FRIEDRICH-LOEFFLER-INSTITUT
seit 1910
FLI
Bundesforschungsinstitut für Tiergesundheit
Federal Research Institute for Animal Health

Epidemiologische Untersuchungen in Tierpopulationen

Ein Leitfaden zur Bestimmung von Stichprobenumfängen

Franz J. Conraths, Jörn Gethmann, Jana Schulz, Susanne Amler, Andreas Fröhlich, Mario Ziller und Nicolai Denzin

Friedrich-Loeffler-Institut, Greifswald-Insel Riems

erstellt auf der Grundlage von Texten und Tabellen von Dr. R. Lorenz, Tübingen (Krankheitsüberwachung in Tierbeständen. Ein Leitfaden zur Bestimmung von Stichprobenumfängen. AID, Bonn, 1990)

3., überarbeitete Auflage, November 2020

Inhaltsverzeichnis

1. Terminologie und Definitionen 3
1.1 Epidemiologische Maßzahlen 3
1.2 Epidemiologische Häufigkeitsmaße 4
1.3 Epidemiologische Studientypen im Rahmen der Krankheitsüberwachung 6
1.3.1 Filteruntersuchung 6
1.3.2. Survey, Erhebung (syn.: Querschnittsuntersuchung) 7
1.3.3. Monitoring 7
1.3.4. Surveillance 7

2. Erhebungen zum Nachweis einer Krankheit 7
2.1 Statistischer Nachweis des Vorhandenseins oder der Abwesenheit einer Krankheit 7
2.2 Krankheits-Monitoring und Seuchen-Überwachung 23

3. Erhebungen zur Bestimmung der Prävalenz einer Krankheit 24
3.1 Schätzung der Häufigkeit kranker Tiere in einer Population 24
3.2 Genauigkeit von Häufigkeitsschätzungen 31

4. Auswahlverfahren 38
4.1 Einfache Zufallswahl *(simple random sampling)* 38
4.2 Systematische Auswahl (*systematic sampling*) 38
4.3 Geschichtete Auswahl (*stratified sampling*) 39
4.4 Klumpen-Auswahl (*cluster sampling*) 40
4.5 Mehrstufige Auswahl (multi-stage sampling) 40

5. Güte diagnostischer Tests 41
5.1 Sensitivität und Spezifität 41
5.2 Diagnostische Voraussagewerte 45
5.3 Serielles und paralleles Testen (Testsysteme) 47
5.4 Auswirkungen imperfekter Tests auf Stichprobenumfänge 49
5.5 Auswirkungen imperfekter Tests auf die Aussagekraft von Schätzwerten 49

1. Terminologie und Definitionen

„Epidemiologie ist die Wissenschaft von den Nennern und stellt als solche ein vernünftiges Gegengewicht zur klinischen Ausbildung dar, die vor allem von den Zählern beherrscht wird. Die Einführung des Nenners hatte für das medizinische Denken etwa dieselbe Bedeutung wie die Erfindung des Rades und sie war ähnlich revolutionär. Der Nenner ist die Grundlage der Rate und damit unseres Verständnisses von Verhältnissen und Prioritäten. Nenner sind zwar langweilig, aber unentbehrlich, wann und wo immer man Schlüsse über Verteilungen, Unterschiede und Dividenden zu finden versucht - sei es auf fiskalischem, sozialem oder medizinischem Gebiet“ (STEWART, 1970[1]).
Wie in vielen wissenschaftlichen Disziplinen, so kann man auch in der Epidemiologie einen gewissen Mangel an terminologischer Präzision beobachten, der Mehrdeutigkeiten zur Folge hat. Ein Grund könnte sein, dass man gelegentlich Ausdrücke aus anderen Anwendungsbereichen entlehnt, ohne ihre genaue Bedeutung zu kennen. Ein anderer Grund liegt in der Semantik: ein Wort kann in der Umgangssprache mehr als eine Bedeutung haben. Andererseits werden gelegentlich auch Worte als Synonyme benutzt, deren Bedeutungen in Wirklichkeit verschieden sind.
Ein Beispiel für vorgenannte terminologische Unklarheiten ist der Gebrauch des Wortes Rate in der Epidemiologie. Es wurde aus der Physik entlehnt, aber falsch interpretiert. Es hat (in der englischen Sprache) mehr als eine Bedeutung, und der häufigste Fehler besteht darin, dass es mit dem Terminus „Anteil“ (engl.: proportion) verwechselt wird, wobei beide Ausdrücke zu allem Überdruss auch noch unkorrekter Weise als Synonyme für „Verhältnis“ (engl. = ratio) verstanden werden.

Im Folgenden werden daher Definitionen für häufig verwendete / im Kontext dieses Leitfadens relevante epidemiologische Maßzahlen gegeben.

1.1. Epidemiologische Maßzahlen

Verhältnis

Das **Verhältnis** (engl.: ratio) ist ein Ausdruck für die Beziehung zwischen einem Zähler und einem Nenner, die beide nichts miteinander zu tun haben müssen und ganz unterschiedliche Quantitäten darstellen können. Keine von beiden muss in der anderen enthalten sein. Zwar werden beide gewöhnlich in denselben Einheiten gemessen, aber dies ist nicht wesentlich.
In der Epidemiologie wird der Begriff Verhältnis aber gewöhnlich nur für Quotienten genutzt, bei denen der Zähler nicht im Nenner enthalten ist (THRUSFIELD, 2018[2]).

<u>Beispiele:</u>

- Geschlechterverhältnis = Anzahl Männer / Anzahl Frauen
- Fetale Sterblichkeitsverhältnis = Anzahl Totgeburten / Anzahl Lebendgeburten

Index

Ein **Index** (engl.: index) ist ein Maß, das zum Vergleich oder zur Klassifikation bestimmt ist. Der Index wird aus mehreren verschiedenen Merkmalen gebildet. Häufig stellt ein Index ein

[1] Stewart, G.T. (1970). The epidemiological approach to the assessment of health. Lancet ii,115-19.
[2] Thrusfield, M. (2018). Veterinary Epidemiology. Wiley-Blackwell, 4. Auflage.

Verhältnis aus den beteiligten Merkmalen dar. So ist der Gewicht-Länge-Index (z.B. Body-Mass-Index, BMI = Körpergewicht (in kg)/(Körpergröße (in m))^2) ein Verhältnis, das als ein Maß für die Bewertung des Körpergewichts herangezogen wird.

Anteil

Ein **Anteil** (engl.: proportion) ist eine relative Häufigkeit. In diesem Fall ist der Zähler im Nenner enthalten, oder genauer: die im Zähler beschriebene Menge ist eine Teilmenge der gesamten Population.

Beispiele:

- Männeranteil = Anzahl Männer / (Anzahl Männer + Anzahl Frauen)
- Totgeburtenanteil = Anzahl Totgeburten / Anzahl Konzeptionen

Relative Häufigkeiten können unter gewissen Voraussetzungen als Schätzwerte für die Wahrscheinlichkeit des Eintretens bestimmter Ereignisse verwendet werden. So ist der empirische Männeranteil in einer bestimmten Population ein Schätzwert für die Wahrscheinlichkeit, dass ein Neugeborenes männlich sein wird.

Rate

Die **Rate** (engl.: rate) ist ein Ausdruck für die Änderung, die eine Größe je Einheit mit Bezug auf eine andere Größe erfährt. Häufig ist die Bezugsgröße die Zeit. Bei vielen biologischen Prozessen ist allerdings nicht die absolute Änderung pro Zeiteinheit von Interesse, sondern die relative Änderung pro Zeiteinheit und pro Einheit jener Organismen (oder der Population), in denen die Änderung stattfindet. Größen mit zeitlichen Bezügen spielen in der Epidemiologie eine herausragende Rolle, da sie der Darstellung der zeitlichen Dynamik eines Infektionsgeschehens dienen.

1.2. Epidemiologische Häufigkeitsmaße

Prävalenz

Die **Prävalenz** (syn.: Krankheitsbestand, engl.: prevalence) beschreibt die Häufigkeit eines Ereignisses, z. B. einer Erkrankung, in einer bestimmten Population innerhalb eines bestimmten Zeitraums. Bei der Prävalenz handelt es sich um einen Anteil (engl.: proportion), s. o., und nicht um eine Rate (s. o.), auch wenn der Begriff Prävalenz oft fälschlich den Zusatz „-rate“ erhält. Richtig wäre, wenn auch nicht erforderlich, der Zusatz „-quote“ [„Quote“ ist ein Synonym für „Anteil“], d. h. man kann von einer Prävalenzquote sprechen (KREIENBROCK et al., 2012[3]).

Punkt-Prävalenz

Die **Punkt-Prävalenz** (engl.: point prevalence) einer Krankheit ist derjenige Anteil der Individuen in einer bestimmten Population, die diese Krankheit zu einem bestimmten Zeitpunkt

[3] Kreienbrock, L., Pigeot, I. und Ahrens, W. (2012). Epidemiologische Methoden. Spektrum Akademischer Verlag, 5. Auflage.

(z. B. Stichtag) haben. In den meisten Felderhebungen ist das, was festgestellt werden soll, die Punkt-Prävalenz.

Perioden-Prävalenz

Die **Perioden-Prävalenz** (engl.: period prevalence**)** einer Krankheit ist derjenige Anteil der Individuen in einer bestimmten Population, die diese Krankheit in einem bestimmten Zeitraum haben. Sie beschreibt die Gesamtzahl aus neuen Fällen, die irgendwann innerhalb einer bestimmten Zeitperiode aufgetreten sind, und alten Fällen, die bereits vor der bestimmten Zeitperiode aufgetreten sind und während ihrer Dauer zumindest zeitweise noch existent waren. Es handelt sich somit um eine Kombination der Punkt-Prävalenz am Beginn dieser Zeitperiode und der Inzidenz (s. u.) neuer Fälle im Verlauf der Zeitperiode. Mit dem Begriff der Periodenprävalenz wird gearbeitet, weil sich Stichproben größerer Studien oft nicht an einem bestimmten Zeitpunkt, z.B. zu einem Stichtag, sondern nur in Zeitintervallen gewinnen lassen, so dass exakte Bestimmungen der Punktprävalenz dann nicht direkt aus den Studienergebnissen abgeleitet werden können. Ähnliches gilt unter angegebenen Bedingungen für die Bestimmung einer Inzidenz.

Inzidenz

Die **Inzidenz** (syn.: Neuerkrankungsrate, engl.: incidence) einer Krankheit bezeichnet die absolute Anzahl der Neuerkrankungen innerhalb eines bestimmten Zeitraumes. Die Inzidenz ist ein Maß für die Geschwindigkeit, mit der die Krankheit entsteht oder sich ausbreitet. Bei der Inzidenz handelt es sich auf Grund des Zeitbezugs um eine Rate (s. o.). Dabei bleibt die Größe der Population zunächst unberücksichtigt.

Relative Inzidenz

Bei der **Relativen Inzidenz** (engl.: relative incidence) erfolgt dagegen auch ein Bezug zur Größe der unterliegenden Population. Sie wird gewöhnlich als Anteil oder Prozentsatz der mittleren Anzahl aller während des betreffenden Zeitraumes in der Gesamtpopulation vorhandenen Tiere ausgedrückt (und ist damit eine Approximation der Inzidenzrate; s. u.) oder aber auch als absolute Anzahl neuer Fälle je z. B. 1000 Tiere je Zeiteinheit. Es hat sich eingebürgert, auch diese Maßzahlen als „Inzidenz“ zu bezeichnen; meistens ergibt sich aus der Art der Darstellung von selbst, ob die relative Inzidenz gemeint ist oder nicht.

Kumulative Inzidenz

Bei der **Kumulativen Inzidenz** (engl.: cumulative incidence) wird die Anzahl der Neuerkrankungen je Zeitraum auf die Größe der Population zu Beginn des Zeitraums bezogen. Die Kumulative Inzidenz entspricht daher wiederum einem Anteil (engl.: proportion) und wird auch als **Risiko** (innerhalb des Zeitraums zu erkranken) bezeichnet. Da sich die kumulative Inzidenz auf die Population zu Beginn des Betrachtungszeitraums bezieht, werden Zugänge und Abgänge aus der Population nicht berücksichtigt.

Inzidenzrate

Bei der **Inzidenzrate** (engl.: incidence rate, auch „Inzidenzdichte“ genannt) werden auch Zu- und Abgänge berücksichtigt, indem der Nenner als Individuumzeit (Tierzeit / Personenzeit) unter Risiko definiert wird. Da sich die Inzidenzrate auf Zeiteinheiten bezieht, handelt es sich um eine echte Rate, die z. B. als Erkrankungen je Individuumwochen / -jahre etc. unter Risiko

ausgedrückt werden kann. Da die tatsächliche Zeit unter Risiko für die einzelnen Individuen einer Population oft schwer zu bestimmen ist, wird der Nenner der Inzidenzrate oft als mittlere Populationsgröße während des Betrachtungszeitraums (Summe von Anfangs- und Endpopulation dividiert durch zwei) approximiert (s. o. „Relative Inzidenz“).

Attackrate

Die Inzidenzrate für ein kurzes Zeitintervall wird auch als **Attackrate** (engl.: attack rate) bezeichnet.

Morbidität

Die **Morbidität** (engl.: morbidity) quantifiziert das Auftreten einer Erkrankung in einer Population und ist ein Überbegriff für die oben vorgestellten Häufigkeitsmaße der Prävalenz und Inzidenz mit ihren Variationen, wobei das Ereignis von Interesse die Erkrankung ist. Häufig wird aber, verkürzt gegenüber o.g. Definition, unter Morbidität der Anteil erkrankter Individuen an der Population (Prävalenz) verstanden.

Mortalität

Die Häufigkeitsmaße der **Mortalität** (syn.: Sterblichkeit, engl.: mortality) entsprechen denen der Inzidenz und ihrer Variationen, wie oben vorgestellt, wobei das Ereignis von Interesse hier nicht die Erkrankung, sondern der Todesfall, ist. Die Kumulative Mortalität, als Beispiel zur Illustration der Analogie gewählt, beschreibt daher die Anzahl der Individuen, die während eines bestimmten Zeitraums an einer Erkrankung sterben, dividiert durch die Gesamtzahl der Individuen zu Beginn des Zeitraums.

Letalität

Die **Letalität** (syn.: Tödlichkeit, engl.: case fatality) beschreibt die Anzahl der Todesfälle (Zähler) innerhalb eines bestimmten Zeitraums nicht bezogen auf die Gesamtzahl der Individuen (im Gegensatz zur Kumulativen Mortalität), sondern auf die Gesamtzahl der erkrankten Individuen (Nenner). Es handelt sich um einen Anteil (engl.: proportion). Man kann die Letalität auch als die Wahrscheinlichkeit interpretieren, an einer bestimmten Krankheit zu sterben, wenn man an ihr erkrankt ist.

1.2. Epidemiologische Studientypen im Rahmen der Krankheitsüberwachung

1.2.1. Filteruntersuchung

Eine **Filteruntersuchung** (syn.: Siebtest, Übersichtsuntersuchung, Grobfilter, Rasteruntersuchung, Vorauswahl; engl.: screening) ist die Untersuchung eines ausgedehnten Querschnitts einer Population mit dem Ziel, neue Fälle einer Krankheit zu entdecken. Ein Screening ist dabei in der Regel eine einmalige, zeitlich begrenzte oder auf einen Stichtag bezogene Studie, während **Monitoring** und **Surveillance** (siehe unten) auf längere Sicht angelegt sind. Im Vordergrund steht nicht so sehr die endgültige Diagnose im Einzelfall. Es kommt vielmehr darauf an, mit großer Sicherheit alle wirklichen Krankheitsfälle auszufiltern; dabei wird ein gewisser Anteil falsch-positiver Diagnosen in Kauf genommen. Ein diagnostisches Verfahren, das bei einer Filteruntersuchung verwendet wird, heißt Screening-

Test. Ein solcher Test ist im Allgemeinen kostengünstig, einfach zu handhaben und erlaubt einen Einsatz in großem Umfang.

1.2.2. Survey, Erhebung

Ein **Survey** / eine **Erhebung** (syn.: Querschnittsuntersuchung, engl.: survey) ist die stichprobenartige Untersuchung einer Population mit dem Ziel, die Prävalenz einer Krankheit - bezogen auf die Gesamtpopulation - zu schätzen.

1.2.3. Monitoring

Das **Monitoring** (engl.: monitoring) einer Tierpopulation ist die kontinuierliche Sammlung von Daten über Gesundheits- und/oder Umweltparameter. Kennzeichnend ist dabei, dass die Daten sich auf klar umrissene Kriterien stützen und geeignet sind, Änderungen der Prävalenz einer Krankheit frühzeitig anzuzeigen, die ihrerseits ein Hinweis auf eine Änderung der Inzidenz sein können. Intervention ist kein primäres Ziel des Monitorings. Beispielsweise wurde im Zusammenhang mit der Blauzungenkrankheit ein entomologisches Monitoring durchgeführt, um die Populationsdynamik von Gnitzen der Gattung *Culicoides* zu bestimmen, welche Vektoren dieser Tierseuche sind. Es ging aber nicht darum, Maßnahmen gegen Gnitzen zu ergreifen oder gegen sie ergriffene Maßnahmen zu überwachen.

1.2.4. Surveillance

Unter **Surveillance** (engl.: surveillance) versteht man die aktive Seuchenüberwachung als laufende Kontrolle einer Tierpopulation mit dem Ziel, Änderungen im Gesundheitsstatus der Population frühzeitig zu erkennen und durch konkrete Interventionen unmittelbar zu steuern. Die Surveillance stützt sich gewöhnlich auf Daten, die durch laufende Beobachtung gewonnen werden, wie z.B. bei Schlachthof- oder Milchuntersuchungen im Zusammenhang mit der Brucellose- oder BHV-1-Bekämpfung. Ein anderes Beispiel sind regelmäßige serologische Bestandsuntersuchungen. Jedoch werden auch alle möglichen anderen Informationen herangezogen, sofern sie für die Bekämpfung einer Seuche bedeutsam sein können, z.B. einmalige gezielte Querschnittsuntersuchung in der Population. Hierzu gehört auch die Analyse von Daten aus anderen relevanten Sachgebieten, wie die Populationsdynamik von Vektoren und der Einfluss von Klima oder Haltungsbedingungen. Schließlich sind Daten über Wirksamkeit und Nebenwirkungen von Überwachungsmaßnahmen selbst (z.B. Impfungen, systematische Rückverfolgsuntersuchungen, Sperrmaßnahmen usw.) von wesentlichem Interesse bei der Seuchenüberwachung.
Die Surveillance ist nicht mit dem Monitoring zu verwechseln. Eine Verwechselung liegt nahe, weil im Deutschen oft für beide Begriffe das Wort „Überwachung“ verwendet wird. Während Surveillance stets im Zusammenhang mit der Steuerung von Bekämpfungsmaßnahmen steht, ist Monitoring unabhängig davon, ob und gegebenenfalls welche Maßnahmen ergriffen werden.

2. Erhebungen zum Nachweis einer Krankheit

2.1. Statistischer Nachweis des Vorhandenseins oder der Abwesenheit einer Krankheit

Der sichere Beweis, dass eine Krankheit in einer Population nicht vorkommt (Defizite der Diagnostik außen vor) - insbesondere dann, wenn normalerweise nur wenige Tiere betroffen sind - erfordert die Untersuchung **aller** empfänglichen Tiere der Population („Zensus“), schließt also die Untersuchung mit Hilfe einer Stichprobe aus. Auch für den sicheren Ausschluss des

Vorkommens von Krankheiten, die in dem betrachteten Gebiet nicht heimisch sind (bei denen die Prävalenz also als Null oder extrem niedrig anzunehmen ist) oder die nur noch selten vorkommen, eignet sich eine Untersuchung mit Hilfe von Stichproben nicht.
Als Beispiel für die Notwendigkeit eines Zensus kann die Untersuchung aller neugeborenen Kälber auf Bovine Virusdiarrhoe (BVD) zur Erkennung persistent infizierter (PI) Tiere genannt werden. In der Endphase der Krankheitsbekämpfung ist die Prävalenz der PI-Tiere sehr niedrig, die Konsequenzen eines Übersehens der Tiere dieses Status sind aber erheblich.

Eine Zensusuntersuchung ist teuer und erfordert viele Ressourcen. Aus diesem Grund wird oft eine Stichprobenuntersuchung durchgeführt. Die Stichprobe ist streng genommen zufällig zu ziehen. Das bedeutet, dass jedes Tier in der betrachteten Population[4] die gleiche Wahrscheinlichkeit haben muss, in die Stichprobe gewählt zu werden (s. auch Seite 41, Auswahlverfahren). Dazu wird zunächst festgelegt (meist auch rechtlich), mit welcher Zuverlässigkeit („Statistische Sicherheit“) die Stichprobenuntersuchung das Erkennen einer bestimmten Prävalenz der Erkrankung in einer Population ermöglichen soll, d. h. wenn diese Prävalenz erreicht ist, findet sich mindestens ein positives Tier in der Stichprobe und die Population wird als infiziert erkannt. Die Prävalenz wird dabei i. d. R. so gewählt, dass sie einem Anteil der Erkrankten in einer Population entspricht, der relativ schnell erreicht wird, sodass eine Infektion durch die Stichprobenuntersuchung frühzeitig erkannt werden kann. Der Umfang einer Stichprobe an Tieren, die getestet werden müssen, um mit einer bestimmten statistischen Sicherheit nachzuweisen, dass eine Krankheit in einer bestimmten Population vorhanden ist, hängt ab

a) von der Größe der betreffenden Population,
b) von der mutmaßlichen Prävalenz der Krankheit in der Population,
c) von der geforderten Zuverlässigkeit der Aussage,
d) von den Güteparametern der Diagnostik.

a) Die Größe der Population, aus der die Stichprobe zu entnehmen ist

Hat eine fortlaufende Untersuchung von Tieren aus einer Population, in der eine Krankheit (Infektion) vorkommt, ausschließlich negative Testergebnisse gebracht, so ist leicht nachvollziehbar, dass die Wahrscheinlichkeit dafür, dass im Folgenden untersuchte Tiere positiv sein werden, merklich zunimmt, sofern die Stichprobe einen großen Anteil an der Gesamtpopulation umfasst. Dieser Sachverhalt gilt ebenso für unendlich große Populationen.

b) Die mutmaßliche Prävalenz der Krankheit in der Population

Falls eine Krankheit in einer Herde oder in einem Bestand vorkommt, ist in den meisten Fällen zu erwarten, dass nicht nur ein einzelnes Tier, sondern ein gewisser Anteil erkrankt (infiziert) ist. Anhaltspunkte über die mutmaßliche Höhe dieses Anteils ergeben sich aus der Epidemiologie der betreffenden Krankheit und aus Erfahrungen mit anderen Herden und Beständen. Je höher die Prävalenz, desto geringer ist der zum Auffinden von erkrankten (infizierten) Tieren erforderliche Stichprobenumfang.
Der Stichprobenumfang kann daher so gewählt werden, dass es im Falle einer negativen Stichprobe unwahrscheinlich ist, dass irgendein Tier in der Herde krank (infiziert) ist.

[4] “Population“ ist als Synonym für das in der Statistik gebräuchliche Wort „Grundgesamtheit“ zu verstehen. Es umfasst begrifflich „Bestand“ oder „Herde“ und bezeichnet diejenige Gesamtheit, aus der gegebenenfalls eine Stichprobe entnommen wird und für welche die Stichprobe repräsentativ sein soll. Es kann sich z.B. auch um eine Gesamtheit aus mehreren Tierbeständen handeln, etwa um die Population aller Ziegen in einem Landkreis.

c) Die Zuverlässigkeit der Aussage

Bei Anwendung eines Stichprobenverfahrens akzeptieren wir (überwiegend aus ökonomischen Gründen), dass auf Grund des Befundes eine absolut sichere Aussage nicht möglich ist. Je größer die Stichprobe, desto größer ist die „statistische Sicherheit".
Die bei Übersichtsuntersuchungen üblicherweise verwendeten statistischen Sicherheiten sind 95 % und gelegentlich 99 %. Die genaue Bedeutung dieser Zahlen soll an einem Beispiel verdeutlicht werden.
Wir nehmen an, dass 10 % aller Tiere, die in den Beständen eines Landes gehalten werden, Antikörper gegen einen bestimmten Infektionserreger besitzen. Die Bestände sollen mittels einer Stichprobe untersucht werden, um herauszufinden, ob sich in den Beständen seropositive Tiere befinden. Der Stichprobenumfang wurde so gewählt, dass eine Wahrscheinlichkeit von 95 % dafür besteht, in der zufällig ausgewählten Stichprobe von Beständen Antikörperträger zu finden. Das bedeutet, dass (im Durchschnitt) 5 von 100 infizierten Beständen (von denen jeder 10 % Antikörperträger enthält) mit diesem Stichprobenverfahren *nicht* als infiziert erkannt werden. Die Formulierung „im Durchschnitt" verweist auf den statistischen Charakter der Argumentation: Man muss sich eine sehr große Anzahl identischer Tierbestände (alle mit 10 % Antikörperträgern) vorstellen; aus jedem dieser Bestände wird eine unabhängige Stichprobe von dem zuvor festgelegten einheitlichen Umfang gezogen. Ein Teil der Stichproben enthält einen oder mehrere Antikörperträger und erkennt somit die Bestände zutreffend als infiziert. Die übrigen Stichproben enthalten zufällig keine Antikörperträger der jeweiligen Tierbestände und führen zu dem (falschen) Schluss, dass die betreffenden Bestände nicht infiziert sind. Die Zuverlässigkeit des Ansatzes besteht darin, dass der Stichprobenumfang so gewählt wurde, dass mindestens 95 % aller dieser Stichproben (oder im „Durchschnitt" 95 von je 100) tatsächlich Antikörperträger enthalten.
Die Beziehung zwischen der Zuverlässigkeit einer Aussage und dem Anteil von Antikörperträgern in der Population lässt sich bei feststehendem Stichprobenumfang wie folgt beschreiben: Für Bestände, die mehr als 10 % Antikörperträger enthalten, besteht eine geringere Irrtumswahrscheinlichkeit als 5 % dafür, dass sie anhand einer Stichprobe (des oben erwähnten Umfangs) nicht als infiziert erkannt werden. Umgekehrt ist bei Beständen mit weniger als 10 % Antikörperträgern eine höhere Fehlerquote (Irrtumswahrscheinlichkeit größer als 5 %, was einer geringeren Zuverlässigkeit entspricht) zu erwarten.

d) Die Güte der verwendeten Diagnoseverfahren

Auch die Güte der verwendeten Diagnoseverfahren beeinflusst die Genauigkeit des statistischen Nachweises des Vorhandenseins einer Krankheit. Ein Diagnoseverfahren, dessen Sensitivität nicht perfekt ist, führt dazu, dass nicht alle Fälle einer Krankheit erkannt werden, so dass das wahre Ausmaß der Verbreitung der Krankheit unterschätzt wird. Ein Diagnoseverfahren, dessen Spezifität nicht perfekt ist, führt dazu, dass eine Krankheit bei Tieren festgestellt wird, die in Wahrheit nicht an der Krankheit leiden. Somit wird das wahre Ausmaß der Verbreitung der Krankheit überschätzt. Wenn die Sensitivität und Spezifität eines Diagnoseverfahrens bekannt sind, können Stichprobenumfänge so korrigiert werden, dass die gewünschte Zuverlässigkeit des statistischen Nachweises des Vorhandenseins einer Krankheit wiederhergestellt wird oder der Schluss gezogen werden muss, dass die gewünschte Zuverlässigkeit nicht erreicht werden kann.

In **Tabelle 1** sind - für verschiedene statistische Sicherheiten - Stichprobenumfänge zusammengestellt, um Bestände mit unterschiedlichen Mindestdurchseuchungsgraden als infiziert zu erkennen. Die Tabelle gibt für perfekte Diagnoseverfahren eine obere Schranke für (notwendige) Mindeststichprobenumfänge an. Hier werden also nicht die minimalen

(notwendigen) Stichprobenumfänge ausgewiesen, sondern nur Näherungen (Approximationsformel, s. u.) derselben, die im Allgemeinen größer als die notwendigen Stichprobenumfänge sind. Dies ist aber nur in wenigen Positionen der Tabelle der Fall und ändert den Wert um nur eine Stichprobeneinheit, d. h. um ein Tier, das mehr beprobt werden muss. Damit wird das jeweils angegebene Sicherheitsniveau stets eingehalten.

Kommen imperfekte Diagnoseverfahren zur Anwendung, so sind in der Regel größere Stichprobenumfänge notwendig. Im Gegensatz zu der hier näher betrachteten perfekten Diagnostik sollte eine gewisse maximale Anzahl positiver Diagnoseresultate als Grenzwert ins Kalkül gezogen werden, wenn das Diagnoseverfahren imperfekt ist. Wendet man beispielsweise eine bezüglich der Spezifität imperfekte Diagnostik auf eine völlig gesunde Population an, muss mit falsch-positiven Resultaten gerechnet werden. Umgekehrt gilt dies auch für eine bezüglich der Sensitivität imperfekte Diagnostik, die in einer Population, in der die diagnostizierte Krankheit vorkommt, falsch-negative Ergebnisse liefern kann. In der Realität kommen fast keine perfekten Diagnoseverfahren vor. Sie sind fast immer bezüglich Sensitivität und Spezifität imperfekt, allerdings oft in unterschiedlichem Maße.

Die Interpretation ist für die perfekte wie für die imperfekte Diagnostik vergleichbar: Überschreitet die Anzahl der positiven Diagnosen in der Stichprobe eine gewisse, maximal zulässige Anzahl (für den perfekten Test ist letztere Null), so ist die Krankheit in der Population nachgewiesen.

Tabelle 2 zeigt den Einfluss unterschiedlicher Stichprobenumfänge auf die Wahrscheinlichkeit, in einer Population mit nur wenigen positiven Tieren mindestens eines dieser positiven Tiere zu finden. Die angegebenen Wahrscheinlichkeiten sind exakt.

In **Tabelle 3** wird hingegen der Einfluss des Stichprobenumfangs auf die Wahrscheinlichkeit aufgezeigt, in einer unendlich großen Population mit einem bestimmten Durchseuchungsgrad kein positives Tier zu finden. Auch hier handelt es sich um exakte Wahrscheinlichkeiten.

Tabelle 1: Nachweis des Vorhandenseins einer Krankheit

Die Tabellen 1 (a), 1 (b) und 1 (c) gelten für Werte der statistischen Mindestsicherheit ε von 90 %, 95 % bzw. 99 %.

Diese Tabellen können auf **zwei Arten** verwendet werden:

<u>Verwendungsmöglichkeit 1:</u>

Zur Bestimmung derjenigen Mindestanzahl von Tieren in einer Stichprobe n, die untersucht werden müssen, um mit einer vorgegebenen Wahrscheinlichkeit (bzw. Sicherheit) ε mindestens ein krankes Tier zu erfassen ($P(X \geq 1)$), wenn die Anzahl kranker Tiere d in der Population N einem bestimmten Prozentsatz entspricht ($p_0 = d/N*100$). Dabei soll folgende Hypothese (als so genannte alternative Hypothese, H_1) geprüft werden: „Die Prävalenz der Krankheit in der zu untersuchenden Population liegt unterhalb der gewählten Prävalenzschwelle d/N*100“, d. h. H_1: $p < p_0$. Die Prüfung der alternativen Hypothese erfolgt dabei gegen die so genannte Nullhypothese, H_0: $p \geq p_0$, die davon ausgeht, dass die Prävalenz größer oder gleich der gewählten Prävalenzschwelle ist.
Die gewählte oder (gegebenenfalls auch rechtlich) geforderte Prävalenzschwelle (häufig als Designprävalenz bezeichnet) stellt eine theoretische Maßzahl dar, die insbesondere für endliche

Populationsgrößen nicht existieren muss, d. h. für eine bestimmte Populationsgröße gibt es gegebenenfalls keine ganzzahlige Anzahl an infizierten/erkrankten Tieren (d), die rechnerisch zu der vorgegebenen Designprävalenz führt. Daher sei auf die Möglichkeit hingewiesen, mit Stichprobenumfängen zu arbeiten, die von der nächst größeren real existierenden Prävalenz abgeleitet werden. Solche Stichprobenumfänge sind möglicherweise von geringerem Umfang. Streng genommen werden bei einem solchen Vorgehen aber etwaige Vorgaben bezüglich der Designprävalenz nicht erreicht. Nutzt man aber die nächstniedrigere real existierende Prävalenz, so ergeben sich möglicherweise Stichprobenumfänge von unnötiger Größe. Aus diesem Grund wurden für die approximativen Stichprobenberechnungen in Tabelle 1 Dezimalzahlen (mit Nachkommastellen) für die Anzahl der kranken Tiere (d) genutzt, sofern sie sich aus den Vorgaben (Populationsgröße, Designprävalenz) ergeben. Ergibt sich für die errechnete Stichprobengröße n eine Dezimalzahl, wird diese auf die nächste ganze Zahl aufgerundet.

Die Tabelle wird wie folgt genutzt:

- Wähle die Tabelle mit der gewünschten Mindestsicherheit (90%, 95% oder 99%).
- Wähle die Spalte mit dem größten tabellierten Prozentsatz (= Prävalenzschwelle d/N*100) infizierter Tiere, der kleiner oder gleich der geforderten Nachweisgrenze ist.
- Suche in dieser Spalte die zur Populationsgröße N gehörende Zeile. Ist der gesuchte Wert von N nicht enthalten, wähle den nächst größeren Wert.
- Das Minimum aus der Zahl in der betreffenden Spalte und Zeile und der Populationsgröße ist der gesuchte Mindeststichprobenumfang n.
- Ziehe die Proben durch zufällige Auswahl der Probanden.
- Werden alle Proben der Stichprobe mit negativem Ergebnis untersucht, so kann die Nullhypothese H_0 verworfen werden, dass die Prävalenz in der Population größer oder gleich der gewählten Prävalenzschwelle ist. Ist dagegen mindestens eine Probe positiv, so kann die formulierte Hypothese H_0 nicht abgelehnt werden.

<u>Verwendungsmöglichkeit 2:</u>

Angenommen, es wurde ein bestimmter prozentualer Anteil von n/N*100 der Tiere einer Population der Größe N untersucht und in dieser Stichprobe wurde <u>kein krankes Tier</u> gefunden. Dann kann mit der angegebenen Sicherheit eine maximale Anzahl der Tiere in einer Population abgeleitet werden, die dennoch erkrankt und übersehen worden sein könnte.

Die Tabelle wird wie folgt genutzt:

- Wähle die Tabelle mit der gewünschten Sicherheit.
- Wähle die Spalte mit dem Prozentsatz untersuchter Tiere.
- Suche die zur Populationsgröße N gehörende Zeile.
- Die Zahl in der betreffenden Spalte und Zeile entspricht der geschätzten maximalen Anzahl erkrankter Tiere in der untersuchten Population, die übersehen wurde.

<u>Umgang mit anderen, in den Tabellen nicht enthaltenen Werten:</u>

Für Werte von N, d/N, ε, die in den Tabellen 1 nicht enthalten sind, können die gesuchten Größen entweder durch Interpolation oder mit Hilfe der unter (1) und (2) aufgeführten Formel approximativ bestimmt werden.

Es bedeuten…

N die Populationsgröße,
d die Anzahl der kranken Tiere in der Population,
n der Stichprobenumfang,
ε die gewünschte statistische Sicherheit, d.h., die Wahrscheinlichkeit, in einer Stichprobe mindestens ein krankes Tier zu finden, sobald der zuvor postulierte Anteil von Erkrankten an der Population d/N in der Realität überschritten wird. Bei einer Mindestsicherheit von z.B. 90 % ist ε = 0,9.

Die in den Tabellen 1 (a-c) gelisteten Näherungswerte ergeben sich aus den folgenden Ungleichungen:

Für Anwendung 1[5]:

$$n \geq \left[1 - (1 - \varepsilon)^{\frac{1}{d}}\right] \times \left[N - \frac{d-1}{2}\right]$$

Ergibt sich für n eine Dezimalzahl, so wird auf die nächst größere ganze Zahl aufgerundet.

Für Anwendung 2[5]:

$$d \geq \left[1 - (1 - \varepsilon)^{\frac{1}{n}}\right] \times \left[N - \frac{n-1}{2}\right]$$

Ergibt sich für n eine Dezimalzahl, so wird auf die nächst größere ganze Zahl aufgerundet.

Die Berechnung der exakten Werte der Tabellen 1 (a-c) wird in den Ausführungen zur Herleitung im Anschluss an die Tabellen erläutert.

[5] In Anlehnung an Cannon und Roe (1982). Livestock Disease Surveys: A field manual for veterinarians. Australian Government Publishing Service, Canberra, Australien.

Tabelle 1 (a) **90 %**

(1) Notwendiger Stichprobenumfang zum Nachweis einer Krankheit
(2) Obere Grenze für die Anzahl kranker Tiere in der Population

	(1) Prozentsatz kranker Tiere in der Population (d/N) oder (2) Prozentsatz ausgewählter und nicht erkrankter Tiere (n/N)												
N	50 %	40 %	30 %	25 %	20 %	15 %	10 %	5 %	2 %	1 %	0,5 %	0,2 %	0,1 %
10	3	4	5	6	7	8	9	10	10	10	10	10	10
20	4	5	6	7	9	11	14	18	20	20	20	20	20
30	4	5	6	8	9	12	16	24	30	30	30	30	30
40	4	5	7	8	10	12	17	28	38	40	40	40	40
50	4	5	7	8	10	13	18	30	45	50	50	50	50
60	4	5	7	8	10	13	19	32	52	59	60	60	60
70	4	5	7	8	10	13	19	34	57	68	70	70	70
80	4	5	7	8	10	14	20	35	61	76	80	80	80
90	4	5	7	8	10	14	20	36	65	84	90	90	90
100	4	5	7	8	10	14	20	37	69	90	100	100	100
120	4	5	7	8	10	14	20	38	74	103	118	120	120
140	4	5	7	8	10	14	21	39	78	113	135	140	140
160	4	5	7	8	11	14	21	40	82	122	152	160	160
180	4	5	7	8	11	14	21	40	85	130	167	180	180
200	4	5	7	8	11	14	21	41	87	137	180	200	200
250	4	5	7	8	11	14	21	42	92	151	211	248	250
300	4	5	7	8	11	14	22	42	95	161	236	294	300
350	4	5	7	8	11	14	22	43	98	169	256	338	350
400	4	5	7	8	11	14	22	43	100	175	274	378	400
450	4	5	7	8	11	14	22	43	101	180	288	416	448
500	4	5	7	8	11	14	22	43	102	184	301	450	496
600	4	5	7	9	11	15	22	44	104	191	321	512	588
700	4	5	7	9	11	15	22	44	106	196	337	565	675
800	4	5	7	9	11	15	22	44	107	200	350	611	756
900	4	5	7	9	11	15	22	44	108	203	360	650	831
1000	4	5	7	9	11	15	22	44	108	205	369	684	900
1200	4	5	7	9	11	15	22	45	109	209	382	740	1024
1400	4	5	7	9	11	15	22	45	110	212	392	785	1130
1600	4	5	7	9	11	15	22	45	111	214	400	821	1221
1800	4	5	7	9	11	15	22	45	111	216	406	850	1299
2000	4	5	7	9	11	15	22	45	111	217	411	875	1368
3000	4	5	7	9	11	15	22	45	112	221	426	956	1607
4000	4	5	7	9	11	15	22	45	113	223	434	1000	1750
5000	4	5	7	9	11	15	22	45	113	224	439	1028	1845
6000	4	5	7	9	11	15	22	45	113	225	443	1047	1912
7000	4	5	7	9	11	15	22	45	114	226	445	1061	1962
8000	4	5	7	9	11	15	22	45	114	226	447	1072	2000
9000	4	5	7	9	11	15	22	45	114	227	448	1080	2031
10000	4	5	7	9	11	15	22	45	114	227	449	1087	2056
∞	4	5	7	9	11	15	22	45	114	230	460	1151	2302

Aus der Tabelle kann abgelesen werden:

(1) der Mindeststichprobenumfang n, der erforderlich ist, um mit 90 % Sicherheit mindestens ein krankes Tier zu erfassen, wenn der Anteil kranker Tiere in der Population mindestens dem gewählten Prozentsatz entspricht.

(2) eine obere Schranke für die maximale Anzahl kranker Tiere in der Population mit der angegebenen Mindestsicherheit (90 %) unter der Voraussetzung, dass ein bestimmter Anteil der Population untersucht und durchweg für gesund befunden wurde.

Beispiele:

(1) Der vermutete Anteil kranker Tiere in der Population sei größer oder gleich 2 %; die Population enthalte 480 Tiere. Man wähle N = 500. Aus der Tabelle ergibt sich ein Mindeststichprobenumfang von n = 102. Dieser ist erforderlich, um mit mindestens 90 % Sicherheit mindestens ein krankes Tier zu erfassen. Ist unter den 102 untersuchten Tieren kein krankes Tier, kann geschlossen werden, dass die Prävalenz in der Population mit einer Mindestsicherheit von 90 % kleiner als 2 % ist.

(2) Aus einer Population mit N = 1000 Tieren wurden 10 % untersucht; alle Tiere waren gesund. Aus der Tabelle ergibt sich, dass mit einer Wahrscheinlichkeit von 90 % behauptet werden kann, dass die Anzahl kranker Tiere in der Population nicht größer als 22 ist.

Tabelle 1 (b) **95 %**

(1) Notwendiger Stichprobenumfang zum Nachweis einer Krankheit
(2) Obere Schranke für die Anzahl kranker Tiere in der Population

	(1) Prozentsatz kranker Tiere in der Population (d/N) oder (2) Prozentsatz ausgewählter und nicht erkrankter Tiere (n/N)												
N	50 %	40 %	30 %	25 %	20 %	15 %	10 %	5 %	2 %	1 %	0,5 %	0,2 %	0,1 %
10	4	5	6	7	8	9	10	10	10	10	10	10	10
20	5	6	7	9	10	13	16	19	20	20	20	20	20
30	5	6	8	9	11	14	19	26	30	30	30	30	30
40	5	6	8	10	12	15	21	31	40	40	40	40	40
50	5	6	8	10	12	16	22	35	48	50	50	50	50
60	5	6	8	10	13	16	23	38	55	60	60	60	60
70	5	6	8	10	13	17	24	40	62	70	70	70	70
80	5	6	9	10	13	17	24	42	68	79	80	80	80
90	5	6	9	10	13	17	25	43	73	87	90	90	90
100	5	6	9	10	13	17	25	45	78	95	100	100	100
120	5	6	9	11	13	18	26	47	86	111	120	120	120
140	5	6	9	11	13	18	26	48	92	124	139	140	140
160	5	6	9	11	13	18	27	49	97	136	157	160	160
180	5	6	9	11	13	18	27	50	101	146	174	180	180
200	5	6	9	11	14	18	27	51	105	155	190	200	200
250	5	6	9	11	14	18	27	53	112	175	228	250	250
300	5	6	9	11	14	18	28	54	117	189	260	299	300
350	5	6	9	11	14	18	28	54	121	201	287	346	350
400	5	6	9	11	14	19	28	55	124	211	311	391	400
450	5	6	9	11	14	19	28	55	127	218	331	434	450
500	5	6	9	11	14	19	28	56	129	225	349	475	499
600	5	6	9	11	14	19	28	56	132	235	379	551	597
700	5	6	9	11	14	19	28	57	134	243	402	618	691
800	5	6	9	11	14	19	28	57	136	249	421	677	782
900	5	6	9	11	14	19	29	57	137	254	437	730	868
1000	5	6	9	11	14	19	29	57	138	258	450	777	950
1200	5	6	9	11	14	19	29	58	140	264	471	856	1102
1400	5	6	9	11	14	19	29	58	141	269	487	920	1236
1600	5	6	9	11	14	19	29	58	142	272	499	972	1354
1800	5	6	9	11	14	19	29	58	143	275	509	1017	1459
2000	5	6	9	11	14	19	29	58	143	277	517	1054	1553
3000	5	6	9	11	14	19	29	58	145	284	542	1179	1895
4000	5	6	9	11	14	19	29	58	146	288	556	1249	2108
5000	5	6	9	11	14	19	29	59	147	290	564	1294	2253
6000	5	6	9	11	14	19	29	59	147	291	569	1325	2358
7000	5	6	9	11	14	19	29	59	147	292	573	1348	2437
8000	5	6	9	11	14	19	29	59	147	293	576	1365	2498
9000	5	6	9	11	14	19	29	59	148	294	579	1379	2548
10000	5	6	9	11	14	19	29	59	148	294	581	1390	2588
∞	5	6	9	11	14	19	29	59	149	299	598	1497	2995

Aus der Tabelle kann abgelesen werden:

(1) der Mindeststichprobenumfang n, der erforderlich ist, um mit 95 % Sicherheit mindestens ein krankes Tier zu erfassen, wenn der Anteil kranker Tiere in der Population mindestens dem gewählten Prozentsatz entspricht.

(2) eine obere Schranke für die maximale Anzahl kranker Tiere in der Population mit der angegebenen Mindestsicherheit (95 %) unter der Voraussetzung, dass ein bestimmter Anteil der Population untersucht und durchweg für gesund befunden wurde.

Beispiele:

(1) Der vermutete Anteil kranker Tiere in der Population sei größer oder gleich 2 %; die Population enthalte 480 Tiere. Man wähle N = 500. Aus der Tabelle ergibt sich ein Mindeststichprobenumfang von n = 129. Dieser ist erforderlich, um mit mindestens 95 % Sicherheit mindestens ein krankes Tier zu erfassen. Ist unter den 129 untersuchten Tieren kein krankes Tier, kann geschlossen werden, dass die Prävalenz in der Population mit einer Mindestsicherheit von 95 % kleiner als 2 % ist.

(2) Aus einer Population mit N = 1000 Tieren wurden 10 % untersucht; alle Tiere waren gesund. Aus der Tabelle ergibt sich, dass mit einer Wahrscheinlichkeit von 95 % behauptet werden kann, dass die Anzahl kranker Tiere in der Population nicht größer als 29 ist.

Tabelle 1 (c) **99 %**

(1) Notwendiger Stichprobenumfang zum Nachweis einer Krankheit
(2) Obere Grenze für die Anzahl kranker Tiere in der Population

	(1) Prozentsatz kranker Tiere in der Population (d/N) oder (2) Prozentsatz ausgewählter und nicht erkrankter Tiere (n/N)												
N	50 %	40 %	30 %	25 %	20 %	15 %	10 %	5 %	2 %	1 %	0,5 %	0,2 %	0,1 %
10	5	6	8	8	9	10	10	10	10	10	10	10	10
20	6	8	10	11	13	15	18	20	20	20	20	20	20
30	7	8	11	13	15	19	23	29	30	30	30	30	30
40	7	9	11	14	16	21	27	36	40	40	40	40	40
50	7	9	12	14	17	22	29	42	50	50	50	50	50
60	7	9	12	15	18	23	31	47	59	60	60	60	60
70	7	9	12	15	18	24	33	51	68	70	70	70	70
80	7	9	12	15	19	24	34	54	76	80	80	80	80
90	7	9	13	15	19	25	35	57	83	90	90	90	90
100	7	9	13	15	19	25	36	59	90	99	100	100	100
120	7	9	13	16	19	26	37	63	102	118	120	120	120
140	7	9	13	16	20	26	38	67	113	135	140	140	140
160	7	9	13	16	20	26	39	69	122	151	160	160	160
180	7	9	13	16	20	27	39	71	129	166	179	180	180
200	7	9	13	16	20	27	40	73	136	180	198	200	200
250	7	10	13	16	20	27	41	76	150	210	244	250	250
300	7	10	13	16	20	28	41	78	160	235	286	300	300
350	7	10	13	16	21	28	42	80	168	256	325	350	350
400	7	10	13	16	21	28	42	81	174	273	360	399	400
450	7	10	13	16	21	28	42	82	179	288	392	448	450
500	7	10	13	16	21	28	42	83	183	300	421	495	500
600	7	10	13	16	21	28	43	84	190	321	470	587	600
700	7	10	13	16	21	28	43	85	195	336	512	674	700
800	7	10	13	16	21	28	43	85	199	349	546	755	798
900	7	10	13	16	21	28	43	86	202	359	576	830	895
1000	7	10	13	16	21	28	43	86	204	368	601	900	990
1200	7	10	13	17	21	29	43	87	208	381	642	1024	1175
1400	7	10	13	17	21	29	44	87	211	391	674	1129	1348
1600	7	10	13	17	21	29	44	88	213	399	699	1220	1510
1800	7	10	13	17	21	29	44	88	215	405	720	1299	1661
2000	7	10	14	17	21	29	44	88	216	410	737	1367	1800
3000	7	10	14	17	21	29	44	89	220	425	792	1607	2353
4000	7	10	14	17	21	29	44	89	222	433	821	1750	2735
5000	7	10	14	17	21	29	44	89	223	438	840	1844	3009
6000	7	10	14	17	21	29	44	90	224	442	852	1911	3214
7000	7	10	14	17	21	29	44	90	225	444	861	1961	3373
8000	7	10	14	17	21	29	44	90	225	446	868	1999	3500
9000	7	10	14	17	21	29	44	90	226	447	874	2030	3604
10000	7	10	14	17	21	29	44	90	226	448	878	2055	3689
∞	7	10	14	17	21	29	44	90	228	459	919	2301	4603

Aus der Tabelle kann abgelesen werden:

(1) der Mindeststichprobenumfang n, der erforderlich ist, um mit 99 % Sicherheit mindestens ein krankes Tier zu erfassen, wenn der Anteil kranker Tiere in der Population mindestens dem gewählten Prozentsatz entspricht.

(2) eine obere Schranke für die maximale Anzahl kranker Tiere in der Population mit der angegebenen Mindestsicherheit (99 %) unter der Voraussetzung, dass ein bestimmter Anteil der Population untersucht und durchweg für gesund befunden wurde.

Beispiele:

(1) Der vermutete Anteil kranker Tiere in der Population sei größer oder gleich 2 %; die Population umfasse 480 Tiere. Man wähle N = 500. Aus der Tabelle ergibt sich ein Mindeststichprobenumfang von n = 183. Dieser ist erforderlich, um mit mindestens 99 % Sicherheit mindestens ein krankes Tier zu erfassen. Ist unter den 183 untersuchten Tieren kein krankes Tier, kann geschlossen werden, dass die Prävalenz in der Population mit einer Mindestsicherheit von 99 % kleiner als 2 % ist.

(2) Aus einer Population mit N = 1000 Tieren wurden 10 % untersucht; alle Tiere waren gesund. Aus der Tabelle ergibt sich, dass mit einer Wahrscheinlichkeit von 99 % angenommen werden kann, dass die Anzahl kranker Tiere in der Population nicht größer als 43 ist.

Die Werte der vorgelisteten Tabellen 1 (a-c) beruhen auf der oben dargelegten Approximationsformel, die fast durchgängig zu denselben Ergebnissen führt wie die exakte Berechnung, deren Grundlagen aber dennoch im Folgenden dargelegt werden sollen:

<u>Unendlich große Population:</u>

Angenommen, die Stichprobengröße n soll so gewählt werden, dass bei einer Durchseuchung der Population von 2 % mit 90 % Wahrscheinlichkeit (ε, Statistische Sicherheit) mindestens ein krankes Tier in die Stichprobe gelangt.
Die genannte Bedingung - mindestens ein Tier erkrankt - wird durch verschiedene Szenarien (ein Tier erkrankt, zwei Tiere erkrankt, etc.) erfüllt. Somit müssten die Wahrscheinlichkeiten für all diese Szenarien errechnet und addiert werden, um zu prüfen, ob die Vorgabe der statistischen Sicherheit von 90 % erreicht wird. Einfacher ist jedoch die Prüfung der Umkehrwahrscheinlichkeit (1- ε): Mit welcher Wahrscheinlichkeit gelangt kein erkranktes Tier in die Stichprobe des Umfangs n. Liegt diese Wahrscheinlichkeit bei 10 % oder niedriger, so muss die Summe aller Wahrscheinlichkeiten der Szenarien mit mindestens einem erkrankten Tier bei 90 % oder höher liegen.

Bei unendlich großen Populationen kann die Wahrscheinlichkeit, dass ein in die Stichprobe gewähltes Tier erkrankt ist, als konstant erachtet werden. Sie entspricht der gewählten Prävalenzschwelle (bzw. Designprävalenz, p_0). Die Wahrscheinlichkeit, dass in eine Stichprobe des Umfangs von z. B. drei kein erkranktes Tier gelangt, ist dann:

$$1 - \varepsilon = (1 - p_0) \times (1 - p_0) \times (1 - p_0),$$

oder allgemein für einen Stichprobenumfang von n:

$$1 - \varepsilon = (1 - p_0)^n$$

Für die Statistische Sicherheit ε folgt dann:

$$\varepsilon = 1 - (1 - p_0)^n$$

<u>Endlich große Population:</u>

Für endlich große Populationen kann nicht davon ausgegangen werden, dass die Wahrscheinlichkeit, dass ein in die Stichprobe gewähltes Tier erkrankt ist, konstant ist. Hier muss vom Prinzip einer Stichprobe „ohne Zurücklegen“ ausgegangen werden.

Nur für das erste in die Stichprobe gewählte Tier gilt für die Wahrscheinlichkeit, dass dieses erkrankt ist, wie bei der unendlichen Population, dass diese Wahrscheinlichkeit der Designprävalenz p_0 (mit p_0 = d/N, hier als Anteil, nicht als Prozentwert ausgedrückt) entspricht. Die Wahrscheinlichkeit, dass in eine Stichprobe des Umfangs von z. B. drei aus einer Population der Größe N = 100 kein erkranktes Tier gelangt, ist bei einer angenommenen Designprävalenz von p_0 = d/N = 5/100 dann:

$$1 - \varepsilon = \left(1 - \frac{5}{100}\right) \times \left(1 - \frac{5}{99}\right) \times \left(1 - \frac{5}{98}\right) \approx 0{,}86$$

Man beachte die Änderung der Wahrscheinlichkeit für das Wählen eines erkrankten Tieres durch das „Nicht-Zurücklegen“.

Allgemein gilt für einen Stichprobenumfang von n und eine endliche Populationsgröße N mit d erkrankten Tieren folgende, auf der so genannten hypergeometrischen Wahrscheinlichkeitsfunktion basierende, Gleichung:

$$1 - \varepsilon = \frac{\binom{d}{0}\binom{N-d}{n}}{\binom{N}{n}} = \frac{\binom{N-d}{n}}{\binom{N}{n}} = \frac{(N-d)!}{n!\,(N-d-n)!} \times \frac{n!\,(N-n)!}{N!} = \frac{(N-d)!}{(N-d-n)!} \times \frac{(N-n)!}{N!}$$

Für die statistische Sicherheit ε folgt dann:

$$\varepsilon = 1 - \frac{(N-d)!}{(N-d-n)!} \times \frac{(N-n)!}{N!}$$

Sowohl für die unendliche als auch die endliche Population ist nun n so zu wählen, dass die gewünschte statistische Sicherheit ε erreicht oder überschritten wird (d. h. bei einer gewünschten Sicherheit von 90 % muss ε ≥ 0,9 sein.

Tabelle 2: Die Wahrscheinlichkeit für die Entdeckung kranker Tiere in Abhängigkeit von der Überwachungsintensität

Die folgenden Tabellen zeigen die Wirkung unterschiedlicher Stichprobengrößen auf die Wahrscheinlichkeit, eine Infektion mit niedriger Prävalenz in einer Population nachzuweisen.

Tabelliert ist die Wahrscheinlichkeit, dass mindestens ein krankes Tier gefunden wird, wenn ein bestimmter Anteil der Population (= Auswahlsatz) getestet wird. Die Tabellen 2 (a) bis 2 (e) gelten für die Auswahlsätze 20 %, 30 %, 40 %, 50 % bzw. 60 %. Es werden zehn verschiedene Populationsgrößen N (von 10 bis 100) angenommen; die Anzahl kranker Tiere in

der Population variiert von 1 bis 8. Für Populationsgrößen über 100 nähern sich die Wahrscheinlichkeiten sehr rasch asymptotisch ihren Grenzen für „unendlich große Populationen“, die ebenfalls genannt werden.

Näherungen für nicht in den Tabellen enthaltene Werte können leicht mit Hilfe der bei Tabelle 1 genannten Approximationen berechnet werden.

Es bedeutet: N die Populationsgröße
n der Stichprobenumfang.

Beispiel: 40 % der Tiere eines Bestandes mit 20 Tieren werden getestet (siehe Tabelle 2 (c)). Wenn im Bestand 6 Tiere krank sind, besteht eine Wahrscheinlichkeit von 97,6 %, dass von diesen mindestens eines in die Stichprobe gelangt.

Tabelle 2(a): Auswahlsatz 20 %

		Anzahl positiver Tiere in der Population							
N	n	1	2	3	4	5	6	7	8
10	2	0.200	0.378	0.533	0.667	0.778	0.867	0.933	0.978
20	4	0.200	0.368	0.509	0.624	0.718	0.793	0.852	0.898
30	6	0.200	0.366	0.501	0.612	0.702	0.773	0.830	0.874
40	8	0.200	0.364	0.498	0.607	0.694	0.764	0.819	0.863
50	10	0.200	0.363	0.496	0.603	0.689	0.758	0.813	0.857
60	12	0.200	0.363	0.495	0.601	0.686	0.755	0.809	0.853
70	14	0.200	0.362	0.494	0.599	0.684	0.752	0.807	0.850
80	16	0.200	0.362	0.493	0.598	0.683	0.751	0.804	0.847
90	18	0.200	0.362	0.492	0.597	0.682	0.749	0.803	0.846
100	20	0.200	0.362	0.492	0.597	0.681	0.748	0.802	0.844
∞	∞	0.200	0.360	0.488	0.590	0.672	0.738	0.790	0.832

Tabelle 2(b): Auswahlsatz 30 %

		Anzahl positiver Tiere in der Population							
N	n	1	2	3	4	5	6	7	8
10	3	0.300	0.533	0.708	0.833	0.917	0.967	0.992	1.000
20	6	0.300	0.521	0.681	0.793	0.871	0.923	0.956	0.976
30	9	0.300	0.517	0.672	0.782	0.857	0.909	0.943	0.965
40	12	0.300	0.515	0.668	0.776	0.851	0.902	0.936	0.960
50	15	0.300	0.514	0.666	0.773	0.847	0.898	0.933	0.956
60	18	0.300	0.514	0.665	0.770	0.844	0.895	0.930	0.954
70	21	0.300	0.513	0.663	0.769	0.842	0.893	0.928	0.952
80	24	0.300	0.513	0.663	0.768	0.841	0.892	0.927	0.951
90	27	0.300	0.512	0.662	0.767	0.840	0.891	0.926	0.950
100	30	0.300	0.512	0.661	0.766	0.839	0.890	0.925	0.949
∞	∞	0.300	0.510	0.657	0.760	0.832	0.882	0.918	0.942

Tabelle 2(c): Auswahlsatz 40 %

		Anzahl positiver Tiere in der Population							
N	n	1	2	3	4	5	6	7	8
10	4	0.400	0.667	0.833	0.929	0.976	0.995	1.000	1.000
20	8	0.400	0.653	0.807	0.898	0.949	0.976	0.990	0.996
30	12	0.400	0.648	0.799	0.888	0.940	0.969	0.984	0.993
40	16	0.400	0.646	0.795	0.884	0.935	0.965	0.981	0.990
50	20	0.400	0.645	0.793	0.881	0.933	0.963	0.980	0.989
60	24	0.400	0.644	0.791	0.879	0.931	0.961	0.978	0.988
70	28	0.400	0.643	0.790	0.878	0.930	0.960	0.977	0.987
80	32	0.400	0.643	0.789	0.877	0.929	0.959	0.977	0.987
90	36	0.400	0.643	0.789	0.876	0.928	0.959	0.976	0.987
100	40	0.400	0.642	0.788	0.876	0.927	0.958	0.976	0.986
∞	∞	0.400	0.640	0.784	0.870	0.922	0.953	0.972	0.983

Tabelle 2(d): Auswahlsatz 50 %

		Anzahl positiver Tiere in der Population							
N	n	1	2	3	4	5	6	7	8
10	5	0.500	0.778	0.917	0.976	0.996	1.000	1.000	1.000
20	10	0.500	0.763	0.895	0.957	0.984	0.995	0.998	1.000
30	15	0.500	0.759	0.888	0.950	0.979	0.992	0.997	0.999
40	20	0.500	0.756	0.885	0.947	0.976	0.990	0.996	0.998
50	25	0.500	0.755	0.883	0.945	0.975	0.989	0.995	0.998
60	30	0.500	0.754	0.881	0.944	0.974	0.988	0.995	0.998
70	35	0.500	0.754	0.880	0.943	0.973	0.988	0.994	0.998
80	40	0.500	0.753	0.880	0.942	0.973	0.987	0.994	0.997
90	45	0.500	0.753	0.879	0.942	0.972	0.987	0.994	0.997
100	50	0.500	0.753	0.879	0.941	0.972	0.987	0.994	0.997
∞	∞	0.500	0.750	0.875	0.938	0.969	0.984	0.992	0.996

Tabelle 2(e): Auswahlsatz 60 %

		Anzahl positiver Tiere in der Population							
N	n	1	2	3	4	5	6	7	8
10	6	0.600	0.867	0.967	0.995	1.000	1.000	1.000	1.000
20	12	0.600	0.853	0.951	0.986	0.996	0.999	1.000	1.000
30	18	0.600	0.848	0.946	0.982	0.994	0.998	1.000	1.000
40	24	0.600	0.846	0.943	0.980	0.993	0.998	0.999	1.000
50	30	0.600	0.845	0.942	0.979	0.993	0.998	0.999	1.000
60	36	0.600	0.844	0.941	0.978	0.992	0.997	0.999	1.000
70	42	0.600	0.843	0.940	0.978	0.992	0.997	0.999	1.000
80	48	0.600	0.843	0.940	0.977	0.992	0.997	0.999	1.000
90	54	0.600	0.843	0.939	0.977	0.991	0.997	0.999	1.000
100	60	0.600	0.842	0.939	0.977	0.991	0.997	0.999	1.000
∞	∞	0.600	0.840	0.936	0.974	0.990	0.996	0.998	0.999

Tabelle 3: Die Wahrscheinlichkeit, kranke Tiere nicht zu entdecken

Die Tabelle gibt die Wahrscheinlichkeit dafür an, dass zufällig kein krankes Tier in die Stichprobe gelangt, obwohl die Krankheit in der Population vorkommt. Werte kleiner 0.0005 werden in der ersten Tabellenzelle, ab der die genannte Bedingung gilt, mit „<0.0005“ ausgewiesen. Die Population wird als „unendlich groß“ angenommen.

	Anzahl getesteter Tiere (= Stichprobenumfang)									
Prävalenz	5	10	25	50	75	100	200	250	500	1000
1 %	0.951	0.904	0.778	0.605	0.471	0.366	0.134	0.081	0.007	<0.0005
2 %	0.904	0.817	0.603	0.364	0.220	0.133	0.018	0.006	<0.0005	
3 %	0.859	0.737	0.467	0.218	0.102	0.048	0.002	<0.0005		
4 %	0.815	0.665	0.360	0.130	0.047	0.017	<0.0005			
5 %	0.774	0.599	0.277	0.077	0.021	0.006	<0.0005			
6 %	0.734	0.539	0.213	0.045	0.010	0.002	<0.0005			
7 %	0.696	0.484	0.163	0.027	0.004	0.001	<0.0005			
8 %	0.659	0.434	0.124	0.015	0.002	<0.0005				
9 %	0.624	0.389	0.095	0.009	0.001	<0.0005				
10 %	0.590	0.349	0.072	0.005	<0.0005					
12 %	0.528	0.279	0.041	0.002	<0.0005					
14 %	0.470	0.221	0.023	0.001	<0.0005					
16 %	0.418	0.175	0.013	<0.0005						
18 %	0.371	0.137	0.007	<0.0005						
20 %	0.328	0.107	0.004	<0.0005						
24 %	0.254	0.064	0.001	<0.0005						
28 %	0.193	0.037	<0.0005							
32 %	0.145	0.021	<0.0005							
36 %	0.107	0.012	<0.0005							
40 %	0.078	0.006	<0.0005							
50 %	0.031	0.001	<0.0005							
60 %	0.010	<0.0005								

Obwohl Tabelle 3 – ähnlich wie Tabelle 2 – die Wirkung des Stichprobenumfangs auf die Wahrscheinlichkeit der Erfassung bzw. Nichterfassung kranker Tiere zeigt, wird das Problem hier auf andere Weise angegangen:

Die Tabelle geht von einem bestimmten Umfang der Stichprobe n aus, die aus einer unendlich großen Population mit einem angenommenen Anteil kranker Tiere p (Prävalenz) entnommen wird. Die Wahrscheinlichkeit, dass zufällig keines der kranken Tiere in die Stichprobe gelangt, ist

$$(1-p)^n$$

Die Werte dieser Funktion sind in Abhängigkeit von p und n tabelliert.

Beispiel:

Aus einer sehr großen Population mit 10 % kranken Tieren werden wiederholt Stichproben von jeweils n = 25 Tieren entnommen und untersucht. Aus Tabelle 3 ergibt sich die Zahl 0,072. Dies bedeutet, dass etwa 7,2 % der Stichproben kein krankes Tier enthalten werden.

2.2. Krankheits-Monitoring und Seuchen-Überwachung

Ein Krankheits-Monitoring wirft ein anderes Problem auf. Als Beispiel sei die Tuberkuloseüberwachung auf Basis von Schlachthofbefunden betrachtet. In diesem Fall weiß man, welcher Anteil der Tiere eines Bestandes anlässlich der Schlachtung untersucht wurde. Wenn all diese Proben Tuberkulose-negativ ausgefallen sind, möchte man wissen, ob dies ausreicht, um sagen zu können, der ganze Bestand sei negativ. Auf diese Weise könnte man auf eine vollständige Bestandsuntersuchung verzichten.

Die Frage, die hier zu beantworten ist, stellt die Umkehrung jener Frage dar, die im Abschnitt 1.1 erörtert wurde. Anstatt nach dem Stichprobenumfang zu fragen, der eine vorgegebene statistische Sicherheit für den Nachweis eines bestimmten Durchseuchungsgrades gewährleistet, wird jetzt nach einem möglichen Durchseuchungsgrad gefragt, der in einem Bestand zu festgelegter Sicherheit maximal vorhanden sein (und übersehen werden) könnte, wenn in der Stichprobe kein Reagent gefunden wurde. Tabelle 1 kann auch zur Beantwortung dieser Frage verwendet werden; sie gibt Obergrenzen für die Anzahl positiver Tiere an, die in einer Population vorhanden sein können, obwohl die Stichprobe frei von positiven Tieren war.

<u>Ein spezielles Wort zu Schlachthofdaten:</u>

Die Nutzung der Tabellen 1 bis 3 setzt voraus, dass eine Zufallsstichprobe untersucht wird. Zur Schlachtung bestimmte Tiere stellen aber keine wirkliche Zufallsstichprobe aus einem Bestand dar. Es sind Tiere, die vom Besitzer vermutlich als in akzeptablem Gesundheitszustand befindlich eingeschätzt wurden. Die meisten dieser Tiere werden daher frei von klinischen Erkrankungen sein. Darüber hinaus sind Schlachttiere nicht repräsentativ für die Altersstruktur und das Geschlechtsverhältnis in der Population, der sie entstammen. Des Weiteren dürften beispielsweise Tiere mit schwacher Leistung, z.B. im Hinblick auf die Zwischenkalbezeit oder die Milchproduktion, unter weiblichen Schlachtrindern überpräsentiert sein.

Der Rückschluss von Schlachthof-Daten auf die Häufigkeit von Erkrankungen (oder von anderen Merkmalen) in einer Population erfordert daher besondere Sorgfalt und Vorsicht. Über die Häufigkeit solcher Merkmale, die keinen Einfluss auf die Bereitstellung eines Tieres zur Schlachtung (oder auf die Wahrscheinlichkeit seiner Aussonderung aus dem Bestand) haben, lassen sich aus sorgfältig erhobenen Schlachthof-Statistiken durchaus Informationen gewinnen. Dennoch sollte man mögliche Verzerrungen in derartigen Untersuchungen schon von vornherein ins Auge fassen und abzuschätzen versuchen.

Andererseits stellen Rückverfolgungsuntersuchungen auf Grund von Schlachthof-Befunden eine kostengünstige Methode zum Aufspüren infizierter Bestände oder Herden dar. Schlachthof-Befunddaten können auch von einer Periode zur nächsten kumuliert werden. Auf diese Weise entstehen Befund-Dateien über die Herkunftsbestände. Auf solche Daten sind die Tabellen selbstverständlich im strengen Sinne nicht korrekt anwendbar, denn die Tiere des Bestandes oder der Herde, für welche die gefundenen Anteilswerte erhoben werden, stellen eine fluktuierende Population dar, denn es werden ständig Tiere ersetzt, sei es durch eigene Nachzucht oder durch Zukauf.

3. Erhebungen zur Bestimmung der Prävalenz einer Krankheit

3.1. Schätzung der Häufigkeit kranker Tiere in einer Population

Für die Schätzung der Prävalenz einer Krankheit gilt dieselbe Stichprobentheorie wie beim Nachweis des Vorhandenseins einer Krankheit. Die Vorgehensweise zum Nachweis einer Krankheit ist eine spezielle Anwendung des Verfahrens, mit dem die Häufigkeit der Krankheit in einer Population geschätzt wird. Auch bei Letzterer hängt der erforderliche Stichprobenumfang vom tatsächlichen Durchseuchungsgrad in der Population sowie von der gewünschten statistischen Sicherheit und der Größe der Population ab – des Weiteren aber auch noch von der gewünschten Genauigkeit (siehe unten und Tabellen 5).

Allerdings lautet die Fragestellung im Zusammenhang mit Erhebungen zur Prävalenzbestimmung gewöhnlich folgendermaßen: „Was lässt sich über den tatsächlichen Durchseuchungsgrad in einer Population von bekannter Größe aussagen, nachdem in einer Zufallsstichprobe gegebener Größe ein bestimmter Anteil positiver Tiere gefunden wurde?“ Bei diesem Ansatz wird nicht die Stichprobengröße auf Grundlage u. a. der gewünschten Genauigkeit, sondern die resultierende Genauigkeit auf Basis der gegebenen Stichprobe ermittelt (s. u.).

Da eine Stichprobe im Allgemeinen einen fehlerbehafteten Punktschätzwert liefert, kann der unbekannte wahre Wert der Prävalenz (bezogen auf die Population) mehr oder weniger stark vom Punktschätzwert abweichen. Eine Quantifizierung der beschriebenen Abweichung (Genauigkeit) von der wahren Prävalenz ist mit Hilfe eines Punktschätzers nicht möglich.
Aus diesem Grund bedient man sich der Schätzung von Konfidenzintervallen (s. u.), die einen Bereich „möglicher“ Werte von Prävalenzen zu einer zuvor festgelegten statistischen Sicherheit beschreiben.

Mit Hilfe von Tabelle 4 lassen sich zu einer statistischen Sicherheit von 95 % bzw. 99 % die exakten zweiseitigen Konfidenzgrenzen für die wahre Prävalenz (unter Annahme eines perfekten Testes mit Sensitivität = Spezifität = 100 %) in einer unendlich großen Population ermitteln (Methode nach CLOPPER-PEARSON[6]) . Unter den Konfidenzgrenzen versteht man zwei Werte ober- bzw. unterhalb des mit Hilfe der Stichprobe gefundenen (Punkt-) Schätzwertes. Diese Grenzen schließen ein Intervall von möglichen (wahren) Prävalenzwerten zu einer zuvor festgelegten statistischen Sicherheit (=Konfidenzniveau) ein, welches als Konfidenzintervall bezeichnet wird. Die genaue Interpretation am Beispiel einer statistischen Sicherheit von 95 % ist die folgende:

Mit einer Wahrscheinlichkeit von 95 % überdeckt das 95 %-Konfidenzintervall den wahren Wert der Prävalenz.

Das Konfidenzintervall muss dabei nicht symmetrisch um die geschätzte Prävalenz liegen, sondern kann auf einer der Seiten (mitunter auch deutlich) größer sein.
Konfidenzintervalle weisen bei kleinen Stichprobenumfängen eine beträchtliche Spanne auf. Darin kommt die Unbestimmtheit zum Ausdruck, mit der ein Schätzwert behaftet ist, falls nur wenige Beobachtungen zur Verfügung stehen. Die Spanne und die Sicherheit, dass der wahre Wert von der Spanne überdeckt wird, stehen in einem direkten Zusammenhang. Hingegen stehen die Spanne und Genauigkeit in einem indirekten Zusammenhang (umgekehrt proportional). Wird eine größere Genauigkeit, d. h. eine kleinere Spanne verlangt, so müssen

[6] beschrieben in Brown, L. D. , T.T. Cai, and A. DasGupta, *Interval Estimation for a Binomial Proportion.* Statistical Science, 2001. **16**(2): p. 101.

weitere zufällig ausgewählte Tiere getestet werden, bis die gewünschte Genauigkeit (gemessen an der Spanne des Konfidenzintervalls) erreicht ist.

Tabelle 4 a1 **Statistische Sicherheit** **95 %**

Untere Grenze (U) und obere Grenze (O) des Konfidenzintervalls für die Schätzung der Prävalenz einer Krankheit bei unendlich großer Population.
Die angegebenen Werte sind entgegen den üblichen Regeln für die unteren Konfidenzgrenzen abgerundet und für die oberen Konfidenzgrenzen aufgerundet, so dass das angegebene Sicherheitsniveau mit den angegebenen Dezimalstellen stets eingehalten wird.

	Anteil positiver Befunde in der Stichprobe					
n	10%		20%		30%	
	U	O	U	O	U	O
10	0.002	0.446	0.025	0.557	0.066	0.653
20	0.012	0.317	0.057	0.437	0.118	0.543
30	0.021	0.266	0.077	0.386	0.147	0.494
40	0.027	0.237	0.090	0.357	0.165	0.466
50	0.033	0.219	0.100	0.338	0.178	0.447
60	0.037	0.206	0.107	0.324	0.188	0.433
70	0.041	0.196	0.113	0.313	0.196	0.422
80	0.044	0.188	0.118	0.305	0.202	0.413
90	0.046	0.182	0.123	0.298	0.207	0.406
100	0.049	0.177	0.126	0.292	0.212	0.400
120	0.052	0.169	0.132	0.283	0.219	0.391
140	0.055	0.163	0.137	0.276	0.225	0.384
160	0.058	0.158	0.141	0.271	0.230	0.378
180	0.060	0.154	0.144	0.266	0.234	0.373
200	0.062	0.151	0.146	0.263	0.237	0.369
250	0.065	0.145	0.152	0.256	0.243	0.361
300	0.068	0.140	0.156	0.250	0.248	0.356
350	0.070	0.137	0.159	0.246	0.252	0.351
400	0.072	0.134	0.161	0.243	0.255	0.348
450	0.073	0.132	0.164	0.240	0.257	0.345
500	0.075	0.130	0.165	0.238	0.260	0.343
600	0.077	0.127	0.168	0.235	0.263	0.339
700	0.078	0.125	0.170	0.232	0.266	0.336
800	0.080	0.123	0.172	0.230	0.268	0.334
900	0.081	0.122	0.174	0.228	0.270	0.332
1000	0.082	0.121	0.175	0.227	0.271	0.330
1200	0.083	0.119	0.177	0.224	0.274	0.327
1400	0.084	0.117	0.179	0.222	0.276	0.325
1600	0.085	0.116	0.180	0.221	0.277	0.324
1800	0.086	0.115	0.181	0.220	0.278	0.322
2000	0.087	0.114	0.182	0.219	0.279	0.321
3000	0.089	0.112	0.185	0.215	0.283	0.317
4000	0.090	0.110	0.187	0.213	0.285	0.315
5000	0.091	0.109	0.188	0.212	0.287	0.313
6000	0.092	0.108	0.189	0.211	0.288	0.312
7000	0.093	0.108	0.190	0.210	0.289	0.311
8000	0.093	0.107	0.191	0.209	0.289	0.311
9000	0.093	0.107	0.191	0.209	0.290	0.310
10000	0.094	0.107	0.192	0.208	0.291	0.310

Tabelle 4 a2 **Statistische Sicherheit** **95 %**

Untere Grenze (U) und obere Grenze (O) des Konfidenzintervalls für die Schätzung der Prävalenz einer Krankheit bei unendlich großer Population.
Die angegebenen Werte sind entgegen den üblichen Regeln für die unteren Konfidenzgrenzen abgerundet und für die oberen Konfidenzgrenzen aufgerundet, so dass das angegebene Sicherheitsniveau mit den angegebenen Dezimalstellen stets eingehalten wird.

	Anteil positiver Befunde in der Stichprobe					
n	40%		50%		60%	
	U	O	U	O	U	O
10	0.121	0.738	0.187	0.813	0.262	0.879
20	0.191	0.640	0.271	0.729	0.360	0.809
30	0.226	0.594	0.312	0.688	0.406	0.774
40	0.248	0.567	0.338	0.662	0.433	0.752
50	0.264	0.549	0.355	0.645	0.451	0.736
60	0.275	0.535	0.368	0.632	0.465	0.725
70	0.284	0.525	0.378	0.622	0.475	0.716
80	0.292	0.516	0.386	0.614	0.484	0.708
90	0.298	0.509	0.392	0.608	0.491	0.702
100	0.303	0.503	0.398	0.602	0.497	0.697
120	0.311	0.494	0.407	0.593	0.506	0.689
140	0.318	0.487	0.414	0.586	0.513	0.682
160	0.323	0.481	0.420	0.580	0.519	0.677
180	0.327	0.476	0.424	0.576	0.524	0.673
200	0.331	0.472	0.428	0.572	0.528	0.669
250	0.338	0.464	0.436	0.564	0.536	0.662
300	0.344	0.458	0.441	0.559	0.542	0.656
350	0.348	0.454	0.446	0.554	0.546	0.652
400	0.351	0.450	0.449	0.551	0.550	0.649
450	0.354	0.447	0.452	0.548	0.553	0.646
500	0.356	0.445	0.455	0.545	0.555	0.644
600	0.360	0.441	0.459	0.541	0.559	0.640
700	0.363	0.438	0.462	0.538	0.562	0.637
800	0.365	0.435	0.464	0.536	0.565	0.635
900	0.367	0.433	0.466	0.534	0.567	0.633
1000	0.369	0.432	0.468	0.532	0.568	0.631
1200	0.372	0.429	0.471	0.529	0.571	0.628
1400	0.374	0.427	0.473	0.527	0.573	0.626
1600	0.375	0.425	0.475	0.525	0.575	0.625
1800	0.377	0.424	0.476	0.524	0.576	0.623
2000	0.378	0.422	0.477	0.523	0.578	0.622
3000	0.382	0.418	0.481	0.519	0.582	0.618
4000	0.384	0.416	0.484	0.516	0.584	0.616
5000	0.386	0.414	0.486	0.514	0.586	0.614
6000	0.387	0.413	0.487	0.513	0.587	0.613
7000	0.388	0.412	0.488	0.512	0.588	0.612
8000	0.389	0.411	0.488	0.512	0.589	0.611
9000	0.389	0.411	0.489	0.511	0.589	0.611
10000	0.390	0.410	0.490	0.510	0.590	0.610

Tabelle 4 a3 **Statistische Sicherheit** **95 %**

Untere Grenze (U) und obere Grenze (O) des Konfidenzintervalls für die Schätzung der Prävalenz einer Krankheit bei unendlich großer Population.
Die angegebenen Werte sind entgegen den üblichen Regeln für die unteren Konfidenzgrenzen abgerundet und für die oberen Konfidenzgrenzen aufgerundet, so dass das angegebene Sicherheitsniveau mit den angegebenen Dezimalstellen stets eingehalten wird.

	Anteil positiver Befunde in der Stichprobe					
n	70%		80%		90%	
	U	O	U	O	U	O
10	0.347	0.934	0.443	0.975	0.554	0.998
20	0.457	0.882	0.563	0.943	0.683	0.988
30	0.506	0.853	0.614	0.923	0.734	0.979
40	0.534	0.835	0.643	0.910	0.763	0.973
50	0.553	0.822	0.662	0.900	0.781	0.967
60	0.567	0.812	0.676	0.893	0.794	0.963
70	0.578	0.804	0.687	0.887	0.804	0.959
80	0.587	0.798	0.695	0.882	0.812	0.956
90	0.594	0.793	0.702	0.877	0.818	0.954
100	0.600	0.788	0.708	0.874	0.823	0.951
120	0.609	0.781	0.717	0.868	0.831	0.948
140	0.616	0.775	0.724	0.863	0.837	0.945
160	0.622	0.770	0.729	0.859	0.842	0.942
180	0.627	0.766	0.734	0.856	0.846	0.940
200	0.631	0.763	0.737	0.854	0.849	0.938
250	0.639	0.757	0.744	0.848	0.855	0.935
300	0.644	0.752	0.750	0.844	0.860	0.932
350	0.649	0.748	0.754	0.841	0.863	0.930
400	0.652	0.745	0.757	0.839	0.866	0.928
450	0.655	0.743	0.760	0.836	0.868	0.927
500	0.657	0.740	0.762	0.835	0.870	0.925
600	0.661	0.737	0.765	0.832	0.873	0.923
700	0.664	0.734	0.768	0.830	0.875	0.922
800	0.666	0.732	0.770	0.828	0.877	0.920
900	0.668	0.730	0.772	0.826	0.878	0.919
1000	0.670	0.729	0.773	0.825	0.879	0.918
1200	0.673	0.726	0.776	0.823	0.881	0.917
1400	0.675	0.724	0.778	0.821	0.883	0.916
1600	0.676	0.723	0.779	0.820	0.884	0.915
1800	0.678	0.722	0.780	0.819	0.885	0.914
2000	0.679	0.721	0.781	0.818	0.886	0.913
3000	0.683	0.717	0.785	0.815	0.888	0.911
4000	0.685	0.715	0.787	0.813	0.890	0.910
5000	0.687	0.713	0.788	0.812	0.891	0.909
6000	0.688	0.712	0.789	0.811	0.892	0.908
7000	0.689	0.711	0.790	0.810	0.892	0.907
8000	0.689	0.711	0.791	0.809	0.893	0.907
9000	0.690	0.710	0.791	0.809	0.893	0.907
10000	0.690	0.709	0.792	0.808	0.893	0.906

Tabelle 4 b1 **Statistische Sicherheit** **99 %**

Untere Grenze (U) und obere Grenze (O) des Konfidenzintervalls für die Schätzung der Prävalenz einer Krankheit bei unendlich großer Population.
Die angegebenen Werte sind entgegen den üblichen Regeln für die unteren Konfidenzgrenzen abgerundet und für die oberen Konfidenzgrenzen aufgerundet, so dass das angegebene Sicherheitsniveau mit den angegebenen Dezimalstellen stets eingehalten wird.

	Anteil positiver Befunde in der Stichprobe					
n	10%		20%		30%	
	U	O	U	O	U	O
10	0.000	0.545	0.010	0.649	0.037	0.736
20	0.005	0.388	0.035	0.507	0.084	0.610
30	0.011	0.321	0.054	0.443	0.114	0.551
40	0.017	0.283	0.068	0.406	0.134	0.515
50	0.022	0.259	0.078	0.381	0.149	0.490
60	0.026	0.241	0.086	0.363	0.160	0.473
70	0.029	0.228	0.093	0.349	0.169	0.459
80	0.033	0.218	0.099	0.338	0.177	0.448
90	0.035	0.209	0.104	0.329	0.183	0.439
100	0.038	0.202	0.108	0.322	0.189	0.431
120	0.042	0.192	0.115	0.310	0.197	0.419
140	0.045	0.183	0.120	0.301	0.205	0.409
160	0.048	0.177	0.125	0.294	0.210	0.402
180	0.050	0.172	0.129	0.288	0.215	0.396
200	0.052	0.167	0.132	0.283	0.219	0.391
250	0.057	0.159	0.139	0.273	0.227	0.381
300	0.060	0.153	0.144	0.266	0.233	0.373
350	0.062	0.149	0.147	0.261	0.238	0.367
400	0.065	0.145	0.151	0.257	0.242	0.363
450	0.066	0.142	0.153	0.253	0.245	0.359
500	0.068	0.140	0.155	0.250	0.248	0.356
600	0.070	0.136	0.159	0.246	0.252	0.351
700	0.072	0.133	0.162	0.242	0.256	0.347
800	0.074	0.131	0.164	0.239	0.258	0.344
900	0.075	0.129	0.166	0.237	0.261	0.341
1000	0.077	0.127	0.168	0.235	0.263	0.339
1200	0.078	0.125	0.171	0.232	0.266	0.336
1400	0.080	0.123	0.173	0.229	0.268	0.333
1600	0.081	0.121	0.174	0.227	0.270	0.331
1800	0.082	0.120	0.176	0.226	0.272	0.329
2000	0.083	0.119	0.177	0.224	0.273	0.328
3000	0.086	0.115	0.181	0.220	0.278	0.323
4000	0.088	0.113	0.183	0.217	0.281	0.320
5000	0.089	0.112	0.185	0.215	0.283	0.317
6000	0.090	0.111	0.186	0.214	0.284	0.316
7000	0.090	0.110	0.187	0.213	0.285	0.315
8000	0.091	0.109	0.188	0.212	0.286	0.314
9000	0.092	0.109	0.189	0.212	0.287	0.313
10000	0.092	0.108	0.189	0.211	0.288	0.312

Tabelle 4 b2 **Statistische Sicherheit** **99 %**

Untere Grenze (U) und obere Grenze (O) des Konfidenzintervalls für die Schätzung der Prävalenz einer Krankheit bei unendlich großer Population.
Die angegebenen Werte sind entgegen den üblichen Regeln für die unteren Konfidenzgrenzen abgerundet und für die oberen Konfidenzgrenzen aufgerundet, so dass das angegebene Sicherheitsniveau mit den angegebenen Dezimalstellen stets eingehalten wird.

	Anteil positiver Befunde in der Stichprobe					
n	40%		50%		60%	
	U	O	U	O	U	O
10	0.076	0.810	0.128	0.872	0.190	0.924
20	0.145	0.701	0.217	0.783	0.299	0.855
30	0.185	0.647	0.264	0.736	0.353	0.815
40	0.210	0.614	0.294	0.706	0.386	0.790
50	0.228	0.591	0.315	0.685	0.409	0.772
60	0.242	0.574	0.331	0.669	0.426	0.758
70	0.253	0.561	0.343	0.657	0.439	0.747
80	0.262	0.550	0.353	0.647	0.450	0.738
90	0.269	0.542	0.361	0.639	0.458	0.731
100	0.276	0.534	0.368	0.632	0.466	0.724
120	0.286	0.522	0.380	0.620	0.478	0.714
140	0.294	0.513	0.389	0.611	0.487	0.706
160	0.301	0.505	0.396	0.604	0.495	0.699
180	0.306	0.499	0.402	0.598	0.501	0.694
200	0.311	0.494	0.407	0.593	0.506	0.689
250	0.320	0.484	0.417	0.583	0.516	0.680
300	0.327	0.476	0.424	0.576	0.524	0.673
350	0.332	0.470	0.430	0.570	0.530	0.668
400	0.337	0.466	0.434	0.566	0.534	0.663
450	0.340	0.462	0.438	0.562	0.538	0.660
500	0.343	0.459	0.441	0.559	0.541	0.657
600	0.348	0.453	0.446	0.554	0.547	0.652
700	0.352	0.449	0.450	0.550	0.551	0.648
800	0.355	0.446	0.453	0.547	0.554	0.645
900	0.358	0.444	0.456	0.544	0.556	0.642
1000	0.360	0.441	0.458	0.542	0.559	0.640
1200	0.363	0.438	0.462	0.538	0.562	0.637
1400	0.366	0.435	0.465	0.535	0.565	0.634
1600	0.368	0.433	0.467	0.533	0.567	0.632
1800	0.370	0.431	0.469	0.531	0.569	0.630
2000	0.371	0.429	0.470	0.530	0.571	0.629
3000	0.376	0.424	0.476	0.524	0.576	0.624
4000	0.380	0.421	0.479	0.521	0.579	0.620
5000	0.382	0.419	0.481	0.519	0.581	0.618
6000	0.383	0.417	0.483	0.517	0.583	0.617
7000	0.384	0.416	0.484	0.516	0.584	0.616
8000	0.385	0.415	0.485	0.515	0.585	0.615
9000	0.386	0.414	0.486	0.514	0.586	0.614
10000	0.387	0.413	0.487	0.513	0.587	0.613

Tabelle 4 b3 **Statistische Sicherheit** **99 %**

Untere Grenze (U) und obere Grenze (O) des Konfidenzintervalls für die Schätzung der Prävalenz einer Krankheit bei unendlich großer Population.
Die angegebenen Werte sind entgegen den üblichen Regeln für die unteren Konfidenzgrenzen abgerundet und für die oberen Konfidenzgrenzen aufgerundet, so dass das angegebene Sicherheitsniveau mit den angegebenen Dezimalstellen stets eingehalten wird.

	Anteil positiver Befunde in der Stichprobe					
n	70%		80%		90%	
	U	O	U	O	U	O
10	0.264	0.963	0.351	0.990	0.455	1.000
20	0.390	0.916	0.493	0.965	0.612	0.995
30	0.449	0.886	0.557	0.946	0.679	0.989
40	0.485	0.866	0.594	0.932	0.717	0.983
50	0.510	0.851	0.619	0.922	0.741	0.978
60	0.527	0.840	0.637	0.914	0.759	0.974
70	0.541	0.831	0.651	0.907	0.772	0.971
80	0.552	0.823	0.662	0.901	0.782	0.967
90	0.561	0.817	0.671	0.896	0.791	0.965
100	0.569	0.811	0.678	0.892	0.798	0.962
120	0.581	0.803	0.690	0.885	0.808	0.958
140	0.591	0.795	0.699	0.880	0.817	0.955
160	0.598	0.790	0.706	0.875	0.823	0.952
180	0.604	0.785	0.712	0.871	0.828	0.950
200	0.609	0.781	0.717	0.868	0.833	0.948
250	0.619	0.773	0.727	0.861	0.841	0.943
300	0.627	0.767	0.734	0.856	0.847	0.940
350	0.633	0.762	0.739	0.853	0.851	0.938
400	0.637	0.758	0.743	0.849	0.855	0.935
450	0.641	0.755	0.747	0.847	0.858	0.934
500	0.644	0.752	0.750	0.845	0.860	0.932
600	0.649	0.748	0.754	0.841	0.864	0.930
700	0.653	0.744	0.758	0.838	0.867	0.928
800	0.656	0.742	0.761	0.836	0.869	0.926
900	0.659	0.739	0.763	0.834	0.871	0.925
1000	0.661	0.737	0.765	0.832	0.873	0.923
1200	0.664	0.734	0.768	0.829	0.875	0.922
1400	0.667	0.732	0.771	0.827	0.877	0.920
1600	0.669	0.730	0.773	0.826	0.879	0.919
1800	0.671	0.728	0.774	0.824	0.880	0.918
2000	0.672	0.727	0.776	0.823	0.881	0.917
3000	0.677	0.722	0.780	0.819	0.885	0.914
4000	0.680	0.719	0.783	0.817	0.887	0.912
5000	0.683	0.717	0.785	0.815	0.888	0.911
6000	0.684	0.716	0.786	0.814	0.889	0.910
7000	0.685	0.715	0.787	0.813	0.890	0.910
8000	0.686	0.714	0.788	0.812	0.891	0.909
9000	0.687	0.713	0.788	0.811	0.891	0.908
10000	0.688	0.712	0.789	0.811	0.892	0.908

3.2. Genauigkeit von Häufigkeitsschätzungen

Aus Tabelle 5 kann für verschiedene Genauigkeiten näherungsweise die notwendige Anzahl zu untersuchender Tiere abgelesen werden, wenn gefordert wird, dass der wahre Anteil kranker Tiere in der Population mit einer vorgegebenen statistischen Sicherheit und Genauigkeit zu schätzen ist. Als Genauigkeit wird dabei die halbe Spanne des zugehörigen Konfidenzintervalls (s. Tabellen 5) festgelegt. Die Genauigkeit als halbe Spanne eines Konfidenzintervalls ist nicht mit der statistischen Sicherheit zu verwechseln.

Zur Berechnung der Stichprobenumfänge in Tabelle 5 wird zunächst unter Annahme einer unendlich großen Population die folgende Näherungsformel[7] verwendet:

$$n = \frac{u_{1-\frac{\alpha}{2}}^2}{g^2} \times p \times (1-p)$$

Anschließend wird der berechnete Stichprobenumfang in Abhängigkeit von der tatsächlichen, endlichen Größe der Population mit Hilfe folgender Korrekturformel[7] (nach unten) angepasst:

$$n_{adj} = \frac{N \times n}{N + n}$$

wobei p – die vermutete Prävalenz,
g – die gewählte minimale Genauigkeit,
u – das entsprechende Quantil der Standardnormalverteilung,
α – 1 – die statistische Sicherheit,
N – die Größe der Population,
n – der Stichprobenumfang (hier Annahme einer unendlich großen Population),
n_{adj} – der adjustierte (endlich große Population) Stichprobenumfang
ist.

Das Quantil u beträgt für eine statistische Sicherheit von…

90 %: 1.645,

95 %: 1.960 und

99 %: 2.576.

[7] Thrusfield, M. (2018). Veterinary Epidemiology. Wiley-Blackwell, 4. Auflage.

Tabelle 5 a **Statistische Sicherheit** **95 %**

Notwendiger Stichprobenumfang zur Schätzung der Prävalenz einer Krankheit bei einer **Genauigkeit von mindestens 10 %:**

	Vermutete Prävalenz								
N	10%	20%	30%	40%	50%	60%	70%	80%	90%
10	8	9	9	10	10	10	9	9	8
20	13	16	17	17	17	17	17	16	13
30	17	21	22	23	23	23	22	21	17
40	19	25	27	28	29	28	27	25	19
50	21	28	31	33	33	33	31	28	21
60	22	31	35	37	37	37	35	31	22
70	24	33	38	40	41	40	38	33	24
80	25	35	41	43	44	43	41	35	25
90	25	37	43	46	47	46	43	37	25
100	26	39	45	48	49	48	45	39	26
120	27	41	49	53	54	53	49	41	27
140	28	43	52	56	57	56	52	43	28
160	29	45	54	59	61	59	54	45	29
180	30	46	56	61	63	61	56	46	30
200	30	48	58	64	65	64	58	48	30
250	31	50	61	68	70	68	61	50	31
300	32	52	64	71	73	71	64	52	32
350	32	53	66	73	76	73	66	53	32
400	32	54	68	75	78	75	68	54	32
450	33	55	69	77	80	77	69	55	33
500	33	55	70	78	81	78	70	55	33
600	33	56	72	80	83	80	72	56	33
700	33	57	73	82	85	82	73	57	33
800	34	58	74	83	86	83	74	58	34
900	34	58	75	84	87	84	75	58	34
1000	34	58	75	85	88	85	75	58	34
1200	34	59	76	86	89	86	76	59	34
1400	34	59	77	87	90	87	77	59	34
1600	34	60	77	88	91	88	77	60	34
1800	34	60	78	88	92	88	78	60	34
2000	34	60	78	89	92	89	78	60	34
3000	35	61	79	90	94	90	79	61	35
4000	35	61	80	91	94	91	80	61	35
5000	35	61	80	91	95	91	80	61	35
6000	35	61	80	91	95	91	80	61	35
7000	35	61	80	91	95	91	80	61	35
8000	35	61	80	92	95	92	80	61	35
9000	35	62	80	92	96	92	80	62	35
10000	35	62	81	92	96	92	81	62	35
∞	35	62	81	93	97	93	81	62	35

Tabelle 5 b **Statistische Sicherheit** **95 %**

Notwendiger Stichprobenumfang zur Schätzung der Prävalenz einer Krankheit bei einer **Genauigkeit von mindestens 5 %:**

	Vermutete Prävalenz								
N	10%	20%	30%	40%	50%	60%	70%	80%	90%
10	10	10	10	10	10	10	10	10	10
20	18	19	19	19	20	19	19	19	18
30	25	27	28	28	28	28	28	27	25
40	32	35	36	37	37	37	36	35	32
50	37	42	44	45	45	45	44	42	37
60	42	49	51	52	52	52	51	49	42
70	47	55	58	59	60	59	58	55	47
80	51	61	65	66	67	66	65	61	51
90	55	66	71	73	73	73	71	66	55
100	59	72	77	79	80	79	77	72	59
120	65	81	88	91	92	91	88	81	65
140	70	90	98	102	103	102	98	90	70
160	75	97	107	112	113	112	107	97	75
180	79	104	116	121	123	121	116	104	79
200	82	111	124	130	132	130	124	111	82
250	90	124	141	149	152	149	141	124	90
300	95	136	156	166	169	166	156	136	95
350	100	145	168	180	184	180	168	145	100
400	103	153	179	192	196	192	179	153	103
450	106	159	188	203	208	203	188	159	106
500	109	165	197	213	218	213	197	165	109
600	113	175	210	229	235	229	210	175	113
700	116	182	221	242	249	242	221	182	116
800	118	189	230	253	260	253	230	189	118
900	120	194	238	262	270	262	238	194	120
1000	122	198	244	270	278	270	244	198	122
1200	125	205	255	283	291	283	255	205	125
1400	126	210	263	292	302	292	263	210	126
1600	128	214	269	300	310	300	269	214	128
1800	129	217	274	307	317	307	274	217	129
2000	130	219	278	312	323	312	278	219	130
3000	133	228	292	329	341	329	292	228	133
4000	134	232	299	338	351	338	299	232	134
5000	135	235	304	344	357	344	304	235	135
6000	136	237	307	348	362	348	307	237	136
7000	136	238	309	351	365	351	309	238	136
8000	136	239	311	353	367	353	311	239	136
9000	137	240	312	355	369	355	312	240	137
10000	137	240	313	356	370	356	313	240	137
∞	139	246	323	369	385	369	323	246	139

Tabelle 5 c **Statistische Sicherheit** **95 %**

Notwendiger Stichprobenumfang zur Schätzung der Prävalenz einer Krankheit bei einer **Genauigkeit von mindestens 1 %:**

	Vermutete Prävalenz								
N	10%	20%	30%	40%	50%	60%	70%	80%	90%
10	10	10	10	10	10	10	10	10	10
20	20	20	20	20	20	20	20	20	20
30	30	30	30	30	30	30	30	30	30
40	40	40	40	40	40	40	40	40	40
50	50	50	50	50	50	50	50	50	50
60	59	60	60	60	60	60	60	60	59
70	69	70	70	70	70	70	70	70	69
80	79	79	80	80	80	80	80	79	79
90	88	89	90	90	90	90	90	89	88
100	98	99	99	99	99	99	99	99	98
120	116	118	119	119	119	119	119	118	116
140	135	137	138	138	138	138	138	137	135
160	153	156	157	158	158	158	157	156	153
180	172	175	177	177	177	177	177	175	172
200	190	194	196	196	196	196	196	194	190
250	234	241	243	244	244	244	243	241	234
300	277	287	290	291	291	291	290	287	277
350	318	332	336	338	338	338	336	332	318
400	359	376	382	384	385	384	382	376	359
450	399	420	427	430	430	430	427	420	399
500	437	463	471	475	476	475	471	463	437
600	512	547	559	564	565	564	559	547	512
700	583	629	645	651	653	651	645	629	583
800	650	708	728	737	739	737	728	708	650
900	715	786	810	820	823	820	810	786	715
1000	776	861	890	903	906	903	890	861	776
1200	891	1004	1045	1062	1067	1062	1045	1004	891
1400	997	1141	1193	1216	1222	1216	1193	1141	997
1600	1094	1270	1336	1364	1372	1364	1336	1270	1094
1800	1184	1393	1472	1506	1516	1506	1472	1393	1184
2000	1268	1509	1603	1644	1656	1644	1603	1509	1268
3000	1607	2016	2187	2264	2286	2264	2187	2016	1607
4000	1855	2424	2675	2790	2824	2790	2675	2424	1855
5000	2044	2758	3087	3242	3289	3242	3087	2758	2044
6000	2194	3037	3441	3635	3693	3635	3441	3037	2194
7000	2315	3273	3748	3979	4049	3979	3748	3273	2315
8000	2415	3476	4017	4284	4365	4284	4017	3476	2415
9000	2498	3653	4255	4555	4647	4555	4255	3653	2498
10000	2570	3807	4466	4797	4899	4797	4466	3807	2570
∞	3458	6147	8068	9220	9604	9220	8068	6147	3458

Tabelle 5 d **Statistische Sicherheit** **99 %**

Notwendiger Stichprobenumfang zur Schätzung der Prävalenz einer Krankheit bei einer **Genauigkeit von mindestens 10 %:**

	Vermutete Prävalenz								
N	10%	20%	30%	40%	50%	60%	70%	80%	90%
10	9	10	10	10	10	10	10	10	9
20	15	17	18	18	18	18	18	17	15
30	20	24	25	26	26	26	25	24	20
40	24	30	32	32	33	32	32	30	24
50	28	34	37	39	39	39	37	34	28
60	30	39	42	44	45	44	42	39	30
70	33	43	47	49	50	49	47	43	33
80	35	46	51	54	54	54	51	46	35
90	36	49	55	58	59	58	55	49	36
100	38	52	59	62	63	62	59	52	38
120	40	57	65	69	70	69	65	57	40
140	42	61	70	75	76	75	70	61	42
160	44	64	75	80	82	80	75	64	44
180	45	67	79	85	87	85	79	67	45
200	46	70	83	89	91	89	83	70	46
250	49	75	90	98	100	98	90	75	49
300	50	79	96	105	107	105	96	79	50
350	52	82	100	110	113	110	100	82	52
400	52	84	104	114	118	114	104	84	52
450	53	86	107	118	122	118	107	86	53
500	54	88	109	121	125	121	109	88	54
600	55	91	114	126	130	126	114	91	55
700	56	93	117	130	135	130	117	93	56
800	56	94	119	133	138	133	119	94	56
900	56	95	121	136	141	136	121	95	56
1000	57	96	123	138	143	138	123	96	57
1200	57	98	125	141	146	141	125	98	57
1400	58	99	127	143	149	143	127	99	58
1600	58	100	129	145	151	145	129	100	58
1800	58	101	130	147	152	147	130	101	58
2000	58	101	131	148	154	148	131	101	58
3000	59	103	134	152	158	152	134	103	59
4000	59	104	135	154	160	154	135	104	59
5000	60	104	136	155	161	155	136	104	60
6000	60	105	137	156	162	156	137	105	60
7000	60	105	137	156	163	156	137	105	60
8000	60	105	137	157	163	157	137	105	60
9000	60	105	138	157	163	157	138	105	60
10000	60	106	138	157	164	157	138	106	60
∞	60	107	140	160	166	160	140	107	60

Tabelle 5 e **Statistische Sicherheit** **99 %**

Notwendiger Stichprobenumfang zur Schätzung der Prävalenz einer Krankheit bei einer **Genauigkeit von mindestens 5 %:**

	Vermutete Prävalenz								
N	10%	20%	30%	40%	50%	60%	70%	80%	90%
10	10	10	10	10	10	10	10	10	10
20	19	20	20	20	20	20	20	20	19
30	27	29	29	29	29	29	29	29	27
40	35	37	38	38	38	38	38	37	35
50	42	45	46	47	47	47	46	45	42
60	48	53	55	55	56	55	55	53	48
70	55	61	63	64	64	64	63	61	55
80	60	68	70	72	72	72	70	68	60
90	66	75	78	79	80	79	78	75	66
100	71	81	85	87	87	87	85	81	71
120	80	94	99	101	102	101	99	94	80
140	89	106	112	115	116	115	112	106	89
160	96	117	125	128	129	128	125	117	96
180	103	127	137	141	142	141	137	127	103
200	109	136	148	153	154	153	148	136	109
250	123	158	173	180	182	180	173	158	123
300	133	176	196	204	207	204	196	176	133
350	142	192	215	226	230	226	215	192	142
400	150	206	233	246	250	246	233	206	150
450	157	219	249	264	269	264	249	219	157
500	162	230	264	281	286	281	264	230	162
600	171	249	289	309	316	309	289	249	171
700	179	265	311	334	341	334	311	265	179
800	184	278	329	355	363	355	329	278	184
900	189	289	345	373	382	373	345	289	189
1000	193	299	358	390	399	390	358	299	193
1200	200	314	381	417	428	417	381	314	200
1400	205	326	399	438	451	438	399	326	205
1600	208	336	414	456	470	456	414	336	208
1800	211	344	426	471	485	471	426	344	211
2000	214	351	436	484	499	484	436	351	214
3000	222	372	471	526	544	526	471	372	222
4000	226	384	490	550	570	550	490	384	226
5000	228	392	502	565	586	565	502	392	228
6000	230	397	510	576	598	576	510	397	230
7000	231	401	517	584	607	584	517	401	231
8000	232	404	522	590	613	590	522	404	232
9000	233	406	525	595	618	595	525	406	233
10000	234	408	528	599	623	599	528	408	234
∞	239	425	558	637	664	637	558	425	239

Tabelle 5 f **Statistische Sicherheit** **99 %**

Notwendiger Stichprobenumfang zur Schätzung der Prävalenz einer Krankheit bei einer **Genauigkeit von mindestens 1 %:**

	Vermutete Prävalenz								
N	10%	20%	30%	40%	50%	60%	70%	80%	90%
10	10	10	10	10	10	10	10	10	10
20	20	20	20	20	20	20	20	20	20
30	30	30	30	30	30	30	30	30	30
40	40	40	40	40	40	40	40	40	40
50	50	50	50	50	50	50	50	50	50
60	60	60	60	60	60	60	60	60	60
70	70	70	70	70	70	70	70	70	70
80	79	80	80	80	80	80	80	80	79
90	89	90	90	90	90	90	90	90	89
100	99	100	100	100	100	100	100	100	99
120	118	119	119	120	120	120	119	119	118
140	137	139	139	139	139	139	139	139	137
160	156	158	159	159	159	159	159	158	156
180	175	177	178	178	179	178	178	177	175
200	194	197	198	198	198	198	198	197	194
250	240	245	246	247	247	247	246	245	240
300	286	292	294	295	295	295	294	292	286
350	331	339	342	343	343	343	342	339	331
400	375	386	389	391	391	391	389	386	375
450	419	432	436	438	439	438	436	432	419
500	462	478	483	485	486	485	483	478	462
600	546	568	576	579	580	579	576	568	546
700	627	657	667	671	672	671	667	657	627
800	706	744	757	762	764	762	757	744	706
900	783	830	846	852	854	852	846	830	783
1000	857	914	934	941	944	941	934	914	857
1200	1000	1079	1105	1116	1120	1116	1105	1079	1000
1400	1135	1237	1273	1287	1292	1287	1273	1237	1135
1600	1262	1391	1436	1454	1460	1454	1436	1391	1262
1800	1384	1540	1595	1618	1624	1618	1595	1540	1384
2000	1499	1683	1749	1777	1785	1777	1749	1683	1499
3000	1997	2340	2469	2525	2541	2525	2469	2340	1997
4000	2396	2906	3108	3197	3223	3197	3108	2906	2396
5000	2722	3400	3680	3806	3842	3806	3680	3400	2722
6000	2993	3834	4194	4358	4407	4358	4194	3834	2993
7000	3223	4219	4660	4863	4923	4863	4660	4219	3223
8000	3420	4563	5083	5325	5398	5325	5083	4563	3420
9000	3590	4871	5469	5751	5835	5751	5469	4871	3590
10000	3739	5150	5822	6143	6239	6143	5822	5150	3739
∞	5972	10616	13934	15924	16588	15924	13934	10616	5972

4. Auswahlverfahren

In den vorausgegangenen Abschnitten wurde, soweit es sich um den erforderlichen Stichprobenumfang handelte, immer unterstellt, dass die in die Stichprobe aufgenommenen Tiere auf Grund einer Zufallsauswahl (Randomisierung) bestimmt werden. Ohne strikte Randomisierung lassen sich aus einer Stichprobe keine Rückschlüsse auf die Population ziehen, für welche die Stichprobe repräsentativ sein soll.

Zufällig bedeutet nicht aufs Geratewohl!

Es ist hinlänglich erwiesen, dass es faktisch unmöglich ist, Verzerrungen zu vermeiden, die - bewusst oder unbewusst - entstehen, wenn die Stichprobenauswahl nach anderen Regeln als denen der Randomisierung erfolgt. Die Auswahl einer einwandfreien Zufallsstichprobe setzt daher die strikte Anwendung eines im Voraus festgelegten Verfahrens der Zufallsauswahl voraus.

Nachfolgend werden die fünf gebräuchlichsten Auswahlverfahren beschrieben. Diese werden oft auch in unterschiedlichen Kombinationen angewendet.

4.1. Einfache Zufallswahl *(simple random sampling)*

Eine einfache Zufallsstichprobe ist dadurch charakterisiert, dass jedes Tier der zu untersuchenden Population die gleiche Chance hat, in die Stichprobe zu gelangen. Alle Tabellen in dieser Schrift, die der Bestimmung des notwendigen (kleinsten) Stichprobenumfangs dienen, setzen voraus, dass einfache Zufallsstichproben gezogen werden. Auf Stichproben, die keine einfachen Zufallsstichproben sind, sind diese Tabellen nicht anwendbar; der erforderliche Stichprobenumfang ist in solchen Fällen in der Regel größer.

Man gewinnt eine Zufallsstichprobe, indem man alle Tiere der Population auf irgendeine Weise nummeriert (z. B. anhand von Ohrenmarken-Nummern), diese Nummern auf Zettel schreibt, in einem Hut durchmischt und die geforderte Zahl von Zetteln - wie Lose - aus dem Hut zieht. Die Tiere, deren Nummern auf diese Weise ausgewählt wurden, werden in die Stichproben aufgenommen. Eine bequemere Methode, Zahlen auszuwählen, besteht in der Benutzung eines Zufallszahlengenerators wie er beispielweise in Microsoft® Excel vorhanden ist. Gibt man in eine Excel-Tabelle in einer Spalte die Kennzeichnung eines jeden Tieres ein, fügt eine Spalte mit dem Befehl „=ZUFALLSZAHL()“ in jeder Zeile der Spalte hinzu, die Einträge in der Nachbarspalte mit den Tierkennzeichnungen enthält, und sortiert die gesamte Tabelle nach der Größe der Zufallszahl, so ergeben sich aus der sortierten Reihenfolge die Kennzeichnungen der Tiere, die zu beproben sind Die Anzahl der zu beprobenden Tiere (Zeile bis zu der die Tierkennzeichnungen bzw. die ihnen entsprechenden Tiere in die Stichprobe eingehen) liefert der berechnete Stichprobenschlüssel.

4.2. Systematische Auswahl (*systematic sampling*)

Bei der systematischen Auswahl stellt man sich vor, dass alle Tiere der Population in einer Reihe angeordnet sind (z.B. gemäß ihrer Aufstellung im Stall oder auf Grund einer Durchnummerierung). Die für die Stichprobe bestimmten Tiere werden in ganz bestimmten Abständen aus dieser Reihe ausgewählt. Soll die Stichprobe den n-ten Teil der gesamten

Population umfassen, so wählt man unter den ersten n Tieren der Reihe eines nach Zufall aus und nimmt von da an jedes weitere *n*-te Tier.

Die Größe der Population muss bei Beginn einer systematischen Auswahl nicht bekannt sein. Dies ist insbesondere von Vorteil, wenn Stichprobenerhebungen über prospektive Ereignisse, wie Lebendgeburten, durchgeführt werden sollen. Will man hingegen eine einfache Zufallsstichprobe erheben, die den *n*-ten Teil der Population umfasst, so muss die Populationsgröße - zumindest näherungsweise - bekannt sein.

Die systematische Auswahl ist in der Durchführung einfacher als die einfache Zufallsauswahl; gelegentlich ist sie die einzig praktikable Methode. Man muss jedoch Acht geben, dass in der oben erwähnten Aufreihung der Population nicht zyklische Schwankungen der Erkrankungswahrscheinlichkeit existieren, deren Frequenz zufällig mit der Schrittweite der systematischen Auswahlprozedur übereinstimmt.

Die systematische Auswahl eignet sich besonders gut bei der Qualitätskontrolle von Produktionsserien.

4.3. Geschichtete Auswahl (*stratified sampling*)

Die zu einem landwirtschaftlichen Betrieb gehörige Population einer Tierspezies wird gewöhnlich in getrennten Teilbeständen oder Herden gehalten. Wenn die auszuwählende Stichprobe nur einen geringen Umfang im Vergleich zur Gesamtpopulation haben soll, kann es - durch Zufälle - leicht geschehen, dass bei einfacher Zufallsauswahl aus der Gesamtpopulation einzelne Teilbestände oder Herden unterrepräsentiert sind oder sogar völlig unberücksichtigt bleiben. So ist bei einer Zufallsstichprobe, bestehend aus einer Menge aus Teilbeständen bei der in Deutschland typischen Herdengrößenverteilung zu erwarten, dass Tiere aus großen Teilbeständen unterrepräsentiert sind, da nur eine vergleichsweise kleine Anzahl von sehr großen Teilbeständen existiert. Demgegenüber werden bei einer Zufallsauswahl von Einzeltieren (aus der Gesamtpopulation) im Allgemeinen kleine Teilbestände unterrepräsentiert sein. Dies kann durch ein geschichtetes Auswahlverfahren vermieden werden, welches sicherstellt, dass jede Gruppe (Teilbestand, Bestandsgrößenklasse, etc.) der Population in der aggregierten Stichprobe adäquat vertreten ist. Geschichtete Teilstichprobenumfänge sind proportional zur Größe der jeweiligen Gruppe in der Population zu wählen.

Beispiel: Über einen Bestand, der sich aus mehreren Herden zusammensetzt, sei eine Aussage auf Einzeltierbasis zu treffen. Nach Tabelle 1.b) würde man in einer Stichprobe von 59 Tieren mit einer Sicherheit von 95 % mindestens einen Merkmalsträger finden, wenn 5 % aller Tiere der Population dieses Merkmal tragen. Bei einer reinen Zufallsauswahl kann nicht ausgeschlossen werden, dass alle zufällig ausgewählten Tiere der Stichprobe genau einer Herde angehören, in der der Durchseuchungsgrad möglicherweise deutlich niedriger liegt als im Durchschnitt (5 %) des Bestandes. Um jede Herde adäquat zu berücksichtigen, ist aus einem Bestand mit z. B. 7800 Schafen in vier Herden die Stichprobe von 59 Tieren mit dem Ziel auszuwählen, dass jede Herde entsprechend ihrer Größe anteilig (proportionale Teilstichprobe) in die Gesamtstichprobe eingeht.

Dies wird im folgenden Rechenbeispiel veranschaulicht:

Altersklasse	Anzahl Tiere in der Herde	Stichprobe	x Anteil Tiere an der Population	Anzahl Tiere in der Stichprobe
Mutterschafe	3000	59	x 3000/7800	= 23
weibliche Zuchttiere	800	59	x 800/7800	= 6
Absatzlämmer	2500	59	x 2500/7800	= 19
Böcke (Hammel)	1500	59	x 1500/7800	= 11
	7800			**59**

4.4. Klumpen-Auswahl (*cluster sampling*)

Bei der Klumpen-Auswahl werden ganze Gruppen (cluster) von Individuen geschlossen ausgewählt. Die Gruppen können dabei zufällig oder systematisch ausgewählt werden. Die Klumpen-Auswahl wird oft angewendet, wenn es keine verlässliche Liste aller Individuen in der Gesamtpopulation gibt, aus der eine Zufallsstichprobe genommen werden soll.

Eine Zufallsstichprobe von Gehöften, in denen alle in Betracht kommenden Tiere untersucht werden, ist ein Beispiel für eine Klumpen-Stichprobe. Die Tiere eines Gehöftes stellen einen „Klumpen“ (cluster) dar. In der praktischen Feldarbeit ist eine solche Auswahl einfach zu bewerkstelligen. Jedoch bedarf es großer Sorgfalt, wenn man daraus Schlussfolgerungen über die Krankheitssituation in der Gesamtpopulation aller Tiere zieht - im Unterschied zu Schlussfolgerungen über die Population der Gehöfte, von denen die Stichproben stammen. Die gewonnene Information über die Krankheitssituation in der Gesamtpopulation aller Tiere ist weniger treffsicher als jene, die man aus einer Stichprobe mit derselben Anzahl von Tieren erhalten würde, die auf Grund einer systematischen oder einer streng zufälligen Auswahl aus der Liste aller Einzeltiere entnommen werden. Der Grund dafür ist, dass Krankheitsstatistiken - insbesondere bei Infektionskrankheiten - im Allgemeinen zwischen den Beständen stärker variieren als innerhalb der Bestände und die Menge der Klumpen im Allgemeinen nicht die innere Struktur der Infektion in der Gesamtpopulation widerspiegelt.

Die Klumpen-Auswahl ist anwendbar bei der Qualitätsüberwachung von Produktchargen, wie z. B. bei der Salmonellen-Überwachung von Fleischprodukten.

4.5. Mehrstufige Auswahl (multi-stage sampling)

Bei diesem Verfahren erfolgt die Auswahl der Stichprobe in mehreren Schritten. Ein Beispiel ist, zunächst eine Zufallsstichprobe von Milchkuhbeständen innerhalb eines Bundeslandes auszuwählen und aus jedem der ausgewählten Bestände eine Zufallsstichprobe von Färsen zu ziehen.

Nach Art und Umfang der Stichproben innerhalb der Bestände können verschiedene Stichprobenstrategien charakterisiert werden. Sie unterscheiden sich sowohl im Gesamtumfang der Untersuchung als auch in der späteren Aussagekraft der Ergebnisse. Es können individuelle, proportionale oder feste Stichprobengrößen innerhalb der Bestände zur Anwendung kommen.

Mit diesem Verfahren kann grundsätzlich sowohl die Prävalenz infizierter Tiere in der betrachteten Population als auch die Häufigkeit infizierter Bestände geschätzt werden. Für die Schätzung von Prävalenzen auf der Ebene von Beständen ist eine geeignete Falldefinition für Bestände erforderlich. Bei einer fachlich sinnvollen Ausschöpfung dieses Spielraums ist eine erhebliche Reduzierung des Gesamtumfangs der erforderlichen Untersuchungen möglich.

Die konkrete Berechnung des jeweils notwendigen Stichprobenumfangs bei einer mehrstufigen Auswahl kann sehr aufwändig sein und erfordert spezielle Computerprogramme. In Abhängigkeit von der gewählten Strategie innerhalb der Bestände können noch weitere Informationen, wie etwa die Bestandsgrößenverteilung im untersuchten Gebiet, zur Berechnung erforderlich sein.

5. Güte diagnostischer Tests

5.1. Sensitivität und Spezifität

Bei den meisten Felduntersuchungen hinsichtlich einer bestimmten Krankheit kommen diagnostische (z. B. serologische) Tests zum Einsatz. Bei der Entscheidung, welche Testprozeduren bei umfangreichen Untersuchungen verwendet werden sollen, muss oft ein Kompromiss zwischen verschiedenen Anforderungen gefunden werden. Natürlich soll ein Test genau sein; er soll aber auch einfach, kostengünstig und allgemein eingeführt sein. Andererseits hat es wenig Sinn, umfangreiche Felduntersuchungen anzustellen und große Datenmengen zu erzeugen, wenn nicht bekannt ist, was die Testergebnisse über die wirkliche Krankheitssituation aussagen. Der Schluss auf die Krankheitssituation hängt von den Eigenschaften des verwendeten diagnostischen Tests ab. Diese Eigenschaften können anhand der Begriffe Wiederholbarkeit, Richtigkeit, Sensitivität und Spezifität definiert werden.

Wiederholbarkeit (engl.: repeatability) - manchmal auch als Präzision bezeichnet - ist die Fähigkeit eines Tests, bei wiederholter Anwendung auf dasselbe Objekt zu übereinstimmenden Ergebnissen zu führen. Wiederholbarkeit ist das Gegenstück zu Variabilität.

Richtigkeit (engl.: accuracy) ist die Fähigkeit eines Tests, ein wahres Maß für die getestete Eigenschaft zu liefern. Richtigkeit und Wiederholbarkeit sind nicht dasselbe! Ein Test kann eine hohe Wiederholbarkeit aufweisen, ohne richtig zu sein. Die Richtigkeit hat zwei voneinander unabhängige Komponenten - die Sensitivität und die Spezifität.

Sensitivität und *Spezifität* sind Wahrscheinlichkeiten und sagen deshalb über den einzeln ausgeführten Test nichts aus. Weiterhin ist zu beachten, dass Sensitivität und Spezifität als Güteeigenschaften des Tests mit dem speziellen Anwendungsfall in Zusammenhang stehen können. So können beispielsweise die Güteeigenschaften für ein und denselben Test voneinander abweichen, wenn der Test auf verschiedene Spezies oder in unterschiedlichen Populationen angewendet wird.

Sensitivität ist die Fähigkeit eines Tests, ein positives Ergebnis zu liefern, wenn das untersuchte Tier eine bestimmte Krankheit hat (allgemein ausgedrückt: eine Merkmalsausprägung, beispielsweise eine Krankheit, zu erkennen, wenn sie vorhanden ist). Die Sensitivität wird durch den Anteil der kranken Tiere mit positivem Testergebnis an der Gesamtheit der kranken Tiere charakterisiert (Tabelle 6).

Spezifität ist die Fähigkeit eines Tests, ein negatives Ergebnis zu liefern, wenn das untersuchte Tier die betreffende Krankheit nicht hat (allgemein ausgedrückt: wenn eine bestimmte Merkmalsausprägung, beispielsweise eine Krankheit, nicht vorhanden ist). Die Spezifität wird durch den Anteil der nicht erkrankten Tiere mit negativem Testergebnis an der Gesamtheit der nicht erkrankten Tiere charakterisiert (Tabelle 6).

Ungenügende Sensitivität (< 100 %) führt zu falsch-negativen, ungenügende Spezifität (< 100 %) zu falsch-positiven Testergebnissen.

Es ist schwierig, die Sensitivität und Spezifität eines Tests exakt zu bestimmen, da für eine hinreichend gute Parameterschätzung eine bezüglich der Population repräsentative Anzahl von Individuen benötigt wird, deren wahrer Krankheits- oder Gesundheitsstatus bekannt sein muss. Das Bestreben bei der Testvalidierung sollte jedoch darin bestehen, bei der Schätzung von Sensitivität und Spezifität auf geeignete Sammlungen von repräsentativen Referenzproben mit bekannter Merkmalsausprägung zurückzugreifen.

Ein anderer praktikabler Weg bei der Bestimmung von Sensitivität und Spezifität eines diagnostischen Tests besteht darin, dass man diesen Test mit einem anderen diagnostischen Test vergleicht, dessen Qualität bekannt ist und der daher als eine Art Standard benutzt wird. Man muss sich aber im Klaren sein, dass damit nicht Sensitivität und Spezifität bestimmt werden, sondern die relative Sensitivität und die relative Spezifität (relativ in Bezug auf den Standardtest). Derartige Vergleiche haben nur dann Wert, wenn die Sensitivität und Spezifität des Standardtests bekannt sind und nahe bei 100 % liegen.

Obwohl Sensitivität und Spezifität grundsätzlich voneinander unabhängige Größen sind, führen spezielle Testkonstruktionen (z. B. cut-off-Festlegungen bei optischen Dichtemessungen etc.) in der Regel dazu, dass die beiden Parameter in einem umgekehrten Verhältnis zu einander stehen. Wenn die Interpretation der Testergebnisse mit der Absicht geändert wird, die Sensitivität zu erhöhen (d. h., die Zahl der falsch-negativen Diagnosen herabzusetzen), so bezahlt man dies meistens mit einer Abnahme der Spezifität (d.h. die Zahl der falsch-positiven Diagnosen wird zunehmen).

Die Beziehung zwischen Sensitivität und Spezifität kann anhand der folgenden Tabelle erläutert werden:

Tabelle 6:

	Wahrer Gesundheitszustand		
	erkrankt	nicht erkrankt	Summe
Testergebnis positiv	a	b	a + b
Testergebnis negativ	c	d	c + d
Summe	a + c	b + d	a + b + c + d = N

a = Anzahl kranker Tiere mit positivem Testergebnis (Richtig-Positive)
b = Anzahl gesunder Tiere mit positivem Testergebnis (Falsch-Positive)
c = Anzahl kranker Tiere mit negativem Testergebnis (Falsch-Negative)
d = Anzahl gesunder Tiere mit negativem Testergebnis (Richtig-Negative)
N = Gesamtanzahl der Tiere in der Population

Sensitivität	=	a / (a+c)
Spezifität	=	d / (b+d)
Prävalenz	=	(a+c) / N
Scheinbare Prävalenz	=	(a+b) / N
Voraussagewert einer positiven Diagnose	=	a / (a+b), vergleiche Abschnitt 5.2
Voraussagewert einer negativen Diagnose	=	d / (c+d), vergleiche Abschnitt 5.2

Tabelle 6 stellt ein Modell der Situation dar, in der man sich bei einer diagnostischen Erhebung befindet. Der tatsächliche Gesundheitszustand der N untersuchten Tiere ist unbekannt, er soll mit Hilfe eines diagnostischen Verfahrens bestimmt werden. Somit sind auch die Werte von a, b, c, d sowie von a+c und b+d nicht bekannt. Alles, was in einer diagnostischen Erhebung zahlenmäßig ermittelt werden kann, sind die beiden Randsummen a+b (im Test positive Tiere) und c+d (im Test negative Tiere). Unter Zuhilfenahme der als bekannt vorausgesetzten Sensitivität und Spezifität lassen sich die nicht beobachtbaren Randsummen in der Tabelle ergänzen.

Das für unsere Zwecke notwendige Gleichungssystem ist unter Verwendung der oben angegebenen Symbolik sowie der Berücksichtigung beobachtbarer Größen wie folgt zu formulieren und nach den Unbekannten (a+c) (Anzahl der tatsächlich erkrankten Tiere) bzw. (b+d) (Anzahl der tatsächlich nicht erkrankten Tiere) aufzulösen:

	Anzahl beobachtet		Anzahl wahr krank		Anzahl wahr gesund
Test - positiv	(a+b)	=	Se x (a+c)	+	(1-Sp) x (b+d)
Test - negativ	(c+d)	=	(1-Se) x (a+c)	+	Sp x (b+d)

Beispielsweise berechnet sich die Anzahl der Kranken (a+c), die zur Schätzung der wahren Prävalenz notwendig ist, nach Lösung des oben dargestellten Gleichungssystems wie folgt:

(1) $$(a+c) = \frac{[Sp \times (a+b) - (1-Sp) \times (c+d)]}{[Se \times Sp - (1-Se) \times (1-Sp)]}$$

Diese Gleichung beschreibt den wesentlichen funktionalen Zusammenhang zwischen den beobachtbaren Größen und der tatsächlichen Anzahl von Erkrankten.

Unter Verwendung von N und (a+c) ergibt sich die Anzahl der Gesunden wie folgt:

(2) $(b + d) = N - (a + c)$.

Wenn $a + b = a + c = a$ ist, d.h. alle Test-positiven Tiere tatsächlich erkrankt sind, und wenn $b + d = c + d = d$ ist, d.h. alle Test-negativen Tiere tatsächlich nicht erkrankt sind, so sind zwingend b und c gleich Null, das bedeutet, dass der Test keine falschen Diagnosen liefert. Jedes positive Testergebnis ist eine absolut sichere Diagnose „erkrankt". Ebenso würde jedes negative Testergebnis die sichere Diagnose „nicht erkrankt" bedeuten. Sensitivität und Spezifität des Tests betragen in diesem Falle jeweils 100 %.

Besitzt der diagnostische Test hingegen keine 100%ige Sensitivität bzw. Spezifität, so werden b und c von Null abweichende Werte haben. Auch wenn wir die Werte von b und c im konkreten Fall nicht kennen, so zeigt das Modell doch die Konsequenzen auf. Es ist offensichtlich, dass die beobachtbare Prävalenz $\frac{(a+b)}{N}$ von der wahren - aber unbekannten – Prävalenz $\frac{(a+c)}{N}$ im Allgemeinen verschieden sein wird. Daher wird die beobachtete Prävalenz auch als scheinbare Prävalenz bezeichnet.

Der Sachverhalt lässt sich also wie folgt zusammenfassen: Die (wahre) Prävalenz ist der Anteil der tatsächlich erkrankten Tiere in einer Population. Da erkrankte und nicht erkrankte Tiere als solche nicht direkt beobachtet werden können, sondern mit Hilfe eines im Allgemeinen nicht ganz fehlerfreien Diagnoseverfahrens identifiziert werden sollen, kann die wahre Prävalenz nicht direkt gemessen werden. Unmittelbar beobachtet wird vielmehr die Anzahl der Test-positiven und Test-negativen Tiere. Der Anteil der Test-Positiven an der Gesamtheit getesteter Tiere ist die scheinbare Prävalenz.

Zur Demonstration, in welchem Ausmaß bei einem diagnostischen Test mit unvollkommener Sensitivität und Spezifität die scheinbare Prävalenz von der wahren Prävalenz abweicht, geben wir das folgende (fiktive) Beispiel (Tabelle 7):

Tabelle 7		Wahrer Gesundheitszustand		
		Erkrankt	nicht erkrankt	Summe
Testergebnis	positiv	80	63	143
	negativ	20	837	857
	Summe	100	900	1000

Wahre Prävalenz	= 100/1000	= 10,0 %
Scheinbare Prävalenz	= 143/1000	= 14,3 %
Sensitivität	= 80/100	= 80,0 %
Spezifität	= 837/900	= 93,0 %
Voraussagewert einer positiven Diagnose	= 80/143	= 55,9 %
Voraussagewert einer negativen Diagnose	= 837/857	= 97,7 %

In einem Bestand mit 1000 Tieren erkennt ein serologischer Test mit einer Sensitivität von 80 % und einer Spezifität von 93 % 143 der 1000 Tiere als sero-positiv und 857 als sero-negativ. Die scheinbare Prävalenz beträgt somit 14,3 %. Aus den vorliegenden Diagnoseergebnissen lässt sich die wahre Prävalenz ermitteln. Mit Hilfe der oben angegebenen Formel (1) ergibt sich für die Anzahl der Antikörperträger (a+c):

$$(a+c) = \frac{[Sp \times (a+b) - (1-Sp) \times (c+d)]}{[Se \times Sp - (1-Se) \times (1-Sp)]} = \frac{[0{,}93 \times 143 - (1-0{,}93) \times 857]}{[0{,}8 \times 0{,}93 - (1-0{,}8) \times (1-0{,}93)]} = 100$$

Für die wahre Prävalenz (P) gilt dann:

$$P = \frac{(a+c)}{N} = \frac{100}{1000} = 0{,}1 = 10\ \%$$

In diesem Beispiel wird die wahre Prävalenz von der scheinbaren Prävalenz um 43 % überschätzt. Dieses verdeutlicht die Problematik, die aus einer unkritischen Verwendung der scheinbaren Prävalenz folgt: Letztere kann bei Anwendung anderer Diagnoseverfahren (mit gleichfalls ungenügenden, aber vom Initialtest abweichenden Testgüteparametern) auf die gleiche Stichprobe völlig andere Werte annehmen. Daraus folgt auch, dass wissenschaftlich fundierte Vergleiche der Prävalenzen in verschiedenen Populationen auf Grundlage scheinbarer Prävalenzen nicht zulässig sind. Es müssen zuvor die scheinbaren Prävalenzen in wahre Prävalenzen umgerechnet werden – wie oben dargelegt.
Aufbauend auf der nun bekannten wahren Prävalenz lassen sich die Voraussagewerte einer positiven beziehungsweise negativen Diagnose ermitteln, auf die im folgenden Abschnitt näher eingegangen wird. Hierbei wird nach der Wahrscheinlichkeit gefragt, mit der ein Testergebnis den tatsächlichen Gesundheitszustand widerspiegelt. Nähere Angaben zur Anwendung der diagnostischen Voraussagewerte enthält das Kapitel 5.2.

5.2. Diagnostische Voraussagewerte

Aus den Tabellen 6 und 7 ergaben sich, wie bereits erwähnt, zwei weitere Kenngrößen, die Voraussagewerte einer positiven bzw. negativen Diagnose (positive bzw. negative predictive value). Die Voraussagewerte, auch prädiktive Werte genannt, geben die Wahrscheinlichkeiten an, mit der positive bzw. negative Diagnoseresultate hinsichtlich des wahren Gesundheitszustandes zutreffend sind. Sie korrigieren gewissermaßen den Blick aus dem Labor auf den tatsächlichen Gesundheitszustand einer umschriebenen Population. So wie die Sensitivität und

Spezifität, die als Güteeigenschaften eines diagnostischen Verfahrens unter experimentellen Bedingungen ermittelt werden und damit die Fehler des Verfahrens charakterisieren, so beschreiben die Voraussagewerte die Glaubwürdigkeit von Messungen an Objekten mit unbekanntem Gesundheitsstatus. Die Voraussagewerte entscheiden mit Blick auf festgelegte Ziele über die Sinnhaftigkeit des Einsatzes eines diagnostischen Verfahrens.

Für das in Abschnitt 5.1 gegebene Beispiel lässt sich der Voraussagewert positiver Diagnosen (*PPV*, positive predictive value) wie folgt ermitteln:

$$PPV = \frac{[Se \times P]}{[(Se \times P) + (1 - Sp) \times (1 - P)]}$$

Somit ergibt sich für das Beispiel ein Voraussagewert positiver Diagnosen von

$$PPV = \frac{[0{,}8 \times 0{,}1]}{[(0{,}8 \times 0{,}1) + (1 - 0{,}93) \times (1 - 0{,}1)]} = 0{,}559 = 55{,}9\ \%$$

Das bedeutet, dass unter allen sero-positiv getesteten Tieren nur knapp 56 % tatsächlich Antikörper-Träger sind, der Rest aller sero-positiv getesteten Tiere besteht aus falsch-positiven Diagnosen.

In Bezug auf negative Diagnosen lässt sich analog der Voraussagewert negativer Diagnosen (*NPV*, negative predictive value) definieren. Er errechnet sich wie folgt:

$$NPV = \frac{[Sp \times (1 - P)]}{[(Sp \times (1 - P)) + (1 - Se) \times P]}$$

Somit ergibt sich für das Beispiel ein Voraussagewert negativer Diagnosen von:

$$NPV = \frac{[0{,}93 \times (1 - 0{,}1)]}{[0{,}93 \times (1 - 0{,}1) + (1 - 0{,}8) \times 0{,}1]} = 0{,}977 = 97{,}7\ \%$$

In diesem Beispiel haben also unter allen sero-negativ getesteten Tieren 97,7 % tatsächlich keine Antikörper, die verbleibenden 2,3 % aller sero-negativ getesteten Tiere sind falsch-negative Diagnosen.

Die Voraussagewerte sind nicht allein von den Güteeigenschaften des Tests abhängen, sondern darüber hinaus auch von der Prävalenz der Krankheit bzw. Infektion in der beobachteten Population. Da die Prävalenz häufig unbekannt ist, wird nach Voruntersuchungen oder auf der Grundlage anderer Schlüsse oder Annahmen von einer vermuteten Prävalenz ausgegangen. Typisch sind folgende Grenzübergänge bei imperfekten Tests, d.h. bei diagnostischen Verfahren, deren Sensitivität oder Spezifität kleiner als 100 % ist:

Sinkt die vermutete Prävalenz Richtung Null, so fällt auch der Voraussagewert einer *positiven* Diagnose (*PPV*) sehr rasch gegen Null. Dies bedeutet, dass bei einer Krankheit, die mit einer sehr geringen Prävalenz vorkommt, der größte Teil der (in Abhängigkeit von der Spezifität möglicherweise vielen verbleibenden) positiven Testergebnisse falsch-positiv ist. Wird ein erfolgreiches Bekämpfungsprogramm gegen eine Krankheit durchgeführt, so senkt man die Prävalenz in der Population. Durch den Abfall der Prävalenz kann unter Beibehaltung des Testsystems die oben beschriebene Situation eintreten, dass der Voraussagewert einer positiven

Diagnose (*PPV*) gegen Null fällt und somit die verbleibenden positiven Befunde mit großer Wahrscheinlichkeit falsch-positiv sind. Für den Voraussagewert einer positiven Diagnose ist im Bereich von Prävalenzen nahe Null stets die Spezifität des diagnostischen Testsystems von entscheidendem Einfluss – nicht, wie oft vermutet, die Sensitivität.

Steigt hingegen die vermutete Prävalenz in Richtung 100 %, so fällt der Voraussagewert einer *negativen* Diagnose (*NPV*) steil gegen Null. Nahe einer Prävalenz von 100% ist die Güteeigenschaft Sensitivität die entscheidende Einflussgröße auf den Voraussagewert einer negativen Diagnose. Das Prinzip der Abhängigkeit der diagnostischen Voraussagewerte von der Prävalenz zeigt Abbildung 1:

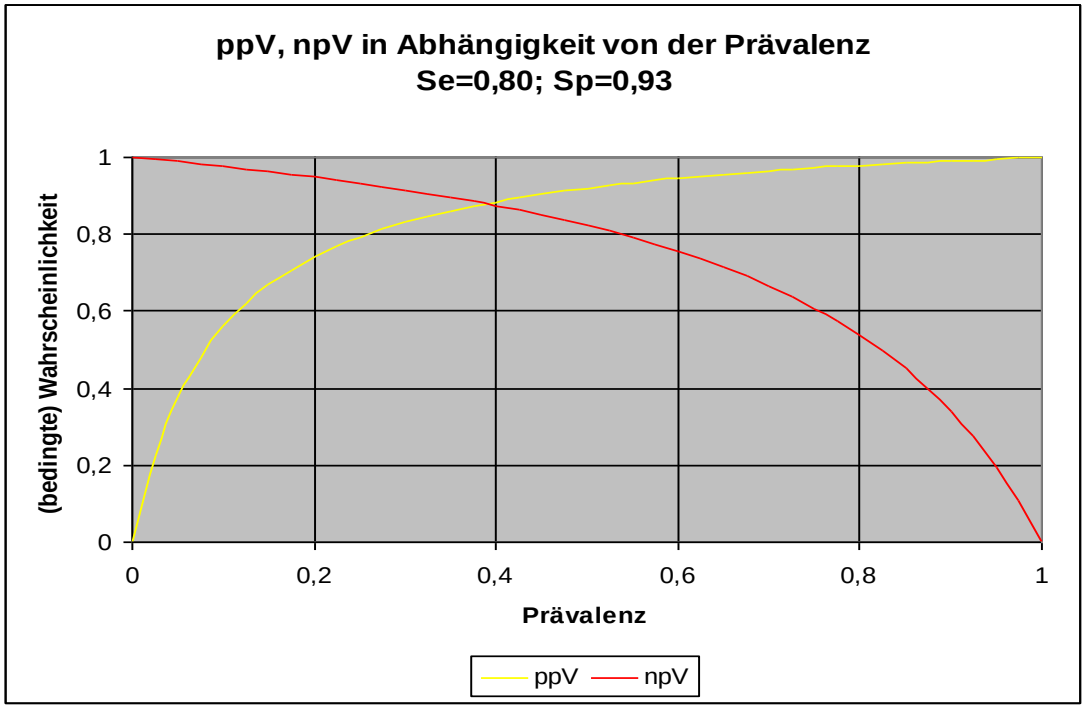

Abbildung 1: Abhängigkeit der Voraussagewerte von der Prävalenz

5.3. Serielles und paralleles Testen (Testsysteme)

Häufig können einzelne diagnostische Verfahren hinsichtlich ihrer Möglichkeit, Aussagen über den Gesundheitszustand von Tieren zu treffen, nicht den gestellten Anforderungen genügen. Aus diesem Grund wird versucht, durch Kombination mehrerer, meist unabhängiger Verfahren die Genauigkeit von Aussagen entsprechend den Erfordernissen zu erhöhen. Die Kombination mehrerer Verfahren soll hier als Testsystem bezeichnet werden. Es ist naheliegend, dass die beteiligten Einzeltests eines Testsystems für ein und dieselbe Probe widersprüchliche Aussagen liefern können. Wird jedoch gefordert, dass das Gesamtresultat eines Testsystems binär (positiv oder negativ; infiziert oder nicht infiziert) sein soll, dann werden zuvor festgelegte Entscheidungsstrategien benötigt, um das Gesamtresultat zu jeder möglichen Kombination aus Einzeltest-Ergebnissen zu ermitteln. Damit muss eine Entscheidungsstrategie stets integraler Bestandteil eines definierten Testsystems sein.

Werden zwei Einzeltests zu einem Testsystem kombiniert (binäres Testsystem), so existieren vier verschiedene Kombinationen aus Einzeltestresultaten (Tabelle 8) und verschiedene Entscheidungsstrategien (Entscheidung für das Urteil positiv bzw. negativ hinsichtlich des

Ergebnisses für das gesamte Testsystem) auf der Grundlage der vorgenannten Kombinationen. Von den möglichen Entscheidungsstrategien sind aber nur zwei sinnvoll und werden häufig verwendet, die „starke Entscheidung“, die das sogenannte „parallele Testen“ beschreibt und die „schwache Entscheidung“, die häufig als das „serielle Testen“ bezeichnet wird (Tabelle 8).

	Einzeltestresultate (T1,T2)			
	- , -	- , +	+ , -	+ , +
Strategie	**Testsystemergebnis**			
„Schwache Entscheidung“	-	-	-	+
„Starke Entscheidung“	-	+	+	+

Tabelle 8: Zwei sinnvolle Entscheidungsstrategien zu einem Testsystem TS, das aus einem geordneten Paar Einzeltests (*T1, T2*) besteht. - bezeichnet ein negatives Einzeltest- bzw. Testsystemresultat, + ein positives Einzeltest- bzw. Testsystemresultat.

Werden zwei an einem Testsystem beteiligte Einzeltests als unabhängig (idealer Weise unabhängig, d. h. möglichst nicht biologisch korreliert wie zwei Antikörpertests, sondern auf verschiedenen biologischen Prinzipien beruhend, z. B. Antigennachweis versus Antikörpernachweis) vorausgesetzt, so lassen sich die charakterisierenden Güteeigenschaften des Testsystems aus den charakterisierenden Güteeigenschaften der beteiligten Einzeltests relativ einfach berechnen. Damit lassen sich Testsysteme, ebenso wie jedes einzelne diagnostische Verfahren, durch ihre Güteeigenschaften beschreiben.

Im Folgenden werde stets die Unabhängigkeit der Einzeltests *T1* und *T2* unterstellt.

Auf ein Testsystem, das aus den beiden Einzeltests *T1* und *T2* besteht, soll die „starke Entscheidung“ angewendet werden. In diesem Fall wird dann, und nur dann, ein positives Testsystemresultat angenommen, wenn mindestens einer der am System beteiligten Einzeltests ein positives Resultat liefert. Diese Strategie hat zur Folge, dass das Testsystem eine höhere Sensitivität gegenüber jedem beteiligten Einzeltest aufweist. Die zugehörige Spezifität fällt jedoch unter den Wert desjenigen Einzeltests mit der schlechteren Spezifität. Damit ist diese Strategie bevorzugt am Beginn von Sanierungsmaßnahmen zu empfehlen, wenn der Durchseuchungsgrad noch relativ hoch ist.

Beispielhaft sei die Berechnung der summarischen Güte des soeben beschriebenen Testsystems *TS* aufgeführt. Hierbei sei $TS = (T1, T2, E = s)$ ein Testsystem, bestehend aus den Einzeltests *T1* mit $G(T1) = (Se1, Sp1)$ und *T2* mit $G(T2) = (Se2, Sp2)$ sowie der starken Entscheidungsstrategie *E=s*. Bezeichnet *G* die Gütefunktion, dann berechnet sich *G(TS)* wie folgt:

$$G(TS) = (Se1 + Se2 - Se1 \times Se2, Sp1 \times Sp2)$$

Eine andere praktizierte Entscheidungsstrategie ist die „schwache Entscheidung“, die dann, und nur dann, das Testsystemresultat als positiv wertet, wenn beide beteiligten Einzeltest-Resultate ein positives Ergebnis zeigen. In allen anderen Fällen wird das Testsystemresultat als negativ gewertet. Diese Strategie hat zur Folge, dass das Testsystem gegenüber allen beteiligten Einzeltests eine höhere Spezifität besitzt, während die Sensitivität des Systems unter die desjenigen Einzeltests mit der kleineren Sensitivität abfällt. Wie bereits erwähnt, bieten sich derartige Testsysteme insbesondere bei niedrigen Prävalenzen an, was am Ende von Sanierungsphasen oder bei der Überwachung der Freiheit von einer Tierseuche der Fall ist.

5.4. Auswirkungen imperfekter Tests auf Stichprobenumfänge

Um eine statistische Aussage mit vorgegebener Präzision treffen zu können, ist stets ein Mindeststichprobenumfang notwendig. Zur Bestimmung von Stichprobenumfängen genutzte Tabellen setzen in der Regel ein perfektes Diagnoseverfahren voraus. Kommen jedoch diagnostische Verfahren zur Anwendung, die nicht in jedem Fall korrekte Resultate liefern, also imperfekt sind, so sind im Vergleich zu perfekten Methoden im Allgemeinen größere Mindeststichprobenumfänge notwendig. Bei sehr „schlechten“ Güteeigenschaften ist es durchaus möglich, dass selbst dann, wenn alle Tiere einer Population getestet werden, die gestellten Erwartungen an den Erkenntnisgewinn nicht erfüllt werden können. Da die Güteeigenschaften eines diagnostischen Verfahrens dieses im statistischen Sinne eindeutig charakterisieren, sind Sensitivität und Spezifität als weitere universelle Parameter bei der Mindeststichprobenberechnung mit zu berücksichtigen. Somit stehen die Größen Sensitivität, Spezifität, Populationsumfang, Konfidenzniveau, Lage und Breite des Konfidenzintervalls der Prävalenz zum Konfidenzniveau als auch Stichprobenumfang und die Anzahl Test-positiver Stichprobenelemente in unmittelbarem Zusammenhang. Die in diesem Dokument vorgestellten Tabellen und Formeln zur Stichprobenberechnung gehen vereinfachend von einem perfekten Test aus und genügen den üblichen Anforderungen in der Tierseuchenbekämpfung. Im Spezialfall müssten die Testgüteparameter erfragt oder bestimmt und die Stichprobengrößen individuell unter Nutzung von Spezialsoftware ermittelt werden.

5.5. Auswirkungen imperfekter Tests auf die Aussagekraft von Schätzwerten

Da imperfekte Messmethoden die wahre Aussage verzerren, lassen sich die klassischen Schätzmethoden für Punktschätzer (wahrscheinlichster Wert) und die zugehörigen Konfidenzintervalle, die gewöhnlich perfekte Messmethoden voraussetzen, beispielsweise für die Schätzung von Prävalenzen aus Stichproben nicht verwenden. Sind jedoch die Güteparameter der Messmethode (d. h. des diagnostischen Verfahrens) bekannt und ist damit die eingangs beschriebene Verzerrung charakterisiert, so kann eine Berechnung der exakten Verteilungsfunktionen unter Einbeziehung von Sensitivität und Spezifität erfolgen, die wiederum eine Ableitung geeigneter Schätzer, beispielsweise für die Prävalenz, gestatten. Der Rechenaufwand für eine Schätzung aus einer konkreten Stichprobe ist in diesem Fall allerdings erheblich, so dass ohne geeignete Computerprogramme praktisch relevante Fragestellungen kaum bearbeitet werden können. Programme, die für große Zahlen innerhalb kürzester Zeit eine Intervallschätzung zurückgeben, verwenden in der Regel Näherungsverfahren, deren Genauigkeit im Vergleich zur exakten Berechnung nur unzulänglich bekannt ist. Bezüglich des Punktschätzers der Prävalenz, d. h. des aus der Stichprobenuntersuchung abgeleiteten wahrscheinlichsten Wertes für die Prävalenz in der Population, ist die Umrechnung der scheinbaren Prävalenz in die wahre Prävalenz allerdings problemlos und wie oben dargestellt (5.1.) möglich, sofern verlässliche Werte für Sensitivität und Spezifität des Testes zur Verfügung stehen.

Zur funktionalen Abhängigkeit geschätzter Parameter von den Güteparametern kann generell Folgendes festgestellt werden: Die geschätzte wahre Prävalenz wächst, wenn die Sensitivität fällt. Fällt hingegen die Spezifität, so fällt auch die geschätzte wahre Prävalenz. Die Konfidenzintervallspanne vergrößert sich generell mit zunehmenden Defiziten hinsichtlich der Testgüte, besitzt aber wie für den perfekten Test ihr Maximum bei 50 % Prävalenz und reduziert sich von ihrem Maximum aus mit anwachsender – bzw. abfallender Prävalenz.