

Zusammenhang zwischen Leistungsmenge und Qualität des Behandlungsergebnisses bei Implantation von Kniegelenk- Totalendoprothesen (Knie-TEPs)

Addendum zu den Projekten V21-01 und V24-10
(Rapid Report)

A decorative horizontal bar spanning the width of the page, composed of 18 small squares in various shades of blue and grey. A dark blue rectangular box is overlaid on the bar, containing the text 'ADDENDUM (RAPID REPORT)'.

ADDENDUM (RAPID REPORT)

Projekt: V25-06

Version: 1.0

Stand: 15.10.2025

IQWiG-Berichte – Nr. 2113

DOI: 10.60584/V25-06

Impressum

Herausgeber

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema

Zusammenhang zwischen Leistungsmenge und Qualität des Behandlungsergebnisses bei Implantation von Kniegelenk-Totalendoprothesen (Knie-TEPs) – Addendum zu den Projekten V21-01 und V24-10

Auftraggeber

Gemeinsamer Bundesausschuss

Datum des Auftrags

10.09.2025

Interne Projektnummer

V25-06

DOI-URL

<https://doi.org/10.60584/V25-06>

Anschrift des Herausgebers

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Siegburger Str. 237
50679 Köln

Tel.: +49 221 35685-0

Fax: +49 221 35685-1

E-Mail: berichte@iqwig.de

Internet: www.iqwig.de

ISSN: 1864-2500

Zitiervorschlag

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Zusammenhang zwischen Leistungsmenge und Qualität des Behandlungsergebnisses bei Implantation von Kniegelenk-Totalendoprothesen (Knie-TEPs); Addendum zu den Projekten V21-01 und V24-10 (Rapid Report) [online]. 2025 [Zugriff: TT.MM.JJJJ]. URL: <https://doi.org/10.60584/V25-06>.

Schlagwörter

Mindestmenge, Arthroplastik, Kniegelenkersatz, Systematische Übersicht

Keywords

Minimum Volume, Arthroplasty, Replacement Knee, Systematic Review

An dem Addendum beteiligte Mitarbeiterinnen und Mitarbeiter des IQWiG

- Eva Höfer
- Carmen Bartel
- Jona Lilienthal
- Claudia-Martina Messow
- Claudia Mischke
- Max Oberste-Frielinghaus
- Corinna Schaefer

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	v
Abkürzungsverzeichnis.....	vi
1 Hintergrund.....	1
2 Methoden	2
3 Bewertung.....	4
3.1 Bewertung der Aussagekraft der Ergebnisse (Rapid Report V21-01; Updaterecherche V24-10)).....	4
3.2 Bewertung der internen Validität gemäß dem aktuellen methodischen Standard	7
4 Aussage zum Zusammenhang der Leistungsmenge und der Behandlungsqualität (Fragestellung 1 des Rapid Reports V21-01) unter Berücksichtigung der internen Validität der Studien.....	12
4.1 Übersicht über die bewertungsrelevanten Zielgrößen	12
4.2 Ergebnisse zu relevanten Zielgrößen	16
4.2.1 Mortalität	16
4.2.2 Morbidität	17
4.2.3 Gesundheitsbezogene Lebensqualität	21
4.2.4 Weitere Zielgrößen.....	23
4.3 Tabellarische Darstellung der Ergebnisse zum Zusammenhang zwischen Leistungsmenge und Qualität des Behandlungsergebnisses (Fragestellung 1)	26
5 Auswirkungen von einer konkret in die Versorgung eingeführten Mindestfallzahlen auf die Qualität des Behandlungsergebnisses (Fragestellung 2 des Rapid Reports V21-01) unter Berücksichtigung der internen Validität der Studien.....	28
5.1 Zielgrößen der Morbidität	28
5.2 Tabellarische Darstellung der Ergebnisse zu Auswirkungen von konkret in die Versorgung eingeführten Mindestfallzahlen auf die Qualität des Behandlungsergebnisses (Fragestellung 2)	29
6 Unklarheiten im Rapid Report V21-01 und der Update-Recherche V24-10 und Korrekturen.....	30
7 Fazit	32
8 Literatur	34

Tabellenverzeichnis

	Seite
Tabelle 1: Studienpool des Rapid Reports V21-01 und des Update V24-10 für die Fragestellung 1 und 2	4
Tabelle 2: Aussagekraft der Ergebnisse (Rapid Report V21-01 und Update-Recherche V24-10)	5
Tabelle 3: Interne Validität der Studien	8
Tabelle 4: Interne Validität der Studien (ausführliche Darstellung)	9
Tabelle 5: Matrix der relevanten Zielgrößen für die Fragestellung 1	13
Tabelle 6: Zusammenfassung der Ergebnisse zum Zusammenhang zwischen Leistungsmenge und Qualität des Behandlungsergebnisses (Fragestellung 1)	26
Tabelle 7: Zusammenfassung der Ergebnisse zu Auswirkungen von konkret in die Versorgung eingeführten Mindestfallzahlen auf die Qualität des Behandlungsergebnisses (Fragestellung 2)	29

Abkürzungsverzeichnis

Abkürzung	Bedeutung
EQ-5D-3L	European-Quality-of-Life-5-Dimensions-3-Level
G-BA	Gemeinsamer Bundesausschuss
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
KH	Krankenhaus
LM	Leistungsmenge
OKS	Oxford-Knee-Score
WOMAC	Western-Ontario-and-McMaster-Universities-Osteoarthritis-Index

1 Hintergrund

Der Gemeinsame Bundesausschuss (G-BA) hat das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) am 10.09.2025 mit einer methodischen Bewertung der Studienqualität zum Rapid Report V21-01 (Zusammenhang zwischen Leistungsmenge und Qualität des Behandlungsergebnisses bei Implantation von Kniegelenk-Totalendoprothesen [Knie-TEPs]) [1] und der Update-Recherche V24-10 [2] beauftragt.

Im Rapid Report V21-01 und der Update-Recherche V24-10 wurden die Studien als Ergebnis der methodischen Bewertung in die 2 Kategorien hohe oder niedrige Aussagekraft der Ergebnisse eingeteilt. Im Jahr 2024 modifizierte das IQWiG das Bewertungsinstrument mit dem Ziel einer besseren Differenzierung der Studien hinsichtlich ihrer methodischen Qualität. Das Bewertungsinstrument entsprechend dem aktuellen methodischen Standard klassifiziert die methodische Studienqualität nun in die 3 Kategorien niedrige, mäßige und hohe interne Validität.

Im Zuge dieser Weiterentwicklung wurde außerdem die Bezeichnung angepasst: „Bewertung der Aussagekraft“ wurde ersetzt durch „Bewertung der internen Validität“.

Der Auftrag umfasst die Bewertung der internen Validität der im Rapid Report V21-01 und der Update-Recherche V24-10 eingeschlossenen Studien entsprechend den gegenwärtigen methodischen Standards des IQWiG. Es wurden die folgenden Fragestellungen bearbeitet:

- 1) Wie ist die interne Validität der in den Rapid Report V21-01 und der Update-Recherche V24-10 eingeschlossenen Studien mit dem Bewertungsinstrument entsprechend dem aktuellen methodischen Standard zu bewerten?
- 2) Hat die Bewertung der internen Validität der Studien Auswirkungen auf die Ergebnisse des Rapid Reports V21-01 und der Update-Recherche V24-10?

Die im Rapid Report V21-01 und der Update-Recherche V24-10 eingeschlossenen Studien wurden in einem abgestuften Verfahren unter Verwendung des Bewertungsinstruments entsprechend dem aktuellen methodischen Standard des IQWiG bewertet und hinsichtlich der internen Validität eingeschätzt.

2 Methoden

Die interne Validität wurde in einem abgestuften Verfahren auf Basis folgender 3 Qualitätskriterien, die speziell für Studien zur Bewertung von Mengen-Ergebnis-Beziehungen entwickelt wurden, eingeschätzt [3-6]:

- Datenqualität: Die Vollständigkeit und die Eignung der Daten für die Beantwortung der Studienfrage sowie die Nachvollziehbarkeit des Patientenflusses (u. a. Angaben zu Ein- und Ausschlusskriterien sowie Drop-outs) werden bewertet.
- Statistische Modellierung: Die Qualität der statistischen Auswertung zur Untersuchung des Zusammenhangs zwischen Leistungsmenge und Behandlungsergebnis werden geprüft, wobei insbesondere die Adjustierung für relevante Risikofaktoren und der Umgang mit fehlenden Daten relevant sind.
- Publikationsqualität: Die Vollständigkeit und Nachvollziehbarkeit der Berichterstattung (z. B. Definition der Endpunkte, Angaben zu Patientencharakteristika sowie ausreichende Angaben zu statistischen Maßen) werden bewertet.

Die Gesamtbewertung ergibt sich in der Regel aus der jeweils niedrigsten Bewertung der Datenqualität und der statistischen Modellierung. Bei Mängeln der Publikationsqualität kann eine Abwertung der Gesamtbewertung erfolgen.

Für das vorliegende Addendum lag für 22 Studien (23 Publikationen) eine Bewertung der Aussagekraft der Ergebnisse vor. Dieser Bewertung lagen 12 Kriterien zugrunde (vergleiche Tabelle 2). Bei der Modifizierung des Bewertungsinstruments wurden diese Kriterien geschärft, drei übergeordneten Qualitätskriterien zugeordnet und für die Ableitung einer Gesamtbewertung der internen Validität unterschiedlich gewichtet. Zum Abgleich der Bewertung der Aussagekraft der Ergebnisse mit der Bewertung der internen Validität entsprechend dem aktuellen methodischen Standard des IQWiG wurden die Studien in einem abgestuften Verfahren eingeschätzt:

- 1) Im 1. Schritt wurde die Datenqualität bewertet. Wurde die Datenqualität als niedrig eingeschätzt, ergab sich gemäß dem Vorgehen zur Ableitung der Gesamtbewertung insgesamt für die Studie eine niedrige interne Validität.
- 2) Wurde die Datenqualität hingegen als mäßig oder hoch eingeschätzt, wurde im 2. Schritt die statistische Modellierung bewertet. Ergab sich für die statistische Modellierung eine niedrige Qualität, hat die Studie gemäß dem Vorgehen zur Ableitung der Gesamtbewertung insgesamt eine niedrige interne Validität.
- 3) Nur wenn die Studie sowohl in der Datenqualität als auch der statistischen Modellierung eine mäßige oder hohe Qualität aufzeigte, wurde die Publikationsqualität bewertet und

die interne Validität auf Basis aller 3 Bewertungskriterien (Datenqualität, statistische Modellierung, Publikationsqualität) bestimmt.

Abschließend wurde dargestellt, ob die Bewertung der internen Validität der Studien entsprechend dem aktuellen methodischen Standard Auswirkungen auf die Ergebnisse des Rapid Reports V21-01 und der Update-Recherche V24-10 hat.

3 Bewertung

3.1 Bewertung der Aussagekraft der Ergebnisse (Rapid Report V21-01; Updaterecherche V24-10))

Im Rapid Report V21-01 und der Update-Recherche V24-10 wurden insgesamt 22 Studien (23 Publikationen) eingeschlossen (siehe Tabelle 1). Die Aussagekraft der Ergebnisse dieser Studien wurde jeweils mit niedrig bewertet (siehe Tabelle 2).

Tabelle 1: Studienpool des Rapid Reports V21-01 und des Update V24-10 für die Fragestellung 1 und 2

Studie	Vollpublikation (in Fachzeitschriften)	Fragestellung	Projekt
Adhia 2020	ja [7]	1	V21-01
Arroyo 2019	ja [8]	1	V21-01
Blum 2013	ja [9]	1	V21-01
Feinglass 2004	ja [10]	1	V21-01
Garriga 2019	ja [11]	1	V21-01
Ghomrawi 2024	Ja [12]	1	V24-10
Hentschker 2018	ja [13]	1	V21-01
Jeschke 2017	ja [14]	1	V21-01
Katz 2004	ja [15,16]	1	V21-01
Maman 2019	ja [17]	1	V24-10
Marashi-Pour 2021	ja [18]	1	V21-01
Marsh 2024	ja [19]	1	V24-10
Menendez 2016	ja [20]	1	V21-01
Meyer 2011	ja [21]	1	V21-01
Muilwijk 2007	ja [22]	1	V21-01
Nimptsch 2017	ja [23]	1	V21-01
Ohmann 2010	ja [24]	1, 2	V21-01
Paterson 2010	ja [25]	1	V21-01
Singh 2011	ja [26]	1	V21-01
Solomon 2006	ja [27]	1	V21-01
Styron 2011	ja [28]	1	V21-01
Varagunam 2015	ja [29]	1	V21-01

Tabelle 2: Aussagekraft der Ergebnisse (Rapid Report V21-01 und Update-Recherche V24-10) (mehrseitige Tabelle)

Studie	Gute Qualität der individuellen Daten ^a	Adäquater Patientenfluss	Analyse der Menge	Plausibles Verfahren zur Bestimmung der Mengengrenzen	Geeignete Modellklasse	Adäquates Verfahren zur Berücksichtigung von Clustereffekten	Adäquate Risikoadjustierung ^a	Adäquater Umgang mit fehlenden Daten	Angaben zur Überprüfung der Modellgüte	Validierung des Modells	Angabe zur Punktschätzung inklusive Präzisionsangabe	Adäquate Berichterstattung relevanter Aspekte	Sonstige Aspekte	Aussagekraft der Ergebnisse
Adhia 2020	nein	unklar	kategorienell	unklar	ja	ja	nein ^b	unklar	nein	unklar	teilweise	ja	–	niedrig
Arroyo 2019	unklar	unklar	kategorienell	unklar	ja	ja	nein ^{b, c}	ja	nein	ja	ja	ja	–	niedrig
Blum 2013	nein	ja	kategorienell	unklar	ja	ja	nein ^b	ja	nein	unklar	ja	unklar	–	niedrig
Feinglass 2004	nein	unklar	kategorienell	ja	ja	ja	nein ^c	unklar	nein	unklar	ja	unklar	–	niedrig
Garriga 2019	unklar	ja	kategorienell	unklar	ja	ja	ja	ja	nein	unklar	ja	ja	–	niedrig
Ghomrawi 2024	nein	ja	kategorienell	nein	ja	ja	nein ^b	ja	nein	unklar	teilweise	nein ^d	–	niedrig
Hentschker 2018	ja	unklar	kategorienell / kontinuierlich	ja	ja	ja	nein ^b	ja	nein	unklar	ja	ja	–	niedrig
Jeschke 2017	ja	unklar	kategorienell	ja	ja	ja	nein ^{b, c}	unklar	nein	unklar	ja	ja	ja ^e	niedrig
Katz 2004 (Publikation Katz 2007)	unklar	unklar	kategorienell	unklar	ja	ja	ja	unklar	nein	unklar	teilweise	ja	–	niedrig
	nein	unklar	kategorienell	unklar	ja	ja	nein ^{b, c}	unklar	nein	unklar	teilweise	ja	–	niedrig
Maman 2019	unklar	unklar	kategorienell	unklar	ja	ja	nein ^b	ja	nein	unklar	ja	unklar ^f	–	niedrig
Marashi-Pour 2021	unklar	ja	kontinuierlich	ja	ja	ja	nein ^{b, c}	ja	ja	unklar	ja	unklar	–	niedrig
Marsh 2024	ja	ja	kategorienell	unklar	ja	ja	nein ^b	ja	nein	unklar	teilweise	ja	–	niedrig
Menendez 2016	unklar	unklar	kategorienell	ja	ja	ja	nein ^b	unklar	nein	unklar	teilweise	ja	–	niedrig
Meyer 2011	unklar	unklar	kategorienell	ja	ja	ja	nein ^{b, c}	unklar	nein	unklar	teilweise	unklar	–	niedrig
Muilwijk 2007	unklar	unklar	kategorienell	ja	ja	ja	nein ^b	ja	nein	unklar	teilweise	ja	ja ^g	niedrig

Tabelle 2: Aussagekraft der Ergebnisse (Rapid Report V21-01 und Update-Recherche V24-10) (mehrsseitige Tabelle)

Studie	Gute Qualität der individuellen Daten ^a	Adäquater Patientenfluss	Analyse der Menge	Plausibles Verfahren zur Bestimmung der Mengengrenzen	Geeignete Modellklasse	Adäquates Verfahren zur Berücksichtigung von Clustereffekten	Adäquate Risikoadjustierung ^a	Adäquater Umgang mit fehlenden Daten	Angaben zur Überprüfung der Modellgüte	Validierung des Modells	Angabe zur Punktschätzung inklusive Präzisionsangabe	Adäquate Berichterstattung relevanter Aspekte	Sonstige Aspekte	Aussagekraft der Ergebnisse
Nimptsch 2017	ja	unklar	kategorien / kontinuierlich	ja	ja	ja	nein ^{b, c}	unklar	ja	unklar	ja	ja	–	niedrig
Ohmann 2010	ja	unklar	kategorien	unklar	ja	ja	nein ^{b, c}	unklar	nein	unklar	teilweise	ja	–	niedrig
Paterson 2010	unklar	ja	kategorien	ja	ja	ja	ja	unklar	nein	unklar	teilweise	ja	–	niedrig
Singh 2011	unklar	unklar	kategorien	unklar	ja	ja	nein ^b	unklar	nein	unklar	ja	unklar	–	niedrig
Solomon 2006	unklar	unklar	kategorien	unklar	ja	ja	nein ^b	unklar	ja	unklar	teilweise	ja	–	niedrig
Styron 2011	unklar	unklar	kategorien	ja	ja	ja	ja	ja	nein	unklar	teilweise	ja	ja ^h	niedrig
Varagunam 2015	unklar	unklar	kontinuierlich	ja	ja	ja	ja	ja	nein	ja	ja	ja	–	niedrig
a. Ein „ja“ oder „nein“ wurde ausschließlich dann vergeben, wenn studienspezifisch eindeutige Angaben vorlagen. b. keine Adjustierung für Risikofaktoren auf Ebene der Ärztinnen und der Ärzte c. keine Adjustierung für Risikofaktoren auf Ebene der KHs d. Eine Zielgröße (Readmission within 90 days) wird im Abstract genannt, aber es werden keine Ergebnisse dazu berichtet. e. Die adäquate Berücksichtigung von Clustereffekten wurde in der Publikation zur Studie nicht angegeben, war aber nachträglich vom WiDO zugesichert worden und im eingereichten statistischen Analyseplan vermerkt. f. Wie die Zielgröße KH-Aufenthaltsdauer dichotomisiert wurde, wird nicht beschrieben. g. Diskrepanzen zwischen Text und Tabellen der Veröffentlichung h. Die angegebenen Patientencharakteristika sind nicht auf die untersuchte Stichprobe bezogen. AOK: Allgemeine Ortskrankenkasse; KH: Krankenhaus; WiDO: Wissenschaftliches Institut der AOK														

3.2 Bewertung der internen Validität gemäß dem aktuellen methodischen Standard

Die Bewertung der Datenqualität mit dem Bewertungsinstrument entsprechend dem aktuellen methodischen Standard ergibt für 7 Studien (Blum 2013, Jeschke 2017, Maman 2019, Marsh 2024, Menendez 2016, Ohmann 2010, Paterson2010) eine mäßige interne Validität. Alle anderen Studien zeigen im Gesamtergebnis eine niedrige interne Validität. Für die Studie Marashi-Pour 2021 wurde die Datenqualität mit hoch, die statistische Modellierung jedoch mit niedrig bewertet. Daher erfolgte keine weitere Bewertung der Publikationsqualität, da dies die Gesamtbewertung nicht geändert hätte. Bei allen weiteren Studien wurde die Datenqualität mit niedrig bewertet und es erfolgte keine weiterführende Bewertung der statistischen Modellierung sowie der Publikationsqualität. Dies hätte die Gesamtbewertung der internen Validität der Studien nicht geändert.

Tabelle 3: Interne Validität der Studien

Studie	Datenqualität	Statistische Modellierung ^a	Publikationsqualität ^b	Sonstige Aspekte ^c	Interne Validität der Studie ^d
Adhia 2020	niedrig	–	–	–	niedrig
Arroyo 2019	niedrig	–	–	–	niedrig
Blum 2013	mäßig	mäßig	mäßig	nein	mäßig
Feinglass 2004	niedrig	–	–	–	niedrig
Garriga 2019	niedrig	–	–	–	niedrig
Ghomrawi 2024	niedrig	–	–	–	niedrig
Hentschker 2018	niedrig	–	–	–	niedrig
Jeschke 2017	mäßig	mäßig	mäßig	nein	mäßig
Katz 2004	niedrig	–	–	–	niedrig
Katz 2007	niedrig	–	–	–	niedrig
Maman 2019	mäßig	hoch	mäßig	nein	mäßig
Marashi-Pour 2021	hoch	niedrig	–	–	niedrig
Marsh 2024	hoch	mäßig	mäßig	nein	mäßig
Menendez 2016	mäßig	mäßig	mäßig	nein	mäßig
Meyer 2011	niedrig	–	–	–	niedrig
Muilwijk 2007	niedrig	–	–	–	niedrig
Nimptsch 2017	niedrig	–	–	–	niedrig
Ohmann 2010	mäßig	mäßig	mäßig	nein	mäßig
Paterson 2010	mäßig	mäßig	mäßig	nein	mäßig
Singh 2011	niedrig	–	–	–	niedrig
Solomon 2006	niedrig	–	–	–	niedrig
Styron 2011	niedrig	–	–	–	niedrig
Varagunam 2015	niedrig	–	–	–	niedrig
a. Aufgrund des abgestuften Verfahrens wurde dieses Qualitätskriterium nur dann bewertet, wenn die Datenqualität mit mäßig oder hoch bewertet wurde. b. Aufgrund des abgestuften Verfahrens wurde dieses Qualitätskriterium nur dann bewertet, wenn sowohl die Datenqualität als auch die statistische Modellierung mit mäßig oder hoch bewertet wurden. c. Aufgrund des abgestuften Verfahrens wurde dieser Aspekt nur dann bewertet, wenn sowohl die Datenqualität als auch die statistische Modellierung und die Publikationsqualität mit mäßig oder hoch bewertet wurden. d. Die Gesamtbewertung der internen Validität beruht maßgeblich auf der Datenqualität und der statistischen Modellierung. Im vorliegenden Addendum wurde die interne Validität abgestuft bewertet: Wenn die Datenqualität als niedrig eingeschätzt wird, hat die Studie eine niedrige interne Validität.					

In der folgenden Tabelle 4 werden die bewerteten Kriterien der internen Validität ausführlich dargestellt.

Tabelle 4: Interne Validität der Studien (ausführliche Darstellung) (mehrseitige Tabelle)

Studie ^a	Datenqualität			Statistische Modellierung					Publikationsqualität					Fehlen sonstiger Aspekte	Gesamtbewertung ^d
	Qualität der individuellen Daten	Adäquater Patientenfluss	Gesamtbewertung Datenqualität ^b	Adjustierung für Risikofaktoren	Adäquater Umgang mit fehlenden Daten	Plausible Bestimmung der Mengengrenzen	Angaben zur Modellgüte / Modellvalidierung	Gesamtbewertung statistische Modellierung ^b	Eindeutige Definition der Zielgrößen	Ausreichende Angaben zur Bewertung der Volume-Outcome-Fragestellung	Ausreichende Angaben zu Patientencharakteristika	Adäquate Berichterstattung relevanter Aspekte	Gesamtbewertung Publikationsqualität ^c		
Adhia 2020	★	★	★	×	×	×	×	×	×	×	×	×	×	×	★
Arroyo 2019	★★	★	★	×	×	×	×	×	×	×	×	×	×	×	★
Blum 2013	★★	★★★★	★★	★★	+	+/-	-	★★	+/-	+/-	+/-	+	★★	+	★★
Feinglass 2004	★	★★	★	×	×	×	×	×	×	×	×	×	×	×	★
Garriga 2019	★★	★	★	×	×	×	×	×	×	×	×	×	×	×	★
Ghomrawi 2024	★	★★★★	★	×	×	×	×	×	×	×	×	×	×	×	★
Hentschker 2018	★★★★	★	★	×	×	×	×	×	×	×	×	×	×	×	★
Jeschke 2017	★★★★	★★	★★	★★	+	+	-	★★	+	+	+/-	+	★★	+	★★
Katz 2004	★	★★	★	×	×	×	×	×	×	×	×	×	×	×	★
Katz 2007	★	★	★	×	×	×	×	×	×	×	×	×	×	×	★
Maman 2019	★★	★★	★★	★★★★	+	?	-	★★★★	+/-	+	+/-	+	★★	+	★★
Marashi-Pour 2021	★★★★	★★★★	★★★★	★	?	+	+	★	×	×	×	×	×	×	★
Marsh 2024	★★★★	★★★★	★★★★	★★★★	+	?	-	★★	+	+/-	+/-	?	★★	+	★★
Menendez 2016	★★	★★	★★	★★★★	+	+	-	★★★★	+	+	+/-	+	★★	+	★★
Meyer 2011	★	★	★	×	×	×	×	×	×	×	×	×	×	×	★

Tabelle 4: Interne Validität der Studien (ausführliche Darstellung) (mehrseitige Tabelle)

Studie ^a	Datenqualität			Statistische Modellierung					Publikationsqualität					Fehlen sonstiger Aspekte	Gesamtbewertung ^d
	Qualität der individuellen Daten	Adäquater Patientenfluss	Gesamtbewertung Datenqualität ^b	Adjustierung für Risikofaktoren	Adäquater Umgang mit fehlenden Daten	Plausible Bestimmung der Mengengrenzen	Angaben zur Modellgüte / Modellvalidierung	Gesamtbewertung statistische Modellierung ^b	Eindeutige Definition der Zielgrößen	Ausreichende Angaben zur Bewertung der Volume-Outcome-Fragestellung	Ausreichende Angaben zu Patientencharakteristika	Adäquate Berichterstattung relevanter Aspekte	Gesamtbewertung Publikationsqualität ^c		
Muilwijk 2007	★	★	★	×	×	×	×	×	×	×	×	×	×	×	★
Nimptsch 2017	★★★	★	★	×	×	×	×	×	×	×	×	×	×	×	★
Ohmann 2010	★★★	★★	★★	★★	+	+	–	★★	+	+/-	+/-	+	★★	+	★★
Paterson 2010	★★	★★★★	★★	★★	+	+	–	★★	+	+/-	+/-	+	★★	+	★★
Singh 2011	★	★★	★	×	×	×	×	×	×	×	×	×	×	×	★
Solomon 2006	★	★	★	×	×	×	×	×	×	×	×	×	×	×	★
Styron 2011	★	★★	★	×	×	×	×	×	×	×	×	×	×	×	★
Varagunam 2015	★	★	★	×	×	×	×	×	×	×	×	×	×	×	★
★: niedrige Qualität ★★: mäßige Qualität ★★★: hohe Qualität +: ja –: nein ?: unklar +/-: teilweise ×: Aufgrund des abgestuften Verfahrens erfolgte keine Bewertung.															

Tabelle 4: Interne Validität der Studien (ausführliche Darstellung) (mehrseitige Tabelle)

Studie ^a	Datenqualität		Statistische Modellierung					Publikationsqualität				Fehlen sonstiger Aspekte	Gesamtbewertung ^d
	Qualität der individuellen Daten	Adäquater Patientenfluss	Gesamtbewertung Datenqualität ^b	Adjustierung für Risikofaktoren	Adäquater Umgang mit fehlenden Daten	Plausible Bestimmung der Mengengrenzen	Angaben zur Modellgüte / Modellvalidierung	Gesamtbewertung statistische Modellierung ^b	Eindeutige Definition der Zielgrößen	Ausreichende Angaben zur Bewertung der Volume-Outcome-Fragestellung	Ausreichende Angaben zu Patientencharakteristika		
a. Die Gesamtbewertung der Datenqualität orientiert sich an der niedrigsten Bewertung in dieser Domäne.													
b. Für die Gesamtbewertung der statistischen Modellierung ist vor allem die Adjustierung maßgeblich.													
c. Für die Gesamtbewertung der Publikationsqualität sind vor allem eine klare und eindeutige Definition der relevanten Zielgrößen, das Vorliegen ausreichender Angaben zur Einschätzung des Zusammenhangs sowie der Patientencharakteristika maßgeblich.													
d. Die Gesamtbewertung beruht maßgeblich auf der Datenqualität und der statistischen Modellierung.													

4 Aussage zum Zusammenhang der Leistungsmenge und der Behandlungsqualität (Fragestellung 1 des Rapid Reports V21-01) unter Berücksichtigung der internen Validität der Studien

Für die Untersuchung des Zusammenhangs zwischen Leistungsmenge und Qualität des Behandlungsergebnisses bei Implantation einer Knie-Totalendoprothese wurden in den Rapid Report V21-01 und die Update-Recherche V24-10 insgesamt 22 Beobachtungsstudien eingeschlossen (siehe Tabelle 1).

Die Bewertung der internen Validität der Studien entsprechend dem aktuellen methodischen Standard führt zu einer differenzierteren Bewertung der Qualität der Studien. 7 der 22 Studien weisen eine andere Qualitätsstufe auf und erreichten bei der Studienbewertung entsprechend dem aktuellen methodischen Standard eine mäßige interne Validität anstelle einer niedrigen. Dies hat insofern Auswirkungen auf die Ergebnisse des Rapid Reports V21-01 und der Update-Recherche V24-10, dass die Qualität der jeweiligen Studien, auf denen die Ergebnisse zu den einzelnen Zielgrößen beruhen, differenzierter dargestellt werden können. Die Ergebnisse zur Ableitung eines Zusammenhangs ändern sich aufgrund dieser Studienbewertungen jedoch nicht.

Im Folgenden werden die Ergebnisse des Rapid Reports V21-01 und der Update-Recherche V24-10 zum Zusammenhang zwischen der Leistungsmenge (LM) pro Krankenhaus (KH), pro Ärztin oder Arzt oder pro Kombination Krankenhaus und Ärztin oder Arzt und der Qualität des Behandlungsergebnisses bei Implantation einer Knie-TEP zu allen Zielgrößen erneut zusammengefasst. Hierzu wurde zunächst analog zum Rapid Reports V21-01 und der Update-Recherche V24-10 geprüft, ob pro Zielgröße ein Zusammenhang abgeleitet wurde. Die Ableitung eines Zusammenhangs erfolgte in der Regel auf Basis von mindestens 1 Studie mit signifikantem Ergebnis zugunsten höherer Leistungsmengen. Wenn die Ergebnisse der Studie(n) in keine eindeutige Richtung für den Zusammenhang zwischen der LM und der Qualität des Behandlungsergebnisses zeigten, wurde kein Zusammenhang abgeleitet. Zusätzlich wurde dargestellt auf welcher internen Validität dieser Zusammenhang basiert.

Es werden nur die Ergebnisse zu den Ebenen der Leistungsmenge (KH, Ärztin oder Arzt, Kombination der LM von KH und Ärztin oder Arzt) berichtet, zu denen Studien mit verwertbaren Ergebnissen vorlagen.

4.1 Übersicht über die bewertungsrelevanten Zielgrößen

Tabelle 5 gibt eine Übersicht über die verfügbaren Daten zu den relevanten Zielgrößen aus den eingeschlossenen Studien. Im Vergleich zu dem Rapid Report V21-01 und der Update-Recherche V24-10 wurde die Darstellung entsprechend der Ergebnisdarstellung in dem vorliegenden Addendum angepasst und enthält alle in den eingeschlossenen Berichten untersuchten relevanten Zielgrößen, für die Daten vorlagen.

Tabelle 5: Matrix der relevanten Zielgrößen für die Fragestellung 1 (mehrsseitige Tabelle)

Studie	Zielgrößen																						
	Mortalität				Morbidität											Andere							
	Versterben im KH	30-Tage-Mortalität	90-Tage-Mortalität	1-Jahres-Mortalität	Postoperativ Komplikationen	Komplikationen während des KH-Aufenthalts	Komplikationen innerhalb eines definierten Zeitraums (30–90 Tage)	Komplikationen nach 6 Monaten	Myokardinfarkt	Wundinfektionen	Postoperative Blutungen oder Hämatom in der Wunde	Pneumonie	Lungenembolie / venöse Thromboembolie	Bluttransfusion	Gesundheitsbezogene Lebensqualität	Zufriedenheit mit der OP	KH-Aufenthaltsdauer	Entlassung aus dem KH nach Hause oder in eine Einrichtung	Prothesenlockerung (Instabilität)	Revisionseingriffe innerhalb von 1 oder 2 Jahren	Revisionseingriffe innerhalb von 5 Jahren	Erneute Operation innerhalb von 5 Jahren, z. B. Synovektomie	Wiederaufnahme in ein KH
Adhia 2020	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	●	●	–	–	–	–	–
Arroyo 2019	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	● ^a
Blum 2013	–	–	–	○	–	○	–	○	○	–	–	–	○	–	–	–	–	–	○	○	●	–	–
Feinglass 2004	○	–	–	–	–	●	–	–	○ ^b	○ ^b	–	○ ^b	○ ^b	–	–	–	–	–	○	–	–	–	–
Garriga 2019	–	–	–	–	–	–	–	●	–	–	–	–	–	–	● ^c	–	●	–	–	–	–	–	–
Ghomrawi 2024	–	–	–	–	●	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Hentschker 2018	●	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Jeschke 2017	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	●	–	–	–
Katz 2004 (Publikation Katz 2007)	–	–	●	–	–	–	● ^d	● ^d	● ^d	–	–	● ^d	● ^d	–	–	● ^e	● ^f	–	–	–	–	–	–
Maman 2019	●	–	–	–	●	–	–	–	–	–	–	–	–	–	–	–	●	–	–	–	–	–	–
Marashi-Pour 2021	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	● ^g

Tabelle 5: Matrix der relevanten Zielgrößen für die Fragestellung 1 (mehrsseitige Tabelle)

Studie	Zielgrößen																						
	Mortalität				Morbidität								Andere										
	Versterben im KH	30-Tage-Mortalität	90-Tage-Mortalität	1-Jahres-Mortalität	Postoperativ Komplikationen	Komplikationen während des KH-Aufenthalts	Komplikationen innerhalb eines definierten Zeitraums (30–90 Tage)	Komplikationen nach 6 Monaten	Myokardinfarkt	Wundinfektionen	Postoperative Blutungen oder Hämatom in der Wunde	Pneumonie	Lungenembolie / venöse Thromboembolie	Bluttransfusion	Gesundheitsbezogene Lebensqualität	Zufriedenheit mit der OP	KH-Aufenthaltsdauer	Entlassung aus dem KH nach Hause oder in eine Einrichtung	Prothesenlockerung (Instabilität)	Revisionseingriffe innerhalb von 1 oder 2 Jahren	Revisionseingriffe innerhalb von 5 Jahren	Erneute Operation innerhalb von 5 Jahren, z. B. Synovektomie	Wiederaufnahme in ein KH
Marsh 2024	–	–	–	○	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	○	●	○	–
Menendez 2016	–	–	–	–	–	–	–	–	–	–	–	–	–	●	–	–	–	–	–	–	–	–	–
Meyer 2011	–	–	–	–	–	–	–	–	–	●	–	–	–	–	–	–	–	–	–	–	–	–	–
Muilwijk 2007	–	–	–	–	–	–	–	–	–	●	–	–	–	–	–	–	–	–	–	–	–	–	–
Nimptsch 2017	●	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Ohmann 2010	–	–	–	–	–	–	–	–	–	●	●	–	–	–	–	–	–	–	–	–	–	–	–
Paterson 2010	–	–	●	–	–	●	–	–	–	–	–	–	–	–	–	–	●	–	–	●	–	–	● ^h
Singh 2011	–	●	–	●	–	–	● ⁱ	–	● ⁱ	● ⁱ	–	–	● ⁱ	–	–	–	–	–	–	–	–	–	–
Solomon 2006	–	–	○	–	–	–	● ^d	–	○ ^d	○ ^d	–	○ ^d	○ ^d	–	–	–	–	–	–	–	–	–	–
Styron 2011	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	●	–	–	–	–	–	–
Varagunam 2015	–	–	–	–	–	–	–	●	–	–	–	–	–	–	● ^{c, j}	–	–	–	–	–	–	–	–

Tabelle 5: Matrix der relevanten Zielgrößen für die Fragestellung 1 (mehrsseitige Tabelle)

Studie	Zielgrößen																						
	Mortalität				Morbidität										Andere								
	Versterben im KH	30-Tage-Mortalität	90-Tage-Mortalität	1-Jahres-Mortalität	Postoperativ Komplikationen	Komplikationen während des KH-Aufenthalts	Komplikationen innerhalb eines definierten Zeitraums (30–90 Tage)	Komplikationen nach 6 Monaten	Myokardinfarkt	Wundinfektionen	Postoperative Blutungen oder Hämatom in der Wunde	Pneumonie	Lungenembolie / venöse Thromboembolie	Bluttransfusion	Gesundheitsbezogene Lebensqualität	Zufriedenheit mit der OP	KH-Aufenthaltsdauer	Entlassung aus dem KH nach Hause oder in eine Einrichtung	Prothesenlockerung (Instabilität)	Revisionseingriffe innerhalb von 1 oder 2 Jahren	Revisionseingriffe innerhalb von 5 Jahren	Erneute Operation innerhalb von 5 Jahren, z. B. Synovektomie	Wiederaufnahme in ein KH
●: Daten wurden berichtet und waren verwertbar. o: Daten wurden berichtet, waren aber nicht für die Nutzenbewertung verwertbar. –: Es wurden keine Daten berichtet (keine weiteren Angaben) oder die Zielgröße wurde nicht erhoben. a. Wiederaufnahme innerhalb von 30 oder 90 Tagen b. während des KH-Aufenthaltes c. Oxford-Knee-Score d. Komplikationen innerhalb von 90 Tagen e. WOMAC-Score (inkl. Functional-Status, Schmerz-Score) f. unklar, ob Erhebungsbogen validiert ist g. Wiederaufnahme innerhalb von 60 Tagen h. Wiederaufnahme in ein KH zur Amputation, Arthrodeese oder Exzision innerhalb 1 Jahres i. Komplikationen innerhalb von 30 Tagen j. EQ-5D-3L EQ-5D-3L: European-Quality-of-Life-5-Dimensions-3-Level; KH: Krankenhaus; OP: Operation; WOMAC: Western-Ontario-and-McMaster-Universities-Osteoarthritis-Index																							

4.2 Ergebnisse zu relevanten Zielgrößen

4.2.1 Mortalität

Versterben im KH

KH-Ebene

Basierend auf 1 Studie mit mäßiger interner Validität (Maman 2019) und 2 Studien mit niedriger interner Validität (Hentschker 2018, Nimptsch 2017) wurde ein Zusammenhang zwischen der LM pro KH und der Qualität des Behandlungsergebnisses zugunsten der KHs mit höherer LM für die Zielgröße Versterben im KH abgeleitet.

Alle 3 Studien, Hentschker 2018, Maman 2019 und Nimptsch 2017, zeigten signifikante Ergebnisse zugunsten einer höheren LM.

Versterben innerhalb von 30 Tagen

KH-Ebene

Basierend auf 1 Studie mit niedriger interner Validität (Singh 2011) war für die Zielgröße Versterben innerhalb von 30 Tagen kein Zusammenhang zwischen der LM pro KH und der Qualität des Behandlungsergebnisses ableitbar.

Die Studie Singh 2011 zeigte keine signifikanten Ergebnisse.

Versterben innerhalb von 90 Tagen

KH-Ebene

Für die Zielgröße Versterben innerhalb von 90 Tagen war auf Basis von 2 Studien (Paterson 2010 [mäßige interne Validität]; Katz 2004 [niedrige interne Validität]) kein Zusammenhang zwischen der LM pro KH und der Qualität des Behandlungsergebnisses ableitbar.

Die Studie Paterson 2010 zeigte keinen statistisch signifikanten Zusammenhang zwischen der LM und der Qualität des Behandlungsergebnisses. Bei der Studie Katz 2004 zeigte sich nur beim Vergleich der KHs mit sehr niedriger LM mit den KHs mit niedriger LM ein statistisch signifikanter Unterschied zugunsten von KHs mit höherer LM. Der p-Wert für den Trend war jedoch nicht statistisch signifikant.

Ebene Ärztin oder Arzt

Basierend auf 1 Studie mit mäßiger interner Validität (Paterson 2010) und 1 Studie mit niedriger interner Validität (Katz 2004) war für die Zielgröße Versterben innerhalb von 90 Tagen kein Zusammenhang zwischen der LM pro Ärztin oder Arzt und der Qualität des Behandlungsergebnisses ableitbar.

Die 2 Studien Paterson 2020 und Katz 2004 zeigten keine signifikanten Ergebnisse.

Versterben innerhalb 1 Jahres

KH-Ebene

Für die Zielgröße 1-Jahres-Mortalität wurde, basierend auf 1 Studie mit niedriger interner Validität (Singh 2011) ein Zusammenhang zwischen der LM pro KH und der Qualität des Behandlungsergebnisses bei einer Teilpopulation (Patientinnen und Patienten ≥ 65 Jahre) zugunsten der KHs mit höherer LM abgeleitet.

Die Studie Singh 2011 zeigte für den Vergleich der KHs mit 26 bis 100 Implantationen von Knie-TEPs pro Jahr im Vergleich zu den KHs mit über 200 Implantationen einen statistisch signifikanten Unterschied zugunsten der KHs mit höherer LM. Für den Vergleich über alle Kategorien konnte kein statistisch signifikanter Trend beobachtet werden. In der Subgruppe der Patientinnen und Patienten im Alter von 65 Jahren und älter zeigte der Vergleich der KHs mit 26 bis 100 Implantationen von Knie-TEPs pro Jahr zu den KHs mit über 200 Implantationen ebenfalls einen statistisch signifikanten Unterschied zugunsten der KHs mit höherer LM. Allerdings zeigte sich für diese Patientenpopulation der Trend für den Vergleich über alle Kategorien als statistisch signifikant.

4.2.2 Morbidität

Komplikationen postoperativ

KH-Ebene

Basierend auf 1 Studie mit mäßiger interner Validität (Maman 2019) wurde ein Zusammenhang zwischen der LM pro KH und der Qualität des Behandlungsergebnisses zugunsten der KHs mit höherer LM für diese Zielgröße abgeleitet.

Die Studie Maman 2019 zeigte signifikante Ergebnisse. Untersucht wurden Komplikationen mit Bezug zu dem intraoperativen Zeitraum oder dem Verfahren, Infektionen, Wundheilungsstörungen sowie pulmonalen, gastrointestinalen, kardiovaskulären und systemischen Problemen sowie Komplikationen des Harntrakts.

Komplikationen während des KH-Aufenthalts

KH-Ebene

Basierend auf 1 Studie mit mäßiger interner Validität (Paterson 2010) und 1 Studie mit niedriger interner Validität (Feinglass 2004) wurde für die Zielgröße Komplikationen im KH keine eindeutige Richtung für den Zusammenhang zwischen der LM pro KH und der Qualität des Behandlungsergebnisses festgestellt und somit kein Zusammenhang abgeleitet.

In der Studie mit mäßiger interner Validität (Paterson 2010) wurde für die Zielgröße Komplikationen im KH kein statistisch signifikanter Unterschied zwischen den KHs mit unterschiedlichen LM aufgezeigt. In 1 Studie mit niedriger interner Validität (Feinglass 2004) wurde für die Zielgröße perioperative Komplikationen lediglich bei dem Vergleich der KHs mit

sehr niedriger LM mit KHs mit mittleren LM ein statistisch signifikanter Unterschied zuungunsten der KHs mit höherer LM festgestellt. Alle anderen Vergleiche der LM-Kategorien ergaben keinen statistisch signifikanten Unterschied.

Ebene Ärztin oder Arzt

Basierend auf 1 Studie mit mäßiger interner Validität (Paterson 2010) wurde auf Arztebene keine eindeutige Richtung für den Zusammenhang zwischen LM und Behandlungsergebnis festgestellt und somit kein Zusammenhang abgeleitet.

In der Studie Paterson 2010 wurde lediglich für den Vergleich der Ärztinnen und Ärzte der niedrigen LM-Kategorie mit der nächsthöheren (mittleren) LM ein statistisch signifikanter Unterschied zuungunsten der höheren LM aufgezeigt. Die anderen Vergleiche der LM-Kategorie waren auf Arztebene nicht statistisch signifikant.

Komplikationen innerhalb eines definierten Zeitraums (30 bis 90 Tagen)

KH-Ebene

Auf Basis von 4 Studien mit niedriger interner Validität (Ghomrawi 2024, Katz 2004, Singh 2011, Solomon 2006) wurde für die Zielgröße Komplikationen innerhalb eines definierten Zeitraums (30 bis 90 Tagen) ein Zusammenhang zwischen der LM pro KH und der Qualität des Behandlungsergebnisses zugunsten der KHs mit höherer LM abgeleitet.

Die Studien Ghomrawi 2024, Solomon 2006 und Katz 2004 zeigten signifikante Ergebnisse zugunsten einer höheren LM, während die Studie Singh 2011 keine signifikanten Ergebnisse zeigte. Die Studie Singh 2011 untersuchte die Zielgröße Komplikationen innerhalb von 30 Tagen und die Studien Ghomrawi 2024, Katz 2004 und Solomon 2006 Komplikationen innerhalb von 90 Tagen.

Ebene Ärztin oder Arzt

Auf Basis 1 Studie mit niedriger interner Validität (Katz 2004) und signifikantem Ergebnis wurde für die Zielgröße Komplikationen innerhalb eines definierten Zeitraums (30 bis 90 Tagen) ein Zusammenhang zwischen der LM pro Ärztin oder Arzt und der Qualität des Behandlungsergebnisses zugunsten einer höheren LM abgeleitet.

Die Studie Katz 2004 zeigte für diese Zielgröße zum Signifikanzniveau von 1 % keinen statistisch signifikanten Unterschied, allerdings war der p-Wert für den Trend statistisch signifikant.

Komplikationen nach 6 Monaten

KH-Ebene und Ebene Ärztin oder Arzt

Basierend auf 2 Studien mit niedriger interner Validität (Garriga 2019, Varaguman 2015) wurde für die Zielgröße Komplikationen innerhalb von 6 Monaten ein Zusammenhang zwischen der LM sowohl pro KH als auch pro Ärztin oder Arzt und der Qualität des Behandlungsergebnisses zugunsten einer höheren LM abgeleitet.

Die Studie Garriga 2019 zeigte für die Zielgröße Komplikationen innerhalb von 6 Monaten auf KH-Ebene und auf Ebene Ärztin oder Arzt signifikante Ergebnisse zugunsten der höheren LM. Die Studie Varagunam 2015 zeigte weder für die KH- noch für die Arzt-Ebene signifikante Ergebnisse.

Myokardinfarkt

Die Studie Katz 2004 untersuchte die Zielgröße akuter Myokardinfarkt innerhalb von 90 Tagen und die Studie Singh 2011 untersuchte die Zielgröße akuter Myokardinfarkt innerhalb von 30 Tagen.

KH-Ebene

Basierend auf 2 Studien mit niedriger interner Validität (Katz 2004, Singh 2011) war für die Zielgröße Myokardinfarkt kein Zusammenhang zwischen der LM pro KH und der Qualität des Behandlungsergebnisses ableitbar.

Die Studien Katz 2004 und Singh 2011 zeigten keine statistisch signifikanten Ergebnisse.

Ebene Ärztin oder Arzt

Basierend auf 1 Studie mit niedriger interner Validität (Katz 2004) war für die Zielgröße Myokardinfarkt kein Zusammenhang zwischen der LM pro Ärztin oder Arzt und der Qualität des Behandlungsergebnisses ableitbar.

Die Studie Katz 2004 zeigte keine statistisch signifikanten Ergebnisse.

Wundinfektionen

KH-Ebene

Basierend auf 1 Studie mit mäßiger interner Validität (Ohmann 2010) und 4 Studien mit niedriger interner Validität (Katz 2004, Meyer 2011, Muilwijk 2007 und Singh 2011) wurde für die Zielgröße Wundinfektionen ein Zusammenhang zwischen der LM pro KH und der Qualität des Behandlungsergebnisses zugunsten KHS mit höherer LM abgeleitet.

Die Studien Ohmann 2010 (mäßige interne Validität) und Meyer 2011 (niedrige interne Validität) zeigten für die Zielgröße Wundinfektionen auf KH-Ebene signifikante Ergebnisse

zugunsten der höheren LM. Die Studien Katz 2004, Muilwijk 2007 und Singh 2011 zeigten keine statistisch signifikanten Ergebnisse.

Ebene Ärztin oder Arzt

Basierend auf 2 Studien mit niedriger interner Validität (Katz 2004, Muilwijk 2007) und signifikanten Ergebnissen wurde für die Zielgröße Wundinfektionen ein Zusammenhang zwischen der LM pro Ärztin oder Arzt und der Qualität des Behandlungsergebnisses zugunsten einer höheren LM abgeleitet.

Die Studien Katz 2004 und Muilwijk 2007 zeigten statistisch signifikante Unterschiede zugunsten einer höheren Leistungsmenge.

Postoperative Blutungen oder Hämatom in der Wunde

KH-Ebene

Basierend auf 1 Studie mit mäßiger interner Validität (Ohmann 2010) wurde für die Zielgröße postoperative Blutungen oder Hämatom in der Wunde keine eindeutige Richtung für den Zusammenhang zwischen der LM pro KH und der Qualität des Behandlungsergebnisses festgestellt und somit kein Zusammenhang abgeleitet.

In der Studie Ohmann 2010 mit mäßiger interner Validität wurde ein statistisch signifikanter Zusammenhang zwischen LM pro KH und dem Behandlungsergebnis für die Zielgröße postoperative Blutungen oder Hämatom in der Wunde für das Jahr 2006 zugunsten der KH mit höheren LM festgestellt. Für die Jahre 2004 und 2005 zeigt sich in derselben Studie kein einheitlicher Zusammenhang zwischen der LM pro KH und der Qualität des Behandlungsergebnisses.

Pneumonie

KH-Ebene und Ebene Ärztin oder Arzt

Basierend auf 1 Studie mit niedriger interner Validität (Katz 2004) wurde für die Zielgröße Pneumonie ein Zusammenhang zwischen der LM sowohl pro KH als auch pro Ärztin oder Arzt und der Qualität des Behandlungsergebnisses zugunsten einer höheren LM abgeleitet.

Die Studie Katz 2004 zeigte für die Zielgröße Pneumonie innerhalb von 90 Tagen statistisch signifikante Unterschiede sowohl zugunsten der KHs mit höherer LM als auch zugunsten der Ärztinnen und Ärzte mit einer höheren LM.

Lungenembolie / venöse Thromboembolie

Die Studie Katz 2004 (niedrige interne Validität) betrachtete die Zielgröße Lungenembolie innerhalb von 90 Tagen und die Studie Singh 2011 (niedrige interne Validität) die Zielgröße venöse Thromboembolie innerhalb von 30 Tagen.

KH-Ebene

Basierend auf 2 Studien mit niedriger interner Validität (Katz 2004, Singh 2011) war für die Zielgröße Lungenembolie / venöse Thromboembolie kein Zusammenhang zwischen der LM pro KH und der Qualität des Behandlungsergebnisses ableitbar.

Die Studien Katz 2004 und Singh 2011 zeigten keine statistisch signifikanten Ergebnisse.

Ebene Ärztin oder Arzt

Basierend auf 1 Studie mit niedriger interner Validität (Katz 2004) war für die Zielgröße Lungenembolie / venöse Thromboembolie kein Zusammenhang zwischen der LM pro Ärztin oder Arzt und der Qualität des Behandlungsergebnisses ableitbar.

Die Studie Katz 2004 zeigte keine statistisch signifikanten Ergebnisse.

Bluttransfusion

KH-Ebene

Basierend auf 1 Studie mit mäßiger interner Validität (Menendez 2016) wurde für die Zielgröße Bluttransfusion ein Zusammenhang zwischen der LM pro KH und der Qualität des Behandlungsergebnisses zugunsten einer höheren LM abgeleitet.

Die Studie Menendez 2016 zeigte signifikante Unterschiede für den Vergleich der KHs der niedrigsten LM-Kategorie mit den KHs der höchsten LM-Kategorie zugunsten einer höheren LM.

4.2.3 Gesundheitsbezogene Lebensqualität

In 2 Studien mit niedriger interner Validität (Garriga 2019, Varagunam 2015) wurde die gesundheitsbezogene Lebensqualität mithilfe des Oxford-Knee-Score (OKS)-Erhebungsbogens erfasst. In 1 Studie mit niedriger interner Validität (Varagunam 2015) wurden zusätzlich Daten zur gesundheitsbezogenen Lebensqualität anhand des generischen Erhebungsbogens EQ-5D-3L (European-Quality-of-Life-5-Dimensions-3-Level) erhoben. In 1 Studie mit niedriger interner Validität (Publikation Katz 2007 der Studie Katz 2004) wurde die patientenberichtete gesundheitsbezogene Lebensqualität in Form des Western-Ontario-and-McMasters-Universities-Osteoarthritis-Index(WOMAC)-Fragebogens erhoben.

KH-Ebene

Basierend auf 2 Studien mit niedriger interner Validität (Garriga 2019, Varagunam 2015) wurde für die Zielgröße gesundheitsbezogene Lebensqualität ein Zusammenhang zwischen der LM pro KH und der Qualität des Behandlungsergebnisses zugunsten der KHs mit höherer LM abgeleitet.

Die Studie Garriga 2019 zeigte für die Zielgröße gesundheitsbezogene Lebensqualität signifikante Ergebnisse zugunsten der höheren LM. Die Studie Varagunam 2015 zeigte keine statistisch signifikanten Ergebnisse.

Ebene Ärztin oder Arzt

Basierend auf 2 Studien mit niedriger interner Validität (Garriga 2019, Varagunam 2015) wurde für die Zielgröße gesundheitsbezogene Lebensqualität ein Zusammenhang zwischen der LM pro Ärztin oder Arzt und der Qualität des Behandlungsergebnisses zugunsten einer höheren LM abgeleitet.

Die Studie Garriga 2019 zeigte statistisch signifikante Ergebnisse zugunsten von Ärztinnen und Ärzten mit höherer LM. Die Studie Varagunam 2015 zeigte keine statistisch signifikanten Ergebnisse.

Ebene Kombination der LM von KH und Ärztin oder Arzt

Basierend auf 1 Studie mit niedriger interner Validität (Publikation Katz 2007 der Studie Katz 2004) wurde für die Zielgröße gesundheitsbezogene Lebensqualität ein Zusammenhang zwischen der LM des KH und Ärztin oder Arzt und der Qualität des Behandlungsergebnisses zugunsten einer höheren LM abgeleitet.

In der Studie Katz 2004 (Publikation Katz 2007) zeigten sich statistisch signifikante Unterschiede zugunsten höherer LM auf Ebene der Kombination von LM auf KH- und Arzzebene bei der Erhebung mittels WOMAC-Fragebogen sowohl für den Funktionsstatus als auch bei den Angaben zur Kniebeugung und zur Kniestreckung. Für die patientenberichtete Erhebung zum Schmerz-Score (WOMAC-Fragebogen) und die Zufriedenheit der Patientinnen und Patienten mit der OP, die ebenfalls in der Studie Katz 2004 (Publikation Katz 2007) erhoben wurden, zeigte sich für die Ebene Kombination der LM von KH und Ärztin oder Arzt kein statistisch signifikanter Unterschied.

Zufriedenheit mit der OP

Ebene Kombination der LM von KH und Ärztin oder Arzt

Basierend auf 1 Studie mit niedriger interner Validität (Publikation Katz 2007 der Studie Katz 2004) war für die Zielgröße Zufriedenheit mit der OP kein Zusammenhang zwischen der LM des KH und Ärztin oder Arzt und der Qualität des Behandlungsergebnisses ableitbar.

Die Studie Katz 2004 (Publikation Katz 2007) zeigte keine signifikanten Ergebnisse.

4.2.4 Weitere Zielgrößen

KH-Aufenthaltsdauer

KH-Ebene

Basierend auf 2 Studien mit mäßiger interner Validität (Maman 2019, Paterson 2010) und 3 Studien mit niedriger interner Validität (Adhia 2020, Garriga 2019, Styron 2011) wurde für die Zielgröße KH-Aufenthaltsdauer ein Zusammenhang zwischen der LM pro KH und der Qualität des Behandlungsergebnisses zugunsten einer höheren LM abgeleitet.

Die Studien Adhia 2020 (niedrige interne Validität), Maman 2019 (mäßige interne Validität) und Styron 2011 (niedrige interne Validität) zeigten für die Zielgröße KH-Aufenthaltsdauer auf KH-Ebene signifikante Ergebnisse zugunsten der höheren LM. Die Studie Garriga 2019 (niedrige interne Validität) zeigte für die verschiedenen Vergleiche zwischen den LM-Kategorien keine einheitliche Richtung für den Zusammenhang zwischen LM und Qualität des Behandlungsergebnisses. Die Studie Paterson 2010 (mäßige interne Validität) zeigte keine signifikanten Ergebnisse.

Ebene Ärztin oder Arzt

Basierend auf 1 Studie mit mäßiger interner Validität (Paterson 2010) und 2 Studien mit niedriger interner Validität (Garriga 2019, Styron 2011) wurde für die Zielgröße KH-Aufenthaltsdauer ein Zusammenhang zwischen der LM pro Ärztin oder Arzt und der Qualität des Behandlungsergebnisses zugunsten einer höheren LM abgeleitet.

Die 3 Studien Garriga 2019, Paterson und Styron 2011 zeigten signifikante Ergebnisse zugunsten einer höheren LM.

Entlassung aus dem KH nach Hause oder in eine Einrichtung

KH-Ebene

Basierend auf 1 Studie mit niedriger interner Validität (Adhia 2020) war für die Zielgröße Entlassung aus dem KH nach Hause oder in eine Einrichtung kein Zusammenhang zwischen der LM pro KH und der Qualität des Behandlungsergebnisses ableitbar.

Die Studie Adhia 2020 zeigte keine statistisch signifikanten Ergebnisse.

Revisionseingriffe

In der Studie Paterson 2010 (mäßige interne Validität) wurde die Zielgröße Revisionseingriffe innerhalb von 1 Jahr untersucht und in der Studie Jeschke 2017 (mäßige interne Validität) die Revisionseingriffe innerhalb von 2 Jahren und in der Studie Blum 2013 (mäßige interne Validität) und Marsh 2024 (mäßige interne Validität) die Revisionseingriffe innerhalb von 5 Jahren.

KH-Ebene

Auf Basis von 2 Studie mit mäßiger interner Validität (Jeschke 2017, Paterson 2010) wurde für die Zielgröße Revisionseingriffe innerhalb von 1 und 2 Jahren ein Zusammenhang zwischen der LM pro KH und der Qualität des Behandlungsergebnisses zugunsten einer höheren LM abgeleitet. Für Revisionseingriffe innerhalb von 5 Jahren war dieser Zusammenhang auf Basis einer Studie mit mäßiger interner Validität (Blum 2013) nicht ableitbar.

Die Studie Paterson 2010 zeigte für Revisionseingriffe innerhalb von 1 Jahr beim Vergleich der KHs mit niedriger LM mit KHs mit hoher sowie mit sehr hoher LM jeweils ein statistisch signifikanten Unterschied zugunsten von KH mit höherer LM. Die Studie Jeschke 2017 zeigte für Revisionseingriffe innerhalb von 2 Jahr beim Vergleich der KHs mit sehr hohe LM mit KHs mit sehr niedriger, niedriger und mittlerer LM jeweils einen statistisch signifikanten Unterschied zugunsten von KH mit höherer LM. Beim Vergleich mit KHs mit hoher LM zeigte sich dagegen kein statistisch signifikanter Unterschied. In der Studie Blum 2013 konnte für die Zielgröße Revisionseingriffe nach 5 Jahren kein statistisch signifikanter Unterschied festgestellt werden.

Ebene Ärztin oder Arzt

Basierend auf 1 Studie mit mäßiger interner Validität (Paterson 2010) war für die Zielgröße Revisionseingriffe innerhalb von 1 und 2 Jahren kein Zusammenhang zwischen der LM pro Ärztin oder Arzt und der Qualität des Behandlungsergebnisses ableitbar. Für Revisionseingriffe innerhalb von 5 Jahren war dieser Zusammenhang zugunsten einer höheren LM auf Basis einer Studie mit mäßiger interner Validität (Marsh 2024) ableitbar.

Die Studie Paterson zeigte für Revisionseingriffe innerhalb von 1 Jahr keine signifikanten Ergebnisse. In der Studie Marsh 2024 zeigte sich für die Zielgröße Revisionseingriffe innerhalb von 5 Jahren ein statistisch signifikanter Unterschied zugunsten der höheren LM auf Arztebene.

Wiederaufnahme in ein KH

Die Studien Arroyo 2019 (niedrige interne Validität) und Marashi-Pour 2021 (niedrige interne Validität) berichteten über die Zielgröße Wiederaufnahme in ein KH zu 3 verschiedenen Zeitpunkten (30, 60 und 90 Tage nach Implantation einer Knie-TEP). Paterson 2010 (mäßige interne Validität) dagegen berichtete über die Zielgröße Wiederaufnahme in ein KH zur Amputation, Arthrodese oder Exzision innerhalb 1 Jahres.

KH-Ebene

Basierend auf 1 Studie mit mäßiger interner Validität (Paterson 2010) und 2 Studien mit niedriger interner Validität (Arroyo 2019, Marashi-Pour 2021) wurde für die Zielgröße

Wiederaufnahme in ein KH ein Zusammenhang zwischen der LM pro KH und der Qualität des Behandlungsergebnisses zugunsten einer höheren LM abgeleitet.

Die Studien Paterson 2010 (mäßige interne Validität) und Arroyo 2019 (niedrige interne Validität) zeigten für die Zielgröße Wiederaufnahme in ein KH auf KH-Ebene signifikante Ergebnisse zugunsten der höheren LM. Die Autoren der Studie Marashi-Pour 2021 stellten die Ergebnisse des Zusammenhangs der LM pro KH und der Qualität des Behandlungsergebnisses lediglich grafisch dar. Die Quantifizierung der Ergebnisse war nicht möglich. Die Autorinnen und Autoren beschrieben, dass kein bedeutsamer Zusammenhang zwischen den Wiederaufnahmeraten und den KHs mit unterschiedlichen LMs vorlag.

Ebene Ärztin oder Arzt

Basierend auf 1 Studie mit mäßiger interner Validität (Paterson 2010) wurde für die Zielgröße KH-Aufenthaltsdauer ein Zusammenhang zwischen der LM pro Ärztin oder Arzt und der Qualität des Behandlungsergebnisses zugunsten einer höheren LM abgeleitet.

Die Studie Paterson 2010 zeigte einen statistisch signifikanten Unterschied für den Vergleich zwischen der Referenzkategorie (niedrige LM) und der mittleren LM zugunsten einer höheren LM.

4.3 Tabellarische Darstellung der Ergebnisse zum Zusammenhang zwischen Leistungsmenge und Qualität des Behandlungsergebnisses (Fragestellung 1)

Tabelle 6: Zusammenfassung der Ergebnisse zum Zusammenhang zwischen Leistungsmenge und Qualität des Behandlungsergebnisses (Fragestellung 1) (mehrsseitige Tabelle)

Zielgrößenkategorie Zielgröße	KH	Ärztin oder Arzt	Kombination	Zusammenfassung
Mortalität				
Versterben im KH	↑**	–	–	Zusammenhang zugunsten einer höheren LM auf KH-Ebene
30-Tage-Mortalität	↔*	–	–	kein Zusammenhang ableitbar
90-Tage-Mortalität	↔**	↔**	–	kein Zusammenhang ableitbar
1- Jahres-Mortalität	↑ ^a *	–	–	Zusammenhang zugunsten einer höheren LM auf KH-Ebene für Patientinnen und Patienten, die mindestens 65 Jahre alt sind
Morbidität				
postoperative Komplikationen	↑**	–	–	Zusammenhang zugunsten einer höheren LM auf KH-Ebene
Komplikationen während des KH-Aufenthalts	↑↓**	↑↓**	–	kein Zusammenhang ableitbar
Komplikationen innerhalb eines definierten Zeitraums (30–90 Tagen)	↑*	↑*	–	Zusammenhang zugunsten einer höheren LM auf KH-Ebene und auf Ebene Ärztin oder Arzt
Komplikationen nach 6 Monaten	↑*	↑*	–	Zusammenhang zugunsten einer höheren LM auf KH-Ebene und auf Ebene Ärztin oder Arzt
Myokardinfarkt	↔*	↔*	–	kein Zusammenhang ableitbar
Wundinfektionen	↑**	↑*	–	Zusammenhang zugunsten einer höheren LM auf KH-Ebene und auf Ebene Ärztin oder Arzt
postoperative Blutungen oder Hämatom in der Wunde	↑↓**	–	–	kein Zusammenhang ableitbar
Pneumonie	↑*	↑*	–	Zusammenhang zugunsten einer höheren LM auf KH-Ebene und auf Ebene Ärztin oder Arzt
Lungenembolie / venöse Thromboembolie	↔*	↔*	–	kein Zusammenhang ableitbar
Bluttransfusion	↑**	–	–	Zusammenhang zugunsten einer höheren LM auf KH-Ebene

Tabelle 6: Zusammenfassung der Ergebnisse zum Zusammenhang zwischen Leistungsmenge und Qualität des Behandlungsergebnisses (Fragestellung 1) (mehrseitige Tabelle)

Zielgrößenkategorie Zielgröße	KH	Ärztin oder Arzt	Kombination	Zusammenfassung
Gesundheitsbezogene Lebensqualität				
gesundheitsbezogene Lebensqualität	↑*	↑*	↑*	Zusammenhang zugunsten einer höheren LM auf KH-Ebene, auf Ebene Ärztin oder Arzt und der Kombination der Ebenen
Zufriedenheit mit der OP	–	–	↔*	kein Zusammenhang ableitbar
Weitere				
KH-Aufenthaltsdauer	↑**	↑**	–	Zusammenhang zugunsten einer höheren LM auf KH-Ebene und auf Ebene Ärztin oder Arzt
Entlassung aus dem KH nach Hause oder in eine Einrichtung	↔*	–	–	kein Zusammenhang ableitbar
Revisionseingriffe innerhalb von 1 Jahr und von 2 Jahren	↑**	↔**	–	Zusammenhang zugunsten einer höheren LM auf KH-Ebene kein Zusammenhang auf Ebene Ärztin oder Arzt ableitbar
Revisionseingriffe innerhalb von 5 Jahren	↔**	↑**	–	kein Zusammenhang ableitbar
Wiederaufnahme in ein KH	↑**	↑**	–	Zusammenhang zugunsten einer höheren LM auf KH-Ebene und auf Ebene Ärztin oder Arzt
↔: kein Zusammenhang ableitbar ↑: Zusammenhang zugunsten höherer LM ↓: Zusammenhang zuungunsten höherer LM ↑↓: keine einheitliche Richtung des Unterschieds zwischen den LM Validität der zugrunde liegenden Studien: ★: Ergebnis basierend auf 1 oder mehreren Studien mit niedriger interner Validität ★★: Ergebnis basierend auf 1 oder mehreren Studien mit niedriger oder mäßiger interner Validität, darunter mindestens 1 mit mäßiger interner Validität ★★★: Ergebnis basierend auf 1 oder mehreren Studien, darunter mindestens 1 mit hoher interner Validität –: keine (verwertbaren) Ergebnisse a. nur für die Teilpopulation Patientinnen und Patienten ≥ 65 Jahre KH: Krankenhaus; LM: Leistungsmenge				

5 Auswirkungen von einer konkret in die Versorgung eingeführten Mindestfallzahlen auf die Qualität des Behandlungsergebnisses (Fragestellung 2 des Rapid Reports V21-01) unter Berücksichtigung der internen Validität der Studien

Für die Untersuchung der Auswirkungen von einer konkret in die Versorgung eingeführten Mindestfallzahl auf die Qualität des Behandlungsergebnisses für die Implantation von Knie-TEPs (Fragestellung 2 des Rapid Reports V21-01) wurde in den Rapid Report V21-01 eine Studie eingeschlossen. In der Update-Recherche V24-10 wurde keine weitere Studie identifiziert.

Die Studie Ohman 2010 untersuchte die Auswirkungen einer Mindestfallzahl auf die Qualität des Behandlungsergebnisses im deutschen Versorgungskontext und wies eine niedrige Aussagekraft der Ergebnisse auf. Bei der Bewertung entsprechend dem aktuellen methodischen Standard zeigte die Studie eine mäßige interne Validität (vergleiche Tabelle 3 und Tabelle 4).

Im Folgenden werden die Ergebnisse des Rapid Reports V21-01 zu den Auswirkungen von konkret in die Versorgung eingeführten Mindestfallzahlen auf die Qualität des Behandlungsergebnisses erneut zusammengefasst. Bei der Darstellung der Ergebnisse wird die interne Validität der zugrunde liegenden Studie ausgewiesen. Es werden nur die Ergebnisse zu den Zielgrößen berichtet, für die verwertbare Ergebnisse vorlagen.

5.1 Zielgrößen der Morbidität

Wundinfektionen

Beim Vergleich der Jahre vor und nach Einführung einer Mindestfallzahl zeigte sich in der Studie Ohmann 2010 (mäßige interne Validität) für die Zielgröße Wundinfektionen ein statistisch signifikanter Unterschied. Dabei war im Jahr nach Einführung einer Mindestfallzahl die Wahrscheinlichkeit für das Auftreten einer Wundinfektion geringer als im Jahr vor der Einführung.

Postoperative Blutungen oder Hämatom in der Wunde

Beim Vergleich des Jahres 2006, nach Einführung einer Mindestfallzahl, mit dem Jahr 2005 zeigte sich in der Studie Ohmann 2010 (mäßige interne Validität) für die Zielgröße postoperative Blutungen oder Hämatom in der Wunde ein statistisch signifikanter Unterschied. Dabei war nach Einführung einer Mindestfallzahl die Wahrscheinlichkeit für das Auftreten von postoperativen Blutungen oder Hämatomen in der Wunde geringer als im Jahr vor der Einführung.

5.2 Tabellarische Darstellung der Ergebnisse zu Auswirkungen von konkret in die Versorgung eingeführten Mindestfallzahlen auf die Qualität des Behandlungsergebnisses (Fragestellung 2)

Tabelle 7: Zusammenfassung der Ergebnisse zu Auswirkungen von konkret in die Versorgung eingeführten Mindestfallzahlen auf die Qualität des Behandlungsergebnisses (Fragestellung 2)

Zielgrößenkategorie Zielgröße	KH	Zusammenfassung
Morbidität		
Wundinfektionen	↑**	Zusammenhang zugunsten des Zeitraums nach Einführung der Mindestfallzahl auf KH-Ebene
postoperative Blutung oder Hämatom in der Wunde	↑**	Zusammenhang zugunsten des Zeitraums nach Einführung der Mindestfallzahl auf KH-Ebene
↑: Basierend auf 1 Studie, die statistisch signifikante Unterschiede hinsichtlich der Zielgröße zugunsten des Zeitraums nach Einführung einer Mindestfallzahl zeigte. Validität der zugrunde liegenden Studien: ★★: Ergebnis basierend auf 1 Studie mit mäßiger interner Validität KH: Krankenhaus		

6 Unklarheiten im Rapid Report V21-01 und der Update-Recherche V24-10 und Korrekturen

Zum Rapid Report V21-01 und der Update-Recherche V24-10 sind zwischenzeitlich Fragen aufgekomen, die auf Unklarheiten in diesen Berichten hinweisen.

Kurzfassung im Rapid Report V21-01

In der Kurzfassung im Rapid Report V21-01 heißt es in Bezug auf die Operationalisierungen „Komplikationen (innerhalb von 30 Tagen, von 90 Tagen und innerhalb von 6 Monaten)“: „Auch konnte keine eindeutige Richtung für den Zusammenhang zwischen Leistungsmenge der Ärztinnen und Ärzte und Qualität des Behandlungsergebnisses abgeleitet werden.“

Diese Zusammenfassung über alle Erhebungszeiträume der Zielgröße Komplikationen gibt ein unklares Bild wieder. Tatsächlich wurde, wie im Ergebnisteil des Rapid Report V21-01 dargestellt, auf Ebene der Ärztin oder des Arztes nur für die Operationalisierung Komplikationen während des KH-Aufenthalts keine eindeutige Richtung für den Zusammenhang zwischen der LM der Ärztinnen und Ärzte und der Qualität des Behandlungsergebnisses festgestellt und kein Zusammenhang abgeleitet (vergleiche Tabelle 23 im Rapid Report V21-01). Demgegenüber wurde für die Erhebungszeiträume Komplikationen innerhalb eines definierten Zeitraums (30 bis 90 Tage) und Komplikationen nach 6 Monaten ein Zusammenhang zwischen der Leistungsmenge der Ärztinnen und Ärzte und der Qualität des Behandlungsergebnisses zugunsten einer höheren LM abgeleitet.

Des Weiteren heißt es in der Kurzfassung im Rapid Report V21-01 in Bezug auf die Operationalisierung „Komplikationen während des KH-Aufenthaltes“: „Für die Zielgröße Komplikationen während des Krankenhausaufenthalts konnte jedoch auf der Arztebene ein Zusammenhang zwischen Leistungsmenge und Qualität des Behandlungsergebnisses zugunsten höherer Leistungsmenge abgeleitet werden.“

Dies ist nicht korrekt, denn für diese Operationalisierung wurde für der Ebene Ärztin oder Arzt, wie im Ergebnisteil des Rapid Report V21-01 dargestellt, keine eindeutige Richtung für den Zusammenhang zwischen LM pro Ärztin oder Arzt und der Qualität des Behandlungsergebnisses festgestellt und somit kein Zusammenhang abgeleitet.

In Tabelle 6 des vorliegenden Addendums sind die Ergebnisse zu den verschiedenen Operationalisierungen und Erhebungszeiträumen der Zielgröße Komplikationen tabellarisch dargestellt und im Fazit erneut zusammengefasst.

Gegenüberstellung der Ergebnisse der Rapid Reports V21-01 und V24-10 (Tabelle 14 der Update-Recherche V24-10)

In Bezug auf die Operationalisierung „Komplikationen während des KH-Aufenthaltes“ und die Arztebene wird in der Tabelle 14 „Gegenüberstellung der Ergebnisse der Rapid Reports V21-01 und V24-10“ der Update-Recherche V24-10 mit der Fußnote d auf die zugrundeliegenden Studien für das Ergebnis „keine eindeutige Richtung des Unterschieds“ hingewiesen. In der Fußnote d heißt es, das Ergebnis beruhe auf 3 Studien.

Diese Fußnote ist nicht korrekt. Die Ableitung des Zusammenhangs zwischen der LM der Ärztinnen und Ärzte und der Qualität des Behandlungsergebnisses für den Erhebungszeitraum KH-Aufenthalt beruht nur auf einer Studie. Ausschließlich die Studie Paterson 2010 untersuchte auf der Arztebene diese Operationalisierung der Zielgröße Komplikationen.

Zusammenfassende tabellarische Darstellung der Ergebnisse

Zudem kam eine Nachfrage dazu, wann bei der tabellarischen Zusammenfassung der Ergebnisse in der Spalte „Zusammenfassung“ kein Zusammenhang ableitbar steht. Dies ist der Fall, wenn in der jeweiligen Ergebnisspalte:

- 1 horizontaler Pfeil ($\leftrightarrow$) oder
- 2 nach oben und unten zeigende Pfeile ($\updownarrow$) stehen.

Im ersten Fall wird kein Zusammenhang abgeleitet, da keine Studien mit signifikantem Ergebnis vorliegen. Im 2. Fall wird trotz statistisch signifikanter Ergebnisse kein Zusammenhang abgeleitet, da die Ergebnisse nicht eindeutig in eine Richtung für den Zusammenhang zwischen Leistungsmenge und Qualität des Behandlungsergebnisses weisen.

7 Fazit

Für die Untersuchung des Zusammenhangs zwischen der Leistungsmenge und der Qualität des Behandlungsergebnisses bei Implantationen von Knie-Totalendoprothesen wurden die in den Rapid Report V21-01 und die Update-Recherche V24-10 eingeschlossenen Studien entsprechend den gegenwärtigen methodischen Standards des IQWiG neu bewertet. Insgesamt wurden 22 Studien (23 Publikationen) neu bewertet. Die Studien waren hinsichtlich der Aussagekraft der Ergebnisse in den Vorgängerprojekten alle mit niedrig bewertet.

Fragestellung 1

Alle 22 Studien (23 Publikationen) untersuchten den Zusammenhang zwischen der Leistungsmenge und der Qualität des Behandlungsergebnisses bei Implantation einer Knie-Totalendoprothese. Die interne Validität von 7 Studien wurde entsprechend des gegenwärtigen methodischen Standards des IQWiG mit mäßig bewertet. 15 Studien (16 Publikationen) zeigten eine niedrige interne Validität.

Auf Krankenhausebene wurde für 2 von 4 untersuchten Zielgrößen der Mortalität ein Zusammenhang zwischen der Leistungsmenge und der Qualität des Behandlungsergebnisses zugunsten höherer Leistungsmenge abgeleitet. Die Ableitung des Zusammenhangs beruht für die Zielgröße Versterben im Krankenhaus u. a. auf den Ergebnissen einer Studie mit mäßiger interner Validität.

Des Weiteren wurde auf Krankenhausebene für 6 von 10 untersuchten Zielgrößen der Morbidität ein Zusammenhang zwischen der Leistungsmenge des Krankenhauses und der Qualität des Behandlungsergebnisses zugunsten höherer Leistungsmenge abgeleitet. Für die Zielgrößen postoperative Komplikationen, Wundinfektionen und Bluttransfusionen beruht die Ableitung des Zusammenhangs u. a. auf den Ergebnissen von Studien mit mäßiger interner Validität. Der Ableitung des Zusammenhangs für die Zielgrößen Komplikationen innerhalb eines definierten Zeitraums, Komplikationen nach 6 Monaten und Pneumonie lagen ausschließlich Studien mit niedriger interner Validität zugrunde. Für die Zielgrößen Komplikationen während des Krankenhausaufenthalts und postoperative Blutungen oder Hämatom in der Wunde zeigte sich in den Studien keine eindeutige Richtung für den Unterschied zwischen den Leistungsmengen. Daher war für diese Zielgrößen kein Zusammenhang zwischen der Leistungsmenge des Krankenhauses und der Qualität des Behandlungsergebnisses ableitbar. Diesen Ergebnissen lagen jeweils mindestens 1 Studie mit mäßiger interner Validität zugrunde.

Auf Ebene der Ärztin und des Arztes wurde für 4 von 7 untersuchten Zielgrößen der Morbidität ein Zusammenhang zwischen Leistungsmenge der Ärztin oder des Arztes und Qualität des Behandlungsergebnisses zugunsten höherer Leistungsmenge abgeleitet. Diese Ergebnisse beruhen ausschließlich auf Studien mit niedriger interner Validität. Allerdings zeigte sich für

die Zielgröße Komplikationen während des Krankenhausaufenthalts auf Ebene der Ärztin und des Arztes keine eindeutige Richtung für den Unterschied und es war kein Zusammenhang ableitbar. Dieses Ergebnis basiert auf 1 Studie mit mäßiger interner Validität.

Für die Zielgröße gesundheitsbezogene Lebensqualität wurde auf Basis der Ergebnisse von Studien mit niedriger interner Validität ein Zusammenhang zwischen der Leistungsmenge und Qualität des Behandlungsergebnisses auf Krankenhaus- und auf Arztebene sowie auf Ebene der Kombination der Leistungsmenge von Krankenhaus und Ärztin oder Arzt zugunsten der höheren Leistungsmenge abgeleitet.

Auf Krankenhausebene wurde unter den weiteren Zielgrößen für die Zielgrößen Krankenhausaufenthaltsdauer, Revisionseingriffe innerhalb von 1 Jahr und 2 Jahren sowie Wiederaufnahme in ein Krankenhaus ein Zusammenhang zwischen Leistungsmenge des Krankenhauses und der Qualität des Behandlungsergebnisses zugunsten höherer Leistungsmenge abgeleitet. Diese Ergebnisse basieren jeweils u. a. auf Studien mit mäßiger interner Validität.

Auf Ebene der Ärztin und des Arztes wurde unter den weiteren Zielgrößen für die Zielgrößen Krankenhausaufenthaltsdauer, Revisionseingriffe innerhalb von 5 Jahren sowie Wiederaufnahme in ein Krankenhaus ein Zusammenhang zwischen Leistungsmenge der Ärztin oder des Arztes und Qualität des Behandlungsergebnisses zugunsten höherer Leistungsmenge abgeleitet. Diese Ergebnisse basieren jeweils u. a. auf Studien mit mäßiger interner Validität.

Fragestellung 2

Zu den Auswirkungen der Einführung einer Mindestfallzahl in die deutsche Versorgung wurde in den Rapid Report V21-01 eine Studie mit niedriger Aussagekraft eingeschlossen. Die Bewertung mit dem gegenwärtigen methodischen Standard des IQWiG ergab für diese Studie eine mäßige interne Validität. Auf Basis dieser Studie konnte für die Zielgrößen Wundinfektionen und postoperative Blutungen oder Hämatom in der Wunde jeweils zugunsten des Zeitraums nach Einführung einer Mindestfallzahl ein Effekt nachgewiesen werden.

8 Literatur

1. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Zusammenhang zwischen Leistungsmenge und Qualität des Behandlungsergebnisses bei Implantation von Kniegelenk-Totalendoprothesen (Knie-TEPs); Rapid Report [online]. 2023 [Zugriff: 11.07.2023]. URL: https://www.iqwig.de/download/v21-01_zusammenhang-leistungsmenge-und-qualitaet-bei-implantation-von-knie-teps_rapid-report_v2-0.pdf.
2. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Zusammenhang zwischen Leistungsmenge und Qualität des Behandlungsergebnisses bei Implantation von Kniegelenk-Totalendoprothesen; Update zum Projekt V21-01; Rapid Report [online]. 2025 [Zugriff: 08.07.2025]. URL: <https://doi.org/10.60584/V24-10>.
3. Bender R, Grouven U. Möglichkeiten und Grenzen statistischer Regressionsmodelle zur Berechnung von Schwellenwerten für Mindestmengen. Z Arztl Fortbild Qualitätssich 2006; 100(2): 93-98.
4. Benchimol EI, Smeeth L, Guttman A et al. Das RECORD-Statement zum Berichten von Beobachtungsstudien, die routinemäßig gesammelte Gesundheitsdaten verwenden. Z Evid Fortbild Qual Gesundhwes 2016; 115-116: 33-48.
<https://doi.org/10.1016/j.zefq.2016.07.010>.
5. Higgins JPT, Morgan RL, Rooney AA et al. Risk Of Bias In Non-randomized Studies - of Exposure (ROBINS-E) [online]. 2024 [Zugriff: 16.10.2024]. URL: <https://drive.google.com/file/d/1bcOVMna3CJJXpsIm5vobFJMP0-13CW8Q/view?usp=sharing>.
6. Wetzel H. Mindestmengen zur Qualitätssicherung: Konzeptionelle und methodische Überlegungen zur Festlegung und Evaluation von Fallzahlgrenzwerten für die klinische Versorgung. Z Arztl Fortbild Qualitätssich 2006; 100(2): 99-106.
7. Adhia AH, Feinglass JM, Suleiman LI. What Are the Risk Factors for 48 or More-Hour Stay and Nonhome Discharge After Total Knee Arthroplasty? Results From 151 Illinois Hospitals, 2016-2018. J Arthroplasty 2020; 35(6): 1466-1473.e1.
<https://doi.org/10.1016/j.arth.2019.11.043>.
8. Arroyo NS, White RS, Gaber-Baylis LK et al. Racial/Ethnic and Socioeconomic Disparities in Total Knee Arthroplasty 30- and 90-Day Readmissions: A Multi-Payer and Multistate Analysis, 2007-2014. Population Health Management 2019; 22(2): 175-185.
<https://doi.org/10.1089/pop.2018.0025>.
9. Blum MA, Singh JA, Lee GC et al. Patient race and surgical outcomes after total knee arthroplasty: an analysis of a large regional database. Arthritis Care Res (Hoboken) 2013; 65(3): 414-420. <https://doi.org/10.1002/acr.21834>.

10. Feinglass J, Amir H, Taylor P et al. How safe is primary knee replacement surgery? Perioperative complication rates in Northern Illinois, 1993-1999. *Arthritis Rheum* 2004; 51(1): 110-116. <https://doi.org/10.1002/art.20072>.
11. Garriga C, Leal J, Sanchez-Santos MT et al. Geographical Variation in Outcomes of Primary Hip and Knee Replacement. *JAMA Network Open* 2019; 2(10): e1914325. <https://doi.org/10.1001/jamanetworkopen.2019.14325>.
12. Ghomrawi HM, Huang LW, Hiredesai AH et al. Using the Stratum-Specific Likelihood Ratio Method to Derive Outcome-Based Hospital Volume Categories for Total Knee Replacement. *Med Care* 2024; 62(4): 250-255. <https://doi.org/10.1097/MLR.0000000000001985>.
13. Hentschker C, Mennicken R, Reifferscheid A et al. Volume-outcome relationship and minimum volume regulations in the German hospital sector - evidence from nationwide administrative hospital data for the years 2005-2007. *Health Econ Rev* 2018; 8(1): 25. <https://doi.org/10.1186/s13561-018-0204-8>.
14. Jeschke E, Citak M, Gunster C et al. Are TKAs Performed in High-volume Hospitals Less Likely to Undergo Revision Than TKAs Performed in Low-volume Hospitals? *Clin Orthop Relat Res* 2017; 475(11): 2669-2674. <https://doi.org/10.1007/s11999-017-5463-x>.
15. Katz JN, Barrett J, Mahomed NN et al. Association between hospital and surgeon procedure volume and the outcomes of total knee replacement. *Journal of Bone & Joint Surgery - American Volume* 2004; 86(9): 1909-1916. <https://doi.org/10.2106/00004623-200409000-00008>.
16. Katz JN, Mahomed NN, Baron JA et al. Association of hospital and surgeon procedure volume with patient-centered outcomes of total knee replacement in a population-based cohort of patients age 65 years and older. *Arthritis Rheum* 2007; 56(2): 568-574. <https://doi.org/10.1002/art.22333>.
17. Maman SR, Andrae MH, Gaber-Baylis LK et al. Medicaid insurance status predicts postoperative mortality after total knee arthroplasty in state inpatient databases. *J Comp Eff Res* 2019; 8(14): 1213–1228. <https://doi.org/10.2217/ce-2019-0027>.
18. Marashi-Pour S, Watson DE, Harris IA. Association between number of replacements and readmissions following total knee and total hip replacements in New South Wales' public hospitals. *ANZ J Surg* 2021; 91(6): 1277-1283. <https://doi.org/10.1111/ans.16882>.
19. Marsh JD, McClure JA, Lanting BA et al. Predictors of undergoing revision within 5 years of total knee arthroplasty: a population-based study. *J Arthroplasty* 2024; 39(3): 689-694.e3. <https://doi.org/10.1016/j.arth.2023.09.006>.
20. Menendez ME, Lu N, Huybrechts KF et al. Variation in Use of Blood Transfusion in Primary Total Hip and Knee Arthroplasties. *J Arthroplasty* 2016; 31(12): 2757-2763.e2. <https://doi.org/10.1016/j.arth.2016.05.022>.

21. Meyer E, Weitzel-Kage D, Sohr D et al. Impact of department volume on surgical site infections following arthroscopy, knee replacement or hip replacement. *BMJ Quality & Safety* 2011; 20(12): 1069-1074. <https://doi.org/10.1136/bmjqs.2011.051250>.
22. Muilwijk J, van den Hof S, Wille JC. Associations between surgical site infection risk and hospital operation volume and surgeon operation volume among hospitals in the Dutch nosocomial infection surveillance network. *Infect Control Hosp Epidemiol* 2007; 28(5): 557-563. <https://doi.org/10.1086/513613>.
23. Nimptsch U, Mansky T. Hospital volume and mortality for 25 types of inpatient treatment in German hospitals: Observational study using complete national data from 2009 to 2014. *BMJ Open* 2017; 7(9): e016184. <https://doi.org/10.1136/bmjopen-2017-016184>.
24. Ohmann C, Verde PE, Blum K et al. Two short-term outcomes after instituting a national regulation regarding minimum procedural volumes for total knee replacement. *Journal of Bone & Joint Surgery - American Volume* 2010; 92(3): 629-638. <https://doi.org/10.2106/jbjs.H.01436>.
25. Paterson JM, Williams JI, Kreder HJ et al. Provider volumes and early outcomes of primary total joint replacement in Ontario. *Can J Surg* 2010; 53(3): 175-183.
26. Singh JA, Kwok CK, Boudreau RM et al. Hospital volume and surgical outcomes after elective hip/knee arthroplasty: a risk-adjusted analysis of a large regional database. *Arthritis Rheum* 2011; 63(8): 2531-2539. <https://doi.org/10.1002/art.30390>.
27. Solomon DH, Chibnik LB, Losina E et al. Development of a preliminary index that predicts adverse events after total knee replacement. *Arthritis Rheum* 2006; 54(5): 1536-1542. <https://doi.org/10.1002/art.21772>.
28. Styron JF, Koroukian SM, Klika AK et al. Patient vs provider characteristics impacting hospital lengths of stay after total knee or hip arthroplasty. *J Arthroplasty* 2011; 26(8): 1418-1426.e1-2. <https://doi.org/10.1016/j.arth.2010.11.008>.
29. Varagunam M, Hutchings A, Black N. Relationship between patient-reported outcomes of elective surgery and hospital and consultant volume. *Med Care* 2015; 53(4): 310-316. <https://doi.org/10.1097/mlr.0000000000000318>.