

Aus dem Zentrum für Innere Medizin der Universität zu Köln
Klinik und Poliklinik für Innere Medizin I
Direktor: Universitätsprofessor Dr. med. M. Hallek

Antibiotikatherapien bei Hospitalisation
—
**Analyse von Prädiktoren auf PatientInnenebene
sowie Score-basierte Risikoabschätzung
(PILGRIM-Score)**

Inaugural-Dissertation zur Erlangung der Doktorwürde
der Medizinischen Fakultät
der Universität zu Köln

vorgelegt von
Moritz Beck
aus Lahnstein

promoviert am 28. November 2024

Gedruckt mit Genehmigung der Medizinischen Fakultät der Universität zu Köln
2025

Dekan: Universitätsprofessor Dr. med. G. R. Fink

1. Gutachter: Professor Dr. med. J. J. Vehreschild
2. Gutachter: Professor Dr. rer. medic. M. Hellmich

Erklärung

Ich erkläre hiermit, dass ich die vorliegende Dissertationsschrift ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskriptes habe ich Unterstützungsleistungen von folgenden Personen erhalten:

Herr Univ.-Prof. Dr. med. Jörg Janne Vehreschild
Frau Dr. med. Annika Claßen
Frau Carolin Koll, PhD

Weitere Personen waren an der Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich nicht die Hilfe einer Promotionsberaterin/eines Promotionsberaters in Anspruch genommen. Dritte haben von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertationsschrift stehen.

Die Dissertationsschrift wurde von mir bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Der dieser Arbeit zugrunde liegende Datensatz wurde im Rahmen der multizentrischen, internationalen PILGRIM-Studie unter der Leitung der Forschungsgruppe *Kohorten in der Infektionsforschung* der Universität zu Köln (Herr Univ.-Prof. Dr. med. Vehreschild) erhoben. Die Datenerhebung fand ohne meine Mithilfe statt.

Die grundlegende Konzeption der statistischen Auswertung erfolgte in Zusammenarbeit mit Frau Koll, PhD und Absprache mit Frau Dr. med. Claßen sowie Herrn Univ.-Prof. Dr. med. Vehreschild. Die gesamte statistische Auswertung sowie Entwicklung des PILGRIM-Scores fanden durch mich statt. Alle statistischen Analysen wurden mit dem *Statistical Package for Social Sciences* (SPSS®) der International Business Machine Corporation (IBM Inc.®) Version 28.0.0.0 durchgeführt.

Regelmäßige Besprechungen mit Frau Koll, PhD, Frau Dr. med. Claßen und Herrn Univ.-Prof. Dr. med. Vehreschild halfen das Modell weiterzuentwickeln und das statistische Ergebnis zu präzisieren.

Erklärung zur guten wissenschaftlichen Praxis:

Ich erkläre hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten (Amtliche Mitteilung der Universität zu Köln AM 132/2020) der Universität zu Köln gelesen habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen.

Köln, den 12.02.2024

Moritz Beck

Danksagung

Hiermit möchte ich mich herzlich bei allen Personen bedanken, die mich im Rahmen meiner Dissertation unterstützt haben.

Dies gilt im Besonderen für Frau Dr. med. Annika Claßen, Frau Carolin Koll, PhD und Herrn Univ.-Prof. Dr. med. Jörg Janne Vehreschild, die mit ihrem offenen Ohr, konstruktiven Vorschlägen und stets ausgezeichneter Betreuung den Abschluss dieser Arbeit ermöglicht haben.

Inhaltsverzeichnis

ABKÜRZUNGSVERZEICHNIS	7
1. ZUSAMMENFASSUNG	8
2. EINLEITUNG	10
2.1. Folgen des Antibiotikaeinsatzes	10
1.1.1. PatientInnen-individuelle Ebene	10
1.1.2. Herausforderungen durch AMR auf globaler Ebene	11
1.1.3. Resistenzmechanismen	12
1.1.4. Resistenz-Surveillance	12
1.2. Antimicrobial Stewardship (AMS)	13
1.3. Prädisponierende Faktoren für eine Antibiotikatherapie	14
1.4. Ziel dieser Arbeit	15
2. MATERIAL UND METHODEN	16
2.1. PILGRIM Studie	17
2.2. Primärkodierung des Datensatzes	18
2.2.1. Datenbereinigung	18
2.2.2. Kodierung zusätzlicher Variablen	19
2.2.3. Endpunkt-Definition	20
2.3. Portionierung des Datensatzes	21
2.4. Deskriptive Statistik	21
2.5. Explorative Datenanalyse: Entwicklung des PILGRIM-Score	22
2.5.1. Identifizierung von Risikofaktoren	22
2.5.2. 10-fache Filterung des Entwicklungsdatensatzes (Implementation der 10-fachen Kreuzvalidierung)	22
2.5.3. Entwicklung des optimalen binär logistischen Regressionsmodells	23
2.5.4. Entwicklung des PILGRIM-Score aus dem logistischen Regressionsmodell	25
2.5.5. Validierung des Pilgrim Score	26

3. ERGEBNISSE	28
3.1. Primärkodierung	28
3.2. Deskriptive Statistiken	28
3.2.1. Deskription des Entwicklungsdatensatzes	28
3.2.2. Deskription des Validierungsdatensatzes	29
3.3. Risikofaktorenanalyse	30
3.3.1. Pearson Chi-Quadrat-Test	30
3.3.2. Univariate Regressionsmodelle	32
3.4. Binäre logistische Regressionsmodelle	35
3.5. Entwicklung der optimalen Kombination	37
3.6. Überprüfung auf Multikollinearität	39
3.7. Validierung des finalen logistischen Regressionsmodells	39
3.8. PILGRIM Score – Berechnung der individuellen Variablen gewichtung	40
3.9. PILGRIM Score – Schwellenwertfindung	42
3.10. Validierung des finalen PILGRIM Scores	43
3.11. Modell mit gerundeten Variablen gewichtungen	45
4. DISKUSSION	47
4.1. Diskussion der identifizierten Risikofaktoren	47
4.1.1. Finale Variablen des PILGRIM-Score	47
4.1.2. Weitere Risikofaktoren	49
4.2. Diskussion der logistischen Regressionsmodelle und des PILGRIM-Score	52
4.2.1. Variablensortierung zum Variableneinschluss	52
4.2.2. Individuelle Variablen gewichtung	53
4.2.3. Multikollinearität	53
4.2.4. Methodik zur Score-Entwicklung	54
4.2.5. Technik der Implementation	55
4.2.6. Validierung der Ergebnisse	55
4.3. Limitationen	56
4.4. Zusammenfassung und Ausblick	58

5. LITERATURVERZEICHNIS	60
6. ANHANG	66
6.1. Literaturrecherche: Risikofaktoren für Antibiotika-Verschreibung	66
6.2. Primärkodierung: Überblick über neu kodierte Variablen	67
6.3. Verlauf der Kennzahlen der binär logistischen Regressionsmodelle	68
6.4. Beschreibung zur Entwicklung der optimalen Variablenkombination	69
6.5. PILGRIM-Score Modelleleistungen nach Matsushita und Schneeweiss	70
6.6. PILGRIM-Score Modelleleistungen (Validierung) nach Matsushita und Schneeweiss	71
6.7. Auszug Studienprotokoll PILGRIM-Studie	72
6.8. Abbildungsverzeichnis	73
6.9. Tabellenverzeichnis	73
7. VORABVERÖFFENTLICHUNG VON ERGEBNISSEN	74

Abkürzungsverzeichnis

Abb.	Abbildung
AIDS	Akquiriertes Immun-Defizienz-Syndrom
AMR	Antimicrobial resistance
AMS	Antimicrobial stewardship
AUC	Area under the curve
BMI	Body mass index
BV	Baseline visit der PILGRIM-Studie
CCI	Charlson-comorbidity-index
CKD	Chronic kidney disease
COPD	Chronic obstructive pulmonary disease
eCRF	Electronic case report form
ESBL	Extended-spectrum- β -lactamase
ICD	International Classification for Diseases
IQR	Interquartil range (Interquartilsabstand)
KI	Konfidenzintervall
MRE	Multiresistente Erreger
MRSA	Methicillin-resistenter <i>Staphylococcus aureus</i>
NPV	Negativer prädiktiver Wert
OC	Observational contact der PILGRIM-Studie
OR	Odds ratio
PPI	Protonen Pumpen Inhibitoren
PPV	Positiver prädiktiver Wert
RDD	Recommended daily doses
ROC	Receiver operating characteristic
Spp.	Species pluralis
SV	Sampling visit der PILGRIM-Studie
WHO	World Health Organization

1. Zusammenfassung

Bakterielle Antibiotikaresistenzen (AMR) werden von der Weltgesundheitsorganisation (WHO) zu den zehn größten globalen Gesundheitsbedrohungen gezählt ¹. In einer aktuellen Studie wurde geschätzt, dass 2019 weltweit etwa 5 Millionen Todesfälle mit AMR in Zusammenhang standen und rund 1,3 Millionen Todesfälle direkt auf AMR zurückzuführen sind. Ursächlich für die zunehmende Resistenzentwicklung sind vor allem der inadäquate Einsatz sowie unverhältnismäßig hohe Verbrauch von Antibiotika in der Veterinär- und Humanmedizin ¹⁻³. Gleichzeitig sind die vorhandenen, neu zugelassenen oder sich in der klinischen Entwicklungsphase befindlichen antimikrobiellen Agenzien unzureichend bei der Bekämpfung der wachsenden Bedrohung durch AMR ⁴. Dies unterstreicht einen Bedarf an neuen antiinfektiven Behandlungsstrategien ⁵. Zur Verringerung der Resistenzentwicklung wurden nationale sowie internationale Strategien zur Koordinierung sektorenübergreifender Maßnahmen entwickelt ^{3,6}. Hierzu zählen auch sogenannte *Antimicrobial Stewardship* (AMS) Initiativen, welche durch verschiedene Maßnahmen die Qualität der Antibiotikaverordnungen zu verbessern versuchen. Die Wirksamkeit von AMS ist durch zahlreiche Studien belegt ^{7,8}. Um AMS-Studien sowie -Maßnahmen durchführen zu können, ist die Identifizierung von Risikokollektiven im Krankenhaus eine wichtige Voraussetzung. PatientInnen-individuelle Faktoren für die Prädiktion einer späteren Antibiotikatherapie nach Krankenhausaufnahme sind bisher jedoch wenig untersucht. Die genauere Analyse möglicher Risikofaktoren und Entwicklung eines Prädiktionsmodells soll Ziel dieser Arbeit sein.

Die vorliegende statistische Analyse wurde im Rahmen der internationalen, multizentrischen und prospektiven Kohortenstudie *Impact of Prescription Quality, Infection Control and Antimicrobial Stewardship on Gut Microbiota Domination by Healthcare-Associated Pathogens* (PILGRIM) durchgeführt (NCT03765528, ClinicalTrials.gov). Ziel der übergreifenden PILGRIM-Studie ist die Untersuchung des Einflusses (nicht-adäquater) Antibiotikatherapien auf das enterische Mikrobiom und der Begünstigung von Resistenzentwicklung. Insgesamt wurden von 2019 bis 2022 an 10 Studienzentren in 6 verschiedenen Ländern gezielt PatientInnen rekrutiert, die nach Einschätzung der StudienärztInnen bei stationärer Krankenhausaufnahme ein hohes Risiko für eine antibiotische Therapie im Verlauf besaßen. Zur Identifikation von Risikofaktoren für eine Antibiotikatherapie in der PILGRIM-Kohorte wurden die erfassten Variablen hinsichtlich ihrer Korrelation mit zukünftigen Antibiotikaverschreibungen analysiert. In zahlreichen Iterationen multivariabler Regressionsmodelle wurde im Folgenden eine Auswahl an besonders aussagekräftigen Variablen getroffen. Auf Basis dieser Auswahl wurden variablenspezifische Gewichtungen berechnet und anschließend der PILGRIM-Score zur Prädiktion zukünftiger Antibiotikatherapien entwickelt.

Die basisdemographischen Daten des Datensatzes zeigten sich insgesamt ausreichend balanciert und normalverteilt. Rund 35% aller PatientInnen wurden in Hämato-onkologischen Abteilungen rekrutiert. Obwohl die Rekrutierung mit dem Ziel erfolgte, möglichst viele Antibiotikatherapien im Verlauf beobachten zu können, erhielten nur knapp die Hälfte aller rekrutierten PatientInnen auch tatsächlich eine antibiotische Therapie. Als signifikante Risikofaktoren konnten herausgestellt werden: Männliches Geschlecht, Alter unter 70 Jahren, bestimmte behandelnde Abteilungen, alle Aufnahmen, die nicht für elektive Operationen bestimmt sind, vergangene Hospitalisierungen oder Antibiotikaeinnahmen, Einnahme von immunsupprimierenden Medikamenten oder jenen, die den gastralen pH anheben sowie hämatologische Malignome. Das beste logistische Regressionsmodell aus den identifizierten Risikofaktoren erreichte in der Validierung eine Sensitivität 50%, eine Spezifität von 88%, einen positiven prädiktiven Wert (PPV) von 75% und einen negativen prädiktiven Wert (NPV) von 70% bei einer Fläche unter der Kurve (*Area under the curve*, AUC) von 75%. Der darauf aufbauende finale Pilgrim-Score besteht aus vier Variablen mit jeweils individueller Gewichtung und einem Maximalwert von zehn Punkten: Hämatologische Malignome (fünf Punkte), immunsupprimierende Medikation (zwei Punkte), geplante elektive Operation (zwei Punkte) und vergangene Hospitalisierung (ein Punkt). Das Modell erreichte hiermit eine Prädiktionsleistung von 92% Spezifität, 41% Sensitivität, 78% im PPV und 69% im NPV bei einem Schwellenwert von 60% (sechs Punkte) in der Validierung.

Arbeiten, die ein ähnliches Modell zur Prädiktion von Antibiotikatherapien beschreiben, konnten nicht gefunden werden. Insgesamt analysieren nur sehr wenige Publikationen explizit Risikofaktoren für Antibiotikatherapien. Die meisten identifizierten Risikofaktoren und finalen Variablen des Pilgrim Scores konnten in der Literatur jedoch zumindest als Risikofaktoren für spätere Infektionen oder beispielsweise ein schlechteres Outcome identifiziert werden und somit wissenschaftlich erklärt werden. Insgesamt limitierend sind der hohe Anteil hämato-onkologischer PatientInnen in der Kohorte, eine damit einhergehende Dokumentationsbias, sowie Multikollinearität zwischen den finalen Variablen. Der entwickelte Score demonstriert dennoch eine solide Prädiktionsleistung und kann zukünftig helfen, PatientInnen unter einem hohen Risiko für Antibiotikatherapien zu Präventions- oder Studienzwecken zu identifizieren. In einem nächsten Schritt könnte der PILGRIM-Score noch mit weiteren Parametern wie Labor- oder Vitalwerten ergänzt oder auf einem externen Datensatz validiert werden.

2. Einleitung

2.1. Folgen des Antibiotikaeinsatzes

Der therapeutische Einsatz von Antibiotika wird als Meilenstein in der Geschichte der Humanmedizin bezeichnet, da er die Behandlung bakterieller Infektionskrankheiten revolutioniert hat ⁹. Obwohl der Antibiotikaverbrauch auf europäischer Ebene zuletzt leicht gesunken ist, bleibt das Verbrauchsniveau insgesamt jedoch hoch ¹⁰. Dies lässt sich unter anderem mit kürzeren Verweildauern, komplexeren Behandlungsmethoden inklusive anspruchsvoller Transplantationsverfahren und damit einhergehenden immunsuppressiven Strategien erklären ⁷. Gleichzeitig ergeben sich jedoch sowohl auf Ebene des einzelnen Individuums als auch auf Bevölkerungs- und Umweltebene Probleme durch den weit verbreiteten Antibiotikaeinsatz. Der *One Health*-Ansatz beschreibt die Annahme, dass die Gesundheit der Menschheit eng mit der Gesundheit von Tier- und Pflanzenwelten einhergeht und die drei Welten stets miteinander interferieren ¹¹. Es ist wissenschaftlicher Konsens, dass der inadäquate und weit verbreitete Einsatz von Antibiotika in der Human- und Veterinärmedizin, Landwirtschaft und Industrie zu einem starken Anstieg in der Ausbreitung endemischer antimikrobieller Resistenzen in den vergangenen 80 Jahren geführt haben ¹². Unterdessen fördern die wachsende Weltbevölkerung, steigende Tier-Populationen, vermehrter internationaler Handel, Reisen und Migrationsbewegungen ihre Verbreitung ¹³. Multi- sowie univariate Analysen konnten einen statistisch signifikanten Zusammenhang zwischen dem Verbrauch von bestimmten Antibiotikaklassen (Carbapenemen, 3.- und 4. Generation Cephalosporinen, Fluoroquinolone) und den zugehörigen Resistenzraten sowohl in der Human- als auch Veterinärmedizin nachweisen. Gleichzeitig ergeben sich Hinweise, dass Resistenzen von Bakterien aus der Humanmedizin mit Resistenzen aus der Veterinärmedizin assoziiert sind, der Einsatz von Antibiotika in der Humanmedizin die Verbreitung von Resistenzen in der Veterinärmedizin fördert und umgekehrt ¹⁴. Die Bedrohung durch AMR kann also nur im Kontext von *One Health* gesehen werden ^{11,13}. Probleme auf individueller Ebene ergeben sich durch die Veränderung des (enterischen) Mikrobioms mit dem Einsatz von Antibiotika sowie durch wirkstoffspezifische Nebenwirkungen.

2.1.1. PatientInnen-individuelle Ebene

Aktuelle Forschungsergebnisse legen nahe, dass PatientInnen unabhängig von vorangegangenen Hospitalisationen oft bereits von Antibiotika-resistenten Erregern besiedelt sind. Dies geschieht entweder durch Kontakt mit dem Gesundheitssystem oder mit Aufnahme durch die Nahrungskette. Die Gabe von antimikrobiell wirksamen Substanzen führt zu einer nachhaltigen Reduktion der Diversität der Darmmikrobiota und zur Entstehung einer ökologischen Nische. Der durch das Antibiotikum ausgeübte Selektionsdruck führt dann

konsekutiv zur klonalen Expansion einer multiresistenten Spezies und zur Ausfüllung der neu entstandenen ökologischen Nische^{9,15}. Hält der Selektionsdruck an, so kann die klonal expandierte Spezies die Mikrobiota dominieren. Sobald eine Barrierestörung der Darmschleimhaut hinzukommt, kann diese Ereigniskette eine Infektion des Gastrointestinaltraktes oder eine Invasion des Erregers in die Blutstrombahn nach sich ziehen^{9,16,17}. Der Prozess wird in Abb. 1.1. verdeutlicht.

Zu den Veränderungen des Mikrobioms kommen weitere Nebenwirkungen, wie Nephro-, Hepato-, Neurotoxizität, kutane Nebenwirkungen, usw., die bei Einnahme von Antibiotika ausgelöst werden können.

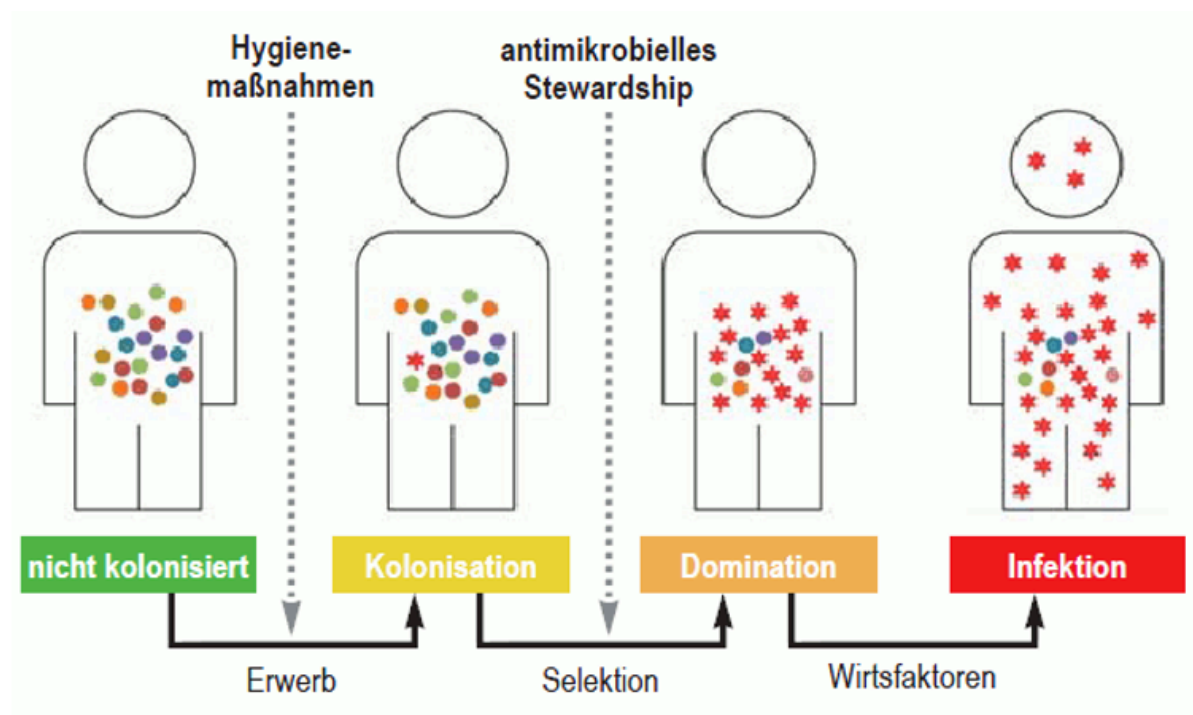


Abbildung 1.1.: Phasen der Infektionspathogenese (Quelle: Khodamoradi Y, Kessel J, Vehreschild JJ, Vehreschild M. The Role of Microbiota in Preventing Multidrug-Resistant Bacterial Infections. Dtsch Arztebl Int 2019; 116(40): 670-6)

2.1.2. Herausforderungen durch AMR auf globaler Ebene

Infektionen mit resistenten Erregern sind aufgrund beschränkter Therapieoptionen insgesamt schwieriger zu behandeln und führen zu einer erhöhten Morbidität und Mortalität^{18,19}. 2019 standen weltweit etwa 5 Millionen Todesfälle mit AMR in Zusammenhang und rund 1,3 Millionen Todesfälle waren in der statistischen Analyse von Murray et al. direkt auf AMR zurückzuführen. Es ergaben sich große regionale Unterschiede mit einer Todesrate von rund 27 pro 100.000 im westlichen Subsahara-Afrika und nur rund 7 pro 100.000 im australischen Raum. Innerhalb der etwa 5 Millionen mit AMR in Zusammenhang stehenden Todesfälle bilden Infektionen der unteren Atemwege mit 1,5 Millionen den größten Anteil. Die drei am häufigsten anzutreffenden Erreger mit AMR sind *Escherichia coli*, *Staphylococcus aureus* sowie *Klebsiella pneumoniae* und die 6 häufigsten AMR-Erreger waren für rund 73% der 1,3

Millionen Todesfälle durch AMR verantwortlich ⁵. Modellrechnungen legen für 2050 Verluste des globalen Bruttoinlandsproduktes zwischen 190 Milliarden und rund 4 Billionen US\$ allein durch AMR nahe ²⁰.

2.1.3. Resistenzmechanismen

Unterschieden werden intrinsische, erworbene und adaptive Resistenzen, wobei die adaptive Resistenz das dritte und neuste Erklärungskonzept darstellt ^{2,21}. Bei vorhandenen und nicht erworbenen, strukturellen sowie funktionellen Charakteristika eines bestimmten Erregers, die eine Widerstandsfähigkeit gegenüber einem bestimmten Antibiotikum ermöglichen, spricht man von intrinsischen Resistenzen. Das Fehlen eines Angriffspunktes für ein Antibiotikum bei einem Erreger ist das einfachste Beispiel einer intrinsischen Resistenz. Gleichzeitig können Resistenzdeterminanten über beispielsweise Plasmide zwischen verschiedenen Spezies transferiert werden, sodass Resistenzmechanismen im Gegensatz zu den bereits vorhandenen Mechanismen auch erworben werden können (siehe unten). Wenn es um die Weitergabe von intrinsischen Resistenzen geht, meint dies nicht die Weitergabe spezifischer Resistenzmechanismen, zumal manche Genloci nicht austauschbar sind. Vielmehr ist hier die Dominanz eines resistenten Stammes unter einem bestimmten Selektionsdruck gemeint, welche im Fortgang dann zur Selektion resistenter Stämme führt ²¹. Instabile und vorübergehende Resistenzen, die durch spezifische Signale oder Umweltbedingungen hervorgerufen werden und meist nach Verlust des Umweltsignals wieder rückgängig gemacht werden, werden als adaptive Resistenzen bezeichnet. Diese entstehen weder durch Mutationen, noch werden sie vertikal vererbt. Stattdessen werden adaptive Resistenzen eher durch epigenetische Phänomene ermöglicht ^{21,22}. Die Formierung eines Biofilms durch *Pseudomonas aeruginosa* und die damit einhergehende verminderte Suszeptibilität ist ein Beispiel für einen adaptiven Resistenzmechanismus ²³.

2.1.4. Resistenz-Surveillance

Der WHO-Bericht *Global Antimicrobial Resistance and Use Surveillance System (GLASS) Report 2022* gibt einen Überblick über die Entwicklung der globalen Resistenzlage. Als größte Bedrohung werden hier steigende Resistenzraten von Blutstrominfektionen-verursachenden Erregern angesehen ³. Für *Escherichia coli* ergaben sich so beispielsweise 2020 15% höhere Resistenzraten gegenüber 2017 für Meropenem und 3.-Generation Cephalosporine. Ähnliches kann für Blutstrominfektionen durch *Salmonella spp.* und die Resistenzraten für Ciprofloxacin beobachtet werden. Obwohl 3.-Generation Cephalosporine als Therapeutikum der Wahl bei Blutstrominfektionen gelten, sind gerade hier steigende Resistenzraten zu beobachten, was wiederum zu einem erhöhten Verbrauch von Reserveantibiotika wie Carbapenemen führt. Als weitere Bedrohung werden sich ausbreitende Carbapenemase-

produzierende *Enterobacterales* sowie steigende Resistenzraten von *Acinetobacter spp.* gegenüber Carbapenemen und Aminoglykosiden angesehen. Carbapenem-resistente Erreger sind meist multiresistent und mit Therapieversagen assoziiert. Für den Methicillin-resistenten *Staphylococcus aureus* oder blutstrominfizierende *Escherichia coli* werden gegenüber 3.-Generation-Cephalosporinen aktuell Resistenzraten zwischen 35-42% angegeben ²⁴.

2.2. Antimicrobial Stewardship (AMS)

Antimicrobial Stewardship (AMS) Programme sollen die rationale Anwendung von Antibiotika sicherstellen. Die AMS-Leitlinie der Deutschen Gesellschaft für Infektiologie fasst die Ziele von AMS Programmen gut zusammen: „[...] die kontinuierliche Verbesserung der Verordnungsqualität von Antiinfektiva [...], um beste klinische Behandlungsergebnisse des Patienten unter Beachtung einer Minimierung von Toxizität zu erreichen“ ⁷. Die Optimierung der Verordnungsqualität soll den vermehrten und unkritischen Einsatz von Antibiotika reduzieren und dadurch die Resistenzentwicklung eindämmen und Kosten senken. AMS-Programme bestehen aus einem Katalog mit verschiedenen Maßnahmen zur Optimierung von Therapien sowie des Verordungsverhaltens. Zu den Maßnahmen zur Verbesserung des Verordungsverhaltens zählen die Erstellung von Antiinfektivalisten, die Implementierung von Freigaberegungen, die Einrichtung von lokalen Behandlungsleitlinien sowie die generelle Beachtung einschlägiger Leitlinien und die Einführung von interdisziplinären AMS-Visiten. Maßnahmen zur Verbesserung der Therapien umfassen Beratungen zur Therapieevaluation, Therapiedauer, Oralisierung, Dosisanpassung, mikrobiologische Diagnostik, Testung auf Penicillinallergie, Befundmitteilung und Substanzwechsel bei multiresistenten Erregern (MRE). Voraussetzung für die Implementierung von AMS-Programmen neben Fachpersonal mit Mandat und Informationstechnologie sind Surveillance Daten zum Antiinfektiva-Verbrauch sowie Resistenzraten. Hierzu zählen auch die Teilnahme an nationalen Surveillance-Systemen zu Benchmarking-Zwecken ⁷.

In der letzten Cochrane-Metaanalyse zur Effektivität von AMS-Interventionsstrategien konnte gezeigt werden, dass durch AMS-Interventionen die Leitlinien-Adhärenz bei Antibiotikaverschreibungen um 15% (95%-KI: 14-16%) von 43% auf 58% im Mittel gesteigert werden konnte. Durch AMS-Interventionen kann die Zeit, in der PatientInnen mit Antibiotika behandelt werden, signifikant verkürzt werden (1,95 Tage bei 95%-KI: 1,67-2,22). Eine Verkürzung der Therapiedauer mit Antibiotika führt *nicht* zu einer Steigerung der Mortalität (0% bei 95%-KI: -1-0%) und liegt sowohl mit als auch ohne AMS-Intervention bei rund 11%. Hinzu kommt eine Verkürzung des gesamt-Krankenhausaufenthaltes um ca. 1 Tag (95%-KI: 0,7-1,54) ⁸.

2.3. Prädisponierende Faktoren für eine Antibiotikatherapie

Eine vor Durchführung dieser Arbeit durchgeführte Literaturrecherche (Pubmed; Suchterm: "risk factor*" OR "predictor*" AND "antibiotic* treatment" OR "antibiotic* prescription" OR "antibiotic* administration"; 15.06.2022) konnte 10 Publikationen identifizieren, die sich mit möglichen Risikofaktoren für eine folgende Verschreibung von Antibiotika beschäftigen (stationär oder ambulant). Insgesamt fanden sich drei retrospektive Kohortenstudien, vier Querschnittsstudien, eine Meta-Analyse, eine Punkt-Prävalenz-Studie sowie eine Beobachtungsstudie²⁵⁻³⁴. Eine detailliertere Übersicht findet sich in Tabelle 6.1. im Anhang. Bereits bestehende und publizierte Prädiktionsmodelle zu Risikofaktoren einer Antibiotikaverschreibung konnten hingegen nicht identifiziert werden.

Zusammenfassend werden sowohl patientInnenabhängige als auch patientInnenunabhängige Risikofaktoren beschrieben. Auf Seite der patientInnenabhängigen Risikofaktoren werden einige Aufnahmediagnosen, wie zum Beispiel die chronisch obstruktive Lungenerkrankung (COPD) oder Infektionen bei Aufnahme als Risikofaktoren beschrieben^{32,34}. PatientInnen mit mehr als einer Komorbidität, Übergewicht, aktivem Tabakkonsum, weiblichem Geschlecht oder niedrigerem Alter scheinen ebenfalls unter einem höheren Risiko zu stehen^{25,28,30,32}. Einkommensverhältnisse und Versicherungsstatus haben ebenfalls Einfluss auf das Risiko einer Antibiotikaverschreibung^{25,28}.

Auf struktureller Ebene werden der geographische Einflüsse (bspw. US-Südstaaten vs. US-Westküste), die Fachrichtung des konsultierten Arztes und beispielsweise die Aufnahmestation als entscheidend bei der Antibiotikaverschreibung verstanden^{25,28,34}. In einer pädiatrischen Studie konnten zudem die Länge des stationären Aufenthaltes und die Größe des Krankenhauses, gemessen an der Bettenzahl, als Risikofaktoren identifiziert werden²⁶. Nicht zuletzt haben die Erwartungshaltung der zu Behandelnden sowie das Wissen und die Qualifikation der Behandelnden auf der anderen Seite ebenfalls einen wesentlichen Einfluss darauf, ob ein Antibiotikum verschrieben wird oder nicht²⁷.

Untersucht man die stationären Verschreibungsraten je nach Krankenhaus oder innerhalb eines Krankenhauses, so zeigen sich auch hier interessante Risikofaktoren: So steigt die Menge verbrauchter Antibiotika mit der Bettenzahl des Krankenhauses. In Universitätsklinika wurden im Median >56 *Recommended Daily Doses* (RDD)/100 Pflgetage verabreicht, in nicht-universitären Krankenhäusern mit >800 Betten etwa 44 RDD und in kleineren Häusern mit <400 Betten nur rund 38 RDD. Besonders hohe Verschreibungsdichten haben hämato-onkologische oder intensivmedizinische Abteilungen (102 bzw. 92 RDD) im Vergleich zu Normalstationen (operativ und nicht-operativ mit 43 bzw. 44 RDD)³⁵. Insgesamt machen

Antibiotikaverschreibungen in der Hämato-Onkologie jedoch nur 12% am gesamt-Antibiotikaverbrauch aus.

2.4. Ziel dieser Arbeit

Der weltweit weit verbreitete Antibiotikaeinsatz zieht einen Anstieg von AMR nach sich, was, wie von der WHO betont, eine der gravierendsten globalen Gesundheitsbedrohungen darstellt. Es ergeben sich Herausforderungen auf patientInnen-individueller- (beispielsweise die Disruption der Darmmikrobiota und die konsekutive Förderung der Ausbreitung von AMR) sowie übergreifender, globaler Ebene. AMR führt dazu, dass verfügbare antimikrobielle Substanzen nicht mehr wirksam sind und Mortalität sowie Morbidität steigen. Die Tatsache, dass global ca. 1,3 Millionen Todesfälle direkt mit AMR in Verbindung stehen, verdeutlicht, dass neue, antiinfektive Behandlungsstrategien dringend erforderlich sind ^{1,4,5}. Die Maßnahmen des AMS haben zum Ziel, dem entgegenzuwirken und große Metanalysen konnten zeigen, dass die Strategien sehr effektiv sein können ^{7,8}. Um Studien durchzuführen, die AMS-Maßnahmen evaluieren und bewerten, muss das Studienkollektiv gut ausgewählt und identifiziert werden. PatientInnen, die unter einem hohen Risiko für zukünftige Antibiotikatherapien stehen, sind hierbei von besonderem Interesse. Bislang sind Risikofaktoren für antibiotische Behandlungen nach erfolgter stationärer Aufnahme jedoch nur wenig untersucht. Aus Antibiotikasurveillance-Daten lassen sich nur allgemeine Risikokonstellationen ableiten ³⁵. Die existierenden Untersuchungen zu Risikofaktoren auf patientInnen-individueller Ebene sind bislang jedoch begrenzt und geben insgesamt wenig Erkenntnisgewinn. Nur sehr wenige Studien beschäftigen sich explizit mit den Risikofaktoren für zukünftige Antibiotikatherapien. Die Identifikation passender Studienkollektive für AMS-Untersuchungen bleibt also eine Herausforderung. Die vorliegende Dissertation hat zum Ziel, diese Lücke zu schließen und ein besseres Verständnis für zugrundeliegende Faktoren zu liefern. Im zur Verfügung stehenden PatientInnenkollektiv der PILGRIM-Studie soll gezielt nach patientInnenabhängigen Risikofaktoren gesucht werden. Mit Hilfe der identifizierten Risikofaktoren soll darauffolgend ein Score-Modell entwickelt werden, welches PatientInnen mit einem hohen Risiko für Antibiotikatherapien identifizieren kann. Durch diese Präzisierung des in Frage kommenden Kollektivs, könnte die Planung und Rekrutierung zukünftiger AMS-Studien erleichtert werden. Langfristig könnte der entwickelte Score also einen Beitrag zur AMS-Forschung leisten und somit zur Senkung von AMR führen. Gleichzeitig könnten sich für betroffene PatientInnen individuelle Vorteile ergeben, da nach Risikoidentifikation gezielt präventive Maßnahmen, wie z.B. die Asservierung des Darm-Mikrobioms vor Antibiotikatherapie, ergriffen werden könnten.

3. Material und Methoden

Die klinischen Daten aller in die PILGRIM-Studie eingeschlossenen StudienpatientInnen wurden in pseudonymisierter Form in einem *electronic case report form* dokumentiert. Ein selektiver Export der Datenbank vom 25.09.2021 mit 1.524 dokumentierten Fällen bildet die Grundlage aller im Folgenden beschriebenen Analysen. Nach der unten beschriebenen Konstruktion des für diese Auswertung spezifisch entwickelten Endpunktes blieben 1.258 auswertbare Fälle übrig. Alle statistischen Analysen wurden mit dem *Statistical Package for Social Sciences* (SPSS®) der International Business Machine Corporation (IBM Inc.®) Version 28.0.0.0 durchgeführt.

Wesentliche Schritte der deskriptiven Statistiken vor der Entwicklung des eigentlichen Scores waren A) die Erstellung des oben genannten Exportes aus der Datenbank der PILGRIM-Studie, B) eine ausführliche Primärkodierung des Datensatzes inklusive der Konstruktion des Endpunktes und C) die Durchführung beschreibender Statistiken zur Charakterisierung der PatientInnenkohorte.

Die darauffolgende explorative Datenanalyse und Konstruktion des PILGRIM-Scores wird in Abb. 2.1. dargestellt: 1) Teilung des Datensatzes in einen Entwicklungs- und einen Validierungsdatensatz, 2) Analyse von Risikofaktoren in Korrelation mit dem Endpunkt, 3) Berechnung multipler Regressionsmodelle als Basis für den folgenden PILGRIM-Score 4) Entwicklung des PILGRIM-Scores als additiven Score und 5) Validierung des PILGRIM-Scores mit dem Validierungsdatensatz.

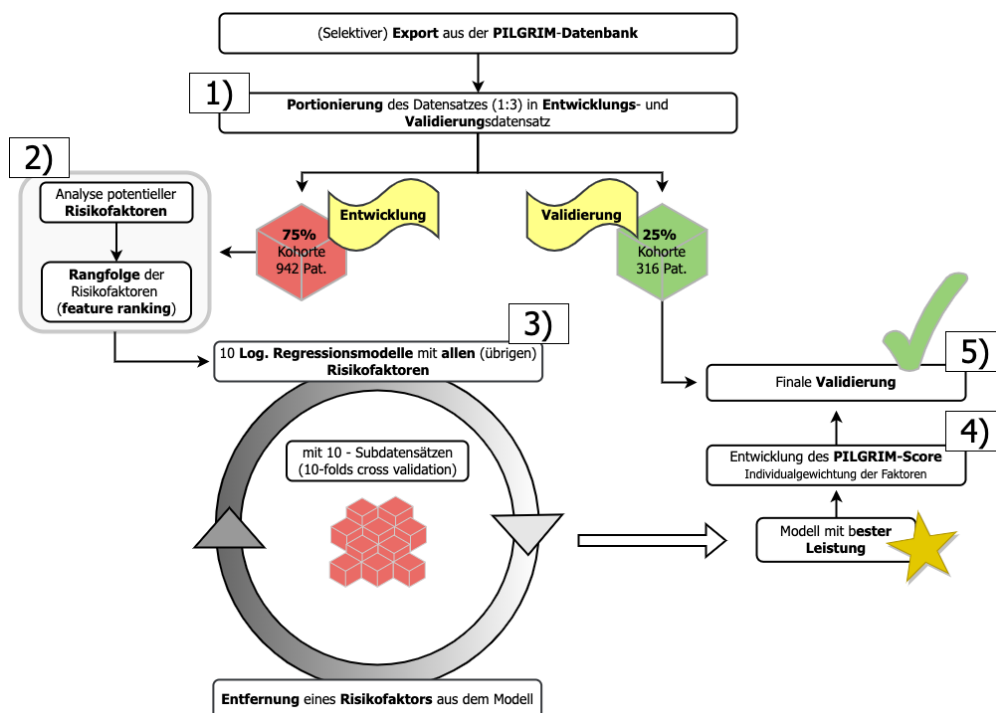


Abbildung 2.1.: Übersicht über den Ablauf der Entwicklung des PILGRIM-Scores

3.1. PILGRIM Studie

Basis für die Entwicklung des PILGRIM-Scores war die Studie mit dem Titel *Impact of Prescription Quality, Infection Control and Antimicrobial Stewardship on Gut Microbiota Domination by Healthcare-Associated Pathogens* (PILGRIM, NCT03765528, ClinicalTrials.gov). Die Datenerhebung der prospektiven und multizentrischen PILGRIM-Studie mit 10 Studienzentren (Kliniken) in Schweden, Lettland, Norwegen, Israel, Kanada und Deutschland startete 2019 und endete im Jahr 2022. Ein positives Votum der Ethikkommission der Universität zu Köln (Februar 2019, EK-Nr. 18-316) liegt vor. Insgesamt wurden mehr als 1600 PatientInnen rekrutiert, welche ihr schriftliches Einverständnis zur Studienteilnahme gegeben haben. Der Fragebogen und die dazugehörigen erhobenen Daten zu jedem Fall wurden online in eine eigens zu diesem Zweck eingerichtete Datenbank der Firma Questback AS (Bogstadveien 54, 0366 Oslo, Norwegen) eingepflegt. Ziel der PILGRIM-Studie ist die Analyse des Einflusses der Antibiotika-Verschreibungsqualität auf die Besiedlung und Infektion von PatientInnen mit *Extended-spectrum beta-lactamase* produzierenden *Enterobacteriaceae* (EPE), Vancomycin-resistenten *Enterococci* (VRE) und *Clostridium difficile*. Die Arbeitshypothese war, dass inadäquate Antibiotikatherapien die intestinale Besiedlung und konsekutive Domination des Mikrobioms sowie nachfolgende Infektion von PatientInnen mit den oben genannten Pathogenen fördern. Einschlusskriterien der PILGRIM-Studie waren 1) Alter \geq 18 Jahre, 2) eine geplante oder ein erwartetes hohes Risiko für eine systemische Antibiotikatherapie im Verlauf eines stationären Aufenthaltes (innerhalb der folgenden 10 Tage für mindestens 5 Tage, ausgenommen sind Prophylaxen mit Trimethoprim/Sulfamethoxazol), 3) die Möglichkeit, dass PatientInnen vor einer ersten Antibiotika-Verabreichung eine Stuhlprobe oder alternativ einen Rektalabstrich abgeben können und 4) eine schriftliche Einverständniserklärung. Ein Ausschlusskriterium war dementsprechend eine Antibiotikatherapie vor der Abgabe einer ersten Stuhlprobe. Darüber hinaus bestanden acht weitere Ausschlusskriterien, welche einem Auszug des Studienprotokoll entnommen werden können (Kapitel 6.7.) Die Suche nach geeigneten PatientInnen fand an den teilnehmenden Studienstandorten nach individuell festgelegten Methoden durch die Studienleitenden statt. Alle PatientInnen gaben danach wie oben beschrieben zu Beginn ihrer Studienteilnahme und nach einem ersten *Baseline Visit* (BV) eine Stuhlprobe oder einen Rektalabstrich ab.

Es folgten wöchentliche *Observational Contacts* (OC), bei denen a) stattgefunden antibakterielle Behandlungen (d.h. Verabreichung von Antibiotika) und b) Fieber oder Zeichen für eine Infektion erfragt wurden. Falls eines dieser drei Kriterien zutreffend war, der oder die PatientIn also ein Antibiotikum bekommen, Fieber oder eine Infektion entwickelt hat, wurde ein *Sampling Visit* (SV) durchgeführt. Beim SV wurden umfassende Informationen über die

bestehende Infektion und den Status der laufenden antibakterielle Therapie dokumentiert, sowie ein weitere Stuhlprobe gewonnen. Ein SV wurde ebenfalls durchgeführt, wenn eine baldige Entlassung geplant war, sodass der oder die PatientIn beim nächsten wöchentlich durchgeführten OC nicht mehr stationär im Krankenhaus gewesen wäre. Das Ende der Studienteilnahme war erreicht, wenn 6 konsekutive OC stattgefunden haben, ohne dass Antibiotika verabreicht wurden oder sich eine Infektion (mit/ohne Fieber) entwickelt hat. In online durchgeführten AMS-Falldiskussionen und -konferenzen wurden durchgeführte antibiotische Therapien hinsichtlich ihrer Indikation geprüft und nicht-adäquate Behandlungen identifiziert. Eine anschließende Analyse der Stuhlproben und Rektalabstriche sollte Aufschluss über den Einfluss von Antibiotika-Verschreibungen auf das intestinale Mikrobiom geben.

3.2. Primärkodierung des Datensatzes

3.2.1. Datenbereinigung

Zunächst wurden alle zur Verfügung stehenden Variablen und deren Antwortmöglichkeiten beschriftet, fehlende Werte je nach Variablentyp mit [-66, -77, -99, 0] kodiert, Skalenniveaus angepasst und nicht benötigte Variablen aussortiert (bspw. Zeitstempel-Variablen oder kaum genutzte Freitext-Variablen)

Es folgte eine ausführliche Datenbereinigung, bei der nicht plausible Fälle oder Fälle, die als Testfälle markiert waren, aus der weitergehenden Analyse entfernt wurden. Fälle mit wenig plausiblen Antworten beispielsweise zum Geburtsjahr, zum Gewicht oder der Körpergröße wurden genauer analysiert und bei $\geq 50\%$ fehlerhaften oder fehlenden Angaben ebenfalls aus der Analyse ausgeschlossen. Variablen mit nur wenigen positiven Ausprägungen ($n \leq 10$) und ansonsten überwiegend negativen Ausprägungen oder vornehmlich fehlenden Werten wurden als solche identifiziert und ebenfalls nicht in die weitergehende Analyse eingeschlossen. Dies gilt beispielsweise für folgende Variablen: Vorangegangene invasive Infektion mit *C. difficile*, MRSA, *Acinetobacter spp.*, *Pseudomonas spp.* (v_479, v_476, v_283, v_284), bekannte Kolonisation mit MRSA, *C. difficile* (v_276, v_279), Demenzerkrankungen (v_5785), Hemiplegie (v_5789) und Akquiriertes Immun-Defizienz-Syndrom (AIDS, v_5793).

PatientInnen mit dem Verdacht auf eine bestehende Infektion bei Studieneinschluss wurden von allen folgenden Analysen ausgeschlossen (v_245, n=25), da ansonsten eine Verzerrung des Bildes bei der Risikofaktorenanalyse erwartet wurde.

3.2.2. Kodierung zusätzlicher Variablen

Nicht binäre Variablen, d.h. Variablen mit mehr als zwei Antwortmöglichkeiten, wurden als Dummy-Variablen mit den Werten 0 und 1 neu kodiert. Hierbei wurde der Wert 1 stets jener Ausprägung zugeordnet, von welcher ein höheres Risiko für das Eintreten des Endpunktes erwartet wurde. So wurde beispielsweise der Einnahme immunsupprimierender Medikation der Wert 1 zugeordnet, andernfalls der Wert 0. Die erwartete weniger risikoreiche Option entspricht somit dem Wert 0 in der dummy-Kodierung.

Weiterhin wurden aus den erfassten Variablen einige weitere Variablen von Interesse konstruiert. So wurden nach Sichtung der Verteilung der Alter innerhalb der PatientInnenkohorte (Normalverteilung), Altersgruppen- und Schwellenwert-Variablen eingeführt. Aus den Variablen für die Körpergröße in Metern und dem Gewicht zum Zeitpunkt des BV wurde der Body-Mass-Index (BMI) berechnet.

Im Rahmen BV wurden auch die jeweiligen Stationen erfasst, auf welchen sich die PatientInnen zu diesem Zeitpunkt aufhielten. Diese entsprechen im Wesentlichen den Aufnahmestationen der PatientInnen. Die Aufnahmestationen wurden folgendermaßen zu neuen Variablen zusammengefasst: Chirurgie (Herz-Thorax-, Allgemein-, Gefäß-, Neuro-, Unfall-, Viszeral-, Thoraxchirurgie, Urologie und Orthopädie), Innere Medizin (Allgemeininternistische, kardiologische, hämato-onkologische, nephrologische, infektiologische und gastroenterologische Stationen) und Sonstige (Hals-Nasen-Ohren-Heilkunde, Gynäkologie/Geburtshilfe und Weitere). Nach einer ersten Sichtung möglicher Korrelationen von behandelnden Abteilungen und dem Endpunkt wurden risikobehaftete Variablen aggregiert und zu Risikoabteilungen zusammengefasst. So wurden in einer ersten Version alle Abteilungen, außer chirurgische Disziplinen, zu einer Risikovariablen zusammengefasst.

Der Charlson-Comorbidity-Index (CCI), welcher auf eine Publikation von 1987 zurückgeht und sich in der Vergangenheit als valides Werkzeug zur Abschätzung der Komorbidität etabliert hat, wurde ebenfalls erfasst ³⁶. Zu den Kriterien des CCI gehören Myokardinfarkte, Herzinsuffizienz, periphere arterielle Verschlusskrankheiten (PAVK), zerebrovaskuläre Insulte/transiente ischämische Attacken, Demenz, COPD, Kollagenosen, gastrointestinale Ulzera, Hemiplegie, moderate bis schwere Niereninsuffizienz, Leukämie, Lymphome, AIDS, Lebererkrankungen (abgestuft nach Schwere der Einschränkung), solide maligne Tumoren (abgestuft nach lokalisierten oder metastasierten Neoplasien) und Diabetes Mellitus (abgestuft je nach Schwere der Endorganschäden). Unter Einbeziehung des Alters ergibt sich eine theoretische maximale Punktzahl von 37 Punkten. Um den CCI für jeden oder jede PatientIn

separat zu berechnen, wurde eine neue Variable unter Berücksichtigung des Alters und spezifischer Gewichtung der Kriterien des CCI berechnet³⁶.

Aufnahmediagnosen und weitere Komorbiditäten neben denen des CCI wurden ebenfalls in Form des *International Classification for Diseases* in der 10. Version (ICD-10) erfragt. Da Kategorie, Diagnose und der Subcode jeder Komorbidität oder Aufnahmediagnose jeweils als eigene Variable erfasst wurden, wurden im Rahmen der Primärkodierung neue, zusammengefasste Variablen gebildet. Häufiger codierte ($n \geq 30$) Diagnosen, welche nicht bereits im CCI erfasst worden sind, wurden identifiziert und in die spätere Risikofaktorenanalyse eingeschlossen.

Thematisch zusammengehörende Variablen wie z.B. Daten zur Reiseanamnese, zur Kolonisation oder zurückliegenden Infektionen mit MRE wurden aggregiert und in der Risikofaktorenanalyse probatorisch gesammelt betrachtet.

3.2.3. Endpunkt-Definition

Unabhängig des Endpunktes der übergreifenden multizentrischen PILGRIM-Studie (Studienendpunkt), wurde für die vorliegende statistische Analyse ein eigener Endpunkt definiert: *antibiotische Therapie im Beobachtungszeitraum*. Alle Variablen wurden hinsichtlich Ihrer Korrelation mit diesem analysiert und je nach dem als Risikofaktoren eingestuft. Auch die nachfolgend berechneten logistischen Regressionsmodelle und der letztendliche PILGRIM-Score beziehen sich stets auf diesen Endpunkt.

Der Endpunkt *abx_endpoint* zur Entwicklung des Regressionsmodells wurde mit Hilfe der Informationen aus den OC zunächst mit 4 Ausprägungen konstruiert. Der Endpunkt wurde als *negativ* ausgeprägt (2) kodiert, wenn in den ersten beiden OC die Frage nach möglicherweise verabreichten Antibiotika mit *nein* beantwortet wurde. Wenn innerhalb des Beobachtungszeitraumes (6 OC), ein Antibiotikum gegeben wurde, wurde der Endpunkt als *positiv* ausgeprägt, d.h. mit (1) kodiert. Wurden die ersten beiden OC mit *nicht-bekannt* und *nein* beantwortet, wurde der Endpunkt als *negativ-unbekannt* ausgeprägt (3) kodiert. Eine perioperative Antibiotikaphylaxe bei den ersten beiden OC wurde beim Endpunkt auch als solche kodiert (4). Wurden die Fragen nach der Verabreichung von Antibiotika in den ersten beiden OC nicht beantwortet, beispielsweise wegen eines frühen Ausschlusses aus der Studie, wurde der Endpunkt auch als *fehlend* (0) kodiert. Für den in der logistischen Regressionsanalyse verwendeten binären Endpunkt wurden die 4 Kategorien des ursprünglichen Endpunktes zusammengefasst: Alle *positiv* ausgeprägten Endpunkte (1) wurden auch im binär kodierten Endpunkt als *positiv* (1) ausgeprägt kodiert. Die Kodierungen

2,3 und 4 wurden beim binär kodierten Endpunkt zu den negativen Ausprägungen (0) gezählt. Perioperative Antibiotikatherapien (*single shot*) haben meist eine prophylaktische Indikation und wurden im Rahmen dieser Analyse also nicht zu den antibiotischen Therapien per se gezählt. Fälle mit fehlenden Angaben zur Antibiotikatherapie (Kodierung 0) wurden wie eingangs beschrieben von der folgenden Analyse ausgeschlossen (n=265).

3.3. Portionierung des Datensatzes

Vor der logistischen Regressionsanalyse wurde eine Teilung des kompletten Datensatzes in einen Entwicklungs- und einen Validierungsdatsatz vorgenommen. Hierzu wurde eine 3:1 Teilung des Datensatzes durchgeführt, sodass ein Entwicklungsdatsatz mit insgesamt circa 75% aller Fälle und ein Validierungsdatsatz mit circa 25% aller Fälle entstanden. Nach den Empfehlungen von Steyerberg³⁷ wurde die Teilung des Datensatzes *nicht*-zufällig vorgenommen, sondern nach Datum des Studieneinschlusses. Die 25% der PatientInnen mit dem jüngsten Studieneinschluss, wurden dem Validierungsdatsatz zugeordnet. Alle älteren Studieneinschlüsse wurden dem Entwicklungsdatsatz zugeordnet. Der Validierungsdatsatz mit 25% aller eingeschlossenen PatientInnen wurde bei allen im Folgenden beschriebenen explorativen Analysen nicht miteinbezogen. Alle Berechnungen fanden lediglich mit dem 75% Entwicklungsdatsatz statt. Der Validierungsdatsatz diente zum Ende der Analyse zur Validierung des Scores und wurde nach der Aufteilung bis zu diesem Punkt, abgesehen von deskriptiven Statistiken, weder nähergehend analysiert noch bearbeitet.

3.4. Deskriptive Statistik

Alle Daten, die im Rahmen des BV der PILGRIM-Studie für alle PatientInnen erhoben wurden, wurden vor der explorativen Datenanalyse deskriptiv analysiert (sowohl für den Entwicklungs- als auch für den Validierungsdatsatz). Diese umfassen Informationen zu basisdemographischen Charakteristika (Alter, Geschlecht, Größe, Gewicht), Aufnahme-Abteilung, nach ICD-10 kodierten Aufnahme Diagnosen, Antibiotika-Therapien in den letzten 6 Monaten, Hospitalisierungen in den letzten 6 Monaten, geplante elektive Operationen, eine Reiseanamnese, aktueller Medikation (Immunsuppressiva oder Antazida), relevante Komorbiditäten nach ICD-10 kodiert und durch Abfrage des *Charlson Comorbidity Index* (CCI), bekannte Infektionen oder Besiedlungen mit MRE/schwierig zu behandelnden Erregern in der Vergangenheit sowie zu bekannten Allergien gegen Antibiotika. Die Berechnung von Mittelwerten, Medianen, Interquartilsabständen sowie Standardabweichungen verschaffte einen Überblick über den Datensatz.

3.5. Explorative Datenanalyse: Entwicklung des PILGRIM-Score

3.5.1. Identifizierung von Risikofaktoren

Im nächsten Schritt wurden vorher selektierte, bereits vorhandene oder neu kodierte Variablen hinsichtlich ihres Zusammenhanges mit dem Endpunkt überprüft und nach der Stärke des jeweiligen Zusammenhanges in einer Rangfolge sortiert. Grundlage der Variablenselektion bildete die eingangs beschriebene Literaturrecherche nach Risikofaktoren sowie die Unterstützung durch anerkannte ExpertInnen im Bereich des AMS. In Kreuztabellen wurde mit Hilfe des Pearson Chi-Quadrat-Tests ein möglicher Zusammenhang zwischen Variable X und dem Endpunkt getestet (Kontingenzanalyse). Bei Stichprobengrößen mit $n \leq 20$ oder erwarteten Zellhäufigkeiten ≤ 5 wurde analog der exakte Test nach Fischer verwendet. Zur Quantifizierung der Stärke des Zusammenhanges zwischen Variable X und dem Endpunkt wurden die Zusammenhangsmaße Phi ϕ für 2x2-Tabellen und Cramérs V für größere Tabellen verwendet. Zusammenhänge mit einem p-Wert $< 0,05$ wurden unter einem Signifikanzniveau von $\alpha = 0,05$ als statistisch signifikant und bei ϕ oder Cramérs V zwischen 0,1 und 0,3 als schwach, zwischen 0,3 und 0,5 als mittel und $\geq 0,5$ als stark angesehen³⁸.

Neben der beschriebenen Kontingenzanalyse wurde jede Variable auch in einer binär-logistischen, univariaten Regressionsanalyse hinsichtlich eines möglichen Zusammenhanges mit dem Endpunkt überprüft. Als abhängige Variable diente der Endpunkt mit der Ausprägung 1 und als unabhängige Variablen dienten in jeweils einzelnen Regressionsanalysen die zu analysierenden potenziellen Risikofaktoren. Von Interesse waren hier bei jeder individuellen Berechnung die Regressionskoeffizienten mit dem dazugehörigen p-Wert, sowie die jeweilige Odds Ratio mit dem dazugehörigen 95% Konfidenzintervall.

Ein möglicher Zusammenhang zwischen Variable X und dem Endpunkt konnte also durch das Durchführen zahlreicher Kontingenzanalysen- und binär-logistischen-Regressionen beschrieben werden. In einer Zusammenschau wurden im Folgenden alle Variablen mit den berechneten Kennzahlen in eine Tabelle eingefügt und nach der Signifikanz in der Kontingenzanalyse sortiert (*feature ranking*). Variablen mit der höchsten Signifikanz wurden in der Zusammenschau an vorderster Stelle geschrieben. Die so entstandene Rangfolge aller Variablen, bildete im Folgenden die Grundlage für die Entwicklung des Regressionsmodells.

3.5.2. 10-fache Filterung des Entwicklungsdatensatzes (Implementation der 10-fachen Kreuzvalidierung)

Vor Berechnung der Regressionsmodelle wurde der Entwicklungsdatensatz zum Zweck der Kreuzvalidierung in 10 Sub-Datensätze portioniert: Es wurden 10 zufällige Filter programmiert,

welche jeweils 90% der Fälle im Entwicklungsdatensatz auswählten und 10% übrigließen. Das Ergebnis bestand aus 10 zufälligen Abkömmlingen des Entwicklungsdatensatzes, die jeweils verschiedene Aufteilungen der Fälle im Verhältnis 9:1 aufwiesen (vgl. untenstehende Abb. 2.2.). Hervorzuheben ist, dass der Entwicklungsdatensatz nicht etwa in 10 kleine Datensätze aufgeteilt wurde, sondern lediglich 10-mal neu und zufällig im Verhältnis 9:1 gefiltert wurde.

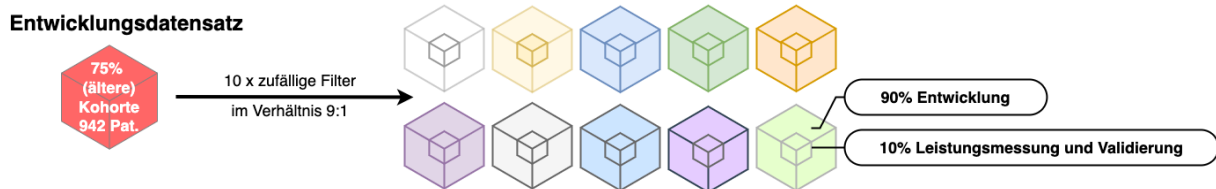


Abbildung 2.2.: 10-fache Kreuzvalidierung – 10-fache Filterung des Entwicklungsdatensatzes im Verhältnis 9:1

3.5.3. Entwicklung des optimalen binär logistischen Regressionsmodells

Im Folgenden wurden die logistischen Regressionsmodelle auf Grundlage der zuvor erstellten Rangfolge aller Variablen berechnet. So wurden in der ersten Iteration alle Variablen der vorher erstellten Rangfolge im logistischen Regressionsmodell als Kovariablen berücksichtigt. Als Methode des Variableneinschlusses wurde *Einschluss (Enter)* gewählt, bei welcher alle Kovariablen in einem Block in das Modell aufgenommen wurden. Die jeweils vermeintlich risikoärmere Option einer Variable wurde als Referenz bei der Berechnung gewählt. Das Ergebnis der ersten Iteration war dementsprechend ein binär logistisches Regressionsmodell mit allen zur Verfüg stehenden Variablen (n). In der zweiten Iteration wurde ein zweites logistisches Regressionsmodell berechnet, bei welchem jedoch jene Variable, mit der in der Kontingenzanalyse geringsten Signifikanz ausgeschlossen wurde (*feature ranking*), sodass das Regressionsmodell nun nur noch mit n-1 Kovariablen berechnet wurde. In einer dritten Iteration wurde ein weiteres Regressionsmodell berechnet, bei welchem wiederum eine Kovariable nach dem gleichen Kriterium ausgeschlossen wurde (s. untenstehende Abb. 2.3.). Dieses Verfahren wurde so oft wiederholt, bis nur noch eine Variable im Regressionsmodell übrigblieb. Die Berechnungen der einzelnen Iterationen wurden wie folgt vorgenommen: Jede Iteration wurde 10-mal mit den jeweils verschieden zusammengesetzten 90%-Teilen der Subdatensätze des Entwicklungsdatensatzes berechnet und dann hinsichtlich der Modelgüte an den zugehörigen 10%-Teilen der Subdatensätze des Entwicklungsdatensatzes bewertet (vgl. obiger Abschnitt mit Abb. 2.2. und Abb. 2.3.). Dieses Verfahren der 10-fachen Berechnung und 10-fachen Validierung wird als 10-fache Kreuzvalidierung (10-fold cross validation) bezeichnet und ist eine anerkannte Methodik bei der Entwicklung komplexer Modelle, vor allem im Bereich des automatisierten maschinellen Lernens. Diese Methodik ermöglicht, dass jeder Fall mit hoher Wahrscheinlichkeit einmal zur Validierung eines Modells

im 10%-Teil eines Subdatensatzes verwendet wurde³⁷. 10-fold cross validation wurde in der Vergangenheit erfolgreich von Jakob et al. und Werfel et al. bei der Entwicklung klinischer Prädiktionsmodelle eingesetzt^{39,40}.

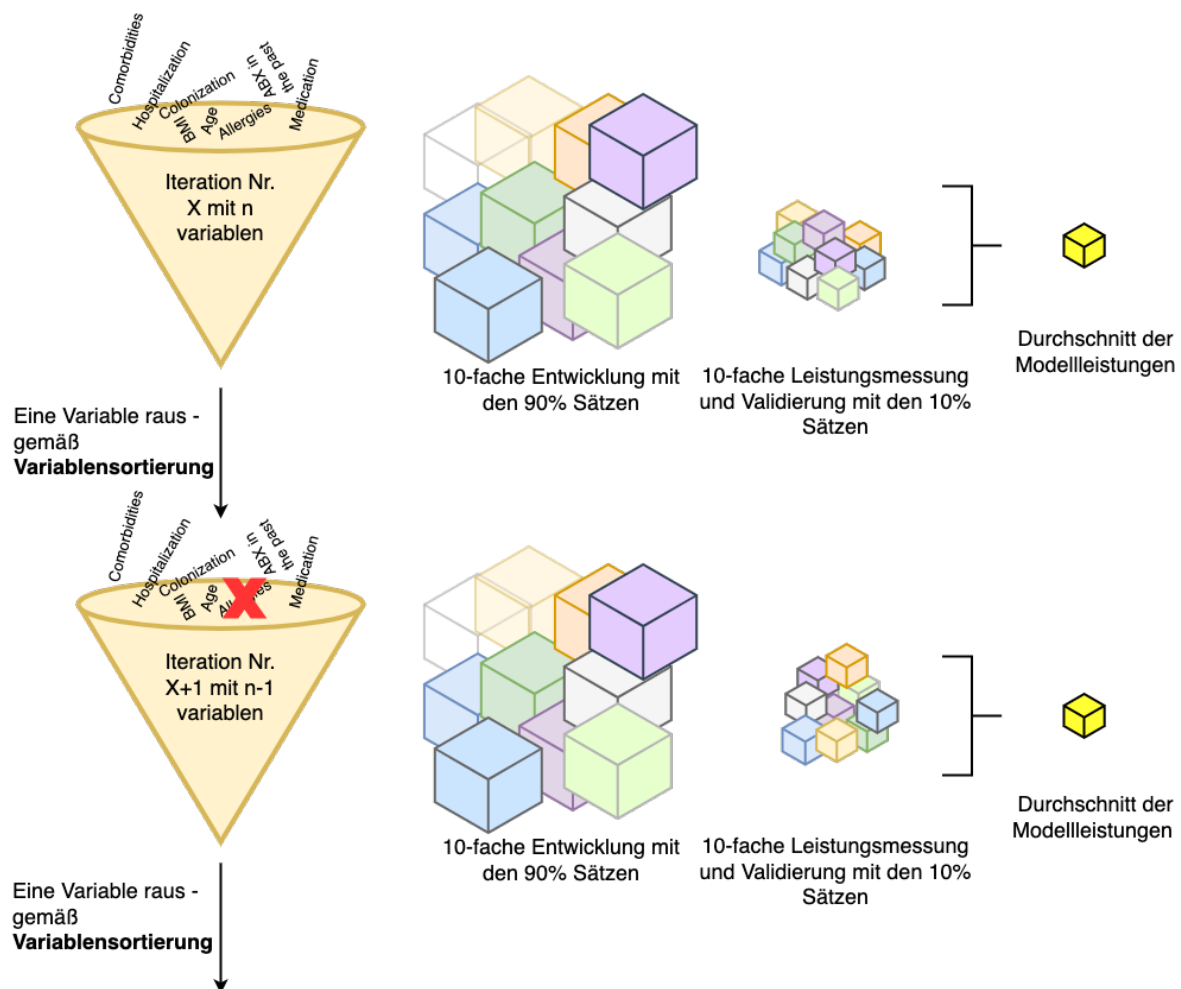


Abbildung 2.3.: Schaubild des Ablaufes mehrerer Iterationen mit 10-facher Kreuzvalidierung

Zur Beurteilung der Modellgüte der Iterationen, wurden verschiedene Kenngrößen mit der 10-fachen Kreuzvalidierung 10-fach berechnet und anschließend gemittelt. Der Mittelwert der Kenngrößen erlaubte schließlich eine Aussage über die Güte des Modells einer Iteration. Folgende Kenngrößen wurden verwendet: Mit der zu jedem einzelnen Modell angezeigten Klassifizierungstabelle (Vier-Felder-Tafel), welche von SPSS standardmäßig mit einem Klassifizierungsschwellenwert von 0,5 ausgegeben wird, wurden bei $p < 0,05$ die Kenngrößen Sensitivität, Spezifität, PPV und NPV berechnet. Zuzüglich wurden *Receiver operating characteristic* (ROC)-Kurven mit Hilfe der Sensitivität (Ordinate), aufgetragen gegen die Falsch-Positiven-Rate (Abszisse), berechnet. Als Kennzahlen der Modellgüte dienten hier die AUC, Standardfehler der Kurve sowie die asymptotische Signifikanz.

Zur Evaluation der optimalen Kombination aus Variablen zur Entwicklung des PILGRIM-Scores wurden die Verläufe der statistischen Kenngrößen der einzelnen Regressionsmodelle

beobachtet. Erwartet wurde zunächst eine Zunahme der Modelleleistungen, da im Verlauf der Iterationen immer mehr störende Variablen entfernt wurden. Mit fortlaufender Anzahl der Iterationen wurde ein Absacken der Performance erwartet, da nun mit nur noch wenig übriggebliebenen Variablen für die Modellgüte wichtige Variablen entfernt wurden. Jene Iteration, die vor dem Absacken der Modellgüte (*drop*) noch eine stabile Modelleleistung erzielt hatte, sollte die Basis für die nachfolgende Entwicklung des eigentlichen Scores bilden. Die in der zuletzt besten Iteration enthaltenen Variablen bildeten also die grundsätzliche Auswahl an möglichen Bestandteilen des PILGRIM-Scores. Das probatorische Entfernen weiterer Variablen abseits der durch die Variablen-Rangfolge vorgegebenen Reihenfolge sollte die zuletzt erreichte Modelleleistung zusätzlich erhöhen und die maximal mögliche Modelleleistung bei gleichzeitig sinnvoller Variablen-Zusammensetzung identifizieren. Bei nicht-Abfallen der Modelleleistung, sollte mit den zuletzt übriggebliebenen 8-10 Variablen die Entwicklung eines Scores versucht werden. Hierbei sollte zur Verbesserung der Leistungen eine Vorwärtsselektion versucht werden: Statt wie vorher kontinuierlich Variablen aus dem Modell zu entfernen, sollten nun wieder sinnige Variablen zum Modell hinzugefügt werden, welche bereits zu Beginn der Iterationen entfernt wurden. Für die Vorwärtsselektion sollten Variablen mit beispielsweise hohem Regressionskoeffizient oder hoher Odds-Ratio ausgewählt werden.

Abschließend wurden die Variablen des finalen logistischen Regressionsmodells hinsichtlich ihrer Multikollinearität nach Kendall-Tau-b analysiert.

3.5.4. Entwicklung des PILGRIM-Score aus dem logistischen Regressionsmodell

Die oben beschriebene Methodik lieferte insgesamt ein validiertes binär logistisches Regressionsmodell mit n Variablen. Die folgende Entwicklung des additiven Scores beinhaltete zunächst die Berechnung der individuellen Gewichtungen der einzelnen Score-Variablen. Zur Gegenüberstellung und Verdeutlichung des Vorteils einer individuellen Gewichtung im additiven System wurde der Score daher ebenfalls mit einer Gewichtung = 1 für alle Variablen berechnet. In einer Literaturrecherche konnten drei verschiedene Arten der Gewichtungsberechnung identifiziert werden. Werfel et al. und Matsushita et al. orientieren sich bei der Gewichtungsberechnung an den Regressionskoeffizienten der einzelnen Variablen in den logistischen Regressionsmodellen (multivariabel) ^{40,41}. Schneeweiss et al. orientierten sich bei Ihrer Publikation an den *Odds Ratios* (OR) ⁴². Die Berechnungen im Einzelnen wurden dann wie folgt durchgeführt (siehe Tabelle 2.1. unten):

Methode

Berechnung der variablenspezifischen Gewichtung

Werfel, Jakob et al.	$\frac{\text{variablenspezifischer Regressionskoeffizient}}{\text{niedrigster Regressionskoeffizient in der log. Regression}}$
Matsushita et al.	$\frac{\text{variablenspezifischer Regressionskoeffizient}}{\emptyset \text{ der 2 niedrigsten Regressionskoeffizienten in der log. Regression}}$
Schneeweiss et al.	$\ln(\text{variablenspezifische OR})$

Tabelle 2.1.: Übersicht über die Methoden der variablenspezifischen Gewichtung

Bei allen für die Berechnungen benötigten Regressionskoeffizienten und Odds Ratios wurden jeweils die Mittelwerte der bei der 10-fold cross validation errechneten 10 verschiedenen Regressionskoeffizienten beziehungsweise Odds Ratios verwendet. Bei der Methode nach Schneeweiss et al. ergab sich die individuelle Gewichtung der Variablen aus dem Logarithmus Naturalis ($\log_e = \ln$) und den folgenden festgelegten Intervallen ⁴²:

if $0 \leq \ln(\text{OR}) \leq 0.15$ then weight = 0
 if $0.15 < \ln(\text{OR}) \leq 0.45$ then weight = 1
 if $0.45 < \ln(\text{OR}) \leq 0.75$ then weight = 2
 if $0.75 < \ln(\text{OR}) \leq 1.05$ then weight = 3
 if $1.05 < \ln(\text{OR}) \leq 1.35$ then weight = 4
 if $1.35 < \ln(\text{OR}) \leq 1.65$ then weight = 5
 if $1.65 < \ln(\text{OR}) \leq 1.95$ then weight = 6

if $0 > \ln(\text{OR}) \geq -0.15$ then weight = 0
 if $-0.15 > \ln(\text{OR}) \geq -0.45$ then weight = -1
 if $-0.45 > \ln(\text{OR}) \geq -0.75$ then weight = -2
 if $-0.75 > \ln(\text{OR}) \geq -1.05$ then weight = -3
 if $-1.05 > \ln(\text{OR}) \geq -1.35$ then weight = -4

Abbildung 2.4.: Gewichtungsintervalle von Schneeweiss (Quelle: Schneeweiss S, Wang PS, Avorn J, Glynn RJ. Improved comorbidity adjustment for predicting mortality in Medicare populations. Health Serv Res 2003; 38(4): 1103-20)

Im Anschluss wurden die individuellen und additiven Score-Werte aller Fälle des Entwicklungsdatensatzes (75%-Datensatz) nach der Erstellung neuer Dummy-Variablen zur Umsetzung der Variablengewichtung berechnet. Dies erfolgte für alle drei Methoden der Gewichtungsberechnung, das heißt nach Werfel, Matsushita und Schneeweiss. Durch die unterschiedliche Variablengewichtung ergaben sich auch unterschiedliche Maximalwerte in der additiven Score Berechnung. Konsekutiv wurden für jede Methode kontinuierlich aufsteigende Schwellenwerte für *PILGRIM-Score = positiv*, das heißt *Endpunkt positiv = PatientIn wird ein Antibiotikum erhalten* simuliert, was verdeutlichte wie sich die Kenngrößen Sensitivität, Spezifität, positiver und negativer prädiktiver Wert je nach Schwellenwert verändern.

3.5.5. Validierung des Pilgrim Score

Zur Validierung der Berechnungen wurde das mit dem 75%-Entwicklungsdatensatz errechnete binär logistische Regressionsmodell mit n verbliebenen Variablen auf dem 25%-Validierungsdatensatz getestet. Zur Validierung wurden auch hier, dieses Mal ohne 10-fold cross validation, die Kenngrößen Sensitivität, Spezifität, positiver und negativer prädiktiver Wert mit Hilfe der ausgegebenen Klassifizierungstabelle berechnet. Auch beim

Validierungsdatensatz wurde mit einem Klassifizierungsschwellenwert von 0,5 gearbeitet. Hinzu kommen ROC-Kurven und die zugehörigen AUC inklusive der Standardfehler sowie der asymptotischen Signifikanz. Die Kenngrößen des Regressionsmodells, berechnet mit dem Validierungsdatensatz verglichen mit den vorherigen Berechnungen auf Grundlage des Entwicklungsdatensatzes erlauben eine Aussage über die Validität der gesamten Berechnungen. Ein gleiches Verfahren wurde zur Validierung des im letzten Schritt berechneten PILGRIM-Scores mit den individuellen Variablen gewichtungen angewendet. Insgesamt lässt sich der in dieser Studie gewählte Ansatz als eine interne, zeitliche Validierung bezeichnen (vgl. Abschnitt Portionierung des Datensatzes) ³⁷.

4. Ergebnisse

4.1. Primärkodierung

Die Möglichkeit der Beantwortung einer Variable mit „Unbekannt“ führte insgesamt zu einem niedrigen Anteil fehlender Werte mit durchweg < 10%. Die fehlenden Werte wurden somit als *missing at random* eingestuft. Einen Überblick über im Rahmen der Primärkodierung neu codierte Variablen gibt Tabelle 6.2. im Anhang.

4.2. Deskriptive Statistiken

4.2.1. Deskription des Entwicklungsdatensatzes

Insgesamt umfasst der Entwicklungsdatensatz nach Bereinigung 942 PatientInnenfälle. 445 (= 52,8%) Fälle im Entwicklungsdatensatz erhielten ein Antibiotikum und 497 Fälle (= 47,2%) erhielten kein Antibiotikum. Bei Fällen mit positivem Endpunkt, trat dieser bei rund 92% aller Fälle innerhalb der ersten drei Wochen nach Studieneinschluss ein. Zur weiteren Deskription des Entwicklungsdatensatzes siehe Tabelle 3.1.:

Variable [Einheit]	Mittelwert / Median	Standardabweichung	Min. / Max.	25. / 75. Perzentile / Interquartilsabstand (IQR)
Alter [Jahre]	61,4 / 64	13,5	18 / 91	55 / 71 / 16
Gewicht [kg]	81,4 / 80	17,8	34 / 181	69,7 / 91 / 21,3
BMI [kg/m ²]	27,3 / 26,6	5,3	14 / 51,1	23,7 / 30,4 / 6,7
Charlson Comorbidity Index (CCI) [Punkte]	4 / 4	2,3	0 / 15	2 / 5 / 3

Tabelle 3.1.: Basisdeskription des Entwicklungsdatensatzes

Innerhalb des Entwicklungsdatensatzes fand sich folgende Verteilung der Abteilungen:

Versorgende Abteilung	Anteil [%]
Hämato-Onkologie	39,4
Kardiochirurgie	25
Kardiologie	6,7
Viszeralchirurgie	5,8
(Allgemein-) Internistisch	5,2
Hals-Nasen-Ohrenheilkunde	3,6

Gastroenterologie	2,9
Thoraxchirurgie	2,4
Urologie	2,3
Sonstige	6

Tabelle 3.2.: Verteilung der versorgenden Abteilungen innerhalb des Entwicklungsdatensatzes

35,9% der Fälle erreichten ≥ 5 Punkte im CCI und können somit als relevant vorerkrankt angesehen werden. Insgesamt bildeten 362 weibliche (= 38,4%) und 580 männliche PatientInnen (= 61,6%) den Entwicklungsdatensatz. Hierbei erhielten rund 15% der Fälle in den vergangenen 6 Monaten ein Antibiotikum, rund 39% waren in den vergangenen 6 Monaten vor Studieneinschluss bereits mindestens einmal hospitalisiert und für 44,5% der eingeschlossenen PatientInnen war eine Operation während des Krankenhausaufenthaltes geplant. Rund ein Drittel der PatientInnen (33,3%) erhielt während des Studieneinschlusses oder in den vorausgegangenen 4 Wochen immunsupprimierende Medikation. PatientInnen der Hämato-Onkologie waren tendenziell etwas jünger (Mittelwert: 57,2 Jahre) als die Gesamtheit des Entwicklungsdatensatzes.

4.2.2. Deskription des Validierungsdatensatzes

Im Validierungsdatensatz waren insgesamt 316 PatientInnenfälle enthalten. 57,6% der PatientInnen (182 Fälle) erhielten kein Antibiotikum, während 134 PatientInnen (= 42,4%) ein Antibiotikum erhielten. PatientInnen, die ein Antibiotikum erhielten, erhielten dieses zu rund 87% innerhalb der ersten drei Wochen nach Studieneinschluss (Woche 1 = 39,4%, Woche 2 = 40,9%).

Variable [Einheit]	Mittelwert / Median	Standard-abweichung	Min. / Max.	25. / 75. Perzentile / Interquartilsabstand (IQR)
Alter [Jahre]	61,2 / 63	14,2	18 / 91	54 / 72 / 18
Gewicht [kg]	81,3 / 80	17,6	41,3 / 140	69,9 / 92 / 23
BMI [kg/m ²]	27,4 / 26,6	5,3	16,2 / 48,3	24 / 29,8 / 5,8
Charlson Comorbidity Index (CCI) [Punkte]	3,8 / 4	2,1	0 / 15	2 / 5 / 3

Tabelle 3.3.: Basisdeskription des Validierungsdatensatzes

Innerhalb des Validierungsdatensatzes fand sich folgende Verteilung der Abteilungen:

Versorgende Abteilung	Anteil [%]
Hämato-Onkologie	30,7
Kardiochirurgie	18,7
Viszeralchirurgie	15,5
(Allgemein-) Internistisch	11,1
Kardiologie	7
Infektiologie	4,1
Gefäßchirurgie	2,5
Gynäkologie	2,5
Sonstige	7,5

Tabelle 3.4.: Verteilung der versorgenden Abteilungen innerhalb des Validierungsdatensatzes

Eine relevante Multimorbidität zeigte sich bei rund 33,1% der inkludierten Fälle im Validierungsdatensatz. Geschlechter teilten sich wie folgt auf: 40,2% (n=127) der Fälle waren weiblich, während männliche Patienten 59,8% (189) der Fälle insgesamt ausmachten. 12,7% der Fälle wurden in den vergangenen 6 Monaten bereits antibiotisch behandelt. Rund 40% waren 6 Monate vor Studieneinschluss bereits mindestens einmal hospitalisiert und für 44% war während des stationären Aufenthaltes eine elektive Operation geplant. Immunsupprimierende Medikation während dem stationären Aufenthalt oder bis zu 4 Wochen vor Aufnahme erhielten 28,2% der PatientInnen. PatientInnen der Hämato-Onkologie waren tendenziell auch hier etwas jünger (Mittelwert: 55 Jahre) als die Gesamtheit des Validierungsdatensatzes.

4.3. Risikofaktorenanalyse

Die Überprüfung eines möglichen Zusammenhangs zwischen einzelnen Variablen und dem Endpunkt ermöglichte im Folgenden die Sortierung aller Variablen nach der Stärke und Signifikanz ihres Einflusses.

4.3.1. Pearson Chi-Quadrat-Test

In der folgenden Tabelle (Nr. 3.5.) wurden alle untersuchten Variablen nach ihrer Signifikanz im Pearson Chi-Quadrat-Test sortiert:

Beschreibung	Variablen- beschriftung	Signifi- kanz p	Phi ϕ
Hämatologische Neoplasien	<i>Leukemialymphoma</i>	<0,001	0,32
Alle Abteilungen außer kardiologische- und kardiochirurgische, Risikoabteilungen nach Kodierung, Version 2	<i>NOcardio_departments</i>	<0,001	0,20
Risikoabteilungen nach Kodierung, Version 3	<i>risk_departments</i>	<0,001	0,15

Immunsupprimierende Medikation aktuell oder vor 4 Wochen	<i>v_224</i>	<0,001	0,15
Antibiotische Behandlung in den letzten 6 Monaten	<i>v_154</i>	<0,001	0,17
Hospitalisierung in den letzten 6 Monaten	<i>v_158</i>	<0,001	0,19
Risikoabteilungen nach Kodierung, Version 1	<i>dummy_department_aggregated_internal ANDother</i>	<0,001	0,23
Keine elektive OP während stationärem Aufenthalt geplant	<i>Noelectivesurgery</i>	<0,001	0,11
Keine Dyslipidämie als Risikofaktor	<i>Nodyslipoproteinemia</i>	<0,001	0,15
Alter unter 70 Jahren	<i>threshold_ageunder70</i>	<0,001	0,13
Keine Herzinsuffizienz als Risikofaktor	<i>Nocardiacinsufficiency</i>	<0,001	0,13
Protonen-Pumpen-Inhibitor oder H₂-Blocker Einnahme	<i>antacida</i>	0,001	0,10
Keine Hypertonie als Risikofaktor	<i>Nohypertonia</i>	0,003	0,10
Kein solides Malignom als Risikofaktor	<i>Nosolidtumor</i>	0,022	0,08
Keine periphere arterielle Verschlusskrankheit als Risikofaktor	<i>Nopavk</i>	0,029	0,07
Bekannte Kolonisation mit ESBL-Bildnern	<i>v_278</i>	0,031	0,07
Kein Nordeuropa/-amerika-Reise anamnestisch als Risikofaktor	<i>Nonothermtravel</i>	0,035	0,07
Männliches Geschlecht	<i>u_gender</i>	0,044	0,07
Z.n. Infektion mit ESBL-Bildnern	<i>v_478</i>	0,076	0,06
Allergie gegen Antibiotika (nicht näher spezifiziert)	<i>v_290</i>	0,092	0,06
Bekannte Kolonisation mit multiresistenten Erregern	<i>mdro_colon</i>	0,097	0,05
Kollagenosen	<i>v_5787</i>	0,102	-0,05
Reise in tropische Länder anamnestisch	<i>v_457</i>	0,129	0,05
Aufnahme wegen Klappenvitien	<i>new.intakes_valvediseases</i>	0,156	-0,05
Diabetes mellitus	<i>dummy_cci.diabetes</i>	0,177	-0,04
Z.n. Myokardinfarkt	<i>v_5781</i>	0,272	-0,04
COPD	<i>copd</i>	0,308	0,03
Untergewicht mit BMI < 18,5 kg/m ²	<i>bmi_underweight</i>	0,310	0,03
Z.n. Infektion mit multiresistenten Erregern	<i>mdro_infect</i>	0,334	0,03
Z.n. Infektion mit VRE	<i>v_477</i>	0,339	0,03
Zerebrovaskuläre Vorerkrankungen	<i>v_5784</i>	0,378	-0,03
Reise in Länder mit endemischen ESBL-E/Carbapenemase-Bildnern	<i>v_456</i>	0,450	0,03
Behandlungsbedürftige Adipositas mit BMI > 30 kg/m ²	<i>fatty</i>	0,451	0,03
Hepatische Insuffizienz	<i>dummy_cci.liver</i>	0,509	-0,02
Multimorbide mit CCI > 5 Punkte	<i>multimorbid</i>	0,593	-0,02
Bekannte Kolonisation mit VRE	<i>v_277</i>	0,625	0,02
Keine Reisen in den letzten 12 Monaten in fremde Länder	<i>v_454</i>	0,638	-0,02
Vorhofflimmern	<i>vhf</i>	0,663	-0,01
Überweisung aus einem anderen Krankenhaus	<i>v_12032</i>	0,667	-0,02

Glomeruläre- und tubulointerstitielle Nierenerkrankungen, CKD	<i>kidney</i>	0,840	0,01
Asthma bronchiale	<i>asthma</i>	0,850	0,01

Tabelle 3.5.: Korrelation der Variablen mit dem Endpunkt im Pearson-Chi-Quadrat-Test

Insgesamt zeigten sich in der Sortierung nach der Signifikanz im Pearson Chi-Quadrat-Test 18 Variablen hinsichtlich ihrer Korrelation mit dem Endpunkt signifikant unter $p < 0,05$. Bei der Betrachtung der Zusammenhangsmaße fiel die Variable *hämatologische Malignome* als einzige mit einem mittelstarken Effekt auf den Endpunkt auf ($\phi = 0,32$). Risikoabteilungen, iatrogene Immunsuppression, Antibiotika oder Hospitalisierung in den letzten 6 Monaten, nicht elektiv-chirurgische PatientInnen, PatientInnen ohne Dyslipidämie, Alter unter 70 Jahren, PatientInnen ohne Herzinsuffizienz und Protonen-Pumpen-Inhibitor / H₂-Blocker-Einnahme, wiesen schwache, aber signifikante Effektstärken in der Korrelation mit dem Endpunkt auf ($0,1 < \phi < 0,3$). Auffällig war ein zumindest statistisch protektiver Effekt einiger Komorbiditäten gegenüber dem Endpunkt der Antibiotikagabe. Zur besseren Nachvollziehbarkeit und der vereinfachten Integration in ein logistisches Regressionsmodell wurden die Ausprägungen jener protektiven Variablen umgekehrt, sodass die gegenteiligen Formulierungen (*nicht-*) analog zum zuvor protektiven Effekt (negative Korrelation) nun in ihrer positiven Korrelation dargestellt werden konnten. Unter dem $p < 0,05$ -Level traf dies auf die Variablen geplante Operation, Dyslipidämie, Herzinsuffizienz, Hypertonie, solide Malignome, peripher arterielle Verschlusskrankheiten und Reisen nach Nordamerika oder -Europa zu. Für das nicht-Vorhandensein einer Dyslipidämie ergab sich im Pearson Chi-Quadrat-Test so beispielsweise eine hochsignifikante Korrelation ($p < 0,001$) mit schwachem Zusammenhangsmaß ($\phi = 0,15$). In einer tiefergehenden Analyse der verschiedenen internistischen Vorerkrankungen mit statistisch protektiven Effekt, konnte festgestellt werden, dass diese hauptsächlich in kardiologischen- und kardiochirurgischen Abteilungen kodiert wurden (durchweg signifikante Zusammenhänge mit $p < 0,001$ für die Variablen Dyslipoproteinämie, ischämische Kardiomyopathie, Vorhofflimmern, Herzinsuffizienz und peripher arterielle Verschlusskrankheit). Eine neu kodierte Variable für eben diese Abteilungen wurde ebenfalls auf ihre Korrelation mit dem Endpunkt überprüft und auch hier ergab sich eine signifikante negative Korrelation bei $p < 0,001$ und $\phi = - 0,2$.

4.3.2. Univariate Regressionsmodelle

Es folgte die Betrachtung der Variablen hinsichtlich ihres Zusammenhanges mit dem Endpunkt in univariaten Regressionsmodellen. In der folgenden Tabelle (Nr. 3.6.) wurden alle berücksichtigten Variablen nach ihrer Odds Ratio und dem Regressionskoeffizienten in den univariaten logistischen Regressionsmodellen sortiert.

Beschreibung	Regressionskoeffizient	Sign. P	OR	95%-KI
Z.n. Infektion mit ESBL-Bildnern	1,73	0,115	5,64	0,66 – 48,43
Hämatologische Neoplasien	1,47	<0,001	4,35	3,20 – 5,90
Keine Dyslipidämie als Risikofaktor	1,42	<0,001	4,12	2,11 – 8,05
Bekannte Kolonisation mit ESBL-Bildnern	1,33	0,044	3,79	1,04 – 13,84
Immunsupprimierende Medikation aktuell oder vor 4 Wochen	1,13	<0,001	3,09	2,33 – 4,10
Risikoabteilungen nach Kodierung, Version 2	0,97	<0,001	2,65	2,00 – 3,51
Antibiotische Behandlung in den letzten 6 Monaten	0,95	<0,001	2,58	1,77 – 3,7
Keine Hypertonie als Risikofaktor	0,94	0,003	2,55	1,36 – 4,77
Alle Abteilungen außer kardiologische- und kardiochirurgische	0,91	<0,001	2,48	1,86 – 3,30
Reise in tropische Länder anamnestisch	0,82	0,139	2,26	0,77 – 6,67
Z.n. Infektion mit VRE	0,81	0,352	2,25	0,41 – 12,32
Hospitalisierung in den letzten 6 Monaten	0,77	<0,001	2,17	1,66 – 2,83
Risikoabteilungen nach Kodierung, Version 1	0,72	<0,001	2,06	1,51 – 2,79
Bekannte Kolonisation mit multiresistenten Erregern	0,66	0,102	1,93	0,88 – 4,27
Keine Herzinsuffizienz als Risikofaktor	0,62	<0,001	1,86	1,26 – 2,56
Alter unter 70 Jahren	0,59	<0,001	1,80	1,35 – 2,39
Keine periphere arterielle Verschlusskrankheit als Risikofaktor	0,56	0,030	1,75	1,06 – 2,90
Z.n. Infektion mit multiresistenten Erregern	0,48	0,339	1,61	0,61 – 4,26
Kein Nordeuropa/-amerika-Reise anamnestisch als Risikofaktor	0,46	0,036	1,59	1,03 – 2,45
Keine elektive OP während stationärem Aufenthalt geplant	0,46	<0,001	1,59	1,22 – 2,06
Untergewicht mit BMI < 18,5 kg/m ²	0,44	0,314	1,56	0,66 – 3,68
Protonen-Pumpen-Inhibitor oder H₂-Blocker Einnahme	0,43	0,001	1,53	1,17 – 2,00
Allergie gegen Antibiotika (nicht näher spezifiziert)	0,40	0,093	1,49	0,94 – 2,36
COPD	0,35	0,310	1,42	0,72 – 2,77
Kein solides Malignom als Risikofaktor	0,34	0,022	1,41	1,05 – 1,90
Bekannte Kolonisation mit VRE	0,30	0,627	1,35	0,41 – 4,44
Männliches Geschlecht	0,27	0,044	1,31	1,01 – 1,71
Reise in Länder mit endemischen ESBL-E/Carbapenemase-Bildnern	0,18	0,450	1,19	0,76 – 1,87
Behandlungsbedürftige Adipositas mit BMI > 30 kg/m ²	0,11	0,451	1,12	0,84 – 1,50
Glomeruläre/tubulointerstitielle Nierenerkrankungen /CKD	0,07	0,840	1,07	0,57 – 2,00
Asthma bronchiale	0,06	0,850	1,06	0,56 – 2,02
Keine Reisen in den letzten 12 Monaten in fremde Länder	-0,65	0,638	0,94	0,71 – 1,23
Multimorbide mit CCI > 5 Punkte	-0,89	0,593	0,92	0,66 – 1,27

Überweisung aus einem anderen Krankenhaus	-0,16	0,667	0,85	0,41 – 1,76
Vorhofflimmern	-0,16	0,663	0,85	0,41 – 1,77
Z.n. Myokardinfarkt	-0,23	0,273	0,79	0,53 – 2,00
Zerebrovaskuläre Vorerkrankungen	-0,28	0,379	0,76	0,41 – 1,40
Aufnahme wegen Klappenvitien	-0,28	0,157	0,76	0,51 – 1,12
Diabetes mellitus	-0,28	0,178	0,75	0,49 – 1,14
Hepatische Insuffizienz	-0,30	0,511	0,74	0,30 – 1,82
Kollagenosen	-0,66	0,108	0,52	0,23 – 1,16

Tabelle 3.6.: Korrelation der Variablen mit dem Endpunkt in den univariaten Regressionsmodellen

In der Betrachtung der univariaten Regressionsmodelle zeigten sich insgesamt ebenfalls 18 Variablen hinsichtlich ihrer Korrelation mit dem Endpunkt unter dem $p < 0,05$ Niveau signifikant. Hierbei reichten die Regressionskoeffizienten von 1,73 (Variable Z.n. Infektion mit ESBL-Bildnern) bis -0,89 (Variable multimorbide nach CCI-Definition ≥ 5 Punkte). Ein statistisch scheinbar protektiven Effekt der Variable Multimorbidität ergab sich auch in der Analyse mit Pearson Chi-Quadrat. Für die Odds Ratios der univariaten Regressionsmodelle ergaben sich ähnliche Ergebnisse mit einem Maximalwert von 5,64 für vorausgegangene Infektionen mit ESBL-Bildnern und 0,52 für Kollagenosen als Vorerkrankungen. Zu den Variablen mit einer Odds Ratio >1 und gleichzeitig einem 95%-Konfidenzintervall >1 (untere Grenze) gehörten: Hämatologische Neoplasien, keine Dyslipidämie, bekannte Kolonisation mit ESBL-Bildnern, immunsupprimierende Medikation, Risikoabteilungen in verschiedenen Kodierungen, antibiotische Behandlung oder Hospitalisierung in den letzten 6 Monaten, keine Hypertonie, PAVK oder Herzinsuffizienz, Alter kleiner 70 Jahre, keine Nordeuropa-/Amerika Reise, keine elektive OP während des stationären Aufenthaltes geplant, antazide Medikation, sowie männliches Geschlecht (s. Abb. X-Y unten).

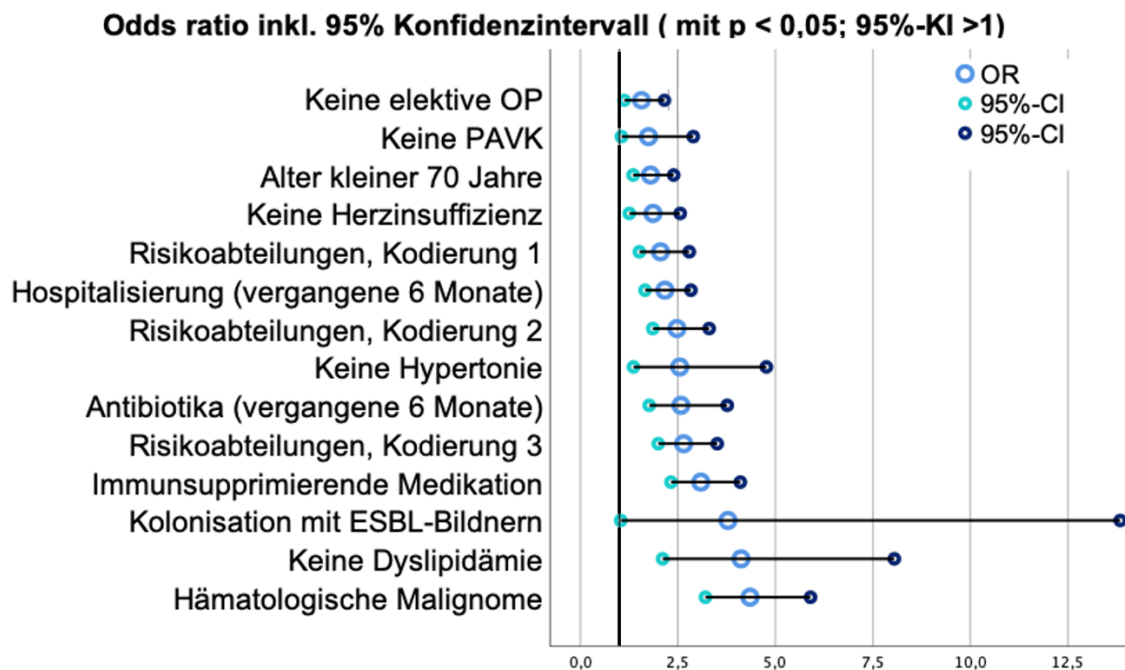


Abbildung 3.1.: Wald-Diagramm der Variablen der univariaten Regressionsmodelle (Auswahl)

Neben den verschieden kodierten Risikoabteilungen sowie den umgekehrt protektiv formulierten Variablen mit hoher Korrelation zu kardiologischen Abteilungen hoben sich vor allem folgende Variablen in der Risikofaktorenanalyse hervor (siehe Tabelle 3.7. unten):

Variable	Univariate Regression			Pearson-Chi	
	p	OR	95%-KI	p	ϕ
Hämatologische Malignome	<0,001	4,4	3,2 – 5,9	<0,001	0,32
Immunsupprimierende Medikation	<0,001	3,1	2,3 – 4,1	<0,001	0,15
Vergangene Hospitalisation	<0,001	2,2	1,6 – 2,8	<0,001	0,19
Vergangene antibiotische Behandlung	<0,001	2,6	1,8 – 3,8	<0,001	0,17
Alter unter 70 Jahren	<0,001	1,8	1,4 – 2,4	<0,001	0,13

Tabelle 3.7.: Synopsis hochsignifikanter Variablen in Pearson-Chi-Quadrat-Test und univariater Regression

4.4. Binäre logistische Regressionsmodelle

In 42 Runden wurden, bedingt durch die 10-fache Kreuzvalidierung, 420 Regressionsmodelle berechnet. Hinzu kommen ca. 200 weitere Iterationen im Rahmen von Probe- und Testberechnungen. Zur Bewertung der Modellleistung wurden die Kennzahlen PPV, NPV, Sensitivität, Spezifität und AUC herangezogen. Eine detaillierte Beschreibung des Verlaufes der Kennzahlen kann dem Anhang entnommen werden und ist in Abbildung 3.2. unten dargestellt.

In der Zusammenschau der Verläufe aller Kennzahlen konnten einige simultane Kurvenveränderungen durch das Entfernen einzelner Variablen beobachtet werden. Das

Entfernen der Variable *multimorbid*, was Multimorbidität mit mehr als 5 Punkten im Charlson Comorbidity Index entspricht, führte so in Iteration 8 zu einem Anstieg der Sensitivität (+ 0,6%), des NPV (+ 1,4%) sowie der Spezifität (+ 0,6%). Nach der 12. Iteration kam es zu einem Abfall von Sensitivität, Spezifität und des NPV. Hierfür verantwortlich könnte das Entfernen der Variable *Zustand nach VRE-Infektion* sein. Bemerkenswert ist jedoch die direkte Erholung der Kennzahlen in der folgenden Iteration 14 durch das Entfernen der Variable *Zustand nach MRE-Infekt*. Für den NPV ergab sich beispielsweise so im Verlauf der 3 Iterationen zunächst ein Abfall von – 4,4% und darauf ein Anstieg um + 4,9%. Ähnliche Verläufe, wenn auch nicht so deutlich, waren für die Sensitivität sowie Spezifität zu erkennen. Plötzliche Anstiege und ein nachhaltig höheres Niveau ergaben sich für Spezifität und den PPV nach dem Entfernen der Variable für Klappenvitien (Iteration 20). Hier sprang der PPV um + 6,4% auf 77,3% und die Spezifität des Regressionsmodells um + 4,9% auf 80,9%. Aber auch Sensitivität (+ 3%), NPV (+ 2,1%) und AUC (+ 0,9%) stiegen durch das Entfernen der Variable simultan an. Während sich PPV, AUC und Spezifität nach dem Anstieg und weiteren Iterationen auf ihrem Niveau halten konnten, bildete Iteration 20 für Sensitivität und NPV den Zenit, nach welchem ein stetiger Abfall mit dem Wegfallen weiterer Variablen zu erkennen war. Der letzte simultane Abfall aller Kurven ergab sich mit der 36. Iteration, bei welcher die Variable *nicht elektiv-chirurgische PatientInnen* entfernt wurde. Das Entfernen der Variable führte für die Sensitivität zu einem Abfall von – 9,4%, die Spezifität sank um – 2,1%, PPV und NPV um – 5,5% und – 4,3% bei der AUC.

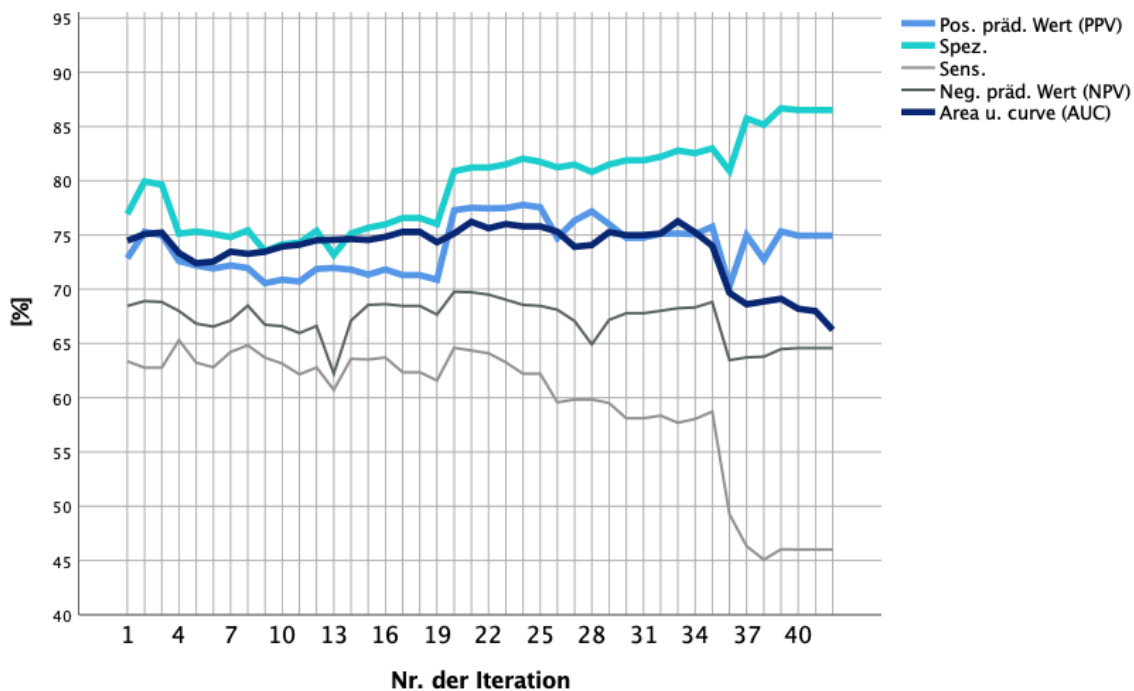


Abbildung 3.2.: Entwicklung der Kennzahlen im Verlauf der Iterationen

Das simultane Abfallen der Leistung der Kennzahlen sollte wie in der Methodik beschrieben die Basis für die konsekutive Entwicklung des PILGRIM-Scores darstellen. Dies traf somit für die 35. Iteration zu, in welcher das Regressionsmodell noch durch die Variablen *leukemialymphoma*, Risiko-Abteilungen Version 1, Risiko-Abteilungen Version 2, *v_224* (immunsupprimierende Medikation), *v_154* (Antibiotika in den letzten 6 Monaten), *v_158* (Hospitalisierung in den letzten 6 Monaten), Risiko-Abteilungen Version 3 und *NOelectivesurgery* gebildet wurde. Für dieses Regressionsmodell ergaben sich mit den noch 8 inkludierten Variablen folgende durchschnittliche Werte der Prädiktionsgenauigkeit: Sensitivität 58,7%, Spezifität 83%, PPV 75,8%, NPV 68,8%, AUC 74%. Die durchschnittliche Signifikanz der Klassifizierungstabellen bei einem Klassifizierungsschwellenwert von 0,5 betrug hier 0,001, für die ROC ergab sich eine mittlere Signifikanz von 0,011.

4.5. Entwicklung der optimalen Kombination

Gemäß der Methodik wurde mit den in Iteration 35 noch inkludierten Variablen die Entwicklung der optimalen Variablen-Kombination gestartet. Die probatorische Berechnung weiterer Regressionsmodelle mit 4 bis 8 der aus Iteration 35 zur Verfügung stehenden Variablen war nun der nächste Schritt und kann in Tabelle 3.8. nachvollzogen werden. Aus der Primärkodierung der deskriptiven Analyse des Entwicklungsdatensatzes hatten sich drei verschiedene Kodierungen möglicher Risikoabteilungen ergeben. In den Kombinationsversuchen zur Optimierung der Modelleistung wurden jedoch immer nur eine der drei Varianten inkludiert. Zum Ausgleich der somit zur Auswahl fehlenden 2 Variablen wurden die in den vorangegangenen beiden Iterationen (33, 34) bereits entfernten Variablen keine-Dyslipidämie (*NOdyslipoproteinemia*) sowie Altersschwelle <70 Jahre (*threshold_ageunder70*) wieder hinzugenommen (Vorwärtsselektion). Beide Variablen hatten sich in der univariaten Betrachtung der Korrelation mit dem Endpunkt als signifikant und risikobehaftet hinsichtlich des Eintretens des Endpunktes herausgestellt. Final zur Auswahl bei der Entwicklung des optimalen Regressionsmodells standen also die Variablen 1. Risikoabteilungen in den Versionen 1-3, 2. Hämatologische Malignome, 3. medikamentöse Immunsuppression, 4. Antibiotika in der Vergangenheit, 5. Hospitalisierung in der Vergangenheit, 6. Nicht-elektiv-chirurgisch, 7. Keine-Dyslipidämie und 8. Alter < 70 Jahre. Mit dem Ziel der einfacheren Implementation des Scores wurde im Folgenden dann eine Kombination aus 4-6 Variablen mit der in Relation besten Modelleistung gesucht. Zur Beurteilung der Modelleistung wurden auch hier wieder Sensitivität, Spezifität, NPV, PPV sowie die AUC berechnet und betrachtet. Die Inklusion der Variablen *NOelectivesurgery* und Risikoabteilungen in den Versionen 1-3 in das gleiche Regressionsmodell wurde vermieden.

Eine detaillierte Beschreibung über die Komposition der Iterationen 35a bis 35m kann dem Anhang entnommen werden.

Variablen im Modell	Sens.	Spez.	PPV	NPV	AUC	Kennung
1. Hämatologische Malignome 2. Medikamentöse Immunsuppression 3. Hospitalisierung <6 Monate 4. ABX <6 Monate 5. Risikoabteilung V1 6. Keine Dyslipoproteinämie 7. Alter < 70 Jahre	60,5	80,6	75,1	68,2	75,4	35a
1. Hämatologische Malignome 2. Medikamentöse Immunsuppression 3. Hospitalisierung <6 Monate 4. ABX <6 Monate 5. Alter < 70 Jahre 6. Risikoabteilung V2	49,2	82	70,9	64,4	69,1	35b
1. Hämatologische Malignome 2. Medikamentöse Immunsuppression 3. Hospitalisierung <6 Monate 4. ABX <6 Monate 5. Alter < 70 Jahre 6. Risikoabteilung V3	46	85	75	63,7	69,4	35c
1. Hämatologische Malignome 2. Medikamentöse Immunsuppression 3. Hospitalisierung <6 Monate 4. ABX <6 Monate 5. Alter < 70 Jahre	46	85	75	63,7	70,2	35d
1. Medikamentöse Immunsuppression 2. Hospitalisierung <6 Monate 3. ABX <6 Monate 4. Alter < 70 Jahre	46,4	80	67,6	62,7	68,4	35e
1. Hämatologische Malignome 2. Medikamentöse Immunsuppression 3. Hospitalisierung <6 Monate 4. ABX <6 Monate 5. Alter < 70 Jahre 6. Keine elektive OP geplant	50,3	85,5	76,3	65,2	73,8	35f
1. Hämatologische Malignome 2. Medikamentöse Immunsuppression 3. Hospitalisierung <6 Monate 4. ABX <6 Monate 5. Keine elektive OP geplant	51,6	84,6	74,8	66,3	73,8	35g
1. Medikamentöse Immunsuppression 2. Hospitalisierung <6 Monate 3. ABX <6 Monate 4. Keine elektive OP geplant	48,6	80,3	69	63,4	67,7	35h
1. Hämatologische Malignome 2. ABX <6 Monate 3. Hospitalisierung <6 Monate 4. Keine elektive OP geplant	47,3	86	75	64,8	71,6	35i
1. Hämatologische Malignome 2. Medikamentöse Immunsuppression 3. Hospitalisierung <6 Monate 4. Keine elektive OP geplant	50,7	86,5	76,9	66,1	73,7	35j
1. Hämatologische Malignome 2. Medikamentöse Immunsuppression 3. ABX <6 Monate	51,5	84,8	75,2	66,2	71,7	35k

4. Keine elektive OP geplant						
1. Hämatologische Malignome	45,3	85,2	72,8	63,9	67,1	35l
2. Medikamentöse Immunsuppression						
3. ABX <6 Monate						
1. Hämatologische Malignome	45,5	86,4	75,3	63,6	68,7	35m
2. Medikamentöse Immunsuppression						
3. ABX <6 Monate						
4. Hospitalisierung <6 Monate						
1. Hämatologische Malignome	50,7	86,5	76,9	66,1	73,7	35ja
2. Medikamentöse Immunsuppression						
3. Hospitalisierung <6 Monate						
4. Elektive OP geplant						

Tabelle 3.8.: Kenngrößen der probatorischen Iterationen (Medikamentöse Immunsuppression = medikamentöse Immunsuppression aktuell oder in den vergangenen 4 Wochen, ABX <6 Monate = Antibiotika in den vergangenen 6 Monaten, Hospitalisierung <6 Monate = Hospitalisierung in den vergangenen 6 Monaten)

Zusammenfassend konnte mit den Modellen 35j und 35g die besten Prädiktionsleistungen erzielt werden. In Anbetracht der Tatsache, dass die Modelleistung unter Entfernung der Variable v_{154} kaum abfiel und sich im Bereich der Spezifität und PPV sogar verbesserte, wurde die Entwicklung des PILGRIM Scores mit dem schlankeren Modell 35j mit 4 übriggebliebenen Variablen fortgesetzt. Inkludierte Variablen waren in diesem Modell: Hämatologische Malignome, medikamentöse Immunsuppression, Hospitalisierung in der Vergangenheit und PatientInnen, die nicht zur Durchführung einer elektiven Operation aufgenommen wurden (*NOelectivesurgery*).

4.6. Überprüfung auf Multikollinearität

Eine Überprüfung der Multikollinearität der finalen Variablenselektion nach Kendall-Tau-b ergab folgende Korrelationskoeffizienten:

	HM	MI	kOP	H
HM	-	0,63*	0,58*	0,35*
MI	0,63*	-	0,56*	0,32*
kOP	0,58*	0,56*	-	0,26*
H	0,35*	0,32*	0,26*	-

Tabelle 3.9.: Multikollinearitätsanalyse nach Kendall-Tau-b. HM = Hämatologische Malignome, MI = Medikamentöse Immunsuppression, kOP = keine elektive OP geplant, H = Vergangene Hospitalisierung. Felder mit * markiert sind signifikant unter $p < 0,05$

Die Analyse zeigt, dass es signifikante Korrelationen zwischen den einzelnen Variablen gibt. Dies trifft vor allem für hämatologische Malignome und medikamentöse Immunsuppression zu, aber auch für hämatologische Malignome / medikamentöse Immunsuppression und die Variable keine elektive OP.

4.7. Validierung des finalen logistischen Regressionsmodells

Das beschriebene logistische Regressionsmodell mit der besten Modelleistung bei gleichzeitig niedriger Anzahl an inkludierten Variablen wurde nun mit den bis dahin nicht

verwendeten Daten des Validierungsdatensatzes neu berechnet. Die Ergebnisse können Tabelle 3.10. entnommen werden. Ziel war eine Bestätigung der mit dem Entwicklungsdatensatz erreichten Modelleleistung zur externen Validierung.

	Entwicklung	Validierung	Differenz
Sensitivität [%]	50,7	<u>50</u>	-0,7%
Spezifität [%]	86,5	<u>87,6</u>	+1,1%
PPV [%]	76,9	<u>75</u>	-1,9%
NPV [%]	66,1	<u>70,1</u>	+4,0%
AUC [%]	73,7	<u>74,8</u>	+1,1%
Standardfehler	0,053	<u>0,029</u>	-0,024
Signifikanz	0,002	<u><0,001</u>	-0,002

Tabelle 3.10.: Ergebnisse der Modelleleistungen bei Berechnung mit dem Entwicklungsdatensatz vs. Berechnung mit dem Validierungsdatensatz

Im Vergleich der Modelleleistungen in der Berechnung mit dem Entwicklungs- und dem Validierungsdatensatz zeigte sich eine verbesserte Leistung bei fast allen Kennzahlen. Lediglich im Bereich der Sensitivität und des PPV ergaben sich leichte Verschlechterungen.

4.8. PILGRIM Score – Berechnung der individuellen Variablen-gewichtung

Im Anschluss an die Validierung des besten Regressionsmodells folgte die Kalkulation der individuellen Gewichtungen der Score-Variablen analog zu den in der Methodik vorgestellten 3 Berechnungsarten (nach Werfel et al., Schneeweiss et al., Matsushita et al.)⁴⁰⁻⁴². Hier wurde stets der Mittelwert des Regressionskoeffizienten beziehungsweise der OR aus der 10-fachen Kreuzvalidierung mit dem Entwicklungsdatensatz verwendet. In Tabelle 3.11. können die durchschnittlichen Regressionskoeffizienten sowie OR und respektiven Individualgewichtungen entnommen werden.

Variablen	Ø Regressions-koeffizient	Ø OR	Werfel	Matsushita	Schneeweiss
Medikamentöse Immunsuppression	0,71	2,04	-0,8	-1,4	2
Hämatologische Malignome	1,57	4,83	-1,8	-3,0	5
Vergangene Hospitalisierung	0,36	1,43	-0,4	-0,7	1

Keine elektive OP geplant	-0,88	0,42	1,0	1,7	-
---------------------------	-------	------	-----	-----	---

Tabelle 3.11.: Variablenbewertungen nach Werfel et al., Matsushita et al. und Schneeweiss et al.

Auffällig war der neu aufgetretene negative durchschnittliche Regressionskoeffizient der Variable *keine elektive OP* bei der Betrachtung der variablenspezifischen Regressionskoeffizienten innerhalb des Regressionsmodells. In der eingangs beschriebenen univariaten Betrachtung zur Sortierung der Risikofaktoren hatte die Variable noch einen positiven Regressionskoeffizienten von 0,46 bei einer OR von 1,59 (95%-KI: 1,22-2,06) erreicht. Alle 3 Berechnungsmethoden liefern Individualbewertungen von Risikofaktoren mit positiver Korrelation mit dem Endpunkt und konnten somit keine plausiblen Bewertungen mit dem nun vorliegenden protektiven Faktor mit negativer Korrelation mit dem Endpunkt darstellen.

Es erfolgte eine probatorische Umstellung des finalen Regressionsmodells und der Austausch der Variable *keine elektive OP* mit der ursprünglichen Variable v_186 (elektive OP geplant), da die Bewertungsberechnung nur sinnvoll erfolgen kann, wenn alle Prädiktorvariablen gleichgerichtet sind (vgl. Iteration 35ja in Tabelle 3.8. oben). Die erneute Berechnung der Individualbewertungen ergab folgendes Ergebnis (Tabelle 3.12.):

Variablen	Ø Regressionskoeffizient	Ø Werfel OR	Matsushita	Schneeweiss	Referenz	
Medikamentöse Immunsuppression	0,71	2,04	2,0	1,3	2	1
Hämatologische Malignome	1,57	4,83	4,4	2,9	5	1
Vergangene Hospitalisierung	0,36	1,43	1,0	0,7	1	1
Elektive OP geplant	0,88	2,42	2,5	1,7	3	1
Maximalpunktzahl:			9,9	6,6	11	4
∑ Einzelbewertungen						

Tabelle 3.12.: Variablenbewertungen nach Werfel et al., Matsushita et al. und Schneeweiss et al. inklusive der Variable „Elektive OP geplant“ sowie Referenzbewertung ohne individuelle Werte

Die Regressionskoeffizienten innerhalb der logistischen Regressionen waren nun für alle 4 Variablen positiv, was die Berechnung der individuellen Bewertungen ermöglichte. Für die Methode nach Werfel und Matsushita wurde die Bewertung im Entwicklungsstadium auf die erste Dezimalstelle gerundet. Insgesamt ergaben sich entsprechend ähnliche Werte für die einzelnen Variablen, wobei die Spanne der Bewertungen bei Schneeweiss mit 4 Punkten am

größten war. Hier ergab sich für Hämatologische Malignome 5 Punkte im Score und für eine vorangegangene Hospitalisierung nur 1 Punkt. Bei der Methode nach Matsushita fällt die Spanne mit 2,2 Punkten deutlich kleiner aus.

4.9. PILGRIM Score – Schwellenwertfindung

Im Folgenden wurde nun für jeden PatientInnenfall des Entwicklungsdatensatzes die theoretisch erreichte Punktzahl des additiven Score-Modells nach Werfel, Schneeweiss und Matsushita berechnet. In Zusammenschau mit den tatsächlich erreichten Endpunkten konnten so verschiedene Schwellenwerte der Modelle simuliert werden, nach denen das Modell einen Endpunkt als positiv oder negativ ausgeprägt vorhersagen würde. Für die Methode nach Werfel ergaben sich so je nach Schwellenwert unterschiedliche Modelleleistungen (Tabelle 3.13.):

Schwellenwert	Mod.+/Ep.+	Mod.+/Ep.-	Mod.-/Ep.+	Mod.-/Ep.-	Sens.	Spez.	PPV	NPV
9	0	1	423	465	0,0	99,8	0,0	52,4
8	0	1	423	465	0,0	99,8	0,0	52,4
7	122	33	301	433	28,8	92,9	78,7	59,0
6	161	53	262	413	38,1	88,6	75,2	61,2
5	181	61	242	405	42,8	86,9	74,8	62,6
4	205	79	218	387	48,5	83,0	72,2	64,0
3	263	154	160	312	62,2	67,0	63,1	66,1
2	382	358	41	108	90,3	23,2	51,6	72,5

Tabelle 3.13.: Ergebnisse der Vier-Felder Tafel in Abhängigkeit des Schwellenwertes des Modells (Werfel) und zugehörige Modelleleistungen („Mod. +“ = Endpunkt laut Modell positiv; „Ep. +“ = Endpunkt im Entwicklungsdatensatz tatsächlich positiv, „Mod. -“, „Ep. -“ = analog)

Generell gilt bei allen Methoden und Schwellenwerten: Je höher der Schwellenwert eines Prädiktionsmodells für die Annahme des positiven Endpunktes, desto höher sind Spezifität sowie PPV und desto niedriger sind Sensitivität sowie NPV. So zeigte sich bei der Methode nach Werfel bei einem Schwellenwert von 6, was rund 60% der Maximalpunktzahl (9,9) entspricht, eine Spezifität von fast 89% und ein PPV von 75%. Sowohl Sensitivität als auch NPV blieben beim 60%-Schwellenwert mit 38% bzw. 61% niedrig. Modelleleistungen oberhalb eines Schwellenwertes von 7 Punkten konnten mit dem Entwicklungsdatensatz und der Variablen gewichtung nach Werfel nicht berechnet werden, da es innerhalb des Datensatzes keine Fälle gab, bei welchen alle zur Berechnung herangezogenen Variablen positiv ausgeprägt waren. Das heißt, kein Fall konnte die Maximalpunktzahl erreichen. Die Berechnungen der individuellen Modelleleistung je nach Schwellenwert für die Gewichtungsmethoden nach Matsushita oder Schneeweiss lieferte ähnliche Ergebnisse. Auch

hier zeigte sich bei einem Schwellenwert von 60% (4 Punkte) eine Spezifität von rund 89%, ein PPV von 75% sowie eine Sensitivität von 38% und ein NPV von 61% bei der Methode nach Matsushita (vgl. Tabellen 6.3. und 6.4. im Anhang).

Die Berechnung der schwellenwertabhängigen Modelleistung ohne variablenindividuelle Gewichtung, d.h. mit einer Variablenengewichtung von 1 für alle Variablen, lieferte folgende Ergebnisse (Tabelle 3.14.):

Schwellenwert	Mod.+/Ep.+	Mod.+/Ep.-	Mod.-/Ep.+	Mod.-/Ep.-	Sens.	Spez.	PPV	NPV
3	125	33	298	433	29,6	92,9	79,1	59,2
2	249	140	174	326	58,9	70,0	64,0	65,2
1	394	387	29	79	93,1	17,0	50,4	73,1

Tabelle 3.14.: Ergebnisse der Vier-Felder Tafel in Abhängigkeit des Schwellenwertes des Modells ohne individuelle Variablenengewichtung (Referenz) und zugehörige Modelleistungen („Mod. +“ = Endpunkt laut Modell positiv; „Ep. +“ = Endpunkt im Entwicklungsdatensatz tatsächlich positiv, „Mod. -“, „Ep. -“ = analog)

Hier fällt auf, dass mit einem Schwellenwert von 3, was 75% der theoretisch zu erreichenden Maximalpunktzahl entspricht, und trotz fehlender individueller Variablenengewichtung eine hohe Prädiktionsleistung des Modells erreicht werden konnte (Spez.: ca. 93%, PPV 79%).

Insgesamt lieferten alle Berechnungsmethoden gute Ergebnisse in der Spezifität und im PPV. Sowohl Sensitivität als auch NPV blieben bei höheren Schwellenwerten stets niedrig und konnten mit Werten <60% keinen wirklichen Beitrag zur Prädiktionsleistung des Modells liefern.

4.10. Validierung des finalen PILGRIM Scores

Analog zur Validierung des finalen logistischen Regressionsmodells folgte nun die Validierung des PILGRIM Scores inklusive der drei verschiedenen Methoden der Variablenengewichtung. Hierzu wurden die gleichen Berechnungen zur Schwellenwertfindung (siehe oben) nun mit dem Validierungsdatensatz durchgeführt. Es ergaben sich folgende Modelleistungen (Tabelle 3.15.):

Schwellenwert	Mod.+/Ep.+	Mod.+/Ep.-	Mod.-/Ep.+	Mod.-/Ep.-	Sens.	Spez.	PPV	NPV
9	0	0	126	178	0,0	100,0	-	58,6
8	0	0	126	178	0,0	100,0	-	58,6
7	34	8	92	170	27,0	95,5	81,0	64,9
6	47	15	79	163	37,3	91,6	75,8	67,4

5	52	17	74	161	41,3	90,4	75,4	68,5
4	62	24	64	154	49,2	86,5	72,1	70,6
3	88	62	38	116	69,8	65,2	58,7	75,3
2	116	118	10	60	92,1	33,7	49,6	85,7

Tabelle 3.15.: Ergebnisse der Vier-Felder Tafel in Abhängigkeit des Schwellenwertes des Modells (Werfel) und zugehörige Modelleleistungen in der Validierung („Mod. +“ = Endpunkt laut Modell positiv; „Ep. +“ = Endpunkt im Entwicklungsdatensatz tatsächlich positiv, „Mod. -“, „Ep. -“ = analog)

Das Score-Modell mit der Gewichtung nach Werfel konnte sich genauso wie das ursprüngliche Regressionsmodell bei der Validierung in seiner Modelleleistung steigern (vgl. Abschnitt Validierung des finalen logistischen Regressionsmodells), sodass bei einem exemplarischen Schwellenwert von 60% der Maximalpunktzahl eine Spezifität von knapp 92% bei einem PPV von rund 76% erreicht werden konnte. Aber auch bei niedrigeren Schwellenwerten (beispielsweise 40% der Maximalpunktzahl) konnten akzeptable Modelleleistungen erreicht werden. Aufgrund der Ähnlichkeit der Berechnungsmethoden nach Schneeweiss und Matsushita zeigten sich analoge Ergebnisse beim Vergleich der Modelleleistungen auf dem Entwicklungsdatensatz bei der Validierung (vgl. Tabellen 6.5. und 6.6. im Anhang).

Die Entscheidung über einen passenden Schwellenwert bleibt eine Individualentscheidung je nach klinischem Setting. Der vorgeschlagene 60%-Schwellenwert liefert eine 4,4-fache positive Likelihood-Ratio bei einem akzeptablen NPV von 67,4%. Youden's J liegt hier bei 0,289 bei der Methode nach Werfel in der Validierung. Niedrigere Schwellenwerte (z.B. 40%-Schwellenwert) zeigen ebenfalls bereits eine hohe Spezifität mit 86,5%, gleichzeitig liegt die Sensitivität jedoch auch bereits unter 50%. Die positive Likelihood-Ratio ist hier 3,6-fach und der Youden's J-Index liegt höher bei 0,357. Je nach Anwendungsgebiet und Zielsetzung eignen sich also unterschiedliche Schwellenwerte.

Das Modell ohne individuelle Variablen Gewichtung konnte seine ohnehin hohe Modelleleistung bei einem 75%-Schwellenwert in der Validierung ebenfalls steigern und erreichte >94% Spezifität, 77,3% im PPV sowie 64,6% im NPV (siehe Tabelle 3.16.).

Schwellenwert	Mod.+/Ep.+	Mod.+/Ep.-	Mod.-/Ep.+	Mod.-/Ep.-	Sens.	Spez.	PPV	NPV
3	34	10	92	168	27,0	94,4	77,3	64,6
2	80	58	46	120	63,5	67,4	58,0	72,3
1	117	128	9	50	92,9	28,1	47,8	84,7

Tabelle 3.16.: Ergebnisse der Vier-Felder Tafel in Abhängigkeit des Schwellenwertes des Modells ohne individuelle Variablen Gewichtung (Referenz) und zugehörige Modelleleistungen in der Validierung („Mod. +“ = Endpunkt laut Modell positiv; „Ep. +“ = Endpunkt im Entwicklungsdatensatz tatsächlich positiv, „Mod. -“, „Ep. -“ = analog)

4.11. Modell mit gerundeten Variablen gewichtungen

Zur besseren Implementation und Erhöhung der Praktikabilität des Scores wurde eine weitere Version mit gerundeten Variablen gewichtungen berechnet (basierend auf Werfel-Gewichtung). Die in Tabelle 3.17. dargestellte ganzzahlige Rundung lieferte die besten Ergebnisse.

Variablen	Variablen gewichtung (gerundet)
Medikamentöse Immunsuppression	2
Hämatologische Malignome	5
Vergangene Hospitalisierung	1
Elektive OP geplant	2
Maximalpunktzahl: Σ Einzelgewichtungen	10

Tabelle 3.17.: Übersicht über die finalen Variablen und ganzzahlig gerundete Variablen gewichtungen (nach Werfel-Gewichtung)

Die Validierung des ganzzahlig gerundeten Modells ergibt folgende Modelleistung abhängig vom Schwellenwert (siehe Tabelle 3.18. sowie Abb. 3.3):

Schwellenwert	Mod.+/Ep.+	Mod.+/Ep.-	Mod.-/Ep.+	Mod.-/Ep.-	Sens.	Spez.	PPV	NPV
9	0	0	126	178	0,0	100,0	-	58,6
8	34	8	92	170	27,0	95,5	81,0	64,9
7	47	15	79	163	37,3	91,6	75,8	67,4
6	52	15	74	163	41,3	91,6	77,6	68,8
5	60	21	66	157	47,6	88,2	74,1	70,4
4	62	24	64	154	49,2	86,5	72,1	70,6
3	88	62	38	116	69,8	65,2	58,7	75,3
2	116	118	10	60	92,1	33,7	49,6	85,7

Tabelle 3.18.: Ergebnisse der Vier-Felder Tafel in Abhängigkeit des Schwellenwertes des Modells (Werfel, ganzzahlig gerundet) und zugehörige Modelleistungen in der Validierung („Mod. +“ = Endpunkt laut Modell positiv; „Ep. +“ = Endpunkt im Entwicklungsdatensatz tatsächlich positiv, „Mod. -“, „Ep. -“ = analog)

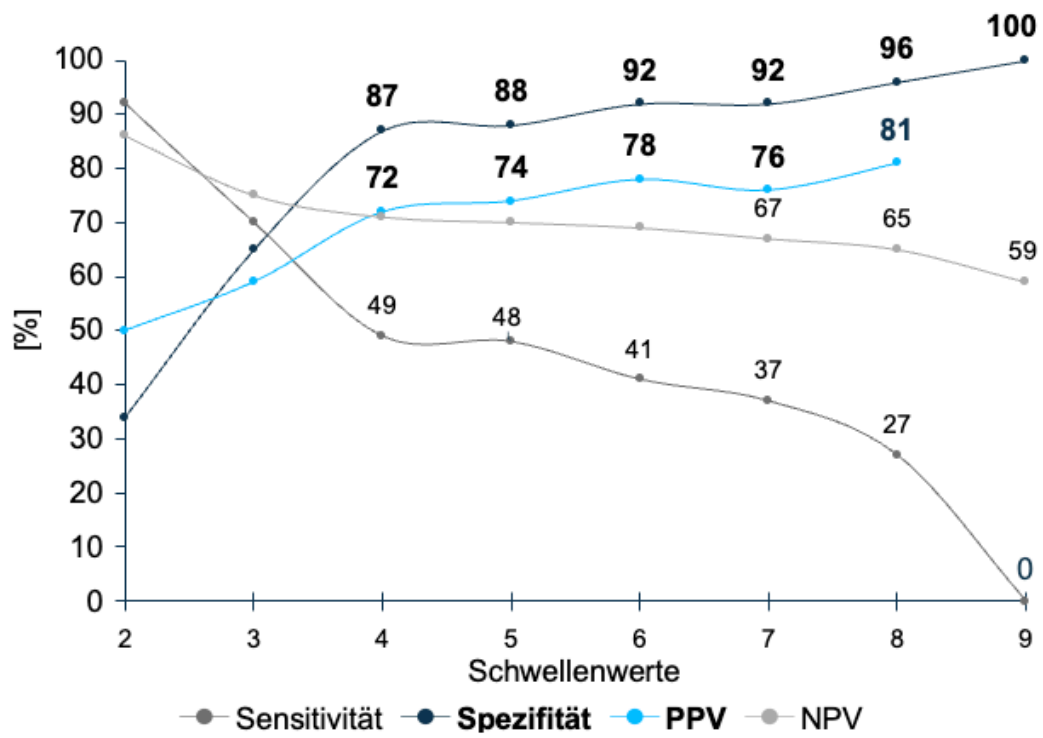


Abbildung 3.3.: Visualisierung der Modelleleistungsparameter in Abhängigkeit der Schwellenwerte

5. Diskussion

Das Ergebnis dieser Arbeit ist ein additiver Score zur Abschätzung des Risikos einer zukünftigen Antibiotikatherapie bei hospitalisierten PatientInnen. In der finalen Version des Scores sind vier Variablen inkludiert: Hämatologische Malignome, medikamentöse Immunsuppression (aktuell oder vergangene vier Wochen), Hospitalisierung (innerhalb der vergangenen sechs Monate) und geplante elektive OP. Insgesamt ergibt sich eine Maximalpunktzahl von 10 Punkten. Ein Schwellenwert von 60% der Maximalpunktzahl im Score erreicht in der Validierung einen PPV von ~78%, einen NPV von ~69% sowie ~92% an Spezifität und ~41% an Sensitivität. Ein Patient oder eine PatientIn mit ≥ 6 Punkten im entwickelten Score erhält also mit einem nahezu 80-prozentigen Risiko eine zukünftige Antibiotikatherapie und der Score bietet hiermit die Möglichkeit, PatientInnen unter hohem Risiko zu Präventions- oder Studienzwecken zu identifizieren.

5.1. Diskussion der identifizierten Risikofaktoren

5.1.1. Finale Variablen des PILGRIM-Score

Hämatologische Malignome zeigten sich als hochsignifikante Risikofaktoren mit der gleichzeitig höchsten Effektstärke unter allen untersuchten Variablen (vgl. Tabellen 3.5. und 3.6.). (Akute) Leukämien und Lymphome führen zu einer Alteration des angeborenen sowie adaptiven Immunsystems. Hierzu zählen qualitative sowie quantitative Defizite der neutrophilen Granulozyten oder Lymphozyten. Diese wiederum sorgen für eine Hypogammaglobinämie und reduzierte zell-vermittelter Immunantwort. Klassische Chemotherapeutika sowie neuere, gezielte Immuntherapeutika können diese Effekte zusätzlich verstärken. Patienten mit (akuten) Leukämien haben also ein höheres Risiko für lebensbedrohliche Infektionen (inklusive MRE) und somit antibiotische Behandlungen sowie deren Nebenwirkungen. Hinzu kommen insgesamt erhöhte Behandlungskosten^{43,44}. Dem entgegen steht die Analyse von Steinberg et al., wo Malignome statistisch eher protektiv wirkten (OR 0,73 bei 95%-KI: 0,52-1,04)³².

Immunsupprimierende Medikamente führen ebenso zu einer Alteration des Immunsystems und prädisponieren dadurch für Infektionen. In der Kohortenstudie von van Delden et al. zu Empfängern solider Organtransplantationen erlitten 55% der beobachteten PatientInnen in einem einjährigen Follow-up eine klinisch relevante Infektion. 63% dieser Infektionen waren bakteriell, ca. 30% viral und ca. 7% durch Pilze bedingt⁴⁵. Auch für PatientInnen mit medikamentöser Immunsuppression bei chronisch-entzündlichen Darmerkrankungen oder rheumatoider Arthritis ergibt sich ein ähnliches Bild und die Inzidenz für klinisch relevante Infektionen ist erhöht. Je nach Status der Erkrankung und je nach immunsupprimierendem

Agens, unterscheidet sich jedoch die Rate an klinisch relevanten Infektionen und die Prädisposition für bestimmte opportunistische Infektionen⁴⁶. Dass sich eine medikamentöse Immunsuppression als Risikofaktor für spätere Antibiotikatherapien zeigte, wird insgesamt hierdurch unterstrichen da davon auszugehen ist, dass sich die erhöhten Infektionsraten auch in erhöhten Raten antibiotischer Behandlungen widerspiegeln. Publikationen, die explizit dies untersuchen, konnten jedoch nicht identifiziert werden.

Es ist wissenschaftlicher Konsens, dass Krankenhausaufenthalte für den Erwerb von MRE prädisponieren und die Länge des Aufenthaltes mit dem Risiko von nosokomialen Infektionen korreliert. Einzelne Studien beschäftigen sich mit dem Zusammenhang zwischen der Epidemiologie einzelner Krankheitsbilder und dem Risikofaktor der vorherigen Hospitalisation. So konnten Halavaara et al. zeigen, dass das Risiko einer ambulant erworbenen infektiösen Endokarditis beispielsweise erhöht ist, wenn man in der Vergangenheit bereits einmal hospitalisiert war (6-monatiger Zeitraum), dialysiert wird oder sich urologischen Prozeduren unterzogen hat⁴⁷. Bei der Mortalitätsanalyse von PatientInnen mit gram-negativen Blutstrominfektionen von Zeng et al. zeigt sich in der multivariablen Cox-Regressionsanalyse eine vorangegangene Hospitalisation (90-tägiger Zeitraum) mit einer Hazard Ratio von 2.1 (95%-KI: 1,1-3,9, p=0,02) als starken Risikofaktor und beispielsweise höher korreliert als der APACHE II Score (Hazard Ratio 1,1 bei 95%-KI: 1,04-1,1, p<0,001)⁴⁸. Insgesamt konnte jedoch kaum Literatur identifiziert werden, die sich mit dem Zusammenhang von vorangegangenen Hospitalisierungen und dem generellen Risiko für spätere Antibiotikaverschreibungen oder Infektionen beschäftigt.

Eine geplante elektive Operation zeigte sich in der Risikofaktorenanalyse als hochsignifikant protektiv. PatientInnen, welche sich einer elektiven Operation unterzogen und wohlmöglich eine perioperative Antibiotikaphylaxe erhalten haben, waren also unter einem signifikant niedrigerem Risiko einer späteren Antibiotikatherapie. Die vorliegenden Ergebnisse könnten also die grundsätzliche Wirksamkeit von Antibiotikaphylaxen unterstreichen, da elektive Operationen mit oder ohne Antibiotikaphylaxe im Folgenden keine weitere Antibiotikagabe benötigten. Die Evidenz zur perioperativen Antibiotikaphylaxe ist hingegen uneindeutig und die Nutzen-Risiko-Analyse bleibt umstritten^{49,50}. Für aseptische Eingriffe wie beispielsweise Mamma- oder Hernienchirurgie konnten in Meta-Analysen zwar eine Reduktion von *Surgical Site Infections* durch perioperative Antibiotikaphylaxen gezeigt werden, dieser Effekt konnte in doppelblinden, randomisierten Kontrollstudien jedoch nicht mehr nachgewiesen werden⁵⁰. Es gibt zahlreiche Publikationen, die das Nutzen-Risikoverhältnis der perioperativen Antibiotikaphylaxe für einzelne Operationen analysieren, große und übergreifende Meta-Analysen fehlen jedoch.

5.1.2. Weitere Risikofaktoren

Die Risikofaktorenanalyse mit univariaten Regressionsmodellen zeigte, dass eine vorausgegangene oder bestehende Infektion / Kolonisation mit MRE ebenfalls risikobehaftet hinsichtlich des Endpunktes sein kann (vgl. Tabelle 3.6., Variablen v_278/478/277/477, mdro_colon/infect). Ein möglicher Zusammenhang wird ebenfalls in der Literatur beschrieben: Eine Publikation von Prado et al. zeigt bei PatientInnen auf Intensivstationen ein signifikant erhöhtes Risiko für bakterielle Infektionen (nicht MRE), wenn bei Aufnahme bereits eine rektale Besiedlung mit MRE vorlag (63,5% versus 48,9%; $p = 0,01$). Das Risiko für eine Infektion mit den besiedelnden MRE selbst ist ebenfalls statistisch signifikant erhöht (36,5% versus 5,9%; $p < 0,0001$)⁵¹. Obwohl der Endpunkt von Prado et al. nicht mit dem für diese Analyse Gewählten identisch ist, ist davon auszugehen, dass eine bestätigte bakterielle Infektion dann konsekutiv ebenfalls eine antibiotische Behandlung nach sich ziehen würde. Gleichzeitig erhöht der Einsatz von Antibiotika das Risiko mikrobiellen Selektion und konsekutiv der MRE-Besiedlung, worauf dann wieder eine antibiotische Behandlung folgen kann⁵².

Ein ähnliches Bild zeigt sich, wenn man nach Literatur für die Assoziation von vorherigen Antibiotikatherapien und dem generellen Auftreten neuer Infektionen oder Antibiotikatherapien sucht. Einzelne Studien untersuchen den Einfluss vorheriger Antibiotikatherapien auf das Auftreten spezieller Erkrankungen und deren Outcome. Eine vorherige Exposition von Antibiotika erhöht so beispielsweise das Mortalitätsrisiko einer ventilatorassoziierten Pneumonie (OR 9,2 bei $p < 0,001$)⁵³. Ruhe et al. konnten β -Lactam- und Makrolidantibiotika (6-monatiger Zeitraum) als unabhängige Risikofaktoren für eine *Streptococcus pneumoniae* Bakteriämie identifizieren und Llor et al. fanden in ihrer Studie mit >280.000 PatientInnen heraus, dass vorherige Antibiotikaeinnahme die schwere einer COVID-19 Infektion signifikant erhöhen kann^{54,55}.

Generell würde man hohes Alter als wesentlichen Risikofaktor für zukünftige Infektionen und somit Antibiotikatherapien erwarten, da Altern auch eine Veränderung des Immunsystems nach sich zieht. Insgesamt verringert sich die Anzahl an Stammzellen, die phagozytische Kapazität von Neutrophilen, natürlichen Killerzellen und Makrophagen nimmt ab, die B-Zell vermittelte Antikörperreaktion schwächt ab und die Produktion von T-Lymphozyten verändert sich⁵⁶. Ein hohes Alter als Risikofaktor konnte jedoch nicht gezeigt werden, stattdessen waren tendenziell eher jüngere PatientInnen unter erhöhtem Risiko (vgl. Tabellen 3.5. und 3.6.). Eine ähnliche Tendenz wurde in den Publikationen von Steinberg und Depew beobachtet^{25,32}. Nachdem in der Untersuchung von Depew et al. ein Alter zwischen 18 und 34 Jahren als Referenz gesetzt wurde, zeigten sich höhere Alter zunehmend protektiv (bspw. Alterskategorie

65+: hochsignifikant ($p < 0,005$) und protektiv mit adjustierte OR: 0,66 mit 95%-KI: 0,57 – 0,77). Gleiches gilt für die Publikation von Steinberg: PatientInnen ≥ 65 Jahren erhielten signifikant weniger Antibiotika (adjustierte OR: 0,59 mit 95%-KI: 0,46-0,77) gegenüber dem Referenzalter (18-24). Keine der beiden Publikationen konnte jedoch eine Erklärung für das beobachtete Phänomen liefern. Guo et al. geben zu bedenken, dass möglicherweise gerade junge PatientInnen mit einem beschleunigten Genesungswunsch (und oberen Atemwegsinfektionen) sowie einer gewissen Erwartungshaltung gegenüber der Verschreibung eines Antibiotikums eine spezielle Risikogruppe darstellen²⁷. Eine weitere Erklärung könnte sein, dass verschreibende ÄrztInnen bei jüngeren PatientInnen tendenziell eher großzügig sind, da sie weniger Nebenwirkungen erwarten und gleichzeitig die Konsequenzen einer verkannten bakteriellen Infektion eher fürchten als bei multimorbiden, älteren PatientInnen. Der große Anteil hämato-onkologischer PatientInnen und das tendenziell eher niedrigere Alter dieser PatientInnen, könnte ebenfalls ein Grund für den statistischen Risikoeffekt des jungen Alters sein. Im Gegensatz dazu finden sich im Zusammenhang mit der stationären Behandlung von Covid-19-PatientInnen zahlreiche Publikationen, in welchen ein zunehmendes Alter signifikant mit höheren Antibiotika-Verschreibungsraten assoziiert ist^{57,58}.

Medikamente, die den gastralen pH anheben (PPI, H₂-Blocker), zeigten sich ebenfalls als signifikante Risikofaktoren (Tabelle 3.6.). Ein möglicher Zusammenhang zwischen der dauerhaften Einnahme von antazider Medikation und einem erhöhten Risiko für (bakterielle) Infektionen wird regelmäßig in der Literatur diskutiert. Eine Publikation der *American Gastroenterological Association* diskutiert die vorhandene Evidenz zu möglichen Zusammenhängen⁵⁹. Insgesamt steigt das Risiko einer bakteriellen intestinalen Überwachsung, einer Salmonellen- oder Campylobacter-Infektion sowie der spontan-bakteriellen Peritonitis mit einem rund 0,5-6-fachen Anstieg des relativen Risikos. Das Risiko einer *Clostridoides difficile*-Infektion bei PPI-Einnahme scheint moderat im Vergleich zu dem Risiko bei Antibiotika-Exposition, jedoch zeigen Studienergebnisse, dass das Risiko bis zu 50% höher sein könnte als in Kontrollgruppen. Dies gilt vor allem für spezielle Risikogruppen wie beispielsweise Kinder^{59,60}. Ein Zusammenhang zwischen PPI-Einnahme (prophylaktisch oder therapeutisch) und der Entwicklung nicht-gastroenterologischer Infektionen (z.B. Pneumonie) konnte bislang nicht belegt werden⁵⁹. Die *American Gastroenterological Association* gibt jedoch zu bedenken, dass PPI wohlmöglich vermehrt in akuten Hospitalisationsphasen verschrieben werden und somit nicht etwa die PPI-Einnahme per se eine Infektion nach sich zieht, sondern dass es sich schlichtweg um einen statistischen Effekt bei zeitgleichem Eintreten handelt⁵⁹. Gleichzeitig könnte ein Grund für den statistischen Risikoeffekt der PPI/H₂-Blocker sein, dass diese primär multimorbiden PatientInnen mit Polypharmazie verschrieben werden, welche ohnehin infektanfälliger sind. Dass

Multimorbidität in der Analyse dieser Studie nicht mit der Verschreibung von Antibiotika assoziiert war, spräche wiederum dagegen (s. Abschnitt unten).

Ein weiterer Grund für den Risikoeffekt der PPI/H₂-Blocker könnte die behandelnde Abteilung sein. Zunächst könnte man vermuten, dass die höchsten Verschreibungsraten für diese in der Gastroenterologie vorliegen. Hier erhielten jedoch nur rund 15% der PatientInnen eine Magensäure-reduzierende Therapie, obwohl fast alle PatientInnen (ca. 90% bei n=27) Antibiotika erhielten. Aufschlussreicher ist das Verschreibungsmuster in hämato-onkologischen Abteilungen. Hier finden sich hohe Raten sowohl für PPI/H₂-Blocker (40%) als auch antibiotische Behandlungen (63%), was ein Teil der Erklärung sein könnte. Auch in der retrospektiven Beobachtungsstudie von Luo et al. konnte gezeigt werden, dass PPI-Verschreibungsraten in gastroenterologischen Abteilungen niedriger waren als in anderen Abteilungen (bspw. chirurgische) ⁶¹.

Die univariaten Analysen legten eine Korrelation einer bekannten Antibiotikaallergie mit dem Endpunkt nahe, jedoch ohne Signifikanz unter dem $p < 0,05$ Niveau (vgl. Tabelle 3.6.). In der Literatur wird die anamnestische, jedoch nicht allergologisch bestätigte Antibiotikaallergie (Penicillinallergie) als Risikofaktor für spätere Wundinfektionen nach einer Operation oder beispielsweise *Clostridoides difficile*-Infektion beschrieben. Statt β -Lactam-Antibiotika inklusive Cephalosporinen werden bei vermeintlichen Allergikern signifikant mehr ($p < 0,001$) Vertreter der Lincosamide, Aminoglykosiden oder Glycopeptid-Antibiotika verabreicht ^{62,63}. Während jedoch eine Penicillinallergie beispielsweise von ca. 10% der US-amerikanischen Bevölkerung berichtet wird, ist die klinisch relevante IgE- oder T-lymphozytär-vermittelte Hypersensitivitätsreaktion in weniger als 5% der Fälle zu finden ⁶⁴. Dies führt auch zur Empfehlung von de With et al., eine anamnestische Penicillinallergie überprüfen zu lassen, um oben genannte Auswirkungen zu verhindern ⁷.

In dieser Analyse (vgl. Tabellen 3.5. und 3.6.) zeigten sich sowohl eine behandlungsbedürftige Adipositas als auch Untergewicht als eher risikobehaftet. Hicks et al. konnten in ihrer Analyse ebenfalls zeigen, dass Übergewicht einen Risikofaktor darstellen kann: In Counties mit einer größeren Zahl übergewichtiger Einwohner (oberhalb des Medians), werden auch signifikant mehr Antibiotika verschrieben (OR 2,1 mit 95%-KI: 1,8-2,4) ²⁸. Dobner et al. beschreiben in ihrer Übersichtsarbeit, dass Untergewicht (besonders in Entwicklungsländern) vor allem für Kinder und Jugendliche ein signifikanter Risikofaktor sein kann. Gleiches gilt für die Übergewichtigkeit in Industrieländern ⁶⁵. Für Erwachsene wird von manchen Studien eine U-förmige Risikokurve postuliert, wonach ebenfalls sowohl für Unter- als auch Übergewicht höhere Risiken vorliegen ⁶⁶. Zu beachten ist für die vorliegende Analyse jedoch, dass die

betrachtete Studienpopulation insgesamt mit einem durchschnittlichen BMI $>27 \text{ kg/m}^2$ eher übergewichtig war und dies möglicherweise die Aussagekraft der Variable beeinflusste (vgl. Tabellen 3.1. und 3.3.).

Multimorbidität im Sinne des CCI mit > 5 Punkten war ebenfalls nicht positiv mit dem Endpunkt korreliert (vgl. Tabellen 3.5., 3.6. sowie 3.1., 3.3.). Der CCI geht auf eine Publikation von 1987 zurück und dient in verschiedenen Modifikationen als valides Messinstrument einer Multimorbiditäts- und Lebenserwartungsabschätzung. Die Publikation von Drosdowsky et al. legt nahe, den Score nicht uneingeschränkt einzusetzen und Aussagen mit Vorsicht zu treffen⁶⁷. In Studien zur Prädiktion des Outcomes bei Infektionen konnte sich der CCI nur teilweise als valides prognostisches Mittel erweisen. So war der CCI beispielsweise in der Publikation von Kuswardhani et al. signifikant assoziiert mit der Mortalität von COVID-19 PatientInnen, wo hingegen Wetter et al. in ihrer Studie zu schweren Atemwegsinfektionen den CCI nicht als prognostischen Marker etablieren konnten^{68,69}. Publikationen, welche den CCI zur Prädiktion von Infektionen oder Antibiotikatherapien verwenden, konnten nicht identifiziert werden. Gleichzeitig zeigten viele chronisch degenerative Erkrankungen in der Analyse keinen signifikanten Zusammenhang mit dem Endpunkt (vgl. Tabellen 3.5. und 3.6.). Diese Beobachtung würde zur Hypothese passen, dass hohes Alter und Multimorbidität zwar zu einer höheren Anfälligkeit für Infektionen führen, dies jedoch aus oben genannten Gründen nicht zwingend zu höheren Verschreibungsraten bei Antibiotika führt. Möglicherweise wird auch die Reduktion des Allgemeinzustandes, welche im klinischen Alltag oft als Indikator einer Infektion gesehen wird, bei multimorbiden PatientInnen vermehrt als Ausdruck der Multimorbidität per se interpretiert.

5.2. Diskussion der logistischen Regressionsmodelle und des PILGRIM-Score

5.2.1. Variablensortierung zum Variableneinschluss

Vergleicht man die Reihenfolge der Variablensortierung nach Signifikanz im Pearson Chi-Quadrat-Test mit der nach OR in den univariaten Regressionsmodellen, so fallen Unterschiede auf (vgl. Tabellen 3.5. und 3.6.). Bei der Entwicklung von Prädiktionsmodellen findet bei der Inklusion von Risikofaktoren meist eine Orientierung an der Signifikanz statt⁷⁰⁻⁷³. In einer Weiterentwicklung des vorliegenden Scores wäre als alternative Herangehensweise interessant, die Variableninklusion nach der Höhe der OR zu sortieren und diese Modelleleistung mit der vorliegenden zu vergleichen. Ein derartiger Ansatz wird ebenfalls in der Handreichung von Steyerberg zur Entwicklung von Prädiktionsmodellen beschrieben³⁷. Das

Modell würde also aus einer Auswahl von hoch-Risiko Variablen bestehen, die jedoch nicht zwingend signifikant mit dem Endpunkt korreliert sind.

5.2.2. Individuelle Variablengewichtung

Wenn man sich an anderen, weit verbreiteten Scores in der klinischen Anwendung orientiert, so findet dort teilweise eine individuelle Variablengewichtung statt^{71,74}. So wurden in einer Literaturrecherche ebenfalls Methoden zur Berechnung einer individuellen Variablengewichtung gesucht. Vergleicht man jedoch jene berechneten Modelle mit individueller Gewichtung mit jenen ohne individuelle Gewichtung, so fallen bei den Modelleistungen keine großen Unterschiede auf. Trotz des unterschiedlichen Einflusses der Variablen innerhalb der logistischen Regressionsmodelle, bringt eine Reflektion dessen in einer individualisierten Gewichtung offenbar keinen großen Vorteil. Den etablierte Wells-Score zur Abschätzung der Wahrscheinlichkeit einer tiefen Beinvenenthrombose bzw. einer Lungenembolie gibt es ebenfalls in einer originalen Version mit individualisierten Variablengewichtungen sowie in einer vereinfachten Version mit gleichen Gewichtungen für alle Faktoren⁷⁵. Eine Metaanalyse mit rund 7300 inkludierten PatientInnen verdeutlichte auch hier, dass der simplifizierte Wells-Score der Originalversion in der prädiktiven Leistung in keiner Weise nachsteht: 73% (Kontingenzkoeffizient) bei 95%-KI: 0,72-0,75 für den gewichteten Score und 72% bei 95%-KI 0,7-0,73 für die ungewichtete Version. Die Autoren empfehlen aufgrund der vereinfachten Handhabung die ungewichtete Version⁷⁶.

5.2.3. Multikollinearität

Die Analyse verdeutlicht, dass die finalen Variablen signifikant miteinander korrelieren (vgl. Abschnitt Überprüfung auf Multikollinearität). Das hauptsächliche Problem einer Multikollinearität zwischen den Prädiktorvariablen ist, dass instabile und verzerrte Standardfehler entstehen können. Diese führen zu instabilen p-Werten, was darin resultiert, dass die statistische Signifikanz von Assoziationen nicht realistisch und haltbar interpretiert werden kann⁷⁷. Durch den sich überschneidenden Effekt der finalen Variablenauswahl, kann nicht zuverlässig bestimmt werden, welche Variablen die wichtigsten unabhängigen Prädiktoren darstellen. Vatcheva et al. betonen in ihrer Publikation jedoch, dass Multikollinearität keinen Einfluss auf die Anpassungsgüte oder die Prädiktionsleistung des Modells hat. Dennoch kann die zuverlässige Interpretation der Regressionskoeffizienten nahezu unmöglich sein. Diese beschreiben nämlich, wie sich die Ergebnisvariable verändert, wenn sich die Prädiktorvariable um eine Einheit vergrößert und alle anderen Variablen konstant bleiben. Dies ist jedoch bei einem hohen Maß an Multikollinearität innerhalb der Variablen nicht möglich⁷⁸. Als Beispiel für den Effekt von Multikollinearität geben die Autoren die Analyse des Einflusses der Prädiktoren BMI und Hüftumfang auf die Ergebnisvariable

Blutdruck (systolisch und diastolisch). Wenn BMI und Hüftumfang einzeln hinsichtlich ihrer Korrelation untersucht werden, zeigen sich signifikante Korrelationen mit Korrelationskoeffizienten von 0,4 bzw. 0,13. Wenn beide Variablen jedoch gleichzeitig in das Regressionsmodell inkludiert werden, ist nur noch die Korrelation des BMI statistisch signifikant, während der Hüftumfang dies nun nicht mehr ist und der Korrelationskoeffizient nun sogar negativ ist (-0,08). Als Strategien zum Umgang mit Multikollinearität werden zum einen die Erhöhung der Stichprobengröße oder die Sammlung weiterer Variablen genannt. Eine weitere Strategie könnte sein, die kollinearen Variablen als Funktionen voneinander zu beschreiben (vorausgesetzt man hat den Zusammenhang vorher quantifiziert) und dann als neue Variable zu berechnen ⁷⁸. Der Grund für den plötzlich negativen Korrelationskoeffizient der Variable *keine elektive OP geplant* in der Gewichtungsberechnung mag ebenfalls in der Multikollinearität der Variablen liegen.

5.2.4. Methodik zur Score-Entwicklung

Der für diesen Score gewählte halbautomatisierte Ansatz der Analyse inklusive 10-facher Kreuzvalidierung versuchte ein weitestgehend standardisiertes Prozedere zu etablieren, um möglichst valide Ergebnisse zu erzielen. Die Möglichkeiten der Automatisierung von Prozessen und der Integration von Zwischenergebnissen für die Weiterverwendung von Ergebnissen sind mit SPSS begrenzt und es stellt sich die Frage, ob ein höherer Grad der Automatisierung bessere Modelleleistungen erbracht hätte. Der von Werfel et al. bei rund 3,3-fach größerer Kohorte (4202 inkludierte Fälle) gewählte Ansatz mit *Machine-Learning (Random forest)* und zudem 20-fold cross validation unterscheidet sich in der automatisierten Umsetzung beispielsweise deutlich vom vorliegenden gewählten Ansatz, ohne jedoch wesentlich bessere Modelleleistungen zu erzielen ⁴⁰. Hier wurden 73% Sensitivität, 72% Spezifität, 31% PPV sowie 94% NPV erreicht. Die Metaanalyse von Christodoulou et al. zum Vergleich von Machine-Learning Ansätzen mit logistischen Regressionsmodellen bei der Entwicklung von klinischen Prädiktionsmodellen zeigt ebenfalls keine signifikanten Unterschiede in der Diskriminationsleistung (AUC) von Modellen nach der Berechnung mit Machine-Learning vs. Regressionsmodellen (Differenz 0,0, 95%-KI: -0,18-0,18). Lediglich bei Studien mit einem hohen Risiko eines Bias zeigten sich Machine-Learning Methoden als vielversprechender (AUC +0,34, 95%-KI: 0,2-0,47) ⁷⁹. Hinsichtlich der zur Verfügung stehenden Datenmenge von 1258 inkludierten Fällen nach Abzug nicht auswertbarer Fälle und 43 verwendeten Prädiktoren landet der PILGRIM-Score nach der Analyse von Christodoulou im Mittelfeld: mittlere Größe des Datensatzes (1250, bei Spannweite 72-3994872) und mittlere Anzahl zur Verfügung stehender Prädiktoren (19, bei Spannweite 5-563) ⁷⁹.

Azimi et al. entwickelten in ihrer Publikation Scores zur Prädiktion von postoperativen Infektionen bei kolorektaler Chirurgie und verglichen hierbei verschiedene Techniken miteinander: *Naïve Bayes*, logistische Regression, *Decision tree*, *Support vector machine*, künstliche neuronale Netzwerke und ebenfalls *Random Forest* ⁸⁰. Auch hier wurden Sensitivität, Spezifität, PPV und Genauigkeit (*Accuracy*, ACC) der jeweiligen Methoden miteinander verglichen. Wenn man die Methodik des PILGRIM-Scores einordnen wollen würde, könnte man es als eine modifizierte Version logistischer Regressionsmodelle interpretieren. Festzuhalten ist, dass bei Azimi et al. keine Methode den anderen Methoden gegenüber in allen Kennzahlen überlegen war. Im Durchschnitt lieferte die *Support vector machine* die besten Leistungen, dicht gefolgt von den logistischen Regressionsmodellen. Im Bereich Spezifität liegen logistische Regressionsmodelle auf Platz 2, bei der Sensitivität jedoch nur im unteren Mittelfeld. Der PPV war über alle Techniken hinweg wenig aussagekräftig. Limitierend bei der Publikation von Azimi ist die niedrige Fallzahl (n=206 mit einer Endpunktverteilung von 1:10) ⁸⁰. Leider wurden auch die Unterschiede in den Modelleistungen je nach Technik nicht in der Diskussion aufgegriffen.

5.2.5. Technik der Implementation

Für den vorliegenden Score wurde ein additives Verfahren gewählt. Der Anwendende kann also die einzelnen Variablen (inklusive der respektiven Gewichtung) zusammenzählen und erhält seine Gesamtpunktzahl, anhand welcher dann eine prädiktive Aussage gemacht werden kann. Das additive Verfahren ist am einfachsten anzuwenden und wird für einen Großteil der klinischen Scores verwendet. Michel et al. haben in ihrer Publikation untersucht, ob innerhalb eines Score-Modells (EuroSCORE) logistische Verfahren additiven Verfahren hinsichtlich der Prädiktion überlegen sind ⁸¹. Im Bereich der Genauigkeit lieferte das logistische Verfahren höhere Werte, die Diskriminationsleistung (AUC der ROC) war jedoch ähnlich. In dieser 2003 erschienenen Publikation wurde damals aufgrund der einfacheren bettseitigen Berechnung noch das additive System für die Praxis empfohlen ⁸¹. Durch die allgegenwärtige Präsenz von Computern und Smartphones wird jedoch mittlerweile der EuroSCORE (II) ausschließlich mit seiner logistischen Berechnung zur Prädiktion herangezogen ⁸².

5.2.6. Validierung der Ergebnisse

Die in dieser Analyse gewählte interne Validierung des Prädiktionsmodells wird in der Literatur teilweise kritisch gesehen. Steyerberg erklärt, dass gerade die zufällige Aufteilung des Datensatzes das Risiko einer signifikanten Dysbalance von Ergebnisvariablen und Prädiktoren bergen kann ³⁷. Es wurde daher eine Teilung des Datensatzes nach Datum des Studieneinschlusses gewählt, da dies im Gegensatz zur zufälligen Teilung eine bessere Aussage über die Generalisierbarkeit des Modells erlaubt ⁸³. Gleichzeitig wird eine 1:1 oder

2:1 Teilung des Datensatzes empfohlen, was jedoch das Risiko instabiler Modellleistungen mit sich bringen kann, da der Entwicklungsdatensatz schlichtweg zu klein ist. Auf der anderen Seite birgt ein zu kleiner Validierungsdatensatz das Risiko, keine gute Aussage über die tatsächliche Validität des Modells liefern zu können. Mit der hier gewählten Aufteilung mit 3:1 wurde versucht ein Mittelweg zu gehen eher auf stabile Modellleistungen in der Entwicklung zu achten. Steyerberg bewertet die interne Validierung insgesamt zwar als die klassische, jedoch eher ineffiziente Methode. Es konnte gezeigt werden, dass eine interne Validierung nur dann Sinn macht, wenn eine wirklich große Validierungskohorte zur Verfügung steht. In diesem Fall ist das Modell jedoch ohnehin mit einer großen Datenmenge entwickelt worden und hat möglicherweise bereits eine gute Validität. Bevor ein Modell extern validiert wird oder falls kein externer Validierungsdatensatz zur Verfügung steht, sollten neuere, aber rechenintensive interne Validierungsmethoden wie Bootstrapping verwendet werden ³⁷. Parallel wird in der Literatur die externe Validierung als besonders aussagekräftig angesehen, da sie Aussagen zum Modell hinsichtlich Generalisierbarkeit und Übertragbarkeit auf andere klinische Situationen erlauben kann. Staffa et al. empfehlen, die interne Validierung als Vorstufe der externen Validierung zu sehen ⁸³. Wenn ein Modell in einer zeitlichen, nicht-zufälligen, internen Validierung bereits wesentlich schlechtere Ergebnisse als in der Entwicklung zeigt, sollte das Modell überdacht werden. Falls die interne Validierung gute Ergebnisse liefert, ist der nächste Schritt, Möglichkeiten der externen Validierung zu erörtern.

5.3. Limitationen

Es wurden rund 1250 Fälle in der Analyse untersucht. Betrachtet man vergleichbare Score-Publikationen gibt es zwar auch Abhandlungen mit niedrigeren Fallzahlen, meistens standen zur Entwicklung jedoch mehr Fälle zur Verfügung. Außerdem fällt auf, dass gerade zu Beginn der Rekrutierungsphase (Fälle des Entwicklungsdatensatzes) überproportional viele PatientInnen aus der Hämatologie/Onkologie sowie Kardiochirurgie inkludiert wurden. Ein Grund hierfür mag in einer zu Beginn starken Rekrutierungsleistung einzelner Studienzentren liegen, welche dann vornehmlich in den oben genannten Fachrichtungen rekrutierten. Insgesamt spiegelt dies keinesfalls einen realen Fallmix wider und könnte für eine Verzerrung in der Analyse potenzieller Risikofaktoren geführt haben. Ähnliches zeigte sich bei der Analyse des Validierungsdatensatzes, der Anteil der beiden oben genannten Fachrichtungen war hier jedoch etwas geringer. Dies mag die Validität des Scores insgesamt erhöhen, da die Validierung auf einem in dieser Hinsicht ausgewogenerem PatientInnenkollektiv stattfand.

Ein Inklusionskriterium der PILGRIM-Studie war außerdem das erwartet hohe Risiko für eine Antibiotikatherapie im Verlauf des stationären Aufenthaltes (innerhalb der folgenden 10 Tage) für mindestens 5 Tage. In dem zur Analyse herangezogenen Datenexport erhielten jedoch nur

ca. 52% der PatientInnen des Entwicklungsdatensatzes ein Antibiotikum (42% im Validierungsdatensatz). Hinsichtlich des oben genannten allgemein formulierten Rekrutierungsziels muss retrospektiv festgesellt werden, dass das erwartete Ereignis (Antibiotikatherapie) weniger häufig eingetroffen ist, als erhofft. Diese Tatsache machte die Berechnung des vorliegenden Scores jedoch erst möglich. Grundsätzlich muss aber von einem gewissen Bias ausgegangen werden, vor allem, wenn man das Ziel des entwickelten Scores (Prädiktion zukünftiger Antibiotikatherapie) betrachtet.

Auffallend waren in der Analyse außerdem die stark negativen Korrelationen mancher (kardiologischen) Vorerkrankungen mit dem Endpunkt (vgl. Tabellen 3.5. und 3.6.). Welche Komorbiditäten der PatientInnen bei der Aufnahme in die PILGRIM-Studie erfasst und dokumentiert wurden, hing mutmaßlich mit der aktuell behandelnden medizinischen Fachrichtung und dem klinischen Fokus zusammen (Dokumentationsbias). Wie bereits beschrieben, wurden vor allem im Entwicklungsdatensatz zahlreiche PatientInnen aus der Herzchirurgie rekrutiert, bei welchen dann wiederum überproportional viele der oben genannten Begleiterkrankungen kodiert wurden. So wurden die Dyslipidämie beispielsweise zu ~ 60% in kardiologischen oder kardiochirurgischen Abteilungen kodiert und nur vereinzelt in anderen Fachgebieten. Zusammenfassend ist die stark negative und partiell hochsignifikante Korrelation mancher Variablen wohl auch auf die uneinheitliche Dokumentation bei der PatientInnenrekrutierung und der in Relation niedrigeren Verschreibungsraten in jenen Abteilungen zurückzuführen. Die genannten Faktoren führten dazu, dass Komorbidität und Multimorbidität nicht in dem Maße als Risikofaktoren identifiziert werden konnten, wie es die vorhandene Literatur vermuten ließ.

Im elektronischen Case Report Form (eCRF) wurden insgesamt weit mehr als 2000 Variablen abgefragt und erfasst. Jene Variablen, welche die Grundlage für die Entwicklung des PILGRIM-Scores bildeten, stellen nur einen Bruchteil der Gesamtzahl der Variablen dar. Der Großteil der zur Verfügung stehenden Variablen dienen der übergreifenden PILGRIM-Studie und erfragen zum Beispiel die Umstände einer Antibiotikatherapie oder Details zur abgegebenen Stuhlprobe und konnten nicht zur Entwicklung des Scores herangezogen werden. Weitere Parameter von Interesse wie zum Beispiel Befunde der klinischen Chemie oder Vitalparameter wurden im eCRF erst nach Eintreten eines klinischen Ereignisses, das heißt bei bestätigter Infektion, Fieber oder Start einer antibiotischen Behandlung erfasst. Generell gab es innerhalb der PatientInnenkohorte kaum PatientInnen, bei welchen eine Infektion oder Fieber ohne die gleichzeitige Initiierung einer antibiotischen Behandlung zu beobachten war. Dies führte dazu, dass oben genannte weitere klinische Parameter von Interesse kaum in die Analyse eingeschlossen werden konnten.

Manche Variablen waren in ihrer positiven Ausprägung nur selten zu finden, sodass die Aussagekraft durch die kleine Fallzahl zweifelhaft sein kann. Dies trifft beispielsweise für die Variablen zur Kolonisation oder vorherigen Infektion mit MRE zu (bspw. v_278 mit n=13).

Neben der patientInnenindividuellen Ebene gibt es zusätzlich noch die Ebene der oder des Behandelnden, welche einen Einfluss auf Antibiotikaverschreibungen haben kann. So können persönliche Einstellungen, Spezialisierung, Erfahrung, kontinuierliche Fortbildung, etc. einen wesentliche Auswirkungen haben. Das Ziel der vorliegenden Analyse war zwar nicht, Einflussfaktoren abseits der PatientInnenebene zu finden, dennoch kann zusammenfassend bemerkt werden, dass die schlussendliche Entscheidung für oder gegen ein Antibiotikum von vielen weiteren Faktoren unterschiedlicher Ebenen abhängig sein kann, welche nicht im PILGRIM-Score reflektiert wurden.

5.4. Zusammenfassung und Ausblick

Die abschließende Modelleistung in der Validierung zeigt ein funktionierendes Prädiktionsmodell, welches zukünftig die Rekrutierung von PatientInnen für AMS- oder Mikrobiomstudien erleichtern und somit einen wissenschaftlichen Beitrag leisten könnte. PatientInnen unter hohem Risiko könnten gefiltert werden und entsprechende präventive Gegenmaßnahmen ergriffen werden.

In einer Weiterentwicklung des Scores, wäre es interessant zusätzliche Variablen wie bspw. Labor- oder Vitalparameter in die Analyse zu integrieren. Möglicherweise könnten diese auch zur Prädiktion einer Antibiotikatherapie beitragen. Obwohl sich in der Literatur andere Methoden der Score-Berechnung mit einem höheren Grad der Automatisierung nicht zwingend überlegen zeigten, wäre es dennoch spannend mit diesen Methoden eine zweite Version des PILGRIM-Score zu entwickeln und mit der bestehenden Modelleistung zu vergleichen. Das Problem der Multikollinearität müsste in einer Weiterentwicklung ebenfalls erörtert werden. Ansätze wie die oben beschriebenen oder eine Residualisierung / Ridge Regression könnten zeigen, ob ein besserer Umgang damit gefunden werden kann.

Bislang wurde lediglich die Diskriminationsleistung des Modells betrachtet. In einem zweiten Schritt sollte unter der Einbeziehung der Prävalenz (für Antibiotikatherapien) eine Kalibrierung des Modells vorgenommen werden, damit das absolute Risiko für PatientInnen prävalenzabhängig besser eingeschätzt werden kann.

Zur Vervollständigung der internen Validierung könnten zusätzlich auch Bootstrapping-Methoden eingesetzt werden. Eine zusätzliche Validierung des PILGRIM-Score auf einem fremden Datensatz des PILGRIM-Studienzentrums in Israel (externe Validierung) ist vorgesehen. Interessant bleibt die Frage, ob eine andere Fallzusammensetzung ohne hohe Anteile aus der Hämato-Onkologie oder der Kardio-/Chirurgie die Modellleistung wesentlich negativ beeinflusst oder ob trotz der starken Gewichtung hämatologischer Neoplasien im Score ähnliche Modellleistungen mit einem externen Datensatz erreicht werden können.

Bei der Literaturrecherche zum übergreifenden Thema fällt auf, dass es kaum Studien gibt, die Ähnliches versuchen, wie die vorliegende Analyse. Wenig Studien untersuchen das generelle Risiko von Antibiotikatherapien auf PatientInnenebene. Es existieren viele einzelne Untersuchungen, die Scores oder Prädiktionsmodelle für das Entstehen oder Outcome einzelner Krankheitsbilder beschreiben. Diese Berechnungen sind jedoch spezifisch und insgesamt fehlt es an übergreifenden, grundlegenden Studien zum Thema, was die Einordnung der vorliegenden Ergebnisse zusätzlich erschwert. Größere Multicenterstudien könnten helfen, Risikofaktoren auf PatientInnenebene weiter zu quali- und quantifizieren. Unseres Kenntnisstandes nach ist dies der erste Score, der das generelle Risiko einer zukünftigen Antibiotikatherapie quantifizieren kann. Als Ausblick könnte man in einer weiterführenden Auswertung der PILGRIM-Kohorte innerhalb der PatientInnen, die zur elektiven OP gekommen sind, jene mit perioperativer Antibiotikaphylaxe mit denen ohne hinsichtlich des Endpunktes vergleichen. Dies würde ein interessantes Bild bezüglich der Wirksamkeit von perioperativen Antibiotikaphylaxen erlauben.

6. Literaturverzeichnis

1. World Health Organization (WHO). Antimicrobial Resistance. 2021. <https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance> (accessed 06-12-2023).
2. Blair JM, Webber MA, Baylay AJ, Ogbolu DO, Piddock LJ. Molecular mechanisms of antibiotic resistance. *Nat Rev Microbiol* 2015; **13**(1): 42-51.
3. World Health Organization (WHO). Global antimicrobial resistance and use surveillance system (GLASS) report 2022. Genf (CH), 2022.
4. World Health Organization (WHO). 2021 Antibacterial agents in clinical and preclinical development: an overview and analysis. Genf (CH); 2022.
5. Murray CJL, Ikuta KS, Sharara F, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet* 2022; **399**(10325): 629-55.
6. Bundesministerium für Gesundheit (BMG). Deutschen Antibiotika-Resistenzstrategie „DART 2030“: Abschlussbericht. Berlin (DE), 2023.
7. de With K, Wilke K, Kern WV, Strauß R, Kramme E. Deutsche Gesellschaft für Infektiologie e.V. (DGI): S3-Leitlinie Strategien zur Sicherung rationaler Antibiotika-Anwendung im Krankenhaus; 2018.
8. Davey P, Marwick CA, Scott CL, et al. Interventions to improve antibiotic prescribing practices for hospital inpatients. *Cochrane Database of Systematic Reviews* 2017.
9. Khodamoradi Y, Kessel J, Vehreschild JJ, Vehreschild M. The Role of Microbiota in Preventing Multidrug-Resistant Bacterial Infections. *Dtsch Arztebl Int* 2019; **116**(40): 670-6.
10. European Center for Disease Prevention and Control (ECDC). Antimicrobial consumption in the EU/EEA (ESAC-Net) - Annual Epidemiological Report for 2021. Stockholm; 2022.
11. McEwen SA, Collignon PJ. Antimicrobial Resistance: a One Health Perspective. *Microbiology Spectrum* 2018; **6**(2): 10.1128/microbiolspec.arba-0009-2017.
12. Christaki E, Marcou M, Tofarides A. Antimicrobial Resistance in Bacteria: Mechanisms, Evolution, and Persistence. *J Mol Evol* 2020; **88**(1): 26-40.
13. Harbarth S, Balkhy HH, Goossens H, et al. Antimicrobial resistance: one world, one fight! *Antimicrobial Resistance and Infection Control* 2015; **4**(1): 49.
14. European Centre for Disease Prevention and Control (ECDC) EFSA/EaEMA. Third joint inter-agency report on integrated analysis of consumption of antimicrobial agents and occurrence of antimicrobial resistance in bacteria from humans and food-producing animals in the EU/EEA, JIACRA III. 2016-2018. Stockholm, Parma, Amsterdam; 2021.
15. Isaac S, Scher JU, Djukovic A, et al. Short- and long-term effects of oral vancomycin on the human intestinal microbiota. *J Antimicrob Chemother* 2017; **72**(1): 128-36.
16. Taur Y, Xavier JB, Lipuma L, et al. Intestinal domination and the risk of bacteremia in patients undergoing allogeneic hematopoietic stem cell transplantation. *Clin Infect Dis* 2012; **55**(7): 905-14.

17. Montassier E, Al-Ghalith GA, Ward T, et al. Pretreatment gut microbiome predicts chemotherapy-related bloodstream infection. *Genome medicine* 2016; **8**(1): 49.
18. Liss BJ, Vehreschild JJ, Cornely OA, et al. Intestinal colonisation and blood stream infections due to vancomycin-resistant enterococci (VRE) and extended-spectrum beta-lactamase-producing Enterobacteriaceae (ESBLE) in patients with haematological and oncological malignancies. *Infection* 2012; **40**(6): 613-9.
19. Ballo O, Tarazzit I, Stratmann J, et al. Colonization with multidrug resistant organisms determines the clinical course of patients with acute myeloid leukemia undergoing intensive induction chemotherapy. *PLoS One* 2019; **14**(1): e0210991.
20. Taylor J, Hafner M, Yerushalmi E, et al. Estimating the economic costs of antimicrobial resistance. *Model and Results (RAND Corporation, Cambridge, UK)* 2014.
21. Lee J-H. Perspectives towards antibiotic resistance: from molecules to population. *Journal of Microbiology* 2019; **57**(3): 181-4.
22. Fernandez L, Breidenstein EB, Hancock RE. Creeping baselines and adaptive resistance to antibiotics. *Drug Resist Updat* 2011; **14**(1): 1-21.
23. Pang Z, Raudonis R, Glick BR, Lin T-J, Cheng Z. Antibiotic resistance in *Pseudomonas aeruginosa*: mechanisms and alternative therapeutic strategies. *Biotechnology Advances* 2019; **37**(1): 177-92.
24. ECDC. European Center for Disease Prevention and Control: Antimicrobial resistance in the EU/EEA (EARS-Net) - Annual Epidemiological Report 2021. Stockholm: ECDC, 2022.
25. Depew RE, Gonzales G. Differences in US antibiotic prescription use by facility and patient characteristics: evidence from the National Ambulatory Medical Care Survey. *Fam Pract* 2020; **37**(2): 180-6.
26. Ge Y, Chipenda Dansokho S, Liao X-P. Advanced neonatal medicine in China: Is newborn ward capacity associated with inpatient antibiotic usage? *PLOS ONE* 2019; **14**(8): e0219630.
27. Guo S, Sun Q, Zhao X, Shen L, Zhen X. Prevalence and risk factors for antibiotic utilization in Chinese children. *BMC Pediatrics* 2021; **21**(1).
28. Hicks LA, Bartoces MG, Roberts RM, et al. US Outpatient Antibiotic Prescribing Variation According to Geography, Patient Population, and Provider Specialty in 2011. *Clinical Infectious Diseases* 2015.
29. Labi A-K, Obeng-Nkrumah N, Nartey ET, et al. Antibiotic use in a tertiary healthcare facility in Ghana: a point prevalence survey. *Antimicrobial Resistance & Infection Control* 2018; **7**(1).
30. Pinto JM, Wagle S, Navallo LJ, Petrova A. Risk Factors and Outcomes Associated With Antibiotic Therapy in Children Hospitalized With Asthma Exacerbation. *J Pediatr Pharmacol Ther* 2022; **27**(4): 366-72.
31. Schulman J, Dimand RJ, Lee HC, Duenas GV, Bennett MV, Gould JB. Neonatal Intensive Care Unit Antibiotic Use. *PEDIATRICS* 2015; **135**(5): 826-33.
32. Steinberg MB, Akincigil A, Kim EJ, Shallis R, Delnevo CD. Tobacco Smoking as a Risk Factor for Increased Antibiotic Prescription. *Am J Prev Med* 2016; **50**(6): 692-8.

33. Steinman MA, Landefeld CS, Gonzales R. Predictors of Broad-Spectrum Antibiotic Prescribing for Acute Respiratory Tract Infections in Adult Primary Care. *JAMA* 2003; **289**(6): 719.
34. Yu KC, Moisan E, Tartof SY, et al. Benchmarking Inpatient Antimicrobial Use: A Comparison of Risk-Adjusted Observed-to-Expected Ratios. *Clinical Infectious Diseases* 2018; **67**(11): 1677-85.
35. Kern WV, Fellhauer M, Hug M, et al. Antibiotika-Anwendung 2012/13 in 109 deutschen Akutkrankenhäusern. *Dtsch Med Wochenschr* 2015; **140**(23): e237-e46.
36. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987; **40**(5): 373-83.
37. Steyerberg EW. *Clinical Prediction Models : A Practical Approach to Development, Validation, and Updating*. Cham, SWITZERLAND: Springer International Publishing AG; 2019.
38. Ellis PD. *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results*. Cambridge: Cambridge University Press; 2010.
39. Jakob CEM, Mahajan UM, Oswald M, et al. Prediction of COVID-19 deterioration in high-risk patients at diagnosis: an early warning score for advanced COVID-19 developed by machine learning. *Infection* 2021: 1-12.
40. Werfel S, Jakob CEM, Borgmann S, et al. Development and validation of a simplified risk score for the prediction of critical COVID-19 illness in newly diagnosed patients. *J Med Virol* 2021.
41. Matsushita A, Tabata M, Mihara W, et al. Risk score system for late aortic events in patients with uncomplicated type B aortic dissection. *The Journal of Thoracic and Cardiovascular Surgery* 2020; **159**(6): 2173-83.e1.
42. Schneeweiss S, Wang PS, Avorn J, Glynn RJ. Improved comorbidity adjustment for predicting mortality in Medicare populations. *Health Serv Res* 2003; **38**(4): 1103-20.
43. Safdar A, Armstrong D. Infections in Patients With Hematologic Neoplasms and Hematopoietic Stem Cell Transplantation: Neutropenia, Humoral, and Splenic Defects. *Clinical Infectious Diseases* 2011; **53**(8): 798-806.
44. Logan C, Koura D, Taplitz R. Updates in infection risk and management in acute leukemia. *Hematology Am Soc Hematol Educ Program* 2020; **2020**(1): 135-9.
45. van Delden C, Stampf S, Hirsch HH, et al. Burden and Timeline of Infectious Diseases in the First Year After Solid Organ Transplantation in the Swiss Transplant Cohort Study. *Clinical Infectious Diseases* 2020; **71**(7): e159-e69.
46. Calvet X, Carpio D, Rodríguez-Lago I, et al. Risk of infection associated with Janus Kinase (JAK) inhibitors and biological therapies in inflammatory intestinal disease and rheumatoid arthritis. Prevention strategies. *Gastroenterología y Hepatología (English Edition)* 2021; **44**(8): 587-98.
47. Halavaara M, Huotari K, Anttila VJ, Järvinen A. Healthcare-associated infective endocarditis: source of infection and burden of previous healthcare exposure. *Antimicrob Steward Healthc Epidemiol* 2023; **3**(1): e152.

48. Zeng Q, Xu F, Jiang S. The impact of previous hospitalization in the preceding 90 days on the outcome in critically ill patients with gram-negative bloodstream infection. *Diagn Microbiol Infect Dis* 2014; **80**(2): 136-40.
49. Reutter F, Reuter DA, Hilgarth H, Heilek AM, Goepfert MS, Punke MA. Perioperative Antibiotikaprophylaxe. *Der Anaesthesist* 2014; **63**(1): 73-86.
50. Leaper D, Ousey K. Evidence update on prevention of surgical site infection. *Current Opinion in Infectious Diseases* 2015; **28**(2): 158-63.
51. Prado V, Hernandez-Tejero M, Mucke MM, et al. Rectal colonization by resistant bacteria increases the risk of infection by the colonizing strain in critically ill patients with cirrhosis. *J Hepatol* 2022; **76**(5): 1079-89.
52. Dethlefsen L, Relman DA. Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proceedings of the National Academy of Sciences of the United States of America* 2011; **108** Suppl 1: 4554-61.
53. Rello J, Ausino V, Ricart M, Castella J, Prats G. Impact of Previous Antimicrobial Therapy on the Etiology and Outcome of Ventilator-associated Pneumonia. *Chest* 1993; **104**(4): 1230-5.
54. Ruhe JJ, Hasbun R. Streptococcus pneumoniae Bacteremia: Duration of Previous Antibiotic Use and Association with Penicillin Resistance. *Clinical Infectious Diseases* 2003; **36**(9): 1132-8.
55. Llor C, Ouchi D, Giner-Soriano M, García-Sangenís A, Bjerrum L, Morros R. Correlation between Previous Antibiotic Exposure and COVID-19 Severity. A Population-Based Cohort Study. *Antibiotics* 2021; **10**(11): 1364.
56. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. *The Lancet* 2013; **381**(9868): 752-62.
57. Langford BJ, So M, Raybardhan S, et al. Antibiotic prescribing in patients with COVID-19: rapid review and meta-analysis. *Clin Microbiol Infect* 2021; **27**(4): 520-31.
58. Hamada S, Tokuda Y, Honda H, et al. Prevalence and characteristics of antibiotic prescription for acute COVID-19 patients in Japan. *Sci Rep* 2022; **12**(1): 22340.
59. Freedberg DE, Kim LS, Yang Y-X. The Risks and Benefits of Long-term Use of Proton Pump Inhibitors: Expert Review and Best Practice Advice From the American Gastroenterological Association. *Gastroenterology* 2017; **152**(4): 706-15.
60. Furuya-Kanamori L, Stone JC, Clark J, et al. Comorbidities, Exposure to Medications, and the Risk of Community-Acquired Clostridium difficile Infection: A Systematic Review and Meta-analysis. *Infection Control & Hospital Epidemiology* 2015; **36**(2): 132-41.
61. Luo H, Fan Q, Xiao S, Chen K. Changes in proton pump inhibitor prescribing trend over the past decade and pharmacists' effect on prescribing practice at a tertiary hospital. *BMC Health Serv Res* 2018; **18**(1): 537.
62. Blumenthal KG, Ryan EE, Li Y, Lee H, Kuhlen JL, Shenoy ES. The Impact of a Reported Penicillin Allergy on Surgical Site Infection Risk. *Clin Infect Dis* 2018; **66**(3): 329-36.
63. Schwiebert R, Sandoe J. Is there a role of penicillin allergy in developing Clostridioides difficile infection? *Current Opinion in Gastroenterology* 2021; **37**(1): 1-3.

64. Shenoy ES, Macy E, Rowe T, Blumenthal KG. Evaluation and Management of Penicillin Allergy: A Review. *Jama* 2019; **321**(2): 188-99.
65. Dobner J, Kaser S. Body mass index and the risk of infection - from underweight to obesity. *Clin Microbiol Infect* 2018; **24**(1): 24-8.
66. Harpsøe MC, Nielsen NM, Friis-Møller N, et al. Body Mass Index and Risk of Infections Among Women in the Danish National Birth Cohort. *American Journal of Epidemiology* 2016; **183**(11): 1008-17.
67. Drosdowsky A, Gough K. The Charlson Comorbidity Index: problems with use in epidemiological research. *J Clin Epidemiol* 2022; **148**: 174-7.
68. Setter NW, Peres ML, de Almeida BMM, Petterle RR, Raboni SM. Charlson comorbidity index scores and in-hospital prognosis of patients with severe acute respiratory infections. *Intern Med J* 2020; **50**(6): 691-7.
69. Tuty Kuswardhani RA, Henrina J, Pranata R, Anthonius Lim M, Lawrensia S, Suastika K. Charlson comorbidity index and a composite of poor outcomes in COVID-19 patients: A systematic review and meta-analysis. *Diabetes Metab Syndr* 2020; **14**(6): 2103-9.
70. Lim WS, van der Eerden MM, Laing R, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax* 2003; **58**(5): 377-82.
71. Nashef SAM, Roques F, Michel P, et al. European system for cardiac operative risk evaluation (EuroSCORE). *European Journal of Cardio-Thoracic Surgery* 1999; **16**(1): 9-13.
72. Roland LT, Gurrola JG, 2nd, Loftus PA, Cheung SW, Chang JL. Smell and taste symptom-based predictive model for COVID-19 diagnosis. *Int Forum Allergy Rhinol* 2020; **10**(7): 832-8.
73. Sun Y, Heng BH, Tay SY, Seow E. Predicting hospital admissions at emergency department triage using routine administrative data. *Acad Emerg Med* 2011; **18**(8): 844-50.
74. Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of Clinical Classification Schemes for Predicting Stroke Results From the National Registry of Atrial Fibrillation. *JAMA* 2001; **285**(22): 2864-70.
75. Wells PS, Anderson DR, Rodger M, et al. Excluding Pulmonary Embolism at the Bedside without Diagnostic Imaging: Management of Patients with Suspected Pulmonary Embolism Presenting to the Emergency Department by Using a Simple Clinical Model and d-dimer. *Annals of Internal Medicine* 2001; **135**(2): 98-107.
76. van Es N, Kraaijpoel N, Klok FA, et al. The original and simplified Wells rules and age-adjusted D-dimer testing to rule out pulmonary embolism: an individual patient data meta-analysis. *Journal of Thrombosis and Haemostasis* 2017; **15**(4): 678-84.
77. Bayman EO, Dexter F. Multicollinearity in Logistic Regression Models. *Anesthesia & Analgesia* 2021; **133**(2): 362-5.
78. Vatcheva KP, Lee M, McCormick JB, Rahbar MH. Multicollinearity in Regression Analyses Conducted in Epidemiologic Studies. *Epidemiology (Sunnyvale)* 2016; **6**(2).
79. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol* 2019; **110**: 12-22.

80. Azimi K, Honaker MD, Chalil Madathil S, Khasawneh MT. Post-Operative Infection Prediction and Risk Factor Analysis in Colorectal Surgery Using Data Mining Techniques: A Pilot Study. *Surg Infect (Larchmt)* 2020; **21**(9): 784-92.
81. Michel P, Roques F, Nashef SAM, Group TEP. Logistic or additive EuroSCORE for high-risk patients? *European Journal of Cardio-Thoracic Surgery* 2003; **23**(5): 684-7.
82. Nashef SA, Roques F, Sharples LD, et al. EuroSCORE II. *Eur J Cardiothorac Surg* 2012; **41**(4): 734-44; discussion 44-5.
83. Staffa SJ, Zurakowski D. Statistical Development and Validation of Clinical Prediction Models. *Anesthesiology* 2021; **135**(3): 396-405.

7. Anhang

7.1. Literaturrecherche: Risikofaktoren für Antibiotika-Verschreibung

Autor	Journal (Jahr)	Studienpopulation	Risikofaktoren	Studiendesign, (Datenquelle)
Depew ²⁵	Fam Pract 2020	Alter >18 Jahre, 140.481 untersuchte, ambulante Konsultationen (alle Ätiologien)	Ort der Konsultation (Notaufnahme), Ethnische Herkunft (Hispanics), Versicherungsstatus (nicht versichert mit höherem Risiko), >1 Komorbidität	Kohortenstudie, Daten aus <i>National Hospital Ambulatory Medical Care Survey</i> (NHAMCS, Register für Konsultationen in Notaufnahmen) und <i>National Ambulatory Medical Care Survey</i> (NAMCS, Register für ambulante Konsultationen) des Centers für Disease Control and Prevention (CDC)
Steinmann ³³	JAMA 2003	1981 Fälle, Rekrutierung zwischen 1997-1999 (akute Infektion der oberen Atemwege ärztlichen Rat suchten (keine Pneumonie)	Konsultierte Fachrichtung (Innere Medizin im Vergleich zu General & Family Physicians, geographische Region (Nordost und Süden; im Vergleich zum Westen der USA). Versicherungsstatus (Selbstzahler im Vergleich zur privaten Versicherung), Black non-Hispanics (im Vergleich zu White-non-Hispanics) zeigen einen protektiven Effekt. Alter, Geschlecht und Ort der Konsultation (rural/urban) zeigen keinen Einfluss	Querschnittsstudie, Daten aus NAMCS
Ge ²⁶	PLOS ONE 2019	Neugeborene auf Säuglingsstationen (2008-2010)	Größe der Säuglingsstation (Stationen mit >100 Betten zeigten einen um 36% höheren Antibiotikaverbrauch als Stationen mit weniger als 50 Betten), numerische Verhältnis von ÄrztInnen zu Pflegenden, Dauer des Krankenhausaufenthaltes sowie das numerische Verhältnis von Pneumonien zu Frühgeborenen	Retrospektive Querschnittsstudie, Daten der <i>National Health and Family Planning Commission of the People's Republic of China</i> (NHFFPC)
Guo ²⁷	BMC Ped 2021	Kinder zwischen 0 und 18 Jahren	Risikofaktoren auf Ebene der Kinder: Art der Krankheit (Infektionen der oberen Atemwege); Risikofaktoren auf Ebene der Verschreibenden: mangelnde Fähigkeiten/Wissen, Erwartungshaltung des Patienten/der Patientin, ineffektive Kommunikation, ökonomische Anreize für das Verschreiben von Antibiotika, mangelnde mikrobiologische Abklärung; Risikofaktoren auf Ebene der Sorgeberechtigten: mangelnde Fähigkeiten/Wissen, Erwartungshaltung, Selbstmedikation (mit höheren oder niedrigeren Dosen als empfohlen); Risikofaktoren auf Ebene der Apotheken: Verkauf ohne Rezept; Risikofaktoren auf Ebene der Krankenhäuser: Stationsgröße	Systematische Übersichtsarbeit (Meta-Analyse), untersucht die von zwischen 2010 und 2020 publizierten Analysen berichtete Prävalenz der Antibiotika-Verschreibung (weltweit) inklusive einer qualitativen Analyse möglicher Risikofaktoren bei chinesischen Kindern.
Hicks ²⁸	Clin Infect Dis 2015	Alle Verschreibungen systemischer, oraler Antibiotika (ambulant) aus dem Jahr 2011	Höchste Verschreibungsraten bei Hausärztinnen und hausärztlich tätigen InternistInnen und Pädiatern und im Süden der USA; Übergewichtigkeit (oberhalb des Medians), weibliches Geschlecht	Retrospektive Querschnittsstudie, Daten aus IMS Health Xponent Datenbank (kommerzielle Datenbank zur Sammlung von

				Verschreibungsdaten auf dem US-Markt, ambulante PatientInnen)
Labi ²⁹	Antimicr Res Infect Contr 2018	677 PatientInnen	Sehr hohe Verschreibungsraten (90,9% bei 95%) in der Kinderchirurgie; 40% der Verschreibungen waren für ambulant erworbene Infektionen; die meisten Verschreibungen waren für Infektionen im urologischen oder geburtshilflichen Bereich	Punkt-Prävalenz-Studie, Daten aus PatientInnenakten des Korle-Bu Teaching Hospital in Accra, Ghana
Schulman ³¹	PED 2015	52061 PatientInnen auf 127 Neugeborenen-Intensivstationen in Kalifornien	Keine Korrelation zwischen Antibiotikaverbrauch und gesicherten Infektionen (z.B. nekrotisierende Enterokolitis, NEC), Operationsvolumen oder Mortalitätsrate. „Die Unterschiede bei der Verschreibung von Antibiotika hängen offenbar davon ab, wie ÄrztInnen klinische Situationen, die unbewiesene Infektionen darstellen, beschreiben, interpretieren und darauf reagieren“	Retrospektive Kohortenstudie, Daten aus California Perinatal Quality Care Collaborative (CPQCC) und California Children's Services (CCS)
Yu ³⁴	Clin Infect Dis 2018	2,7 Millionen untersuchte Aufnahmen	Diagnosis-related groups (DRG), Infektion bei Aufnahme, PatientInnengruppe und Aufnahmestation sind Prädiktoren für den Einsatz von Antibiotika	Retrospektive Kohortenstudie, Daten aus PatientInnenakten der Kaiser Permanente Southern and Northern California (KP) Versicherung
Steinberg ³²	Am J Prev Med 2016	Alter ab 18 Jahren, 8300 untersuchte ambulante Arztkontakte	Tabak-Konsum; Alter über 65 Jahren, bekanntes Asthma oder maligne Erkrankung wurden als protektiv identifiziert. Niedriger sozioökonomische oder gesetzliche Versicherung (Medicaid) auch eher protektiv.	Querschnittsstudie, Daten aus NAMCS (2006-2010)
Pinto ³⁰	J Pediatr Pharmacol Ther	Pädiatrische Normal- und Intensivstation: Alle ICD-9 codierten PatientInnen mit Asthma	Alter, weibliches Geschlecht, niedrige Sauerstoffstättigung, Sauerstoffpflichtigkeit, Aufnahme auf Intensivstation (Zusätzlich: Verlängerung des durchschnittlichen stationären Aufenthaltes durch die Verschreibung von ABX bei PatientInnen ohne bakterielle Infektion oder Fieber)	Retrospektive Querschnittsstudie, Daten aus PatientInnenakten eines Kinderkrankenhauses in New-Jersey (2012-2015)

Tabelle 6.1.: Literaturrecherche zu Risikofaktoren für Antibiotikatherapien

7.2. Primärkodierung: Überblick über neu kodierte Variablen

Komorbidität	Kodierung	ICD-10-Code
Diabetes mellitus	dm	E10-E14
Dyslipidämien	dyslipoproteinemia	E78
Hypertonie	hypertonia	I10-I15
Ischämische Herzerkrankungen	icm	I20-I25
Vorhofflimmern	vhf	I48
Herzinsuffizienz	cardiacinsuff	I50
Chronisch obstruktive Lungenkrankheit	copd	J44
Asthmatische Erkrankungen	asthma	J45
Glomeruläre und tubulointerstitielle Nierenkrankheiten inklusive Niereninsuffizienz	Kidney	N00-N19
Altersgruppen (18-14/ 25-64/ 65-84/ ≥85 Jahre)	age_group	-

Alters-Schwellenwerte (40, 50, 60, 70, 80, 90)	threshold_ageunder	-
Aggregation der behandelnden Stationen (Chirurgie/Innere/Sonstige)	department_aggregated	-
Aggregation aller behandelnden Stationen außer chirurgische (Risikoabteilung Version 1)	department_aggregated_ internalANDother	-
Alle Abteilungen außer kardiologische- und kardiochirurgische (Risikoabteilungen Version 2)	NOcardio_departments	-
Kombination der Kodierungen Risikoabteilungen Versionen 1 & 2 (Risikoabteilung Version 3)	risk_departments	-
CCI-Gesamtpunktzahl	CCI	-
Übergewicht mit BMI > 30	fatty	-
Einnahme von H2-Blockern oder Protonen-Pumpen-Inhibitoren	antacida	-
Reise in Länder mit endemischen ESBL-E/Carbapenemasen oder endemischen tropischen Infektionskrankheiten	foreigntravel	-
Bekannte Kolonisation mit MRE	mdro_colon	-
Bekannter Zustand nach invasiver Infektion mit MRE in den vergangenen 6 Monaten	mdro_infect	-

Tabelle 6.2.: Neu kodierte Variablen im Rahmen der Primärkodierung

7.3. Verlauf der Kennzahlen der binär logistischen Regressionsmodelle

Entgegen der Erwartung fiel die AUC der ROC nach dem Entfernen der ersten (nicht signifikanten) Variablen von rund 75% auf fast 72% ab, um sich in den folgenden Iterationen nach dem Entfernen weiterer insignifikanter Variablen langsam zu erholen. Insgesamt zeigte sich eine leicht steigende Tendenz der AUC mit zunehmendem Ausschluss von Variablen, ab Iteration 34 (9 Variablen übrig) sackte die AUC jedoch kontinuierlich bis auf ihr Minimum von 66,3% in der letzten Iteration (Nr. 42, 1 Variable übrig) ab. Die beste durchschnittliche AUC der 10-fachen Kreuzvalidierung wurde in Iteration 33 mit 10 übriggebliebenen Variablen erreicht (76,3%). Wenn man alternativ zu den errechneten Mittelwerten einzelne Werte der Kreuzvalidierung betrachtet, wurden beispielsweise in Iteration 17 und 18 AUC von rund 85% erreicht. Der negative prädiktive Wert (NPV) veränderte sich entgegen der Erwartung im Verlauf kaum und zeigte am Ehesten eine rückläufige Tendenz. Zu Beginn der Iterationen wurde im logistischen Regressionsmodell mit allen Variablen ein NPV von 68,5% erreicht. Bei der 42. Iteration wurden noch rund 64% im NPV erreicht. 70% der Fälle ohne positiven

Endpunkt konnten in Iteration 20 auch als PatientInnen ohne Antibiotikum identifiziert werden (NPV = 70%). Bei der Betrachtung der Sensitivität in den Vier-Felder-Tafeln der Iterationen ergab sich im Verlauf ebenfalls eine abfallende Tendenz. In der ersten Iteration konnten noch 63,4% erreicht werden, in Iteration Nr. 38 ergab sich mit 45,1% der Tiefpunkt der errechneten Sensitivität. Das Entfernen weiterer Variablen führte also dazu, dass weniger Fälle aus der Grundgesamtheit derer mit Antibiotikum (positiver Endpunkt) auch als solche erkannt wurden. Für die Spezifität der Modelle ergab sich im Verlauf der Iterationen eine positive Tendenz. Während die Modelle zu Beginn mit 76,9% relativ unspezifisch waren, konnten zuletzt mit wenigen übriggebliebenen Variablen immerhin rund 87% der PatientInnen ohne Antibiotikum auch als solche markiert werden. Bei der Betrachtung einzelner Iterationen der 10-fachen Kreuzvalidierung konnten sogar über 95% Spezifität mit den Regressionsmodellen erreicht werden. Der PPV oscillierte im Verlauf der Iterationen stets um seinen Ausgangswert, um dann in der letzten Iteration in der Nähe desselben bei 75% zu landen.

7.4. Beschreibung zur Entwicklung der optimalen Variablenkombination

Mit 7 Variablen ergab sich in Iteration 35a eine solide Modellleistung und der im Vergleich besten Sensitivität. In dieser ersten Version für den zukünftigen Score war die Variable *NOdyslipoproteinemia* noch vorhanden. Wie bereits beschrieben, wurde *NOdyslipoproteinemia* durch überarbeitete Versionen von Risikoabteilungen (Version 2 und 3) ersetzt und bei den folgenden Kombinationen nicht mehr integriert. Für die neu kodierten Risikoabteilungen hatten sich in den univariaten Betrachtungen bessere Korrelationswerte als für die Nicht-Dyslipidämie allein gezeigt. In einem zweiten Kombinationsversuch folgte eine Berechnung mit 6 Variablen. Dies führte zwar zu einer Verbesserung der Spezifität des Modells, alle anderen statistischen Kenngrößen waren in dieser neuen Berechnung jedoch niedriger. Die Entwicklung einer alternativen Kodierung der Risikoabteilung als Kombination aus den in den Iterationen 35a und 35b verwendeten Kodierungen für Risikoabteilungen (vgl. Tabelle 6.2.), ergaben sich wieder höhere Modellleistungsparameter. Da im klinischen Alltag, beispielsweise in der Notaufnahme, die Erhebung der aktuell behandelnden medizinischen Abteilung für einen Score nicht immer zuverlässig realisierbar ist, wurde in Iteration 35d eine Kombination aus Variablen ohne die Integration von Risikoabteilungen gewählt. Hierunter zeigte sich im Vergleich zur vorangegangenen Iteration eine stabile Prädiktionsgenauigkeit des Modells und eine Verbesserung der AUC. In den univariaten Betrachtungen der Risikofaktoren hatten sich *hämatologische Malignome* als besonders hoch und stark mit dem Endpunkt korreliert herausgestellt. Um den eigentlichen prädiktiven Effekt der Variable im Regressionsmodell abschätzen zu können, wurden hämatologische Malignome als Variable aus dem folgenden Modell (Iteration 35e) herausgestrichen. Dies führte zu einer teils

deutlichen Reduktion der Prädiktionsleistungsparameter (Spezifität und PPV), Sensitivität, NPV und AUC blieben jedoch auch ohne die Variable *leukemialymphoma* verhältnismäßig stabil. Da die Betrachtung der Verläufe der Kenngrößen im Verlauf der Iterationen eine deutliche Absenkung des Leistungsniveaus nach der Entfernung der Variable *keine elektive OP* ergeben hatte, wurde diese Variable nun wieder in die folgenden Modelle aufgenommen. Die Variable *hämatologische Malignome* wurde aufgrund ihres starken Effekts auf die Prädiktionsgenauigkeit ebenfalls wieder aufgenommen. In Iteration 35f ergab sich so eine deutliche verbesserte Modelleleistung. Die vorher in einer Vorwärtsselektion hinzugefügte Variable Altersschwelle <70 Jahre wurde in einer folgenden weiteren Kombination wieder entfernt, worunter die Prädiktionsleistung des Modells annähernd gleich blieb. Analog zu Iteration 35e wurde in Iteration 35h nochmals die Variable *leukemialymphoma* ausgelassen, was erwartungsgemäß wieder zu einer deutlichen Reduktion der Modelleleistung führte. Da sich die beiden Variablen *hämatologische Malignome* und *keine elektive OP* in den berechneten Modellen als Leistungsträger herausgestellt hatten, wurden in den darauffolgenden drei Kombinationsversuchen 35i, 35j, 35k jeweils eine der drei Variablen *v_224* (medikamentöse Immunsuppression), *v_154* (Antibiotikum in der Vergangenheit) und *v_158* (Hospitalisierung in der Vergangenheit) ausgelassen. Die Entfernung der Variable *v_224* führte hier zum größten Leistungsabfall, die Entfernung der Variable *v_154* führte hingegen kaum zu einer Verschlechterung im Vergleich zu den bis dato besten Modelleleistungen der Iterationen 35f und 35g. Die Berechnung eines weiteren Modells ohne die Variable der elektiven Operation ergab als weitere Kombinationsmöglichkeit vor allem einen Abfall in der AUC.

7.5. PILGRIM-Score Modelleleistungen nach Matsushita und Schneeweiss

Schwellenwert	Mod.+ / Ep.+	Mod.+ / Ep.-	Mod.- / Ep.+	Mod.- / Ep.-	Sens.	Spez.	PPV	NPV
6	0	1	423	465	0,0	99,8	0,0	52,4
5	1	2	422	464	0,2	99,6	33,3	52,4
4	161	53	262	413	38,1	88,6	75,2	61,2
3	191	65	232	401	45,2	86,1	74,6	63,3
2	263	154	160	312	62,2	67,0	63,1	66,1
1	382	358	41	108	90,3	23,2	51,6	72,5

Tabelle 6.3.: Ergebnisse der Vier-Felder Tafel je nach Schwellenwert des Modells (Matsushita) und zugehörige Modelleleistungen („Mod. +“ = Endpunkt laut Modell positiv; „Ep. +“ = Endpunkt im Entwicklungsdatensatz tatsächlich positiv, „Mod. -“, „Ep. -“ = analog)

Schwellenwert	Mod.+/Ep.+	Mod.+/Ep.-	Mod.-/Ep.+	Mod.-/Ep.-	Sens.	Spez.	PPV	NPV
10	0	1	423	465	0,0	99,8	0,0	52,4
9	1	2	422	464	0,2	99,6	33,3	52,4
8	122	34	301	432	28,8	92,7	78,2	58,9
7	161	53	262	413	38,1	88,6	75,2	61,2
6	181	61	242	405	42,8	86,9	74,8	62,6
5	205	79	218	387	48,5	83,0	72,2	64,0
4	250	135	173	331	59,1	71,0	64,9	65,7
3	372	328	51	138	87,9	29,6	53,1	73,0
2	382	358	41	108	90,3	23,2	51,6	72,5

Tabelle 6.4.: Ergebnisse der Vier-Felder Tafel je nach Schwellenwert des Modells (Schneeweiss) und zugehörige Modelleleistungen („Mod. +“ = Endpunkt laut Modell positiv; „Ep. +“ = Endpunkt im Entwicklungsdatensatz tatsächlich positiv, „Mod. -“, „Ep. -“ = analog)

7.6. PILGRIM-Score Modelleleistungen (Validierung) nach Matsushita und Schneeweiss

Schwellenwert	Mod.+/Ep.+	Mod.+/Ep.-	Mod.-/Ep.+	Mod.-/Ep.-	Sens.	Spez.	PPV	NPV
6	0	0	126	178	0,0	100,0	-	58,6
5	1	0	125	178	0,8	100,0	100,0	58,7
4	47	15	79	163	37,3	91,6	75,8	67,4
3	54	20	72	158	42,9	88,8	73,0	68,7
2	88	62	38	116	69,8	65,2	58,7	75,3
1	116	118	10	60	92,1	33,7	49,6	85,7

Tabelle 6.5.: Ergebnisse der Vier-Felder Tafel je nach Schwellenwert des Modells (Matsushita) und zugehörige Modelleleistungen in der Validierung („Mod. +“ = Endpunkt laut Modell positiv; „Ep. +“ = Endpunkt im Entwicklungsdatensatz tatsächlich positiv, „Mod. -“, „Ep. -“ = analog)

Schwellenwert	Mod.+/Ep.+	Mod.+/Ep.-	Mod.-/Ep.+	Mod.-/Ep.-	Sens.	Spez.	PPV	NPV
10	0	0	126	178	0,0	100,0	-	58,6
9	1	0	125	178	0,8	100,0	100,0	58,7
8	34	9	92	169	27,0	94,9	79,1	64,8
7	47	15	79	163	37,3	91,6	75,8	67,4
6	52	17	74	161	41,3	90,4	75,4	68,5

5	62	24	64	154	49,2	86,5	72,1	70,6
4	80	57	46	121	63,5	68,0	58,4	72,5
3	115	111	11	67	91,3	37,6	50,9	85,9
2	116	118	10	60	92,1	33,7	49,6	85,7

Tabelle 6.6.: Ergebnisse der Vier-Felder Tafel je nach Schwellenwert des Modells (Schneeweiss) und zugehörige Modelleleistungen in der Validierung („Mod. +“ = Endpunkt laut Modell positiv; „Ep. +“ = Endpunkt im Entwicklungsdatensatz tatsächlich positiv, „Mod. -“, „Ep. -“ = analog)

7.7. Auszug Studienprotokoll PILGRIM-Studie

Einschlusskriterien:

1. Alter \geq 18 Jahre.
2. Geplante antibiotische Behandlung oder hohe Wahrscheinlichkeit einer systemischen antibiotischen Behandlung innerhalb der nächsten zehn Tage für eine Dauer von \geq fünf Tagen. Ausgenommen sind prophylaktische Behandlungen mit Trimethoprim/Sulfamethoxazol.
3. PatientInnen müssen in der Lage sein vor oder innerhalb von 6 Stunden nach erster Antibiotikagabe eine Stuhlprobe abzugeben oder sind mit einem Rektalabstrich einverstanden, falls die Abgabe einer Stuhlprobe nicht möglich ist.
4. Schriftliche Einverständniserklärung liegt vor Studieneinschluss vor.

Ausschlusskriterien:

1. PatientInnen, die systemische Antibiotikatherapien für \geq sieben Tage innerhalb der vergangenen zwei Monate erhalten haben.
2. PatientInnen, die Antibiotika (außer Trimethoprim/Sulfamethoxazol) innerhalb der letzten 14 Tage vor Studieneinschluss erhalten haben (außer letzte antibiotische Therapie ist weniger als vier vor Studieneinschluss her, s.o.).
3. PatientInnen, die an klinischen Studien zur Evaluation antibakterieller Substanzen teilnehmen.
4. PatientInnen mit Diarrhoe bei Aufnahme (\geq drei ungeformte Stühle innerhalb 24 Stunden).
5. PatientInnen mit Stoma (Jejunostoma, Ileostoma, Colostoma) bei Studienaufnahme.
6. PatientInnen mit enteraler (Magensonde oder perkutane endoskopische Gastrostomie) oder parenteraler Ernährung.
7. Jegliche soziale oder logistische Gegebenheiten, die in Augen des Untersuchers mit der Durchführung der Studie interferieren könnten, wie beispielsweise beschränkte Auffassungsgabe, mangelnder Willen zur Kollaboration oder mangelhafte Kontaktmöglichkeiten nach Entlassung.
8. Ambulante PatientInnen ohne vorangegangene Hospitalisation.
9. Vorangegangene Partizipation in der Studie.

7.8. Abbildungsverzeichnis

Nr.	Seite	Beschreibung
1.1.	13	Phasen der Infektionspathogenese mit AMR
2.1.	18	Ablauf der PILGRIM-Score Entwicklung
2.2.	25	Schema 10-fache Kreuzvalidierung
2.3.	26	Schema Ablauf Iterationen mit 10-facher Kreuzvalidierung
2.4.	28	Gewichtungsintervalle nach Schneeweiss
3.1.	36	Wald-Diagramm der Variablen der univariaten Regressionsmodelle (Auswahl)
3.2.	38	Entwicklung der Kennzahlen im Verlauf der Iterationen
3.3.	47	Visualisierung der Modelleleistungsparameter in Abhängigkeit der Schwellenwerte

7.9. Tabellenverzeichnis

Nr.	Seite	Beschreibung
2.1.	28	Übersicht Methoden der variablenspezifischen Gewichtungen
3.1.	30	Basisdeskription des Entwicklungsdatensatzes
3.2.	30	Verteilung der versorgenden Abteilungen innerhalb des Entwicklungsdatensatzes
3.3.	31	Basisdeskription des Validierungsdatensatzes
3.4.	31	Verteilung der versorgenden Abteilungen innerhalb des Validierungsdatensatzes
3.5.	32	Korrelation der Variablen mit dem Endpunkt im Pearson-Chi-Quadrat-Test
3.6.	34	Korrelation der Variablen mit dem Endpunkt in den univariaten Regressionsmodellen
3.7.	37	Synopsis hochsignifikanter Variablen im Pearson-Chi-Quadrat-Test und univariater Regression
3.8.	40	Kenngößen der probatorischen Iterationen
3.9.	41	Multikollinearitätsanalyse nach Kendall-Tau-b.
3.10.	41	Ergebnisse der Modelleleistungen bei Vergleich Entwicklungsdatensatz mit Validierungsdatensatz
3.11.	42	Variablenweightungen nach Werfel, Matsushita und Schneeweiss
3.12.	43	Variablenweightungen nach Werfel, Matsushita und Schneeweiss inkl. der Variable „Elektive OP geplant“
3.13.	43	Ergebnisse der Vier-Felder Tafel in Abhängigkeit des Schwellenwertes des Modells (Werfel)
3.14.	44	Ergebnisse der der Vier-Felder Tafel in Abhängigkeit des Schwellenwertes des Modells ohne individuelle Variablenweightung (Referenz)
3.15.	45	Ergebnisse der Vier-Felder Tafel in Abhängigkeit des Schwellenwertes des Modells in der Validierung (Werfel)
3.16.	46	Ergebnisse der Vier-Felder Tafel in Abhängigkeit des Schwellenwertes des Modells ohne individuelle Variablenweightung (Referenz) in der Validierung
3.17.	46	Übersicht der finalen Variablen und ganzzahliger Variablenweightungen

3.18.	47	Ergebnisse der Vier-Felder Tafel in Abhängigkeit des Schwellenwertes des Modells nach Werfel mit ganzzahlig gerundeten Variablen gewichtungen
6.1.	67	Literaturrecherche zu Risikofaktoren für Antibiotikatherapien
6.2.	68	Neu kodierte Variablen im Rahmen der Primärkodierung
6.3.	71	Ergebnisse der Vier-Felder Tafel in Abhängigkeit des Schwellenwertes des Modells (Matsushita)
6.4.	71	Ergebnisse der Vier-Felder Tafel in Abhängigkeit des Schwellenwertes des Modells (Schneeweiss)
6.5.	72	Ergebnisse der Vier-Felder Tafel in Abhängigkeit des Schwellenwertes des Modells in der Validierung (Matsushita)
6.6.	72	Ergebnisse der Vier-Felder Tafel in Abhängigkeit des Schwellenwertes des Modells in der Validierung (Schneeweiss)

8. Vorabveröffentlichung von Ergebnissen

Es wurden keine Ergebnisse vorab veröffentlicht.