

Aus der Klinik für Anästhesie und Intensivtherapie

Geschäftsführender Direktor: Prof. Dr. med. Hinnerk Wulf

des Fachbereichs Medizin der Philipps-Universität Marburg

in Zusammenarbeit mit dem Universitätsklinikum Gießen und Marburg GmbH,
Standort Marburg

**Vorhersagemodelle für das Risiko von postoperativer
Übelkeit und Erbrechen –
Eine qualitative Analyse aller verfügbaren Vorhersage-
modelle für das Auftreten von Übelkeit und Erbrechen nach
Narkosen (PONV)**

Inaugural-Dissertation zur Erlangung des Doktorgrades der gesamten

Humanmedizin

dem Fachbereich Medizin der Philipps-Universität Marburg

vorgelegt von

Theresa Monika Bohlender aus Ingolstadt

Marburg, 2022

**Angenommen vom Fachbereich Medizin der Philipps-Universität Marburg am:
23.09.2022**

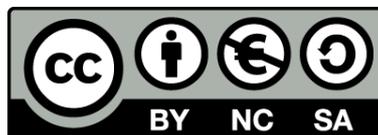
Gedruckt mit Genehmigung des Fachbereiches Medizin

Dekanin: Frau Prof. Dr. D. Hilfiker-Kleiner

Referent: Herr Prof. Dr. L. Eberhart

1.Korreferent: Herr PD Dr. H. Sitter

Originaldokument gespeichert auf dem Publikationsserver der
Philipps-Universität Marburg
<http://archiv.ub.uni-marburg.de>



Dieses Werk bzw. Inhalt steht unter einer
Creative Commons
Namensnennung
Keine kommerzielle Nutzung
Weitergabe unter gleichen Bedingungen
3.0 Deutschland Lizenz.

Die vollständige Lizenz finden Sie unter:
<http://creativecommons.org/licenses/by-nc-sa/3.0/de/>

Inhaltsverzeichnis

Abkürzungsverzeichnis	10
Abbildungsverzeichnis	12
Tabellenverzeichnis	13
1 Einleitung.....	14
1.1 Allgemein	14
1.2 Fragestellung	16
2 Material und Methoden	17
2.1 Rahmenbedingungen	17
2.1.1 Definitionen.....	17
2.1.2 Messung des Endpunkts.....	18
2.2 Ein- und Ausschlusskriterien	19
2.2.1 Einschlusskriterien	19
2.2.2 Ausschlusskriterien	19
2.3 Literatursuche	19
2.3.1 Suchstrategie.....	19
2.3.2 Selektion der Studien	20
2.4 Datenmaske	21
2.5 Qualitätskriterien	22
2.5.1. Generalisierbarkeit.....	22
2.5.2 Praktikabilität	23
2.5.3 Präzision	24
2.5.4 Validierung.....	25
2.5.5 Breite Anwendung in der Praxis und Extrapunkte	26
2.5.6 Tabelle der Qualitätskriterien	27
3 Ergebnisse.....	32
3.1 Gesamtauswertung des Scores Junger 2001	32
3.1.1 Modellentwicklung.....	32
3.1.1.1 Junger 2001	32

3.1.2 Interne Validierung mit Update.....	33
3.1.2.1 Engel 2006.....	33
3.1.3 Qualitätskriterien des Scores Junger 2001.....	34
3.2 Gesamtauswertung des Scores Palazzo & Evans 1993	36
3.2.1 Modellentwicklung.....	36
3.2.1.1 Palazzo 1993.....	36
3.2.2 Externe Validierung ohne Update	36
3.2.2.1 Toner 1996.....	36
3.2.2.2 Eberhart 2000a.....	37
3.2.2.3 Apfel 2002	38
3.2.2.4 Eberhart 2004.....	38
3.2.2.5 Wu 2015.....	39
3.2.3 Impact Assessment.....	40
3.2.3.1 Thomas 2002	40
3.2.4 Qualitätskriterien des Scores Palazzo & Evans 1993.....	41
3.3 Gesamtauswertung des Scores Sinclair 1999	43
3.3.1 Modellentwicklung mit interner Validierung	43
3.3.1.1 Sinclair 1999.....	43
3.3.2 Externe Validierung mit Update.....	44
3.3.2.1 Engel 2006	44
3.3.2.2 Sarin 2012.....	45
3.3.3 Externe Validierung ohne Update	46
3.3.3.1 Apfel 2002	46
3.3.3.2 Pierre 2002.....	46
3.3.3.3 Eberhart 2004.....	47
3.3.4 Impact Assessment.....	47
3.3.4.1 Thomas 2002	47
3.3.5 Qualitätskriterien des Scores Sinclair 1999.....	48
3.4 Gesamtauswertung des Scores Apfel 2012 PDNV	50
3.4.1 Modellentwicklung mit interner Validierung	50

3.4.1.1 Apfel 2012	50
3.4.2 Interne Validierung mit Update.....	51
3.4.2.1 Odom Forren 2013	51
3.4.3 Externe Validierung ohne Update	52
3.4.3.1 Wallden 2016	52
3.4.4 Qualitätskriterien des Scores Apfel 2012	54
3.5 Gesamtauswertung des Scores Eberhart 2004a (POVOC-Score).....	56
3.5.1 Modellentwicklung mit interner Validierung	56
3.5.1.1 Eberhart 2004a.....	56
3.5.2 Externe Validierung ohne Update	57
3.5.2.1 Kranke 2007	57
3.5.2.2 Bassanezi 2013.....	58
3.5.3 Impact Assessment	58
3.5.3.1 Wolf 2016	58
3.5.4 Qualitätskriterien des Scores Eberhart 2004a	61
3.6 Gesamtauswertung des Scores Van den Bosch 2005a.....	63
3.6.1 Modellentwicklung mit interner Validierung	63
3.6.1.1 Van den Bosch 2005a	63
3.6.2 Externe Validierung mit Update.....	65
3.6.2.1 Kappen 2012	65
3.6.3 Impact Assessment.....	66
3.6.3.1 Kappen 2014	66
3.6.3.2 Kappen 2015	67
3.6.3.3 Kappen 2016	68
3.6.4 Qualitätskriterien des Scores Van den Bosch 2005a.....	70
3.7 Gesamtauswertung des Scores Koivuranta-Scores.....	72
3.7.1 Modellentwicklung.....	72
3.7.1.1 Koivuranta 1997	72
3.7.2 Externe Validierung mit Update.....	73
3.7.2.1 Engel 2006	73

3.7.3 Externe Validierung ohne Update	74
3.7.3.1 Eberhart 2000a.....	74
3.7.3.2 Trope 2000.....	74
3.7.3.3 Apfel 2002.....	75
3.7.3.4 Traeger 2003.....	75
3.7.3.5 Eberhart 2004.....	76
3.7.3.6 Van den Bosch 2005	76
3.7.3.7 Peng 2007.....	77
3.7.3.8 Wu 2015.....	78
3.7.4 Impact Assessment.....	78
3.7.4.1 Thomas 2002	78
3.7.4.2 Kumar 2012.....	79
3.7.5.1 Qualitätskriterien des original Scores Koivuranta 1997	80
3.7.5.2 Qualitätskriterien des vereinfachten Scores Koivuranta 1997.....	82
3.8 Gesamtauswertung des Scores Apfel.....	84
3.8.1 Modellentwicklung mit interner Validierung	84
3.8.1.1 Apfel 1998.....	84
3.8.1.2 Apfel 1998a.....	84
3.8.1.3 Apfel 1998b.....	85
3.8.1.4 Apfel 1999.....	86
3.8.2 Externe Validierung mit Update.....	87
3.8.2.1 Engel 2006	87
3.8.2.2 Rodseth 2010.....	88
3.8.2.3 Sarin 2012.....	89
3.8.2.4 Sawatzky 2014	89
3.8.2.5 DaSilva 2015.....	90
3.8.2.6 Schaefer 2015.....	91
3.8.2.7 Lee 2015.....	92
3.8.2.8 Klenke 2018	93
3.8.3 Externe Validierung ohne Update	93

3.8.3.1 Eberhart 1999.....	93
3.8.3.2 Eberhart 2000a.....	94
3.8.3.3 Trope 2000.....	94
3.8.3.4 Apfel 2002.....	94
3.8.3.5 Traeger 2003.....	95
3.8.3.6 Pierre 2002.....	95
3.8.3.7 Apfel 2004.....	95
3.8.3.8 Eberhart 2004.....	96
3.8.3.9 Van den Bosch 2005.....	96
3.8.3.10 Weilbach 2006.....	97
3.8.3.11 Peng 2007.....	97
3.8.3.12 White 2008.....	98
3.8.3.13 Roberts 2010.....	98
3.8.3.14 Pluta 2017.....	99
3.8.3.15 Wu 2015.....	99
3.8.4 Impact Assessment.....	100
3.8.4.1 Rüsçh 1999.....	100
3.8.4.2 Thomas 2002.....	100
3.8.4.3 Biedler 2004.....	101
3.8.4.4 Pierre 2004.....	102
3.8.4.5 Rusch 2005.....	102
3.8.4.6 Kredel 2006.....	103
3.8.4.7 Kapoor 2008.....	103
3.8.4.8 Franck 2010.....	104
3.8.4.9 Latz 2011.....	105
3.8.4.10 Kooji 2012.....	106
3.8.4.11 Ahmed 2014.....	107
3.8.4.12 Kaneda 2004.....	107
3.8.4.13 Pospiech 2014.....	108
3.8.4.14 Sundqvist 2014.....	108

3.8.4.15 Agarkar 2015.....	109
3.8.4.16 Wolf 2016.....	110
3.8.4.17 Gan 2017.....	110
3.8.4.18 Halliday 2017.....	111
3.8.5.1 Qualitätskriterien des original Scores Apfel1998.....	113
3.8.5.2 Qualitätskriterien des vereinfachten Scores Apfel 1999.....	115
3.9 Gesamtauswertung aller Scores.....	117
3.9.1 Gesamtzahl der Publikationen pro Score.....	117
3.9.2 Patientenkollektive pro Score.....	118
3.9.3 Punktwerte der Scores bei den Qualitätskriterien.....	119
4 Diskussion.....	124
4.1 Positive Punkte der Arbeit.....	124
4.2 Einschränkungen der durchgeführten Analyse.....	124
4.3 Diskussion der Ergebnisse.....	125
4.3.1 Umfrage PONV-SOPs in Kliniken.....	125
4.3.1.1 PONV-SOPs für Erwachsene.....	127
4.3.1.2 PONV-SOPs für Erwachsene & Kinder.....	128
4.3.1.3 Vergleich SOP Ergebnisse mit Literaturanalyse der Scores.....	129
4.4 Schlussfolgerungen für die klinische Praxis.....	131
5 Zusammenfassung.....	133
6 Summary.....	135
7 Literatur.....	137
8 Anhang.....	146
8.1 Suchbaum für Datenbanken bei Literatursuche.....	146
8.2 Inhalt der Datenmaske aus Excel.....	149
8.3 Lebenslauf.....	155
8.4 Verzeichnis der akademischen Lehrer.....	156
8.5 Danksagung.....	157
8.6 Ehrenwörtliche Erklärung.....	158

Abkürzungsverzeichnis

Abb.	Abbildung
AIC	Akaike Information Criterion (engl.), Akaike Informationskriterium
ASA	Klassifizierung nach Empfehlung der American Society of Anaesthesiology zur Einschätzung des Narkoserisikos
AUC	Area under the ROC-Curve (engl.), Integral der ROC-Kurve
BA	balancierte Anästhesie
BIC	Bayesian Information Criterion (engl.), Bayesianisches Informationskriterium
BMI	Body Mass Index (engl.), Körpermassenindex
CINV	chemotherapie induced nausea and vomiting (engl.), Chemotherapie induzierte Übelkeit und Erbrechen
CSICU	cardiac surgery intensive care unit (engl.), Intensivstation für Herzchirurgie
EPV	Events per variable (engl.), Ereignisse pro Variable
ICU	intensive care unit (engl.), Intensivstation
KI	Konfidenzintervall
KH	Krankenhaus
KNN	künstliches neuronales Netzwerk
PACU	Post anesthesia care unit (engl.), Aufwachraum
PDNV	Post discharge nausea and vomiting (engl.), Übelkeit und Erbrechen nach Entlassung
PE	Postoperatives Erbrechen
PONV	Postoperativ nausea and vomiting (engl.), postoperative Übelkeit und Erbrechen
POVOC	Postoperative Vomitingin Children (engl.), postoperatives Erbrechen bei Kindern
PN	Postoperative nausea (engl.), postoperative Übelkeit
PR	Postoperatives retching (engl.), postoperatives Würgen

PV	Postoperative vomiting (engl.), postoperatives Erbrechen
RCT	randomized controlled trial (engl.), randomisierte kontrollierte Studie
RF	Risikofaktor/-en
ROC	Receiver Operating Characteristic (engl.), grafische Methode zur Grenzwertoptimierung
RR	relatives Risiko
Sens.	Sensitivität
SOP	Standard Operating Procedure (engl.), Standardverfahren
Spez.	Spezifität
Tab.	Tabelle
TIVA	Totale intravenöse Anästhesie

Abbildungsverzeichnis

Abbildung 1: Flussdiagramm zur Literatursuche	21
Abbildung 2: PONV-Risiko und Inzidenz [96]	60
Abbildung 3: Scorchart Van den Bosch 2005a [88]	64
Abbildung 4: Normogramm Van den Bosch 2005a [88]	64
Abbildung 5: Scorechart zur Erbrechenswahrscheinlichkeit in % [5]	85
Abbildung 6: Anzahl der Publikationen pro Score	118
Abbildung 7: Patientenkollektive pro Score	119
Abbildung 8: Gesamtpunktzahl der Scores in der Gesamtauswertung	120
Abbildung 9: einzelne Punktwerte der Scores pro Qualitätskriterium in der Gesamtauswertung	121
Abbildung 10: Antwortverhalten der Kliniken bei der Umfrage bezüglich PONV SOPs	125
Abbildung 11: Scores bei PONV-SOPs	130
Abbildung 12: SOPs – generelle Prophylaxe vs. risikoadaptiertes Vorgehen	131

Tabellenverzeichnis

Tabelle 1 Tabelle der Qualitätskriterien	31
Tabelle 2 Qualitätskriterien des Scores Junger 2001.....	35
Tabelle 3: Qualitätskriterien des Scores Palazzo & Evans 1993	42
Tabelle 4: Qualitätskriterien des Scores Sinclair 1999.....	49
Tabelle 5: Wahrscheinlichkeit für das Auftreten von PDNV: Apfelscore 2012 [11].....	51
Tabelle 6: Qualitätskriterien des Scores Apfel 2012	55
Tabelle 7: Wahrscheinlichkeit für das Auftreten von PV: Eberhartscore 2004a [25] ...	57
Tabelle 8: Qualitätskriterien des Scores Eberhart 2004a.....	62
Tabelle 9: Qualitätskriterien des Scores Van den Bosch 2005a	71
Tabelle 10: Wahrscheinlichkeit für das Auftreten von PON bzw. PV: Koivurantascore1997 [46]	73
Tabelle 11: Qualitätskriterien des original Scores Koivuranta 1997	81
Tabelle 12: Qualitätskriterien des vereinfachten Scores Koivuranta 1997	83
Tabelle 13: Wahrscheinlichkeit für das Auftreten von PONV: vereinfachter Apfelscore 1999 [7].....	87
Tabelle 14: Wahrscheinlichkeit für das Auftreten von PONV: Score von Da Silva 2015 [19]	91
Tabelle 15: Qualitätskriterien des original Scores Apfel1998.....	114
Tabelle 16: Qualitätskriterien des vereinfachten Scores Apfel 1999.....	116

1 Einleitung

1.1 Allgemein

Die Inzidenz von Übelkeit und Erbrechen (PONV = postoperative nausea and vomiting) nach Operationen bei Nutzung von Inhalationsanästhetika beträgt ohne zusätzliche Prophylaxe im Schnitt 30%[79]. Somit ist PONV neben postoperativen Schmerzen eine sehr häufige Komplikation nach Operationen in Allgemeinanästhesie[74].

Die Pathophysiologie von PONV ist bislang noch nicht vollständig geklärt. Jedoch konnten in den letzten Jahrzehnten diverse Risikofaktoren für das Auftreten identifiziert werden. So gibt es beispielsweise eine Publikation zu diesem Thema bereits aus dem Jahr 1973, in der die einzelnen Risikofaktoren: Alter, Geschlecht, Anästhesiedauer, Narkosemittel und Schwere der Operation benannt werden.[54]

Da ein einzelner Risikofaktor alleine keine zufriedenstellende PONV-Vorhersage liefert, wurden im Laufe der Jahre viele verschiedene Prognosesysteme entwickelt. Die PONV-Scoringssysteme bieten eine individuelle, orientierende Risikoeinschätzung.

Dieses Wissen kann für eine risikoadaptierte oder eine fixe Kombinationsprophylaxe genutzt werden.[74] So sind die Medikamente Dexamethason, Droperidol und Ondansetron etwa gleich potent. Sie erzielen eine relative Risikoreduktion von 26%. Eine TIVA führt zu einer relativen Risikoreduktion von 31%.[9] Bei einer Kombination dieser Maßnahmen addieren sich die Effekte.

PONV erhöht das Risiko für postoperative Komplikationen und steigert das Risiko für postoperative Naht Dehiszenz[28]. Es können schwerwiegende Folgen wie Trachealruptur[35] mit einhergehendem Pneumothorax[16] oder ausgeprägte Hautemphyseme [78] entstehen. In Einzelfällen sind das Boerhaave-Syndrom [12] oder ein Visusverlust [98] nach PONV beschrieben. Bei pädiatrischen Patienten stellt PONV den häufigsten Grund für ungeplante stationäre Aufnahmen nach ambulanten Operationen dar [15,61,74].

Auch ökonomische Aspekte spielen eine Rolle, da PONV zu einer signifikanten Verlängerung der Verweildauer im Krankenhaus bzw. dem Aufwachraum führt. Des Weiteren kommt es zu erhöhtem personellen Betreuungsaufwand. Eine Prozesskostenanalyse zeigt, die durchschnittliche Arbeitszeit des Personals für die Behandlung einer Übelkeit und Erbrechen Episode beträgt 26 Minuten. In Deutschland entspricht dies durchschnittlichen Personalkosten von 18 €. Die Materialkosten im Wert von 13€, v.a. für Antiemetika kommen ebenfalls hinzu. Folglich hat Übelkeit und Erbrechen Auswirkungen auf die Arbeitsbelastung des Krankenhauspersonals und eine wirtschaftliche Belastung von 31€ pro Episode. [21] Somit spielt PONV eine

wichtige Rolle für die Ökonomie und die Patientenzufriedenheit.[59]Zudem sind effektive und sichere Medikamente zur Prophylaxe verfügbar.

Dies zeigt, PONV ist häufig, relevant und kann vorgebeugt werden. Es stellt sich die Frage, sollen alle Patienten eine Prophylaxe erhalten, oder muss eine Selektion erfolgen. Für eine Differenzierung der Patienten ist eine Prognose durch ein Vorhersagesystem nötig. Die Prognosesysteme dienen der Risikoerfassung, der Klassifizierung von Krankheitsbildern, oder um Patienten zu identifizieren, die stationär aufgenommen werden müssen. Für alle Risikomodelle gilt, dass sie definierte Punktwerte vergeben, diese verrechnen und das Ergebnis auf einer einheitlichen Skala abbilden.[95]

Vorhersagemodelle sind nützliche Werkzeuge, haben jedoch Einschränkungen, die ihre Nutzung limitieren oder zu Fehlinterpretationen führen können. Das Auffinden eines Zusammenhangs zwischen einem klinischen Faktor und einem Endpunkt kann immer noch fehleranfällig sein. Beispiele hierfür sind: fehlerhafte Datensammlung, falsche Interpretation von Ergebnissen oder die Beschreibung eines Zusammenhangs, der auf verfälschende Variablen zurückzuführen ist.[18] Bevor eine Vorhersageregul angenommen wird, muss ihre Anwendbarkeit auf die Patienten geprüft werden. [92]

Die Qualität einer klinischen Vorhersageregul ist durch diverse Merkmale gekennzeichnet. Ein gutes Modell besitzt eine geringe Anzahl an Variablen und ist leicht anzuwenden. Je komplizierter ein Vorhersagemodell ist, desto unwahrscheinlicher wird es dauerhaft, in der täglichen Praxis verwendet. Ein genau definierter Endpunkt oder ein eindeutiges Ergebnis sind elementar. Dieser Fixpunkt muss leicht zu messen, klinisch relevant und nicht von extern beeinflussbar sein. Die Regul sollte aus einer großen Datenbank abgeleitet werden, damit sie auf eine Vielzahl von Patienten anwendbar ist. Bei einem kleinen Kollektiv könnte es zu Verzerrungen kommen. Fehler sind bei allen Vorhersageinstrumenten unvermeidlich, aber mit fortlaufender Verfeinerung und Aktualisierung können sie auf ein akzeptables Maß reduziert werden. Eine ständige Anpassung der Modelle ist durch neue Krankheiten, neuen Therapien und neuen Behandlungsansätzen für das Patientenmanagement nötig. [18]

Diese Erkenntnisse sollen Einzug in die alltägliche Krankenhauspraxis und Praxisleitfäden finden sowie weitere Forschung nach sich ziehen[74].

1.2 Fragestellung

Das Anliegen dieser Arbeit ist es die Vorhersagemodelle für das Risiko von postoperativer Übelkeit und Erbrechen zu detektieren– eine qualitative Analyse aller verfügbaren Vorhersagemodelle für das Auftreten von PONV erfolgt im Anschluss. Es ergeben sich folgende Fragen: Welche Modelle haben die größte Validität und Generalisierbarkeit? Welche Modelle sind auf verschiedene Patientengruppen übertragbar und sollten bevorzugt im klinischen Alltag bzw. in der Forschung angewendet werden? Bei welchen Scores war die Anwendung mit einem messbaren bzw. mit einem nachweisbaren Nutzen verbunden? Das Ziel ist den „besten“ Score zu detektieren und eine konkrete Empfehlung auszusprechen.

Um diese Fragen zu beantworten wurden eigene Qualitätskriterien entwickelt und die jeweiligen Publikationen eines Scores zu einer Gruppe zusammengefasst und bewertet. Da die Effizienz spezifischer PONV-Scoringmodelle sehr von der Risikoverteilung in dem jeweiligen Patientenkollektiv abhängt, scheint eine einfache Empfehlung für einen bestimmten Algorithmus schwierig. Diese Arbeit soll zeigen, welche Systeme sich bereits etabliert haben. Abschließend werden die gewonnenen Ergebnisse mit den SOPs aus dem Klinikalltag verglichen. Hierzu wird eine exemplarische, stichprobenartige Befragung durchgeführt.

2 Material und Methoden

2.1 Rahmenbedingungen

Diese Arbeit entstand initial im Rahmen eines „Systematic Review“ der Cochrane Collaboration mit dem Titel: „Prediction models for the risk of postoperative nausea and vomiting“ [59]. Als federführende Mitglieder der Arbeitsgruppe dieses Projekt aufgrund von fehlenden Kapazitäten verließen stagnierte der Fortschritt des Reviews. Der zuvor geplante große Umfang war mit der reduzierten Teamgröße schwer zu realisieren. Diese Doktorarbeit wurde aus dem Projekt entkoppelt und als eigenständige qualitative Analyse weitergeführt. Die Literatursuche sowie die Einschluss- und Ausschlusskriterien für Studien stimmen mit dem Protokoll der ursprünglich angestrebten Übersichtsarbeit überein. Alle weiteren Schritte wurden autonom durchgeführt.

Aufgabe dieser Arbeit ist es, alle Prognosemodelle zu detektieren, die entwickelt wurden, um das Risiko von PONV abzuschätzen. Anschließend erfolgt eine Charakterisierung der Prognosemodelle, indem die Vorhersageparameter oder die Entwicklungsdatensätze beschrieben werden. Im nächsten Schritt geht es um die Bewertung und Beurteilung der Vorhersagegenauigkeit der Modelle anhand von externen Validierungen sowie der Diskriminations- und Kalibrierungsfähigkeit. Abschließend wird ein "Impact Assessment" d.h. der Anwendungsbereich und der direkte Vorteil durch die Anwendung des Scores aufgeführt. Das Ziel ist es, den „besten“ Score zu identifizieren.

2.1.1 Definitionen

PONV ist per Definition ein zusammengesetzter Endpunkt, der postoperative Übelkeit (PN), postoperatives Erbrechen (PV) oder das Vorhandensein beider einschließt. Gelegentlich wird auch postoperatives Würgen (PR) unter PONV zusammengefasst. Es werden Prognosemodelle eingeschlossen, die einen oder die Kombination dieser Punkte berücksichtigen. [59]

In dieser Arbeit wird ein Prognosefaktor für PONV als ein Faktor definiert, der mit einem höheren oder niedrigeren Risiko für PONV assoziiert ist. [69]

Ein Prognosemodell für PONV ist eine zahlenmäßige Darstellung des geschätzten Likelihood Ratio für das Event PONV, basierend auf einem Patientenwert für eine Reihe von Prognosefaktoren. [81]

Als Studienteilnehmer wird jeder Patient, d.h. Kinder und Erwachsene, die sich einer Anästhesie (Lokal-, Regional- und Allgemeinanästhesie) und einer OP jeglicher Art unterziehen, akzeptiert [59].

Primärer Endpunkt ist der am häufigsten genutzte Endpunkt für prognostische Forschungsmodelle, der das Auftreten von PONV innerhalb der ersten 24 Stunden nach einer OP erfasst [59].

Die sekundären Endpunkte beinhalten PN und PV, die einzeln genutzt werden. Weitere sekundäre Endpunkte sind das späte Auftreten von PONV bis 72 Stunden nach einer OP und PONV nach Entlassung aus dem Krankenhaus (PDNV). Manche Scores nutzen nur die schweren Fälle von PONV oder dichotome Endpunkte wie der Gebrauch von Antiemetika oder verspätete Patientenentlassungen. All diese Modelle werden eingeschlossen. [59]

2.1.2 Messung des Endpunkts

PONV kann gemessen werden in der Intensität der Symptome oder der Häufigkeit des Auftretens des Ereignisses. Für diese Arbeit muss der Endpunkt dichotom sein, d.h. vorhanden oder nicht vorhanden, um in die Prognosemodelle eingesetzt zu werden.

Der Zeitpunkt der Erfassung des Endpunkts kann an verschiedenen Stationen zwischen postoperativer Betreuung und Genesung bewertet werden. Prognosemodelle, die folgende Zeitintervalle nutzen, wurden berücksichtigt:

- in dem Aufwachraum
- während einer fixen Zeitspanne
- die gesamte Zeit des postoperativen Krankenhausaufenthaltes
- die anfänglichen Tage nach Entlassung aus der Einrichtung

Während der Routineversorgung kann der Endpunkt von medizinischen Fachkräften freiwillig gemeldet werden. In prospektiven Forschungsstudien werden spezielle Beobachter genutzt. Des Weiteren kann eine direkte Befragung des Patienten erfolgen. Prognosemodelle, die das Auftreten des Endpunktes in jeglicher Art und Weise erfassen, wurden eingeschlossen. [59]

2.2 Ein- und Ausschlusskriterien

Die Einschluss- und Ausschlusskriterien dieser Arbeit sind identisch mit den Kriterien des Systematic Reviews der Cochrane Collaboration.

2.2.1 Einschlusskriterien

Gesucht wurden prognostische Forschungsmodelle, die sich mit der Entwicklung, Validierung und den direkten Vorteilen, wenn PONV-Modelle eingesetzt werden, befassen. Dies beinhaltet folgende Studientypen:

- Vorhersagemodelle mit/ohne internerValidierung
- externe Modellvalidierung mit/ ohne Modellüberarbeitung
- Studien zum ImpactAssessment [59]

2.2.2 Ausschlusskriterien

Ausschlusskriterien waren folgende falsche Endpunkte:

- nicht postoperative Übelkeit
- nicht postoperatives Würgen
- nicht postoperatives Erbrechen
- nicht postoperative Übelkeit und Erbrechen

Eine Studie mit falschem Patientenkollektiv wurde nicht berücksichtigt. Somit wurden Patienten mit voraussichtlicher postoperativer Verlegung auf die Intensivstation ausgeschlossen, da sie dort oft mechanisch beatmet werden.

Ebenso wurde ein falsches Studiendesign, z.B. reine Beobachtungsstudien mit Risikofaktorsuche ohne anschließende Modellentwicklung nicht einbezogen. [59]

2.3 Literatursuche

Die Literatursuche entspricht dem Vorgehen des Systematic Reviews mit dem Titel: „Prediction models for the risk of postoperative nausea and vomiting“

2.3.1 Suchstrategie

Ziel war es, alle Publikationen in den folgenden elektronischen Datenbanken zu finden:

- The Cochrane Library
- Ovid MEDLINE
- Ovid EMBASE
- Web of Science des Institute of Scientific Information (ISI)

- Cumulative Index to Nursing and Allied Health Literature (CINAHL)(EBSCO host)
- Latin American and Caribbean Health Science Information Database(LILACS) (BIREME)
- database of prognostic studies betreut von the Cochrane Prognosis Methods Group (PMG)

Die Suchstrategien wurden modifiziert, hierbei wurden die von Geersinget al empfohlen Suchfilter benutzt [30].Die exakten Kombinationen der Suchbegriffe pro Datenbank befinden sich im Anhang unter Punkt 8.1. Dabei gab es keine Einschränkungen für Sprache oder Erscheinungsdatum der Publikation. Anschließend wurde im Literaturverzeichnis der eingeschlossenen Studien nach weiteren interessanten Publikationen gesucht.[59]

2.3.2 Selektion der Studien

Nach Überprüfung der sieben elektronischen Datenbanken im Jahre 2014 wurden die 1384 Publikationen der ersten Literatursuche auf die Plattform Rayyan [68] hochgeladen. Dies ist eine Webanwendung, mit der Autoren systematische Überprüfungen der Publikationen anhand von Überschrift und Abstract durchführen können. Nach dem ersten Screeningvorgang wurden 183 Publikationen eingeschlossen. Hierunter befanden sich einige Duplikate, sodass schließlich 175 Publikationen zum Volltextscreening verblieben. Alle Volltexte wurden auf die Plattform google drive [31] hochgeladen und komplett gesichtet. Abschließend wurden 59 Publikationen eingeschlossen. Dies war der letzte gemeinsame Arbeitsschritt mit dem Team der Cochrane Gruppe.

Da die Literatursuche bereits ein paar Jahre zurücklag, wurde im April 2018 eigenständig ein Literaturupdate durchgeführt. Nach erneutem Durchsuchen der sieben Datenbanken ergab dies 391 Publikationen, deren Abstracts zu überprüfen waren. 326 Publikationen blieben nach Entfernung der Duplikate erhalten. Nach dem Screening der Abstracts blieben 76 Publikationen zur Volltextüberprüfung. Hiervon wurden 26 eingeschlossen. Zusammenfassend ergibt dies aus beiden Literatursuchen 85 Volltexte. Da die Publikationen in diversen Fremdsprachen z.B. Schwedisch, Norwegisch, Bulgarisch, Japanisch und Tschechisch vorlagen, wurden diese mithilfe von google translate[32] und Bekannten, die diese Sprachen beherrschen, übersetzt.

Das nachfolgende Flussdiagramm veranschaulicht die gesamte Literatursuche.

Literatursuche 2014

Literaturupdate April2018

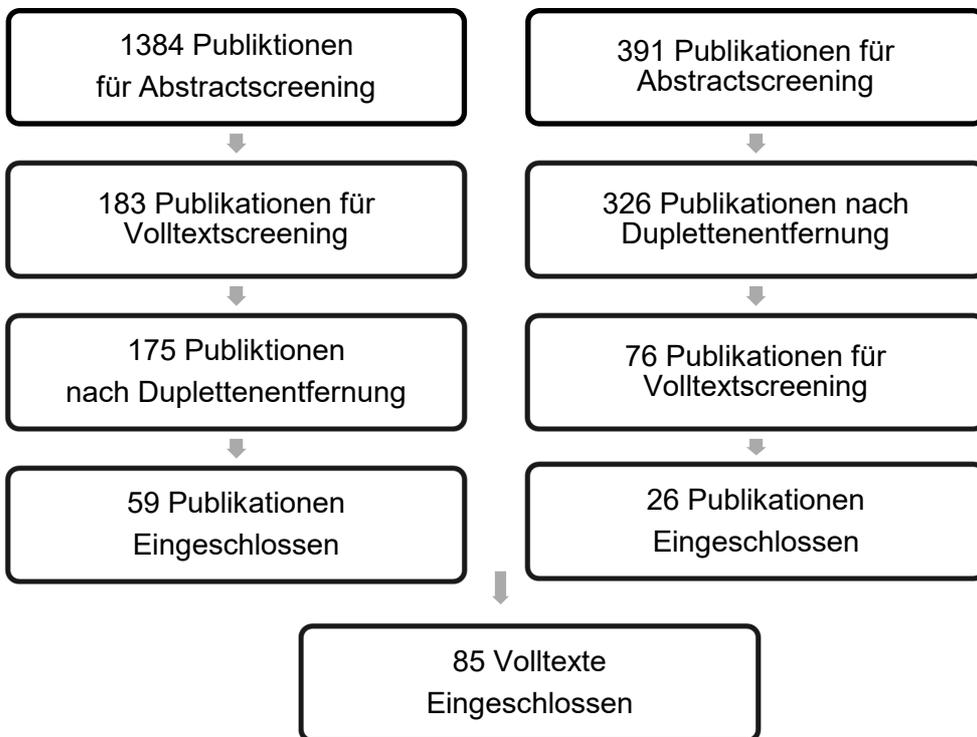


Abbildung 1: Flussdiagramm zur Literatursuche

2.4 Datenmaske

Zur Verwaltung der Daten, die den Publikationen zu entnehmen sind, wurde eine Exceltabelle als Datenmaske erstellt. Eine einheitliche, ausführliche Oberfläche für alle Publikationstypen wurde entworfen. Die meisten Felder der Datenmaske waren zur Vereinfachung der Datenauswertung mit ja/ nein als geschlossene Fragen zu beantworten. Zusätzlich gab es einige Freitextfelder. Das Ziel war eine großzügige Datengewinnung. Eine genaue Auflistung der einzelnen Punkte und Unterpunkte findet sich im Anhang unter Punkt 8.2.

Initial wurde unterschieden, ob es sich um eine Modellentwicklung, Validierung oder ImpactAssessment handelt. Im Anschluss wurde sich mit der Datenquelle (retrospektiv vs. prospektiv), der Qualität der Daten sowie dem Setting der Studie befasst. Ein- und Ausschlusskriterien, die Definition der Endpunkte sowie Ort, Zeit und Art und Weise der Endpunkterhebung wurden erfasst. Der nächste große Abschnitt waren die Prädiktoren eines Modells. Hier wurde pro Prädiktor einzeln die Anzahl, die genaue Definition, die Reliabilität und die Art und Weise der Datenerfassung dokumentiert.

Zusätzlich wurde kodiert, ob der Prädiktor ohne Kenntnis des Outcomes erfasst wurde. Außerdem wurde der Zeitpunkt der Endpunkterfassung und die Handhabung der Prädiktoren im Modell (z.B. binär vs. linear) untersucht. Weitere Unterpunkte waren die Stichprobengröße, die Anzahl und der Umgang mit fehlenden Daten. Die Methodik bei der Modellbildung sowie die Auswahlkriterien bzw. Ein- und Ausschlusskriterien der Prädiktoren wurden dokumentiert. Extrahiert wurden die Modellkalibrierung, die Diskriminationsfähigkeit sowie die Maßzahlen der Modellklassifizierung. Zudem wurde untersucht, ob es eine Modellprüfung mit einem Trainingsdatensatz oder externen Daten gab. Außerdem wurde z.B. eine Überarbeitung eines schlecht validierten Systems berücksichtigt. Das endgültige, als auch ein alternatives Vorhersagemodell wurden gespeichert. Zusätzlich gab es die Punkte „Anwendbarkeit in der täglichen Praxis“ sowie „Auswirkungen auf die Forschung“ zur Dokumentation des Impact Assessments.

2.5 Qualitätskriterien

Aus den einzelnen Publikationen wurden Cluster gebildet. Alle Publikationen zu einem Score wurden zu einer Gruppe zusammengefasst und bewertet. Hierfür wurden vorab Qualitätskriterien festgelegt. Das Bewertungsinstrument besitzt folgende fünf Untergruppen:

- Generalisierbarkeit
- Praktikabilität
- Präzision
- Validierung
- Breite Anwendung in der Praxis und Extrapunkte

Die Maximale Punktzahl, die insgesamt erreicht werden konnte, sind 60 Punkte.

2.5.1. Generalisierbarkeit

Die Generalisierbarkeit umfasst die Größe des Patientenkollektivs, in dieser Kategorie wurden 3 Punkte vergeben. Hier wurde unterteilt in Scoreentwicklung und Scoreupdate. Je mehr Patienten eingeschlossen wurden, desto mehr Punkte wurden vergeben. Durch ein großes Patientenkollektiv kann dem Selektionsbias vorgebeugt werden. Dieser entsteht, wenn der ursprüngliche Datensatz des Modells nicht groß genug ist, um auf eine neue Zielpopulation angewendet zu werden. Unvorhergesehene oder ungewöhnliche Eigenschaften des ursprünglichen Patientenkollektivs können die Variablen beeinflussen. Die Fehlerrate einer Vorhersageregeln und deren potentielle

Folgen muss bekannt sein. Wenn ein Modell bestimmte Patienten nicht berücksichtigt, die anschließend versterben, wäre das Modell fehlerhaft. Eine Anpassung der Regel wäre hier notwendig. [18]

2.5.2 Praktikabilität

Die Praktikabilität umfasst die Risikofaktoren, die Risikoberechnung und die Endpunkte. Es konnten maximal 19 Punkte erreicht werden.

Zuerst wurden die Risikofaktoren in dem Modell bewertet. Je weniger Faktoren das Modell enthält und je einfacher sie zu erheben sind, desto handlicher ist das Modell und umso mehr Punkte wurden vergeben. Ein weiterer Aspekt war die Reliabilität der Risikofaktoren. Somit war eine genaue Definition der Risikofaktoren wichtig. Des Weiteren wurde beurteilt, ob die Faktoren ohne Kenntnis des Outcomes beurteilt wurden, um eine Verfälschung zu vermeiden. Die Genauigkeit, mit der die Risikofaktoren dokumentiert wurden, ist von Bedeutung. Kontinuierliche Variablen beziehen sich auf Faktoren, die genau gemessen werden können, um ein Spektrum von Abnormalitäten widerzuspiegeln. Bei nichtkontinuierlichen Variablen, z.B. für eine körperliche Untersuchung, ist eine größere Variabilität der Ergebnisse vorhanden. Vorhersagemodelle, die auf diesen Einschätzungen beruhen, sind bei schlechter Reproduzierbarkeit fehleranfälliger. [18]

Als nächstes war zu betrachten, ob zur Risikoberechnung Hilfsmittel wie z.B. in Form von Software, Taschenrechner oder Tabellenwerk nötig sind. Falls ein vereinfachter Score vorlag, wurde dies ebenfalls berücksichtigt. Unter einem vereinfachten Risikoscore wurde verstanden, dass alle Faktoren des Modells gleich gewichtet werden und dass durch die Anzahl an Risikofaktoren ohne Rechnung das Patientenrisiko ablesen werden kann. Je handlicher der Score war, desto mehr Punkte wurden vergeben. [18]

Der Endpunkt des Scores, wie z. B. PONV als kombinierter Endpunkt, nur Teilaspekte wie Postoperative nausea (PN), Postoperative vomiting (PV), Postoperatives retching (PR), oder keine genaue Definition wurden bewertet. Die Art und Weise der Endpunkterhebung spielte ebenfalls eine Rolle. Dies ging von routinemäßiger Datensammlung, die leicht in den klinischen Alltag implementierbar ist, bis zur zeitaufwendigeren, aber genaueren persönlichen Visite des Patienten. Der Erfolg eines prädiktiven Instruments ist abhängig von einer genauen Datensammlung. Die Erfassung, Interpretation und Eingabe der Daten unterliegen menschlichen Fehlern. Die Interobserver-Zuverlässigkeit wird durch geschulte Datensammler erhöht.

Idealerweise würde ein Vorhersagemodell Daten verwenden, die direkt von Überwachungsgeräten stammen um diese Fehlerquelle zu eliminieren. Jedoch verfügen die medizinischen Informationssysteme in den meisten Krankenhäusern nicht über diese Funktion. [18]

2.5.3 Präzision

Die Vorhersagegenauigkeit des Modells wurde anhand der Diskriminationsfähigkeit und der Kalibrierung bewertet. Die maximal erreichbare Punktzahl waren 10 Punkte.

Die Diskriminationsfähigkeit zeigt, wie gut das Modell zwischen Patienten mit PONV und ohne PONV unterscheiden kann. Dies wird grafisch in einer Grenzwertoptimierungskurve (ROC = Receiver Operating Characteristic) dargestellt. Um die Sensitivität und die Spezifität für die ROC-Kurve zu berechnen, wird eine Vierfeldertafel mit den vorhergesagten und den tatsächlichen Häufigkeiten von PONV benötigt. Auf der x-Achse des Graphen wird 1-Spezifität aufgetragen und auf der y-Achse die Sensitivität. Somit ergibt sich ein Streudiagramm, in dem für jeden einzelnen cut-off-Wert die richtig positiven Testergebnisse mit dem Anteil der falsch positiven Testergebnisse korrelieren. Die Fläche unter dieser Kurve ist die Area under the curve (AUC). Die AUC ist eine einheitslose Maßzahl, mit einem Intervall von 0 bis 1.

Eine AUC < 0,5 hat eine Vorhersagekraft schlechter als der Zufall. Eine AUC von 0,5 hat die gleiche Präzision wie der Zufall und eine AUC > 0,5 hat eine Vorhersagekraft besser als Zufall. Die maximale Vorhersagekraft besitzt eine AUC = 1. [34][1]

Zur Beurteilung der Prädiktion mittels Diskriminationsfähigkeit (AUC der ROC-Kurve) bei mehreren ROC-Kurven wurde eine Berechnung durchgeführt. Zuerst wurde zwischen internen und externen Patientendaten unterschieden. Für beide Gruppen wurden getrennt gewichtete Mittelwerte der Fläche unter der Kurve erstellt (je nach Fallzahl der Studie). Der erreichte Punktwert für externe Daten wurde mit 1,1 multipliziert, da diese für gewöhnlich schlechter abschneiden. Anschließend wurde der Durchschnitt aus interner und externer Fläche zu allen AUC-Werten eines Scores gebildet. Je größer diese Fläche war, desto mehr Punkte wurden abschließend vergeben.

Die Kalibrierung ist die statistische Übereinstimmung der Prognosewahrscheinlichkeit mit dem tatsächlichen Ergebnis. Häufig wird sie grafisch mit einer Geradengleichung oder einem Kalibrierungsplot dargestellt. Die prognostizierten Wahrscheinlichkeiten werden nach Risikogruppen sortiert. Anschließend erfolgt die Berechnung der beobachteten relativen Häufigkeiten von PONV bei den jeweiligen Risikogruppen. Bei

der Geradengleichung wird das prognostizierte Risiko der Gruppe gegen das beobachtete Risiko aufgetragen. Eine Ausgleichsgerade wird zwischen die Punkte gelegt. Bei linearen Beziehungen ist eine Steigung von 45 Grad und ein Achsenabschnitt von 0 ideal. Hingegen wird bei dem Kalibrierungsplot die relative Häufigkeit des Endpunkts berechnet und gegen die mittlere Prognosewahrscheinlichkeit der jeweiligen Gruppe aufgetragen. [34]

Zusätzlich kann die Kalibrierung durch den Hosmer-Lemeshow-Test bewertet werden. Die Stärke der Abweichung von der geschätzten Wahrscheinlichkeit zur beobachteten Häufigkeit wird durch diesen Test quantifiziert. Falls dieser Test durchgeführt wurde, erfolgte die Vergabe eines Punktes. [34]

Zum Vergleich mehrerer Geradengleichungen eines Scores wurde eine Berechnung durchgeführt. Folgende Grundüberlegungen gehen hierfür voraus: Die ideale Geradengleichung der Kalibrierungsgeraden lautet: $Y=1x+0$, somit entspricht der x-Wert dem y-Wert. Von der „Geradengleichung des Scores der jeweiligen Publikation“ wurde die „Idealgerade“ subtrahiert. Auf diese Weise entstand eine 3. Geradengleichung, die „Differenzgerade“. Im Anschluss wurde die Fläche unter der Differenzgeraden und der x-Achse berechnet. Diese Fläche war die Einheit zur Bewertung. Je kleiner die Fläche zwischen Differenzgerade und x-Achse ist, desto näher ist die Beispielgerade an der Idealgeraden und umso besser war die Kalibrierung. Für alle Geradengleichungen eines Scores wurden diese Berechnungen durchgeführt und der Durchschnitt gebildet. Dieser Abschlusswert wurde dann bepunktet.

2.5.4 Validierung

Bei der Validierung wurde zwischen interner und externer Validierung unterschieden, maximal wurden 15 Punkte vergeben.

Eine interne Validierung des Scores z.B. durch Splitten der Daten wurde bepunktet. Die Gesamtzahl des hierfür genutzten Patientenkollektivs fand ebenfalls Berücksichtigung.

Bei externen Validierungen eines Scores zählte die Anzahl der Validierungen. Zusätzlich wurde die Lokalisation der Validierungen z.B. auf einem anderen Kontinent berücksichtigt, da andere Ethnien unterschiedliche Risiken aufweisen können. Ebenso wurde die Gesamtzahl des Patientenkollektivs berücksichtigt.

2.5.5 Breite Anwendung in der Praxis und Extrapunkte

Diese Rubrik umfasst das Impact Assessment sowie Extrapunkte, es konnten maximal 13 Punkte erreicht werden.

Es gab eine große Zunahme von Publikationen zu Bewertungssystemen, ohne dass bisher ein eindeutiger klinischer Nutzen untersucht wurde [17]. Somit wurde hier das Impact Assessment eines Scores geprüft. Bepunktet wurde, ob dessen Anwendung einen Benefit z.B. in Form von geringerer PONV Inzidenz, höherer Patientenzufriedenheit, kürzerer Krankenhausverweildauer oder weniger Komplikationen aufwies. Darüber hinaus werden die Einsatzbereiche bisheriger Nutzung und Anwendung des Scores zur Identifizierung von Hochrisikopatienten, Vergleich von Studienteilnehmern und Herstellung von Kohortengleichheit bepunktet. Die Gesamtzahl des Patientenkollektivs wird berücksichtigt.

Abschließend gibt es noch vier mögliche Extrapunkte zur Modifikation der erreichten Punktzahl, für Aspekte, die in der Tabelle nicht erwähnt wurden, beispielsweise besondere statistische Werte.

2.5.6 Tabelle der Qualitätskriterien

Kriterium	Untergruppe	Unterpunkt	Max. erreichte Punktzahl
Generalisierbarkeit	Anzahl des Patientenkollektivs	Scoreentwicklung & -update	<p>3 P > 2500</p> <p>2,5 P > 2000 und ≤ 2500</p> <p>2 P > 1500 und ≤ 2000</p> <p>1,5 P > 1000 und ≤ 1500</p> <p>1 P > 500 und ≤ 1000</p> <p>0,5 P > 0 und ≤ 500</p>
Praktikabilität	Risiko-faktoren in dem Modell	Anzahl der RF in dem Modell	<p>3P 1-3 RF</p> <p>2P 4-6 RF</p> <p>1P 7-9 RF</p> <p>0P ≥ 10 RF</p>
		einfach zu erheben (ideal: ohne technische Hilfsmittel, geringer Zeitaufwand, sind durch Fragen erhebbar, die sich klar mit ja/nein beantworten lassen)	<p>3P alle ohne technische Hilfsmittel, mit ja/nein Antwort, geringer Zeitaufwand</p> <p>2P Mehrheit der RF ohne technische Hilfsmittel mit ja/nein Antwort, mittlerer Zeitaufwand</p> <p>1P Mehrheit der RF ohne technische Hilfsmittel mit diversen Antwortmöglichkeiten, höherer Zeitaufwand</p> <p>0P Mehrheit der RF Erhebung mittels Hilfsmittel, hoher Zeitaufwand</p>
		RF sind exakt definiert, hoch reliabel	<p>3P alle RF sind exakt definiert → alle RF sind hoch reliabel</p>

			<p>2P Mehrheit der RF sind exakt definiert → Mehrheit der RF sind hoch reliabel</p> <p>1P Minderheit der RF sind exakt definiert → Minderheit der RF sind hoch reliabel</p> <p>0P keine Definition der RF → RF sind nicht hoch reliabel</p>
		RF ohne Kenntnis des Outcomes beurteilt	<p>1P ja</p> <p>0P nein/ unklar</p>
Praktikabilität	Risiko- berechnung/ Handhabung des Scores	Hilfsmittel nötig (Software, Taschenrechner, Tabellenwerk...)	<p>3P keine Hilfsmittel</p> <p>2P Taschenrechner/mit Handy (z.B. als App möglich)</p> <p>1P spezielles Tabellenwerk/Schaubild</p> <p>0P Software</p>
		vereinfachter Score	<p>1P vereinfachter Score vorhanden</p> <p>0P kein vereinfachter Score</p>
Praktikabilität	Endpunkt des Scores	PONV kombiniert, nur Teilaspekte (PN, PV,PR) oder unklar	<p>2 P zusammengesetzter Endpunkt PONV (max. PN+PV+(PR))</p> <p>1 P einzelner Endpunkt PN/PV/PR</p> <p>0 P unklarer Endpunkt</p>
		Art und Weise der Endpunkterhebung	<p>3P via Routinedaten</p> <p>2P via extra Fragebogen</p>

			<p>1P via Gespräch (persönliche Visite/ Telefonat)</p> <p>0P unklarer Modus</p>
Präzision	Vorhersagegenauigkeit/ Beurteilung der Prädiktion mittels:	Diskriminationsfähigkeit (AUC der ROC-Kurve)	<p>Punktevergabe anhand der AUC unter der ROC-Kurve:</p> <p>5 P $\geq 0,95$ und < 1</p> <p>4,5 P $\geq 0,90$ und $< 0,95$</p> <p>4 P $\geq 0,85$ und $< 0,90$</p> <p>3,5P $\geq 0,80$ und $< 0,85$</p> <p>3 P $\geq 0,75$ und $< 0,80$</p> <p>2,5 P $\geq 0,70$ und $< 0,75$</p> <p>2 P $\geq 0,65$ und $< 0,70$</p> <p>1,5 P $\geq 0,60$ und $< 0,65$</p> <p>1 P $\geq 0,55$ und $< 0,60$</p> <p>0,5 P $\geq 0,50$ und $< 0,55$</p> <p>0 P $< 0,5$</p>
		<p>Kalibrierung</p> <ul style="list-style-type: none"> ▪ Geradengleichung (Calibration Slope) Fläche 	<p>Punktevergabe anhand der Flächeneinheiten:</p> <p>3 P > 0 und ≤ 1</p> <p>2,5 P > 1 und ≤ 2</p> <p>2 P > 2 und ≤ 3</p> <p>1,5 P > 3 und ≤ 4</p> <p>1 P > 4 und ≤ 5</p> <p>0,5 P > 5 und ≤ 6</p> <p>0 P > 6</p>
		<p>Kalibrierung</p> <ul style="list-style-type: none"> ▪ Calibration Plot 	<p>1 P Calibration Plot</p> <p>1 P Hosmer-Lemeshow-Test</p>

		<ul style="list-style-type: none"> ▪ Hosmer-Lemeshow-Test 	
Validierung	Interne Validierung des Scores	Methode der Validierung	1P Splitting der Daten 1P Bootstrapping 1P Kreuzvalidierung 0P keine Validierung
		Gesamtzahl des Patientenkollektivs	3 P > 2500 2,5 P > 2000 und ≤ 2500 2 P > 1500 und ≤ 2000 1,5 P > 1000 und ≤ 1500 1 P > 500 und ≤ 1000 0,5 P > 0 und ≤ 500
Validierung	Externe Validierung des Scores	Anzahl der Validierungen	3 P dreimalige externe Validierung oder mehr 2P zweimalige externe Validierung 1 P einmalige externe Validierung 0P keine externe Validierung
		Lokalisation der Validierungen	3 P externe Validierung auf einem anderen Kontinent wie Modellentwicklung 2 P externe Validierung in einem anderen Land wie Modellentwicklung 1 P externe Validierung im selben Land wie Modellentwicklung 0P keine genaue Ortsangabe

		Gesamtzahl des Patientenkollektivs	3 P > 2500 2,5 P > 2000 und ≤ 2500 2 P > 1500 und ≤ 2000 1,5 P > 1000 und ≤ 1500 1 P > 500 und ≤ 1000 0,5 P > 0 und ≤ 500
Breite Anwendung in der Praxis	Impact Assessment des Scores	Anwendung des Scores zeigt Benefit:	1 P geringere PONV Inzidenz 1 P höhere Patientenzufriedenheit 1 P kürzere KH Verweildauer 1 P weniger Komplikationen
		bisherige Nutzung und Anwendung des Scores:	1 P Identifizierung von Hochrisikopatienten 1 P Vergleich von Studienteilnehmern/ Herstellung von Kohortengleichheit
		Gesamtzahl des Patientenkollektivs	3 P > 2500 2,5 P > 2000 und ≤ 2500 2 P > 1500 und ≤ 2000 1,5 P > 1000 und ≤ 1500 1 P > 500 und ≤ 1000 0,5 P > 0 und ≤ 500
Extra-punkte	Extrapunkte	Modifikation der erreichten Punktzahl:	4 P extra sind möglich, für Aspekte, die in der Tabelle nicht berücksichtigt wurden
Insgesamt			60 Punkte

Tabelle 1 Tabelle der Qualitätskriterien

3 Ergebnisse

3.1 Gesamtauswertung des Scores Junger 2001

3.1.1 Modellentwicklung

3.1.1.1 Junger 2001

Bei dieser Publikation handelt es sich um eine eigenständige Modellentwicklung zu PONV im Aufwachraum. Zusätzlich erfolgt eine interne Validierung durch Splitten der Daten in einen Evaluierungs- & Validierungsdatensatz.

Die Studie findet in der Abteilung für Anästhesiologie und Intensivmedizin an der Justus-Liebig-Universität Gießen statt. Die Patientendaten werden aus einem Anästhesie-Informationsmanagementsystem gewonnen. Als Endpunkt wird PONV im Aufwachraum (PACU) definiert. Die Endpunkterhebung erfolgt durch die Visite des Patienten. Die Einschlusskriterien beinhalten Erwachsene und Kinder. Es ist unklar, ob es sich um ambulante oder stationäre Patienten handelt. Eine prophylaktische präoperative und intraoperative antiemetische Therapie wird nach Einschätzung des Anästhesisten durchgeführt.

Die Stichprobe beinhaltet insgesamt 2154 Patienten, die aufgeteilt in einen Entwicklungsdatensatz mit 854 Patienten sowie einen Validierungsdatensatz mit 1300 Teilnehmern ist. Insgesamt leiden 168 Patienten an PONV, 66 im Entwicklungs- und 102 im Validierungsdatensatz.

Es ist ein Kalibrierungsdiagramm vorhanden und die Geradengleichung lautet: $y=1,0535x-0,014$. Die Diskriminationsfähigkeit für den Validierungssatz beschreibt eine Fläche unter der Kurve von 0,76 bei 1300 Patienten. Zusätzlich werden als Maßzahlen zur Modellklassifizierung für den Validierungsdatensatz die Sensitivität mit 0,513 und die Spezifität mit 0,816 berechnet.

Die endgültige Formel lautet wie folgt:

$$p(\text{PONV}) = 1 / (1 + e^{-z})$$

mit $z = 0,8949 * (\text{weibliches Geschlecht}) - 0,6294 * (\text{Raucher}) - 0,9170 * (\text{i.v. Anästhesie}) + 0,8078 * (\text{N}_2\text{O}) + 1,4301 * (\text{intraoperative Opioide}) - 0,0050 * (\text{Alter (Jahre)}) + 0,0051 * (\text{Operationsdauer (Minuten)}) - 4,9682$ [36]

3.1.2 Interne Validierung mit Update

3.1.2.1 Engel 2006

Dieses Paper handelt von einer internen Validierung mit Update, da dieser Score wie das Originalmodell in Gießen entwickelt wurde. Zusätzlich beschreibt es eine externe Validierung und Anpassung der Koeffizienten von folgenden Scores: original Koivuranta, Apfel und Sinclair. Da die ursprünglichen Risikoscores unzureichende Ergebnisse liefern, erfolgt eine Überarbeitung der Scoringsysteme. Es werden die Risikofaktoren (RF) der jeweiligen Modelle beibehalten, die Koeffizienten sowie der Achsenabschnitt angepasst und erneut validiert. Im Folgenden wird auf das Modell in upgedateter Form eingegangen.

Für die Modelupdates und deren Validierung dient der gleiche Datensatz wie bei der ersten Validierung. Die prognostische Leistung der neu angepassten Modelle wird durch eine Kreuzvalidierung unter Verwendung der Leave-One-Out-Technik getestet.

Der Studienort ist Gießen. Als Endpunkt wird PONV definiert, wenn mindestens eine Episode von Übelkeit und/oder Erbrechen innerhalb der ersten 24 Stunden nach einer Operation auftritt. Dies wird bei der Visite des Patienten festgestellt. Einschlusskriterien sind erwachsene, stationäre Patienten, die sich HNO-OPs unterziehen.

Ausgeschlossen werden Patienten mit einem Gewicht >150kg bzw. einem BMI > 40kg/m².

Die Stichprobengröße des Validierungsdatensatzes beträgt 748 Patienten. Hiervon erleiden 84 PONV. Somit ergibt sich bei sieben Risikofaktoren ein Ereignis pro Variablen Quotient von 12.

Für die Kalibrierung wird der Hosmer-lesmeshow-Test durchgeführt. Bezüglich der Diskriminationsfähigkeit ergibt sich für das upgedatete Modell eine AUC von 0,729 im Validierungsdatensatz.

Bei identischer Gleichung lauten die Werte nach dem Update wie folgt:

$$p(\text{PONV}) = 1 / (1 + e^{-z})$$

mit $z = 0,812^* (\text{weibliches Geschlecht}) - 0,714^* (\text{Raucher}) - 0,577^* (\text{i.v. Anästhesie}) + 0,167^* (\text{N}_2\text{O}) + 0,746^* (\text{intraoperative Opioide}) - 0,023^* (\text{Alter (Jahre)}) + 0,001^* (\text{OP-Dauer (Minuten)}) - 1,682$ [26]

3.1.3 Qualitätskriterien des Scores Junger 2001

Kriterium	Untergruppe	Unterpunkt	Max. erreichte Punktzahl
Generalisierbarkeit	Anzahl des Patientenkollektivs	Scoreentwicklung & -update: 1602 <ul style="list-style-type: none"> ▪ Entwicklung: 854 ▪ Scoreupdate: 748 	2
	Praktikabilität	Risiko-faktoren in dem Modell	Anzahl der RF in dem Modell: 7
Mehrheit der RF ohne technische Hilfsmittel mit ja/nein Antwort, mittlerer Zeitaufwand			2
Mehrheit der RF exakt definiert			2
RF ohne Kenntnis des Outcomes beurteilt: nein/unklar			0
Risiko-berechnung/ Handhabung des Scores		Hilfsmittel nötig: spezielles Tabellenwerk/Schaubild	1
		kein vereinfachter Score	0
Endpunkt des Scores		zusammengesetzter Endpunkt PONV	2
		Art und Weise der Endpunkterhebung: via Gespräch (persönliche Visite/ Telefonat mit dem Patienten)	1
Präzision	Vorhersagegenauigkeit/ Beurteilung der Prädiktion mittels:	Diskriminationsfähigkeit (AUC der ROC-Kurve) <ul style="list-style-type: none"> ▪ 0,749 (Intern) 	2,5
		Kalibrierung <ul style="list-style-type: none"> ▪ Geradengleichung Fläche: 0,264 	3
		Kalibrierung <ul style="list-style-type: none"> ▪ Calibration Plot ▪ Hosmer-Lemeshow-Test 	1 1

Validierung	Interne Validierung des Scores	Methode der Validierung <ul style="list-style-type: none"> ▪ Splitting der Daten ▪ Kreuzvalidierung 	1
		Gesamtzahl des Patientenkollektivs: <ul style="list-style-type: none"> ▪ 2048 	1
	Externe Validierung des Scores	Anzahl der Validierungen: 0	0
		Lokalisation der Validierungen: keine genaue Ortsangabe	0
		Gesamtzahl des Patientenkollektivs: 0	0
	Breite Anwendung in der Praxis & Extrapunkte	Impact Assessment des Scores	Anwendung des Scores zeigt Benefit: keine
bisherige Nutzung und Anwendung des Scores: keine			0
Gesamtzahl des Patientenkollektivs: 0			0
Extrapunkte		Modifikation der erreichten Punktzahl: Maßzahlen zur Modellklassifizierung	1
Insgesamt			24 Punkte

Tabelle 2 Qualitätskriterien des Scores Junger 2001

3.2 Gesamtauswertung des ScoresPalazzo &Evans 1993

3.2.1 Modellentwicklung

3.2.1.1Palazzo1993

Der Score wurde von Palazzo und Evans in Großbritannien im Charing Cross Hospital in London und in der Abteilung für Anästhesie in Oxford im Jahre 1993 entwickelt.

Als Endpunkt wird PONV im Zeitraum von 24 Stunden nach einer Operation festgelegt. PONV wird definiert als Erbrechen entsprechend einem Reflex, der zum Ausstoßen des Mageninhalts führt. Zusätzlich beinhaltet PONV Würgen, d.h. ein nicht-produktives Erbrechen, oder Übelkeit, welche gleichzusetzen ist mit dem Gefühl, erbrechen zu wollen. In der Datenauswertung wird jeder Patient gewertet, der mindestens eine Folge von Übelkeit, Würgen oder Erbrechen hat. Dies wird durch direkte Befragung oder freiwillige Rückmeldungen der Patienten dokumentiert. Einschlusskriterien für die Studie sind Erwachsene, der ASA Klasse 1, die kleine orthopädische Operationen unter stationären Bedingungen erhalten. Ausschlusskriterien sind schwangere oder stillende Frauen. Zusätzlich werden Patienten ausgeschlossen, die während der Narkose abgesaugt werden, einen oropharyngealen Atemweg eingeführt haben oder einen zusätzlichen Eingriff in die Atemwege im Aufwachraum innerhalb von 24 Stunden erhalten.

Die Stichprobengröße beträgt 147 Patienten mit 40 PONV-Ereignissen. Der Score hat fünf Koeffizienten, wobei einer die Kombination aus zwei Risikofaktoren ist.

Die abschließend entwickelte Gleichung lautet:

$$p(\text{PONV}) = e^z / (1 - e^z)$$

mit $z = -5,03 + 2,24^* (\text{postoperative Opiode}) + 3,97^* (\text{PONV in der Vorgeschichte}) + 2,4^* (\text{weibliches Geschlecht}) + 0,78^* (\text{Reisekrankheit}) - 3,2^* (\text{weibliches Geschlecht UND PONV in der Vorgeschichte})$ [60]

3.2.2 Externe Validierung ohne Update

3.2.2.1 Toner 1996

Dies ist die Erste von insgesamt fünf externen Validierungen des Scores von Palazzo und Evans.

Der genaue Standort der Validierung wird nicht erwähnt, jedoch weisen die Kontaktdaten der Autoren auf das Krankenhaus des Medical College in London sowie auf das Royal London Hospital Whitechapel in London hin. Als Endpunkt dient PONV 24 Stunden nach der Operation. Dabei wird Erbrechen als ein Reflex definiert, der zur

Austreibung des Mageninhaltes führt. Würgen ist ein nicht-produktives Erbrechen und Übelkeit wird als Gefühl, sich übergeben zu wollen, beschrieben. Die Endpunkterhebung erfolgt durch die Visite des Patienten. Einschlusskriterien für die Studie sind Erwachsene sowie Kinder ab 17 Jahren der ASA-Klasse 1 bis 3. Diverse elektive OP-Verfahren (HNO-, Mund-, Kiefer- und Gesichtschirurgie sowie orthopädische-, gynäkologische-, allgemein- und urologische Chirurgie) sind berücksichtigt. Das Patientenkollektiv ist mit 400 Patienten größer als die Initialstudie. Der Score hat fünf Koeffizienten, wobei einer die Kombination aus zwei Risikofaktoren ist.

Ein Kalibrierungsdiagramm ist vorhanden. Es zeigt, dass die Gesamtinzidenz von PONV höher war als die vorhergesagte PONV-Rate. Somit unterschätzt das Modell das Patientenrisiko. Zusätzlich werden folgende Maßzahlen zur Modellklassifizierung genannt: Sensitivität: 41%, Spezifität 88%, positiver Prädiktiver Wert 65% und negativer Prädiktiver Wert 73%. Dies bedeutet, dass das Modell relativ schlecht die Risikopatienten erkennt und die detektierten Patienten ein hohes Risiko haben.[84]

3.2.2.2 Eberhart 2000a

Dies ist die zweite externe Validierung. In dieser Publikation werden drei verschiedene Scores (Palazzo und Evans von 1993, original Apfel-Score aus dem Jahre 1998 und Koivuranta von 1997) zur Vorhersage von postoperativem Erbrechen oder postoperativer Übelkeit und Erbrechen überprüft.

Die Studienteilnehmer werden in der Abteilung für Anästhesiologie am Universitätsklinikum in Ulm rekrutiert. Mittels Visite des Patienten 24 Stunden nach der Operation werden die Endpunkte PV und PONV getrennt erfasst. Einschlusskriterien sind erwachsene Patienten, die sich unter Vollnarkose einer Vielzahl von stationären, elektiven, chirurgischen Eingriffen unterziehen. Es erfolgt keine antiemetische Prophylaxe. Die Verletzung des Studienprotokolls (z.B. Aufrechterhaltung der Anästhesie mit Propofol oder Applikation von perioperativen Steroiden) oder unvollständige Daten sind Ausschlusskriterien.

Das Studienkollektiv enthält 1444 Patienten. Davon leiden 247 Patienten an PV und 540 Patienten an PONV.

Die Kalibrierungsgeradengleichung zum vorhergesagten Risiko für PONV lautet: $Y = 0,64x + 22,7$. Die Diskriminationsfähigkeit des Scores wird mit einer AUC von 0,68 beschrieben.[23]

3.2.2.3 Apfel 2002

Die dritte externe Validierung ist in einer Publikation, die sechs verschiedene Scores (Apfel 1999, Palazzo 1993, Koivuranta 1997, Sinclair 1999, Scholz 1999 und Gan 2000) vergleicht.

Das Patientenkollektiv entspricht größtenteils der Kohorte aus dem Uniklinikum Ulm, das bereits in der Publikation Eberhart 2000a angeführt wurde. Als Endpunkt gilt PONV, wenn mindestens einmal im Zeitraum von 24 Stunden nach der OP Übelkeit und/oder eine emetische Episode auftritt. Erbrechen oder Würgen wird als emetische Episode angesehen. Die Einschlusskriterien beinhalten Erwachsene, stationäre Patienten, die sich diversen chirurgischen Eingriffen (orthopädische Verfahren, ophthalmologische-, HNO-, gastrointestinale- sowie gynäkologische Operationen) unterziehen. Es erfolgt keine antiemetische Prophylaxe.

Die Stichprobengröße umfasst 1566 Patienten mit 600 dokumentierten Events.

Ein Kalibrierungsdiagramm ist vorhanden. Die Geradengleichung lautet: $y = 0,30x + 0,30$. Für die Diskriminationsfähigkeit wird eine AUC von 0,64 angegeben. [8]

3.2.2.4 Eberhart 2004

Es handelt sich um eine externe Validierung von fünf Scores (Palazzo 1993, Koivuranta 1997, Apfel 1998, Sinclair 1999, Apfel 1999) an pädiatrischen Patienten.

Die Rekrutierung der Studienteilnehmer erfolgt an drei Zentren: der Klinik für Anästhesiologie und Intensivmedizin der Philipps-Universität Marburg, dem Olga-Kinderkrankenhaus Stuttgart und dem ambulanten Chirurgischen Center Ulm-Söflingen. PV 24 Stunden nach der OP wird als Hauptendpunkt der Befragung gewählt, da Übelkeit ein subjektives Phänomen ist, das Kleinkinder nicht beschreiben können. Bei stationären Patienten wird postoperatives Erbrechen durch Befragung erhoben. Mit einem strukturierten Interview via Telefon wird dieser Endpunkt bei ambulanten Patienten gewonnen. Eingeschlossen werden Kinder (0-12 Jahre) mit ambulanten oder stationären Eingriffen (HNO-, Ophthalmische-, Urologische-, Bauch-, Knochen- und Gelenkchirurgie, zahnärztliche Verfahren, diagnostische OPs, plastische Eingriffe, sowie Katheterimplantationen). Ausschlusskriterien sind eine i.v. Anästhesie mit Propofol, prophylaktische Antiemetika (einschließlich Kortikosteroiden) oder Patienten mit unvollständigen Daten für das Follow-up.

Die Stichprobengröße umfasst 983 Kinder mit 326 Events.

Die Geradengleichung der Kalibrierung wird beschrieben mit: $y=0,55x+30$. Als Diskriminationsfähigkeit ergibt sich eine AUC von 0,56. Dies ist nur gering besser als eine Zufallsvorhersage mit einer Vergleichs-AUC von 0,5.

Als Schlussfolgerung zeigt sich, dass die bestehenden Scores nicht auf Kinderübertragbar sind und ein Extrascore für Minderjährige zu entwickeln ist.[24]

3.2.2.5Wu2015

Die letzte externe Validierung erfolgt in Taiwan. Es werden fünf Scoringsysteme (vereinfachter Apfel, Koivuranta, vereinfachter Koivuranta, Palazzo & Evans sowie original Apfelscore) verglichen. Des Weiteren werden zwei neue Modellentwicklungen vorgestellt, die sich jedoch nicht durchsetzten.

Studienort ist die Abteilung für Anästhesie des Chia-Yi Christian Hospital in Taiwan. Die Häufigkeit von Übelkeit oder Erbrechen in den ersten 24 Stunden nach der Operation wird durch Befragung der Patienten erfasst. Jede Episode von Übelkeit, Würgen oder Erbrechen wird als PONV registriert. Die Einschlusskriterien erfassen stationäre Erwachsene (> 20 Jahre), die gut kommunizieren können und in Vollnarkose Eingriffe (Schilddrüsen-, Brustkrebs-, HNO- und Wirbelsäulenoperationen sowie intraabdominelle oder intrathorakale Eingriffe) erhalten. Als Ausschlusskriterien zählen Notfalloperationen und Patienten mit Sprachbarriere. Es wird keine antiemetische Prophylaxe verabreicht.

Die Stichprobengröße beträgt 992 Patienten mit 416 Ereignissen.

Für die Kalibrierung wird der Hosmer-Lemeshow-Test durchgeführt. Die Diskriminationsfähigkeit beschreibt eine AUC von 0,659. Zusätzlich werden folgende Maßzahlen zur Modellklassifizierung genannt: Sensitivität: 80,6%, Spezifität: 45,4%, positiver prädiktiver Wert: 51,6%, negativer Prädiktiver Wert 76,4%, und Youden's Index: 26%. [97]

3.2.3 Impact Assessment

3.2.3.1 Thomas 2002

Diese Publikation fällt unter die Rubrik Impact Assessment. Während einer randomisierten, kontrollierten Studie über die Auswirkungen von kombinierten Antiemetika werden Patientengruppen prospektiv durch Scoring-Systeme für PONV und POV verglichen. Für jeden Patienten wird mit vier Scoring-Systemen (vereinfachter Koivuranta, Palazzo&Evans, Sinclair, vereinfachter Apfel) das prädiktive Risiko von PONV berechnet. Zusätzlich wird mit zwei Scoring-Systemen (Koivuranta, nicht vereinfachter Apfel) auf POV untersucht. Im Anschluss wird nach der Methode von Bland und Altman bewertet, ob zwischen den Scoring-Systemen eine Übereinstimmung besteht. Die Kongruenz zwischen den Bewertungssystemen ist unzureichend.

Der Studienort ist das North Hampshire Hospital in Großbritannien. Als Endpunkte dienen sowohl PONV als auch POV. Die Patientinnen werden 24 Stunden nach der OP angerufen, um zu erfragen, ob sie die „Notfallmedikation“ bei Beschwerden eingenommen haben. Es werden Erwachsene, der ASA-Klasse I oder II, im Alter von 19-53 Jahren, die sich einer ambulanten, gynäkologischen Operation unterziehen eingeschlossen.

Die Stichprobengröße beträgt 177 Patientinnen.

Aus der Studie geht hervor, dass die fehlende Übereinstimmung der Scores nicht bedeutet, dass alle Systeme unbefriedigend für die Vorhersage von PONV oder POV sind. Die Autoren der Studie berichten, die Hauptaufgabe der Scoring-Systeme liege im klinischen Bereich, für den sie entwickelt wurden. Für den Vergleich der verschiedenen Arbeitsgruppen in dieser Publikation werden keine Scoring-Systeme in den demographischen Tabellen verwendet, da keine ausreichende Genauigkeit für diese erweiterte Anwendung der Risikosysteme garantiert werden kann. [83]

3.2.4 Qualitätskriterien des Scores Palazzo & Evans 1993

Kriterium	Untergruppe	Unterpunkt	Max. erreichte Punktzahl
Generalisierbarkeit	Anzahl des Patientenkollektivs	Scoreentwicklung: 147	0,5
Praktikabilität	Risiko-faktoren in dem Modell	Anzahl der RF in dem Modell: 5	2
		alle ohne technische Hilfsmittel, mit ja/nein Antwort, geringer Zeitaufwand	3
		Mehrheit der RF exakt definiert	2
		RF ohne Kenntnis des Outcomes beurteilt: nein/ unklar	0
	Risiko-berechnung/ Handhabung des Scores	Hilfsmittel nötig: spezielles Tabellenwerk/Schaubild	1
		Kein vereinfachter Score	0
	Endpunkt des Scores	zusammengesetzter Endpunkt PONV	2
		Art und Weise der Endpunkterhebung: unklarer Modus	0
Präzision	Vorhersagegenauigkeit/ Beurteilung der Prädiktion mittels:	Diskriminationsfähigkeit (AUC der ROC-Kurve) <ul style="list-style-type: none"> ▪ Intern: keine ▪ Extern: 0,704 (AUC) 	2,5
		Kalibrierung <ul style="list-style-type: none"> ▪ Geradengleichung Fläche: 17,49 	0
		Kalibrierung <ul style="list-style-type: none"> ▪ Calibration Plot ▪ Hosmer-Lemeshow-Test 	1 1

Validierung	Interne Validierung des Scores	Methode der Validierung: keine	0
		Gesamtzahl des Patientenkollektivs: 0	0
	Externe Validierung des Scores	Anzahl der Validierungen: 5	3
		Lokalisation der Validierungen: auf einem anderen Kontinent wie Modellentwicklung	3
		Gesamtzahl des Patientenkollektivs: 5385	3
	Breite Anwendung in der Praxis & Extrapunkte	Impact Assessment des Scores	Anwendung des Scores zeigt einen Benefit: keine
bisherige Nutzung und Anwendung des Scores: Vergleich von Studienteilnehmern/ Herstellung von Kohortengleichheit			1
Gesamtzahl des Patientenkollektivs: 177			0,5
Extrapunkte		Modifikation der erreichten Punktzahl:	1
		<ul style="list-style-type: none"> ▪ Maßzahlen zur Modellklassifizierung ▪ Vergleich der Scoringsysteme mit der Methode von Bland und Altman 	1
Insgesamt			27,5 Punkte

Tabelle 3: Qualitätskriterien des Scores Palazzo & Evans 1993

3.3 Gesamtauswertung des Scores Sinclair1999

3.3.1 Modellentwicklung mit interner Validierung

3.3.1.1 Sinclair1999

In dieser Publikation wird eine eigenständige neue Modelentwicklung mit interner Validierung dargestellt. Die Studie findet im Toronto Hospitalin Kanada statt. Eine Patientenpopulation wird in zwei Hälften, in einen Modellentwicklungssatz und einen Modellvalidierungssatz, aufgeteilt. Nachdem das endgültige Modell entwickelt wurde, wird für jeden Patienten die PONV-Wahrscheinlichkeit im Validierungsset berechnet.

PONV wird im Aufwachraum, in der ambulanten ChirurgieAbteilung und 24 Stunden nach der OP mit einem standardisierten Check-off-Formular erfasst. Dokumentiert wird ein freiwilliger Bericht über Übelkeit, ein beobachtetes aktives Würgen oder Erbrechen, das Antiemetika erfordert. Zusätzlich erfolgt ein Telefonanruf 24 Stunden nach der OP unter Verwendung eines standardisierten Fragebogens.

Als prophylaktische antiemetische Behandlung erhalten alle Patienten Dimenhydrinat 25-50 mg intravenös. Eingeschlossen werden über einen Zeitraum von 3 Jahren erwachsene Patienten, die ambulant einen Eingriff aus einer der folgenden Disziplinen erhalten: Orthopädie, Urologie, Allgemeinchirurgie, HNO, Gynäkologie, Ophtalmologie, Zahnchirurgie, plastische Chirurgie oder Neurochirurgie. Ausgeschlossen sind Patienten, die sich weigern, vor der OP dem Telefoninterview zuzustimmen, kein Englisch sprechen oder beim zweiten Versuch nicht erreicht werden.

Die Stichprobengröße beträgt insgesamt 17638 Patienten. Diese werden gesplittet in 8819 Patienten im Entwicklungsdatensatz und 8819 im Validierungsdatensatz. 24 Stunden nach der OP haben insgesamt 820 Patienten (397 im Entwicklungs-, 423 im Validierungsdatensatz) PONV erlitten.

Ein Kalibrierungsdiagramm sowie ein Hosmer-Lemeshow-Test werden durchgeführt. Die Diskriminationsfähigkeit des Modells zeigt eine AUC von 0,785 für den Validierungsdatensatz.

Die Modellprüfung mit einem Trainingsdatensatz, entsprechend einer internen Validierung, erfolgt durch Splitting des Datensatzes.

Die Gleichung lautet:

$$p(\text{PONV}) = e^z / (1 - e^{-z})$$

mit $z = -5,97 - 0,14 * (\text{Alter in Jahrzehnten}) - 1,03 * (\text{männliches Geschlecht}) - 0,42 * (\text{Raucher}) + 1,14 * (\text{PONV Anamnese}) + 0,46 * (\text{OP- Dauer in 30 Minuten Schritten}) + 2,36 * (\text{Allgemeinanästhesie}) + 1,48 * (\text{HNO-OP}) + 1,77 * (\text{Augen-OP}) + 1,90 * (\text{Plastische-OP}) + 1,20 * (\text{Gynäkologische-OP ohne Dilatation \& Kürettage}) + 1,04 * (\text{Knie-OP}) + 1,78 * (\text{Schulter-OP}) + 0,94 * (\text{andere orthopädische Eingriffe})$ [80]

3.3.2 Externe Validierung mit Update

3.3.2.1 Engel 2006

Auf diese Publikation wird bereits unter Punkt 3.1.2.1 eingegangen. Es handelt sich um externe Validierungen der Originalscores von Koivuranta, Apfel und Sinclair sowie eine interne Validierung von Junger.

PONV 24 Stunden nach der Operation wird als Endpunkt definiert. Bei dem Update werden die RF der jeweiligen Modelle beibehalten, die Koeffizienten plus der Achsenabschnitt angepasst und erneut mit dem gleichen Datensatz durch Kreuzvalidierung getestet. Es wird nun auf das überarbeitete Modell eingegangen.

Die Stichprobengröße des Validierungsdatensatzes beträgt 748 Patienten. Hiervon erleiden 84 PONV.

Für die Modellkalibrierung wird der Hosmer-Lemeshow-Test durchgeführt. Die Diskriminationsfähigkeit des neu aktualisierten Sinclairscores ergibt eine AUC von 0,776.

Die Werte des überarbeiteten Modells lauten wie folgt:

$$p(\text{PONV}) = e^z / (1 - e^{-z})$$

mit $z = -2,11 - 0,29 * (\text{Alter in Jahrzehnten}) - 0,76 * (\text{männliches Geschlecht}) - 0,78 * (\text{Raucher}) + 1,88 * (\text{PONV Anamnese}) + 0,22 * (\text{OP- Dauer in 30 Minuten Schritten}) + 0,34 * (\text{Allgemeinanästhesie})$ [26]

3.3.2.2 Sarin 2012

Hier handelt es sich um die Entwicklung eines eigenständigen Scoringsystems mit interner Validierung, welches als experimentelles Modell bezeichnet wird. Dieses Modell wird nicht weiterverfolgt, bzw. von anderen Autoren aufgegriffen. Des Weiteren erfolgt eine externe Validierung folgender Scores: original Apfelmodel und Update dieses Scores zum refitted Apfelmodel, sowie Validierung des simplified Apfelriscores. Zusätzlich findet eine Validierung des refitted Sinclairmodel statt. Auf dieses Modell ist nun näher einzugehen. Die neuen β -Koeffizienten des refitted Sinclairmodels sowie des refitted Apfelmodels werden unter Verwendung der Variablen aus dem ursprünglichen Modell mithilfe einer logistischen Regression berechnet.

Der Studienstandort ist das Brigham and Women's Hospital der Harvard Medical School in Boston in der USA.

Als Endpunkt wird PONV definiert, wobei das Ende des Beobachtungszeitraumes nicht genau festgelegt ist. Einschlusskriterien erfüllen erwachsene, ambulante Patienten, die sich einer der folgenden Operationen unterziehen: Brustbiopsie, Lumpektomie, Brustplastik, Hysteroskopie, Dilatation, Kürettage, Pelviskopie, Tubenligatur, laparoskopische Cholezystektomie und Herniorrhaphie. Ausgeschlossen werden Patienten mit ungültigen oder fehlenden Daten. Dem Anästhesisten ist es freigestellt, prophylaktische Antiemetika zu geben. Postoperativ gab es Standardvorlagen für die Verabreichung von Antiemetika und Analgetika.

Die Stichprobengröße beträgt insgesamt 2505 Patienten. Diese sind aufgeteilt in den Entwicklungsdatensatz mit 1670 Patienten sowie den Validierungsdatensatz, auch Testsatz genannt, mit 835 Patienten. Die Anzahl an Events (PONV) wird nicht berichtet.

Bezüglich der Kalibrierung wird der Hosmer-Lemeshow-Test für alle Scores durchgeführt. Zusätzlich liegt eine Grafik zur Beschreibung der Kalibrierung vor, dort kann die Steigung abgelesen werden. Jedoch wird die Geradengleichung selbst nicht veröffentlicht. Die AUC für den Validierungsdatensatz mit 835 Patienten beträgt 0,711 für das refitted Sinclairmodel.

Fazit der Autoren ist, dass das refitted Sinclairmodel eine gute Diskriminierung und Kalibrierung für die Daten des Krankenhauses hat. Dieses kann durch die Tatsache erklärt werden, dass sowohl die Entwicklung als auch eine Modellvalidierung an einem ambulanten Patientenkollektiv stattgefunden hat.

Die angepassten Werte für das refitted Sinclairmodell lauten wie folgt:

$$p(\text{PONV}) = e^z / (1 + e^{-z})$$

mit $z = -2,41 - 0,21 \cdot (\text{Alter in Jahrzehnten}) - 0,70 \cdot (\text{männliches Geschlecht}) - 0,12 \cdot (\text{Raucher}) + 0,98 \cdot (\text{PONV Anamnese}) + 0,30 \cdot (\text{OP-Dauer in 30 Minuten Schritten}) + 1,22 \cdot (\text{Allgemeinanästhesie}) - 0,20 \cdot (\text{Brust-OP}) + 0,07 \cdot (\text{Gynäkologische-OP ohne Dilatation \& Kürettage})$ [75]

3.3.3 Externe Validierung ohne Update

3.3.3.1 Apfel2002

Dieses Paper wurde bereits unter 3.2.2.3 ausführlich beschrieben. Es handelt sich um eine Publikation, die sechs verschiedene Scores (Apfel 1999, Palazzo 1993, Koivuranta 1997, Sinclair 1999, Scholz 1999 und Gan 2000) in einer externen Validierung vergleicht.

Aus der Publikation geht hervor, dass von den 1566 Patienten 600 an PONV leiden. Ein Kalibrierungsdiagramm ist vorhanden. Die Geradengleichung lautet: $y = 0,49x + 0,29$. Für die Diskriminationsfähigkeit wird eine AUC von 0,68 berichtet.[8]

3.3.3.2 Pierre2002

Diese Publikation schildert eine externe Validierung des vereinfachten Apfel-Scores und des Sinclair-Scores.

Die Studie findet in einem Krankenhaus mit dem Schwerpunkt Krebsleiden in Toulouse in Frankreich statt. Als Endpunkt wird PONV klassifiziert, wenn Übelkeit und/oder Erbrechen innerhalb der ersten 24 postoperativen Stunden auftreten. Für die Endpunkterhebung werden die Formulare und das Anästhesieprotokoll überprüft. Zusätzlich visitiert der Autor den Patienten und das betreuende Pflegepersonal. Einschlusskriterien sind stationäre, erwachsene Patienten der ASA Klasse I bis III. Folgende Operationen werden durchgeführt: Hals-, Schilddrüsen-, Brust- oder gynäkologische Eingriffe unter allgemeiner Inhalationsanästhesie. Ausgeschlossen werden Patienten mit unvollständiger, intraoperativer Dokumentation. Patienten, die Antihistaminika erhalten werden nicht berücksichtigt, da eine prophylaktische Antiemetikagabe nicht gestattet ist.

Die Stichprobengröße des Validierungssatzes beträgt 428 Patienten. Hiervon erleiden 212 Patienten PONV.

Für die Kalibrierung ist ein Kalibrierungsplot sowie die Geradengleichung vorhanden. Sie lautet: $y = 0,93x + 0,27$. Bezüglich der Diskriminationsfähigkeit beträgt die AUC 0,64.[64]

3.3.3.3 Eberhart 2004

Dieser Score wurde bereits unter 3.2.2.4 ausführlich beschrieben. Es handelt sich um eine externe Validierung von fünf Scores (Palazzo1993, Koivuranta1997, Apfel1998, Sinclair 1999, Apfel 1999) an pädiatrischen Patienten. Als Endpunkt dient postoperatives Erbrechen innerhalb der ersten 24 Stunden nach der OP, da Übelkeit bei Kleinkindern nicht erfasst werden kann. Abschließend wird festgestellt, dass der bestehende Score nicht auf Kinder übertragbar ist.

Bezüglich des Scores von Sinclair werden folgende Zusatzfakten veröffentlicht. Der Score besitzt 7 Risikofaktoren und es erleiden 326 KinderPONV.

Die Geradengleichung für die Kalibrierung lautet $y = 0,64x + 13$. Kalibrierungskurven liegen vor. Abschließend wird die Diskriminationsfähigkeit mit einer AUC von 0,65 beschrieben.[24]

3.3.4 Impact Assessment

3.3.4.1 Thomas2002

Diese Publikation wurde bereits unter Punkt 3.2.3.1 ausführlich beschrieben und fällt unter die Rubrik Impact Assessment. Während einer randomisierten, kontrollierten Studie zu den Effekten von kombinierten Antiemetika werden Patientengruppen durch Scoring-Systeme für PONV und POV verglichen. Es wird das vorhergesagte Risiko für PONV mit folgenden Scores berechnet: vereinfachter Koivuranta, Palazzo und Evans, Sinclair, vereinfachter Apfel. Zusätzlich wird das prädiktive Risiko für POV mit den Scores von Koivuranta und dem original Apfelscore berechnet. Im Anschluss wird die Übereinstimmungsquote zwischen den Scoringsystemen bewertet. Die Übereinstimmung zwischen den einzelnen Scores ist unzureichend. Das Kollektiv dieser Studie sind, wie in dem Originalscore von Sinclair, ambulante Patienten. Jedoch hat Thomas2002 ein rein weibliches Patientenkollektiv. Dies könnte zu einer Verzerrung geführt haben.[83]

3.3.5 Qualitätskriterien des Scores Sinclair1999

Kriterium	Untergruppe	Unterpunkt	Max. erreichte Punktzahl
Generalisierbarkeit	Anzahl des Patientenkollektivs	Scoreentwicklung & -update: 9567 <ul style="list-style-type: none"> Entwicklung: 8819 Scoreupdate: 748 	3
		Risiko-faktoren in dem Modell	Anzahl der RF in dem Modell: 7
Praktikabilität	Risiko-berechnung/ Handhabung des Scores	Mehrheit der RF ohne technische Hilfsmittel mit ja/nein Antwort, mittlerer Zeitaufwand	2
		Minderheit der RF exakt definiert	1
		RF ohne Kenntnis des Outcomes beurteilt: ja	1
		Hilfsmittel nötig: spezielles Tabellenwerk/Schaubild	1
	Endpunkt des Scores	kein vereinfachter Score	0
		zusammengesetzter Endpunkt PONV	2
	Endpunkt des Scores	Art und Weise der Endpunkterhebung: via Gespräch (persönliche Visite/ Telefonat mit dem Patienten)	1
		Vorhersagegenauigkeit/ Beurteilung der Prädiktion mittels:	Diskriminationsfähigkeit (AUC der ROC-Kurve) <ul style="list-style-type: none"> Endwert: (Intern & Extern) AUC 0,777
Kalibrierung <ul style="list-style-type: none"> Geradengleichung Fläche: 4,4 	1		
Kalibrierung <ul style="list-style-type: none"> Calibration Plot Hosmer-Lemeshow-Test 	1 1		

Validierung	Interne Validierung des Scores	Methode der Validierung: Splitting der Daten	1
		Gesamtzahl des Patientenkollektivs: 8819	3
	Externe Validierung des Scores	Anzahl der Validierungen: 5 <ul style="list-style-type: none"> ▪ 2x mit Update 	3
		Lokalisation der Validierungen: externe Validierung auf einem anderen Kontinent wie Modellentwicklung	3
		Gesamtzahl des Patientenkollektivs: 4560	3
Breite Anwendung in der Praxis & Extrapunkte	Impact Assessment des Scores	Anwendung des Scores zeigt Benefit: keine	0
		bisherige Nutzung und Anwendung des Scores: Vergleich von Studienteilnehmern	1
		Gesamtzahl des Patientenkollektivs: 177	0,5
	Extrapunkte	Modifikation der erreichten Punktzahl: keine	0
Insgesamt			32,5 Punkte

Tabelle 4: Qualitätskriterien des Scores Sinclair1999

3.4 Gesamtauswertung des Scores Apfel 2012 PDNV

3.4.1 Modellentwicklung mit interner Validierung

3.4.1.1 Apfel 2012

Bei dieser Publikation handelt es sich um die Entwicklung und interne Validierung eines vereinfachten Scores bei Patienten mit ambulanten Operationen zu Post discharge Nausea and vomiting (PDNV). In das Deutsche übersetzt, ein Score zu Übelkeit und Erbrechen nach Entlassung aus dem Krankenhaus. Ziel dieser Arbeit ist die Bestimmung des Anteiles der Patienten mit Übelkeit, Erbrechen und/oder Würgen nach Entlassung aus dem Krankenhaus bis 48 Stunden nach einer Narkose.

Die Studienorte sind die folgenden zwölf Zentren für ambulante Chirurgie in den Vereinigten Staaten: Medical College of Georgia in Augusta, Brigham & Womens Hospital in Boston, Universität von Virginia in Charlottesville, St. Luke's Episcopal Hospital in Houston; Anderson Center in Houston; University of Kansas Medical Center in Kansas City; Universität von Kentucky in Lexington; Loma Linda Universität in Loma Linda; Universität Louisville in Louisville; San Francisco Zentrum für ambulante Chirurgie in San Francisco, Medical Center in Mt. Zion in San Francisco, Mayo Clinic Scottsdale in Scottsdale.

Als Endpunkt wird PDNV 48 Stunden nach einer OP festgelegt. Für die Endpunkterhebung schreiben die Patienten Beschwerden in ein standardisiertes Tagebuch auf. Dies wird 24 und 48 Stunden nach der OP durch einen Anruf abgefragt. Die Einschlusskriterien erfüllen Erwachsene mit ambulanten OPs unter Vollnarkose. Ausschlusskriterien sind die stationäre Aufnahme in das Krankenhaus der ehemals ambulanten Patienten oder, wenn ein Patient wegen Komplikationen über Nacht intubiert blieb. Ebenfalls nicht berücksichtigt werden Patienten, die für die postoperative Nachsorge nicht erreicht werden. Bezüglich der antiemetischen Behandlung wird die Standardbehandlung des jeweiligen Zentrums und ggf. Prophylaxe oder Rescue-Medikation beibehalten.

Die Stichprobengröße beträgt insgesamt 2170 Patienten. Diese werden für eine interne Validierung in einen Entwicklungsdatensatz mit 1913 Patienten und einen Validierungsdatensatz mit 257 Teilnehmern gesplittet. PDNV erlitten insgesamt 803 Personen.

Es ist ein Kalibrierungsdiagramm vorhanden. Die Geradengleichungen lauten:
Entwicklungsdatensatz (n=1913): $y = 0,942x + 0,006$
Validierungsdatensatz (n=257): $y = 1,075x - 0,044$.

Die Diskriminationsfähigkeit beschreibt für den Entwicklungsdatensatz eine AUC von 0,706 und für den Validierungsdatensatz eine Fläche unter der Kurve von 0,721.

Das finale, vereinfachte Modell besitzt fünf gleichgewichtete Risikofaktoren: weibliches Geschlecht, Alter unter 50 Jahren, Vorgeschichte von Übelkeit und/oder Erbrechen bei früherer Anästhesie, Übelkeit in der Postanästhesie-Behandlungseinheit, Opioid-Verabreichung in der Postanästhesie-Station. Die verabreichten Opiode werden in Morphinäquivalente umgerechnet.

Das Risiko für PDNV beim Vorhandensein von null, eins, zwei, drei, vier oder fünf dieser Faktoren ist in Tabelle 4 dargestellt.[11]

Anzahl Risikofaktoren	PDNV-Risiko (%)
0	7 %
1	20 %
2	28 %
3	53 %
4	60 %
5	89 %

Tabelle 5: Wahrscheinlichkeit für das Auftreten von PDNV: Apfelscore 2012[11]

3.4.2 Interne Validierung mit Update

3.4.2.1 Odom Forren 2013

Bei dieser Publikation handelt es sich um eine interne Validierung mit Update des PDNV-Scores von Apfel aus dem Jahre 2012. Eine Subgruppe des initialen Patientenkollektivs wird statt zwei, sieben Tage nach der OP verfolgt, um PDNV zu beobachten. Abschließend wird publiziert, dass der PDNV-Score von Apfel auch auf den Zeitraum bis 7 Tage nach einer OP angewendet werden kann. Somit handelt es sich um ein Update, bei dem der Endpunkt erweitert wird. Zusätzlich wird probeweise ein Score entwickelt, der den Zeitraum drei bis sieben Tage nach einer OP abdeckt. Wichtigster Faktor ist in diesem Modell der Schmerz. Dieses Risikosystem wird jedoch nicht weiterverfolgt.

Die Studie findet an zwei Zentren, der Universität von Kentucky und dem Medical College von Georgia, statt. Als Endpunkt wird PDNV festgelegt. Übelkeit wird auf

einer numerischen Bewertungsskala gemessen. Erbrechen wird als die Anzahl der Episoden erfasst, die mindestens eine Minute auseinander liegen. Da eine geringere Inzidenz von Würgen auftritt, wird dies mit Erbrechen kombiniert. Die Endpunkterhebung erfolgt an den Tagen eins, zwei und sieben aus der Krankenakte sowie aus dem Patiententagebuch, das an das Krankenhaus zurückgesendet wird. Eingeschlossen sind Erwachsene, die sich einem ambulanten chirurgischen Eingriff mit Vollnarkose und einem Endotrachealtubus oder einer Larynxmaske unterziehen. Die Ausschlusskriterien beinhalten fehlende Sprachkenntnisse in Englisch, Operationen, die zu einem stationären Aufenthalt führen, Schwangerschaft, anhaltende bzw. wiederkehrende Übelkeit und/oder Erbrechen vor der Narkose sowie Patienten mit Regionalanästhesie. Die antiemetische Behandlung liegt im Ermessen des Anästhesisten.

Die Stichprobengröße beträgt 248 Patienten. Diese Teilnehmerdaten werden einerseits für die Validierung des Original PDNV-Scores von Apfel als auch für die Entwicklung des Scores für Tag drei bis sieben verwendet. 141 Patienten leiden während des Zeitraums von sieben Tagen an PONV.

Bezüglich der Kalibrierung liegt ein Kalibrierungsdiagramm für den original Apfel PDNV-Score vor. Die Geradengleichung lautet: $y=1,04x+0,13$ (bezogen auf alle 2170 Patienten der Gesamtkohorte). Für die Diskrimination des PDNV-Apfelscore zeigt sich eine Fläche unter der Kurve von 0,766 an 248 Patienten.[58]

3.4.3 Externe Validierung ohne Update

3.4.3.1 Wallden 2016

Bei dieser Publikation handelt es sich um eine externe Validierung des vereinfachten sowie des koeffizientenbasierten PDNV-Scores von Apfel. Jedoch setzt sich der vereinfachte Score durch. Auf ihn wird im weiteren Verlauf eingegangen. Es wird Übelkeit und Erbrechen nach Entlassung bei einer Kohorte ambulanter, chirurgischer Patienten unter Allgemeinanästhesie bis drei Tage nach einer OP beobachtet.

Die Studie findet in zwei Krankenhäusern in Sundsvall und Sunderbyn in Schweden statt. Die Definition des Endpunktes beschreibt PDNV als jede Episode von Übelkeit und/oder Erbrechen, die ab der Entlassung bis zum dritten postoperativen Tag auftritt. Postoperativ werden den Patienten standardisierte Fragen zu ihren Erfahrungen mit Übelkeit (ja/nein), Erbrechen (ja/nein), Anzahl der Erbrechenschübe und Schmerzen gestellt. Nach der Entlassung beantworten die Patienten die Fragen selbst und schicken die Antworten per Post zurück. Die Einschlusskriterien enthalten

Erwachsene, die sich ambulanten, chirurgischen Eingriffen in Vollnarkose folgender Teilbereiche unterziehen: Gynäkologie, Orthopädie und Allgemeinchirurgie. Des Weiteren sind Forschungsressourcen sowie ein Verständnis des Studienverfahrens und Kommunikation mit dem Patienten nötig. Ein Ausschlusskriterium sind fehlende Daten.

Fast alle Patientinnen erhalten mindestens ein prophylaktisches Antiemetikum, wobei Betamethason am häufigsten eingesetzt wird.

Die Stichprobengröße umfasst 431 Patienten in der Validierungskohorte. PDNV bis einschließlich zum dritten postoperativen Tag erleiden 122 Patienten.

Bezüglich der Kalibrierung ist ein Kalibrierungsdiagramm und eine Geradengleichung: $y = 0,893x + 0,021$ vorhanden. Die Diskriminationsfähigkeit des vereinfachten PDNV-Scores beschreibt eine AUC von 0,693.[90]

3.4.4 Qualitätskriterien des Scores Apfel 2012

Kriterium	Untergruppe	Unterpunkt	Max. erreichte Punktzahl
Generalisierbarkeit	Anzahl des Patientenkollektivs	Scoreentwicklung & -update: 2161 <ul style="list-style-type: none"> Entwicklung: 1913 Update: 248 	2,5
Praktikabilität	Risiko-faktoren in dem Modell	Anzahl der RF in dem Modell: 5	2
		Mehrheit der RF ohne technische Hilfsmittel mit ja/nein Antwort, mittlerer Zeitaufwand	2
		Mehrheit der RF exakt definiert	2
		RF ohne Kenntnis des Outcomes beurteilt: ja	1
	Risiko-berechnung/ Handhabung des Scores	Hilfsmittel nötig: keine Hilfsmittel	3
		vereinfachter Score vorhanden	1
	Endpunkt des Scores	zusammengesetzter Endpunkt PDNV	2
		Art und Weise der Endpunkterhebung: via Gespräch (persönliche Visite/ Telefonat mit dem Patienten)	1
Präzision	Vorhersagegenauigkeit/ Beurteilung der Prädiktion mittels:	Diskriminationsfähigkeit (AUC der ROC-Kurve) <ul style="list-style-type: none"> Endwert: (Intern & Extern) AUC 0,721 	2,5
		Kalibrierung <ul style="list-style-type: none"> Geradengleichung Fläche: 0,059 	3
		Kalibrierung <ul style="list-style-type: none"> Calibration Plot Hosmer-Lemeshow-Test 	1 0

Validierung	Interne Validierung des Scores	Methode der Validierung: Splitting der Daten	1
		Gesamtzahl des Patientenkollektivs: 505	1
	Externe Validierung des Scores	Anzahl der Validierungen: 1	1
		Lokalisation der Validierung: externe Validierung auf einem anderen Kontinent wie Modellentwicklung	3
		Gesamtzahl des Patientenkollektivs: 431	0,5
Breite Anwendung in der Praxis & Extrapunkte	Impact Assessment des Scores	Anwendung des Scores zeigt einen Benefit: keine	0
		bisherige Nutzung und Anwendung des Scores: keine	0
		Gesamtzahl des Patientenkollektivs: 0	0
	Extrapunkte	Modifikation der erreichten Punktzahl: Update des Endpunktes mit Verlängerung auf 7 Tage post OP bei Odom-Forren	1
Insgesamt			30,5 Punkte

Tabelle 6: Qualitätskriterien des Scores Apfel 2012

3.5 Gesamtauswertung des Scores Eberhart 2004a (POVOC-Score)

3.5.1 Modellentwicklung mit interner Validierung

3.5.1.1 Eberhart 2004a

Diese Publikation beschreibt die Entwicklung und interne Validierung eines Risiko-Scores zur Vorhersage von PV bei pädiatrischen Patienten. Das Risikomodell wird auch Postoperative Vomitingin Children-Score (POVOC-Score) genannt.

An vier Zentren in Deutschland werden Patienten rekrutiert. Hiervon sind zwei Unikliniken (u.a. die Klinik für Anästhesie und Intensivmedizin der Philipps-Universität Marburg), ein Kinderkrankenhaus sowie ein ambulantes Zentrum für Operationen. Als Endpunkt wird PV 24 Stunden nach der OP festgelegt. Da Übelkeit ein subjektives Phänomen ist, das kleinere Kinder oft nicht beschreiben können wird sich auf Erbrechen beschränkt. Somit handelt es sich um einen einzelnen Endpunkt, der durch Visite des Patienten oder bei Entlassung via Telefon erhoben wird. Die Einschlusskriterien beinhalten ambulante und stationäre Eingriffe bei Kindern zwischen null bis 14 Jahren unter Vollnarkose. Folgende Fachbereiche werden berücksichtigt: HNO-, Augen-, Bauchchirurgie, Orthopädie, Urologie, zahnmedizinische Verfahren, Diagnoseverfahren und plastische- bzw. ästhetische Chirurgie. Ein Ausschlusskriterium ist die Gabe von prophylaktischem Antiemetikum, einschließlich Kortikosteroiden, sowie unvollständige Daten.

Die Stichprobengröße umfasst insgesamt: 1257 Kinder. Diese werden für eine interne Validierung in zwei Datensätze gesplittet. Der Entwicklungsdatensatz des Scores enthält 657 Patienten und der Validierungsdatensatz 600 Kinder.

Bezüglich der Diskriminationsfähigkeit beträgt die Fläche unter der Kurve im Validierungsdatensatz 0,72 für den vereinfachten Score.

Der finale, vereinfachte Score enthält folgende Risikofaktoren: Operationsdauer > 30 min, Alter > 3 Jahre, Strabismusoperation, positive Vorgeschichte von PV bei Kindern oder PV / postoperative Übelkeit und Erbrechen bei Angehörigen (Mutter, Vater, oder Geschwister)

Pro vorhandenen Risikofaktor wird ein Punkt vergeben, es ergeben sich folgende Vorhersagewerte: [25]

Anzahl Risikofaktoren	PV-Risiko (%)
0	9%
1	10%
2	30%
3	55%
4	70%

Tabelle 7: Wahrscheinlichkeit für das Auftreten von PV: Eberhartscore 2004a[25]

3.5.2 Externe Validierung ohne Update

3.5.2.1 Kranke 2007

Dieses Paper ist die externe Validierung des POVOC-Scores in der Abteilung für Anästhesiologie und Intensivmedizin des Universitätsklinikums Würzburg.

Endpunkt ist das Auftreten von PV während der ersten 24 postoperativen Stunden. Für die Endpunkterhebung werden die Kinder, wenn möglich ihre Eltern, und das Pflegepersonal nach dem Auftreten von PV befragt. Außerdem werden relevante Krankenakten gescreent. Daten von ambulanten Patienten werden durch ein standardisiertes und strukturiertes Telefoninterview 24 Stunden nach Extubation erfragt. Die Einschlusskriterien beinhalten ambulante und stationäre Eingriffe bei Kindern zwischen null bis 16 Jahren. Folgende OPs werden berücksichtigt: Herniotomien, endoskopische Eingriffe (Gastrointestinaltrakt & Rektalchirurgie), kleine und oberflächliche Chirurgie (Laserresektion, Exzisionen, Einschnitte), urologische Operationen, orthopädische Chirurgie einschließlich Unfallchirurgie, Verbrennungen, Otoplastiken, Oralchirurgie und Bronchoskopien. Ebenso werden elektive OPs oder ungeplante Verfahren bei Aufrechterhaltung der Anästhesie mit inhalativem Anästhetikum von Kindern der ASA-Klasse I bis III berücksichtigt. Ausschlusskriterien sind bereits vorher bestehende Übelkeit oder Erbrechen, der Einsatz von Medikamenten mit bekanntem antiemetischen Potenzial in der perioperativen Phase, der postoperative Einsatz einer Nasensonde sowie die postoperative Beatmung oder unvollständige Patientendaten.

Der Validierungsdatensatz besteht aus 524 Kindern, von denen 106 PV erleiden. Es ist ein Kalibrierungsdiagramm vorhanden und die Geradengleichung lautet $y=0,78x+2,37$. Bezüglich der Diskrimination zeigt sich eine Fläche unter der Kurve von 0,72.[49]

3.5.2.2 Bassanezi 2013

Bei dieser Publikation handelt es sich initial um die Entwicklung eines neuen eigenständigen Fuzzy-Logic-Modells zur Vorhersage von postoperativem Erbrechen nach 24 Stunden bei pädiatrischen, onkologischen Patienten. Auf dieses Modell wird hier nicht näher eingegangen. Zusätzlich wird die Prädiktion des neu entwickelten Scores mit dem POVOC-Score verglichen. Somit zählt dieses Paper als externe Validierung für den Eberhart-Score.

Die Studie findet im Boldrini-Kinderkrebszentrum in Campinas in Brasilien statt. Endpunkt ist das postoperative Erbrechen bis 24 Stunden nach der OP. Die Einschlusskriterien umfassen alle stationären OPs bei onkologischen Kindern. Als Ausschlusskriterium zählt die frühzeitige Entlassung aus dem Krankenhaus vor der Endpunkterfassung. Ebenfalls nicht berücksichtigt werden Kinder mit Aufrechterhaltung der Sedierung oder Endotrachealtubus auf der Intensivstation. Antiemetische Prophylaxe sowie der intravenöse Einsatz von Dexamethason in der Neurochirurgie sind nicht gestattet.

Die Stichprobengröße umfasst 198 Kinder. Hiervon erleiden 86 PV.

Die Diskriminationsfähigkeit des POVOC-Scores bei dem brasilianischen Patientenkollektiv ergibt eine AUC von 0,62.[13]

3.5.3 Impact Assessment

3.5.3.1 Wolf 2016

Diese Publikation ist einzuordnen unter Impact Assessment des POVOC-Scores und des vereinfachten Apfel-Scores bei Kindern, die eine Strabismusoperation erhalten. Präoperativ werden die Risikofaktoren für PONV mittels der zwei etablierten Scores erhoben. Bei Kindern im Alter von vier bis 15 Jahren werden diese durch den POVOC-Score ermittelt. Bei Jugendlichen und Erwachsenen ≥ 16 Jahre wird der Apfel-Score herangezogen.

Die Behandlung der Patienten richtet sich nach der Anzahl der Risikofaktoren im Score:

1. Bei null bis zwei Risikofaktoren erfolgt die Narkose als balancierte Anästhesie.
2. Liegendrei bis vier Risikofaktoren vor, wird die Narkose als totale intravenöse Anästhesie (TIVA) aufrechterhalten. Zusätzlich werden hier zwei Antiemetika verabreicht: Dexamethason direkt nach Einleitung und Ondansetron kurz vor OP-Ende.

Die Studie findet an der Klinik für Anästhesiologie, Intensivtherapie und Schmerztherapie des Universitätsklinikums Homburg/Saar statt. Als Endpunkt wird das Auftreten von PONV und die Symptomintensität gewertet. Die Endpunkterfassung erfolgt jeweils zwei, sechs und 24 Stunden postoperativ. Die Einschlusskriterien erfassen elektive, stationäre Strabismus-OPs bei Kindern \geq vier Jahren der ASA Klassen I–III. Außerdem ist die Durchführung der Operation in Allgemeinanästhesie nötig. Die Compliance des Patienten bzw. der erziehungsberechtigten Person für die Befragung und eine unterschriebene Einverständniserklärung werden vorausgesetzt. Ausschlusskriterien sind Schwangerschaft und Teilnahme an einer anderen klinischen Studie.

Die Stichprobengröße sind insgesamt 92 Studienteilnehmer. Bei 45 Kindern wird der POVOC-Score und bei 47 Patienten der vereinfachte Apfel-Score herangezogen. In der kompletten Stichprobe erleiden 16 Kinder PONV.

45 Patienten erhalten eine TIVA mit antiemetischer Prophylaxe, während 47 inbalancierter Anästhesie operiert werden. In den Hochrisikogruppen wird durch die Kombination aus TIVA und Antiemetika die PONV-Inzidenz gegenüber den erwarteten Werten deutlich gesenkt. Die Gesamt-PONV-Inzidenz liegt unter 20 %.

Die Abbildung 2 zeigt eine Gegenüberstellung der tatsächlich aufgetretenen PONV-Inzidenzen (y-Achse und 95 %-Konfidenzintervall) zu den prognostizierten (x-Achse). Die Risikogruppen sind nach den Scoringsystemen aufgeteilt: POVOC (P:2, P:3, P:4) und Apfel (A:0, A:1, A:2, A:3, A:4). „P:2“ entspricht demnach 2 Risikofaktoren nach POVOC Score. Die Breite der Balken ist proportional zur jeweiligen Patientengruppe.

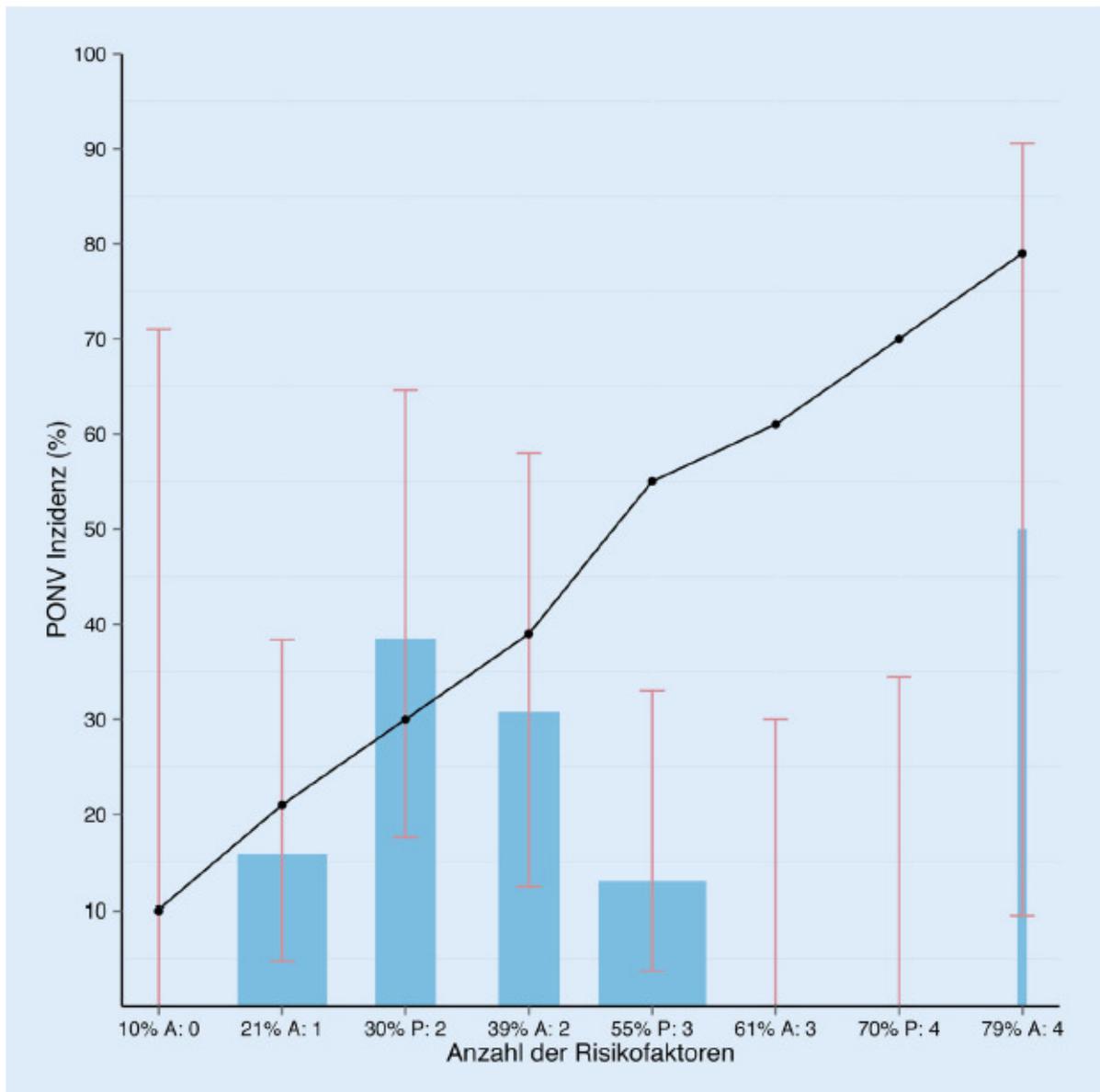


Abbildung 2: PONV-Risiko und Inzidenz[96]

Die PONV-Inzidenz in der gesamten Stichprobe liegt bei 17 %. Die Inzidenz steigt nur bis zu einem Risiko von 30 % (POVOC) bzw. 39 % (Apfel) an. Danach nimmt sie nicht wie prognostiziert weiter zu, sondern fällt ab. Dieser Verlauf ergibt sich aus der Zuteilung der Patienten zu dem entsprechenden Anästhesieverfahren. Ab hier liegen mind. drei Risikofaktoren vor, sodass eine TIVA und zwei Antiemetika appliziert werden.

Die hohe Inzidenz bei Patienten mit zwei Risikofaktoren, der Zeitaufwand für die Risikoeinstufung und etwa ein Drittel der Fälle, die nicht korrekt behandelt und somit von der Analyse ausgeschlossen werden, sprechen für eine generelle PONV-Prophylaxe bei allen Strabismuspatienten.[96]

3.5.4 Qualitätskriterien des Scores Eberhart 2004a

Kriterium	Untergruppe	Unterpunkt	Max. erreichte Punktzahl	
Generalisierbarkeit	Anzahl des Patientenkollektivs	Scoreentwicklung: 657	1	
Praktikabilität	Risiko-faktoren in dem Modell	Anzahl der RF in dem Modell: 4	2	
		alle ohne technische Hilfsmittel, mit ja/nein Antwort, geringer Zeitaufwand	3	
		Mehrheit der RF sind exakt definiert	2	
		RF ohne Kenntnis des Outcomes beurteilt: ja	1	
	Risiko-berechnung/ Handhabung des Scores	keine Hilfsmittel nötig	3	
		vereinfachter Score vorhanden	1	
	Endpunkt des Scores	einzelner Endpunkt PV	1	
		Art und Weise der Endpunkterhebung: via Gespräch (persönliche Visite/ Telefonat mit dem Patienten)	1	
	Präzision	Vorhersagegenauigkeit/ Beurteilung der Prädiktion mittels:	Diskriminationsfähigkeit (AUC der ROC-Kurve) <ul style="list-style-type: none"> Endwert: (Intern & Extern) AUC 0,743 	2,5
			Kalibrierung <ul style="list-style-type: none"> Geradengleichung Fläche: 2,26 	2
Kalibrierung <ul style="list-style-type: none"> Calibration Plot Hosmer-Lemeshow-Test 			1 0	

Validierung	Interne Validierung des Scores	Methode der Validierung: Splitting der Daten	1
		Gesamtzahl des Patientenkollektivs: 600 Kinder	1
	Externe Validierung des Scores	Anzahl der Validierungen: 2	2
		Lokalisation der Validierungen: externe Validierung auf einem anderen Kontinent wie Modellentwicklung	3
		Gesamtzahl des Patientenkollektivs: 722 Kinder	1
Breite Anwendung in der Praxis & Extrapunkte	Impact Assessment des Scores	Anwendung des Scores zeigt einen Benefit: geringere PONV Inzidenz	1
		bisherige Nutzung und Anwendung des Scores: Identifizierung von Hochrisikopatienten	1
		Gesamtzahl des Patientenkollektivs: 45 Kinder	0,5
	Extrapunkte	Modifikation der erreichten Punktzahl: keine	0
Insgesamt			31 Punkte

Tabelle 8: Qualitätskriterien des Scores Eberhart 2004a

3.6 Gesamtauswertung des Scores Van den Bosch 2005a

3.6.1 Modellentwicklung mit interner Validierung

3.6.1.1 Van den Bosch2005a

Bei dieser Publikation handelt es sich um eine erstmalige Modellentwicklung mit interner Validierung mittels Bootstrapping in der Klinik für perioperative Versorgung und Notfallmedizin in dem Universitätsklinikum Utrecht.

Der EndpunktPONV wird24 Stunden nach der Operation erfasst. Er istdefiniert als mindestens eine Episode von Übelkeit, Erbrechen, Würgen oder eine Kombination dieser Symptome.Die Einschlusskriterien beinhalten Erwachsene im Alter von 18-80 Jahren, die sich stationär diversen OP-Eingriffen (Laparoskopien, Ober- &Unterbauchchirurgie, Mittelohr- & Schieloperationen)unterziehen. Als Ausschlusskriterien gelten die Bereiche Herzchirurgie, intrakranielle Neurochirurgie und Notfalloperationen. Des Weiteren werden Patienten mit Schwangerschaft, ASA Klasse IV, morbider Adipositas (Gewicht 120 kg) und schweren Nieren- oder Lebererkrankungen ausgeschlossen. Nicht gestattet sind die Anwendung von antiemetischen Mitteln zwei Wochen vor der Operation.Regionale Anästhesietechniken werden ausgeschlossen, außer im Fall von Oberbauchoperationen mit epiduraler Analgetika-Ergänzung.

Somit ergibt sich eine Gruppe von 1389 randomisierten Studienteilnehmern, die 667 Events im Entwicklungsdatensatz erleiden.

Zur Modellkalibrierung wird der Hosmer-Lemeshow-Test durchgeführt. Bezüglich der Diskriminationsfähigkeit ergibt sich eine AUC von 0,72.

Die abschließend entwickelte Gleichung lautet:

$$p(\text{PONV}) = 1 / (1 + e^{-z})$$

mit $z = 0,15 + 0,46 * (\text{weibliches Geschlecht}) - 0,022 * (\text{Alter}) - 0,63 * (\text{aktuell Raucher}) + 0,76 * (\text{Vorgeschichte von PONV oder Reisekrankheit}) + 0,61 * (\text{Chirurgie des Unterbauchs- oder Mittelohrs}) + 0,72 * (\text{Isoflurananästhesie})$

Es wird zusätzlich ein Normogramm sowie ein Scorchart für das Modell zur einfacheren Berechnung des Risikospräsentiert.[88]

Variable	Points ^a	Variable	Points ^a	Total points	Probability of PONV
Sex		History of PONV or motion sickness		2	0.10
Male	0	No	0	12	0.20
Female	6	Yes	10	19	0.30
				25	0.40
Age (yr)		Current smoking		31	0.50
15-19	20	No	8	36	0.60
20-24	19	Yes	0	42	0.70
25-29	17			49	0.80
30-34	16	Surgery type		59	0.90
35-39	14				
40-44	13	Lower abdominal or middle ear	8		
45-49	11	Other	0		
50-54	10				
55-59	9	Anesthetic technique			
60-64	7				
65-69	6	Propofol	0		
70-74	4	Isoflurane	9		
75-79	3				
80-84	1				
85	0				

^a The points are assigned by multiplying the adjusted regression coefficient (Table 2, second column) of the predictor times 14 and rounded to the nearest integer. The model's intercept (constant) is already taken into account in the score assigned to the predictor age.

Abbildung 3: Scorchart Van den Bosch 2005a[88]

Dargestellt ist die Punkteverteilung pro Prädiktor des endgültigen Vorhersagemodells fürPONV.

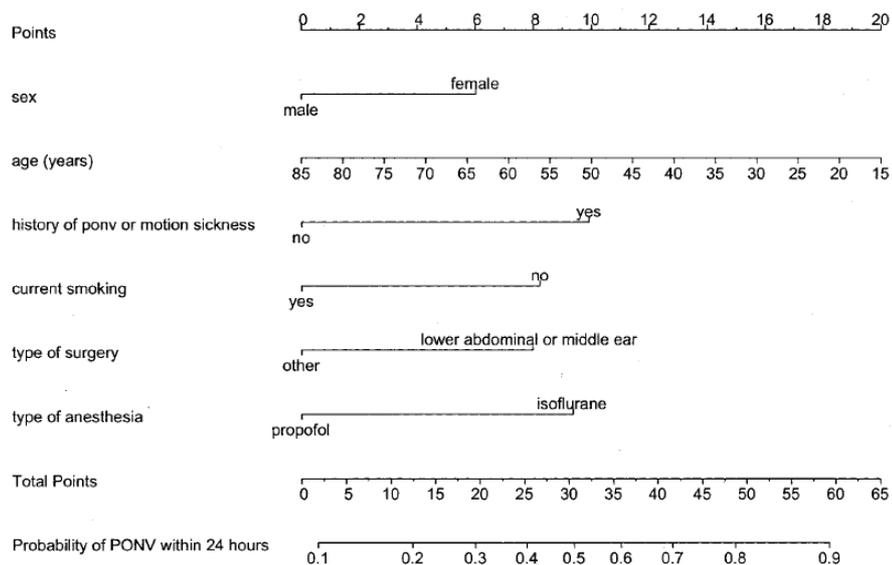


Abbildung 4 Normogramm Van den Bosch 2005a[88]

Diese Abbildung zeigt die Prädiktoren des endgültigen Modells von PONV innerhalb der ersten 24 Stunden nach einer OP. Die Länge der Linie jedes Prädiktors im Normogramm gibt den relativen Beitrag des Prädiktors zur Wahrscheinlichkeit von PONV an.

3.6.2 Externe Validierung mit Update

3.6.2.1 Kappen2012

Dies ist die erste externe Validierung des ursprünglichen Modells „van den Bosch20005a“ mit Update. Das Vorhersagemodell wird mit einem Datensatz von 1847 Patienten („update data set“) aktualisiert und mit 3822 Patienten validiert. Des Weiteren wird zu Vergleichszwecken eine neue Modellentwicklung mit interner Validierung durchgeführt. Dieses neue Modell setzt sich nicht durch und wird im Folgenden nicht behandelt.

Studienort ist das Universitätsklinikum Utrecht in den Niederlanden. Als Endpunkt gilt das Auftreten von PONV innerhalb von 24 Stunden nach der OP. Eingeschlossen werden ambulante und stationäre Behandlungen bei Erwachsenen. Die Stichprobe beinhaltet insgesamt 5669 Patienten und teilt sich auf in einen Updatedatensatz von 1847 Patienten (1203 stationär, 644 ambulant) sowie einen Validierungsdatsatz von 3822 Patienten. PONV erleiden in der Studiengruppe insgesamt 1993 Patienten, gesplittet in 653 im Updatedatensatz und 1340 im Validierungsdatsatz.

Die Kalibrierungsgerade des überarbeiteten Modells im externem Validierungsdatsatz lautet: $y = 1,00x + 0,34$. Die Diskriminationsfähigkeit des „updated Modells“ im externen Validierungsdatsatz zeigt eine AUC von 0,68.

Die Überarbeitung des Modells besteht aus mehreren Komponenten. Der Risikofaktor Unterbauch- oder Mittelohrchirurgie wird ersetzt durch Bauch- oder Mittelohrchirurgie. Der Faktor Isofluran und/oder Lachgasanästhesie wird ausgetauscht durch Inhalationsnarkose. Als neuer protektiver Punkt wird ambulante Operation hinzugefügt. Alle Risikofaktoren erhalten neue Regressionskoeffizienten und der Achsenabschnitt wird neu berechnet.

Als endgültiges upgedatetes Modell entstand:

$$p(\text{PONV}) = 1 / (1 + e^{-z})$$

mit $z = 0,12 + 0,36 * (\text{weibliches Geschlecht}) - 0,017 * (\text{Alter}) - 0,50 * (\text{aktuell Raucher}) + 0,60 * (\text{Vorgeschichte von PONV oder Reisekrankheit}) + 0,48 * (\text{Chirurgie des Bauchs- oder Mittelohrs}) + 0,35 * (\text{Inhalationsnarkose}) - 1,16 * (\text{ambulante OP})$ [39]

3.6.3 Impact Assessment

3.6.3.1 Kappen2014

Diese Publikation beschreibt ein Impact Assessment für das upgedatete Modell, wie es in Kappen2012 vorgestellt wird. Der Studienort ist das Universitätsklinikum in Utrecht in den Niederlanden.

Der Ablauf der Studie ist wie folgt: Die Anästhesisten der Klinik werden randomisiert in eine Interventionsgruppe, die das vorhergesagte PONV-Risiko durch die Formel angezeigt bekommt. Als zweites Team existiert die „Care-as-usual-Gruppe“.

Primärer Endpunkt ist die Auswirkung des Vorhersagemodells auf das Patientenergebnis, die PONV-Inzidenz. Das sekundäre Ergebnis ist die Änderung des Arztverhaltens und die Verabreichung einer risikoabhängigen PONV-Prophylaxe.

PONV wird definiert als das Auftreten mindestens eines der folgenden Ereignisse innerhalb der ersten 24 Stunden nach der Operation: Übelkeit, Erbrechen oder die Verabreichung eines Rettungsantiemetikums. Hierbei wird Erbrechen als Ausschuss von Mageninhalt festgelegt. Einschlusskriterien erfüllen ambulante und stationäre erwachsene Patienten, die sich in Allgemeinanästhesie elektiven, nichtkardialen Operationen unterziehen. Ausschlusskriterien für Patienten sind Schwangerschaft, postoperative Aufnahme auf der Intensivstation oder Beatmung über Nacht auf der Postanästhesie-Station. Außerdem werden Patienten mit Sprachbarriere (Niederländisch oder Englisch) ausgeschlossen. Anästhesisten, die weniger als 50 eingeschriebene Patienten während des gesamten Untersuchungszeitraums behandeln, werden ebenso nicht berücksichtigt.

Die Studiengröße umfasst insgesamt 11613 Patienten. Diese sind aufgeteilt in eine Interventionsgruppe mit 5471 und „Treatment-as-usual-Gruppe“ mit 6142 Patienten.

Die Datenerfassung erfolgt durch medizinisches Pflegepersonal anhand eines validierten Fragebogens. Die Daten werden auf der Station gesammelt oder, wenn die Patienten bereits entlassen wurden, per Telefon eingeholt.

In der Hauptanalyse verabreichen Anästhesisten der Interventionsgruppe mehr Antiemetika-Prophylaxe als Anästhesisten der „Fürsorge-wie-sonst-Gruppe“. Anästhesisten der Interventionsgruppe verabreichen Patienten mit höherem Risiko mehr prophylaktische Antiemetika und weniger Antiemetika für Patienten mit geringem Risiko. Die gesamt- als auch die risikoabhängigen Unterschiede im ärztlichen Verhalten sind statistisch signifikant. Die maßgeschneiderte Verschreibung von

Antiemetika in der Interventionsgruppe führt jedoch nicht zu einer wesentlich geringeren Inzidenz von PONV.

Zusammenfassend ergibt die Einführung eines Vorhersagemodells trotz Zunahme der „risikoadaptierten“ Anwendung prophylaktischer antiemetischer Strategien keine relevante Abnahme der PONV-Inzidenz. Diese Studie demonstriert die Grenzen der Anwendung von Scores. Die Diskrepanz der Ergebnisse zeigt die Notwendigkeit einer Wirkungsanalyse, die das Patientenergebnis einschließt, bevor ein Vorhersagemodell in die klinische Praxis implementiert wird.[40]

3.6.3.2 Kappen 2015

Diese Studie beschreibt ebenfalls Impact Assessment für das upgedatete Modell, wie es in Kappen 2012 vorgestellt wurde. Der Studienort ist die anästhesiologische Abteilung des Universitätskrankenhauses Utrecht in den Niederlanden.

Diese Vorher-Nachher-Studie baut auf die Publikation „Kappen 2014“ auf. Sie beschreibt einen Interventionszeitraum, indem Anästhesisten das berechnete PONV Risiko und ein konkreter Prophylaxe vorschlag in der Prämedikation angezeigt wird. Die Auswirkungen auf die Inzidenz von PONV sowie die Verabreichung einer risikoabhängigen PONV-Prophylaxe wird quantifiziert.

PONV wird definiert als das Auftreten von mindestens einem der folgenden Ereignisse 24 Stunden nach der Operation: eine Episode von Übelkeit, eine Episode von Erbrechen oder die Verabreichung eines Antiemetikums zur Rettung.

Einschlusskriterien sind elektive, stationäre OPs in Vollnarkose bei Erwachsenen. Ausgeschlossen werden ambulante Verfahren, Schwangere, Patienten, die eine postoperative Aufnahme auf der Intensivstation oder nächtliche Beatmung auf der Post-Anästhesie-Station benötigen. Ein weiterer Ausschlussgrund sind mangelnde Sprachkenntnisse in Niederländisch oder Englisch.

Das Prognosesystem funktioniert wie folgt:

Basierend auf den Berechnungen des Software-Tools wird das prognostizierte PONV-Risiko mit einer anfänglichen Ampelfarbe eingestuft:

- „grün“: → keine Antiemetika, 26% prognostiziertes Risiko
- „gelb“: → ein Antiemetikum, 41% prognostiziertes Risiko
- „orange“: → zwei Antiemetika, unter 62% prognostiziertes Risiko

- „rot“: →drei Antiemetika, für 62% oder mehr prognostiziertes Risiko

Die Anästhesisten entscheiden, ob sie den Rat befolgen und entsprechend prophylaktische Antiemetika verabreichen. Die Farbe ändert sich erst, wenn ein Antiemetikum verabreicht wird. Die Software adaptiert das vorhergesagte Risiko um eine relative Risikoreduktion von 26% pro Antiemetikum. Wenn das Risiko unter einen Schwellenwert fällt, wird die Ampel grün und es werden keine weiteren prophylaktischen Antiemetika empfohlen.

Die Studie umfasst insgesamt 1480 Patienten, 1022 Patienten in der „Care-as-usual-Phase“ und 458 Patienten im Interventionszeitraum.

Forschungspersonal und ausgebildete Medizinstudenten sammeln Daten über das Auftreten von postoperativer Übelkeit mit Hilfe eines validierten Fragebogens.

Als Fazit zeigt sich, dass während des Interventionszeitraums die Anästhesisten mehr Antiemetika für Patienten verabreichen, bei denen ein erhöhtes PONV-Risiko identifiziert wurde. Dieser direkte Ansatz führt zu einer Verringerung der PONV-Inzidenz mit einer noch stärkeren Reduktion des PONV bei Hochrisikopatienten.[41]

3.6.3.3 Kappen 2016

Diese Studie mit dem Titel „Barrieren und Unterstützung, die von Ärzten bei der Verwendung von Vorhersage-Modellen in der Praxis wahrgenommen werden“ behandelt ebenfalls Impact Assessment. Anlass war der Score von "Kappen 2012" sowie die Nachfolgestudien aus den Jahren 2014 und 2015. Sie behandelt die Frage, warum die klinische Implementierung eines validierten und rekalierten Vorhersagemodells die Patientenergebnisse in einem Cluster einer randomisierten, klinischen Studie (RCT) nicht verbesserte.

Die Untersuchung findet in der Abteilung für Anästhesiologie, Intensiv- und Notfallmedizin des Universitätsklinikums Utrecht in den Niederlanden statt. Als Endpunkt dient die PONV-Inzidenz nach Implementierung des Scores. Das Setting dieser Studie besteht aus einer sequentiellen, qualitativen und einer quantitativen Phase, die am Ende des Cluster-RCT durchgeführt wird.

Zuerst erfolgt die Sammlung qualitativer Daten über mögliche Hilfen und Barrieren für die risikoabhängige, prophylaktische Behandlung von PONV durch persönliche, semistrukturierte Interviews mit einer Stichprobe von acht Ärzten. Das Interview besteht aus vier Blöcken:

1. die vom Arzt wahrgenommene Belastung der Patienten durch PONV
2. die Risikostratifizierung: Wie identifiziert der Arzt Patienten mit PONV-Risiko?
3. der Entscheidungsprozess bezüglich präventiver antiemetischer Strategien
4. die Einstellung zur Verwendung von Prädiktionsmodellen und Entscheidungshilfen für die prophylaktische Behandlung von PONV

In einem zweiten Schritt werden strukturierte, webbasierte Umfragen unter allen 57 Ärzten, die an der RCT teilgenommen haben, durchgeführt. Die Interviews ergeben Unterschiede zwischen den Ärzten der Interventionsgruppe und der „Care-as-usual-Gruppe“.

Die Stichprobengröße enthält insgesamt 57 Anästhesisten, die an der webbasierten Umfrage teilnehmen. Das Interview führen acht Ärzte.

Obwohl das Prognoseinstrument die Ärzte auf die PONV-Prävention aufmerksam gemacht hat, berichten die Ärzte über drei Hemmnisse für die Verwendung von Scoringsystemen:

1. PONV wird nicht als Ergebnis von äußerster Wichtigkeit betrachtet.
2. Die Entscheidung über die PONV-Prophylaxe ist eher intuitiv als risikobasiert.
3. Prognosemodelle gewichten nicht den Nutzen und das Risiko prophylaktischer Medikamente.

Somit ergibt sich als Anregung für eine erfolgreiche Implementierung eines Vorhersagemodells in der klinischen Praxis das Hinzufügen einer umsetzbaren Empfehlung zum vorhergesagten Risiko. Die Darstellung des vorhergesagten Risikos sollte automatisiert in den Arbeitsablauf integriert werden. Die Forschungsergebnisse des zugrunde liegenden Vorhersagemodells sollten für die Ärzte verfügbar sein. Prinzipiell ist ein Score besser, wenn er Ergebnisse vorhersagt, die für Ärzte und ihre Patienten relevant sind.[42]

3.6.4 Qualitätskriterien des Scores Van den Bosch 2005a

Kriterium	Untergruppe	Unterpunkt	Max. erreichte Punktzahl	
Generalisierbarkeit	Anzahl des Patientenkollektivs	Scoreentwicklung & -update: 2042 <ul style="list-style-type: none"> ▪ Entwicklung: 1389 ▪ Update: 653 	2,5	
		Risiko-faktoren in dem Modell <ul style="list-style-type: none"> ▪ 6 RF im Original ▪ 7 RF nach Update 	2	
Praktikabilität	alle ohne technische Hilfsmittel, mit ja/nein Antwort, geringer Zeitaufwand	Minderheit der RF exakt definiert	3	
		RF ohne Kenntnis des Outcomes beurteilt: ja	1	
		Risiko-berechnung/ Handhabung des Scores	Hilfsmittel nötig: spezielles Tabellenwerk	1
		kein vereinfachter Score	0	
	Endpunkt des Scores	zusammengesetzter Endpunkt PONV	2	
		Art und Weise der Endpunkterhebung: via Routinedaten	3	
	Präzision	Vorhersagegenauigkeit/ Beurteilung der Prädiktion mittels	Diskriminationsfähigkeit (AUC der ROC-Kurve) <ul style="list-style-type: none"> ▪ Endwert: (Intern & Extern) AUC: =0,734 	2,5
			Kalibrierung <ul style="list-style-type: none"> ▪ Geradengleichung Fläche: 0,24 	3
Kalibrierung <ul style="list-style-type: none"> ▪ Calibration Plot ▪ Hosmer-Lemeshow-Test 			0 1	

Validierung	Interne Validierung des Scores	Methode der Validierung: Bootstrapping	1
		Gesamtzahl des Patientenkollektivs: nicht berichtet	0
	Externe Validierung des Scores	Anzahl der Validierungen: 1	1
		Lokalisation der Validierungen: externe Validierung im selben Land wie Modellentwicklung	1
		Gesamtzahl des Patientenkollektivs: 3822	3
	Breite Anwendung in der Praxis & Extrapunkte	Impact Assessment des Scores	Anwendung des Scores zeigt Benefit: geringere PONV Inzidenz
bisherige Nutzung und Anwendung des Scores: keine			0
Gesamtzahl des Patientenkollektivs: 5929			3
Extrapunkte		Modifikation der erreichten Punktzahl: Abschlusspublikation Kappen2016: Hindernisse bei der Nutzung bzw. Ursachen für die fehlende PONV Reduktion trotz Score Nutzung	1
Insgesamt			33 Punkte

Tabelle 9: Qualitätskriterien des Scores Van den Bosch 2005a

3.7 Gesamtauswertung des Scores Koivuranta-Scores

3.7.1 Modellentwicklung

3.7.1.1 Koivuranta1997

Bei dieser Studie handelt es sich um eine Modellentwicklung ohne externe Validierung. Es wird sowohl ein Score mit Gleichung als auch ein vereinfachtes Modell mit fünf gleichgewichteten Risikofaktoren entwickelt. Die Studie findet in dem Universitätsklinikum in Oulu in Finnland statt.

Der Endpunkt PONV wird für Übelkeit und Erbrechen jeweils getrennt erfasst. Übelkeit wird anhand des subjektiven Gefühls des Patienten, sich krank zu fühlen oder sich übergeben zu wollen, beurteilt. Emetische Episoden werden getrennt als Würgen oder Erbrechen aufgezeichnet. Die Einschlusskriterien umfassen stationäre Patienten beider Geschlechter im Alter von vier bis 86 Jahren der ASA Klassen 1-3, die sich einer elektiven Operation in Allgemein- oder Regionalanästhesie unterziehen. Folgende Fachbereiche werden einbezogen: Gynäkologie, HNO-Heilkunde, Augenheilkunde, Orthopädie und Intraabdominalchirurgie. Des Weiteren ist ein Follow-up für die ersten zwei Stunden im Aufwachraum nötig. Ausgeschlossen sind schwangere oder intensivpflichtige Patienten.

Die Stichprobengröße umfasst 1107 Patienten. Davon erleiden 576 Übelkeit und 277 Erbrechen.

Die Diskriminierungsfähigkeit ergibt eine AUC von 0,719 für Übelkeit und eine AUC von 0,695 für Erbrechen.

Es wird folgende Gleichung präsentiert:

$$p(\text{PON/PV}) = e^z / (1 + e^z)$$

mit $z = 0,93 * (\text{weibliches Geschlecht}) + 0,82 * (\text{PONV in der Vorgeschichte}) + 0,75 * (\text{OP-Dauer} > 60 \text{ Minuten}) + 0,61 * (\text{Nichtraucher}) + 0,59 * (\text{Reisekrankheit in der Vorgeschichte})$. Der Wert für den Achsenabschnitt wird nicht veröffentlicht.

Ein vereinfachter Score, indem alle Risikofaktoren gleich gewichtet sind wird publiziert:[46]

Anzahl Risikofaktoren	PON-Risiko (%)	PV-Risiko (%)
0	17 %	7 %
1	18 %	7 %
2	42 %	17 %
3	54 %	25 %
4	74 %	38 %
5	87 %	61 %

Tabelle 10:Wahrscheinlichkeit für das Auftreten von PON bzw. PV: Koivurantascore1997[46]

3.7.2 Externe Validierung mit Update

3.7.2.1 Engel2006

Dieses Paper wurde bereits unter 3.1.2.1 ausführlich erläutert. Es handelt sich um eine externe Validierung mit Anpassung der Koeffizienten der folgenden Scores in Gießen: original Koivuranta, Apfel und Sinclair sowie um eine interne Validierung von Junger.

PONV 24 Stunden nach der Operation wird als Endpunkt definiert. Bei dem Update werden die Risikofaktoren der jeweiligen Modelle beibehalten. Die Koeffizienten und der Achsenabschnitt werden angepasst und erneut mit dem gleichen Datensatz durch Kreuzvalidierung getestet. Es wird nun auf das überarbeitete Modell eingegangen.

Die Stichprobengröße des Validierungsdatensatzes beträgt 748 Patienten. Hiervon erleiden 84 PONV.

Für die Kalibrierung wird der Hosmer-Lemeshow-Test durchgeführt. Bezüglich der Diskriminationsfähigkeit ergibt sich eine AUC von 0,648 für den upgedaten Koivuranta Score.

Bei identischer Geradengleichung lauten die Werte nach dem Update wie folgt:

$$p(\text{PON/PV}) = e^z / (1 + e^z)$$

mit $z = -2,67 + 0,67 * (\text{weibliches Geschlecht}) + 0,77 * (\text{PONV in der Vorgeschichte}) + 0,21 * (\text{OP-Dauer} > 60 \text{ Minuten}) - 0,88 * (\text{Raucher}) + 1,96 * (\text{Reisekrankheit in der Vorgeschichte})$ [26]

3.7.3 Externe Validierung ohne Update

3.7.3.1 Eberhart2000a

Dieses Paper wurde bereits unter 3.2.2.2 abgehandelt, da es sich mit dreiverschiedenen Scores zur Vorhersage von postoperativem Erbrechen bzw. postoperativer Übelkeit und Erbrechen befasst und sie extern validiert. Die Scores sind von folgenden Autoren: Palazzo und Evans, der original Apfel-Score aus dem Jahre 1998 und der vereinfachte Score von Koivuranta.

Das Studienkollektiv enthält 1444 Patienten. Davon leiden 247 Patienten an PV und 540 Patienten an PONV.

Ein Kalibrierungsdiagramm liegt vor. Die Kalibrierungsgeradengleichung zum vorhergesagten Risiko für PONV lautet: $Y = 0,99x - 5,1$. Die Geradengleichung für PV lautet: $Y = 1,71x + 4,4$. Die Diskriminationsfähigkeit des Scores für PONV ergibt eine AUC von 0,71 und für PV wird eine AUC von 0,73 beschrieben. [23]

3.7.3.2 Trope2000

Diese Publikation beschreibt die Entwicklung eines eigenständigen Scores ohne Validierung sowie eine externe Validierung vom vereinfachten Apfelscore sowie dem vereinfachten Koivurantascore. Das selbst entwickelte Scoringsystem für PONV heißt Ullevalscore und besteht aus zehn Fragen, die mit ja oder nein zu beantworten sind. Auf dieses Modell wird hier nicht näher eingegangen.

Der Studienort wird nicht genannt. Als Endpunkt dient PONV innerhalb von 24 Stunden nach der OP. Die Einschlusskriterien beinhalten Erwachsene, die sich verschiedenen elektiven OPs unterziehen. Folgende Gebiete finden Berücksichtigung: gynäkologische Laparoskopien, Laparotomien, Gastrochirurgie, Krebsoperationen der Mamma, HNO-Chirurgie, Schilddrüsenoperationen und orthopädische Chirurgie. Ausschlusskriterien werden nicht genannt.

Die Stichprobengröße umfasst 120 Patienten, von denen 72 an PONV erkrankten.

Die Kalibrierungs- und Diskriminationsfähigkeit wird nicht näher beschrieben.[86]

3.7.3.3 Apfel 2002

Auf diese Veröffentlichung wurde unter dem Punkt 3.2.2.3 ausführlich eingegangen, da es sich um eine Publikation handelt, die sechs verschiedene Scores (Apfel 1999, Palazzo 1993, vereinfachter Koivuranta 1997, Sinclair 1999, Scholz 1999 und Gan 2000) in einer externen Validierung vergleicht.

Als Zusatzinformation wird berichtet, dass von den 1566 Patienten 600 an PONV leiden.

Ein Kalibrierungsdiagramm existiert und die Geradengleichung lautet $y = 1.13x - 0.10$. Für die Diskriminationsfähigkeit wird eine AUC von 0,70 berichtet.[8]

3.7.3.4 Traeger 2003

In dieser Studie wird die Modellentwicklung eines neuronalen Netzwerkes zur PONV-Vorhersage vorgestellt. Zusätzlich erfolgt eine externe Validierung dieses Modells sowie der Vergleich mit den vereinfachten, etablierten Scores von Koivuranta und Apfel. Auf das künstliche neuronale Netzwerk geht diese Arbeit nicht näher ein.

Die Daten für die Modellentwicklung werden am Universitätsklinikum in Marburg erhoben. Die 400 Patienten für die Validierung stammen aus einem anderen Zentrum. Als Endpunkt dient PONV 24 Stunden nach der OP. Eingeschlossen werden Erwachsene, die sich einem elektiven, stationären Eingriff unterziehen.

Ausschlusskriterien sind eine antiemetische Prophylaxe oder Substanzen mit antiemetischen Effekten.

Die Stichprobengröße umfasst 2164 Patienten. Diese sind aufgeteilt in einen Entwicklungsdatensatz von 1764 (aufgesplittet in 1364 Patienten für das Training des KNN und 400 zur Überwachung) und 400 Patienten für die Validierung des Modells sowie der bereits etablierten Scores.

Es existiert ein Kalibrierungsdiagramm sowie eine Geradengleichung. Diese lautet für das vereinfachte Koivuranta Modell: $y = 0,86x - 5$. Die Diskriminationsfähigkeit des Modells wird mittels der AUC von 0,69 beschrieben.[85]

3.7.3.5 Eberhart2004

Dieser Score wurde bereits unter dem Punkt 3.2.2.4 behandelt, da es sich um eine externe Validierung ohne Update von fünf Scores (Palazzo1993, vereinfachter Score von Koivuranta1997, Apfel1998, Sinclair 1999, Apfel 1999) an pädiatrischen Patienten handelt. Als Endpunkt dient postoperatives Erbrechen innerhalb der ersten 24 Stunden nach der OP. Als Resümee ergibt sich, dass der bestehende Score nicht auf Kinder übertragbar ist.

Als Zusatzinformationen zu dem Score von Koivuranta ist eine Stichprobengröße von 983 Kindern erwähnt. Hiervon erleiden 326 PV.

Bezüglich der Kalibrierung liegt sowohl ein Kalibrierungsdiagramm für PONV als auch für PV vor. Die Geradengleichungen des vereinfachten Scores lauten für PONV: $Y = 0,64x + 12$ und für PV: $Y = 1,18x + 8$. Die Diskriminationsfähigkeit wird durch die Fläche unter der Kurve gemessen und ergibt zur Vorhersage von PV eine AUC von 0,61.[24]

3.7.3.6VandenBosch2005

Bei dieser Publikation handelt es sich um eine externe Validierung ohne Update zu den Risikomodellen von Apfel (original und vereinfachter Score) sowie Koivuranta (original und vereinfachter).

Die Studie findet am Akademisch-Medizinischen Zentrum der Universität Amsterdam, in den Niederlanden statt. Als Endpunkt wird PONV 24 Stunden nach der OP festgelegt. Dieser Endpunkt wird registriert, wenn mindestens eine Episode von postoperativer Übelkeit, Würgereiz, Erbrechen oder eine Kombination dieser Symptome auftritt. Die Einschlusskriterien erfüllen Erwachsene, die sich stationären Eingriffen unterziehen. Es werden alle Arten von Operationen außer Herzchirurgie und intrakranielle Prozeduren berücksichtigt. Ausschlusskriterien sind Notfalloperationen, eine mögliche Schwangerschaft, ASA Status 4, morbid Adipositas (Gewicht > 120 kg), Nieren- oder Lebererkrankungen vor Anwendung einer der beiden Anästhesietechniken. Weitere Ausschlusskriterien sind die Einnahme von antiemetischen oder pro-emetischen Medikamenten 2 Wochen vor der OP, eine vorherige Einschreibung in die gleiche Studie, unzureichende Beherrschung der niederländischen Sprache oder unvollständige Daten.

Die Stichprobengröße des Validierungsdatensatzes beträgt 1388 Teilnehmer. Davon erleiden 666 PONV.

Es gibt ein Kalibrierungsdiagramm für das original Koivurantamodell. Für die Gleichung der Kalibrierungsgeraden des original Koivurantamodells wird nur die Steigung mit einem Wert von 0,58 genannt, nicht der Achsenabschnitt. Für das ursprüngliche Koivurantamodell zeigt sich ein signifikanter Hosmer-Lemeshow-Test. Die Diskriminationsfähigkeit wird durch die Fläche unter der Kurve erfasst und ergibt für den ursprünglichen Koivurantascore eine AUC von 0,66 sowie für den vereinfachten Koivurantascore ebenfalls 0,66.[87]

3.7.3.7 Peng2007

Diese Publikation hat initial als Ziel die Neuentwicklung eines Vorhersagetools für PONV mittels neuronalen Netzwerks. Es wird intern validiert und enthält sieben Risikofaktoren. Auf diesen Teil der Studie wird hier jedoch nicht näher eingegangen. Zusätzlich erfolgt in diesem Paper eine externe Validierung des vereinfachten Koivurantascores sowie des vereinfachten Apfelscores.

Diese Studie findet im Shin Kong Wu Ho-Su Memorial Hospital in Taiwan statt. Als Endpunkt gilt PONV innerhalb von 24 Stunden nach der OP. PONV wird definiert als Patienten, die während der postoperativen Periode trockenen Husten, Übelkeit, Würgen oder Erbrechen verspüren. Einschlusskriterien beinhalten Erwachsene, die sich elektiven, stationären Operationen in Vollnarkose ohne routinemäßige Antiemetikatherapie unterziehen. Ausgeschlossen werden Kinder (Alter <18 Jahre), Patienten mit postoperativer Beatmungsunterstützung oder Intensivpflege sowie Patienten mit Vollnarkose, Spinal- oder Epiduralanästhesie. Patienten, die augenärztliche Eingriffe erhalten, werden ebenfalls nicht berücksichtigt.

Die Stichprobengröße beinhaltet insgesamt 1086 Patienten. Diese sind aufgeteilt in einen Entwicklungsdatensatz von 656 und einen Validierungsdatensatz von 430 Patienten. In der Validierungskohorte erleiden 104 Patienten PONV.

Die Diskriminationsfähigkeit wird durch den AUC-Wert von 0,626 angegeben. Als sonstige Maßzahlen für den Koivurantascore werden die Sensitivität mit 66,3%, die Spezifität mit 57,4% sowie ein positiver prädiktiver Wert von 33,2% und ein negativer Prädiktiver Wert von 84,2% angeführt.[62]

3.7.3.8Wu2015

Diese Publikation wurde bereits unter dem Punkt 3.2.2.5 beschrieben, da es sich um eine externe Validierung ohne Update von fünf Scoringsystemen in Taiwan handelt. Sie lauten: original Apfel- und vereinfachter Apfelscore, original Koivuranta- und vereinfachter Koivurantascore sowie Palazzo & Evans-Score.

Die Stichprobengröße beträgt 992 Patienten. Hiervon erleiden 416 Patienten PONV innerhalb von 24 Stunden nach der OP.

Für die Kalibrierung wird der Hosmer-Lemeshow-Test durchgeführt. Die Diskriminationsfähigkeit beschreibt eine AUC von 0,674 für den original Koivurantascore und 0,632 für den vereinfachten.

Zusätzlich werden als Maßzahlen zur Modellklassifizierung folgende Werte genannt: Sensitivität (71,2% Originalscore/77,6% vereinfachter Koivurantascore), Spezifität (59,0% Originalscore / 42,9% vereinfachter Koivurantascore), positiver prädiktiver Wert (55,6% Originalscore/ 49,5% vereinfachter Koivurantascore), negativer prädiktiver Wert (73,9% Originalscore / 72,6% vereinfachter Koivurantascore), Youden's Index (30,2% Originalscore/ 20,5% vereinfachter Koivurantascore).[97]

3.7.4 Impact Assessment

3.7.4.1 Thomas2002

Diese Publikation wurde unter dem Punkt 3.2.3.1 beschrieben und zählt zu der Rubrik Impact Assessment.

Während eines RCTs zu den Effekten von kombinierten Antiemetika werden Patientengruppen mithilfe von Scoringsystemen verglichen. Es werden folgende Scores zur Hilfe gezogen: vereinfachter Koivuranta, Palazzo und Evans, Sinclair, vereinfachter Apfel. Die Übereinstimmungsquote zwischen den einzelnen Scoringsystemen ist unzureichend. Mögliche Ursache ist das ambulante, rein weibliche Patientenkollektiv dieser Studie.[83]

3.7.4.2Kumar2012

Diese Studie beschreibt ein Impact Assesment des vereinfachten Koivurantascores.

Der Studienort ist die Abteilung für Anästhesiologie des Aberdeen Royal Krankenhauses in Großbritannien. Die Studiendaten stammen aus intermittierenden Überprüfungen der Unterlagen aus den Jahren 2006-2010. Als Endpunkt gilt PONV innerhalb von 24 Stunden nach der OP. Die Einschlusskriterien erfassen Erwachsene. Ausschlusskriterien werden nicht genannt.

Mithilfe des Scores werden Patienten in drei Risikoklassen eingeteilt und es wird risikoadaptiert eine Prophylaxeempfehlung ausgesprochen. Diejenigen Patienten mit null bis einem Punkt im Koivurantascore gelten als geringes Risiko. Von einem mittleren Risiko wird bei zwei bis drei Punkten und von einem hohen Risiko bei vier bis fünf Punkten im Score gesprochen. Hochrisikopatienten sollen nach der Richtlinie mindestens drei Antiemetika erhalten. Menschen mit mittlerem Risiko sollen mindestens zwei und Patienten mit geringem Risiko null oder ein Antiemetika bekommen.

Die Stichprobengröße für das Impact Assessment enthält 749 Patienten.

Die PONV-Rate in der Gruppe mit niedrigem Risiko ist gemittelt bei 22%. Die Gruppe mit mittlerem Risiko hat eine durchschnittliche PONV Inzidenz von 34%. Die PONV Inzidenz in dieser Gruppe mit Guideline beträgt 28%, ohne Guideline 38%. Die Hochrisikogruppe erleidet im Mittel zu 55% PONV, mit Guideline sind es 33%, ohne Guideline erkranken 62%.

Die vereinfachte Bewertung der Patienten nach dem PONV-Risiko ist effektiv. Die Ergebnisse bestätigen, dass Gruppen mit hohem Risiko und mittlerem Risiko von der Richtlinie profitieren. Leider ist die Einhaltung der Richtlinien gering. Nur 37% der Patienten mit mittlerem und hohem Risiko erhalten die angegebene Prophylaxe.[51]

3.7.5.1 Qualitätskriterien des original Scores Koivuranta 1997

Kriterium	Untergruppe	Unterpunkt	Max. erreichte Punktzahl	
Generalisierbarkeit	Anzahl des Patientenkollektivs	Scoreentwicklung & -update: 1855 <ul style="list-style-type: none"> ▪ Entwicklung: 1107 ▪ Update: 748 	2	
		Risiko-faktoren in dem Modell	Anzahl der RF in dem Modell: 5	2
Praktikabilität	Risiko-berechnung/ Handhabung des Scores	Mehrheit der RF ohne technische Hilfsmittel mit ja/nein Antwort, mittlerer Zeitaufwand	2	
		Mehrheit der RF sind exakt definiert	2	
		RF ohne Kenntnis des Outcomes beurteilt: nein/ unklar	0	
		Hilfsmittel nötig: Taschenrechner/mit Handy (z.B. als App möglich)	2	
	Endpunkt des Scores	kein vereinfachter Score	0	
		einzelne Endpunkte PN/PV/PR	1	
	Präzision	Vorhersagegenauigkeit/ Beurteilung der Prädiktion mittels:	Art und Weise der Endpunkterhebung: via Gespräch (persönliche Visite)	1
			Diskriminationsfähigkeit (AUC der ROC-Kurve) <ul style="list-style-type: none"> ▪ Endwert (Intern & Extern) AUC: 0,719 	2,5
Kalibrierung <ul style="list-style-type: none"> ▪ Geradengleichung 			0	
		Kalibrierung <ul style="list-style-type: none"> ▪ Calibration Plot ▪ Hosmer-Lemeshow-Test 	1 1	

Validierung	Interne Validierung des Scores	Methode der Validierung: keine	0
		Gesamtzahl des Patientenkollektivs: 0	0
	Externe Validierung des Scores	Anzahl der Validierungen: 3x <ul style="list-style-type: none"> ▪ 1x mit Update ▪ 2x ohne Update: 	3
		Lokalisation der Validierungen: externe Validierung auf einem anderen Kontinent wie Modellentwicklung	3
		Gesamtzahl des Patientenkollektivs: 3128 <ul style="list-style-type: none"> ▪ ohne Update: 2380 ▪ mit Update: 748 	3
	Breite Anwendung in der Praxis & Extrapunkte	Impact Assessment des Scores	Anwendung des Scores zeigt Benefit: keine
bisherige Nutzung und Anwendung des Scores: keine			0
Gesamtzahl des Patientenkollektivs: 0			0
Extrapunkte		Modifikation der erreichten Punktzahl: keine	0
Insgesamt			25,5 Punkte

Tabelle 11: Qualitätskriterien des original Scores Koivuranta 1997

3.7.5.2 Qualitätskriterien des vereinfachten Scores Koivuranta 1997

Kriterium	Untergruppe	Unterpunkt	Max. erreichte Punktzahl	
Generalisierbarkeit	Anzahl des Patientenkollektivs	Socrentwicklung: 1107	1,5	
Praktikabilität	Risiko-faktoren in dem Modell	Anzahl der RF in dem Modell: 5	2	
		Mehrheit der RF ohne technische Hilfsmittel mit ja/nein Antwort, mittlerer Zeitaufwand	2	
		Mehrheit der RF sind exakt definiert	2	
		RF ohne Kenntnis des Outcomes beurteilt: nein/ unklar	0	
	Risiko-berechnung/ Handhabung des Scores	keine Hilfsmittel nötig	3	
		vereinfachter Score vorhanden	1	
	Endpunkt des Scores	einzelne Endpunkte PN/PV/PR	1	
		Art und Weise der Endpunkterhebung: via Gespräch (persönliche Visite)	1	
	Präzision	Vorhersagegenauigkeit/ Beurteilung der Prädiktion mittels:	Diskriminationsfähigkeit (AUC der ROC-Kurve) <ul style="list-style-type: none"> Endwert (Intern & Extern) AUC 0,738 	2,5
			Kalibrierung <ul style="list-style-type: none"> Geradengleichung aller vereinfachten Scores Fläche: 5,81 	0,5
Kalibrierung <ul style="list-style-type: none"> Calibration Plot Hosmer-Lemeshow-Test 			1 1	

Validierung	Interne Validierung des Scores	Methode der Validierung: keine	0
		Gesamtzahl des Patientenkollektivs: 0	0
	Externe Validierung des Scores	Anzahl der Validierungen: 8	3
		Lokalisation der Validierungen: externe Validierung auf einem anderen Kontinent wie Modellentwicklung	3
		Gesamtzahl des Patientenkollektivs: 7979	3
Breite Anwendung in der Praxis & Extrapunkte	Impact Assessment des Scores	Anwendung des Scores zeigt Benefit: geringere PONV Inzidenz	1
		bisherige Nutzung und Anwendung des Scores <ul style="list-style-type: none"> ▪ Identifizierung von Hochrisikopatienten ▪ Vergleich von Studienteilnehmern/ Herstellung von Kohortengleichheit 	1
			Gesamtzahl des Patientenkollektivs: 926
	Extrapunkte	Modifikation der erreichten Punktzahl: extra für sonstige Maßzahlen	1
Insgesamt			32,5 Punkte

Tabelle 12: Qualitätskriterien des vereinfachten Scores Koivuranta 1997

3.8 Gesamtauswertung des Scores Apfel

3.8.1 Modellentwicklung mit interner Validierung

3.8.1.1 Apfel1998

Diese Studie beschreibt die Entwicklung und interne Validierung des originalen Apfelscores. Die Studie findet am Universitätsklinikum in Würzburg statt. Als Endpunkt dient postoperatives Erbrechen innerhalb der ersten 24 Stunden nach der OP.

Die Einschlusskriterien erfassen erwachsene, männliche und weibliche Patienten mit einem Body Mass Index (BMI) zwischen $> 15 / \text{kgm}^2$ und $< 40 \text{kg} / \text{m}^2$ sowie einer ASA-Klasse I-III. Alle Patienten unterziehen sich stationären, elektiven HNO-OPs.

Ausschlusskriterien sind bekannte Kontraindikationen für die in der Studie verwendeten Medikamente. Des Weiteren ist die prä- oder intraoperative Anwendung von Medikamenten mit möglichen antiemetischen Eigenschaften sowie eine Protokollverletzung (z. B. Verwendung von Propofol, Aufhebung der Muskelrelaxation) und unvollständige oder inkonsistente Daten nicht gestattet.

Die Stichprobengröße beträgt insgesamt 1137 Patienten. Diese sind aufgeteilt in einen Entwicklungsdatensatz (553 Patienten) und einen Validierungsdatensatz (584 Patienten). Es erleiden 245 Patienten PV. Davon sind 116 in der Entwicklungsgruppe und 129 in der Validierungskohorte.

Ein Kalibrierungsdiagramm ist vorhanden und die Geradengleichung lautet: $Y = 0,98x + 1,18$. Die Diskriminationsfähigkeit ergibt eine AUC von 0,77.

Das jeweilige Risiko eines Patienten an PV zu erkranken wird durch folgende Gleichung berechnet:

$$p(\text{PV}) = 1 / (1 + e^{-z})$$

mit $z = 1,28 * (\text{weibliches Geschlecht}) - 0,029 * (\text{Alter in Jahren}) - 0,74 * (\text{Raucher}) + 0,63 * (\text{Reisekrankheit oder PV in der Vorgeschichte}) + 0,26 * (\text{OP-Dauer in Stunden}) - 0,92$ [4]

3.8.1.2 Apfel1998a

Diese Veröffentlichung beschreibt eine Weiterentwicklung des Scores von Apfel, indem eine handliche Risikotabelle vorgestellt wird. Hier lässt sich das Erbrechensrisiko direkt in %-Werten ablesen (siehe Abbildung 5). Zusätzlich erfolgt eine Validierung des Scores an Patienten, die sich ebenfalls im Uniklinikum Würzburg operieren lassen. Diese Eingriffe finden jedoch in den Fachrichtungen Augenheilkunde und Chirurgie (Abdominal-, Gefäß-, Knochen-, Plastische Chirurgie) statt.

Weitere Einschlusskriterien beinhalten erwachsene Patienten und stationäre Eingriffe. Ausschlusskriterien sind perioperative prophylaktische Gabe von antiemetischen Medikamenten, Nachbeatmung der Patienten nach der OP sowie unvollständige oder inkonsistente Daten. Als Endpunkt dient postoperatives Erbrechen 24 h nach der OP, da es zur Erfassung von Übelkeit keinen akzeptierten Standard gibt.

Die gesamte Stichprobengröße umfasst 2220 Patienten. Hier werden auch die 1137 Patienten aus der HNO-Abteilung miteingefasst, die initial zur Scoreentwicklung herangezogen wurden. Nun erfolgt die Validierung an 1083 Patienten der Chirurgie und Augenklinik.

Für die Validierungskohorte gibt es ein Kalibrierungsdiagramm. Die Geradengleichung lautet $y=1,15x-0,02$. Die Diskriminationsfähigkeit des Modells wird nicht beschrieben.[5]

Dauer (h)	ohne anamnestische PÜ&E oder RK								mit anamnestischer PÜ&E oder RK								
	Nichtraucher				Raucher				Nichtraucher				Raucher				
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	
Männliche Patienten																	
Alter (J)	20	22	27	33	39	12	15	19	23	35	41	48	54	21	25	30	36
	30	18	22	27	32	9	12	15	18	29	35	41	47	16	20	25	30
	40	14	17	21	26	7	9	12	14	23	28	34	40	13	16	20	24
	50	11	14	17	21	5	7	9	11	19	23	28	33	10	12	15	19
	60	8	11	13	17	4	5	7	9	15	18	22	27	8	10	12	15
	70	6	8	10	13	3	4	5	7	11	14	18	22	6	7	9	12
	80	5	6	8	10	2	3	4	5	9	11	14	17	4	6	7	9
Weibliche Patienten																	
Alter (J)	20	51	57	64	69	33	39	46	52	66	72	77	81	48	55	61	67
	30	44	50	57	63	27	33	38	45	59	65	71	76	41	48	54	60
	40	37	43	50	56	22	27	32	38	52	59	65	70	34	40	47	53
	50	30	36	42	49	17	21	26	31	45	51	58	64	28	34	40	46
	60	25	30	35	42	13	17	21	25	38	44	51	57	23	27	33	39
	70	20	24	29	35	10	13	16	20	31	37	44	50	18	22	27	32
	80	15	19	24	28	8	10	13	16	26	31	37	43	14	18	22	26

Abbildung 5: Scorechart zur Erbrechenswahrscheinlichkeit in % [5]

3.8.1.3 Apfel1998b

Diese Publikation arbeitet mit denselben Patientendaten, die bereits unter Apfel 1998a geschildert wurden. Als Endpunkt dient PV 24 Stunden nach der Operation. Die Ein- und Ausschlusskriterien sind identisch zur Publikation von Apfel1998a.

Daten von insgesamt 1091 Patienten des Universitätsklinikums Würzburg aus den Abteilungen Chirurgie und Augenheilkunde werden analysiert. 542 Patienten dienen zur Validierung des bereits etablierten original Apfelscores. Die Daten der verbleibenden 549 Patienten (Evaluationsset) werden verwendet, um einen neuen Score zu entwickeln. Dieser setzt sich jedoch nicht durch.

Im Validierungsdatensatz erleiden 127 Patienten PV. Für die Validierungskohorte liegt ein Kalibrierungsdiagramm vor und die Geradengleichung lautet: $y=0,99x-2,56$. Die Diskriminationsfähigkeit beschreibt eine AUC von 0,769.[6]

3.8.1.4 Apfel1999

Diese Publikation beschreibt die Entwicklung und Kreuzvalidierung des vereinfachten, kombinierten Apfelkoivurantascores. Im weiteren Verlauf der Arbeit wird er lediglich vereinfachter Apfelscore genannt.

Die Originalscores von Koivuranta aus Oulu in Finnland und Apfel aus Würzburg werden vereinfacht. Anschließend werden die Originalscores und die vereinfachten Scores kreuzvalidiert. Diese setzen sich jedoch nicht durch. Abschließend wird mithilfe eines gemischten Datensatzes der kombinierte, vereinfachte Score entwickelt. Dieser Datensatz von 1040 Patienten besteht für eine gleiche Gewichtung der zwei Zentren aus 520 Patienten aus Oulu und 520 zufällig ausgewählten Patienten aus dem Würzburger Kollektiv.

Als Endpunkt wird PONV innerhalb von 24 Stunden nach der OP beschrieben. Hierzu werden die Patienten befragt und jede postoperative Übelkeit auf der 11-Punkte-Skala größer als Null berichtet. Auch die Anzahl der Episoden von postoperativem Erbrechen werden aufgezeichnet. Die Einschlusskriterien beinhalten Erwachsene, die sich stationären, elektiven Operationen unter Vollnarkose unterziehen. Der BMI der Patienten muss $> 15 \text{ kg/m}^2$ und $< 40 \text{ kg/m}^2$ sein. Es dürfen keine Kontraindikationen für die in der Studie verwendeten Medikamente vorliegen. Die Ausschlusskriterien umfassen die präoperative- oder intraoperative Verwendung von Antiemetika, unvollständige oder inkonsistente Daten, eine alleinige Regionalanästhesie oder diese in Kombination mit einer Vollnarkose.

Es liegen zwei Stichproben vor. In Oulu erleiden von 520 Patienten 289 Menschen PONV. Aus dem Würzburger Kollektiv bekommen von den 520 Patienten 689 PONV. Zum Schluss wird ein kombinierter Datensatz von 1040 Patienten erstellt.

Die Diskriminationsfähigkeit des kombinierten, vereinfachten Scores ergibt bei den 1040 Patienten beider Zentren eine AUC von 0,75. Eine Geradengleichung für die Kalibrierung dieses endgültigen Scores bei dem zusammengesetzten Patientenkollektiv liegt nicht vor.

Als Vorstufe wird folgende Gleichung für den kombinierten Score publiziert:

$$\text{Gleichung: } P = 1 / (1 + e^{-z})$$

Mit $z = -2,88 + 1,27^* (\text{weibliches Geschlecht}) + 0,65^* (\text{PONV oder MS in der Vorgeschichte}) + 0,72^* (\text{Nichtraucher}) + 0,78^* (\text{postoperative Opiode})$

Der finale vereinfachte Score beinhaltet folgende gleichgewichtete Risikofaktoren: weibliches Geschlecht, Reisekrankheit oder PONV in der Vorgeschichte, Nichtraucher, Einsatz postoperativer Opiode.

Die Wahrscheinlichkeit für das Auftreten von PONV in Abhängigkeit von der Anzahl der vorhandenen Risikofaktoren lässt sich folgender Tabelle entnehmen:[7]

Anzahl Risikofaktoren	PONV-Risiko (%)
0	10%
1	21%
2	39%
3	61%
4	79%

Tabelle 13: Wahrscheinlichkeit für das Auftreten von PONV: vereinfachter Apfelscore 1999[7]

3.8.2 Externe Validierung mit Update

3.8.2.1 Engel2006

Diese Veröffentlichung wurde bereits unter 3.1.2.1 ausführlich erläutert. Eine externe Validierung und Anpassung der Koeffizienten folgender Scores wird beschrieben: original Koivuranta, Apfel und Sinclair.

Die Risikofaktoren der Modelle werden bei der Überarbeitung beibehalten. Die Koeffizienten sowie der Achsenabschnitt werden angepasst und mit demselben Datensatz via Kreuzvalidierung getestet. Im Folgenden wird auf das überarbeitete Modell eingegangen. Als Endpunkt dient PONV 24 Stunden nach der Operation. Die Stichprobengröße des Validierungsdatensatzes beträgt 748 Patienten. Hiervon erleiden 84 PONV.

Für die Kalibrierung wird der Hosmer-LESmeshow-Test durchgeführt. Bezüglich der Diskriminationsfähigkeit ergibt sich eine AUC von 0,627 im Validierungsdatensatz für das upgedatete Modell.

Bei identischer Gleichung lauten die Werte nach dem Update wie folgt:

Gleichung: $P = 1 / (1 + e^{-z})$

Mit $z = -2,77 + 0,76 * (\text{weibliches Geschlecht}) + 0,94 * (\text{PONV oder MS in der Vorgeschichte}) - 0,91 * (\text{Raucher}) + 0,86 * (\text{postoperative Opiode})$ [26]

3.8.2.2 Rodseth2010

Dies ist eine externe Validierung des vereinfachten Apfelscores mit Update. Die PONV Vorhersagegenauigkeit des Apfelscore wird in einer Gruppe von Schwarzafrikanern vs. Kaukasiern getestet. Abschließend wird festgestellt, dass der Faktor Schwarzafrikaner das PONV Risiko senkt.

Die Studie findet an zwei Zentren in Südafrika statt: das Zentralkrankenhaus Inkosi Albert Luthuli in Durban und das St. Aidan's Hospital ebenfalls in Durban. Der Endpunkt ist festgelegt mit Übelkeit, Erbrechen oder Würgen sowie die Kombination der Ereignisse innerhalb von 24 Stunden nach der OP. Die Einschlusskriterien umfassen erwachsene, stationäre Patienten, die sich elektiven Operationen in allgemeiner, volatiler Anästhesie mit Propofolinduktion unterziehen. Die Ausschlusskriterien beinhalten nicht unmittelbare postoperative Extubation und intraoperatives Lachgas.

Die Stichprobengröße umfasst insgesamt 800 Patienten. Diese sind aufgesplittet in 400 Afrikaner und 400 Kaukasier (= Nicht-Afrikaner). Die Patienten werden gebeten, sich selbst als Afrikaner (indigener schwarzer Südafrikaner) oder Nichtafrikaner einzustufen. In dem Patientenkollektiv erleiden 288 Patienten PONV.

Für die Kalibrierung liegen keine genauen Daten vor. Die Diskriminationsfähigkeit des upgedateten Modells ergibt eine AUC von 0,67.

Der überarbeitete Score beinhaltet drei Risikofaktoren: nicht-afrikanische Ethnizität, weibliches Geschlecht und eine Vorgeschichte von Reisekrankheit oder PONV.[71]

3.8.2.3 Sarin2012

Diese Publikation wurde bereits unter Punkt 3.3.2 beschrieben. Das Hauptziel der Studie ist die Entwicklung eines eigenständigen Scoring-Systems. Zusätzlich erfolgt eine externe Validierung des refitted Sinclair-Modells, des originalen Apfelmodells und ein Update dieses Scores zum refitted Apfelmodell sowie eine Validierung des simplified Apfelrisikoscores. Auf das refitted Apfelmodell und den vereinfachten Apfelscore wird nun näher eingegangen.

Als Endpunkt wird PONV definiert. Das Ende des Beobachtungszeitraumes wird nicht publiziert. Die Stichprobengröße beträgt 2505 Patienten. Diese sind aufgeteilt in einen Entwicklungsdatensatz mit 1670 Patienten und den Validierungsdatensatz mit 835 Patienten. Der Hosmer-Lemeshow-Test für die Kalibrierung wird bei allen Scores durchgeführt. Zusätzlich liegt ein Kalibrierungsdiagramm vor, hier kann die Steigung abgelesen werden. Die AUC für den Validierungsdatensatz mit 835 Patienten beträgt 0,629 für das refitted Apfelmodell und 0,626 für den vereinfachten Apfelscore.

Die angepassten Werte für das „refitted Apfelmodell“ unter Beibehaltung der Gleichung lauten wie folgt:

$$\text{Gleichung: } P = 1 / (1 + e^{-z})$$

Mit $z = -3,48 + 1,19 * (\text{weibliches Geschlecht}) + 0,90 * (\text{PONV oder MS in der Vorgeschichte}) + 0,14 * (\text{Nichtraucher}) + 0,13 * (\text{postoperative Opiode})$ [75]

3.8.2.4 Sawatzky2014

Diese Studie beschreibt eine externe Validierung des vereinfachten Apfelscores mit Update. Ziel ist es, Prädiktoren für schweres PONV (≥ 2 Episoden) in der Herzchirurgie zu finden. Initial wurde der Apfelscore ohne den Risikofaktor Opiode zur Risikoeinschätzung der Population genutzt, da alle herzchirurgischen Patienten Opiode erhalten. Somit erfolgt die Identifizierung von zwei Gruppen: hohes Risiko von PONV ($n = 74$ Risikofaktorwert ≥ 3) und geringes Risiko für PONV-Gruppen ($n = 76$; Risikofaktorwert < 3). Abschließend wird für die spezielle herzchirurgische Patientenpopulation ein neuer Score namens cardiac surgery intensive care unit Instrument (CSICU), der einem Update des Apfelscores entspricht, vorgestellt.

Die Studie findet in den Abteilungen für Chirurgie, Anästhesie und Physiologie der Universität Manitoba in Winnipeg in Kanada statt. Der Endpunkt sind Patienten mit schwerem PONV ≥ 2 Episoden. Sie werden anhand der Daten der Krankenhausapotheke identifiziert, indem die Patienten, die Granisetron innerhalb von

24 Stunden nach der OP erhalten herausgefiltert werden. Die Einschlusskriterien umfassen stationäre Eingriffe bei Erwachsenen im Bereich der Herzchirurgie (Koronararterien-Bypass-Operation, Klappen-OP oder eine Kombination der beiden Eingriffe).

Die Stichprobengröße umfasst 150 Patienten. Hiervon erleiden 50 PONV.

Genauere Angaben über die Kalibrierungs- und Diskriminationsfähigkeit des upgedateten Modells werden nicht gemacht.

Das CSICU -Instrument um schweres PONV (≥ 2 Episoden) bei Herzchirurgischen Patienten zu detektieren hat als maximale Gesamtpunktzahl 7, wenn alle möglichen Risikofaktoren vorhanden sind. Die Risikofaktoren lauten: weibliches Geschlecht (2 Punkte), Vorgeschichte von PONV / Reisekrankheit (2 Punkte), Nichtraucher (1 Punkt), Extubiert bei Aufnahme in die Intensivstation für Herzchirurgie (CSICU cardiac surgery intensive care unit) (1 Punkt) und Verabreichung intraoperativer Steroide (1 Punkt). [76]

3.8.2.5 DaSilva2015

Diese Publikation beschreibt eine externe Validierung des vereinfachten Apfelscores mit Update. Das Hinzufügen des Risikofaktors Chemotherapie-induzierte Übelkeit und Erbrechen (CINV) zu dem Score verbessert die Vorhersage von PONV bei Patienten mit Krebs-OP.

Die Studie findet in der Abteilung für Anästhesie im Krebsinstitut des Staates Sao Paulo in Brasilien statt. Als Endpunkt dient PONV in den ersten 24 Stunden postoperativ. Dies kann durch ein Patientengespräch erfasst, in der Krankenakte dokumentiert sein, vom Pflegepersonal berichtet oder durch die Verabreichung antiemetischer Medikamente zur Symptomlinderung detektiert werden. Die Einschlusskriterien umfassen stationäre, elektive, mittlere bis große Krebsoperationen bei Erwachsenen. Ausgeschlossen werden Patienten, die sich nicht auf Portugiesisch ausdrücken oder in den ersten 24 Stunden nach der Operation nicht verständigen können. Weitere Ausschlusskriterien sind Orotrachealtuben, psychische Erkrankungen, Verwirrung, Unruhe, Delirium, unvollständige Daten, Tod und unerwartete Aufnahme auf der Intensivstation (ICU) vor Abschluss des 24-Stunden-Protokolls.

Die Stichprobengröße umfasst 1491 Patienten, davon erleiden 388 PONV.

Bezüglich der Kalibrierung wird der Hosmer-Lemeshow-Test durchgeführt. Die Diskriminationsfähigkeit des überarbeiteten Modells ergibt eine AUC von 0,705.

Das upgedatete Modell zur Vorhersage von PONV bei Krebs-OPs beinhaltet folgende Risikofaktoren: weibliches Geschlecht, Chemotherapie-induzierte Übelkeit und Erbrechen (CINV), Vorgeschichte von PONV oder Reisekrankheit, Nichtraucherstatus, voraussichtlicher postoperativer Opioidkonsum

Die Tabelle veranschaulicht die PONV Inzidenz abhängig von der Anzahl an Risikofaktoren.[19]

Anzahl Risikofaktoren	PONV-Risiko (%)
0	11,4%
1	16,2%
2	19,4%
3	32,1%
4	53,2%
5	66,6%

Tabelle 14: Wahrscheinlichkeit für das Auftreten von PONV: Score von Da Silva 2015[19]

3.8.2.6 Schaefer2015

Dieses Paper beschreibt eine externe Validierung mit Update des vereinfachten Apfelscores. Bei Operationen mit Xenon-basierter Anästhesie zeigen sich folgende drei Risikofaktoren des Apfelscores als signifikant und erhalten neue Odd Ratios: weibliches Geschlecht, junges Alter, lange Anästhesiedauer.

Es handelt sich hier um eine multizentrische Studie in Deutschland, folgende Häuser nehmen teil: Universitätsklinikum Düsseldorf, Evangelisches Waldkrankenhaus Berlin-Spandau, Berufsgenossenschaftliche Kliniken BergmannstrostHalle, Asklepios Klinik St. Georg Hamburg, Klinikum Links der Weser Bremen und Uniklinik RWTH Aachen.

Als Endpunkt dient Übelkeit, Erbrechen oder beides innerhalb von 24 Stunden nach der Operation. Hierfür erfolgt die Einsichtnahme in das Krankenblatt bezüglich einer postoperativen antiemetischen Medikation ("Notfallmedikation") und eine persönliche Befragung der Patienten. Die Einschlusskriterien umfassen Erwachsene der ASA-Klasse 1-2, die sich einer xenonbasierten Vollnarkose unterziehen. Ausgeschlossen

werden Patienten mit erhöhtem intrakraniellm Druck, Lungenkrankheit, Koronararterienkrankheit und eingeschränkter linksventrikulärer Funktion.

Die Stichprobe beinhaltet 488 Patienten. Hiervon erleiden 136 PONV.

Genauere Angaben zur Kalibrierung und Diskriminationsfähigkeit des überarbeiteten Scores werden nicht gemacht.

Eine vollständige Gleichung liegt nicht vor. Die drei folgenden Risikofaktoren mit angepassten Odds Ratios stellen das Modell dar: weibliches Geschlecht (Odds-Ratio 1,76), junges Alter (Odds Ratio 0,82 pro 10 Jahre) und längerer Anästhesiedauer (Odds Ratio 1,36 pro Stunde).[77]

3.8.2.7 Lee2015

Diese Veröffentlichung beschreibt eine externe Validierung des vereinfachten Apfelscores mit Update. Initial wird der Apfelscore in Singapur angewendet. Es wird geprüft, welche Risikofaktoren in dieser Population vorliegen und wie die tatsächliche Inzidenz von PONV mit dem vorhergesagten Risiko übereinstimmt. Abschließend wird der Risikofaktor Vollnarkose hinzugefügt.

Die Studie findet in der Yong Loo Lin School of Medicine der Universität von Singapur statt. Als Endpunkt dient PONV 24 Stunden nach der OP. Es wird die Intensität der Übelkeit (d.h. das Gefühl, sich übergeben zu wollen) und die Anzahl des Erbrechens erfasst. Die Einschlusskriterien umfassen elektive, stationäre Eingriffe bei Erwachsenen mit Vollnarkose und/oder Regionalanästhesie. Ausschlusskriterien beinhalten die unmittelbar postoperative Verlegung auf die Intensivstation, die Neurochirurgie oder die psychiatrische Station. Ebenfalls nicht berücksichtigt werden Patienten mit kognitiver Beeinträchtigung oder Schwierigkeiten mit der verbalen Kommunikation. Patienten, die eine andere Sprache als Englisch, Mandarin, chinesische Dialekte oder Malaiisch sprechen, werden nicht berücksichtigt.

Die Stichprobengröße umfasst 707 Patienten. Hiervon erleiden 271 PONV.

Angaben zu Kalibrierung und Diskriminationsfähigkeit des Scores werden nicht gemacht.

Es wird kein abschließender Score angegeben, lediglich neue Odds für die Risikofaktoren: weibliches Geschlecht (1,74), Nichtraucher (1,72), Vorgeschichte von PONV und/oder Reisekrankheit (OR 3,45), Gabe von postoperativen Opioiden (1,39), Vollnarkose (1,76).[53]

3.8.2.8 Klenke2018

Diese Publikation stellt eine externe Validierung des vereinfachten Apfelscores mit Update dar. Die Studie beschreibt, dass PONV eine signifikante Assoziation zu dem Genotyp CHRM3 rs2165870 SNP hat. Ein univarianter, paarweiser Vergleich der Probanden ergibt ein 30% höheres Risiko für PONV beim homozygoten AA-Genotypen. Heterozygote GA-Genotypträger weisen verglichen mit dem GG-Genotyp ein 20% höheres Risiko für PONV auf. Dies deutet auf einen Gendosis-Effekt hin. Somit wird der Risikofaktor „CHRM3 rs2165870 SNP mit jeweiligem Genotyp“ zum Modell hinzugefügt.

Die Studie findet am Universitätsklinikum Essen statt. Als Endpunkt wird Übelkeit, Würgen und Erbrechen zwei, sechs und 24 Stunden nach der OP aufgezeichnet. Das Modell sagt frühes PONV zwei bis sechs Stunden nach der OP voraus. Die Einschlusskriterien beinhalten stationäre, elektive Eingriffe aus den Bereichen Gynäkologie, HNO und Schilddrüsenoperationen. Ausgeschlossen werden Patienten mit Herzschrittmacher oder implantiertem Defibrillator, Patienten mit einer Nickel- oder Chromallergie und Patienten der ASA Klassifizierung > 3.

Die Stichprobengröße umfasst 454 Patienten. Hiervon erleiden 169 PONV.

Genaue Fakten zur Kalibrierung bzw. Diskriminationsfähigkeit des Modells werden nicht genannt.

Das abschließende Modell zeigt 5 signifikante Risikofaktoren, in Klammern jeweils das relative Risiko (RR): Alter (pro Jahrzehnt) (0,9), Nichtraucher (1,4), postoperative Opioidanwendung (1,3), hypotisches Induktionsmittel Thiopental oder Etomidate (1,4), CHRM3 rs2165870 SNP: GG (1), GA (1,5), AA (1,4). Eine Gleichung wird nicht veröffentlicht. [43]

3.8.3 Externe Validierung ohne Update

3.8.3.1 Eberhart1999

Diese Veröffentlichung beschreibt eine externe Validierung des originalen Apfelscores unter veränderten operativen und anästhesiologischen Rahmenbedingungen. Zusätzlich erfolgt die Erweiterung des Endpunktes von postoperativem Erbrechen (PE) auf PONV.

Die Studie findet am Universitätsklinikum Ulm statt. Der Endpunkt PONV wird als Übelkeit und/oder Würgen oder Erbrechen 24 Stunden nach der Operation definiert. Die

Einschlusskriterien erfassen Erwachsene, die sich stationären, elektiven, orthopädischen Operationen in Allgemeinanästhesie unterziehen.

Die Stichprobengröße beinhaltet 226 Patienten. Hiervon erleiden 44 PE und 75 PONV.

Die Diskriminationsfähigkeit des original Apfelscores mit Geradengleichung ergibt eine AUC von 0,72 für PONV und eine AUC von 0,73 für PE. Somit prädiziert der Score PONV ebenfalls sehr gut, obwohl er initial für PE entwickelt wurde.[22]

3.8.3.2 Eberhart2000a

Diese Veröffentlichung wurde bereits unter 3.2.2.2 beschrieben. Es handelt sich um eine externe Validierung dreiverschiedener Scores zur Vorhersage von PE und PONV. Die Scores sind von folgenden Autoren: Palazzo und Evans, der original Apfelscore aus dem Jahre 1998 und der vereinfachte Score von Koivuranta.

Das Studienkollektiv enthält 1444 Patienten. Davon erleiden PV 247 Patienten und PONV 540 Patienten.

Ein Kalibrierungsdiagramm liegt vor. Die Kalibrierungsgerade für PONV lautet: $Y=0,94x+13,4$. Die Diskriminationsfähigkeit des Scores für PONV ergibt eine AUC von 0,70. Für PV wird eine AUC von 0,73 beschrieben.[23]

3.8.3.3 Trope2000

Diese Publikation wurde unter Punkt 3.7.3.2 bereits dargestellt. Es handelt sich um die Entwicklung eines eigenständigen Scores ohne Validierung sowie eine externe Validierung des vereinfachten Apfel- und vereinfachten Koivurantascores.

Als Endpunkt dient PONV innerhalb von 24 Stunden nach der OP. Die Stichprobengröße umfasst 120 Patienten, von denen 72 an PONV erkrankten.

Die Kalibrierungs- und Diskriminationsfähigkeit wird nicht näher beschrieben.[86]

3.8.3.4 Apfel2002

Diese Veröffentlichung wurde unter 3.2.2.3 ausführlich beschrieben. Es werden sechs verschiedene Scores (vereinfachter Apfelscore 1999, Palazzo 1993, Koivuranta 1997, Sinclair 1999, Scholz 1999 und Gan 2000) in einer externen Validierung verglichen.

Von den 1566 Patienten der Studie erleiden 600 PONV. Ein Kalibrierungsdiagramm ist vorhanden. Die Geradengleichung lautet: $y=0,82x+0,01$. Für die Diskriminationsfähigkeit ergibt sich eine AUC von 0,68[8].

3.8.3.5 Traeger2003

Auf diese Studie wurde bereits im Unterpunkt 3.7.3.4 eingegangen. Die Modellentwicklung eines neuronalen Netzwerkes zur PONV-Vorhersage sowie eine externe Validierung der vereinfachten Scores von Koivuranta und Apfel werden dargestellt.

Als Endpunkt dient PONV 24 Stunden nach der OP. Die Stichprobengröße umfasst 400 Patienten für die Validierung des etablierten Scores.

Es existiert ein Kalibrierungsdiagramm sowie eine Geradengleichung. Diese lautet für das vereinfachte Apfelmodell: $y=0,71x+1$. Die Diskriminationsfähigkeit des Modells wird mittels der AUC von 0,66 beschrieben.[85]

3.8.3.6 Pierre2002

Diese Publikation ist unter Punkt 3.3.3.2 beschrieben und stellt eine externe Validierung des vereinfachten Apfelscores und des Sinclairscores dar.

Als Endpunkt wird PONV innerhalb der ersten 24 postoperativen Stunden klassifiziert. Die Stichprobengröße des Validierungssatzes beträgt 428 Patienten. Hiervon erleiden 212 Patienten PONV.

Für die Kalibrierung ist ein Kalibrierungsplot sowie eine Geradengleichung vorhanden. Sie lautet: $y = 1,08x - 0,07$. Bezüglich der Diskriminationsfähigkeit beträgt die AUC 0,71.[64]

3.8.3.7 Apfel2004

In dieser Publikation wird der vereinfachte kombinierte Apfelscore extern validiert. Er wird mit der alleinigen Betrachtung der Risikofaktoren PONV in der Vorgeschichte und dem Operationsgebiet verglichen. Die einzelnen Faktoren sind dem Score unterlegen und werden verworfen. Somit wird im Folgenden auf den vereinfachten Apfelscore eingegangen.

Als Endpunkt werden Übelkeit, emetische Episoden oder beides innerhalb der ersten 24 Stunden nach der OP erfasst. Die Einschlusskriterien umfassen stationäre Eingriffe bei

Erwachsenen aus den Bereichen: Ophthalmologie, HNO, Gynäkologie, Orthopädie sowie abdominelle und andere Operationen. Das Patientenkollektiv sind 1566 Patienten aus Ulm, die initial für die Publikationen Eberharts in den Jahren 1999 und 2000 erhoben wurden. Im weiteren Verlauf werden diese Daten von Apfel ebenfalls genutzt. Es erleiden 600 Patienten des Kollektivs PONV.

Die Diskriminationsfähigkeit des vereinfachten Apfelscores ergibt eine AUC von 0,68. Zusätzlich werden statistische Größen des Risikomodells für Scorewerte von zwei, drei und vier angegeben. Beispielsweise beträgt für einen Scorewert von drei die Sensitivität 58%, die Spezifität 70%, der positive prädiktive Wert 54,5% und der negative prädiktive Wert 72,8.[10]

3.8.3.8 Eberhart2004

Diese Veröffentlichung wurde bei Unterpunkt 3.2.2.4 beschrieben. Es handelt sich um eine externe Validierung des original sowie des vereinfachten Apfelscores an pädiatrischen Patienten. Insgesamt werden in dieser Publikation fünf Scores (Palazzo 1993, vereinfachter Score von Koivuranta 1997, Apfel 1998, Sinclair 1999, Apfel 1999) validiert.

Als Endpunkt dient postoperatives Erbrechen 24 Stunden nach der OP. Bezüglich der Scores von Apfel ergibt sich eine Stichprobengröße von 983 Kindern. Hiervon erleiden 326 PV.

Für die Kalibrierung liegt sowohl ein Kalibrierungsdiagramm für den vereinfachten Score mit PONV als auch für den Originalscore mit PV vor. Die Geradengleichung des vereinfachten Apfelscores für PONV lautet: $Y = 0,62x + 14$. Die Gleichung des Originalscores für PV ist: $Y = 0,51x + 18$. Die Diskriminationsfähigkeit wird durch die Fläche unter der Kurve gemessen und ergibt bei Vorhersage durch den Originalscore eine AUC von 0,59. Die Diskriminationsfähigkeit des vereinfachten Apfelmodells liefert eine AUC von 0,58.[24]

3.8.3.9 VandenBosch2005

Diese Veröffentlichung ist unter Punkt 3.7.3.6 ausführlich beschrieben und handelt von einer externen Validierung der Risikomodelle von Apfel (original und vereinfachter Score) sowie Koivuranta (original und vereinfacht).

Als Endpunkt wird PONV 24 Stunden nach der OP festgelegt. Die Stichprobengröße des Validierungsdatensatzes beträgt 1388 Teilnehmer. Davon erleiden 666 PONV.

Es gibt ein Kalibrierungsdiagramm für das original Apfelmodell. Eine vollständige Geradengleichung zur Beschreibung der Kalibrierung existiert nicht. Es wird nur die Steigung mit einem Wert von 0,39 genannt, nicht der Achsenabschnitt. Für das Originalmodell wird der Hosmer-Lemeshow-Test durchgeführt. Die Diskriminationsfähigkeit ergibt für den ursprünglichen Score eine AUC von 0,62 sowie für den vereinfachten Score 0,63.[87]

3.8.3.10 Weilbach 2006

Das Ziel dieser Publikation ist es, den vereinfachten Apfelscore bei Patienten mit hohem PONV-Risiko extern zu validieren.

Die Studie findet im St. Josefs-Krankenhaus in Cloppenburg statt. Als Endpunkt dient PONV 24 Stunden nach der OP. Dies wird mit standardisierten Fragebögen erfasst. Übelkeit wird auf einer numerischen 100-Punkte-Skala bewertet. Eingeschlossen werden Erwachsene, die sich stationär elektiven OPs unterziehen. Ausschlusskriterien sind Erbrechen vor der Anästhesie und schwere, postoperative Komplikationen.

Es werden 93 Patienten mit einem hohen präoperativen Risiko für PONV (Apfel Score III und IV) analysiert. Die Anzahl der an PONV Erkrankten und die Diskriminations- sowie Kalibrierungsfähigkeit wird nicht berichtet.[93]

3.8.3.11 Peng 2007

Diese Studie wird unter Punkt 3.7.3.7 ausführlich beschrieben. Sie handelt von der externen Validierung des vereinfachten Koivurantascores und des vereinfachten Apfelscores.

Als Endpunkt gilt PONV innerhalb von 24 Stunden. Die Stichprobengröße des Validierungsdatensatzes umfasst 430 Patienten. In der Validierungskohorte erleiden 104 Patienten PONV.

Die Diskriminationsfähigkeit wird durch den AUC-Wert von 0,624 angegeben. Als sonstige Maßzahlen für den Apfelscore werden die Sensitivität mit 59,6%, die Spezifität mit 61,0% sowie ein positiver prädiktiver Wert von 32,8% und ein negativer prädiktiver Wert von 82,5% angeführt.[62]

3.8.3.12 White2008

Diese Publikation ist eine externe Validierung des vereinfachten Apfelscores. Hier wird untersucht, ob der vereinfachte Apfelscore bei Hochrisikopatienten, die eine antiemetische Prophylaxe bekommen, aussagekräftig ist. Außerdem ist herauszufinden, ob die Zeitspanne für eine PONV Vorhersage von 24 auf 72 Stunden nach der OP erweitert werden kann. Es zeigt sich, dass ein hoher Apfelscore mit einer höheren Inzidenz von PONV in den ersten 24 Stunden nach der OP verbunden und nicht auf ein Intervall von 24 -72 Stunden erweiterbar ist.

Der Studienort ist das Southwestern Medical Center der Universität von Texas in Dallas. Als Endpunkte werden frühes (0 -24 h) und spätes (24 -72 h) PONV erhoben. Brechreiz wird definiert als eine beliebige Anzahl von Erbrechen oder Würgen in einer einzigen 5-minütigen Periode oder eine Folge von Erbrechen, die nicht durch eine 2-minütige Entspannungsphase gelindert wird. Bei Übelkeit wird der Patient gebeten, den Schweregrad anhand einer vierstufigen Likert-Skala zu quantifizieren. Die Einschlusskriterien umfassen Erwachsene, die selektiv, großen laparoskopischen oder plastischen OPs unter Vollnarkose unterziehen. Postoperativ erhalten sie Opioid-Analgetika und müssen mindestens einen Apfelscore von zwei aufweisen. Ausgeschlossen werden Patienten, wenn sie innerhalb von 24 Stunden nach ihrer OP antiemetische Medikamente erhalten, instabile Vorerkrankungen haben, innerhalb von 30 Stunden nach ihrer Operation ein Prüfpräparat benötigen oder in den letzten sechs Monaten Alkohol- oder Drogenmissbrauch hatten.

Die Stichprobengröße des Validierungsdatensatzes beinhaltet 130 Patienten. Die absolute Anzahl der Events wird nicht genannt. Die Kalibrierungs- bzw. Diskriminierungsfähigkeit wird nicht näher beschrieben.[94]

3.8.3.13 Roberts2010

Diese Studie ist eine externe Validierung des vereinfachten Apfelscores. Der vereinfachte Apfelscore wird bzgl. der prädiktiven Vorhersagekraft mit dem Bauchgefühl der Anästhesisten verglichen. Die Anästhesisten werden gebeten, die Patienten in drei Risikogruppen einzuteilen. Der vereinfachte Apfelscore hat eine bessere Vorhersagekraft als die subjektive Einschätzung der Anästhesisten. Diese unterschätzen das PONV Risiko.

Diese Untersuchung findet am Liverpool Women's Hospital in Großbritannien statt. Als Endpunkt dient postoperative Übelkeit und Erbrechen, wobei ein genauer Zeitpunkt der Erhebung nicht genannt wird. Die Einschlusskriterien umfassen erwachsene Frauen,

die in Allgemeinanästhesiegynäkologische Operationen bekommen. Die Stichprobengröße umfasst 125 Frauen. Die Anzahl der Patienten, die an PONV erkranken, sowie die Diskriminations- und Kalibrierungsfähigkeit werden nicht genannt.[70]

3.8.3.14 Pluta2017

Bei dieser Publikation handelt es sich um eine externe Validierung des vereinfachten Apfelscores.

Die Studie findet an dem Universitätsklinikum in Katowice in Polen statt. Als Endpunkt dient PONV. Der genaue Beobachtungszeitraum wird nicht definiert. Die Einschlusskriterien umfassen Erwachsene, die sich v.a. gastrointestinalen Operationen unterziehen. Es nehmen insgesamt 101 Patienten an der Studie teil. Hiervon erleiden 9 PONV. Bezüglich der Diskriminationsfähigkeit ergibt sich eine AUC von 0,734. [66]

3.8.3.15 Wu2015

Diese Publikation wurde bereits unter dem Punkt 3.2.2.5 beschrieben, da es sich um eine externe Validierung ohne Update von fünf Scoringsystemen in Taiwan handelt. Sie lauten: original Apfel- und vereinfachter Apfelscore, original Koivuranta- und vereinfachter Koivurantascore sowie Palazzo & Evansscore.

Die Stichprobengröße beträgt 992 Patienten. Hiervon erkrankten 416 Patienten.

Für die Kalibrierung wird der Hosmer-Lemeshow-Test durchgeführt. Die Diskriminationsfähigkeit beschreibt eine AUC von 0,668 für den original Apfelscore und 0,614 für den vereinfachten.

Zusätzlich werden als Maßzahlen zur Modellklassifizierung folgende Werte genannt: Sensitivität (68% Originalscore/ 40,7% vereinfachter Apfelscore), Spezifität (60,5% Originalscore / 78,5 % vereinfachter Apfelscore), positiver prädiktiver Wert (55,5% Originalscore/ 57,7 % vereinfachter Apfelscore), negativer prädiktiver Wert (72,3% Originalscore / 64,6 % vereinfachter Apfelscore), Youden's Index (28,5m% Originalscore/ 19,1% vereinfachter Apfelscore). [97]

3.8.4 Impact Assessment

3.8.4.1 Rüschi1999

Diese Studie zählt zum Impact Assessment des vereinfachten Apfelscores. Ein Patientenkollektiv wird in drei Gruppen eingeteilt: Dolasetron, Metoclopramid und Placebo. Der Apfelscore dient dazu, nachzuweisen, dass die Medikamente wirksam sind, da die tatsächliche PONV-Inzidenz niedriger ist als das vorhergesagte Risiko. Des Weiteren zeigt der Score, dass die Gruppen bezüglich der Risikofaktorenvergleichbar sind.

Die Studie findet an der Klinik für Anästhesiologie der Universität Kiel statt. Als Endpunkt dient PONV im Zeitraum bis 24 Stunden nach der OP. Die Einschlusskriterien umfassen erwachsene Patientinnen der ASA-Klasse 1 und 2, die sich stationären, elektiven diagnostischen Pelviskopien unterziehen. Es gibt diverse Ausschlusskriterien wie Schwangerschaft, Stillzeit, Allergien gegen eine der eingesetzten Substanzen, gastrointestinaler Infekt innerhalb der letzten zwei Wochen, Antiemetikaeinnahme innerhalb der letzten 24 Stunden, extremes Unter- oder Übergewicht und Kommunikationsschwierigkeiten.

Die Stichprobengröße beinhaltet 97 Patientinnen, die in drei Gruppen aufgeteilt werden: Dolasetron 32, Placebo 32 und Metoclopramid 33. Insgesamt erleiden 47 Patientinnen PONV.

Abschließend wird publiziert, dass mit Hilfe des Scores die Wirkweise von Medikamenten sowie die homogene Zusammensetzung von Patientenkollektiven überprüfen werden kann.[72]

3.8.4.2 Thomas2002

Dieses Paper wurde unter Punkt 3.2.3.1 ausführlich beschrieben und gehört zu der Gruppe Impact Assessment.

Während eines RCTs zu den Effekten von kombinierten Antiemetika werden Patientengruppen durch Scoringssysteme verglichen. Es wird das PONV-Risiko mit folgenden Scores berechnet: vereinfachter Koivuranta, Palazzo und Evans, Sinclair und vereinfachter Apfel. Das Risiko für PV wird mit den Scores von Koivuranta und dem original Apfelscore ermittelt. Abschließend wird die Übereinstimmung zwischen den Scoringssystemen bewertet. Diese war nicht zufriedenstellend.[83]

3.8.4.3 Biedler2004

Diese Publikation zählt zur Gruppe des Impact Assessments vom vereinfachten Apfelscore. Ziel der Untersuchung ist es, mit Hilfe des vereinfachten Scoringsystems die Durchführbarkeit und Wirksamkeit einer risikoadaptierten PONV-Prophylaxe zu bewerten. Die Studie findet an drei deutschen Zentren statt: Universitätskliniken des Saarlandes, Uniklinikum Würzburg und Universitätsklinikum Aachen.

Als Endpunkt gilt PONV innerhalb der ersten 24 Stunden nach der OP. Eine emetische Episode wird als Erbrechen oder Würgen, d.h. unproduktive Emesis, definiert. Die Anzahl der emetischen Episoden, die Häufigkeit von Übelkeit sowie die Notwendigkeit einer antiemetischen Behandlung mit Ondansetron oder eines Rescue-Antiemetika werden erfasst. Eingeschlossen werden 162 Erwachsene der ASA Klasse 1 bis 3, die für eine elektive, stationäre Operation vorgesehen sind. Ausgeschlossen werden Patienten, die an einer anderen klinischen Studie innerhalb der letzten 30 Tage teilnehmen, bekannte Überempfindlichkeiten gegen Ondansetron haben oder an Phenylketonurie leiden. Zudem sind Fehlfunktionen der gastrointestinalen Motilität oder Diabetes Ausschlusskriterien. Des Weiteren werden Patienten mit schweren inneren oder neurologischen Erkrankungen, akuten lebensbedrohlichen Zuständen oder Übelkeit und Erbrechen nicht berücksichtigt. Zusätzliche Ausschlusskriterien sind eine antiemetische Behandlung innerhalb der letzten 24 Stunden vor der OP, Alkohol- oder Drogenmissbrauch in der Vergangenheit sowie Schwangerschaft und Stillzeit.

Vor der OP werden die Patienten mit null bis einen Risikofaktor in die Gruppe mit geringem Risiko eingeteilt und erhalten keine Prophylaxe. Als hohes Risiko gelten Patienten mit zwei bis vier Risikofaktoren. Sie erhalten 4 mg Ondansetron intravenös. Insgesamt haben während der 24 Stunden 58 Patienten PONV (36,5%). Dies ist signifikant niedriger als die erwartete PONV-Inzidenz laut Apfelscore von 47,2%. Somit zeigt die Verwendung des vereinfachten Risikoscores mit Prophylaxe-Strategie eine allgemeine Reduktion der PONV-Rate, indem die Inzidenz von PONV in der Hochrisikogruppe verringert wird. Zusätzlich sind über 85% der Patienten und 80% des Pflegepersonals mit dem antiemetischen Management zufrieden.[14]

3.8.4.4 Pierre2004

Diese Veröffentlichung gehört zum Impact Assessment des vereinfachten Apfelscores. Es wird überprüft, ob eine risikoadaptierte PONV-Prophylaxe die Inzidenz senkt. Dies bestätigt sich.

Die Studie findet am Universitätskrankenhaus von Toulouse in Frankreich statt. Als Endpunkt dient PONV, wenn ein Patient innerhalb der ersten 24 postoperativen Stunden Übelkeit oder Erbrechen schildert. Die Einschlusskriterien beinhalten Erwachsene, die sich stationären OPs unter Vollnarkose aus den Bereichen HNO oder Gynäkologie unterziehen. Ausgeschlossen werden Revisions-OPs innerhalb der ersten 24 Stunden wegen Hämatomen, Protokollverletzungen sowie Patienten mit präoperativ antiemetischer Medikation.

Die Stichprobengröße umfasst 428 Patienten. Davon erleiden 62 PONV.

Die Gesamtinzidenz von PONV in demselben Krankenhaus sinkt von 49,5% ohne Score Anwendung auf 14,3% nach Etablierung des Scores mit Prophylaxe. Die prophylaktische antiemetische Strategie reduziert die Inzidenz von PONV in der mittleren Risikogruppe um etwa 50% und in der Hochrisikogruppe um 75%. [65]

3.8.4.5 Rusch2005

Diese Veröffentlichung wird dem Impact Assessment des vereinfachten Apfelscores zugeordnet. Das Ziel der Studie ist es, die Wirksamkeit von prophylaktischem Ondansetron allein mit Ondansetron plus Dexamethason bei Patienten mit hohem PONV-Risiko zu vergleichen. Aus ethischen Gründen wird keine Placebogruppe erstellt. Stattdessen wird die tatsächliche Inzidenz mit der errechneten Inzidenz des vereinfachten Apfelscores verglichen.

Mit Hilfe des Scores werden Patienten in eine risikoarme Gruppe und eine Hochrisikogruppe unterteilt. Allein die Hochrisikogruppe erhält Antiemetika, entweder nur Ondansetron oder Ondansetron plus Dexamethason. Bei Patienten, die beide Medikamente erhalten, ist die durchschnittliche Inzidenz von PN, PV und PONV sowie die durchschnittliche, maximale Intensität der Episoden signifikant geringer. Zusätzlich besteht während des Beobachtungszeitraums ein geringerer Bedarf an notfallmäßiger antiemetischer Therapie. Vor allem die Hochrisikogruppen profitieren von der Prophylaxe, da die tatsächliche Inzidenz beider Gruppen signifikant geringerausfällt als vorhergesagt.

Als Endpunkt wird postoperative Übelkeit, postoperatives Erbrechen und PONV für den Zeitraum null bis 24 Stunden nach Vollnarkose festgelegt. Die Einschlusskriterien beinhalten Erwachsene, die sich stationären, elektiven Eingriffen unter Vollnarkose unterziehen. Es gibt diverse Ausschlusskriterien: Allergie gegen verwendetes Arzneimittel, schwere Beeinträchtigung der Darmmotilität, insulinabhängiger Diabetes mellitus und Phenylketonurie. Patienten, die innerhalb von 24 Stunden vor Studieneintritt Übelkeit oder Erbrechen, eine antiemetische Behandlung oder systemische Behandlung mit Steroiden erhalten, werden nicht berücksichtigt. Drogenmissbrauch, Schwangerschaft oder Stillzeit und Teilnahme an anderen klinischen Prüfungen innerhalb von 30 Tagen vor Studieneintritt sind ebenso Ausschlussgründe.

Die Stichprobengröße umfasst 453 Patienten. Hiervon erleiden 106 PONV. [73]

3.8.4.6 Kredel2006

Diese Publikation zählt zum Impact Assessment des vereinfachten Apfelscores. Es wird untersucht, ob das Wetter oder die Mondphase Einfluss auf die Inzidenz von PONV hat. Das individuelle PONV-Risiko wird anhand des Apfelscores berechnet. Zudem wird beobachtet, ob es zu Abweichungen der Inzidenz kommt, die darauf zurückzuführen sind. Dies bestätigt sich nicht. Somit dient der bereits etablierte Apfelscore zum Ausschluss potentieller Risikofaktoren.

Die Studie findet am Universitätsklinikum in Würzburg statt. Als Endpunkt wird PONV 24 Stunden nach der OP definiert. Die Einschlusskriterien beinhalten Erwachsene, die sich stationären, elektiven Eingriffen aus den Bereichen Urologie, Kopf-Hals oder allgemeinen Operationen unterziehen. Ausschlusskriterien sind Notoperationen sowie die Notwendigkeit einer postoperativen Beatmung.

Als Stichprobe werden 1801 Patienten rekrutiert. Hiervon erleiden 555 PONV. [50]

3.8.4.7 Kapoor2008

Dieses Paper zählt zum Impact Assessment des vereinfachten Apfelscores.

Es werden zwei Scoringssysteme, der vereinfachte Apfelscore und ein selbstentwickeltes Modell aus 16 Fragen, zeitlich versetzt in einem Krankenhaus angewendet und verglichen. Abhängig von dem erreichten Scorewert wird eine Therapieempfehlung ausgesprochen. Null bis maximal drei Medikamente der im folgenden aufgelisteten Medikamente sollen appliziert werden: Scopolamin 1,5 mg

transdermal präoperativ, Dexamethason 4–8 mg i.v. präoperativ, Prochlorperazin 5 oder 10 mg i.v. oder i.m. am Ende der Operation oder TIVA. Im Anschluss wird der Prozentsatz der ausgefüllten Formulare, diePONV-Inzidenz, die Anzahl der PONV-Risikofaktoren, die Patientendaten und die Angemessenheit der Prophylaxe bewertet.

Die Studie findet am Saint Barnabas Medical Center in Livingston in den USA statt. Als Endpunkt wird das Auftreten von PONV bis 48 Stunden nach der OP festgelegt. PONV wird als dokumentierte Episode von Übelkeit oder Erbrechen bei gleichzeitiger, medizinischer Intervention definiert. Eingeschlossen werden Erwachsene und Kinder. Die Stichprobengröße umfasst insgesamt 100 Patienten, wobei 50 in der Gruppe vereinfachter Apfelscore und 50 in der Gruppe hauseigener Score eingeteilt sind.

Das Risikostratifizierungsformular wird für 38% der Patienten mit dem hauseigenen Score und bei 64% der Patienten nach dem Apfelscore ordnungsgemäß ausgefüllt. Ein großer Prozentsatz der Patienten in beiden Gruppen erhält weniger prophylaktische Interventionen als empfohlen. Dies tritt jedoch bei deutlich mehr Patienten auf, die nach dem Apfelscore beurteilt werden. Die Häufigkeit von PONV unterscheidet sich nicht signifikant zwischen den Gruppen und die Behandlungsstrategien sind ähnlich.[38]

3.8.4.8 Franck2010

Diese Veröffentlichung wird dem Impact Assessment des vereinfachten Apfelscores zugeschrieben. Das Risikomodell wird in Standard Operating Procedures (SOPs) mit risikoadaptierter Antiemetikagabe eingepflegt. Als Ergebnis zeigt sich, dass die effektive PONV-Inzidenz unter der Vorhergesagten liegt.

Die Studie findet in der Klinik für Anästhesiologie der Charité in Berlin statt. Als Endpunkt gilt die Gesamt-PONV-Inzidenz über 24 Stunden. Eingeschlossen werden Erwachsene und Kinder ab 14 Jahren, die in Allgemeinanästhesie operiert und während der Regelarbeitszeit im Aufwachraum behandelt werden. Die Stichprobengröße umfasst 2729 Patienten.

Gemäß der Risikoklassifizierung werden 725 Patienten (26,6%) in die Gruppen mit niedrigem Risiko, 1050 Patienten (38,5%) in die mittlere Kohorte und 954 Patienten (35,0%) in die Hochrisikogruppe eingeteilt. Ein SOP-konformes Regime kommt bei 668 Patienten (92,1%) der Niedrigrisikogruppe, 373 Patienten (35,5%) der Gruppe mit mittlerem Risiko und 177 Patienten (18,6%) der Gruppe mit hohem Risiko zur Anwendung. Die abschließende Analyse zeigt eine errechnete Inzidenz in Gegenüberstellung zu der gemessenen Inzidenz: in der niedrig Risikogruppe von

17,6% vs.18,1% sowie in der mittleren Risikogruppe von 33,7% im Verhältnis zu 22,4% und in der Hochrisikogruppe 47,9% vs.45,5%.

Als Schlussfolgerung wird veröffentlicht, dass Patienten mit mittlerem und hohem Risiko durch die Einhaltung der SOP eine niedrigere PONV-Inzidenz als erwartet haben. Allerdings wird die klinikintern empfohlene Vorgehensweise unzureichend umgesetzt. Zudem ist zu hinterfragen, ob sich eine generelle PONV-Prophylaxe unter Berücksichtigung der Implementierungsproblematik als vorteilhaft erweisen wird.[27]

3.8.4.9 Latz2011

Diese Publikation gehört in die Rubrik Impact Assessment für den vereinfachten Apfelscore und benennt Kraniotomie als weiteren Risikofaktor.

Ziel der Arbeit ist es, den Einfluss der Operation, hier Kraniotomie, auf PONV zu bewerten. Es wird der vereinfachte Apfelscore verwendet, um das Patientenrisiko für PONV vorherzusagen. Dieses berechnete Risiko wird dann mit der tatsächlichen Inzidenz von PONV in der Patientengruppe verglichen.

Der Studienort ist die Universitätsklinik in Mainz. Als Endpunkt wird PONV in den ersten 24 Stunden nach der Operation festgelegt. Übelkeit wird als unangenehmes Gefühl im Zusammenhang mit dem Drang, zu erbrechen, definiert. Erbrechen wird als Ausstoßen von Mageninhalt festgelegt. Würgen wird wegen seiner pathophysiologischen Ähnlichkeit wie Erbrechen behandelt. Einschlusskriterien sind erwachsene Patienten, die sich stationären, elektiven supratentoriellen oder infratentoriellen Kraniotomien unterziehen. Ausgeschlossen werden Patienten unter 18 Jahren, Schwangere, Notfalloperationen und Patienten mit ASA-Klasse 5. Ebenfalls nicht berücksichtigt werden Eingriffe bei akustischem Neurinom, transsphenoidale Chirurgie und Verfahren, die nur eine Bohrung erfordern. Falls eine mechanische Beatmung > 2 Stunden nach Ende der Operation nötig ist werden die Patienten nicht berücksichtigt. Antiemetikagabe, unvollständige Daten, TIVA mit Propofol und keine regelmäßige Kommunikationsfähigkeit des Patienten sind auch Ausschlusskriterien. Die Stichprobe beinhaltet 229 Patienten. Davon erleiden 107 PONV.

Die beobachtete Inzidenz von PONV innerhalb von 24 Stunden nach der Operation ist höher (47%) als die mit dem Apfel-OP-unabhängigen Risiko-Score (31%) vorhergesagte Inzidenz. Ein möglicher Grund ist, dass intrakranielle Operationen einen zusätzlichen und unabhängigen Risikofaktor darstellen. Abschließend wird

veröffentlicht, dass Patienten, die sich einer Kraniotomie unterziehen als Risikopatienten für PONV gewertet werden.[52]

3.8.4.10 Kooji2012

Diese Veröffentlichung gehört zum Impact Assessment des vereinfachten Apfelscores. Es wird ein automatisiertes Erinnerungssystem für das PONV-Management, das den vereinfachten Risikoscore von Apfel enthält, getestet.

Der Studienort ist das akademische medizinische Zentrum der Abteilung für Anästhesiologie in Amsterdam. Als primärer Endpunkt dient PONV innerhalb der ersten 24 Stunden nach der OP. Sekundäre Endpunkte sind die Inzidenz von frühem und spätem PONV. Frühes PONV besteht während der Periode im Aufwachraum und spätes PONV danach bis zu 24 Stunden. Die Einschlusskriterien beinhalten Erwachsene, die sich einer elektiven nicht-kardialen Operation sowohl stationär als auch ambulant unter Vollnarkose unterziehen. Ausgeschlossen werden Schwangere, Patienten mit Allergien gegen Antiemetika sowie Patienten, die nicht fähig sind, zu kommunizieren. Antiemetische Therapie vor der Operation ist ein Ausschlusskriterium.

Es gibt zwei Zeiträume: die Kontrollperiode ohne die Verwendung des Erinnerungssystems und eine Phase mit Supportsystem, die Unterstützungsperiode. Das Support System schreibt ab ein bis drei Risikofaktoren eine konkrete Prophylaxe vor. Die Anästhesisten nehmen die Empfehlungen an.

Die Stichprobengröße umfasst insgesamt 2662 Patienten, die in einen Kontrollzeitraum mit 981 und in die Unterstützungsphase mit 1681 Patienten aufgeteilt sind. An PONV erkrankten insgesamt 642 Patienten. Diese sind aufgesplittet in 264, die während dem Kontrollzeitraum erkrankten, und 378 in der Unterstützungsphase.

Als Ergebnis zeigt sich eine Reduktion der Gesamt-PONV-Inzidenz in der Hochrisikopatientengruppe, in der die PONV-Inzidenz von 47% auf 30% zurückgeht. Darüber hinaus verringert sich die Häufigkeit der Inzidenzen, wobei die Häufigkeit von PONV im Kontrollzeitraum zwischen 18% und 36% und im Interventionszeitraum zwischen 17% und 27% liegt. Die Gesamtanwendung von Dexamethason, Granisetron und Metoclopramid unterscheidet sich nicht zwischen der Interventions- und Kontrollperiode. Im Kontrollzeitraum wird eine Verringerung des Droperidol-Einsatzes beobachtet. Die Gabe von Prophylaxemedikamenten an Hochrisikopatienten erhöht sich mit Hilfe des Supportsystems für Dexamethason und Granisetron signifikant. Im Gegensatz dazu erhalten Patienten mit niedrigem Risiko während der Interventionsphase weniger PONV-Prophylaxe.

Das wichtigste Ergebnis dieser Studie ist, dass ein automatisiertes patientenspezifisches Erinnerungssystem in der täglichen Routine eine verringerte Inzidenz von PONV durch eine bessere Einhaltung der Leitlinien herbeiführt.[47]

3.8.4.11 Ahmed2014

Diese Veröffentlichung zählt zu Impact Assessment des vereinfachten Apfelscores. Es handelt sich um eine eintägige Momentaufnahme, in der im Rahmen einer Umfrage folgende Punkte erfasst wurden: der errechnete Apfelscore jedes Patienten, die verwendeten intraoperativen Antiemetika, PONV und Rettungsantiemetika.

Die Studie findet am Royal Alexandra Hospital in Paisley in Großbritannien statt. Als Endpunkt dient PONV. Ein Zeitraum wird nicht angegeben. Eingeschlossen werden Erwachsene, die sich Operationen aus folgenden Bereichen unterziehen: Gynäkologie, HNO, Orthopädie, Allgemeinchirurgie, Traumatologie und Urologie. Die Stichprobengröße umfasst 27 Patienten. Hiervon erleiden zwei PONV.

Als Fazit zeigt sich, dass die Gesamtinzidenz von PONV bei dieser Datenerhebung viel niedriger (7,4%) als die national angegebene (30%) ist. Im Text wird angemerkt, dass der antiemetische Gebrauch zunehme, wenn der Apfel-Risiko-Score angewendet wird. Es ist jedoch fraglich, ob mehr Antiemetika bei dieser geringen Inzidenz angemessen sind.[3]

3.8.4.12 Kaneda2004

Diese Studie beschreibt das Impact Assessment des vereinfachten Apfelscores.

Sie findet in den Kliniken für Anästhesiologie und Mund-, Kiefer- Gesichtschirurgie der Universität von Osaka in Japan statt. Als Endpunkt dient PONV 24 Stunden nach der OP. Die Einschlusskriterien umfassen Erwachsene und Kinder ab 11 Jahren, die sich mund- und kieferchirurgischen Eingriffen unterziehen.

Die Stichprobe beinhaltet insgesamt 57 Patienten. Diese werden in drei Gruppen aufgeteilt: Gruppe C mit 20 Patienten ohne antiemetische Behandlung, Gruppe D mit 19 Patienten und intravenös verabreichtem Droperidol (0,025 mg/kg), sowie Gruppe M mit 18 Teilnehmern und intravenös verabreichtem Metoclopramid (0,15 mg/kg).

Dann wird die PONV-Inzidenz von Gruppe C mit dem vorhergesagten Ergebnis anhand des Apfelscores verglichen. Die tatsächliche Inzidenz von PONV bei Patienten mit einem Risikofaktor beträgt 20% vs. 29% in der Vorhersage. Bei zwei Risikofaktoren liegt die Inzidenz bei 75% und ist höher als die vorhergesagten 39%. In dieser Studie

werden aus dem bestehenden Modell nur die Risikofaktoren Frauen und Nichtraucher bestätigt.

Abschließend werden weibliche Nichtraucherinnen bei oralen oder maxillofazialen Eingriffen als Risikopatienten eingestuft, da sie im Vergleich zur Allgemeinchirurgie häufiger als vorhergesagt an PONV erkranken.[37]

3.8.4.13 Pospiech 2014

Diese Veröffentlichung zählt zu Impact Assessment des vereinfachten Apfelscores.

Die Studie findet an der Universitätsklinik Saint Luc in Brüssel in Belgien statt. Als Endpunkt wird PONV in der unmittelbaren, postoperativen Phase im Aufwachraum definiert. Die Einschlusskriterien beinhalten stationäre, elektive und gynäkologische Eingriffe. Ausgenommen sind Brustoperationen bei Erwachsenen.

Die Stichprobe umfasst insgesamt 612 Patienten. Diese werden aufgeteilt in Gruppe 1 „treatment as usual“ mit 318 Patienten und Gruppe 2 mit 294 Patienten. Hier wird bereits der vereinfachte Apfelscore durch die Anästhesisten berechnet. Es gibt keine Handlungsanweisungen oder ähnliches. Das Team wird lediglich für das Thema sensibilisiert. Nach der Umsetzung des vereinfachten Apfelscores steigt die präventive Medizin um das Zweifache. Dementsprechend erhalten im Jahr 2011 ohne Score 82 Patienten (25,8% von 318) im Vergleich zum Jahr 2012 mit Apfelscore 166 Patienten (56% von 294) eine Prophylaxe. Des Weiteren werden zweimal weniger Medikamente zur PONV-Notfallbehandlung eingesetzt. Die PONV Rescue-Medikation erhalten im Jahre 2011 150 Patienten 15,7% von 318 im Verhältnis zu 24 Patienten 8,2% von 294 im Jahre 2012.

Als Fazit wird publiziert, dass der vereinfachte Apfelscore für alle Patienten während der Prämedikation berechnet werden sollte, um das Team für die PONV-Präventivmedizin zu sensibilisieren.[67]

3.8.4.14 Sundqvist 2014

Diese Publikation gehört zu Impact Assessment des vereinfachten Apfelscores.

Die Studie findet an der Universität Umeå in dem Krankenhaus Sundsvall in Schweden statt. Als Endpunkt dient PONV. Dieses wird definiert als Übelkeit oder Erbrechen bzw. Würgen. Der Beobachtungszeitraum zieht sich bis zum dritten postoperativen Tag. Die Einschlusskriterien umfassen Erwachsene, die sich stationären,

laproskopischen, bariatrischen OPs in vollständiger i.v. Anästhesie unterziehen. Die Stichprobengröße umfasst 39 Patienten. Davon erleiden 28 PONV.

Ziel der Studie ist die Untersuchung der PONV-Inzidenz sowie die Wirksamkeit der PONV-Prophylaxe bei bariatrischen Patienten. Die Prophylaxe wird unabhängig von der Berechnung des Apfelscores für jeden Patienten nach individuellem Risiko und lokaler Routine verabreicht. Eine Prophylaxe gilt als optimal, wenn die Anzahl der verabreichten Prophylaxemedikamente um eins größer als der Apfelscore ist.

Es wird kein signifikanter Unterschied in der Inzidenz von PONV bei denjenigen gefunden, die eine suboptimale PONV-Prophylaxe im Vergleich zu Patienten, die eine optimale oder mehr als optimale PONV-Prophylaxe erhalten. Als Schlussfolgerung wird publiziert, dass Patienten mit laparoskopischen, bariatrischen Operationen eine hohe Inzidenz von PONV mit einem relativ hohen Grad an Beschwerden haben. Somit kann die derzeitige Prophylaxe- und Risikobewertung von PONV für Patienten, die sich einer bariatrischen Operation unterziehen, möglicherweise nicht geeignet sein.[82]

3.8.4.15 Agarkar2015

Diese Veröffentlichung gehört zu Impact Assessment des vereinfachten Apfelscores.

Die Studie findet im Tata Memorial Hospital Parel in Mumbai in Indien statt. Der primäre Endpunkt ist die PONV-Inzidenz 24 Stunden nach der OP. Der sekundäre Endpunkt ist der Schweregrad von Übelkeit sowie die Notwendigkeit von Notfallmedikamenten. Der Beobachtungszeitraum wird in zwei Intervalle, früh 0-6 Stunden und spät 6-24 Stunden nach der OP, unterteilt. Erbrechen wird als Ausstoßen des Mageninhalts und Würgen als unfreiwilliger, nicht produktiver Versuch, zu erbrechen, definiert. Übelkeit wird festgelegt, als der Wunsch zu erbrechen. Die Einschlusskriterien beinhalten Erwachsene mit mindestens zwei Risikofaktoren des Apfelscores und ASA-Status I - II, die sich Brust-, Parotis-, Schilddrüsen- oder gynäkologischen Operationen unterziehen.

Ausschlusskriterien sind schwangere bzw. menstruierende Patientinnen, Patienten mit gastroösophagealem Reflux in der Vorgeschichte sowie Patienten, die zuvor unter Übelkeit und Erbrechen leiden oder 24 Stunden vor geplanter OP Antiemetika erhalten. Nicht berücksichtigt werden folgende OP-Arten: große intraabdominale-, intrathorakale- bzw. pelvine OPs sowie rekonstruktive Eingriffe. Die Stichprobe beinhaltet 206 Patienten. Insgesamt erleiden 81 Patienten PONV.

Der Apfelscore wird genutzt, um Hochrisikopatienten zu detektieren. Sie werden auf zwei Medikamente (Ramosetron und Ondansetron) zufällig gesplittet. Die Verteilung der Risikofaktoren unter den Patienten zeigt, dass es keine signifikanten Unterschiede bei der Risikostratifizierung des Kollektivs gibt. Des Weiteren sind die beiden Medikamente gleich potent bezüglich der antiemetischen Prophylaxe. Zudem wird kein signifikanter Unterschied in der Inzidenz von PONV, der Schwere der Übelkeit und der Notwendigkeit eines Rettungsantiemetikums zwischen den beiden Gruppen in den ersten 24 Stunden nach der Operation gefunden.[2]

3.8.4.16 Wolf2016

Diese Publikation zählt zu Impact Assessment des vereinfachten Apfelscores sowie des POVOC-Scores und wurde bereits ausführlich unter Punkt 3.5.3.1 beschrieben.

Präoperativ werden die Risikofaktoren für PONV durch die etablierten Scores erhoben. Bei Kindern im Alter von vier bis 5 Jahren wird der POVOC-Score, bei Jugendlichen und Erwachsenen ≥ 16 Jahre wird der Apfelscore genutzt. Die Behandlung der Patienten richtet sich nach der Anzahl der Risikofaktoren im Score.

In den Hochrisikogruppen wird durch die Prophylaxe die PONV-Inzidenz gegenüber den erwarteten Werten deutlich gesenkt. Die Gesamt-PONV-Inzidenz liegt unter 20 %.[96]

3.8.4.17 Gan2017

Diese Studie ist bei Impact Assessment des vereinfachten Apfelscores einzuordnen. Es handelt sich um eine multizentrische, doppelblinde placebokontrollierte Studie zur Vorbeugung von PONV via Amisuprid.

Die Untersuchungen finden in 19 Zentren (4x Frankreich, 6x Deutschland, 9x USA) statt. Als Endpunkt dient das Auftreten und der Schweregrad von PONV innerhalb der 72 Stunden nach Abschluss der Operation. Erbrechen wird definiert als Auswurf auch kleinster Mengen an Mageninhalt durch den Mund. Würgen wird beschrieben als muskulöse Aktionen wie Erbrechen, jedoch ohne Ausstoßung des Mageninhalts. Übelkeit wird festgelegt als Drang zum Erbrechen ohne das Vorhandensein von ausstoßenden Muskelbewegungen. Die Einschlusskriterien beinhalten stationäre, elektiv Eingriffe unter Vollnarkose mit einer Mindestdauer von einer Stunde von Einleitung der Anästhesie bis zum Wundverschluss. Zusätzlich müssen mindestens zwei der vier Apfel-Risikofaktoren für PONV

vorliegen. Ausgeschlossen werden ambulante, intrathorakale Transplantations- oder Zentralnervensystem-Operationen. Wenn erwartet wird, dass der Patient für eine Zeit nach der Operation beatmet bleibt oder eine Nasen- oder Magensonde in situ benötigt, wird er ausgeschlossen. Ebenfalls nicht berücksichtigt werden Patienten mit einer Vorgeschichte von Schwindel, Übelkeit oder Erbrechen 24 Stunden vor der Operation oder, wenn sie mit regelmäßiger antiemetischer Therapie einschließlich systemischer Kortikosteroide behandelt werden. Falls vor oder während der Operation ein Antiemetikum (außer dem Studienmedikament) oder postoperativ ein Antiemetikum außer als Rettungsmaßnahme verabreicht wird, ist keine Berücksichtigung möglich. Auch vorbestehende vestibuläre Störungen waren ein Ausschlusskriterium.

Bei Einleitung der Narkose wird eine 5-mg-Dosis Amisulprid oder ein passendes Placebo verabreicht. Die Daten von insgesamt 689 Probanden (342 in den USA und 347 in Europa) werden randomisiert. Die Protokolle von 626 Teilnehmern sind auswertbar.

Die Studienarme sind hinsichtlich des PONV-Risikos, wie es durch prospektiv ermittelte Risikofaktoren geschätzt wurde, gut ausbalanciert. In beiden Armen hat mehr als ein Viertel der Patienten zwei, knapp die Hälfte drei und ein Viertel aller vier Risikofaktoren. Übelkeit tritt bei Patienten, die Amisulprid erhalten, seltener auf als bei Patienten, die Placebo bekommen. Eine konsistente relative Risikoreduktion von etwa 20 % wird sowohl für die zusammengesetzte PONV-Messung als auch einzeln für Erbrechen, Übelkeit und den Bedarf an Notfallmedikamenten festgestellt.[29]

3.8.4.18 Halliday2017

Diese Publikation zählt zu Impact Assessment des vereinfachten Apfelscores.

Es handelt sich um eine Beobachtungsstudie bei bariatrischen OPs. Untersucht wird die Inzidenz von PONV sowie die Verwendung und Wirksamkeit von PONV-Prophylaxe. Die Patienten erhalten eine Prophylaxe auf der Grundlage veröffentlichter Leitlinien und eines vereinfachten Apfelscores.

Die Studie findet im Sundsvall Hospital in Schweden statt. Als Endpunkte zählen das Auftreten von PONV, seine Schwere und Häufigkeit sowie die notfallmäßige Antiemetikagabe. Der Beobachtungszeitraum beträgt 72 Stunden nach der OP. Die Einschlusskriterien umfassen bei erwachsenen Patienten stationäre Eingriffe in TIVA mit bariatrischen Operationen (laparoskopischer Magenbypass oder Magenmanschette). Die Ausschlusskriterien beinhalten mangelnde Sprachkenntnisse, unzureichende Forschungsressourcen und die Umstellung auf eine offene Operation.

Die Stichprobe besteht aus 74 Patienten. Hiervon erleiden 48 PONV. Das ergibt bei vier Risikofaktoren 12 Ereignisse pro Variable.

Allen Patienten wird eine TIVA mit Propofol und Remifentanyl sowie Betamethason als Antiemetikum verabreicht. 92% der Patienten nehmen Ondansetron und 24% Droperidol ein. Ein Patient erhielt eine suboptimale Behandlung, 23% der Patienten erhielten eine optimale und 76% eine supra-optimale Behandlung nach den aktuellen Richtlinien.

Die PONV-Rate bei bariatrischen OPS ist trotz eines Prophylaxe-Regimes nach aktuellen Richtlinien hoch.[33]

3.8.5.1 Qualitätskriterien des original Scores Apfel1998

Kriterium	Untergruppe	Unterpunkt	Max. erreichte Punktzahl	
Generalisierbarkeit	Anzahl des Patientenkollektivs	Scoreentwicklung & -update: 2136 <ul style="list-style-type: none"> ▪ Entwicklung: 553 ▪ Scoreupdate: 1583 	2,5	
		Risiko-faktoren in dem Modell	Anzahl der RF in dem Modell: 5	2
Praktikabilität	Risiko-berechnung/ Handhabung des Scores	Mehrheit der RF ohne technische Hilfsmittel mit ja/nein Antwort, mittlerer Zeitaufwand	2	
		Mehrheit der RF exakt definiert	2	
		RF ohne Kenntnis des Outcomes beurteilt: nein/ unklar	0	
		Hilfsmittel nötig: spezielles Schaubild	1	
	Endpunkt des Scores	vereinfachter Score vorhanden	1	
		einzelner Endpunkt PN/ (initial bei Apfel)	1	
	Präzision	Vorhersagegenauigkeit/ Beurteilung der Prädiktion mittels:	Art und Weise der Endpunkterhebung: via Gespräch (persönliche Visite/ Telefonat mit dem Patienten)	1
			Diskriminationsfähigkeit (AUC der ROC-Kurve) <ul style="list-style-type: none"> ▪ Endwert: (Intern & Extern) AUC: 0,736 	2,5
Kalibrierung <ul style="list-style-type: none"> ▪ Geradengleichung ▪ Fläche: 6,1 			0	
		Kalibrierung <ul style="list-style-type: none"> ▪ Calibration Plot ▪ Hosmer-Lemeshow-Test 	1 1	

Validierung	Interne Validierung des Scores	Methode der Validierung: Splitting der Daten	1
		Gesamtzahl des Patientenkollektivs: 1667	2
	Externe Validierung des Scores	Anzahl der Validierungen: 9 <ul style="list-style-type: none"> ▪ 3x mit Update ▪ 6x ohne Update 	3
		Lokalisation der Validierungen: externe Validierung auf einem anderen Kontinent wie Modellentwicklung	3
		Gesamtzahl des Patientenkollektivs: 8182 <ul style="list-style-type: none"> ▪ mit Update: 1583 Patienten ▪ ohne update: 6599 Patienten 	3
	Breite Anwendung in der Praxis & Extrapunkte	Impact Assessment des Scores	Anwendung des Scores zeigt Benefit: keine
bisherige Nutzung und Anwendung des Scores: Vergleich von Studienteilnehmern/ Herstellung von Kohortengleichheit			1
Gesamtzahl des Patientenkollektivs: 177			0,5
Extrapunkte		Modifikation der erreichten Punktzahl: Maßzahlen zur Modellklassifizierung	1
Insgesamt			31,5 Punkte

Tabelle 15: Qualitätskriterien des original Scores Apfel1998

3.8.5.2 Qualitätskriterien des vereinfachten Scores Apfel 1999

Kriterium	Untergruppe	Unterpunkt	Max. erreichte Punktzahl
Generalisierbarkeit	Anzahl des Patientenkollektivs	Scoreentwicklung & -update: 5130 <ul style="list-style-type: none"> Entwicklung: 1.040 Update: 4090 	3
		Risiko-faktoren in dem Modell	Anzahl der RF in dem Modell: 4
Praktikabilität	Risiko-berechnung/ Handhabung des Scores	Mehrheit der RF ohne technische Hilfsmittel mit ja/nein Antwort, mittlerer Zeitaufwand	2
		Mehrheit der RF exakt definiert	2
		RF ohne Kenntnis des Outcomes beurteilt: nein/ unklar	0
		keine Hilfsmittel nötig	3
	Endpunkt des Scores	vereinfachter Score vorhanden	1
		zusammengesetzter Endpunkt PONV	2
	Art und Weise der Endpunkterhebung: via Gespräch (persönliche Visite/ Telefonat mit dem Patienten)		1
Präzision	Vorhersagegenauigkeit/ Beurteilung der Prädiktion mittels:	Diskriminationsfähigkeit (AUC der ROC-Kurve) <ul style="list-style-type: none"> Endwert (Intern & Extern) AUC: 0,721 	2,5
		Kalibrierung <ul style="list-style-type: none"> Geradengleichung Fläche: 4,67 	1
		Kalibrierung <ul style="list-style-type: none"> Calibration Plot Hosmer-Lemeshow-Test 	1 1

Validierung	Interne Validierung des Scores	Methode der Validierung: Kreuzvalidierung	1
		Gesamtzahl des Patientenkollektivs: 1040	1,5
	Externe Validierung des Scores	Anzahl der Validierungen: 18 <ul style="list-style-type: none"> ▪ 6x mit Update ▪ 12x ohne Update 	3
		Lokalisation der Validierungen: externe Validierung auf einem anderen Kontinent wie Modellentwicklung	3
		Gesamtzahl des Patientenkollektivs: 11681 <ul style="list-style-type: none"> ▪ mit Update: 4925 ▪ ohne Update: 6756 	3
	Breite Anwendung in der Praxis & Extrapunkte	Impact Assessment des Scores	Anwendung des Scores zeigt Benefit: <ul style="list-style-type: none"> ▪ geringere PONV Inzidenz ▪ höhere Patientenzufriedenheit ▪ weniger Komplikationen
bisherige Nutzung und Anwendung des Scores: <ul style="list-style-type: none"> ▪ Identifizierung von Hochrisikopatienten ▪ Vergleich von Studienteilnehmern/ Herstellung von Kohortengleichheit 			1 1
Gesamtzahl des Patientenkollektivs: 9190			3
Extrapunkte		Modifikation der erreichten Punktzahl: <ul style="list-style-type: none"> ▪ Maßzahlen zur Modellklassifizierung ▪ Impact Assessment: Ausschluss oder Einschluss neuer Risikofaktoren ▪ Impact Assessment: Bewertung gleicher Potenz zweier Antiemetika 	1 1 1
Insgesamt			44 Punkte

Tabelle 16: Qualitätskriterien des vereinfachten Scores Apfel 1999

3.9 Gesamtauswertung aller Scores

3.9.1 Gesamtzahl der Publikationen pro Score

Initial wurden 85 Studien in die Excel Datenmaske eingespeist. Im Anschluss wurden alle Scores berücksichtigt, die in mindestens zwei Publikationen behandelt wurden. Somit waren 62 Publikationen relevant. Sie umfassen acht Scores und ein Kollektiv von insgesamt 81834 Patienten. Diese Scores lauten:

- Junger Score 2001 mit 2 Publikationen (1 Modellentwicklung, 1 interne Validierung)
- Palazzo & Evans Score 1993 mit 7 Publikationen (1 Modellentwicklung, 5 externe Validierungen, 1 Impact Assessment)
- Sinclair Score 1999 mit 7 Publikationen (1 Modellentwicklung, 5 externe Validierungen, 1 Impact Assessment)
- Apfel 2012 PDNV Score mit 3 Publikationen (1 Modellentwicklung, 1 interne Validierung, 1 externe Validierung)
- Eberhart 2004a POVOC-Score mit 4 Publikationen (1 Modellentwicklung, 2 externe Validierungen, 1 Impact Assessment)
- Van den Bosch Score 2005a mit 5 Publikationen (1 Modellentwicklung, 1 externe Validierung, 3 Impact Assessment)
- Koivuranta Score 1997 mit 12 Publikationen (1 Modellentwicklung, 9 externe Validierungen, 2 Impact Assessment)
- Apfel Score 1998 mit 45 Publikationen (4 Modellentwicklungen, 23 externe Validierungen, 18 Impact Assessment)

Den genauen Verlauf dieses Prozesses zeigt Abbildung 6.

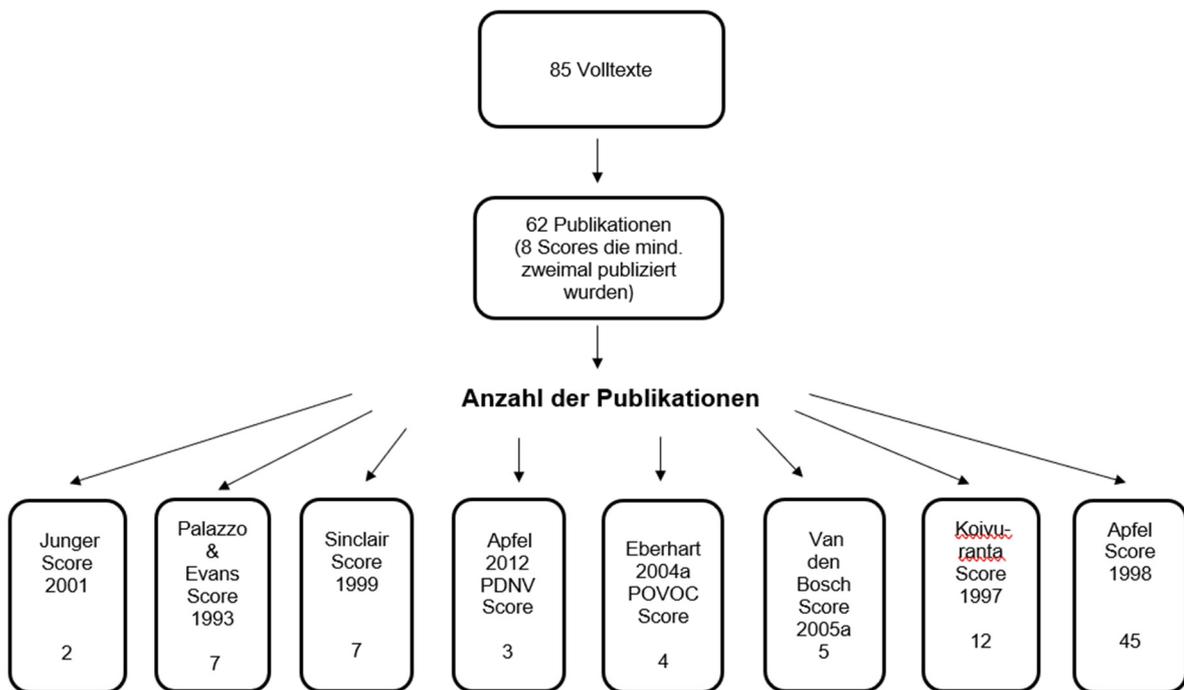


Abbildung 6: Anzahl der Publikationen pro Score

3.9.2 Patientenkollektive pro Score

Nachdem die Publikationen den acht Scores zugeordnet wurden, ist in diesem Punkt auf die Gesamtzahl der Patientenkollektive einzugehen. Für die beiden Scores Apfel 1998 und Koivuranta 1997, die jeweils eine vereinfachte Scoreversion vorstellen, wurden die Subgruppen: original und vereinfachter Score zusätzlich differenziert.

Es zeigte sich, dass für die meisten Scores alle drei wichtigen Gruppen: Modellentwicklung, externe Validierung sowie Impact Assessment vorlagen. Im Bereich der Erwachsenen-Scores waren die Kollektive für den vereinfachten Apfelscore, den vereinfachten Koivurantascoresowie der Score von Van den Bosch 2005a am ausgewogensten. Im Bereich der Scores für Kinder war der Eberhart 2004a POVOC-Score der einzige etablierte. Es lagen hier ebenfalls aus allen drei Teilbereichen Publikationen vor.

Nachfolgende Abbildung verdeutlicht die Anzahl und Verteilung der Patienten.

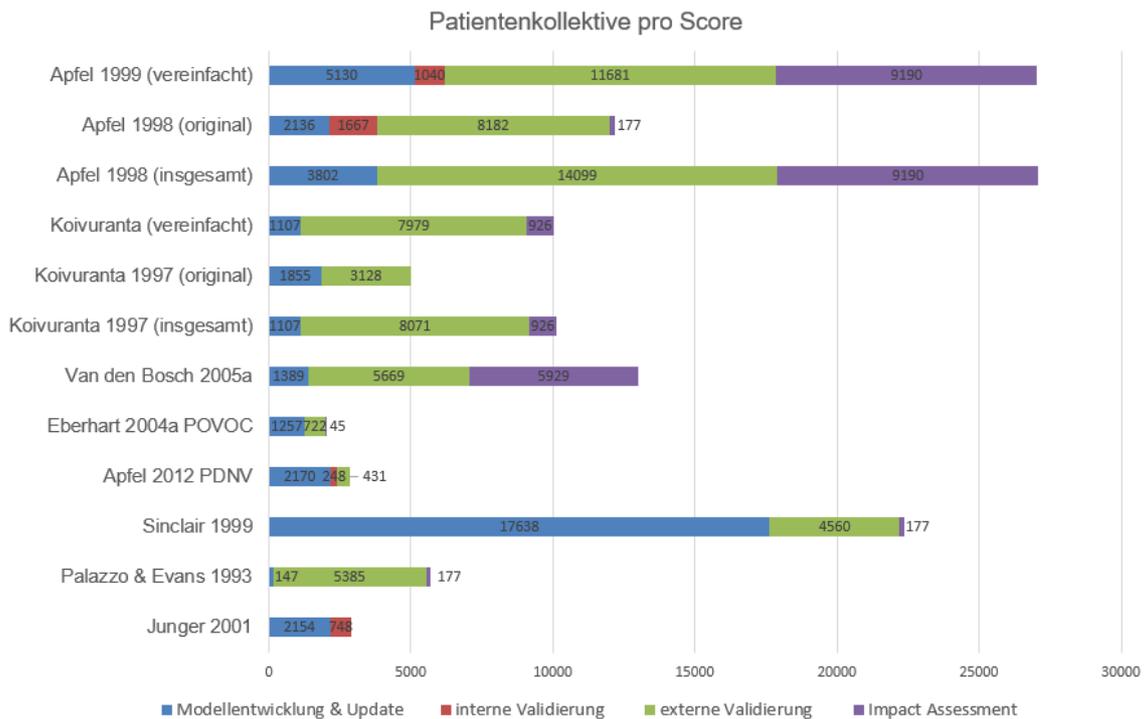


Abbildung 7 Patientenkollektive pro Score

3.9.3 Punktwerte der Scores bei den Qualitätskriterien

Bei der Bewertung der verschiedenen Scoringssysteme warnicht nur die absolute Anzahl an Publikationen und Patienten relevant, da die alleineige Masse kein Qualitätsgarant ist.

Somit wurden Qualitätskriterien für die Punktvergabe festgelegt. Die Höchstpunktzahl bei der entwickelten Datenmaske in Excel betrug 60 Punkte. Folgende Rubriken wurden berücksichtigt: die Anzahl der Risikofaktoren in dem Modell sowie die Risikoberechnung. Dies ist für die Handhabung im klinischen Alltag wichtig. Die Größe des Patientenkollektivsin den Teilbereichen Modellentwicklung, interne Validierung, externe Validierung und Impact Assessment ist ebenso wesentlich. Wenn das Modell an vielen Patienten entwickelt wurde, an einer breiten Masse mehrfach validiert wurde sowie ggf. eine Anpassung erfolgte, dann ist davon auszugehen, dass es robuster ist. Die Beurteilung der Vorhersagegenauigkeit mit Diskriminationsfähigkeit und Kalibrierung wurde bepunktet. Die Aufgabe des Modells ist es, keine Über- oder Unterschätzung des Risikos vorzunehmen und Risikopatienten gut zu identifizieren.

Auch die bisherige Nutzung der Scores sowie der direkte Benefit für den Patienten wurde bewertet. Schließlich gab es Extrapunkte zur Modifikation der erreichten Punktzahl für Besonderheiten, die noch nicht berücksichtigt wurden.

Die nachfolgenden Abbildungen veranschaulichen das Abschneiden der jeweiligen Scores bei den Qualitätskriterien.

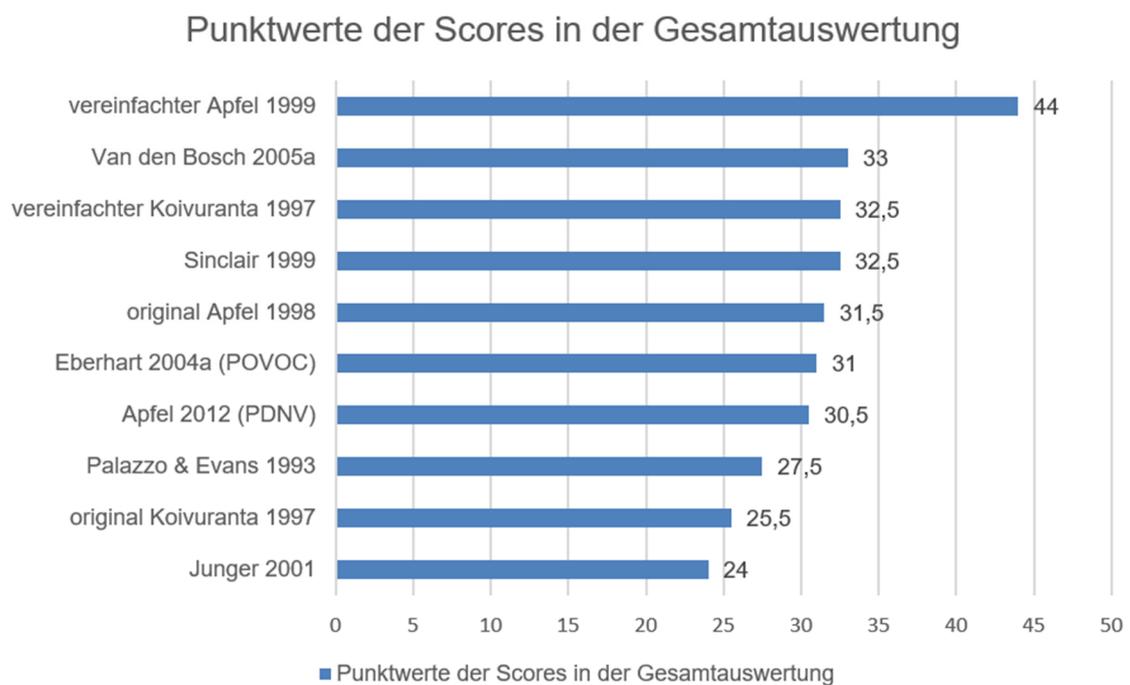


Abbildung 8: Gesamtpunktzahl der Scores in der Gesamtauswertung

Score	Bewertungskriterium und erreichte Punktzahl									Gesamtpunktzahl
	Generalisierbarkeit	Praktikabilität			Präzision	Validierung		Breite Anwendung in der Praxis & Extrapunkte		
		Patientenkollektiv	Risikofaktoren	Risikoberechnung		Endpunkt	Diskrimination & Kalibrierung	Interne Validierung	Externe Validierung	
Vereinfachter Apfel 1999	3	6	4	3	5,5	2,5	9	8	3	44
Van den Bosch 2005a	2,5	7	1	5	6,5	1	5	4	1	33
Vereinfachter Koivuranta 1997	1,5	6	4	2	5	0	9	4	1	32,5
Sinclair 1999	3	5	1	3	6	4	9	1,5	0	32,5
Original Apfel 1998	2,5	6	2	2	4,5	3	9	1,5	1	31,5
Eberhart 2004a	1	8	4	2	5,5	2	6	2,5	0	31
Apfel 2012	2,5	7	4	3	6,5	2	4,5	0	1	30,5
Palazzo & Evans 1993	0,5	7	1	2	4,5	0	9	1,5	2	27,5
Original Koivuranta 1997	2	6	2	2	4,5	0	9	0	0	25,5
Junger 2001	2	5	1	3	7,5	4,5	0	0	1	24

Abbildung 9 einzelne Punktwerte der Scores pro Qualitätskriterium in der Gesamtauswertung

Im Folgenden wird auf die ersten drei Plätze der Erwachsenen-Scores sowie den einzigen Score für Kinder eingegangen.

Spitzenreiter war der vereinfachte Apfelscore. Er war mit seinen vier gleichgewichteten Risikofaktoren einfach anzuwenden. Initial wurde er an der europäischen Bevölkerung entwickelt. In den letzten Jahren erfolgte eine Vielzahl von Validierungen unter anderem auch auf anderen Kontinenten z.B. Asien oder Südamerika. Hierbei wurde überprüft, ob der Score auf andere Ethnien übertragbar ist. Des Weiteren bestehen viele Publikationen zu dem Bereich Impact Assessment.

Den zweiten Platz belegte der Score Van den Bosch 2005a. Dieser Score ist eine komplizierte Gleichung. Ein Normogramm zur besseren Handhabung im Alltag wurde vorgestellt. Anschließend wurde der Score in einer Folgepublikation upgedatet. Das Besondere an der Gruppe der Publikationen war das ausführliche Impact Assessment. Es war sehr nahe an der klinischen Praxis. Initial wurde gezeigt, dass die alleinige Nutzung des Scores keine Senkung der PONV-Ereignisse lieferte. In einem nächsten Schritt wurde mit einer Ampel das jeweilige Patientenrisiko visualisiert und ein konkreter Prophylaxe vorschlag angeboten. In der letzten Publikation wurden die Anästhesisten über die Hindernisse bei der Verwendung von Scoringsystemen interviewt.

Punktgleich waren der vereinfachte Score von Koivuranta sowie der Score von Sinclair 1999.

Der vereinfachte Score von Koivuranta mit den fünf gleichgewichteten Risikofaktoren war der erste vereinfachte Score, der publiziert wurde. Der Datensatz aus Finnland spielte bei der Entwicklung des vereinfachten Apfelscores eine wichtige Rolle. Zusätzlich besaß der vereinfachte Score von Koivuranta gemittelt eine etwas bessere Diskriminationsfähigkeit mit einer AUC von 0,738 in den Berechnungen als der vereinfachte Apfelscore mit einer AUC von 0,721.

Das Besondere an dem Score von Sinclair aus dem Jahre 1999 war, dass er für ambulante Eingriffe entwickelt wurde. Die Endpunkterhebung erfolgte via Telefoninterview. Die Patientenrekrutierung lief über drei Jahre und schloss ein sehr großes Kollektiv mit 17638 Patienten ein. Die gemittelte AUC aus allen Publikationen war mit 0,777 höher als bei den vereinfachten Scores von Apfel oder Koivuranta. Allerdings ist die Gleichung mit sieben Risikofaktoren unhandlich und es liegt kein Scorechart für den Alltag vor.

Mit dem Risiko von postoperativem Erbrechen bei Kindern befasste sich als einziger der POVOC-Score von Eberhart2004a. Als Endpunkt wurde postoperatives Erbrechen definiert, Übelkeit wurde nicht vorhergesagt, das sie bei Kleinkindern nicht objektiv ermittelbar ist. Das Patientenkollektiv für Entwicklung und Validierung war groß, gemessen an der kleineren Fallzahl an pädiatrischen Patienten. Eine externe Validierung in Brasilien sowie eine Studie zum Impact Assessment lag vor. Somit waren alle zu untersuchenden Teilbereiche gut beschrieben.

4 Diskussion

4.1 Positive Punkte der Arbeit

Diese Arbeit stellt eine Systematische Zusammenstellung aller Scores zu PONV dar. Diverse Literaturdatenbanken wurden durchsucht und ein Literaturupdate durchgeführt, da die initiale Suche zu weit zurück lag. Um ein möglichst umfangreiches Bild zu erhalten, gab es keine Ausschlusskriterien bezüglich des Erscheinungsjahrs oder der Sprache einer Publikation.

Die Bewertung der Studien erfolgte anhand einer zuvor rational entwickelten Methode mit Punktsystem. Somit ist die Objektivität, nach Vollziehbarkeit und Transparenz der Bewertung vorhanden. Die Umfrage in den Krankenhäusern bezüglich SOPs zu der Vorbeugung und Behandlung von PONV verleiht der Arbeit eine Praxisnähe. Das Ziel dieser Arbeit den Klinikalltag durch Steigerung der Patientensicherheit und Zufriedenheit zu erleichtern wurde verfolgt.

4.2 Einschränkungen der durchgeführten Analyse

Ein möglicher Kritikpunkte ist, dass es sich bei dieser Arbeit um eine rein deskriptive Herangehensweise handelt. Es wurden keine statistischen Berechnungen angestellt. Dies ist größtenteils der inhomogenen Gruppe an Publikationen geschuldet. Diese lassen sich nur miteinander vergleichen. Der Ansatz, die jeweils gepoolten Gruppen (Scoreentwicklung, Update, Validierung und Impact Assesment) pro Score zu bilden diente der Objektivierung der Ergebnisse. Die Herausforderung war es allen 62 Publikationen gerecht zu werden, sich auf die wesentlichen Aussagen zu beschränken und trotzdem eine Übersichtlichkeit beizubehalten.

Als Leitfaden zur Datengewinnung und für die Qualitätskriterien dienten zwei Vorlagen aus dem Protokoll des Cochrane Reviews, sie wurden dort lediglich Appendix 7 und 8 genannt[59]. Die Erste Vorlage, Anhang sieben genannt, ist eine sehr ausführliche Datenmaske und somit etwas unübersichtlich. Vorteilhaft waren die geschlossenen Fragen und die vordefinierten Antwortmöglichkeiten zum Ankreuzen. Die zweite Vorlage, Charms 2014 genannt hatte den Titel: „relevant items to extract from individual studies in a systematic review of prediction models“[57]. Sie war eine gute Orientierungshilfe, gut überschaubar, mit vielen Freitextantworten und somit schwer auszuwerten, bei einer großen Anzahl an Publikationen. Diese Maske deckte nicht den Bereich Impact Assesment ab, er ist wichtig für die Beurteilung des klinischen Nutzens. Im Anschluss wurde aus beiden Vorlagen und eigenen Überlegungen eine Excel Datenmaske erstellt. Hier kodierte die Zahl 1 ja und 0

nein. Somit war eine leichtere Auswertung möglich. Es erfolgte jeweils eine Probedatengewinnung mit einigen Publikationen und anschließend ein Feinschliff, bevor alle Studien bearbeitet wurden.

In dem zweiten Schritt wurden die gruppierten Publikationen via Qualitätskriterien bewertet. Hier wurde versucht den verschiedenen Unterpunkten ein gleichmäßiges Gewicht zu geben. Alltagsrelevante Faktoren zur Scorenutzung wurden ebenso berücksichtigt wie Qualitätskriterien an sich. Ziel war es eine Aussage für die klinische Arbeit machen zu können. [56,59]

4.3 Diskussion der Ergebnisse

4.3.1 Umfrage PONV-SOPs in Kliniken

Da es sich bei dieser Arbeit bereits um eine literaturbasierte Arbeit handelt, ist es schwer sie erneut mit Literatur zu vergleichen. Somit entstand die Idee, die gewonnenen Ergebnisse mit der Praxis im klinischen Alltag zu vergleichen. Stimmen die gewonnen Erkenntnisse mit den alltäglich genutzten SOPs in den Kliniken überein.

Hierfür wurden am 10.04.2020 26 beliebige Krankenhäuser deutschlandweit angeschrieben, vom Kreiskrankenhaus bis zum Universitätsklinikum. Als Antwort haben acht Kliniken ihre SOPs übermittelt. Vier Handlungsanweisungen behandelten nur Erwachsene, die vier anderen Erwachsene und Kinder. Dies ist in dem folgenden Tortendiagramm dargestellt.

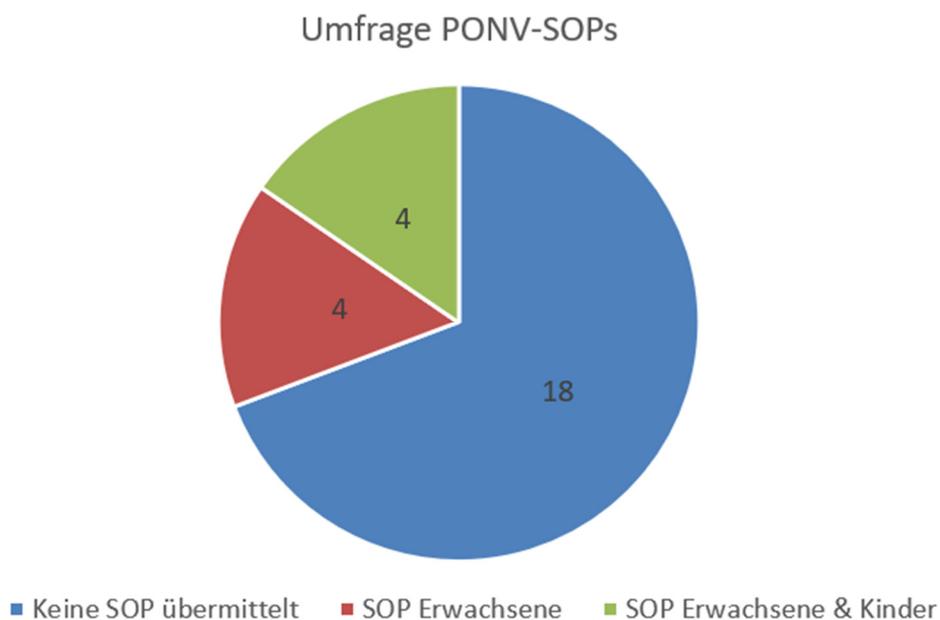


Abbildung 10 Antwortverhalten der Kliniken bei der Umfrage bezüglich PONV SOPs

4.3.1.1 PONV-SOPs für Erwachsene

Die vier Kliniken mit SOPs für Erwachsene lauten: Universitätsklinikum Marburg, Klinikum Augsburg, Klinikum Würzburg Mitte Standort Juliusspital und Diakonie Klinikum Stuttgart.

Am Universitätsklinikum in Marburg gilt, jeder erwachsene Patient soll prophylaktisch 2 Maßnahmen erhalten. Diese sind in einer Liste mit fix vorgegebenen Prozeduren, auch i.v. Anästhesie mit Propofol, vorgegeben. Eine Reduktion bzw. Ausweitung der Prophylaxe ist möglich. Die hier berücksichtigten Risikofaktoren entsprechen dem vereinfachten Apfelscore von 1999.[20]

Am Klinikum Augsburg dient der vereinfachte Apfelscore von 1999 zur Risikostratifizierung bei Erwachsenen. Es wird ein risikoadaptiertes Vorgehen empfohlen.

Patienten mit maximal einem Risikofaktor erhalten keine medikamentöse Prophylaxe. Bei zwei Risikofaktoren erhalten Erwachsene eine Prophylaxe mit einem Präparat. Ab drei Risikofaktoren gibt es eine medikamentöse 2-fach Kombination zur Prophylaxe. Eine Regionalanästhesie soll erwogen werden. Patienten mit vier Risikofaktoren erhalten eine 3-fache medikamentöse Prophylaxe. Eine TIVA oder Regionalanästhesie muss in Erwägung gezogen werden und auf emetogene Substanzen wird verzichtet. [44]

Im Klinikum Würzburg Mitte Standort Juliusspital werden für Erwachsene diverse Risikofaktoren genannt, jedoch kein Score explizit angewendet. Eine generelle PONV-Prophylaxe Empfehlung für alle Patienten wird nicht ausgesprochen.

Bei bekanntem PONV soll, wenn möglich, eine Regionalanästhesie oder eine TIVA erfolgen. Verabreicht werden kann: Dexamethason in Absprache mit dem Operateur, ein Schmerztropf mit Nalbuphin statt Tramadol, bzw. MCP oder Ondansetron. Als Standard wird empfohlen Propofol als Einleitungshypnotikum mit ggf. Bolus in der Ausleitphase sowie Dimehydrinat-Infusion. [45]

Das Diakonie Klinikum in Stuttgart bringt die Risikoeinschätzung mithilfe des vereinfachten Apfelscores von 1999. Es erfolgt ein Vorgehen in Abhängigkeit vom Risikoscore. Jeder Patient erhält eine Prophylaxe mit Dexamethason. Ab zwei Risikofaktoren wird Ondansetron hinzugenommen. Bei drei Faktoren soll eine TIVA erwogen werden und eine duale Prophylaxe erfolgen. Vier Risikofaktoren haben eine TIVA plus duale PONV-Prophylaxe zur Folge.[55]

4.3.1.2 PONV-SOPs für Erwachsene & Kinder

Die vier Kliniken, die sowohl SOPs für Erwachsene und Kinder haben, werden im Folgenden vorgestellt: Klinikum Stuttgart, Universitätsklinikum Würzburg, Universitätsklinikum Ulm und Universitätsklinikum Jena.

Im Klinikum Stuttgart ähneln die Arbeitsanweisungen für PONV bei Erwachsenen, den SOPs des Uniklinikum Marburg. Jeder Erwachsene erhält prophylaktisch 2 Maßnahmen aus einer Liste von 5 fix vorgegeben Prozeduren. Eine Anpassung d.h. Reduktion bzw. Ausweitung der Prophylaxe ist möglich. Auch hier werden die Risikofaktoren entsprechend dem vereinfachten Apfelscore von 1999 angewendet.[91]

Die SOP zur Prophylaxe von PONV bei Kindern zeigen ein fixes Prophylaxe Schema. Jedes Kind erhält vorsorglich Dexamethason, außer es liegen Kontraindikationen vor. Bei erhöhtem Risiko wird zusätzlich Ondansetron oder Dimenhydrinat appliziert. Als Risikofaktoren werden die Punkte des POVOC Scores von Eberhart aus dem Jahre 2004 angeführt. Der Risikofaktor Strabismus-OP wird um Adenotomien und Tonsillektomien erweitert. [91]

Am Universitätsklinikum in Würzburg liegt eine SOP vor, die Erwachsene und Kinder berücksichtigt.

Die generelle Prophylaxe Empfehlung besagt, sofern keine Kontraindikationen vorliegen, erhalten alle Patienten zwei festgelegte Antiemetika (Dexamethason & Ondansetron). Bei Patienten mit erhöhtem Risiko wird zusätzlich eine TIVA durchgeführt. Soll bei diesen Risikopatienten keine TIVA durchgeführt werden, wird ein drittes Antiemetikum appliziert.

Alternativ kann bei der Prophylaxe ein individualisiertes Vorgehen unter Berücksichtigung des Patientenrisikos erfolgen. Bei Erwachsenen werden die Risikofaktoren entsprechend dem vereinfachten Apfelscore von 1999 berücksichtigt. Für Kinder werden die Risikofaktoren des Eberhart-POVOC Scores angewendet. Als Hilfestellung gelten folgende Regeln: Patienten mit einem Risiko bis 20 % und Patienten mit Regionalanästhesie benötigen keine Prophylaxe. Ab einem Risiko von 40 – 60% erhalten Patienten ein bis zwei Therapieoptionen aus einer Tabelle mit sechs verschiedenen Interventionen. Patienten mit einem Risiko von 80% erhalten mindestens drei Therapieoptionen aus der Tabelle.[48]

In der SOP der Universitätsklinik in Ulm wird für Erwachsene der vereinfachte Apfelscore von 1999 als System genannt. Des Weiteren wird der PDNV Score von

Apfel aus dem Jahre 2012 aufgeführt. Bei Erwachsenen wird ein risikoadaptiertes Vorgehen nach Anzahl der Risikofaktoren empfohlen.

Bei null bis einem Risikofaktor gibt es keine Prophylaxe. Ab zwei bis drei Risikofaktoren wird eine Regionalanästhesie oder eine TIVA ohne weitere medikamentöse Prophylaxe durchgeführt. Alternativ werden zwei vorgegebene Antiemetika verabreicht. Bei vier oder mehr Risikofaktoren erfolgt eine Regionalanästhesie ggf. plus medikamentöse Prophylaxe. Falls dies nicht realisierbar ist wird eine TIVA plus zwei Antiemetika appliziert.

Zur Risikoabschätzung bei Kindern dient der POVOC Score von Eberhart aus dem Jahre 2004. Bei Kindern ≥ 2 Jahren erfolgt routinemäßig eine Prophylaxe. Insgesamt liegt hier eine großzügigere Indikationsstellung vor. Es werden drei mögliche Substanzen zur Prophylaxe angegeben. [63]

Am Universitätsklinikum Jena erfolgt für Erwachsene eine medikamentöse Prophylaxe nach dem Risikoprofil, keine generelle Prophylaxe. Die Risikoeinschätzung erfolgt mittels des vereinfachten Apfelscore von 1999. Zusätzlich werden zwei weitere Risikofaktoren berücksichtigt: Lebensalter < 50 und/oder PONV- Risikoeingriff (Ösophagus-/Magenresektionen, Kiefferverdrahtung, Patienten mit erhöhtem Hirndruck).

Als fixes Handlungsschema gilt, Patienten mit einem Risikofaktor erhalten regulär keine Prophylaxe. Dexamethason soll in Erwägung gezogen werden, das Lebensalter < 50 ist oder es sich um einen Risikoeingriff handelt. Patienten mit zwei Risikofaktoren erhalten Dexamethason, falls die Zusatzkriterien greifen, kann Garnisetron zusätzlich appliziert werden. Ab drei Risikofaktoren wird Dexamethason plus Garnisetron verabreicht. Falls die Extrakriterien zutreffen, wird eine TIVA oder eine höhere Dexamethasondosis empfohlen. Liegen vier Risikofaktoren vor, wird eine Kombination aus höherer Dosis Dexamethason plus Garnisetron empfohlen. Alternativ kann eine TIVA durchgeführt werden.

Für Kinder wird ein modifizierter POVOC-Score angewendet. Adaptierte Risikofaktoren sind das Alter, hier gilt ≥ 2 Jahren als prädisponierend und die OP-Art wird von Strabismus zusätzlich erweitert auf Adenotomie und Tonsillektomie. Als Prophylaxe wird allen Kindern ab zwei Jahren eine Doppelmedikation verabreicht.[89]

4.3.1.3 Vergleich SOP Ergebnisse mit Literaturanalyse der Scores

Insgesamt nutzen die acht PONV-SOPs der Erwachsenen sieben Mal den vereinfachten Apfelscore aus dem Jahre 1999, einmal wird zusätzlich der PDNV-Score von 2012 genutzt und einmal wird kein Scoringssystem herangezogen. Bei den vier

SOPs die Kinder behandeln, wird immer der POVOC Score von Eberhart aus dem Jahr 2004 herangezogen. Diese Fakten sind in folgendem Diagramm dargestellt.

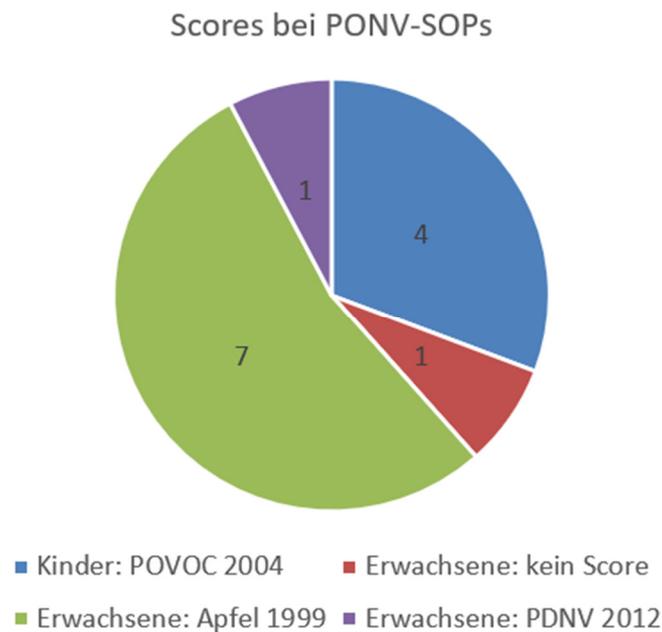


Abbildung 11: Scores bei PONV-SOPs

Diese Ergebnisse passen zu den Fakten, die diese literaturbasierte Übersichtsarbeit geliefert hat. Der am weitesten verbreitete und handlichste Score im Erwachsenenbereich ist der vereinfachte Score von Apfel aus dem Jahre 1999. Für Kinder war der POVOC Score von Eberhart ebenfalls Spitzenreiter in dieser Altersklasse. In den SOPs wurden gelegentlich die Scores leicht angepasst, z.B. das Alter bei Kindern oder Hochrisiko-OPs hinzugefügt. Dies ist ebenfalls bei den Scoringsystemen in der Literatursuche erfolgt. Somit decken sich die Ergebnisse.

Als bisher nicht berücksichtigter Fakt in der literaturbasierten Übersichtsarbeit ist eine generelle Prophylaxe für jeden Patienten mit ggf. zusätzlichem risikoadaptiertem Vorgehen zu nennen. In den Studien war stets ein rein scorebasiertes Procedere beschrieben. In den SOPs im Erwachsenenbereich empfehlen vier von acht Zentren eine generelle Prophylaxe mit anschließender zusätzlicher Risikoprüfung zur weiteren Therapieerweiterung. Die andere Hälfte (4/8) der Kliniken nutzt nur das scorebasierte Procedere. Bei den vier SOPs, die Kinder behandeln, empfehlen alle Zentren eine generelle Prophylaxe. Jedoch schränken zwei der vier Zentren die generelle Prophylaxe ein, sie gilt erst für Kinder ≥ 2 Jahren. Diese Fakten sind im nachfolgenden Diagramm dargestellt.

SOPs - generelle Prophylaxe vs. risikoadaptiertes Vorgehen

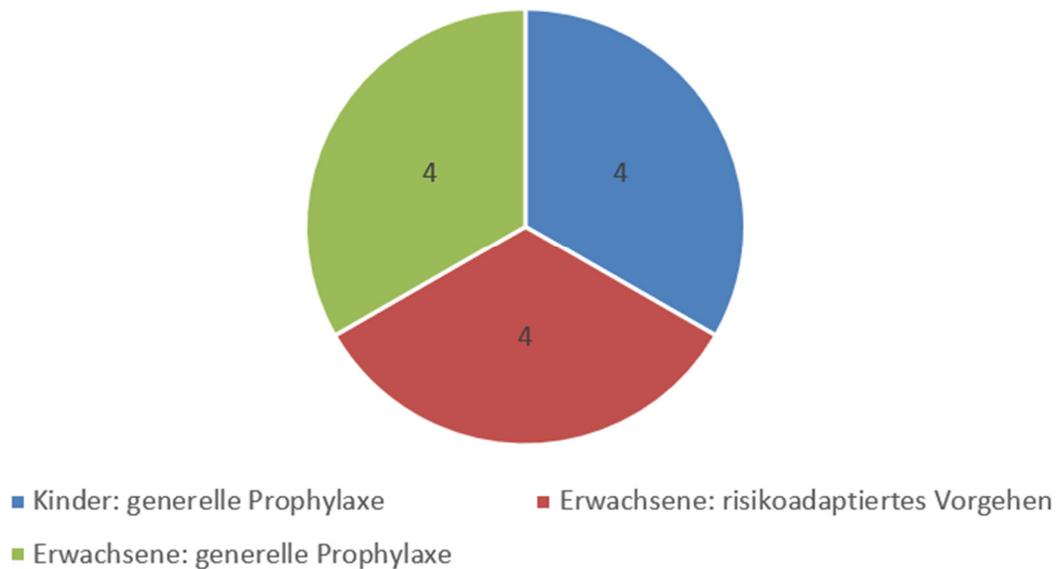


Abbildung 12: SOPs – generelle Prophylaxe vs. risikoadaptiertes Vorgehen

4.4 Schlussfolgerungen für die klinische Praxis

Nach Auswertung aller Studien hat sich gezeigt, dass eine Wirkungsanalyse, die das Patientenergebnis einschließt, bevor ein Modell in der Klinik angewendet wird am wichtigsten ist.

Die Studie von Kappen aus dem Jahr 2014 in Utrecht demonstriert, dass beim alleinigen anzeigen des berechneten PONV Risikos im Patientenverwaltungssystem keine relevante Abnahme der PONV-Inzidenz entsteht. Es liegt lediglich eine Zunahme der risikoadaptierten PONV Prophylaxe vor[40].

Erst in der Folgestudie Kappen 2015, in der im Interventionszeitraum den Anästhesisten das berechnete PONV Risiko und ein konkreter Prophylaxe vorschlag elektronisch angezeigt wird, verringert sich die PONV-Inzidenz. Die Software berechnet das Risiko anhand des Scores von Van den Bosch 2005 und das PONV-Risiko wird mit einer anfänglichen Ampelfarbe eingestuft. Die Ampelfarbe ändert sich, wenn ein Antiemetikum verabreicht wird. Die Software adaptiert das vorhergesagte Risiko neu. Dieser direkte Ansatz führt zu einer Verringerung der PONV-Inzidenz mit einer starken Reduktion des PONV bei Hochrisikopatienten[41].

Ein ähnliches Vorgehen liegt bei der Publikation von Kooji aus dem Jahre 2012 in Amsterdam vor. Es wird ein automatisiertes Erinnerungssystem für das PONV-Management, das den vereinfachten Risiko-Score von Apfel enthält, getestet. Das

Supportsystem schreibt ab einem Risikofaktor eine konkrete Prophylaxe vor und die Anästhesisten nehmen die Empfehlungen an. Eine Reduktion der Gesamt-PONV-Inzidenz in der Hochrisikopatientengruppe von 47% auf 30% zeigt sich. Die Gesamtanwendung der Prophylaxemedikamente unterscheidet sich nicht zwischen der Interventions- und Kontrollperiode. Sie werden lediglich effektiver eingesetzt[47].

Somit kann man diese niederländischen Studien als Vorbild für konkrete, automatisierte Empfehlungen im Patientenverwaltungssystem auf der Basis eines Risikoscores sehen. Bis dies technisch im Klinikalltag zu realisieren ist, sind SOPs mit genauen Handlungsanweisungen eine gute Alternative. [40,41,47]

5 Zusammenfassung

Ziel dieser Arbeit ist es, herauszufinden, welche Modelle zur Vorhersage von PONV die größte Validität und Generalisierbarkeit aufweisen. Sie sollten im klinischen Alltag bzw. in der Forschung eingesetzt werden. Gesucht wurden prognostische Modelle, die sich mit der Entwicklung und Validierung befassen. Zusätzlich wurden Publikationen, die die breite Anwendung und Praxisrelevanz von Scoringsystemen dokumentieren, eingeschlossen. Diese Veröffentlichungen beschreiben eine höhere Patientenzufriedenheit, eine kürzere Verweildauer im Krankenhaus oder weniger ungeplante Aufnahmen. In der Forschung werden die Modelle zur Herstellung von Kohortengleichheit in klinischen Studien verwendet und bieten die Möglichkeit Patientenkollektive zu vergleichen. Im Verlauf wurden nur Scores berücksichtigt, die in mindestens zwei Publikationen behandelt wurden. Dies traf auf 62 Publikationen zu. Sie umfassten acht Scores und ein Kollektiv von 81834 Patienten.

Bei der Überprüfung von sieben elektronischen Datenbanken im Jahre 2014 wurden 1384 Publikationen detektiert. Ein Literaturupdate im April 2018 ergab zusätzliche 391 Treffer. Nach Entfernen der Duplikate, dem Screening der Abstracts und lesen der Volltexte wurden 85 Volltexte eingeschlossen. Diese Studien wurden zur Datenverwaltung in eine Datenmaske eingepflegt.

Alle Publikationen eines Scores wurden zu einer Gruppe zusammengefasst. Zur Bewertung jedes Clusters wurden standardisierte Qualitätskriterien festgelegt. Das selbstentwickelte Bewertungstool prüfte folgende Punkte: die Generalisierbarkeit, dargestellt durch die Größe des Patientenkollektivs. Die Praktikabilität spiegelt sich wieder in der Anzahl der Risikofaktoren in dem Modell, der Risikoberechnung und dem Endpunkt des Scores. Zusätzlich wurde die Präzision d.h. die Vorhersagegenauigkeit des Modells anhand der Diskriminationsfähigkeit und der Kalibrierung bewertet. Auch interne und externe Validierung der Modelle wurden berücksichtigt. Des Weiteren wurden Punkte vergeben für die breite Anwendung in der Praxis, wenn nachgewiesen werden konnte, dass die Nutzung des Scores einen Vorteil für den Patienten oder die Forschung hatte. Abschließend gab es Extrapunkte für Aspekte die bisher nicht gewertet wurden wie z. B spezielle Maßzahlen. Die Maximale Punktzahl, die in dem Bewertungssystem erreicht werden konnte, sind 60 Punkte.

Am besten hat in den Qualitätskriterien der vereinfachter Apfelscore aus dem Jahre 1999 mit 44 Punkten abgeschnitten. Den zweiten Platz belegte der Score von Van den Bosch 2005 mit 33 Punkten. Punktgleich waren der vereinfachter Koivurantascore von 1997 sowie der Sinclairscore von 1999 mit 32,5 Punkten. Für Kinder erbrachte der POVOC-Score von Eberhart aus dem Jahr 2004 mit 31 Punkten die beste Leistung.

Abschließend wurde eine Umfrage bezüglich den SOPs zu PONV an diversen Kliniken durchgeführt. Sie teilten mit, dass bei Erwachsenen v.a. der vereinfachte Apfelscore von 1999 angewendet wird. Zusätzlich empfiehlt die Hälfte aller Zentren für Erwachsene eine generelle PONV Prophylaxe. Bei den Kindern nutzen Klinik SOPs v.a. den POVOC Score von 2004 und es wird eine generelle Prophylaxeempfehlung ausgesprochen.

6 Summary

The aim of this work is to assess which predicting models and scores for PONV possess the greatest validity and generalizability and should therefore be used in clinical practice or research. We examine prognostic models that deal with development and validation. In addition, publications describing the broad application and practical relevance of scoring systems were considered. These articles report, a higher patient satisfaction, a shorter stay in hospital or fewer unplanned admissions. In research, the models are used to establish cohort equality in clinical studies and offer the opportunity to compare patient groups. Out of this group, only the scores that were subject of at least two publications were considered for the examination. This applied to 62 publications. The studies included a total of 81834 patients and were based on eight scores.

In 2014 there were 1384 corresponding publications on seven electronic medical databases detected. A literature update in April 2018 revealed 391 additional publications. After removing the duplicates, screening the abstracts and reading the full texts 85 scientific papers from both literature searches were chosen for the next step in the selection process. These studies were entered into a data mask for further data analysis.

All publications dealing with the same score were grouped together. To analyse each cluster, standardized quality criteria were established. The self-developed evaluation tool examined and awarded points for the following properties of each publication: the generalizability, represented by the size of the patient population. The practicability is reflected in the number of the risk factors in the model, the risk calculation and the end point of each score. In addition the predictive accuracy of the model based on discrimination and calibration was assessed. The internal or external validations of a score were evaluated. Furthermore, points were awarded for broad application in practice, if it could be shown that the use of the score had an advantage for the patient or for research. In addition, extra points for aspects that have not been considered so far were assigned to each score. The maximum number of points that can be achieved in this evaluation system is 60.

Based on the quality criteria the simplified Apfelscore from 1999 achieved the best score with 44 points. The Score of Van den Bosch 2005a took the second place with 33 points. The simplified Koivuranta Score from 1997 and the Sinclair Score from 1999 were rated as equivalent with 32,5 points each. Regarding the children population the best Score was the Eberhart's POVOC Score from 2004 with 31 points.

Finally, a survey regarding the standard operating protocols (SOP) for PONV was sent to several clinics. They revealed that for adults procedures most of them use the simplified Apfelscore from 1999. Half of the centers recommend general PONV prophylaxis for adults. In the treatment of children, all available SOPs used the POVOC score from 2004. A general prophylaxis recommendation for children is provided by all clinics participating in the survey.

7 Literatur

1. Acomed Statistik (o.J.) ROC-Kurve (receiver operating characteristics). <https://www.acomed-statistik.de/roc-kurve.html>. Zugriff am 26.8.2020
2. Agarkar S, Chatterjee AS (2015) Comparison of ramosetron with ondansetron for the prevention of post-operative nausea and vomiting in high-risk patients. *Indian journal of anaesthesia* 59(4): 222–227
3. Ahmed R, Edwards J (2014) A survey of antiemetic use at Royal Alexandra Hospital, Paisley. *Anaesthesia Conference AAGBI GAT Annual Scientific Meeting 2014*. Newcastle United Kingdom. Conference Publication: 61
4. Apfel CC, Greim CA, Haubitz I et al. (1998) A risk score to predict the probability of postoperative vomiting in adults. *Acta Anaesthesiologica Scandinavica* 42(5): 495–501
5. Apfel CC, Greim CA, Goepfert C et al. (1998) Postoperatives Erbrechen. Ein Score zur Voraussage des Erbrechensrisikos nach Inhalationsanaesthesien. *Anaesthesist* 47(9): 732–740
6. Apfel CC, Greim CA, Haubitz I et al. (1998) The discriminating power of a risk score for postoperative vomiting in adults undergoing various types of surgery. *Acta Anaesthesiologica Scandinavica* 42(5): 502–509
7. Apfel CC, Läärä E, Koivuranta M et al. (1999) A Simplified Risk Score for Predicting Postoperative Nausea and Vomiting. *Anesthesiology* 91(3): 693
8. Apfel CC, Kranke P, Eberhart LHJ et al. (2002) Comparison of predictive models for postoperative nausea and vomiting. *British Journal of Anaesthesia* 88(2): 234–240
9. Apfel CC, Korttila K, Abdalla M et al. (2004) A Factorial Trial of Six Interventions for the Prevention of Postoperative Nausea and Vomiting. *New England Journal of Medicine* 350(24): 2441–2451
10. Apfel CC, Kranke P, Eberhart LHJ (2004) Comparison of surgical site and patient's history with a simplified risk score for the prediction of postoperative nausea and vomiting. *Anaesthesia* 59(11): 1078–1082
11. Apfel CC, Philip BK, Cakmakkaya OS et al. (2012) Who is at risk for postdischarge nausea and vomiting after ambulatory surgery? *Anesthesiology* 117(3): 475–486

12. Atallah FN, Riu BM, Nguyen LB et al. (2004) Boerhaave's syndrome after postoperative vomiting. *Anesthesia and Analgesia* 98(4): 1164–1166
13. Bassanezi BSB, Oliveira-Filho AG de, Jafelice RSM et al. (2013) Postoperative vomiting in pediatric oncologic patients: prediction by a fuzzy logic model. *Paediatric anaesthesia* 23(1): 68–73
14. Biedler A, Wermelt J, Kunitz O et al. (2004) A risk adapted approach reduces the overall institutional incidence of postoperative nausea and vomiting. *Canadian journal of anaesthesia = Journal canadien d'anesthesie* 51(1): 13–19
15. Blacoe DA, Cuning E, Bell G (2008) Paediatric day-case surgery: an audit of unplanned hospital admission Royal Hospital for Sick Children, Glasgow. *Anaesthesia* 63(6): 610–615
16. BREMNER WGM, KUMAR CM (1993) Delayed surgical emphysema, pneumomediastinum and bilateral pneumothoraces after postoperative vomiting. *British Journal of Anaesthesia* 71(2): 296–297
17. Challener DW, Prokop LJ, Abu-Saleh O (2019) The Proliferation of Reports on Clinical Scoring Systems: Issues About Uptake and Clinical Utility. *JAMA: The Journal of the American Medical Association* 321(24): 2405–2406
18. Cowen JS, Kelley MA (1994) Errors and Bias in Using Predictive Scoring Systems. *Critical Care Clinics* 10(1): 53–72
19. Da,Silva,H.B.G., Sousa AM, Guimaraes GMN et al. (2015) Does previous chemotherapy-induced nausea and vomiting predict postoperative nausea and vomiting? *Acta Anaesthesiologica Scandinavica* 59(9): 1145–1153
20. Eberhart L, Rüsç D, Artner J (2019) SOP - Prophylaxe und Therapie von PONV. *Universitätsklinikum Marburg, Marburg*
21. Eberhart L, Koch T, Kranke P et al. (2014) Activity-based cost analysis of opioid-related nausea and vomiting among inpatients. *Journal of Opioid Management* 10(6): 415–422
22. Eberhart LH, Seeling W, Staack AM et al. (1999) Validierung eines Risikoscores zur Vorhersage von Erbrechen in der postoperativen Phase. *Anaesthesist* 48(9): 607–612
23. Eberhart LH, Högel J, Seeling W et al. (2000) Evaluation of three risk scores to predict postoperative nausea and vomiting. *Acta Anaesthesiologica Scandinavica* 44(4): 480–488

24. Eberhart LHJ, Morin AM, Guber D et al. (2004) Applicability of risk scores for postoperative nausea and vomiting in adults to paediatric patients. *British Journal of Anaesthesia* 93(3): 386–392
25. Eberhart LHJ, Geldner G, Kranke P et al. (2004) The development and validation of a risk score to predict the probability of postoperative vomiting in pediatric patients. *Anesthesia and Analgesia* 99(6): 1630-7, table of contents
26. Engel JM, Junger A, Hartmann B et al. (2006) Performance and customization of 4 prognostic models for postoperative onset of nausea and vomiting in ear, nose, and throat surgery. *Journal of Clinical Anesthesia* 18(4): 256–263
27. Franck M, Radtke FM, Baumeyer A et al. (2010) Einhaltung der Behandlungsrichtlinien für postoperative Übelkeit und Erbrechen. Wie gut gelingt der Wissenstransfer hin zu einer besseren klinischen Versorgung? *Der Anaesthesist* 59(6): 524–528
28. Gan TJ, El-Molem H, Ray J et al. (1999) Patient-controlled Antiemesis. *Anesthesiology* 90(6): 1564–1570
29. Gan TJ, Kranke P, Minkowitz HS et al. (2017) Intravenous Amisulpride for the Prevention of Postoperative Nausea and Vomiting: Two Concurrent, Randomized, Double-blind, Placebo-controlled Trials. *Anesthesiology* 126(2): 268–275
30. Geersing G-J, Bouwmeester W, Zuithoff P et al. (2012) Search filters for finding prognostic and diagnostic prediction studies in Medline to enhance systematic reviews. *PloS one* 7(2): e32844
31. google (o.J.) google drive. https://www.google.com/intl/de_ALL/drive/. Zugriff am 26.8.2020
32. google (o.J.) google translate. <https://translate.google.com/?hl=de>. Zugriff am 26.8.2020
33. Halliday TA, Sundqvist J, Hultin M et al. (2017) Post-operative nausea and vomiting in bariatric surgery patients: an observational study. *Acta Anaesthesiologica Scandinavica* 61(5): 471–479
34. Held L, Rufibach K, Seifert B (2013) *Medizinische Statistik*. Pearson, München
35. Irefin SA, Farid IS, Senagore AJ (2000) Urgent Colectomy in a Patient with Membranous Tracheal Disruption After Severe Vomiting. *Anesthesia & Analgesia* 91(5): 1300–1302

36. Junger A, Hartmann B, Benson M et al. (2001) The use of an anesthesia information management system for prediction of antiemetic rescue treatment at the postanesthesia care unit. *Anesthesia and Analgesia* 92(5): 1203–1209
37. Kaneda K, Sakuma Y, Terano T, Matsuda Y, Kan Y, Murata K, Yokoyama Y, Fukui T, Kotani J (2004) Postoperative Nausea and Vomiting in Oral and Maxillofacial Surgery Operations Using VIMA. *Journal of Japanese Dental Society of Anesthesiology* 32(3): 305–310
38. Kapoor R, Hola ET, Adamson RT et al. (2008) Comparison of two instruments for assessing risk of postoperative nausea and vomiting. *Am J Health Syst Pharm* 65(5): 448–453
39. Kappen TH, Vergouwe Y, van Klei WA et al. (2012) Adaptation of clinical prediction models for application in local settings. *Medical decision making : an international journal of the Society for Medical Decision Making* 32(3): E1-10
40. Kappen TH, Moons KGM, Van, Wolfswinkel, L. et al. (2014) Impact of risk assessments on prophylactic antiemetic prescription and the incidence of postoperative nausea and vomiting. *Anesthesiology* 120(2): 343–354
41. Kappen TH, Vergouwe Y, VanWolfswinkel L et al. (2015) Impact of adding therapeutic recommendations to risk assessments from a prediction model for postoperative nausea and vomiting. *British Journal of Anaesthesia* 114(2): 252–260
42. Kappen TH, Van, Loon, K., Kappen MAM et al. (2016) Barriers and facilitators perceived by physicians when using prediction models in practice. *Journal of clinical epidemiology* 70: 136–145
43. Klenke S, de, Vries, G.J., Schiefer L et al. (2018) CHRM3 rs2165870 polymorphism is independently associated with postoperative nausea and vomiting, but combined prophylaxis is effective. *British Journal of Anaesthesia* 121(1): 58–65
44. Klinikum Augsburg (2013) Klinik für Anästhesie und Operative Intensivmedizin - PONV. Klinikum Augsburg, Augsburg
45. Klinikum Würzburg Mitte Standort Juliusspital (unbekannt) Verfahrensanweisung Postoperative Übelkeit und Erbrechen. Klinikum Würzburg Mitte Standort Juliusspital, Würzburg
46. Koivuranta M, Läärä E, Snåre L et al. (1997) A survey of postoperative nausea and vomiting. *Anaesthesia* 52(5): 443–449

47. Kooij FO, Vos N, Siebenga P et al. (2012) Automated reminders decrease postoperative nausea and vomiting incidence in a general surgical population. *British Journal of Anaesthesia* 108(6): 961–965
48. Kranke P (unbekannt) SOP PONV Prophylaxe und Therapie Frauenklinik. Universitätsklinikum Würzburg, Würzburg
49. Kranke P, Eberhart LH, Toker H et al. (2007) A prospective evaluation of the POVOC score for the prediction of postoperative vomiting in children. *Anesthesia and Analgesia* 105(6): 1592-7, table of contents
50. Kredel M, Goepfert C, Bassi D et al. (2006) The influence of the weather and the phase of the moon on post-operative nausea and vomiting. *Acta Anaesthesiologica Scandinavica* 50(4): 488–494
51. Kumar A, Brampton W, Watson S et al. (2012) Postoperative nausea and vomiting: simple risk scoring does work. *European journal of anaesthesiology* 29(1): 57-9; author reply 59-60
52. Latz B, Mordhorst C, Kerz T et al. (2011) Postoperative nausea and vomiting in patients after craniotomy: incidence and risk factors. *Journal of Neurosurgery* 114(2): 491–496
53. Lee YZ, Lee RQ, Thinn KK et al. (2015) How patients fare after anaesthesia for elective surgery: A survey of postoperative nausea and vomiting, pain and confusion. *Singapore Medical Journal* 56(1): 40–46
54. Low C. (1973) Prediction of immediate postoperative vomiting. *Abstracts of hospital management studies*: 11042–11054
55. Lucke K (2016) SOP PONV - Prophylaxe und Therapie in der Anästhesie. Diakonie Klinikum Stuttgart, Stuttgart
56. Moons K, de Groot J, Bouwmeester W, Vergouwe Y, Mallett S, Altman D, Reitsma J, Collins G (o.J.) Guidelines and Guidance Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist. <https://journals.plos.org/plosmedicine/article/file?id=10.1371/journal.pmed.1001744&type=printable>. Zugriff am 3.12.2020
57. Moons KGM, Groot JAH de, Bouwmeester W et al. (2014) Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med* 11(10): e1001744

58. Odom-Forren J, Jalota L, Moser DK et al. (2013) Incidence and predictors of postdischarge nausea and vomiting in a 7-day population. *Journal of Clinical Anesthesia* 25(7): 551–559
59. Pace NL, Carlisle J, Eberhart LH et al. (2014) Prediction models for the risk of postoperative nausea and vomiting. *Cochrane Database of Systematic Reviews* 2014(9): 1–26
60. Palazzo M, Evans R (1993) Logistic regression analysis of fixed patient factors for postoperative sickness: a model for risk assessment. *British Journal of Anaesthesia* 70(2): 135–140
61. Patel RI, Hannallah RS (1988) Anesthetic Complications following Pediatric Ambulatory Surgery. *Anesthesiology* 69(6): 1009–1012
62. Peng SY, Wu KC, Wang JJ et al. (2007) Predicting postoperative nausea and vomiting with the application of an artificial neural network. *British Journal of Anaesthesia* 98(1): 60–65
63. Pfenninger E, von Normann V, Gauss A (unbekannt) SOP - Postoperative Übelkeit und Erbrechen. Universitätsklinikum Ulm, Ulm
64. Pierre S, Benais H, Pouymayou J (2002) Apfel's simplified score may favourably predict the risk of postoperative nausea and vomiting. *Canadian journal of anaesthesia = Journal canadien d'anesthésie* 49(3): 237–242
65. Pierre S, Corno G, Benais H et al. (2004) A risk score-dependent antiemetic approach effectively reduces postoperative nausea and vomiting — a continuous quality improvement initiative. *Canadian Journal of Anesthesia/Journal canadien d'anesthésie* 51(4): 320–325
66. Pluta M, Miszczenkow H, Liberski P et al. (2017) Usefulness of Apfel score in predicting postoperative nausea and vomiting (PONV): Single-center experience. *Anaesthesiology Intensive Therapy Conference 19. Międzynarodowy Zjazd Polskiego Towarzystwa Anestezjologii i Intensywnej Terapii*. Poland. 49: 31
67. Pospiech A, Scholtes J-L, De,Kock,M. (2014) Should we systematically calculate the simplified Apfel score? *European journal of anaesthesiology Conference European Anaesthesiology Congress, EUROANAESTHESIA 2014*. Stockholm Sweden. Conference Publication: 18
68. rayyan (o.J.) Rayyan QCRI. <https://rayyan.qcri.org/welcome>. Zugriff am 26.8.2020

69. Riley RD, Hayden JA, Steyerberg EW et al. (2013) Prognosis Research Strategy (PROGRESS) 2: Prognostic Factor Research. *PLoS Med* 10(2): e1001380
70. Roberts A, Barclay P, Stott M (2010) Improving the accuracy of risk assessment in postoperative nausea and vomiting. *Anaesthesia* 65(8): 863–864
71. Rodseth RN, Gopalan PD, Cassimjee HM et al. (2010) Reduced incidence of postoperative nausea and vomiting in black South Africans and its utility for a modified risk scoring system. *Anesthesia and Analgesia* 110(6): 1591–1594
72. Rüsç D, Bernhardt J, Wulf H (1999) Prophylaxe von Übelkeit und Erbrechen nach Pelviskopien. Dolasetron oder MCP im Vergleich zu Plazebo. *Anaesthesist* 48(10): 705–712
73. Rüsç D, Eberhart L, Biedler A et al. (2005) Prospective application of a simplified risk score to prevent postoperative nausea and vomiting. *Canadian journal of anaesthesia = Journal canadien d'anesthesie* 52(5): 478–484
74. Rüsç D, Eberhart LHJ, Wallenborn J et al. (2010) Nausea and vomiting after surgery under general anesthesia: an evidence-based review concerning risk assessment, prevention, and treatment. *Dtsch Arztebl Int* 107(42): 733–741
75. Sarin P, Urman RD, Ohno-Machado L (2012) An improved model for predicting postoperative nausea and vomiting in ambulatory surgery patients using physician-modifiable risk factors. *J Am Med Inform Assoc* 19(6): 995–1002
76. Sawatzky JAV, Rivet M, Ariano RE et al. (2014) Post-operative nausea and vomiting in the cardiac surgery population: Who is at risk? *Heart and Lung: Journal of Acute and Critical Care* 43(6): 550–554
77. Schaefer MS, Apfel CC, Sachs H-J et al. (2015) Predictors for postoperative nausea and vomiting after xenon-based anaesthesia. *British Journal of Anaesthesia* 115(1): 61–67
78. Schumann R, Polaner DM (1999) Massive Subcutaneous Emphysema and Sudden Airway Compromise After Postoperative Vomiting. *Anesthesia & Analgesia* 89(3): 796
79. Schwenk W (2018) SOP Prophylaxe und Therapie von postoperativer Übelkeit und Erbrechen (PONV). *Allgemein- und Viszeralchirurgie up2date* 12(04): 326–328
80. Sinclair DR, Chung F, Mezei G (1999) Can Postoperative Nausea and Vomiting Be Predicted? *Anesthesiology* 91(1): 109–118

81. Steyerberg EW, Moons KGM, van der Windt DA et al. (2013) Prognosis Research Strategy (PROGRESS) 3: Prognostic Model Research. *PLoS Med* 10(2): e1001381
82. Sundqvist J, Wallden J (2014) PONV in bariatric surgery. *European journal of anaesthesiology Conference European Anaesthesiology Congress, EUROANAESTHESIA 2014*. Stockholm Sweden. Conference Publication: 16–17
83. Thomas R, Jones NA, Strike P (2002) The value of risks scores for predicting postoperative nausea and vomiting when used to compare patient group in a randomised controlled trial. *Anaesthesia* 57(11): 1119–1128
84. Toner CC, Broomhead CJ, Littlejohn IH et al. (1996) Prediction of postoperative nausea and vomiting using a logistic regression model. *British Journal of Anaesthesia* 76(3): 347–351
85. Traeger M, Eberhart A, Geldner G et al. (2003) Vorhersage von Ubelkeit und Erbrechen in der postoperativen Phase durch ein künstliches neuronales Netz. *Der Anaesthesist* 52(12): 1132–1138
86. Tropé A, Raeder JC (2000) Kan postoperativ kvalme eller oppkast forutsies? *Tidsskrift for den Norske laegeforening* 120(20): 2423–2426
87. van den Bosch JE, Kalkman CJ, Vergouwe Y et al. (2005) Assessing the applicability of scoring systems for predicting postoperative nausea and vomiting. *Anaesthesia* 60(4): 323–331
88. van den Bosch JE, Moons KG, Bonse GJ et al. (2005) Does measurement of preoperative anxiety have added value for predicting postoperative nausea and vomiting? *Anesthesia and Analgesia* 100(5): 1525-32, table of contents
89. von Loeffelholz C, Geil D (2020) *Verfahrensanweisung PONV- Prophylaxe und Therapie bei Erwachsenen und Kindern*. Universitätsklinikum Jena, Jena
90. Wallden J, Flodin J, Hultin M (2016) Validation of a prediction model for post-discharge nausea and vomiting after general anaesthesia in a cohort of Swedish ambulatory surgery patients. *European journal of anaesthesiology* 33(10): 743–749
91. Walther (2019) *SOP - Prophylaxe und Therapie von PONV*. Klinikum Stuttgart, Stuttgart
92. Wasson JH, Sox HC, Neff RK et al. (1985) Clinical Prediction Rules. *New England Journal of Medicine* 313(13): 793–799

93. Weilbach C, Rahe-Meyer N, Raymondos K et al. (2006) Postoperative nausea and vomiting (PONV) : usefulness of the Apfel-score for identification of high risk patients for PONV. *Acta Anaesthesiologica Belgica* 57(4): 361–363
94. White PF, Sacan O, Nuangchamnonng N et al. (2008) The relationship between patient risk factors and early versus late postoperative emetic symptoms. *Anesthesia and Analgesia* 107(2): 459–463
95. wikipedia (o.J.) Scoring (Medizin). [https://de.wikipedia.org/wiki/Scoring_\(Medizin\)](https://de.wikipedia.org/wiki/Scoring_(Medizin)). Zugriff am 17.8.2021
96. Wolf R, Morinello E, Kestler G et al. (2016) PONV after strabismus surgery: Risk adapted prophylaxis? [German]. *Anaesthesist* 65(7): 507–513
97. Wu YH, Sun HS, Wang ST et al. (2015) Applicability of risk scores for postoperative nausea and vomiting in a Taiwanese population undergoing general anaesthesia. *Anaesthesia and intensive care* 43(4): 473–478
98. Zhang GS, Mathura JR (2005) Painless Loss of Vision after Vomiting. *New England Journal of Medicine* 352(17): e16

8 Anhang

8.1 Suchbaum für Datenbanken bei Literatursuche

The Cochrane Library search strategy[59]

#1 MeSH descriptor: [Risk Assessment] explode all trees

#2 MeSH descriptor: [Models,Statistical] explode all trees

#3(((predict*ormulticomponent or multivariable) near model*) or (predict* near (outcome* or risk*or model*)) or ((history or variable*or criteria or scor* or characteristic* or finding* or factor* orvalue*) near (predict* or model* or decision* or identif* or prognos*)) or (decision* near (model* or clinical* or logistic model*)) or (prognostic near (history or variable* or criteria or scor* or characteristic* or finding* or factor* or model*)) or (observ*near(variation or model*)))

#4 (#1or #2 or #3) not pain

#5 MeSH descriptor: [Postoperative Nausea and Vomiting] explore all trees

#6 ((postoperative or postdischarge or discharge* or post?an?esth*) near (nausea or vomit*)) or PONV

#7 #5 or #6

#8 #4 and #7

OvidMEDLINESearchStrategy[59]

1. (Risk Assessment/ or Models, Statistical/ or (predict* or validat* or rule* or scor*).ti,ab. or ((predict* or multicomponent or multivariable) adj3 model*).mp. or (predict* adj5 (outcome* or risk* or model*)).ti,ab. or ((history or variable* or criteria or scor* or characteristic* or finding* or factor* or value*) adj5 (predict* or model* or decision* or identif* or prognos*)).ti,ab. or (decision* adj5 (model* or clinical* or logistic model*)).ti,ab. or (prognostic adj5 (history or variable* or criteria or scor* or characteristic* or finding* or factor* or model *)).ti,ab. or (observ* adj3 (variation or model *)).ti,ab.) not pain.mp.

2. "Postoperative Nausea and Vomiting"/ or ((postoperative or postdischarge or discharge* or post?an?esth*) adj5 (nausea or vomit*)) .ti,ab. or PONV.mp.

3. 1 and 2

Ovid EM BASE search strategy[59]

1. (risk assessment/ or statistical model/ or prediction/ or predictive value/ or predictive validity/ or (predict* or validat* or rule* or scor*).ti,ab. or ((predict* or multicomponent or multivariable) adj3 model*).mp. or (predict* adj5 (outcome* or risk* or model *)).ti,ab. or

((history or variable* or criteria or scor* or characteristic* or finding* or factor* or value*) adj5 (predict* or model* or decision* or identif* or prognos*)).ti,ab. or (decision* adj5 (model* or clinical* or logistic model*)).ti,ab. or (prognostic adj5 (history or variable* or criteria or scor* or characteristic* or finding* or factor* or model *)).ti,ab. or (observ* adj3 (variation or model*)).ti,ab.) not pain .mp.

2. postoperative nausea/ or postoperative vomiting/ or "postoperative nausea and vomiting"/ or ((postoperative or postdischarge or discharge* or post?an?esth*) adj5 (nausea or vomit*)).ti,ab. or PONV.mp.

3. 1 and 2

ISI Web of Science search strategy[59]

1 TS=(risk assessment or (model* SAME statistical) or (predict* or validat* or rule* or scor*) or ((predict* or multicomponent or multivariable) SAME model*) or (predict* SAME (outcome* or risk* or model*)) or ((history or variable* or criteria or scor* or characteristic* or finding* or factor* or value*) SAME (predict* or model* or decision* or identif* or prognos*)) or (decision* SAME (model* or clinical* or logistic model*)) or (prognostic SAME (history or variable* or criteria or scor* or characteristic* or finding* or factor* or model *)) or (observ* SAME (variation or model*))) not TS=pain

#2 TS=(((postoperative or postdischarge or discharge* or post?an?esth*) SAME (nausea or vomit*)) or PONV)

#3 # 1 and #2

CINAHL(EBSCOhost)searchstrategy[59]

SI ((MM "Risk Assessment") OR (MM "Models, Statistical") or TI ((predict* or validat* or rule* or scor*) or ((predict* or multicomponent or multivariable) N3 model *) or (predict* N3 (outcome* or risk* or model *)) or ((history or variable* or criteria or scor* or characteristic* or finding* or factor* or value*) N3 (predict* or model * or decision* or identif* or prognos*)) or (decision* N3 (model* or clinical* or logistic model *)) or (prognostic N3 (history or variable* or criteria or scor* or characteristic* or finding* or factor* or model*)) or (observ* N3 (variation or model *))) or AB ((predict* or validat* or rule* or scor*) or ((predict* or multicomponent or multivariable) N3 model *) or (predict* N3 (outcome* or risk* or model *)) or ((history or variable* or criteria or scor* or characteristic* or finding* or factor* or value*) N3 (predict* or model * or decision* or identif* or prognos*)) or (decision* N3 (model* or clinical* or logistic model*)) or

(prognostic N3 (history or variable* or criteria or scor* or characteristic* or finding* or factor* or model *)) or (observ* N3 (variation or model *)))) NOT TX pain

S2 MM "Nausea and Vomiting+" OR ((postoperative or postdischarge or discharge* or post?an?esth*) N3 (nausea or vomit*)) OR TX PONV

S3 S1 AND S2

LILACS(BIREME)searchstrategy[59]

(risk assessment\$ or statistical model\$ or predict\$ or validat\$ or rule\$ or scor\$ or ((predict\$ or multicomponent or multivariable) and model\$) or (predict\$ and (outcome\$ or risk\$ or model\$)) or ((history or variable\$ or criteria or scor\$ or characteristic\$ or finding\$ or factor\$ or value\$) and (predict\$ or model\$ or decision\$ or identif\$ or prognos\$)) or (decision\$ and (model\$ or clinical\$ or logistic model\$)) or (prognostic and (history or variable\$ or criteria or scor\$ or characteristic\$ or finding\$ or factor\$ or model\$)) or (observ\$ and (variation or model\$))) and (((postoperative or postdischarge or discharge\$ or post?an?esth\$) and (nausea or vomit\$)) or PONV)

8.2 Inhalt der Datenmaske aus Excel

Art der Publikation

- Modellentwicklung mit interner Validierung
- Modellentwicklung ohne interne Validierung
- externe Validierung mit Update
- externe Validierung ohne Update
- Impact Assessment

Datenquelle

- Kohorte
- Fallkontrollstudie
- randomisierte Studienteilnehmer
- Registerdaten
- andere Quelle (Freitext)

Qualität der Daten

- prospektive Daten
- Convenient sample (= praktische/zweckmäßige Probe)
- retrospektive Daten (chart review)
- unklar (Freitext)

Studiendaten

- Standort der Teilnehmer (Freitext)
- Anzahl an Zentren (Freitext)
- Studiensetting: stationäre Patienten
- Studiensetting: ambulante Patienten
- Studiensetting: unklar

Einschlusskriterien

- Erwachsene
- Kinder
- Art der eingeschlossenen Operationen (Freitext)
- sonstige Zusatzinformationen (Freitext)

Ausschlusskriterien

- (Freitext)

zusätzliche Beschreibung der Patienten

- (Freitext)

antiemetische Behandlung

- ja
- ja: Medikament (Freitext)
- ja: Dosis (Freitext)
- ja: Art der Verabreichung (Freitext)
- ja: Zeit der Verabreichung (Freitext)
- nein
- unklar
- ein Teil der Patienten

Definitionen der Endpunkte

- (Freitext)

Art des Endpunktes

- einzelner Endpunkt
- zusammengesetzter Endpunkt
- unklar

Ort der Endpunkterhebung

- Aufwachstation
- stationär (noch im KHK)
- ambulant (nach Entlassung)
- unklar

Zeitpunkt der Endpunkterhebung

- 6 Stunden nach OP
- 12 Stunden nach OP
- 24 Stunden nach OP
- 48 Stunden nach OP
- anderer Zeitpunkt nach der OP (Freitext)

Art und Weise der Patientenbewertung für die Endpunkte

- routinemäßig erhobene Daten
- Fragebogen
- Visite des Patienten
- unklar

Erfassung folgender Prädiktoren

- demografische Prädiktoren
- Prädiktoren aus der Patientengeschichte

- Prädiktoren aus der körperlichen Untersuchung
- Laborwerte, -tests, OP-Daten
- andere Prädiktoren

Erfassung folgender Fakten pro Prädiktor

- Anzahl
- Benennung/ namentlich (Freitext)
- genaue Definition (Freitext)
- Reliabilität: keine, niedrig, mittel, hoch
- Art und Weise der Datenerfassung: Fragebogen, mündlich, unklar
- Zeitpunkt der Erfassung des Prädiktors: präop, intraop, postop, unklar
- Beurteilung des Prädiktors ohne Kenntnis des Outcomes: ja, nein, unklar
- Handhabung von Prädiktoren in dem Model: kategorisch (binär), kategorisch (nominal), kontinuierlich, linear, nicht-lineare Transformation, unklar

Stichprobengröße

- Anzahl an Teilnehmern insgesamt (Freitext)
- Anzahl an Teilnehmern Entwicklungsdatensatz (Freitext)
- Anzahl an Teilnehmern Validierungsdatensatz (Freitext)
- Anzahl der Teilnehmer mit Outcome (Events) insgesamt (Freitext)
- Anzahl der Teilnehmer mit Outcome (Events) Entwicklungsdatensatz (Freitext)
- Anzahl der Teilnehmer mit Outcome (Events) Validierungsdatensatz (Freitext)
- Anzahl an Prädiktoren (Variablen) (Freitext)
- Events pro Variable (EPV) (Freitext)

Anzahl an Teilnehmern mit fehlenden Daten

- Anzahl (%) mit irgendwelchen fehlenden Daten (Freitext)
- Anzahl (%) pro Prädiktor (Freitext)
- nicht berichtet

Angaben über den Umgang mit fehlenden Daten

- Complete-case analysis (vollständige Fallanalyse)
- Imputation (Anrechnungsverfahren)
- andere Methoden (Freitext)
- unbekannt

Methodik bei der Modellbildung

- logistische Regression
- survival (Fortbestehen)
- neuronales Netzwerk

- maschinelles Lernen
- andere Methode (Freitext)

Auswahl der Prädiktoren bei der Modellbildung

- alle Prädiktoren
- univariablen Screening
- content Expertise (inhaltliche Kompetenz)
- andere (Freitext)

Einschluss/Ausschluss von Prädiktoren während der Modellbildung

- beibehalten aller Prädiktoren
- backward Selection (Rückwärtsselektion)
- forward Selection (Vorwärtsselektion)
- forward/backward Selection (Vorwärts/Rückwärtsselektion)
- andere (Freitext)

Kriterien für die Prädiktoreliminierung

- p- Wert
- Akaike Information Criterion (AIC)
- Bayesian Information Criterion (BIC)
- c-Index
- andere (Freitext)

Modellkalibrierung

- Calibration Plot (Kalibrierungsdiagramm)
- Calibration Slope (Kalibrierungssteigung)
- Hosmer-Lemeshow-Test
- andere (Freitext)

Diskriminationsfähigkeit des Modells

- C-Statistik /Index (AUC of ROC-Curve)
- D-Statistik
- Logrank-Test
- andere (Freitext)

Maßzahlen zur Modell Klassifizierung

- A-priori Cutpoints (grundsätzlicher Trennwert)
- Sensitivität
- Spezifität
- positiver prädiktiver Wert (PPV= Positiv Predictive Value)

- negativer prädiktiver Wert (NPV= Negative Predictive Value)
- andere (Freitext)

Modellprüfung mit einem Trainingsdatensatz

- keine Testung
- Splitting der Daten
- Bootstrapping
- Kreuzvalidierung
- andere (Freitext)

Modellprüfung mit externen Daten

- zeitliche Validierung
- geographische Validierung
- Validierung mit einem anderen Setting
- Validierung mit einem anderen Untersucher
- andere (Freitext)

Überarbeitung des schlecht validierten Modells

- keine Anpassung
- Achsenabschnitt rekaliert
- Anpassung des Prädiktor-Effekts
- neue Prädiktoren hinzugefügt
- andere (Freitext)

Präsentation des endgültigen Modells

- ausgewählte Prädiktoren (Freitext)
- Regressionskoeffizienten (Freitext)
- gerundete (vereinfachte) Prädiktorwerte (Freitext)
- Achsenabschnitt (Freitext)
- Maßzahlen für die Modelleistung (Enter via Info Box)
- vollständige Gleichung (Freitext)

alternative endgültige Modellpräsentation

- keine
- vereinfachter Vorhersagescore (Freitext)
- Nomogramm (=Rechentafel)
- Score chart (=Schaubild)
- Prädiktoren für Risikountergruppen (Freitext)
- andere (Freitext)

Anwendbarkeit in der täglichen Praxis/Konsequenz für die Praxis(persönliche Meinung)

- ja (Freitext)
- nein (Freitext)
- andere (Freitext)

Auswirkungen für die Forschung (persönliche Meinung)

- ja (Freitext)
- nein (Freitext)
- andere (Freitext)

8.3 Lebenslauf

Diese Seite enthält persönliche Daten, sie ist deshalb nicht Bestandteil der Veröffentlichung.

8.4 Verzeichnis der akademischen Lehrer

Meine akademischen Lehrenden waren in Marburg:

Adamkiewicz, Albert, Alter, Arabin, Aust, Barth, Bartsch, Bauer, Baum, Baumann, Becker, Best, Bette, Bien, Boekhoff, Bolander, Bonaterra, Bösner, Bücking, Burchert, Cetin, Cordes, Czubayko, Daut, Decher, Del Rey, Dietz, Dodel, Donner-Banzhoff, Duda, Efe, El-Zayat, Fendrich, Feuser, Fuchs-Winkelmann, Geks, Görg, Greene, Gress, Greulich, Grimm, Grundmann, Halaszovich, Hegele, Hertl, Heverhagen, Heyse, Höffken, Holland, Hoyer, Hundt, Jerrentrup, Josephs, Kalder, Kann, Kanngießer, Kill, Kinscherf, Kircher, Kirschbaum, Klose, Koczulla, Koehler, Köhler, König, Koolmann, Kroh, Kühnert, Kwee, Lang, Librizzi, Lill, Lohoff, Mahnken, Maier, Maschuw, Maurer, Metzelder, Moll, Moosdorf, Mueller, Müller, Mutters, Neff, Neubauer, Neumüller, Nikolaizik, Nimphius, Nimsky, Oberkircher, Oberwinkler, Olbert, Oliver, Opitz, Örtel, Pagenstecher, Pankatz, Patrascan, Peterlein, Pfestroff, Plant, Preisig-Müller, Printz, Ramaswamy, Reese, Renke, Renz, Riera Knorrenschild, Rolfes, Rost, Rothmund, Ruchholz, Sahmland, Schäfer, Schieffer, Schierl, Schmidt, Schneider, Schrott, Schütz, Schwarz, Seitz, Sekundo, Sevinc, Sommer, Steininger, Strik, Swaid, Tackenberg, Teymoortash, Thieme, Timmesfeld, Vannucchi, Vogelmeier, Vogt, Wagner, Weihe, Weiß, Werner, Westermann, Wiegand, Wilhelm, Wissniowski, Worzfeld, Wroklage, Wulff, Zaraca, Zemlin, Ziller, Zoremba, Zovko, Zwiorek

8.5 Danksagung

Ein besonderer Dank gilt Herrn Prof. Dr. Leopold Eberhart, der mir stets freundlich und sehr engagiert mit Rat und Tat zur Seite stand und auch nach vielen Jahren die Geduld und das Interesse an dieser Arbeit nichtverlor.

Diese Arbeit wäre ohne die mentale und finanzielle Unterstützung meiner Eltern, auf die ich mich immer verlassen konnte, nicht möglich gewesen.

8.6 Ehrenwörtliche Erklärung

„Ich erkläre ehrenwörtlich, dass ich die dem Fachbereich Medizin Marburg zur Promotionsprüfung eingereichte Arbeit mit dem Titel „Vorhersagemodelle für das Risiko von postoperativer Übelkeit und Erbrechen – Eine qualitative Analyse aller verfügbaren Vorhersagemodelle für das Auftreten von Übelkeit und Erbrechen nach Narkosen (PONV)“ in der Klinik für Anästhesie und Intensivtherapie unter Leitung von Prof. Dr. med. Hinnerk Wulf mit Unterstützung durch Prof. Dr. med. Leopold Eberhart ohne sonstige Hilfe selbst durchgeführt und bei der Abfassung der Arbeit keine anderen als die in der Dissertation aufgeführten Hilfsmittel benutzt habe. Ich habe bisher an keinem in- oder ausländischen Medizinischen Fachbereich ein Gesuch um Zulassung zur Promotion eingereicht, noch die vorliegende oder eine andere Arbeit als Dissertation vorgelegt.

Ich versichere, dass ich sämtliche wörtlichen oder sinngemäßen Übernahmen und Zitate kenntlich gemacht habe.

Mit dem Einsatz von Software zur Erkennung von Plagiaten bin ich einverstanden.

Ort, Datum, Unterschrift Doktorandin

„Die Hinweise zur Erkennung von Plagiaten habe ich zur Kenntnis genommen.“*2

Ort, Datum, Unterschrift Referent
