
ORIENTATE

Podcast Forschung MIT jedem
- eine Frage der Perspektive

*Episode 2:
Gesundheitsforschung & Statistik*

Entstanden im Rahmen
des Projektes ORIENTATE
- Citizen Science:
Patient*innenorganisationen
als (Mitbe-)Gründer von
Datenregistern.

Gefördert von
DFG & NFDI4Health



Eine Podcast-Reihe im Rahmen vom Projekt ORIENTATE

Forschung MIT jedem - eine Frage der Perspektive

*Episode 2:
Gesundheitsforschung
& Statistik*

Titelmusik

1. EINLEITUNG

Dieser Podcast ist Teil des Forschungsprojektes ORIENTATE, das eine Forschungsgruppe der Hochschule für Angewandte Wissenschaften Hamburg gemeinsam mit der Patientenorganisation PRO RETINA Deutschland e.V. durchführt. Das Projekt wird von Prof. Dr. Sabine Wöhlke und Dr. Henk Jasper van Gils-Schmidt geleitet. Dieser Podcast ist ein Weiterbildungsformat, das im Rahmen des Projektes von Lea Koop-Meyer entwickelt wurde und möchte einen Einblick in die Welt der Gesundheitsforschung geben. Dabei steht in dieser Episode vor allem der informierte Umgang mit Forschungsergebnissen und Statistiken im Mittelpunkt. Aus Gründen der besseren Verständlichkeit wird auf die gleichzeitige Verwendung der Sprachformen männlich, weiblich und divers (m/w/d) verzichtet. Sämtliche Personenbezeichnungen gelten gleichermaßen für alle Geschlechter.

Musik

2. ANMODERATION

Sabine: Herzlich Willkommen zu unserem Podcast! Mein Name ist Sabine Wöhlke und möchte in dieser Podcast Folge gemeinsam mit Ihnen in die Welt der Gesundheitsforschung & Statistik eintauchen. Wir beschäftigen uns heute mit der Rolle, die Daten in der Gesundheitsforschung und für unsere Gesundheit und unseren Alltag spielen. Wir möchten Orientierung geben, wie Daten und Statistiken aus Forschungsstudien und anderswo zu verstehen sind. Wir möchten alle Zuhörenden dazu ermutigen, sich zu trauen Daten zu hinterfragen und mitzudenken.ww

Hierzu haben wir Leonie Rosenberg und Thomas Duda zu einem Gespräch eingeladen, die gerade ge-



meinsam mit mir im Projekt ORIENTATE arbeiten. Wir geben Einblicke in unsere Forschungsarbeit und Erfahrungen in der Zusammenarbeit zwischen Wissenschaftlern und Patientinnen und Patienten. Gemeinsam wollen wir die Bedeutung von Daten in der Gesundheitsforschung diskutieren. Dabei soll vor allem die Rolle der Patienten, in der sich verändernden Welt der Gesundheitsforschung, im Fokus stehen: Warum sind meine Daten eigentlich so wichtig? Warum ist es wichtig, Statistiken und Ergebnisse von Studien zu verstehen und kritisch zu hinterfragen? Wie kann ich erreichen, dass Daten mich bei meiner persönlichen Entscheidungsfindung unterstützen können? Das alles sind Fragen, die wir heute hier gemeinsam besprechen wollen.

Musik

Leonie: Vielen Dank für die freundliche Begrüßung, Sabine! Ich freue mich, dass ich heute hier sein darf, um gemeinsam mit euch diesen spannenden Podcast aufzunehmen.

Thomas: Ja auch ich bin wieder mit dabei. Mein Name ist Thomas Duda. Ich bin Mitglied in der Patientenorganisation PRO RETINA und darf heute die Perspektive der Betroffenen und der Patienten einmal darstellen.

3. GESUNDHEITSFORSCHUNG – WAS BEDEUTET DAS?

Sabine: Heute soll sich ja alles um das Thema Daten & Gesundheit drehen. Um erstmal zu verstehen, welchen Stellenwert Daten in der Gesundheitsforschung einnehmen, möchte ich kurz mit euch über euer Verständnis der Gesundheitsforschung sprechen. Dabei wollen wir auch darüber sprechen, wie sich die beiden Forschungsansätze der qualitativen und quantitativen

Forschung unterscheiden. Dabei bin ich besonders gespannt darauf, von euren Erfahrungen zu hören.

Leonie: Dann gebe ich einen kurzen Einblick in die quantitative Forschung. Quantitative Forschung kann grundsätzlich erstmal als zahlenbasierte Forschung verstanden werden.

Ziel der quantitativen Forschung ist es, die vom Menschen geschaffene Welt in der Gesellschaft besser zu verstehen. Es geht uns also darum, Vermutungen über Zusammenhänge zwischen verschiedenen Ereignissen in der Realität zu überprüfen. Diese Vermutungen, die aus Theorien abgeleitet sind, müssen so umgewandelt werden, dass wir sie in Zahlen messen können. Ein Beispiel wäre die Annahme, dass es einen Zusammenhang zwischen der täglichen Lernzeit der Schülerinnen und Schüler und ihren Prüfungsergebnissen gibt. Wenn wir diese Vermutung in einer quantitativen Studie überprüfen wollen, müssen wir die Zeit, die die Schülerinnen und Schüler für das Lernen aufwenden, in Zahlen messen. Wir könnten dies tun, indem wir die Stunden aufzeichnen, die sie pro Tag mit Lernen verbringen. Mit mathematischen Methoden suchen wir dann nach Mustern und Regeln zwischen den verschiedenen Ereignissen.

Um zu dem Beispiel zurückzukommen: Die gemessenen Lernzeiten der Schülerinnen und Schüler können dann in Zahlen umgewandelt und mit ihren Prüfungsergebnissen in Verbindung gebracht werden, um zu sehen, ob es einen statistisch signifikanten Zusammenhang gibt. Einfach gesagt, wir suchen hier zum Beispiel nach einem möglichen Zusammenhang zwischen mehr Lernzeit und besseren Noten. In der quantitativen Forschung machen wir also soziale Dinge messbar, indem wir sie in Zahlen umwandeln. Mit diesen Zahlen können wir dann verschiedenste statistische Analysen durchführen. Das Wort "quantitativ" bedeutet also, dass wir Dinge messen, zählen und wiegen, und die Ergebnisse dann am Ende auch in Form von Zahlen präsentieren.

Sabine: Ich selber forsche hauptsächlich qualitativ, dort gibt es doch einige Unterschiede. Der Begriff "qualitativ" bedeutet, dass wir nicht versuchen, Dinge in Zahlen umzuwandeln oder mathematisch zu analysieren. Stattdessen versucht qualitative Forschung, die besonderen Eigenschaften und Merkmale eines sozialen Bereichs und menschlichen Handelns genau und detailliert zu verstehen, ohne sie zu messen oder zu erklären. Dabei ist es uns immer wichtig, dass wir den Kontext des sozialen Handelns berücksichtigen. Ein Beispiel für ein qualitatives Forschungsprojekt könnte es sein – um bei deinem Beispiel der Schülerinnen und Schüler zu bleiben - zu fragen, wie die Schülerinnen und Schüler das Lernen für ihre Prüfung erleben und welchen Einfluss dieses Erleben auf ihren Lernprozess und ihre Prüfungsergebnisse haben kann. Es geht uns also darum, ein tiefes Verständnis für die subjektiven Wahrnehmungen der einzelnen Schülerinnen und Schüler zu erhalten.

Thomas: Vielen Dank für den Einblick in deine Forschungspraxis und in quantitative Studien. Ich habe den Eindruck, als würde Partizipation eher bei qualitativen Studien stattfinden als bei quantitativen. Ich hatte schon die Möglichkeit, bei beiden teilzunehmen, also als Mitglied von PRO RETINA, sowohl in qualitativen Studien, also in quantitativen Studien. Und ich denke mir, das eine oder andere werden wir heute auch noch besprechen.

4. DIE RELEVANZ VON DATEN FÜR DIE GESUNDHEITSFORSCHUNG

Sabine: Vielen Dank für deine Eindrücke, Thomas! Gerade die Zusammenarbeit mit Patienten und anderen Nicht-Wissenschaftlern wird auch im Hinblick auf die zunehmende Digitalisierung der Medizin und Forschung und auch im Bereich der Persönlichkeitsmedizin immer wichtiger.

Leonie: Als Beispiel, denken Sie doch einmal an all die Gesundheitsdaten, die jeden Tag gesammelt werden: Krankenakten, Labortests, Umfragen und mehr. Bei jeder Untersuchung und Behandlung fallen medizinische Daten an. Diese Daten bilden die Grundlage für medizinische Forschung und ermöglichen es uns Wissenschaftlern, Krankheiten genauer zu verstehen und Faktoren zu identifizieren, die wichtig für die Behandlung von unterschiedlichsten Erkrankungen sind. Ein Beispiel wäre zum Beispiel die Testung eines neuen Medikaments, das die Behandlung verbessert. Statistiken und mathematische Analysen helfen uns dann wiederum zu verstehen, ob ein neues Medikament wirklich besser ist als das, was wir bereits haben, und ob es unerwünschte Nebenwirkungen gibt. So können wir sicherstellen, dass Patienten die bestmögliche Behandlung erhalten. Wir können also durchaus sagen „Daten helfen heilen“.

Sabine: Daten spielen also eine entscheidende Rolle in der Gesundheitsforschung. Sie sind nicht nur Zahlen und Diagramme, sondern ein mächtiges Werkzeug zur Verbesserung unseres Verständnisses von Krankheiten, zur Entwicklung effektiverer Behandlungen und zur Optimierung der Gesundheitsversorgung.

5. WARUM IST MEINE STIMME FÜR DIE FORSCHUNG WICHTIG?

Leonie: Und warum ist es euch so wichtig, dass sich Patienten in die Gesundheitsforschung einbringen und beispielsweise auch an quantitativen Projekten teilhaben?

Thomas: Für mich ist es aus zwei Gründen wichtig. Der eine ist natürlich der wissenschaftliche Aspekt. Und der andere stellt natürlich die Patienten und die Betroffenen in den Vordergrund. Ich möchte natürlich als Patient, als Betroffener schon wissen, was mit meinen Daten geschieht. Wie werden sie verarbeitet, Wer verarbeitet sie und wo werden sie vielleicht auch gespe-

ichert? Das heißt aber auch, man muss erst einmal eine gewisse Bereitschaft haben, Daten von sich preiszugeben. Es sind ja sehr sensible Daten, und das heißt eben auch eine gewisse, ein gewisses Vertrauen zu denjenigen Leuten, die diese Daten dann nachher verarbeiten. Und ich glaube, Vertrauen ist hier ein ganz wichtiger Punkt. Wie werden diese Daten gehandhabt? Liegen sie lose auf dem Schreibtisch? Hat jeder Zugriff auf diese Daten oder sind sie in einem geschützten Bereich? Ist extrem wichtig und wirkt eben auch dann dazu, die Daten mehr oder weniger preiszugeben. Dann möchte ich natürlich auch wissen, was geschieht anschließend mit diesen Daten? Wie werden sie ausgewertet? Ja, welche Statistiken ergeben sich daraus? Also sprich das Ganze dahinter, sozusagen, Was hinter den Kulissen passiert mit meinen Daten, ist ein extrem wichtiger Punkt aus Sicht eines Patienten, eines Betroffenen.

Leonie: Sabine, wie siehst du das? Warum ist es aus deiner Sicht wichtig, dass Patienten ihre Stimme einbringen?

Sabine: Naja, vieles davon hat Thomas bereits angesprochen. In erster Linie ist es wichtig für mich, da wir die Daten der Patienten benötigen, um statistische Tests durchzuführen. Gleichzeitig denke ich, dass gerade quantitative Forschung auf die Menschen oft wie eine Black Box wirkt: Ich fülle einen Fragebogen aus und am Ende steht eine Ergebnistabelle mit Begriffen, die mir oft fremd sind. Was dazwischen passiert und wie die Autoren zu den Ergebnissen kommen, bleibt vielen dabei unklar.

Leonie: Was auch daran liegt, dass dies in den Ergebnisberichten der Studien nicht niedrigschwellig beschrieben wird. Ich denke, dass hier wirklich Interventionsbedarf besteht. Wir müssen versuchen eine Brücke zu bauen zwischen der Zahlenwelt der Statistik und den Lebenswelten der Menschen. Wenn wir nicht versuchen unsere Arbeitsweise und die Ergebnisse einfach zu vermitteln, dann können wir unsere Ergebnisse

auch nicht mit denen diskutieren, für die sie möglicherweise Folgen haben.

Sabine: Ja, da gehe ich voll und ganz mit Thomas und mit dir mit, das ist natürlich eine geteilte Verantwortung. Ich kann nicht von jedem erwarten, dass er Statistik Experte wird, allerdings finde ich deshalb unser Gespräch heute so wichtig: ich möchte wenigstens dazu ermutigen, die Angst vor Zahlen und komplizierten Ausdrücken zu überwinden. Am Ende können Patienten nur so kritische Nachfragen stellen und mit den Wissenschaftlern diskutieren. So werden sie dann schließlich wieder zu den Hauptakteuren über ihre Gesundheitsentscheidungen und müssen sich nicht auf die Schlussfolgerungen anderer verlassen.

Leonie: Diesen letzten Punkt finde ich sehr wichtig. Ich denke auch, dass wir mündige Patienten brauchen, die mitdenken und ihre Meinung vertreten.

Thomas: Da gebe ich dir vollkommen recht. Es ist extrem wichtig, dort auch als mündiger Patient eben aufzutreten. Und es ist auch wichtig, dass sich die quantitativen Forschungen eben öffnen für eine Partizipation mit Betroffenen.

6. STATISTIKEN „RICHTIG“ LESEN UND BESSER VERSTEHEN

Leonie: Jetzt haben wir schon viel über die Bedeutung von Daten für die Forschung gesprochen. Wie sieht es aber konkret aus? Wie kann ich mir einen Überblick über Statistiken verschaffen und Studien quasi „richtig“ lesen? Was gilt es da zu beachten, damit ich nicht den Überblick verliere?

Thomas: Wir haben hier im Prinzip tatsächlich etwas mit Statistiken zu tun. Manchmal bewusst, manchmal unbewusst. Und ich glaube, es fängt schon sehr früh an. Beispielsweise in der Schule mit einem Notenspiegel. Das ist ja auch nichts anderes als eine Statistik. Dann, wenn wir tagtäglich die Nachrichten sehen, wo es dann

heißt, wir haben eine Regenwahrscheinlichkeit von X Prozent. Es werden die Trockentage gezählt. Wir haben aber auch einmal im Monat die Arbeitslosenstatistik und es gibt viele Sachen im alltäglichen Bereich, sei es jetzt im Sportbereich beispielsweise oder auch woanders, wo wir einfach Statistiken sehen. Aber natürlich, Statistiken muss man auch lesen können und auch entsprechend einordnen können, weil die können ganz unterschiedlich sein. Sei es, dass Sachen in Prozent ausgedrückt werden, sei es an der Skalierung beispielsweise, sei es auch in der Art der Darstellung in Form von Tortendiagramm, beispielsweise in einem Balkendiagramm oder in Trendkurven. Insofern, hier bedarf es einer gewissen Kompetenz, diese Daten selbst für sich zu interpretieren und auszuwerten.

Leonie: Für mich wäre es an dieser Stelle wichtig noch einmal einen Schritt zurückzumachen und zu verdeutlichen, wie wir in der Forschung von den Daten, die beispielsweise von Patienten bereitgestellt werden, zu einer Statistik kommen (siehe Infobox).

Sabine: Also, Statistiken sind - einfach ausgedrückt - eine zahlenmäßige Erfassung von Erscheinungen und in der Regel Ergebnisse empirischer Erhebungen und Untersuchungen, häufig im Rahmen der empirischen Sozialforschung (Empirie kommt übrigens von griechisch und stammt von dem Wort *empireia* ab und bedeutet etwas, wie Erfahrung, Erfahrungswissen). Diese Daten können unterschiedlich erhoben werden, ein Beispiel wären Volkszählungen. Die so gewonnenen Zahlen werden ausgewertet und - häufig in Form von Tabellen oder Diagrammen - veröffentlicht (bpb, 2010).

Thomas: Tabellen und Diagramme findet man ja wirklich sehr häufig, sei es in Artikeln oder auch in Newslettern oder auch in Nachrichtenformaten. Hast du eventuell einen Tipp, wie man diese Tabellen richtig deuten kann, wenn man richtig lesen kann?

Statistische Erhebungen



Statistische Erhebungen werden mit dem Zweck durchgeführt, von einer Grundgesamtheit (Population) an Erhebungseinheiten (z. B. der Grundgesamtheit der Wahlberechtigten) Informationen über interessierende Merkmale (z. B. das Wahlverhalten) zu erhalten. Ziel ist eine Bündelung von Informationen, die in den erhobenen Daten steckt. So wird erst durch die Auszählung der einzelnen abgegebenen Wahlzettel (Daten) und die Präsentation der Ergebnisse dieser Auszählung in Tabellenform und/oder als Säulen oder Kreisdiagramm (Bündelung der Informationen) ein nachvollziehbares Wahlergebnis.

Unter dem Begriff „Statistik“ werden alle Methoden der Datenanalyse und ihre Ergebnisse zusammengefasst. Die Statistik gliedert sich in drei große Bereiche.

- 1) Der erste ist die sogenannte beschreibende Statistik. Diese wird betrieben, wenn die Daten von einer interessierenden Grundgesamtheit wie einer Bevölkerung vollständig vorliegen, die man dann z. B. durch einen Mittelwert beschreiben möchte.
- 2) Der zweite Bereich ist die schließende Statistik. Diese hat die Aufgabe, auf Basis der Daten aus einem sorgfältig aus gewählten Teil einer Grundgesamtheit, der sogenannten Stichprobe, auf die interessierenden Kennzahlen der Grundgesamtheit rückzuschließen.
- 3) Die Wahrscheinlichkeitsrechnung schließlich wird für diese Rückschlüsse von Stichproben auf Grundgesamtheiten benötigt, da sie klarerweise mit gewissen Unsicherheiten behaftet sind.

Leonie: Ja, fangen wir doch zunächst einmal mit Tipps zum Lesen und Verstehen einer Tabelle an (bpb, 2010). Es empfiehlt sich

- Die Überschrift der Tabelle genau zu lesen: was wurde ausgewertet und zu welchem Zweck? In Verbindung mit der Kopf- und Randspalte der Tabelle enthält die Überschrift in der Regel die Informationen, die zum Verständnis der Tabelle über die Merkmale oder Items (= Fragen) und Antwortmöglichkeiten, deren Daten in der Tabelle angezeigt werden, erforderlich sind.
- Sich die Art der Zahlen bewusst zu machen: Werden absolute Zahlen (die Mengen, Größen, Häufigkeiten angeben) oder relative Zahlen (die einen Zahlenwert in Beziehung zu einer anderen Größe, meist der Grundgesamtheit, setzen) verwendet?
 - Ein Beispiel für die Verwendung von absoluten und relativen Zahlen wäre eine Umfrage zur Kundenzufriedenheit in einem Einkaufszentrum. Die Grundgesamtheit in diesem Fall wäre die Gesamtheit der Kunden, die das Einkaufszentrum besuchen.
 - o Absolute Zahlen würden die konkreten Angaben zur Zufriedenheit von Kunden wiedergeben, z. B. die Anzahl der Kunden, die sehr zufrieden, zufrieden, neutral, unzufrieden oder sehr unzufrieden sind.
 - o Relative Zahlen hingegen setzen diese Ergebnisse in Beziehung zur Grundgesamtheit. Zum Beispiel könnte man angeben, dass 80% der Kunden sehr zufrieden sind. Hierbei bezieht sich die relative Zahl (80%) auf die Grundgesamtheit aller Kunden im Einkaufszentrum.

Sabine: Und wie sieht es mit Diagrammen aus? Denen begegnet man gefühlt ja noch häufiger und sie fassen Informationen noch stärker zusammen, das kann manchmal schnell herausfordernd werden.

Leonie: Auch bei Diagrammen ist es wichtig sich Zeit zu nehmen und erst einmal zu orientieren.

- Was wird hier eigentlich dargestellt? Welche Zahlen sollen miteinander in Verbindung gebracht werden?
- Handelt es sich bei den Werten um absolute oder prozentuale Zahlen?
- Wird das gesamte Diagramm abgebildet oder nur ein Ausschnitt? Was ist der Maßstab? Diagramme dienen dazu, wichtige Informationen in statistischen Ergebnissen hervorzuheben, aber manchmal werden sie übermäßig verändert und sind daher nicht korrekt. Solche irreführenden Diagramme sind weit verbreitet in den Medien. Es gibt die Möglichkeit, die Skalierung zu manipulieren, indem Intervalle verkleinert werden oder Balken nicht bei Null beginnen (Statista, o.D).

Oft hilft auch ein Blick auf die Quellenangabe, die Auskunft über die Herkunft - und damit meist auch über die Glaubwürdigkeit - und das Alter der Daten gibt (bpb, 2010).

Thomas: Ich glaube, es wäre manchmal ganz hilfreich, gerade für Menschen mit einer Seheinschränkung, dass man zu einer Statistik auch einen entsprechenden Begleittext hätte, die eben dann auch das beschreibt, was man in einer Statistik dann eben sieht. Das heißt, wie ist es aufgebaut, was ist auf der X Achse abgebildet, was ist auf der Y Achse abgebildet? Über welche Sachen sprechen wir hier? Sind es ganze Werte? Sind es Prozentuale Werte? Und ich glaube, das würde ganz gut helfen, eine Statistik besser zu verstehen.

Leonie: Das ist ein guter Punkt, Thomas! Diese Texte geben z.B. Auskunft über die Grundgesamtheit, d.h. bei Befragungen über die Gesamtzahl der Befragten, und weitere Hintergrundinformationen, die für die richtige Einordnung und Bewertung der Ergebnisse wichtig sein können (bpb, 2010).

Sabine: Ich finde es wirklich wichtig, dass unsere Zuhörerinnen und Zuhörer verstehen, dass Statistiken zwar häufig wie wahre und objektive Fakten wirken mögen, aber die Sichtweise beeinflussen oder irreführend sein können, wenn sie Daten falsch darstellen oder wichtige Informationen fehlen (bpb, 2010).

Ich habe hierzu mal ein passendes Zitat mitgebracht:

„Statistik ist für mich das Informationsmittel der Mündigen. Wer mit ihr umgehen kann, kann weniger leicht manipuliert werden. Der Satz „Mit Statistik kann man alles beweisen“ gilt nur für die Bequemen, die keine Lust haben, genau hinzusehen.“ (Elisabeth Noelle-Neumann) (Statista, o.D)

Thomas: Das zeigt mir aber auch, dass Mittel Statistiken, teilweise falsche Aussagen verbreitet und untermauert werden können. Mit anderen Worten Man muss Statistiken stets mit einer gewissen Skepsis begegnen und natürlich auch Statistiken hinterfragen. Und hier gilt natürlich auch das alte Sprichwort Traue keiner Statistik, die du nicht selbst—naja, gefälscht würde ich nicht sagen, aber zumindestens manipuliert hast, um eben die Aussagen zu bekommen, die auch erreicht werden sollen.

Leonie: Ich finde es wichtig, noch einmal häufige Fehlerquellen von Statistiken hervorzuheben, auf die man beim Lesen wirklich immer achten sollte:

- Fehlerhafter Sprachgebrauch bei der Beschreibung der Zahlen im Begleittext: Wenn im Text z.B. „41% aller Befragten“ steht, die Prozentzahlen sich aber nur auf die Gruppe

der weiblichen Befragten beziehen, ist die Formulierung im Text falsch und führt zu Missverständnissen (insbesondere bei der Beschreibung von Zahlen in Kreuztabellen ist darauf zu achten, auf welche Basisgröße sich die Prozentwerte beziehen!)

- Eine andere wäre das Fehlerhafter Sprachgebrauch bei der Beschreibung von statistischen Begriffen: was genau wird mit Prozentangaben gemeint und was ist eigentlich ein Mittelwert?
- Außerdem ist ein häufiger Fehler bei der Dateninterpretation die Verwechslung von Kausalität und Korrelation. Eine Korrelation bedeutet eben nicht unbedingt, dass es eine Ursache-Wirkung-Beziehung gibt (bpb, 2010).

Thomas: Könntest du bitte noch einmal erklären, was genau eine Korrelation bedeutet?

Leonie: Korrelation bedeutet, dass zwei Dinge gemeinsam auftreten oder sich gemeinsam verändern können, aber nicht unbedingt, dass das eine das andere verursacht. Es ist wie bei einer Beziehung, bei der zwei Dinge gleichzeitig geschehen können, aber das eine nicht zwangsläufig das andere verursacht.

- Beispiel: Eine Studie findet eine positive Korrelation zwischen dem Verzehr von Eiscreme und der Anzahl der Badeunfälle im Sommer. Es wäre ein Fehler zu schlussfolgern, dass der Eiscreme-Verzehr Badeunfälle verursacht. In Wirklichkeit könnte die höhere Temperatur im Sommer sowohl den Eiscreme-Konsum als auch das Schwimmen im Freien erhöhen, was zu der beobachteten Korrelation führt. Hier besteht eine Korrelation, aber keine Kausalität (bpb, 2010).

Weitere wichtige Punkte beim Lesen von Statistiken sind:

- Fehlerhafte Verallgemeinerungen. Zum Beispiel werden die Schüler*innen aus meinem Anfangsbeispiel zu den Stunden, die sie in die Prüfungsvorbereitung investieren, befragt. Die Ergebnisse, die sich im Zusammenhang mit ihren Noten ergeben, können nicht für die gesamte Schule verallgemeinert werden.
- Oft wird auch eine Darstellungsform für die Daten gewählt, die den Betrachter bereits in eine bestimmte Richtung lenkt (bpb, 2010).

Definition statistischer Grundbegriffe



1. **Stichprobe:**
 - *Definition: Eine Teilmenge einer Population, die für statistische Untersuchungen ausgewählt wurde.*
2. **Population:**
 - *Definition: Die Gesamtheit aller Elemente oder Individuen, auf die sich eine statistische Studie bezieht.*
3. **Mittelwert:**
 - *Definition: Die Summe aller Werte in einer Datenmenge geteilt durch die Anzahl der Werte.*
4. **Standardabweichung:**
 - *Definition: Ein Maß für die Streuung oder Variabilität von Datenpunkten in einer Datenmenge.*
5. **Varianz:**
 - *Definition: Die durchschnittliche quadratische Abweichung der Werte von ihrem Mittelwert.*
6. **Regression:**
 - *Definition: Eine statistische Analysemethode, die die Beziehung zwischen einer abhängigen Variablen und einer oder mehreren unabhängigen Variablen untersucht.*
7. **Korrelation:**
 - *Definition: Ein statistisches Maß, das den Grad und die Richtung des Zusammenhangs zwischen zwei oder mehr Variablen beschreibt.*
8. **Signifikanzniveau:**
 - *Definition: Der Schwellenwert, bei dem ein statistischer Test als signifikant oder nicht signifikant betrachtet wird.*
9. **Hypothese:**
 - *Definition: Eine aufgestellte Annahme, die in statistischen Tests überprüft wird.*

7. FEHLER AUFSPÜREN & ERKENNEN

Sabine: Eine Person, die sich sehr ausführlich mit den Fallstricken im Verstehen von Statistiken beschäftigt, ist der Psychologe Gerd Gigerenzer, der selbst schon in zahlreichen Podcasts darüber gesprochen hat, wie schwer es uns allen fällt, Statistiken richtig zu deuten und eigene Schlussfolgerungen aus den Daten zu ziehen (RWI Essen, 2023a).

Leonie: Du willst sicherlich auf die Unstatistik des Monats hinaus, richtig?

Thomas: Unstatistik des Monats. Davon habe ich im Prinzip noch nie etwas gehört. Könnt ihr etwas noch mehr dazu sagen?

Sabine: Gerne! Ich habe hier ein Unstatistik des Monats mitgebracht:

Unstatistik des Monats



„Der Berliner Psychologe Gerd Gigerenzer, der Bochumer Ökonom Thomas Bauer und der Dortmunder Statistiker Walter Krämer haben im Jahr 2012 die Aktion „Unstatistik des Monats“ ins Leben gerufen. Sie werden jeden Monat sowohl jüngst publizierte Zahlen als auch deren Interpretationen hinterfragen. Die Aktion will so dazu beitragen, mit Daten und Fakten vernünftig umzugehen, in Zahlen gefasste Abbilder der Wirklichkeit korrekt zu interpretieren und eine immer komplexere Welt und Umwelt sinnvoller zu beschreiben.“

Alle Unstatistiken finden Sie im Internet auch unter www.unstatistik.de und unter dem Twitter-Account [@unstatistik](https://twitter.com/unstatistik).“ (RWI Essen, 2023a)

Thomas: Habt ihr auch ein Beispiel für eine Unstatistik des Monats und wie man sie am besten entdeckt und wie sie klar wird?

Sabine: Ich habe als Beispiel einmal die Unstatistik des Monats Juni 2023 mitgebracht. Das ZDF und die FAZ titeln: Studie der Hamburger Uni-Klinik: Klimawandel führt zu mehr Frühgeburten. Und diese Schlussfolgerung basiert auf den Daten der Studie Erhöhtes Risiko einer späten Frühgeburt aufgrund von Hitzestress (Yüzen et al., 2022), die in der Fachzeitschrift The Lancet erschienen ist.

Informationen zur Studie

In den Pressemeldungen heißt es: Die Studie analysierte 26.000 Geburten zwischen 1999 und 2021 in den Sommermonaten (Mai bis September) und berücksichtigte die Temperaturen. Die Ergebnisse zeigten, dass das Risiko von Frühgeburten um 45 Prozent anstieg, wenn es in den letzten beiden Tagen vor der Geburt außergewöhnlich heiß war, im Vergleich zu Geburten, bei denen höchstens ein sehr warmer Tag voranging. „Außergewöhnlich heiß“ bedeutet, dass die gefühlte Temperatur höher war als in 99 Prozent der Fälle, was etwa 35 Grad entspricht. Dabei wurde nicht nur die gemessene Temperatur, sondern auch die Luftfeuchtigkeit berücksichtigt.

Kritische Analyse:

In den Daten sind insgesamt 2.181 Frühgeburten erfasst. Die Studie gibt jedoch nicht genau an, wie viele davon an heißen oder sehr heißen Tagen auftraten. Sie erwähnt jedoch, dass 32 Frühgeburten nach zwei aufeinanderfolgenden sehr heißen Tagen auftraten. Das sind 45 Prozent mehr als zu erwarten gewesen wäre, wenn die Temperatur keinen Einfluss gehabt hätte. In absoluten Zahlen bedeutet dies 10 zusätzliche Frühgeburten in 22 Jahren.

Die Frühgeburtenrate stieg damit von etwa 8,5 Prozent auf 12,3 Prozent, was einer absoluten Zunahme von 3,8 Prozentpunkten entspricht. Zusammenfassend lässt sich sagen, dass nach einem Prozent der wärmsten Tage 3 bis 4 Frauen mehr pro 100 Schwangere eine Frühgeburt hatten (RWI Essen, 2023b).

The screenshot shows the website 'Unstatistik des Monats' with the following content:

- Header:** HARDING-ZENTRUM FÜR RISIKOKOMPETENZ. Navigation: Das Harding-Zentrum, Forschung, Transfer und Nutzen. Social media icons and language options (DE, EN).
- Breadcrumbs:** Startseite > unstatistik > Unstatistik des Monats: Der Klimawandel führt zu mehr Frühgeburten.
- Title:** Unstatistik des Monats: Der Klimawandel führt zu mehr Frühgeburten.
- Main Text:**
 - Die Unstatistik des Monats Juni sind Meldungen des ZDF und der FAZ zu einer Studie des Universitätsklinikums Hamburg-Eppendorf (UKE). Die in der angesehenen Fachzeitschrift „The Lancet“ veröffentlichte Studie verwendet Daten von knapp 26.000 Geburten aus den Sommermonaten Mai bis September der Jahre 1999 bis 2021, die mit Temperaturdaten verknüpft wurden. Die Ergebnisse legen nahe, dass das Risiko von Frühgeburten um 45 Prozent anstieg, wenn die letzten zwei Tage vor der Geburt außergewöhnlich heiß waren, verglichen mit Geburten, denen höchstens ein besonders warmer Tag vorangegangen war. „Außergewöhnlich heiß“ bedeutet dabei, dass die gefühlten Temperaturen höher waren als in 99 Prozent der Fälle. Das dürfte den in der Pressemitteilung genannten 35 Grad entsprechen. Bei der gefühlten Temperatur wird neben der tatsächlich gemessenen Temperatur auch die Luftfeuchtigkeit berücksichtigt.
 - Wichtige Zusatzinformationen kann man dem [Anhang der Studie](#) entnehmen. Insgesamt gab es 2.181 Frühgeburten in den Daten des UKE. Leider schlüsseln die Studienautoren nicht exakt auf, wie viele dieser Frühgeburten an wie vielen Tagen mit durchschnittlichen, heißen oder außergewöhnlich heißen Tagen auftraten.
 - Die Studie gibt allerdings an, dass 32 Frühgeburten nach zwei aufeinanderfolgenden, außergewöhnlich heißen Tagen auftraten. Dies sind 45 Prozent mehr, als man erwarten würde, wenn es keinen Zusammenhang mit den Temperaturen gäbe, d.h. wenn an den 1 Prozent heißesten Tagen auch 1 Prozent der Frühgeburten stattgefunden hätten. Es geht also in absoluten Zahlen um 10 zusätzliche Frühgeburten in 22 Jahren.
- Image:** A small image of a newborn baby's hand being held.
- Caption:** Bild von Lisa Fotios auf pexels.
- Sidebox: Die Unstatistik des Monats**
 - Der Berliner Psychologe Gerd Gigerenzer, der Bochumer Ökonom Thomas Bauer und der Dortmunder Statistiker Walter Krämer haben im Jahr 2012 die Aktion „Unstatistik des Monats“ ins Leben gerufen. Sie werden jeden Monat sowohl jüngst publizierte Zahlen als auch deren Interpretationen hinterfragen. Die Aktion will so dazu beitragen, mit Daten und Fakten vernünftig umzugehen, in Zahlen gefasste Abbilder der Wirklichkeit korrekt zu interpretieren und eine immer komplexere Welt und Umwelt sinnvoller zu beschreiben. Seit August 2018 ist Katharina Schüller, Geschäftsführerin und Gründerin von STAT-UP, Mitglied des Unstatistik-Teams. Alle Unstatistiken finden Sie im Internet auch unter www.unstatistik.de und unter dem Twitter-Account [@unstatistik](https://twitter.com/unstatistik).

Abbildung 1: Screenshot Unstatistik des Monats, RWI Essen (2023)

Leonie: Ich übernehme einmal und lese weiter:

- „Der Klimawandel führt zu mehr Frühgeburten“, „bei andauernden Temperaturen von über 35 Grad steige das Risiko auf 45 Prozent“ warnt der ZDF-Artikel. Der Studie zufolge führe „Hitzestress ausgelöst durch Temperaturen von 30 Grad zu einem 20-prozentigen Anstieg von Frühgeburten, ab 35 Grad ist beinahe jede zweite Schwangere betroffen.“
- Auch die FAZ schreibt in ihrer Meldung „Hitze führt vermehrt zu Frühgeburten“: „Hitzestress von 30 Grad erhöhte das Risiko einer Frühgeburt, also vor der 37. Schwangerschaftswoche, um 20 Prozent; bei Temperaturen von mehr als 35 Grad lag das Risiko sogar bei fast 45 Prozent.“

Hier möchte ich einmal ganz deutlich den Fehler hervorheben:

- Ein Anstieg des relativen Risikos um 45 Prozent, ist kein Anstieg des Risikos auf 45 Prozent. Es bedeutet nicht, dass beinahe jede zweite Schwangere betroffen ist. ZDF und FAZ haben in ihren Meldungen das relative Risiko mit einem absoluten Risiko verwechselt. Es sind 3 bis 4 mehr Schwangere von je hundert, nicht 45 Prozent, wohlgeachtet: nach extremen, seltenen Hitzeereignissen.

Sabine: Das hier finde ich auch spannend:

- Die Aussage, dass laut der Studie „in zehn Jahren bei fortschreitender Häufigkeit von heißen Sommertagen bereits jedes 6. Kind in Deutschland zu früh auf die Welt kommt - das wären doppelt so viele Frühgeburten wie heute“, mag zwar von den UKE-Forschern prognostiziert worden sein, führt aber ebenfalls weitgehend in die Irre.
- Denn es handelt sich eben erstmal nur um bloße Annahmen, wie sich die Sommertemperaturen in Zukunft entwickeln werden. Außerdem kommen von den zu früh geborenen Kindern die allermeisten (65 Prozent) in den letzten 1,5 Monaten vor dem errechneten Geburtstermin zur Welt. Von reif geborenen Kindern unterscheiden sie sich im Hinblick auf Gewicht und Größe nur wenig, auch wenn ihnen wichtige Entwicklungszeit im Mutterleib fehlt. Die hoch emotionalen Bilder und Beispielgeschichten, die das ZDF in seinem Beitrag verwendet, bilden jedenfalls nicht die Gruppe der späten Frühgeborenen ab, sondern schüren Angst und Panik.

Thomas: Ich habe den Eindruck, als würde ein relatives Risiko zu einem absoluten Risiko gemacht werden

und damit kommt es zu einer deutlichen Verzerrung. Und hier wird ein Problem quasi oder hier entsteht ein Problem, das im Prinzip gar nicht in der Größenordnung existiert.

Sabine: Leider sagen Gerd Gigerenzer und Kollegen, dass das UKE an dieser dramatisierenden Berichterstattung nicht ganz unschuldig ist. In der Pressemitteilung zur Studie wird der Effekt der Hitze nur als relative Zahl (45 Prozent mehr) und nicht als absoluter Anstieg (3 bis 4 von je 100 Frauen mehr) berichtet.

Leonie: Und da hat man eben genau diesen Effekt: Relative Risiken sind große Zahlen und machen dann auch schnell unnötige Angst; absolute Risiken sind kleine Zahlen und können dabei helfen, das tatsächliche Risiko zu verstehen. Es bietet sich natürlich für eine Schlagzeile an, diese beeindruckenderen Zahlen zu nutzen.

Thomas: Und wer steht hierfür in der Verantwortung? Wir haben ja heute schon sehr viel über das Thema Verantwortung mit Daten und mit Statistiken gehört. Aber wie sieht das eigentlich aus in der Kompetenz beispielsweise von Akademikern oder natürlich auch von Mitarbeitern in der Mediengesellschaft? Wie schaut es dort mit Kompetenz aus? Was die Darstellung und die Verbreitung von Statistiken anbetrifft?

Sabine: Der Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e. V. (ADM) meint: „Die Verantwortung für eine korrekte Kommunikation liegt bei denjenigen, die die Ergebnisse veröffentlichen. Gleichzeitig sollten die Medien die Ergebnisse kritisch hinterfragen und sich über ihre Entstehung informieren, insbesondere wenn eine große öffentliche Aufmerksamkeit vermutet wird oder es sich um kontroverse Themen handelt. Dies erfordert natürlich ein hohes Maß an Datenkompetenz. Wenn es Zweifel an der Aussagekraft einer Studie gibt, muss von einer weiteren Verbreitung abgesehen werden.“ (RWI Essen, 2023b)

Thomas: Das heißt also, dass in diesem Falle sowohl das ZDF als auch die FAZ auf diese Art der Berichterstattung eigentlich hätte verzichten müssen.

Leonie: Genau, das heißt es. Das zeigt einmal wieder, dass Datenkompetenz und der verantwortungsvolle Umgang mit Statistiken uns alle angeht. Umso wichtiger, dass sich unsere Zuhörer in Zukunft trauen, auch einmal kritisch nachzufragen, was denn hier eigentlich beschrieben wird.

8. ABSCHLUSS

Leonie: Vielen Dank für diesen Einblick in eure Gedanken zu Partizipationsmöglichkeiten. Ich denke, es ist wirklich wichtig, Patienten und Betroffene dazu zu ermutigen, sich die Frage zu stellen: wo kann ich mich hier einbringen, um letztlich auch zu verstehen, was mit meinen eigenen Daten passiert. Gleichzeitig sollten natürlich auch Forschende immer wieder überlegen, wo die Perspektive der Betroffenen wertvoll ist und eine Zusammenarbeit Sinn macht.

Sabine: Es braucht natürlich auch das Bemühen von uns Forschenden, Betroffenen Einblicke in Forschungsstudien zu ermöglichen und Barrieren für die Zusammenarbeit abzubauen. Zum Beispiel, dass man sich bemüht, die einzelnen Schritte in einem Forschungsprojekt zu erklären und dass ein gemeinsames Verständnis von Forschung entstehen kann.

Thomas: Ich glaube, es ist extrem wichtig, dass wir als Patienten oder Beteiligte an einer Studie auch nachher natürlich wissen, was geschieht mit den Daten, wie werden sie ausgewertet? Und das ist wichtig auch für das Vertrauen zueinander und füreinander.

Leonie: Das sind noch mal wichtige Gedanken, die einen guten Abschluss für unsere heutige Podcast-Episode bilden und gleichzeitig auch schon eine schöne Überleitung für unsere nächste Episode sind: da wollen wir

uns nämlich genauer anschauen, welche Rolle Daten in der Forschung spielen und darüber sprechen, was mit den Daten in einem Forschungsprojekt passiert und wie man die tollen Grafiken und Statistiken, die dann am Ende oft veröffentlicht werden, besser und leichter verstehen kann. Damit man am Ende eben beurteilen kann: Ist das für mich relevant? Und was sagt diese Statistik eigentlich aus? Wir würden uns sehr freuen, wenn Sie, liebe Zuhörerinnen und Zuhörer, auch bei der nächsten Folge dabei sind. Vielen Dank fürs Zuhören und bis zum nächsten Mal.

Sabine: Vielen Dank und Tschüss!

Thomas: Ja, vielen Dank, dass ich heute hier dazu beitragen konnte, die Sicht von Patienten und meine persönliche Sicht und auch aus Sicht der PRO RETINA noch einmal vorzutragen. Ich freue mich auf die nächste Episode und dann viel Spaß beim Zuhören.

Schlussmusik

9. ZUSAMMENFASSUNG UND SHOWNOTES

Links

Bundesagentur für Arbeit. (2019). Statistical Literacy - Statistik verstehen und richtig interpretieren. <https://statistik.arbeitsagentur.de/DE/Navigation/Grundlagen/Statistical-Literacy/Statistical-Literacy-Nav.html>

Bundeszentrale für politische Bildung (bpb). 2010. M 03.05 "Wie liest man eine Statistik?". <https://www.bpb.de/lernen/angebote/grafstat/klassencheckup/46358/m-03-05-wie-liest-man-eine-statistik/>

Statista. o.D. Definition Lügen mit Statistiken. https://de.statista.com/statistik/lexikon/definition/8/luegen_mit_statistiken/#:~:text=Mit%20den%20Worten%20der%20deutschen,%E2%80%9C

RWI Leibniz Institut für Wirtschaftsforschung (RWI). 2023a. Unstatistik des Monats. <https://www.rwi-essen.de/presse/wissenschaftskommunikation/unstatistik>

RWI Leibniz Institut für Wirtschaftsforschung (RWI). 2023b. Der Klimawandel führt zu mehr Frühgeburten. <https://www.rwi-essen.de/presse/wissenschaftskommunikation/unstatistik/detail/der-klimawandel-fuehrt-zu-mehr-fruehgeburten>

10. LITERATUR

Benesch, T. (2013). Grundbegriffe der Statistik. In: Schlüsselkonzepte zur Statistik. Spektrum Akademischer Verlag, Heidelberg. https://doi.org/10.1007/978-3-8274-2772-4_1

Bortz, J. & Döring, N. (2015). Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler: Mit 87 Tabellen (Limitierte Sonderausg., 4., überarb. Aufl.). Springer-Lehrbuch. Springer.

Bundeszentrale für politische Bildung (bpb). (2010). M 03.05 "Wie liest man eine Statistik?". <https://www.bpb.de/lernen/angebote/grafstat/klassencheckup/46358/m-03-05-wie-liest-man-eine-statistik/>

Cohen. (2022). Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences (3. Auflage). Routledge; Taylor & Francis Group. <https://doi.org/10.4324/9780203774441>

Field, A. (2013). Discovering Statistics Using IBM SPSS Statistics: And Sex and Drugs and Rock "N" Roll, 4th Edition, Sage, Los Angeles, London, New Delhi.

Flick, U. (2019). Gütekriterien qualitativer Sozialforschung. In: Baur, N., Blasius, J. (eds) Handbuch Methoden der empirischen Sozialforschung. Springer VS, Wiesbaden. https://doi.org/10.1007/978-3-658-21308-4_33

Montgomery, D. C. (2017). Design and analysis of experiments (Ninth edition). John Wiley & Sons, Inc.

Perkhofer, S., Stamm, T., Ritschl, V., Hirmann, E., Huber, A., Unterhumer, G., Oberhauser, H., Weigl, R., Jocham, A., Moser, D., Ameshofer, L. & Neurer, S. (2016). Quantitative Forschung. In: Ritschl, V., Weigl, R., Stamm, T. (eds) Wissenschaftliches Arbeiten und Schreiben. Studium Pflege, Therapie, Gesundheit. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-49908-5_7

Ott, R. L. & Longnecker, M. (2001). An Introduction to statistical methods and data analysis (5. ed.). Thomson.

Statista. (o.D). Definition Lügen mit Statistiken. https://de.statista.com/statistik/lexikon/definition/8/luegen_mit_statistiken/#:~:text=Mit%20den%20Worten%20der%20deutschen,%E2%80%9C

RWI Leibniz Institut für Wirtschaftsforschung (RWI). (2023a). Unstatistik des Monats. <https://www.rwi-essen.de/presse/wissenschaftskommunikation/unstatistik>

RWI Leibniz Institut für Wirtschaftsforschung (RWI). (2023b). Der Klimawandel führt zu mehr Frühgeburten. <https://www.rwi-essen.de/presse/wissenschaftskommunikation/unstatistik/detail/der-klimawandel-fuehrt-zu-mehr-fruehgeburten>

Yüzen, D., Graf, I., Tallarek, A.-C., Hollwitz, B., Wiessner, C., Schleussner, E., Stammer, D., Padula, A., Hecher, K., Arck, P. C. & Diemert, A. (2023). Increased late preterm birth risk and altered uterine blood flow upon exposure to heat stress. EBioMedicine, 93, 104651. <https://doi.org/10.1016/j.ebiom.2023.104651>

