

## Ausschuss, mangelnde Reproduzierbarkeit und Betrug in der Wissenschaft

Der Zeitgeist des Überangebots mit all seinen negativen Auswirkungen auf unser Denken, Verhalten und die Umwelt macht auch vor der medizinischen Wissenschaft nicht halt. In der Pubmed-Datenbank werden derzeit rund 38 Mio. Publikationen gelistet. Allein im Jahr 2024 sind rund 1,6 Mio. hinzugekommen <sup>[1]</sup>. Täglich erscheinen > 4.000 Originalarbeiten, Reviews, Letters, Kommentare und Korrekturen in rund 30.000 biomedizinischen Journalen, darunter eine wachsende Zahl von Verlagen mit Open Access-Policy (OA). Nicht wenige dieser Journale sind „Raubtiere“ („Predatory Journals“), die qualitativ minderwertig sind und fast alles gegen Geld abdrucken (vgl. <sup>[2]</sup>). Viele Artikel stammen aus sog. „Paper mills“, d.h. kommerziellen Schreibbüros und haben nicht selten fiktive oder rein nominelle Autoren („Honorary Authorship“).

Es liegt auf der Hand, dass sich unter diesem Berg an Veröffentlichungen auch viel Unwichtiges und Uninteressantes befindet, also Ausschuss. Etwa 15-20% der in Pubmed gelisteten Publikationen werden niemals und 30-40% nur 1-5mal zitiert <sup>[3]</sup>. Sehr beunruhigend ist, dass viele Ergebnisse aus Originalarbeiten gar nicht oder nur teilweise reproduziert werden können. Das Center for Open Science (COS) ist eine gemeinnützige Organisation, die sich für Transparenz, Reproduzierbarkeit und offene Wissenschaft einsetzt. Im Rahmen des sog. „Reproducibility Project“ wurden 193 Experimente aus der Krebsbiologie aus 53 Publikationen in hochrangigen Journalen der Jahre 2010-2012 überprüft <sup>[4]</sup>. Nur 50 dieser Experimente konnten nochmals in gleicher Weise durchgeführt werden. Bei den übrigen war das veröffentlichte Protokoll unvollständig oder die Autoren antworteten nicht oder nur unzureichend auf Nachfragen zu ihrer Methodik. Wurden die Experimente wiederholt, konnten die Ergebnisse nur bei knapp der Hälfte reproduziert werden, und bei 85% waren die Effektstärken kleiner als in den Veröffentlichungen.

Vor diesem Hintergrund und mit den stetig wachsenden Möglichkeiten der sog. „digitalen Forensik“ wundert es nicht, dass die Zahl zurückgezogener Artikel („Retractions“) seit Jahren stark zunimmt: von 38 im Jahr 2000 auf zuletzt > 10.000 pro Jahr <sup>[5]</sup>. Wahrscheinlich müssten noch sehr viel mehr Artikel zurückgezogen oder zumindest mit Korrekturen versehen werden. Dem stehen jedoch psychologische, praktische und wahrscheinlich auch rechtliche Hürden im Wege.

„Retraction Watch“ ist ein durch Spenden und öffentliche Fördergelder finanziertes journalistisches Projekt, das sich auf die Dokumentation und Analyse zurückzogener wissenschaftlicher Publikationen spezialisiert hat. Auf deren Webseite finden sich bereits > 50.000 Einträge über alle Wissenschaftsdisziplinen hinweg. Unter den zurückgezogenen Artikeln sind auch solche, die > 1.000mal zitiert wurden, von Nobelpreisträgern stammen und aus sehr renommierten Zeitschriften <sup>[6]</sup>. Die Mehrzahl der „Retractions“ findet sich aber in Zeitschriften mit niedrigem „Impact factor“.

Nach Analysen von „Retraction Watch“ gibt es hauptsächlich drei Gründe, warum Artikel zurückgezogen werden:

1. *Fehler*: z.B. unbeabsichtigte („Honest Error“) oder methodisch unzuverlässige oder nicht nachvollziehbare Daten (Rechenfehler, fehlerhafte Statistiken etc.);
2. *Wissenschaftliches Fehlverhalten*: B.: Manipulation und Fälschung von Daten („Fraud“: Erfindungen, Verfälschungen und Bildmanipulationen); fehlendes Ethikvotum oder Verstoß dagegen; nachträgliche Veränderungen des Protokolls oder der Endpunkte; selektive Berichterstattung; Mehrfach- (duplication) und Salami-Veröffentlichungen („slicing“); fehlerhaftes Zitieren (Plagiarismus) und Urheberrechtsverletzungen; „Ghostwriting“ und „Honorary Authorships“; Verschweigen von finanziellen oder immateriellen Interessenkonflikten bei den Autoren oder die vollständige Prozesssteuerung durch eine Firma oder ein pharmazeutisches Unternehmen („Ghost Management“);
3. *Fehlverhalten beim Herausgeber*: B.: nicht deklarierte Interessenkonflikte; Bevorzugung bestimmter Themen, Institutionen oder Autoren; manipulierte oder komplett gefälschte Peer Reviews; Beeinflussung der Gutachter; Manipulation bei den Zitierungen („Coercive Citation“).

Wissenschaftliches Fehlverhalten ist derzeit für etwa 60% aller Rücknahmen verantwortlich, Fehler für 17% und Fehlverhalten beim Herausgeber für 9%. Bei den übrigen werden die Gründe nicht eindeutig dargelegt <sup>[7]</sup>. Eine ältere Analyse von 228 „Retractions“ wegen wissenschaftlichen Fehlverhaltens ergab, dass die beteiligten Autoren überproportional häufig Männer waren und dass das gesamte Spektrum vertreten ist, vom Praktikanten bis zum leitenden Wissenschaftler <sup>[8]</sup>.

Etwa 2% der Wissenschaftler räumen in anonymen Antworten auf Umfragen ein, mindestens schon einmal Daten oder Ergebnisse frei erfunden, gefälscht oder verändert zu haben <sup>[9]</sup>. Es werden auch regelmäßig Fälle von hochproduktiven und respektierten Autoren bekannt, die über Jahre systematisch Daten fälschen. Bemerkenswert ist, dass die Kollegen im Umfeld und die oft zahlreichen Koautoren den Betrug nicht erkennen oder nicht verhindern konnten – oder wollten. Die Betrügereien werden nur selten als Systemversagen, sondern meist als Taten eines Einzelnen missverstanden, und die, die den Betrug aufgedeckt haben, oft eingeschüchtert und diffamiert.

Die Ursachen von wissenschaftlichem Fehlverhalten sind vielfältig. Nach Umfragen unter chinesischen Forschenden sind entscheidend: hoher Publikationsdruck („publish or perish“), Karriereaspekte (wachsende Konkurrenz und Globalisierung der Forschung) und moralische Mängel <sup>[10]</sup>. Als „Währung“ der Wissenschaft und als Maß von Expertise gilt nach wie vor die Zahl der Publikationen und die Häufigkeit deren Zitierung in anderen Publikationen (Impact- und H-Faktor). Dabei sind diese Parameter nur mäßig aussagekräftig für die translationale Wirkung von Forschung <sup>[11]</sup>. Andere, komplexere Modelle, die Themen, Abstracts und Metadaten einschließen, sind deutlich genauer (sog. inhaltsbasierte Impact-Modelle). In der „San Francisco Declaration on Research Assessment“ (DORA) wurden 2012 entsprechende Vorschläge unterbreitet <sup>[12]</sup>, aber leider haben sich diese noch nicht durchgesetzt.

Alle sollten ein großes Interesse daran haben, dass wissenschaftliches und publizistisches Fehlverhalten eingedämmt wird. Doch wie kann die Situation verbessert werden?

Die Verantwortlichen an den Universitäten und Forschungseinrichtungen müssen dafür sorgen, dass ihre Forscher eine Ausbildung in guter wissenschaftlicher Praxis („Good Scientific Practice“ = GSP) erhalten und diese Standards auch durchgesetzt werden. Dazu zählen die Ertüchtigung interner Strukturen wie Mentoring, Fehler- und Konfliktmanagement, Ombudsstellen und ein System, in dem beobachtete oder vermutete Fehler ohne persönliche Nachteile angezeigt und überprüft werden können. Außerdem müssen Compliance- und Ethikboards präsent sein und die Möglichkeit erhalten, Stichproben zu machen.

Unabhängige und hierzu befähigte externe Institutionen wie Ethikkommissionen, Forschungsgesellschaften und Kammern müssen die Einhaltung der GSP überprüfen und Verstöße sanktionieren. Veröffentlichte Forschungsdaten sollten stichprobenartig überprüft werden. Dazu ist es erforderlich, die angewendete Methodik und die Daten vollständig offenzulegen (Transparenz). Auch automatisierte Überprüfungen von Publikationen sollten ausgeweitet werden.

Wissenschaftler sollten nur noch in Journalen veröffentlichen, die bestimmte Standards einhalten, wie z.B. die von COPE (Committee on Publication Ethics; <sup>[13]</sup>). Diese Verlage müssen die Transparenz der Daten gewährleisten und das Peer-Review-System ertüchtigen (Honorierung, Beschleunigung). Korrekturen und Retraktionen müssen da, wo Fehler erkannt worden sind, als wissenschaftlich notwendiger Bestandteil des Diskurses verstanden werden.

## Fazit

Das System der wissenschaftlichen Publikationen befindet sich in einer Krise. Zurzeit gilt das Prinzip Masse statt Klasse. Die Reproduzierbarkeit von Forschungsergebnissen ist mangelhaft, und die Fälle mit wissenschaftlichem und publizistischem Fehlverhalten nehmen erheblich zu. Die Zahl der Publikationen und Zitate (Impact- und H-Faktor) als Maß für die translationale Wirkung von Forschung und als Währung für Karrieren in der Medizin hat ausgedient. Das System muss dringend reformiert werden.

## Literatur

1. Sayers, E.W., et al.: Nucleic Acids Res. 2025, 53, D20. ([Link zur Quelle](#))
2. AMB 2018, 52, 64DB01. ([Link zur Quelle](#))
3. Meho, L.I.: Physics World 2007, 20, 32. ([Link zur Quelle](#))
4. <https://www.cos.io/rpcb> ([Link zur Quelle](#))
5. Jaime, A., et al.: Learned Publishing 2025, 38, e1667. ([Link zur Quelle](#))
6. <https://retractionwatch.com> ([Link zur Quelle](#))
7. Hwang, S.Y., et al.: J. Korean Med. Sci. 2023, 38, e333. ([Link zur Quelle](#))
8. Fang, F.C., et al.: mBio 2013, 4, e00640. ([Link zur Quelle](#))
9. Fanelli, D.: PLoS One 2009, 4, e5738. ([Link zur Quelle](#))
10. Yu, L, et al.: Account. Res. 2021, 28, 95. ([Link zur Quelle](#))
11. Nelson, A.P.K., et al.: Patterns 2022, 3, 100483. ([Link zur Quelle](#))
12. <https://sfdora.org/read/> ([Link zur Quelle](#))

13. <https://publicationethics.org/members> ([Link zur Quelle](#))