Tierversuche in der Forschung: Das 3R-Prinzip und die Aussagekraft wissenschaftlicher Forschung

Handreichung der Ständigen Senatskommission für tierexperimentelle Forschung der DFG zur Planung und Beschreibung tierexperimenteller Forschungsprojekte

DFG

Stand: Juli 2019

Ansprechpartnerin:

Dr. Sonja Ihle
Gruppe Lebenswissenschaften 1
Tel.: 0228/885-2362
E-Mail: sonja.ihle@dfg.de

Inhaltsverzeichnis

1 Zusammenfassung ... 3

2 Zusammenhang zwischen Tierschutz in der Forschung und wissenschaftlicher Qualität ... 5

3 Mehr als 3R – 3R-Prinzip und wissenschaftliche Aussagekraft ... 7

3.1 Legitimation von tierexperimenteller Forschung durch Güterabwägung ... 7

3.2 Das 3R-Prinzip als normatives Kriterium für den Tierschutz in der Forschung ... 8

3.3 Das 3R-Prinzip im Kontext von Validität und Replizierbarkeit ... 8

3.4 Aspekte der wissenschaftlichen Aussagekraft von Tierversuchen ... 9

4 Spannungsfelder zwischen der wissenschaftlichen Validität und den 3R ... 12

4.1 Wahl der Tierart ... 12

4.2 Festlegung der Tierzahl ... 14

4.3 Standardisierung ... 15

4.4 Pilot- und Explorationsversuche ... 16

4.5 Wiederholungsversuche ... 17

5 Anhang ... 19

6 Mitglieder der Ständigen Senatskommission für tierexperimentelle Forschung der Deutschen Forschungsgemeinschaft ... 20

1 Zusammenfassung

Das im Grundgesetz verankerte Recht auf Freiheit der Forschung und das im Grundgesetz benannte Staatsziel Tierschutz sind zwei bedeutende Güter, die für die Belange der tierexperimentellen Forschung eine hohe Relevanz haben.

Die Verwendung von Tieren zu wissenschaftlichen Zwecken ist nur dann ethisch zu rechtfertigen, wenn ein Tierversuch „unerlässlich“ ist und der zu erwartende wissenschaftliche Erkenntnisgewinn die Belastung der Tiere im Versuch rechtfertigt.

Wesentlicher ethischer Leitgedanke in der tierexperimentellen Forschung ist das 3R-Prinzip (Replace, Reduce, Refine)[1]: Tierversuche dürfen nur dann durchgeführt werden, wenn zur Aufklärung der wissenschaftlichen Fragestellung keine anderen geeigneten Methoden zur Verfügung stehen und die Anzahl und Belastung der eingesetzten Tiere auf das unerlässliche Maß beschränkt werden. Die konsequente Umsetzung des 3R-Prinzips ist im ureigenen Interesse der Wissenschaft, da eine Beeinträchtigung des Wohlergehens der Versuchstiere auch die Aussagekraft der Forschungsergebnisse beeinflussen kann. Der Anspruch, ein Höchstmaß an wissenschaftlicher Validität und Replizierbarkeit der Ergebnisse zu erzielen, bei gleichzeitiger Berücksichtigung von tierschutzrelevanten Maßnahmen, muss die Grundlage jeder tierexperimentellen Versuchsplanung sein.

In der Praxis kann es zwischen den Maßnahmen zur Umsetzung des Tierschutzes in der Forschung und den Anforderungen an die wissenschaftliche Validität zu Spannungen kommen. Aus der Wechselwirkung der beiden Aspekte ergibt sich die Notwendigkeit, Maßnahmen zur Umsetzung des 3R-Prinzips nicht separat zu betrachten, sondern als integralen Bestandteil wissenschaftlicher Projektplanung zu verankern und im Rahmen der Beschreibung von Forschungsprojekten zu berücksichtigen.

---

1 Russell, W. M. S., Burch, R. L. (1959): The principles of humane experimental technique. London: Methuen.

Die Ständige Senatskommission für tierexperimentelle Forschung der DFG möchte mit dieser Handreichung einen Beitrag zur Debatte über Qualität in der biomedizinischen Forschung liefern und die spezifischen Anforderungen an die Durchführung von Tierversuchen herausarbeiten. Die Handreichung bietet darüber hinaus Hilfestellung zur validen Planung und Beschreibung von tierexperimentellen Forschungsprojekten.

2 Zusammenhang zwischen Tierschutz in der Forschung und wissenschaftlicher Qualität

Wissenschaft basiert auf dem Anspruch, ein Höchstmaß an Qualität und Aussagekraft der Forschungsergebnisse anzustreben. Im Rahmen der Diskussionen über eine „Replikationskrise“ in der Wissenschaft[2,3] hat die DFG in ihren Stellungnahmen zur Replizierbarkeit von Forschungsergebnissen die Sicherung der Qualität von Forschung als grundlegenden Anspruch an die Wissenschaft hervorgehoben und stößt einen Diskurs zur fachspezifischen Reflexion über Qualität und Replizierbarkeit sowie verschiedene Aspekte der Qualitätssicherung an.[4,5]

In der tierexperimentellen Forschung besteht die Notwendigkeit, die wissenschaftliche Aussagekraft von Ergebnissen und zugleich einen adäquaten und verantwortlichen Umgang mit Versuchstieren zu gewährleisten. Hohe Tierschutzstandards stellen dabei eine Voraussetzung für qualitativ hochwertige Forschung dar, und umgekehrt sind Tierversuche nur zu rechtfertigen, wenn die Qualität der Forschung und damit die Aussagekraft von Forschungsergebnissen sichergestellt sind.

Aufgrund der grundsätzlichen Schutzwürdigkeit von Tieren tragen Forschende eine besondere moralische Verantwortung bei der Verwendung von Tieren. Eine kriteriale Orientierung für Maßnahmen zum Tierschutz beschreibt das 3R-Prinzip, nämlich die Verfeinerung (refinement), Verringerung (reduction)

---

2 Ioannidis, J. (2005): Why most published research findings are false. PLoS Medicine, 2(8): 696–701.

3 McLeod, M. R. et al. (2014): Increasing value, reducing waste. The Lancet, 383(9912): 101–104.

4 Deutsche Forschungsgemeinschaft (2017): Replizierbarkeit von Forschungsergebnissen – Stellungnahme.

5 Deutsche Forschungsgemeinschaft, Ständige Senatskommission für Grundsatzfragen in der Klinischen Forschung, Arbeitsgruppe „Qualität in der Klinischen Forschung“ (2018): Replizierbarkeit von Ergebnissen in der Medizin und Biomedizin – Stellungnahme.

und Vermeidung (replacement) von Tierversuchen. Maßnahmen im Sinne des 3R-Prinzips und das bei Forschungsprojekten inhärente Ziel, ein Maximum an wissenschaftlicher Aussagekraft zu erzielen, stellen zwei Aspekte dar, die bei der Planung und Beschreibung wissenschaftlicher Projekte nicht losgelöst voneinander betrachtet werden dürfen. Auch stellen die zu erwartende Belastung der Tiere und der angestrebte Erkenntnisgewinn aus dem Forschungsprojekt die Grundlage für eine Güterabwägung dar, ob der Einsatz von Tieren zu Forschungszwecken zu rechtfertigen ist.[6]

Insofern ist es im ureigenen Interesse der Wissenschaft, hohe wissenschaftliche Standards in tierexperimentellen Forschungsprojekten anzustreben und Maßnahmen zur Sicherstellung des Tierschutzes unmittelbar in der Projektplanung zu verankern.

Um diesem Anliegen nachzukommen, beschränkt sich die Handreichung bewusst auf Aspekte des 3R-Prinzips bei der Ausarbeitung, Planung und Durchführung von tierexperimentellen Forschungsprojekten. Konkrete Spannungsfelder werden beispielhaft vorgestellt. Übergeordnete Aspekte, die die Qualitätssicherung in der tierexperimentellen Forschung allgemein betreffen, werden in der Stellungnahme der Arbeitsgruppe „Qualität in der Klinischen Forschung“ thematisiert. Dieser Stellungnahme schließt sich die Senatskommission für tierexperimentelle Forschung an.

6 Deutsche Forschungsgemeinschaft, Ständige Senatskommission für tierexperimentelle Forschung (2016): Tierversuche in der Forschung.

3 Mehr als 3R – 3R-Prinzip und wissenschaftliche Aussagekraft

3.1 Legitimation von tierexperimenteller Forschung durch Güterabwägung

Die Verwendung von Tieren in der Forschung richtet sich nach dem Grundsatz der Verhältnismäßigkeit. Dieser kommt als rechtsstaatliches Prinzip immer dann zur Anwendung, wenn mehrere im Grundgesetz verankerte Grundrechte, Rechtsgüter oder -prinzipien (hier: Allgemeines Persönlichkeitsrecht, Integrität von Leib und Leben, Forschungsfreiheit und Tierschutz) kollidieren. Demnach sind Versuche an Tieren legitim, wenn die zu erwartenden Erkenntnisse bedeutsam sind und den Tieren dabei keine vermeidbaren Schmerzen, Leiden oder Schäden zugefügt werden. Das Instrument zur Überprüfung der Verhältnismäßigkeit eines Tierversuchsvorhabens ist die Güterabwägung im Sinne einer Schaden-Nutzen-Analyse. Dabei wird überprüft, ob ein legitimes Versuchsvorhaben (a) geeignet, (b) erforderlich und (c) angemessen ist, um den angestrebten Erkenntnisgewinn zu erzielen.[7]

Eignung und Erforderlichkeit sind notwendige Voraussetzungen für eine Güterabwägung; hier wird nach der wissenschaftlichen Begründung des Versuchsvorhabens gefragt (Eignung) und nach möglichen Alternativen zur Verwendung von Tieren und belastenden Methoden (Erforderlichkeit). Nur wenn ein Versuchsvorhaben wissenschaftlich sinnvoll und geeignet erscheint und keine Alternativen zu Tierversuchen zur Verfügung stehen, kann eine Güterabwägung im Sinne der Schaden-Nutzen-Analyse erfolgen. Die Güterabwägung konzentriert sich dann auf die Frage der Angemessenheit von Versuchsvorhaben und überprüft, ob der zu erwartende Nutzen des Versuchsvorhabens die Belastungen der Versuchstiere rechtfertigt. Nur dann kann das Versuchsvorhaben genehmigt werden.

---

7 Wienbracke, M. (2013): Der Verhältnismäßigkeitsgrundsatz. Zeitschrift für das Juristische Studium (ZJS). Nr. 2: 148 –155.

3.2 Das 3R-Prinzip als normatives Kriterium für den Tierschutz in der Forschung

Jeder Tierversuch muss nach dem deutschen Tierschutzgesetz (TierSchG) vor Durchführung bei der zuständigen Genehmigungsbehörde beantragt werden (§ 8 TierSchG). Im Zuge des Genehmigungsverfahrens wird ausdrücklich geprüft, ob das 3R-Prinzip (Replace, Reduce, Refine) in seinen unterschiedlichen Aspekten sinnvoll und bestmöglich umgesetzt wird. Hier ist nachzuweisen, dass der angestrebte Erkenntnisgewinn nicht ohne Verwendung empfindungsfähiger Tiere (Replace), mit weniger Tieren (Reduce) oder mit weniger belastenden Methoden (Refine) erzielt werden kann. Neben dem Einsatz von Methoden ohne Versuchstiere (sogenannte Alternativmethoden) ist zu prüfen, ob Tiere mit geringerer Empfindungsfähigkeit eingesetzt werden können („relatives replacement“). Die Minimierung der Tierzahl erfordert eine sorgfältige Versuchsplanung mit Stichprobenberechnung (zum Beispiel Poweranalyse). Zur Minimierung der Belastung der Tiere stehen zahlreiche Verfahren zur Verfügung. So können zum Beispiel eine Verbesserung der Haltungsbedingungen durch Umweltanreicherung, die schrittweise Gewöhnung der Tiere an die Versuchsbedingungen, die Verwendung nicht invasiver Techniken sowie optimierte Anästhesie- und Analgesieverfahren die Gewichtung von Schaden und Nutzen in der Güterabwägung zugunsten eines Tierversuchs verschieben.

3.3 Das 3R-Prinzip im Kontext von Validität und Replizierbarkeit

Eine gewissenhafte Umsetzung der 3R ist eine notwendige, aber nicht hinreichende Voraussetzung für die ethische Vertretbarkeit von Tierversuchen. Unabhängig vom Belastungsgrad der Tiere muss das Studiendesign stringenten wissenschaftlichen Qualitätsanforderungen an Objektivität, Validität und Replizierbarkeit entsprechen. Fundamental für die Verantwortung eines jeden Forschenden ist die Sicherstellung wissenschaftlicher Qualität, um maximale wissenschaftliche Aussagekraft zu erzielen. Dieses Ziel muss ebenso immer zentrales Kriterium des wissenschaftlichen Begutachtungsverfahrens von Forschungsprojekten sein.

Eine positive Bewertung der wissenschaftlichen Aussagekraft (siehe Abschnitt 3.4) eines Versuchsvorhabens und die Anwendung des 3R-Prinzips ist Voraussetzung für die abschließende Prüfung der Angemessenheit im Rahmen der finalen Güterabwägung. Gemäß dem Grundsatz der Verhältnismäßigkeit kann vor allem unter Berücksichtigung des nur in Kollision mit anderen Grundrechten einschränkbaren Grundrechts der Forschungsfreiheit das Kriterium der wissenschaftlichen Aussagekraft dem nach 3R aufgestellten Tierschutz vorgeordnet werden.[8] Das 3R-Prinzip kann seine Funktion nur bei Vorliegen eines validen Versuchsplans erfüllen und das Reduce-Prinzip darf nicht so missverstanden werden, dass die Tierzahl in einer Studie auf Kosten der Aussagekraft gesenkt wird (siehe Abschnitt 4.2). Ob ein Tierversuch unter Berücksichtigung seiner wissenschaftlichen Qualität und seiner Erforderlichkeit angemessen und damit genehmigungsfähig ist, muss die abschließende Güterabwägung ergeben.

3.4 Aspekte der wissenschaftlichen Aussagekraft von Tierversuchen

Im Gegensatz zum 3R-Prinzip als zentraler Komponente bei der Beurteilung der ethischen Vertretbarkeit der Durchführung von Tierversuchen steht für die Beurteilung der Eignung von Tierversuchen kein ähnlich griffiges Konzept zur Verfügung. Ganz generell lässt aber die Prüfung der wissenschaftlichen Aussagekraft von Versuchsergebnissen eine gute Orientierung zu. Die wesentlichen Kriterien wissenschaftlicher Aussagekraft von tierexperimentellen Forschungsprojekten lassen sich drei unterschiedlichen Aspekten wissenschaftlicher Validität zuordnen, die sich auf die Güte des Tier- beziehungsweise Versuchsmodells, die Güte der postulierten Ursache-Wirkung-Relation und auf den Grad der Generalisierbarkeit beziehen. Diese werden beispielsweise nach Würbel (2017) in Konstruktvalidität, interne Validität und externe Validität untergliedert:[9]

---

8 Wienbracke, M. (2013): a. a. O.

9 Würbel H. (2017): More than 3Rs: The importance of scientific validity for harm-benefit analysis of animal research. Lab Animal, 46(4): 164–166.

Konstruktvalidität bezieht sich auf die Güte des Tier- beziehungsweise Versuchsmodells und der verwendeten Methoden zur Untersuchung der konkreten Fragestellung eines Versuchsvorhabens. Sie bezeichnet die Genauigkeit, mit der eine Versuchsanordnung das misst, was sie zu messen vorgibt.[10] Die Beurteilung der Konstruktvalidität sollte sich deshalb auf empirische Evidenz zur biologischen Übereinstimmung zwischen Eigenschaften der Versuchspopulation und Eigenschaften der Zielpopulation und auf die Aussagekraft der Zielgröße des Versuchs für die damit zu messende Funktion beziehen.[11,12]

Interne Validität bezeichnet die Güte der postulierten Ursache-Wirkung-Relation. Interne Validität ist gegeben, wenn eine Veränderung in der abhängigen Variable (ein Behandlungseffekt) nachweislich auf Variation in der Versuchsbehandlung zurückgeführt werden kann. Eine Beurteilung der internen Validität muss deshalb neben grundlegenden Aspekten der Versuchsplanung und der Verwendung von Kontrollgruppen auch alle anwendbaren Kriterien guter wissenschaftlicher Praxis zur Vermeidung von Verzerrungen (bias) berücksichtigen (zum Beispiel Randomisierung, Verblindung, Stichprobenberechnung, Definition der primären und sekundären Zielvariablen, Definition von Ein- und Ausschlusskriterien, Definition des statistischen Auswertungsplans etc.).

Externe Validität bezeichnet den Grad der Generalisierbarkeit von Versuchsergebnissen über die spezifischen Bedingungen des konkreten Versuchs hinaus. In der empirischen Forschung steht die Generalisierbarkeit der Befunde im Vordergrund. Allein für die Replizierbarkeit von Ergebnissen ist ein

---

10 Koob, G. F., Heinrichs, S. C., Britton K. (1998): Animal models of anxiety disorders. The American Psychiatric Press Textbook of Psychopharmacology. (2. Schatzberg AF, Nemeroff CB, editor). Washington DC-London: American Psychiatric Press: 133–144.

11 Willner, P. (1984): The validity of animal models of depression. Psychopharmacology, 83(1): 1–16.

12 Belzung, C., Lemoine, M. (2011): Criteria of validity for animal models of psychiatric disorders: focus on anxiety disorders and depression. Biology of Mood & Anxiety Disorders, 1(1): 9.

Mindestmaß an externer Validität erforderlich, da sich die Versuchsbedingungen zwischen Wiederholungsversuchen (sowohl innerhalb eines Versuchslabors als auch zwischen verschiedenen Versuchslaboren) zwangsläufig unterscheiden. Ausgehend davon sollte die Beurteilung der externen Validität Aspekte des Versuchsplans berücksichtigen, die Rückschlüsse auf die Replizierbarkeit und Generalisierbarkeit (zum Beispiel auf andere Tiermodelle, andere Versuchsbedingungen) zulassen. Dazu gehören unter anderem die Berücksichtigung der Geschlechter, die Aufteilung eines Versuchs in mehrere unabhängige Replikate, die systematische Variation einer oder mehrerer unabhängiger Variablen (zum Beispiel mehrere Mauslinien, Haltungsbedingungen, Versuchsanordnungen oder Experimentatoren) oder die Durchführung von Multi-Labor-Studien.

Mit diesen drei Aspekten der Validität tierexperimenteller Forschungsergebnisse können wesentliche Kriterien zur Beurteilung der wissenschaftlichen Aussagekraft und Replizierbarkeit von Tierversuchen abgedeckt werden. Analog zum 3R-Prinzip, welches eine Beurteilung der Erforderlichkeit von Tierversuchen ermöglicht, kann die Anwendung dieser drei Aspekte der Validität die Beurteilung der Eignung von Tierversuchsvorhaben zur Erzielung des angestrebten Nutzens erleichtern.

4 Spannungsfelder zwischen der wissenschaftlichen Validität und den 3R

Grundsätzlich haben Forschende nicht nur aus ethischer Sicht ein großes Interesse daran, Tierschutzmaßnahmen in ihren Projekten umzusetzen. So können tierschutzrelevante Belastungen (Schmerzen, Leiden, Schäden) der Versuchstiere die Aussagekraft der Ergebnisse beeinträchtigen. Deshalb ist es nicht nur aus Tierschutzgründen, sondern auch mit Blick auf die Qualität der Forschung ein vorrangiges Anliegen, die Belastungen der Tiere soweit wiemöglich zu minimieren. Das 3R-Prinzip und wissenschaftliche Aussagekraft gehen in den meisten Fällen Hand in Hand.

In manchen Bereichen der Versuchsplanung bestehen jedoch grundlegende Spannungsfelder zwischen der Umsetzung des 3R-Prinzips und der Maximierung der wissenschaftlichen Aussagekraft. Dies kann sich auch auf die Beurteilung von Versuchsvorhaben im Rahmen des Genehmigungsverfahrens auswirken. Die Anwendung verbindlicher Kriterien wissenschaftlicher Validität (zum Beispiel Konstruktvalidität, interne und externe Validität) helfen, die Auswirkungen von 3R-Maßnahmen auf wesentliche Aspekte der wissenschaftlichen Aussagekraft zu prüfen.

Spannungsfelder bestehen bezüglich aller 3R; bei der Verwendung bestimmter Tierarten, bei der Festlegung der Tierzahl (Stichprobengröße) und bei der Anwendung bestimmter tierexperimenteller Techniken. Um Forschenden bei der Planung, Beschreibung und Begründung ihrer Projekte eine Hilfestellung zu geben, werden einige dieser Spannungsfelder exemplarisch dargestellt und mögliche Lösungsansätze vorgeschlagen.

4.1 Wahl der Tierart

Der Begriff „Replace“ meint nicht nur den Ersatz von Tierversuchen durch Alternativmethoden (in silico, in vitro usw.), sondern umfasst auch „relatives replacement“, das heißt das Ausweichen auf sogenannte „niedere“ Tiere (Wirbellose) oder „niederere“ Wirbel- bzw. Säugetiere (zum Beispiel Mäuse statt

Primaten, Fische statt Säugetiere). Hintergrund für das Verständnis des „relativen replacements“ ist die verbreitete Auffassung, dass sogenannte „höhere“ Wirbeltierarten (zum Beispiel Primaten, aber auch Hunde, Katzen) unter den Versuchseinwirkungen stärker leiden als „niederere“ Arten[13]. Diese Auffassung spiegelt sich auch in der EU-Richtlinie 2010/63/EU zum Schutz der für wissenschaftliche Zwecke verwendeten Tiere wider, die Primaten und andere Tierarten unter einen besonderen Schutz stellt. Die postulierte Hierarchie des Leidens „höherer“ versus „niederer“ Arten unter gegebenen Versuchsbedingungen ist aus wissenschaftlicher Sicht fragwürdig und wird auch aus ethischer Sicht kontrovers beurteilt.[14,15] Das Tierschutzgesetz geht grundsätzlich bei allen Wirbeltieren (und einigen Wirbellosen) von Empfindungsfähigkeit aus und schließt diese für andere Tierarten nicht aus[16]. Für eine ethische Hierarchisierung innerhalb der Wirbeltiere anhand unterschiedlicher Leidensfähigkeit fehlen bisher verbindliche Kriterien. Die Anwendung des „relativen replacement“ kann auch aus wissenschaftlicher Sicht durchaus kritisch sein, wenn die Wahl des Tiermodells einen Einfluss auf die wissenschaftliche Aussagekraft hat. Der Nutzen der Forschung (Erkenntnisgewinn) und damit auch deren ethische Legitimation kann dadurch unter Umständen beeinträchtigt werden.

- *Neben einer ethischen und juristischen Klärung der postulierten ethischen Hierarchie innerhalb der vom Tierschutzgesetz geschützten Tierarten braucht es diesbezüglich auch eine größere Sensibilisierung für die Problematik. Forschende sollten sich bei der Versuchsplanung bewusst mit der Wahl ihrer Tiermodelle auseinandersetzen und diese im Rahmen der Güterabwägung auch anhand wissenschaftlicher*

13 So schreibt das Tierschutzgesetz (§ 7(1)1) vor, dass „Tierversuche im Hinblick auf […] die artspezifische Fähigkeit der verwendeten Tiere, unter den Versuchseinwirkungen zu leiden, auf das unerlässliche Maß zu beschränken“ sind.

14 Rippe, K. P. (2003): Tierethik. Bioethik, hrsg. v. M. Düwell & K. Steigleder, Frankfurt a. M.

15 Ach, J. S., Borchers, D. (Hg.): Handbuch Tierethik. Grundlagen – Kontexte – Perspektiven. Stuttgart: J. B. Metzler 2018.

16 Tierschutzgesetz § 8(a)(4).

Kriterien der Validität und der artspezifischen Leidensfähigkeit unter den gegebenen Versuchsbedingungen begründen.

4.2 Festlegung der Tierzahl

Der Begriff „Reduce“ bezieht sich hauptsächlich auf die Anzahl von Tieren, die für einen Versuch verwendet werden (Gesamtstichprobe). Aus wissenschaftlicher Sicht wären große Stichproben wünschenswert, aus Sicht des Tierschutzes hingegen möglichst kleine. Die Berechnung der erforderlichen Stichprobengröße (zum Beispiel mittels Poweranalyse) ist in vielen Fällen ein geeignetes Mittel, um die minimale Tierzahl festzulegen.[17] Da der vermeintliche Fortschritt im Tierschutz in der Forschung primär anhand der jährlichen Versuchstierstatistik gemessen und öffentlich kommuniziert wird, besteht zudem erheblicher politischer Druck auf die Verminderung der Tierzahlen. Aufgrund der oben genannten verbreiteten Ansichten einer ethischen Hierarchisierung innerhalb der Wirbeltiere trifft dies im besonderen Maße auf Versuche an sogenannten „höheren“ Säugetieren zu.

Übermäßiges Minimieren der Stichprobengröße unter dem Titel 3R (Reduce) wird dann problematisch, wenn dies durch subkritische wissenschaftliche Aussagekraft „erkauft“ wird. Dies ist besonders dann der Fall, wenn auf wichtige Vergleichs- oder Kontrollgruppen verzichtet wird oder zu kleine Stichproben verwendet werden. Dabei ist zu bedenken, dass die ethischen „Kosten“ von zu geringen Stichproben höher ausfallen können als die von zu großen, da Tierversuche unterhalb einer bestimmten Stichprobengröße aufgrund mangelnder statistischer Power im ungünstigsten Fall wertlos werden können.

Forschende müssen deshalb die Tierzahlen für ihre Versuche anhand wissenschaftlicher Kriterien (etwa der internen und externen Validität der angestrebten Ergebnisse) begründen können. Es gibt zahlreiche Hilfsmittel zur Versuchsplanung (zum Beispiel der Experimental Design Assistant,

17 Festing, M. F. W. (2018): On determining sample size in experiments involving laboratory animals. Laboratory Animals, 52(4): 341–350.

https://eda.nc3rs.org.uk/) oder zur Berechnung und Begründung der Stichprobengröße (etwa G*Power, www.gpower.hhu.de/). Bei vielen Versuchen besteht zudem die Möglichkeit eines adaptiven Versuchsdesigns (beispielsweise gruppensequentielles Design, Zwischenanalyse mit Anpassung der Stichprobengröße, adaptive Randomisierung usw.), um die Tierzahl ohne Verlust an wissenschaftlicher Aussagekraft zu minimieren. Bei komplexeren Versuchsansätzen empfiehlt es sich jedoch, professionelle Unterstützung durch Biostatistiker mit Expertise in der Tierversuchsplanung in Anspruch zu nehmen.

- *Die Angaben zur Anzahl von Versuchstieren bedürfen einer sorgfältigen Abwägung zwischen Bestrebungen zur Reduktion der Tierzahlen und der wissenschaftlichen Aussagekraft und sollten bei der Beschreibung des Projekts erläutert werden.*

4.3 Standardisierung

Tierversuche werden meist im Versuchslabor unter kontrollierten Bedingungen durchgeführt. Kontrollierte Versuchsbedingungen sind in der Grundlagenforschung oft wünschenswert, um allfällige Störvariablen auszuschließen. Dabei werden die Tiere und die Haltungsbedingungen meist rigoros standardisiert, indem genetisch weitgehend identische Tiere (zum Beispiel Inzuchtlinien) unter identischen Bedingungen (bezüglich Käfig, Futter, Management) aufgezogen und gehalten werden. Standardisierung wird zudem oft als Mittel zur Minimierung der Tierzahl (Reduce) empfohlen, da sich Behandlungseffekte bei geringerer Variation der Versuchsergebnisse mit kleineren Stichproben statistisch nachweisen lassen.

Standardisierung kann zwar die Präzision von Versuchsergebnissen verbessern (geringere Variation), übermäßige Standardisierung beeinträchtigt allerdings die externe Validität und damit die Generalisierbarkeit der Versuchsergebnisse.[18] Durch rigorose Standardisierung der Versuchsbedingungen lässt

18 Richter, S. H. et al. (2009): Environmental standardization: cure or cause of poor reproducibility in animal experiments? Nature Methods, 6: 257 – 261.

sich deshalb allenfalls die Tierzahl pro Versuch minimieren, dafür sind mehr unabhängige Versuche – und damit letztlich mehr Tiere – erforderlich, um die Reproduzierbarkeit bei variierenden Bedingungen und damit die Generalisierbarkeit nachzuweisen.[19]

Forschende sollten deshalb Versuche möglichst so planen, dass auch Rückschlüsse auf die externe Validität (das heißt auf die Generalisierbarkeit über die Geschlechter sowie genetische und umweltbedingte Variation) möglich sind. Dazu eignet sich die kontrollierte Heterogenisierung der Versuchspopulation mittels mehrfaktorieller Versuchsdesigns, mit denen gezielt die Geschlechter, mehrere Genotypen und/oder mehrere Umwelt- beziehungsweise Versuchsbedingungen berücksichtigt werden können, ohne dass dafür mehr Tiere erforderlich sind.[20] Besonders robuste Ergebnisse können mit Multi-Labor-Studien erzielt werden.[21] Obwohl nicht immer umsetzbar, erscheint dies insbesondere in großen Forschungsverbünden eine sinnvolle Ergänzung der Forschungsmethodik.

- *Forschende sollten bei der Versuchsplanung externe Validität, das heißt die Replizierbarkeit und Generalisierbarkeit berücksichtigen und das Studiendesign entsprechend begründen.*

4.4 Pilot- und Explorationsversuche

Viele Versuchsvorhaben sind Teil umfangreicher Forschungsprogramme. Dabei werden in einem ersten Schritt oft explorative Studien zur Generierung von Hypothesen durchgeführt, gefolgt von Pilotstudien zur Optimierung des Versuchsplans im Hinblick auf konfirmatorische Studien zur Über-

19 Voelkl, B., Vogt, L., Sena, E. and Würbel, H. (2018): Reproducibility of preclinical animal research improves with heterogeneity of study samples. PLoS Biology, 16(2), e2003692.

20 Shaw, R., Festing, M. F. W., Peers, I. and Furlong, L. (2002): Use of Factorial Designs to Optimize Animal Experiments and Reduce Animal Use. ILAR Journal, 43(4): 223 – 232.

21 Voelkl, B., Vogt, L., Sena, E. and Würbel, H. (2018): a. a. O.

prüfung konkreter Hypothesen. Für Explorations- und Pilotstudien sowie in manchen Bereichen der Grundlagenforschung können keine genauen Stichprobenberechnungen durchgeführt werden.[22] Die Studien dienen jedoch nicht nur der Generierung vielversprechender Hypothesen und der Optimierung von Versuchsdesigns, sondern tragen auch zur Minimierung der Tierzahl im Rahmen sich anschließender umfangreicher Forschungsprogramme bei. Andernfalls besteht das Risiko, unergiebige Hypothesen zu testen oder Hypothesen mit unausgereiften Versuchsdesigns und ungeeigneten Methoden zu prüfen, wodurch die wissenschaftliche Aussagekraft erheblich beeinträchtigt werden kann.

- *Explorations- und Pilotstudien sollten so gestaltet werden, dass sie im Hinblick auf weiterführende Studien beziehungsweise umfassende Forschungsprogramme zu einer Fokussierung und damit zu einer Minimierung der insgesamt benötigten Tierzahl führen.*

4.5 Wiederholungsversuche

Gemäß geltendem Tierschutzgesetz sind Tierversuche nur dann unerlässlich, wenn sie den jeweiligen Stand der wissenschaftlichen Erkenntnisse erweitern und der Erkenntnisgewinn die Belastungen der Tiere überwiegt. Dies ist auch bei der Legitimation von Wiederholungsversuchen zur Überprüfung der Replizierbarkeit von Versuchsergebnissen zu berücksichtigen.

Replizierbarkeit ist jedoch ein Grundpfeiler des wissenschaftlichen Erkenntnisgewinns. Ohne Überprüfung von Forschungsergebnissen besteht ein hohes Risiko unergiebiger Folgeversuche auf der Basis vielversprechender, jedoch nicht replizierbarer Ergebnisse. Dies spricht einerseits für die Verwendung von Versuchsansätzen, die eine Abschätzung der externen Validität und Replizierbarkeit von Versuchsergebnissen ermöglichen (beispielsweise mehrfaktorielle Designs, Multi-Labor-Studien; siehe Abschnitt 4.3). Anderer-

22 Mayer, B., Muche, R. (2013): Formal sample size calculation and its limited validity in animal studies of medical basic research. Tierärztliche Praxis, Ausg K Kleintiere Heimtiere, 41(6): 367–374.

seits muss jedoch auch eine unabhängige Überprüfung der Replizierbarkeit bestehender Erkenntnisse möglich sein.

- *Forschende sollten Wiederholungsversuche (zum Beispiel zur Überprüfung der Replizierbarkeit von Versuchsergebnissen) in Erwägung ziehen und begründen. Unter Umständen können notwendige Wiederholungsversuche als Bestandteil neuer Versuche in Form von Positiv- oder Negativkontrollen eingeplant werden.*

5 Anhang

Reporting Guidelines

ARRIVE Guideline (NC3Rs)
Kilkenny, C. et al. (2010): *Improving bioscience research reporting: The ARRIVE guidelines for reporting animal research.* PLoS Biology, 8(6), e1000412.
https://www.nc3rs.org.uk/arrive-guidelines

HARRP Guidelines (ICLAS)
Osborne, N. et al. (2018): *Improving animal reporting standards.* EMBO Reports, 19(5), e46069.
http://iclas.org/

PREPARE Guidelines (NORECOPA)
Smith, A. et al. (2017): *PREPARE: Guidelines for planning animal research and testing.* Laboratory Animals, 52(2).
https://norecopa.no/prepare

Übersicht über weitere Reporting Guidelines:
https://www.nlm.nih.gov/services/research_report_guide.html

Datenbanken zur Prä-Registrierung von tierexperimentellen Studien

Animal Study Registry (Bundesinstitut für Risikobewertung)
https://www.animalstudyregistry.org/

PreClinicalTrials.EU
https://www.preclinicaltrials.eu/

Weitere Information und Links

What are the 3Rs? (NC3Rs)
https://www.nc3rs.org.uk/the-3rs

Experimental Design Assistant (NC3Rs)
https://eda.nc3rs.org.uk/

G*Power (Heinrich-Heine Universität, Düsseldorf)
http://www.gpower.hhu.de/

6 Mitglieder der Ständigen Senatskommission für tierexperimentelle Forschung der Deutschen Forschungsgemeinschaft

Professorin Dr. Brigitte Vollmar, Rostock (Vorsitz)

Professor Dr. Peter Dabrock, Erlangen

Professor Dr. Bernd Fleischmann, Bonn

Professorin Dr. Christiane Herden, Gießen

Professor Dr. Nils Hoppe, Hannover

Professor Dr. Lars Lewejohann, Berlin

Professorin Dr. Heidrun Potschka, München

Professor Dr. Stefan Schlatt, Münster

Professor Dr. René H. Tolba, Aachen

Professor Dr. Stefan Treue, Göttingen

Professor Dr. Hanno Würbel, Bern

Kommissionsekretariat

Dr. Cornelia Exner, Marburg

Dr. Valeska Stephan, Rostock

DFG-Geschäftsstelle

Dr. Sonja Ihle, Bonn

Deutsche Forschungsgemeinschaft

Kennedyallee 40 • 53175 Bonn

Postanschrift: 53170 Bonn

Telefon: +49 228 885-1

Telefax: +49 228 885-2777

postmaster@dfg.de

www.dfg.de