

# Sag Adieu zu hohem Blutdruck: Positive Beziehung zwischen dunkler Schokolade und Bluthochdruck

VON ERIC D. NORDMOE

Übersetzt, bearbeitet und ergänzt von Manfred Borovcnik und Gerhard König

(Original: Kiss High Blood Pressure Goodbye: The Relationship between Dark Chocolate and Hypertension. In: Teaching Statistics Volume 30, Number 2, Summer 2008, S. 34-38)

*Zusammenfassung: In diesem Artikel wird eine in Deutschland durchgeführte kontrollierte randomisierte Studie vorgestellt nach der regelmäßiger Verzehr dunkler Schokolade den Blutdruck senken kann. Diese Studie wird zuerst kurz mit ihren statistischen Grundlagen vorgestellt, danach wird diskutiert, wie diese im Statistik-Unterricht eingesetzt werden kann, um wichtige Themen wie Versuchsplanung oder statistisches Schließen in der Praxis kennenzulernen.*

## Einleitung

Regelmäßiges Essen dunkler Schokolade kann den Blutdruck senken. Das betont der Pharmakologe Dirk Taubert von der Universitätsklinik Köln als Folgerung aus einer klinischen Studie im „Journal of the American Medical Association“ (JAMA; Bd. 298, S. 49). Bei weißer Schokolade stellte sich diese Wirkung nicht ein. Im Folgenden soll diese Studie kurz vorgestellt werden, danach wollen wir diskutieren, wie diese produktiv im Unterricht eingesetzt werden kann.

## Hintergrundinformationen

Der ideale Blutdruck liegt nach Angaben der Hochdruckliga bei 120/80 mmHg. Von Bluthochdruck (Hypertonie oder Hypertonus) spricht man, wenn der Druck in den Arterien krankhaft auf einen systolischen Wert von über 140 mmHg und einen diastolischen Wert über 90 mmHg gesteigert ist. ([www.hochdruckliga.de/Leitlinien-Therapie2007.pdf](http://www.hochdruckliga.de/Leitlinien-Therapie2007.pdf), S. 5)

Die Entscheidung, ob der Blutdruck behandlungsbedürftig ist oder nicht, hängt nicht nur von der Druckhöhe ab, sondern vom Gesamtrisiko für einen Herzinfarkt oder Schlaganfall. Eine entscheidende Rolle spielen weitere Krankheiten, beispielsweise das Metabolische Syndrom oder ein Typ-2-Diabetes.

Klassifikation	Blutdruck (mmHg)	
	systolisch	diastolisch
Optimal	< 120	< 80
Normal	< 130	< 85
Hoch normal	130 – 139	85 – 89
Bluthochdruck	≥ 140	≥ 90
Stufe 1 (Hypertonie leicht)	140-159	90-99
Stufe 2 (Hypertonie mittelschwer)	160-179	100-109
Stufe 3 (Hypertonie schwer)	≥180	≥110

**Tabelle 1: Definitionen Bluthochdruck**

Aus mehreren neueren Berichten und Studien (z. B. Kanavos et al. 2007, [1]) ist Bluthochdruck weltweit als gefährliche Krankheit anerkannt. Hohe Anteile von Bluthochdruck bei Erwachsenen werden aus mehreren Industrieländern berichtet, z. B. 23% in den USA, 39% in England, 45% in Spanien und 55% in Deutschland. Neuere Forschungsberichte prognostizieren einen dramatischen Anstieg der Erkrankung Bluthochdruck mit ungefähr 1,56 Milliarden Erwachsener bis zum Jahr 2025.

Zur Senkung des Bluthochdrucks werden Ernährungsumstellungen, regelmäßiger Ausdauersport und die Einnahme geeigneter Medikamente empfohlen. Während Bluthochdruckpatienten noch geneigt sind, Medikamente einzunehmen, sind die Empfehlungen, mehr Sport zu treiben und/oder die Ernährung umzustellen, meist weniger effektiv, weil diese erfahrungsgemäß von den Patienten nicht dauerhaft eingehalten werden. Gesundheitspolitiker und Forscher sind daher bestrebt, nach weiteren geeigneten Methoden zu suchen, die auch von Patienten langfristig angenommen werden.

Frühere Untersuchungen des Pharmakologen Taubert aus dem Jahr 2003 ([2]) hatten bereits darauf hingewiesen, dass große Mengen Kakaoprodukte den Blutdruck senken können. Als Ursache gelten die im Kakao enthaltenen Polyphenole, eine Gruppe von Farb- und Geschmacksstoffen, die in zahlreichen Pflanzen vorkommen. In der klinischen Studie des Jahres 2003 wurden jedoch stets sehr große Mengen Schokolade (bis 100g täglich) für nur relativ kurze Zeit (max. 2 Wochen) verabreicht. Aus ernährungstechnischer Sicht sind diese Mengen Schokolade natürlich als dauerhafte Gabe nicht vertretbar: Der positive Effekt

würde durch die meist hohen Zucker- und Fettanteile in Kakaoprodukten wieder aufgehoben. Unklar war Taubert zufolge, wie sich eine niedrig dosierte Gabe von dunkler Schokolade mit reichlich Polyphenolen auf den Blutdruck auswirken würde. Im Folgenden soll daher die neuere klinische Studie aus dem Jahr 2007, s. [3], besprochen und analysiert werden.

## Dunkle Schokolade und Bluthochdruck – Die Studie: Bitterschokolade statt bitterer Pillen

Das hauptsächliche Anliegen der neueren, hier vorgestellten Folgestudie, war, die Einschränkungen der früheren Studie aufzuheben und deren Ergebnisse zu verallgemeinern; ferner lagen viele weitere Berichte anderer Studien mit ähnlichen gesundheitlichen Vorteilen des Kakaoverzehrs vor. Um allgemeinere Ergebnisse und Aussagen zu gewinnen, führten die Autoren ein neues kontrolliertes, randomisiertes Experiment durch.

Für eine „klinische Studie unter ärztlicher Leitung“ zum Thema „Wie wirkt sich diätetisches Essverhalten auf die Gesundheit aus“ wurden dazu im Jahr 2006 in Duisburg Frauen und Männer im Alter über 55 Jahre gesucht. Es meldeten sich 119 Interessenten. Von diesen wurden die Teilnehmer an der Untersuchung in Duisburg wie folgt ausgewählt:

1. Sie sollten weiteres Interesse an einer Teilnahme haben;
2. im Alter zwischen 55 und 75 sein;
3. gesund sein, bis auf die Ausnahme eines leichten Bluthochdrucks zwischen 130/85 und 160/100; sie sollten keinerlei Medikamente gegen hohen Blutdruck einnehmen;
4. zu einer wie auch immer bestimmten höheren sozioökonomischen Schicht angehören (ein Kriterium war ein Haushaltsnettoeinkommen von mindestens 20.000 €/Jahr);
5. Nichtraucher sein.

Diese Kriterien (Genauerer s. S. 50 der Originalstudie) sollten die Probanden untereinander besser vergleichbar machen. Von den gemeldeten Interessenten blieben schließlich 44 Teilnehmer übrig. Diese wurden randomisiert, das heißt zufällig aufgeteilt auf zwei Gruppen zu je 22 Teilnehmern, in eine Versuchsgruppe (Behandlungsgruppe) und eine Kontrollgruppe.

An dem Versuch nahmen also 44 Probanden (24 Frauen, 20 Männer) mit leicht erhöhtem Blutdruck teil.

Die Versuchsteilnehmer aßen 18 Wochen lang täglich 6,3 Gramm dunkle Schokolade (Ritter Sport, Halb-bitter), die anderen bekamen 5,8 g weiße Schokolade (Milka weiße Schokolade). In der ersten Gruppe sank der Blutdruck im Mittel um 2,9 Millimeter Quecksilbersäule (mmHg) in der Pumpphase des Herzens (systolischer Wert) und um 1,9 mmHg, während das Herz sich wieder mit Blut füllt (diastolischer Wert). Weiße Schokolade veränderte die Blutdruckwerte nicht. Dagegen hatten nach 18 Wochen alle Patienten der Behandlungsgruppe einen tieferen systolischen oder diastolischen Blutdruckwert. Keiner der Patienten konnte allerdings in eine tiefergradige hochnormale Blutdrucksituation oder gar normale Blutdrucksituation gebracht werden.

In der Studie habe sich nun ein kommerzielles Kakaoprodukt in kleiner Menge als ebenso wirksam erwiesen wie bewährte Ernährungsumstellungen, berichten die Kölner Ärzte. „Auch wenn die Blutdrucksenkung klein ausfiel, ist sie doch klinisch bemerkenswert“, betont Taubert. Schätzungen zufolge verringert nämlich eine durchschnittliche Senkung des Blutdrucks um 3 mmHg in der Bevölkerung das Todesrisiko durch Schlaganfall um acht Prozent und jenes durch Herzinfarkt um fünf Prozent.

Nun genauere Daten und Analysen aus der Studie. Der Blutdruck aller Patienten wurde zu Beginn (das ist der Basiswert, im Jargon auch Baseline genannt) und nach 6, 12 und 18 Wochen gemessen. Dass die Randomisierung, die computergeneriert durchgeführt wurde, erfolgreich war, zeigt sich auch darin, dass keine signifikanten Unterschiede ( $\alpha = 0,05$ ) im gemessenen Ausgangswert für den Blutdruck zwischen den Teilnehmern der Versuchsgruppe und der Kontrollgruppe gefunden wurden. Es gab auch keine Geschlechtsunterschiede. (s. Tabelle 2)

<b>Versuchs- personen</b>	<b>Mittelwert (Standardabweichung)</b>	
	<b>Dunkle Schokolade (n = 22)</b>	<b>Weiß Schokolade (n = 22)</b>
Systolisch	147,7 (7,1)	147,5 (8,0)
Frauen	147,3 (8,5)	146,4 (9,0)
Männer	148,3 (5,3)	148,9 (6,9)
Diastolisch	86,4 (4,1)	86,7 (3,8)
Frauen	86,0 (4,4)	87,3 (4,3)
Männer	86,8 (3,8)	86,0 (3,1)

Tabelle 2: Ausgangsblutdruck in Versuchsgruppe und Kontrollgruppe (in mmHg), nach Taubert et al. (2007)

Es wurden Zweistichproben t-Tests (s. Henze, [4], S. 318) für paarige Beobachtungen verwendet, um die

Mittelwerte des Blutdrucks zu Beginn mit jenen nach 12 und 18 Wochen zu vergleichen. Dabei fanden die Forscher signifikante ( $p < 0,001$ ) Blutdrucksenkungen für die Teilnehmer der Versuchsgruppe, jedoch nicht für die Kontrollgruppenteilnehmer. (s. Tabelle 3)

Versuchspersonen	Mittelwert (Standardabweichung)	
	Dunkle Schokolade (n = 22)	Weißer Schokolade (n = 22)
Systolisch	-2,9 (1,6)	0,1 (1,6)
Diastolisch	-1,9 (1,0)	0,0 (1,8)

Tabelle 3: Änderung im Blutdruck vom Ausgangsniveau zur Woche 18 (in mmHg) für Behandlungs- und Kontrollgruppe aus Taubert et al. (2007)

Bemerkung: Normalerweise wählt man die Irrtumswahrscheinlichkeit  $\alpha$  (Fehler 1. Art) eines Tests vor der Datenerhebung und berechnet – davon abhängig – die kritische Zahl, bei deren Überschreiten die Nullhypothese abgelehnt wird. Wenn jetzt diese Ablehnzahl „weit“ übertroffen wird, geht etwa die

Information verloren, dass auch ein viel kleineres Signifikanzniveau  $\alpha$  zur Ablehnung der Nullhypothese geführt hätte. Diesen Informationsverlust vermeidet man heute durch die Angabe von p-Werten: Wie groß wäre das Signifikanzniveau, das genau den beobachteten Wert der Testgröße zur Ablehnzahl berechnet hätte. Der beobachtete Wert der Testgröße wird also zum „Wendepunkt“ der Entscheidung: bei Unterschreiten dieses Wertes erfolgt keine Ablehnung, bei dessen Überschreiten wird die Nullhypothese abgelehnt.

Der p-Wert gibt also indirekt an, wie extrem der beobachtete Wert der Teststatistik ist. Je kleiner der p-Wert, desto eher sollte die Nullhypothese verworfen werden. Ist der p-Wert, wie hier, kleiner als 0,001, so hätte auch die Wahl von  $\alpha = 0,001$  zur Ablehnung der Nullhypothese geführt.

Mittels linearer Regression und Korrelation wurde zusätzlich festgestellt, dass Blutdrucksenkungen in der Versuchsgruppe tendenziell höher bei Patienten

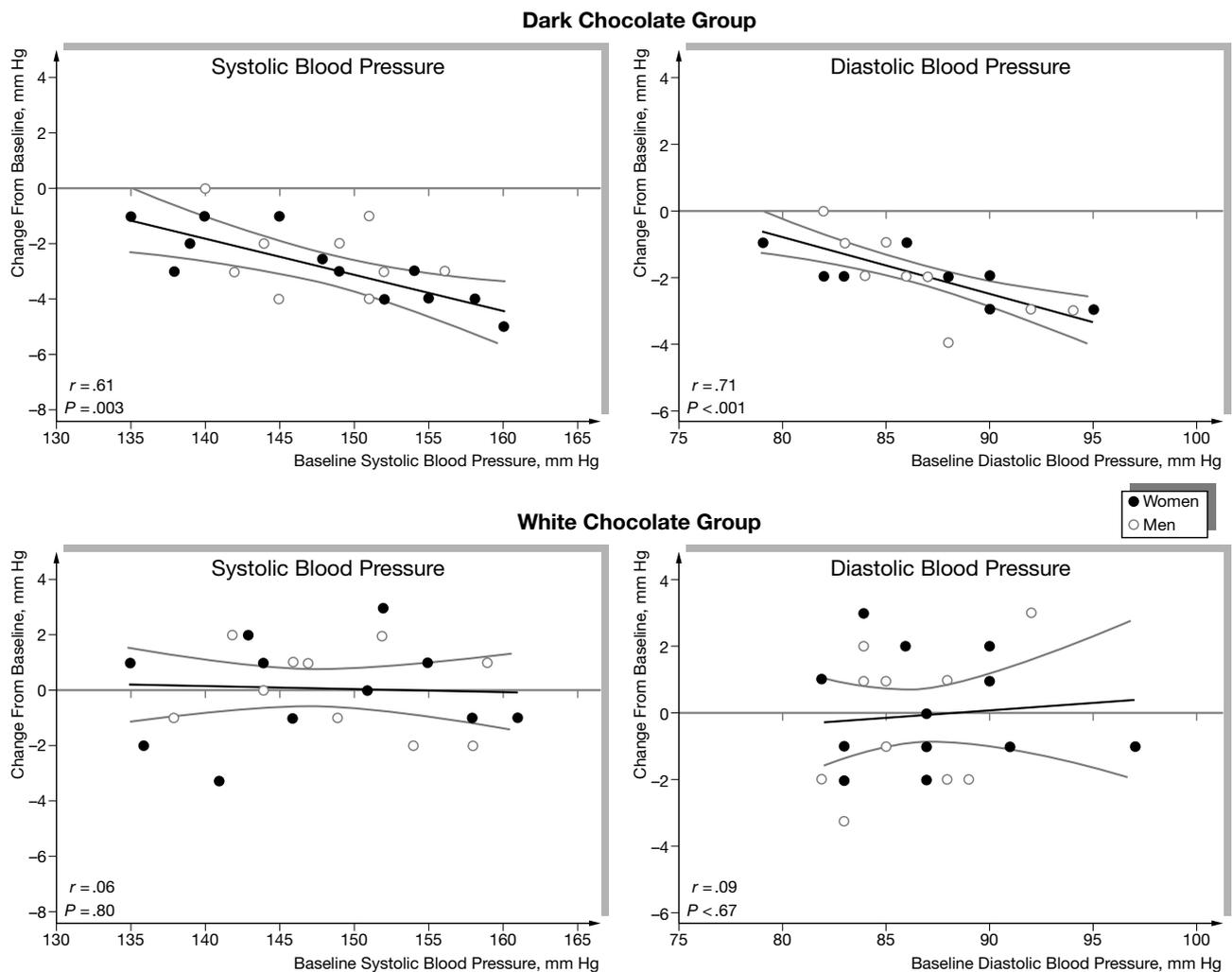


Fig. 1

mit höherem Blutdruck ausfielen. Die beiden Grafiken aus der Originaluntersuchung von Taubert zeigen den Zusammenhang; dabei sind die Korrelationskoeffizienten  $r = -0,61$  für systolischen Blutdruck ( $p = 0,003$ ) und  $r = -0,71$  für diastolischen Blutdruck ( $p < 0,001$ ) signifikant. Änderungen im Blutdruck waren hingegen in der Kontrollgruppe nicht signifikant korreliert ( $r = -0,06$  bzw.  $0,09$ ) mit den Anfangswerten ( $p$ -Werte von  $0,80$  bzw.  $0,67$ ).

Die Autoren schlussfolgern, dass eine niedrig dosierte Einnahme von dunkler Schokolade und der darin enthaltenen Polyphenole eine blutdrucksenkende Wirkung hat. Gestützt auf die Ergebnisse dieser Studie könnte gesagt werden, dass die diätetische Empfehlung zur Blutdrucksenkung im Falle von leichtem Bluthochdruck um eine zumindest niedrig dosierte Menge an dunkler Schokolade erweitert werden könnte, was sicherlich viele der Betroffenen freuen dürfte. Eine exakte Definition von „niedriger Dosierung“ müsste aber auf jeden Fall noch festgelegt werden, um gegenteilige Effekte wie zum Beispiel Gewichtszunahmen, zu verhindern.

## Diskussionsvorschläge für den Unterricht

Wenn Blutdruckpatienten von dieser Untersuchung erfahren, werden sie sich ohne Zweifel bald an ihre Krankenkasse wenden, um zu erfragen, ob die Kosten für ihren Schokoladenverzehr übernommen werden könnten. Wir als Unterrichtende im Fach Stochastik können uns überlegen, welche Relevanz diese Studie für den Unterricht hat. Während Bluthochdruck im Allgemeinen kein Problem für unsere Schüler und Studenten selbst ist, können gleichwohl nahe Verwandte davon betroffen sein. Die Einfachheit der Versuchsplanung verbunden mit der Gelegenheit, den lebenslangen Verzehr dunkler Schokolade gesundheitlich zu rechtfertigen, könnte für die Motivation im Unterricht ausreichen.

In diesem Abschnitt sollen einige Probleme andiskutiert werden, die ihren Schülern und Studenten einen bitteren Geschmack in ihrem Mund (im Sinn dunkler Schokolade) hoffentlich zurücklassen werden. Dazu werden die relevanten Aspekte in der Studie [3] von Taubert et al. skizziert. Studierende sollten Gelegenheit bekommen, kritische Fragen zum Design der Studie zu stellen, wobei man auch die praktischen und finanziellen Einschränkungen, unter denen Forscher arbeiten, nicht vergessen darf.

## 1. Wie lautet die Forschungsfrage in der Studie?

Frühere Untersuchungen, ebenfalls von Taubert im Jahre 2003 ([2]), hatten bereits darauf hingewiesen, dass große Mengen (etwa 100g) dunkler Schokolade den Blutdruck senken können. Der Effekt von niedrigen Dosen von Kakao blieb unklar. Die Forscher waren daher daran interessiert, zu untersuchen, ob eine niedrig dosierte Einnahme dunkler Schokolade über eine längere Zeit hinweg ebenfalls blutdrucksenkend wirkt. Dazu wurde eine sog. randomisierte, kontrollierte Studie, ein nicht nur in der medizinischen Forschung übliches Studiendesign, durchgeführt.

## 2. Zielgruppe der Untersuchung?

Um den unerwünschten Einfluss von Störfaktoren (im Fachjargon auch confounding factors genannt), welche die eigentlich untersuchte Zielgröße (hier: der Blutdruck) auch beeinflussen, möglichst klein zu halten, waren die Forscher bestrebt, eine ziemlich homogene Gruppe für die Studie auszuwählen: Nichtraucher im Alter von 55–75 Jahren, in gutem Allgemeinzustand, welche einen leicht erhöhten Blutdruck aufweisen. Ferner sollten die Probanden einer höheren sozialökonomischen Schicht angehören (s. S.7). Die Versuchspersonen sollten auch nicht regelmäßig Schokolade verzehren. Im Unterricht könnte man diskutieren, ob diese zusätzliche Einschränkung bei der Auswahl der Versuchsteilnehmer tatsächlich notwendig war (das wäre ein solcher Störfaktor, weil diese Personen die – nachzuweisende – Wirkung ja gewöhnt sind!). Auch die Diskussion, wie die Auswahl der Zielgruppe eine Erweiterung der Studienergebnisse auf andere Populationen ermöglicht, scheint begründet. Auch die Autoren geben in ihrem Aufsatz zu: „Hence our results may be valid only for individuals who are older and mildly hypertensive but otherwise healthy“. Im Unterricht soll aufgefordert werden, alternative Populationen zu finden, bei denen dieselben Ergebnisse möglich erscheinen.

Zu Beginn qualifizierten sich von den interessierten Teilnehmern 76 Patienten für die Teilnahme an der Studie. Als diesen dann berichtet wurde, worum es genau ging, nämlich um den Verzehr von dunkler Schokolade versus heller Schokolade, verweigerten 32 Patienten die Teilnahme an der Studie. Eine Erörterung dieses Problems kann sich darauf konzentrieren, wie diese relativ hohe Verweigerungsrate eine Verallgemeinerung der Ergebnisse beeinflusst bzw. wie sich dies auf solche Personen, die später den Diät-Empfehlungen folgen sollten, auswirkt.

### 3. Wie wurde die Kontrollgruppe festgelegt?

Die Untersuchungsfrage war, inwieweit die Einnahme von dunkler Schokolade eine blutdrucksenkende Wirkung hat. Eine Alternative für die „Behandlung“ der Kontrollgruppe wäre, dass diese ihr gewöhnliches Essverhalten beibehielten. Dann aber könnten weitere Nebenwirkungen des Verzehrs dunkler Schokolade, die aus der zusätzlichen Kalorienaufnahme, dem Fettverzehr und ähnlichem resultieren, die eigentliche Wirkung als Störgröße überlagern. Der Verzehr einer etwa gleichen Menge weißer Schokolade (mit den genannten „Nebenwirkungen“) in der Kontrollgruppe eliminiert solche möglichen Störfaktoren (confounding factors). Die Schüler und Studenten sollten kritisch weitere mögliche Behandlungsschemata für die Kontrollgruppe erörtern.

### 4. Verblindungstechniken

Blindversuche sind in der medizinischen Forschung weit verbreitet. Es sind Formen des Experimentes bei denen u.a.:

- die Versuchspersonen nicht wissen, ob sie der Experimental- oder Kontrollgruppe angehören,
- die Person, die den Test durchführt, diese Zugehörigkeit nicht kennt,
- Personen, welche die Messgrößen bestimmen, diese Zugehörigkeit der Versuchsperson nicht kennen.

Die Verblindung auf Seiten der Probanden soll (positive oder negative) Placebo-Effekte ausschließen. Eine Behandlung kann allein aufgrund des Wissens, dass man behandelt wird, Effekte zeitigen. Eine Verblindung auf Seiten der Behandler (solche die behandeln oder nur messen) soll Beeinflussung der Behandlung oder Messung verhindern. Der Standard in der Forschung ist das sogenannte doppelblinde Experiment, Behandler und Behandelte wissen nicht, womit tatsächlich behandelt wird. Hier war ein doppelblindes Experiment aus nahe liegenden Gründen nicht wirklich möglich.

Die Patienten wussten genau, welche Art von Schokolade sie verzehrten. Die Autoren sagen dazu aus, dass keine polyphenolreiche und polyphenolfreie Schokolade ähnlichen Aussehens im freien Handel verfügbar wären. Die Studierenden könnten sich fragen, warum die dunkle Schokolade nicht in Pillenform oder in einer anderen versteckten Form angeboten wurde.

Wichtig bei einem Blindversuch ist nämlich, dass die Arzneimittel von den Probanden nicht unterschieden werden können. Die Arzneimittel sollten also optisch identisch erscheinen (inklusive der Verpackung), und auch Geruch und Farbe spielen eine Rolle. Nichtsdestotrotz waren die Teilnehmer in einer gewissen Art „verblindet“, weil sie keine Informationen bis zum Ende der 18 Wochen über die Erreichung der Untersuchungsziele (ob ihr Blutdruck tatsächlich gesenkt wurde) erhielten.

Es handelte sich jedoch um einen Blindversuch bezüglich der untersuchenden Personen (sog. untersucherverblindete Studie): Die behandelnden Personen (Mediziner) wussten nämlich nicht, welche Person zu welcher Behandlungsgruppe gehört. Blutdruckmessungen, zum Beispiel, sind im Allgemeinen mit einem nicht zu vernachlässigenden Messfehler behaftet. Die untersuchenden Personen könnten je nach ihren Vorurteilen die Messungen beeinflussen; durch die Befangenheit eines Untersuchers könnten die gemessenen Blutdruckwerte in seinem Sinne manipuliert werden. Taubert hatte seine Teilnehmer an der Studie verpflichtet, den Untersuchern ihre Zuordnung zu einer der Gruppen zu verheimlichen.

### 5. Wie wurden die Teilnehmer auf die Untersuchungsgruppen verteilt?

Trotz Randomisierung wurde darauf geachtet, dass Männer und Frauen in ausgewogenem Verhältnis in der Untersuchungs- und in der Kontrollgruppe vertreten waren. Von den 44 Teilnehmern, die letztendlich ausgewählt wurden, wurden jeweils Untergruppen mit vier Mitgliedern desselben Geschlechts gebildet. Aus jeder Gruppe wurden zwei Personen zufällig in die Untersuchungsgruppe zugeteilt und die anderen zwei in die Kontrollgruppe. Im Unterricht sollte diskutiert werden, warum dieses Verfahren der eingeschränkten Randomisierung gewählt wurde und warum die Untersucher es für wichtig erachteten, die Randomisierung nach Geschlecht getrennt durchzuführen. Welche anderen Möglichkeiten homogener „Versuchsblöcke“ sind möglich?

### 6. Wie wurde die Größe der Untersuchungsgruppe bestimmt?

Die Zahl der Probanden in der Studie erscheint im ersten Augenblick sehr gering. Im Originalartikel beschreiben die Autoren ausführlich, wie die nachzuweisende Größe der Wirkung (effect size), Machtüberlegungen

(englisch power; Macht ist die Gegenwahrscheinlichkeit zum Fehler 2. Art) und Information aus vorhergehenden Studien die erforderliche Stichprobengröße für jede der Behandlungsgruppen ergeben. Die Autoren wollten den so genannten „gepaarten t-Test“ (paired t test) benutzen; sie wollten dabei in der Lage sein, eine Absenkung des Blutdrucks um 1,5 mmHg (systolisch) bzw. 1,0 mmHg (diastolisch) mit einer Macht von 0,8 zu erkennen, wenn der  $\alpha$ -Fehler des Tests 0,05 beträgt. Dazu ist es notwendig, die Streuung der Blutdrucksenkung zu kennen; aus der vorhergehenden Studie schätzten sie die Standardabweichung der Senkung mit 2,0 (systolisch) bzw. 1,5 (diastolisch).

Aus diesen Forderungen ergibt sich ein minimal erforderlicher Stichprobenumfang von 20 Teilnehmern für jede Gruppe. (Taubert et al. wendeten sicherlich noch Berichtigungen an, die notwendig sind für die anstehenden multiplen Vergleiche, welche mehrere Tests simultan erforderlich machen – so kommentiert der Autor des Aufsatzes über die Taubertsche Studie in Teaching Statistics seine eigenen Versuche, diese Berechnungen nachzuvollziehen.) Falls diese Thematik in ihr Curriculum passt, können sie die Studenten auffordern, diese Rechnungen für die Mindeststichprobengröße nachzuvollziehen.

## 7. Zur Blutdruckmessung

Die Blutdruckmessung erfolgte indirekt mit Hilfe der üblichen aufblasbaren Manschette am Arm. Um Messfehler zu minimieren wurden jeweils drei Messungen im Abstand von 5 Minuten unter kontrollierten Bedingungen durchgeführt. Besonders interessant für Statistikstudenten sollte die Tatsache sein, dass zur Beurteilung der arteriellen Druckhöhe jeweils der Mittelwert der drei gemessenen Blutdruckwerte genommen wurde. Um weiter erhöhte Werte aufgrund momentaner Nervosität und wegen des Weißkittelsyndroms zu minimieren, wurde der Blutdruck in einem extra Raum abseits der Ärztezimmer gemessen. Wie bereits vorne erwähnt, wussten die messenden Personen nicht, welcher Gruppe die Patienten zugeordnet waren, damit sollte eine mögliche Beeinflussung ausgeschlossen werden.

## 8. Welche Verfahren einführender Statistiklehrgänge werden in der Studie benutzt?

Die Autoren des medizinischen JAMA-Fachartikels benutzen die Methode der Varianzana-

lyse (ANOVA, engl. analysis of variance, mit Messwiederholungen). Damit wird geprüft, ob verschiedene Gruppen sich hinsichtlich ihres Erwartungswertes einer Zielvariablen unterscheiden. Die dadurch ermöglichten Aussagen sind: – einerseits, es können keine Unterschiede gefunden werden, und – andererseits, die beobachteten Unterschiede sind signifikant, d. h., die Gruppen haben signifikant verschiedene Erwartungswerte. Die Varianzanalyse selbst gibt nur eine globale „Signifikanz“ an, sie sagt nicht, *welche* Gruppen unterschiedlich sind. In der Studie wurden die Gruppen durch den Beobachtungszeitpunkt gebildet: es wurden die 4 Gruppen (gemäß dem Zeitpunkt der Messung) der Versuchsgruppe miteinander verglichen; ebenso ging man in der Kontrollgruppe vor.

Wurden global Unterschiede als signifikant ausgewiesen, schlossen die Forscher einen gepaarten t-Test an. Dieser Test vergleicht zwei Gruppen auf Unterschiede im Erwartungswert. Die zwei Merkmale sind jedoch „gepaart“, hier wird der Blutdruck ein und derselben Person an zwei Zeitpunkten gemessen. Geprüft wird, ob die Differenz den Erwartungswert 0 (kein Unterschied) hat, oder ob die Daten einen signifikanten Unterschied ausweisen.

Wegen der kleinen Zufallsstichprobe kommt zum Test des Unterschiedes zwischen beiden Zufallsstichproben die t-Verteilung zum Zug. Bei diesen paarweisen Vergleichen werden eine Reihe von Tests durchgeführt. Bei 4 Messzeitpunkten macht das  $\binom{4}{2} = 6$  Vergleiche, also 6 Tests, nur innerhalb der Versuchsgruppe. Das macht eine weitere Korrektur der p-Werte bzw. der  $\alpha$ -Fehler erforderlich. Man kann sich das ganz einfach in der Wiederholung ein und desselben Tests klar machen: Wird ein Test mit  $\alpha = 5\%$  auf eine richtige Nullhypothese 1000-mal angewendet, so erhält man „im Durchschnitt“ 50 irrtümliche Ablehnungen. Ein reines Artefakt, bei häufiger Wiederholung „muss“ man einfach einen  $\alpha$ -Fehler vorfinden. Wendet man viele, *verschiedene* Tests in einer Studie an, so hat man mit einem ähnlichen Artefakt zu rechnen: Es werden ganz einfach einige Ablehnungen der entsprechenden Nullhypothese, sprich signifikante Ergebnisse, dabei sein – ohne dass irgendwelche Abweichungen vorhanden sind. Dazu gibt es in der Literatur Modifikationen, ältere von Bonferoni, neuere von Bonferoni und Holm.

Die im Originalartikel aufgeführten p-Werte wurden nach Holm (1979, s.[5]) bearbeitet; diese Korrektur gehört zwar nicht in einen einführenden Statistiklehrgang, sie ist aber in der Praxis ungemein wichtig, will man sich gegen die angesprochenen Artefakte absi-

chern. Eine weitere Frage, die sich aus Unterrichtsdiskussionen ergeben könnte, ist die: welche Personen profitieren am meisten? Taubert und sein Stab benutzen die Methode von Korrelation und Regression, um diese Frage zu beantworten. Sie berichten von einer signifikanten Korrelation zwischen Blutdrucksenkung und Basiswert, d. h. je höher der Ausgangsblutdruck, desto größer die festgestellte Blutdrucksenkung in der Versuchsgruppe. Für die Kontrollgruppe gab es keinen signifikanten Zusammenhang.

Hier ist noch ein wesentlicher Gesichtspunkt anzusprechen, der für viele empirische Studien erörtert werden muss, wenngleich das technische Rüstzeug dazu kompliziert erscheint. Bleiben wir zunächst bei der schon angesprochenen Frage, wer von der „Behandlung mit dunkler Schokolade“ am meisten profitiert. Eine teilweise Antwort wurde schon weiter oben gegeben: Bei jenen, die schon einen relativen hohen, noch normalen Blutdruck haben, ist die Senkung des Blutdrucks im Verlauf des Experiments (in der Versuchsgruppe) am höchsten. Da in der Studie der Blutdruck am Beginn (der Basiswert) aufgezeichnet wurde (war ja ganz wesentlich Teil der Studie), kann man den Effekt der dunklen Schokolade bereinigen um die Auswirkungen des Basiswerts. Solche Variable, die den Behandlungserfolg mit beeinflussen können, nennt man Kovariate. Man kann die Prognose eines Behandlungseffekts für eine neue Person verbessern, wenn man ihren Basiswert für den Blutdruck kennt, siehe die Streudiagramme in Fig. 1.

Ein weiterer nützlicher Diskussionspunkt wäre, die Logik zu hinterfragen, die hinter der Erfassung der Auswirkung von Schokoladenverzehr auf den Blutdruck über das Merkmal der Veränderung des Blutdrucks steckt. Dazu lohnt es sich zu betrachten, welche Art der Korrelation für die beiden Behandlungsgruppen zwischen Basiswert und den verschiedenen Messzeitpunkten (aber nicht der Blutdruckänderungen) zu erwarten ist und was dies, wenn überhaupt, für die Wirksamkeit der Behandlung bedeuten würde. Hier sollte man darauf hinweisen, dass aufeinanderfolgende Messungen immer korreliert sind, auch ohne Vorliegen eines allfälligen Effekts einer Behandlung.

## 9. Welche ethischen Probleme sind mit dieser Studie verbunden?

Es kann vorkommen, dass klinische Studien große ethische Probleme aufwerfen. Beispiel: Onkologen testen ein neues Chemotherapeutikum zur Bekämpfung eines Krebstumors. Die eine Hälfte der Patien-

ten bekäme das (vermutlich nützliche) Heilmittel, während die andere Gruppe nicht behandelt würde, da sie das bei klinischen Studien übliche Placebo erhielten. Eine solche Vorgehensweise kann zu erheblichen ethischen Problemen führen, selbst wenn die Personen in der Kontrollgruppe die nach dem aktuellen medizinischen Standard gültige Behandlung statt des Placebos erhielten, weil etwa die neue Behandlung in Vorversuchen sich als besonders gut erwiesen hat.

Wo können hier ethische Probleme auftreten? Ein wichtiger ethischer Aspekt dieser Studie ist, dass Patienten mit einem Blutdruck von 170/100 eher zum Gang zu einem Kardiologen als zur Teilnahme an dieser Studie bewegt werden sollten. Ein anderer Aspekt ist die Gefahr einer Gewichtszunahme, die den möglichen Effekt der Blutdrucksenkung weniger interessant macht. Eine weitere ethisch relevante Frage ist, ob den der Kontrollgruppe zugelosten Patienten durch die Gabe der weißen Schokolade ein Nachteil entsteht: diese nehmen nur zusätzliche unnötige Kalorien und Fette auf; was ist ihr Nutzen?

## Schlussbetrachtungen

Die hier vorgestellte Studie ist eine ausgezeichnete Fallstudie für Diskussionen über viele Aspekte und Themen einführender Statistiklehrgänge, wie z. B. Design und Planung einer Studie, Ethik, statistisches Schließen, sowie Korrelation und Regression. Der Unterrichtende kann die angesprochenen Themen schon früh in seinem Unterricht verwenden, indem er z. B. das Thema Versuchsplanung aufgreift. Alternativ kann später Versuchsplanung und Auswahl dazu geeigneter statistischer Verfahren – z. B. verbundene Stichproben (matched pairs) versus unabhängige Stichproben und die entsprechenden t-Tests – thematisiert werden. Auf Deutsch kann man für diese Themen etwa Sachs [6] als Hilfe heranziehen.

Natürlich kann man auch den Originalbeitrag im medizinischen Fachjournal lesen und in Seminaren vortragen lassen. Man kann sich dabei auch lediglich auf die Rezeption der Studie in den Massenmedien beziehen oder die Hintergrundinformation aufarbeiten lassen.

Auf alle Fälle ist es mein aufrechter Wunsch, dass die Vorliebe von Studenten für Süßigkeiten ihre Zuneigung für statistische Reflexionen nähren möge.

## Literatur

- [1] Kanavos, P., Östergren, J. und Weber, M. (2007): High Blood Pressure and Health Policy: Where We Are and Where We Need to Go Next. New York: Ruder Finn Inc.
- [2] Taubert, D., Berkels, R., Roesen, R. und Klaus W. (2003): Chocolate and Blood Pressure in Elderly Individuals with Isolated Systolic Hypertension. In: Journal of the American Medical Association (JAMA), 290, 1029–30.
- [3] Taubert, D.; Roesen, R.; Lehmann; C.Jung, N.; Schomig, G. (2007): Effects of Low Habitual Cocoa Intake on Blood Pressure and Bioactive Nitric Oxide: A Randomized Controlled Trial. In: Journal of the American Medical Association (JAMA), 298(4. Juli 2007)1, S. 49-60
- [4] Henze, N. (2008): Stochastik für Einsteiger. Wiesbaden: Vieweg, 2008 (7., überarbeitete und erweiterte Auflage)
- [5] Holm, S. (1979): A Simple Sequentially Rejective Multiple Test Procedure. In: Scandinavian Journal of Statistics, 6, S. 65–70.

- [6] Sachs, L. (2004): Angewandte Statistik. Anwendungen statistischer Methoden. Berlin, Heidelberg: Springer (11. Auflage)

*Bemerkung der Bearbeiter: Wie in der Mathematik als Wissenschaft so ist auch in der medizinischen Forschung Englisch die wissenschaftliche Kommunikationssprache. Es zählen nur Publikationen, die in Englisch verfasst sind, weil nur diese von Kollegen gelesen und erfasst werden. Auf eine Nachfrage beim Autor in Köln, ob eine deutsche Fassung seines Beitrages oder auch nur Teile der Arbeit auf Deutsch verfügbar seien, kam prompt die Antwort: Der Beitrag wurde original in Englisch verfasst.*

## Anschrift des Verfassers

Eric D. Nordmoe  
Kalamazoo College, Michigan, USA  
e-mail: enordmoe@kzoo.edu

## Was tun bei Mammutbäumen?

STEFAN BARTZ, MECKEL

**Zusammenfassung:** *Wahrscheinlichkeitsaufgaben lassen sich im Schulbereich in der Regel über Baumdiagramme anschaulich und sicher lösen. Manchmal stößt man jedoch auf riesige, unüberschaubare Bäume (Mammutbäume). Anhand einer dem Sammelbildproblem entsprechenden Aufgabe wird gezeigt, wie man solche Wahrscheinlichkeitsprobleme systematisch angehen kann.*

## Aufgabe<sup>[1]</sup>

Laut Mitteilung des Polizeipräsidiums Leverkusen gab es in 110 Tagen 329 Meldungen an die Presse (Verkehrsunfälle mit erheblichem Sach- oder Personenschaden, Serienunfälle, Unfälle mit Fahrerflucht, schwerere Einbruchsdelikte Schaden, Einbruchsserien). Wenn man 329 Meldungen zufällig auf 110 Tage verteilt, ist es dann außergewöhnlich, dass es Tage ohne Meldungen gibt? Wie groß ist die Wk, dass zufällig 0, 1, 2, ... meldungsfreie Tage entstehen? Ab wie vielen Meldungen treten mit 95%iger Sicherheit keine meldungsfreien Tage mehr auf?

## 1. Standardlösungsansatz<sup>[2]</sup>

Bei diesem Ansatz versucht man, Wks-Aufgaben erst einmal mit Hilfe von Baumdiagrammen zu lösen. Wir nehmen uns zunächst ein reduziertes Problem vor: Wie groß ist die Wk, dass beim zufälligen Verteilen von 6 Meldungen auf 4 Tage 0, 1, 2 bzw. 3 Tage nicht besetzt werden?

(1) *Ereignis E:* Beim zufälligen Verteilen von 6 Meldungen auf 4 Tage bleiben genau 0 (1, 2, 3) Tage unbesetzt.

(2) *Baumdiagramm:* Um Mehrfacheinträge in einzelnen Knoten zu vermeiden, weist man den 6 Meldungen jeweils bestimmte Tage zu und nicht umgekehrt den Tagen Meldungen. D. h. man stellt sich eine Urne mit 4 nummerierten Kugeln vor, aus der man für jede Meldung eine Tageszahl mit Zurücklegen zieht. Mögliche Pfade wären demnach:

