

Comment assurer la véracité d'un plus grand nombre de recherches publiées

John P. A. Ioannidis*

Article original : Ioannidis, J. P. A. (2014). How to Make More Published Research True. *PLoS Med*, 11(10), e1001747.
<https://doi.org/10.1371/journal.pmed.1001747>

RÉSUMÉ

- Actuellement, de nombreuses conclusions de recherche publiées sont inexactes ou exagérées, et on estime que 85 % des ressources de recherche sont gaspillées.
- Pour accroître la fiabilité des recherches publiées, il est possible de transposer à d'autres domaines certaines pratiques ayant déjà amélioré la crédibilité et l'efficacité dans des secteurs spécifiques. Ces pratiques incluent : la mise en place de collaborations à grande échelle, l'adoption d'une culture de la réplication, l'enregistrement des études, le partage des données, la mise en œuvre de pratiques favorisant la reproductibilité, le recours à de meilleures méthodes statistiques, la normalisation des définitions et des analyses, l'utilisation de seuils statistiques plus appropriés (souvent plus stricts), ainsi que l'amélioration des normes de conception des études, du processus d'évaluation par les pairs, de la communication et de la diffusion des résultats de recherche, et de la formation des scientifiques.
- Le choix des interventions visant à améliorer les pratiques de recherche doit s'appuyer sur un examen rigoureux et des tests expérimentaux dans la mesure du possible.
- Les interventions les plus efficaces doivent prendre en compte et exploiter les motivations des différentes parties prenantes de la recherche scientifique, qui varient selon leur intérêt pour la production de résultats publiables, finançables, applicables ou rentables.
- Des réformes sont nécessaires dans le système de récompense en science, en ajustant les « taux de change » des différentes « devises académiques » (par exemple, les publications et subventions) et des « biens académiques » acquis (comme les promotions ou d'autres formes de pouvoir académique ou administratif), tout en introduisant de nouvelles « devises » mieux alignées sur les objectifs de reproductibilité et d'applicabilité des recherches.

Les réalisations de la recherche scientifique sont remarquables. La science est passée d'une profession réservée à quelques dilettantes à une industrie mondiale dynamique, mobilisant plus de 15 millions de personnes et produisant plus de 25 millions d'articles scientifiques entre 1996 et 2011 seulement (Boyack et al., 2013). Cependant, les découvertes majeures, véritablement innovantes et directement

* L'affiliation de l'auteur se trouve à la fin de l'article.

applicables demeurent relativement rares. De nombreuses associations ou effets nouvellement proposés se révèlent faux ou largement exagérés (Ioannidis, 2008; Ioannidis, 2005), et la transformation des connaissances en applications concrètes s'avère souvent lente et potentiellement inefficace (Contopoulos-Ioannidis et al., 2008). Compte tenu de l'abondance des données disponibles, la recherche sur la recherche (ou métarecherche) peut fournir des estimations empiriques sur la prévalence des facteurs de risque liés à un taux élevé de faux positifs : études insuffisamment puissantes; faibles tailles d'effet; faibles probabilités pré-étude; flexibilité dans les conceptions, définitions, résultats et analyses; biais et conflits d'intérêts; effets de mode; et manque de collaboration (Ioannidis, 2005). On estime actuellement qu'environ 85 % des ressources allouées à la recherche sont gaspillées (Macleod et al., 2014).

Interventions efficaces

Nous avons besoin d'interventions efficaces pour améliorer la crédibilité et l'efficacité de la recherche scientifique. Certains facteurs de risque menant à des résultats erronés sont immuables, comme les faibles tailles d'effet, mais d'autres peuvent être modifiés. Nous devons réduire les biais, les conflits d'intérêts et la fragmentation des efforts, et promouvoir une recherche impartiale, transparente et collaborative, avec une normalisation accrue. Cependant, il convient également de prendre en compte la possibilité que certaines interventions visant à améliorer l'efficacité scientifique puissent entraîner des effets secondaires indésirables ou consommer inutilement des ressources. Prenons un exemple extrême : on pourrait éliminer tous les faux positifs en écartant systématiquement toute étude présentant même un biais minimal, en formulant des questions de recherche si anodines qu'elles n'intéresseraient personne (et ne susciteraient aucun conflit), et en attendant que tous les scientifiques d'un domaine se mettent d'accord sur un protocole et un plan d'analyse unique et normalisé. Dans ce cas, le taux d'erreur serait effectivement réduit à zéro, mais au prix de l'arrêt complet de la production scientifique. Ainsi, quelles que soient les solutions proposées, elles doivent être pragmatiques, applicables et, idéalement, susceptibles d'être soumises à des tests fiables pour évaluer leur efficacité.

Actuellement, les grandes décisions concernant la manière dont la recherche est menée reposent encore trop souvent sur des conventions établies et une certaine inertie, plutôt que sur une approche véritablement imaginative ou fondée sur des preuves solides (Macleod et al., 2014; Nicholson et Ioannidis, 2012; Wenneras et Wold, 1997; Nickerson, 1998; Mynatta et al., 1977; Greenhalgh et al., 2014; Stamatakis et al., 2013; Chalmers et al., 2014; Rennie et Flanagan, 2014; Danthi et al., 2014; Ioannidis, 2011). Par exemple, il a été démontré que ceux et celles qui évaluent les demandes de subvention ont généralement des CV relativement modestes et que la plupart des scientifiques ayant le plus d'influence ne participent pas à l'évaluation de ces demandes ni ne bénéficient de financements publics; et ce, même aux États-Unis (Nicholson et Ioannidis, 2012), un pays qui à l'heure actuelle exerce probablement l'impact scientifique le plus important au monde (par exemple, en termes de citations cumulées). Des pratiques non méritocratiques, telles que le népotisme, le sexisme et un conservatisme excessif, semblent être largement répandues (Wenneras et Wold, 1997). Par ailleurs, les biais d'allégeance et de confirmation exercent une influence considérable sur les processus scientifiques (Nickerson, 1998; Mynatta et al., 1977). Dans le domaine des soins de santé et de la pratique clinique,

bien que la médecine fondée sur les preuves se soit renforcée avec le temps, on soutient qu'elle traverse aujourd'hui une crise (Greenhalgh et al., 2014), et que le terme « fondé sur les preuves » a été détourné pour justifier des croyances de spécialistes ou pour servir les intérêts de l'industrie (Stamatakis et al., 2013). Nous disposons de peu de données expérimentales sur la manière optimale de procéder à l'évaluation par les pairs et sur le moment où celle-ci devrait avoir lieu (par exemple, sur la base de protocoles, de manuscrits ou après publication) (Macleod et al., 2014; Chalmers et al., 2014; Rennie et Flanagin, 2014), ou encore sur la façon dont les fonds de recherche devraient être attribués (Danthi et al., 2014; Ioannidis, 2011). De nombreuses structures scientifiques dominantes trouvent leurs origines au Moyen Âge (par exemple, les hiérarchies académiques) ou au XVIIe siècle (par exemple, les sociétés professionnelles et la publication dans des revues savantes), mais leur pertinence pour répondre aux besoins actuels du développement scientifique reste incertaine. Par ailleurs, une tension évidente existe entre l'ambition de rendre les décisions à la fois plus créatives et plus fondées sur des preuves. La bureaucratie et la pratique de la science pourraient nécessiter des personnes dotées de compétences différentes, et il est également possible qu'un système trop axé sur l'élimination des discriminations injustes aboutisse à supprimer également les discriminations raisonnables nécessaires pour prendre des décisions éclairées. Si certaines réformes pourraient effectivement nuire à la science, nous avons également la possibilité d'introduire, de manière délibérée, des changements qui l'amélioreraient.

Une option consiste à transposer dans le plus grand nombre possible de disciplines scientifiques des pratiques de recherche qui ont démontré leur efficacité lorsqu'elles ont été appliquées ailleurs. L'encadré 1 présente quelques exemples, détaillés ici.

L'adoption de recherches collaboratives à grande échelle, avec une forte culture de la réplication (NCI-NHGRI Working Group on Replication in Association Studies, 2007), a été un succès dans plusieurs domaines biomédicaux, notamment en épidémiologie génétique et moléculaire. Ces techniques ont contribué à transformer l'épidémiologie génétique, autrefois considérée comme un domaine douteux (Ioannidis et al., 2011), en une discipline hautement crédible (Panagiotou et al., 2013). Ces pratiques pourraient être appliquées à d'autres champs de la recherche observationnelle et au-delà (Houry et al., 2013).

Encadré 1 : Certaines pratiques de recherche susceptibles d'augmenter la proportion de résultats de recherche fiables

- Recherche collaborative à grande échelle
- Adoption d'une culture de réplication
- Enregistrement (des études, protocoles, codes d'analyse, ensembles de données, données brutes et résultats)
- Partage (des données, protocoles, matériaux, logiciels et autres outils)
- Pratiques visant à renforcer la reproductibilité
- Limitation des conflits d'intérêts entre parties qui financent, auteurs et autrices
- Utilisation de méthodes statistiques plus appropriées
- Normalisation des définitions et des analyses
- Seuils plus stricts pour revendiquer des découvertes ou des « succès »
- Amélioration des normes de conception des études

- Amélioration du processus d'évaluation par les pairs, de l'élaboration des rapports et de la diffusion des recherches
- Meilleure formation des scientifiques en méthodologie et en culture statistique

La notion de réplication diffère selon les contextes et les modèles d'étude. Dans les études de base en laboratoire et les études précliniques, la réplication devrait par défaut être réalisable. Cependant, même dans ces cas, il est crucial de définir a priori les caractéristiques essentielles qui doivent être reproduites et de déterminer le degré d'hétérogénéité acceptable (Bissell, 2013). Dans certaines recherches cliniques, la réplication est difficile, notamment pour les études très vastes, de longue durée et coûteuses. La perspective de réplication doit être envisagée et intégrée dès le départ, lors de la conception du programme de recherche dans un domaine donné (Chalmers et al., 2014). Faut de cela, certaines questions ne sont jamais explorées ou ne le sont que par des études uniques jamais reproduites, tandis que d'autres font l'objet de réplifications multiples inutiles, voire de méta-analyses redondantes les combinant (Siontis et al., 2013).

L'enregistrement des essais randomisés (Zarin et al., 2007) (et, plus récemment, l'enregistrement de leurs résultats [Zarin et al., 2011]) a renforcé la transparence dans la recherche clinique et a permis d'analyser les biais de rapport sélectif (Dwan et al., 2013; Chan et al., 2014), bien que cela ne les élimine pas complètement. Cet enregistrement peut révéler des redondances et permettre de mieux visualiser l'évolution de l'ensemble du corpus de recherche dans un domaine donné. Actuellement, l'enregistrement est proposé pour de nombreux autres types de recherche, y compris les études observationnelles humaines (Dal-Ré et al., 2014) et les études non humaines (Macleod, 2011).

Le partage des données, des protocoles, des matériaux et des logiciels a été encouragé dans plusieurs domaines « -omiques », fournissant ainsi une base pour des pratiques de données reproductibles (Stodden et al., 2013; Donoho, 2010; Peng, 2011; Peng et al., 2006). De même, la promotion du partage des données dans les essais cliniques pourrait renforcer la crédibilité de la recherche clinique (Doshi et al., 2013). Cependant, certains inconvénients ont été soulevés, comme la possibilité que plusieurs analystes effectuent des analyses contradictoires, les défis liés à l'anonymisation des personnes participantes, et le risque que certaines parties introduisent des incertitudes pour des résultats qui nuisent à leurs intérêts, comme cela a été observé dans le cas des gaz d'échappement diesel et des risques de cancer (Montfortin, 2006).

Dissocier certains types de recherche des commanditaires ou des auteurs et autrices présentant des conflits d'intérêts a été proposé (non sans controverse) pour des approches aussi diverses que les analyses coût-efficacité (Kassirer et Angell, 1994), les méta-analyses (Jørgensen et al., 2006; Gøtzsche et Ioannidis, 2012) et les lignes directrices (Institute of Medicine, 2011). Pour toutes ces catégories de recherche, il a été démontré que l'implication de commanditaires présentant des conflits d'intérêts tend à orienter les conclusions de manière plus favorable.

L'adoption de méthodes statistiques plus appropriées (Nuzzo, 2014), de définitions et d'analyses normalisées, ainsi que de seuils plus stricts pour revendiquer des découvertes ou des « succès »

(Johnson, 2013), pourrait réduire les taux de faux positifs dans des disciplines jusqu'ici trop laxistes, comme l'épidémiologie (Young et Karr, 2011), la psychologie (Pashler et Harris, 2012; Simmons et al., 2011) ou l'économie (Ioannidis et Doucouliagos, 2013). Ces réformes pourraient rehausser la crédibilité de ces domaines, les rapprochant des disciplines traditionnellement plus rigoureuses en la matière, comme les sciences physiques (Fanelli, 2010).

Améliorer les normes de conception des études pourrait également renforcer la fiabilité des résultats (Poste, 2012). Par exemple, pour les études animales portant sur des interventions, cela inclurait la randomisation et l'insu des chercheurs et chercheuses (Macleod, 2011). Avec l'intérêt croissant que suscitent les listes de contrôle pour la réalisation des études devant être approuvées (Landis et al., 2012; Collins et Tabak, 2014), il faut s'assurer que ces critères soient réellement essentiels et que les affirmations de conformité soient vérifiables.

La manière dont la recherche est rapportée, évaluée, publiée, diffusée et examinée après publication joue un rôle essentiel dans sa fiabilité. De nombreux efforts sont actuellement en cours pour améliorer et normaliser ces processus (comme le répertorie l'initiative EQUATOR [Simera et al., 2010]), et des discussions se poursuivent sur la manière de réformer l'évaluation par les pairs (déterminer qui doit la réaliser, comment et à quel moment) et la diffusion des informations (Chan et al., 2014; Nosek et Bar-Anand, 2012; Glasziou et al., 2014; Al-Shahi Salman et al., 2014).

Enfin, une formation adéquate et une éducation continue des scientifiques sur les méthodes de recherche et la maîtrise des concepts statistiques restent des éléments cruciaux (Collins et Tabak, 2014).

Parties prenantes

Lors de la conception, du test et de la mise en œuvre d'interventions sur les pratiques de recherche, il est essentiel de comprendre qui est concerné par la recherche et qui l'influence (Macleod et al., 2014; Khoury et al., 2012; Al-Shahi Salman et al., 2014). Les scientifiques ne représentent qu'un groupe au sein d'un réseau plus large (tableau 1), où différentes parties prenantes ont des attentes variées. Certaines valorisent la recherche pour son caractère publiable, finançable, applicable ou rentable. Cependant, ces attentes ne sont pas nécessairement alignées entre elles. Les scientifiques peuvent continuer à publier et à obtenir des subventions sans faire de réels progrès si la quantité de publications et de subventions est la seule chose qui compte. Si la science est principalement financée par des investissements privés visant brevets et profits, cela peut accélérer l'application et mener à des découvertes qui fonctionnent (ou semblent fonctionner), mais aussi créer des obstacles à la transparence et au partage d'informations. L'influence des entreprises peut détourner la science à des fins publicitaires, transformant les articles dans des revues influentes, les congrès prestigieux des sociétés savantes, et même un système professoral de meneurs d'opinion devenant des extensions du département de communication (Stamatakis et al., 2013; Krumholz et al., 2011). La géographie de la production scientifique évolue rapidement. Par exemple, il y aura bientôt plus d'articles en anglais provenant de Chine que d'Europe et des États-Unis réunis (Van Noorden, 2014). Les efforts de recherche sont enracinés dans des sociétés plus larges qui ont façonné des développements scientifiques variant selon les périodes historiques et

les contextes géographiques. Que peut-on faire pour renforcer la capacité de la science à prospérer, et pour évaluer et promouvoir cette capacité dans des cultures qui diffèrent par leur attitude envers le scepticisme, la curiosité et le raisonnement contradictoire ? Les différentes parties prenantes ont leurs propres préférences quant au moment où la reproductibilité devrait être encouragée ou évitée. Les équipes de l'industrie pharmaceutique ont soutenu la reproductibilité dans la recherche préclinique (Begley et Ellis, 2012; Prinz et al., 2011), car elles dépendent de la recherche universitaire préclinique pour identifier avec précision des cibles médicamenteuses utiles. À l'inverse, l'industrie se montre réticente à partager les données issues des essais cliniques (Peng, 2011), car à ce stade du développement d'un produit, de nouvelles analyses pourraient, à tort ou à raison (Christakis et Zimmerman, 2013), invalider les preuves soutenant des médicaments dans lesquels elle a déjà investi massivement.

Tableau 1 | Quelques principales parties prenantes de la science et leur degré d'intérêt pour la recherche et ses résultats selon diverses perspectives; des tendances typiques sont présentées (des exceptions existent)

	Degré d'intérêt pour les résultats de la recherche			
	Publiables	Finançables	Applicables	Rentables
Scientifiques	+++	+++	+	
Industrie - ventes et marketing				+++
Industrie - R & D			+++	+++
Investisseurs privés, y compris fonds spéculatifs			++	+++
Bailleurs publics – ouverts (par ex. NIH, NSF)	++		+	
Bailleurs publics – fermés (par ex. militaire)			+++	
Bailleurs à but non lucratif/philanthropes	++		+++	
Éditeurs de revues	+++			+
Éditeurs à but lucratif	+			+++
Sociétés professionnelles et scientifiques	+			
Universités	+	+++		+
Institutions de recherche à but non lucratif	+++	+++	+	+
Personnel non scientifique de soutien		+++		
Hôpitaux et autres installations professionnelles offrant des services liés à la science			+	+++
Autres entités financières affectées par ces services (par ex. assurances)				+++
Gouvernements et autorités étatiques/fédérales				++
Consommateurs de produits et services			+++	

doi:10.1371/journal.pmed.1001747.t001

Les dynamiques entre les différentes parties prenantes sont complexes. Par ailleurs, une même personne peut parfois jouer plusieurs rôles : des scientifiques peuvent également être rédacteurs ou rédactrices de revue, propriétaires d'une entreprise dérivée, responsables d'une société professionnelle, conseillères ou conseillers gouvernementaux et/ou bénéficiaires de l'industrie.

Devises de la recherche

Les publications et les subventions constituent des « devises » essentielles dans le domaine scientifique (tableau 2). Elles permettent « d'acheter » des « biens » académiques comme des promotions ou d'autres formes de pouvoir. Les titres académiques et le pouvoir renforcent encore davantage la « richesse » de la personne qui les détient. Le taux de change exact de ces devises et le prix des biens académiques (Young et al., 2008) peuvent varier selon les microenvironnements institutionnels, les disciplines scientifiques et les circonstances. Ils dépendent également du degré d'équité ou d'injustice dans chaque microenvironnement (par exemple, népotisme, favoritisme ou corruption). Le pouvoir administratif, le réseautage et le lobbying dans les universités, les sociétés savantes endogames et les académies déforment encore davantage ce tableau. Ce statu quo tend à favoriser ceux et celles qui excellent dans l'art de contourner le système, produisant une recherche prolifique mais médiocre et/ou non reproductible; contrôlant l'évaluation par les pairs dans les revues et les comités d'évaluation; prospérant dans une bureaucratie stérile, le lobbying et les manœuvres politiques; et faisant la promotion des personnes qui partagent leur manière de penser et d'agir.

Cependant, comprendre l'importance de ces devises clés crée aussi des opportunités. Par exemple, l'enregistrement des essais cliniques a été efficace parce que toutes les grandes revues scientifiques l'ont adopté comme condition préalable à la publication (Laine et al., 2007), faisant de cette pratique une devise de référence majeure dans la chaîne des récompenses. À l'inverse, des initiatives intéressantes de revue post-publication, comme PubMed Commons (Witten et Tibshirani, 2013), n'ont pas encore pleinement réalisé leur potentiel en tant qu'outils progressistes pour évaluer la recherche. Cela est probablement dû au fait qu'il n'existe actuellement aucune récompense associée à l'évaluation post-publication par les pairs.

Modification du système de récompense

Le système de récompense peut être modifié de manière systématique (Ioannidis et Khoury, 2014). Ces modifications peuvent aller de simples ajustements à des transformations innovantes. Le tableau 2 compare le statu quo (première colonne) à deux modifications potentielles du système de récompense, où le « Changement 2 » est plus radical que le « Changement 1 ».

Le système actuel valorise principalement les publications, les subventions, les titres académiques et le pouvoir accumulé. Les chercheurs et chercheuses occupant des postes supérieurs disposent généralement de plus de publications et de subventions. Cependant, les universitaires situés au sommet de la hiérarchie (par exemple, la haute direction des universités) affichent souvent des bilans modestes,

médiocres, voire négligeables en termes de publications et de citations (Ioannidis, 2010). Cela pourrait s'expliquer par leur habileté au lobbying, qui compense un manque de réalisations académiques, leur succès se faisant parfois au détriment de candidats plus méritants, capables d'apporter une rigueur intellectuelle accrue et une meilleure valeur stratégique aux décisions de haut niveau. À l'inverse, cela pourrait refléter leur excellence dans la gestion bureaucratique nécessaire pour maintenir en fonctionnement la gigantesque machine académique, leurs compétences permettant à des collègues scientifiquement plus talentueux de se consacrer pleinement à la recherche. Le système actuel ne récompense pas la répliation – il pénalise même souvent ceux et celles qui cherchent à reproduire rigoureusement des travaux antérieurs. Il encourage les scientifiques à présenter leurs travaux comme étant hautement novateurs et significatifs (Nosek et al., 2012). Le partage (de données, de protocoles, de codes d'analyse, etc.) n'est pas encouragé ni exigé, à quelques rares exceptions près (Hayden, 2014; Krumholz et al., 2014; *Data sharing will pay dividends*, 2014). En l'absence de ressources de soutien, et dans un climat de concurrence intense (« mes concurrents voleront mes données, mes idées, et finiront par obtenir mon financement »), le partage est encore plus dissuadé. D'autres aspects de la citoyenneté scientifique, comme la réalisation d'évaluations par les pairs de haute qualité, ne sont pas valorisés. L'évaluation par les pairs peut être un processus bénéfique, agissant comme un filet de sécurité et un mécanisme pour améliorer la qualité. Cependant, elle peut également être superficielle, apporter des améliorations limitées aux travaux évalués, et permettre l'acceptation d'articles manifestement erronés (Bohannon, 2013; Hopewell et al., 2014). Que ce processus soit si peu valorisé et récompensé limite ses avantages et n'aide pas à réduire ses inconvénients.

Tableau 2 | Illustration des différents taux de conversion pour diverses « devises » et éléments de richesse dans la recherche

	Différents exemples de systèmes de récompense		
	Actuel	Changement 1	Changement 2
DEVICES			
Publication (par unité)	Gain 1	Aucune valeur	Aucune valeur
Publication répliquée (par unité)	Gain 1	Gain 2	Gain 2
Publication appliquée avec succès (par unité)	Gain 1	Gain 5	Gain 5
Publication réfutée (par unité)	Gain 1	Perte 1	Perte 1
Partage de données, protocoles, codes d'analyse (par unité)	Aucune valeur	Gain 2	Gain 2
Contribution à l'évaluation par les pairs (par unité)	Aucune valeur	Gain 2	Gain 2
Contribution à l'éducation/formation (par unité)	Aucune valeur	Gain 1	Gain 1
Financement de projet (par unité R01)	Gain 5	Gain 5	Perte 5
AUTRES ÉLÉMENTS DE RICHESSE			
Poste de professeur adjoint dans une grande université	Gain 3	Gain 3	Aucune valeur

Poste de professeur agrégé dans une grande université	Gain 10	Gain 10	Aucune valeur
Poste de professeur titulaire dans une grande université	Gain 20	Gain 20	Aucune valeur
Chef d'équipe/direction			
Par doctorant/post-doc (par unité)	Gain 2	Gain 2	Perte 2
Pouvoir administratif, mise en réseau, lobbying	Gain jusqu'à 200	Aucune valeur	Perte jusqu'à 200

doi:10.1371/journal.pmed.1001747.t002

Les valeurs « monétaires » présentées dans le tableau 2 sont fournies à titre illustratif afin de susciter une réflexion sur les types de récompenses susceptibles d'influencer le processus du travail scientifique. Ces valeurs monétaires varient selon les microenvironnements, les champs disciplinaires et les contextes spécifiques. Par exemple, une valeur hypothétique de 1 pour une unité de publication (comme un article signé par une autrice ou un auteur principal ou senior dans une revue hautement respectée du domaine), de 5 pour une subvention de recherche importante (par exemple, un R01 aux États-Unis) et de 2 pour une chercheuse ou un chercheur postdoctoral, signifierait qu'on jugerait équivalent de publier cinq articles de ce type en tant qu'autrice ou auteur principal ou senior et d'obtenir une subvention R01 en tant que chercheuse ou chercheur principal, ou encore qu'on jugerait équivalent de publier deux articles similaires et de recruter une chercheuse ou un chercheur postdoctoral travaillant sous sa direction. Par ailleurs, ce qui constitue une unité de publication peut également varier selon les disciplines : dans les domaines où les publications sont rares, un seul article peut suffire à définir une unité de publication, tandis que dans les champs où il est courant que les scientifiques figurent comme co-auteurs et co-autrices sur des centaines d'articles, souvent marqués par un multi-autorat extrême, il peut falloir jusqu'à dix articles de ce type pour constituer une unité de publication équivalente. Des tendances inflationnistes, comme les publications redondantes ou « saucissonnées » (Schein et Paladugu, 2001) et le multi-autorat injustifié, ont entraîné une dévaluation relative de la « devise » des publications dans de nombreuses disciplines. Des ajustements pour le multi-autorat sont facilement réalisables (Hagen, 2008; Aziz et Rozing, 2013). Une meilleure connaissance des contributions individuelles à chaque article permettrait une attribution encore plus précise des crédits (Yank et Rennie, 1999).

Dans le premier exemple de modification proposée du système de récompense présenté dans le tableau 2, le pouvoir d'achat des publications est principalement différencié en fonction de leur statut de répliation et d'applicabilité. La valeur est accordée aux idées et résultats solides, répliqués et reproductibles (Wagenmakers et Forstman, 2014), plutôt qu'à la simple publication en tant que telle. Une valeur supplémentaire est attribuée aux publications débouchant sur des applications concrètes telles que des traitements efficaces, des tests diagnostiques ou des outils pronostiques qui améliorent manifestement d'importantes issues dans les essais cliniques. Une valeur additionnelle est également accordée au partage des connaissances et à une participation significative à l'évaluation par les pairs et aux activités éducatives ayant prouvé leur efficacité. Une personne en charge de l'évaluation ou de

l'édition peut, à certaines occasions, apporter une contribution équivalente à celle d'un auteur ou d'une autrice.

Le deuxième exemple de modification proposé, présenté dans le tableau 2, introduit des changements encore plus radicaux dans le système de récompenses. En plus des ajustements adoptés dans le premier exemple, l'obtention de subventions, de distinctions ou d'autres formes de pouvoir est perçue négativement, à moins qu'un chercheur ou une chercheuse ne produise une science de haute qualité en proportion. Les ressources et le pouvoir sont considérés comme des opportunités, et les chercheurs et chercheuses doivent aligner leur production sur les opportunités qui leur ont été offertes : plus les opportunités sont nombreuses, plus la production attendue (répliquée et, idéalement, appliquée) doit être élevée. Dans ce modèle, les grades académiques n'ont aucune valeur et pourraient même être supprimés : les scientifiques doivent simplement maintenir un équilibre non négatif entre leur production et les opportunités qu'ils ont reçues. Dans ce scénario délibérément provocateur, les chercheurs et chercheuses hésiteraient à demander des subventions ou à occuper des positions de pouvoir (au sens actuel), car cela serait perçu comme une contrainte. Les effets secondaires potentiels pourraient inclure de dissuader des candidatures à des subventions ambitieuses et une réduction des aspirations à des rôles de leadership.

De tels compromis soulignent qu'en matière de modification de la structure des carrières scientifiques, tout comme lorsqu'on modifie la physiopathologie pour tenter de lutter contre une maladie, les interventions peuvent engendrer des effets aussi néfastes que bénéfiques. Compte tenu de la complexité de la situation, il est crucial d'évaluer les impacts réels des interventions de manière équitable et fiable.

Perspectives

L'ampleur des améliorations possibles dans l'efficacité actuelle des pratiques de recherche reste inconnue. Cependant, compte tenu des inefficacités massives existantes, des améliorations substantielles sont presque certainement réalisables. Perfectionner des politiques en place ainsi que mener des interventions plus perturbatrices et radicales devrait être envisagé, mais ni la présence ni l'absence d'intentions révolutionnaires ne doivent être considérées comme un substitut fiable de l'impact réel. De nombreux scénarios différents peuvent être envisagés pour l'évolution de la recherche biomédicale et de l'investigation scientifique en général, chacun étant plus ou moins aligné sur la quête de la vérité et du bien-être humain. Les interventions visant à modifier le système actuel ne devraient pas être acceptées sans un examen rigoureux, même lorsqu'elles semblent raisonnables et bien intentionnées. Idéalement, ces interventions devraient faire l'objet d'une évaluation expérimentale. Les réalisations de la science sont remarquables, mais la majorité des efforts de recherche sont actuellement gaspillés. Réduire ce gaspillage et rendre la science plus efficace pourrait offrir des bénéfices considérables pour notre santé, notre confort et notre compréhension de la vérité, tout en aidant la recherche scientifique à atteindre plus efficacement ses nobles objectifs.

AFFILIATION DE L'AUTEUR

John P. A. Ioannidis, MD, DSc

Meta-Research Innovation Center at Stanford (METRICS) | Stanford University, Stanford, Californie

Département de Médecine | Stanford Prevention Research Center, Stanford, Californie

Département de la recherche et des politiques en santé | Stanford University School of Medicine, Stanford, Californie

Département de Statistique | Stanford University School of Humanities and Sciences, Stanford, Californie

FINANCEMENT (Article original)

Le Meta-Research Innovation Center de Stanford bénéficie du financement de la Fondation Laura et John Arnold. Les financeurs n'ont joué aucun rôle dans la décision de publication ni dans la préparation du manuscrit.

CONFLITS D'INTÉRÊTS (Article original)

JPAI est membre du comité de rédaction de PLOS Medicine. L'auteur a déclaré n'avoir aucun conflit d'intérêts.

TRADUCTION

L'article a été traduit en suivant ces étapes : 1) Utilisation de la traduction automatisée avec le logiciel DeepL Pro; 2) Amélioration de la traduction automatisée à l'aide d'un modèle GPT développé par Mathieu Ouimet; 3) Révision humaine par Mathieu Ouimet, Julie Desnoyers et Aurélie Hot.

RÉFÉRENCES

Al-Shahi Salman, R., Beller, E., Kagan, J., Hemminki, E., Phillips, R.S. et al. (2014). Increasing value and reducing waste in biomedical research regulation and management. *Lancet*, 383, 176–185.

Aziz, N.A. et Rozing, M.P. (2013). Profit (p)-index: the degree to which authors profit from co-authors. *PLoS ONE*, 8, e59814.

Begley, C.G. et Ellis, L.M. (2012). Drug development: Raise standards for preclinical cancer research. *Nature*, 483, 531–533.

Bissell, M. (2013). Reproducibility: The risks of the replication drive. *Nature*, 503, 333–334.

Bohannon, J. (2013). Who's afraid of peer review? *Science*, 342, 60–65.

- Boyack, K.W., Klavans, R., Sorensen, A.A. et Ioannidis, J.P. (2013). A list of highly influential biomedical researchers, 1996–2011. *Eur J Clin Invest*, 43, 1339–1365.
- Chalmers, I., Bracken, M.B., Djulbegovic, B., Garattini, S., Grant, J., Gülmezoglu, A. M., Howells, D.W., Ioannidis, J.P.A. et Oliver, S. (2014). How to increase value and reduce waste when research priorities are set. *Lancet*, 383, 156–165.
- Chan, A.W., Song, F., Vickers, A., Jefferson, T., Dickersin, K. et al. (2014) Increasing value and reducing waste: addressing inaccessible research. *Lancet*, 383, 257–266.
- Christakis, D.A. et Zimmerman, F.J. (2013). Rethinking reanalysis. *JAMA*, 310, 2499–2500.
- Collins, F.S. et Tabak, L.A. (2014). NIH plans to enhance reproducibility. *Nature*, 505, 612–613.
- Contopoulos-Ioannidis, D.G., Alexiou, G.A., Gouvas, T.C. et Ioannidis, J.P. (2008). Life cycle of translational research for medical interventions. *Science*, 321, 1298–1299.
- Dal-Re, R., Ioannidis, J.P., Bracken, M.B., Buffler, P.A., Chan, A.W., Franco, E.L., La Vecchia, C. et Weiderpass, E. (2014). Making prospective registration of observational research a reality. *Sci Transl Med*, 6, 224cm1.
- Danthi, N., Wu, C.O., Shi, P. et Lauer, M. (2014). Percentile ranking and citation impact of a large cohort of national heart, lung, and blood institute-funded cardiovascular R01 grants. *Circ Res*, 114, 600–606.
- Data sharing will pay dividends.* (2014). *Nature*, 505, 131.
- Donoho, D.L. (2010). An invitation to reproducible computational research. *Biostatistics*, 11, 385–388.
- Doshi, P., Goodman, S.N. et Ioannidis, J.P. (2013). Raw data from clinical trials: within reach? *Trends Pharmacol Sci*, 34, 645–647.
- Dwan, K., Gamble, C., Williamson, P.R., Kirkham, J.J. et Reporting Bias Group (2013). Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. *PLoS ONE*, 8, e66844.
- Fanelli, D. (2010). "Positive" results increase down the Hierarchy of the Sciences. *PLoS ONE*, 5, e10068.
- Glasziou, P., Altman, D.G., Bossuyt, P., Boutron, I., Clarke, M. et al. (2014). Reducing waste from incomplete or unusable reports of biomedical research. *Lancet*, 383, 267–276.
- Gøtzsche, P.C. et Ioannidis, J.P. (2012). Content area experts as authors: helpful or harmful for systematic reviews and meta-analyses? *BMJ*, 345, e7031.
- Greenhalgh, T., Howick, J., Maskrey, N. et Evidence Based Medicine Renaissance Group (2014). Evidence based medicine: a movement in crisis? *BMJ*, 348, g3725.

- Hagen, N.T. (2008). Harmonic allocation of authorship credit: source-level correction of bibliometric bias assures accurate publication and citation analysis. *PLoS ONE*, 3, e4021.
- Hayden, E.C. (2014). Cancer-gene data sharing boosted. *Nature*, 510, 198.
- Hopewell, S., Collins, G.S., Boutron, I., Yu, L.M., Cook, J. et al. (2014). Impact of peer review on reports of randomised trials published in open peer review journals: retrospective before and after study. *BMJ*, 349, g4145.
- Institute of Medicine (2011). *Clinical Practice Guidelines We Can Trust*. National Academies Press.
- Ioannidis, J.P. (2005). Why most published research findings are false. *PLoS Med*, 2, e124.
- Ioannidis, J.P. (2008). Why most discovered true associations are inflated. *Epidemiology*, 19, 640–648.
- Ioannidis, J.P. (2010). Is there a glass ceiling for highly cited scientists at the top of research universities? *FASEB J*, 24, 4635–4638.
- Ioannidis, J.P. (2011). More time for research: fund people not projects. *Nature*, 477, 529–531.
- Ioannidis, J.P., Tarone, R., McLaughlin, J.K. (2011). The false-positive to false-negative ratio in epidemiologic studies. *Epidemiology*, 22, 450–456.
- Ioannidis, J.P. et Doucouliagos, C. (2013). What's to know about the credibility of empirical economics. *J Economic Surveys*, 27, 997–1004.
- Ioannidis, J.P. et Khoury, M.J. (2014). Assessing value in biomedical research: The PQRST of appraisal and reward. *JAMA*, 312, 483–484. doi:10.1001/jama.2014.6932
- Johnson, V.E. (2013). Revised standards for statistical evidence. *Proc Natl Acad Sci U.S.A.*, 110, 19313–19317.
- Jørgensen, A.W., Hilden, J. et Gøtzsche, P.C. (2006). Cochrane reviews compared with industry-supported meta-analyses and other meta-analyses of the same drugs: systematic review. *BMJ*, 333, 782.
- Kassirer, J.P. et Angell, M. (1994). The journal's policy on cost-effectiveness analyses. *N Engl J Med*, 331, 669–670.
- Khoury, M.J., Gwinn, M., Dotson, W.D. et Schully, S.D. (2012). Knowledge integration at the center of genomic medicine. *Genet Med*, 14, 643–647.
- Khoury, M.J., Lam, T.K., Ioannidis, J.P., Hartge, P., Spitz, M.R. et al. (2013). Transforming epidemiology for 21st century medicine and public health. *Cancer Epidemiol Biomarkers Prev*, 22, 508–516.
- Krumholz, S.D., Egilman, D.S. et Ross, J.S. (2011). Study of Neurontin: titrate to effect, profile of safety (STEPS) trial. A narrative account of a gabapentin seeding trial. *Arch Intern Med*, 171, 1100–1107.

- Krumholz, H.M., Gross, C.P., Blount, K.L., Ritchie, J.D., Hodshon, B. et al. (2014). Sea change in open science and data sharing: leadership by industry. *Circ Cardiovasc Qual Outcomes*, 7, 499–504.
- Laine, C., Horton, R., DeAngelis, C.D., Drazen, J.M., Frizelle, F.A. et al. (2007). Clinical trial registration: looking back and moving ahead. *JAMA*, 298, 93–94.
- Landis, S.C., Amara, S.G., Asadullah, K., Austin, C.P., Blumenstein, R. et al. (2012). A call for transparent reporting to optimize the predictive value of preclinical research. *Nature*, 490, 187–191.
- Macleod, M. (2011). Why animal research needs to improve. *Nature*, 477, 511.
- Macleod, M.R., Michie, S., Roberts, I., Dirnagl, U., Chalmers, I. et al. (2014). Biomedical research: increasing value, reducing waste. *Lancet* 383: 101–104.
- Montfortin, C. (2006). Weight of the evidence or wait for the evidence? Protecting underground miners from diesel particulate matter. *Am J Public Health*, 96, 271–276.
- Mynatta, C.R., Doherty, M.E. et Tweneya, R.D. (1977). Confirmation bias in a simulated research environment: An experimental study of scientific inference. *Quarterly J Exp Psychol*, 29, 85–95.
- NCI-NHGRI Working Group on Replication in Association Studies (2007). Replicating genotype-phenotype associations. *Nature*, 447(7145), 655–660.
- Nicholson, J.M. et Ioannidis, J.P.A. (2012). Research grants: Conform and be funded. *Nature*, 492, 34–36.
- Nickerson, R.S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Rev Gen Psychol*, 2, 175–220.
- Nosek, B.A. et Bar-Anand, Y. (2012). Scientific utopia: I. Opening scientific communication. *Psychological Inquiry*, 23, 217–223.
- Nosek, B.A., Spies, J.R., Motyl, M. (2012). Scientific Utopia: II. Restructuring incentives and practices to promote truth over publishability. *Persp Psychological Sci*, 7, 615–631.
- Nuzzo, R. (2014). Scientific method: statistical errors. *Nature*, 506, 150–152.
- Panagiotou, O.A., Willer, C.J., Hirschhorn, J.N. et Ioannidis, J.P. (2013). The power of meta-analysis in genome-wide association studies. *Annu Rev Genomics Hum Genet*, 14, 441–465.
- Pashler, H. et Harris, C.R. (2012). Is the replicability crisis overblown? Three arguments examined. *Persp Psychol Sci*, 7, 531–536.
- Peng, R.D., Dominici, F. et Zeger, S.L. (2006). Reproducible epidemiologic research. *Am J Epidemiol*, 163, 783–789.
- Peng, R.D. (2011). Reproducible research in computational science. *Science*, 334, 1226–1227.

- Poste, G. (2012). Biospecimens, biomarkers, and burgeoning data: the imperative for more rigorous research standards. *Trends Mol Med*, 18, 717–722.
- Prinz, F., Schlange, T. et Asadullah, K. (2011). Believe it or not: how much can we rely on published data on potential drug targets? *Nat Rev Drug Discov*, 10, 712.
- Rennie, D. et Flanagan, A. (2014). Research on peer review and biomedical publication: furthering the quest to improve the quality of reporting. *JAMA*, 311, 1019–1020.
- Schein, M., Paladugu, R. (2001). Redundant surgical publications: tip of the iceberg? *Surgery*, 129, 655–661.
- Simera, I., Moher, D., Hoey, J., Schulz, K.F. et Altman, D.G. (2010). A catalogue of reporting guidelines for health research. *Eur J Clin Invest*, 40, 35–53.
- Simmons, J.P., Nelson, L.D. et Simonsohn, U. (2011). False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol Sci*, 22, 1359–1366.
- Siontis, K.C., Hernandez-Boussard, T. et Ioannidis, J.P. (2013). Overlapping meta-analyses on the same topic: survey of published studies. *BMJ*, 347, f4501.
- Stamatakis, E., Weiler, R. et Ioannidis, J.P. (2013). Undue industry influences that distort healthcare research, strategy, expenditure and practice: a review. *Eur J Clin Invest*, 43, 469–475.
- Stodden, V., Guo, P. et Ma, Z. (2013). Toward reproducible computational research: an empirical analysis of data and code policy adoption by journals. *PLoS ONE*, 8, e67111.
- Van Noorden, R. (2014). China tops Europe in R&D intensity. *Nature*, 505, 144–145.
- Wagenmakers, E.J. et Forstman, B.U. (2014). Rewarding high-power replication research. *Cortex*, 51, 105–106.
- Wenneras, C. et Wold, A. (1997). Nepotism and sexism in peer-review. *Nature*, 387, 341–343.
- Witten, D.M. et Tibshirani, R. (2013). Scientific research in the age of omics: the good, the bad, and the sloppy. *J Am Med Inform Assoc*, 20, 125–127.
- Yank, V., Rennie, D. (1999). Disclosure of researcher contributions: a study of original research articles in The Lancet. *Ann Intern Med*, 130, 661–670.
- Young, N.S., Ioannidis, J.P. et Al-Ubaydli, O. (2008). Why current publication practices may distort science. *PLoS Med*, 5, e201.
- Young, S.S. et Karr, A. (2011). Deming, data, and observational studies: a process out of control and needing fixing. *Significance*, 8, 116–120.
- Zarin, D.A., Ide, N.C., Tse, T., Harlan, W.R., West, J. et Lindberg, D.A.B. (2007). Issues in the registration of clinical trials. *JAMA*, 297, 2112–2120.

Zarin, D.A., Tse, T., Williams, R.J., Califf, R.M. et Ide, N.C. (2011). The ClinicalTrials.gov results database—update and key issues. *N Engl J Med*, 364, 852–860.

CITATION SUGGÉRÉE

Ioannidis, J. P. A. (2024). Comment assurer la véracité d'un plus grand nombre de recherches publiées (traduction). *Revue sur le transfert et l'utilisation des connaissances*, 9(1).
<https://doi.org/10.18166/tuc.2025.9.1.53> (Article original publié en 2014)



ISSN | 2369-8896

www.revue-tuc.ca



Cet article est mis à disposition selon les termes de la Licence Creative Commons
Attribution - Partage dans les Mêmes Conditions 4.0 International