

USO RESPONSÁVEL DE MÉTRICAS E INDICADORES

Solange Maria dos Santos¹

 <https://orcid.org/0000-0002-5067-6362>

1 INTRODUÇÃO

Raramente mudanças acontecem da noite para o dia e menos ainda mudanças repentinas, profundas e fundamentais na maneira como conduzimos nossas vidas. A pandemia do coronavírus (COVID-19) parece ser um desses casos incomuns que além de ter um impacto devastador em todos nós também deixou claro que a ciência tem um papel fundamental para a sobrevivência da espécie humana.

Beneficiado-se de práticas da ciência aberta e da infraestrutura para apoiar a colaboração, reutilização e o compartilhamento de dados. O mundo da pesquisa mostrou que é possível avançar na descoberta de medicamentos e vacinas em velocidades jamais (Cross, 2020; Lurie et al., 2020; Meng, et al., 2020). A pandemia do COVID-19 impactou não somente aquilo que já sabíamos como ampliou nossa consciência sobre a importância do Acesso Aberto, da Ciência Aberta, dos dados abertos e da colaboração distribuída ante a necessidade de uma resposta coletiva.

Em 31 de janeiro de 2020, o *Wellcome Trust* chamou o coronavírus de “ameaça significativa e urgente à saúde global” e apelou a “pesquisadores, periódicos e financiadores para garantir que os resultados de pesquisa e os dados relevantes para este surto fossem compartilhados rápida e abertamente para informar a resposta à saúde pública e ajudar a salvar vidas.” (Carr, D., 2020). Os signatários conjuntos desta declaração (incluindo as principais editoras como: *Elsevier*, *Springer Nature* e *Taylor & Francis*, além de vários financiadores e sociedades científicas) se comprometeram a abrir imediatamente todas as pesquisas e dados sobre o surto: em Repositórios de *Preprints*² para os artigos ainda não revisados pelos pares e em plataformas de periódicos para os artigos que já sido submetidos à avaliação. Os preprints e outros tipos de compartilhamentos online tornaram-se norma quando o atraso de algumas semanas na publicação pode significar vidas perdidas.

A crise sanitária e socioeconômica de proporção global demonstrou como é possível mudar rapidamente culturas e mentalidades previamente arraigadas. Estamos vendo a colaboração global em escala sem precedentes entre países e disciplinas, pois milhares de especialistas se concentram prioritariamente em um único problema. O sigilo e acumulação habituais de dados que podem levar a subsídios e promoções foram colocados em segundo

¹ Programa SciELO-FAPESP e CiMetrias - Grupo de Pesquisa em Métricas da Ciência e Tecnologia da ECA-Universidade de São Paulo.

² Vimos um aumento extraordinário no uso de servidores de preprints como *bioRxiv*, *medRxiv* e outros, utilizados para compartilhar rapidamente novas ideias sobre o coronavírus. Durante o surto, os servidores preprints foram muito utilizados como uma opção para compartilhar rapidamente os resultados de novas pesquisas antes de passar por uma revisão por pares (que geralmente leva meses até a publicação formal num periódico) possibilitando que outros pesquisadores da área avaliem rapidamente os resultados.

plano dada a urgência do momento. Algumas das falhas fundamentais das tradicionais práticas fechadas de pesquisa estão sendo reveladas e questionadas. Muitos dos usuais imperativos do sistema de crédito acadêmico disfuncional estão sendo deixados de lado à medida que os resultados e os dados de pesquisa são compartilhados imediatamente como defende o movimento de Ciência Aberta³.

Um relatório publicado recentemente pela European Commission destaca que um dos obstáculos mais significativos para os objetivos da ciência aberta está nas estruturas de incentivo da pesquisa acadêmica, que muitas vezes deixam de reconhecer, valorizar e recompensar os esforços para abrir o processo científico (European Commission, 2017). Como resultado dessa falta de reconhecimento, o avanço na carreira dos pesquisadores pode ser dificultado se eles adotarem novas maneiras de trabalhar e publicar sem que sejam feitos os ajustes nos sistemas de avaliação existentes.

As métricas desempenham um papel importante em qualquer sistema de pesquisa, incluindo um aberto, mas é crucial entender também os limites inerentes às métricas. Nos últimos anos, o debate sobre o uso de métricas na avaliação da pesquisa se intensificou. Alguns vêem o uso de indicadores como uma contribuição para o desenvolvimento de uma cultura excessivamente gerencial e orientada por auditoria nas universidades (Collini, 2016; Martin, 2016, European Commission, 2017). As preocupações tendem a se concentrar em três questões. Primeiro, um estreitamento da atenção gerencial ou de financiadores para aspectos que podem ser medidos, em detrimento de aspectos que não podem ser medidos facilmente. Segundo, leva a uma redução na diversidade, com ênfase em indicadores ou tabelas de classificação específicas, que levam as instituições a adotarem prioridades estratégicas semelhantes e aos pesquisadores individuais a se concentrarem no trabalho incremental de menor risco, destinado a periódicos de maior impacto (Hicks et al., 2015; European Commission, 2017). E o terceiro, se deve a uma distorção dos incentivos, que por sua vez agrava os problemas de qualidade, integridade e reprodutibilidade da pesquisa (Benedictus, Miedema e Ferguson, 2016; Sarewitz, 2016; European Commission, 2017).

Todas estas são preocupações apresentadas pertinentes, no entanto, a alternativa não é de que não haja avaliação. À medida que a avaliação da pesquisa se torna mais importante e impacta significativamente a carreira dos pesquisadores, suas deficiências devem ser discutidas para possibilitar a proposição de novas formas de avaliação que possibilitem mensurar e recompensar adequadamente a relevância acadêmica, social, cultural e econômica das pesquisas.

³ Esse modo de fazer ciência aberta engloba muito mais do que a ciência no âmbito da academia, das instituições de pesquisa e do universo dos periódicos que a dissemina. Trata-se da ciência de e para a comunidade acadêmica, o setor produtivo, as políticas públicas e a sociedade, instâncias que passaram a participar abertamente do debate sobre os aspectos da produção, avaliação, gestão e disseminação da pesquisa e inovação científica (SANTOS e CALÓ, 2020).

2 AVALIAÇÃO DA PESQUISA

À medida que a pesquisa se institucionalizava nas universidades, a avaliação se tornou um instrumento cada vez mais complexo, projetado para selecionar pesquisadores, avaliar instituições ou conceder subsídios para projetos de pesquisa. Sem dúvida, um elemento chave que transformou o processo de avaliação acadêmica foi a publicação centrada na indexação dos periódicos. A relevância que a avaliação das publicações adquiriu no mundo acadêmico durante o século XX e até agora no século 21, deu origem ao surgimento de uma disciplina, a bibliometria, que se tornou cada vez mais central nos processos de avaliação. Com o tempo, foram produzidas distorções, usos e abusos de métricas, que foram observadas em um campo de estudos que examina a prática da avaliação científica (Gingras, 2016).

Um número crescente de líderes científicos acredita que o atual sistema de incentivos e recompensas dos pesquisadores não está alinhado com as necessidades da sociedade (Moher, et al., 2018). Os regimes de avaliação acadêmica estabelecidos para quantificar a qualidade da pesquisa, acadêmicos e instituições têm sido amplamente criticados por causa dos efeitos prejudiciais que exercem sobre os ambientes acadêmicos e sobre a própria produção de conhecimento.

No entanto, sabe-se que a avaliação dos pesquisadores é um ponto central das decisões sobre sua contratação, promoção e mandato. Elaborar, escrever, apresentar, avaliar, priorizar e selecionar o curriculum é um processo muitas vezes demorado para candidatos a concessão de financiamentos, professores e comitês de avaliação (Moher et al., 2018). As instituições precisam tomar decisões dentro de determinados limites de tempo e orçamento. Muitos esforços de avaliação avaliam principalmente o que é facilmente identificável, como o número e o montante de concessões de financiamento e o número e citações de artigos publicados. Segundo Moher e colaboradores (2018) mesmo para aspectos prontamente mensuráveis, os critérios usados para avaliação e decisões variam entre ambientes e instituições e não são necessariamente aplicados de forma consistente, mesmo dentro da mesma instituição.

Soma-se a isso o fato de que embora exista farta literatura sobre os problemas e as alternativas ao fator de impacto do periódico para avaliar o impacto da citação (Hammarfelt, 2017). Muitas instituições ainda o usam como indicador na avaliação do corpo docente como proxy qualidade dos trabalhos publicados, ou mesmo para determinar recompensas monetárias (Quan, Chen e Shu, 2017).

A necessidade de modificar os critérios de avaliação de pesquisas e publicações tem sido há anos uma demanda da comunidade científica que acabou culminando numa variedade de pronunciamentos públicos muito críticos em relação ao monopólio exercido pelo fator de impacto para avaliar as publicações. Estes posicionamentos sugerem, entre outras medidas, incorporar pontos de vista qualitativos, avaliação no nível do artigo (não centrada no fator de impacto do periódico) e também ampliação do espectro de indicadores

a serem incorporadas nas avaliações e a recomendação do uso de métricas responsáveis robustas, transparentes, diversas e reflexivas.

3 POSICIONAMENTOS EM PROL DO USO RESPONSÁVEL DE INDICADORES NA AVALIAÇÃO DA PESQUISA

Nos últimos anos houve um interesse crescente da comunidade científica em discutir e recomendar o uso responsável de indicadores na avaliação. Como parte dos esforços desse movimento, uma série de documentos com posicionamento de alto nível foram publicados principalmente nos Estados Unidos e Europa.

Alguns desses posicionamento foram feitos por grandes grupos (por exemplo, o Manifesto de Leiden), outros esforços foram de grupos menores ou individuais (por exemplo, Ioannidis e Khoury, 2014); ou iniciativas de periódicos (por exemplo, Nature, 2016).

O aparecimento cada vez mais frequente de manifestos e declarações como: Declaração de San Francisco para Avaliação da Pesquisa (DORA), Manifesto de Leiden, o relatório The Metric Tide ou, mais recentemente, os Princípios de Hong Kong expuseram as várias distorções que o sistema de publicação e avaliação haviam gerado na vocação universalista da ciência e destacaram a necessidade de buscar as melhores práticas e reduzir os abusos no emprego de métricas e indicadores.

3.1 Declaração sobre Avaliação da Pesquisa (DORA)

A Declaração de San Francisco para Avaliação da Pesquisa (DORA), foi desenvolvida por um grupo de editores e representantes de editoras de revistas acadêmicas em 16 de dezembro 2012 durante a Reunião Anual da American Society for Cell Biology (ASCB) em San Francisco e desde então se tornou uma iniciativa mundial que abrange todas as disciplinas. O grupo elaborou um conjunto de recomendações e a declaração inclui princípios fundamentais para financiadores, editores, sociedades profissionais, instituições e pesquisadores.

Embora contenha várias recomendações o principal alvo da DORA é o uso abusivo do Fator de impacto. O Fator de Impacto⁴ é frequentemente usado como o parâmetro na comparação da produção científica de indivíduos e instituições. Essa métrica originalmente concebida como uma ferramenta para auxiliar bibliotecários no processo de aquisição de periódicos mais importantes para melhor atender suas comunidades dentro de uma

⁴ As deficiências do Fator de Impacto enquanto ferramenta para a avaliação da pesquisa são amplamente conhecidas e muito bem documentadas na literatura: a) algumas distribuições de citações dentro de cada periódico podem ser muito desiguais [1-3]; B) as propriedades do Fator de Impacto são dependentes do campo: trata-se de um conjunto de muitos tipos de artigos, muito diferentes, incluindo artigos resultado pesquisa primária e resenhas, editoriais [1-4]; c) Fatores de Impacto podem ser distorcidos por políticas e práticas editoriais questionáveis [5]; e d) os dados usados para calcular o Fator de Impacto não são transparentes nem abertamente acessíveis ao público [4,6,7], dentre outros tipos de disfunção.

disciplina específica, mas que tem não qualquer utilidade como medida da qualidade científica das pesquisas publicadas. Logo, é conceitualmente errado utilizar o Fator de Impacto para avaliar a importância de um artigo para fins contratação, promoção, avaliação de pesquisadores individuais ou grupos de pesquisa de pesquisa.

Nesse sentido, a DORA defende claramente que qualquer outro uso do fator de impacto é mau uso.

Quadro 1: Recomendações da Declaração DORA a Agências de fomento,

RECOMENDAÇÕES DA DECLARAÇÃO DORA	
Geral	1. Não usar métricas baseadas em revistas, como Fatores de Impacto, como medida substitutiva da qualidade de artigos de pesquisa individuais, para avaliar as contribuições de um cientista individual, ou em decisões relativas a contratações, promoções ou financiamentos.
Para Agências de fomento	<p>2. Explicitar os critérios usados na avaliação da produtividade de solicitantes de bolsas e enfatizar claramente, em especial para pesquisadores em início de carreira, que o conteúdo científico de um artigo é muito mais importante que a métrica ou a identidade da revista em que é publicado.</p> <p>3. Para fins de avaliação da pesquisa, considerar o valor e o impacto de todos os resultados (inclusive conjuntos de dados e software) além das publicações, e considerar um amplo espectro de medidas de impacto, inclusive indicadores qualitativos de impacto da pesquisa, tais como as influências em políticas e práticas. Para instituições</p> <p>4. Explicitar os critérios usados nas decisões de contratação, concessão de estabilidade (tenure), e promoção, enfatizando claramente, em especial para pesquisadores novatos, que o conteúdo científico de um artigo é muito mais importante que a métrica ou a identidade da revista em que é publicado.</p> <p>5. Para fins de avaliação da pesquisa, considerar o valor e o impacto de todos os resultados (inclusive conjuntos de dados e software) além das publicações, e considerar um amplo espectro de medidas de impacto, inclusive indicadores qualitativos de impacto da pesquisa, tais como as influências em políticas e práticas.</p>
Para Editoras	<p>6. Reduzir drasticamente a ênfase no Fator de Impacto como recurso promocional, idealmente deixando de promover o Fator de Impacto, ou apresentando-o no contexto de uma diversidade de métricas referentes a revistas (p. ex., fator de impacto quinquenal, EigenFactor, SCImago], índice-h, prazos editoriais e de publicação, etc.) que fornecem uma visão mais rica do desempenho da revista.</p> <p>7. Disponibilizar uma variedade de métricas referentes a artigos para encorajar a mudança em direção a avaliações baseadas no conteúdo científico de cada artigo, em vez de na métrica da revista em que é publicado.</p> <p>8. Encorajar práticas responsáveis de autoria e o fornecimento de informações sobre as contribuições específicas de cada autor.</p> <p>9. Quer uma revista seja de acesso aberto, quer baseada em assinaturas, remover todas as limitações sobre o reuso em listas de referências em artigos de pesquisa e disponibilizá-las de acordo com a Creative Commons Public Domain Dedication.</p> <p>10. Remover ou reduzir as restrições ao número de referências em artigos de pesquisa e, onde for adequado, tornar obrigatória a citação da literatura primária no lugar de resenhas para dar crédito ao grupo, ou grupos, que primeiro relataram uma descoberta.</p>
Para organizações que fornecem métricas	11. Ser aberto e transparente deixando claros os dados e métodos usados para calcular todas as métricas.

	<p>12. Fornecer os dados sob uma licença que permita reuso irrestrito, e acesso computacional aos dados, quando possível.</p> <p>13. Deixar claro que a manipulação inadequada de métricas não será tolerada; explicitar os critérios do que constitui manipulação inadequada, e que medidas serão tomadas para combatê-la.</p> <p>14. Levar em conta a diversidade nos tipos de artigo (p. ex., resenhas versus artigos de pesquisa), e nas áreas do conhecimento quando métricas são usadas, agregadas ou comparadas.</p>
Para os pesquisadores	<p>15. Ao participar de comissões deliberativas sobre o financiamento, contratação, concessão de estabilidade (tenure) ou promoção, fazer avaliações baseadas no conteúdo científico, em vez de em métricas de publicação.</p> <p>16. Quando apropriado, citar a literatura primária em que observações são relatadas, em vez de resenhas, a fim de dar crédito a quem o merece.</p> <p>17. Usar um espectro de métricas de artigo e indicadores em currículos, como evidência do impacto de artigos individuais publicados e outros resultados da pesquisa.</p> <p>18. Questionar práticas de avaliação da pesquisa baseadas impropriamente em Fatores de Impacto, promover e ensinar a prática mais adequada, centrada no valor e influência de resultados de pesquisa específicos.</p>

Fonte: Adaptado de DORA (2012)

A DORA também oferece a possibilidade de "assiná-la" - transformando-a em uma possível declaração pública que confirma que um indivíduo ou instituição está comprometida com os princípios nela contidos. Em Julho de 2020, a DORA conta com mais de 1900 intuições signatárias e mais de 16.000 assinaturas individuais.

Após a DORA, outros manifestos e políticas internacionais publicadas seguiram direção semelhante.

3.2 **Manifesto de Leiden**

O Manifesto de Leiden para Métricas de Pesquisa foi publicado pela primeira vez na forma de comentário no periódico Nature em 2015. Cinco especialistas liderados por Diana Hicks, professora da Escola de Políticas Públicas do *Georgia Institute of Technology*, e Paul Wouters, diretor do *Centre for Science and Technology Studies (CWTS)* da Universidade de Leiden que reuniram o pensamento coletivo dos participantes da conferência onde a iniciativa se originou⁵.

Os dez princípios propostos no manifesto não são novidade para os pesquisadores das áreas de bibliometria e cientometria, mas esses especialistas não estão presentes quando os avaliadores se reportam aos gestores universitários que também não são especialistas no tema. Além disso, a literatura necessária para contestar ou questionar o

⁵ Science and Technology indicators conference (STI) realizada de 3-5 setembro de 2014 em Leiden na Holanda pelo Centre for Science and Technology Studies (CWTS) em colaboração com a European Network of Indicator Developers (ENID) <http://sti2014.cwts.nl/Home>.

uso inadequado de indicadores nas avaliações encontra-se normalmente publicada em periódicos especializados e nem sempre de modo acessível.

Desse modo, Manifesto de Leiden sintetiza algumas das melhores práticas de avaliação da pesquisa baseada em métricas, para que os pesquisadores avaliados possam confiar em seus avaliadores, e para que os avaliadores possam confiar em bem empregar os indicadores (Hicks et al., 2015)

Quadro 2: 10 princípios propostos pelo Manifesto de Leiden

PRINCÍPIOS DO MANIFESTO DE LEIDEN	
1	<p>A avaliação quantitativa deve apoiar especialistas qualitativos avaliação.</p> <p>Os indicadores quantitativos podem corrigir tendências enviesadas da avaliação por pares e facilitar a deliberação. Nesse sentido, devem fortalecer a revisão por pares já que emitir julgamentos sobre colegas é difícil sem apoio de uma série de informações relevantes. No entanto, os avaliadores não devem ceder à tentação de basear suas decisões apenas em números. Os indicadores não devem substituir o juízo informado. Os tomadores de decisão têm plena responsabilidade por suas avaliações.</p>
2	<p>Medir o desempenho em relação às missões de pesquisa da instituição, grupo ou pesquisador.</p> <p>Os objetivos de um programa de pesquisa devem ser indicados no início, e os indicadores utilizados para avaliar o desempenho devem estar claramente vinculados aos objetivos. A escolha dos indicadores e de como eles são utilizados deve levar em conta o contexto socioeconômicos e culturais mais amplo. Os cientistas têm diversas missões de pesquisa. A pesquisa que avança as fronteiras do conhecimento acadêmico difere da pesquisa que é focada em proporcionar soluções para os problemas da sociedade. A avaliação pode ser baseada em méritos relevantes para as políticas públicas, para a indústria ou para os cidadãos em geral, em vez de méritos baseados em noções acadêmicas de excelência. Não existe um modelo único de avaliação que se aplique a todos os contextos.</p>
3	<p>Proteger a excelência em pesquisas relevantes localmente.</p> <p>Em muitas partes do mundo, a excelência da pesquisa é associada à publicação no idioma inglês. O Fator de Impacto é calculado na Web of Science, que indexa principalmente os periódicos com base nos Estados Unidos e em língua inglesa. Este viés é particularmente problemático para as Ciências Sociais e Humanidades, áreas mais orientadas para a pesquisa de temas regionais e nacionais. Muitas outras áreas possuem uma dimensão nacional ou regional.</p> <p>O pluralismo e a relevância para a sociedade tendem a ser anulado quando se produzem artigos de interesse para os defensores do alto impacto: os periódicos em inglês. Os indicadores baseados nos periódicos de qualidade publicados em outros idiomas que não o inglês deve identificar e privilegiar as áreas de pesquisa de interesse local.</p>
4	<p>Manter a coleta de dados e os processos analíticos abertos, transparente e simples.</p> <p>A construção das bases de dados necessárias para a avaliação deve observar regras claramente definidas e fixadas antes da conclusão da pesquisa. Esta era a prática comum entre os grupos acadêmicos e comerciais que desenvolveram metodologias de avaliação bibliométrica ao longo de muitas décadas. Tais grupos referenciaram protocolos publicados na literatura revisada por pares. Esta transparência possibilitou o escrutínio das metodologias. Por exemplo, em 2010, o debate público sobre as propriedades técnicas de um importante indicador utilizado por um dos nossos grupos (o Centro de Estudos de Ciência e Tecnologia - Centre for Science and Technology Studies, CWTS, da Universidade de Leiden, na Holanda) levou a uma revisão no cálculo deste indicador (6). Os novos operadores do setor privado devem seguir os mesmos padrões; ninguém deve aceitar avaliações saídas de uma caixa-preta.</p>
5	<p>Permitir que os avaliados verifiquem os dados e as análises.</p> <p>Para garantir a qualidade dos dados, todos os pesquisadores incluídos em estudos bibliométricos deveriam poder verificar se suas produções foram corretamente identificadas. Todos os que dirigem e administram os processos de avaliação devem assegurar a precisão dos dados, através de verificação própria ou auditoria de terceiros. As universidades poderiam implementar esse princípio em seus sistemas de informação sobre a pesquisa, o que deveria ser um princípio norteador na seleção de fornecedores desses sistemas. A coleta e processamento de dados precisos e de alta qualidade demandam tempo e dinheiro e devem ser considerados no orçamento institucional.</p>
6	<p>Considerar as diferenças entre áreas nas práticas de publicação e citação.</p> <p>A melhor prática de avaliação é selecionar um conjunto de possíveis indicadores e permitir que as distintas áreas escolham aqueles que lhes são mais adequados. Há alguns anos, um grupo europeu de historiadores recebeu uma classificação relativamente baixa em uma avaliação nacional por pares, porque escreviam livros em vez de artigos em revistas indexadas na WoS. Estes historiadores tiveram o azar de fazer parte de um departamento de Psicologia. Historiadores e cientistas sociais precisam que os livros e</p>

	<p>a literatura publicada no idioma nacional sejam incluídos na contagem de publicações; já os cientistas da computação esperam que seus trabalhos apresentados em eventos e conferências sejam levados em conta.</p> <p>Os valores de citações variam por área: os periódicos melhor avaliados em Matemática têm Fator de Impacto por volta de 3; já as revistas melhor avaliadas em Biologia Celular tem Fator de Impacto em torno de 30. Portanto, é necessário o uso de indicadores normalizados, e o método de normalização mais confiável é baseado em percentuais: cada artigo é ponderado segundo o percentual a que pertence na distribuição de citações em sua área (os melhores 1%, 10% ou 20%, por exemplo). Uma única publicação altamente citada melhora ligeiramente a posição de uma universidade em um ranking baseado em indicadores percentuais, mas pode impulsionar a universidade de uma posição mediana para as primeiras posições em um ranking baseado em médias de citação (7).</p>
7	<p>Basear a avaliação de pesquisadores individuais no juízo qualitativo da sua carreira.</p> <p>Quanto mais idade você tem, maior será o seu Índice h, mesmo que não publique novos artigos. O Índice h varia por área: os pesquisadores das Ciências da Vida chegam ao topo com 200; os físicos com 100 e cientistas sociais com 20 a 30 (8). Depende da base de dados: há pesquisadores em Ciência da Computação que têm um Índice h de cerca de 10 na WoS, mas de 20 a 30 no Google Scholar (9). Ler e julgar o trabalho de um pesquisador é muito mais adequado do que depender de um número. Mesmo quando se compara um grande número de pesquisadores, uma abordagem que considere informações diversas sobre o conhecimento, experiência, atividades e influência de cada indivíduo é a melhor.</p>
8	<p>Evite objetividade mal colocada e falsa precisão.</p> <p>Indicadores de ciência e tecnologia são propensos à ambiguidade conceitual e à incerteza, e demandam fortes suposições que não são universalmente aceitas. O significado das contagens de citações, por exemplo, tem sido amplamente discutido. Assim, a melhor prática de avaliação utiliza indicadores múltiplos para fornecer uma imagem mais robusta e plural da pesquisa. Se as incertezas e os erros podem ser quantificados, esta informação deve acompanhar os valores dos indicadores publicados, usando barras de erro, por exemplo. Se isso não for possível, os produtores de indicadores deveriam, pelo menos, evitar a falsa precisão. Por exemplo, o Fator de Impacto de revistas é publicado com três casas decimais para evitar empates. No entanto, dada a ambiguidade conceitual e a variabilidade aleatória das contagens de citações, não faz sentido distinguir as revistas com base em diferenças mínimas do Fator de Impacto. Evite a falsa precisão: apenas uma casa decimal se justifica.</p>
9	<p>Reconhecer os efeitos sistêmicos da avaliação e dos indicadores.</p> <p>Os indicadores mudam o sistema da pesquisa por meio dos incentivos que estabelecem. Estes efeitos devem ser previstos. Isto significa que um conjunto de indicadores é sempre preferível - um único indicador convida a burlas ou a desvios do objetivo (em que a medida se torna um fim em si). Por exemplo, na década de 1990, a Austrália financiou a pesquisa universitária através de uma fórmula baseada sobretudo no número de artigos publicados pelas instituições. As universidades poderiam calcular o "valor" de um artigo em uma revista revisada por pares; em 2000, o valor era de estimados AUS\$ 800 (em torno de US\$ 480) para o financiamento da pesquisa. Previsivelmente, o número de artigos publicados por pesquisadores australianos subiu, mas em revistas menos citadas, sugerindo uma queda na qualidade dos artigos (10).</p>
10	<p>Examinar e atualizar os indicadores regularmente.</p> <p>A missão da pesquisa e os objetivos da avaliação mudam, e o próprio sistema de pesquisa evolui junto. Medidas que anteriormente eram úteis se tornam inadequadas e surgem novos indicadores. Os sistemas de indicadores têm de ser revistos e talvez modificados. Percebendo os efeitos de sua fórmula simplista, em 2010 a Austrália introduziu a iniciativa "Excelência na Pesquisa para a Austrália" (Excellence in Research for Australia), mais complexa e com ênfase na qualidade.</p>

Fonte: Adaptado de HICKS, D. *et al.*, (2015)

Os pontos 6 a 10 do Manifesto de Leiden fornecem exemplos relevantes sobre a variação das métricas chamando atenção para o fato de que métricas flutuam e isso é normal.

3.3 The Metric Tide

O relatório "*The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*", publicado em julho de 2015, foi coordenado pelo professor James Wilsdon, professor de ciência e democracia da Universidade de Sussex, apoiado por um grupo independente e multidisciplinar de especialistas em cientometria, financiamento de pesquisa, política de pesquisa, publicação, gestão universitária e administração de pesquisa.

O relatório explora o uso de métricas em diferentes disciplinas e avalia sua potencial contribuição para o desenvolvimento da excelência, da avaliação, e do impacto da pesquisa. Também analisa a aplicabilidade e os efeitos negativos ou não intencionais das métricas em vários aspectos da cultura de pesquisa incluindo gerenciamento por desempenho, igualdade, diversidade e interdisciplinaridade.

O grupo de trabalho responsável pela elaboração do *The Metric Tide* identificou 20 recomendações que demandam atenção e ação das partes interessadas em todo o sistema de pesquisa do Reino Unido, mas que podem ser de muita utilidade para outros sistemas de nacionais de pesquisa. O propõem ações nas seguintes áreas: apoiar a liderança eficaz, governança e gestão das culturas de pesquisa; melhorar a infraestrutura de dados que suporta o gerenciamento de informações de pesquisa; aumentar a utilidade dos dados e fontes de informação existentes; usando métricas na próxima *Research Excellence Framework* (REF)⁶; e coordenando atividades e construindo evidências.

Estas recomendações são sustentadas pela noção de métricas responsáveis como uma maneira de estruturar usos apropriados de indicadores quantitativos na governança, gestão e avaliação da pesquisa.

Quadro 3: Recomendações do relatório independente *The Metric Tide*

THE METRIC TIDE – RECOMENDAÇÕES		
OBJETIVOS	RECOMENDAÇÕES	DESTINATÁRIOS
Apoiar a liderança, governança e gerenciamento eficazes das culturas de pesquisa	1. A comunidade de pesquisa deve desenvolver uma abordagem mais sofisticada e diferenciada da contribuição e das limitações dos indicadores quantitativos. É necessário cuidado com a linguagem e a terminologia adotada. O termo "métricas" não é o mais adequado; o termo preferencial é "indicadores" que reflete o reconhecimento de que os dados podem não ter relevância específica, mesmo que sejam úteis num contexto geral.	Instituições de Ensino Superior, Financiadores, Gestores, pesquisadores
	2. No nível institucional, os líderes das instituições de ensino superior devem desenvolver uma declaração clara de princípios sobre sua abordagem a gestão e avaliação da pesquisa, incluindo o papel dos indicadores quantitativos. Com base nesses princípios, eles devem selecionar cuidadosamente	Dirigentes de instituições, dirigentes de pesquisa, gestores de IES

⁶ O REF é o sistema do Reino Unido para avaliar a qualidade da pesquisa nas instituições de ensino superior. O REF é realizado pelos quatro organismos de financiamento do ensino superior do Reino Unido: Research England, the Scottish Funding Council (SFC), o Higher Education Funding Council for Wales (HEFCW), e o Department for the Economy, Northern Ireland (DfE).

	<p>indicadores quantitativos adequados aos seus objetivos e contexto institucionais. Nos casos onde as instituições usam tabelas de classificação, elas devem explicar por que as estão usando como um meio para alcançar fins específicos. Sempre que possível, indicadores alternativos que apoiam a igualdade e a diversidade devem ser identificados e incluídos. A comunicação clara da justificativa para a seleção de indicadores específicos e como eles serão usados como uma ferramenta de gerenciamento é fundamental. Como parte desse processo, as instituições de ensino superior devem considerar assinar a DORA ou basear-se em seus princípios e adaptá-los aos seus contextos institucionais.</p>	
	<p>3. Os gestores e administradores de pesquisa devem defender esses princípios e o uso de métricas responsáveis em suas instituições. Eles devem prestar a devida atenção às implicações de igualdade e diversidade das opções de avaliação da pesquisa; envolver-se com especialistas externos, ajudar a facilitar uma infraestrutura de dados mais aberta e transparente; defesa do uso de identificadores exclusivos, como ORCID iDs; trabalhar com financiadores e editores na interoperabilidade de dados; explorar indicadores para aspectos da pesquisa que eles desejam avaliar, em vez de usar os indicadores existentes apenas por estarem prontamente disponíveis; aconselhar os líderes seniores em métricas significativas para o seu contexto institucional ou departamental; e compartilhar boas práticas por órgãos setoriais.</p>	<p>Gestores, Administradores de pesquisa, Associações de Gestores e administradores de pesquisa</p>
	<p>4. Os gerentes de RH e os painéis de recrutamento ou promoção nas instituições de ensino superior devem ser explícitos sobre os critérios usados nas decisões de nomeação e promoção acadêmica. Esses critérios devem ser fundamentados em especialistas julgamento e pode refletir tanto a qualidade acadêmica dos resultados quanto as contribuições mais amplas para políticas, indústria ou sociedade. Às vezes, os julgamentos podem ser utilmente guiados por métricas, se relevantes para o critérios em questão e utilizados de forma responsável; as métricas de citação no nível do artigo, por exemplo, podem ser indicadores úteis de impacto acadêmico, desde que sejam interpretadas à luz de normas disciplinares e levando em consideração suas limitações. Métricas no nível do periódico, como o FI, não devem ser usadas.</p>	<p>Gerentes de RH, painéis de recrutamento e promoção</p>
	<p>5. Pesquisadores individuais devem estar atentos às limitações de indicadores específicos na maneira como apresentam seus próprios currículos e avaliam o trabalho dos colegas. Quando os indicadores padrão são inadequados, os pesquisadores individuais devem procurar uma variedade de fontes de dados para documentar e apoiar reivindicações sobre o impacto do seu trabalho.</p>	<p>Todos os pesquisadores</p>
	<p>6. Instituições como IES e financiadores de pesquisa devem desenvolver seus próprios princípios, em seus contextos específicos, sobre o uso de indicadores quantitativos na avaliação e gestão da pesquisa e devem garantir que sejam claramente comunicados, fáceis de localizar e entender. Eles devem buscar abordagens para a coleta de dados que sejam transparentes, acessíveis e permitam maior interoperabilidade em uma diversidade de plataformas.</p>	<p>Órgãos de Financiamento ensino superior, Conselhos de Pesquisa, outros financiadores de pesquisa</p>
	<p>7. Os provedores de dados, analistas e produtores de rankings universitários e tabelas de classificatórias devem se esforçar para obter maior transparência e interoperabilidade entre diferentes sistemas de medição. Alguns rankings universitários têm adotando medidas louváveis para serem mais abertos sobre a escolha de indicadores e os pesos a eles atribuídos, mas outros rankings ainda permanecem uma "na caixa preta".</p>	<p>Fornecedores de dados, analistas e produtores de rankings universitários e tabelas de classificatórias</p>
	<p>8. Os editores devem reduzir a ênfase nos fatores de impacto de periódicos como uma ferramenta de promoção e usá-los apenas no contexto de uma variedade de métricas baseadas</p>	<p>Editores</p>

	<p>em periódicos que fornecem uma visão mais rica do desempenho. Conforme sugerido pela DORA, esse conjunto mais amplo de indicadores pode incluir fator de impacto de 5 anos, EigenFactor, SCImago, tempos de edição e publicação. Os editores, com a ajuda do Committee on Publication Ethics (COPE), devem incentivar práticas de autoria responsáveis e o fornecimento de informações mais detalhadas sobre as contribuições específicas de cada autor. Os editores também devem disponibilizar uma variedade de métricas no nível do artigo para incentivar uma mudança na direção da avaliação com base na qualidade acadêmica de um artigo, e não nos Fatores de impacto.</p>	
<p>Melhorar a infraestrutura de dados que suporta o gerenciamento de informações de pesquisa</p>	<p>9. É necessário maior transparência e abertura na infraestrutura de dados de pesquisa. Um conjunto de princípios deve ser desenvolvido para tecnologias, práticas e culturas que possam apoiar o gerenciamento de informações de pesquisa abertas e confiáveis. Esses princípios devem ser adotados por financiadores, provedores de dados, administradores e pesquisadores como base para trabalhos futuros.</p>	<p>Órgãos de fomento de IES, Conselhos de Pesquisa, provedores de dados, gestores, administradores</p>
	<p>10. O sistema de pesquisa nacional deve aproveitar ao máximo do identificador ORCID como seu sistema preferido de identificadores únicos. Os IDs do ORCID devem ser obrigatórios para todos os pesquisadores. Os financiadores e as instituições de ensino superior devem usar o ORCID para solicitar subsídios, gerenciamento e plataformas de relatórios, e os benefícios do ORCID precisam ser melhor comunicados aos pesquisadores.</p>	<p>IES, Organismos de Financiamento, gestores</p>
	<p>11. Identificadores também são necessários para as instituições. O uso de identificadores persistentes para instituições deve ser mais amplamente empregado em processos internos de gestão de instituições de ensino superior e financiadores. Um componente da solução será mapear os vários sistemas de identificadores organizacionais para permitir a interoperação entre os vários sistemas existentes.</p>	<p>Órgãos de Financiamento, Instituições de Ensino Superior, IES, financiadores, editores</p>
	<p>12. Os editores devem exigir que os ORCID IDs e identificadores de instituições e referência aos financiadores durante o processo de submissão de artigos garantindo esses metadados durante todo o ciclo de vida da publicação. Isso facilitará o intercâmbio de informações sobre as atividades de pesquisa e ajudará a fornecer dados e métricas com carga mínima para pesquisadores e administradores.</p>	<p>Editores e provedores de dados</p>
	<p>13. O uso de identificadores de objetos digitais (DOIs) deve ser estendido para abranger todos os resultados da pesquisa. Os DOIs também devem ser mais amplamente adotados nos processos internos das IES e financiadores de pesquisas. Os DOIs já predominam na esfera de publicação de periódicos - eles devem ser estendidos para cobrir outros formatos para os quais não existam um sistema identificador, como capítulos de livros e conjuntos de dados.</p>	<p>Organismos Nacionais de financiamento, instituições de ensino superior</p>
	<p>14. É necessário mais investimento na infraestrutura de informações de pesquisa. Financiadores devem explorar oportunidades para investimentos estratégicos adicionais, principalmente para melhorar a interoperabilidade dos sistemas de gerenciamento de pesquisa.</p>	<p>Financiadores, IES</p>
<p>Aumentar a utilidade dos dados e fontes de informação existentes</p>	<p>15. Conselhos de Financiamento do Ensino Superior, os financiadores, as instituições de ensino superior devem explorar como aproveitar os dados mantidos nas plataformas existentes para apoiar os processos de avaliação da pesquisa e ensino superior, e vice-versa. Também é necessário um debate adicional sobre os méritos das coletas nas IES e coleta de dados em nível nacional.</p>	<p>Órgãos de finanças públicas, Departamentos de inovação, Conselhos de Pesquisa, Organismos Nacionais de Financiamento, Associações de</p>

		Gestores de pesquisa
	16 Departamentos de inovação devem identificar maneiras de vincular os dados coletados das plataformas relacionadas à pesquisa que sejam mais diretamente relacionadas aos processos dos departamentos de inovação e dos outros departamentos, especialmente em relação à previsão, exploração de horizonte e prioridades de pesquisa.	Departamentos de inovação, outros departamentos governamentais, Organismos de Financiamento do Ensino Superior
Usar métricas nos próximos ciclos avaliativos	<p>17. Para o próximo ciclo de a avaliação do <i>Research Excellence Framework (REF)</i>, são feitas algumas recomendações específicas ao <i>Higher Education Funding Council for England (HEFCE)</i> e aos outros organismos de financiamento, como segue: a) ao avaliar os resultados, recomendamos que os dados quantitativos - principalmente os resultados publicados - continuem a ter um lugar para informar os julgamentos da avaliação por pares sobre a qualidade da pesquisa. Essa abordagem foi usada com sucesso no REF2014 e recomendamos que seja continuada e aprimorada em exercícios futuros.</p> <p>b. Ao avaliar o impacto, recomendamos que o HEFCE e os organismos de financiamento Ensino Superior desenvolvam a análise dos estudos de caso de impacto do ciclo e avaliação REF 2014 para desenvolver diretrizes claras para o uso de indicadores quantitativos em estudos de caso de impacto futuro. Embora não sejam prescritivas, essas diretrizes devem fornecer dados sugeridos para evidenciar tipos específicos de impacto. Eles devem incluir padrões para a coleta de metadados para garantir que as características da pesquisa descrita sejam capturadas sistematicamente.</p> <p>c. Ao avaliar o ambiente de pesquisa, recomendamos que haja espaço para aprimorar o uso de dados quantitativos, mas esses dados precisam ser fornecidos com contexto suficiente para permitir sua interpretação. No mínimo, é precisa incluir informações sobre o tamanho total da cobertura em termos de unidades de avaliação as quais os dados se referem. Em alguns casos, a coleta de dados especificamente relacionada à equipe submetida ao exercício é preferível, ainda que possa ser mais dispendiosa. Além disso, dados sobre a estrutura e o uso de sistemas de informação digital para apoiar a pesquisa (ou pesquisa e ensino) podem ser cruciais para o desenvolvimento de excelentes ambientes de pesquisa.</p>	Organismos de Financiamento do Ensino Superior
Coordenar atividades e construir evidências	18 A comunidade de pesquisa precisa de um mecanismo para levar adiante a agenda estabelecida neste relatório. Propomos o estabelecimento de um Fórum de Métricas Responsáveis, que reuniria financiadores de pesquisa, IES e seus órgãos representativos, editores, provedores de dados e outros envolvidos para trabalhar em questões de padrões de dados, interoperabilidade, abertura e transparência. Órgãos de financiamento devem coordenar este fórum, contando com o apoio e a experiência de outros financiadores e órgãos do setor, conforme apropriado. O fórum deve se preparar para o futuro ciclo de avaliação do REF dentro de suas atribuições, mas também deve examinar mais amplamente o uso de métricas na gestão de IES e por outros financiadores.	Organismos de Financiamento do Ensino Superior, Associações de Gestores e administradores de pesquisa
	19. Os financiadores de pesquisas precisam aumentar o investimento na política da ciência da ciência. É necessário maior pesquisa e inovação nessa área, para desenvolver e aplicar insights de computação, estatística, ciências sociais e economia para entender melhor a relação entre pesquisa, suas qualidades e impactos mais amplos.	Organismos de financiamento da pesquisa
	20. Um aspecto positivo desta relatório foi o debate que o documento gerou. Como uma iniciativa legada, o grupo diretor criou um blog (www.ResponsibleMetrics.org) como	

	<p>um fórum para discussão contínua das questões levantadas por este relatório. O site celebra práticas responsáveis, mas também apresenta as más práticas quando elas ocorrerem. Os pesquisadores são incentivados a enviar tanto exemplos de melhores práticas quanto de exemplos mais flagrantes de usos inadequado de indicadores quantitativos no gerenciamento de pesquisas.</p>	<p>Comitê de Acompanhamento</p>
--	---	--

Fonte: Adaptado de WILSDON, J., et al. (2015)

Essas recomendações são sustentadas pela noção de 'métricas responsáveis' como uma maneira de estruturar usos apropriados de indicadores quantitativos na governança, gestão e avaliação da pesquisa. Métricas responsáveis podem ser entendidas em termos das seguintes dimensões:

- **Robustez:** basear as métricas nos melhores dados possíveis em termos de precisão e escopo;
- **Humildade:** reconhecer que a avaliação quantitativa deve apoiar - mas não suplantam - a avaliação qualitativa e especializada;
- **Transparência:** manter a coleta de dados e os processos analíticos abertos e transparentes, para que aqueles que estão sendo avaliados possam testar e verificar os resultados;
- **Diversidade:** Considerar a variação por campo e empregar uma série de indicadores para refletir e apoiar uma pluralidade de carreiras de pesquisadores e em todo o sistema
- **Reflexividade:** reconhecer e antecipar os efeitos potenciais e sistêmicos dos indicadores e atualizá-los sempre considerando esses efeitos;

Além dessas recomendações o *The Metric Tide* inclui também algumas conclusões importantes:

- Há um ceticismo considerável entre pesquisadores, universidades, órgãos representativos e sociedades instruídas sobre o uso mais amplo de métricas na avaliação e gestão de pesquisas.
- A revisão por pares, apesar de suas falhas, continua a comandar amplo apoio como a principal base para avaliar resultados de pesquisas, propostas e indivíduos. No entanto, uma minoria significativa está entusiasmada com o maior uso de métricas, desde que sejam tomados os devidos cuidados.
- Indicadores cuidadosamente selecionados podem complementar a tomada de decisão, mas será necessária uma "geometria variável" de julgamento de especialistas, indicadores quantitativos e medidas qualitativas que respeitem a diversidade da pesquisa.
- Existe uma preocupação legítima de que alguns indicadores possam ser mal utilizados ou usados: fatores de impacto de periódicos, classificações universitárias e citação são três exemplos importantes.
- A infraestrutura de dados que sustenta o uso de métricas e informações sobre pesquisa permanece fragmentada, com interoperabilidade insuficiente entre sistemas.
- A análise concluiu que, atualmente, nenhuma métrica pode fornecer uma substituição idêntica à revisão por pares REF.
- Na avaliação dos resultados da pesquisa no REF, atualmente não é viável avaliar os resultados ou impactos da pesquisa no REF usando apenas indicadores quantitativos.

- Na avaliação do impacto no REF, atualmente não é possível usar indicadores quantitativos no lugar de estudos de caso narrativos. No entanto, existe um escopo para aprimorar o uso de dados na avaliação de ambientes de pesquisa.

Embora as recomendações e principais conclusões do *The Metric Tide* apresentem uma contribuição importante para avançarmos na discussão e implementação de processos avaliativos centrados no uso responsável de métricas, é importante destacar que embora o relatório seja independente ele foi desenvolvido por solicitação do *Minister for Universities and* em parte para informar o *Research Excellence Framework* (REF) usado pelas universidades de todo o Reino Unido que nos últimos anos têm encabeçado um movimento pelo uso responsável de indicadores nas avaliações conduzidas dentro universidades.

3.4 Princípios de Hong kong

Nesse cenário do uso responsável métrica e indicadores, um dos mais recentes posicionamentos coletivos foi a publicação, em 2019, dos *Hong Kong principles* (HKP) que foram desenvolvidos como parte da *6th World Conference on Research Integrity* realizada em Hong Kong no período de 2 a 5 de junho de 2019, que teve como foco a integridade⁷ na pesquisa buscando valorizar pesquisadores que se comprometem com práticas robustas, rigorosas e transparentes que impulsionam a melhoria da pesquisa, garantindo que sejam devidamente reconhecidos e recompensados por adotarem praticas e comportamentos que levam a pesquisas confiáveis.

Os HKPs foram desenvolvidos considerando que sua implementação poderia auxiliar tanto na avaliação e avanço dos pesquisadores em suas carreiras, com o objetivo de fortalecer a integridade da pesquisa. Para tanto, os HKPs apresentam cinco princípios: 1) Práticas de pesquisa responsáveis; 2) relatórios transparentes; 3) ciência aberta (pesquisa aberta); 4) valorização da diversidade de tipos de pesquisa; e 5) reconhecimento da variedade de contribuições à pesquisa e à atividade acadêmica.

O quadro 2 apresenta cada princípio, assim como, a justificativa para sua consideração nos modelos de avaliação de grupos de pesquisadores ou indivíduos.

Quadro 2: Princípios Hong Kong

⁷ Embora existam alguns mecanismos para apoiar os pesquisadores a garantir a transparência em todos os estágios do projeto, execução e geração de relatórios, não há adoção generalizada dessas práticas em todas as disciplinas. Um fator que contribui para isso é que pouca ênfase colocada no rigor da pesquisa ao contratar, avaliar e promover pesquisadores (Moher, 2020).

PRINCÍPIOS DE HONG KONG

1	<p>Princípio 1: Avaliar os pesquisadores sobre práticas responsáveis desde a concepção até a publicação, incluindo o desenvolvimento da ideia de pesquisa, desenho da pesquisa, metodologia, execução e disseminação eficaz.</p> <p>Fundamentação O número de publicações, citações e o volume total de concessão bolsas ainda são frequentemente os indicadores predominantemente empregados pelas instituições de pesquisa para avaliar e recompensar seus pesquisadores. O fornecimento de bônus aos acadêmicos para publicação em determinados periódicos (isto é, pagamento por mérito) também é comum em muitas partes do mundo. Esses critérios de avaliação dizem pouco aos avaliadores sobre os pesquisadores e o rigor de seu trabalho; portanto, elas não são métricas particularmente "responsáveis". Essas métricas também podem ser indevidamente influenciadas por práticas de campo e citação e fornecem pouca informação sobre as contribuições de uma publicação (e, portanto, de um pesquisador) para pesquisa e sociedade. Outros critérios são necessários para fornecer uma visão mais ampla dos indicadores das melhores práticas: por exemplo, até que ponto um pesquisador desenvolve problemas de pesquisa com o envolvimento de membros da sociedade.</p>
2	<p>Princípio 2: Valorize os relatórios precisos e transparentes de todas as pesquisas, independentemente dos resultados.</p> <p>Fundamentação A não publicação de todas as descobertas de todos os estudos distorce seriamente a base de evidências para a tomada de decisão. Por exemplo, uma revisão sistemática de estudos de reboxetina para o tratamento da depressão descobriu que quase três quartos dos pacientes incluídos estavam em estudos não publicados. A publicação seletiva de pesquisas com resultados positivos (isto é, viés de publicação) distorce a base de evidências da ciência e foi demonstrada em uma variedade de disciplinas, incluindo economia, psicologia e pesquisa clínica e pré-clínica em saúde. Além disso, a frequência de outros vieses de relatórios (por exemplo, resultados primários alternados sem divulgação e rotação) é de cerca de 30%. Isso é inaceitavelmente alto e diminui a confiabilidade e a integridade da pesquisa. Parece também que os Comitês de Avaliação e Promoção geralmente não dão importância suficiente ao registro de protocolos e planos de análise de dados, publicação completa de estudos concluídos ou disponibilização de dados, códigos e materiais.</p>
3	<p>Princípio 3: Valorizar as práticas da ciência aberta (pesquisa aberta) - como métodos, materiais e dados abertos.</p> <p>Fundamentação A abertura na pesquisa é mais do que apenas o acesso à pesquisa - ela traz igualdade ao processo de pesquisa. Ela abrange uma série de práticas em todo o ciclo de vida da pesquisa. O acesso à pesquisa não deve ser sobre quem tem os recursos para pagar acesso às pesquisas (paywall), normalmente os periódicos de assinatura. As decisões em saúde e políticas sociais devem ser tomadas com base no acesso a todo conhecimento de pesquisa e não apenas a uma parte dele. Uma quantidade considerável de fundos públicos é usada para pesquisa e seus resultados podem ter um impacto social profundo. Os cientistas pré-clínicos estão se comprometendo a compartilhar abertamente seus cadernos de laboratório para otimizar a pesquisa, promover colaborações e reduzir a duplicação desnecessária. Em um esforço para dissuadir práticas questionáveis de autoria, os Consórcios que avançam os padrões em informações de administração de pesquisa apóiam a taxonomia CRediT como uma maneira de os autores da pesquisa descreverem mais abertamente como cada pessoa contribuiu para um projeto de pesquisa.</p> <p>O compartilhamento de dados é outro exemplo de abertura, mas não é uma prática comum em pesquisas clínicas (com algumas exceções, como a genética), embora os pacientes pareçam favoráveis ao compartilhamento de seus dados, pelo menos em ensaios randomizados nos quais eles participaram. O compartilhamento de dados também não é considerado padrão em muitas outras disciplinas. Sem o compartilhamento de dados, é difícil verificar a seletividade dos relatórios; o compartilhamento de dados é essencial para enfrentar a crise de reprodutibilidade e criar confiança. Existem estimativas variáveis de qual proporção de pesquisa é disponibilizada por meios de acesso aberto, como periódicos de acesso aberto, repositórios ou preprints, mas está longe de 100%.</p>
4	<p>Princípio 4: Valorizar uma ampla gama de pesquisas e bolsas de estudos, como replicação, inovação, tradução, síntese e meta-pesquisa.</p> <p>Fundamentação Um sistema busca beneficiar a sociedade e incentivar a pesquisa confiável é importante e precisa levar em consideração os diferentes tipos de pesquisa: criação de novas ideias; testando-os; replicando as principais descobertas; síntese de pesquisas existentes; desenvolvimento e validação de novas ferramentas; medidas ou métodos; etc. Diferentes indicadores e critérios precisam ser desenvolvidos e</p>

	para esses diferentes tipos e estágios de pesquisa. Isso inclui diferentes prazos de avaliação para diferentes tipos de pesquisa.
5	<p>Princípio 5: Avalie uma série de outras contribuições à pesquisa responsável e à atividade acadêmica, como revisão por pares de subsídios e publicações, orientação, divulgação e troca de conhecimentos.</p> <p>Fundamentação</p> <p>Conforme discutido ao lado do Princípio 1, as avaliações de pesquisa freqüentemente se concentram em uma gama estreita de métricas fáceis de medir, incluindo publicações, citações e receita de financiamento. Para que o ecossistema de pesquisa funcione de maneira ideal, outras atividades de pesquisa também são essenciais. A revisão por pares continua sendo a pedra angular da avaliação da qualidade de bolsas concedidas, publicações e conferências. A qualidade das contribuições da revisão por pares para os periódicos e financiadores também deve fazer parte das avaliações de promoção e posse, assim como as contribuições para várias infra-estruturas, supervisão ou regulamentos de pesquisa. Da mesma forma, as contribuições para melhorias que vão além de uma abordagem centrada no indivíduo para avaliação devem ser consideradas. Atualmente, essas atividades estão em falta nos Comitês de avaliação e Promoção. As contribuições para o desenvolvimento de carreiras de outras pessoas em todas as etapas de sua carreira são críticas, assim como as contribuições de vários comitês relacionados à pesquisa (por exemplo, assumindo o papel de editor científico). A melhor maneira de fazer isso sem criar mais barreiras e burocracia, no entanto, é debatida há muito tempo.</p> <p>Qualquer sistema de recompensa que tenha em mente toda a empresa de pesquisa e que vise promover um clima propício à pesquisa confiável e útil, com o mais alto respeito à integridade, precisa encontrar maneiras de incorporar essas funções vitais em sua estrutura geral de avaliação.</p>

Os HKPs centram-se na promoção de práticas de avaliação que buscam valorizar a integridade da pesquisa, concentrando-se principalmente no que as instituições de pesquisa podem fazer para modificar os critérios usados pelos Comitês de seleção e promoção nas avaliações de suas carreiras. Os cinco princípios que formulados visam como as instituições de pesquisa buscam incentivar, recompensar e avaliar pesquisadores individuais em suas respectivas organizações.

Os HKPs enviam uma mensagem clara de que comportamentos que promovem a integridade da pesquisa precisam ser reconhecidos e recompensados. Os cinco princípios que formulados pelos pesquisadores responsáveis pelo desenvolvimento dos HKPs chamam atenção para o fato de que as instituições de pesquisa devem incentivar, recompensar e avaliar pesquisadores individuais por comportamentos que promovem a integridade da pesquisa em suas respectivas organizações. Esses temas precisam ser considerados com atenção durante uma avaliação de um grupo de pesquisadores no momento de tomar decisões sobre alocações de fundos ou políticas de recursos humanos. Além disso, essas questões dizem respeito à justiça social e à relevância social da pesquisa, e não à integridade da pesquisa (MOHER, ET AL 2020).

De acordo com Moher e colaboradores (2020), se implementados, os HKPs podem desempenhar um papel crítico nas avaliações baseadas em evidências que colocam o rigor da pesquisa no centro da avaliação, além de abrir a pesquisa para o benefício mais amplo da sociedade.

Assim como no caso da DORA para ajudar a facilitar a adoção os HKPs pelos envolvidos na avaliação da pesquisa e dos pesquisadores, tais como: instituições acadêmicas, financiadores, outros grupos e indivíduos podem endossar a proposta dos HKPs.

5. CONSIDERAÇÕES FINAIS

As atividades acadêmico-científicas são financiadas, na maioria dos países do mundo, total ou parcialmente por recursos públicos. Mesmo universidades privadas são, pelo menos em parte, financiadas com recursos públicos. Sendo assim, nosso trabalho se beneficia de financiamento público, desse modo, por razões fundamentais de prestação de contas à sociedade, a partir da academia podemos e devemos discutir as formas e os métodos empregados na avaliação de projetos e carreiras acadêmicas, mas não a sua necessidade.

Avaliações como as conduzidas pela (Coordenação de Aperfeiçoamento de Pessoal em Nível Superior (CAPES) tem desempenhado papel importante no desenvolvimento da pós-graduação brasileira, mas, como qualquer processo avaliativo de longa duração, deve ser periodicamente repensada para o aprimoramento dos mecanismos de atribuição de valor e diferenciação entre objetos avaliados (Iriart, et al., 2015). Essa revisão é necessária, sobretudo, para que se evite a cristalização de parâmetros e critérios e para que se discutam criticamente as consequências que produzem, evitando distorções (Franceschet e Costantin, 2011; Moed, 2008).

A avaliação tem um papel indutor importante no desenvolvimento de uma área, pois os critérios utilizados orientam os caminhos para a organização dos programas de pesquisa e apontam tendências. Portanto, se desejamos incentivar ou induzir maior alinhamento com práticas da ciência aberta, devemos mudar a maneira como a pesquisa é avaliada de modo a recompensar a adoção dessas práticas pelos pesquisadores. Se buscamos melhorar a igualdade e a diversidade na academia precisamos modificar os critérios de recrutamento e promoção para incluir uma gama mais ampla de qualidades, habilidades e resultados. A maneira como a pesquisa é avaliada tem um efeito significativo em todo o esforço acadêmico. Isso se chama Lei de Campbell muitas vezes definida como "a maneira de medir alguém afeta a maneira como eles se comportam" (Campbell, 1979).

Os compromissos, posicionamentos coletivos e manifestos aqui apresentados destacam a necessidade mudar a forma como estamos "medindo" a pesquisa e nos convida a empregar as melhores práticas de avaliação para coibir abusos e promover o uso responsável de métricas e indicadores nos processos avaliativos.

6. REFERÊNCIAS

BENEDICTUS, R., Miedema, F., & Ferguson, M. W. Fewer numbers, better science. *Nature*, 538(7626), 453-455, 2016.

CAMPBELL, Donald T. "Assessing the impact of planned social change. *Evaluation and Program Planning*, v. 2, n.1, p.67-90, (1979). DOI: [https://doi.org/10.1016/0149-7189\(79\)90048-X](https://doi.org/10.1016/0149-7189(79)90048-X).

CARR, D. Sharing research data and findings relevant to the novel coronavirus (COVID-19) outbreak [online]. Wellcome Trust. 2020 [viewed 12 March 2020]. Available from: <https://wellcome.ac.uk/press-release/sharing-research-data-and-findings-relevant-novel-coronavirus-covid-19-outbreak>

COLLINI, S., Who are the spongers now? *London Review of Books*, 38(2), 2016.

CROSS, R. COVID-19 vaccines and antibodies advance even faster than expected. **Chemical & Engineering News**, v.98, n.24. Disponível em: <https://cen.acs.org/pharmaceuticals/drug-development/COVID-19-vaccines-antibodies-advance/98/i24>. Acesso em: 14 jul 2020.

DORA - SAN FRANCISCO DECLARATION OF RESEARCH ASSESSMENT (DORA). San Francisco, 2012. Disponível em: <https://sfdora.org/read/es/>. Acesso em: 5 julho 2020.

EUROPEAN COMMISSION. Study on Open Science. Impact, Implications and Policy Options. 2015. Disponível em: https://ec.europa.eu/research/innovation-union/pdf/expert-groups/rise/study_on_open_science-impact_implications_and_policy_options-salmi_072015.pdf. Acesso em: 23 Jul. 2019

GADD, Elizabeth. 2019. "Influenciar o mundo em mudança da avaliação da pesquisa". *Insights 32* (1): 6. DOI: <http://doi.org/10.1629/uksg.456>.

GINGRAS, Y. (2016). *Bibliometrics and research evaluation. Uses and abuses*. MIT Press.

HICKS, D. *et al.* "The Leiden Manifesto for research metrics". **Nature**, v. 520, p. 429-431, 2015. Disponível em: <https://www.nature.com/news/bibliometrics-the-leiden-manifesto-for-research-metrics-1.17351>. Acesso em: 5 julho 2020.

HAMMARFELT B. Recognition and reward in the academy: valuing publication oeuvres in biomedicine, economics and history. *Aslib J Inform Manag* 2017; 69(5):607-23.

IOANNIDIS JP, KHOURY MJ. Assessing value in biomedical research: the PQRST of appraisal and reward. *JAMA* 2014; 312:483±4. <https://doi.org/10.1001/jama.2014.6932>

LURIE, N. SAVILLE, M.; HATCHETT, R.; HALTON, J. Desenvolvimento de vacinas COVID-19 em velocidade pandêmica. **N. Engl. J.**, n.382, p.1969-1973, 2020. Doi: 10.1056/NEJMp2005630.

MENG, Y; HEJUN, L.; NICHOLAS, C. W.; CHANG-CHUN, D. L.; XUEYONG, Z., FANGZHU, Z.; DELI, H.; WENLI, Y.; YUANZI, H.; HENRY, T.; THOMAS, F. R.; ELISE, L.; DEVIN, S.; JOSEPH, G. J.; DENNIS, R. B.; IAN, A. W. Base estrutural de uma resposta compartilhada de anticorpos para SARS-CoV-2. **Science**, 13 julho 2020, eabd2321. DOI: 10.1126 / science.abd2321. Disponível em: <https://science.sciencemag.org/content/early/2020/07/10/science.abd2321.full>. Acessado em: 14 jul 2020.

MOHER D, NAUDET F, Cristea IA, Miedema F, Ioannidis JPA, Goodman SN (2018) Assessing scientists for hiring, promotion, and tenure. *PLoS Biol* 16(3): e2004089. <https://doi.org/10.1371/journal.pbio.2004089>.

MOHER D, BOUTER L, KLEINERT S, GLASZIOU P, SHAM MH, BARBOUR V, et al. (2020) Os Princípios de Hong Kong para avaliar pesquisadores: Promovendo a integridade da pesquisa. *PLoS Biol* 18 (7): e3000737. <https://doi.org/10.1371/journal.pbio.3000737>.

MUNAFO M, Nosek B, Bishop D, Button K, Chambers C, Percie du Sert N, et al. A manifesto for reproducible science. *Nature Human Behaviour*. 2017;1[1]:0021.

NATURE. Time to remodel the journal impact factor. Nature 2016;535(7613):466. Disponível em: <https://www.nature.com/news/time-to-remodel-the-journal-impact-factor-1.20332>. Acesso em 22 jul 2020.

QUAN W, CHEN B, SHU F. Publish or impoverish: An investigation of the monetary reward system of science in China (1999-2016). 2017. Disponível em: <https://arxiv.org/ftp/arxiv/papers/1707/1707.01162.pdf>. Acesso em: 22 jul 2020.

SANTOS, S. ; CALÓ, L. N. Gestão editorial: tendências e desafios na transição para a ciência aberta. In: SILVEIRA, L.; SILVA, F. C. C. (orgs.). Gestão editorial de periódicos científicos: tendências e boas práticas. 1. ed. – Florianópolis : BU Publicações/UFSC : Edições do Bosque/UFSC, 2020.

SAREWITZ, D. (2016). The pressure to publish pushes down quality. Nature, 533(7602), 147-147.

WILSDON, J., et al.. The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management. 2015. DOI: 10.13140/RG.2.1.4929.1363