

4 Ergebnisse

4.1 Literatursuche

4.1.1 Zur Bewertung gelangte Publikationen

Insgesamt erfüllten 287 Publikationen zur Fortpflanzungsmedizin des Hundes die im Vorfeld definierten Einschlusskriterien (siehe Kapitel 3.2.1). Diese Fachartikel wurden in die Literaturliste aufgenommen und gelangten im Anschluss zur Bewertung.

Die Literaturrecherche erfolgte einerseits in den zwei bibliographischen Datenbanken PubMed und Vet-CD des Internets. Auf der anderen Seite wurden weitere Fachartikel über eine Recherche in den Literaturverzeichnissen bereits gefundener Publikationen gesichtet. Eine Übersicht über die Quelle der bibliographischen Daten der jeweiligen Fachartikel ist in Tabelle 5 dargestellt.

Tabelle 5: Anzahl der Fachartikel bezogen auf die Quelle der Literaturrecherche

Quelle	Anzahl der Publikationen	Anteil an der Gesamtzahl
Datenbanken	226	78,7 %
Literaturverzeichnisse	61	21,3 %
Summe	287	100,0 %

Die Verteilung der Fachartikel auf die beiden Datenbanken PubMed und Vet-CD ergibt sich aus Tabelle 6.

Tabelle 6: Anzahl der Fachartikel bezogen auf die Datenbanken

Datenbank	Anzahl der Publikationen	Anteil an der Gesamtzahl
PubMed	164	72,6 %
Vet-CD	62	27,4 %
Summe	226	100,0 %

Die 287 gefundenen Fachartikel entstammten 64 Fachzeitschriften und einem Kongressbericht (Proceeding). Eine Übersicht über die Anzahl der Fachartikel aus den jeweiligen Fachzeitschriften bzw. dem Kongressbericht befindet sich im Anhang (siehe Anhang 9.3). Eine Auflistung der Fachzeitschriften, denen mindestens zehn Publikationen entstammten, ist in Tabelle 7 dargestellt.

Tabelle 7: Fachzeitschriften mit mindestens zehn Publikationen

Fachzeitschrift	Herausgeber, Land und Zeitraum der Erscheinung	Impaktfaktor nach ISI Web of Knowledge SM (2007)
Journal of the American Animal Hospital Association	American Animal Hospital Association, USA, 1968-2007	0.818
Journal of the American Veterinary Medical Association	American Veterinary Medical Association, USA, 1915-2007	1.266
Journal of Reproduction and Fertility Supplement	Journals of Reproduction and Fertility, Ltd., England, 1966-2007	2.958
The Indian Veterinary Journal	Rao Sahib T Vinayaka Mudaliar, Indien, 1924-2007	0.052
The Journal of Small Animal Practice	Blackwell Scientific Publications, England, 1960-2007	0.717
The Journal of Veterinary Medical Science	Japanese Society of Veterinary Science, Japan, 1991-2007	0.827
Theriogenology	Elsevier, USA, 1974-2007	1.898
The Veterinary Record	British Veterinary Medicine, England, 1888-2007	1.168

Gemäß dem Einschlusskriterium Sprache waren 271 Publikationen englischsprachig (94,4 %) und 16 Artikel deutschsprachig (5,6 %). Eine Kategorisierung nach dem Einschlusskriterium Publikationstyp ergab für PubMed und Vet-CD insgesamt 199 Fallberichte (69,3 %) und 88 klinische Studien (30,7 %). Von den 88 klinischen Studien wurden 27 Studien (30,7 %) über die Recherche in der Datenbank PubMed ermittelt, während weitere 61 Studien (69,3 %) aus der Recherche in Literaturverzeichnissen hervorgingen. In der Handsuche wurden gezielt nur klinische Studien ausgewählt, da diese nach der Recherche in den Literaturdatenbanken des Internets unterrepräsentiert waren.

Die Fachartikel wurden über das Internet, veterinärmedizinische Bibliotheken und per Fernleihe beschafft. Tabelle 8 zeigt die Anzahl der Fachartikel geordnet nach dem jeweiligen Modus der Beschaffung.

Tabelle 8: Anzahl der Fachartikel bezogen auf den Modus der Beschaffung

Quelle der Beschaffung	Anzahl der Publikationen	Anteil an der Gesamtzahl
Download aus dem Internet	128	44,6 %
Bibliothek Berlin Düssel, Fachbereich Veterinärmedizin	111	38,7 %
Bibliothek des Instituts für Zoo- und Wildtierkunde	16	5,6 %
Fernleihe mit Hilfe der Bibliothek der Freien Universität Berlin	32	11,1 %
Summe	287	100,0 %

Alle zur Bewertung gelangten Publikationen wurden einem Themenkomplex zugeteilt (siehe Kapitel 3.3.4). Im Bereich „Andrologie“ gelangten 88 Publikationen (30,7 %) zur Bewertung. Während dem Gebiet „Gynäkologie“ 129 Publikationen (44,9 %) zugeordnet werden konnten, repräsentierten 70 Publikationen (24,4 %) das Thema „Trächtigkeit und Geburt“. Aus Gründen der Übersichtlichkeit wurden die Themenkomplexe „Andrologie“ und „Gynäkologie“ in weitere untergeordnete Themenbereiche aufgegliedert (siehe Kapitel 3.3.4). Eine Aufstellung der thematischen Zuordnung und Anzahl der zugehörigen Fachartikel zeigt Tabelle 9.

Tabelle 9: Anzahl der Fachartikel nach Zuordnung zu Themenkomplex und Themenbereich

Themenkomplex	Untergeordneter Themenbereich	Anzahl der Publikationen	Anteil an der Gesamtzahl
Andrologie	Hoden und Nebenhoden	29	10,1 %
	Penis und Präputium	9	3,1 %
	Prostata	24	8,4 %
	Sonstige	13	4,5 %
	Spermatologie	13	4,5 %
Gynäkologie	Ovar	18	6,3 %
	Sonstige	31	10,8 %
	Uterus	56	19,5 %
	Vagina und Vulva	24	8,4 %
Trächtigkeit und Geburt	-	70	24,4 %
Summe	-	287	100,0 %

4.1.2 Von der Bewertung ausgeschlossene Publikationen

Im Laufe der Internetrecherche wurden insgesamt 109 Publikationen (27,5 %) gesichtet, die von der Bewertung ausgeschlossen werden mussten. Sie erfüllten die vorab definierten Ausschlusskriterien und wurden nach anfänglicher Aufnahme in eine Vorauswahlliste im Anschluss von der Bewertung ausgeschlossen. Eine Übersicht über diese 109 Publikationen befindet sich im Anhang (siehe Anhang 9.4).

Fachartikel, die in PubMed zur Physiologie und Pathologie des Harnapparates ohne Bezug zum Geschlechtsapparat in die Suche eingeschlossen worden waren, wurden ausselektiert. Es handelte sich um insgesamt 64 Publikationen (16,2 %).

Das Ausschlusskriterium humanmedizinische Forschung traf auf 17 Publikationen (4,3 %) zu. Ihr Inhalt diene lediglich dem Erkenntnisgewinn in der Humanmedizin. Alle 17 Publikationen wurden bei der Recherche in der Datenbank PubMed gelistet.

Bei 14 Veröffentlichungen (3,5 %) handelte es sich um nicht verfügbare Publikationen. Sie stammten alle aus der Datenbank Vet-CD.

Das Ausschlusskriterium Indexierungsfehler/ falscher Publikationstyp traf für neun Publikationen zu (2,3 %). Es handelte sich um fünf Reviews (1,3 %), drei Briefe (0,8 %) und eine Metaanalyse (0,3 %), die fälschlicherweise in den Suchergebnissen von PubMed erschienen. Der Fachartikel, der zunächst als Metaanalyse gelistet war, stellte sich nach seiner Lektüre als Review ohne statistische Komponente heraus. Diese Publikationen gelangten nicht zur Bewertung.

Das Ausschlusskriterium Indexierungsfehler/ falsche Spezies erfüllten drei Publikationen (0,8 %). In einer Publikation (0,3 %) aus PubMed wurde über die Katze berichtet, während sich zwei Publikationen (0,5 %) aus Vet-CD mit dem Rind beschäftigten.

Außerdem wurden zwei Publikationen (0,5 %) aufgrund des Ausschlusskriteriums Indexierungsfehler/ falsche Sprache von der Bewertung ausgeschlossen. Beide stammten aus der Suche in Vet-CD. Ihre Zusammenfassungen waren zwar in englischer Sprache verfasst, die Volltexte jedoch auf koreanisch geschrieben.

4.2 Literaturbewertung

4.2.1 Ergebnisse der Bewertung in der Kategorie „Material und Methodik“

Die Ergebnisse für die einzelnen Bewertungskriterien der Kategorie „Material und Methodik“ ergeben sich aus Tabelle 10. Die am häufigsten angekreuzten Beurteilungen für die einzelnen Bewertungskriterien werden im Anschluss kurz angeführt.

Tabelle 10: Qualität der Publikationen bezogen auf die Kategorie „Material und Methodik“

Bewertungskriterium	Antwort					
	trifft voll und ganz zu	trifft zu	trifft mäßig zu	trifft nicht zu	trifft gar nicht zu	nicht zu be- urteilen
Untersuchungsziel	34 11,8 %	101 35,2 %	58 20,2 %	24 8,4 %	42 14,6 %	28 9,8 %
Materialbeschreibung	0 0,0 %	114 39,7 %	149 51,9 %	24 8,4 %	0 0,0 %	0 0,0 %
Untersuchungen	14 4,9 %	82 28,6 %	113 39,4 %	73 25,4 %	5 1,7 %	0 0,0 %
Behandlungen	36 12,5 %	72 25,1 %	33 11,5 %	88 30,7 %	10 3,5 %	48 16,7 %
Nachbeobachtungsphase	6 2,1 %	99 34,5 %	65 22,6 %	63 22,0 %	29 10,1 %	25 8,7 %
Ergebnisdarstellung	31 10,8 %	152 53,0 %	70 24,4 %	34 11,8 %	0 0,0 %	0 0,0 %
Ergebnisdiskussion	10 3,5 %	117 40,8 %	56 19,5 %	77 26,8 %	27 9,4 %	0 0,0 %

Die Aussage „das Untersuchungsziel ist konkret formuliert“ konnte in vielen Fällen mit „trifft zu“ beantwortet werden (35,2 %). Eine zusammenfassende Bewertung der Materialbeschreibung, die sich aus den Kriterien Probandenanzahl, Auswahlkriterien, Haltung, Rasse, Alter, Geschlecht sowie Vorerkrankungen und Vorbehandlungen ergab, führte in der Mehrzahl der Fälle zur Bewertung „trifft mäßig zu“ (51,9 %). Die Angabe der Probandenanzahl erfolgte in nahezu allen Publikationen eindeutig und genau, so dass hier die Bewertung „trifft voll und ganz zu“ angekreuzt werden konnte (99,7 %). Die Auswahlkriterien der Tiere waren in der Mehrzahl der Publikationen mäßig beschrieben (62,0 %). Häufig erfolgten keine Angaben zur Haltung der Tiere, so dass diese Angabe überwiegend mit „trifft gar nicht zu“ beurteilt wurde (79,8 %). Die Rasse der Tiere wurde in der überwiegenden Anzahl der Fachartikel explizit genannt (72,1 %). Dies führte zur Bewertung „trifft voll und ganz zu“. Ebenso wurde das Alter der Tiere in der überwiegenden Anzahl der Fachartikel eindeutig angegeben, so dass die Bewertung „trifft voll und ganz zu“ angekreuzt werden konnte (70,4 %). Schließlich gelangte auch in der Mehrzahl der Fälle die Aussage „das Geschlecht der Tiere ist angegeben“ zur Bewertung „trifft voll und ganz zu“ (88,5 %). Auf Informationen zu Vorerkrankungen und Vorbehandlungen der Tiere wurde nur selten hingewiesen. Hier wurde in mehr als der Hälfte der Publikationen „trifft gar nicht zu“ angekreuzt (55,1 %). Die Beschreibung der Untersuchungen wurde oft als mäßig eingestuft (39,4 %). Für die Bewertung der Beschreibung durchgeführter Behandlungen waren einerseits vollständige Angaben zum eingesetzten Arzneimittel bzw. zur vorgenommenen Interventionsmethode wichtig. Diese Angaben erfolgten oft nur mäßig (35,2 %). Auf der anderen Seite spielte die genaue Schilderung der Anwendung des Prüfpräparates bzw. Durchführung der Intervention eine Rolle. Hier konnte für viele Publikationen „trifft zu“ angekreuzt werden (22,3 %). Insgesamt wurde die Aussage „die Behandlungen sind ausführlich beschrieben“ für viele Publikationen mit „trifft nicht zu“ bewertet (30,7 %). Hingegen erlangte ein großer Anteil der Publikationen für die Ausführlichkeit der Beschreibung der Nachbeobachtungsphase die Bewertung „trifft zu“ (34,5 %). Die Ergebnisse wurden bezüglich ihrer Darstellung und Diskussion beurteilt. Für über die Hälfte der Publikationen konnte die Vollständigkeit der Ergebnisdarstellung mit „trifft zu“ bewertet werden (53,0 %). Des Weiteren nahmen viele Publikationen auch eine kritische Diskussion ihrer Ergebnisse vor, so dass dieses Kriterium mit „trifft zu“ bewertet wurde (40,8 %).

4.2.2 Ergebnisse der Bewertung in der Kategorie „Studiendesign“

Für das Vorhandensein von Kontrollgruppe, Randomisierung und Verblindung wurde ein nur dreistufiges Antwortschema vorgegeben. Bei Fallberichten bzw. Expertenmeinungen wurde für diese Kriterien „nicht zu beurteilen“ angekreuzt (68,6 %). Bezogen auf die Gesamtheit aller Fachartikel fanden nur 20,9 % der Untersuchungen unter kontrollierten Bedingungen statt. Während eine Randomisierung nur in 8,0 % stattfand, waren lediglich 2,8 % aller Untersuchungen verblindet (siehe Tabelle 11).

Tabelle 11: Qualität der Publikationen bezogen auf die Kategorie „Studiendesign“

Bewertungskriterium	Antwort		
	trifft voll und ganz zu	trifft gar nicht zu	nicht zu beurteilen
Kontrollgruppe	60 20,9 %	30 10,5 %	197 68,6 %
Randomisierung	23 8,0 %	67 23,3 %	197 68,6 %
Verblindung	8 2,8 %	82 28,6 %	197 68,6 %

Bezogen auf den Anteil klinischer Studien an allen Fachartikeln (31,4 %) handelte es sich bei 66,7 % der Studien um Untersuchungen mit Kontrollgruppe. Eine Randomisierung nahmen nur 25,6 % der klinischen Studien vor. Noch seltener wurden die durchgeführten Studien verblindet (8,9 %) (siehe Tabelle 12). Für fünf klinische Studien (5,6 %) trafen alle drei Kriterien gleichzeitig zu. Es handelte sich dann um randomisierte kontrollierte klinische Studien mit Verblindung.

Tabelle 12: Studiendesign bezogen auf die Gesamtzahl klinischer Studien

Studiendesign	Anteil an der Gesamtzahl klinischer Studien
Klinische Studie mit Kontrollgruppe	66,7 %
Klinische Studie mit Randomisierung	25,6 %
Klinische Studie mit Verblindung	8,9 %
Klinische Studie mit Kontrollgruppe, Randomisierung und Verblindung	5,6 %

Für die Beurteilung der Beschreibung des Studiendesigns insgesamt spielten die einzelnen Bewertungskriterien Pro-/ Retrospektivität, Angemessenheit der Kontrollgruppe sowie die Beschreibung der Art der Randomisierung und Verblindung eine Rolle (siehe Tabelle 13).

Tabelle 13: Qualität der Publikationen bezogen auf die Kategorie „Studiendesign“

Bewertungskriterium	Antwort					
	trifft voll und ganz zu	trifft zu	trifft mäßig zu	trifft nicht zu	trifft gar nicht zu	nicht zu beurteilen
Studiendesign insgesamt	0 0,0 %	24 8,4 %	55 19,2 %	11 3,8 %	0 0,0 %	197 68,6 %
Pro-/ Retrospektivität	11 3,8 %	79 27,5 %	0 0,0 %	0 0,0 %	0 0,0 %	197 68,6 %
Art der Kontrollgruppe	2 0,7 %	6 2,1 %	17 5,9 %	29 10,1 %	5 1,7 %	228 79,4 %
Art der Randomisierung	0 0,0 %	1 0,3 %	0 0,0 %	3 1,0 %	20 7,0 %	263 91,6 %
Art der Verblindung	1 0,3 %	7 2,4 %	0 0,0 %	0 0,0 %	1 0,3 %	278 96,9 %

Die zusammenfassende Bewertung der Beschreibung des Studiendesigns gelangte in den meisten Fällen zur Antwort „nicht zu beurteilen“ (68,6 %). War eine Beurteilung möglich, so wurde die ausführliche Beschreibung des Studiendesigns insgesamt häufig mit „trifft mäßig zu“ beurteilt (19,2 %). Das Kriterium der Pro-/ Retrospektivität war für Fallberichte und Expertenmeinungen erneut nicht zu beurteilen (68,6 %). In den meisten klinischen Studien konnte jedoch hier ein „trifft zu“ angekreuzt werden (27,5 %). Die Angemessenheit der Kontrollgruppe war nur in den wenigsten Fällen gegeben. Bei der Mehrzahl der Publikationen war dieses Kriterium nicht zu beurteilen (79,4 %). Bezogen auf den Anteil kontrollierter klinischer Studien (66,7 %) wurde nur für 3,3 % der Studien mit Kontrollgruppe das Kriterium „die Kontrollgruppe ist angemessen gewählt“ mit „trifft voll und ganz zu“ bewertet. Auch bezüglich der Beschreibung der Art der Randomisierung wurde in den meisten Fällen „nicht zu beurteilen“ angekreuzt (91,6 %). Bezogen auf den Anteil randomisierter klinischer Studien (25,6 %) konnte für die Beschreibung der Art der Randomisierung in keinem Fall „trifft voll und ganz zu“ angekreuzt werden (0,0 %). Bei der Beschreibung der Art der Verblindung ergab sich eine vergleichbare Situation. Für die meisten Fachartikel wurde „nicht zu beurteilen“ angekreuzt (96,9 %). In klinischen Studien mit Verblindung (8,9 %) war die Art der Verblindung in 12,5 % der Fachartikel angemessen beschrieben, so dass ein „trifft voll und ganz zu“ angekreuzt werden konnte.

4.2.3 Ergebnisse der Bewertung in der Kategorie „Statistik“

Auch in der Kategorie „Statistik“ konnte für Fallberichte keine Beurteilung vorgenommen werden (68,6 %). Für Publikationen, in denen jedoch eine Bewertung möglich war (siehe Tabelle 14), sind deren Ergebnisse im Folgenden kurz genannt.

Tabelle 14: Qualität der Publikationen bezogen auf die Kategorie „Statistik“

Bewertungskriterium	Antwort					
	trifft voll und ganz zu	trifft zu	trifft mäßig zu	trifft nicht zu	trifft gar nicht zu	nicht zu be- urteilen
Statistische Methoden	50 17,4 %	9 3,1 %	4 1,4 %	0 0,0 %	27 9,4 %	197 68,6 %
Stichprobenumfang	9 3,1 %	37 12,9 %	15 5,2 %	24 8,4 %	5 1,7 %	197 68,6 %
Signifikanzniveau	59 20,6 %	3 1,0 %	0 0,0 %	0 0,0 %	28 9,8 %	197 68,6 %
Umgang mit Daten	2 0,7 %	13 4,5 %	17 5,9 %	26 9,1 %	32 11,1 %	197 68,6 %
Ergebnisanalyse	3 1,0 %	57 19,9 %	25 8,7 %	4 1,4 %	1 0,3 %	197 68,6 %
Statistik insgesamt	2 0,7 %	44 15,3 %	17 5,9 %	25 8,7 %	2 0,7 %	197 68,6 %

Die Anwendung korrekter statistischer Methoden traf für 17,4 % aller Fachartikel voll und ganz zu. Bezogen auf den Anteil klinischer Studien (31,4 %) waren die statistischen Methoden zu 55,6 % korrekt beschrieben. Ein angemessener Stichprobenumfang traf für viele klinischen Studien ebenfalls zu (41,1 %). Bei der Betrachtung aller Fachartikel traf die Beurteilung „trifft zu“ für lediglich 12,9 %. Auch die Wahl eines angemessenen Signifikanzniveaus traf für die Mehrzahl der klinischen Studien voll und ganz zu (65,6 %). Bezogen auf die Gesamtanzahl der Fachartikel war das Signifikanzniveau zu 20,6 % angemessen gewählt worden. Ein nachvollziehbarer Umgang mit fehlenden, nicht

verwendeten oder fraglichen Daten traf für 11,1 % aller Fachartikel und 35,6 % aller klinischen Studien gar nicht zu. Dagegen erfolgte die Ergebnisanalyse für 19,9 % aller Fachartikel und 63,3 % aller klinischen Studien angemessen. Insgesamt konnte das Kriterium „die Statistik ist konkret und nachvollziehbar beschrieben“ zu 15,3 % aller Fachartikel bzw. der Hälfte aller klinischen Studien (48,9 %) mit „trifft zu“ beurteilt werden.

4.2.4 Ergebnisse der Bewertung in der Kategorie „Präsentation und Informationsgehalt“

Die Ergebnisse der Bewertung in der Kategorie „Präsentation und Informationsgehalt“ sind in Tabelle 15 dargestellt.

Tabelle 15: Qualität der Publikationen bezogen auf die Kategorie „Präsentation und Informationsgehalt“

Bewertungskriterium	Antwort					
	trifft voll und ganz zu	trifft zu	trifft mäßig zu	trifft nicht zu	trifft gar nicht zu	nicht zu be- urteilen
Objektivität des Autors	5 1,7 %	267 93,0 %	13 4,5 %	2 0,7 %	0 0,0 %	0 0,0 %
Zusammenfassung	63 22,0 %	117 40,8 %	20 7,0 %	6 2,1 %	81 28,2 %	0 0,0 %
Bisherige Studien	21 7,3 %	172 60,0 %	70 24,4 %	22 7,7 %	2 0,7 %	0 0,0 %
Literaturverzeichnis	10 3,5 %	102 35,5 %	95 33,1 %	59 20,6 %	21 7,3 %	0 0,0 %

In der Kategorie „Präsentation und Informationsgehalt“ wurde zunächst die Objektivität des Autors bei seiner Schilderung des Themas bewertet. Sie traf für den überwiegenden Anteil der Autoren zu (93,0 %). Auch der Aussage „es ist eine Zusammenfassung vorhanden, die über die wesentlichen Aspekte des Artikels informiert“ konnte in vielen Fällen mit einem „trifft zu“ zugestimmt werden (40,8 %). Des Weiteren traf auch zu, dass bisherige Studien bzw. Untersuchungen zum jeweiligen Thema in den meisten Publikationen angesprochen worden waren (60,0 %). In vielen Fällen war ein angemessenes Literaturverzeichnis vorhanden, so dass hier „trifft zu“ angekreuzt wurde (35,5 %).

4.2.5 Ergebnisse der Bewertung in der Kategorie „Praktische Anwendbarkeit“

Die Qualität der Publikationen in der Kategorie „Praktische Anwendbarkeit“ wurde anhand von vier Bewertungskriterien ermittelt. Die Ergebnisse der Bewertung gehen aus Tabelle 16 hervor.

Tabelle 16: Qualität der Publikationen bezogen auf die Kategorie „Praktische Anwendbarkeit“

Bewertungskriterium	Antwort					
	trifft voll und ganz zu	trifft zu	trifft mäßig zu	trifft nicht zu	trifft gar nicht zu	nicht zu beurteilen
Praxisrelevanz	13 4,5 %	204 71,1 %	59 20,6 %	11 3,8 %	0 0,0 %	0 0,0 %
Praktische Umsetzung	0 0,0 %	0 0,0 %	29 10,1 %	237 82,6 %	21 7,3 %	0 0,0 %
Therapiealternativen	0 0,0 %	23 8,0 %	46 16,0 %	47 16,4 %	165 57,5 %	6 2,1 %
Nutzen-Risiko-Analyse	0 0,0 %	40 13,9 %	82 28,6 %	72 25,1 %	93 32,4 %	0 0,0 %

Eine Relevanz der Informationen für die Praxis oder Wissenschaft traf für die Mehrzahl der Publikationen zu (71,1 %). Eine ausreichende Schilderung der praktischen Umsetzung der Informationen traf jedoch meistens nicht zu (82,6 %). Auch eine ausführliche und kritische Diskussion von Therapiealternativen traf für die Mehrzahl der Publikationen gar nicht zu (57,5 %). Ebenso traf der ausreichende Hinweis auf Nebenwirkungen, Misserfolge und Komplikationen für viele der Publikationen gar nicht zu (32,4 %).

4.2.6 Ergebnisse der Bewertung in der Kategorie „Abschließende Bewertung“

In der abschließenden Bewertung wurde jeder Fachartikel bezüglich seiner Gesamtqualität kritisch betrachtet. Die Bewertung, ob die Publikation allgemeingültige und wissenschaftlich abgesicherte Schlussfolgerungen rechtfertigt, musste für viele Publikationen mit „trifft nicht zu“ (49,8 %) beantwortet werden (siehe Tabelle 17).

Tabelle 17: Qualität der Publikationen bezogen auf die Kategorie „Abschließende Bewertung“

Bewertungskriterium	Antwort				
	trifft voll und ganz zu	trifft zu	trifft mäßig zu	trifft nicht zu	trifft gar nicht zu
Rechtfertigung gesicherter Schlussfolgerungen	1 0,3 %	35 12,2 %	56 19,5 %	143 49,8 %	52 18,1 %

Außerdem wurde jede Publikation einer Evidenzstufe zugeordnet. Eine Übersicht über die Anzahl der Fachartikel bezogen auf die Evidenzstufe ist in Tabelle 18 dargestellt. In der Mehrzahl der Fälle handelte es sich um Fallberichte (67,9 %).

Tabelle 18: Anzahl der Fachartikel bezogen auf die Evidenzstufe

Evidenzstufe	Anzahl der Publikationen	Anteil an der Gesamtzahl
Metaanalyse	0	0,0 %
Randomisierte, kontrollierte Studie	21	7,3 %
Kontrollierte Studie	37	12,9 %
Unkontrollierte Studie	18	6,3 %
Beobachtungsstudie	14	4,9 %
Fallbericht	195	67,9 %
Berichte, Meinungen oder klinische Erfahrungen von Experten	2	0,7 %
Nicht zu beurteilen	0	0,0 %
Summe	287	100,0 %

Die Verteilung verschiedener Evidenzstufen auf die drei Themenkomplexe zeigt Tabelle 19.

Tabelle 19: Anzahl der Fachartikel einzelner Evidenzstufen bezogen auf das Thema

Evidenzstufe	Andrologie	Gynäkologie	Trächtigkeit/ Geburt
Randomisierte, kontrollierte Studie	4 1,4 %	4 1,4 %	13 4,5 %
Kontrollierte Studie	9 3,1 %	15 5,2 %	13 4,5 %
Unkontrollierte Studie	4 1,4 %	9 3,1 %	5 1,7 %
Beobachtungsstudie	3 1,0 %	7 2,4 %	4 1,4 %
Fallbericht	68 23,7 %	93 32,4 %	34 11,8 %
Berichte, Meinungen oder klinische Erfahrungen von Experten	0 0,0 %	1 0,3 %	1 0,3 %
Summe	88 30,7 %	129 44,9 %	70 24,4 %

4.3 Pretest

Von 13 angeschriebenen Teilnehmern füllten sechs Personen die Fragebögen aus. Es handelte sich um vier Fachleute der Universitäten in Giessen, Hannover, Leipzig und Zürich sowie zwei Mitarbeiter des eigenen Instituts. Damit ergibt sich eine Rücklaufquote von 46,2 %. Für die Auswertung des Pretests wurden zusätzlich die Ergebnisse der Artikelbewertung durch die Doktorandin hinzugezogen. Die Bewertungsergebnisse der Befragung aller Pretest-Teilnehmer wurden verglichen, um Übereinstimmungen und Abweichungen im Ankreuzverhalten der Bewerter festzustellen. Wo die verschiedenen Bewerter ihre Kreuzchen bei der Beurteilung der drei Fachartikel setzten, ist dem Anhang zu entnehmen (siehe Anhang 9.5). Die Auswertungen erfolgten bezüglich der drei Bereiche Zustimmung bzw. Ablehnung der Aussage und Wahl der neutralen Mitte. Die Antwortmöglichkeiten „trifft voll und ganz zu“ oder „trifft zu“ wurden dazu dem Bereich Zustimmung der Aussage zugeordnet. Eine Ablehnung der Aussage traf zu, sofern „trifft nicht zu“ oder „trifft gar nicht zu“ angekreuzt wurden. Die Antwort „trifft mäßig zu“ repräsentierte die neutrale Mitte.

4.3.1 Übereinstimmungen im Ankreuzverhalten aller Bewerter des Pretests

Für den ersten Artikel („Artificial Insemination with Canine Semen Stored at a Low Temperature“, siehe Anhang 9.5.1) stimmten in der Beurteilung von Probandenanzahl, Nachbeobachtungsphase, Ergebnisdarstellung, Vorhandensein von Kontrollgruppe und Randomisierung, Signifikanzniveau, Objektivität des Autors, Zusammenfassung, Relevanz der Informationen und Rechtfertigung von Schlussfolgerungen alle Bewerter in ihrem Ankreuzverhalten überein (maximale Übereinstimmung aller Bewerter von 100,0 %). Bei der Einordnung in eine Evidenzstufe waren ebenfalls alle Teilnehmer des Pretests einer Meinung. Jeder ordnete dem Fachartikel die Evidenzstufe „kontrollierte klinische Studie“ zu. In den Kriterien Untersuchungsziel, Haltung, Geschlecht, Vorerkrankungen und Vorbehandlungen, Beschreibung der Untersuchungen, Prüfpräparat/ Interventionsmethode, Beschreibung der Behandlungen insgesamt, Verblindung, statistische Methoden, Stichprobenumfang sowie bisherige Studien/ Untersuchungen war eine maximale Übereinstimmung aller Bewerter von 85,7 % gegeben, d.h. maximal sechs der sieben Bewerter setzten ihre Kreuzchen im gleichen Antwortbereich. Die Kriterien Anwendung des Prüfpräparates/ Durchführung der Intervention, Ergebnisdiskussion, Art der Randomisierung und Verblindung, Beschreibung des Studiendesigns insgesamt, Ergebnisanalyse, Beschreibung der Statistik insgesamt sowie Literaturverzeichnis beurteilten die Bewerter mit einer maximalen Übereinstimmung von 71,4 %. Des Weiteren stimmten maximal vier von sieben Fachleuten in der Bewertung der Auswahlkriterien, Rasse und Alter der Tiere, Materialbeschreibung insgesamt sowie der Pro- bzw. Retrospektivität und Angemessenheit der Kontrollgruppe in ihren Meinungen überein (maximale Übereinstimmung von 57,1 %).

Für den zweiten Artikel („Evaluation of Cabergolin and Buserelin Efficacy for Oestrous Induction in the Bitch“) wurde das Ankreuzverhalten ebenfalls ausgewertet (siehe Anhang 9.5.2). In der Hälfte aller Bewertungskriterien stimmten die Bewerter zu 100,0 % in ihrem Ankreuzverhalten miteinander überein (Untersuchungsziel, Probandenanzahl, Haltung, Rasse, Alter, Geschlecht, Materialbeschreibung insgesamt, Behandlungen insgesamt inklusive Angaben zu Präparat/ Interventionsmethode und Anwendung/ Durchführung, Nachbeobachtungsphase, Ergebnisdarstellung, Randomisierung, statistische Methoden, Signifikanzniveau, Statistik insgesamt, Objektivität, Zusammenfassung, Relevanz der Informationen, Rechtfertigung von Schlussfolgerungen). In den Bewertungskriterien Untersuchungen, Ergebnisdiskussion, Pro- bzw. Retrospektivität, Studiendesign insgesamt, Stichprobenumfang, Ergebnisanalyse, bisherige Studien/ Untersuchungen, und Literaturverzeichnis lagen jeweils sechs der sieben Teilnehmer in ihrer Bewertung im gleichen Antwortbereich (maximale Übereinstimmung von 85,7 %). Der Fachartikel wurde von ebenfalls sechs Bewertern einheitlich der Evidenzklasse „randomisierte, kontrollierte klinische Studie“ zugeordnet. Während die Kriterien Auswahlkriterien der Tiere und Vorhandensein einer Verblindung von fünf Bewertern gleich beurteilt wurde (maximale Übereinstimmung von 71,4 %), lag die maximale Übereinstimmung aller Teilnehmer in den Kriterien Vorhandensein einer Kontrollgruppe, Art von Randomisierung und Verblindung sowie praktische Umsetzung der Informationen bei 57,1 %.

Der dritte Artikel („Cystic endometrial hyperplasia and endometritis in a dog following prolonged treatment of medroxyprogesterone acetate“) kam zu folgenden Ergebnissen (siehe Anhang 9.5.3). Insgesamt zehn Bewertungskriterien wurden in 100 %iger Übereinstimmung aller Bewerter miteinander beurteilt (Rasse, Alter, Geschlecht, Art von Kontrollgruppe, Randomisierung und Verblindung, Stichprobenumfang, Umgang mit fehlenden/ nicht verwendeten/ fraglichen Daten, Therapiealternativen, Zuordnung zu einer Evidenzstufe). Alle Bewerter ordneten diesen Fachartikel in die Evidenzklasse „Fallbericht“ ein. Bei acht weiteren Kriterien setzten die Teilnehmer des Pretests ihre Kreuzchen in maximaler Übereinstimmung von 85,7 % (Probandenanzahl, statistische Methoden, Signifikanzniveau, Ergebnisanalyse, Statistik insgesamt, Zusammenfassung, praktische Umsetzung,

Rechtfertigung von Schlussfolgerungen). In der Bewertung der Auswahlkriterien der Tiere, Vorhandensein von Kontrollgruppe, Randomisierung und Verblindung, Pro-/ Retrospektivität, Studiendesign insgesamt sowie Informationsrelevanz waren maximal fünf der Fachleute einer Meinung (maximale Übereinstimmung 71,4 %). Des Weiteren lag für die Bewertungskriterien Haltung der Tiere, Materialbeschreibung und Behandlungen insgesamt, Ergebnisdiskussion sowie Objektivität des Autors und Nutzen-Risiko-Analyse eine maximale Übereinstimmung von 57,1 % vor.

Eine detaillierte Darstellung der Übereinstimmungen aller Bewerter im Ankreuzverhalten der Artikelbewertung im Pretest ist dem Anhang 9.5 zu entnehmen.

4.3.2 Abweichungen im Ankreuzverhalten aller Bewerter des Pretests

Unter Berücksichtigung aller 40 Bewertungskriterien kam es pro Fragebogen durchschnittlich 34 mal vor, dass mehr als die Hälfte aller Teilnehmer des Pretests in ihrem Ankreuzverhalten im gleichen Bereich lag. In den Fällen, in denen maximal drei bzw. zwei Bewerter (maximale Übereinstimmung von 42,9 % oder 28,6 %) in ihrem Antwortverhalten übereinstimmten, wurde die Stärke des Abweichverhaltens betrachtet.

Für den ersten Artikel (siehe Anhang 9.5.1) lagen bei den Kriterien praktische Umsetzung der Informationen, Therapiealternativen und Nutzen-Risiko-Analyse nur maximal drei Bewerter im gleichen Antwortbereich (maximale Übereinstimmung von 42,9 %). Die Beschreibung der praktischen Umsetzung der Informationen empfanden beispielsweise jeweils zwei Bewerter als zufrieden stellend bzw. unzureichend, während drei eine neutrale Haltung einnahmen. Die Beurteilung des Umgangs mit fehlenden/ nicht verwendeten/ fraglichen Daten ergab die geringsten Übereinstimmungen. Nur maximal zwei der Pretest-Teilnehmer waren einer Meinung (maximale Übereinstimmung von 28,6 %). Dabei wichen sie jedoch insofern voneinander ab, als jeweils zwei Bewerter den Umgang mit Daten als nachvollziehbar, nicht nachvollziehbar bzw. nicht zu beurteilen einstufen.

Bei dem zweiten Artikel (siehe Anhang 9.5.2) kreuzten maximal drei der sieben Teilnehmer in den folgenden Bewertungskriterien im gleichen Antwortbereich an (maximale Übereinstimmung von 42,9 %): Vorerkrankungen und Vorbehandlungen, Angemessenheit der Kontrollgruppe, Umgang mit fehlenden/ nicht verwendeten/ fraglichen Daten, Therapiealternativen und Nutzen-Risiko-Analyse. Dabei wichen die Antworten beispielsweise für das Kriterium Vorerkrankungen und Vorbehandlungen besonders stark voneinander ab. Während die Angaben zu Vorerkrankungen und Vorbehandlungen für drei Bewerter ausreichend waren, empfanden ebenso viele Bewerter diese als unzureichend.

Die wenigsten Übereinstimmungen aller Bewerter ergaben sich für den dritten Artikel (siehe Anhang 9.5.3). In der Beurteilung der Kriterien Untersuchungsziel, Vorerkrankungen und Vorbehandlungen, Untersuchungen, Angaben/ Anwendung von Prüfpräparat/ Intervention, Nachbeobachtungsphase, Ergebnisdarstellung, bisherige Studien/ Untersuchungen und Angemessenheit des Literaturverzeichnisses wurde eine maximale Übereinstimmung von 42,9 % erreicht. Damit stimmten in neun von 40 Bewertungskriterien maximal drei Bewerter in ihrem Ankreuzverhalten überein. So gingen beispielsweise die Meinungen über die Angemessenheit des Literaturverzeichnisses auseinander. Während drei Bewerter dieses als angemessen empfanden, war die Angabe der Literaturstellen für weitere drei Bewerter weder in Umfang noch Aktualität ausreichend. Ein weiterer Teilnehmer nahm eine neutrale Haltung gegenüber diesem Kriterium ein.

Für alle drei Fachartikel ist eine detaillierte Darstellung der Abweichungen im Ankreuzverhalten aller Teilnehmer des Pretests im Anhang einsehbar (siehe Anhang 9.5).