

Aus dem Deutschen Diabetes-Zentrum (DDZ) des
Leibniz-Zentrum für Diabetes-Forschung
an der Heinrich-Heine-Universität Düsseldorf
Institut für Biometrie und Epidemiologie
Direktor: Univ.-Prof. Dr. sc. hum. Oliver Kuß

**Absolute Behandlungseffekte der primären Endpunkte und der
Gesamtmortalität in den kardiovaskulären Endpunktstudien der
neuen Antidiabetika:
Meta-Analyse digitalisierter Patientendaten**

Dissertation

zur Erlangung des Grades eines Doktors der Medizin der
Medizinischen Fakultät der
Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Cihan Akbulut
2024

Als Inauguraldissertation gedruckt mit Genehmigung der
Medizinischen Fakultät der Heinrich-Heine-Universität Düsseldorf

gez.:

Dekan: Univ.-Prof. Dr. med. Nikolaj Klöcker

Erstgutachter: Prof. Dr. sc. hum. Oliver Kuß

Zweitgutachterin: Prof. Dr. med. Bernd Richter

Für meine Eltern:

In Liebe und Dankbarkeit

Teile dieser Arbeit wurden publiziert:

Oliver Kuss, Cihan Akbulut, Sabrina Schlesinger, Asen Georgiev, Malte Kelm, Michael Roden & Georg Wolff

Absolute treatment effects for the primary outcome and all-cause mortality in the cardiovascular outcome trials of new antidiabetic drugs: a meta-analysis of digitalized individual patient data

Acta Diabetologica, 59: 1349–1359 (2022)

I Zusammenfassung

Als Reaktion auf den Rosiglitazon-Skandal im Jahr 2008 entwickelte die amerikanische Zulassungsbehörde *US Food and Drug Administration* (FDA) neue Zulassungsrichtlinien für neue antidiabetische Medikamente. Der Nachweis der Nicht-Unterlegenheit gegenüber Placebo hinsichtlich harter kardiovaskulärer Endpunkte wurde zur Zulassungsvoraussetzung. Als Reaktion führte die pharmazeutische Industrie eine Reihe von großen randomisierten Studien (CVOTs) durch, in denen unter anderem Dipeptidylpeptidase-4 (DPP-4) Inhibitoren, Glukagon-like-peptide-1 (GLP-1) Rezeptoragonisten und Natrium-Glukose-Cotransporter-2 (SGLT-2) Inhibitoren untersucht wurden. In diesen Studien konnte in allen Fällen die kardiovaskuläre Sicherheit, teilweise sogar die Überlegenheit für kardiovaskulären Endpunkte nachgewiesen werden.

Die Behandlungseffekte in den CVOTs wurden fast ausschließlich als relative Effektmaße kommuniziert, obwohl viele internationale Leitlinien und die FDA empfehlen, Behandlungseffekte ebenfalls als absolute Effektmaße, wie z. B. als Number-needed-to-treat (NNT) zu kommunizieren. In der vorliegenden Arbeit wurden NNTs für die primären Endpunkte und die Gesamtmortalität in 19 CVOTs geschätzt. Dazu wurden die Kaplan-Meier-Überlebenszeitkurven aus den publizierten CVOTs digitalisiert und individuelle *Time-to-event* Datensätze gebildet. Für jede Studie und jedes Ergebnis wurden Weibull-Modelle geschätzt und monatliche NNTs abgeleitet. Die NNTs wurden für alle CVOTs und innerhalb der Wirkstoffklassen mit Methoden der zufälligen Effektverteilung in Meta-Analysen aggregiert.

Die Behandlungseffekte von CVOTs erscheinen geringer, wenn sie als absolute Effektmaße, anstatt als relative Behandlungseffekte dargestellt werden. Insbesondere zeigt sich in der Meta-Analyse des primären Endpunkts nur eine Meta-NNT von 100 (95%-KI: 60 - 303) nach 29 Behandlungsmonaten und für die Gesamtmortalität von 128 (95%-KI: 85 - 265) nach 39 Behandlungsmonaten. Hinsichtlich der Wirkstoffklasseneffekte zeigen sich ähnliche Effekte der Meta-NNT für SGLT-2 Inhibitoren und GLP-1 Rezeptoragonisten. Die Meta-NNT der DPP-4 Inhibitoren zeigt für beide untersuchten Endpunkte die geringsten absoluten Behandlungseffekte. Darüber hinaus konnte im Rahmen dieser Arbeit die Validität der angewandten Methode zur Bildung von Meta-Analysen und zur Schätzung von NNTs aus digitalisierten, patientenindividuellen *Time-to-event* Datensätzen auf mehreren Ebenen nachgewiesen werden.

Zusammenfassend lässt sich aus den Ergebnissen der einzelnen NNTs und der Meta-NNTs für die drei Wirkstoffklassen sowie für den primären Endpunkt und die Gesamtmortalität ableiten, dass die Behandlungseffekte weniger eindrucksvoll erscheinen, wenn sie als absolute Behandlungseffekte kommuniziert werden. Wir stimmen daher mit anderen Autoren überein, dass für ein gutes Verständnis der Studienergebnisse sowohl relative als auch absolute Behandlungseffekte kommuniziert werden sollten.

II Summary

Following the rosiglitazone scandal in 2008, the US Food and Drug Administration (FDA) issued a revised version of its approval guidelines for new anti-diabetic drugs. A requirement for approval was to demonstrate non-inferiority to placebo in hard cardiovascular endpoints. In response, the pharmaceutical industry has conducted a number of large, randomized cardiovascular outcome trials (CVOTs), including dipeptidyl peptidase-4 (DPP-4) inhibitors, glucagon-like peptide-1 (GLP-1) receptor agonists and sodium-glucose cotransporter-2 (SGLT-2) inhibitors. In these studies, cardiovascular safety was demonstrated in all cases, and in some cases superiority in cardiovascular endpoints.

The CVOTs almost exclusively communicated findings as relative treatment effects, although many international guidelines and the FDA itself recommend including absolute treatment effects, such as number-needed-to-treat (NNT). In the present study, NNTs were estimated for the primary endpoints and all-cause mortality in 19 CVOTs. Individual time-to-event datasets were created by digitizing Kaplan-Meier curves from published CVOTs. For each trial and each endpoint, Weibull models were estimated and monthly NNTs were derived. The NNTs were aggregated for all CVOTs and within drug classes using random-effects meta-analysis methods.

The treatment effects of the anti-diabetic drugs in the CVOTs appear to be smaller when presented as absolute treatment effects, such as NNTs, rather than relative treatment effects. In particular, the meta-analysis of the primary endpoint for all included CVOTs only shows a meta-NNT of 100 (95%-CI: 60 - 303) after 29 months of treatment and for all-cause mortality of 128 (95%-CI: 85 - 265) after 39 months of treatment. In terms of class effects, the meta-NNT for SGLT-2 inhibitors and GLP-1 receptor agonists showed similar effects. The meta-NNT for DPP-4 inhibitors showed the lowest absolute treatment effect for both endpoints analyzed.

In addition, this study demonstrated the validity of the method used to generate meta-analyses and estimate NNTs from digitized patient-specific time-to-event datasets at multiple levels.

In summary, the results of the individual NNTs and meta-NNTs for all three drug classes, as well as for the primary endpoint and all-cause mortality, suggest that treatment effects are less impressive when reported as absolute treatment effects.

Therefore, we agree with other authors that it is essential to communicate both relative and absolute treatment effects to ensure a comprehensive understanding of the study results.

III Abkürzungsverzeichnis

Abb.	Abbildung
ACS	Akutes Koronarsyndroms
ATP	Adenosintriphosphat
BMI	Body Mass Index
Bzw.	Beziehungsweise
ca.	circa
cAVK	cerebrale arterielle Verschlusskrankheit
CVOT	Cardiovascular Outcome Trials
DDP-4 Inhibitor	Dipeptidylpeptidase-4 Inhibitor
DM-Typ 2	Diabetes mellitus Typ 2
eGFR	Geschätzte glomeruläre Filtrationsrate
GLP-1 RA	Glukagon-like peptide-1 Rezeptoragonist
GLUT-2	Glukosetransporter-2
GLUT-4	Glukosetransporter-4
HbA1c	Hämoglobin A1c
HR	Hazard Ratio
i.d.R.	in der Regel
KHK	Koronare Herzerkrankung
KI	Konfidenzintervall
MACE	Major Adverse Cardiovascular Event
NNT	Number-needed-to-treat
oGTT	Oraler Glukosetoleranztest
pAVK	periphere arterielle Verschlusskrankheit
P- Wert	Signifikanzniveau
SD	Standardabweichung
SGLT-2 Inhibitor	Sodium-glucose-linked-transporter-2 Inhibitor
U.S. FDA	U.S. Food and Drug Administration

IV Inhaltsverzeichnis

I	Zusammenfassung	
II	Summary	
II	Abkürzungsverzeichnis	
IV	Inhaltsverzeichnis	

1	Einleitung	1
1.1	Hintergrund der vorliegenden Arbeit	1
1.2	Ergebnisse der Cardiovascular Outcome Trials (CVOTs)	3
1.3	Diabetes mellitus	3
1.3.1	Definition	3
1.3.2	Klassifikation	4
1.3.3	Epidemiologie	5
1.3.4	Physiologie der Insulinwirkung und Pathophysiologie des DM-Typ 2	5
1.3.5	Klinische Manifestation	7
1.3.6	Diagnostik	8
1.3.7	Therapie des Diabetes mellitus Typ 2	9
1.4	Absolute und relative Effektmaße	12
1.5	Ziele der Arbeit	15
2	Material und Methoden	16
2.1	Material	16
2.2	Methoden	16
2.2.1	Studiendesign	16
2.2.2	Digitalisierung und Datenrekonstruktion	17
2.2.3	Statistische Analyse	18
2.2.4	Meta-Analyse	20
2.2.5	Lebenszeitperspektive	21
2.2.6	Validität der Methode	21
2.2.7	Definition und Interpretation der Number-needed-to-treat (NNT)	22
2.3	Ethikvotum	22
3	Ergebnisse	23
3.1	Studienauswahl	23
3.2	Studiencharakteristika	23
3.3	NNT-Ergebnisse nach 12, 24, 36 und 48 Behandlungsmonaten	27
3.4	Meta-NNT: DDP-4 Inhibitoren, GLP-1 Rezeptoragonisten und SGLT-2 Inhibitoren	32

3.5 Meta-NNT: Primärer Endpunkt und die Gesamtmortalität	32
3.6 Meta-NNT: Gesamtmortalität für 30 Jahre (Lebenszeitperspektive)	32
3.7 Validität	37
4. Diskussion	45
4.1 Limitationen	53
4.2 Schlussfolgerungen.....	55
5 Literatur- und Quellenverzeichnis.....	57
6 Anhang	64
6.1 Guyot et al. [64]: Flowchart of the algorithm (all information case).....	64
6.2 Guyot et al. [64]: Additional file 1: The algorithm (R coding)	64
6.3. Vollständige Datensätze der digitalisierten Patientendaten in einer öffentlichen Datenbank [70]	64
7 Danksagung.....	69

1 Einleitung

1.1 Hintergrund der vorliegenden Arbeit

Die *Cardiovascular Outcomes Trials* (CVOTs) sind randomisierte kontrollierte Studien, die das kardiovaskuläre Sicherheitsprofil neuer antidiabetischer Medikamente untersuchen [1]. Die CVOTs wurden entwickelt, nachdem die US.-Arzneimittelbehörde (Food and Drug Administration [FDA]) 2008 neue Richtlinien für die Entwicklung von antidiabetischen Medikamenten vorgab [2]. Die neuen Richtlinien waren eine Reaktion auf den „Rosiglitazon-Skandal“ [3].

Rosiglitazon war zu diesem Zeitpunkt ein gängiges Medikament zur Behandlung von Diabetes mellitus Typ 2 und ein wichtiger Bestandteil der medikamentösen Stufentherapie. Im Jahr 2007 wurde im *New England Journal of Medicine* eine Studie von Nissen et al. [3] veröffentlicht, die das kardiovaskuläre Risikoprofil von Rosiglitazon untersuchte. Nissen et al. [3] konzipierten eine Meta-Analyse aus 42 Studien mit über 27.000 Patienten (Rosiglitazon-Gruppe n = 15.565; Kontrollgruppe n = 12.282). In der Behandlungsgruppe mit Rosiglitazon wurden 86 Myokardinfarkte und 39 kardiovaskuläre Todesursachen beobachtet, während in der Kontrollgruppe nur 72 Myokardinfarkte und 22 kardiovaskuläre Todesursachen im gleichen Beobachtungszeitraum festgestellt werden konnten. Somit lag die Odds-Ratio für Myokardinfarkte bei 1,43 (95%-Konfidenzintervall [KI]: 1,03 - 1,98; P=0.03) und für die kardiovaskuläre Mortalität bei 1,64 (95%-KI: 0.98 - 2.74; P=0.06) zu Ungunsten von Rosiglitazon. Die Meta-Analyse von Nissen et al.[3] warf die Frage auf, ob Rosiglitazon mit einem erhöhten Risiko für kardiovaskuläre Ereignisse assoziiert war. Obwohl die Studie von Nissen et al. [3] einige Limitationen aufwies, wie z. B. niedrige Ereignisraten, wurde die Frage der kardiovaskulären Sicherheit von Antidiabetika in den wissenschaftlichen Fokus gerückt.

Vor dem „Rosiglitazon-Skandal“ basierten die Empfehlungen für die Entwicklung neuer antidiabetischer Medikamente hauptsächlich auf dem Nachweis einer Blutzuckersenkung. Lange Zeit war umstritten, ob die alleinige Messung des HbA1c-Werts ausreicht, um Rückschlüsse auf den kardiovaskulären Gesamtnutzen zu ziehen. Während eine Senkung des HbA1c-Werts nachweislich zu einer Risikoreduktion für

mikrovaskuläre Folgeschäden führt [4], konnte für makrovaskuläre Endpunkte kein eindeutiger Effekt nachgewiesen werden [5-9].

Ein möglicher Grund für die eingeschränkte Aussagekraft der erfassten Studien zur kardiovaskulären Gesamtrisikoreduktion waren die niedrigen Ereignisraten. In den meisten Fällen handelte es sich um klinische Studien der Phase 2 oder 3 mit einem kurzen Beobachtungszeitraum von sechs bis zwölf Monaten. Zudem waren die Probanden im Durchschnitt jung und hatten selten kardiovaskuläre Vorerkrankungen. Aufgrund dieser Umstände traten kardiovaskuläre Ereignisse im Studienzeitraum nur selten auf [10-12].

Zusammenfassend lässt sich sagen, dass der Wissensstand über das kardiovaskuläre Sicherheitsprofil der neuen antidiabetischen Medikamente gering war und der „Rosiglitazon-Skandal“ den Druck auf die Zulassungsbehörden erhöhte, so dass 2008 neue Richtlinien für die Arzneimittelentwicklung geschaffen wurden.

Die neuen Richtlinien der FDA beinhalteten sogenannte harte kardiovaskuläre Endpunkte. Der primäre Endpunkt setzte sich aus mehreren obligatorischen und optionalen Endpunkten zusammen. Zu den obligatorischen Endpunkten zählten die Endpunkte kardiovaskuläre Mortalität, nicht-tödlicher Myokardinfarkt und nicht-tödlicher Schlaganfall (*major adverse cardiovascular event* [MACE]). Optional konnten auch die Endpunkte Hospitalisierung aufgrund eines akuten Koronarsyndroms (ACS), dringende Eingriffe zur Revaskularisation und Hospitalisierung aufgrund einer dekompensierten Herzinsuffizienz eingeschlossen werden. Für den Nachweis der Nicht-Unterlegenheit gegenüber der Kontrollgruppe sollte die Obergrenze des 95%-Konfidenzintervalls auf der relativen Hazard Ratio Skala nicht größer als 1,3 betragen. Auf der Grundlage der neuen Richtlinien wurden für alle neuen antidiabetischen Medikamente große randomisierte kontrollierte Studien, die sogenannten *Cardiovascular Outcomes Trials* entwickelt. Entscheidend für die Studiendauer war eine ausreichend hohe Anzahl von Ereignissen für den definierten primären Endpunkt. Um dieses Ziel zu erreichen, wurden insbesondere Patienten mit einem erhöhten kardiovaskulären Risikoprofil in die Studien eingeschlossen [1, 10].

1.2 Ergebnisse der Cardiovascular Outcome Trials (CVOTs)

Alle seit 2008 publizierten CVOTs konnten die Nicht-Unterlegenheit im primären Endpunkt nachweisen. Damit kann für alle neuen antidiabetischen Medikamente die kardiovaskuläre Sicherheit nach den Richtlinien der FDA bestätigt werden.

Als erste CVOT der Wirkstoffklasse der SGLT-2 Inhibitoren hat die EMPA-REG Studie im Jahr 2015 viel Aufmerksamkeit erregt. Der kombinierte primäre Endpunkt (nicht-tödlicher Herzinfarkt, nicht-tödlicher Schlaganfall und kardiovaskuläre Mortalität [3-Punkt-MACE]) war mit einer Hazard Ratio (HR) von 0,86 (95%-KI: 0,74 - 0,99; $P < 0,001$ für Nicht-Unterlegenheit und $P = 0,04$ für Überlegenheit) signifikant reduziert und konnte erstmalig die Überlegenheit für den primäre Endpunkt nachweisen. Ausschlaggebend für die relative Risikoreduktion des primären Endpunktes war die Risikoreduktion der kardiovaskulären Mortalität (HR 0,62 [95%-KI: 0,49 - 0,77; $P < 0,001$]). Außerdem war die Anzahl der Patienten, die aufgrund einer Herzinsuffizienz ins Krankenhaus eingewiesen wurden, in der Interventionsgruppe mit Empagliflozin signifikant um relative 35% reduziert (HR 0,65 [95%-KI: 0,50 - 0,85; $P = 0,002$]) [13]. Der Nachweis eines Überlebensvorteils kam zu diesem Zeitpunkt einem Paradigmenwechsel gleich. Insgesamt konnten zwei CVOTs aus der Wirkstoffklasse der SGLT-2 Inhibitoren (EMPA-REG und CANVAS) und vier aus der Wirkstoffklasse der GLP-1 Rezeptorantagonisten (AMPLITUDE-O, LEADER, SUSTAIN-6 und REWIND) einen kardiovaskulären Vorteil nachweisen. In Folge dieser Studienergebnisse wurden internationale Leitlinien zur Behandlung von Diabetes mellitus Typ 2 und der Herzinsuffizienz angepasst [14, 15].

1.3 Diabetes mellitus

1.3.1 Definition

Der Begriff Diabetes mellitus leitet sich vom altgriechischen Wort „*diabetes*“ für hindurchfließen und dem lateinischen Wort „*mellitus*“ für honigsüß ab. Unter dem modernen medizinischen Begriff Diabetes mellitus werden ätiologisch unterschiedliche Erkrankungen zusammengefasst, denen eine Hyperglykämie gemeinsam ist [16].

1.3.2 Klassifikation

Die Krankheit Diabetes mellitus kann in vier Klassen eingeteilt werden.

Der Diabetes mellitus Typ 1 umfasst etwa 5-10 % aller Diabetiker. Diese Erkrankung wurde lange Zeit als juveniler Diabetes bezeichnet, da sie häufig bereits im Kindesalter auftritt und i. d. R. mit einem absoluten Insulinmangel einhergeht. Der Diabetes mellitus Typ 1 kann in einen immunologischen und einen idiopathischen Subtyp differenziert werden. Der immunologische Subtyp wird durch eine autoimmungeleitete Zerstörung der Beta-Zellen in der Bauchspeicheldrüse definiert. Die häufigsten Autoantikörper sind Zytoplasmatische Inselzell-Antikörper (ICA), Anti-GAD-Antikörper (GADA), Insulin-Auto-Antikörper (IAA), Anti-IA-2-Antikörper (IA-2A) sowie Anti-ZnT8-Antikörper (ZnT8A). Außerdem besteht eine starke Assoziation mit dem humanen Leukozytenantigen-System (HLA). Die Dynamik der Zerstörung der insulinproduzierenden Beta-Zellen ist variabel, sodass die Erkrankung auch erst im Erwachsenenalter sich klinisch ausprägen kann. In diesem Fall handelt es sich um den sogenannten *late-onset autoimmune diabetes in adults* (LADA). Der idiopathische Diabetes mellitus Typ 1 wird durch den fehlenden Nachweis der bekannten Autoantikörper definiert. Aufgrund der vollständigen Zerstörung der Beta-Zellen mit einem daraus resultierenden absoluten Insulinmangel wird eine medikamentöse Insulintherapie notwendig [17].

Der Diabetes mellitus Typ 2 stellt mit 90% aller Diabetiker die größte Gruppe dar und wurde früher auch als Altersdiabetes bezeichnet. Pathophysiologisch geht die Erkrankung mit einem relativen Insulinmangel und einer peripheren Insulinresistenz einher. Die Erkrankung wird nicht autoimmun vermittelt, ist aber mit einer starken genetischen Prädisposition verbunden. Zusätzlich können Risikofaktoren wie stammbetonte Adipositas, Bewegungsmangel, Nikotinabusus und eine hochkalorische Ernährung (westliche Diät) die Erkrankung auslösen, die unter dem Begriff metabolisches Syndrom zusammengefasst wird. Da eine hyperglykämische Stoffwechsellage lange asymptomatisch bleiben kann, erfolgt die Erstdiagnose nicht selten im Rahmen einer Routineuntersuchung [17].

Zur Gruppe des Diabetes mellitus Typ 3 gehören noch weitere Grunderkrankungen, die zu einer hyperglykämischen Stoffwechsellage führen können. Dazu gehören die Mukoviszidose, die Hämochromatose, das Cushing-Syndrom, die Akromegalie oder das

Phäochromozytom, aber auch iatrogene Beeinträchtigungen des Pankreas, wie z. B. nach viszeralen Operationen [17].

Diabetes mellitus Typ 4, auch Schwangerschaftsdiabetes genannt, ist eine Störung der Glukosetoleranz, die erstmals in der Schwangerschaft auftritt. In der Regel normalisiert sich der Glukosestoffwechsel nach der Schwangerschaft wieder. Allerdings besteht nach der Schwangerschaft ein erhöhtes Risiko für die Entwicklung eines Diabetes mellitus Typ 2 [17, 18].

1.3.3 Epidemiologie

Sowohl in Industrie- als auch in Entwicklungsländern nimmt die Prävalenz von Diabetes mellitus Typ 2 stetig zu. Epidemiologische Schätzungen gehen davon aus, dass bis zum Jahr 2045 weltweit mehr als 700 Millionen Menschen an Diabetes mellitus Typ 2 erkrankt sein werden [19]. Auch in Entwicklungsländern ist ein Anstieg der Prävalenz von Menschen mit Diabetes mellitus Typ 2 zu beobachten. Im Jahr 2023 lag die Prävalenz von Diabetes mellitus Typ 2 in China bei ca. 95 Millionen und in Indien bei ca. 65 Millionen Menschen. Damit weisen die beiden bevölkerungsreichsten Nationen der Welt auch die höchsten Prävalenzen von Menschen mit Diabetes mellitus Typ 2 auf [20].

1.3.4 Physiologie der Insulinwirkung und Pathophysiologie des DM-Typ 2

Für die physiologische Stoffwechselregulation des Blutzuckerspiegels spielt das Peptidhormon Insulin eine Schlüsselrolle. Das Vorläuferpeptid Proinsulin wird in den Betazellen der Bauchspeicheldrüse synthetisiert und durch die proteolytische Abspaltung des C-Peptids gebildet. Daraus entsteht das stoffwechselaktive Hormon Insulin [21]. Die Freisetzung von Insulin und dem abgespaltenen C-Peptid erfolgt mittels Exozytose und wird primär durch die Glukosekonzentration im Blut bestimmt. Der Anstieg der Blutglukosekonzentration führt über den insulinunabhängigen Glukosetransporter-2 (GLUT-2) zum Einstrom von Glukose in die Betazellen des Pankreas. In der Beta-Zelle wird Glukose durch das Enzym Glukokinase zu Glukose-6-Phosphat phosphoryliert. Durch den damit verbundenen Konzentrationsanstieg von Adenosintriphosphat (ATP) werden die ATP-abhängigen Kaliumkanäle geschlossen, was zu einer Depolarisation der Zellmembran führt. Daraufhin öffnen sich die

spannungsabhängigen Calciumkanäle und Calcium diffundiert entlang des Gradienten in die Zellen hinein. Durch den Anstieg der Calciumkonzentration in der Beta-Zelle wird durch die Exozytose Insulin und das biologisch inaktive C-Peptid sezerniert [22].

Das sezernierte Insulin entfaltet seine Wirkung in der Peripherie durch die Bindung an transmembrane Tyrosinkinase-Rezeptoren in verschiedenen Organen. Dadurch werden weitere Signalkaskaden ausgelöst, die zur Aktivierung von Stoffwechsellenzymen führen [23]. Durch die Translokation des insulinabhängigen Glukosetransporter-4 (GLUT-4) in die Zellmembranen wird zusätzlich die intrazelluläre Glukoseaufnahme in Fett- und Skelettmuskelgewebe stimuliert, was zu einer postprandialen Senkung der Blutglukosekonzentration führt [24, 25]. In der Skelettmuskulatur und in der Leber wird Glukose in Glykogen umgewandelt sowie die Glykogenolyse und die Glukoneogenese aus Aminosäuren gehemmt. Im Fettgewebe aktiviert Insulin die Lipoproteinlipase, die ihrerseits die Aufnahme von freien Fettsäuren aus dem Blut in die Fettzellen fördert und deren Umwandlung in Triglyceride stimuliert. Damit einher geht eine Hemmung der Lipolyse in den Fettzellen [26]. Bei absolutem Insulinmangel verliert die Lipolyse ihren Gegenspieler im Stoffwechsel und es kommt zur vermehrten Bildung von Ketonkörpern. Im weiteren Verlauf kann eine lebensbedrohliche Ketoazidose ausgelöst werden. Zusätzlich supprimiert Insulin die Ausschüttung des Stoffwechselhormons Glukagon aus den Alphazellen der Bauchspeicheldrüse, welches sich in seiner Wirkung antagonistisch zu Insulin verhält [27].

Zusammenfassend hat Insulin eine anabole Stoffwechselwirkung, die zur Senkung des Blutzuckerspiegels, zur Umwandlung und Energiespeicherung von Glukose in Fettsäuren und zur Aufnahme von Aminosäuren in die Zellen führt.

Die Entstehung von Diabetes mellitus Typ 2 ist multifaktoriell bedingt. Eine starke genetische Veranlagung und Umwelteinflüsse spielen eine wesentliche Rolle. Wichtige beeinflussbare Faktoren sind Übergewicht, hochkalorische Ernährung und Bewegungsmangel. Ein nicht beeinflussbarer Faktor ist die genetische Veranlagung, die eine wichtige Rolle in der Pathophysiologie der Erkrankung spielt. Es besteht eine periphere Insulinresistenz, die zu einem verminderten insulinabhängigen Einbau von GLUT-4 in die Zellmembran von Muskel- und Fettzellen führt. Dadurch wird weniger Glukose aus dem Blut in die Zellen aufgenommen. Dem kann vorübergehend durch eine

vermehrte Insulinausschüttung aus den Betazellen entgegengewirkt werden. Langfristig erschöpfen sich die Betazellen (Apoptose) und es wird weniger Insulin ausgeschüttet [28].

1.3.5 Klinische Manifestation

Die allgemeinen Symptome sind auf die hyperglykämische Stoffwechsellage zurückzuführen. Das erste Symptom ist häufig eine durch Glykosurie verursachte Polyurie. Als Begleitsymptom der Glykosurie entwickeln die Patienten ein vermehrtes Durstgefühl (Polydipsie). Die Kaskade von Glykosurie und Polyurie wird durch die osmotische Wirkung der Glukose im Urin ausgelöst. Die Ultrafiltration in den Glomeruli der Nierenrinde ist für Glukose durchlässig, so dass Glukose aus dem Primärharn erst in den proximalen Tubuli rückresorbiert wird. Ab Blutzuckerwerten von ca. 180 mg/dl im Primärharn reichen die Rückresorptionskapazitäten der proximalen Tubuli nicht mehr aus und die Glukose wird über den Harn ausgeschieden.

Patienten mit Diabetes mellitus Typ 1 entwickeln aufgrund des absoluten Insulinmangels eine rasche Symptomausprägung, die sich häufig als Ketoazidose mit möglicher Vigilanzminderung manifestiert und in diesem Fall als *Coma diabeticum* bezeichnet werden kann. Beim Diabetes mellitus Typ 2 verläuft dieser Prozess hingegen schleichend, so dass die Veränderungen von den Betroffenen nicht unmittelbar wahrgenommen werden [16].

Die Langzeitkomplikationen lassen sich in makro- und mikrovaskuläre Schäden unterteilen. Anhaltende Hyperglykämie kann atherosklerotische Gefäßveränderungen verursachen, die zu makrovaskulären Komplikationen führen. Je nachdem, welche arteriellen Gefäße betroffen sind, können sich unterschiedliche Endorganschäden entwickeln. Zu den häufigsten gehören die koronare Herzkrankheit (KHK), atherosklerotische Veränderungen der hirnversorgenden Gefäße (cAVK) oder periphere arterielle Verschlüsse (pAVK). Die Verengung der großen Gefäße kann sich in einer Minderdurchblutung des entsprechenden Stromgebietes äußern. Hier sind Patienten mit Diabetes mellitus Typ 2 aufgrund des häufig begleitenden metabolischen Syndroms besonders betroffen. Das Risiko für kardiovaskuläre Ereignisse ist bei Patienten mit Diabetes mellitus Typ 2 doppelt so hoch wie in der Allgemeinbevölkerung [29].

Kardiovaskuläre Erkrankungen sind daher die häufigste Todesursache [30] und verursachen mehr als die Hälfte der Gesundheitskosten bei Typ-2-Diabetikern [31].

Bei allen Formen des Diabetes mellitus entwickeln sich nach längerer hyperglykämischer Stoffwechsellage mikrovaskuläre Folgeschäden. Der erhöhte Glukosespiegel im Blut erhöht die Wahrscheinlichkeit einer nicht-enzymatischen Glykierung von Proteinen und Lipiden. Dies führt zu einer arteriellen Minderversorgung im Bereich der Arteriolen und Kapillaren. Die häufigsten mikrovaskulären Schädigungen sind die diabetische Retino-, Nephro- und Polyneuropathie sowie das diabetische Fußsyndrom [16].

1.3.6 Diagnostik

Die Diagnose Diabetes mellitus Typ 2 kann anhand klinischer Symptome und laborchemischer Untersuchungen gestellt werden. Bei Vorliegen der klassischen Symptome ist ein Nüchternblutzuckerwert von > 200 mg/dl ausreichend. Ein Nüchternblutzuckerwert von >126 mg/dl an zwei verschiedenen Tagen kann ebenfalls zur Diagnose führen. In unklaren Fällen mit hohem klinischem Verdacht und unklaren Nüchternblutzuckerwerten kann die Diagnose durch einen oralen Glukosetoleranztest (oGTT) gestellt werden. Nach 16 Stunden Nahrungskarenz werden 75 g Glukose getrunken und nach 2 Stunden der Blutzucker gemessen. Ein Blutzuckerwert von > 200 mg/dl spricht für einen Diabetes mellitus. In der klinischen Praxis wird die Diagnose meist durch die Bestimmung des HbA1c-Werts gestellt, da dieser Wert weniger anfällig für kurzfristige Blutzuckerschwankungen ist [32].

Hämoglobin A1 ist das häufigste Hämoglobin beim erwachsenen Menschen. Es kommt in den Erythrozyten vor und dient dem Sauerstofftransport im Blut. Das im Blut zirkulierende Hämoglobin kommt mit freien Zuckermolekülen in Kontakt, wobei es zu einer irreversiblen, nichtenzymatischen Bindung von Glukose an das Hämoglobin kommt. Dieser Vorgang ist an sich physiologisch und findet auch bei Menschen ohne Diabetes mellitus statt. Die Wahrscheinlichkeit einer solchen Glykierung ist jedoch bei Diabetikern aufgrund des durchschnittlich höheren Blutzuckerspiegels erhöht. Die physiologische Lebensdauer von Erythrozyten und dem darin gebundenen Hämoglobin beträgt 100-120 Tage. Der HbA1c-Wert spiegelt somit den durchschnittlichen Blutzuckerspiegel der letzten 3 bis 4 Monate wider. Ab einem HbA1c-Wert von $\geq 6,5\%$

kann die Diagnose gestellt werden. Im Bereich zwischen 5,7% und 6,5% wird von einer prädiabetischen Stoffwechsellage ausgegangen [33].

	Physiologisch	Prädiabetes	Diabetes mellitus
Serumglukose in mg/dl (nüchtern)	<100	100-125	≥ 126
HbA1c in %	<5,7	5,7-6,4	≥ 6,5
oGTT* nach 2 Stunden in mg/dl	<140	140-199	≥ 200

Abb. 1: **Laborchemische Kriterien für die Diagnose Diabetes mellitus Typ 2 [16]**. oGTT= oraler Glukosetoleranztest mit 75 g Glukose.

1.3.7 Therapie des Diabetes mellitus Typ 2

Aufgrund der multifaktoriellen Genese des Diabetes mellitus Typ 2 ist die Therapie komplex und muss für jeden Patienten individuell angepasst werden. Darüber hinaus müssen Risikofaktoren identifiziert und berücksichtigt werden. Hierbei sind insbesondere die beeinflussbaren Risikofaktoren wie Übergewicht, körperliche Inaktivität, hochkalorische Ernährung, Dyslipidämie, Alkohol- und Tabakkonsum relevant [32, 34]. Im Vordergrund der Therapieplanung steht die partizipative Entscheidungsfindung, auf deren Basis die individuellen Therapieziele für die wichtigsten Parameter (Lebensstil, Ernährung, Blutzucker, Gewicht, Blutdruck und Lipide) gemeinsam festgelegt werden. Die Therapieziele sollten jedoch auch kritisch unter Berücksichtigung von Lebenserwartung, Komorbiditäten und Lebensqualität festgelegt werden [15, 32].

Die Grundlage jeder Behandlung ist die Änderung des Lebensstils. In mehreren Studien konnte nachgewiesen werden, dass eine Änderung des Lebensstils zu einer signifikanten Reduktion der diabetischen Komplikationen führt und teilweise eine Remission erreicht werden kann [35, 36]. Generell wird für Patienten mit Diabetes mellitus Typ 2 eine ausgewogene Ernährung im Sinne einer mediterranen Ernährung empfohlen [37, 38]. Darüber hinaus werden aerobes Ausdauertraining (150 Minuten pro Woche) und Krafttraining zum Erhalt der Muskulatur empfohlen [39, 40]. Außerdem ist die aktive und passive Exposition gegenüber Nikotin zu vermeiden [41]. Aufgrund des bereits erhöhten kardiovaskulären Risikos bei Patienten mit Diabetes mellitus Typ 2 ist die

antihypertensive Therapie ein wichtiger Baustein zur Senkung des kardiovaskulären Gesamtrisikos. Ein weiterer therapeutischer Ansatzpunkt ist die Behandlung der Dyslipidämie. In Abhängigkeit vom individuellen kardiovaskulären Risikoprofil gelten für Patienten mit Diabetes mellitus Typ 2 LDL-Cholesterin-Zielwerte zwischen < 55 mg/dl und < 100 mg/dl [42].

Im Vordergrund der antidiabetischen Therapie steht die Stabilisierung des Blutzuckerspiegels in einem tolerierbaren Bereich, der in Abhängigkeit von Alter, Multimorbidität und Lebensqualität individuell zu erarbeiten ist. Die Verlaufskontrolle einer Therapie erfolgt über den HbA1c-Wert. Eine Senkung des HbA1c-Werts vermindert nachweislich die Entstehung und das Fortschreiten mikrovaskulärer Komplikationen [6, 7, 43-45]. Allerdings muss berücksichtigt werden, dass eine zu strenge antidiabetische Therapie das Risiko von Hypoglykämien birgt, die zu einer verminderten Vigilanz führen können und nachweislich das Risiko für Arrhythmien und myokardiale Ischämien erhöhen [46]. Nach den aktuellen Leitlinien wird ein HbA1c-Wert zwischen 6,5 und 7 % empfohlen. Bei fehlenden Anzeichen von Hypoglykämien können bei jüngeren Patienten auch HbA1c-Werte von 6,0 % bis 6,5 % angestrebt werden. Bei älteren und multimorbiden Patienten kann eine weniger strenge Blutzuckereinstellung individuell festgelegt werden [15, 47].

Die medikamentöse Therapie des Diabetes mellitus Typ 2 sollte - wie in verschiedenen internationalen Leitlinien empfohlen - nach einem Stufenschema erfolgen [15, 48].

Die erste und grundlegende Behandlungsstufe gilt für alle Patienten und beinhaltet die Änderung der Lebensstilfaktoren. In der zweiten Stufe kann ein Erstlinienmedikament verordnet werden. Unter den Erstlinienmedikamenten wird Metformin aufgrund seiner Kosteneffektivität und Behandlungssicherheit bevorzugt. [45, 49]. In der dritten Stufe können Patienten mit einer Kombinationstherapie behandelt werden, wenn die Wirkung der Monotherapie nicht ausreicht. Auch das Fortschreiten der Erkrankung kann im Laufe der Zeit eine Kombinationstherapie erforderlich machen. Liegt bereits der erste gemessene HbA1c-Wert $\geq 1,5$ % über dem Zielwert, kann auch direkt mit einer Kombinationstherapie begonnen werden [48]. In den meisten Fällen wurde eine Kombinationstherapie mit Metformin und einem zusätzlichen Antidiabetikum wie SGLT-2 Inhibitoren, DDP-4 Inhibitoren, GLP-1 Rezeptoragonisten oder Basalinsulinen durchgeführt. Patienten mit Diabetes mellitus Typ 2 und hohem kardiovaskulärem

Risiko profitieren hinsichtlich der kardiovaskulären Mortalität von SGLT-2 Inhibitoren und GLP-1 Rezeptoragonisten [15, 48]. Konventionell erfolgt die Steigerung der oralen antidiabetischen Therapie durch eine schrittweise Dosissteigerung des Erstlinienmedikamentes und die anschließende Therapieerweiterung durch ein Kombinationspräparat. Dieses Vorgehen hat offensichtliche Vorteile, wie z. B. das frühzeitige Erkennen von Nebenwirkungen. Es gibt jedoch erste Studienergebnisse, die die Vorteile einer initialen oralen Kombinationstherapie unterstreichen [50].

Auf der vierten Stufe kann eine Kombinationstherapie aus drei Wirkstoffklassen erweitert werden, wenn die individuellen Therapieziele mit einer Zweifachkombinationstherapie nicht erreicht werden[48].

In der letzten und fünften Stufe kann eine Insulintherapie eingeleitet werden, die gegebenenfalls durch ein orales Antidiabetikum ergänzt wird. Sobald die individuellen Therapieziele durch Lebensstiländerungen und orale Antidiabetika nicht mehr erreicht werden können, sollte eine Erweiterung der Therapie mit Insulinen erwogen werden. Der Einsatz von Insulinen muss hinsichtlich der Vor- und Nachteile kritisch abgewogen werden, da die Therapie mit relevanten Nebenwirkungen wie Gewichtszunahme oder gefährlichen Hypoglykämien einhergehen kann. Die konventionelle Insulintherapie beinhaltet ein festes Therapieschema mit lang- und kurzwirksamen Insulinen. Der Vorteil liegt in der einfachen Anwendung für den Patienten. Bei reduzierter Nahrungsaufnahme können jedoch Hypoglykämien auftreten. Patienten mit guter Compliance können eine intensiviertere Insulintherapie erhalten, bei der ein langwirksames Insulin fest appliziert wird und die kurzwirksamen Insuline mahlzeitenabhängig verabreicht werden. In Einzelfällen kann bei ausgeprägter Adipositas permagna ($\text{BMI} \geq 40 \text{ kg/m}^2$) eine bariatrische Operation mit dem Ziel der Reduktion der Nahrungsaufnahme durchgeführt werden [48].

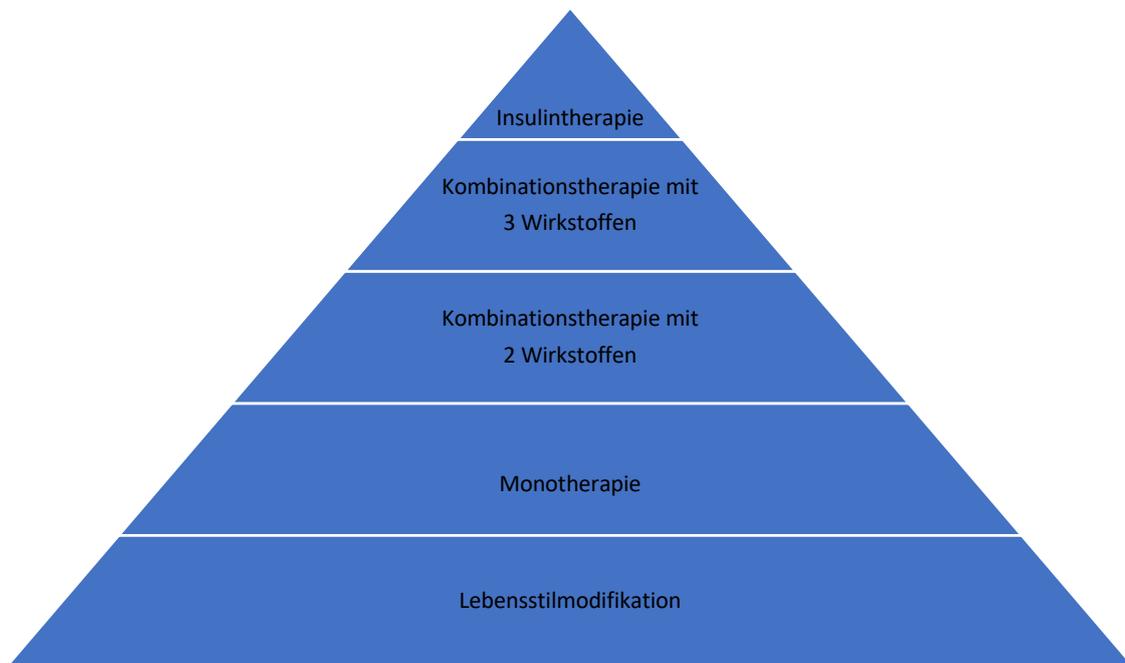


Abb. 2: Stufenschema zur Behandlung des Diabetes mellitus Typ 2 [48]

1.4 Absolute und relative Effektmaße

In der Arzneimittelforschung haben sich randomisierte kontrollierte Studien als Goldstandard durchgesetzt. Für die Darstellung von Behandlungsunterschieden in randomisierten, kontrollierten Studien werden überwiegend nur relative Effektmaße dargestellt [51-53]. Zu diesem Thema analysierten King et al. [51] 344 Artikel aus den zehn führenden Fachzeitschriften des Jahres 2009 zum Themengebiet: Gesundheitliche Ungleichheiten in der Gesellschaft. King et al. [51] konnten in ihrer Arbeit aufzeigen, dass nur 2% (3/344) der Artikel im Abstract und nur 7% (25/344) im Haupttext über relative und absolute Effektmaße berichteten.

Die Studienautoren der CVOTs veröffentlichten die Ergebnisse ebenfalls überwiegend nur mit relativen Effektmaßen, wie der Hazard Ratio. Die Interpretation von Studienergebnissen hat großen Einfluss auf die Entscheidungsfindung von Ärzten und anderen Entscheidungsträgern im Gesundheitswesen. Angemessene Kenntnisse über Nutzen und Risiken sind für eine informierte Entscheidung unerlässlich. In diesem Zusammenhang wurden internationale Leitlinien entwickelt, die die Verwendung absoluter und relativer Effektmaße zur Darstellung von Behandlungsunterschieden betonen, wie z. B. die CONSORT- und GRADE-Leitlinien [54, 55].

Sprenger et al. [56] haben sich in diesem Zusammenhang mit der Bedeutung absoluter Effektmaße und deren Vorteilen gegenüber relativen Effektmaßen aus wissenschaftsphilosophischer Perspektive auseinandergesetzt. Dabei wird das Argument der kognitiven Verzerrung erläutert. Es besagt, dass der Nutzen einer medizinischen Intervention überschätzt werden kann, wenn nur relative Effektmaße in die Entscheidungsfindung einfließen [57]. Je nachdem, welches Basisrisiko in einer Kohorte vorliegt, können sich relative und absolute Effektmaße in ihrer Darstellung erheblich unterscheiden und damit zu Fehlinterpretationen verleiten.

Sprenger et al. [56] untermauern ihre These mit den Studienergebnissen von Bobbio et al. [58]. In dieser Studie legten Bobbio et al. [58] Ärzten die Studienergebnisse von zwei vermeintlich unterschiedlichen Medikamenten vor. In der Studie sollten sich die Ärzte aufgrund der Studienergebnisse für ein Medikament A oder B entscheiden. Die Studienergebnisse wurden in Gruppe A als relative Risikomaße und in Gruppe B als absolute Risikomaße dargestellt. In beiden Gruppen wurden jedoch die Studiendaten desselben Medikaments aus der *Helsinki-Heart* Studie [59] präsentiert. In der *Helsinki-Heart* Studie wurde der Einfluss von Gemfribozil auf die Inzidenz der koronaren Herzerkrankung untersucht. Nach einer Behandlungsdauer von 5 Jahren wurde eine relative Risikoreduktion von 34 % festgestellt. Unter Berücksichtigung des geringen Basisrisikos in der untersuchten Kohorte betrug die absolute Risikoreduktion nur 1,4%. Im Ergebnis entschieden sich die Ärzte deutlich häufiger für die Verschreibung von Medikament A, also der Gruppe mit der relativen Risikoreduktion, obwohl beide Gruppen das gleiche Medikament quantifizierten.

Vergleichbare Studien wurden auch mit Patienten durchgeführt und zeigten ähnliche Entscheidungsmuster [60]. An dieser Stelle ist anzumerken, dass in der Studie von Bobbio et al. [58] den Ärzten keine Informationen über das Basisrisiko zur Verfügung gestellt wurden, so dass aus statistischer Sicht kein sinnvoller Vergleich zwischen relativer und absoluter Risikoreduktion möglich wäre. Die Ergebnisse anderer Studien, in denen die Entscheidungsträger über das Basisrisiko informiert wurden, haben jedoch auch aufzeigen können, dass die alleinige Darstellung der relativen Risikoreduktion anfällig für kognitive Verzerrungen ist [61].

Im klinischen Alltag kann die Interpretation von Studienergebnissen mühselig sein und Ärzte werden dazu verleitet, die Interpretation von Studiendaten auf wenige prägnante Informationen zu stützen. Ferrannini et al. [62] stellen diese Problematik in den Kontext der CVOTs und betonen ebenfalls die starke Abhängigkeit der relativen Risikoreduktion vom bestehenden Basisrisiko und der damit veränderten Wahrnehmung der Medikamentenwirksamkeit. Ferrannini et al. [62] betonen in ihrer Arbeit erneut, dass für die realistische Einschätzung der Studienergebnisse die alleinige Information der relativen Risikoreduktion nicht ausreichend ist.

Zur Veranschaulichung dieses Zusammenhangs ist das folgende Beispiel ausgeführt:

In einer Kohorte mit einer sehr hohen Basisrisikorate (100 Ereignisse pro 1.000 Patientenjahre) entspricht eine relative Risikoreduktion von 50% einer absoluten Risikoreduktion von 50 Ereignissen pro 1.000 Patientenjahre. Die absolute Risikoreduktion beträgt somit 5 Prozentpunkte pro Jahr. Bei einer niedrigeren Basisrisikorate (20 Ereignisse in 1.000 Patientenjahren) und der gleichen relativen Risikoreduktion von 50% ergibt sich eine absolute Risikoreduktion von 10 Ereignissen in 1.000 Patientenjahren. Die absolute Risikoreduktion beträgt also 1 Prozentpunkt pro Jahr statt 5 Prozentpunkte pro Jahr wie im obigen Beispiel.

Der Kehrwert der absoluten Risikoreduktion, die Number-needed-to-treat (NNT), beschreibt die Anzahl der Personen, die im Interventionsarm der Studie behandelt werden müssen, um ein Ereignis zu verhindern. In den obigen Beispielen würde die NNT im ersten Beispiel 20 und im zweiten Beispiel 100 Personen betragen, obwohl in beiden Kohorten eine relative Risikoreduktion von 50% vorliegt. Daher wird die Darstellung von relativen und absoluten Effektmaße empfohlen, um Therapieentscheidungen auf einer guten Informationsbasis treffen zu können.

1.5 Ziele der Arbeit

Die Darstellung der Studienergebnisse kann die Interpretation erheblich beeinflussen, daher wird, wie in den CONSORT- und GRADE-Leitlinien [54, 55], die Angabe von relativen und absoluten Effektmaßen empfohlen. Die Autoren der *Cardiovascular Outcomes Trials* zu den neuen Antidiabetika berichten die Behandlungseffekte überwiegend in Form von relativen Effektmaßen. Absolute Effektmaße wie die Number-needed-to-treat (NNT) werden dagegen kaum berichtet. Die zentralen Ziele und Fragestellungen dieser Dissertation sind wie folgt definiert:

- 1) Im Rahmen dieser Arbeit werden die absoluten Behandlungseffekte und die Number-needed-to-treat für den primären Endpunkt und für die Gesamtmortalität der eingeschlossenen CVOTs geschätzt. Welche klinische Interpretation und Schlussfolgerungen ergeben sich aus den Ergebnissen?
- 2) Im Rahmen dieser Arbeit werden Meta-NNTs für die drei relevanten Wirkstoffklassen und Meta-NNTs für den primären Endpunkt und die Gesamtmortalität gebildet. Gibt es Wirkstoffklasseneffekte für die Meta-NNT?
- 3) Im Rahmen dieser Arbeit wird eine Methode zur Schätzung der absoluten Risikoreduktion und der NNT in Endpunktstudien aus digitalisierten, individuellen Patientendaten entwickelt. Welche Möglichkeiten stehen zur Verfügung, um die Validität zu beurteilen?
- 4) Welche Empfehlungen für die Darstellung und Interpretation von Studiendaten in Endpunktstudien lassen sich aus dieser Arbeit ableiten?

2 Material und Methoden

2.1 Material

Für die Arbeit wurden folgende Materialien verwendet:

- Computer (Betriebssystem: Windows)
- Kaplan-Meier-Überlebenszeitkurven aus den Voll- und Ergänzungstexteten der CVOTs (Tabelle 1)
- WebPlotDigitizer (Version 4.2) [63]
- R-Algorithmus von Guyot et al. [64]
- Programmiersprache R und die Benutzeroberfläche R-Studio
- SAS (SAS Institute Inc., Cary, NC, USA), Version 9.4

2.2 Methoden

2.2.1 Studiendesign

In der vorliegenden Arbeit wurde eine Meta-Analyse der digitalisierten, individuellen Patientendaten aus den *Cardiovascular Outcome Trials* durchgeführt. Die Auswahl der einbezogenen CVOTs erfolgte auf der Grundlage der Übersichtsarbeit von Cefalu et al. [1]. Zur Datenerhebung wurden im September 2020 alle zur Verfügung stehenden Voll- und Ergänzungstexte zu den CVOTs gesichtet und die Kaplan-Meier-Überlebenszeitkurven überprüft. Der Schwerpunkt wurde auf die Wirkstoffklassen DDP-4 Inhibitoren, SGLT-2 Inhibitoren und GLP-1 Rezeptoragonisten gelegt. Die Wirkstoffklassen Insuline, Glitazone und Alpha-Glukosidase Inhibitoren sind in der Meta-Analyse nicht berücksichtigt worden. Studien ohne Placebo-Kontrollgruppe wurden ebenfalls ausgeschlossen. Entsprechend der festgelegten Ein- und Ausschlusskriterien konnten vier CVOTs mit DDP-4 Inhibitoren, acht CVOTs mit SGLT-2-Inhibitoren und sieben CVOTs mit GLP-1 Rezeptoragonisten einbezogen werden (Tabelle 1). Insgesamt wurden 19 publizierte CVOTs in diese Arbeit eingeschlossen. Von diesen 19 CVOTs konnten alle primären Endpunkte eingeschlossen werden, jedoch nur 13 für den Endpunkt Gesamtmortalität (Tabelle 1). In 6 CVOTs wurden die Endpunkte zur

Gesamtmortalität beschrieben, jedoch nicht graphisch in Form einer Kaplan-Meier-Überlebenszeitkurve dargestellt.

Die Kaplan-Meier-Überlebenszeitkurven wurden unter Zuhilfenahme des Programms WebPlotDigitizer (Version 4.2) [63] digitalisiert und unter Verwendung des R-Algorithmus von Guyot et al. [64] konnten die digitalisierten, individuellen Patientendaten geschätzt werden. Beide Verfahren haben sich bereits in früheren Arbeiten als reliabel und valide erwiesen [65-67]. Zusätzlich wurden die originalen Hazard Ratios für die beiden untersuchten Endpunkte mit dem 95%-Konfidenzintervall extrahiert. Mit den geschätzten individuellen Patientendaten wurden parametrische Regressionsmodelle gebildet und anschließend zeitabhängige absolute Risikodifferenzen berechnet. Aus den zeitabhängigen absoluten Risikodifferenzen wurden die Number-needed-to-treat für die untersuchten Endpunkte berechnet.

2.2.2 Digitalisierung und Datenrekonstruktion

In den eingeschlossenen CVOTs wurden die Behandlungseffekte in Form von Kaplan-Meier-Überlebenszeitkurven graphisch dargestellt. Zur weiteren Verarbeitung und Extraktion der darin enthaltenen Informationen mussten diese digitalisiert werden.

Die Digitalisierung erfolgte mit der Software WebPlotDigitizer [63], welche aufgrund ihrer guten Reliabilität und Validität in den Analysen von Drevon et al. [66] und Moeyaert et al. [65] ausgewählt wurde. Die Digitalisierung mit dem WebPlotDigitizer [63] erfolgte nach einem zuvor festgelegten Verfahren. Die Kaplan-Meier-Überlebenszeitkurven wurden aus den Voll- oder Ergänzungstexten digital ausgeschnitten und in ein PDF-Dokument konvertiert. Das PDF-Dokument wurde in die Software WebPlotDigitizer [63] geladen.

Im ersten Schritt wurden die Koordinatenpunkte für den Nullpunkt sowie für die X- und Y-Achse festgelegt. Anschließend wurden die Markierungspunkte für die Behandlungs- und Placebogruppe erstellt. Für die automatische Digitalisierung der Kurven konnten die Parameter Bildausschnitt, detektierte Farbe und Pixeldichte der Markierungspunkte festgelegt werden. Anschließend erfolgte die automatische Platzierung der Markierungspunkte durch die Software. In den Fällen, in denen das automatische Setzen der Markierungspunkte nicht möglich war, z. B. aufgrund einer zu geringen

Pixelauflösung der Originalgrafiken oder einer zu großen Überlappung der Kurven, konnten die Markierungspunkte manuell gesetzt werden. Schließlich konnten zwei separate Datensätze mit den Koordinatenpunkten als CSV-Dateien für die Behandlungsgruppe und die Placebogruppe heruntergeladen werden.

Für die Schätzung der absoluten Behandlungseffekte für den primären Endpunkt und die Gesamtmortalität wurden die individuellen *Time-to-Event* Daten aus den CVOTs benötigt. Da diese Daten in der Regel nicht in der Primärliteratur veröffentlicht werden, standen sie für die Meta-Analyse nicht zur Verfügung. Um dennoch die notwendigen *Time-to-Event* Daten zu ermitteln, wurde ein von Guyot et al. [64] entwickelter R-Algorithmus verwendet. In einer Übersichtsarbeit von Saluja et al. [67] wurden verschiedene Methoden zur Schätzung von *Time-to-Event* Daten hinsichtlich ihrer Validität verglichen. In dieser Analyse zeichnete sich der R-Algorithmus von Guyot et al. [64] mit einer sehr geringen Abweichung von den Originaldaten aus.

Für die Implementierung des Algorithmus von Guyot et al. [64] wurde die Programmiersprache R verwendet. Für die Durchführung wurden die Datensätze der digitalisierten Kaplan-Meier-Überlebenszeitkurven, die Anzahl der Risikopersonen zu mehreren Zeitpunkten und die Gesamtzahl der Ereignisse in beiden Behandlungsarmen am Ende der Beobachtungszeit benötigt. Anschließend konnten mit Hilfe des Algorithmus die individuellen *Time-to-Event* Daten geschätzt werden.

2.2.3 Statistische Analyse

In der vorliegenden Arbeit wurden Kaplan-Meier-Überlebenszeitkurven für den primären Endpunkt und die Gesamtmortalität für alle bis September 2020 veröffentlichten *Cardiovascular Outcome Trials* (CVOTs) untersucht.

Zur Schätzung der absoluten Risikoreduktion in einer *Time-to-Event* Studie müssen für beide Behandlungsgruppen getrennte Überlebensfunktionen geschätzt werden. Da dies mit einem proportionalen Cox-Wahrscheinlichkeitsmodell nicht möglich ist, wurde ein Weibull-Regressionsmodell verwendet. Weibull-Modelle sind parametrische, proportionale Hazard-Modelle, mit denen Weibull-Hazard-Verhältnisse geschätzt werden können. Diese Weibull-Hazardraten können mit den ursprünglichen Hazardraten verglichen werden, um die Validität der Methode zu überprüfen [68].

Für die beiden untersuchten Endpunkte und jede eingeschlossene Studie wurden die monatlichen Wahrscheinlichkeitsdifferenzen zwischen Behandlungs- und Placebogruppe mit einem parametrischen Weibull-Modell geschätzt. Sowohl für die Behandlungs- als auch für die Placebogruppe wurde die Wahrscheinlichkeit für das Ausbleiben des Endpunkts vom ersten Monat bis zum maximalen Beobachtungsmontat berechnet. Die Wahrscheinlichkeitsdifferenz für das Ausbleiben des Endpunkts zwischen der Behandlungs- und der Placebogruppe bezeichnet die absolute Risikoreduktion. Durch die Inversion der absoluten Risikoreduktion kann die Number-needed-to-treat für jeden Monat berechnet werden. Für die Darstellung der absoluten Behandlungseffekte der Substanzklassen wurden standardisierte Inverse-Varianz-Methoden mit festen Effekten verwendet [68].

Ein Beispiel für die Berechnung der NNT durch Inversion der absoluten Risikoreduktion in Bezug auf die Gesamtmortalität in der EMPA-REG-Studie ist in Abbildung 3 dargestellt. Auf der X-Achse erfolgte die zeitliche Darstellung anhand der Behandlungsmonate. Auf der Y-Achse ist die Überlebenswahrscheinlichkeit aufgetragen.

Die blaue Kurve stellt den graphischen Verlauf der Überlebenswahrscheinlichkeit für die Behandlungsgruppe (Empagliflozin) und die rote Kurve für die Placebogruppe dar. In beiden Gruppen nimmt die Überlebenswahrscheinlichkeit mit der Zeit ab. Für die Berechnung der NNT nach 36 Behandlungsmonaten wurde die Differenz der Überlebenswahrscheinlichkeiten voneinander subtrahiert ($p = 0,94478$ in der Behandlungsgruppe (Empagliflozin) und $p = 0,91912$ in der Placebogruppe; $0,94478 - 0,91912 = 0,02566$). Der Kehrwert der absoluten Risikodifferenz entspricht der Number-needed-to-treat für den gegebenen Zeitpunkt ($NNT = 1 / \text{Risikodifferenz} = 1 / 0,02566 = 38,97116$). Für den Zeitraum nach 36 Behandlungsmonaten ergibt sich somit für Empagliflozin eine NNT von ca. 39.

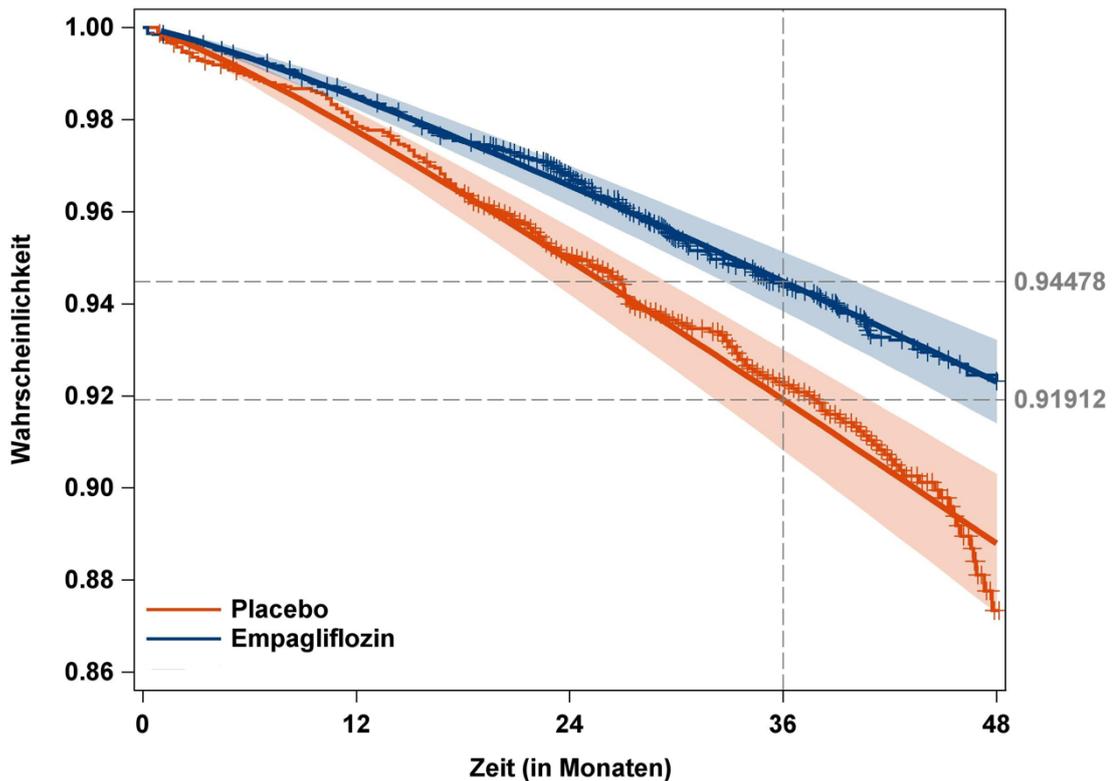


Abb. 3: Weibull-Überlebenszeitkurven für den Endpunkt Gesamtmortalität in der EMPA-REG Studie [69]. Auf der X-Achse ist die Zeit in Monaten dargestellt und auf der Y-Achse die Überlebenswahrscheinlichkeit.

2.2.4 Meta-Analyse

Für alle eingeschlossenen CVOTs wurden die NNT für den primären Endpunkt und die Gesamtmortalität berechnet und in den 3 Wirkstoffklassen meta-analytisch zusammengefasst. Für die Meta-Analyse wurde eine inverse Varianzmethode verwendet. Der Anteil einer einzelnen Studie innerhalb der Meta-Analyse wurde über den Kehrwert der Varianz bestimmt. Auf diese Weise werden größere Studien stärker gewichtet, da sie in der Regel geringere Standardfehler aufweisen. Darüber hinaus wurde für die Meta-Analyse ein Modell mit zufälligen Effekten verwendet, bei dem angenommen wird, dass die Behandlungseffekte zwischen den Studien variieren können. Alle Berechnungen wurden auf der Wahrscheinlichkeitsdifferenz-Skala durchgeführt und nur für die Darstellung der Ergebnisse in Abbildungen und Diagrammen in die NNT-Skala transformiert. Für die Datenverarbeitung und -analyse wurde SAS (SAS Institute Inc., Cary, NC, USA) Version 9.4, verwendet.

Die vollständigen Datensätze sind in einer öffentlichen Datenbank verfügbar und enthalten die vollständigen digitalisierten Patientendaten [70].

2.2.5 Lebenszeitperspektive

Um die Perspektive der Langzeitwirkung der untersuchten Arzneimittel abzubilden, wurden für die Gesamtmortalität die Number-needed-to-treat bis zu einem Zeitpunkt von 360 Behandlungsmonaten (30 Behandlungsjahre) unter der Annahme einer kontinuierlichen Behandlung berechnet. Der Endpunkt Gesamtmortalität wurde ausgewählt, da er im Gegensatz zu anderen Endpunkten keine konkurrierenden Endpunkte aufweist. Für den primären Endpunkt nimmt die Relevanz der konkurrierenden Endpunkte (insbesondere Tod) mit zunehmender Behandlungsdauer zu, da die Wahrscheinlichkeit anderer konkurrierender Endpunkte mit der Zeit zunimmt. Alle Schätzungen basieren auf der Annahme, dass die Patienten dieselbe Behandlung über den langen Zeitraum beibehalten und der Behandlungseffekt nach Abschluss der Studie konstant bleibt.

2.2.6 Validität der Methode

Die Validität der einzelnen Analyseschritte (Digitalisierung mit dem WebPlotDigitizer [63], Berechnung der *Time-to-Event* Daten mit dem R-Algorithmus von Guyot et al. [64] und die Umrechnung in Weibull-Hazardraten [68]) konnte bereits in früheren Arbeiten nachgewiesen werden. Um die Validität der Kombination der einzelnen Analyseschritte zu überprüfen, wurden die Weibull-Hazardraten mit den Hazardraten der Originaldaten verglichen und Intra-Class-Korrelationskoeffizienten berechnet. Die Intra-Class-Korrelationskoeffizienten konnten graphisch mit Streudiagrammen in Abbildung 10 dargestellt werden. Darüber hinaus konnte die Validität der angewandten Methoden durch den Vergleich mit anderen Arbeiten, in denen die NNT von CVOTs mit anderen analytischen Ansätzen bestimmt wurden, überprüft werden.

2.2.7 Definition und Interpretation der Number-needed-to-treat (NNT)

Die Number-needed-to-treat wurde erstmals von Laupacis et al. [71] 1988 zur Beschreibung des Nutzens einer therapeutischen Intervention eingeführt. Die NNT gibt die Anzahl der Patienten an, die in einem bestimmten Zeitraum behandelt werden müssen, um ein Ereignis oder einen Endpunkt in der Behandlungsgruppe zu verhindern. Ein positiver Wert für die NNT bedeutet, dass die Behandlung oder die Intervention im Vergleich zur Kontrollgruppe einen Vorteil bietet. Ein negativer Wert für die NNT bedeutet, dass die Behandlung oder die Intervention im Vergleich zur Kontrollgruppe einen Nachteil hat. Im Falle eines negativen Effekts des Interventionsarms kann auch von einem Number-needed-to-harm (NNH) gesprochen werden. Ein niedriger Wert für die NNT zeigt einen großen Behandlungseffekt für das untersuchte Ereignis an, ein hoher Wert einen kleinen Behandlungseffekt. Ein NNT-Wert von 1 bedeutet, dass jeder Patient in Bezug auf das untersuchte Ereignis von der Behandlung profitiert. Ein absoluter Behandlungseffekt ist signifikant, wenn das 95%-Konfidenzintervall nicht unendlich (∞) [oder - unendlich ($-\infty$)] umfasst (mit einer zweiseitigen α -Wahrscheinlichkeit von 5%).

2.3 Ethikvotum

Für die Dissertation wurde im Juli 2019 ein Ethikvotum bei der Ethikkommission der Medizinischen Fakultät der Heinrich-Heine-Universität Düsseldorf beantragt. Die Kommission konnte bestätigen, dass für die Dissertation kein Ethikvotum erforderlich ist, da in der Arbeit ausschließlich bereits publizierte Daten verwendet werden.

3. Ergebnisse

3.1 Studienauswahl

Die Auswahl der Studien basiert auf der Arbeit von Cefalu et al.[1], die einen Überblick über alle geplanten und abgeschlossenen CVOTs zwischen 2013 und 2020 bietet. Die Datenerhebung für diese Arbeit wurde im September 2020 unter Berücksichtigung der Ein- und Ausschlusskriterien abgeschlossen. Insgesamt wurden 19 CVOTs eingeschlossen und die *Time-to-event* Daten zum primären Endpunkt und zur Gesamtmortalität ermittelt. Die *Carolina* Studie wurde aufgrund des Fehlens einer Placebo-Kontrollgruppe ausgeschlossen. Die ebenfalls im Review von Cefalu et al. [1] beschriebenen Studien FREEDOM-CVO und EMPEROR-PRESERVED waren zum Zeitpunkt der Datenerhebung noch nicht publiziert.

Die CVOTs können anhand der untersuchten Medikamente in 3 Wirkstoffklassen eingeteilt werden. Für die Meta-Analyse wurden 4 CVOTs mit DDP-4 Inhibitoren (CARMELINA [72], EXAMINE [73], SAVOR-TIMI-53 [74], TECOS [75]), 7 CVOTs mit GLP-1 Rezeptoragonisten (ELIXA [76], EXSCEL [77], HARMONY [78], LEADER [79], PIONEER-6 [80], REWIND [81], SUSTAIN-6 [82]) und acht CVOTs mit SGLT-2 Inhibitoren (CANVAS [83], CREDENCE [84], DAPA-CKD [85], DAPA-HF [86], DECLARE-TIMI-58 [87], EMPA-REG [13], EMPEROR-REDUCED [88], VERTIS-CV [89]) eingeschlossen.

3.2 Studiencharakteristika

Tabelle 1 gibt einen Überblick über die Charakteristika der eingeschlossenen CVOTs und die Definitionen des jeweiligen primären Endpunkts in jeder CVOT. Für die Meta-Analyse des primären Endpunkts wurden Daten aus 19 CVOTs von 159.265 untersuchten Patienten mit 17.501 Ereignissen und für die Meta-Analyse der Gesamtmortalität Daten aus 13 CVOTs von 112.524 untersuchten Patienten mit 8.888 Ereignissen eingeschlossen. Die mediane Nachbeobachtungszeit beträgt für den primären Endpunkt 28,7 Monate und 39,3 Monate für die Gesamtmortalität. Für alle primären Endpunkte der CVOTs lagen die Hazard Ratios zwischen 0,61 und 1,02 mit einer medianen Hazard Ratio von 0,87. Für die Gesamtmortalität wurden Hazard Ratios zwischen 0,68 und 1,01

mit einer medianen Hazard Ratio von ebenfalls 0,87 ermittelt. Alle eingeschlossenen Studien weisen eine hohe Anzahl an eingeschlossenen Patienten auf. Die Studie PIONEER-6 [80] weist mit 3.183 die geringste und die Studie DECLARE-TIMI-58 [87] mit 17.160 die höchste Anzahl an Patienten auf. Das mittlere Alter in den eingeschlossenen Studien lag zwischen 60,3 und 66,8 Jahren. Das Geschlechterverhältnis war in allen Studien zugunsten der männlichen Studienteilnehmer verschoben. Der geringste Anteil männlicher Studienteilnehmer fand sich in der REWIND-Studie [81] mit 50,3 %, der höchste in der DAPA-HF-Studie [86] mit 77,7 %. Die mittleren HbA1c-Werte lagen zwischen 7,2 % (TECOS-Studie [75]) und 8,7 % (HARMONY [78], LEADER [79] und SUSTAIN-6 [82]). Für die Studien DAPA-HF [86], DAPA-CKD [85] und EMPEROR-REDUCED [88] wurden keine Daten zum medianen HbA1c-Wert veröffentlicht. Der niedrigste mittlere Body-Mass-Index (BMI) lag in der EMPEROR-REDUCED-Studie [88] bei 27,9 kg/m², der höchste in der EXAMINE-Studie [73] bei 33,8 kg/m². In einigen Studien wurden nur Mediane und/oder Quartile und/oder Minima/Maxima angegeben, statt Mittelwerte und Standardabweichungen. In diesen Fällen wurde die Formel von Wan et al. [90] zur Berechnung der entsprechenden Mittelwerte und Standardabweichungen verwendet.

Studie (Wirkstoff)	Fachjournal (Jahr)	Anzahl der beobachteten Patienten	Mittelwert Alter (SD) in Jahren	Anteil Männer (%)	Mittelwert (SD) HbA1c (%)	Mittelwert (SD) BMI (kg/m ²)	Mittelwert (SD) Diabetes-Dauer (Jahre)	Definition: Kombiniertes primäres Endpunkt
DDP-4 Inhibitor								
CARMELINA [72] (Linagliptin)	JAMA (2018)	6.991	65,9 (9,1)	62,9	7,9 (1,0)	31,4 (5,3)	14,8 (9,5)	Kardiovaskuläre Mortalität, nicht-tödlicher Myokardinfarkt, nicht-tödlicher Schlaganfall
EXAMINE [73] (Alogliptin)	NEJM (2013)	5.380	61,0 (--)	67,8	8,0 (1,1)	33,8 (6,6)	7,7 (8,2)	Kardiovaskuläre Mortalität, nicht-tödlicher Myokardinfarkt, nicht-tödlicher Schlaganfall
SAVOR-TIMI53 [74] (Saxagliptin)	NEJM (2013)	16.492	65,1 (8,5)	66,9	8,0 (1,4)	31,1 (5,6)	10,7 (8,7)	Kardiovaskuläre Mortalität, nicht-tödlicher Myokardinfarkt, nicht-tödlicher Schlaganfall
TECOS [75] (Sitagliptin)	NEJM (2015)	14.671	65,5 (8,0)	70,7	7,2 (0,5)	30,2 (5,7)	11,6 (8,1)	Kardiovaskuläre Mortalität, nicht-tödlicher Myokardinfarkt, nicht-tödlicher Schlaganfall, Hospitalisierung bei instabiler Angina pectoris
GLP-1 Rezeptoragonist								
ELIXA [76] (Lixisenatide)	NEJM (2015)	6.068	60,3 (9,7)	69,4	7,7 (1,3)	30,2 (5,7)	9,3 (8,3)	Kardiovaskuläre Mortalität, nicht-tödlicher Myokardinfarkt, nicht-tödlicher Schlaganfall, Hospitalisierung bei instabiler Angina pectoris
EXSCEL [77] (Exenatide)	NEJM (2017)	14.752	62,0 (8,9)	62,0	8,1 (1,2)	32,0 (5,9)	12,2 (7,8)	Kardiovaskuläre Mortalität, nicht-tödlicher Myokardinfarkt, nicht-tödlicher Schlaganfall
HARMONY [78] (Albiglutide)	Lancet(2018)	9.463	64,2 (8,7)	69,4	8,7 (1,5)	32,3 (5,9)	14,2 (8,8)	Kardiovaskuläre Mortalität, nicht-tödlicher Myokardinfarkt, nicht-tödlicher Schlaganfall
LEADER [79] (Liraglutide)	NEJM (2016)	9.340	64,3 (7,2)	64,2	8,7 (1,5)	32,5 (6,3)	12,9 (8,1)	Kardiovaskuläre Mortalität, nicht-tödlicher Myokardinfarkt, nicht-tödlicher Schlaganfall
PIONEER-6 [80] (Semaglutide)	NEJM (2019)	3.183	66,0 (7,0)	68,4	8,2 (1,6)	32,3 (6,5)	14,9 (8,5)	Kardiovaskuläre Mortalität, nicht-tödlicher Myokardinfarkt, nicht-tödlicher Schlaganfall

REWIND [81] (Dulaglutide)	Lancet(2019)	9.901	66,2 (6,5)	53,7	7,4 (1,1)	32,3 (5,8)	10,6 (7,2)	Kardiovaskuläre Mortalität, nicht-tödlicher Myokardinfarkt, nicht-tödlicher Schlaganfall
SUSTAIN-6 [82] (Semaglutide)	NEJM (2016)	3.297	64,6 (7,4)	60,7	8,7 (1,5)	32,8 (6,2)	13,9 (8,1)	Kardiovaskuläre Mortalität, nicht-tödlicher Myokardinfarkt, nicht-tödlicher Schlaganfall
SGLT-2 Inhibitor								
CANVAS [83] (Canagliflozin)	NEJM (2017)	10.142	63,3 (8,3)	64,2	8,2 (0,9)	31,9 (5,9)	13,6 (7,7)	Anhaltende Verringerung der eGFR um mindestens 40 %, Einleitung einer Nierenersatztherapie, renale Todursache
CREDENCE [84] (Canagliflozin)	NEJM (2019)	4.401	63,0 (9,2)	66,0	8,3 (1,3)	31,4 (6,2)	15,7 (8,7)	Nierenerkrankung im Endstadium, Verdopplung des Serumkreatinin, renale oder kardiovaskuläre Mortalität
DAPA-CKD [85] (Dapagliflozin)	NEJM (2020)	4.304	61,9 (12,2)	66,9	-	29,5 (6,2)	-	Anhaltender Rückgang der eGFR um mindestens 50 %, terminale Niereninsuffizienz, renale oder kardiovaskuläre Mortalität
DAPA-HF [86] (Dapagliflozin)	NEJM (2019)	4.744	66,3 (10,9)	77,7	-	28,2 (6,0)	-	Kardiovaskuläre Mortalität, Verschlechterung der Herzinsuffizienz (Krankenhausaufenthalt oder intravenöse Therapie der Herzinsuffizienz),
DECLARE-TIMI-58 [87] (Dapagliflozin)	NEJM (2018)	17.160	63,9 (6,8)	62,6	8,3 (1,2)	32,1 (6,0)	10,9 (7,4)	Kardiovaskuläre Mortalität, nicht-tödlicher Myokardinfarkt, nicht-tödlicher Schlaganfall
EMPA-REG [13] (Empagliflozin)	NEJM (2015)	7.020	63,1 (8,7)	71,5	8,1 (0,8)	30,6 (5,2)		Kardiovaskuläre Mortalität, nicht-tödlicher Myokardinfarkt, nicht-tödlicher Schlaganfall
EMPEROR- REDUCED [88] (Empagliflozin)	NEJM (2020)	3.730	66,8 (11)	76,0	-	27,9 (5,4)	-	Kardiovaskuläre Mortalität, Hospitalisierung aufgrund einer Herzinsuffizienz
VERTIS-CV [89] (Ertugliflozin)	NEJM (2020)	8.246	64,4 (8,1)	70,0	8,2 (1,0)	31,9 (5,4)	13,0 (8,3)	Kardiovaskuläre Mortalität, nicht-tödlicher Myokardinfarkt, nicht-tödlicher Schlaganfall

Tabelle 1: **Übersicht der Studiencharakteristika und die Definition des primären Endpunkts [69].** (Abkürzungen: SD = Standardabweichung, BMI = Body Mass Index, eGFR= geschätzte glomeruläre Filtrationsrate)

3.3 NNT-Ergebnisse nach 12, 24, 36 und 48 Behandlungsmonaten

Die Tabellen 2 und 3 zeigen die NNT-Ergebnisse für 12, 24, 36 und 48 Behandlungsmonate für jede eingeschlossene CVOT und getrennt für den primären Endpunkt und die Gesamtmortalität. Die NNTs wurden für jeden Zeitpunkt separat mit dem parametrischen Weibull-Modell berechnet. Alle Berechnungen wurden auf einer Wahrscheinlichkeitsskala durchgeführt und erst für die graphische oder tabellarische Darstellung in die NNT umgerechnet. Für ein besseres Verständnis der zeitlichen Entwicklung der NNT ist die graphische Darstellung besser geeignet. Abbildung 4 zeigt die NNT-Ergebnisse für die primären Endpunkte von 19 CVOTs und Abbildung 5 die NNT-Ergebnisse für die Gesamtmortalität von 13 CVOTs. In beiden Abbildungen ist auf der x-Achse die Zeit in Monaten und auf der y-Achse die NNT mit einer logarithmischen Skala aufgetragen. Alle NNT-Schätzungen und Konfidenzintervalle wurden aus Gründen der Übersichtlichkeit ab einem NNT-Wert von 100.000 abgeschnitten. Die farbige Schattierung auf jeder Kurve gibt das 95%-Konfidenzintervall an. Zur besseren Übersicht wurden die drei Wirkstoffklassen farblich getrennt dargestellt. Die CVOTs mit einem DDP-4 Inhibitor sind blau, die CVOTs mit einem GLP-1-Rezeptoragonisten sind gelb und alle CVOTs mit einem SGLT-2 Inhibitor sind rot. Die Kurvenform ist für die meisten CVOTs und für beide Endpunkte ähnlich. Die NNT ist zu Beginn der Behandlung hoch und sinkt mit zunehmender Behandlungsdauer, wobei die Dynamik der NNT-Senkung mit der Zeit abflacht. Für den primären Endpunkt zeigen die NNT-Kurven für die Studien CARMELINA und ELIXA eine horizontale Linie, da die NNT zu keinem Zeitpunkt unter den obersten Punkt der Skalierung auf der Y-Achse von 100.000 fällt.

Dasselbe kann in der TECOS-Studie für die Gesamtmortalität beobachtet werden. Die anderen Abbildungen zeigen den oben beschriebenen Kurvenverlauf. Die Abbildungen 4 und 5, in denen die NNT-Ergebnisse für die primären Endpunkte und die Gesamtmortalität dargestellt sind, zeigen trotz ähnlicher Kurvenverläufe große Unterschiede hinsichtlich der erreichten NNT und des zugehörigen Konfidenzintervalls. Hierbei ist zu beachten, dass sich die Definitionen des primären Endpunktes leicht unterscheiden.

Studien	NNT (95%-KI) nach 12 Monaten	NNT (95%-KI) nach 24 Monaten	NNT (95%-KI) nach 36 Monaten	NNT (95%-KI) nach 48 Monaten
Primärer Endpunkt				
DDP-4 Inhibitor				
CARMELINA	- 898 (181 - [- 129])	- 446 (90 - [- 64])	- 304 (62 - [- 44])	
EXAMINE	255 (58 - [- 107])	167 (38 - [- 70])		
SAVOR-TIMI-53	6.765 (237 - [- 255])	3.487 (122 - [- 131])		
TECOS	1.026 (207 - [- 347])	540 (109 - [- 183])	378 (76 - [- 128])	296 (60 - [- 100])
GLP-1-Rezeptoragonist				
ELIXA	- 1.816 (108 - [- 97])	- 1.112 (66 - [- 59])	- 849 (51 - [- 45])	
EXSCEL	326 (159 - [- 7.005])	162 (79 - [- 3.577])	110 (54 - [- 2.433])	84 (41 - [- 1.867])
HARMONY	83 (52 - 197)	43 (27 - 103)		
LEADER	207 (117 - 881)	103 (58 - 432)	69 (39 - 291)	53 (30 - 223)
PIONEER-6	136 (55 - [- 299])			
REWIND	434 (227 - 4.938)	198 (104 - 2.163)	127 (67 - 1.366)	93 (49 - 999)
SUSTAIN-6	80 (44 - 464)			
SGLT-2 Inhibitor				
CANVAS	229 (119 - 3.170)	125 (65 - 1.732)	89 (46 - 1.236)	70 (36 - 981)
CREDENCE	103 (70 - 191)	35 (24 - 63)	19 (13 - 35)	
DAPA-CKD	58 (43 - 92)	20 (15 - 31)		
DAPA-HF	28 (20 - 52)	17 (12 - 30)		
DECLARE-TIMI-58	683 (266 - [- 1.208])	343 (134 - [- 609])	232 (91 - [- 412])	
EMPA-REG	170 (89 - 1.733)	84 (44 - 822)	57 (30 - 545)	43 (23 - 413)
EMPEROR-REDUCED	22 (15 - 43)	14 (10 - 27)		
VERTIS-CV	549 (147 - [- 317])	280 (75 - [- 162])	192 (52 - [- 111])	149 (40 - [- 86])

Tabelle 2: Ergebnisse der Number-needed-to-treat für den primären Endpunkt nach 12, 24, 36 und 48 Behandlungsmonaten [69].

Studien	NNT (95%-KI) nach 12 Monaten	NNT (95%-KI) nach 24 Monaten	NNT (95%-KI) nach 36 Monaten	NNT (95%-KI) nach 48 Monaten
Gesamtmortalität				
DDP-4 Inhibitor				
CARMELINA	1.070 (165 - [- 239])	467 (72 - [- 104])	296 (46 - [- 66])	
EXAMINE	223 (75 - [- 230])	129 (43 - [- 133])		
TECOS	- 9.238 (445 - [- 406])	- 4.221 (204 - [- 186])	- 2.703 (130 - [- 119])	- 1.988 (96 - [- 87])
GLP-1-Rezeptoragonist				
EXSCEL	497 (271 - 2.948)	202 (110 - 1.168)	121 (66 - 694)	85 (46 - 486)
LEADER	423 (232 - 2.444)	170 (93 - 943)	101 (56 - 555)	71 (39 - 387)
REWIND	740 (372 - 57.774)	288 (145 - 15.415)	168 (85 - 8.117)	116 (58 - 5.383)
SGLT-2 Inhibitor				
CANVAS	475 (237 - [- 96.864])	210 (105 - [- 50.993])	131 (66 - [- 27.830])	95 (47 - [- 17.270])
CREDENCE	252 (123 - [- 5.131])	96 (47 - [- 2.356])	56 (27 - [- 1.405])	
DAPA-CKD	137 (82 - 419)	55 (33 - 163)		
DAPA-HF	76 (40 - 735)	37 (19 - 353)		
DECLARE-TIMI58	2.019 (596 - [- 1.455])	808 (239 - [- 583])	477 (141 - [- 344])	
EMPA-REG	137 (91 - 280)	62 (41 - 122)	39 (26 - 77)	28 (19 - 56)
EMPEROR-REDUCED	138 (44 - [-124])	69 (22 - [- 62])		

Tabelle 3: Ergebnisse der Number-needed-to-treat für die Gesamtmortalität nach 12, 24, 36 und 48 Behandlungsmonaten [69].

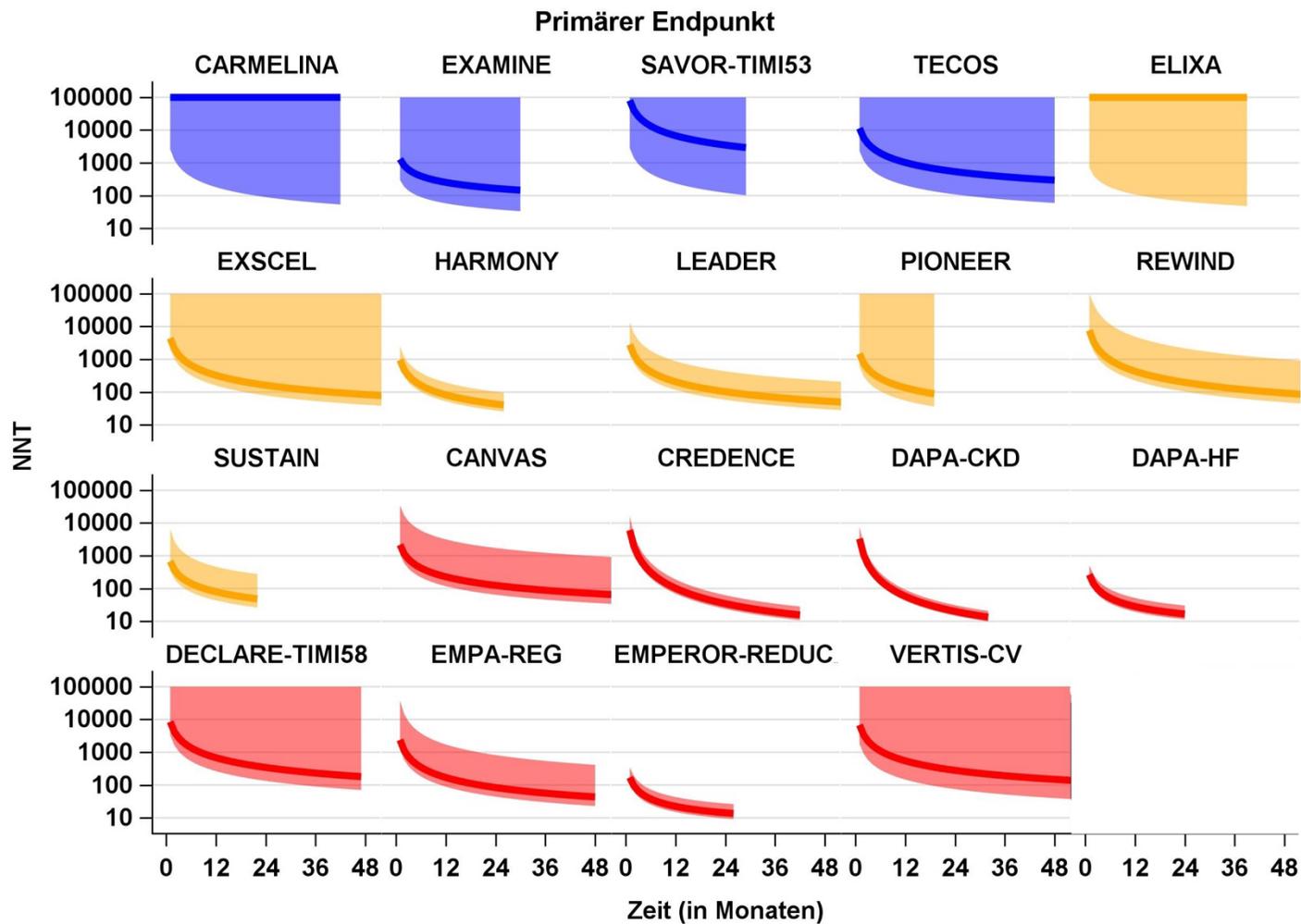


Abb. 4: Graphische Darstellung der Number-needed-to-treat für den primären Endpunkt (blau: DPP-4 Inhibitoren, gelb: GLP-1 Rezeptor-Agonisten, rot: SGLT-2 Inhibitoren) [69]. Die farbliche Schattierung stellt das 95%-Konfidenzintervall dar. Auf der x-Achse ist die Zeit in Monaten und auf der y-Achse die NNT mit einer logarithmischen Skala aufgetragen. Alle NNT-Schätzungen und Konfidenzintervalle wurden aus Gründen der Übersichtlichkeit ab einem NNT-Wert von 100.000 abgeschnitten.

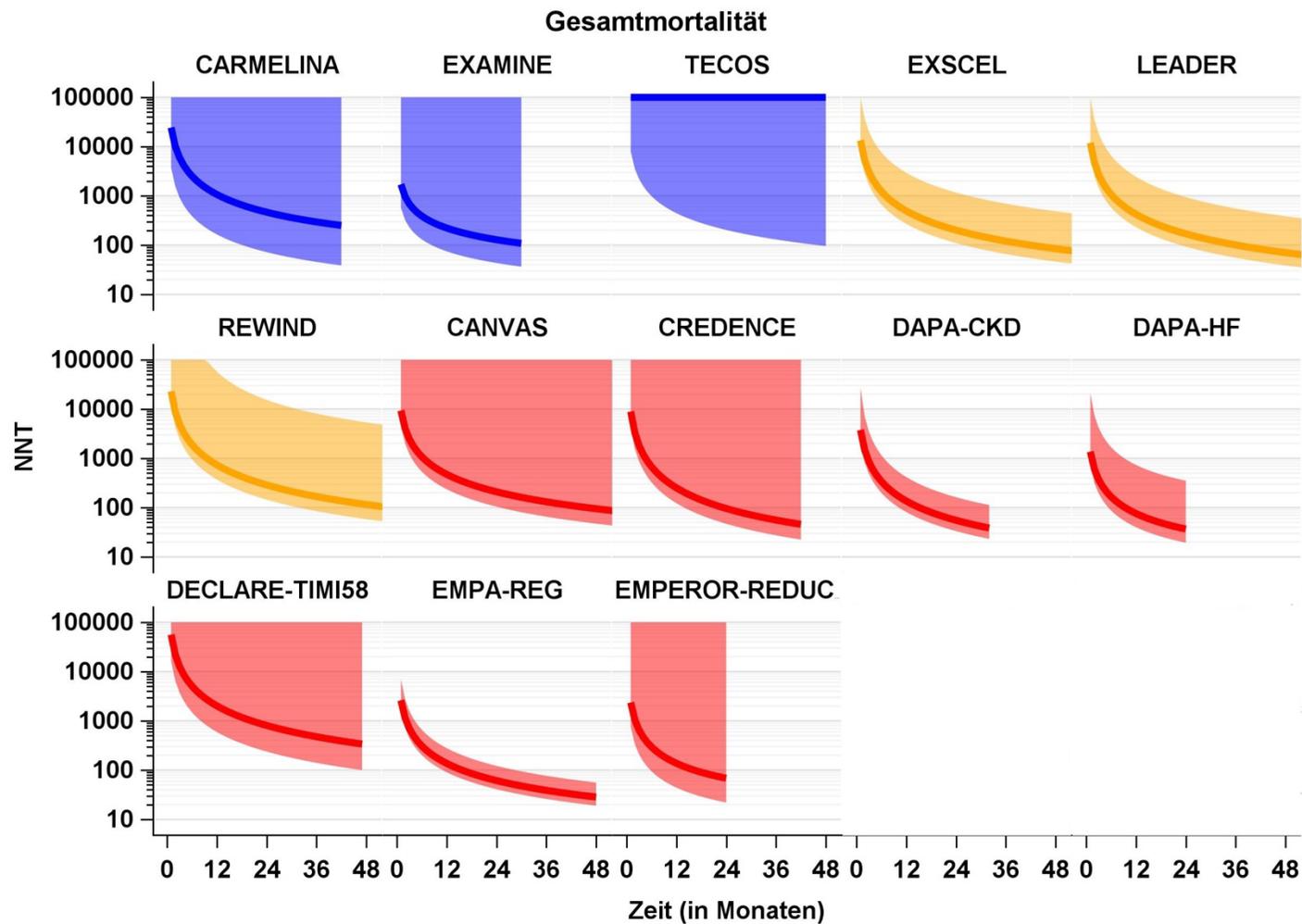


Abb. 5: Graphische Darstellung der Number-needed-to-treat für die Gesamtmortalität (blau: DPP-4 Inhibitoren, gelb: GLP-1 Rezeptor-Agonisten, rot: SGLT-2 Inhibitoren) [69]. Die farbliche Schattierung stellt das 95%-Konfidenzintervall dar. Auf der x-Achse ist die Zeit in Monaten und auf der y-Achse die NNT mit einer logarithmischen Skala aufgetragen. Alle NNT-Schätzungen und Konfidenzintervalle wurden aus Gründen der Übersichtlichkeit ab einem NNT-Wert von 100.000 abgeschnitten.

3.4 Meta-NNT: DDP-4 Inhibitoren, GLP-1 Rezeptoragonisten und SGLT-2 Inhibitoren

Aus den einzelnen NNT-Ergebnissen für jede CVOT und beide Endpunkte wurden Meta-NNTs für die drei Wirkstoffklassen (DDP-4 Inhibitoren, GLP-1 Rezeptoragonisten und SGLT-2 Inhibitoren) berechnet. Die Meta-NNTs sind in den Abb. 6 und 7 graphisch dargestellt. Die Meta-Analysen wurden mit den monatlich gepoolten Überlebensdaten aus den Weibull-Modellen und einer zufälligen Effektverteilung der inversen Varianzanalyse gebildet. Die Kurvenverläufe der Meta-NNT für SGLT-2 Inhibitoren und GLP-1 Rezeptoragonisten sind für beide Endpunkte ähnlich. Geringere absolute Behandlungseffekte zeigen die Meta-NNT für die DDP-4 Inhibitoren.

3.5 Meta-NNT: Primärer Endpunkt und die Gesamtmortalität

In Abb. 8 sind die Meta-NNTs für beide Endpunkte für alle eingeschlossenen CVOTs meta-analytisch zusammengefasst und graphisch dargestellt. Nach einer medianen Nachbeobachtungszeit von 29 Monaten beträgt die Meta-NNT für den primären Endpunkt 100 (95%-KI: 60-303). Die Meta-NNT für die Gesamtmortalität beträgt nach einer medianen Nachbeobachtungszeit von 39 Monaten 128 (95%-KI: 85-265).

3.6 Meta-NNT: Gesamtmortalität für 30 Jahre (Lebenszeitperspektive)

Um eine Langzeitperspektive für den absoluten Behandlungseffekt der neuen Antidiabetika zu erhalten, wurde getrennt für die drei Wirkstoffklassen eine Meta-NNT für die Gesamtsterblichkeit über 360 Behandlungsmonate (30 Behandlungsjahre) berechnet. Das Ergebnis ist in Abb. 9 dargestellt. Dabei wurde von einem konstanten Behandlungseffekt über den Zeitraum ausgegangen. Naturgemäß steigt das Sterberisiko in beiden Behandlungsgruppen mit der Zeit an. Dies hat zur Folge, dass der absolute Behandlungseffekt zum Zeitpunkt der niedrigsten Meta-NNT sein Maximum erreicht und dann mit der Zeit wieder ansteigt. Die Meta-NNT für SGLT-2 Inhibitoren und GLP-1 Rezeptoragonisten verhalten sich ähnlich. Die Meta-NNT für DDP-4 Inhibitoren zeigt einen geringeren Klasseneffekt für die Gesamtmortalität.

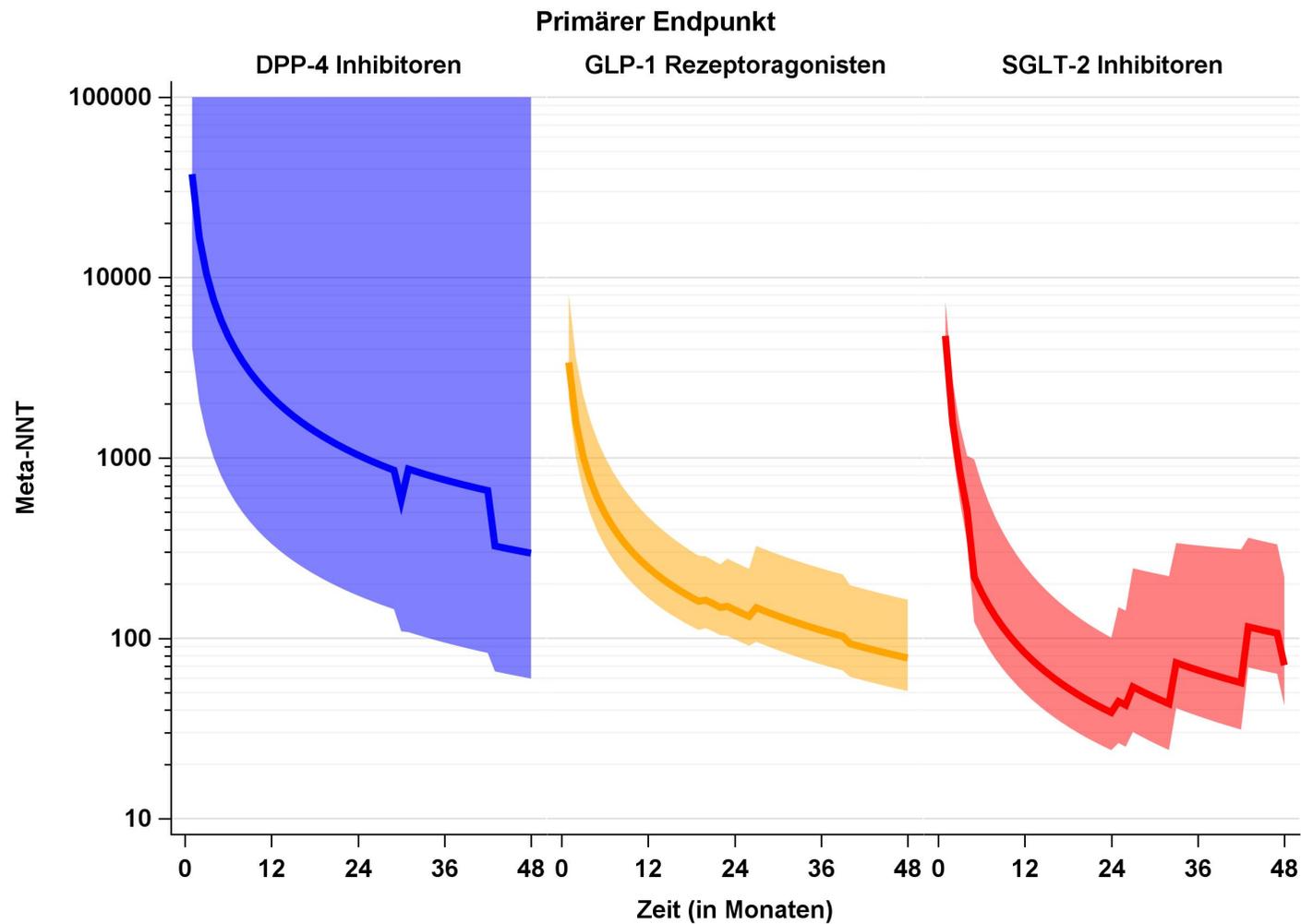


Abb. 6: Meta-NNT der drei Wirkstoffklassen für den primären Endpunkt (blau: DPP-4 Inhibitoren, gelb: GLP-1 Rezeptor-Agonisten, rot: SGLT-2 Inhibitoren) [69]. Die farbliche Schattierung stellt das 95%-Konfidenzintervall dar. Auf der x-Achse ist die Zeit in Monaten und auf der y-Achse die NNT mit einer logarithmischen Skala aufgetragen. Alle NNT-Schätzungen und Konfidenzintervalle wurden aus Gründen der Übersichtlichkeit ab einem NNT-Wert von 100.000 abgeschnitten. Für den primären Endpunkt wurden die Meta-NNTs aus den primären Endpunkten der einzelnen Studien berechnet, die sich in ihrer Definition leicht unterscheiden (Tabelle 1).

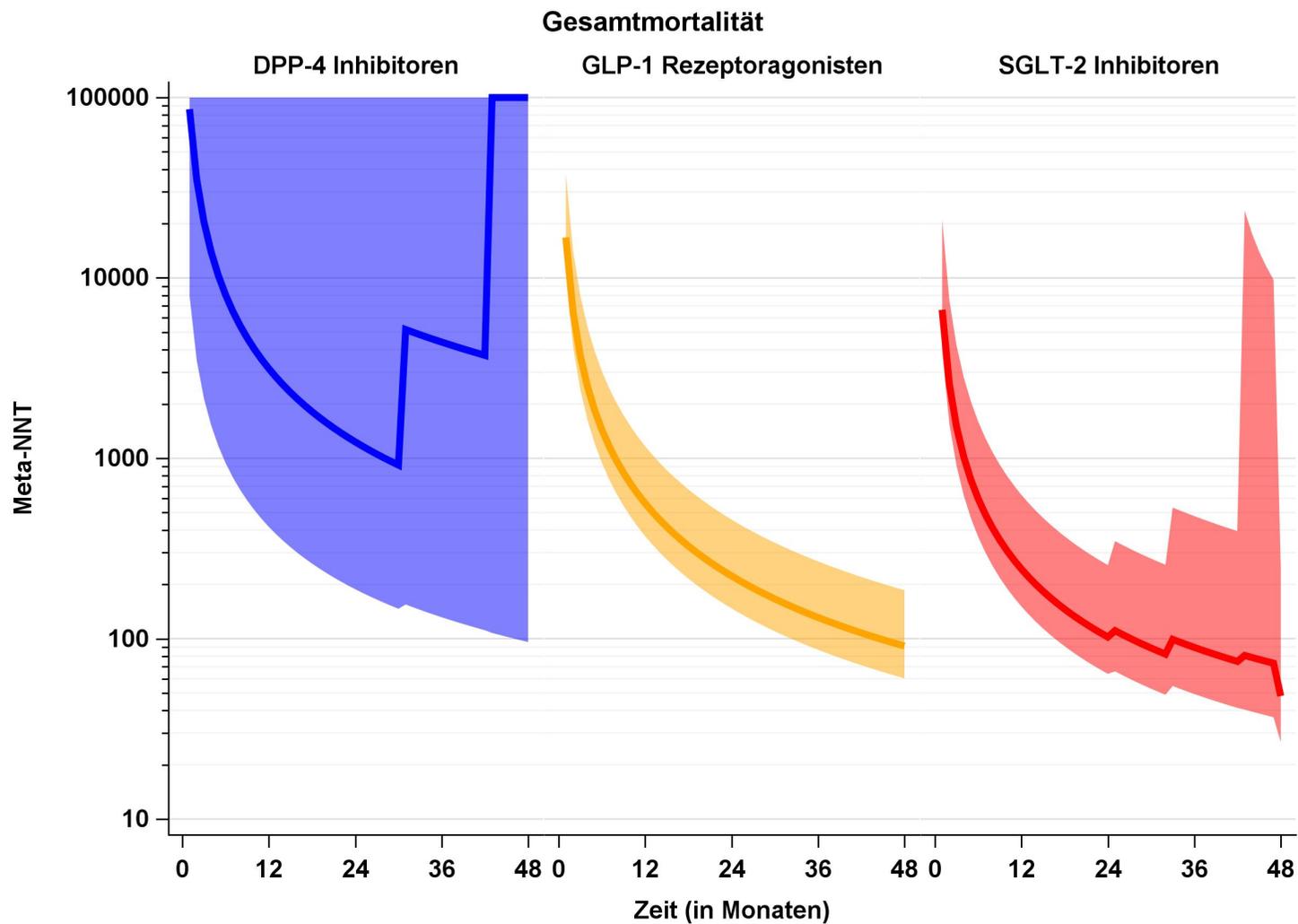


Abb. 7: Meta-NNT der drei Wirkstoffklassen für die Gesamtmortalität (blau: DPP-4 Inhibitoren, gelb: GLP-1 Rezeptor-Agonisten, rot: SGLT-2 Inhibitoren) [69]. Die farbliche Schattierung stellt das 95%-Konfidenzintervall dar. Auf der x-Achse ist die Zeit in Monaten und auf der y-Achse die NNT mit einer logarithmischen Skala aufgetragen. Alle NNT-Schätzungen und Konfidenzintervalle wurden aus Gründen der Übersichtlichkeit ab einem NNT-Wert von 100.000 abgeschnitten.

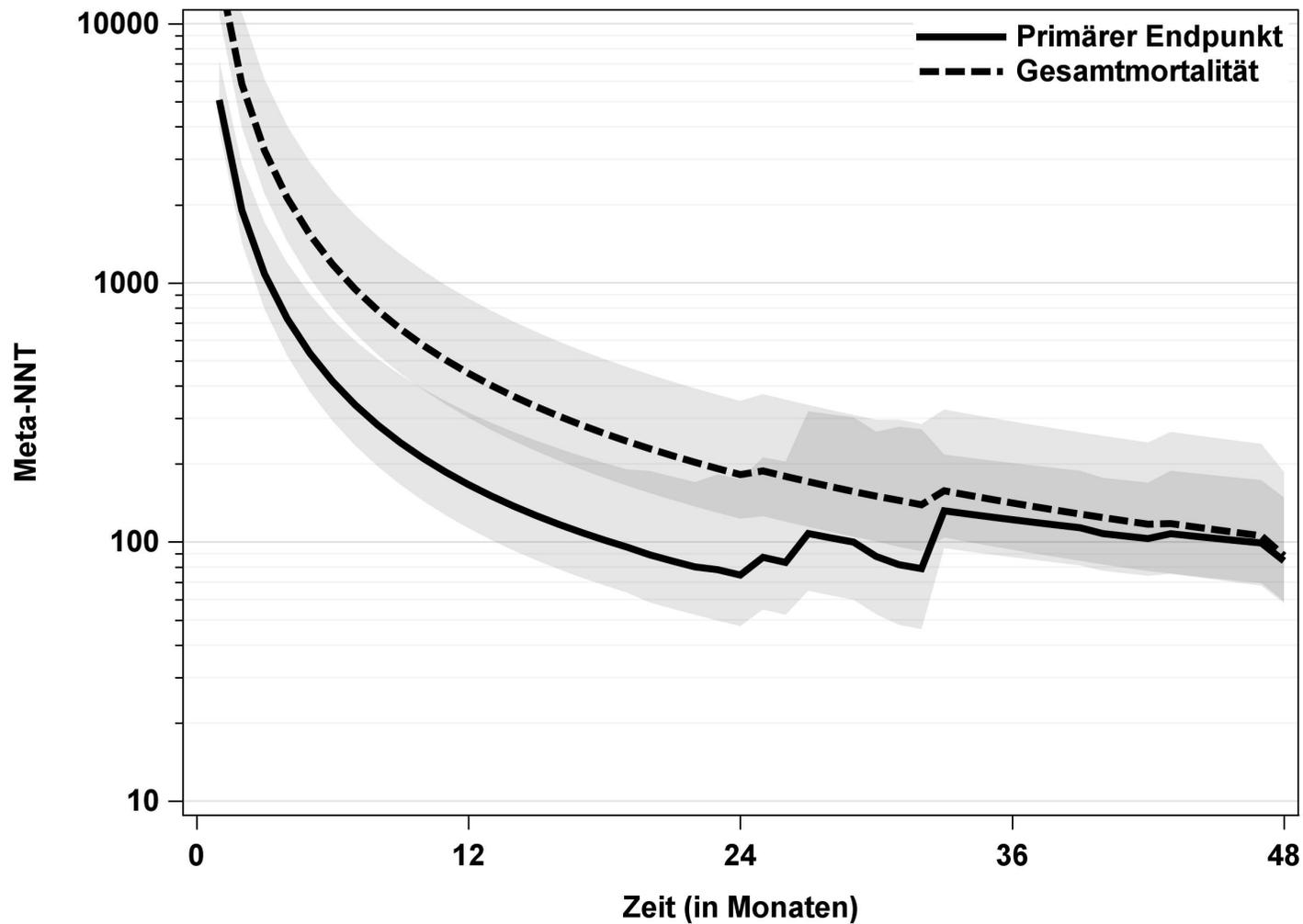


Abb.8: **Meta-NNT für die beiden Endpunkte: Primärer Endpunkt und Gesamtmortalität [69]**. In der Legende ist die Gesamtmortalität als gestrichelte Linie und der primäre Endpunkt als durchgezogene Linie dargestellt. Die graue Schattierung stellt das 95%-Konfidenzintervall dar. Auf der x-Achse ist die Zeit in Monaten und auf der y-Achse die NNT mit einer logarithmischen Skala aufgetragen. Für den primären Endpunkt wurden die Meta-NNTs aus den primären Endpunkten der einzelnen Studien berechnet, die sich in ihrer Definition leicht unterscheiden (Tabelle 1).

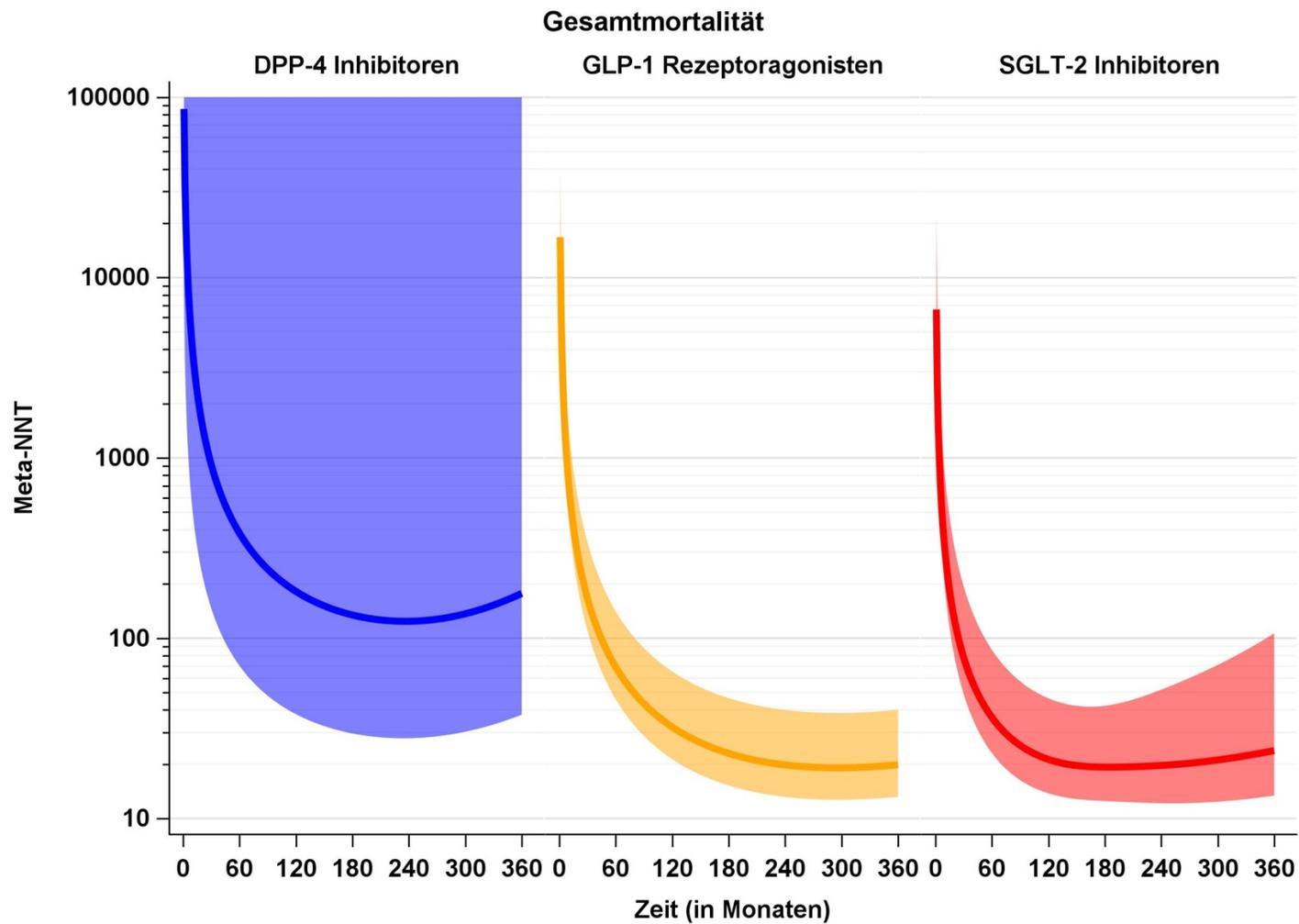


Abb. 9: **Lebenszeitperspektive: Meta-NNT Gesamtmortalität für 360 Monate (blau: DPP-4 Inhibitoren, gelb: GLP-1 Rezeptor-Agonisten, rot: SGLT-2 Inhibitoren)** [69]. Die farbliche Schattierung stellt das 95%-Konfidenzintervall dar. Auf der x-Achse ist die Zeit in Monaten und auf der y-Achse die NNT mit einer logarithmischen Skala aufgetragen. Alle NNT-Schätzungen und Konfidenzintervalle wurden aus Gründen der Übersichtlichkeit ab einem NNT-Wert von 100.000 abgeschnitten. Für den primären Endpunkt wurden die Meta-NNTs aus den primären Endpunkten der einzelnen Studien berechnet, die sich in ihrer Definition leicht unterscheiden (Tabelle 1).

3.7 Validität

Um die Validität der verwendeten Methode zu überprüfen, wurden die originalen Hazard Ratios aus den veröffentlichten Studienergebnissen mit den Hazard Ratios aus den angepassten Weibull-Modellen verglichen und in den Tabellen 4 und 5 dargestellt. Zusätzlich sind in den Tabellen 4 und 5 für alle eingeschlossenen Studien die Anzahl der beobachteten Patienten, die absoluten Ereigniszahlen, die prozentualen Ereignisraten und die mediane Nachbeobachtungszeit in Monaten dargestellt. Der Intra-Class-Korrelationskoeffizient zwischen den originalen Hazard Ratios und denjenigen aus den angepassten Weibull-Modellen wird als Maß für die Validität der verwendeten Methode angesehen. Die Intra-Class-Korrelationskoeffizienten wurden in Abb. 10 in Form eines Streudiagramms dargestellt. Es zeigt sich eine sehr geringe Abweichung zwischen den originalen Hazardraten und den Weibull-Hazardraten. Der Intra-Class-Korrelationskoeffizient für den primären Endpunkt beträgt 99,8 % (95 %-KI: 99,5 % - 100 %) und für den Endpunkt Gesamtmortalität 99,5 % (95 %-KI: 98,9 % - 100 %). Ergänzend ist anzumerken, dass die oberen Grenzen der Konfidenzskalen bei 100 % abgeschnitten wurden. Zur Beurteilung der Anpassung des Weibull-Modells wurden die Kaplan-Meier-Schätzer für beide Behandlungsgruppen aller untersuchten Studien und beide Endpunkte zusammen mit den 95%-Konfidenzintervallen der angepassten Weibull-Überlebensfunktionen in den Abb. 11 und 12 graphisch dargestellt. Hier zeigen sich keine relevanten Unterschiede, die die Interpretation der NNTs beeinträchtigen würden.

Darüber hinaus konnten die geschätzten NNT-Werte aus dieser Arbeit mit den Ergebnissen anderer Arbeiten verglichen werden, die andere Methoden zur Schätzung der NNTs verwendeten. Davies et al. [91] veröffentlichten 2020 ihre Arbeit, in der 13 CVOTs mit GLP-1 Rezeptoragonisten und DDP-4 Inhibitoren hinsichtlich ihrer klinischen Relevanz untersucht wurden. Wie in dieser Arbeit wurden auch die absolute Risikoreduktion und die NNT geschätzt. Ebenfalls wurde für die Schätzung der NNT-Werte aus den Originalpublikationen extrahiert. Im Gegensatz zu Davies et al. [91] wurden in dieser Arbeit nicht nur die Medikamentenklassen der GLP-1 Rezeptoragonisten und DDP-4 Inhibitoren, sondern auch die SGLT2-Inhibitoren untersucht. Davies et al. [91] schlossen in ihrer Arbeit die primären Endpunkte von 13

CVOTs ein, während in der vorliegenden Arbeit die primären Endpunkte von 19 CVOTs eingeschlossen wurden. Darüber hinaus wurde in dieser Arbeit auch die Gesamtmortalität analysiert und Metaanalysen für die drei Medikamentenklassen und für beide Endpunkte durchgeführt. Davies et al. [91] verwendeten die gleichen publizierten Daten der CVOTs für die Extraktion der *Time-to-event* Daten: Kaplan-Meier-Überlebenszeitkurven, die Anzahl der Ereignisse in jedem Studienarm und die Anzahl der Risikopersonen zu mehreren Zeitpunkten. Es wurde jedoch eine andere Digitalisierungssoftware (Engauge Digitizer Software, Version 10.11) verwendet.

Anschließend wurden die *Time-to-event* Daten für jedes Individuum mit dem Royston-Parmer-Überlebenszeitmodell rekonstruiert und die Berechnungen mit der STATA-Software (v. 16.0) durchgeführt. In Tabelle 6 werden die Ergebnisse der Studie von Davies et al. [91] mit den Ergebnissen der vorliegenden Arbeit für die Studien EXSCEL [77], HARMONY [78], LEADER [79], PIONEER [80], REWIND [81] und SUSTAIN-6 [82] für die Nachbeobachtungszeit von 1 und 3 Jahren verglichen.

Ebenfalls im Jahr 2020 publizierten Ludwig et al. [92] eine Arbeit, in der NNTs für publizierte CVOTs geschätzt wurden. Der Schwerpunkt dieser Arbeit lag auf der angewandten Methodik zur Schätzung von NNTs und der korrekten Interpretation von *Time-to-event* Studien, zu denen auch die CVOTs gehören. Im Rahmen der Arbeit wurden CVOTs der Wirkstoffklassen GLP-1 Rezeptoragonisten und SGLT-2 Inhibitoren analysiert. Im Gegensatz zur vorliegenden Arbeit wurden in der Arbeit von Ludwig et al. [92] nicht alle CVOTs einer Wirkstoffklasse systematisch analysiert, sondern nur einzelne Stichproben durchgeführt. Ein wesentlicher methodischer Unterschied besteht darin, dass Ludwig et al. [92] im Gegensatz zur vorliegenden Arbeit nicht die publizierten Originaldaten der CVOTs verwendeten, sondern bereits aggregierte Daten. Dadurch haben Ludwig et al. [92] nicht die Möglichkeit, die digitalisierten *Time-to-event* Daten mit den Hazard Ratios aus den Originalarbeiten zu vergleichen. In Tabelle 7 sind die Studienergebnisse von Ludwig et al. [92] für die Studien HARMONY [78], LEADER [79], REWIND [81], SUSTAIN-6 [82], DECLARE-TIMI-58 [87] und EMPA-REG [13] den Ergebnissen dieser Arbeit gegenübergestellt. Ludwig et al. [92] beschreiben in ihrer Arbeit die NNT nur für die Nachbeobachtungszeit der entsprechenden Studien. Der Vergleich mit den Ergebnissen dieser Arbeit erfolgt daher nicht zu einem einheitlichen Zeitpunkt im Nachbeobachtungszeitraum.

Studien	Anzahl der Ereignisse	Patientenanzahl	Ereignisrate (%)	Mediane Nachbeobachtungszeit (Monate)	Hazard Ratio (95%-KI) Originale Daten	Hazard Ratio (95%-KI) Digitalisierte Daten: Cox-Modell	Hazard Ratio (95%-KI) Digitalisierte Daten: Weibull-Modell
Primärer Endpunkt							
DDP-4 Inhibitor							
CARMELINA	852	6.979	12,2	25,9	1,02 (0,89 - 1,17)	1,02 (0,90 - 1,17)	1,02 (0,89 - 1,16)
EXAMINE	621	5.380	11,5	18,4	0,96 (-)	0,96 (0,82 - 1,12)	0,95 (0,80 - 1,10)
SAVOR-TIMI53	1.217	16.492	7,4	24,8	1,00 (0,89 - 1,12)	1,00 (0,89 - 1,11)	1,00 (0,88 - 1,11)
TECOS	1.686	14.671	11,5	34,0	0,98 [0,89 - 1,08]	0,99 (0,90 - 1,09)	0,98 (0,88 - 1,07)
GLP-1 Rezeptoragonist							
ELIXA	790	6.068	13,0	24,5	1,02 (0,89 - 1,17)	1,00 (0,87 - 1,15)	1,01 (0,87 - 1,15)
EXSCEL	1.708	14.752	11,6	35,2	0,91 (0,83 - 1,00)	0,91 (0,83 - 1,00)	0,91 (0,83 - 1,00)
HARMONY	766	9.463	8,1	19,0	0,78 (0,68 - 0,90)	0,79 (0,68 - 0,91)	0,78 (0,67 - 0,89)
LEADER	1.298	9.340	13,9	45,9	0,87 (0,78 - 0,97)	0,87 (0,78 - 0,96)	0,87 (0,77 - 0,96)
PIONEER-6	135	3.183	4,2	15,9	0,79 (0,57 - 1,11)	0,79 (0,56 - 1,11)	0,79 (0,52 - 1,06)
REWIND	1.243	9.901	12,6	64,8	0,88 (0,79 - 0,99)	0,88 (0,79 - 0,99)	0,88 (0,79 - 0,98)
SUSTAIN-6	221	3.297	6,7	22,1	0,74 (0,58 - 0,95)	0,73 (0,56 - 0,95)	0,72 (0,53 - 0,92)
SGLT-2 Inhibitoren							
CANVAS	1.036	10.142	10,2	28,9	0,86 (0,75 - 0,97)	0,86 (0,76 - 0,97)	0,87 (0,77 - 0,98)
CREDENCE	583	4.401	13,2	30,6	0,70 (0,59 - 0,82)	0,69 (0,59 - 0,82)	0,69 (0,58 - 0,81)
DAPA-CKD	493	4.304	11,5	25,5	0,61 (0,51 - 0,72)	0,60 (0,50 - 0,72)	0,60 (0,49 - 0,71)
DAPA-HF	885	4.744	18,7	17,5	0,74 (0,65 - 0,85)	0,75 (0,65 - 0,85)	0,75 (0,65 - 0,84)
DECLARE-TIMI 58	1.447	17.160	8,4	47,3	0,93 (0,84 - 1,03)	0,94 (0,84 - 1,04)	0,94 (0,84 - 1,03)
EMPA-REG	753	7.020	10,7	36,2	0,86 (0,74 - 0,99)	0,85 (0,73 - 0,98)	0,84 (0,72 - 0,97)
EMPEROR-REDUCED	808	3.730	21,7	14,2	0,75 (0,65 - 0,86)	0,75 (0,65 - 0,86)	0,75 (0,65 - 0,85)
VERTIS-CV	959	8.238	11,6	32,8	0,97 (0,85 - 1,11)	0,95 (0,83 - 1,09)	0,95 (0,82 - 1,08)

Tabelle 4: Primärer Endpunkt: Beschreibung der Studien und Vergleich der originalen Hazard Ratios mit den Hazard Ratios aus den digitalisierten Kaplan-Meier-Überlebenszeitkurven (Cox- und Weibull-Modell) [69].

Studien	Anzahl der Ereignisse	Patientenanzahl	Ereignisrate (%)	Mediane Nachbeobachtungszeit (Monate)	Hazard Ratio (95%-KI) Originale Daten	Hazard Ratio (95%-KI) Digitalisierte Daten: Cox-Modell	Hazard Ratio (95%-KI) Digitalisierte Daten: Weibull-Modell
Gesamtmortalität							
DDP-4 Inhibitor							
CARMELINA	739	6.979	10,6	26,8	0,98 (0,84 - 1,13)	0,97 (0,84 - 1,12)	0,97 (0,83 - 1,11)
EXAMINE	316	5.380	5,9	19,5	0,88 (0,71 - 1,09)	0,89 (0,72 - 1,11)	0,89 (0,70 - 1,09)
TECOS	1.042	14.671	7,1	36,2	1,01 (0,90 - 1,14)	1,01 (0,89 - 1,14)	1,01 (0,88 - 1,13)
GLP-1 Rezeptoragonist							
EXSCEL	1.051	14.752	7,1	40,1	0,86 (0,77 - 0,97)	0,86 (0,77 - 0,98)	0,86 (0,76 - 0,97)
LEADER	824	9.340	8,8	46,5	0,85 (0,74 - 0,97)	0,84 (0,74 - 0,97)	0,85 (0,73 - 0,96)
REWIND	1.109	9.901	11,2	65,4	0,90 (0,80 - 1,01)	0,89 (0,79 - 1,00)	0,89 (0,78 - 0,99)
SGLT-2 Inhibitor							
CANVAS	701	10.142	6,9	29,6	0,87 (0,74 - 1,01)	0,85 (0,73 - 0,99)	0,86 (0,73 - 0,99)
CREDENCE	366	4.401	8,3	31,5	0,83 (0,68 - 1,02)	0,82 (0,67 - 1,01)	0,82 (0,65 - 0,99)
DAPA-CKD	234	4.304	5,4	27,5	0,69 (0,53 - 0,88)	0,68 (0,52 - 0,88)	0,68 (0,50 - 0,85)
DAPA-HF	598	4.744	12,6	18,2	0,83 (0,71 - 0,97)	0,84 (0,71 - 0,98)	0,84 (0,70 - 0,97)
DECLARE-TIMI 58	954	17.160	5,6	47,3	0,93 (0,82 - 1,04)	0,95 (0,84 - 1,08)	0,95 (0,83 - 1,07)
EMPA-REG	461	7.020	6,6	37,7	0,68 (0,57 - 0,82)	0,67 (0,56 - 0,81)	0,67 (0,55 - 0,80)
EMPEROR-REDUCED	493	3.730	13,2	15,6	0,92 (0,77 - 1,10)	0,91 (0,76 - 1,09)	0,92 (0,76 - 1,08)

Tabelle 5: Gesamtmortalität: Beschreibung der Studien und Vergleich der ursprünglichen Hazard Ratios mit den Hazard Ratios aus den digitalisierten Kaplan-Meier-Überlebenszeitkurven (Cox- und Weibull-Modell) [69].

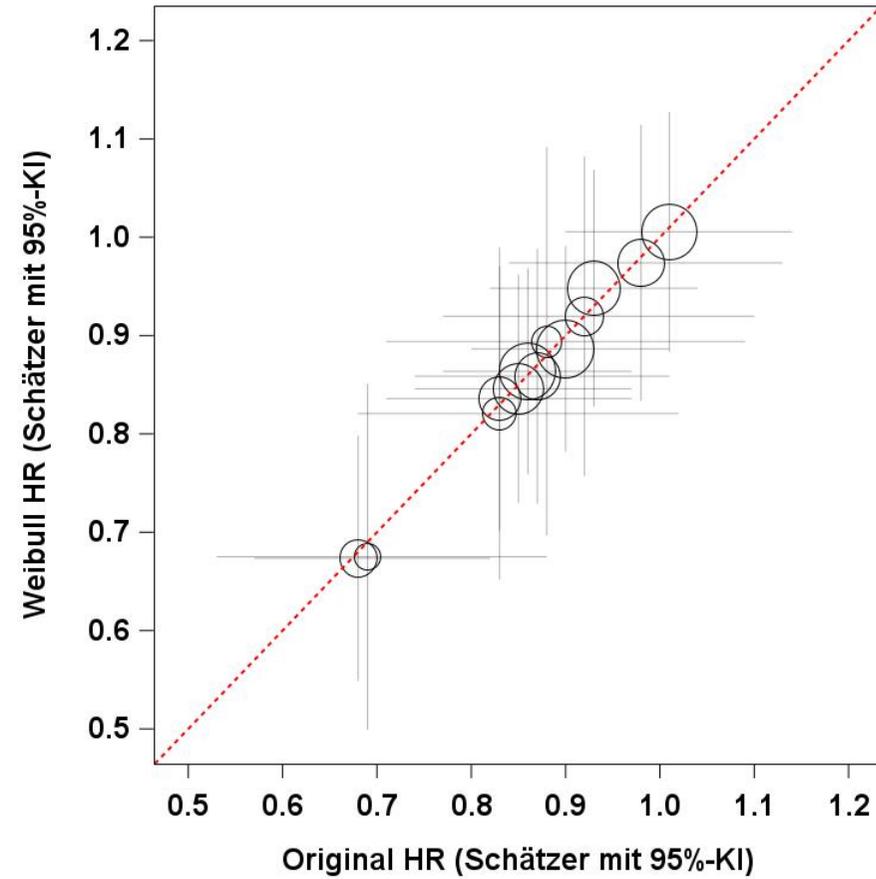
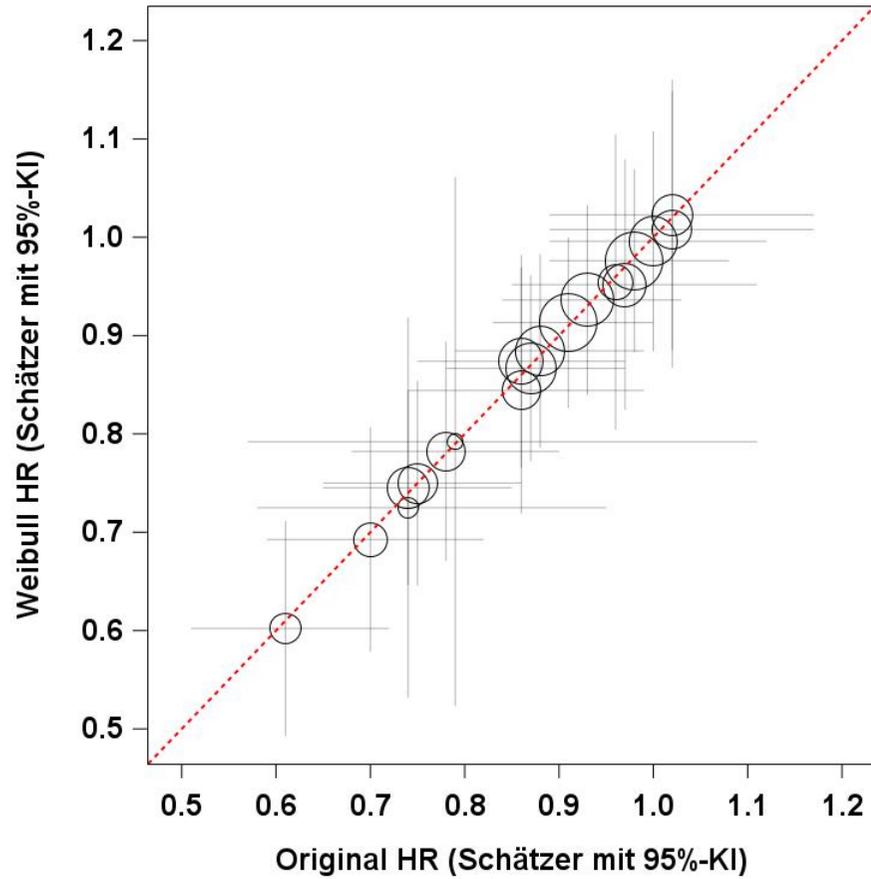


Abb.10: Streudiagramme für die Intra-Class-Korrelationskoeffizienten [69]. Links: Primäre Endpunkte. Rechts: Gesamtmortalität. Vergleich der Hazard Ratios aus den Originalpublikationen (x-Achse) mit den Hazard Ratios aus dem angepassten Weibull-Modell der digitalisierten Daten (y-Achse).

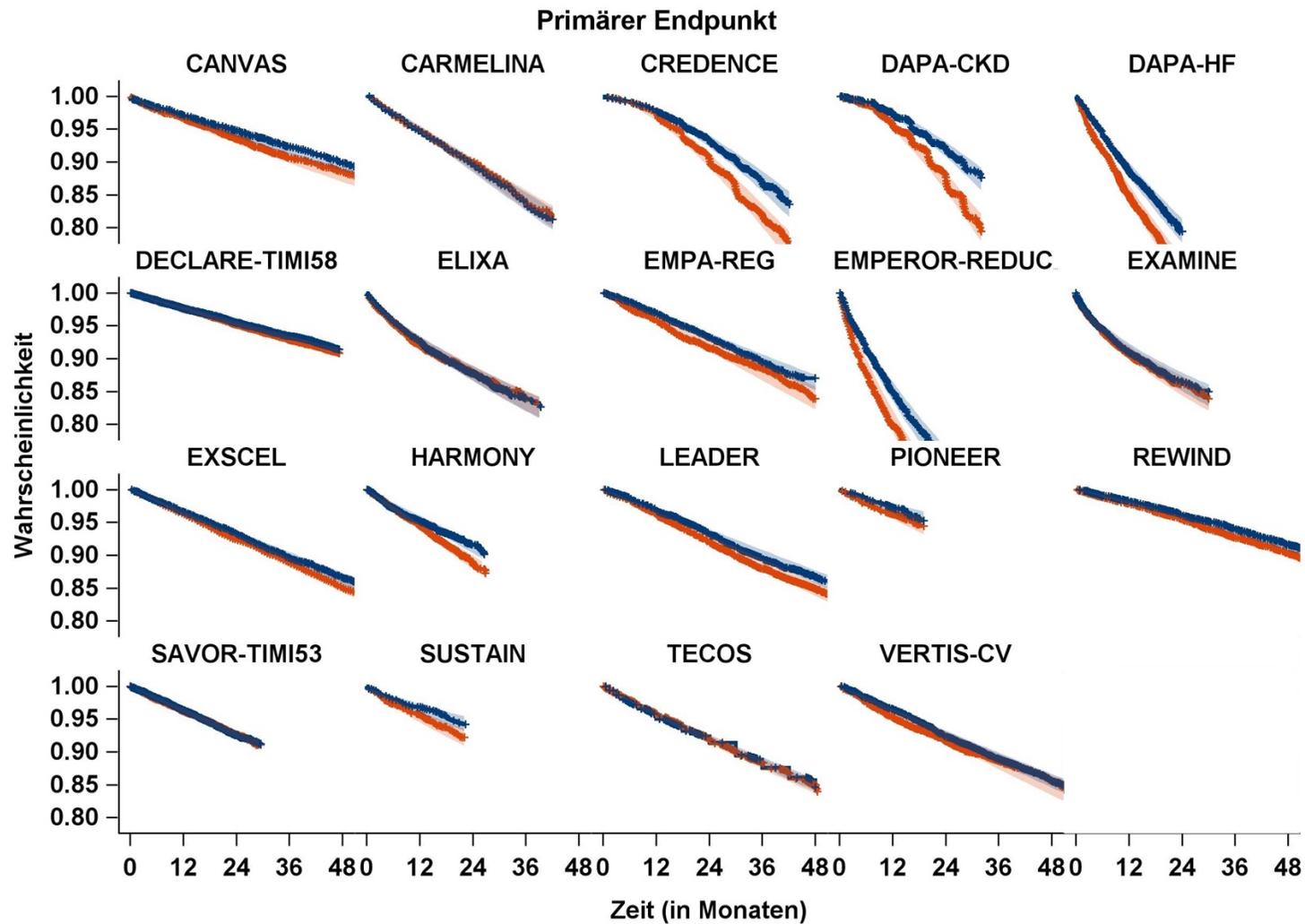


Abb. 11: Primäre Endpunkte: Kaplan-Meier-Schätzungen für die extrahierten Daten mit den 95%-Konfidenzintervallen der Weibull-Anpassung für die jeweilige maximale Beobachtungszeit der Studie [69]. Auf der X-Achse ist die Zeit in Monaten abgebildet und auf der Y-Achse die Überlebenswahrscheinlichkeit.

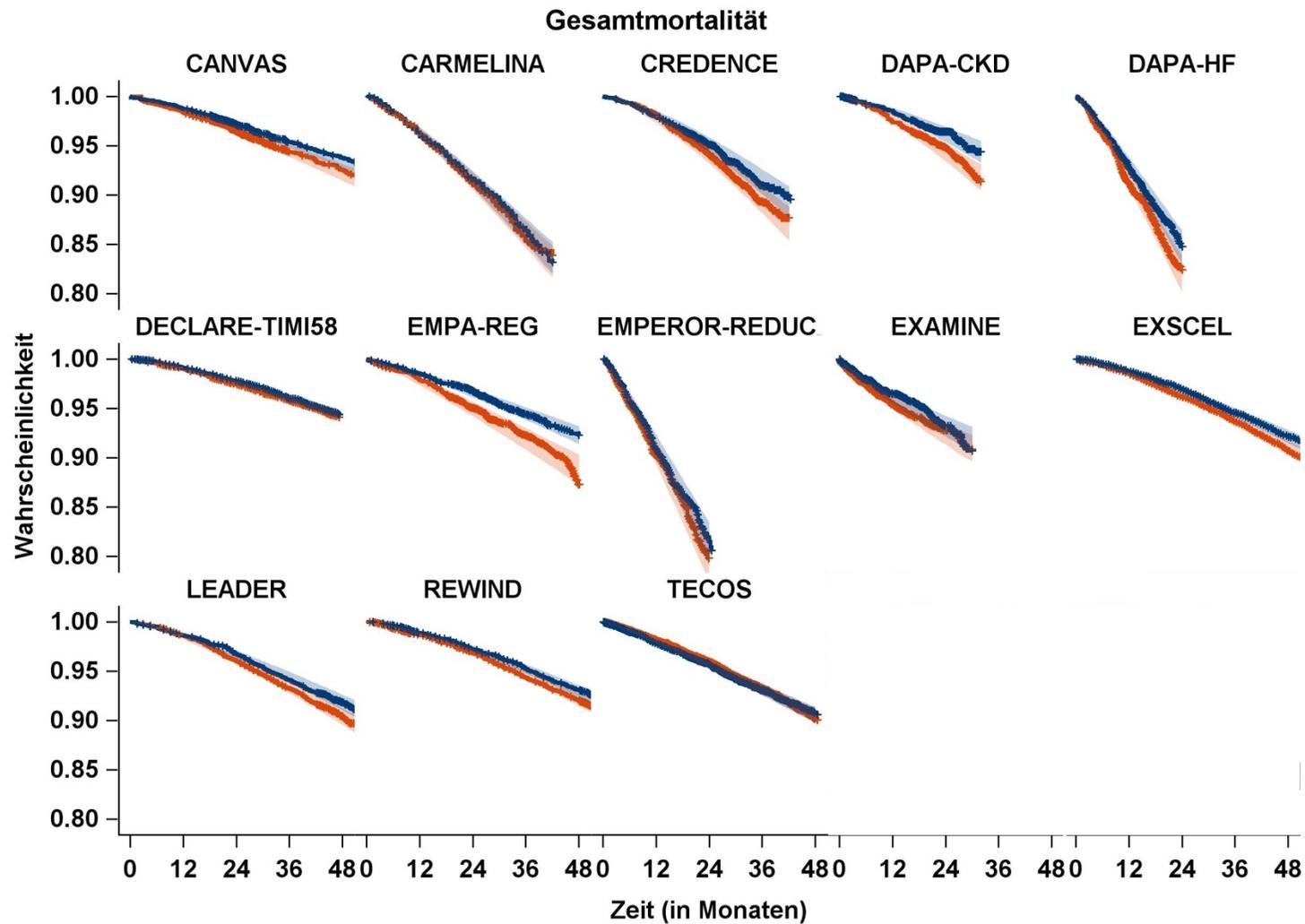


Abb. 12: Gesamt mortalität: Kaplan-Meier-Schätzungen für die extrahierten Daten mit den 95%-Konfidenzintervallen der Weibull-Anpassung für die jeweilige maximale Beobachtungszeit der Studie [69]. Auf der X-Achse ist die Zeit in Monaten abgebildet und auf der Y-Achse die Überlebenswahrscheinlichkeit.

Studien	Nachbeobachtungszeit (Jahre)	Davies et al.[91] NNT (95%-KI)	Digitalisierte Daten [69] NNT (95%-KI)
EXSCel	1	365 (801 - [- 71])	326 (159 - [-7.005])
HARMONY	1	83 (35 - 131)	83 (52 - 197)
LEADER	1	209 (53 - 366)	207 (117 - 881)
PIONEER	1	135 (331 - [-61])	136 (55 - [-299])
REWIND	1	429 (41 - 816)	434 (227 - 4.938)
SUSTAIN-6	1	88 (15 - 161)	80 [44 - 464)
EXSCel	3	126 (275 - [-24])	110 (54 - [-2.433])
HARMONY	3	30 (13 - 48)	30 (19 - 72)
LEADER	3	65 (17 - 113)	69 (39 - 291)
PIONEER	3	50 (124 - [-23])	50 (20 - [-108])
REWIND	3	129 (13 - 246)	127 (67- 1.366)
SUSTAIN-6	3	31 (6 - 57)	33 (18 - 188)

Tabelle 6: Vergleich mit den Studienergebnissen von Davies et al. [91].

Studien	Nachbeobachtungszeit (Jahre)	Ludwig et al. [92] NNT (95% KI)	Digitalisierte Daten [69] NNT (95% KI)
HARMONY	1,6	53 (36 - 116)	53 (34 - 126)
LEADER	3,8	56 (33 - 243)	56 (32 - 234)
REWIND	5,1	67 (38 - 803)	73 (38 - 775)
SUSTAIN-6	2,1	45 (28 - 235)	43 (24 - 248)
DECLARE-TIMI58	4,2	104 (66 - 355)	169 (66 - [-300])
EMPA-REG	3,1	63 (34 - 882)	55 (29 - 528)

Tabelle 7: Vergleich mit den Studienergebnissen der Studie von Ludwig et al. [92].

4. Diskussion

In dieser Arbeit wurden die absoluten Behandlungseffekte der neuen Antidiabetika für den primären Endpunkt und die Gesamtmortalität untersucht.

Die Schätzung absoluter Behandlungseffekte in Form der Number-needed-to-treat bietet eine neue Perspektive zur Beurteilung von Behandlungseffekten im Vergleich zur alleinigen Darstellung relativer Risikomaße. Diese Analyse zeigt, dass insbesondere die Wirkstoffklassen der GLP-1 Rezeptoragonisten und der SGLT-2 Inhibitoren günstige Effekte auf den primären Endpunkt und die Gesamtmortalität zeigen.

Die neuen Richtlinien der US.-Arzneimittelbehörde (*Food and Drug Administration* [FDA]) [2] von 2008 in Folge des „Rosiglitazon-Skandals“ [3] haben zur Entwicklung und Durchführung der sogenannten *Cardiovascular Outcomes Trials* (CVOTs) geführt. Das Leitmotiv für die Durchführung dieser großen randomisiert kontrollierten Studien hat im Laufe der Zeit einen Paradigmenwechsel erfahren. Ursprünglich sollten die neuen Antidiabetika auf ihr kardiovaskuläres Sicherheitsprofil hin untersucht werden. Aus dieser defensiven Vorgabe der FDA [2] entstanden die Studienprotokolle der CVOTs. Zwischenzeitlich wurde in Fachkreisen offen diskutiert, ob die Fragestellung hinsichtlich der kardiovaskulären Sicherheit, die Ressourcen und der enorme finanzielle Aufwand, der mit der Durchführung der CVOTs verbunden ist, gerechtfertigt waren [1]. Diese Frage kann aus der Sicht von 2008, als es um den reinen Nachweis der kardiovaskulären Sicherheit ging, nicht vollständig beantwortet werden.

Letztlich wurden die CVOTs als große internationale und multizentrische Studien mit Tausenden von Patienten konzipiert und durchgeführt. Dadurch konnten im Laufe der Zeit viele medizinische Daten zu den neuen Antidiabetika gesammelt und ausgewertet werden, die über die ursprüngliche Fragestellung hinausgingen. Diese umfangreichen Daten haben die Frage aufgeworfen, ob einige der neuen Antidiabetika nicht nur zur Behandlung des Diabetes mellitus Typ 2 eingesetzt werden können, sondern auch einen krankheitsunabhängigen protektiven Effekt auf das kardiovaskuläre Risikoprofil haben.

Die ersten abgeschlossenen CVOTs waren die Studien Savor-Timi-53 [74] (Saxagliptin) und EXAMINE [73] (Alogliptin), die beide zu den DDP-4 Inhibitoren gehören. Beide

Studien konnten die von der FDA geforderte Nicht-Unterlegenheit gegenüber Placebo nachweisen. Als 2015 die Ergebnisse der abgeschlossenen EMPA-REG Studie [13] (Empagliflozin Cardiovascular Outcome Event Trial in Type 2 Diabetes Mellitus Patient) im *New England Journal of Medicine* publiziert wurden, kamen sie einem Paradigmenwechsel gleich. Erstmals konnte in einer CVOT ein signifikanter Vorteil für den primären Endpunkt (Kombinierter Endpunkt: nicht-tödlicher Myokardinfarkt, nicht-tödlicher Schlaganfall und kardiovaskuläre Mortalität [3-Punkt-MACE]) nachgewiesen werden. Mit einer Hazard Ratio von 0,86 (95%-KI: 0,74 - 0,99; $P < 0,001$ für die Nicht-Unterlegenheit und $P = 0,04$ für die Überlegenheit) war der primäre Endpunkt signifikant reduziert und konnte somit nicht nur die Nicht-Unterlegenheit, sondern auch die Überlegenheit aufzeigen. Ausschlaggebend für die Reduktion des kombinierten primären Endpunkts war die Reduktion der kardiovaskulären Mortalität (HR 0,62 [95%-KI: 0,49 - 0,77; $P < 0,001$]). Darüber hinaus zeigten Subgruppenanalysen signifikante Vorteile für Patienten mit Herzinsuffizienz [13].

Spätestens zu diesem Zeitpunkt wurde die Frage aufgeworfen, ob die neuen Medikamente nicht nur zur Behandlung des Diabetes mellitus Typ 2 eingesetzt werden können, sondern auch einen kardioprotektiven Ansatz bieten. Mit dem Nachweis der kardiovaskulären Sicherheit wandte sich die pharmazeutische Industrie in der Folge neuen Studien zu, die sich mit protektiven Effekten auf das kardiovaskuläre und das renale System befassten. Schließlich wurde ein Teil der in den CVOTs untersuchten Medikamente in die Leitlinienempfehlungen großer Fachgesellschaften zur Behandlung des Diabetes mellitus Typ 2 [32] und zur Therapie von Patienten mit Herzinsuffizienzen [93] aufgenommen.

Die Studiendaten der CVOTs haben zu einem Paradigmenwechsel in der Behandlung des Diabetes mellitus Typ 2 geführt und sind zu einem wichtigen Baustein in der Therapie der Herzinsuffizienz geworden. Umso wichtiger ist es, sie in einen wissenschaftlichen Kontext mit bestmöglicher Informationsqualität zu stellen.

Die ersten aufsehenerregenden Ergebnisse der EMPA-REG Studie [13] wurden mit einer Risikoreduktion der kardiovaskulären Mortalität von 38% betitelt. Die 38% beziehen sich auf die relative Risikoreduktion für den Endpunkt der kardiovaskulären Mortalität, die in der Studie durch die folgende Hazard Ratio ausgedrückt wird: HR 0,62 (95%-KI: 0,49 -

0,77; $P < 0,001$). Doch wie ist diese Information einzuordnen? Auf den ersten Blick zeigt die Aussage eine beeindruckende Reduktion der kardiovaskulären Mortalität. Viele Patienten und andere medizinische Laien würden diese Zahl so interpretieren, dass bei einer Einnahmedauer von 3 Jahren (mediane Dauer der EMPA-REG Studie betrug 3,1 Jahre) etwa jede dritte Person hinsichtlich der kardiovaskulären Mortalität profitieren würde [62]. Jedoch handelt es sich bei dieser Schlussfolgerung um eine häufige Überschätzung der präsentierten Daten, die mit einer Fehlinterpretation der relativen Risikomaße einhergeht. Tatsächlich würde eine von 45 Personen (NNT 45 [95%-KI: 30 - 86]) nach 3 Jahren Behandlung von dem Medikament hinsichtlich der kardiovaskulären Mortalität profitieren. Zur korrekten Interpretation relativer Risikomaße wie der Hazard Ratio ist zusätzlich die Information über das bestehende Basisrisiko der untersuchten Population notwendig, um eine valide Aussage über die klinische und ökonomische Relevanz treffen zu können. Hierzu analysierten Ferrannini et. al [62] die starke Abhängigkeit der relativen Risikoreduktion vom bestehenden Basisrisiko. Die zusätzliche Kenntnis der absoluten Risikoreduktion oder ihres Kehrwerts, der Number-needed-to-treat, ermöglicht den Entscheidungsträgern im Gesundheitswesen die bessere Interpretation der Studiendaten, insbesondere für die Bewertung von Kosten-Nutzen-Analysen.

In diesem Zusammenhang ist auch hervorzuheben, dass relative Risikomaße (z.B. Hazard Ratio) den Vorteil haben, dass sie die Behandlungseffekte auf das Ausgangsrisiko der Placebo-Population normieren und damit zu relativ stabilen Ergebnissen über das gesamte Risikospektrum führen. Diese Tatsache erleichtert den Vergleich verschiedener Studienpopulationen. Aus diesen Gründen wird in wissenschaftlichen Fachkreisen auf die Bedeutung von absoluten und relativen Risikomaße für eine umfassende Interpretation von Studiendaten hingewiesen [55, 94].

Die CVOT-Studienprotokolle wurden in Übereinstimmung mit den Richtlinien der FDA [2] entwickelt. Diese Richtlinien betonten zwar die Bedeutung absoluter Risikomaße, schrieben diese aber nicht zwingend vor. Aus diesem Grund haben nur wenige Studienautoren absolute Risikomaße, wie z.B. NNTs, zu ihren Studiendaten veröffentlicht. Nur zwei CVOTs (CREDESCENCE [84] und DAPA-CKD[85]) kommunizierten aktiv NNTs im Hauptbericht und nur zwei CVOTs (LEADER [95] und EMPA-REG [96]) kommunizierten NNTs in nachfolgenden Artikeln.

Die Literaturrecherche in der vorliegenden Arbeit stützt sich auf die bestehende Übersichtsarbeit von Cefalu et al. [1]. Bei allen CVOTs handelt es sich um große und gut konzipierte internationale Studien. Im Sinne der Fragestellung wurde in den CVOTs ein Patientenkollektiv mit hohem kardiovaskulärem Risiko eingeschlossen. So wurden bis auf zwei Studien (EMPEROR-REDUCED [88] und DAPA-HF [86]) ausschließlich Patienten mit langjährigem Diabetes mellitus Typ 2 eingeschlossen, die häufig bereits kardiovaskuläre Vorerkrankungen aufwiesen (Tabelle 1). Aufgrund der hohen Rate an kardiovaskulären Vorerkrankungen wollten die Studienautoren eine ausreichend hohe Anzahl an Ereignissen im relevanten Beobachtungszeitraum sicherstellen. Aus diesem Grund wird in dieser Arbeit von einer validen Schätzung der NNT auf der Basis der CVOT-Daten ausgegangen.

In dieser Arbeit wurden die absolute Risikoreduktion und die NNT zu zwei Endpunkten aus den CVOTs geschätzt. Der zusammengesetzte primäre Endpunkt wird in den CVOTs aus 3 oder 4 Endpunkten gebildet. Diese unterschiedliche Definition des primären Endpunktes in den CVOTs ist darauf zurückzuführen, dass die FDA in ihren Richtlinien [2] die Zusammensetzung des primären Endpunktes nicht genau definiert hatten, sondern nur eine Auswahl an verpflichtenden und optionalen kardiovaskulären Endpunkten zur Auswahl stellten. Dies führte zu Unterschieden in der Zusammensetzung des primären Endpunktes zwischen den CVOTs (Tabelle 1). Trotz dieser Schwäche ist der primäre Endpunkt für die Bewertung der kardiovaskulären Sicherheit im Sinne der FDA-Richtlinien wichtig und wurde daher in unsere Meta-Analyse einbezogen. Der Endpunkt Gesamtmortalität wurde ebenfalls eingeschlossen, da es keine Verzerrung zwischen den CVOTs aufgrund unterschiedlicher Definitionen gibt und dieser Endpunkt den einzigen Endpunkt ohne konkurrierende Risiken darstellt und ebenfalls von hoher klinischer Relevanz ist.

Zum Zeitpunkt der Datenerhebung hatten nur wenige Studienautoren Angaben zur absoluten Risikoreduktion in Erst- oder Folgepublikationen veröffentlicht, so dass eine valide Meta-Analyse zur absoluten Risikoreduktion oder zu NNTs für CVOTs allein auf der Basis der veröffentlichten Studiendaten nicht möglich war. Für die Berechnung der absoluten Risikoreduktion bzw. deren Kehrwert, ausgedrückt als NNT, wurden die *Time-*

to-event Daten der Studien benötigt. Aus den publizierten Studienergebnissen, die die relative Risikoreduktion in Form der Hazard Ratio beschreiben und unter der Annahme eines Cox-Modells berechnet wurden, ist die Bestimmung der absoluten Risikoreduktion nicht möglich. Aus diesen Gründen wurden die *Time-to-event* Daten auf der Grundlage der Ergebnisse der veröffentlichten Studienergebnisse mit Hilfe der oben beschriebenen Methode geschätzt. Die folgenden publizierten Informationen der CVOTs wurden hierfür verwendet: Die Kaplan-Meier-Überlebenszeitkurven für die untersuchten Endpunkte, die Informationen zu den absoluten Ereigniszahlen in beiden Studiengruppen und die Anzahl der Risikopatienten zu mehreren Zeitpunkten im Studienintervall.

Die Kaplan-Meier-Überlebenszeitkurven wurden mit dem WebPlotDigitizer (Version 4.2) digitalisiert [63]. Anschließend wurden aus den digitalisierten Informationen der Kaplan-Meier-Überlebenszeitkurven mit Hilfe des R-Algorithmus von Guyot et al. [64] die *Time-to-event* Daten für jede Studiengruppe extrahiert. Beide Verfahren erwiesen sich in unserer Literaturrecherche als valide und reliabel [65-67].

Für die Berechnung der absoluten Risikoreduktion wurde das parametrische Weibull-Modell gewählt. Entsprechend dem Weibull-Modell wurden monatliche absolute Risikodifferenzen zwischen den Studienarmen berechnet. Durch Inversion der monatlichen Risikodifferenzen konnten die monatlichen NNT-Werte geschätzt werden. Es gibt auch andere parametrische Modelle zur Berechnung der erforderlichen Daten, wie z. B. das Royston-Parmar-Modell, die auch flexiblere Überlebenskurven zulassen. In dieser Arbeit wurde sich für das Weibull-Modell entschieden, da dieses Modell ebenfalls die Berechnung von Hazard Ratios ermöglicht und somit einen direkten Vergleich der digitalisierten Hazardraten mit den originalen Hazardraten erlaubt.

Letztendlich handelt es sich bei dieser Methode um eine Schätzung der *Time-to-event* Daten der CVOTs. Jedoch konnte ein systematischer Vergleich zwischen den originalen Hazard Ratios mit den Hazard Ratios der digitalisierten *Time-to-event* Daten und mit den Hazard Ratios der digitalisierten *Time-to-event* Daten unter Annahme eines Weibull-Modells gebildet werden. Dieser Vergleich ist in Tabelle 4 für den primären Endpunkt und in Tabelle 5 für die Gesamtmortalität dargestellt. Zusätzlich wurden Streudiagramme für beide Endpunkte mit den ursprünglichen Hazard Ratios und den Hazard Ratios aus den extrahierten Daten unter Annahme eines Weibull-Modells

erstellt, um die Intra-Class-Korrelationskoeffizienten graphisch zu vergleichen (Abb. 10). Hier zeigt sich, dass die Abweichungen der Hazard Ratios aus den digitalisierten Daten im Vergleich zu den originalen Hazard Ratios vernachlässigbar gering sind, was die Validität der angewandten Methode bestätigt.

Die NNT wurden für jeden Zeitpunkt direkt aus der Weibull-Überlebensfunktion abgeleitet. Somit wird auch die funktionale Form der NNT-Dynamik auf der Zeitachse durch die parametrische Form bestimmt, was zu einer glatten und plausiblen Funktion führt. Es zeigt sich, dass die Weibull-Überlebensfunktionen sehr gut an die digitalisierten Überlebensdaten angepasst sind, was ebenfalls für die Validität der Methode spricht.

Die originalen Hazard Ratios aus den CVOTs wurden alle mit einem Cox-Modell unter der Annahme berechnet, dass die Hazards in den Behandlungsgruppen proportional sind. Diese Annahme, die dem Design der ursprünglichen Studien zugrunde lag, könnte potenziell falsch sein [97]. Wir sind der Ansicht, dass das von uns verwendete parametrische Weibull-Modell nicht so restriktiv ist wie die Annahme proportionaler Hazards und darüber hinaus die direkte Schätzung von Überlebenswahrscheinlichkeiten und NNTs ermöglicht.

Es wurden bereits Arbeiten veröffentlicht, in denen NNTs auf der Basis von digitalisierten Datensätzen geschätzt wurden [91, 92]. Mit dieser Arbeit werden diese Daten um weitere Punkte ergänzt und eine noch umfassendere Meta-Analyse geboten. Darüber hinaus kann die in dieser Arbeit verwendete Methode mit den Methoden bereits veröffentlichter Arbeiten verglichen werden, um ihre Validität zu bewerten.

Zusammenfassend ist festzuhalten, dass Davies et al. [91] und Ludwig et al. [92] andere Methoden zur Extraktion der *Time-to-event* Daten und andere Softwareprogramme zur Berechnung der NNTs verwendet haben. Die publizierten NNTs sind jedoch mit unseren Ergebnissen vergleichbar und unterstreichen damit die Konsistenz und Validität unserer Methode der numerischen Extraktion individueller Patientendaten aus den publizierten Kaplan-Meier-Überlebenszeitkurven der CVOTs.

In der vorliegenden Arbeit wurden 19 bis September 2020 publizierte CVOTs der drei Wirkstoffklassen DDP-4 Inhibitoren, GLP-1 Rezeptoragonisten und SGLT-2 Inhibitoren

analysiert und die *Time-to-event* Daten extrahiert. Aus den extrahierten Daten wurden die NNTs für den primären Endpunkt und die Gesamtmortalität berechnet und Meta-Analysen für jede Wirkstoffklasse durchgeführt.

In Abb. 8 wird die Meta-NNT für den primären Endpunkt und die Gesamtmortalität dargestellt. Bei einer medianen Gesamtnachbeobachtungszeit von 29 Monaten für den primären Endpunkt und 39 Monaten für die Gesamtmortalität liegen die geschätzten Meta-NNTs bei 100 (95%-KI: 60 - 303) bzw. 128 (95%-KI: 85 - 265).

Für die weitere Analyse wurden für jede der drei Wirkstoffklassen separate Meta-NNTs für beide Endpunkte gebildet. In Abb. 6 ist der primäre Endpunkt für die Wirkstoffklassen DDP-4 Inhibitoren, GLP-1 Rezeptorantagonisten und SGLT-2 Inhibitoren aufgeschlüsselt. Es zeigt sich, dass die Meta-NNT für die GLP-1 Rezeptorantagonisten und die SGLT-2 Inhibitoren einen vergleichbaren Kurvenverlauf innerhalb der Nachbeobachtungszeit aufweisen, während die Meta-NNT für die DDP-4 Inhibitoren einen geringeren Behandlungseffekt in Form einer höheren NNT und eines größeren 95%-Konfidenzintervalls zeigt. Abb. 7 zeigt die drei Medikamentenklassen DDP-4 Inhibitoren, GLP-1 Rezeptorantagonisten und SGLT-2 Inhibitoren für die Gesamtmortalität. Die Meta-NNT für GLP-1 Rezeptorantagonisten und SGLT-2 Inhibitoren zeigen für die Gesamtmortalität ebenfalls einen vergleichbaren Kurvenverlauf. Die Meta-NNT der DPP-4 Inhibitoren ist auch für diesen Endpunkt bedeutend schlechter.

In diesem Zusammenhang ist zu erwähnen, dass sich die Zusammensetzung der drei Meta-NNT hinsichtlich der Anzahl der Studien und der Anzahl der beobachteten Patienten unterscheidet (Tabelle 1). Dennoch ist festzuhalten, dass keine der 4 eingeschlossenen CVOTs mit DPP-4 Inhibitoren (CARMELINA [72], EXAMINE [73], SAVOR-TIMI53 [74] und TECOS [75]) einen signifikanten Vorteil für den primären Endpunkt oder die Gesamtmortalität nachweisen konnte. Dagegen konnten von den 7 CVOTs mit GLP-1 Rezeptoragonisten (ELIXA [76], EXSCEL [77], HARMONY [78], LEADER [79], PIONEER [80], REWIND [81] und SUSTAIN [82]) insgesamt vier (HARMONY [78], LEADER[79], REWIND[81] und SUSTAIN[82]) einen signifikanten Vorteil aufzeigen. Von den acht CVOTs mit SGLT-2 Inhibitoren (CANVAS [83], CREDENCE [84], DAPA-CKD [85], DAPA-HF [86], DECLARE-TIMI58 [87], EMPA-REG[13], EMPEROR-REDUCED [88] und VERTIS-CV [89]), die für den primären Endpunkt in unsere Arbeit eingeschlossen

wurden, zeigten sechs (CANVAS [83], CREDENCE [84], DAPA-CKD [85], DAPA-HF [86], EMPA-REG[13] und EMPEROR-REDUCED [88]) einen signifikanten Vorteil (Tabelle 1). Für den Endpunkt Gesamtmortalität ergibt sich für die eingeschlossenen Studien ein ähnliches Bild. Keine der drei CVOTs mit DDP-4 Inhibitoren konnte für die Gesamtmortalität einen signifikanten Vorteil demonstrieren. Von den drei eingeschlossenen CVOTs mit GLP-1 Rezeptoragonisten konnten zwei einen signifikanten Vorteil nachweisen (EXSCEL [77] und LEADER [79]) und von den 7 CVOTs mit einem SGLT-2 Inhibitor konnten drei (DAPA-CKD [85], DAPA-HF [86] und EMPA-REG [13]) einen signifikanten Vorteil für den Endpunkt Gesamtmortalität nachweisen. Die besseren Ergebnisse für die Meta-NTT der SGLT-2 Inhibitoren und GLP-1 Rezeptoragonisten waren unter diesem Gesichtspunkt zu erwarten. Allerdings erscheinen die Meta-NTT der SGLT-2 Inhibitoren und GLP-1 Rezeptorantagonisten für beide Endpunkte für sich genommen nicht so beeindruckend, wie sie als relative Risikoreduktion kommuniziert wurden.

Um die geschätzte NNT in einer Kosten-Nutzen-Analyse besser einordnen zu können, sind Vergleiche mit anderen Arzneimitteln, die ebenfalls zur Standardtherapie gehören, hilfreich. Beispielsweise können die berechneten NNTs mit Metformin verglichen werden, das seit Jahrzehnten als Standardmedikament in der oralen Stufentherapie des Diabetes mellitus Typ 2 eingesetzt wird.

Die *UK Prospective Diabetes Study Group* (UKPDS) publizierte 1998 Daten zur intensivierten Blutzuckerkontrolle mit Metformin versus konventioneller Insulintherapie bei übergewichtigen Patienten mit Diabetes mellitus Typ 2. In der prospektiven Studie zeigte sich im 10-Jahres-Follow-up für die Metformin-Interventionsgruppe eine NNT von 14 (95%-KI: 8 - 67) für die Gesamtmortalität und eine NNT von 16 (95%-KI: 9 - 78) für harte kardiovaskuläre Endpunkte wie Myokardinfarkte [98]. Fairerweise muss bei diesem Vergleich der vorhandene Zeithorizont berücksichtigt werden. In der Regel sind Medikamente, die sich bereits als Standardtherapie für eine Erkrankung etabliert haben, schon seit längerer Zeit unter Verwendung. Daraus ergibt sich im direkten Vergleich mit neuen Medikamenten ein Vorteil für die älteren Medikamente, da neue Medikamente häufig in Studien zusammen mit der bestehenden Standardtherapie gegen ein Placebo untersucht werden. Es ist davon auszugehen, dass

die absolute Risikoreduktion durch ein zusätzliches neues Medikament umso geringer ist, desto besser die bestehende Standardtherapie für eine Erkrankung bereits ist. In diesem Zusammenhang sind die relativen Effektmaße wie die Hazard Ratio im Vorteil, weil sie die Behandlungseffekte auf das Ausgangsrisiko der Placebo-Population normieren und damit zu relativ stabilen Ergebnissen über das gesamte Risikospektrum führen. Dies erleichtert den Vergleich verschiedener Studienpopulationen [94].

Darüber hinaus haben wir mit dem projizierten Lebenshorizont von 30 Jahren einen weiteren Aspekt zur Beurteilung der absoluten Behandlungseffekte hinzugefügt. Die Gesamtmortalität wurde gewählt, da sie keine konkurrierenden Endpunkte aufweist und zudem aus Patientensicht wahrscheinlich der wichtigste Endpunkt ist. In Abb. 9 ist der projizierte Verlauf der Meta-NNT für die drei Wirkstoffklassen dargestellt. Für alle drei Meta-NNT zeigt sich zu Beginn der Behandlung zunächst ein steiler Abfall der Meta-NNT, also eine Zunahme des absoluten Behandlungseffekts auf die Gesamtmortalität. Irgendwann erreicht der Behandlungseffekt sein Maximum, d. h. eine minimale Meta-NNT. Da die Sterbewahrscheinlichkeit naturgemäß mit der Zeit zunimmt, hat die Meta-NNT auch unter der Annahme eines konstanten Behandlungseffekts keinen monotonen zeitlichen Verlauf. Zum Zeitpunkt des maximalen Behandlungseffekts ist die Meta-NNT am niedrigsten und steigt im Zeitverlauf wieder an, da die Mortalitätswahrscheinlichkeit zunimmt.

4.1 Limitationen

An dieser Stelle wird auf einige Einschränkungen dieser Arbeit hingewiesen.

Die Auswahl der Studien erfolgte nicht auf der Grundlage einer systematischen Literaturrecherche oder eines zuvor veröffentlichten Studienprotokolls. Für diese Arbeit wurde auf eine systematische Literaturrecherche verzichtet, da bereits mehrere Übersichtsarbeiten zu diesem Thema vorliegen [1] und das Ziel dieser Arbeit die Schätzung der absoluten Behandlungseffekte darstellte und zusätzlich einen experimentellen Ansatz für die verwendete Methode verfolgte.

Die größte Einschränkung und Herausforderung dieser Arbeit bestand darin, dass kein Zugang zu den Primärdaten der CVOTs bestand und somit keine gepoolten Primärdaten

zur Verfügung standen. Die CVOTs haben zwar gemeinsam, dass es sich um große, multizentrische und gut konzipierte Studien handelt, jedoch sind sie aufgrund der flexiblen Anforderungen der Zulassungsbehörden nicht völlig einheitlich konzipiert worden. Unterschiede in den Studiencharakteristika führen daher zwangsläufig zu Verzerrungen in der Meta-Analyse. Zusätzlich ist zu berücksichtigen, dass das Basisrisiko der eingeschlossenen Studien für die untersuchten Endpunkte natürlich vom zugrunde liegenden kardiovaskulären Risiko abhängt und für die Beurteilung der absoluten Behandlungseffekte eine Rolle spielt. Auch hier gibt es Unterschiede zwischen den Studien, die auf die unterschiedliche Zusammensetzung der untersuchten Kohorten zurückzuführen sind (Tabelle 1). Letztlich handelt es sich jedoch um eine Verzerrung, die in jeder Meta-Analyse aufgrund des fehlenden *Poolings* von Primärdaten auftritt. Darüber hinaus ist es aufgrund der fehlenden Primärdaten der CVOTs nicht möglich, eine Assoziation zwischen Patientencharakteristika und den Endpunkten herzustellen. Allerdings wurden die verwendeten Studiendaten auch von anderen Studienautoren verwendet und in mehreren Meta-Analysen gepoolt. Insofern wird das Ausmaß der Heterogenität der CVOTs auch von vielen Studienautoren als akzeptabel angesehen.

Eine weitere Einschränkung ergibt sich aus dem zugrundeliegenden Design der CVOTs. Für die Abschätzung der absoluten Risikoreduktion oder der NNT sind Studien mit multifaktoriellen Therapieansätzen wahrscheinlich besser geeignet [99]. Die Bewertung anderer Risikofaktoren, die einen Einfluss auf das kardiovaskuläre Gesamtrisikoprofil haben könnten, wurde im Studiendesign der CVOTs nicht genau charakterisiert. Aus diesem Grund konnte keine Adjustierung für parallel konkurrierende Risiken für die Gesamtmortalität vorgenommen werden. Bei der Interpretation der Ergebnisse muss daher berücksichtigt werden, dass es sich um das durchschnittliche kardiovaskuläre Risiko aller in die Studien eingeschlossenen Patienten handelt.

Die Kommunikation mit Patienten und anderen medizinischen Laien über Behandlungseffekte von Therapien ist für Ärzte oft problematisch, da meist große Wissenslücken zu überbrücken sind und die behandelnden Ärzte selbst nicht selten mit der Interpretation von Studienergebnissen überfordert sind, wenn diese nur in Form von relativen Risikoreduktionen berichtet werden [100].

Aus diesem Grund besteht ein großes wissenschaftliches Interesse an der Erforschung von Faktoren, die eine gute Interpretation von Studiendaten und eine gute Therapietreue der Patienten bedingen können. Frühere Arbeiten haben auch gezeigt, dass auch Patienten Schwierigkeiten haben können, NNTs zu verstehen und zu interpretieren [101, 102]. Daher muss in diesem Zusammenhang auch kritisch angemerkt werden, dass die Einordnung absoluter Behandlungseffekte auch für Patienten und andere medizinische Laien ebenfalls schwierig sein kann. Einige Patienten könnten z. B. die NNT als Lotterie interpretieren [103]. Eine NNT von z. B. 50 für einen Endpunkt bedeutet, dass eine von 50 Personen einen Nutzen in Bezug auf den untersuchten Endpunkt hat. Patienten könnten diese Definition so interpretieren, dass 49 Patienten überhaupt keinen Behandlungseffekt haben werden und nur eine Person einen Behandlungseffekt haben wird. Bei kardiovaskulären Endpunkten wie z. B. dem primären Endpunkt der CVOT ist es aus klinischer Sicht realistischer, dass die meisten Patienten einen gewissen Nutzen von der Behandlung haben werden, da die antiglykämische Wirkung der untersuchten neuen Antidiabetika aus klinischer Sicht nicht nur bei einem Patienten, sondern bei allen 50 behandelten Patienten wirksam ist und somit ein gewisser Vorteil bei allen Patienten vorliegt.

Aus diesen Gründen stimmen wir mit anderen Autoren überein, dass die NNT in erster Linie in der Kommunikation mit Fachleuten verwendet werden sollte [104, 105].

4.2 Schlussfolgerungen

In dieser Meta-Analyse der digitalisierten individuellen Patientendaten aus den *Cardiovascular Outcome Trials* berichten wir für die drei relevanten Wirkstoffklassen (DDP-4 Inhibitoren, GLP-1 Rezeptoragonisten und SGLT-2 Inhibitoren) über absolute Behandlungseffekte in Form der Number-needed-to-treat. Im Ergebnis erfüllten alle neuen antidiabetischen Medikamente die initialen Anforderungen der FDA [2] an die kardiovaskuläre Sicherheit und einige CVOT konnten mit ihren Studienergebnissen zur Änderung mehrerer internationaler Leitlinien und zu einem Paradigmenwechsel in der Behandlung von Patienten mit Diabetes mellitus Typ 2 und Herzinsuffizienzen beitragen. Allerdings publizierten die meisten Studienautoren der CVOTs ihre Studienergebnisse

nur mit relativen Risikomaßen, obwohl die neuen FDA-Richtlinien [2] auch die Berichterstattung absoluter Risikomaße empfehlen.

In dieser Arbeit wurden Schätzungen der NNT für den primären Endpunkt und die Gesamtmortalität für jede eingeschlossene Studie, für eine Meta-NNT der drei Wirkstoffklassen, für eine Meta-NNT der beiden Endpunkte und für eine projizierte Meta-NNT der Gesamtmortalität für die drei Wirkstoffklassen durchgeführt. Es konnte gezeigt werden, dass die absoluten Behandlungseffekte, ausgedrückt als NNT, weniger beeindruckende Ergebnisse liefern als die Darstellung als relative Behandlungseffekte. Die Meta-NNT der drei Wirkstoffklassen zeigen deutliche Klasseneffekte. Die Wirkstoffklassen der SGLT-2 Inhibitoren und GLP-1 Rezeptoragonisten weisen im Vergleich zu den DDP-4 Inhibitoren eine signifikant niedrigere Meta-NNT auf. Dieser Klasseneffekt konnte auch bei der Darstellung der Studienergebnisse in Form von relativen Risikomaßen nachgewiesen werden. Für die Interpretation der Wirksamkeit zwischen den drei untersuchten Wirkstoffklassen (SGLT-2 Inhibitoren, GLP-1 Rezeptoragonisten und DDP-4 Inhibitoren) sind die Meta-NNT in dieser Hinsicht nicht hilfreich, wohl aber im Vergleich zu anderen Wirkstoffklassen.

Im Rahmen dieser Arbeit wurde eine Methode zur Schätzung der absoluten Risikoreduktion und der NNT aus digitalisierten Patientendaten in *Time-to-event* Studien entwickelt. Dazu wurden Informationen aus publizierten Studienergebnissen verwendet. Die Digitalisierungssoftware WebPlotDigitizer [63] und der R-Algorithmus von Guyot et al. [64] hatten bereits in früheren Arbeiten ihre Validität und Reliabilität unter Beweis gestellt. Auch im Rahmen dieser Arbeit konnte die Validität der Methode zur Schätzung der absoluten Risikoreduktion aus digitalisierten Daten auf verschiedenen Ebenen bestätigt werden.

Zusammenfassend glauben wir, dass es für Ärzte, andere Gesundheitsberufe und politische Entscheidungsträger notwendig ist, die absoluten Behandlungseffekte zu kennen, um die Studienergebnisse vollständig verstehen und interpretieren zu können. Daher sollten Studienautoren ihre Ergebnisse auch in absoluten Effektmaßen angeben und die Zulassungsbehörden sollten weiterhin auf der Veröffentlichung von relativen und absoluten Effektmaßen bestehen.

5 Literatur- und Quellenverzeichnis

1. Cefalu, W.T., et al., *Cardiovascular Outcomes Trials in Type 2 Diabetes: Where Do We Go From Here? Reflections From a Diabetes Care Editors' Expert Forum*. Diabetes Care, 2018. **41**(1): p. 14-31.
2. Administration, U.S.F.a.D. *Guidance for industry: diabetes mellitus—evaluating cardiovascular risk in new antidiabetic therapies to treat type 2 diabetes* 2008; Available from: <https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071627.pdf>.
3. Nissen, S.E. and K. Wolski, *Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes*. N Engl J Med, 2007. **356**(24): p. 2457-71.
4. Diabetes, C., et al., *The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus*. N Engl J Med, 1993. **329**(14): p. 977-86.
5. Terry, T., et al., *Does aggressive glycemc control benefit macrovascular and microvascular disease in type 2 diabetes? Insights from ACCORD, ADVANCE, and VADT*. Curr Cardiol Rep, 2012. **14**(1): p. 79-88.
6. *Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33)*. UK Prospective Diabetes Study (UKPDS) Group. Lancet, 1998. **352**(9131): p. 837-53.
7. *Effect of intensive blood-glucose control with metformin on complications in overweight patients with type 2 diabetes (UKPDS 34)*. UK Prospective Diabetes Study (UKPDS) Group. Lancet, 1998. **352**(9131): p. 854-65.
8. Gerstein, H.C., et al., *Nine-Year Effects of 3.7 Years of Intensive Glycemic Control on Cardiovascular Outcomes*. Diabetes Care, 2016. **39**(5): p. 701-708.
9. Hayward, R.A., P.D. Reaven, and N.V. Emanuele, *Follow-up of Glycemic Control and Cardiovascular Outcomes in Type 2 Diabetes REPLY*. New England Journal of Medicine, 2015. **373**(10): p. 978-978.
10. Regier, E.E., M.V. Venkat, and K.L. Close, *More Than 7 Years of Hindsight: Revisiting the FDA's 2008 Guidance on Cardiovascular Outcomes Trials for Type 2 Diabetes Medications*. Clin Diabetes, 2016. **34**(4): p. 173-180.
11. Holman, R.R., H. Sourij, and R.M. Califf, *Cardiovascular outcome trials of glucose-lowering drugs or strategies in type 2 diabetes*. Lancet, 2014. **383**(9933): p. 2008-17.
12. John, M., et al., *Cardiovascular outcome trials for anti-diabetes medication: A holy grail of drug development?* Indian Heart J, 2016. **68**(4): p. 564-71.
13. Zinman, B., et al., *Empagliflozin, Cardiovascular Outcomes, and Mortality in Type 2 Diabetes*. N Engl J Med, 2015. **373**(22): p. 2117-28.
14. Buse, J.B., et al., *2019 Update to: Management of Hyperglycemia in Type 2 Diabetes, 2018. A Consensus Report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD)*. Diabetes Care, 2020. **43**(2): p. 487-493.

15. Cosentino, F., et al., *2019 ESC Guidelines on diabetes, pre-diabetes, and cardiovascular diseases developed in collaboration with the EASD*. Eur Heart J, 2020. **41**(2): p. 255-323.
16. American Diabetes, A., *2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2018*. Diabetes Care, 2018. **41**(Suppl 1): p. S13-S27.
17. American Diabetes, A., *2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2021*. Diabetes Care, 2021. **44**(Suppl 1): p. S15-S33.
18. Harreiter, J. and M. Roden, [*Diabetes mellitus-Definition, classification, diagnosis, screening and prevention (Update 2019)*]. Wien Klin Wochenschr, 2019. **131**(Suppl 1): p. 6-15.
19. Saeedi, P., et al., *Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9(th) edition*. Diabetes Res Clin Pract, 2019. **157**: p. 107843.
20. Guariguata, L., et al., *Global estimates of diabetes prevalence for 2013 and projections for 2035*. Diabetes Res Clin Pract, 2014. **103**(2): p. 137-49.
21. Weiss, M., D.F. Steiner, and L.H. Philipson, *Insulin Biosynthesis, Secretion, Structure, and Structure-Activity Relationships*, in *Endotext*, K.R. Feingold, et al., Editors. 2000: South Dartmouth (MA).
22. Rutter, G.A., et al., *Pancreatic beta-cell identity, glucose sensing and the control of insulin secretion*. Biochem J, 2015. **466**(2): p. 203-18.
23. Kiselyov, V.V., et al., *Harmonic oscillator model of the insulin and IGF1 receptors' allosteric binding and activation*. Mol Syst Biol, 2009. **5**: p. 243.
24. Ho, C.K., G. Sriram, and K.M. Dipple, *Insulin sensitivity predictions in individuals with obesity and type II diabetes mellitus using mathematical model of the insulin signal transduction pathway*. Molecular Genetics and Metabolism, 2016. **119**(3): p. 288-292.
25. Kavanagh Williamson, M., et al., *Upregulation of Glucose Uptake and Hexokinase Activity of Primary Human CD4+ T Cells in Response to Infection with HIV-1*. Viruses, 2018. **10**(3).
26. Topp, B.G., L.L. Atkinson, and D.T. Finegood, *Dynamics of insulin sensitivity, beta-cell function, and beta-cell mass during the development of diabetes in falfa rats*. American Journal of Physiology-Endocrinology and Metabolism, 2007. **293**(6): p. E1730-E1735.
27. Faerch, K., et al., *Insulin Resistance Is Accompanied by Increased Fasting Glucagon and Delayed Glucagon Suppression in Individuals With Normal and Impaired Glucose Regulation*. Diabetes, 2016. **65**(11): p. 3473-3481.
28. Czech, M.P., *Insulin action and resistance in obesity and type 2 diabetes*. Nat Med, 2017. **23**(7): p. 804-814.
29. Rawshani, A., et al., *Mortality and Cardiovascular Disease in Type 1 and Type 2 Diabetes*. N Engl J Med, 2017. **376**(15): p. 1407-1418.
30. Matheus, A.S., et al., *Impact of diabetes on cardiovascular disease: an update*. Int J Hypertens, 2013. **2013**: p. 653789.
31. American Diabetes, A., *Economic costs of diabetes in the U.S. In 2007*. Diabetes Care, 2008. **31**(3): p. 596-615.

32. American Diabetes Association Professional Practice, C., et al., 9. *Pharmacologic Approaches to Glycemic Treatment: Standards of Medical Care in Diabetes-2022*. Diabetes Care, 2022. **45**(Suppl 1): p. S125-S143.
33. Weykamp, C., *HbA1c: a review of analytical and clinical aspects*. Ann Lab Med, 2013. **33**(6): p. 393-400.
34. Ziegler, D., et al., *Screening, diagnosis and management of diabetic sensorimotor polyneuropathy in clinical practice: International expert consensus recommendations*. Diabetes Res Clin Pract, 2022. **186**: p. 109063.
35. Balk, E.M., et al., *Combined Diet and Physical Activity Promotion Programs to Prevent Type 2 Diabetes Among Persons at Increased Risk: A Systematic Review for the Community Preventive Services Task Force*. Ann Intern Med, 2015. **163**(6): p. 437-51.
36. Lean, M.E., et al., *Primary care-led weight management for remission of type 2 diabetes (DiRECT): an open-label, cluster-randomised trial*. Lancet, 2018. **391**(10120): p. 541-551.
37. Martinez-Gonzalez, M.A., E. Ros, and R. Estruch, *Primary Prevention of Cardiovascular Disease with a Mediterranean Diet Supplemented with Extra-Virgin Olive Oil or Nuts*. N Engl J Med, 2018. **379**(14): p. 1388-1389.
38. Locke, A., J. Schneiderhan, and S.M. Zick, *Diets for Health: Goals and Guidelines*. Am Fam Physician, 2018. **97**(11): p. 721-728.
39. Jabardo-Camprubi, G., et al., *Drop-out ratio between moderate to high-intensity physical exercise treatment by patients with, or at risk of, type 2 diabetes mellitus: A systematic review and meta-analysis*. Physiol Behav, 2020. **215**: p. 112786.
40. Sluik, D., et al., *Physical Activity and Mortality in Individuals With Diabetes Mellitus: A Prospective Study and Meta-analysis*. Arch Intern Med, 2012. **172**(17): p. 1285-95.
41. Kar, D., et al., *Association of smoking and cardiometabolic parameters with albuminuria in people with type 2 diabetes mellitus: a systematic review and meta-analysis*. Acta Diabetol, 2019. **56**(8): p. 839-850.
42. Nathan, D.M. and D.E.R. Group, *The diabetes control and complications trial/epidemiology of diabetes interventions and complications study at 30 years: overview*. Diabetes Care, 2014. **37**(1): p. 9-16.
43. *The effect of intensive diabetes treatment on the progression of diabetic retinopathy in insulin-dependent diabetes mellitus. The Diabetes Control and Complications Trial*. Arch Ophthalmol, 1995. **113**(1): p. 36-51.
44. *Effect of intensive diabetes treatment on the development and progression of long-term complications in adolescents with insulin-dependent diabetes mellitus: Diabetes Control and Complications Trial*. Diabetes Control and Complications Trial Research Group. J Pediatr, 1994. **125**(2): p. 177-88.
45. Holman, R.R., et al., *10-year follow-up of intensive glucose control in type 2 diabetes*. N Engl J Med, 2008. **359**(15): p. 1577-89.
46. Mellbin, L.G., et al., *Prognostic implications of hypoglycaemic episodes during hospitalisation for myocardial infarction in patients with type 2 diabetes: a report from the DIGAMI 2 trial*. Heart, 2009. **95**(9): p. 721-7.

47. ElSayed, N.A., et al., *9. Pharmacologic Approaches to Glycemic Treatment: Standards of Care in Diabetes-2023*. Diabetes Care, 2023. **46**(Suppl 1): p. S140-S157.
48. American Diabetes Association Professional Practice, C., *9. Pharmacologic Approaches to Glycemic Treatment: Standards of Medical Care in Diabetes-2022*. Diabetes Care, 2022. **45**(Suppl 1): p. S125-S143.
49. Maruthur, N.M., et al., *Diabetes Medications as Monotherapy or Metformin-Based Combination Therapy for Type 2 Diabetes: A Systematic Review and Meta-analysis*. Ann Intern Med, 2016. **164**(11): p. 740-51.
50. Matthews, D.R., et al., *Glycaemic durability of an early combination therapy with vildagliptin and metformin versus sequential metformin monotherapy in newly diagnosed type 2 diabetes (VERIFY): a 5-year, multicentre, randomised, double-blind trial*. Lancet, 2019. **394**(10208): p. 1519-1529.
51. King, N.B., S. Harper, and M.E. Young, *Use of relative and absolute effect measures in reporting health inequalities: structured review*. Bmj-British Medical Journal, 2012. **345**.
52. Hasan, H., K. Goddard, and A.F. Howard, *Utility of the number needed to treat in paediatric haematological cancer randomised controlled treatment trials: a systematic review*. BMJ Open, 2019. **9**(2): p. e022839.
53. Heneghan, C. and K.R. Mahtani, *Absolute effects of statins in the elderly*. BMJ Evid Based Med, 2019. **24**(5): p. 200-202.
54. Schulz, K.F., et al., *CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials*. Int J Surg, 2011. **9**(8): p. 672-7.
55. Guyatt, G.H., et al., *GRADE guidelines: 12. Preparing summary of findings tables-binary outcomes*. J Clin Epidemiol, 2013. **66**(2): p. 158-72.
56. Sprenger, J. and J. Stegenga, *Three Arguments for Absolute Outcome Measures*. Philosophy of Science, 2017. **84**(5): p. 840-852.
57. Stegenga, J., *Measuring effectiveness*. Stud Hist Philos Biol Biomed Sci, 2015. **54**: p. 62-71.
58. Bobbio, M., B. Demichelis, and G. Giustetto, *Completeness of reporting trial results: effect on physicians' willingness to prescribe*. Lancet, 1994. **343**(8907): p. 1209-11.
59. Frick, M.H., et al., *Helsinki Heart Study: primary-prevention trial with gemfibrozil in middle-aged men with dyslipidemia. Safety of treatment, changes in risk factors, and incidence of coronary heart disease*. N Engl J Med, 1987. **317**(20): p. 1237-45.
60. Sorensen, L., et al., *Laypersons' understanding of relative risk reductions: randomised cross-sectional study*. BMC Med Inform Decis Mak, 2008. **8**: p. 31.
61. Malenka, D.J., et al., *The framing effect of relative and absolute risk*. J Gen Intern Med, 1993. **8**(10): p. 543-8.
62. Ferrannini, E. and J. Rosenstock, *Clinical Translation of Cardiovascular Outcome Trials in Type 2 Diabetes: Is There More or Is There Less Than Meets the Eye?* Diabetes Care, 2021. **44**(3): p. 641-646.
63. Rohatgi, A. *WebPlotDigitizer [Internet]*. <https://automeris.io/WebPlotDigitizer>.
64. Guyot, P., et al., *Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves*. BMC Med Res Methodol, 2012. **12**: p. 9.

65. Moeyaert, M., D. Maggin, and J. Verkuilen, *Reliability, Validity, and Usability of Data Extraction Programs for Single-Case Research Designs*. Behav Modif, 2016. **40**(6): p. 874-900.
66. Drevon, D., S.R. Fursa, and A.L. Malcolm, *Intercoder Reliability and Validity of WebPlotDigitizer in Extracting Graphed Data*. Behav Modif, 2017. **41**(2): p. 323-339.
67. Saluja, R., et al., *Estimating hazard ratios from published Kaplan-Meier survival curves: A methods validation study*. Res Synth Methods, 2019. **10**(3): p. 465-475.
68. Collett, D., *Modelling survival data in medical research*. Third edition. ed. Chapman & Hall/CRC texts in statistical science series. 2015, Boca Raton: CRC Press, Taylor & Francis Group. pxvi, 532 pages.
69. Kuss, O., et al., *Absolute treatment effects for the primary outcome and all-cause mortality in the cardiovascular outcome trials of new antidiabetic drugs: a meta-analysis of digitalized individual patient data*. Acta Diabetol, 2022. **59**(10): p. 1349-1359.
70. Akbulut C. , K.O., *Data set from "Absolute treatment effects for the primary outcome and all-cause mortality in the cardiovascular outcome trials of new antidiabetic drugs—a meta-analysis of digitalized individual patient data" [Data set]*
<https://doi.org/10.5281/zenodo.6630421>. 2022.
71. Laupacis, A., D.L. Sackett, and R.S. Roberts, *An assessment of clinically useful measures of the consequences of treatment*. N Engl J Med, 1988. **318**(26): p. 1728-33.
72. Rosenstock, J., et al., *Effect of Linagliptin vs Placebo on Major Cardiovascular Events in Adults With Type 2 Diabetes and High Cardiovascular and Renal Risk The CARMELINA Randomized Clinical Trial*. Jama-Journal of the American Medical Association, 2019. **321**(1): p. 69-79.
73. White, W.B., et al., *Alogliptin after Acute Coronary Syndrome in Patients with Type 2 Diabetes*. New England Journal of Medicine, 2013. **369**(14): p. 1327-1335.
74. Scirica, B.M., et al., *Saxagliptin and cardiovascular outcomes in patients with type 2 diabetes mellitus*. N Engl J Med, 2013. **369**(14): p. 1317-26.
75. Green, J.B., et al., *Effect of Sitagliptin on Cardiovascular Outcomes in Type 2 Diabetes*. N Engl J Med, 2015. **373**(3): p. 232-42.
76. Pfeffer, M.A., et al., *Lixisenatide in Patients with Type 2 Diabetes and Acute Coronary Syndrome*. N Engl J Med, 2015. **373**(23): p. 2247-57.
77. Holman, R.R., et al., *Effects of Once-Weekly Exenatide on Cardiovascular Outcomes in Type 2 Diabetes*. N Engl J Med, 2017. **377**(13): p. 1228-1239.
78. Hernandez, A.F., et al., *Albiglutide and cardiovascular outcomes in patients with type 2 diabetes and cardiovascular disease (Harmony Outcomes): a double-blind, randomised placebo-controlled trial*. The Lancet, 2018. **392**(10157): p. 1519-1529.
79. Marso, S.P., et al., *Liraglutide and Cardiovascular Outcomes in Type 2 Diabetes*. N Engl J Med, 2016. **375**(4): p. 311-22.
80. Husain, M., et al., *Oral Semaglutide and Cardiovascular Outcomes in Patients with Type 2 Diabetes*. N Engl J Med, 2019. **381**(9): p. 841-851.

81. Gerstein, H.C., et al., *Dulaglutide and cardiovascular outcomes in type 2 diabetes (REWIND): a double-blind, randomised placebo-controlled trial*. The Lancet, 2019. **394**(10193): p. 121-130.
82. Marso, S.P., et al., *Semaglutide and Cardiovascular Outcomes in Patients with Type 2 Diabetes*. N Engl J Med, 2016. **375**(19): p. 1834-1844.
83. Neal, B., et al., *Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes*. New England Journal of Medicine, 2017. **377**(7): p. 644-657.
84. Perkovic, V., et al., *Canagliflozin and renal outcomes in type 2 diabetes: results from the CANVAS Program randomised clinical trials*. The Lancet Diabetes & Endocrinology, 2018. **6**(9): p. 691-704.
85. Heerspink, H.J.L., et al., *Dapagliflozin in Patients with Chronic Kidney Disease*. N Engl J Med, 2020. **383**(15): p. 1436-1446.
86. McMurray, J.J.V., K.F. Docherty, and P.S. Jhund, *Dapagliflozin in Patients with Heart Failure and Reduced Ejection Fraction. Reply*. N Engl J Med, 2020. **382**(10): p. 973.
87. Wiviott, S.D., et al., *Dapagliflozin and Cardiovascular Outcomes in Type 2 Diabetes*. N Engl J Med, 2019. **380**(4): p. 347-357.
88. Packer, M., et al., *Cardiovascular and Renal Outcomes with Empagliflozin in Heart Failure*. N Engl J Med, 2020. **383**(15): p. 1413-1424.
89. Cannon, C.P., et al., *Cardiovascular Outcomes with Ertugliflozin in Type 2 Diabetes*. N Engl J Med, 2020. **383**(15): p. 1425-1435.
90. Wan, X., et al., *Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range*. BMC Med Res Methodol, 2014. **14**: p. 135.
91. Davies, M.J., et al., *Number needed to treat in cardiovascular outcome trials of glucagon-like peptide-1 receptor agonists: A systematic review with temporal analysis*. Diabetes Obes Metab, 2020. **22**(9): p. 1670-1677.
92. Ludwig, L., P. Darmon, and B. Guerci, *Computing and interpreting the Number Needed to Treat for Cardiovascular Outcomes Trials : Perspective on GLP-1 RA and SGLT-2i therapies*. Cardiovasc Diabetol, 2020. **19**(1): p. 65.
93. McDonagh, T.A., et al., *2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: Developed by the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) With the special contribution of the Heart Failure Association (HFA) of the ESC*. Rev Esp Cardiol (Engl Ed), 2022. **75**(6): p. 523.
94. Schulz, K.F., D.G. Altman, and D. Moher, *CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials*. J Pharmacol Pharmacother, 2010. **1**(2): p. 100-7.
95. Verma, S., et al., *Effects of Liraglutide on Cardiovascular Outcomes in Patients With Type 2 Diabetes Mellitus With or Without History of Myocardial Infarction or Stroke*. Circulation, 2018. **138**(25): p. 2884-2894.
96. Fitchett, D., et al., *Heart failure outcomes with empagliflozin in patients with type 2 diabetes at high cardiovascular risk: results of the EMPA-REG OUTCOME(R) trial*. Eur Heart J, 2016. **37**(19): p. 1526-34.
97. Gregson, J., et al., *Nonproportional Hazards for Time-to-Event Outcomes in Clinical Trials: JACC Review Topic of the Week*. J Am Coll Cardiol, 2019. **74**(16): p. 2102-2112.

98. Gerstein, H.C., *Metformin reduced diabetes-related end points and all-cause mortality in overweight patients with type 2 diabetes*. Lancet, 1998.
99. Sasso, F.C., et al., *Efficacy and durability of multifactorial intervention on mortality and MACEs: a randomized clinical trial in type-2 diabetic kidney disease*. Cardiovasc Diabetol, 2021. **20**(1): p. 145.
100. Weir, I.R., et al., *Interpretation of time-to-event outcomes in randomized trials: an online randomized experiment*. Ann Oncol, 2019. **30**(1): p. 96-102.
101. Christensen, P.M., et al., *A randomized trial of laypersons' perception of the benefit of osteoporosis therapy: number needed to treat versus postponement of hip fracture*. Clin Ther, 2003. **25**(10): p. 2575-85.
102. Albarqouni, L., J. Doust, and P. Glasziou, *Patient preferences for cardiovascular preventive medication: a systematic review*. Heart, 2017. **103**(20): p. 1578-1586.
103. Hansen, M.R., et al., *Postponement of Death by Statin Use: a Systematic Review and Meta-analysis of Randomized Clinical Trials*. J Gen Intern Med, 2019. **34**(8): p. 1607-1614.
104. Zipkin, D.A., et al., *Evidence-based risk communication: a systematic review*. Ann Intern Med, 2014. **161**(4): p. 270-80.
105. Kloecker, D.E., et al., *Uses and Limitations of the Restricted Mean Survival Time: Illustrative Examples From Cardiovascular Outcomes and Mortality Trials in Type 2 Diabetes*. Ann Intern Med, 2020. **172**(8): p. 541-552.

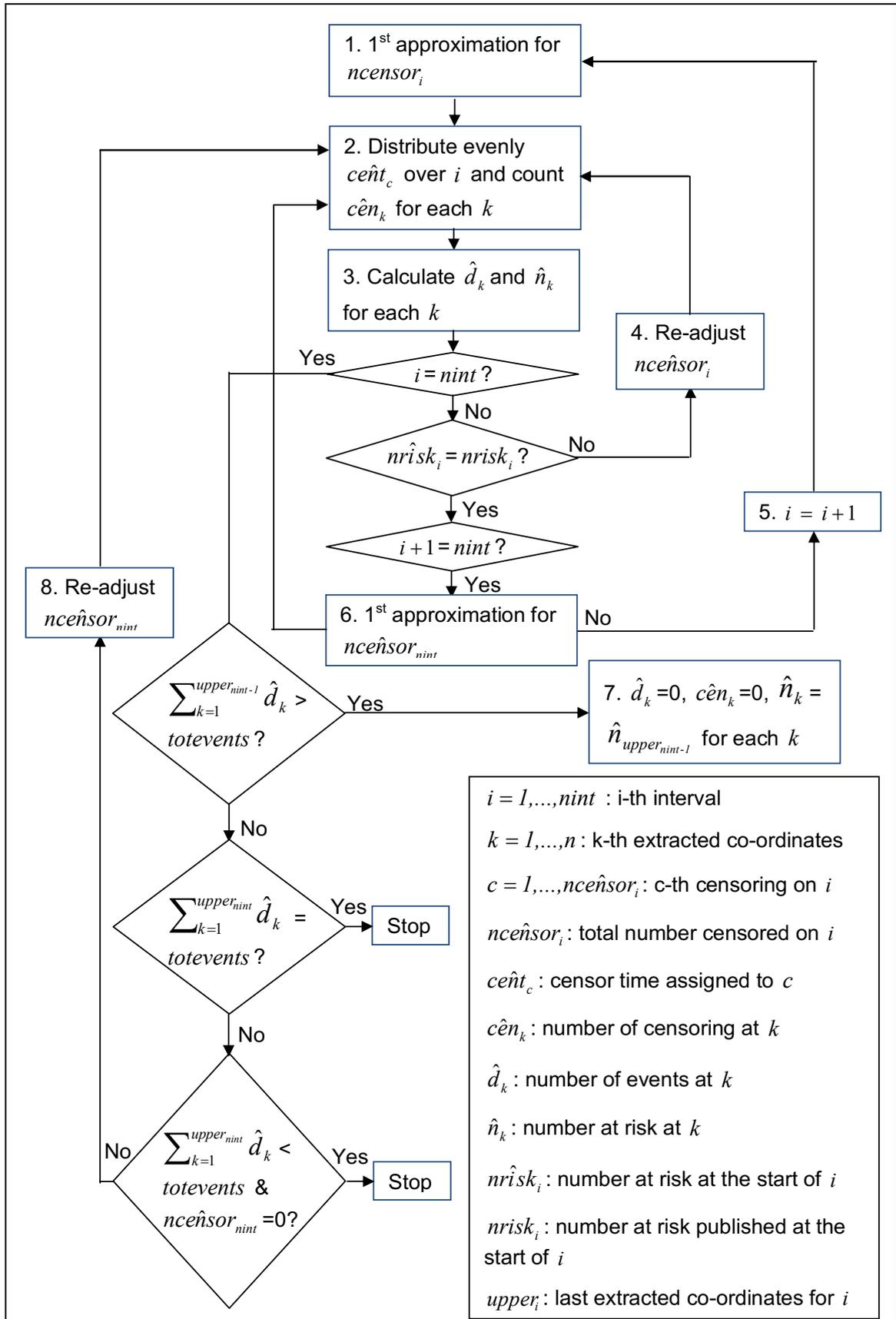
6 Anhang

6.1 Guyot et al. [64]: Flowchart of the algorithm (all information case)

6.2 Guyot et al. [64]: Additional file 1: The algorithm (R coding)

6.3. Vollständige Datensätze der digitalisierten Patientendaten in einer öffentlichen Datenbank [70]

6.1 Guyot et al. [64]: Flowchart of the algorithm (all information case)



6.2 Guyot et al. [64]: Additional file 1: The algorithm (R coding)

```

#Algorithm to create a raw dataset from Digizeit readings from a Kaplan-Meier curve

library("MASS")
library("splines")
library("survival")

###FUNCTION INPUTS
path<-"C:\PHD\algorithm\reliability exercice\"
digisurvfile<-"data initials study2 figA arm1 time1.txt"           #Input survival times from graph reading
nriskfile<-"nrisk study2 figA arm1 time1.txt"                   #Input reported number at risk
KMdatafile<-"KMdata study2 figA arm1 time1 ne.txt"             #Output file events and cens
KMdataIPDfile<-"KMdataIPD study2 figA arm1 time1 ne.txt"       #Output file for IPD
tot.events<-"NA"                                               #tot.events = total no. of events reported. If not reported, then tot.events="NA"
arm.id<-1 #arm indicator
###END FUNCTION INPUTS

#Read in survival times read by digizeit
digizeit<- read.table(paste(path,digisurvfile,sep=""),header=TRUE)
t.S<-digizeit[,2]
S<-digizeit[,3]

#Read in published numbers at risk, n.risk, at time, t.risk, lower and upper
# indexes for time interval
pub.risk<-read.table(paste(path,nriskfile,sep=""),header=TRUE)
t.risk<-pub.risk[,2]
lower<-pub.risk[,3]
upper<-pub.risk[,4]
n.risk<-pub.risk[,5]
n.int<-length(n.risk)
n.t<- upper[n.int]

#Initialise vectors
arm<-rep(arm.id,n.risk[1])
n.censor<- rep(0,(n.int-1))
n.hat<-rep(n.risk[1]+1,n.t)
cen<-rep(0,n.t)
d<-rep(0,n.t)
KM.hat<-rep(1,n.t)
last.i<-rep(1,n.int)
sumdL<-0

if (n.int > 1){
#Time intervals 1,...,(n.int-1)
for (i in 1:(n.int-1)){
#First approximation of no. censored on interval i
n.censor[i]<- round(n.risk[i]*S[lower[i+1]]/S[lower[i]]- n.risk[i+1])

#Adjust tot. no. censored until n.hat = n.risk at start of interval (i+1)
while((n.hat[lower[i+1]]>n.risk[i+1])|((n.hat[lower[i+1]]<n.risk[i+1])&&(n.censor[i]>0))){
if (n.censor[i]<=0){
cen[lower[i]:upper[i]]<-0
n.censor[i]<-0
}
if (n.censor[i]>0){
cen.t<-rep(0,n.censor[i])
for (j in 1:n.censor[i]){
cen.t[j]<- t.S[lower[i]] +
j*(t.S[lower[(i+1)]]-t.S[lower[i]])/(n.censor[i]+1)
}
#Distribute censored observations evenly over time. Find no. censored on each time interval.
cen[lower[i]:upper[i]]<-hist(cen.t,breaks=t.S[lower[i]:lower[(i+1)]],
plot=F)$counts
}
}

#Find no. events and no. at risk on each interval to agree with K-M estimates read from curves
n.hat[lower[i]]<-n.risk[i]
last<-last.i[i]
for (k in lower[i]:upper[i]){
if (i==1 & k==lower[i]){
d[k]<-0
KM.hat[k]<-1
}
else {
d[k]<-round(n.hat[k]*(1-(S[k]/KM.hat[last])))
}
}
}
}

```

```

                                KM.hat[k]<-KM.hat[last]*(1-(d[k]/n.hat[k]))
                                }
                                n.hat[k+1]<-n.hat[k]-d[k]-cen[k]
if (d[k] != 0) last<-k
                                }
                                n.censor[i]<- n.censor[i]+(n.hat[lower[i+1]]-n.risk[i+1])
                                }
if (n.hat[lower[i+1]]<n.risk[i+1]) n.risk[i+1]<-n.hat[lower[i+1]]
last.iff(i+1)<-last
}
}

#Time interval n.int
if (n.int>1){
#Assume same censor rate as average over previous time intervals.
n.censor[n.int]<- min(round(sum(n.censor[1:(n.int-1)]*(t.S[upper[n.int]]-
t.S[lower[n.int]])/(t.S[upper[(n.int-1)]]-t.S[lower[1]])), n.risk[n.int])
}
if (n.int==1){n.censor[n.int]<-0}
if (n.censor[n.int] <= 0){
cen[lower[n.int]:(upper[n.int]-1)]<-0
n.censor[n.int]<-0
}
if (n.censor[n.int]>0){
cen.t<-rep(0,n.censor[n.int])
for (j in 1:n.censor[n.int]){
cen.t[j]<- t.S[lower[n.int]] +
j*(t.S[upper[n.int]]-t.S[lower[n.int]])/(n.censor[n.int]+1)
}
cen[lower[n.int]:(upper[n.int]-1)]<-hist(cen.t,breaks=t.S[lower[n.int]:upper[n.int]],
plot=F)$counts
}

#Find no. events and no. at risk on each interval to agree with K-M estimates read from curves
n.hat[lower[n.int]]<-n.risk[n.int]
last<-last.iff(n.int)
for (k in lower[n.int]:upper[n.int]){
if (KM.hat[last] !=0){
d[k]<-round(n.hat[k]*(1-(S[k]/KM.hat[last]))) else {d[k]<-0}
KM.hat[k]<-KM.hat[last]*(1-(d[k]/n.hat[k]))
n.hat[k+1]<-n.hat[k]-d[k]-cen[k]
#No. at risk cannot be negative
if (n.hat[k+1] < 0) {
n.hat[k+1]<-0
cen[k]<-n.hat[k] - d[k]
}
if (d[k] != 0) last<-k
}
}

#If total no. of events reported, adjust no. censored so that total no. of events agrees.
if (tot.events != "NA"){
if (n.int>1){
sumdL<-sum(d[1:upper[(n.int-1)]])
#If total no. events already too big, then set events and censoring = 0 on all further time intervals
if (sumdL >= tot.events){
d[lower[n.int]:upper[n.int]]<- rep(0,(upper[n.int]-lower[n.int]+1))
cen[lower[n.int]:(upper[n.int]-1)]<- rep(0,(upper[n.int]-lower[n.int]))
n.hat[(lower[n.int]+1):(upper[n.int]+1)]<- rep(n.risk[n.int],(upper[n.int]+1-lower[n.int]))
}
}
#Otherwise adjust no. censored to give correct total no. events
if ((sumdL < tot.events)|| (n.int==1)){
sumd<-sum(d[1:upper[n.int]])
while ((sumd > tot.events)||((sumd< tot.events)&&(n.censor[n.int]>0))){
n.censor[n.int]<- n.censor[n.int] + (sumd - tot.events)
if (n.censor[n.int]<=0){
cen[lower[n.int]:(upper[n.int]-1)]<-0
n.censor[n.int]<-0
}
if (n.censor[n.int]>0){
cen.t<-rep(0,n.censor[n.int])
for (j in 1:n.censor[n.int]){
cen.t[j]<- t.S[lower[n.int]] +
j*(t.S[upper[n.int]]-t.S[lower[n.int]])/(n.censor[n.int]+1)
}
}
}
}
}

```

```

        cen[lower[n.int]:(upper[n.int]-1)]<-hist(cen.t,breaks=t.S[lower[n.int]:upper[n.int]],
        plot=F)$counts
    }
    n.hat[lower[n.int]]<-n.risk[n.int]
    last<-last.i[n.int]
    for (k in lower[n.int]:upper[n.int]){
        d[k]<-round(n.hat[k]*(1-(S[k]/KM.hat[last])))
        KM.hat[k]<-KM.hat[last]*(1-(d[k]/n.hat[k]))
        if (k != upper[n.int]){
            n.hat[k+1]<-n.hat[k]-d[k]-cen[k]
            #No. at risk cannot be negative
            if (n.hat[k+1] < 0) {
                n.hat[k+1]<-0
                cen[k]<-n.hat[k] - d[k]
            }
        }
        if (d[k] != 0) last<-k
    }
    sumd<- sum(d[1:upper[n.int]])
}
}

write.table(matrix(c(t.S,n.hat[1:n.t],d,cen),ncol=4,byrow=F),paste(path,KMdatafile,sep=""),sep="\t")

### Now form IPD ###
#Initialise vectors
t.IPD<-rep(t.S[n.t],n.risk[1])
event.IPD<-rep(0,n.risk[1])
#Write event time and event indicator (=1) for each event, as separate row in t.IPD and event.IPD
k=1
for (j in 1:n.t){
    if(d[j]!=0){
        t.IPD[k:(k+d[j]-1)]<- rep(t.S[j],d[j])
        event.IPD[k:(k+d[j]-1)]<- rep(1,d[j])
        k<-k+d[j]
    }
}
#Write censor time and event indicator (=0) for each censor, as separate row in t.IPD and event.IPD
for (j in 1:(n.t-1)){
    if(cen[j]!=0){
        t.IPD[k:(k+cen[j]-1)]<- rep(((t.S[j]+t.S[j+1])/2),cen[j])
        event.IPD[k:(k+cen[j]-1)]<- rep(0,cen[j])
        k<-k+cen[j]
    }
}

#Output IPD
IPD<-matrix(c(t.IPD,event.IPD,arm),ncol=3,byrow=F)
write.table(IPD,paste(path,KMdataIPDfile,sep=""),sep="\t")

#Find Kaplan-Meier estimates
IPD<-as.data.frame(IPD)
KM.est<-survfit(Surv(IPD[,1],IPD[,2])~1,data=IPD,type="kaplan-meier",)
KM.est
summary(KM.est)

#Find Cox hazard ratio
data0<- read.table("C:\\PHD\\algorithm\\reliability exercise\\KMdataIPD study2 figA arm0 time2 no.txt",header=TRUE)
data1<- read.table("C:\\PHD\\algorithm\\reliability exercise\\KMdataIPD study2 figA arm1 time2 no.txt",header=TRUE)
data<-merge(data0,data1,all=T)
cox<-coxph(formula = Surv(data[,1], data[,2]) ~ data[,3], data=data)
summary(cox)

```

7 Danksagung

An erster Stelle möchte ich mich herzlich bei Herrn Professor Kuß für die Überlassung des Dissertationsthemas und für die freundliche und geduldige Betreuung - auch in schwierigen Phasen - bedanken. Die gemeinsame wissenschaftliche Arbeit und der Austausch über viele andere Themen haben mir immer viel Freude bereitet. Darüber hinaus möchte ich mich für die Unterstützung durch das Team des Instituts für Biometrie und Epidemiologie bedanken.

Mein besonderer Dank gilt meiner Partnerin Alia Firat, die mich bei der Fertigstellung dieser Arbeit emotional unterstützt hat.

Ich danke meinen Großeltern für ihre Mühen und Opfer, die sie für ihre Kinder und Enkelkinder auf sich genommen haben. Ohne sie hätte ich diese Möglichkeiten nicht gehabt. Auch meiner Familie bin ich dankbar für ihre tatkräftige Unterstützung auf meinem Weg. Ebenso bedanke ich mich bei meinen Freunden und Weggefährten, die nicht nur unterstützend zur Seite standen, sondern auch immer ein offenes Ohr für mich hatten. Besonders erwähnenswert ist Özden Altunbas, der mich seit der ersten Klasse als sehr guten Freund begleitet und mir in allen Lebenslagen zur Seite steht.

Abschließend möchte ich mich bei meinen Eltern Gülsen und Yüksel Akbulut sowie meiner Schwester Cigdem bedanken, die mich immer unterstützen und mir Mut und Kraft geben, um schwierige Wege zu meistern. Dank Ihrer Unterstützung konnte ich meinen Traum verwirklichen.