(citing_pmid)
        ref_list = extract_pmids(row.Referencias)
        if ref_list:
            n_with_refs += 1
            total_refs_found += len(ref_list)
        for cited_pmid in ref_list:
            cited_pmid = str(cited_pmid).strip()
            cited_year = pmid_to_year.get(cited_pmid) # pode ser NaN se o citado
não estiver na planilha
            citations.append({

```

```

        "CitingPublicationId": citing_pmid,
        "CitedPublicationId": cited_pmid,
        "CitingYear": citing_year,
        "CitedYear": cited_year
    })
    citations_df = pd.DataFrame(citations)
    # ===== REMOVER DUPLICATAS DE ARESTAS
    =====
    if not citations_df.empty:
        before = len(citations_df)
        citations_df = citations_df.drop_duplicates(subset=["CitingPublicationId",
"CitedPublicationId"])
        after = len(citations_df)
    else:
        before = after = 0
    # ===== SALVAR =====
    citations_df.to_csv(output_file, index=False)
    # ===== RESUMO =====
    missing_cited_year = citations_df["CitedYear"].isna().sum() if after > 0 else 0
    print("\n✅ Rede de citações criada com sucesso!")
    print(f"Artigos analisados: {len(df):,}".replace(",","."))
    print(f"Artigos com referências: {n_with_refs:,}".replace(",","."))
    print(f"Total de referências extraídas (antes de deduplicar):
{total_refs_found:,}".replace(",","."))
    print(f"Arestas antes de deduplicar: {before:,}".replace(",","."))
    print(f"Arestas após deduplicar: {after:,}".replace(",","."))
    print(f"Arestas sem ano do citado (citado fora da planilha):
{missing_cited_year:,}".replace(",","."))
    print(f"📁 Arquivo salvo em: {output_file}")

```

ANEXO E

```

import pandas as pd
import numpy as np
import os
from tqdm import tqdm
tqdm.pandas() # habilita barras em operações pandas, se precisar
# ===== CAMINHOS =====
dados_path = r"C:\Users\apoli\OneDrive\Desktop\Larissa\16-12\Dados_oficiais_sem_duplicados.xlsx"
citations_path = r"C:\Users\apoli\OneDrive\Desktop\Larissa\16-12\rede_citacoes_com_anos.csv"
# ===== PARÂMETROS =====
JANELA_CITACAO = 5
ANO_LIMITE_FOCO = 2020
# ===== LER DADOS =====
print("📂 Lendo arquivos...")
dados = pd.read_excel(dados_path)
citations = pd.read_csv(citations_path)
# ===== NORMALIZAÇÃO =====
dados["PMID"] = dados["PMID"].astype(str).str.strip()
dados["Ano"] = pd.to_numeric(dados["Ano"], errors="coerce")
citations["CitingPublicationId"] =
citations["CitingPublicationId"].astype(str).str.strip()
citations["CitedPublicationId"] =
citations["CitedPublicationId"].astype(str).str.strip()
# ===== GARANTIR ANOS NA REDE =====
has_years = ("CitingYear" in citations.columns) and ("CitedYear" in
citations.columns)
pmid_to_year = dict(zip(dados["PMID"], dados["Ano"]))

```

```

if not has_years:
    print("🧩 Mapeando anos para a rede de citações...")
    citations["CitingYear"] = citations["CitingPublicationId"].map(pmid_to_year)
    citations["CitedYear"] = citations["CitedPublicationId"].map(pmid_to_year)
    citations["CitingYear"] = pd.to_numeric(citations["CitingYear"], errors="coerce")
    citations["CitedYear"] = pd.to_numeric(citations["CitedYear"], errors="coerce")
    citations = citations.dropna(subset=["CitingYear", "CitedYear"]).copy()
    citations["CitingYear"] = citations["CitingYear"].astype(int)
    citations["CitedYear"] = citations["CitedYear"].astype(int)
    # ===== APLICAR JANELA DE CITAÇÃO
=====

    print("⌚ Aplicando janela de citação de 5 anos...")
    citations = citations[
        (citations["CitingYear"] - citations["CitedYear"] >= 0) &
        (citations["CitingYear"] - citations["CitedYear"] <= JANELA_CITACAO)
    ].copy()
    print(f"🔗 Citações após filtro de janela: {len(citations):,}".replace(",", "."))
    # ===== AGRUPAMENTOS =====
    reference_groups = citations.groupby("CitingPublicationId", sort=False)
    ["CitedPublicationId"]
    citation_groups = citations.groupby("CitedPublicationId", sort=False)
    ["CitingPublicationId"]
    def get_references(pid):
        try:
            return reference_groups.get_group(pid).values
        except KeyError:
            return np.array([])
    def get_citations(pid):
        try:
            return citation_groups.get_group(pid).values
        except KeyError:
            return np.array([])
    # ===== FILTRAR ARTIGOS FOCALIS =====

    print("📌 Aplicando critérios de elegibilidade dos artigos focais...")
    dados_foco = dados[dados["Ano"].notna() & (dados["Ano"] <=
ANO_LIMITE_FOCO)].copy()
    focus_ids = []
    for pid in tqdm(dados_foco["PMID"].values, desc="Filtrando artigos focais"):
        if len(get_references(pid)) == 0:
            continue
        if len(get_citations(pid)) == 0:
            continue

```

```

    focus_ids.append(pid)
print(f"\n✅ Artigos focais elegíveis: {len(focus_ids):,}".replace(",","."))
# ===== CALCULAR ÍNDICE DE DISRUPÇÃO
=====
results = []
for focusid in tqdm(focus_ids, desc="Calculando Índice de Disrupção"):
    focus_refs = get_references(focusid)
    focus_citings = set(get_citations(focusid))
    cite_to_refs = set()
    for refid in focus_refs:
        cite_to_refs.update(get_citations(refid))
    nj = len(focus_citings.intersection(cite_to_refs))
    ni = len(focus_citings) - nj
    nk = len(cite_to_refs) - nj
    soma = ni + nj + nk
    di = (ni - nj) / soma if soma != 0 else np.nan
    results.append([focusid, ni, nj, nk, soma, di])
# ===== RESULTADO FINAL =====
disruption_results = pd.DataFrame(
    results,
    columns=["PublicationId", "ni", "nj", "nk", "Soma", "DisruptionIndex"]
)
disruption_results = disruption_results.merge(
    dados[["PMID", "Ano"]],
    left_on="PublicationId",
    right_on="PMID",
    how="left"
).drop(columns="PMID")
disruption_results = disruption_results[
    ["PublicationId", "Ano", "ni", "nj", "nk", "Soma", "DisruptionIndex"]
]
# ===== SALVAR =====
output_path = r"C:
\Users\apoli\OneDrive\Desktop\Larissa\16-12\indice_disrupcao_janela_5anos_ate20
20.xlsx"
disruption_results.to_excel(output_path, index=False)
print("\n PROCESSO FINALIZADO COM SUCESSO!")
print(f" Arquivo salvo em:\n{output_path}")

```