

Stichprobengröße und Probennahme zur PRRS Herdenüberwachung mittels IDEXX Herdcheck ELISA

Fa. IDEXX, 2004

Einleitung

Die Serologie ist ein wichtiges Instrument für das Monitoring von Vakzinen und maternalen Antikörpern, zur Festlegung geeigneter Impfzeitpunkte, Detektion von Infektionen und Feststellung der Erkrankungsprävalenzen. Zur Maximierung der Vorteile dieses Monitoring-Instruments sollten Unternehmen dieses in ein umfassendes, präventivmedizinisches Programm aufnehmen.

Genau definierte Ziele und korrekte Interpretationen von Daten sind notwendig, um wirklich von der Serologie zu profitieren. Zusätzlich zur Information über die Testspezifikationen (z.B. Sensitivität, Spezifität, prädiktiver Wert) und die historischen Daten der Population sollte ein Monitoring-Programm erstellt werden.

Ein Monitoring-Programm hängt von der Stichprobengröße und der Häufigkeit der Probennahme ab. Die häufigsten Bedenken im Feld betreffen die Anzahl der Proben, die genommen werden sollen. Die Festlegung der Mindestprobenzahl ist entscheidend für die Gültigkeit der Ergebnisse.

Aufgrund ihres statistischen Wertes ist eine Stichprobengröße von 30 in der Veterinärmedizin, wie auch auf anderen Gebieten, weit verbreitet. Diese Stichprobengröße wurde in Monitoring-Programmen häufig verwendet, bis zu dem Zeitpunkt, an dem der Kostensenkungsdruck für Unternehmen zu einer Priorität wurde. In der Folge wurde die Zahl der genommenen, serologischen Proben reduziert. Die Kernfragen sind:

- **Wieviel Information geht verloren?**
- **Welche Risiken bringt die Reduktion der Probenzahl?**

In dieser kurzen Diskussion werden wir versuchen die grundlegenden Konzepte der Probennahme zu vereinheitlichen, zeigen warum die Stichprobengröße 30 empfohlen wird und feststellen, welche Implikationen bestehen, wenn mit verschiedenen Stichprobengrößen gearbeitet wird.

Konzepte der Probenziehung

Ein *Stichprobe* ist irgendein Teil einer Population, wohingegen die *Probenziehung* der Prozess der Gewinnung von Proben aus einer Population ist.¹

Der Idee für die Verwendung der Probenziehung bei diagnostischen Tests ist, dass die Gewinnung von relativen Daten einiger Elemente der Population und die entsprechende Analyse derselben, relevante Informationen liefern sollte, die auf die gesamte Population anwendbar sind.

Die Probenziehung ist eng verbunden mit der Grundlage des Prozesses, durch den Antikörper oder Antigene mit äußerster Genauigkeit überprüft werden. Es geht darum, nur einen Teil der Population zu überprüfen, um Schlussfolgerungen für die Gesamtheit zu ziehen, anstatt mit der Gesamtheit zu arbeiten, was dann ein Zensus wäre.

Die Probenziehung beruht auf zwei Prämissen: 1) Die Ähnlichkeit unter den Elementen der Population gestaltet sich so, dass eine bestimmte Anzahl von ihnen die Charakteristika der gesamten Population korrekt repräsentiert. 2) Die Diskrepanzen zwischen den Werten der Populationsvariablen (Parameter) und den Werten der gleichen Variablen aus der Stichprobe (statistisch) sind minimiert, weil einige der Messungen den Wert der Parameter unterschätzen, während andere ihn überschätzen. Wenn die Stichprobe richtig gewonnen wurde, (repräsentative Stichprobe), so neigen die Variationen dieser Werte dazu, sich auszugleichen und sich gegenseitig aufzuheben, was zu Stichproben-Messungen führt, die im Allgemeinen jener der Population nahe kommen.³

Charakteristika einer guten Stichprobe

Das Wesentliche einer guten Stichprobe besteht in der Erstellung eines Mittelwertes, um so genau wie möglich die Charakteristika einer Population von der Messung der Charakteristika in der Stichprobe ableiten zu können. Zu einer guten Stichprobe gehören:

1. **Präzision:** Die Übereinstimmung der Ergebnisse aus der Stichprobe (statistisch) und der entsprechenden Ergebnisse, die in der gesamten Population gemessen werden würden (Parameter). Präzision ist die Maßeinheit des Probenziehungs-Fehlers. Je kleiner der Probenziehungs-Fehler, umso größer die Präzision der Stichprobe.²
2. **Effizienz:** Die vergleichende Messung zwischen verschiedenen Probenziehungs-Projekten. Man sagt, ein bestimmtes Projekt ist unter spezifischen Bedingungen effizienter, als andere, wenn es verlässlichere und wirtschaftlichere Ergebnisse mit der gleichen Präzision oder einer höheren Präzision zu gleichen Kosten liefert. Eine Stichprobe muss präzise und effizient sein ([Abbildung 1](#)).²
3. **Genauigkeit:** Der Grad an Abwesenheit von nicht bei der Probenziehung entstandenen Fehlern in der Stichprobe. Eine Stichprobe wird als genau erachtet, wenn die überschätzten und unterschätzten Messungen sich unter den Komponenten der Stichprobe untereinander ausgleichen.²

Abbildung 1.



Schritte zur Auswahl von Stichproben:

- Schritt 1** : Definieren der gewünschten Population
- Schritt 2** : Bestimmung der Stichproben-Größe
- Schritt 3** : Bestimmung der spezifischen Prozedur der Proben-Auswahl
- Schritt 4** : Ziehung der Proben gemäß der vorgenannten Schritte

Arten der Probenziehung

Es gibt eine Vielzahl von Probenziehungs-Arten, aber es sollte unterschieden werden zwischen zufälligen Auswahlverfahren und nicht-zufälligen Auswahlverfahren.

Zufällige Auswahlverfahren: Jedes Element der Population hat eine bekannte (nichtnull) Chance, als Teil der Stichprobe gewählt zu werden. Dies nennt man auch Zufalls-Probenziehung.

Nicht-zufällige Auswahlverfahren: Die Auswahl der Elemente, die in die Stichprobe einbezogen werden, hängen von der Entscheidung des Prüfers ab.

Man beachte, dass es in jeder Population viele mögliche Stichproben jeder Größe gibt. Wichtig ist, dass **klassische, statistische Inferenz auf dem beruht, was passiert, wenn verschiedene Stichproben der gleichen Größe in der gleichen Population wiederholt ausgewählt werden.**

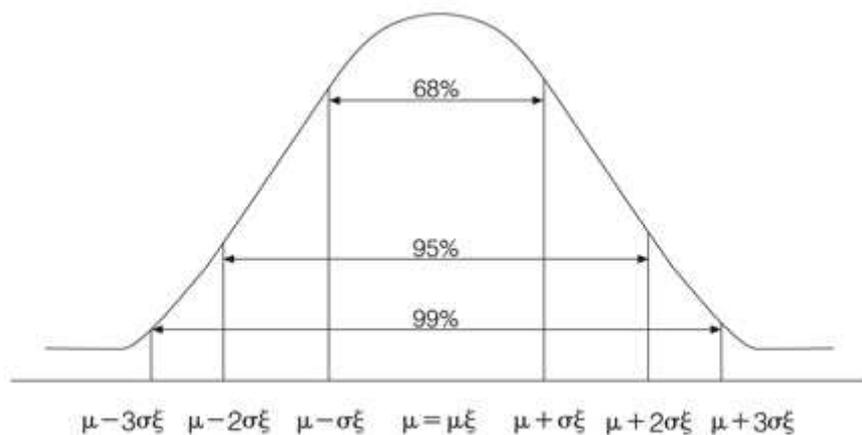
Betrachten Sie die Analyse dreier Probenziehungen von jeweils fünf Proben ($n=5$) aus einer bekannten Population und kalkulieren Sie deren Mittelwerte. Die Mittelwerte werden ähnlich sein, und doch unterschiedlich. Man kann sich vorstellen, dass es drei unterschiedliche Parameter in der Population gibt. Die statistische Theorie empfiehlt jedoch, es nicht bei drei Probenziehungen bewenden zu lassen, sondern weitere Stichproben zu nehmen, bis bestimmte Mittelwerte mit größerer Häufigkeit wiederholt

werden.

Dann wird sich zeigen, dass sich Stichproben-Mittelwerte näher am Populationsmittelwert häufiger wiederholen werden, als jene, die weiter davon entfernt sind. Werden diese Werte auf einem System aus zwei Achsen aufgetragen, entsteht eine Gauß'sche Kurve (Normalkurve). Diese Verteilung der Stichproben-Mittelwerte ist bekannt als die Probenziehungs-Verteilung des Mittelwertes oder *Probenziehungs-Verteilung*.^{1,2}

Symmetrische Intervalle um den wahrscheinlichsten Mittelwert werden Vertrauensintervalle (Konfidenzintervalle) des Mittelwertes genannt. Gemeinhin wird auf drei Intervalle Bezug genommen. Der erste Intervall ist 68%, der zweite 95% und der dritte 99%. [Abbildung 2](#) beschreibt die Gauß'sche Kurve und die entsprechenden Vertrauensintervalle.

Abbildung 2. Bereich unter der Normalkurve für 1,2 und 3 Standardabweichungen vom Mittelwert.²



Zum Verständnis der Bedeutung dieser Intervalle, betrachten Sie den 68%igen Vertrauensintervall. Es ist klar, dass die Verteilung des Stichproben-Mittelwertes dem Mittelwert der Population gleich ist und dass in der Praxis von all den möglichen Stichproben der Größe n nur eine genommen wird und dass ihr Mittelwert als Schätzer des Mittelwertes der Population (der unbekannt ist) verwendet wird. Beachten Sie, dass der Probenziehungs-Mittelwert innerhalb des berechneten Vertrauensintervalls sein kann oder auch nicht, in diesem Falle 68%. Im 68%-Bereich des Vertrauensintervalls zu sein, heißt nicht, dass es für den Probenziehungs-Mittelwert 68 von 100 Chancen gibt, in einen solchen Intervall inkludiert zu sein. Es heißt dagegen, dass, wenn 100 verschiedene Zufalls-Stichproben aus einer Population genommen und für jedes der 68% Vertrauensintervall gebildet würde, zu erwarten ist, dass der Populations-Mittelwert innerhalb von 68 dieser Intervalle fällt.

Stichproben-Grösse und Präzision der Daten

Für gleich große Stichproben gilt, je höher das Vertrauensniveau, desto höher die Präzision. Die Präzision steigt auch, wenn die Anzahl der Elemente in der Stichprobe steigt, aber die Erhöhung der Proben-Anzahl

ist nicht proportional zur Datengenauigkeit.^{1,2}

Es gibt Tabellen, die drei Fehlerkomponenten im Bereich von 1% bis 10% und Vertrauensniveaus von 68%, 95% und 99,7% (Tabelle 1) berücksichtigen.

Tabelle 1: Korrelation von Irrtum (Error), Vertrauensniveau und Anzahl der Elemente für eine Stichprobe unendlicher (infiniter), dichotomischer Populationen ($n > 3000$).²

Error(e)	$n = PQ/e^2$ 68%	$n = 4PQ/e^2$ 95%	$n = 9PQ/e^2$ 99%
$P = Q = 0,50$			
0,01	2500	10000	22500
0,02	625	2500	5625
0,03	278	1112	2502
0,04	156	624	1404
0,05	100	400	900
0,06	70	280	630
0,07	51	204	459
0,08	39	156	351
0,09	31	124	279
0,10	25	100	225

P = Möglichkeit, dass diese Erkrankung eintreten wird

Q = Möglichkeit, dass diese Erkrankung nicht eintreten wird

e = Error

n = Proben-Anzahl

Dies ist eine allgemeine Tabelle, die für Stichproben-Berechnungen von Untersuchungen verschiedenster Arten verwendet werden kann. Gemäß der Tabelle sollten für eine infinite Population mit 9% Irrtum (Error) und einem Vertrauensniveau von 68% in einem zufälligen Auswahlverfahren 31 Proben genommen werden, um das Vorhandensein einer bestimmten Krankheit und/oder einer Impf-Immunantwort in mindestens 68% der Population zu bestimmen.

Probenziehungs-Strategie entsprechend dem Ziel der Untersuchung

In Bezug auf das Ziel der Serologie und der Art der Probenziehung können zwei Situationen

berücksichtigt werden:

Krankheitserkennung und Bestimmung der Krankheitsprävalenz.

Wir konzentrieren uns auf die **Krankheitserkennung** bei geimpften Tieren, eine häufig durchgeführte Untersuchung in Privatlaboratorien.⁵

Die Formel zur Kalkulation der Stichproben-Größe hinsichtlich der Krankheitsprävalenz in einer infiniten Population (>3000) ist:

$$n = \log(\text{Error}) \div \log(1 - \text{Krankheitsprävalenz})$$

Beispiel: Was wäre die Probenziehungs-Strategie zur Erkennung einer bestimmten Krankheit in einer Population mit 10% Prävalenz und 95% Konfidenz (Vertrauen)?

$$n = \log 0,05 \div \log(1 - 0,10)$$

Antwort: $n = 29$

Das heißt, wenn wir ein Minimum von 29 Tieren überprüfen, so haben wir die Chance, mit einem Vertrauenslevel von 95% mindestens ein infiziertes Tier zu entdecken. [Tabelle 2](#) veranschaulicht die nötige Stichproben-Größe zur Erkennung einer Krankheit bei verschiedenen Vertrauens- und Prävalenzniveaus.

Tabelle 2. Geschätzte Stichproben-Größen zur Krankheitserkennung bei Populationen mit einer hohen Anzahl (2000) von Individuen.²

Vertrauen	Prävalenz in Prozent (%)		
	5%	10%	15%
99%	90	44	28
95%	58	28	18
90%	45	22	14
85%	37	18	12
80%	31	15	10
75%	27	13	9

sicherzustellen (das sind 95% Wahrscheinlichkeit der Krankheitserkennung), unter Berücksichtigung einer Population mit einer Infektionsrate von 10%, beträgt 28. Der Unterschied zwischen 28 und 29 Proben kann durch die auf einer Population von 2000 beruhenden Tabelle nachgewiesen werden und dadurch, dass die Kalkulation auf Basis einer Population gleich oder größer als 3000 Individuen durchgeführt worden ist (infini).

Aus [Tabelle 2](#) ist zu entnehmen, dass Sie, wenn die Anzahl der Proben abnimmt, entweder mit einem niedrigeren Prozentsatz an Vertrauen arbeiten oder warten müssen, bis die Krankheit an Prävalenz zunimmt, um erkannt werden zu können. **Die Präventivmedizin hat das Ziel, die Krankheit so früh wie möglich und mit dem höchsten Vertrauens-Index zu erkennen..** Daher ist die empfohlene Stichproben-Größe um eine Infektion mit 10% Prävalenz zu erkennen, bei einem 95%igen Vertrauen, 28.

Im Feld ist die Häufigkeit der Probenziehung eine gebräuchliche Rechtfertigung für die Verwendung einer kleinen Stichproben-Größe. Vorsicht ist in diesem Falle angebracht, da Fehler durch häufigere Probenziehung kleiner Stichproben-Größen nicht korrigiert werden. Nehmen wir etwa an, ein Unternehmen hat sich dafür entschieden, eine Stichproben-Größe von 10 anstatt 29 anzuwenden, in der Hoffnung, 15% Prävalenz in einer Herde der Größenordnung von 2000 zu detektieren. Gemäß oben angeführter [Tabelle 2](#) reduziert diese Entscheidung ihr Vertrauen von 99% auf 80%. Wenn 13 Proben von jedem Haus alle fünf Wochen genommen werden, so wird dies zu einem 75%igen Vertrauen führen, 10% Prävalenz zu detektieren. Eine andere Art, die Konsequenzen einer Reduktion der Stichproben-Größe zu verdeutlichen, ist in der folgende Tabelle erläutert.

Eine andere Art, die Konsequenzen einer Reduktion der Stichproben-Größe zu verdeutlichen, ist in der folgende Tabelle erläutert. [Tabelle 3](#) zeigt die aufgrund der Herdengröße benötigte Anzahl an Proben, um Infektionen bei mindestens einem Tier ($n=1$) in einer Population mit 5% Prävalenz und 95% Vertrauen zu detektieren. Vergleichen Sie den Prozentsatz des Vertrauens, wenn die Anzahl der Proben auf 10 reduziert wird.

Tabelle 3. Anzahl der Proben entsprechend der Populationsgröße, um eine Infektion mit 95% Vertrauen in einer Population mit 5% Krankheitsprävalenz und Grad an Detektionsvertrauen zu erkennen, wenn die Anzahl der Proben auf 10 reduziert wird.⁴

Populationsgröße	Empfohlene Proben-Anzahl	Vertrauen	Reduzierte Proben-Anzahl	Reduziertes Vertrauen
20	19	95%	10	50%

40	31	95%	10	44%
60	38	95%	10	43%
80	42	95%	10	42%
100	45	95%	10	42%
120	47	95%	10	41%
160	49	95%	10	41%
200	51	95%	10	41%
300	54	95%	10	40%
400	55	95%	10	40%
500	56	95%	10	40%
1000	57	95%	10	40%
2000	58	95%	10	40%
>=3000	59	95%	10	40%

Mit nur 10 Proben pro Population sinkt die Wahrscheinlichkeit, irgendeine Infektion zu detektieren, auf 40% (bei Annahme von 5% Prävalenz).

Tabelle 4. Anzahl der Proben zur Erreichung von 95% Wahrscheinlichkeit, ein (1) oder mehr positive Proben in einer infizierten Population zu detektieren.²

Populations Größe	Prävalenz in Prozent (%)											
	50	40	30	25	20	15	10	5	2	1	0.5	0.1
20	4	6	7	9	10	12	16	19	20	20	20	20
30	4	6	8	9	11	14	19	26	30	30	30	30
40	5	6	8	10	12	15	21	31	40	40	40	40
50	5	6	8	10	12	16	22	35	46	50	50	50
60	5	6	8	10	12	16	23	38	55	60	60	60
70	5	6	8	10	13	17	24	40	62	70	70	70
80	5	6	8	10	13	17	24	42	68	79	80	80
90	5	6	8	10	13	17	25	43	73	87	90	90

100	5	6	9	10	13	17	25	45	78	96	100	100
150	5	6	9	11	13	18	27	49	95	130	148	150
200	5	6	9	11	13	18	27	51	105	155	190	200
500	5	6	9	11	14	19	28	56	129	225	349	500
1000	5	6	9	11	14	19	29	57	138	258	450	950
5000	5	6	9	11	14	19	29	59	147	290	564	2253
10000	5	6	9	11	14	19	29	59	148	294	581	2588
∞	5	6	9	11	14	19	29	59	149	299	596	2995

Schlußfolgerung

Die Grundlagen der Stichprobenziehung sollten bekannt sein, bevor ein Monitoring-Programm etabliert wird. Vor der Änderung der Stichproben-Größe sollte stets Einsicht in statistische Tabellen genommen werden. Wenn ein Unternehmen beschließt, die Kosten zu senken, so sollte man sich im Klaren sein, dass die Reduktion der Probenzahl zu einem Verlust an Information und zu einer potentiellen Fehlinterpretation der Ergebnisse führen könnte, was zu noch höheren Produktionskosten führen kann.

Der Vorteil der Serologie, vor allem von ELISA, ist die Krankheits-Früherkennung. Die Verwendung von 23 Proben (zwischen 29 und 19) für geimpfte Tiere, die zu 95% Vertrauen (oder Wahrscheinlichkeit) bei der Krankheitserkennung mit 10–15% Prävalenz führt, ist äußerst empfehlenswert. Die Proben-Anzahl für die Detektion von Krankheiten mit langsamer Übertragung und geringer Prävalenz (<5%) kann aus Tabelle 4 entnommen werden.

Literaturhinweise

1. Bhattacharya GK, et al. *Statistical Concepts and Methods*. New York: John Wiley & Sons. 1977.
2. Martin SW, Meek AH, Willeberg P. *Veterinary Epidemiology—principles and methods*. Ames, Ia: Iowa State University Press, 1987.
3. Mattar NF. Pesquisa de Marketing. Amostragem, confiança e numero de elementos da amostra. *Editores Atlas Sao Paulo*. 2001.
4. Ridpath HD. Personal communication, 1994.
5. Snelson H. Serology as an aid to disease diagnosis and vaccine management. Technical Services Update, Schering-Plough Animal Health. 2003