



mitp

ALEX REINHART

# STATISTICS DONE WRONG

STATISTIK RICHTIG ANWENDEN  
UND GÄNGIGE FEHLER VERMEIDEN

DEUTSCHE  
AUSGABE



no starch  
press



## **Hinweis des Verlages zum Urheberrecht und Digitalen Rechtemanagement (DRM)**

Der Verlag räumt Ihnen mit dem Kauf des ebooks das Recht ein, die Inhalte im Rahmen des geltenden Urheberrechts zu nutzen. Dieses Werk, einschließlich aller seiner Teile, ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Dies gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und Einspeicherung und Verarbeitung in elektronischen Systemen.

Der Verlag schützt seine ebooks vor Missbrauch des Urheberrechts durch ein digitales Rechtemanagement. Bei Kauf im Webshop des Verlages werden die ebooks mit einem nicht sichtbaren digitalen Wasserzeichen individuell pro Nutzer signiert.

Bei Kauf in anderen ebook-Webshops erfolgt die Signatur durch die Shopbetreiber. Angaben zu diesem DRM finden Sie auf den Seiten der jeweiligen Anbieter.

# Rezensionen

»Ein kostbares kleines Buch ... erstaunlich und auch für Laien zugänglich.«

– *Alberto Cairo, Visualization Program Director,  
Zentrum für Informatik der Universität Miami*

»Wenn Sie regelmäßig Daten auswerten und sich nicht sicher sind, es richtig zu machen, sollten Sie dieses Buch lesen.«

– *Nathan Yau, FlowingData*

»Ein herrlicher und informativer Leitfaden ... ein Ausbund an Klarheit.«

– *John A. Wass, Scientific Computing*

»Ich werde dieses Buch zweifellos nicht nur Leuten empfehlen, die ein Interesse an medizinischen Statistiken haben, sondern erst recht Ärzten, die Statistiken nicht leiden können!«

– *Dr. Catey Bunce, Leiterin der Statistikabteilung,  
Moorfield Eye Hospital NHS Foundation Trust*

»Ich schätze dieses Buch sehr und werde es meinen Studenten empfehlen ... verblüffend.«

– *Dr. Nicole Radziwill, Assistenzprofessorin,  
Department of Integrated Science and Technology,  
James Madison Universität*

»Ich wünschte, alle Ärzte würden das lesen.«

– *Dr. Eric LaMotte, Internal Medicine Residency Program,  
Universität Washington*

»Ein gewagtes Buch, das gleichzeitig auch fasziniert ... wirklich unterhaltsam, es wird Ihre Denkweise über Statistik dauerhaft ändern.«

– *Ben Rothke, Information Security Professional*

»Ein gut geschriebener, amüsanter und nützlicher Leitfaden, der sich der gebräuchlichsten Probleme annimmt, die heutzutage bei statistischen Auswertungen praktiziert werden.«

– *Civil Statistician*

»Wer Forschungsergebnisse auch nur ansatzweise verstehen möchte, sollte dieses Buches als wertvollen Leitfaden in Betracht ziehen, um es richtig anzustellen.«

– *Sandra Henry-Stocker, ITworld*

»Eine wichtige Ergänzung der Bibliothek eines jeden Wissenschaftlers, der Statistik betreibt. Der prägnante Stil hält Ihr Interesse wach und wird Ihre zukünftigen Projekte inspirieren. Ich kann das Buch nur wärmstens empfehlen.«

– *Inside BIGDATA*



*Der oberste Grundsatz lautet, dass man sich nichts vormachen darf  
– und sich selbst macht man am leichtesten etwas vor.*

Richard P. Feynman

*Einen Statistiker nach Abschluss eines Experiments um Rat zu fragen, bedeutet oftmals nichts anderes, als ihn um eine Art Autopsie zu bitten – möglicherweise ist er in der Lage festzustellen, woran das Experiment gestorben ist.*

R.A. Fisher

Alex Reinhart

# Statistics Done Wrong

Statistik richtig anwenden und  
gängige Fehler vermeiden

Übersetzung aus dem Amerikanischen  
von Knut Lorenzen



**mitp**

## **Bibliografische Information der Deutschen Nationalbibliothek**

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

ISBN 978-3-95845-253-4

1. Auflage 2016

[www.mitp.de](http://www.mitp.de)

E-Mail: [mitp-verlag@sigloch.de](mailto:mitp-verlag@sigloch.de)

Telefon: +49 7953 / 7189 - 079

Telefax: +49 7953 / 7189 - 082

Übersetzung der englischen Originalausgabe:

Copyright © 2015 by Alex Reinhart. Title of English-language original: Statistics Done Wrong, ISBN 978-1-59327-620-1, published by No Starch Press. German-language edition copyright © 2016 by mitp Verlags GmbH & Co. KG. All rights reserved.

© 2016 mitp Verlags GmbH & Co. KG

Dieses Werk, einschließlich aller seiner Teile, ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Dies gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Lektorat: Sabine Schulz

Fachkorrektur: Prof. Dr. Stefanie Schöberl

Sprachkorrektur: Maren Feilen

Coverbild: Josh Ellingson

Satz: III-satz, Husby, [www.drei-satz.de](http://www.drei-satz.de)

# Inhaltsverzeichnis

	<b>Rezensionen</b> .....	1
	<b>Über den Autor</b> .....	10
	<b>Vorwort</b> .....	11
	<b>Einleitung</b> .....	16
<b>1</b>	<b>Eine Einführung in die statistische Signifikanz</b> .....	23
1.1	p-Werte .....	23
1.1.1	»Übernatürliche« Statistik .....	27
1.1.2	Neyman-Pearson-Tests .....	28
1.2	Intervallen vertrauen .....	31
<b>2</b>	<b>Teststärke und Stichprobenumfang</b> .....	35
2.1	Die Gütefunktion .....	35
2.2	Risiken unzureichender Teststärke .....	39
2.2.1	Weshalb unzureichende Teststärke? .....	41
2.2.2	Abbiegen bei Rot .....	43
2.3	Beeinflussung des Konfidenzintervalls .....	46
2.4	Überbewertung der Tatsachen .....	47
2.4.1	Kleine Extrema .....	51
2.5	Tipps .....	54
<b>3</b>	<b>Pseudoreplikation: Datenauswahl</b> .....	57
3.1	Pseudoreplikation in Aktion .....	58
3.2	Pseudoreplikation berücksichtigen .....	60
3.3	Biologie wie am Fließband .....	61
3.4	Synchronisierte Pseudoreplikation .....	63
3.5	Tipps .....	66
<b>4</b>	<b>p-Wert und Prävalenzfehler</b> .....	69
4.1	Prävalenzfehler .....	71
4.1.1	Ein Kurztest .....	72
4.1.2	Prävalenzfehler bei medizinischen Untersuchungen .....	73

4.1.3	Wie man mit Raucherstatistiken lügt . . . . .	74
4.1.4	Maßnahmen zur Vermeidung von Prävalenzfehlern . . . . .	77
4.2	Wenn es nicht sofort klappt, immer wieder probieren . . . . .	80
4.3	Falsche Fährten bei Kernspintomografien des Gehirns . . . . .	85
4.4	Erkennung fehlerhafter Zuordnungen . . . . .	86
4.5	Tipps . . . . .	88
<b>5</b>	<b>Unzureichende Beurteilung der Signifikanz</b> . . . . .	<b>89</b>
5.1	Nicht signifikante Unterschiede der Signifikanz . . . . .	89
5.2	Visueller Vergleich der Signifikanzen . . . . .	94
5.3	Tipps . . . . .	98
<b>6</b>	<b>Doppelte Datennutzung</b> . . . . .	<b>99</b>
6.1	Zirkuläre Analysen . . . . .	100
6.2	Regression zur Mitte . . . . .	104
6.3	Abbruchregeln . . . . .	107
6.4	Tipps . . . . .	111
<b>7</b>	<b>Fortbestehende Fehler</b> . . . . .	<b>113</b>
7.1	Überflüssige Zweiteilung . . . . .	114
7.2	Statistische Verdunkelung . . . . .	116
7.3	Verwirrende Störfaktoren . . . . .	117
7.4	Tipps . . . . .	120
<b>8</b>	<b>Modellmissbrauch</b> . . . . .	<b>121</b>
8.1	Wassermelonen . . . . .	123
8.2	Korrelation und Kausalität . . . . .	128
8.3	Simpsons Paradoxon . . . . .	130
8.4	Tipps . . . . .	134
<b>9</b>	<b>Forschungsfreiheit: Alles in Butter?</b> . . . . .	<b>135</b>
9.1	Ein wenig Spielraum kann gefährlich sein . . . . .	137
9.2	Voreingenommenheit vermeiden . . . . .	140
9.3	Tipps . . . . .	144
<b>10</b>	<b>Fehler begeht jeder</b> . . . . .	<b>145</b>
10.1	Unreproduzierbare Genetik . . . . .	146
10.2	Reproduzierbarkeit erleichtern . . . . .	149
10.3	Experimentieren, aufräumen, wiederholen . . . . .	152
10.4	Tipps . . . . .	154

II	<b>Daten verbergen</b> .....	155
II.1	Eingesperrte Daten .....	156
II.1.1	Hindernisse für das Teilen der Daten .....	157
II.1.2	Datenverfall .....	159
II.2	Einzelheiten weglassen .....	162
II.2.1	Bekannte Unbekannte .....	162
II.2.2	Voreingenommene Berichterstattung .....	164
II.3	Wissenschaft im Aktenschrank .....	167
II.3.1	Unveröffentlichte klinische Studien .....	167
II.3.2	Voreingenommene Berichterstattung erkennen .....	169
II.3.3	Erzwungene Offenlegung .....	171
II.4	Tipps .....	173
12	<b>Was ist zu tun?</b> .....	175
12.1	Statistikausbildung .....	178
12.2	Veröffentlichung wissenschaftlicher Arbeiten .....	182
12.3	Ihre Aufgaben .....	185
	<b>Anmerkungen</b> .....	189
	<b>Stichwortverzeichnis</b> .....	207

# Über den Autor

Alex Reinhart ist Doktorand an der Carnegie Mellon University in Pittsburgh, Pennsylvania, und unterrichtet dort Statistik. Seinen Bachelor of Science in Physik erlangte er an der University of Texas in Austin. Er befasst sich mit Forschungsarbeiten zum Aufspüren radioaktiver Objekte unter Zugrundelegung physikalischer und statistischer Gesetzmäßigkeiten.

# Vorwort



Als ich vor einigen Jahren mein Hauptstudium im Fachbereich Physik an der University of Texas in Austin absolvierte, wurde ich im Rahmen eines Seminarkurses aufgefordert, mir ein Thema für einen individuellen 25-minütigen Vortrag auszusuchen, den alle Seminarteilnehmer halten sollten. »Ich würde gern etwas über Verschwörungstheorien machen«, sagte ich zu dem Dozenten Dr. Brent Iverson, der sich damit jedoch nicht zufriedengab. Das Thema sei zu weitreichend, meinte er, und ein engagierter Vortrag müsse fokussiert und ausführlich sein. Also sah ich mir die Liste der Themenvorschläge an. »Wie wäre es denn mit wissenschaftlichem Fehlverhalten?«, fragte er, und ich willigte ein.

Rückblickend bin ich gar nicht mal sicher, inwieweit Falschangaben und fehlerhafte Schlussfolgerungen in der Wissenschaft ein begrenzteres Thema als Verschwörungstheorien sind, das spielte aber auch keine Rolle. Nach einigen Stunden doch ziemlich intensiver Recherche wurde mir klar, dass das wissenschaftliche Fehlverhalten an sich kein allzu interessantes Thema ist – jedenfalls nicht in Anbetracht all der Fehler, die Wissenschaftler *unabsichtlich* begehen.

Ungeachtet meiner beklagenswert mangelhaften Qualifikation in punkto Statistik suchte ich mir also einige Dutzend Forschungsarbeiten über die zahllosen statistischen Fehler zusammen, die Wissenschaftlern regelmäßig unterlaufen. Ich las sie, resümierte sie und arbeitete schließlich einen Vortrag aus, der Dr. Iverson zufriedenstellte. Und als angehender Wissenschaftler (und nun auch noch selbsternannter Statistikexperte) kam ich letztendlich zu der Überzeugung, dass ich vielleicht auch mal einen Statistikkurs besuchen sollte.

Gesagt, getan: Zwei Jahre und zwei Statistikkurse später immatrikulierte ich an der Carnegie Mellon University als Doktorand im Fachbereich Statistik – und bis heute bereitet es mir ein diebisches Vergnügen, Gelegenheiten zu finden, um mit statistischen Werten »herumzupfuschen«.

Dieses Buch ist ein Leitfaden, der den eher hanebüchenden statistischen Fehlern und Irrtümern gewidmet ist, die regelmäßig im Namen der Wissenschaft begangen werden. Da viele Wissenschaftler keine formale Statistikausbildung genießen – und weil ich meine Leserschaft nicht auf ohnehin an diesem Thema Interessierte beschränken möchte –, werden hier keine besonderen Vorkenntnisse in Sachen Statistik vorausgesetzt. Manch ein Leser mag daher möglicherweise versucht sein, das erste Kapitel zu überspringen, ich empfehle jedoch, es zumindest zu überfliegen, um sich mit meinem Erklärungsstil vertraut zu machen.

Mein Ziel besteht nicht nur darin, die oft begangenen Fehler beim Namen zu nennen und entsprechende Beispiele zu liefern, über die Sie schmunzeln können. Soweit es ohne ausführliche mathematische Exkurse möglich ist, erläutere ich außerdem, *warum* besagte statistische Fehler eigentlich Fehler sind und berufe mich in diesem Kontext auf Studien, die belegen, wie weit die meisten dieser Fehler tatsächlich verbreitet sind. Die Lektüre dieses Buches wird dadurch zwar ein wenig erschwert, ich bin jedoch der Ansicht, dass der hieraus resultierende Erkenntnisgewinn die Mühe durchaus wert ist – ein gründliches Verständnis von elementarer Statistik ist schließlich für jeden Wissenschaftler unverzichtbar.

Für Leser, die alltäglich statistische Analysen durchführen, finden sich am Ende der meisten Kapitel Tipps, die Aufschluss darüber geben, welche statistischen Verfahren verwendet werden können, um die häufigsten Fallstricke zu umgehen. Das vorliegende Werk ist allerdings kein Lehrbuch, daher werde ich auch nicht näher auf die Einzelheiten der genannten Verfahren eingehen. Ich hoffe jedoch, Sie auf die gängigsten Probleme aufmerksam machen und somit in die Lage versetzen zu können, das für eine gegebene Aufgabe am besten geeignete statistische Verfahren auszuwählen.

Für den Fall, dass ein oder mehrere der hier angesprochenen Themen Ihre Neugier wecken, finden Sie am Ende des Buches ein ausführliches Literaturverzeichnis. Außerdem sind alle statistischen Fehler und Irrtümer mit Quellverweisen versehen. Zugunsten eines grundsätzlichen Verständnisses habe ich sehr viel Mathematisches weggelassen. Sollten Sie jedoch an formal strengeren Abhandlungen interessiert sein, empfehle ich die Lektüre der diesem Buch zugrundeliegenden Originalarbeiten.

Bevor Sie dieses Buch lesen, hier noch ein kurzer Warnhinweis: Sobald man sich Wissen angeeignet hat, über das nur wenige andere Menschen verfügen, ist es verlockend, dies bei jeder sich bietenden Gelegenheit zur Schau zu stellen. Sollte dieses Buch auf wundersame Weise zum Bestseller werden, erwarte ich als Reaktion auf jegliche wissenschaftliche Meldungen in der Boulevardpresse das, was Paul Graham als »Ablehnung durch den Durchschnittsmenschen« bezeichnet – denn anstatt sich die Zeit zu nehmen, die interessanten Teile wissenschaftlicher Forschung zu verstehen, verwenden Möchtegern-Statistiker lieber die von irgendeiner übereifrigen Universitätspressestelle wiedergekäute ungenaue Beschreibung einer Studie, um sich auf solche Artikel einzuschießen und den statistischen Aufbau der Forschungsarbeit zu kritisieren.\*

Auf den meisten Websites, die wissenschaftliche Meldungen diskutieren, geschieht dies bereits und es würde mich ungemein ärgern, wenn dieses Buch als Rechtfertigung dafür erhalten sollte. Die ersten Kommentare zu neuen Newsartikeln und Nachrichtenmeldungen sind fast immer Beschwerden wie »diese Variable wurde nicht kontrolliert« oder »die Stichprobe ist viel zu klein« – und in 9 von 10 Fällen hat der Kommentierende die wissenschaftliche Arbeit überhaupt nicht gelesen,

---

\* Nebenbei bemerkt denke ich, dass Verschwörungstheorien aus eben diesem Grund so populär sind: Wenn man glaubt, etwas zu wissen, das niemand sonst weiß (die Regierung ist hinter uns her!), nimmt man jede Gelegenheit wahr, diese Erkenntnisse zur Schau zu stellen – und schließlich reagiert man auf alle Nachrichtenmeldungen damit, Gründe hervorzubringen, warum die Regierung sie fälscht. Das sollte Ihnen bei statistischen Fehlern jedoch bitte nicht passieren.

denn sonst wüsste er, dass seinem Einwand im dritten Absatz begegnet wird.

Das Ganze ist absolut töricht: Einige wenige Statistikenkenntnisse sind keine Entschuldigung dafür, die gesamte moderne Wissenschaft abzulehnen. Die statistischen Methoden einer Forschungsarbeit dürfen nur im Detail und im Kontext der anderen Methoden beurteilt werden: Aufbau der Studie, Messverfahren, Kostenbeschränkungen und Aufgabenstellung. Setzen Sie Ihre Statistikenkenntnisse ein, um die Stärken, die Einschränkungen und die potenzielle Voreingenommenheit einer Untersuchung besser zu verstehen, nicht um jede Forschungsarbeit niederzumachen, die unsachgemäßen Gebrauch von einem  $p$ -Wert macht oder Ihren persönlichen Ansichten widerspricht. Bedenken Sie außerdem, dass auch eine Schlussfolgerung, die durch mangelhafte Statistiken belegt wird, dennoch richtig sein kann – statistische und logische Fehler führen nicht automatisch zu einem Fehlschluss, sie belegen ihn nur nicht.

Kurzum: Gehen Sie bitte verantwortungsvoll mit Statistiken um. Ich kann mir nur wünschen, dass Sie danach streben, die Wissenschaft zu verbessern, von der wir alle abhängig sind.

## Danksagungen

Danke an James Scott, dessen Statistikkurse am Anfang meiner Laufbahn im Fachbereich Statistik standen und mir das zum Schreiben dieses Buches nötige Hintergrundwissen vermittelten. Weiterhin bedanke ich mich bei Raye Allen, der James' Hausaufgaben so viel spaßiger machte, bei Matthew Watson und Moriel Schottlander, die mir wertvolle Anregungen und unschätzbare Feedback zu meinen Textentwürfen gaben, bei meinen Eltern für ihre Denkanstöße und Rückmeldungen sowie bei Dr. Brent Iverson, dessen Seminar mich überhaupt erst dazu veranlasste, mich mit dem unsachgemäßen Gebrauch von Statistiken zu befassen. Und natürlich gilt mein Dank ebenso allen Wissenschaftlern und Statistikern, die gegen die Regeln verstoßen haben und mir somit einen Grund zum Schreiben lieferten.

Auch meine Freunde an der Carnegie Mellon University haben viele Ideen beigesteuert und jede Menge Fragen beantwortet. Sie haben mir stets geduldig zugehört, wenn ich mal wieder versuchte, einen neuen statistischen Fehler zu erläutern. Das nötige Hintergrundwissen hierfür verdanke ich meinen Professoren, insbesondere Jing Lei, Valérie Ventura und Howard Seltman. Howard fungierte außerdem als fachlicher Redakteur und spürte in dieser Eigenschaft einige peinliche Fehler in meinem Buchentwurf auf. Sollten immer noch welche übrig sein, bin ich selbst dafür verantwortlich – allerdings werde ich frech behaupten, dass sie lediglich dem Titel des Buches Rechnung tragen.

Und meine Herausgeber bei No Starch haben dem Manuskript schließlich den nötigen Feinschliff verpasst und es damit drastisch verbessert. Greg Poulos gab sich nach der sorgfältigen Lektüre der ersten Kapitel erst zufrieden, als er jedes einzelne Konzept verstanden hatte. Leslie Shen sorgte für die Glättung meiner Polemik in den letzten Kapiteln, und das gesamte Team hat maßgeblich dazu beigetragen, dass sich dieses Projekt insgesamt verblüffend einfach gestaltete.

Darüber hinaus bin ich vielen weiteren Leuten zu Dank verpflichtet, die mir Anregungen und Kommentare per E-Mail geschickt haben, nachdem der Leitfaden online gestellt wurde. Ich bedanke mich (in keiner bestimmten Reihenfolge) bei: Axel Boldt, Eric Franzosa, Robert O’Shea, Uri Bram, Dean Rowan, Jesse Weinstein, Peter Hozák, Chris Thorp, David Lovell, Harvey Chapman, Nathaniel Graham, Shaun Gallagher, Sara Alspaugh, Jordan Marsh, Nathan Gouwens, Arjen Noordzij, Kevin Pinto, Elizabeth Page-Gould und David Merfield, ohne deren Feedback meine Erläuterungen zweifelsohne weniger vollständig wären.

Vielleicht gesellen auch Sie sich bald zu dieser Liste hinzu, denn trotz meiner größten Bemühungen wird dieser Leitfaden auch jetzt unweigerlich noch einige Fehler und Versäumnisse aufweisen. Sollten Sie also eine Unstimmigkeit finden, eine Frage haben oder ein gängiges statistisches Fehlverhalten kennen, das ich übersehen habe, können Sie mich auf Englisch unter [alex@refsmmat.com](mailto:alex@refsmmat.com) per E-Mail erreichen.

# Einleitung



Im letzten Kapitel seines bekannten Buches *How to Lie with Statistics* (»Wie man mit Statistiken lügt«) schreibt Darrell Huff, dass »alles, was nach dem medizinischen Berufsstand riecht« oder von wissenschaftlichen Laboratorien und Universitäten mitgetragen wird, unser Vertrauen genießt – zwar kein bedingungsloses Vertrauen, aber sicherlich ein größeres als das den Medien oder Politikern entgegengebrachte. (Kein Wunder, denn schließlich enthüllt Huffs Buch eine Vielzahl der irreführenden statistischen Tricksereien, die in den Medien und der Politik Anwendung finden.) Doch nur die wenigsten Leute beklagen sich über die von ausgebildeten Wissenschaftlern erstellten Statistiken – denn schließlich streben besagte Akademiker ja nach Erkenntnisgewinn und nicht nach irgendwelchem Material, das sich gegen politische Gegner verwenden lässt.

Statistische Datenanalysen sind ein Grundpfeiler der Wissenschaft. Schlagen Sie irgendeine beliebige Seite in einem der Fachmagazine auf, die medizinische Forschungsergebnisse veröffentlichen, und Sie werden mit Statistiken überhäuft:  $t$ -Tests,  $p$ -Werte (Signifikanzwerte), Cox-Regressionen, Propensity-Score-Analysen (Verwirklichungstendenzen), logistische (Logit-Modell) oder lineare Regressionen mit der Methode der kleinsten Quadrate und Konfidenzintervalle (Erwartungsbereiche). Statistiker haben Wissenschaftlern außerordentlich mächtige Werkzeuge an die Hand gegeben, um Sinn und Ordnung in den kompliziertesten Datensammlungen aufzuspüren – und die Wissenschaftler haben sie mit Freuden entgegengenommen.

Sie haben sich allerdings weder einer *Statistikausbildung* unterzogen, noch ist in vielen wissenschaftlichen Studiengängen eine wie auch immer geartete Schulung für den Einsatz von Statistiken vorgesehen.

Seit den 1980er Jahren dokumentieren Forscher immer wieder statistische Irrtümer und Missverständnisse in der einschlägigen wissenschaftlichen Literatur und beanstanden, dass viele wissenschaftliche Publikationen (womöglich mehr als die Hälfte) diesen Fehlschlüssen zum Opfer fallen. Statistisches Unvermögen führt häufig dazu, dass die eigentlich gesuchten Erkenntnisse in vielen Studien gar nicht erst gefunden bzw. ermittelt werden können: Mehrfache Vergleiche und fehlinterpretierte Signifikanzwerte verursachen reihenweise falsche Positive, anpassungsfähige Datenanalysen begünstigen das Auffinden eigentlich gar nicht vorhandener Korrelationen, und die Auswahl ungeeigneter Modelle sorgt für die Verzerrung maßgeblicher Ergebnisse. Die meisten dieser Fehler werden von Fachkollegen und Redakteuren nicht bemerkt, weil sie oftmals nicht über eine spezielle statistische Ausbildung verfügen – denn nur wenige Fachmagazine beschäftigen »echte« Statistiker, um die eingereichten Beiträge zu überprüfen. Abgesehen davon liefern ohnehin nur die wenigsten Veröffentlichungen hinreichend genaue statistische Daten, um sie überhaupt korrekt bewerten zu können.

Das Problem besteht nicht etwa in betrügerischen Absichten, sondern vielmehr in mangelhaften statistischen Kenntnissen – und zwar derart mangelhaften, dass einige Wissenschaftler mutmaßen, dass die meisten veröffentlichten Forschungsergebnisse vermutlich falsch sind.<sup>1</sup> In den führenden Fachmagazinen finden sich regelmäßig Rezensionen und Leitartikel, die ein höheres statistisches Niveau und strengere Überprüfungen der Publikationen verlangen. Diesen Forderungen schenken allerdings nur wenige Wissenschaftler Gehör, und auch die von den Fachzeitschriften verfügbaren Standards werden häufig missachtet. Da die Anleitungen und Empfehlungen zum Umgang mit statistischen Daten auf oftmals irreführende Lehrbücher, Rezensionen in den verschiedensten Fachmagazinen und nur schwer verständliche Forschungsarbeiten zum Thema Statistik verteilt sind, haben die meisten Wissenschaftler größte Mühe, ihre Statistikkenntnisse zu verbessern.

Die Komplexität der Verfahrensweisen moderner Forschung bringt es mit sich, dass Wissenschaftler ohne umfassende Ausbildung im Bereich der Statistik womöglich einen Großteil der in ihren jeweiligen

Fachbereichen veröffentlichten Forschungsarbeiten gar nicht wirklich umfassend begreifen können. Im Bereich Medizin beispielsweise hätte ein Arzt, der einen Standardeinführungskurs in Statistik absolviert hat, gerade mal ausreichende Kenntnisse, um lediglich ein Fünftel der im *New England Journal of Medicine* veröffentlichten Forschungsartikel komplett zu verstehen.<sup>2</sup> Die meisten Ärzte verfügen allerdings noch nicht einmal über dieses statistische Basiswissen – viele der in Ausbildung befindlichen Ärzte erwerben ihre dahingehenden Kenntnisse in Journal Clubs oder kurzen Lehrgängen statt in den eigentlich erforderlichen Kursen.<sup>3</sup> Zudem werden die den Medizinstudenten dort vermittelten Inhalte oft kaum wirklich verstanden – im Durchschnitt werden weniger als 50% der Fragen zu den in der Medizin gebräuchlichen statistischen Methoden korrekt beantwortet.<sup>4</sup> Selbst an Lehreinrichtungen mit medizinischem Fachbereich, die Schulungen für die Forschung anbieten, werden weniger als 75% richtige Antworten erreicht.

Die Situation ist tatsächlich so verfahren, dass sogar die Autoren von Studien über Statistikkenntnisse selbst nicht über das erforderliche Statistikwissen verfügen, um die Fragen korrekt zu formulieren – die soeben genannten Zahlen sind daher irreführend, denn die Befragung der in Ausbildung befindlichen Ärzte enthielt eine Frage, bei der es um die Definition eines  $p$ -Wertes ging, mit einer Multiple-Choice-Vorgabe, die allerdings vier fehlerhafte Antworten zur Auswahl stellte.<sup>5</sup> Wir können hier jedoch etwas Nachsicht mit den Autoren üben, weil viele einführende Lehrbücher zum Thema Statistik dieses grundlegende Konzept ebenfalls mangelhaft oder falsch definieren.

Wenn die Entwickler wissenschaftlicher Studien bei der Verwendung statistischer Werte nicht ausreichend Vorsicht walten lassen, werden womöglich jahrelange Anstrengungen und zig Tausende Euros in Forschungsprojekte gesteckt, die letztlich gar keine Antworten auf die Fragen liefern, zu deren Klärung sie eigentlich beitragen sollen. Der Psychologe Paul Meehl beschwerte sich darüber wie folgt:

*Zwischenzeitlich hat unser bienenfleißiger Forscher, unbeeindruckt von wissenschaftlicher Logik und angesichts der »Exaktheit« der statistischen Überprüfung von Hypothesen glücklich, eine lange Liste von Veröffentlichungen produziert und wurde zum Professor*

*berufen. Was seine Beiträge zu den nachhaltigen Erkenntnissen der Psychologie angeht: Er hat praktisch nichts geleistet. Tatsächlich stellt er einen zwar potenten, aber unfruchtbaren intellektuellen Lebemann dar, der auf seinem Lebensweg fröhlich reihenweise geschändete Jungfern hinter sich lässt, jedoch keinerlei brauchbaren wissenschaftlichen Nachwuchs zeugt.<sup>6</sup>*

Nun ist es vielleicht nicht ganz fair, einem Großteil der Wissenschaftler intellektuelle Unfruchtbarkeit zu unterstellen, denn schließlich hängen die meisten wissenschaftlichen Forschungsbereiche nicht von ein paar fehlerinterpretierten  $p$ -Werten ab. Allerdings haben solche fehlerhaften Rückschlüsse im wahren Leben erhebliche Auswirkungen: Klinische medizinische Studien beeinflussen unser Gesundheitswesen und definieren, wie sicher wirksame verschreibungspflichtige Medikamente sind, Kriminologen werten auf der Grundlage statistischer Werte verschiedene Strategien zur Eindämmung von Verbrechen aus, Epidemiologen versuchen anhand entsprechender Daten die Ausbreitung neuer Krankheiten zu verhindern, und Vermarkter und Geschäftsleute suchen nach den besten Möglichkeiten, ihre Produkte zu verkaufen – letztendlich dreht sich alles um Statistiken. Um fehlerbehaftete Statistiken.

Wer sich schon mal über Ärzte beschwert hat, die sich keine Gedanken über das Wohlergehen der Patienten machen, kennt das Ausmaß dieses Problems. Mittlerweile überwiegt eine ablehnende Haltung gegenüber neuen Studien, die behaupten, dass uns ein bestimmtes Lebensmittel, eine bestimmte Ernährungsweise oder eine körperliche Ertüchtigungsform schadet – inzwischen warten wir lieber erst mal auf die einige Monate später unweigerlich erscheinende zweite Studie, die dann höchstwahrscheinlich das genaue Gegenteil zum Ergebnis hat. Ein prominenter Epidemiologe merkte dazu einmal an: »Wir werden zusehends zu einem gesellschaftlichen Ärgernis. Die Leute nehmen uns nicht mehr ernst – und wenn sie es doch tun, richten wir womöglich unabsichtlich mehr Schaden als Nutzen an.«<sup>7</sup> Unsere Instinkte funktionieren. In vielen Fachbereichen fallen die Ergebnisse der ursprünglichen und dann nachfolgend durchgeführten Untersuchungen zu ein und demselben Thema tendenziell widersprüchlich aus. Es scheint,