

Homburger Kolloquium – Tierschutz und Versuchstierkunde | 26.4.2018

Fallzahlplanung für Tierversuche: Biometrische Ansätze zur Lösung grundlegender Probleme

PD Dr. Benjamin Mayer

Institut für Epidemiologie und Medizinische Biometrie, Universität Ulm
benjamin.mayer@uni-ulm.de

Übersicht

- Rechtliche und methodische Aspekte
- Notwendige Vorinformationen zur Fallzahlplanung
- Probleme der Ergebnisvalidität
- Biometrische Ansätze zur Lösung dieser Probleme
- Fazit für die Praxis

Rechtliche und methodische Aspekte

- Sowohl die deutsche (Tierschutz-Versuchstierverordnung), als auch die europäische Gesetzgebung (EU-Direktive 2010/63) fordern eine biometrische Begründung der für Tierversuche eingesetzten Tierzahl
- In Deutschland kann diese Begründung entweder über ein biometrisches Gutachten oder über ein entsprechendes Formblatt bei der Antragstellung gegeben werden
- Ziel: so wenige Tiere wie möglich einschließen einerseits (um dem Reduction-Aspekt der 3R gerecht zu werden), andererseits muss jedoch eine ausreichend große Fallzahl eingeschlossen werden, um tatsächlich bestehende Effekte auch entdecken zu können (ausreichende statistische Power)
- Problem: aufgrund verschiedener Umstände sind die Ergebnisse einer statistischen Fallzahlplanung bei Tierversuchen oftmals in ihrer Validität eingeschränkt

Rechtliche und methodische Aspekte

- Im Wesentlichen umfassen die Überlegungen zur statistischen Methodik in Tierversuchen die Fragen...
 1. Welche Zielgrößen werden im Rahmen des Projektes erhoben?
 2. Welche Methoden müssen für die Auswertung dieser Zielgrößen verwendet werden?
 3. Welche Methoden zur Fallzahlschätzung müssen daher angewandt werden?

- Die Reihenfolge dieser formulierten Fragen ist keineswegs willkürlich gewählt, sondern spiegelt vielmehr das praktische Vorgehen wieder

- Die Art der zu analysierenden Variablen bestimmt das angemessene Auswertungsverfahren, welches wiederum maßgebend für die Methode zur Fallzahlschätzung ist (Planungs- und Analyseverfahren müssen übereinstimmen)

Notwendige Informationen für die Fallzahlplanung

- ...was waren nochmal die essentiellen Informationen, die für eine Fallzahlplanung zur Verfügung stehen sollten?
 - Einschätzung des zu erwartenden/klinisch/biologisch relevanten Effektes (z.B. Mittelwertsunterschiede, Anteilsunterschiede)
 - Im Falle stetiger (messbarer) Variablen: Einschätzung der zu erwartenden Streuung in den Daten
 - Grundlegende Annahmen aus statistischer Sicht (Typ I Fehler, Art der Hypothese, statistische Power)
 - Anzahl der zu untersuchenden Faktoren
 - Logistische/finanzielle/ethische Rahmenbedingungen
- Eine formale Planung kann also eigentlich nur dann durchgeführt werden, wenn entsprechende Vordaten verfügbar sind

Planungsdaten

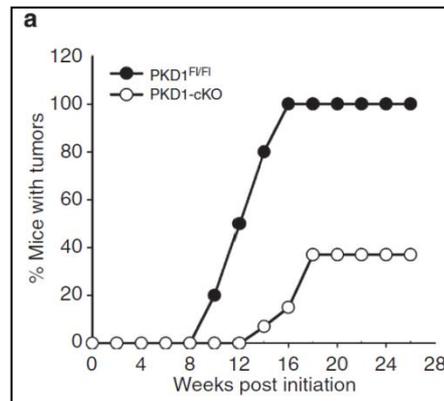
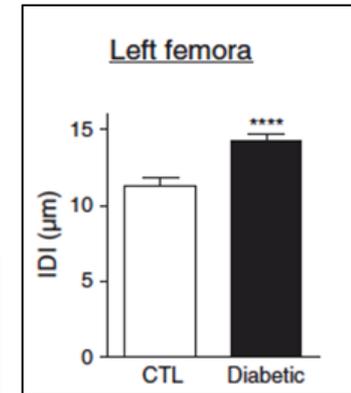
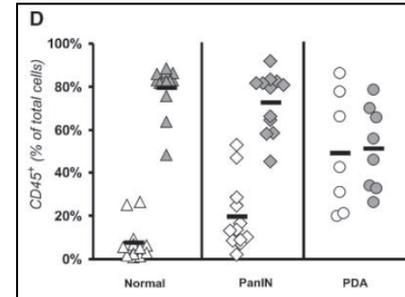
- Informationen können folgenden Quellen entstammen
 - Eigenen Vorarbeiten
 - Publizierten Studien zu ähnlichen Forschungsinhalten

- Sofern keine konkreten Werte verfügbar sind, müssen aus inhaltlicher Sicht relevante Gruppenunterschiede definiert werden (z.B. ein Mittelwertunterschied von 30mg/dl wäre relevant)

- Falls tatsächlich keinerlei Vordaten (eigene oder publizierte) existieren, deren Übertragbarkeit für OK befunden wird, bleibt als Alternative nur der Pilotversuch

Vordaten? Wo sollen die herkommen?

Biomechanik (3-Punkt-Biegung): Hauptzielgröße Biegesteifigkeit			
Altes Projekt #1			
	Biegesteifigkeit		Biegesteifigkeit
Kontrolle	re/li in%	Thorax-Trauma	re/li in%
Maus 1	141,14	Maus 1	110,68
Maus 2	114,60	Maus 2	63,33
Maus 3	117,72	Maus 3	68,60
Maus 4	164,07	Maus 4	82,50
Maus 5	107,86	Maus 5	102,16
Maus 6	94,66	Maus 6	137,26
Maus 7	118,77	Maus 7	81,68
		Maus 8	67,82
Mittelwert	122,69	Mittelwert	89,25
SD	22,96	SD	25,61



IL-6		
3h nach Fraktur	Fraktur (n=8)	Fraktur + TXT (n=6)
MW	46,53	137,64
SD	9,92	68,25

„...aus der Abbildung in der Ghilardi-Studie komm ich in der Vehicle-Gruppe auf eine kalzifizierte Kallusfläche (mm²) von 6,2±0,4 und in der Behandlungsgruppe auf 8,4±0,6. Die Ergebnisse sind in MW±SEM dargestellt...“

Einige praktische Probleme

- Die hauptsächlichen Probleme in der Praxis sind:
 - Eingeschränkte Übertragbarkeit (Vordaten basieren auf anderen Mauslinien, Erhebungszeitpunkten, Dosisstufen, Knock-out-Genen, etc.)
 - Fehlen entsprechend verwertbarer Vordaten
 - Unklarheit darüber, ob bestimmte statistische Voraussetzungen erfüllt sind (z.B. Normalverteilung)
 - Unklarheit darüber, wie mit multiplen Endpunkten (im Tierversuch fast immer der Fall) umgegangen werden soll => explorative vs. konfirmatorische Hypothesentests
 - Oftmals keinerlei Vordaten verfügbar: wie also soll die Fallzahl für Pilotversuche sinnvoll festgelegt werden?
 - Was muss ich denn in die Formblätter eintragen?

Biometrische Lösungsansätze

- Drei Kernprobleme aus methodischer Sicht:
 - A) Validierung des verfügbaren Vorwissens:
Resampling als vielversprechender Ansatz bei Stichproben mit kleinen Fallzahlen
 - B) Berücksichtigung des multiplen Testproblems:
Gatekeeping als anschaulicher und flexibler Ansatz zum Umgang mit zahlreichen Endpunkten
 - C) Festlegung der Fallzahl in Pilotversuchen:
 - (i) Markov Chain Monte Carlo Simulation zur Erzeugung von Zufallsstichproben, die fehlende Vordaten ersetzen ODER
 - (ii) „Resource Equation Approach“

A) Resampling - Die Jackknife-Methode

- “Jackknife is promising since it can provide a less biased estimate of an originally biased parameter of interest” (Gladen, 1979)
- Ein interessierender Parameter e (z.B. mean oder SD) kann geschätzt werden auf der Basis einer Gleichung, die ausschließlich auf der Berechnung von Mittelwerten von Teilmengen der Stichprobe basiert (Manly, 2007):

$$\hat{e}^* = \frac{\sum_{i=1}^n \hat{e}_i^*}{n}$$

- Diese Gleichung berechnet das arithmetische Mittel einer Menge von Werten $\hat{e}_i^* = n \cdot \hat{e} - (n-1) \cdot \hat{e}_{-i}$, $i=1, \dots, n$, wobei $\hat{e} = \hat{e}(y_1, \dots, y_n)$ eine beliebige Funktion der Stichprobenwerte y_1, \dots, y_n ist, und \hat{e}_{-i} die betreffende Funktion darstellt unter Ausschluss des Elementes i der Stichprobe

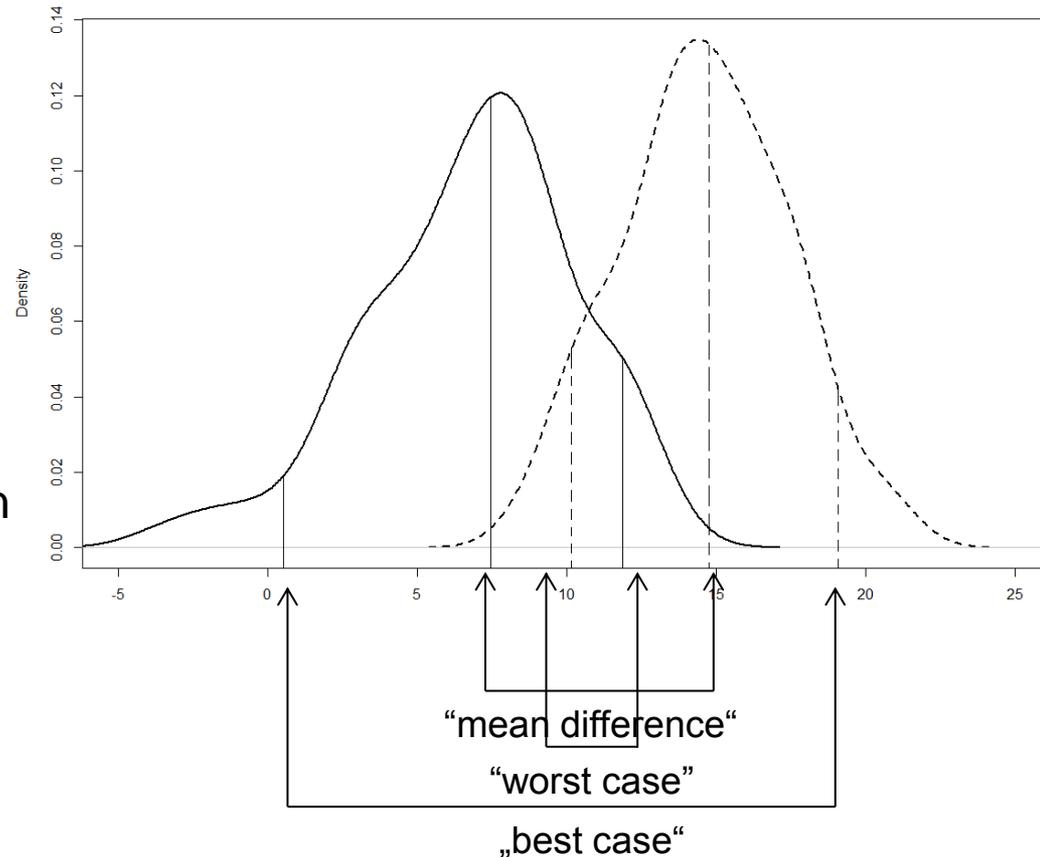
A) Resampling – Der Bootstrapping-Ansatz

- “A random sample of size n from an unknown population is the best guide to the distribution in the population” (Manly, 2007)
- Sei $x=(x_1, \dots, x_n)$ eine Stichprobe einer zugrunde liegender Population mit unbekannter Verteilungsfunktion F
- Bootstrapping bedeutet nun: zufälliges Ziehen von den beobachteten (Stichproben-)Daten x
- Genereller Algorithmus zur Schätzung eines interessierenden Verteilungsparameters (z.B. mean):

- Generiere Bootstrap-Stichproben x_1, \dots, x_B , $x_b=(x_1^*, \dots, x_n^*)$, $b=1, \dots, B$ durch zufälliges Ziehen mit Zurücklegen von x
- Bestimme für jede Bootstrap-Stichprobe den Parameter θ^*
- Die empirische Verteilung von $\theta_1^*, \dots, \theta_B^*$ approximiert die (wahre) Verteilung von θ

A) Resampling – Der Einsatz bei Fallzahlplanung

- Generiere Jackknife/Bootstrap-Stichproben der zur Verfügung stehenden Originaldaten
- Nutze die empirischen Verteilungen, um die Fallzahlplanung durchführen zu können
- Ziel: Erfassung der Sensitivität der Ergebnisse der Fallzahlplanung gegenüber Abweichungen der ursprünglichen Annahmen



$$\text{Effektgröße: } \left| q_{0.05 \cdot i}^{Exp} - q_{1 - (0.05 \cdot i)}^{Con} \right|$$

A) Resampling – Resultate der Simulationen

BOOTSTRAPPING

JACKKNIFE

Sample size per group		SD			
		q _{0.05}	q _{0.5}	q _{0.95}	
Best case scenario	i=1	3	4	5	
	i=2	3	4	6	
	i=3	3	4	6	
	i=4	3	5	6	
	i=5	4	5	7	
	i=6	4	5	7	
	Effect	i=7	4	5	7
		i=8	4	5	7
		i=9	4	6	8
		i=10	4	6	8
		i=11	4	6	9
		i=12	4	6	9
		i=13	5	7	9
		i=14	5	7	10
		i=15	5	7	11
		i=16	5	8	12
	Worst case scenario	i=17	6	9	13
		i=18	7	10	15
		i=19	8	12	18

Sample size per group		SD			
		q _{0.05}	q _{0.5}	q _{0.95}	
Best case scenario	i=1	2	4	6	
	i=2	2	4	6	
	i=3	2	4	7	
	i=4	2	5	8	
	i=5	3	5	9	
Effect	i=6	3	6	10	
	i=7	3	6	11	
	i=8	3	7	12	
	i=9	3	7	13	
	i=10	3	8	14	
	i=11	4	9	16	
	i=12	4	10	19	
	i=13	4	12	21	
	i=14	5	14	26	
	i=15	6	17	32	
	i=16	7	22	41	
	i=17	9	30	57	
	Worst case scenario	i=18	14	51	91
		i=19	41	226	256

Beispiel 1:
n=6

Sample size per group		SD			
		q _{0.05}	q _{0.5}	q _{0.95}	
Best case scenario	i=1	5	5	6	
	i=2	5	6	6	
	i=3	5	6	6	
	i=4	5	6	6	
	i=5	6	6	6	
	i=6	6	6	6	
	Effect	i=7	6	6	6
		i=8	6	6	6
		i=9	6	6	6
		i=10	6	6	6
		i=11	6	6	6
		i=12	6	6	6
		i=13	6	6	7
		i=14	6	6	7
		i=15	6	6	7
		i=16	6	7	7
	Worst case scenario	i=17	6	7	7
		i=18	6	7	7
		i=19	6	7	7

Sample size per group		SD			
		q _{0.05}	q _{0.5}	q _{0.95}	
Best case scenario	i=1	5	7	7	
	i=2	5	7	8	
	i=3	5	8	9	
	i=4	5	8	9	
	i=5	6	8	9	
Effect	i=6	6	8	10	
	i=7	6	9	10	
	i=8	6	9	11	
	i=9	7	10	11	
	i=10	7	10	11	
	i=11	7	10	11	
	i=12	7	10	12	
	i=13	7	11	12	
	i=14	8	11	13	
	i=15	8	12	14	
	i=16	9	13	15	
	i=17	9	14	17	
	Worst case scenario	i=18	10	16	18
		i=19	11	17	20

Beispiel 2:
n=10

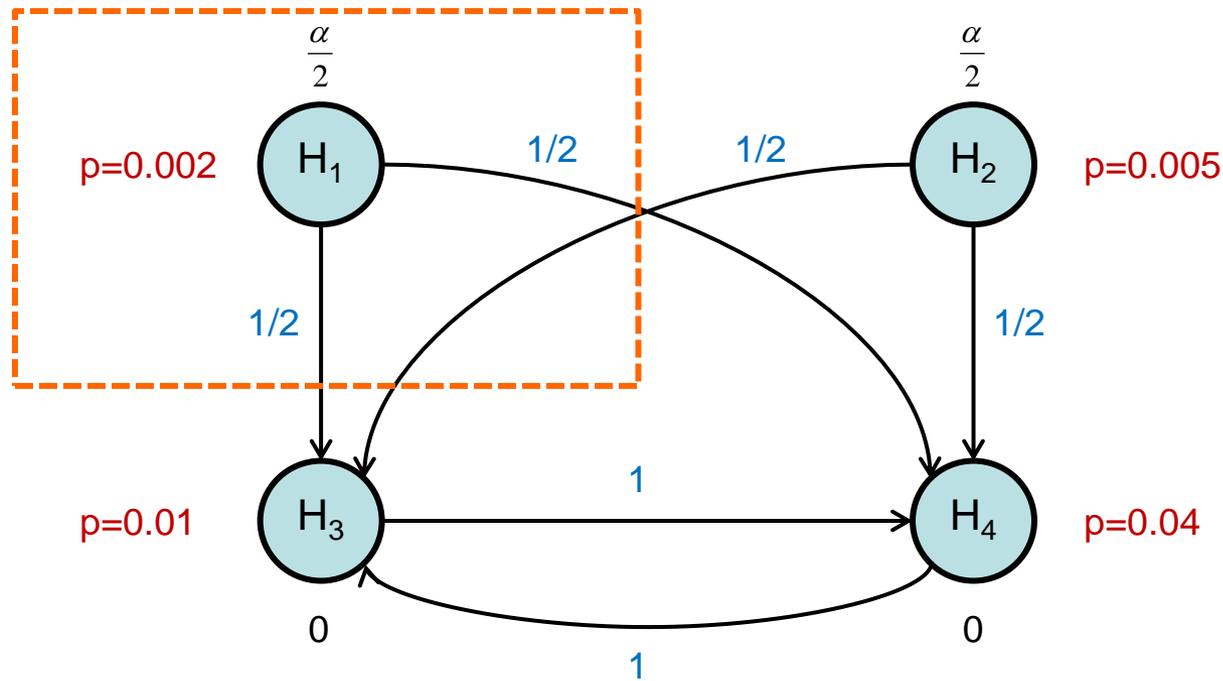
$$|q_{0.05}^{Exp} - q_{1-(0.05 \cdot i)}^{Con}|$$

B) Gatekeeping - Kernidee

- Unter der Annahme, dass konfirmatorische Aspekte in Tierversuchen überwiegen, muss das so genannte multiple Testproblem angemessen berücksichtigt werden
- Wahrscheinlichkeit für mindestens ein falsch signifikantes Testergebnis im Falle von k zu testenden Hypothesen (Zielgrößen) ist $1-(1-\alpha)^k$, unter der Annahme dass alle Nullhypothesen wahr sind
- Verschiedene Ansätze zur Kontrolle des α -Niveaus wurden entwickelt, z.B. Bonferroni, viele davon sind jedoch zu konservativ mit der Konsequenz höherer Fallzahlen
- Gatekeeping:
 - Prinzip der α -Umverteilung
 - Manche Hypothesen fungieren als “Wächter” für nachfolgende
 - Vorgehen: weise das lokale α -Level einer abgelehnten Hypothese einer oder mehrerer verbleibender Hypothesen zu gemäß einer vorab festgelegten Regel (welche Kanten? welche Gewichte?) und fahre mit dem Testen fort

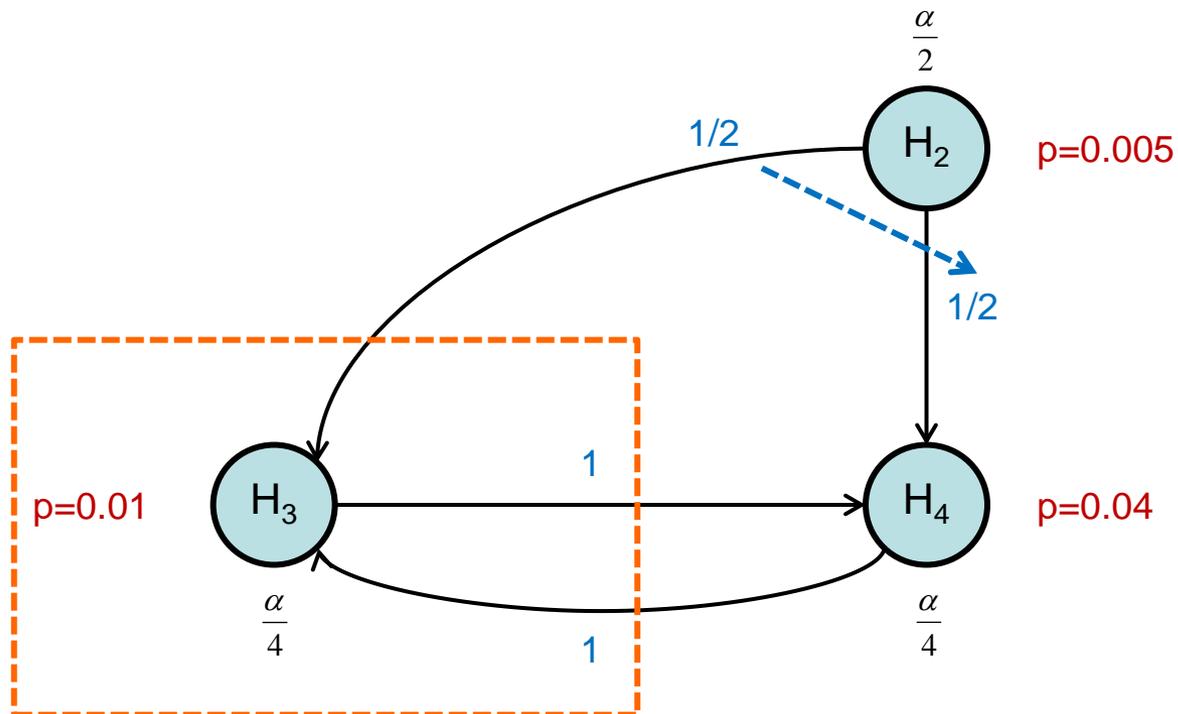
B) Gatekeeping- Beispiel

- $H = \{H_1, H_2, H_3, H_4\}$, $\mathbf{p} = \{0.002, 0.005, 0.01, 0.04\}$, $\boldsymbol{\alpha} = \{\alpha/2, \alpha/2, 0, 0\}$ mit $\alpha = 5\%$ (global)



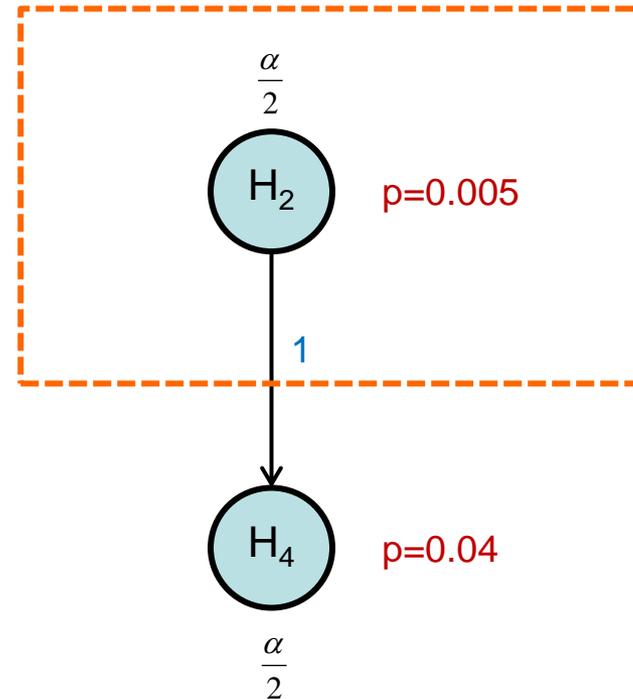
B) Gatekeeping- Beispiel

- $H = \{H_1, H_2, H_3, H_4\}$, $\mathbf{p} = \{0.002, 0.005, 0.01, 0.04\}$, $\boldsymbol{\alpha} = \{1/2, 1/2, 0, 0\}$



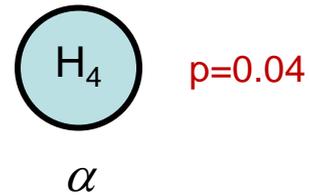
B) Gatekeeping- Beispiel

- $H = \{H_1, H_2, H_3, H_4\}$, $\mathbf{p} = \{0.002, 0.005, 0.01, 0.04\}$, $\boldsymbol{\alpha} = \{1/2, 1/2, 0, 0\}$

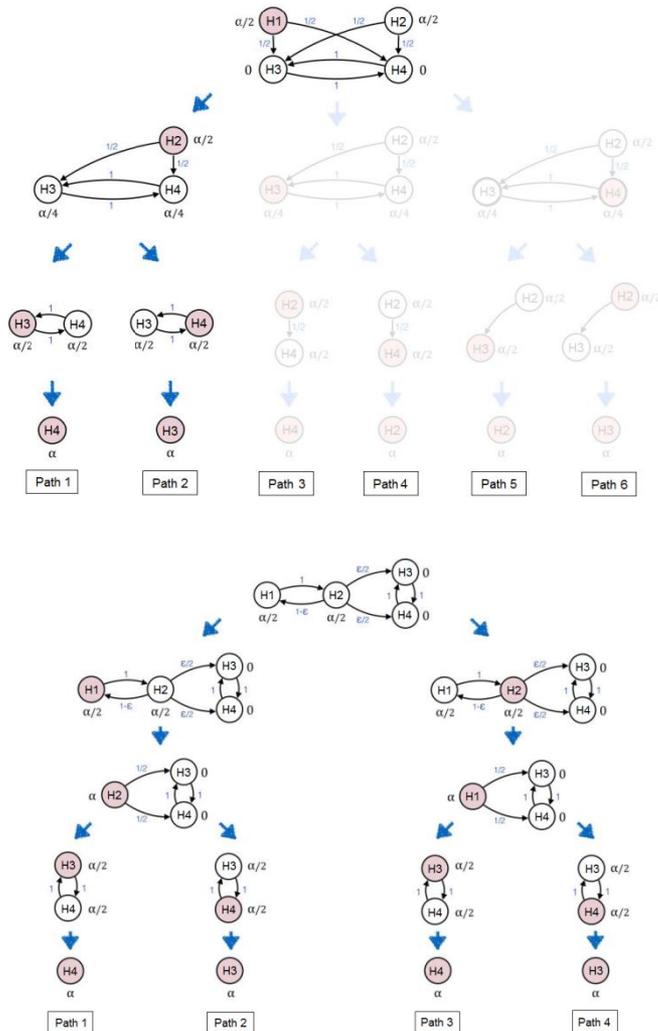


B) Gatekeeping- Beispiel

- $H = \{H_1, H_2, H_3, H_4\}$, $p = \{0.002, 0.005, 0.01, 0.04\}$, $\alpha = \{1/2, 1/2, 0, 0\}$



B) Gatekeeping – Resultate der Simulation



	No adjustment of α				Bonferroni				Parallel gatekeeping approach							
									Path 1				Path 2			
	H ₁	H ₂	H ₃	H ₄	H ₁	H ₂	H ₃	H ₄	H ₁	H ₂	H ₃	H ₄	H ₁	H ₂	H ₃	H ₄
effect size ²	1.4	2.5	1.3	4.9	1.4	2.5	1.3	4.9	1.4	2.5	1.3	4.9	1.4	2.5	1.3	4.9
local α	α	α	α	α	$\alpha/4$	$\alpha/4$	$\alpha/4$	$\alpha/4$	$\alpha/2$	$\alpha/2$	$\alpha/2$	α	$\alpha/2$	$\alpha/2$	α	$\alpha/2$
n	10	5	12	7	14	6	17	10	12	5	15	7	12	5	12	9
N total	12				17 (+42%)				15 (+25%)				12 (+0%)			
Serial gatekeeping approach																
	Path 1				Path 2				Path 3				Path 4			
	H ₁	H ₂	H ₃	H ₄	H ₁	H ₂	H ₃	H ₄	H ₁	H ₂	H ₃	H ₄	H ₁	H ₂	H ₃	H ₄
effect size ²	1.4	2.5	1.3	4.9	1.4	2.5	1.3	4.9	1.4	2.5	1.3	4.9	1.4	2.5	1.3	4.9
local α	$\alpha/2$	α	$\alpha/2$	α	$\alpha/2$	α	α	$\alpha/2$	α	$\alpha/2$	$\alpha/2$	α	α	$\alpha/2$	α	$\alpha/2$
n	12	5	15	7	12	5	12	9	10	5	15	7	10	5	12	9
N total	15 (+25%)				12 (+0%)				15 (+25%)				12 (+0%)			

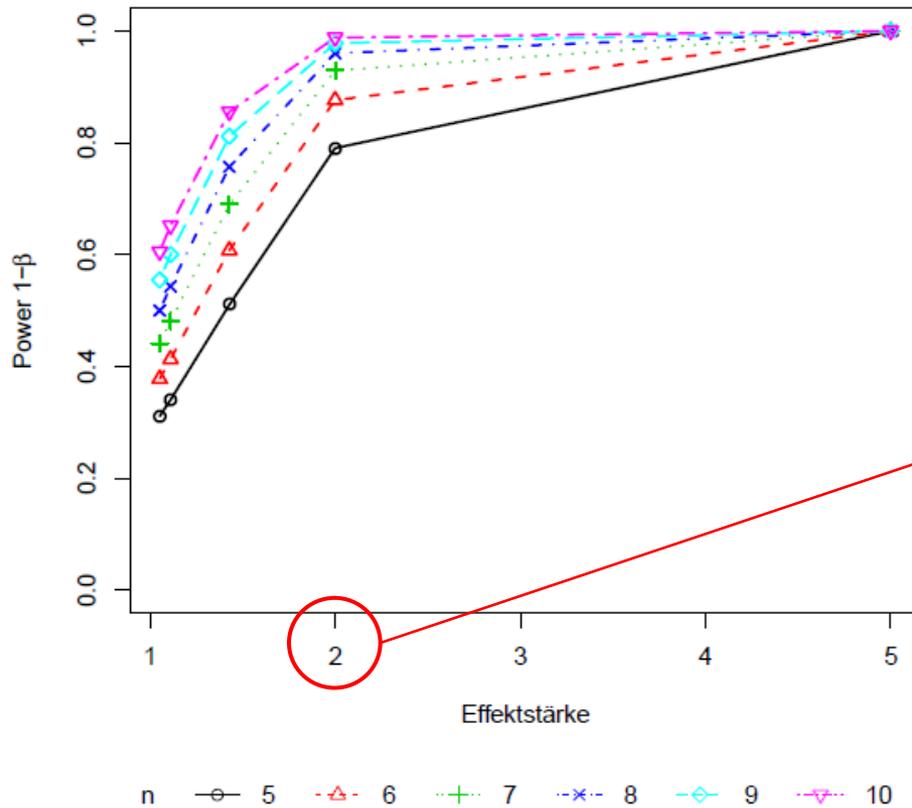
C) Pilotversuche – MCMC Simulation

- Markov Chain Monte Carlo (MCMC)-Methoden werden genutzt, um Zufallsstichproben auf Basis von Verteilungsannahmen zu simulieren
- Grundlage für die Simulationen bilden verschiedene Algorithmen, wie beispielsweise der Metropolis-Hastings-Algorithmus
- Zur Abschätzung der notwendigen Fallzahlen bei Pilotversuchen, in denen aufgrund nicht verfügbarer Vordaten keine Schätzung mit den üblichen Formeln gemacht werden kann
- Im Rahmen der Simulationen wurden verschiedene Verteilungsannahmen getroffen, um häufige Situationen in der Praxis abbilden zu können:
 - Simulation einer Binomialverteilung für kategoriale Zielgrößen
 - Simulation einer trunkeierten Normalverteilung für stetige Zielgrößen

C) Pilotversuche – MCMC Simulation

- Zudem wurden verschiedene Algorithmen (Random Walk sampler vs. Independent sampler) evaluiert, die auf der Gleichverteilung bzw. trunkierten Normalverteilung basierten
- Notwendig für die Durchführung der Simulationen war eine Festlegung der nachzuweisenden Effekte
 - Für kategoriale Endpunkte: $p_1=1/n$ vs. $p_2=n-1/n$
 - Für stetige Endpunkte: $\Delta = \mu_2 - \mu_1 = \{0.3, 0.5, 0.7\}$ mit einer Standardabweichung σ von 20, 50, 70, 90 bzw. 95% von Δ
- Fallzahlplanung basierte auf
 - χ^2 -Test mit nicht-zentraler χ^2 -Verteilung oder Normalverteilung bzw. dem exakten Test nach Fisher für kategoriale Endpunkte und
 - t-Test für Parallelgruppen und gleichen Varianzen mit nicht-zentraler t-Verteilung bzw. entsprechender Approximationsformel für stetige Endpunkte

C) Pilotversuche – MCMC Simulation

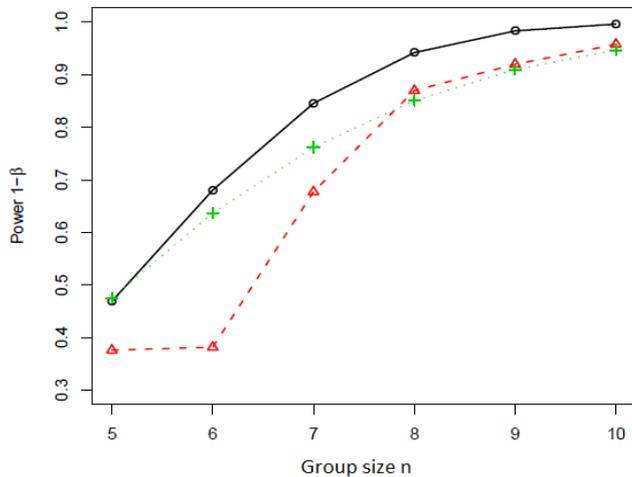


Kleinstes $\Delta=0.3$ mit moderater Streuung von 50%

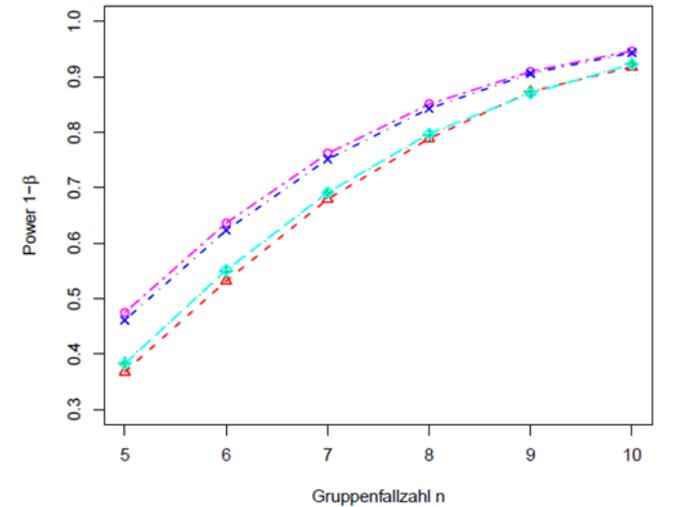
n=5-6 als
Empfehlung

C) Pilotversuche – MCMC Simulation

		n					
		5	6	7	8	9	10
Annahme	p_1	0.20000	0.16667	0.14286	0.12500	0.11111	0.10000
	p_2	0.80000	0.83333	0.85714	0.87500	0.88889	0.90000
	Δ	0.60000	0.66667	0.71429	0.75000	0.77778	0.80000
uis	\hat{p}_1	0.24449	0.20526	0.17601	0.15431	0.13814	0.12584
	\hat{p}_2	0.75748	0.79432	0.82452	0.84497	0.86959	0.87536
	$\hat{\Delta}$	0.51299	0.58906	0.64851	0.69066	0.73147	0.74953
nis	\hat{p}_1	0.23666	0.19893	0.17156	0.15194	0.13459	0.12159
	\hat{p}_2	0.76315	0.80076	0.82851	0.84854	0.86533	0.87864
	$\hat{\Delta}$	0.52649	0.60183	0.65695	0.69660	0.73074	0.75705
urw	\hat{p}_1	0.20582	0.17015	0.14722	0.12911	0.11131	0.10327
	\hat{p}_2	0.79478	0.82687	0.85270	0.87075	0.88400	0.89670
	$\hat{\Delta}$	0.58896	0.65673	0.70548	0.74164	0.77270	0.79343
nrw	\hat{p}_1	0.23695	0.19717	0.17164	0.15084	0.13600	0.12124
	\hat{p}_2	0.76226	0.80063	0.82917	0.84917	0.86357	0.87842
	$\hat{\Delta}$	0.52531	0.60347	0.65753	0.69833	0.72757	0.75718



n=8-9 als Empfehlung



—○— $\chi^2 \sim N$
-△- Fisher
·+· nichtzentraler χ^2
-◇- Annahme
-△- uis
·+· nis
-×- urw
-◇- nrw

C) Pilotversuche – Resource Equation Approach

Brief Communication

Sample Size Calculation in Animal Studies Using Resource Equation Approach

Wan Nor ARIFIN¹, Wan Mohd ZAHIRUDDIN²

¹ *Unit of Biostatistics and Research Methodology, School of Medical Sciences, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia*

² *Department of Community Medicine, School of Medical Sciences, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia*

Submitted: 24 May 2017

Accepted: 16 Jul 2017

Online: 26 Oct 2017

To cite this article: Arifin WN, Zahiruddin WM. Sample size calculation in animal studies using resource equation approach. *Malays J Med Sci.* 2017;**24(5)**:101–105. <https://doi.org/10.21315/mjms2017.24.5.11>

To link to this article: <https://doi.org/10.21315/mjms2017.24.5.11>

Diskussion und Fazit für die Praxis

- Methodische Konzepte zur Adressierung der 3 genannten Kernprobleme im Rahmen der biometrischen Planung von Tierversuchen wurden vorgestellt
- Alle Ansätze zeigten ein unmittelbares bzw. implizites Potential im Sinne der 3R
- Die Methoden der Medizinstatistik können gewinnbringend eingesetzt werden für den Tierschutz bzw. die Optimierung des Versuchswesens
- Allerdings müssen diese Methoden speziell für den Einsatz bei kleinen Fallzahlen geeignet sein
- In den zuständigen Genehmigungsbehörden braucht es zudem die Expertise, den Einsatz bestimmter Konzepte und Methoden angemessen bewerten zu können

Diskussion und Fazit für die Praxis

- Im Wesentlichen sollte in der Praxis der biometrischen Planung bedacht werden:
 - Dass die Anwendung „klassischer“ Test auf Normalverteilung (Kolmogorov-Smirnov, Shapiro-Wilk, etc.) bei Fallzahlen <50 eine extrem schlechte Power haben und demnach nicht dafür geeignet sind, die Anwendung parametrischer Tests zu rechtfertigen
 - Dass eine formale Fallzahlschätzung nur möglich ist, wenn seitens des Antragstellers vernünftige Vordaten zur Verfügung gestellt werden können
 - Dass metrische Zielgrößen bevorzugt genutzt werden sollten
 - Dass es zumindest derzeit noch keinen allgemein-akzeptierten Konsens darüber gibt, ob eher der confirmatorische oder der explorative Anspruch im Tierversuch überwiegt, was allerdings während der Planung zu berücksichtigen wäre
 - Dass auch „unüblichere“ Methoden der Medizinstatistik gewinnbringend eingesetzt werden können

Diskussion und Fazit für die Praxis

- Im Wesentlichen sollte auf Seiten der genehmigenden/beratenden Institutionen erörtert werden:
 - Ob die entsprechenden Gremien (zwingend) auf die Expertise von (Medizin-) Statistikern zurückgreifen sollten
 - Welche Maßnahmen zur Unterstützung von Antragstellern ggf. angeboten werden können/müssen (Core Facilities, Methodenberatung, etc.)
 - Ob die derzeitigen Antragsformulare den Prozess des Tierversuchswesen in der Weise abbilden, wie es gewünscht ist
 - Ob es eine grundlegende Neustrukturierung des gesamten Prozesses geben sollte
 - Ob es einer weitergehenden Vereinheitlichung bedarf

Diskussion und Fazit für die Praxis

1.1.5.2 Vorgesehene Anzahl und Begründung für die Anzahl der Tiere einschließlich Angaben zur biometrischen Planung (§ 31 Abs. 1 Satz 2 Nr. 1c TierSchVersV)

-ggf. Anlage „Statistisches Gutachten“ beifügen-
-ggf. Anlage Formblatt „Angaben zur biometrischen Planung“ verwenden-

Tierart	Gesamtanzahl (inkl. Reservetiere)

Versuchs- und Kontrollgruppen

Erläuterungen:

Angaben zur biometrischen/statistischen Planung

Erläuterungen:

Hauptzielgröße(n):

Nebenzielgröße(n):

Studientyp

- a) Orientierungsstudie
 b) Vergleichsstudie

Es werden folgende biometrische Verfahren zur Auswertung eingesetzt:

Die vorgesehene Tierzahl und Gruppengröße ist zur statistischen Absicherung mit

- einer Wahrscheinlichkeit für den Fehler 1. Art von _____
- einer Wahrscheinlichkeit für den Fehler 2. Art von _____
- einer biologisch relevanten Differenz _____
- Varianz oder Effektstärke (mit Angabe des genutzten Parameters, z.B. Effektstärke nach Cohen) _____

RP Tübingen



2.3.8.3 Vorgesehene Anzahl und Begründung für die Anzahl der Tiere einschließlich Angaben zur biometrischen Planung (§ 31 Abs. 1 Satz 2 Nr. 1c TierSchVersV)

- bitte Gutachten einer Statistikerin/eines Statistikers oder ein Formblatt zur biometrischen Planung (Anlage 3) für jeden Teilversuch beifügen -

Tierart(en)	Gesamtanzahl pro Tierart (inkl. Reservetiere)

2.3.8.4 Versuchstyp

- A Technisch erforderlicher Vorversuch, bei dem die Tiere der Gewinnung des Materials dienen und selbst nicht im Versuch verwendet werden.
- B Hypothesen generierender Versuch mit geringen Tierzahlen, bei dem noch keine spezifizierten Hypothesen geprüft werden können. (Pilot-Versuch, Grundlagenklärung)
- C Hypothesen überprüfender Versuch. → in diesem Fall je ein Formblatt zur biometrischen Planung pro Teilversuch beifügen!

Tabellarische Übersicht über die Versuchs- und Kontrollgruppen inkl. Tierzahlen

Im Falle von Auswahl A (Materialgewinnung) oder B (Pilotstudie) Erläuterung der beantragten Tierzahl



Saarland

Diskussion und Fazit für die Praxis

Stand: Juli 2016

Dieses Formblatt bitte für jeden hypothesen-überprüfenden Teilversuch ausfüllen!

Vorgesehene Anzahl und Begründung für die Anzahl der Tiere einschließlich Angaben zur biometrischen Planung (§ 31 Abs. 1 Satz 2 Nr. 1c TierSchVersV)

Teilversuchs-Nr.
 Tierart
 Gesamttierzahl Anzahl Gruppen

Bezeichnung des Teilversuches:

1 Primäre Zielgröße (zu untersuchendes Haupt-Zielmerkmal des Versuches, anhand dessen die Stichprobe berechnet wurde, ggf. inkl. der Transformation dieser Zielgröße).

2 Verteilungsform der Zielgröße (soweit anhand der Vordatendichte Angabe möglich; in diesem Fall inkl. des Testverfahrens mit dem die Verteilungsform bestimmt wurde)

3 Planungsgrößen (erwarteter Unterschied (z.B. erwartete relevante Mittelwertsdifferenz und Standardabweichung oder zu detektierende Veränderung der Ereignisrate, i.A. mit Begründung für diese Erwartung anhand von Vordaten) & verwendete Effektgröße mit Angabe ihrer Berechnung)

4 Statistisches Test-Verfahren mit welchem die Tierzahlplanung durchgeführt wird (z.B. t-Test, Anova, Wilcoxon-Test).

5 Annahmen zum Fehler 1. Art (alpha) und 2. Art (beta) (falls alpha ungleich 0,05 und beta ungleich 0,2 bitte begründen)

6 Falls verwendet: Bei der Berechnung durchgeführte Korrekturen (z.B. Bonferroni-Korrektur → in diesem Fall inkl. Angabe, welche Gruppen jeweils gegeneinander getestet werden und Angabe des geänderten alpha-Werts).

7 Richtung von Hypothese und Test

(ein- oder zweiseitig? → 2-seitiges Testen bitte begründen).

8 Name (und soweit nicht aus dem Antrag ersichtlich Anschrift) des Verfassers der biometrischen Planung

Diskussion und Fazit für die Praxis

Hilfsmittel zur Fallzahlplanung

Der Gruppenbesatz kann leicht abgeschätzt werden bei experimentellen Gruppenvergleichen, die auf die Forderung „beobachtbare Differenz eines Zielparameters mit bekannter oder abschätzbarer Standard - Abweichung“ zurückgeführt werden können. Zunächst sollte definiert werden: welcher Zielparameter (z.B. Blutdruck, Tumorwachstum) kann als quantitativer „Schlüsselparameter“ für die Abschätzung verwendet werden? Wie groß muss eine Verschiebung / Abweichung („Blutdrucksenkung“) dieses Parameters sein, um ein Experiment als erfolgreich zu kennzeichnen? („relevante Differenz – Δ “) und mit welcher Standard – Abweichung σ ist dieser Parameter ungefähr messbar? Der Quotient aus beiden Größen ist die „Effektstärke“ ε . Weiter ist die Zahl der zu vergleichenden Gruppen (ANOVA ja/nein), ggfs. die erwartete Richtung der Abweichung (einseitiger oder zweiseitiger Test) und die Art des Vergleichs (gepaarter Test / unabhängige Gruppen) zu berücksichtigen.

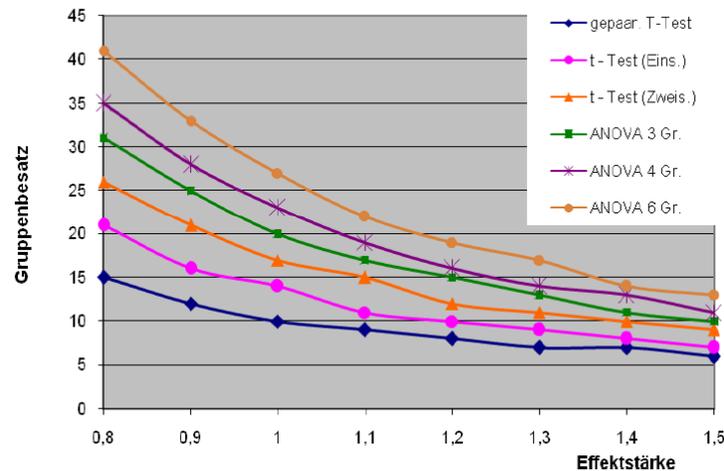
Gruppengrößen bei:

Effekt - Stärke	Gepaa t-Test	t-Test (eins.)	t- Test (zweis.)	ANOVA 3 Gr.	ANOVA 4 Gr.	ANOVA 6 Gr.
0,8	15	21	26	31	35	41
0,9	12	16	21	25	28	33
1	10	14	17	20	23	27
1,1	9	11	15	17	19	22
1,2	8	10	12	15	16	19
1,3	7	9	11	13	14	17
1,4	7	8	10	11	13	14
1,5	6	7	9	10	11	13

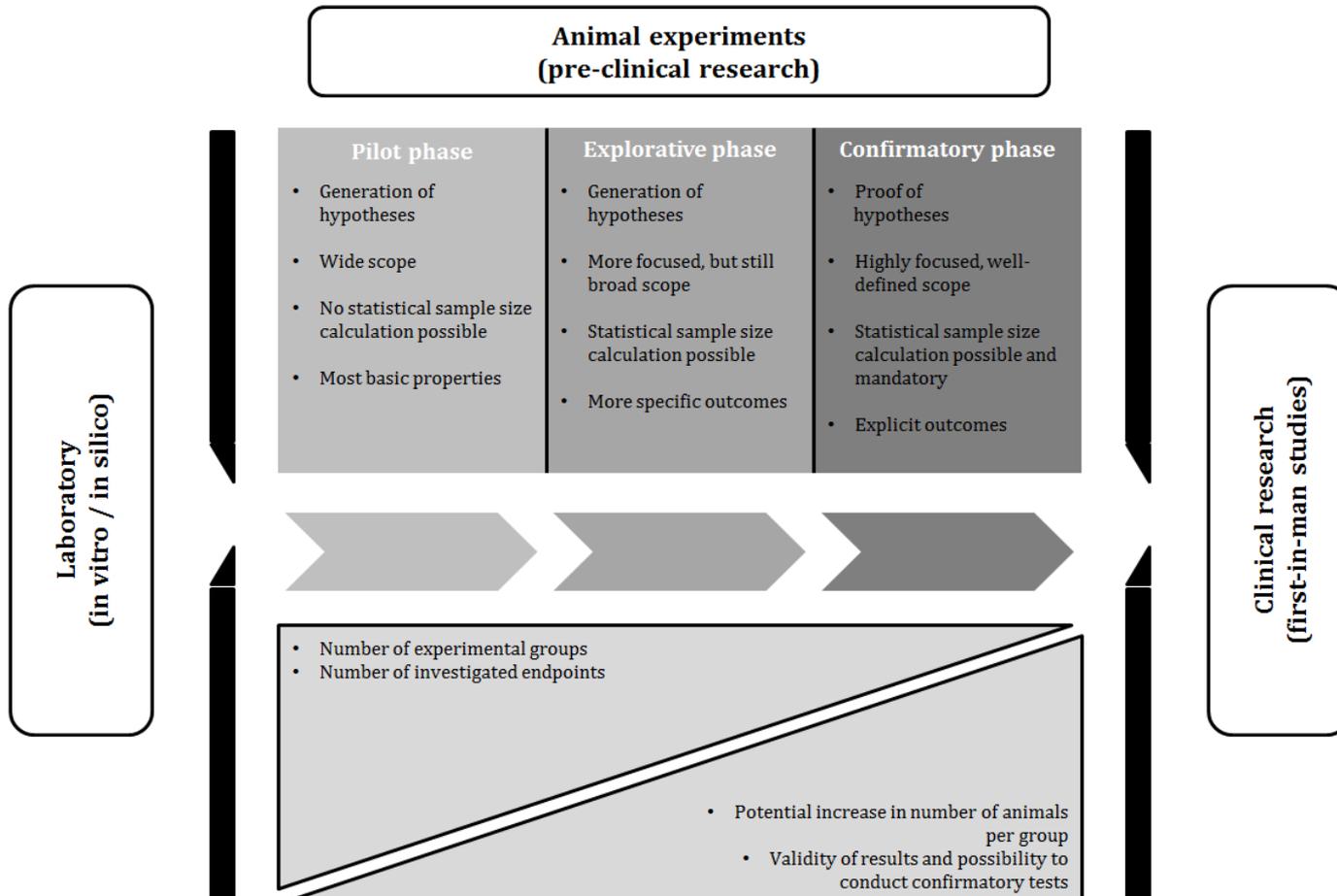
Planungen von Einzelgruppen - Größe mit folgenden Annahmen:

- Standard - Abweichungen σ in allen Experimental – Gruppen gleich groß und gleich der Kontrollgruppe
- $p < 0,05$;(Fehler erster Art)
- Fehler zweiter Art (β) $< 0,2$;
- "Effekt - Stärke" : die erwartete / geforderte Differenz bezogen auf die Standard - Abweichung. ($\varepsilon = \Delta / \sigma$), Werte normal verteilt

Fallzahlplanung mit "gPower" bzw. "Primer of Biostatistics"



Diskussion und Fazit für die Praxis



Literatur

- Allgoewer A, Mayer B. Sample size estimation for pilot experiments in animal research using a Markov Chain Monte Carlo approach. *ATLA* 2017, 45(2): 83-90
- Arifin WN, Zahiruddin WM. Sample size calculation in animal studies using resource equation approach. *Malaysian Journal of Medical Science* 2017, 24(5): 101–105
- Bretz F, Maurer W, Brannath W, Posch M. A graphical approach to sequentially rejective multiple test procedures. *Statistics in Medicine* 2009, 28: 586-604
- Efron B, Tibshirani RJ. *An introduction to the bootstrap*. Chapman & Hall, Boca Raton, 1994
- Gietz L, Mayer B. A resampling approach to validate sample size estimation for animal experiments. *Journal of International Translational Medicine* 2017, 5(2): 53-62
- Hastings W. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 1970, 57(1): 97-109
- Manly BFJ. *Randomization, bootstrap and monte carlo methods in biology (3rd edition)*. Chapman & Hall, Boca Raton, 2007
- Mayer B, Allgöwer A, Muche R. Notwendige Standards der biometrischen Fallzahlplanung von Tierversuchen im Sinne der 3R. *Berliner und Münchner Tierärztliche Wochenschrift* 2018, in press
- Mayer B, Stahl V, Kron M. Statistical planning of animal experiments using gatekeeping procedures. *ATLA* 2017, 45: 317-328
- Qumsiyeh M. Using the bootstrap for estimating the sample size in statistical experiments. *Journal of Modern Applied Statistical Methods* 2013, 12(1): 45-53
- Russell WMS, Burch RL. *The Principles of Humane Experimental Technique*. Methuen, London, 1959
- Yu CH. Resampling methods: concepts, applications, and justification. *Practical Assessment, Research & Evaluation* 2003, 8(19)

Vielen Dank für die Einladung!



"I go home today. They cured me using this new miracle drug. I'm afraid it'll be years before it's approved for humans."