



programa de apoio à ciência

guia de boas práticas em ciência aberta e reprodutível



guia de boas práticas em ciência aberta e reprodutível

Dentre os valores básicos da ciência está a ideia de que o conhecimento é um patrimônio coletivo da humanidade.¹ O avanço da ciência depende da veracidade e robustez de seus resultados, bem como da possibilidade de sua verificação e reprodução independente.

No entanto, particularidades do ambiente científico impõem obstáculos ao acesso e à confiabilidade do que é publicado. Devido a um sistema de incentivos baseado na publicação em periódicos especializados, geralmente associados a editoras comerciais, menos da metade dos artigos publicados estão disponíveis em acesso aberto.² Uma fração ainda menor disponibiliza inteiramente seus dados para verificação independente.^{3,4} Talvez por isso, levantamentos recentes em áreas como psicologia e ciência biomédica sugerem que grande parte dos resultados publicados não sejam reprodutíveis,⁵⁻⁷ o que acarreta desperdício de recursos e atraso no progresso científico.

Como uma instituição de apoio à ciência do Brasil, o Serrapilheira reafirma seu compromisso e expectativa de que os achados de pesquisas realizadas com o financiamento do instituto sejam divulgados de forma transparente para que sejam acessíveis, verificáveis e reprodutíveis. As diretrizes listadas abaixo constituem sugestões de boas práticas em diferentes etapas do processo científico para que estes objetivos sejam atingidos.

É natural que nem todas as recomendações se apliquem diretamente a todas as áreas de pesquisa. Físicos teóricos e matemáticos, por exemplo, não precisam se ater a questões sobre a condução de experimentos. Por outro lado, orientações sobre temas como publicação científica são válidas para todas as áreas, ainda que com particularidades próprias.

Ao elaborar um projeto:

- Realize uma revisão cuidadosa da literatura antes de iniciar um projeto, seja ele experimental ou teórico. Isto ajuda a evitar a duplicação de esforços e a assegurar que a autoria de ideias e de resultados prévios seja adequadamente atribuída. O uso de protocolos sistemáticos de revisão, com metodologias de busca explícitas, torna o processo mais isento, reprodutível e robusto como contribuição científica.⁸



- Registrar seu protocolo antes da coleta de dados acrescenta transparência ao processo e permite verificar se a obtenção e a análise dos dados foram realizadas conforme o planejado, a fim de diferenciar as análises confirmatórias das exploratórias.^{9,10} Além disso, a prática permite a identificação de trabalhos registrados e ainda não publicados.¹¹ Tal procedimento é obrigatório para ensaios clínicos^{12,13} e pode ser adotado de formas variadas por outras áreas de pesquisa. Repositórios usados para o pré-registro de protocolos incluem a [Open Science Framework](#) e o [AsPredicted](#). Os protocolos podem ser disponibilizados publicamente no momento do registro ou colocados em embargo até a divulgação dos dados.
- Uma opção para receber feedback antes da realização do trabalho – momento em que críticas e sugestões têm maior possibilidade de contribuir para o projeto – é submeter os protocolos à revisão por pares, seja como artigo independente, seja como submissão preliminar em revistas que aceitem o formato de Registered Report, em que se realiza a revisão sobre os métodos, gerando um aceite preliminar do trabalho se o protocolo for seguido, independentemente do resultado.¹⁴ Uma lista de periódicos que aceitam o formato está disponível na página do [Center for Open Science](#).
- Antes de coletar dados, faça um plano de análise estatística que inclua um cálculo de tamanho amostral. Isto previne a realização de experimentos com poder estatístico baixo, que desperdiçam recursos e geram resultados menos confiáveis.¹⁵ Inúmeras calculadoras online realizam estes cálculos de maneira fácil, como as disponíveis em [Power and Sample Size](#).
- Ao desenhar seu projeto, lembre-se de incluir medidas e controles para obter conclusões robustas, independentemente do resultado. É importante estabelecer critérios para diferenciar um resultado “negativo” (em que uma determinada intervenção experimental não gera o resultado esperado) de um experimento falho devido a problemas metodológicos, de forma a garantir que os primeiros sejam confiáveis o suficiente para serem publicados. A publicação de resultados negativos é uma medida fundamental para prevenir a distorção do conhecimento científico pelo viés de publicação.^{16,17}



- Estabeleça um plano de gerenciamento de dados que preveja onde eles serão armazenados durante o projeto e após sua conclusão. Inclua uma previsão de cópias de segurança e garantia de acessibilidade a longo prazo, independentemente de mudanças no laboratório/instituição. Diversos repositórios podem ser utilizados para este fim, sejam eles atrelados a instituições específicas ou não – para uma lista, acesse FAIRsharing.org. O site da [FAPESP](https://FAPESP.org) também fornece orientações de como elaborar um plano de gestão de dados.

Ao realizar experimentos:

- Utilize medidas de controle de viés, como cegamento dos experimentadores em relação à identidade dos grupos durante a realização de experimentos e avaliação de desfechos. Ainda que a importância de tais procedimentos seja reconhecida há décadas, a adoção por diversas áreas de pesquisa, como a ciência de laboratório com animais, ainda é inadmissivelmente baixa para um procedimento que deveria ser corriqueiro.¹⁸
- Adote um sistema de registro de protocolos e resultados, como um caderno de laboratório ou uma ferramenta eletrônica que cumpra esta função. Treine os responsáveis pela coleta de dados, deixando claro que o registro é obrigatório e representa a documentação do que foi realizado, não podendo ser apagado ou adulterado. Para mais orientações sobre cadernos de laboratório, acesse o site do [National Institutes of Health](https://www.nih.gov).
- Tente conhecer a reprodutibilidade intra e interlaboratorial dos métodos que você utiliza. Caso esta não seja conhecida, considere a possibilidade de replicar dados-chave de seu projeto entre diferentes membros de seu laboratório, ou entre diferentes laboratórios. Ainda que não seja possível fazer isso para todos os dados, esta pode ser uma prática valiosa para experimentos confirmatórios importantes.¹⁹
- Estabeleça protocolos de controle de qualidade e checagem de erros metodológicos comuns de forma a avaliar periodicamente a qualidade dos dados gerados. Caso tais procedimentos ocasionem a exclusão de dados ou experimentos, estabeleça critérios para que isto aconteça antes de obter os resultados e mencione-os explicitamente em suas

descrições de métodos, para evitar que essas exclusões possam enviesar os resultados.

Ao analisar seus dados:

- Mantenha seus bancos de dados organizados e documente as análises realizadas da mesma forma que você faria com resultados experimentais. Caso use código próprio para análise, mantenha-o documentado e compreensível para terceiros. Considere integrar o código a seus resultados com ferramentas como o [Jupyter](#), [R Markdown](#), [knitr](#) ou [Sweave](#), que permitem repetir análises e simulações facilmente.
- Ao utilizar estatística inferencial, lembre que tal processo assume pressupostos sobre os dados e hipóteses previamente estabelecidas – e idealmente registradas a priori. Não nos opomos ao uso de análises exploratórias; estas, porém, devem ser claramente descritas como tal e diferenciadas de análises desenhadas para testar hipóteses a priori.²⁰ Modelos estatísticos construídos a partir dos dados, como aqueles gerados por *machine learning*, devem ser testados em conjuntos de dados distintos daqueles em que foram gerados para evitar circularidade nas análises.^{21,22}
- Evite a dicotomização corriqueira entre resultados “significativos” e “não-significativos”. Caso você utilize estatística frequentista, forneça valores de p exatos, e interprete-os à luz da plausibilidade de sua hipótese e do poder estatístico do experimento antes de considerar um resultado como provavelmente verdadeiro.^{23,24} Se desejado, este tipo de lógica pode ser formalizado em análises bayesianas que levem em conta a probabilidade a priori atribuída a uma hipótese.
- Leve em conta que significância estatística e magnitude de efeito são conceitos distintos e considere alternativas para expressar ambas as grandezas de forma clara ao apresentar seus dados, como o uso de medidas de tamanho de efeito associadas a intervalos de confiança.²⁵
- Independentemente do resultado de suas análises, considere que toda inferência a partir de uma amostra tem possibilidade de erro. Exponha honestamente a dúvida, tanto em relação a resultados positivos quanto a negativos, e aborde as limitações de seus métodos e análises ao discutir resultados.

Ao publicar seus resultados:

- Utilize formas de publicação em acesso aberto. Elas podem incluir tanto periódicos *open access* (via dourada) como o depósito de versões dos artigos em repositórios públicos (via verde). Para mais informações sobre acesso aberto, acesse o site da [Budapest Open Access Initiative](#) e da [cOAlition S](#). Em relação à via verde, a maioria das editoras possui políticas compatíveis com o depósito de versões pre- e postprint – para consultar as políticas de uma editora ou periódico em particular, acesse [SHERPA/RoMEO](#). Ao publicar ou depositar um artigo, utilize opções de compartilhamento que garantam que ele possa ser utilizado e redistribuído, como as licenças [Creative Commons](#).
- Caso você opte por periódicos de acesso aberto que cobrem taxas de publicação (ressaltando que esta não é a única forma de garantir o acesso aberto), busque certificar-se de que os mesmos possuem um sistema de revisão por pares real e não se enquadram na categoria dos “periódicos predatórios”.²⁶ Uma lista de periódicos de acesso aberto que atendem requisitos mínimos de controle de qualidade pode ser encontrada no [Directory of Open Access Journals](#).
- Deposite seu artigos como preprints antes ou no momento da submissão em repositórios como o [arXiv](#), [bioRxiv](#), [chemRxiv](#), [PsyArXiv](#), [medRxiv](#) e outros (uma lista ampla é atualizada periodicamente em [Research Preprints](#)). A utilização de preprints acelera o processo científico e é compatível com a submissão do artigo para a maioria dos periódicos (para pesquisar políticas específicas, veja [SHERPA/RoMEO](#) ou a [Wikipedia](#)). Após a revisão por pares, atualize seu preprint para que ele esteja o mais próximo possível da publicação final – o que também serve como uma forma de alcançar o acesso aberto pela via verde.
- A citação de preprints em artigos científicos e o engajamento em comentários sobre os mesmos (i.e. revisão por pares pós-publicação) também é estimulada. Além disso, preprints devem ser considerados como produção científica válida na avaliação de projetos e pesquisadores. Para mais informações sobre preprints, especialmente na área de ciências da vida, acesse o site do [ASAPbio](#).

- Descreva sua metodologia o mais extensivamente possível para que outros pesquisadores possam replicar seus resultados. Caso o periódico apresente alguma limitação de espaço, utilize material suplementar ou deposite o protocolo em um repositório, usando idealmente plataformas específicas para este fim, como o protocols.io. Caso existam *reporting guidelines* para sua área de pesquisa, utilize-os como um checklist do que descrever. Para uma coleção de *guidelines* na área da saúde e ciências da vida, acesse a [Equator Network](https://equator-network.org).
- Compartilhe integralmente os dados de sua pesquisa no momento da submissão de seu artigo. Os mesmos podem ser incluídos como material suplementar ou depositados em repositórios, sejam eles pertencentes a sua instituição, repositórios gerais (e.g. [Zenodo](https://zenodo.org), [Dryad](https://dryad.org), [figshare](https://figshare.com), [Dataverse](https://dataverse.org)) ou plataformas especializadas em um tipo particular de dados. Uma lista de repositórios por área está disponível no [FAIRsharing.org](https://fairsharing.org). Se houver consenso sobre standardização de dados em sua área, dê preferência a estes formatos estruturados. Caso contrário, inclua um manual descrevendo dados e metadados que os tornem acessíveis a buscas automatizadas. Em todos os casos, certifique-se de que seus dados tenham um DOI para citação. A mera informação de que os dados estão disponíveis sob contato com os autores não é suficiente, já que tal prática não assegura o acesso.^{27,28} Para mais informações sobre princípios de depósito de dados científicos, consulte os [princípios FAIR](https://fairsharing.org).
- No caso de dados sensíveis que não possam ser compartilhados integralmente, como informações pessoais de voluntários ou dados que requeiram proteção por questões de propriedade intelectual, considere se os mesmos não podem ser disponibilizados parcialmente (e.g. com remoção de informações para garantir o anonimato de voluntários) ou após um período de embargo (e.g. depois que a propriedade intelectual for estabelecida). Para mais informações sobre privacidade de dados, consulte a [Lei Geral de Proteção de Dados](https://www.gov.br/legis/pt/leis/lei-13709-2018) e as [diretrizes da European Commission](https://ec.europa.eu/commission/presscorner/detail/en/ip-18-171) sobre proteção de dados em pesquisa.
- Caso sua pesquisa utilize código elaborado por você ou por terceiros, inclua-o como material suplementar ou deposite-o em um repositório, como o [GitHub](https://github.com). Dê preferência a linguagens de código aberto, como [R](https://www.r-project.org/) ou [Python](https://www.python.org/), mas disponibilize seu código mesmo que você trabalhe em uma linguagem proprietária. Novamente, considere integrar o código



aos resultados com ferramentas como o [Jupyter](#), [R Markdown](#), [knitr](#) ou [Sweave](#). Em suas publicações, cite códigos e bancos de dados de terceiros com seus respectivos DOIs, a fim de reconhecê-los como formas válidas de produção científica.

- Atribua a autoria de publicações de forma justa e compatível com os critérios de sua área de pesquisa. Ainda que estes possam variar conforme o campo, diretrizes gerais podem ser encontradas no site do [Committee on Publication Ethics \(COPE\)](#), e diversas áreas e periódicos utilizam os critérios do [International Committee of Medical Journal Editors](#). Para maior clareza, recomendamos que as contribuições de cada autor sejam detalhadas em uma seção à parte, idealmente utilizando taxonomias padronizadas, como a [CRediT](#).
- Caso você tenha qualquer conflito de interesse relacionado a seus resultados – seja ele financeiro ou não –, declare-o de forma transparente ao publicá-los e apresentá-los. Ainda que por vezes conflitos de interesse sejam inevitáveis, o esforço para minimizá-los contribui para a isenção e integridade de seus resultados de pesquisa. Para mais informações e definições sobre conflitos de interesse, acesse o site do [COPE](#).

Em todos os momentos:

- Engaje-se na discussão sobre ciência aberta e reprodutível, especialmente na formação de novos cientistas. Forneça os incentivos certos a seus estudantes e colaboradores, deixando claro que eles serão valorizados pelo esforço e rigor que investirem na pesquisa, e não pelo resultado obtido. Traga questões relacionadas à confiabilidade científica para discussão em seu grupo de pesquisa e instituição, e tome contato com a literatura sobre o tema.^{29,30}
- Ao avaliar seus pares, como editor, revisor, membro de banca ou parecerista, leve em conta o rigor científico, a transparência e a aderência à ciência aberta do pesquisador que está sendo avaliado. Exerça seu ceticismo e avalie se os resultados que você está avaliando são sólidos e confiáveis antes de perguntar-se sobre o impacto dos mesmos.

- Considere a ciência uma prática que necessita de revisão permanente. Valorize, discuta ou realize replicações e estudos confirmatórios em sua área de pesquisa.³¹ Registre e publique suas tentativas de replicação, tenham elas sucesso ou não. Engaje-se em práticas de avaliação da ciência publicada, participando de fóruns de revisão pós-publicação, como o [PubPeer](#), e as seções de comentários de periódicos e repositórios de preprints. Ao atuar como revisor para periódicos, dê preferência àqueles que divulgam as revisões,³² seja de forma assinada ou anônima. Seja cortês e respeitoso em todas as instâncias, de forma a providenciar críticas construtivas que contribuam para o avanço da ciência.
- Da mesma forma, seja aberto a críticas de colegas e procure não interpretar críticas aos seus dados como ataques pessoais. Esteja aberto à possibilidade de correção e reanálise de seus dados, colaborando com as mesmas quando possível. Caso erros ou inconsistências sejam detectados, seja transparente ao admiti-los, utilizando os mecanismos de correção e retratação existentes na literatura. Lembre que admitir erros honestamente não é uma mácula na reputação de um pesquisador, e que as consequências éticas de não corrigi-los são mais graves.³³

As recomendações deste guia foram elaboradas por **Olavo Amaral** (Instituto de Bioquímica Médica Leopoldo de Meis, UFRJ) e revisadas por **grantees do Instituto Serrapilheira** atuantes em diferentes áreas de pesquisa. Coordenação: **Cristina Caldas**. Edição e revisão: **Maria Emilia Bender**.

Instituto Serrapilheira

Diretor-Presidente: Hugo Aguilaniu

Diretora de Pesquisa Científica: Cristina Caldas

Diretora de Divulgação Científica: Natasha Felizi

Referências

1. Merton RK. **The normative structure of Science.** In: Merton, RK. *The Sociology of Science: Theoretical and Empirical Investigations.* Chicago, University of Chicago Press, 1942.
<https://doi.org/10.2307/2576098>
2. Piwowar H, Priem J, Larivière V et al. **The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles.** *PeerJ* 2018; 6: e4375.
<https://doi.org/10.7717/peerj.4375>
3. Alsheikh-Ali AA, Qureshi W, Al-Mallah MH, Ioannidis JPA. **Public availability of published research data in high-impact journals.** *PLoS One* 2011; 6: e23457.
<https://doi.org/10.1371/journal.pone.0024357>
4. Iqbal SA, Wallach JD, Khoury MJ, Schully SD, Ioannidis JPA. **Reproducible research practices and transparency across the biomedical literature.** *PLoS Biol* 2016; 14: e1002333
<https://doi.org/10.1371/journal.pbio.1002333>
5. Open Science Collaboration. **Estimating the reproducibility of psychological science.** *Science* 2015; 349: aac4716.
<https://doi.org/10.1126/science.aac4716>
6. Begley CG, Ellis LM. **Drug development: Raise standards for preclinical cancer research.** *Nature* 2012; 483: 531–533.
<https://doi.org/10.1038/483531a>
7. Prinz F, Schlange T, Asadullah K. **Believe it or not: how much can we rely on published data on potential drug targets?** *Nat Rev Drug Discov* 2012; 10: 712. <https://doi.org/10.1038/nrd3439-c1>
8. Higgins JPT, Green S (eds). *Cochrane Handbook for Systematic Reviews of Intervention.* Version 5.1.0. The Cochrane Collaboration, 2011. Disponível em <http://www.handbook.cochrane.org>.
<https://doi.org/10.1002/9780470712184>

9. Kimmelman J, Mogil JS, Dirnagl U. **Distinguishing between exploratory and confirmatory preclinical research will improve translation.** *PLoS Biol* 2014; 12: e1001863.
<https://doi.org/10.1371/journal.pbio.1001863>
10. Nosek BA, Ebersole CR, DeHaven AC, Mellor DT. **The preregistration revolution.** *Proc Nat Acad Sci USA* 2018; 115: 2600-2606.
<https://doi.org/10.1073/pnas.1708274114>
11. Goldacre B, DeVito N, Heneghan C, Lane S, Stephens R. **Introducing unreported clinical trial of the week.** *BMJ blog*, disponível em <https://blogs.bmj.com/bmj/2018/03/29/it-is-time-to-fix-the-problem-of-unreported-clinical-trials/>
12. Zarin D, Tse T, Williams RJ, Califf RM. **The ClinicalTrials.gov results database – update and key issues.** *N Engl J Med* 2011; 364: 852-860.
<https://doi.org/10.1056/nejmsa1012065>
13. World Health Organization. **International Clinical Trials Registry Platform (ICTRP).** Disponível em <https://www.who.int/ictcp/en/>
14. Chambers C, Munafò M et al. **Trust in science would be improved by study pre-registration.** *Guardian*, 5/6/2013. Disponível em: <https://www.theguardian.com/science/blog/2013/jun/05/trust-in-science-study-pre-registration>
15. Button KS, Ioannidis JP, Mokrysz C et al. **Power failure: why small sample size undermines the reliability of neuroscience.** *Nat Rev Neurosci* 2013; 14: 365-376. <https://doi.org/10.1038/nrn3475>
16. Dwan K, Altman DG, Arnaiz JA et al. **Systematic review of the empirical evidence of study publication bias and outcome reporting bias.** *PLoS One* 2008; 3: e3081.
<https://doi.org/10.1371/journal.pone.0003081>
17. Moorthy VS, Karam G, Vannice KS, Kieny MP. **Rationale for WHO's new position calling for prompt reporting and public disclosure of interventional clinical trial results.** *PLoS Med* 2015; 12: e1001819.
<https://doi.org/10.1371/journal.pmed.1001819>



18. Macleod MR, McLean AL, Kyriakopoulou A. et al. **Risk of bias in reports of in vivo research: a focus for improvement.** *PLoS Biol* 2015; 13: e1002273. <https://doi.org/10.1371/journal.pbio.1002273>
19. Mogil JS, Macleod MR. **No publication without confirmation.** *Nature* 2017; 542: 409-411. <https://doi.org/10.1038/542409a>
20. Kerr NL. **HARKing: hypothesizing after the results are known.** *Pers Soc Psychol Rev* 1998; 2: 196-217. https://doi.org/10.1207/s15327957pspr0203_4
21. Kriegeskorte N, Simmons WK, Bellgowan PSF, Baker CI. **Circular analysis in systems neuroscience: the dangers of double dipping.** *Nat Neurosci* 2009; 12:535-540. <https://doi.org/10.1038/nn.2303>
22. Baştanlar Y, Özuysal M. **Introduction to machine learning.** *Methods Mol Biol* 2014; 1107:105-128. https://doi.org/10.1007/978-1-62703-748-8_7
23. Gigerenzer G. **Statistical rituals: the replication delusion and how we got there.** *Adv Meth Pract Psychol Sci* 2018; 1: 198-218. <https://doi.org/10.1177/2515245918771329>
24. Nuzzo R. **Statistical errors.** *Nature* 2014; 506: 150-152. <https://doi.org/10.1038/506150a>
25. Cumming G. **The new statistics: why and how.** *Psychol Sci* 2014; 25: 7-29. <https://doi.org/10.1177/0956797613504966>
26. Clark J, Smith R. **Firm action needed on predatory journals.** *BMJ* 2015; 350. <https://doi.org/10.1136/bmj.h210>
27. Hardwicke TE, Ioannidis, JPA. **Populating the Data Ark: an attempt to retrieve, preserve, and liberate data from the most highly-cited psychology and psychiatry articles.** *PLoS One* 2018; 13: e0201856. <https://doi.org/10.1371/journal.pone.0201856>



28. Stodden V, Seller J, Ma Z. **An empirical analysis of journal policy effectiveness for computational reproducibility.** *Proc Natl Acad Sci USA* 2018; 115: 2584-2589.
<https://doi.org/10.1073/pnas.1708290115>
29. Munafò MR, Nosek BA, Bishop DVM et al. **A manifesto for reproducible science.** *Nat Hum Behav* 2017 1: 0021.
<https://doi.org/10.1038/s41562-016-0021>
30. Ioannidis JPA. **How to make more published research true.** *PLoS Med* 2014; 11: e1001747.
<https://doi.org/10.1371/journal.pmed.1001747>
31. Nature. **Go forth and replicate!** *Nature* 2016; 536: 373.
<https://doi.org/10.1038/536373a>
32. Polka JK, Kiley R, Konforti B, Stern B, Vale RD. **Publish peer reviews.** *Nature* 2018; 560: 545-547.
<https://doi.org/10.1038/d41586-018-06032-w>
33. Rohrer JM, DeBruine LM, Heyman T, et al. **Putting the self in self-correction.** *PsyArXiv* 2018. <https://doi.org/10.31234/osf.io/exmb2>