

LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
INSTITUT FÜR STATISTIK

BACHELORTHESES

REAL-DATA-BEISPIELE IN BIostatistik- UND BioINFORMATIK-FACHZEITSCHRIFTEN: EIN SURVEY



Germaine Wenzler

Betreuerin: Prof. Dr. Anne-Laure Boulesteix

Sommersemester 2013
18. Septemeber 2013

Zusammenfassung

In der Biostatistik und Bioinformatik werden laufend Artikel veröffentlicht, in welchen neue Methoden vorgestellt und bereits vorhandenen Methoden gegenüber gestellt werden. Meist ist dabei zu beobachten, dass die neuen Methoden besser bewertet werden. In dieser Arbeit wird versucht zu überprüfen worauf sich die Hervorhebung neuer Methoden in wissenschaftlichen Arbeiten stützt.

Um dies statistisch untersuchen zu können, wurde ein Datensatz erzeugt, für welchen vier wissenschaftliche Fachzeitschriften genauer betrachtet wurden: „Bioinformatics“; „Biometrics“; „BMC Bioinformatics“ und „Biostatistics“. Es wurden jeweils die 29 ersten Artikel aus dem Jahr 2013 auf bestimmte Kriterien hin untersucht. Hierbei wurde z.B. darauf geachtet, ob reale Daten und neue Methoden verwendet, wie viele Methoden durchgeführt und ob Methoden anderen gegenüber besser bewertet werden. Für die Bewertung einer Methode wurden die Artikel auf verschiedene Schlüsselwörter, wie z.B. „advantage“, „better“ und weitere hin untersucht. Die deskriptiven Analysen zeigen, dass in fast 78% der Veröffentlichungen eine Methode positiver bewertet wird als andere Methoden und 95.8% der Berichte reale Daten verwenden. Lediglich in 36.2% der Artikel konnten signifikante Ergebnisse gefunden werden.

Die Ergebnisse der Auswertungen bestätigen, dass Aussagen wie „our method outperforms the other methods“ nicht für allgemeingültig angenommen werden können.

Inhaltsverzeichnis

1	Einleitung	1
2	Theorie	2
3	Erhebung der Daten und Auswertung	5
3.1	Beschreibung des Datensatzes	5
3.2	Deskriptive Statistiken	8
3.3	Zusammenfassung der Ergebnisse	19
4	Fazit	20
5	Anhang	21

Kapitel 1

Einleitung

Jedes Jahr erscheinen mehrere hundert Artikel in biostatistischen und bioinformatischen Fachzeitschriften, in denen neue Verfahrensweisen zur Auswertung von Daten präsentiert werden. Dabei werden neu entwickelte Methoden nicht nur vorgestellt, sondern zusätzlich mit bereits existierenden Methoden verglichen. Solche Vergleichsstudien sind notwendig, um einen Überblick über die angemessene Anwendung von Methoden zu erhalten. In den meisten Fällen zeigt sich, dass die neuen Methoden gegenüber den bestehenden als wirksamer dargestellt werden. Hierbei stellt sich die Frage, ob Autoren von vorneherein von ihren eigenen Methoden überzeugt sind oder ob sie ihre Aussagen und Ergebnisse auf statistische Beweise gründen. Der Vergleich von Methoden kann auf verschiedene Art und Weise erfolgen, gängigste Vorgehensweisen sind dabei die Durchführung statistischer Tests und die Betrachtung von Konfidenzintervallen. Dabei kommt es oft vor, dass Hypothesen nicht richtig formuliert werden und der Fehler 1. Art sowie der Fehler 2. Art vernachlässigt werden (Boulesteix et al., 2013a). Zusätzlich spielt es eine Rolle, ob die Ergebnisse signifikant sind und an wievielen realen Datensätzen die Methoden gemessen werden.

Das Ziel dieser Arbeit ist, herauszufinden worauf sich Aussagen wie „our method performs best“ stützen. Hierfür werden Artikel von vier Fachzeitschriften („Bioinformatics“, „Biometrics“, „BMC Bioinformatics“ und „Biostatistics“) genauer auf die Anzahl verwendeter realer Datensätze in den Berichten und auf Erwähnung signifikanter Ergebnisse überprüft.

Der Aufbau dieser Abhandlung gliedert sich wie folgt:

Kapitel 2 wird sich mit der Theorie beschäftigen, wohingegen Kapitel 3 anschließend das Vorgehen zur Erstellung des Datensatzes erläutert und stellt einige statistische Auswertungen vor. Den Abschluss dieser Arbeit bildet ein kurzes Fazit in Kapitel 4.

Kapitel 2

Theorie

Biostatistik kann, ebenso wie Biometrie, vereinfacht definiert werden als „Zweig der Statistik, der sich mit der Anwendung mathematisch-statistischer Methoden auf die belebte Natur befasst“ (Koehler et al., 2012). Die Biostatistik ist ein wesentlicher Baustein der Biowissenschaften und findet neben der Biologie und der Biotechnik auch Anwendung in der Medizin oder den Agrarwissenschaften.

Den Begriff der Bioinformatik zu erklären gestaltet sich schon etwas schwieriger, da der Bioinformatik viele Aufgaben zugesprochen werden. Bioinformatik versucht mittels computergestützten Methoden Lösungen für Probleme der Biowissenschaften zu finden. Unter anderem zählen dazu die Entwicklung von mathematischen Algorithmen zur Auswertung von DNA- und Proteinsequenzen sowie deren Interpretation (Hütt und Dehnert, 2006).

Kurz gesagt zählen Biostatistik und Bioinformatik zu den rechenintensiven Wissenschaften und behandeln Probleme der Lebenswissenschaften.

Vergleichsstudien können in zwei Kategorien aufgeteilt werden und zwar in die sogenannten repräsentativen sowie illustrativen Vergleichsstudien, wobei der Unterschied in der Auswahl der Datensätze liegt. Im ersten Fall wird eine randomisierte Auswahl getroffen, die versucht das Anwendungsgebiet sowie die Teststärke zu berücksichtigen. Hierbei wird versucht den Bias, also eine systematische Verzerrung, zu vermeiden. Bei illustrativen Studien hingegen werden bestimmte Datensätze ausgewählt, welche spezielle Eigenschaften einer Methode verdeutlichen sollen (Boulesteix, 2013c).

Studien, die Vergleiche zwischen Methoden betreiben, sind ohne Frage wichtig und können einen Überblick liefern, welche Vorteile gewisse Methoden bringen und wann diese angemessen sind. Jedoch kann von keiner gängigen Methode erwartet werden, dass sie in jeder Situation die besten Ergebnisse liefert (siehe Wolpert, 2001, „no-free-lunch theorem“) und es sollte immer genau überprüft werden, wel-

che Methode für den aktuellen Fall am besten passt. Bei solchen Vergleichsstudien kommt es überwiegend vor, dass eine Methode als klarer Favorit hervor geht, meist die neu entwickelte. Das Problem ist somit nicht, dass Vergleiche von Methoden angestellt werden, sondern die Art und Weise der Durchführung und Interpretation dieser.

Ein Kritikpunkt ist, dass bei Auswertungen nicht berücksichtigt wird, ob die Ergebnisse signifikant sind, also nicht zufällig auftreten, sondern ein überzufälliger Zusammenhang besteht und „im statistischen Sinn bedeutend sind“ (Fahrmeier et al., 2010; Seite 418). Zur Überprüfung der Signifikanz können statistische Tests verwendet werden, bei denen eine Alternativ- und eine Nullhypothese aufgestellt werden muss. Hierbei können Fehler beim Formulieren der Hypothesen auftreten. Zudem ist die Wahl eines Signifikanzniveaus α nötig, welches die sogenannte Irrtumswahrscheinlichkeit bezeichnet und meist auf den Wert 0.01, 0.05 oder 0.1 gesetzt wird. Die Irrtumswahrscheinlichkeit gibt an, in wieviel Prozent der Fälle die Nullhypothese fälschlicherweise verworfen wird (Fehler 1. Art). Mit dem Signifikanzniveau wird also versucht die Auftrittswahrscheinlichkeit für den Fehler 1. Art zu kontrollieren.

„Je größer der Stichprobenumfang, desto größer die Power des Tests.“

Dieser Satz aus Janczyk und Pfister (2013) führt zu einem weiteren Kritikpunkt in den Vergleichen. Die Power des Tests bezeichnet die Wahrscheinlichkeit, dass die Nullhypothese verworfen werden kann. Diese Wahrscheinlichkeit nimmt mit dem Stichprobenumfang, hier der Anzahl verwendeter realer Datensätze, zu. In den meisten Artikeln liegt der Median für die Anzahl realer Datensätze gerade einmal bei fünf (Boulesteix et al., 2013b). Das Problem dabei ist, dass sich Methoden für jeden Datensatz anders verhalten. Um eine genaue Aussage über die Leistungsfähigkeit von Methoden treffen zu können, müsste eine angemessene Zahl Datensätze betrachtet werden. Daraus könnte dann abgeleitet werden, welche Methode für die jeweiligen Datensätze die beste Leistung erzielt. Je öfter eine Methode für verschiedene reale Datensätze positive Ergebnisse liefert, desto höher ist die Wahrscheinlichkeit, dass sie valide ist. Die angemessene Größe eines Stichprobenumfangs kann mittels statistischer Verfahren bestimmt werden.

Es ist zu beachten, dass diese Arbeit Ergebnisse simulierter Daten vernachlässigt, da diese Daten künstlich erzeugt werden. Dabei wird zwar darauf geachtet, eine ähnliche Verteilung zu treffen wie bei realen Daten, allerdings muss schon im Vorfeld eine Verteilungsannahme getroffen werden. Dabei besteht die Schwierigkeit die Verteilung der Originaldaten korrekt wieder zu spiegeln und deren Komplexität vollends zu erfassen. Daraus resultiert die Möglichkeit, dass künstliche Daten speziell für das erhoffte Ergebniss erzeugt werden können. So kann es sein, dass

neue Methoden bevorzugt werden, weil die simulierten Daten für diese Methoden vorteilhaft sind.

Reale Daten können ebenso Nachteile aufweisen. Zum einen kann die Streuung der Beobachtungen innerhalb des Datensatzes variieren. Zum anderen folgt jeder Datensatz, der auf realen Daten basiert, seiner eigenen Verteilung. So kann es beim Vergleich von Methoden anhand mehrerer Datensätze vorkommen, dass diesen Datensätzen unterschiedliche Verteilungsannahmen zugrunde liegen.

Kapitel 3

Erhebung der Daten und Auswertung

Im Folgenden wird die in Kapitel 2 erläuterte Thematik anhand eines eigens erzeugten Datensatzes statistisch untersucht. Zunächst wird die Erstellung des Datensatzes genauer erläutert, anschließend werden einige deskriptive Statistiken vorgestellt.

3.1 Beschreibung des Datensatzes

Um einen möglichst aktuellen Bezug zu den Daten zu haben, wurden Artikel betrachtet die Anfang des Jahres 2013 erschienen. In der nachstehenden Tabelle 3.1 sind die Fachblätter mit ihren jeweiligen Ausgaben vermerkt:

Journal	Ausgabe
Bioinformatics (BI)	Volume 29, Nr.1; Januar 2013
	Volume 29, Nr.2; 15.Januar 2013
Biometrics (BM)	Volume 69; März 2013
BMC Bioinformatics (BMC)	Volume 14; Januar-März 2013
Biostatistics (BS)	Volume 14, Nr.1; Januar 2013
	Volume 14, Nr.2; April 2013

Tabelle 3.1: Verwendete Journals und Ausgaben

Im weiteren Verlauf dieser Arbeit werden der Einfachheit halber lediglich die in der Tabelle aufgeführten Abkürzungen für die Journals verwendet (BI, BM, BMC und BS). Ursprünglich sollten die ersten 20 Artikel jedes Journals betrachtet werden. Allerdings hätte dies in BI zu einer Überrepräsentierung von thematischen Rubriken geführt, weshalb von diesem Journal alle Berichte der ersten beiden Ausgaben

von 2013 betrachtet wurden. Dies führte schließlich zur Untersuchung der ersten 29 Artikel je Journal und einem Stichprobenumfang von $n=116$. Es ist anzumerken, dass bei BMC ausschließlich *methodology article* betrachtet wurden.

Jeder Artikel wurde auf bestimmte Kriterien untersucht, wobei die zwei folgenden Aspekte die wichtigsten waren:

1. Welche Daten werden für die Überprüfung der Methoden verwendet?
2. Findet eine bessere Beurteilung von Methoden gegenüber anderen statt?

Zusätzlich wurde u.a. überprüft an welcher Stelle im Text diese Bewertungen auftreten, wie viele Methoden verwendet und ob neue Methoden eingeführt werden. In Tabelle 3.2 sind die Variablen des Datensatzes, welcher als Excel-Tabelle vorliegt, alphabetisch aufgelistet. Zur Beurteilung der Methoden treten die Bezeichnungen „advantage“, „better“, „effective/efficient“, „improve“ und „outperform“ am häufigsten auf, daher wurden diese fünf als eigenständige Variablen in den Datensatz aufgenommen und werden im weiteren Verlauf als Schlüsselwörter bezeichnet, alle anderen Begrifflichkeiten, die zur Bewertung verwendet werden, wurden als „sonstiges“ zusammengefasst.

Hinter der Variable „advantage“ verbirgt sich zudem noch das Wort *advantageous*, explizit verdeutlicht die Variable, dass eine neue Methode Vorteile bzw. vorteilhafte Ergebnisse liefert. Die Variable „better“ beinhaltet ebenso weitere Stichpunkte wie *better performance*, *better accuracy* oder auch *significantly better*, was bedeutet, dass die neue Methode u.a. bessere Genauigkeit, Leistung oder Ergebnisse erbringt. „Effective/efficient“ bezeichnet die höhere Wirksamkeit neuer Methoden, mittels „outperform“ wird ausgedrückt, dass eine neue Methode die anderen in verschiedenen Punkten übertrifft. „Improve“ verdeutlicht in den Artikeln die Verbesserungen, die eine Methode erzeugt, zur Variable zählt ebenfalls *improvement* dazu. Unter sonstigen Begriffen verbergen sich z.B. Wörter wie *accurate*, *faster*, *more practical/powerful* oder auch *simpler*. Die Betrachtung der Bewertung erfolgte ausschließlich unter dem Aspekt, ob eine Methode gegenüber anderen besser bewertet wird, das Gegenteil wurde nicht berücksichtigt. Zusätzlich wurde betrachtet, ob in den Ergebnissen der Begriff der Signifikanz erwähnt wird (Variable „signifikant“). Es ist darauf hinzuweisen, dass sich die Variablen nur auf die Abschnitte im Text beziehen, die einen Bezug zu den realen Daten haben. Abschnitte in denen ausschließlich simulierte Daten behandelt werden, wurden vernachlässigt. So kann z.B. der Begriff „outperform“ im Text vorkommen, aber im Datensatz trotzdem mit 0 vermerkt worden sein, wenn sich diese Wertung auf ein simuliertes Beispiel bezieht. Verwendet ein Artikel ausnahmslos simulierte Daten, wurden alle aufgeführten Dummyvariablen, außer „neue_methode“, im Datensatz auf 0 gesetzt.

Variable	Beschreibung
advantage	Dummyvariable; $\text{advantage} \begin{cases} 1 & \text{falls „advantage“ im Artikel verwendet wird,} \\ 0 & \text{sonst} \end{cases}$
Ausgabe	Erscheinungsdatum des Journals
better	Dummyvariable; $\text{better} \begin{cases} 1 & \text{falls "better“ im Artikel verwendet wird,} \\ 0 & \text{sonst} \end{cases}$
Bewertung	Dummyvariable; $\text{bewertung} \begin{cases} 1 & \text{falls pos. im Artikel bewertet wird, unabhängig von der Stelle,} \\ 0 & \text{sonst} \end{cases}$
Daten	Art der verwendeten Daten (real, simuliert oder beides)
effective_efficient	Dummyvariable; $\text{effective_efficient} \begin{cases} 1 & \text{falls „effective/efficient“ im Artikel verwendet wird,} \\ 0 & \text{sonst} \end{cases}$
ID	Kennung des Papers (BI_1-BI_29; BMC_1-BMC_29; BM_1-BM_29 und BS_1-BS_29)
improve	Dummyvariable; $\text{improve} \begin{cases} 1 & \text{falls „improve“ im Artikel verwendet wird,} \\ 0 & \text{sonst} \end{cases}$
Kategorie	Art des Journals (BI, BM, BMC, BS)
Methode	Anzahl der auf reale Daten angewendeten Methoden
Neue_Methode	Dummyvariable; $\text{neue_methode} \begin{cases} 1 & \text{falls eine neue Methode im Artikel vorgestellt wird,} \\ 0 & \text{sonst} \end{cases}$
outperform	Dummyvariable; $\text{outperform} \begin{cases} 1 & \text{falls „outperform“ im Artikel verwendet wird,} \\ 0 & \text{sonst} \end{cases}$
Reale_Daten	Anzahl der verwendeten realen Datensätze
Signifikant	Dummyvariable; $\text{signifikant} \begin{cases} 1 & \text{falls signifikante Ergebnisse erwähnt werden,} \\ 0 & \text{sonst} \end{cases}$
Sonstiges	Dummyvariable; $\text{sonstiges} \begin{cases} 1 & \text{falls ein sonstiger Begriff zur Bewertung verwendet wird,} \\ 0 & \text{sonst} \end{cases}$
Stelle	Stelle im Text, an der positiv bewertet wird (abstract, abstr.+disc., abstr.+sonst, discussion, disc.+sonst., keine, sonstige, überall)

Tabelle 3.2: Variablen des Datensatzes

3.2 Deskriptive Statistiken

Die folgenden Auswertungen wurden mit Hilfe von R (Version 2.11.0) erstellt, wobei in allen Auswertungen 0 für „nein“ und 1 für „ja“ steht. Die ersten Abbildungen 3.1 bis 3.4 sollen einen Überblick über den Datensatz ermöglichen.

Graphik A in Abbildung 3.1 zeigt, dass in 76 der 116 Artikel beide Datentypen und in sechs Berichten einzig simulierte Daten verwendet werden. In Graphik B ist zu erkennen, dass meistens ein bis zwei reale Datensätze betrachtet werden und lediglich in vier Artikeln mehr als fünf, zudem findet in 72 Artikeln die Einführung mindestens einer neuen Methode statt, was Graphik C verdeutlicht. In 90 Berichten kann beobachtet werden, dass eine Methode gegenüber anderen vorteilhaft bewertet wird, siehe Graphik D.

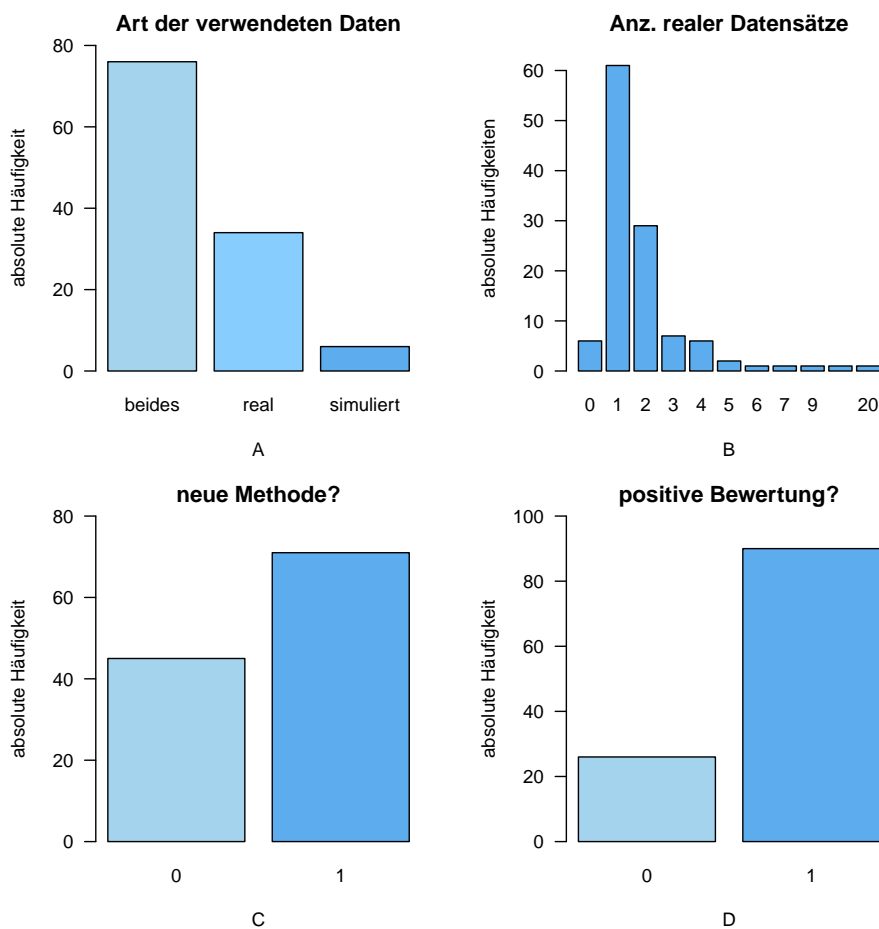


Abbildung 3.1: Überblick über die Daten

In Abb. 3.2 wird die Verwendung einer neuen Methode je nach Journal dargestellt. Auffällig ist, dass in 24 von 29 BMC-Artikeln neue Methoden eingeführt werden. Bei BS werden zwölfmal neue Methoden eingeführt, in BM waren es in etwa gleich viele neue wie alte Methoden. BI-Veröffentlichungen verwenden etwas mehr neue Methoden, in diesen werden auch am häufigsten nur reale Daten betrachtet, was

aus Abbildung 3.3 hervorgeht. Die anderen drei Journals benutzen überwiegend sowohl reale als auch simulierte Datensätze. Von den 29 BM-Artikeln verwendet keiner ausschließlich simulierte Datensätze, wohingegen in den BS-Berichten jeweils zweimal nur reale bzw. nur simulierte Datensätze auftreten.

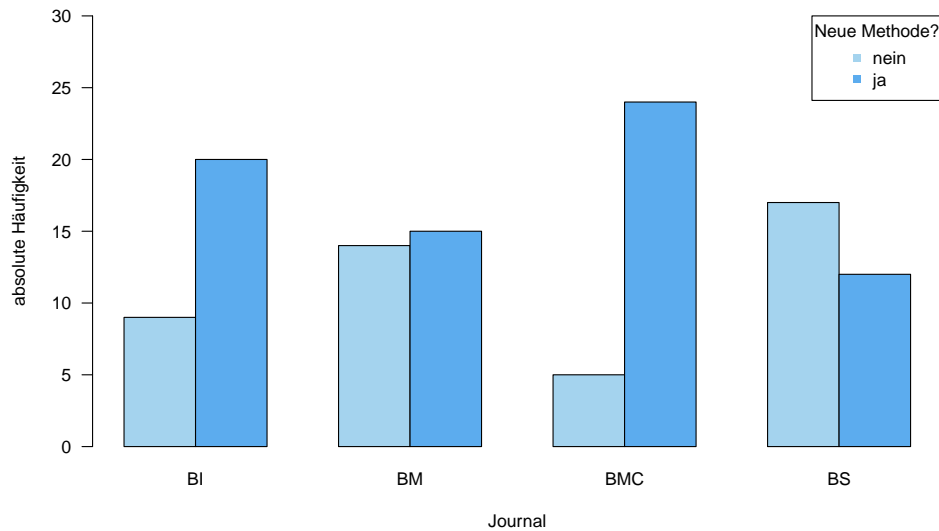


Abbildung 3.2: Verwendung einer neuen Methode je nach Journal

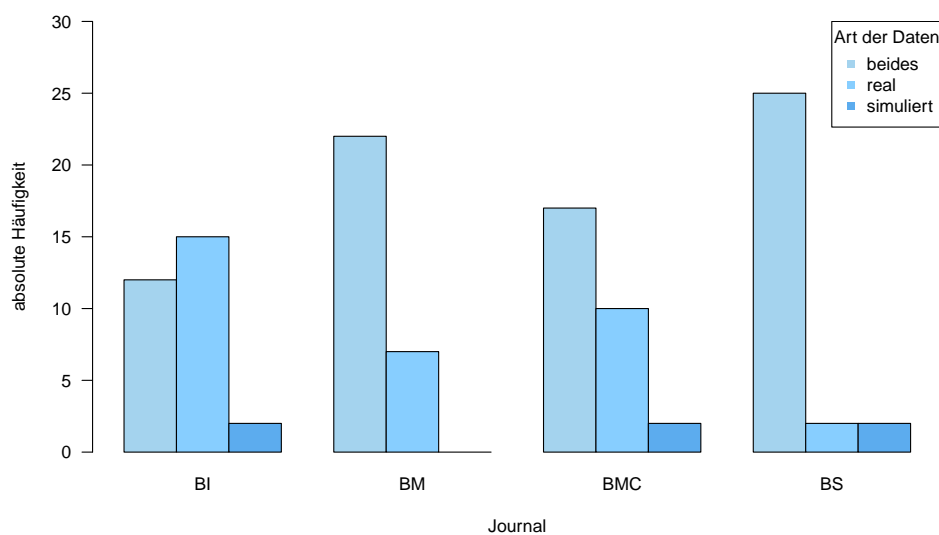


Abbildung 3.3: Art der Daten je nach Journal

In Abbildung 3.4 ist dargestellt, wie häufig in den jeweiligen Passagen des Textes eine positive Bewertung einer Methode auftritt. In 27 Veröffentlichungen erfolgt dies im gesamten Text, also „überall“. Beurteilungen im Abschnitt „discussion“

kommen 21 Mal vor, im Teil „abstr.+disc.“ 19 Mal. In 26 Artikeln taucht an keiner Stelle eine bessere Beurteilung einer Methoden auf.

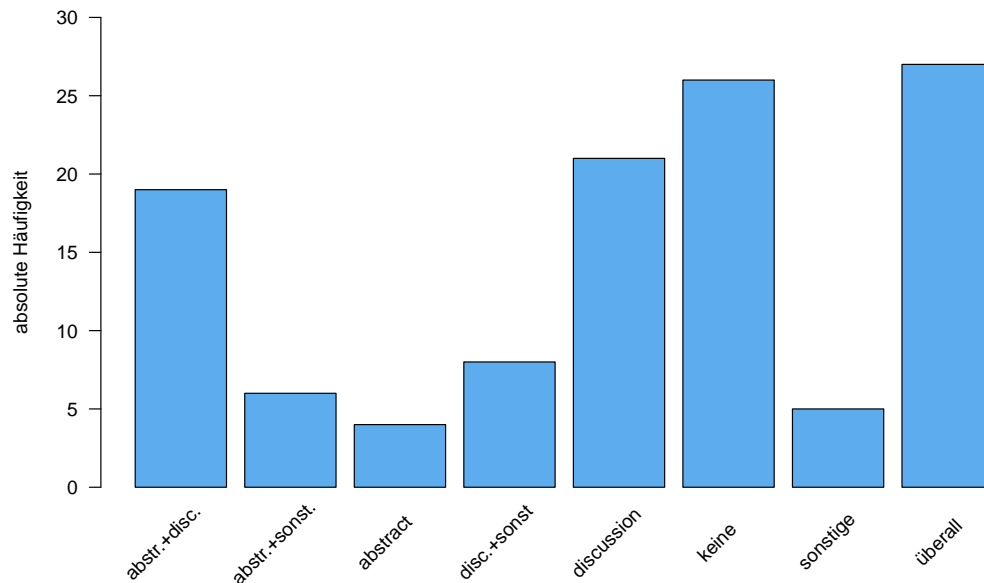


Abbildung 3.4: Stelle an der positive Bewertung auftritt

Abbildung 3.5 zeigt verschiedene Boxplots für die Anzahl verwendeter realer Datensätze (Graphiken A,C und E) bzw. für die Anzahl verwendeter Methoden (Graphiken B,D und E).

In den Graphiken A und B werden die genannten Aspekte je nach Fachzeitschrift betrachtet. Auffällig ist, dass BMC die größte Spannweite mit den meisten Ausreißern aufweist. Die anderen drei Journals verwenden meist unter fünf reale Datensätze. In BMC wurden zwischen 0 und 20 realer Datensätze ausgewertet, sowie zwischen 0 und 11 Methoden verwendet. Wobei null verwendete Methoden so zu verstehen sind, dass keine Methode auf einen realen Datensatz angewendet wird, sondern ausschließlich auf simulierte. Bei Betrachtung des Datensatzes zeigt sich, dass es sich bei dem Artikel mit den elf Methoden auch um den Artikel mit den zwanzig Datensätzen handelt. Wie bereits erwähnt, werden bei BM nie ausschließlich simulierte Daten verwendet, daher reichen die Werte für die Methoden von 1 bis 8. Für die Anzahl verwendeter Methoden liegt der Median für BI und BS bei drei, für BM bei zwei und für BMC bei vier.

Graphiken C und D beziehen sich darauf, ob eine neue Methode eingeführt wird. Dabei zeigt sich, dass bei einer neuen Methode eine höhere Anzahl realer Datensätze zu beobachten ist (Ausreißer). 50% der Artikel, bei denen keine neue Methode eingeführt wird, verwenden zwischen zwei und vier Methoden. Wird eine neue Methode vorgestellt so verwenden 50% der Artikel zwischen drei und vier Methoden.

Der Median liegt hier in beiden Fällen bei drei.

Graphiken E und F konzentrieren sich auf die Art der verwendeten Daten, hierbei können die Ergebnisse für simulierte Daten vernachlässigt werden. Für die Anzahl der Datensätze ist die Spannweite bei realen Daten größer und der Median liegt über dem von beiden Datentypen. Jedoch finden sich bei Verwendung beider Datentypen eine größere Anzahl realer Datensätze. Bei der Anzahl Methoden zeigt sich sowohl für beide Datentypen als auch für reale Daten, dass 50% der Artikel zwischen zwei und vier Methoden verwenden und der Median bei drei liegt. Eine höhere Anzahl Methoden wird bei Artikeln verwendet, in welchen beide Arten von Daten verwendet werden.

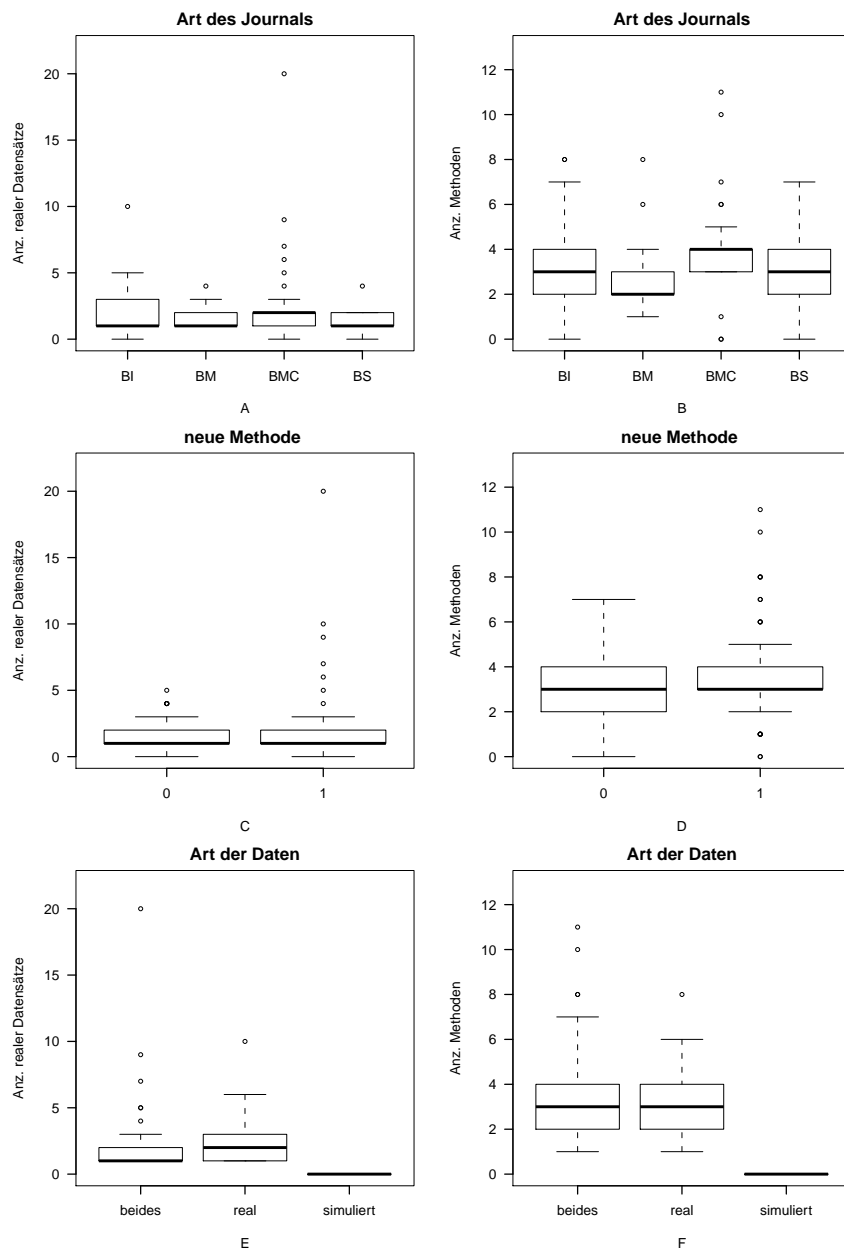


Abbildung 3.5: Boxplots bezüglich der Anzahl realer Datensätze und Methoden

Abbildung 3.6 stellt Boxplots für die Anzahl angewendeter realer Datensätze getrennt nach Verwendung der fünf Schlüsselwörter (Graphiken A-E) und sonstigen Begriffen (Graphik F) dar. Graphik A konzentriert sich z.B. darauf, ob der Begriff „outperform“ verwendet wurde oder nicht, für die Verwendung des Begriffs ist die Spannweite größer, der Median liegt bei zwei.

Für den Begriff „improve“, Graphik B, weisen beide Boxplots eine ähnliche Spannweite auf und der Median liegt beidesmal bei eins. Für die Verwendung von „im-

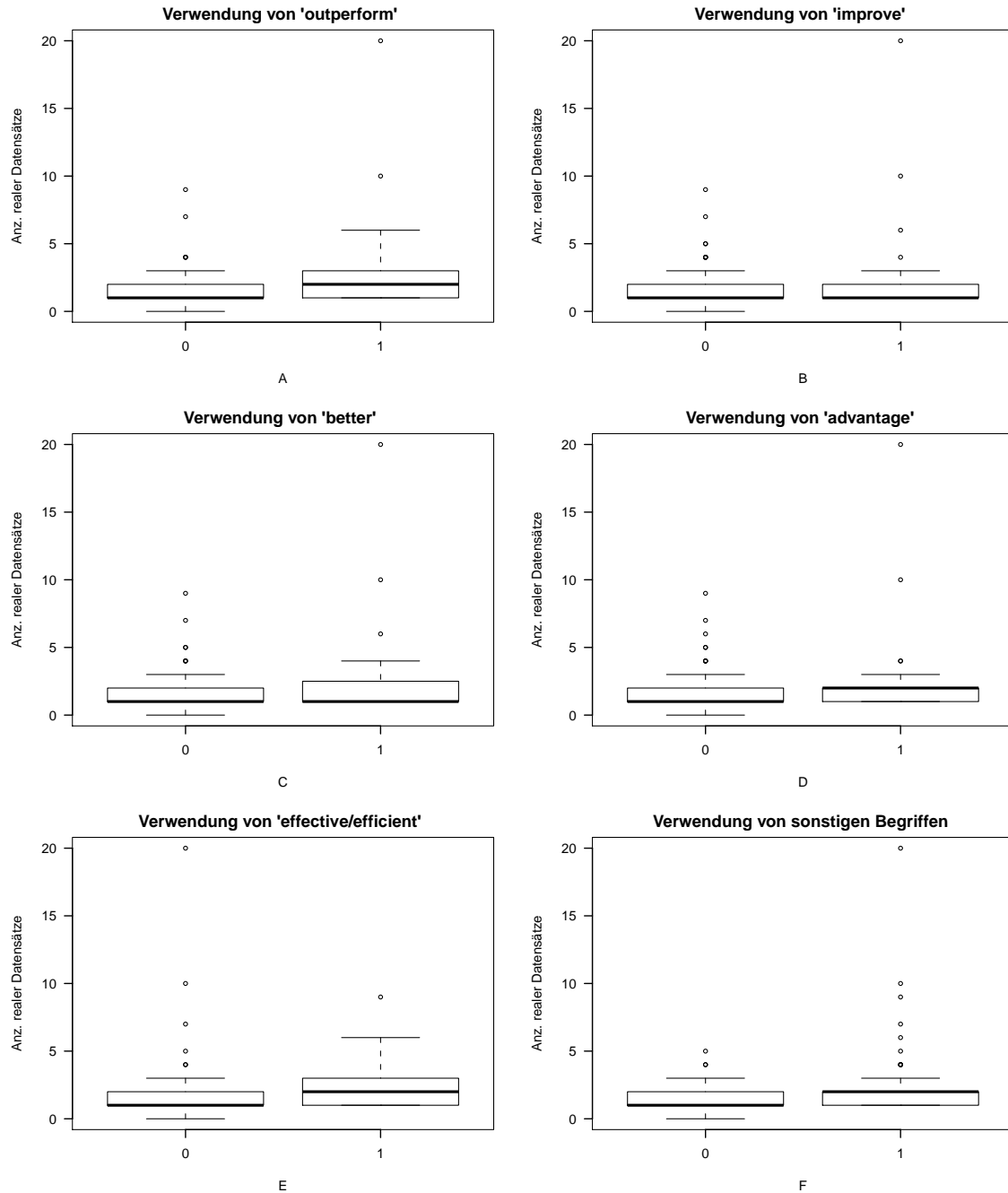


Abbildung 3.6: Boxplots für die Schlüsselwörter bezüglich der Anzahl realer Datensätze

prove“ kann jedoch der Maximalwert von 20 Datensätzen beobachtet werden. Bei näherer Betrachtung aller Graphiken fällt auf, dass der Wert für 20 reale Datensätze in fast jeder Graphik bei eins, also bei „ja“ auftaucht. Nur bei „effective/efficient“ taucht er bei 0, also „nein“ auf. Im gesamten Datensatz gibt es nur einen Artikel in dem 20 reale Datensätze verwendet werden, es handelt sich also in jeder Graphik um dieselbe Ausprägung. Somit werden bei diesem Artikel in besonders hohem Maße Methoden hervorgehoben, im Datensatz zeigt sich zudem, dass in diesem Fall zusätzlich der Begriff „signifikant“ auftaucht.

Bei Graphik F ist beim Auftreten von sonstigen Begriffen ein höherer Median für die Anzahl Datensätze zu beobachten, zudem ist die Spannweite größer im Vergleich zum Nichtverwenden sonstiger Begriffe .

Graphik A, Abbildung 3.7, zeigt, dass positive Bewertungen bei einer hohen Anzahl verwendeter realer Datensätze auftauchen. Die Spannweite hier reicht von 1 bis 20 und zudem wird ersichtlich, dass Artikel mit mehr als fünf realen Datensätzen Methoden positiv bewerten.

Graphik B stellt Boxplots für die Anzahl verwendeter Methoden je nach Beurteilung dar und zeigt für die Abstinenz von positiven Bewertungen eine höhere Streuung im Vergleich zu positiven Bewertungen. Für eine vorteilhaftere Beurteilung liegt der Median bei drei, somit verwenden in diesem Fall 50% der Artikel weniger als drei und die anderen 50% mehr als drei Methoden. In 50% der Berichte, die keine positive Einschätzung geben, werden zwischen einer und vier Methoden gebraucht. Zudem ist an den Ausreißern zu erkennen, wenn positiv bewertet wird, taucht auch eine höhere Anzahl Methoden auf.

Graphik C bezieht sich auf die Anzahl realer Datensätze je nach signifikanten Ergebnissen. Es zeigt sich, dass die Spannweite der Anzahl realer Datensätze für signifikante Ergebnisse größer ist. Signifikante Ergebnisse werden somit bei Artikeln mit einer höheren Anzahl realer Datensätze erwähnt.

Graphik D behandelt die Anzahl Methoden je nach signifikanten Ergebnissen. Die Streuung für nicht signifikant und signifikant ist in etwa gleich hoch, für beide Fälle werden in 50% der Berichte zwischen zwei und vier Methoden verwendet. Die Ausreißer zeigen, dass bei Verwendung von mehr als acht Methoden signifikante Ergebnisse erwähnt werden. Die Spannweite für die Anzahl Methoden reicht bei nicht signifikanten Ergebnissen von null bis acht. Es wird deutlich, dass positive Bewertungen bei Artikeln mit größerer Anzahl realer Datensätze auftauchen. Allerdings kommt es in dieser Arbeit nur siebenmal vor, dass fünf oder mehr Datensätze verwendet werden (Graphik B, Abb. 3.1).

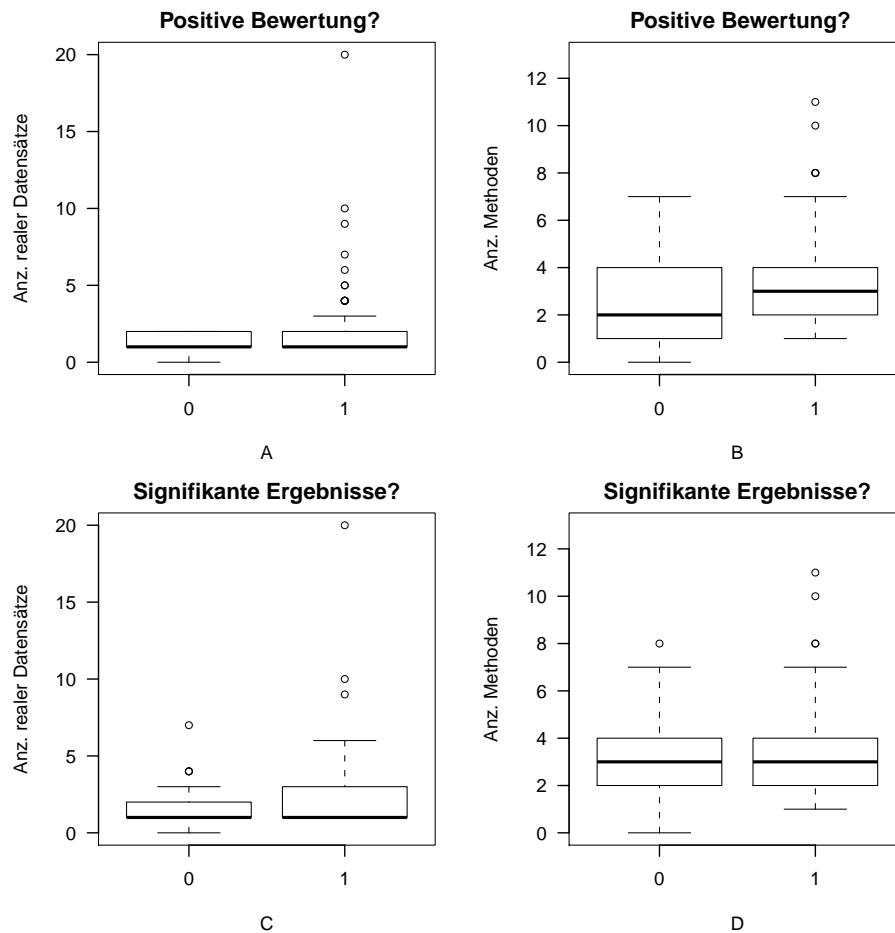


Abbildung 3.7: Boxplots je nach Bewertung und Signifikanz

Mosaicplots für das Auftreten der fünf Schlüsselwörter und sonstiger Begriffe je nach Journal sind in Abbildung 3.8 dargestellt, die Zahlen bezeichnen die absoluten Häufigkeiten. In Graphik A-E lässt sich erkennen, dass die Schlüsselwörter überwiegend nicht verwendet werden. Diese fünf am häufigsten genannten Schlagwörter kommen zwar in mehreren Artikeln vor, allerdings tritt jedes in nur etwa einem Viertel aller Artikel auf, BI und BMC verwenden bei der Beurteilung häufiger diese Wörter. In etwa gleich vielen Artikeln der Journals, mit Ausnahme von BMC, wird ein sonstiger Begriff zur Hervorhebung verwendet bzw. nicht verwendet (Graphik F). Aus dieser Abbildung 3.8 könnte der Eindruck hervorgehen, dass in den Veröffentlichungen selten Methoden bevorzugt werden. Allerdings wird hier nur die Verwendung der fünf Schlüsselwörter und sonstiger Begriffe betrachtet. Wird sich jetzt aber nur darauf konzentriert, ob ein Artikel Methoden hervorhebt, zeigt sich ein anderes Bild.

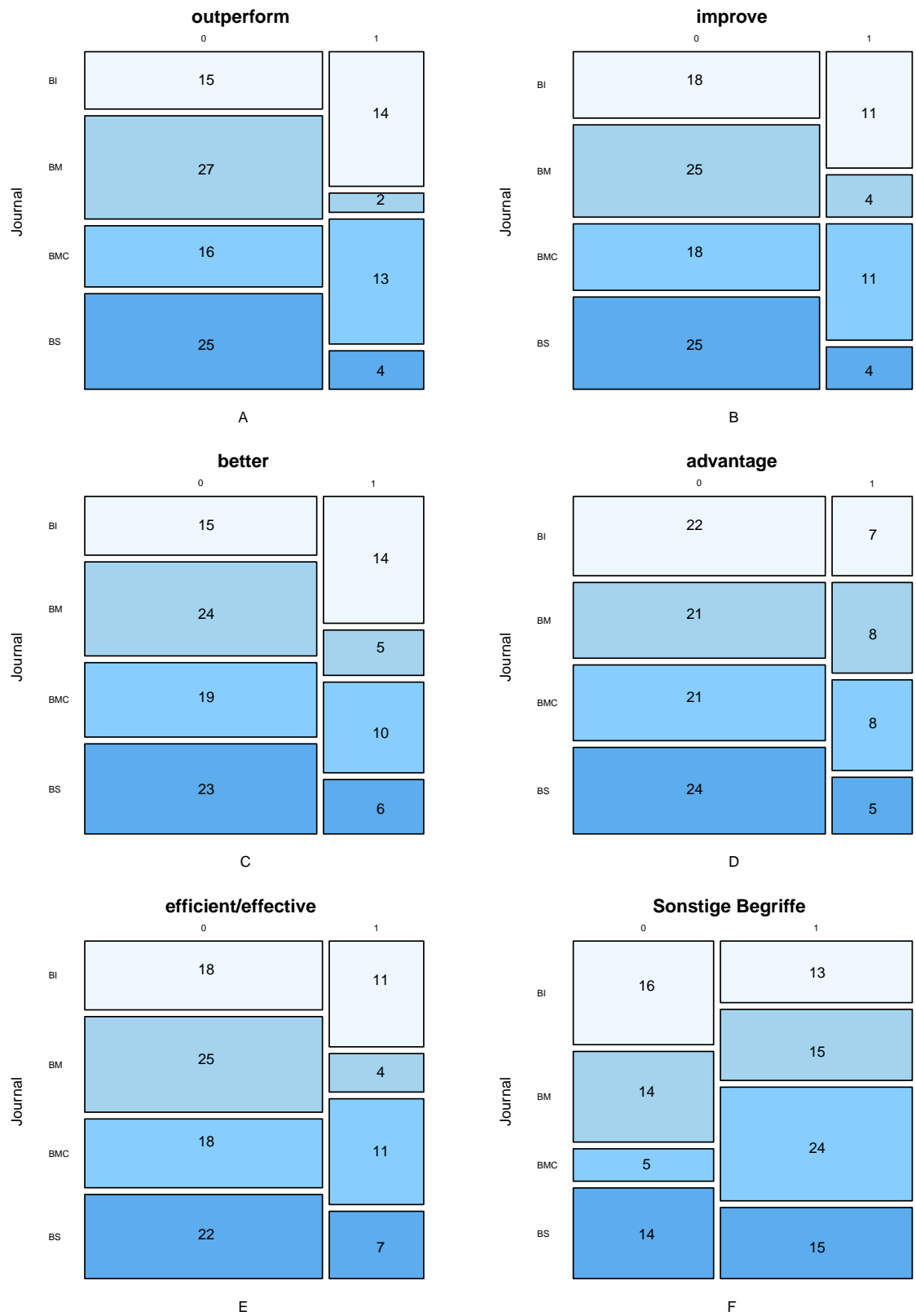


Abbildung 3.8: Verwendung der verschiedenen Schlüsselwörter je nach Journal

In Abbildung 3.9 wird sich nicht darauf fokussiert wie bewertet wird, sondern ob bewertet wird. Dabei ist deutlich zu erkennen, dass in allen Fachzeitschriften häufig eine Bevorzugung von Methoden geschildert wird. In Zahlen ausgedrückt bedeutet dies, dass in 90 von 116 Artikeln positive Bewertungen ausgesprochen werden, also in mehr als Dreiviertel der Veröffentlichungen. Bei der Betrachtung der einzelnen Zeitschriften zeigt sich, dass in 89.7% der BMC-Artikel eine bessere Beurteilung auftritt, gefolgt von BI mit 82.8% , bei BM sind es ca. 72.4% und BS liefert mit 65.5% den niedrigsten Wert. Somit zeigt sich, dass die BMC-Artikel im Vergleich zu den anderen optimistischer gegenüber den Leistungen von Methoden sind. Wobei auch die anderen Berichte diesbezüglich nicht pessimistisch sind.

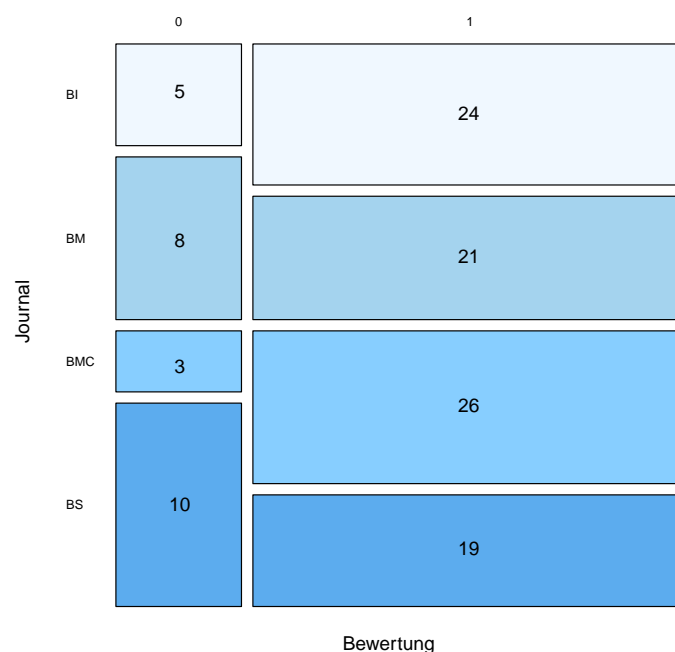


Abbildung 3.9: Bewertung je nach Journal

Wie in Kapitel 2 bereits erwähnt, ist es auch von Bedeutung darauf zu achten, ob Signifikanz in den Artikeln berücksichtigt wird. Daher wird sich in den nächsten zwei Abbildungen (3.10 & 3.11) mit der Variable „signifikant“ beschäftigt.

Abbildung 3.10 veranschaulicht, ob und wie häufig der Begriff „signifikant“ in den Arbeiten verwendet wird, auf der y-Achse ist jeweils die absolute Häufigkeit abgetragen. Graphik A konzentriert sich auf die Art der Daten. Für beide Datentypen tauchen 30 Mal signifikante Ergebnisse auf, in 46 Fällen kann keine Signifikanz festgestellt werden. Bei der ausschließlichen Verwendung realer Datensätze wurde in etwa gleich vielen Fällen Signifikanz berücksichtigt bzw. nicht berücksichtigt,

der letztere Fall ist etwas häufiger.

Ob ein Bezug zur Signifikanz beobachtet werden konnte ist in Graphik B dargestellt. In 70 der 116 Artikel wurden keine signifikanten Erwähnungen gefunden.

Graphik C veranschaulicht, wie häufig signifikante Ergebnisse je nach Bewertung vorkommen. Hier ist wiederum zu erkennen, dass in den meisten Fällen eine bessere Bewertung von Methoden auftritt. Allerdings zeigt sich in 43 von 90 Fällen eine positive Bewertung ohne Erwähnung signifikanter Verbesserungen.

Graphik D zeigt, ob eine neue Methode verwendet wird und wie häufig dabei signifikante Ergebnisse genannt werden. Es ist zu sehen, dass meistens mehr neue Methoden eingeführt werden, dabei überwiegend aber keine Signifikanz erwähnt wird. Insgesamt zeichnet sich der Eindruck ab, dass die positive Hervorhebung von Methoden nicht auf die Grundgesamtheit übertragen werden kann.

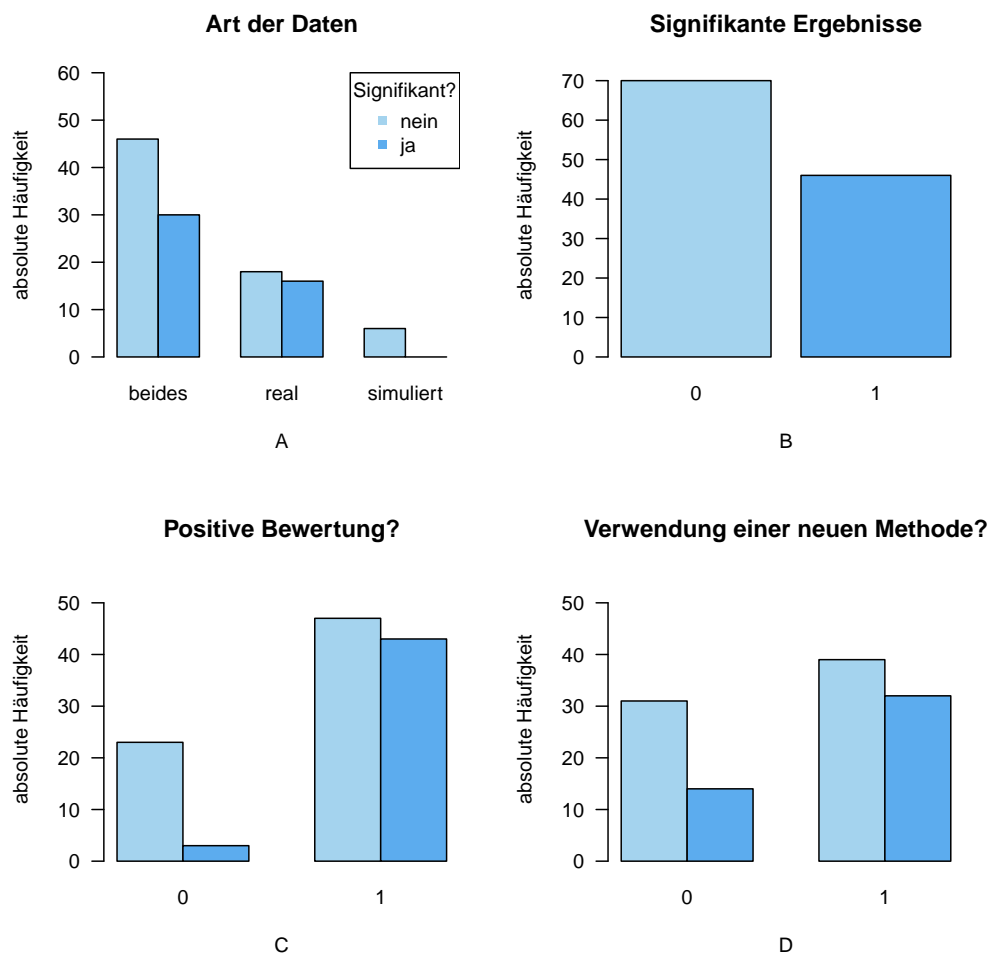


Abbildung 3.10: Deskriptive Auswertungen für die Variable „signifikant“

Das Auftreten signifikanter Ergebnisse je nach Journal ist in Abbildung 3.11 verdeutlicht. BI und BMC beziehen sich eher auf signifikante Ergebnisse als die anderen beiden Magazine. Am häufigsten weist BI signifikante Erwähnungen auf, dieses ist nach BMC das Journal mit den meisten positiven Bewertungen, was aus Abb. 3.9 hervorgeht. In fünf von 29 BS-Artikeln können signifikante Ergebnisse gefunden werden. Zudem ist dies die Kategorie in der die wenigsten positiven Wertungen auftraten, was aus Abbildung 3.9 hervorgeht. In nur etwa einem Drittel aller Artikel wird Signifikanz ausdrücklich erwähnt.

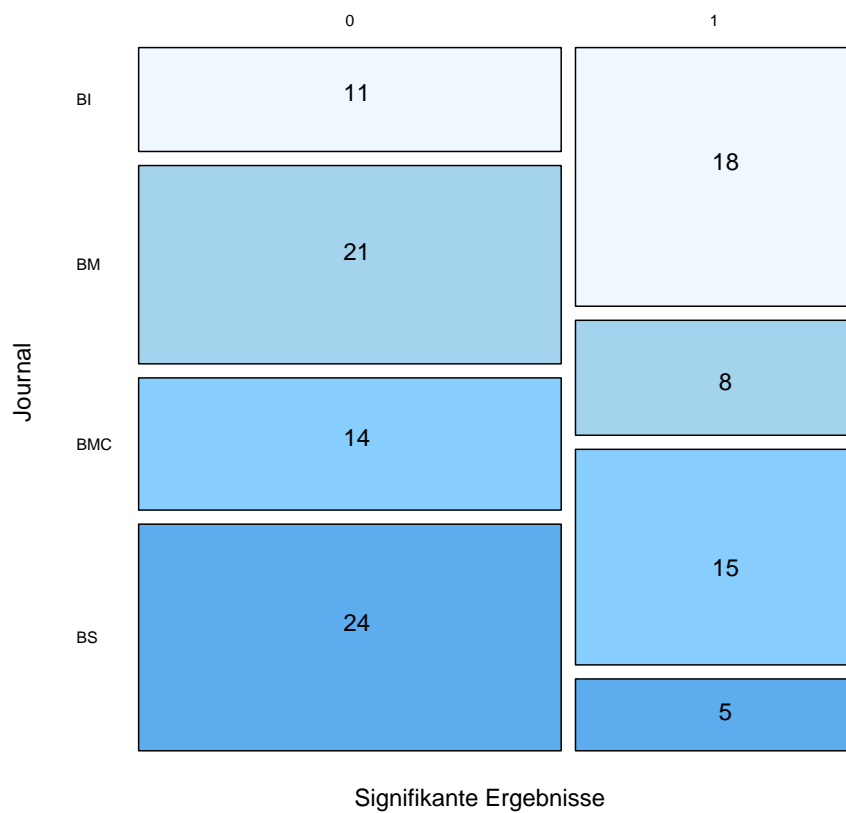


Abbildung 3.11: Auftreten von signifikanten Ergebnissen je nach Journal

Im Anhang finden sich weitere Auswertungen sowie der R-Code für alle erzeugten Abbildungen.

3.3 Zusammenfassung der Ergebnisse

Bei den betrachteten Artikeln liegt die Anzahl der verwendeten realen Datensätze im Mittel bei 1.9, nur siebenmal wurden mehr als vier Datensätze verwendet. Im Schnitt werden drei Methoden auf reale Daten angewendet und insgesamt 72 neue Methoden eingeführt. 90 der 116 Artikel bewerten eine Methode anderen gegenüber für besser.

In allen vier Magazinen werden bestimmte Methoden optimistisch dargestellt, was sich daran zeigt, dass in etwa 75% der Artikel diese Methoden positiv hervorgehoben werden, besonders häufig geschah dies bei BMC. In fast Zweidrittel aller Artikel konnte jedoch keine Erwähnung von Signifikanz beobachtet werden.

Die Ergebnisse zeigen, dass Artikel mit einer größeren Anzahl realer Datensätze und Methoden sowohl eher positive als auch signifikante Bewertungen aussprechen. Hierbei sei noch darauf hingewiesen, dass der Artikel mit der höchsten Anzahl realer Datensätze die meisten Methoden verwendet. Ausserdem treten dort positive Bewertungen mittels vier der fünf Schlüsselwörter und sonstiger Begriffe auf und weist einen Bezug zur Signifikanz auf. Allerdings kann es sich auch um Zufall handeln, da für genauere Schlüsse Artikel mit ähnlich hoher Anzahl realer Datensätze fehlen.

Der Vollständigkeit halber sei erwähnt, dass in einer geringen Anzahl von Artikeln auch die Größe der Datensätze in die Bewertung miteinbezogen wird. So wird in vier Artikeln erwähnt, dass Methoden bei größeren Datensätzen eine bessere Leistung erbringen. Ein Artikel hingegen besagt, dass die vorgestellte Methode bei kleinen Datensätzen besser funktioniert. Diese Aussagen beruhen auf der Analyse von maximal zwei realen Datensätzen. Weshalb nicht gesichert ist, dass große Datensätze für eine Methode geeigneter sind. Zudem ist nur eine dieser Aussagen signifikant.

Kapitel 4

Fazit

In biostatistischen und bioinformatischen Fachzeitschriften ist es gängig, dass in Arbeiten Methoden einander gegenüber gestellt und verglichen werden. Meist handelt es sich bei einer dieser Methoden um eine Neuentwicklung. Dabei werden in den Aussagen der Berichte häufig neue Methoden den anderen gegenüber als besser dargestellt. Es ist jedoch zu hinterfragen, ob Autoren objektiv vergleichen. Neue Methoden können durchaus Erleichterungen mit sich bringen, wie z.B. eine schnellere Laufzeit oder einfachere Anwendung. In dieser Arbeit jedoch hat sich herausgestellt, dass Aussagen wie „our method outperforms existing methods“ oder „our method shows a better performance“ mit Skepsis betrachtet werden sollten. Häufig wird ein Ergebnis akzeptiert ohne darauf zu achten, ob es sich um ein signifikantes handelt. Zudem stützen sich die Ergebnisse meist nur auf wenige reale Datensätze, weshalb es nicht möglich ist eine Methode im Allgemeinen als besser zu bezeichnen. Denn die Wahrscheinlichkeit, dass eine Methode valide ist, steigt mit dem Stichprobenumfang und der Häufigkeit der erfolgreichen Anwendung. Selbstverständlich kann keine Methode für jeden Datensatz optimal funktionieren. Dass eine Methode für einen Datensatz ein besseres Ergebnis erbringt, heißt nicht, dass sie dies auch für andere Datensätze erbringt. Es sollte für jeden Datensatz bzw. für jede Fragestellung abgewägt werden, welche Methode am geeignetsten ist. Aus den genannten Gründen sollte den Ergebnissen und Aussagen von wissenschaftlichen Artikeln nicht bedingungslos vertraut und diese nicht für allgemeingültig anerkannt werden.

Kapitel 5

Anhang

Weitere Graphiken

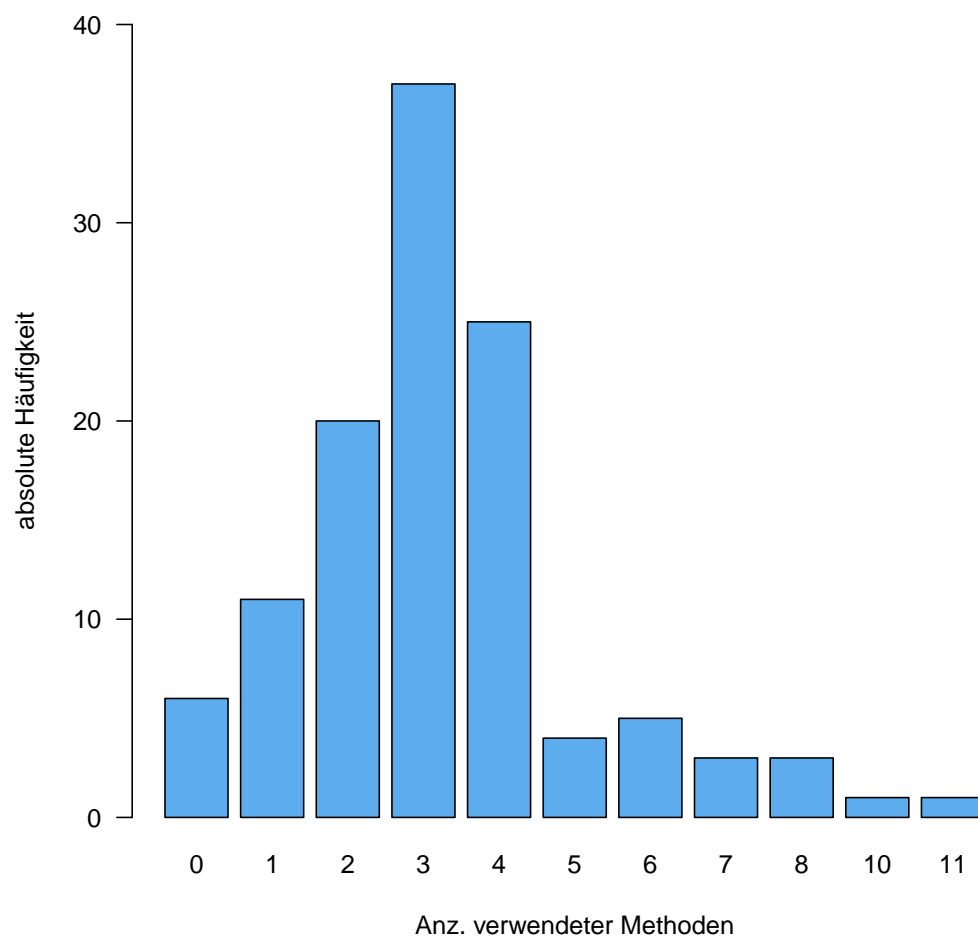


Abbildung 5.1: Häufigkeiten der verwendeten Methoden

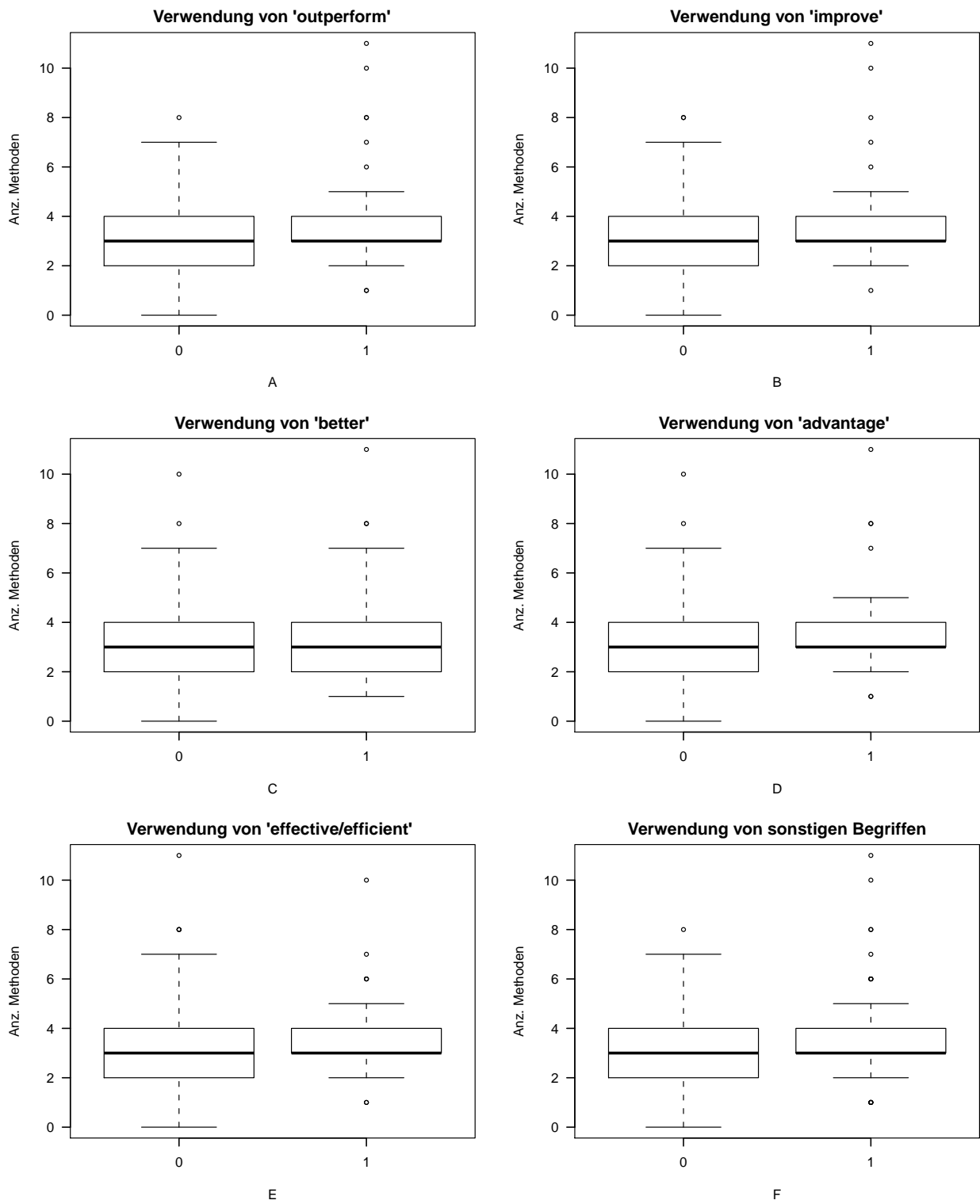


Abbildung 5.2: Boxplots für die Schlüsselwörter bezogen auf die Anzahl Methoden

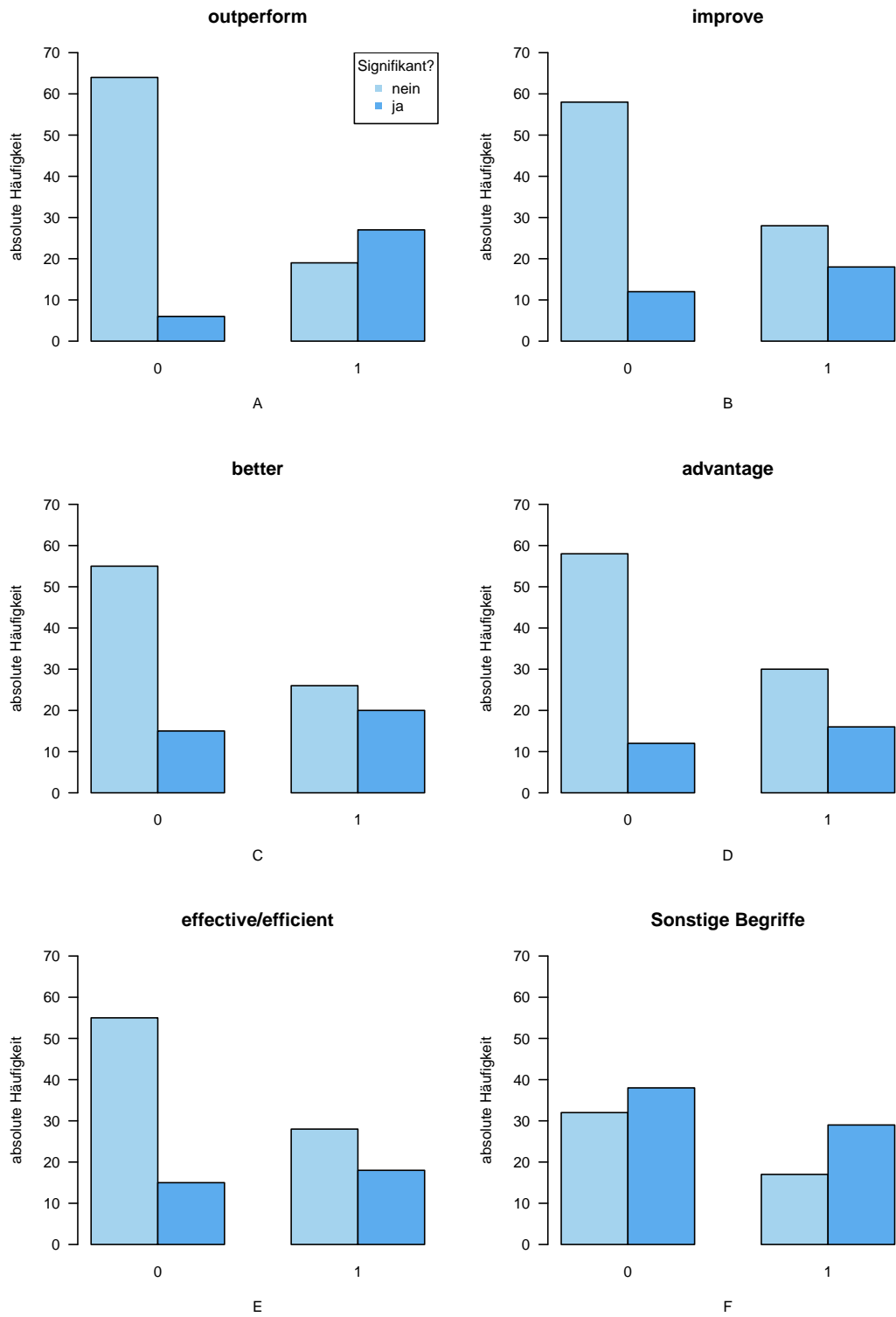


Abbildung 5.3: Verwendung von signifikant je nach Schlüsselwort

Abkürzungsverzeichnis

Abkürzung		Beschreibung
abstr.+disc.	abstract plus discussion	Zusammenfassung und Fazit; bezeichnet Stelle an der im Text bewertet wird
abstr.+sonst.	abstract plus sonstige	Zusammenfassung und sonstige Stelle; bezeichnet Stelle an der im Text bewertet wird
Anz.	Anzahl	
BI	„Bioinformatics“	Journal
BM	„Biometrics“	Journal
BMC	„BMC Bioinformatics“	Journal
BS	„Biostatistics“	Journal
disc.+sonst.	discussion plus sonstige	Fazit und sonstige Stelle; bezeichnet Stelle an der im Text bewertet wird

Abbildungsverzeichnis

3.1	Überblick über die Daten	8
3.2	Verwendung einer neuen Methode je nach Journal	9
3.3	Art der Daten je nach Journal	9
3.4	Stelle an der positive Bewertung auftritt	10
3.5	Boxplots bezüglich der Anzahl realer Datensätze und Methoden . . .	11
3.6	Boxplots für die Schlüsselwörter bezüglich der Anzahl realer Daten- sätze	12
3.7	Boxplots je nach Bewertung und Signifikanz	14
3.8	Verwendung der verschiedenen Schlüsselwörter je nach Journal . . .	15
3.9	Bewertung je nach Journal	16
3.10	Deskriptive Auswertungen für die Variable „signifikant“	17
3.11	Auftreten von signifikanten Ergebnissen je nach Journal	18
5.1	Häufigkeiten der verwendeten Methoden	21
5.2	Boxplots für die Schlüsselwörter bezogen auf die Anzahl Methoden .	22
5.3	Verwendung von signifikant je nach Schlüsselwort	23

Tabellenverzeichnis

3.1	Verwendete Journals und Ausgaben	5
3.2	Variablen des Datensatzes	7

Literaturverzeichnis

Boulesteix, A.-L. (2013c). On representative and illustrative comparisons with real data in bioinformatics: response to the letter to the editor by smith et al.

Boulesteix, A.-L., Hable, R., Lauer, S. and Eugster, M. J. (2013a). A statistical framework for hypothesis testing in real data comparison studies.

URL: <http://epub.ub.uni-muenchen.de/14324/>

Boulesteix, A.-L., Lauer, S. and Eugster, M. J. (2013b). A plea for neutral comparison studies in computational sciences.

URL: <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0061562>

Diverse (2013a). *Bioinformatics* **29**(1&2).

URL: <http://bioinformatics.oxfordjournals.org>

Diverse (2013b). *Biometrics* **69**(1).

URL: <http://www.biometrics.tibs.org/>

Diverse (2013c). *BMC Bioinformatics* **14**.

URL: <http://www.biomedcentral.com/bmcbioinformatics/content>

Diverse (2013d). *Biostatistics* **14**(1&2).

URL: <http://biostatistics.oxfordjournals.org>

Fahrmeir, L., Kuenstler, R., Pigeot, I. and Tutz, G. (2010). *Statistik, Der Weg zur Datenanalyse*, Springer-Verlag.

Huett, M.-T. and Dehnert, M. (2006). *Methoden der Bioinformatik: Eine Einführung*, Springer.

Janczyk, M. and Pfister, R. (2013). *Inferenzstatistik verstehen*, Springer Spektrum.

Koehler, W., Schachtel, G. and Voleske, P. (2012). *Biostatistik: Eine Einführung für Biologen und Agrarwissenschaftler*, Springer Spektrum.

Wolpert, D. (2001). The supervised learning no-free-lunch theorems, *Proceedings of the 6th Online World Conference on Soft Computing in Industrial Applications*

R-Code

```
### Einlesen des Datensatzes ###
data<-read.csv("C:\\Users\\Administrator\\Desktop\\bachelorarbeit\\datensatz2.csv",
header=TRUE,sep=";",dec=".")
head(data)

###Umkodieren der Variable daten ###
daten_rec<-c()
daten_rec[data$daten %in% c("beides", "beides ")]<-"beides"
daten_rec[data$daten %in% c("real")]<-"real"
daten_rec[data$daten %in% c("simuliert")]<-"simuliert"
data$daten_rec <- daten_rec

### Abbildung 3.1; Überblick der Daten ###
par(mfrow=c(2,2),las=1,mar=c(4.5,4.5,2,2))
barplot(table(data$daten_rec), ylim=c(0,80),xlab="A",
col=c("lightskyblue2","skyblue1","steelblue2"), main="Art der verwendeten Daten",
ylab="absolute Häufigkeit" )

barplot(table(data$reale_daten), ylim=c(0,65), col=c("steelblue2"), xlab="B",
main = "Anz. realer Datensätze", ylab="absolute Häufigkeiten")

barplot(table(data$neue_methode), main="neue Methode?", xlab="C",
col=c("lightskyblue2","steelblue2"), ylim=c(0,80),ylab="absolute Häufigkeit")

barplot(table(data$bewertung),main="positive Bewertung?", xlab="D",
ylim=c(0,100), col=c("lightskyblue2","steelblue2"),ylab="absolute Häufigkeit")

dev.off()

### Abbildung 3.2; Verwendung einer neuen Methode je nah Journal ###
par(las=1)
barplot(xtabs(~data$neue_methode+data$katgorie), ylim=c(0,30),beside=TRUE,
col=c("lightskyblue2","steelblue2"), xlab="Journal",ylab="absolute Häufigkeit")
legend("topright",legend=c("nein", "ja"),col=c("lightskyblue2","steelblue2"),
pch=15, title="Neue Methode?")

dev.off()
```



```

### Abbildung 3.3; Art der Daten je nach Journal ###
par(las=1)
barplot(xtabs(~data$daten_rec+data$katgorie),ylim=c(0,30), beside=TRUE,
col=c( "lightskyblue2","skyblue1","steelblue2"), xlab="Journal",
ylab="absolute Häufigkeit")
legend("topright",legend=c("beides", "real", "simuliert"),
col=c( "lightskyblue2","skyblue1","steelblue2"), pch=15, title="Art der Daten")

dev.off()

### Abbildung 3.4; Stelle an der positive Bewertung auftritt ###
par(las=1)
barplot(table(data$Stelle), ylim=c(0,30),names.arg="", col="steelblue2",
ylab="absolute Häufigkeit")
text(seq(0.5,9, length.out=8), par("usr")[3] -3.5, srt = 45, adj = 0.5,
labels = levels(data$Stelle), xpd = TRUE)

dev.off()

### Abbildung 3.5;Boxplots für reale Datensätze und Methoden ###

par(mfrow=c(3,2),las=1, mar=c(4.2,4,2,2), lwd=0.75)

boxplot(data$reale_daten~data$katgorie, main="Art des Journals", xlab= " A ",
ylim=c(0,22), ylab="Anz. realer Datensätze")
boxplot(data$methoden~data$katgorie, main ="Art des Journals", xlab="B",
ylim=c(0,13),ylab="Anz. Methoden")
boxplot(data$reale_daten~data$neue_methode,xlab="C", ylim=c(0,22),
ylab="Anz. realer Datensätze", main="neue Methode ")
boxplot(data$methoden~data$neue_methode,xlab="D", ylim=c(0,13),
ylab="Anz. Methoden", main="neue Methode ")
boxplot(data$reale_daten~data$daten_rec,xlab="E", ylim=c(0,22),
ylab="Anz. realer Datensätze", main="Art der Daten")
boxplot(data$methoden~data$daten_rec,xlab="F", ylim=c(0,13),
ylab="Anz. Methoden", main="Art der Daten")

dev.off()

```

```

### Abbildung 3.6;Boxplots für die Schlüsselwörter bezüglich der
# Anzahl realer Datensätze ###

par(mfrow=c(3,2),las=1,mar=c(5,4.2,2,2),lwd=0.75)
boxplot(data$reale_daten~data$outperform,main="Verwendung von 'outperform' ",
ylab="Anz. realer Datensätze",xlab="A")

boxplot(data$reale_daten~data$improve,main="Verwendung von 'improve' ",
ylab="Anz. realer Datensätze",xlab="B")

boxplot(data$reale_daten~data$better,main="Verwendung von 'better' ",
ylab="Anz. realer Datensätze",xlab="C")

boxplot(data$reale_daten~data$advantage,main="Verwendung von 'advantage' ",
ylab="Anz. realer Datensätze",xlab="D")

boxplot(data$reale_daten~data$effective_efficient,xlab="E" ,
main="Verwendung von 'effective/efficient' ",ylab="Anz. realer Datensätze" )

boxplot(data$reale_daten~data$sonstiges,ylab="Anz. realer Datensätze",xlab="F",
main="Verwendung von sonstigen Begriffen ")

dev.off()

### Abbildung 3.7; Boxplots für positive Bewertung und Signifikanz ###

par(mfrow=c(2,2),las=1,mar=c(4.2,4,2,2),lwd=0.75)
boxplot(data$reale_daten~data$bewertung,xlab="A",main=" Positive Bewertung?",
main="Anzahl realer Datensätze (n=116)",ylab="Anz. realer Datensätze")

boxplot(data$methoden~data$bewertung,xlab="B",main=" Positive Bewertung?",
ylim=c(0,13),main="Anzahl Methoden (n=116)",ylab="Anz. Methoden")

boxplot(data$reale_daten~data$signifikant,main="Signifikante Ergebnisse?",
xlab="C",ylab="Anz. realer Datensätze")

boxplot(data$methoden~data$signifikant,main="Signifikante Ergebnisse?",
ylim=c(0,13),xlab="D",ylab="Anz. Methoden")

dev.off()

```

```

### Abbildung 3.8; Verwendung der Schlüsselwörter je nach Journal ###
par(mfrow=c(3,2), las=1, mar=c(3,3,2,2))
mosaicplot(xtabs(~data$outperform+data$kategorie), xlab="A", main="outperform ",
ylab="Journal", col=c("aliceblue", "lightskyblue2", "skyblue1", "steelblue2"))
text (x=0.4, y = 0.9, c("15"))   text (x=0.87, y = 0.8, c("14"))
text (x=0.4, y = 0.63, c("27"))  text (x=0.87, y = 0.54, c("2"))
text (x=0.4, y = 0.38, c("16"))  text (x=0.87, y = 0.3, c("13"))
text (x=0.4, y = 0.1, c("25"))   text (x=0.87, y = 0.02, c("4"))

mosaicplot(xtabs(~data$improve+data$kategorie), main="improve", xlab=" B",
ylab="Journal", col=c("aliceblue", "lightskyblue2", "skyblue1", "steelblue2"))
text (x=0.4, y = 0.9, c("18"))   text (x=0.87, y = 0.8, c("11"))
text (x=0.4, y = 0.63, c("25"))  text (x=0.87, y = 0.54, c("4"))
text (x=0.4, y = 0.38, c("18"))  text (x=0.87, y = 0.3, c("11"))
text (x=0.4, y = 0.1, c("25"))   text (x=0.87, y = 0.02, c("4"))

mosaicplot(xtabs(~data$better+data$kategorie), main="better", xlab=" C",
ylab="Journal", col=c("aliceblue", "lightskyblue2", "skyblue1", "steelblue2"))
text (x=0.4, y = 0.9, c("15"))   text (x=0.87, y = 0.8, c("14"))
text (x=0.4, y = 0.63, c("24"))  text (x=0.87, y = 0.53, c("5"))
text (x=0.4, y = 0.38, c("19"))  text (x=0.87, y = 0.27, c("10"))
text (x=0.4, y = 0.1, c("23"))   text (x=0.87, y = 0.04, c("6"))

mosaicplot(xtabs(~data$advantage+data$kategorie), main="advantage", xlab=" D",
ylab="Journal", col=c("aliceblue", "lightskyblue2", "skyblue1", "steelblue2"))
text (x=0.4, y = 0.9, c("22"))   text (x=0.88, y = 0.87, c("7"))
text (x=0.4, y = 0.63, c("21"))  text (x=0.88, y = 0.57, c("8"))
text (x=0.4, y = 0.38, c("21"))  text (x=0.88, y = 0.3, c("8"))
text (x=0.4, y = 0.1, c("24"))   text (x=0.88, y = 0.04, c("5"))

mosaicplot(xtabs(~data$effective_efficient+data$kategorie), xlab="E", ylab="Journal",
main="efficient/effective ", col=c("aliceblue", "lightskyblue2", "skyblue1", "steelblue2"))
text (x=0.4, y = 0.9, c("18"))   text (x=0.87, y = 0.87, c("11"))
text (x=0.4, y = 0.63, c("25"))  text (x=0.87, y = 0.59, c("4"))
text (x=0.4, y = 0.38, c("18"))  text (x=0.87, y = 0.35, c("11"))
text (x=0.4, y = 0.1, c("22"))   text (x=0.87, y = 0.06, c("7"))

mosaicplot(xtabs(~data$sonstiges+data$kategorie), main="Sonstige Begriffe ",
xlab="F", ylab="Journal", col=c("aliceblue", "lightskyblue2", "skyblue1", "steelblue2"))
text (x=0.27, y = 0.85, c("16"))  text (x=0.73, y = 0.89, c("13"))
text (x=0.27, y = 0.53, c("14"))  text (x=0.73, y = 0.65, c("15"))
text (x=0.27, y = 0.31, c("5"))   text (x=0.73, y = 0.36, c("24"))
text (x=0.27, y = 0.1, c("14"))   text (x=0.73, y = 0.06, c("15"))

dev.off()

```

```
### Abbildung 3.9; Bewertung je nach Journal###
```

```
par(las=1)

mosaicplot(xtabs(~data$bewertung+data$kategorie), main="", xlab="Bewertung",
ylab="Journal", col=c("aliceblue", "lightskyblue2", "skyblue1", "steelblue2"))
text (x=0.16, y = 0.9, c("5"))
text (x=0.16, y = 0.63, c("8"))
text (x=0.16, y = 0.4, c("3"))
text (x=0.16, y = 0.15, c("10"))
text (x=0.63, y = 0.86, c("24"))
text (x=0.63, y = 0.6, c("21"))
text (x=0.63, y = 0.33, c("26"))
text (x=0.63, y = 0.07, c("19"))

dev.off()
```

```
### Abbildung 3.10; Auswertungen für die Variable "signifikant"###
```

```
par(mfrow=c(2,2), las=1)
barplot(xtabs(~data$signifikant+data$daten_rec), ylim=c(0,60), beside=TRUE,
col=c("lightskyblue2", "steelblue2"), ylab="absolute Häufigkeit",
main="Art der Daten", xlab="A")
legend("topright", legend=c("nein", "ja"), col=c("lightskyblue2", "steelblue2"),
pch=15, title="Signifikant?")
barplot(table(data$signifikant), ylim=c(0,72), main="Signifikante Ergebnisse",
col=c("lightskyblue2", "steelblue2"), ylab="absolute Häufigkeit", beside=TRUE, xlab="B")
barplot(xtabs(~data$signifikant+data$bewertung), ylim=c(0,55), beside=TRUE,
col=c("lightskyblue2", "steelblue2"), ylab="absolute Häufigkeit",
main="Positive Bewertung?", xlab="C")
barplot(xtabs(~data$signifikant+data$neue_methode), ylim=c(0,55),
col=c("lightskyblue2", "steelblue2"), ylab="absolute Häufigkeit", beside=TRUE,
main="Verwendung einer neuen Methode?", xlab="D")

dev.off()
```

```
### Abbildung 3.11, Verwendung von signifikant je nach Journal
```

```
par(las=1)
mosaicplot(xtabs(~data$signifikant+data$kategorie), xlab="Signifikante Ergebnisse",
ylab="Journal", col=c("aliceblue", "lightskyblue2", "skyblue1", "steelblue2"), main="")
text (x=0.34, y = 0.92, c("11"))
text (x=0.34, y = 0.69, c("21"))
text (x=0.34, y = 0.42, c("14"))
text (x=0.34, y = 0.15, c("24"))
text (x=0.82, y = 0.82, c("18"))
text (x=0.82, y = 0.5, c("8"))
text (x=0.82, y = 0.23, c("15"))
text (x=0.82, y = 0.02, c("5"))
dev.off()
```

```

### Abbildung 5.1; Häufigkeit der verwendeten Methoden ###
par(las=1)
barplot(table(data$methoden), ylim=c(0,40), col="steelblue2",
xlab="Anz. verwendeter Methoden", ylab="absolute Häufigkeit")

dev.off()

### Abbildung 5.2; Boxplots für Schlüsselwort bezogen auf die Anzahl Methoden###
par(mfrow=c(3,2), las=1, mar=c(5,4.2,2,2), lwd=0.75)
boxplot(data$methoden~data$outperform, main="Verwendung von 'outperform'",
ylab="Anz. Methoden", xlab="A")
boxplot(data$methoden~data$improve, main="Verwendung von 'improve'",
ylab="Anz. Methoden", xlab="B")
boxplot(data$methoden~data$better, main="Verwendung von 'better'",
ylab="Anz. Methoden", xlab="C")
boxplot(data$methoden~data$advantage, main="Verwendung von 'advantage' ",
ylab="Anz. Methoden", xlab="D")
boxplot(data$methoden~data$effective_efficient, ylab="Anz. Methoden",
xlab="E", main="Verwendung von 'effective/efficient' ")
boxplot(data$methoden~data$sonstiges, main="Verwendung von sonstigen Begriffen",
ylab="Anz. Methoden", xlab="F")

dev.off()

### Abbildung 5.3, Verwendung von signifikant je nach Schlüsselwort ###

par(mfrow=c(3,2), las=1)
barplot(xtabs(~data$outperform+data$signifikant), ylim=c(0,70), xlab="A",
beside=TRUE, col=c("lightskyblue2", "steelblue2"),
ylab="absolute Häufigkeit", main="outperform")
legend("topright", legend=c("nein", "ja"), col=c("lightskyblue2", "steelblue2"),
pch=15, title="Signifikant?")
barplot(xtabs(~data$improve+data$signifikant), ylim=c(0,70), xlab="B", beside=TRUE,
col=c("lightskyblue2", "steelblue2"), ylab="absolute Häufigkeit", main="improve")
barplot(xtabs(~data$better+data$signifikant), ylim=c(0,70), xlab="C", beside=TRUE,
col=c("lightskyblue2", "steelblue2"), main="better", ylab="absolute Häufigkeit")
barplot(xtabs(~data$advantage+data$signifikant), ylim=c(0,70), xlab="D",
beside=TRUE, main="advantage", ylab="absolute Häufigkeit",
col=c("lightskyblue2", "steelblue2"))
barplot(xtabs(~data$effective_efficient+data$signifikant), ylim=c(0,70), xlab="E",
beside=TRUE, col=c("lightskyblue2", "steelblue2"), main="effective/efficient",
ylab="absolute Häufigkeit")
barplot(xtabs(~data$sonstiges+data$signifikant), ylim=c(0,70), xlab="F",
beside=TRUE, col=c("lightskyblue2", "steelblue2"), main="Sonstige Begriffe",
ylab="absolute Häufigkeit")

dev.off()

```

Elektronischer Anhang

Inhaltsverzeichnis der CD

Ordner/Datei	Inhalt	Kommentar
Bachelorarbeit	Bachelorarbeit.pdf Bachelorarbeit.tex Literatur.bib figneu (Ordner)	Pdf-Version der Bachelorarbeit Latex-Code für die Bachelorarbeit Literatur-File für den Latex-Code 21 Graphiken für den Latex-Code
Informationen	Bioinformatics.txt Biometrics.txt Biostatistics.txt BMC Bioinformatics.txt	Stichpunkte für diese Artikel “ “ “
Paper	bioinformatics (Ordner) biometrics (Ordner) biostatistics (Ordner) BMC bioinformatics (Ordner)	29 Artikel aus Bioinformatics 29 Artikel aus Biometrics 29 Artikel aus Biostatistics 29 Artikel aus BMC Bioinformatics
R	R-Code.r Graphiken (Ordner)	R-Code der Auswertungen 14 erzeugte Graphiken
datensatz.xls datensatz.csv		Datensatz mit Anmerkungen und Liste der Schlüsselwörter Datensatz für R

Eidesstaatliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Bachelorarbeit selbständig und lediglich unter der Benutzung der angegebenen Hilfsmittel angefertigt habe.

München, den 18. September 2013

(Germaine Wenzler)