



Einsichten eines Wissenschaftsnarren (66)

PubPeer – Forum für persönliche Vendettas oder Zukunft des Peer Review?

Der Wissenschaftsbetrieb hat zu wenig Fehlerkultur, um die wissenschaftliche Literatur adäquat zu korrigieren. Deshalb ist PubPeer derzeit das beste Medium dafür.

Der deutsch-amerikanische Hirnforscher und Nobelpreisträger 2013 Thomas Südhof macht derzeit Schlagzeilen. Südhof forscht an der Stanford University und ist Berlin unter anderem durch eine langjährige Fellowship der Charité verbunden. In den Zeitungen war er vor kurzem allerdings nicht durch Entdeckungen auf seinem Fachgebiet – der komplexen

Interaktion von Nervenzellen –, sondern vielmehr durch Kommentare auf der Paper-Diskussionsplattform PubPeer. Dort wird ihm bei einer Vielzahl von Artikeln aus seiner Arbeitsgruppe Schlampigkeit, wenn nicht sogar Datenmanipulation vorgeworfen. In der Folge hat er bei einigen der dort kritisierten Artikel Korrekturen veröffentlichen lassen, einer davon wurde mittlerweile zurückgezogen. Dass nun Südhof im Rampenlicht steht, ist in mehrfacher Hinsicht bemerkenswert: Als Nobelpreisträger ist Südhof ein wissenschaftlicher Leuchtturm, seit Jahren ist er aber auch einer der prominentesten Kritiker des kommerziellen akademischen Publikationswesens und ein Streiter für eine robustere, reproduzierbarere Forschung.

Wir erinnern uns: Erst kürzlich war die deutsche Krebsforscherin Simone Fulda in die Schlagzeilen geraten, nachdem ihr ebenfalls auf PubPeer Datenmanipulation in einer Reihe ihrer Veröffentlichungen vorgeworfen wurde. Sie legte daraufhin im Februar ihr Amt als Präsidentin der Universität Kiel nieder, der Fall befindet sich bei der Deutschen Forschungsgemeinschaft (DFG) noch im Stadium der Vorprüfung. Die Liste lässt sich fast beliebig fortsetzen; vor Simone Fulda war zum Beispiel der Neurowissenschaftler und Präsident der Stanford University, Marc Tessier-Lavigne, auf den internationalen Titelseiten. Auch er trat zurück, nachdem – ebenfalls via PubPeer – offensichtliche Manipulationen in Abbildungen aus seiner Arbeitsgruppe aufgedeckt worden waren.

Wieso spielt PubPeer bei all diesen Vorgängen eine derart zentrale Rolle? PubPeer ermöglicht jeder und jedem, wissenschaftliche Publikationen zu kommentieren. Die Plattform ermöglicht folglich einen Post-Publication-Review, also eine Begutachtung nach Publikation der Ergebnisse. Diese werden hier also ein zweites Mal begutachtet, da sie ja bereits vor der Veröffentlichung im eigentlichen Peer-Review-Prozess begutachtet wurden. Bei Letzterem evaluieren in der Regel zwei, manchmal auch bis zu vier Fachexperten („Peers“) das Manuskript. Wegen der hochgradigen Komple-

xität der Fragestellungen, der speziellen Methodik sowie der Analysen und des oft massiven Umfangs der Manuskripte kann dieser Prozess, wenn überhaupt, nur die allergrößten Fehler und Probleme identifizieren. Von einer wirklichen Qualitätskontrolle kann deshalb nicht mehr gesprochen werden.

»Resultat eines methodischen Artefaktes oder eines „ehrlichen“ Copy-Paste-Fehlers?«

Hier setzt das beim wissenschaftlichen Publizieren bisher kaum eingesetzte Prinzip des Post-Publication-Review an: Es erlaubt der gesamten Leserschaft die Begutachtung und Diskussion von bereits veröffentlichten Artikeln – also in der Regel einer großen Anzahl von Expertinnen und Experten. Die PubPeer-Plattform, die von einer US-amerikanischen, nichtkommerziellen Stiftung betrieben wird, ermöglicht nun die Kommentierung beliebiger wissenschaftlicher Publikationen, unabhängig davon, in welchem Journal sie veröffentlicht wurden. Damit treten die Leser in einen Diskurs mit den Autoren, die auf die dort geposteten Kommentare reagieren können. Um durch Hinweise auf Fehler in den Arbeiten oder Datenmanipulationen nicht Sanktionen durch mächtige Autoren fürchten zu müssen, kann auch anonym kommentiert werden. Die Kommentare und die Reaktionen der Autoren darauf werden von den Betreibern der Plattform nur moderiert, wenn sie nicht den Anforderungen der „Commenting Guidelines“ entsprechen.

Häufig werden auf PubPeer, wie auch bei Südhof, Fulda oder Tessier-Lavigne geschehen, Abbildungen von biochemischen, molekularbiologischen oder histologischen Versuchsergebnissen diskutiert beziehungsweise in Frage gestellt. Mit großem Abstand werden hier am häufigsten Duplikationen, Verschleppungen und andere Auffälligkeiten in fluoreszenzmikroskopischen Abbildungen oder Wes-



Foto: BIH/Thomas Rafalzyk

Ulrich Dirnagl

ist experimenteller Neurologe an der Berliner Charité und Gründungsdirektor des QUEST Center for Responsible Research am Berlin Institute of Health. Für seine Kolumne schlüpft er in die Rolle eines „Wissenschaftsnarren“ – um mit Lust und Laune dem Forschungsbetrieb so manche Nase zu drehen.

tern Blots moniert. Das liegt ganz einfach daran, dass es sich hier um experimentelle Standardtechniken handelt, die jeder Experte auch ohne Fachwissen auf dem Gebiet des Artikels, ja sogar ohne weitere Kenntnis des sonstigen Inhalts einer Studie gut visuell inspizieren und interpretieren kann. Es liegt aber auch daran, dass das menschliche Gehirn – mittlerweile oft unterstützt durch künstliche Intelligenz (KI) – sehr gut Muster erkennen kann. Dem geschulten Auge oder der KI gelingt es so, Abbildungen zu identifizieren, in denen Teile verdoppelt, gespiegelt oder verzerrt wurden.

In solchen Fällen steht dann meist die Frage im Raum, ob der beanstandete Sachverhalt das Resultat eines methodischen Artefaktes oder eines „ehrlichen“ Copy-Paste-Fehlers bei der Erstellung der Abbildung oder deren Beschriftung war. Oder ob er eben in betrügerischer Absicht erfolgt ist. Etwa um Resultate vorzuspiegeln, die sich im Experiment so nicht einstellten (aber für die „Story“ relevant waren); oder um die Studienergebnisse „spektakulärer“ erscheinen zu lassen, als sie waren, und damit eine prestigeträchtige Publikation wahrscheinlicher zu machen; oder einfach nur als „Abkürzung“ nach dem Motto

„Ähnliche Kontrollen haben wir doch schon in anderen Experimenten gemacht – lasst uns also die wieder nehmen, jetzt wäre ja auch nichts anderes rausgekommen.“

»Als anonyme Plattform eignet sich PubPeer trefflich für private Vendettas und Kampagnen.«

Tatsächlich wurde durch Diskussionen auf PubPeer bereits eine Vielzahl von Fehlern in wissenschaftlichen Publikationen verschiedener Disziplinen aufgedeckt. Durch Korrektur der Veröffentlichung oder im Extremfall Zurückziehen der Arbeit wurde der Wissenschaft so ein großer Dienst erwiesen. Viele dieser Fehler waren „ehrlich“. Die Wissenschaft, ihre Methoden und Analysen sind komplex und kompliziert. Dass Fehler gemacht werden, ist nachgerade notwendig und normal. Leider gibt es aber in den meisten wissenschaftlichen Disziplinen keine entwickelte Fehlerkultur (siehe hierzu *LJ* 1/2019: 22-23). Fehler werden aus Angst vor Reputationsverlust oder Sanktio-

nen lieber unter den Tisch gekehrt, als diese zu kommunizieren, zu analysieren und damit eine Wiederholung unwahrscheinlicher zu machen.

Und hier liegt leider eine Schattenseite von PubPeer: Paradoxerweise verstärkt die Möglichkeit, dort mit Fehlern offen umzugehen, die Angst der Wissenschaftler, vor ihren Kollegen bloßgestellt zu werden. Denn die pure Tatsache, die eigene Arbeit mit dem Hinweis auf einen möglichen Fehler bei PubPeer widerzufinden, kann eben wegen des Fehlens einer wissenschaftlichen Fehlerkultur zu Reputationsverlust und bei Konkurrenten zu Schadenfreude führen.

Und es gibt noch ein anderes Problem: Als anonyme Plattform eignet sich PubPeer trefflich für private Vendettas und Kampagnen, die nichts mit Wissenschaft, dafür aber umso mehr mit Neid oder persönlicher Animosität zu tun haben.

Ein instruktives und interessantes Beispiel für das Niveau der Auseinandersetzungen auf PubPeer zeigt die Diskussion einiger weniger Abbildungen in einem Paper aus Südhoofs Gruppe (*Neuron*, doi: 10.1016/j.neuron.2017.04.011). In wenigen Wochen haben



INTEGRA

SCHNELLE AUFREINIGUNG

REINIGUNG MIT MAGNETISCHEN BEADS LEICHT GEMACHT

MAG UND HEATMAG

Optimierte Aufreinigung mit magnetischen Beads auf INTEGRAs Benchtop-Pipettierplattformen

NEU



MAG



HEATMAG



VIAFLO 96 | VIAFLO 384



ASSIST PLUS

integra-biosciences.com

sich auf PubPeer vierzig Kommentare und Antworten angehäuft, die sich teilweise über Seiten ziehen sowie häufig mit bildanalytischen Auswertungen und hochtechnischen Auslassungen über die verwendete Gerätschaft und Software garniert sind (Stand 16.5.2024). Sicher erkennt man dabei die Ernsthaftigkeit und auch die Kompetenz der meisten Kommentatoren. Es schleicht sich aber auch das Gefühl ein, dass sich manche Kritiker wiederholen oder sich in irrelevanten Details ergehen. Bei dem schlichten Versuch, sich ein Bild der Kontroverse zu machen, verliert man irgendwann unweigerlich den Überblick – und je nach persönlichem Bias und eigener Expertise denkt man sich entweder: „Irgendwie geht das zu weit, hier wird eine Sau durchs Dorf getrieben.“ Oder man insistiert: „Wenn schon diese paar Abbildungen in einer Arbeit problematisch sind, wie steht es dann eigentlich mit dem Rest? Richtig und wichtig, dass auch ein Nobelpreisträger an den Standards der guten wissenschaftlichen Praxis gemessen wird!“

»Ich meine, wir haben nicht zu viele „Data Sleuths“, sondern viel zu wenige.«

Auf der von Südhof eigens für die Auseinandersetzung mit den PubPeer-Vorwürfen aufgesetzten Website (<https://med.stanford.edu/sudhoflab/integrity---pubpeer.html>; diese und weitere Zitate und Links finden sich wie immer unter <http://dirnagl.com/lj>) beschwert er sich: „PubPeer und andere soziale Medien sind intransparent, zensieren Antworten und verwenden anonyme Kommentatoren mit ständig wechselnden Alias-Namen. Zudem werden PubPeer-Beiträge sofort an Zeitschriften, Journalisten und Universitätsverwaltungen weitergeleitet. Nicht alle Kommentare auf PubPeer sind aufrichtig, und einige scheinen eine Agenda zu verfolgen, die nichts mit wissenschaftlicher Integrität zu tun hat. Dennoch reagieren wir auf solche Kommentare, weil die Zeitschriften, an die die Kommentatoren ihre Anschuldigungen übermitteln, manchmal Anschuldigungen akzeptieren, ohne deren Plausibilität zu überprüfen. Darüber hinaus besteht eine gängige Strategie einiger Kommentatoren darin, Anschuldigungen auf verschiedene Weise zu wiederholen, um die Wirkung zu verstärken – oft begleitet von Grafiken und Animationen, die einen Anschein von Ernsthaftigkeit vermitteln. Selbst wenn eine Anschuldigung eher unwahrscheinlich ist, schafft diese Strategie eine Aura von ‚Hier stimmt etwas nicht!‘, die für die Beschuldigten schwer zu widerlegen ist.“

Die wenigen, die auf PubPeer mit offenem Visier diskutieren und daher nicht anonym bleiben, sind häufig das, was man derzeit als „Professional Data Sleuths“ bezeichnet. Hierzu zählen zum Beispiel Leonid Schneider, Sholto David – und nicht zuletzt Elisabeth Bik, die über die Vielzahl der von ihr mitaufgedeckten Fälle von Bildmanipulationen mittlerweile selbst über Fachkreise hinaus prominent geworden ist. Data Sleuths leben vom Aufspüren von schlechter Wissenschaft, indem sie sich oft zumindest teilweise über Social-Payment-Services wie etwa Patreon oder über Crowdsourcing finanzieren – und das vermutlich mehr schlecht als recht. Dennoch wirft Thomas Südhof ihnen genau das vor: nämlich ein „finanzielles Interesse“ zu haben. Und er ist nicht allein damit.

Allerdings liegen er und alle anderen damit voll daneben. Zum einen ist das ein unbrauchbares Argument gegen die von den Sleuths ja immer ausführlich belegten Auffälligkeiten. Die Auffälligkeiten sind entweder real oder sie sind es nicht – das hat nichts mit dem finanziellen Hintergrund derer zu tun, die sie aufzeigen. Zum anderen ist es traurig, dass es so einen „Berufsstand“ offensichtlich braucht. Denn wenn das schon so ist, dann sollte er von den Wissenschaftsinstitutionen, Unis und Verlagen finanziert werden. Ich meine vielmehr, wir haben nicht zu viele Data Sleuths, die sich dieser Aufgabe professionell widmen, sondern viel zu wenige. Mittlerweile ist es doch leider eher so, dass es sich lohnen würde, einen entsprechenden Studiengang zur Ausbildung von kompetentem Nachwuchs aufzubauen und dafür dauerfinanzierte Stellen in den Forschungsabteilungen der Institutionen einzurichten.

Festzuhalten bleibt daher, dass PubPeer das derzeit beste Medium ist, das die Korrektur der wissenschaftlichen Literatur erlaubt. Der etablierte Wissenschaftsbetrieb – einschließlich unehrlicher Forscher, ängstlicher Herausgeber und klageunwilliger Institutionen – ist jedenfalls bisher nicht in der Lage dazu.

Ebenso zeigt der „Fall Südhof“ eindrücklich die Stärken des Post-Publication-Review-Prozesses und einer Plattform wie PubPeer – und wie wichtig es ist, dass sich Wissenschaftler sowohl als Kommentatoren wie auch als kritisierte Autoren auf dieses Format einlassen. Thomas Südhof und seine Koautoren haben – im Gegensatz zu vielen anderen, deren Arbeiten auf PubPeer kritisiert wurden – zeitnah auf jeden Kommentar konstruktiv und wissenschaftlich reagiert. Sie konnten einen erheblichen Teil der Kritik aufklären beziehungsweise ausräumen – und haben tatsächliche Fehler eingestanden und korrigiert. Ausdrücklich bedankt sich Südhof bei denen, die konstruktiv kommentiert hatten, und lobt das Prinzip

von PubPeer. Nach gegenwärtigem Stand hatten die nachgewiesenen Fehler relevante Aussagen der Arbeiten nicht beeinflusst. Absichtliche Manipulationen oder Täuschung wurden meiner Ansicht nach nicht überzeugend nachgewiesen – wobei die Diskussion hierüber noch nicht abgeschlossen ist. Sehr wohl jedoch weist die Häufung von Fehlern im Südhof-Labor auf allgemeine Qualitäts- und Supervisionsprobleme seiner vermutlich sehr großen Arbeitsgruppe hin.

»Wir sollten PubPeer mit kühlem Kopf und wissenschaftlicher Etikette weiter nutzen.«

Sicher scheint mir bei alledem aber, dass die vermehrt auch in der Öffentlichkeit diskutierten PubPeer-Fälle von möglichem Wissenschaftsbetrug nur die kleine Spitze eines gigantischen Eisberges darstellen. Denn diese sind oftmals nur wegen der Prominenz der beteiligten Wissenschaftler besonders sichtbar. Große Teile der Wissenschaft haben ein Qualitäts- beziehungsweise Integritätsproblem, deren Hauptursachen – wie in dieser Kolumne schon häufiger angeprangert – in einem falschen akademischen Anreizsystem liegen. PubPeer ist ein Indikator hierfür – ich glaube aber auch, dass es Teil der Lösung ist.

Im Übrigen lehrt uns PubPeer, dass jeder Wissenschaftler sich darüber bewusst sein sollte:

- » dass die eigenen Arbeiten nach deren Veröffentlichung Gegenstand einer intensiven Auseinandersetzung mit einer breiten Community werden können;

- » dass die Einhaltung aller Regeln der guten wissenschaftlichen Praxis sowie die gute Dokumentation der Resultate und Analysen (in Originalformaten) sich demnach langfristig auszahlt;

- » und dass es daher nicht nur kurzfristig auf eine tolle Publikation ankommt.

Deshalb empfiehlt der Wissenschaftsnarr auch die Installation des PubPeer-Browser-Plugins, das auf existierende PubPeer-Diskussionen hinweist, wenn man Artikel sucht oder liest. Wir sollten PubPeer mit kühlem Kopf und wissenschaftlicher Etikette weiter passiv wie aktiv – idealerweise mit offenem Visier – nutzen, und es zu dem machen, als das es sich selbst bezeichnet: einem internationalen „Online-Journal-Club“.

Dieser Artikel basiert auf einem Gastbeitrag des Autors im Berliner Tagesspiegel vom 12.4.2024. Weiterführende Literatur und Links finden sich wie immer unter: <http://dirnagl.com/lj>.