

Evidências empíricas para uso do *software* iThenticate por editores de revistas científicas – relato de experiência

Empirical evidence for iThenticate software usage by scientific journal editors – experience report

Evidencia empírica para uso del *software* iThenticate por editores de revistas científicas – informe de experiencia

João de Deus Barreto Segundo^{1,3,a}

jao.barreto@gmail.com | <https://orcid.org/0000-0002-4802-6045>

Ana Paula de Oliveira Villalobos^{1,b}

anap.villalobos@terra.com.br | <https://orcid.org/0000-0003-3151-156X>

Katia Nunes Sá^{2,3,c}

katia.sa@bahiana.edu.br | <https://orcid.org/0000-0002-0255-4379>

¹ Universidade Federal da Bahia, Programa de Pós-graduação em Ciência da Informação. Salvador, BA, Brasil.

² Escola Bahiana de Medicina e Saúde Pública, Programa de Pós-graduação em Medicina e Saúde Humana. Salvador, BA, Brasil.

³ Escola Bahiana de Medicina e Saúde Pública, Núcleo de Comunicação Científica. Salvador, BA, Brasil.

^a Mestrado em Comunicação e Cultura Contemporâneas pela Universidade Federal da Bahia.

^b Doutorado em Educação pela Universidade Federal da Bahia.

^c Doutorado em Medicina e Saúde Humana pela Escola Bahiana de Medicina e Saúde Pública.

RESUMO

Introdução: O plágio é responsável pela retirada de metade dos artigos da literatura. *Softwares* de detecção de similaridade textual como o iThenticate apoiam os editores na prevenção ao plágio através de relatórios automáticos. Contudo, relatos voltados a editores, explanando como interpretar esses relatórios, são escassos. **Objetivo:** Compartilhar a experiência dos editores do portal BAHIANA Journals na interpretação dos relatórios do iThenticate em cinco anos de operação do portal. **Casuística:** São apresentados as incidências de similaridade presentes nos textos analisados pela equipe do portal e os percursos analíticos para compreensão de escores totais e parciais de não originalidade. É discutido também como analisar os escores em função da seção de métodos dos relatos e da demanda contemporânea por mais transparência na comunicação de achados de pesquisas científicas. **Considerações finais:** Os relatórios automáticos do *software* subsidiam indícios de plágio no geral, mas solicitam análises qualitativas que levem em consideração contextos específicos de cada texto.

Palavras-chave: Periódicos eletrônicos; Editoração científica; Indicadores; Autoria; Plágio.

ABSTRACT

Introduction: Plagiarism accounts for half of the articles retracted from the literature. Textual similarity detection software like iThenticate support editors in preventing plagiarism. However, accounts intended for editors, explaining how to interpret their similarity reports, are scarce. **Objective:** To share the experience of the editors of the BAHIANA Journals portal in interpreting iThenticate reports in the past five years. **Casuistic:** The incidence of similarity in the analyzed texts and a rationale to analyze total and partial non-originality scores are presented. How to analyze the reports respecting the methods section of the texts and observing the contemporary demand for more scientific transparency is discussed. **Final remarks:** The software's automatic reports provide evidence of plagiarism in general, but the final verdict requires qualitative analysis that take into consideration the specifics of each text.

Keywords: Electronic journals; Scholarly publishing; Indicators; Authorship; Plagiarism.

RESUMEN

Introducción: El plagio representa la mitad de las retractaciones de artículos científicos. *Softwares* de detección de similitudes de texto como iThenticate ayudan a los editores a prevenir el plagio. Artículos dirigidos a los editores, que expliquen cómo interpretar los informes de similitud del iThenticate, son escasos. **Objetivo:** Compartir la experiencia de los editores del portal BAHIANA Journals en la interpretación de informes iThenticate en cinco años de funcionamiento. **Casuístico:** Se presenta la incidencia de similitud en los textos analizados y un racional para entender sus puntuaciones de no originalidad total y parcial. Se discute cómo analizar las puntuaciones en función de la sección de métodos de reportes y de la demanda contemporánea de mayor transparencia en la comunicación de los resultados de las investigaciones científicas. **Consideraciones finales:** Los informes automáticos del *software* proporcionan evidencia de plagio en general, pero requieren análisis cualitativos que tomen en consideración los contextos específicos de cada texto.

Palabras clave: Revistas electrónicas; Publicación académica; Indicadores; Autores; Plagio.

INFORMAÇÕES DO ARTIGO

Contribuição dos autores:

Concepção e desenho do estudo: Ana Paula de Oliveira Villalobos; João de Deus Barreto Segundo; Katia Nunes Sá.

Aquisição, análise ou interpretação dos dados: João de Deus Barreto Segundo; Katia Nunes Sá.

Redação do manuscrito: Ana Paula de Oliveira Villalobos; João de Deus Barreto Segundo; Katia Nunes Sá.

Revisão crítica do conteúdo intelectual: Ana Paula de Oliveira Villalobos; Katia Nunes Sá.

Declaração de conflito de interesses: João de Deus Barreto Segundo é editor executivo do portal BAHIANA Journals. Katia Nunes Sá é coordenadora do portal BAHIANA Journals e editora-chefe na Revista Pesquisa em Fisioterapia (RPF), vinculada ao portal. Ambos são colaboradores na Escola Bahiana de Medicina e Saúde Pública. A Fundação Bahiana para Desenvolvimento das Ciências, fundação privada sem fins lucrativos que é mantenedora da Escola Bahiana de Medicina e Saúde Pública, não interferiu ou participou da concepção e produção deste relato, que é parte integrante da pesquisa de doutoramento de João de Deus Barreto Segundo, sob orientação de Katia Nunes Sá e Ana Paula de Oliveira Villalobos. Não há nenhum conflito financeiro, legal ou político envolvendo terceiros a ser declarado para nenhum aspecto deste trabalho.

Fontes de financiamento: não houve.

Considerações éticas: todos os dados utilizados foram anonimizados para preservar autores, avaliadores e editores; taxas de submissão, avaliação, aceite e recusa de artigos são dados públicos e ficam disponíveis no portal BAHIANA Journals na respectiva página de cada revista. O uso dos dados relacionados a plágio foi autorizado pela instituição mantenedora do portal BAHIANA Journals, responsável pelo contrato com a Turnitin e com a Associação Brasileira de Editores Científicos, conforme carta de anuência anexada à submissão.

Agradecimentos/Contribuições adicionais: A Izabel Lima, por sugerir a necessidade de um relato de experiência apoiando editores de revistas científicas no uso do *software* antiplágio. A Gustavo Araújo, pela revisão crítica da primeira versão deste texto.

Histórico do artigo: submetido: 13 fev. 2021 | aceite: 10 nov. 2021 | publicado: 25 fev. 2022.

Apresentação anterior: não houve.

Licença CC BY-NC atribuição não comercial. Com essa licença é permitido acessar, baixar (*download*), copiar, imprimir, compartilhar, reutilizar e distribuir os artigos, desde que para uso não comercial e com a citação da fonte, conferindo os devidos créditos de autoria e menção à Reciis. Nesses casos, nenhuma permissão é necessária por parte dos autores ou dos editores.

INTRODUÇÃO

Nos últimos trinta anos, o estudo sobre práticas questionáveis em pesquisa e publicação assumiu uma posição de destaque no debate científico nas ciências biomédicas (OLIVEIRA, 2015; O'BOYLE JR.; BANKS; GONZALEZ-MULÉ, 2017). Essa renovação mostrou-se necessária para melhorar a integridade científica do campo, buscando compreender o fenômeno do desperdício em pesquisa e dos resultados espúrios, que podem induzir a erros em condutas clínicas (IOANNIDIS *et al.*, 2014; MACLEOD *et al.*, 2014).

A metaciência, como disciplina emergente, vem explorando caminhos para incremento da eficiência da pesquisa em saúde (MACLEOD *et al.*, 2014; PETERSON; PANOFSKY, 2020). Evidências empíricas metacientíficas têm demonstrado a baixa confiabilidade de estudos acarretada por mau uso de ferramentas estatísticas e metodológicas (ALTMAN, 1994; IOANNIDIS, 2005; IOANNIDIS *et al.*, 2014; IOANNIDIS; TRIKALINOS, 2005; PFEIFFER; BERTRAM; IOANNIDIS, 2011; WILSON; HARRIS; WIXTED, 2020). A evidência acumulada sugere um ecossistema científico de pesquisa e publicação em saúde que incentiva a quantidade de publicações em detrimento da qualidade (BIANCHETTI; ZUIN; FERRAZ, 2018; OLIVEIRA, 2015; IOANNIDIS *et al.*, 2014; HAVEN *et al.*, 2021; MACLEOD *et al.*, 2014; VILAÇA, 2015).

Problemas de escrita científica, tais como redação de conclusões que não são apoiadas pelos achados de pesquisa e vieses de publicação selecionando achados estatisticamente significantes, também têm sido documentados (ALVES *et al.*, 2020; OLIVEIRA, 2015; IOANNIDIS *et al.*, 2014; NISSEN *et al.*, 2016). Essas distorções podem ser perpetuadas na literatura, a partir de práticas de citação desatentas (GREENBERG, 2009; OLIVEIRA, 2015; IOANNIDIS *et al.*, 2014).

A identificação *a posteriori* desses problemas resulta em artigos retirados da literatura. A retratação é a punição máxima imputada a um artigo, quando a alegação de má conduta se confirma. As retratações acumuladas revelam os casos extremos, envolvendo falsificação e fabricação de dados na esfera da produção, e plágio na esfera da publicação (ADAM; ANDERSON; SPRONKEN-SMITH, 2017; BRAINARD; YOU, 2018; OLIVEIRA, 2015; WATANABE, 2014). A depender da legislação do país de origem, a descoberta da má conduta pode gerar repercussões civis e criminais também (OLIVEIRA, 2015; PEMBERTON *et al.*, 2019; WATANABE, 2014). Artigos retirados podem continuar a ser citados, contaminando a literatura com informações exageradas ou falsas (OLIVEIRA, 2015; GREENBERG, 2009; NISSEN *et al.*, 2016). Citações de artigos retratados foram recentemente apelidadas de 'citações-zumbi': aquelas que ocorrem após a 'morte' do artigo por retratação (BUCCI, 2019). O plágio é atualmente responsável pela retirada de metade dos artigos encontrados na literatura (BRAINARD; YOU, 2018).

Plágio: nuances e incertezas

Conceitualmente, plágio pode ser definido como o ato de copiar parcial ou totalmente o trabalho de terceiros sem a devida atribuição de autoria. Pode consistir em assumir ideias, conceitos, métodos descritos por outros; é uma prática que pode ser intencional ou não, configurando má conduta ou erro (LYKKES-FELDT, 2016; WORLD ASSOCIATION OF MEDICAL EDITORS, 2016). Na comunicação científica, a atribuição de autoria acontece por citação ao trabalho que primeiro concebeu uma ideia ou documentou um fenômeno natural. Portanto, independentemente de intencionalidade, o consenso é que se trata de uma prática que agride o sistema científico de recompensas (OLIVEIRA, 2015; WATANABE, 2014; WORLD ASSOCIATION OF MEDICAL EDITORS, 2016). Há evidências de que o plágio está associado a más condutas mais graves que podem gerar resultados falsos (EL-TAHAN, 2019). Presume-se que o plágio pode acontecer por diversos motivos, quais sejam: pressão acadêmica por reconhecimento e promoção; disputa para obtenção de recursos; falta de treinamento em escrita; escassez de integridade científica; dificuldade

com o idioma (HIGGINS; LIN; EVANS, 2016; YAHAGHI; SOROOSHIAN; YAHAGHI, 2017). O autoplágio consiste em plágio de trabalhos anteriores realizados pelo próprio autor (PEMBERTON *et al.*, 2019).

As práticas questionáveis em pesquisa e a publicação são valoradas de modo diferente em diversas culturas. Em função de traços culturais, alguns comportamentos são considerados como mais ou menos graves do que outros (ADAM; ANDERSON; SPRONKEN-SMITH, 2017; LI; CORNELIS, 2021). Infelizmente, nem sempre as práticas questionáveis em relato são passíveis de serem corrigidas durante o processo editorial. Uma intervenção visando provocar maior adesão a guias de completude de relatos disponibilizados pela iniciativa Equator Network (UK EQUATOR NETWORK CENTRE, 2021) obteve bons resultados em autores da revista *BMJ Open* (BLANCO *et al.*, 2020). Em contrapartida, iniciativas no intuito de reduzir a incidência de outras práticas questionáveis não obtiveram nenhum efeito significativo (BRUTON *et al.*, 2019; GHANNAD *et al.*, 2021).

Os experimentos de Blanco *et al.*, 2020, Bruton *et al.*, 2019 e Ghannad *et al.*, 2021, observados conjuntamente, indicam que é possível que as motivações dos autores possam variar expressivamente em função do sistema de incentivos e punições a que estão sujeitos (OLIVEIRA, 2015; HAVEN *et al.*, 2021; MACLEOD *et al.*, 2014) e da compreensão que eles têm acerca da produção, da escrita e da publicação científicas (ADAM; ANDERSON; SPRONKEN-SMITH, 2017; BRUTON *et al.*, 2019; PEMBERTON *et al.*, 2019). Esses aspectos interagem de forma complexa, e ainda não há consenso baseado em evidência de como isso acontece. A cultura organizacional do ambiente acadêmico e a pressão para obter publicações aparentam ter papel indutor de más práticas na pesquisa e escrita científicas (HAVEN *et al.*, 2021; VILAÇA, 2015).

Uma revisão de literatura sobre integridade científica e de relato no Brasil encontrou uma discussão incipiente. Debatem-se eminentemente plágio e conflitos de interesses, enquanto demais questões de integridade ficam à margem (PÁDUA; GUILHEM, 2015). Apesar do papel catalisador da educação e da cultura organizacional, uma análise das diretrizes curriculares e dos códigos de ética nacionais de biomedicina, enfermagem, farmácia, medicina, nutrição e odontologia observou lacunas. Tanto as diretrizes quanto os códigos discutiam termos relacionados à integridade científica e à integridade de relato de modo insuficiente (BARBOSA; RODRIGUES; NOVAES, 2019). A falta de diretrizes curriculares pode ser um indício de uma formação deficitária para o profissional e o pesquisador das ciências da saúde no Brasil.

Especificamente sobre plágio, há indícios de que as instituições de pesquisa no Brasil focam em medidas normativas e corretivas, mas não em diagnóstico (KROKOSZ, 2011). No cenário global, essa abordagem moralizadora e punitiva também domina a resposta da comunidade científica às más condutas. É aplicada mais por instituições de pesquisa do que por instituições de fomento (ŠČEPANOVIĆ *et al.*, 2021). Porém, falta evidência robusta sobre a eficácia dessa abordagem, que direciona o esforço para moralizar e punir o indivíduo, mas não trata de problemas estruturais como a cultura organizacional científica e o produtivismo acadêmico (BIANCHETTI; ZUIN; FERRAZ, 2018; OLIVEIRA, 2015; HAVEN *et al.*, 2021; VILAÇA, 2015).

No que diz respeito à legislação brasileira, o artigo 5 da Constituição Federal estabelece o direito de autor no ordenamento jurídico nacional (COSTA, 2016). O crime de violação de direitos de autor é tipificado no artigo 184 do Código Penal (BRASIL, 1940, 1984), em redação atualizada pela Lei 10.695 (BRASIL, 2003), que também complementa e atualiza a Lei 9.610 (BRASIL, 1998). Contudo, o ordenamento jurídico brasileiro não menciona formal e explicitamente o plágio acadêmico por erro. Na letra da lei, a infração aos direitos autorais é caracterizada como uma ação deliberada, intencional, objetivando o lucro direto ou indireto. Conceitualmente, é possível argumentar que o capital simbólico (BORDIEU, 2003, 2004, 2011a, 2011b) é um tipo de lucro. Essa abstração não está explicitada na letra da lei, dá margem a interpretações subjetivas e, portanto, não tem valor jurídico. Soma-se a isso que o princípio da anterioridade da lei estabelece que não há crime sem lei que o defina (BRASIL, 1940, 1984). Assim, a tipificação do plágio e o seu apenamento ficam sujeitos à jurisprudência mediante confirmação de intencionalidade.

A possibilidade de judicialização se torna ainda mais incerta quando se leva em conta o artigo 13 do Código Penal, que trata da omissão (BRASIL, 1940, 1984). Os avaliadores da revista e o editor são responsáveis por omissão, posto que poderiam identificar o plágio e não o fazem? Havendo tantas lacunas e ambiguidades na letra da lei, a judicialização de um plágio em uma revista científica é um caminho cheio de incertezas, com custos diretos e indiretos a todos os envolvidos na situação (OLIVEIRA, 2015; PEMBERTON *et al.*, 2019).

Não há lei internacional que regulamente a profissão de editor científico em ciências da saúde, mas diversas entidades internacionais possuem orientações no intuito de auxiliar os editores a prevenir e dirimir conflitos, inclusive no que se refere à suspeita de plágio. Além do WAME (WORLD ASSOCIATION OF MEDICAL EDITORS, c2016), destacam-se: o International Committee of Medical Journal Editors – ICMJE (INTERNATIONAL COMMITTEE OF MEDICAL JOURNAL EDITORS, c2021), o Committee on Publication Ethics – COPE (COMMITTEE ON PUBLICATION ETHICS, c2021) e o Office of Research Integrity – ORI (OFFICE OF RESEARCH INTEGRITY, c2021).

O ICMJE define quando é aceitável a duplicidade de publicações, mas recomenda o debate de alegações de plágio no fórum do COPE, para que se estabeleça intencionalidade e se decidam as consequências (INTERNATIONAL COMMITTEE OF MEDICAL JOURNAL EDITORS, c2021). O ORI possui fluxogramas de como escrever eticamente sem plágio, mas se exige de defini-lo de forma estrita, sugerindo que as instituições de origem dos autores conduzam investigações de alegações de plágio (OLIVEIRA, 2015; OFFICE OF RESEARCH INTEGRITY, 2021). Inferem-se um olhar de cautela por partes dessas entidades e fóruns consultivos na condução das alegações.

Do ponto de vista do ordenamento jurídico brasileiro, há muitas questões em aberto. Observando as orientações dos órgãos consultivos internacionais de editoração em saúde, observam-se cautela e uma preferência por autorregulação e autocontrole institucional (OLIVEIRA, 2015). Em um cenário de produção científica tensionado por pressões para publicar (BIANCHETTI; ZUIN; FERRAZ, 2018; OLIVEIRA, 2015; HAVEN *et al.*, 2021; MACLEOD *et al.*, 2014; VILAÇA, 2015), prevenir plágio e autoplagio parece ser o caminho mais eficiente para os editores de revistas científicas se protegerem de passivos jurídicos e evitarem desperdício de tempo e outros recursos.

Softwares como o iThenticate podem auxiliar na identificação de similaridades e reuso de textos (OLIVEIRA, 2015), mas há escassez de evidências sobre como usá-lo adequadamente. Uma busca nas bases de dados Latin American and Caribbean Health Sciences Literature (Lilacs) e Scientific Electronic Library Online (SciELO) sobre a operação do iThenticate indicou inexistência de relatos sobre o assunto. Na base de dados MEDLINE, o mesmo descritor retorna alguns relatos empíricos testando hipóteses e alguns editoriais. Essa lacuna justifica o relato de uma experiência em publicação científica nacional para apoiar editores brasileiros a operarem o iThenticate.

Diante disso, compartilhamos a experiência do Núcleo de Comunicação Científica (NUCC) da Escola Bahiana de Medicina e Saúde Pública, na operação do *software*. O portal publica seis títulos, que cobrem os seguintes temas: enfermagem, educação médica, fisioterapia, medicina, odontologia e psicologia (BAHIANA JOURNALS, c2018).

CASUÍSTICA

A política de detecção de similaridade do portal BAHIANA Journals foi implantada em 2016 em consonância com as recomendações do ICMJE e do WAME (INTERNATIONAL COMMITTEE OF MEDICAL JOURNAL EDITORS, c2021; WORLD ASSOCIATION OF MEDICAL EDITORS, c2016). Após revisão de literatura, o *software* escolhido foi o iThenticate (HIGGINS; LIN; EVANS, 2016; LYKKESFELDT, 2016).

Após a contratação do serviço, foi necessário decidir em que momento do fluxo editorial o texto seria analisado através do *software*. Se fosse analisado logo após a sua submissão à revista, problemas poderiam emergir no texto após avaliação por pareceristas e editores, quando os autores fizessem as correções solicitadas pelas revistas. Com isso, decidiu-se que o texto seria analisado para identificação de similaridades após avaliação por pareceristas, correção por autores e liberação para publicação por editores. Este fluxo é aplicado até hoje e está documentado em estatuto público disponível no portal em português e em inglês (BAHIANA JOURNALS, 2018; SÁ *et al.*, 2018).

A análise através do *software* é realizada antes do contato com os autores que confirma o aceite ou a recusa definitiva do texto. Os editores têm, assim, flexibilidade para recusar o texto por ele não ter atingido o padrão científico da revista ou por ter indícios de plágio. Essa decisão objetivou também diminuir o tempo desde a submissão até a publicação, reduzindo os custos e aumentando a eficiência da operação do portal. Além da anuidade pela assinatura do *software*, cada relatório emitido pelo iThenticate tem um custo individual em dólares, e as revistas do portal são totalmente subsidiadas pela instituição mantenedora, sem cobrança de nenhum tipo de taxa aos autores (BAHIANA JOURNALS, 2018; SÁ *et al.*, 2018).

O *software* iThenticate

O iThenticate compara o texto submetido à análise com 60 bilhões de páginas da internet e 150 milhões de conteúdos, dos quais 49 milhões são trabalhos acadêmicos publicados por cerca de 800 casas publicadoras (BASKARAN *et al.*, 2019). Com isso, identifica a repetição de uma quantidade predeterminada de palavras numa mesma ordem, mesmo que algumas palavras sejam removidas da frase ou substituídas por outras. Essas repetições compõem percentagens parciais de similaridade textual entre o documento analisado e os textos disponíveis na internet. Esses textos são apontados nos relatórios como fontes individuais. A soma desses percentuais compõe o escore total de similaridade (BASKARAN *et al.*, 2019; HIGGINS; LIN; EVANS, 2016; TURNITIN, 2021). A rigor, o iThenticate é um *software* de detecção de não originalidade textual. Sua operação não garante que se estabeleçam em definitivo intencionalidade e, portanto, infração de direitos autorais na acepção criminal, que é cercada de ambiguidades (PEMBERTON *et al.*, 2019). A avaliação de cada relatório individual por parte dos editores é fundamental (HIGGINS; LIN; EVANS, 2016; LYKKESFELDT, 2016).

Há evidência na literatura de que textos com mais de 20% de similaridade no total, após acionamento de filtros do próprio *software*, costumam apresentar baixa originalidade (BASKARAN *et al.*, 2019; EL-TAHAN, 2019; HIGGINS; LIN; EVANS, 2016; LYKKESFELDT, 2016; PARK *et al.*, 2017). Esses filtros ficam disponíveis em cada relatório, após ativação do botão “Texto-Only Report” no canto direito inferior (Figura 1). Servem para excluir falsos positivos contendo trechos insignificantes para a análise. O acionamento dos filtros faz com que o relatório fique mais acurado, ao desconsiderar citações diretas entre aspas (“Exclude Quotes”) e referências bibliográficas (“Exclude Bibliography”). Também é possível determinar, usando um dos filtros, quantas palavras constituem uma repetição significativa para compor o relatório (a opção “Limiting match size”). No BAHIANA Journals, configuramos os relatórios para detecção de repetições de até 10 palavras, porque na literatura a repetição de seis a vinte palavras foi independentemente documentada por estudos distintos como um indicativo de irregularidade no relato analisado (BASKARAN *et al.*, 2019; HIGGINS; LIN; EVANS, 2016; PARK *et al.*, 2017).

The screenshot displays the iThenticate software interface. At the top, it shows the date '09-May-2016 05:29PM', the document title '930 - EQUATOR Network resources have also been made available in', and the author 'BY JOÃO DE DEUS BARRETO SEGUNDO'. The interface includes a 'Match Overview' table on the right side, which lists seven matches, all with a 1% similarity score. The matches are categorized as 'Internet' and include details such as the number of words, the crawling date, and the source URL. A 'Text-Only Report' button is visible in the bottom right corner of the interface.

| Match | Words | Crawled On | Source | Similarity |
|-------|-------|-------------|-------------------------|------------|
| 1 | 36 | 09-Sep-2015 | journals.plos.org | 1% |
| 2 | 35 | 10-Feb-2015 | www.medioevoeuropeo.org | 1% |
| 3 | 22 | 22-Dec-2015 | neolaki.net | 1% |
| 4 | 22 | 09-May-2016 | www.paho.org | 1% |
| 5 | 18 | 08-Feb-2014 | www.amwa-doc.org | 1% |
| 6 | 17 | 06-May-2014 | www.ncbi.nlm.nih.gov | 1% |
| 7 | 15 | | | 1% |

Figura 1 – Captura de tela de relatório do iThenticate, mostrando o botão “Text-Only Report” no canto direito inferior. Fonte: imagem capturada pelos autores. Reprodução autorizada pela instituição contratante do serviço e pelo autor do texto analisado.

Entre 2016 e 2020, 972 documentos foram analisados através do *software*, após liberação para publicação pelos editores das seis revistas do portal. Destes, 847 foram publicados. Então, 87% dos textos avaliados por pareceristas, corrigidos por autores e liberados por editores foram definitivamente aprovados, após análise através do iThenticate. Equivale a uma incidência de plágio de 13%, ou 125 submissões. Essas 125 submissões foram recusadas.

Retrospectivamente, quando se projeta essa incidência no total de submissões do portal durante cinco anos ($n=2.146$), estima-se que 278 textos poderiam ter recebido diagnóstico de plágio, caso tivessem sido analisados através do *software*, antes da avaliação pelos pareceristas. Isso informa também que pelo menos metade dos textos com plágio poderiam não ter gerado suspeitas nos pareceristas e editores científicos, durante a avaliação, posto que 125 chegaram à aprovação final com problemas de não originalidade na escrita. Isso reforça a necessidade do uso de ferramentas automatizadas de detecção de similaridade como o iThenticate. A seguir, há o que aprendemos sobre o reúso de texto e a escrita científica nesses cinco anos empregando o *software* como ferramenta de apoio ao nosso trabalho.

Escore de similaridade

Após o acionamento dos filtros, a maioria dos textos liberados para publicação apresentou menos de 20% de similaridade com textos previamente publicados. A mediana dos escores totais foi de 11,5%, a variância foi 165%, o menor escore foi 0%, e o maior, 89%. A nossa experiência (Figura 2), portanto, no que diz respeito aos escores totais, mostrou-se em consonância com a literatura (BASKARAN *et al.*, 2019; EL-TAHAN, 2019; HIGGINS; LIN; EVANS, 2016; LYKKESFELDT, 2016; PARK *et al.*, 2017).

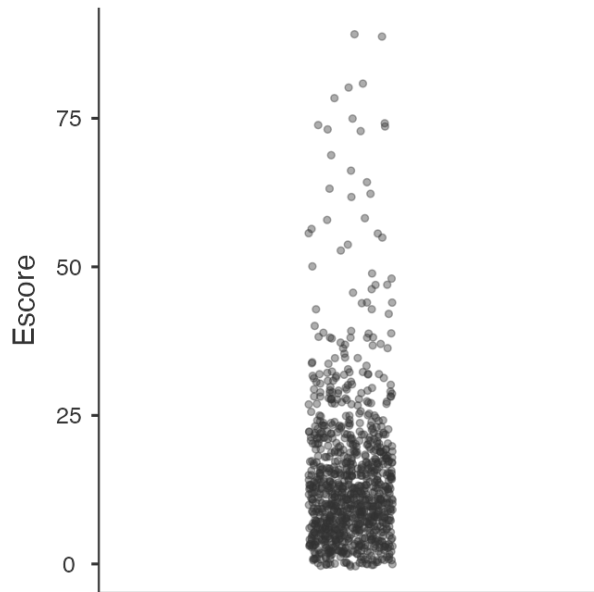


Figura 2 – Densidade de escores totais de similaridade textual em intervalos percentuais em cinco anos de uso do *software* iThenticate no portal BAHIANA Journals
Fonte: elaboração dos autores.

Observando os relatórios individuais emitidos pelo iThenticate, alguns textos exibiram indícios de reuso textual expressivo, apesar de apresentarem menos de 20% no escore total de similaridade. Por simetria, o inverso poderia ser verdadeiro: textos acima de 20% de similaridade no total, mas essencialmente originais. Percebeu-se, assim, que somente considerar o percentual total de similaridade disponível na literatura poderia resultar em má interpretação dos relatórios, prejuízo aos autores e dano ao portal. Portanto, notou-se a necessidade de critérios adicionais para complementar a análise. As porcentagens parciais das fontes individuais de similaridade foram incluídas obrigatoriamente nas análises.

A tabela a seguir agrupa a mediana dos escores totais em função do escore parcial da primeira fonte individual que apresentou o maior índice de similaridade no relatório (Tabela 1). Há aproximação entre principal escore parcial e escore total para os relatos que tiveram pelo menos uma fonte parcial igual ou acima de 30%. Contudo, textos com escores totais abaixo disso apresentaram, em alguns casos, uma única fonte com 10% a 25% de não originalidade. Textos com menos de 30% no escore total poderiam, assim, conter uma única fonte parcial com indício de plágio, em conformidade com a literatura (EL-TAHAN, 2019; HIGGINS; LIN; EVANS, 2016; LYKKESFELDT, 2016).

Muitas fontes parciais diluem o escore total. Mas, quanto menor é o escore total, mais sensível ele se mostra a fontes parciais elevadas. Inversamente, quanto maior a fonte parcial principal, mais ela impacta no escore total. Ou seja, há indícios de plágio quando os textos se apoiam apenas em uma ou duas fontes para sua redação, reutilizando expressivamente trechos dessas fontes, entre 10% e 40% de reuso (EL-TAHAN, 2019).

Tabela 1 – Medianas dos escores totais em função de fonte parcial de maior índice de similaridade

| Escore parcial | Mediana do escore total | n |
|-----------------------|--------------------------------|----------|
| ≤5% | 9% | 755 |
| ≤10% | 21% | 126 |
| ≤15% | 28% | 37 |
| ≤20% | 32% | 16 |
| ≤25% | 35% | 11 |
| ≤30% | 31% | 8 |
| ≤40% | 54% | 3 |
| ≤50% | 47% | 3 |
| ≤60% | 55% | 2 |
| ≤70% | 71% | 4 |
| ≤80% | 74% | 5 |
| ≤90% | 89% | 2 |

Fonte: elaboração dos autores.

Para compor a tabela, os escores parciais foram transformados em variáveis categóricas dicotomizadas. Considerou-se a maior fonte de similaridade parcial de cada relatório único para tal. Ou seja, um artigo cuja maior fonte parcial continha 55% de similaridade foi incluído na categoria ≤60%. Artigos com os escores parciais de similaridade zerados foram incluídos em ≤5%. Como a densidade de similaridade rareou expressivamente a partir de ≤30% (Figura 2), para fins de didatismo, os autores optaram por não usar ≤35%, ≤45% etc. como categorias. Esta escolha foi necessária em particular pela escassez de parciais elevadas em categorias acima de ≤30% (n=8). Portanto, todos os escores parciais depois de ≤25% foram aproximados.

Observando os escores parciais como condicionantes dos escores totais (EL-TAHAN, 2019), percebemos que textos com menos de 30% de similaridade no total e, simultaneamente, menos de 5% de similaridade na maior fonte parcial costumam ser textos sem indícios de plágio.

A seção de métodos

A análise da seção de métodos dos relatos trouxe complicadores. Existem inúmeros desenhos de estudo na pesquisa em saúde, mas não existem muitas formas de descrever uma seleção e um cálculo amostrais – ou seja, o processo de randomização de um ensaio clínico ou suas estratégias de cegamento para evitar vieses de aferição e análise (PEMBERTON *et al.*, 2019). Os guias de relato do Equator Network servem para padronizar os relatos, de modo que a diversidade cultural não vire barreira para a completude e a transparência do relato. Criam, assim, uma homogeneidade na escrita científica em saúde, estabelecendo os elementos mínimos que um relato deve conter (UK EQUATOR NETWORK CENTRE, c2021).

A reprodução de estudos prévios aumenta a confiabilidade e a robustez da literatura biomédica. Para isso, métodos não só podem como devem ser copiados em estudos que objetivam acumular evidências sobre determinada intervenção. Esse processo é fundamental para que as magnitudes de efeito estimadas estatisticamente sejam próximas da efetividade da intervenção no mundo real (BOUTER; RIET, 2021; IOANNIDIS; TRIKALINOS, 2005). Entretanto, a reprodução de estudos prévios, ou a mera reprodução de métodos comuns na pesquisa, em geral, eleva os valores individuais e absolutos de similaridade do texto, conforme nossa experiência. Em alguns casos, o reuso textual na seção de métodos objetiva clareza e concisão para evitar enviesamento dos resultados e das conclusões (PEMBERTON *et al.*, 2019).

Com isso, percebemos a necessidade de desconsiderar os índices de similaridade da seção de métodos em alguns casos, quando o estudo anuncia expressamente que é uma reprodução de outro estudo prévio ou quando replica, em métodos, informações já elencadas no protocolo pré-registrado em bases como a ClinicalTrials e afins. Isso está em consonância com as recomendações de estudos prévios (HIGGINS; LIN; EVANS, 2016; LYKKESFELDT, 2016).

Repositórios institucionais

Também em consonância com as práticas de transparência e a integridade científicas, dissertações e teses são costumeiramente postadas em repositórios institucionais (LEITE; ASSIS; MELO, 2015; WATANABE, 2014). Em ciências da saúde, é comum que dissertações e teses originem os artigos que compõem etapas adjuntas à defesa para obtenção do título de mestre ou doutor. Artigos que resultam de dissertações e teses elaboradas nesses contextos, uma vez analisados no iThenticate, podem resultar em percentuais elevados de similaridade parcial e total, mesmo com todos os filtros acionados.

A seguir, na Tabela 2, percebe-se que 10% das fontes parciais de maior índice vieram de reuso de trechos de textos depositados previamente em repositórios institucionais. Sendo assim, esse tipo de reuso aparenta ser uma prática comum na amostra. Em geral, não ocasionou incremento expressivo da fonte parcial de similaridade. Mas, nas poucas vezes em que provocou essa elevação, a principal fonte parcial oscilou entre 50% e 80%. A comparação da lista de autores do texto analisado no iThenticate com a lista de autores da dissertação ou da tese pôde, nesses casos, oferecer mais pistas aos editores sobre indícios de plágio ou não.

Em alguns casos, os autores informaram na própria submissão que os seus textos eram oriundos de suas dissertações e teses e que estas se encontravam depositadas em repositório. Conforme estatuto do portal e políticas editoriais de todas as revistas, o portal estimula as práticas de ‘ciência aberta’ e, com isso, o depósito em repositórios (BAHIANA JOURNALS, 2018; SÁ *et al.*, 2018).

Tabela 2 – Coincidência de fonte parcial de maior índice de similaridade com texto depositado em repositório institucional

| Parcial | Não | Sim |
|---------|-----|-----|
| ≤5% | 681 | 74 |
| ≤10% | 110 | 16 |
| ≤15% | 34 | 3 |
| ≤20% | 13 | 3 |
| ≤25% | 7 | 4 |
| ≤30% | 6 | 2 |
| ≤40% | 3 | 0 |
| ≤50% | 2 | 1 |
| ≤60% | 0 | 2 |
| ≤70% | 0 | 4 |
| ≤80% | 2 | 3 |
| ≤90% | 2 | 0 |

Fonte: elaboração dos autores.

Repositórios são fontes informacionais, mas não constituem uma publicação científica em sentido estrito (LEITE; ASSIS; MELO, 2015). Não possuem ISSN (International Standard Serial Number) ou ISBN (International Standard Book Number). Há evidência de que pode não haver consenso entre editores científicos sobre o que configura reuso aceitável e o que configuraria autoplágio (PEMBERTON *et al.*, 2019). Com isso,

amparados pelo estatuto, pelas políticas editoriais e pela comparação de autoria entre texto analisado e texto de repositório, decidimos avaliar caso a caso.

Recusar artigos originários de dissertações e teses já defendidos e postados em repositórios pelo mesmo autor poderia, em alguns casos, constituir uma perda em transparência científica. Comparar a dissertação ou tese postada em repositório com o artigo final aceito pela revista pode ser inclusive uma forma indireta de estimar a qualidade e a profundidade da revisão por pares empregada pela revista. A expectativa é que a revisão por pares melhore a escrita do relato e aprofunde a discussão dos achados da pesquisa.

Vale ressaltar também que a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), fazendo uso de suas atribuições, por meio do seu sistema de avaliação, tem moldado e direcionado o comportamento dos docentes e coordenadores de programas de mestrado e doutorado (BIANCHETTI; ZUIN; FERRAZ, 2018). Os produtos que foram frutos da formação de cientistas egressos dos programas, conforme recomendações da agência, devem ser tornados públicos para que o esforço dos envolvidos sirva ao desenvolvimento da ciência. Um caminho para isso tem sido o depósito de produções de programas de pós-graduação em repositórios institucionais (WATANABE, 2014). Portanto, considerar autoplágio a publicação em artigos de dados de uma dissertação ou tese seria um contrassenso. O uso de trechos de dissertações e teses de terceiros contudo segue, no entendimento dos editores do portal, condicionado ao uso racional desses trechos e às boas práticas de citação e atribuição de autoria (OLIVEIRA, 2015).

A comunicação com os autores

De todo modo, compreendemos que a função dos editores científicos e executivos em uma instituição de ensino deve ser servir à formação científica e tecnológica da sociedade, a começar pelos autores do portal. Por esse motivo, a identificação de um plágio, ainda que não intencional, é tratada como uma oportunidade formativa para contribuir para o avanço científico, através do apoio à capacitação de autores em escrita científica com integridade (SANCHES, 2019; SILVA, 2008).

Textos que apresentaram leve incremento nos indicadores de similaridade puderam ser devolvidos aos autores para reescrita e ajustes. Alguns puderam ser publicados depois de terem as similaridades ajustadas e as devidas citações incluídas. Na comunicação com os autores, o arquivo de texto foi devolvido com os trechos que necessitavam de reescrita destacados. O contrato com a Turnitin, fornecedora do iThenticate, orienta o não compartilhamento do relatório do *software*. O compartilhamento sujeitaria a instituição contratante a passivos jurídicos.

A palavra plágio também não é usada na comunicação com o autor. Primeiramente, porque a análise qualitativa e quantitativa do relatório não subsidia que se estabeleça intencionalidade, segundo porque os organismos internacionais de apoio à atividade editorial científica nas ciências biomédicas pregam cautela e, terceiro, para evitar passivos jurídicos relacionados a crimes de calúnia, injúria e difamação tipificados no Código Penal brasileiro nos artigos 138, 139 e 140 (BRASIL, 1940). Os autores têm o benefício da dúvida e o apoio dos editores na reescrita. A comunicação do índice elevado de similaridade textual informa que aqueles trechos destacados no arquivo anexado devem ser reescritos como condição obrigatória para publicação. Autores que não respondem a três sucessivas cobranças de correção de texto por *e-mail* têm a sua submissão cancelada na respectiva revista. Cada cobrança delimita um prazo de dez dias para as correções finais. Assim, o autor recebe 30 dias de prazo no total para correções.

Textos que apresentaram grande variação positiva nos indicadores de similaridade, sem as devidas referências, sem configurarem reprodução de estudos prévios ou partes de dissertações e teses, tiveram a submissão recusada sem contato com os autores orientando reescrita. O entendimento dos editores é que esses textos precisariam ser totalmente reescritos. Conclui-se, assim, que é oportuno que o autor seja o último a ser acionado. Isso deve acontecer com cautela, sem precipitação e somente após haver deliberação

por parte dos editores baseada nas evidências do relatório emitido pela ferramenta, associada à análise qualitativa do texto e do contexto. O cuidado na interpretação do relatório é fundamental (HIGGINS; LIN; EVANS, 2016; LYKKESFELDT, 2016).

CONSIDERAÇÕES FINAIS

Observando-se o escore total de similaridade, o escore por fontes principais, o contexto em que aquelas similaridades ocorreram, e observando-se também o reúso racional e justificado de trechos de textos e as boas práticas de citação e atribuição de autoria, é possível fazer diagnósticos precisos e eficientes. Percebe-se que menos de 30% de escore total de similaridade associado a menos de 5% de similaridade parcial nas fontes individuais, quando ocorrem simultaneamente, estão associados a textos originais e bem redigidos. Mas isso é só o começo da análise quantitativa. Associada à análise quantitativa, deve ser feita a análise qualitativa do relato em função de critérios, como a análise da seção de métodos e de possíveis depósitos em repositório institucional.

O que se percebeu nesses cinco anos é que a incidência de similaridade é pouco frequente. Isso é compatível com a literatura sobre práticas questionáveis em publicação, que traz evidência de que erros tendem a ser mais frequentes do que fraudes, havendo em geral uma baixa propensão a deturpar a literatura científica por execução deliberada de atos censuráveis de grande monta (BRAINARD; YOU, 2018; LYKKESFELDT, 2016; STEEN, 2011; STEEN; CASADEVALL; FANG, 2013).

Na experiência do portal, aqueles autores que, apesar de terem boas ideias e pesquisas interessantes, sofriam com dificuldades na escrita, ao receberem a devolutiva que identificava as similaridades e os trechos onde deveria ocorrer a reescrita, devolveram textos de melhor qualidade. Quando a não originalidade textual é dessa natureza e, portanto, muito mais uma dificuldade educacional do que moral, não há por que não apoiar os autores na reescrita. Com isso, ganham os autores, as revistas e a ciência. Resta, então, que as instituições de pesquisa e de fomento se empenhem em realizar treinamentos em integridade de relato, ensinando os autores a escreverem melhor nas ciências da saúde.

Diante desta reflexão, recomendamos que outras revistas reproduzam a iniciativa de implantação de políticas de detecção de similaridade com abordagens cautelosas na condução desse momento crítico junto aos autores e que compartilhem o que foi aprendido nessas interações. Recomendamos que os editores estejam sempre atentos às nuances e executem a avaliação do texto baseado na melhor evidência disponível, almejando sempre trazer os autores de volta às melhores práticas, quando for possível. Editores e revistas podem fortalecer qualidade e confiabilidade no ecossistema de publicação científica nas ciências da saúde e multiplicar as melhores práticas para outras áreas da pesquisa e da publicação científica.

REFERÊNCIAS

ADAM, Lee; ANDERSON, Vivienne; SPRONKEN-SMITH, Rachel. 'It's not fair': policy discourses and students' understandings of plagiarism in a New Zealand university. **Higher Education**, Amsterdã, v. 74, n. 1, p. 17-32, 2017. DOI: <https://doi.org/10.1007/s10734-016-0025-9>. Disponível em: <https://link.springer.com/article/10.1007/s10734-016-0025-9>. Acesso em: 26 jul. 2021.

ALTMAN, Douglas. The scandal of poor medical research. **The BMJ**, Londres, v. 398, p. 283, Jan. 29, 1994. DOI: <https://doi.org/10.1136/bmj.308.6924.283>. Disponível em: <https://www.bmj.com/content/308/6924/283>. Acesso em: 26 jul. 2021.

ALVES, Charles Phillipe de Lucena *et al.* Spin: modificações na redação científica que escondem fragilidades metodológicas com impacto social negativo. **Journal of Evidence-based Healthcare**, Salvador, v. 2, n. 1, p. 97-105, 2020. DOI: <http://www.doi.org/10.17267/2675-021Xevidence.v2i1.2840>. Disponível em: <https://www5.bahiana.edu.br/index.php/evidence/article/view/2840/3481>. Acesso em: 26 jul. 2021.

BAHIANA JOURNALS. **Homepage**. Salvador: BAHIANA Journals, c2018. Disponível em: www.journals.bahiana.edu.br. Acesso em: 31 jan. 2021.

BARBOSA, Quésia Ferreira; RODRIGUES, Camila Serra; NOVAES, Maria Rita Carvalho Garbi. Integridade científica na educação de profissionais de saúde. **Revista Bioética**, Brasília, DF, v. 27, n. 1, p. 120-126, 2019. DOI: <https://doi.org/10.1590/1983-80422019271294>. Disponível em: <https://www.scielo.br/j/bioet/a/6kpntFmWcgWPRVWcvd9jbrP/?lang=pt>. Acesso em: 26 jul. 2021.

BASKARAN, Saradha *et al.* Is there plagiarism in the most influential publications in the field of andrology?. **Andrologia**, [s. l.], v. 51, n. 10, p. e13405, 2019. DOI: <https://doi.org/10.1111/and.13405>. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/and.13405>. Acesso em: 26 jul. 2021.

BIANCHETTI, Lucídio; ZUIN, Antônio A.S.; FERRAZ, Obdália. **Publique, apareça ou pereça**: produtivismo acadêmico, “pesquisa administrada” e plágio nos tempos da cultura digital. Salvador: EDUFBA, 2018.

BLANCO, David *et al.* Effect of an editorial intervention to improve the completeness of reporting of randomised trials: a randomised controlled trial. **BMJ Open**, Londres, v. 10, n. 5, p. e036799, 2020. DOI: <http://dx.doi.org/10.1136/bmjopen-2020-036799>. Disponível em: <https://bmjopen.bmj.com/content/10/5/e036799>. Acesso em: 26 jul. 2021.

BOURDIEU, Pierre. **A economia das trocas simbólicas**. São Paulo: Perspectiva, 2011a.

BOURDIEU, Pierre. **Os usos sociais da ciência**: por uma sociologia clínica do campo científico. São Paulo: UNESP, 2003.

BOURDIEU, Pierre. **O poder simbólico**. Lisboa: Edições 70, 2011b.

BOURDIEU, Pierre. **Para uma sociologia da ciência**. Lisboa: Edições 70, 2004.

BOUTER, Lex; RIET, Gerben. Replication Research Series-Paper 2: empirical research must be replicated before its findings can be trusted. **Journal of Clinical Epidemiology**, Oxford, v. 129, p. 188-190, 2021. DOI: <https://doi.org/10.1016/j.jclinepi.2020.09.032>. Disponível em: [https://www.jclinepi.com/article/S0895-4356\(20\)31118-5/fulltext](https://www.jclinepi.com/article/S0895-4356(20)31118-5/fulltext). Acesso em: 26 jul. 2021.

BRAINARD, Jeffrey; YOU, Jia. What a massive database of retracted papers reveals about science publishing's 'death penalty'. **Science**, Washington (DC), 25 out. 2018. Scientific Community. DOI: <https://doi.org/10.1126/science.aav8384>. Disponível em: <https://www.sciencemag.org/news/2018/10/what-massive-database-retracted-papers-reveals-about-science-publishing-s-death-penalty>. Acesso em: 11 jul. 2021.

BRASIL. Presidência da República. Casa Civil. Decreto-Lei nº 2.848, de 7 de dezembro de 1940. **Diário Oficial da União**, Brasília, DF, p. 2391, 31 dez. 1940. Disponível em: http://www.planalto.gov.br/ccivil_03/decreto-lei/del2848compilado.htm. Acesso em: 18 jul. 2021.

BRASIL. Presidência da República. Casa Civil. **Lei nº 7.209, de 11 de julho de 1984**. Altera dispositivos do Decreto-Lei nº 2.848, de 7 de dezembro de 1940 - Código Penal, e dá outras providências. Brasília, DF: DOFC, 1984. Disponível em: http://www.planalto.gov.br/ccivil_03/LEIS/1980-1988/L7209.htm#art13. Acesso em: 18 jul. 2021.

BRASIL. Presidência da República. Casa Civil. Lei nº 9.610, de 19 de fevereiro de 1998. Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências. **Diário Oficial da União**, Brasília, DF, p. 3, 20 fev. 2021. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/19610.htm. Acesso em: 18 jul. 2021.

BRASIL. Presidência da República. Casa Civil. Lei nº 10.695, de 1º de julho de 2003. Altera e acresce parágrafo ao art. 184 e dá nova redação ao art. 186 do Decreto-Lei no 2.848, de 7 de dezembro de 1940 – Código Penal, alterado pelas Leis nos 6.895, de 17 de dezembro de 1980, e 8.635, de 16 de março de 1993, revoga o art. 185 do Decreto-Lei no 2.848, de 1940, e acrescenta dispositivos ao Decreto-Lei no 3.689, de 3 de outubro de 1941 – Código de Processo Penal. **Diário Oficial da União**, Brasília, DF, p. 1, 02 jul. 2003. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/2003/10.695.htm. Acesso em: 18 jul. 2021.

BRUTON, Samuel *et al.* Testing an active intervention to deter researchers' use of questionable research practices. **Research Integrity and Peer Review**, Londres, v. 4, p. 24, 2019. DOI: <https://doi.org/10.1186/s41073-019-0085-3>. Disponível em: <https://researchintegrityjournal.biomedcentral.com/articles/10.1186/s41073-019-0085-3>. Acesso em: 26 jul. 2021.

BUCCI, Enrico. On zombie papers. **Cell Death & Disease**, Londres, v. 10, n. 3, p. 189, 2019. DOI: <https://doi.org/10.1038/s41419-019-1450-3>. Disponível em: <https://www.nature.com/articles/s41419-019-1450-3>. Acesso em: 26 jul. 2021.

COMMITTEE ON PUBLICATION ETHICS. **Homepage**. Hampshire: COPE, c2021. Disponível em: <https://publicationethics.org>. Acesso em: 16 jul. 2021.

COSTA, Rosa Maria Cardoso Dalla. Plágio acadêmico: a responsabilidade das associações científicas. **Intercom: Revista Brasileira de Ciências da Comunicação**, São Paulo, v. 39, n. 3, p. 187-200, set.-dez. 2016. DOI: <https://doi.org/10.1590/1809-58442016311>. Disponível em: <https://www.scielo.br/j/intercom/a/gcJYXmNvb8qrJ3fk5B6kRmH/?lang=pt>. Acesso em: 26 jul. 2021.

EL-TAHAN, Mohamed R. Can the similarity index predict the causes of retractions in high-impact anesthesia journals? A bibliometric analysis. **Saudi Journal of Anaesthesia**, Riade, v. 13, n. 5, p. 2-8, 2019. DOI: https://dx.doi.org/10.4103%2Fsja.SJA_709_18. Disponível em: <https://www.saudija.org/article.asp?issn=1658-354X;year=2019;volume=13;issue=5;spage=2;epage=8;aulast=El%2DTahan>. Acesso em: 26 jul. 2021.

GHANNAD, Mona *et al.* A randomized trial of an editorial intervention to reduce spin in the abstract's conclusion of manuscripts showed no significant effect. **Journal of Clinical Epidemiology**, Oxford, v. 130, p. 69-77, 2021. DOI: <https://doi.org/10.1016/j.jclinepi.2020.10.014>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0895435620311537>. Acesso em: 09 fev. 2021.

GREENBERG, Steven. How citation distortions create unfounded authority: analysis of a citation network. **The BMJ**, Londres, v. 339, p. b2680, 2009. DOI: <https://doi.org/10.1136/bmj.b2680>. Disponível em: <https://www.bmj.com/content/339/bmj.b2680>. Acesso em: 01 fev. 2021.

HAVEN, Tamarinde *et al.* Explaining variance in perceived research misbehavior: results from a survey among academic researchers in Amsterdam. **Research Integrity and Peer Review**, Londres, v. 6, p. 7, 2021. DOI: <https://doi.org/10.1186/s41073-021-00110-w>. Disponível em: <https://researchintegrityjournal.biomedcentral.com/articles/10.1186/s41073-021-00110-w>. Acesso em: 26 jul. 2021.

HIGGINS, Janet; LIN, Feng-Chang; EVANS, James P. Plagiarism in submitted manuscripts: incidence, characteristics and optimization of screening: case study in a major specialty medical journal. **Research Integrity and Peer Review**, Londres, v. 1, n. 13, p. 1-8, 2016. DOI: <https://doi.org/10.1186/s41073-016-0021-8>. Disponível em: <https://researchintegrityjournal.biomedcentral.com/articles/10.1186/s41073-016-0021-8>. Acesso em: 26 jul. 2021.

INTERNATIONAL COMMITTEE OF MEDICAL JOURNAL EDITORS. **Recommendations**. [S. l.]: Annals of Internal Medicine: American College of Physicians, c2021. Disponível em: <http://www.icmje.org/recommendations/>. Acesso em: 18 jul. 2021.

IOANNIDIS, John. Why most published research findings are false. **PLoS Medicine**, São Francisco, v. 2, n. 8, p. e124, 2005. DOI: <https://doi.org/10.1371/journal.pmed.0020124>. Disponível em: <https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0020124>. Acesso em: 26 jul. 2021.

IOANNIDIS, John; TRIKALINOS, Thomas. Early extreme contradictory estimates may appear in published research: the Proteus phenomenon in molecular genetics research and randomized trials. **Journal of Clinical Epidemiology**, Nova York, v. 58, n. 6, p. 543-549, 2005. DOI: <https://doi.org/10.1016/j.jclinepi.2004.10.019>. Disponível em: [https://www.jclinepi.com/article/S0895-4356\(05\)00023-5/abstract](https://www.jclinepi.com/article/S0895-4356(05)00023-5/abstract). Acesso em: 26 jul. 2021.

IOANNIDIS, John *et al.* Increasing value and reducing waste in research design, conduct, and analysis. **The Lancet**, Londres, v. 383, n. 9.912, p.166-175, 2014. DOI: [https://doi.org/10.1016/S0140-6736\(13\)62227-8](https://doi.org/10.1016/S0140-6736(13)62227-8). Disponível em: [https://www.thelancet.com/pdfs/journals/lancet/PIIS0140-6736\(13\)62227-8.pdf](https://www.thelancet.com/pdfs/journals/lancet/PIIS0140-6736(13)62227-8.pdf). Acesso em: 09 fev. 2021.

KROKOSCZ, Marcelo. Abordagem do plágio nas três melhores universidades de cada um dos cinco continentes e do Brasil. **Revista Brasileira de Educação**, Rio de Janeiro, v. 16, n. 48, p. 745-818, 2011. Disponível em: <https://www.scielo.br/j/rbedu/a/tKsDQfr6xgRGbNTghvQRFnK/?lang=pt&format=pdf>. Acesso em: 26 jul. 2021.

LEITE, Fernando César Lima; ASSIS, Tainá Batista de; MELO, Bianca Amaro de. Gestão de teses e dissertações eletrônicas no Brasil: sobre bibliotecas digitais de teses e dissertações e repositórios institucionais. **Informação & Informação**, Londrina, v. 20, n. 3, p. 529-543, 2015. DOI: <https://www.doi.org/10.5433/1981-8920.2015v20n3p529>. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/21090/17675>. Acesso em: 26 jul. 2021.

LI, Dan; CORNELIS, Gustaaf. Differing perceptions concerning research misconduct between China and Flanders: a qualitative study. **Accountability in Research**, [s. l.], v. 28, n. 2, p. 63-94, 2021. DOI: <https://www.doi.org/10.1080/08989621.2020.1802586>. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/08989621.2020.1802586?journalCode=gacr20>. Acesso em: 26 jul. 2021.

LYKKESFELDT, Jens. Strategies for using plagiarism software in the screening of incoming journal manuscripts: recommendations based on a recent literature survey. **Basic & Clinical Pharmacology & Toxicology**, Copenhagen, v. 119, n. 2, p. 161-164, 2016. DOI: <https://www.doi.org/10.1111/bcpt.12568>. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1111/bcpt.12568>. Acesso em: 26 jul. 2021.

MACLEOD, Malcolm *et al.* Biomedical research: increasing value, reducing waste. **The Lancet**, Londres, v. 383, n. 9.912, p. 101-104, 2014. DOI: [https://doi.org/10.1016/S0140-6736\(13\)62329-6](https://doi.org/10.1016/S0140-6736(13)62329-6). Disponível em: [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(13\)62329-6/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(13)62329-6/fulltext). Acesso em: 31 jan. 2021.

NISSEN, Silas Boye *et al.* Research: publication bias and the canonization of false facts. **eLife**, Cambridge, v. 5, p. e21451, 2016. DOI: <https://www.doi.org/10.7554/eLife.21451>. Disponível em: <https://elifesciences.org/articles/21451>. Acesso em: 26 jul. 2021.

O'BOYLE JR., Ernest Hugh; BANKS, George Christopher; GONZALEZ-MULÉ, Erik. The Chrysalis Effect: how ugly initial results metamorphosize into beautiful articles. **Journal of Management**, [s. l.], v. 43, n. 2, p. 376-399, 2017. DOI: <https://doi.org/10.1177/0149206314527133>. Disponível em: <https://journals.sagepub.com/doi/abs/10.1177/0149206314527133?journalCode=joma>. Acesso em: 31 jan. 2021.

OFFICE OF RESEARCH INTEGRITY. **ORI Policy on Plagiarism**. Rockville: U.S. Department of Health and Human Services, c2021. Disponível em: <https://ori.hhs.gov/ori-policy-plagiarism>. Acesso em: 16 jul. 2021.

OLIVEIRA, Marcos Barbosa de. A epidemia de más condutas na ciência: o fracasso do tratamento moralizador. **Scientiæ Studia**, São Paulo, v. 13, n. 4, p. 867-897, 2015. DOI: <https://doi.org/10.1590/S1678-31662015000400007>. Disponível em: <https://www.scielo.br/j/ss/a/hMs7Y4FVZvbJwHkmYKtn3Ny/?lang=pt>. Acesso em: 26 jul. 2021.

PÁDUA, Gabriela Cristina Cantisani; GUILHEM, Dirce. Integridade científica e pesquisa em saúde no Brasil: revisão da literatura. **Revista Bioética**, Brasília, DF, v. 23, n. 1, p. 124-138, 2015. DOI: <http://dx.doi.org/10.1590/1983-80422015231053>. Disponível em: <https://www.scielo.br/j/bioet/a/Y4PhdYmqhqq6wScFYPBdSsr/?lang=pt>. Acesso em: 26 jul. 2021.

PARK, Soyoungh *et al.* Similarity analysis of Korean medical literature and its association with efforts to improve research and publication ethics. **Journal of Korean Medical Science**, Seul, v. 32, n. 6, p. 887-892, 2017. DOI: <https://doi.org/10.3346/jkms.2017.32.6.887>. Disponível em: <https://jkms.org/DOIx.php?id=10.3346/jkms.2017.32.6.887>. Acesso em: 26 jul. 2021.

PEMBERTON, Michael *et al.* Text recycling: views of North American journal editors from an interview-based study. **Learned Publishing**, [s. l.], v. 32, n. 4, p. 355-366, 2019. DOI: <https://doi.org/10.1002/leap.1259>. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/leap.1259>. Acesso em: 26 jul. 2021.

PETERSON, David; PANOFKY, Aaron. Metascience as a scientific social movement. **SocArXiv Papers**, [College Park], p. 1-33, 2020. DOI: <https://doi.org/10.31235/osf.io/4dsqa>. Disponível em: <https://osf.io/preprints/socarxiv/4dsqa/>. Acesso em: 26 jul. 2021.

PFEIFFER, Thomas; BERTRAM, Lars; IOANNIDIS, John. Quantifying selective reporting and the Proteus phenomenon for multiple datasets with similar bias. **PLoS ONE**, São Francisco, v. 6, n. 3, p. e18362, 2011. DOI: <https://doi.org/10.1371/journal.pone.0018362>. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0018362>. Acesso em: 31 jan. 2021.

SÁ, Katia Nunes *et al.* **Estatuto BAHIANA Journals**. Salvador: Escola Bahiana de Medicina e Saúde Pública, 2018. Disponível em: <https://journalsbahiana.files.wordpress.com/2018/10/dq-nucc-03-estatuto-bahiana-journals-v00-cc.pdf>. Acesso em: 31 jan. 2021.

SANCHES, Tatiana. Citar e referenciar: uma estratégia formativa para o uso ético da informação e prevenção do plágio em meio acadêmico. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 24, n. 3, p. 59-72, 2019. DOI: <https://doi.org/10.1590/1981-5344/3214>. Disponível em: <https://www.scielo.br/j/pci/a/hNDBzcbRwyy9bYGTFVt7wDS/?lang=pt>. Acesso em: 26 jul. 2021.

ŠČEPANOVIĆ, Rea *et al.* Practices for research integrity promotion in research performing organizations and research funding organizations: a scoping review. **Science and Engineering Ethics**, [s. l.], v. 27, n. 4, p. 1-20, 2021. DOI: <https://doi.org/10.1007/s11948-021-00281-1>. Disponível em: <https://link.springer.com/article/10.1007%2Fs11948-021-00281-1>. Acesso em: 26 jul. 2021.

SILVA, Obdália Santana Ferraz. Entre o plágio e a autoria: qual o papel da universidade?. **Revista Brasileira de Educação**, Rio de Janeiro, v. 13, n. 38, p. 357-414, 2008. DOI: <https://doi.org/10.1590/S1413-24782008000200012>. Disponível em: <https://www.scielo.br/j/rbedu/a/PK7VSKjhMWTqCrspQrVYTDb/abstract/?lang=pt>. Acesso em: 26 jul. 2021.

STEEN, R. Grant. Retractions in the scientific literature: is the incidence of research fraud increasing?. **Journal of Medical Ethics**, Londres, v. 37, n. 4, p. 249-253, 2011. DOI: <http://www.doi.org/10.1136/jme.2010.040923>. Disponível em: <https://jme.bmj.com/content/37/4/249>. Acesso em: 01 fev. 2021.

STEEN, R. Grant; CASADEVALL, Arturo; FANG, Ferric C. Why has the number of scientific retractions increased?. **PLoS ONE**, São Francisco, v. 8, n. 7, p. e68397, 2013. DOI: <http://doi.org/10.1371/journal.pone.0068397>. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0068397>. Acesso em: 01 fev. 2021.

TURNITIN. **iThenticate**. Oakland: Turnitin, c2021. Disponível em: <https://www.ithenticate.com/>. Acesso em: 26 jul. 2021.

UK EQUATOR NETWORK CENTRE. **Equator Network**. Oxford: The EQUATOR Network, c2021. Disponível em: <https://www.equator-network.org/>. Acesso em: 31 jan. 2021.

VILAÇA, Murilo Mariano. Más condutas científicas uma abordagem crítico-comparativa para *in-formar* uma reflexão sobre o tema. **Revista Brasileira de Educação**, Rio de Janeiro, v. 20, n. 60, p. 245-269, 2015. DOI: <https://doi.org/10.1590/S1413-24782015206012>. Disponível em: <https://www.scielo.br/j/rbedu/a/CkvpTffVyh7SS9LKPLZQ4gJ/abstract/?lang=pt>. Acesso em: 26 jul. 2021.

WATANABE, Edson H. A não linearidade entre a reação de quem copia e de quem é copiado. **Estudos Avançados**, v. 28, n. 80, p. 199-212, 2014. DOI: <https://doi.org/10.1590/S0103-40142014000100017>. Disponível em: <https://www.scielo.br/j/ea/a/nK8hJW4sGDqpYTNHbJjv5Ns/?lang=pt>. Acesso em: 26 jul. 2021.

WILSON, Brent; HARRIS, Christine; WIXTED, John. Science is not a signal detection problem. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, DC, v. 117, n. 11, p. 5.559-5.567, 2020. DOI: <https://doi.org/10.1073/pnas.1914237117>. Disponível em: <https://www.pnas.org/content/117/11/5559>. Acesso em: 31 jan. 2021.

WORLD ASSOCIATION OF MEDICAL EDITORS. **Recommendations on publication ethics policies for medical journals**. Bellagio: WAME, c2016. Disponível em: <https://wame.org/recommendations-on-publication-ethics-policies-for-medical-journals>. Acesso em: 26 jul. 2021.

YAHAGHI, Hossein; SOROOSHIAN, Shahryar; YAHAGHI, Javad. Unethical postgraduate supervision. **Science and Engineering Ethics**, [s. l.], v. 23, p. 629-630, 2017. DOI: <https://doi.org/10.1007/s11948-016-9789-7>. Disponível em: <https://link.springer.com/article/10.1007/s11948-016-9789-7>. Acesso em: 01 fev. 2021.