

Aus dem Fachbereich Medizin
der Johann Wolfgang Goethe-Universität
Frankfurt am Main
Institut für Allgemeinmedizin

Der systematische Leitlinienreview

**Eine Methode zur Entwicklung evidenzbasierter Leitlinien,
dargestellt am Beispiel der Entwicklung
der DEGAM-Leitlinie 'Herzinsuffizienz'**

Dissertation

zur Erlangung des Doktorgrades der Medizin des Fachbereichs Medizin
der Johann Wolfgang Goethe-Universität Frankfurt am Main

vorgelegt von
Christiane Muth, MPH
aus Berlin

Frankfurt am Main 2008

Dekan: Prof. Dr. Josef Pfeilschifter

Referent: Prof. Dr. Ferdinand M. Gerlach, MPH

Koreferent: Prof. Dr. Sebastian Harder

Tag der mündlichen Prüfung: 29.10.2008

Meinen Eltern

Inhaltsverzeichnis

A - Abstrakts.....	ix
1 Deutschsprachiger Abstrakt	ix
2 English Abstract	x
B - Ausführliche Zusammenfassungen.....	xi
1 Ausführliche Zusammenfassung (deutsch)	xi
2 Executive Summary	xv
C - Hauptdokument	1
1 Hintergrund	1
1.1 Bezugsrahmen der Arbeit.....	1
1.2 Leitlinien	2
1.2.1 Definition und Ziele von Leitlinien in der gesundheitlichen Versorgung.....	2
1.2.2 Leitlinien und andere Formen evidenzbasierter Entscheidungsunterstützung	3
1.2.3 Anforderungen an Leitlinien	6
1.2.3.1 Methodische Qualität von Leitlinien.....	7
1.2.3.2 Validität von Leitlinien.....	8
1.2.4 Leitlinienentwicklung	10
1.2.4.1 Grundsätze im methodischen Vorgehen zur Entwicklung valider Leitlinien	10
1.2.4.2 Leitlinienentwicklung als mehrstufiger Prozess	10
1.2.4.3 Der DEGAM-Zehnstufenplan	13
1.2.5 Offene Fragen in der Leitlinienentwicklung	14
1.2.5.1 Leitlinienstruktur: Initialalgorithmus und Framework.....	14
1.2.5.2 Gewinnung von Evidenz.....	16
1.2.5.3 Graduierung von Evidenz und Empfehlungen.....	16
1.2.5.4 Hoher Ressourcenaufwand in der de novo-Entwicklung	17
1.2.5.5 Der Prozess der Formulierung von Leitlinienempfehlungen.....	19
1.2.5.5.1 Ableitung normativer aus deskriptiven Aussagen.....	19
1.2.5.5.2 Ebenen der Normativität in Leitlinien.....	22
1.2.5.5.3 Weitere Einflüsse auf die Formulierung von Empfehlungen mit Beispielen aus der Leitlinienentwicklung	26
1.2.5.5.2 Modell von Einflussfaktoren auf die Formulierung von Empfehlungen	28
1.2.5.2.5 Schlussfolgerungen für Leitlinienentwicklung und -adaptation	30
1.2.5.2.6 Implikation für die vorliegende Arbeit	30
2 Forschungsfragen	31

3	Methodisches Vorgehen.....	32
3.1	Prämissen	32
3.1.1	Beschreibung und Eingrenzung der Zielerkrankung.....	33
3.1.1.1	Definition und Beschreibung der Herzinsuffizienz	33
3.1.1.1.1	Definition der Herzinsuffizienz (ICD-10: I50)	33
3.1.1.1.2	Ätiologie	33
3.1.1.1.3	Pathophysiologie	34
3.1.1.1.4	Formen der Herzinsuffizienz	36
3.1.1.1.5	Stadien der chronischen Herzinsuffizienz	37
3.1.1.2	Eingrenzung der Zielerkrankung	38
3.1.2.	Definition des Zielkontextes	38
3.1.3	Veränderungsbedarf.....	39
3.1.3.1	Epidemiologie.....	39
3.1.4	Versorgungssituation.....	41
3.1.4	Festlegung der Ergebnisparameter	41
3.2	Verfahrensschritte im systematischen Leitlinienreview	42
3.2.1	Systematische Leitlinienrecherche	43
3.2.2	Ein- und Ausschlusskriterien	44
3.2.3	Bewertung der formalen methodischen Qualität	45
3.2.4	Framework-Entwicklung	46
3.2.5	Datenextraktion	46
3.2.6	Konsistenzanalyse	47
3.2.7	Informationssynthese	49
3.2.8	Analyse der Belegtheit (Validierung)	49
3.2.8.1	Recherchen	50
3.2.8.2	Bewertung von Primärquellen	50
3.2.8.2.1	Reevaluation der in Leitlinien zitierten Sekundärstudien	50
3.2.8.2.2	Evaluation neu identifizierter HTA-Berichte	50
3.2.8.2.3	Reevaluation der in Leitlinien zitierten Primärstudien.....	51
3.2.9	Überarbeitung des Erstentwurfs und Identifikation weiteren Forschungsbedarfs.....	51
4	Ergebnisse	52
4.1	Ergebnisse der systematischen Literaturrecherche.....	52
4.2	Prüfung der Einschlussfähigkeit	52
4.2.1	Berücksichtigte Leitlinien	52
4.2.2	Ausgeschlossene Leitlinien	56
4.2.3	Untersucherübereinstimmung	57
4.3	Ergebnisse der Bewertung der formalen methodischen Qualität	57
4.4	Ergebnis der Framework-Entwicklung.....	59
4.5	Ergebnisse der Datenextraktion	60
4.5.1	Kurzberichte der berücksichtigten Leitlinien	60
4.5.2	Datenextraktion in Evidenztabelle	62
4.6	Ergebnisse der Konsistenzanalyse	63

4.7	Ergebnisse der Informationssynthese	65
4.8	Ergebnisse der Analyse der Belegtheit	66
4.8.1	Rechercheergebnisse	67
4.8.2	Ergebnisse der Bewertung von Primärquellen und HTA-Berichten ..	69
4.8.2.1	Konsistente Empfehlung, Typ (1), Fallbeispiel	
	ACE-Hemmertherapie	69
4.8.2.2	Konsistente Empfehlungen vom Typ (2), Fallbeispiel	
	Aldosteron-Antagonisten bei Herzinsuffizienz, NYHA III-IV.....	75
4.8.2.3	Konsistente Empfehlung, Typ (4), Fallbeispiel Diuretikatherapie	
	bei symptomatischer Herzinsuffizienz	78
4.8.2.4	Inkonsistenz Typ (B), Fallbeispiel: β -Blockertherapie bei	
	asymptomatischen Patienten ohne Infarktereignis	80
4.9	Weiterer Forschungsbedarf zur Entwicklung der Leitlinie.....	84
4.10	Zusammenfassung der Ergebnisse	86
5	Diskussion	89
5.1	Diskussion des Bezugsrahmens	89
5.2	Diskussion der Methode.....	90
5.2.1	Schrittfolge im systematischen Leitlinienreview im Vergleich	
	zur transkulturellen Adaptation von Leitlinien	91
5.2.1.1	Ähnliche Vorgehensweisen in beiden Verfahren	91
5.2.1.1.1	Definitionsphase.....	91
5.2.1.1.2	Recherche nach Quelleitlinien	94
5.2.1.1.3	Bewertung der formalen Leitlinienqualität	94
5.2.1.2	Grundlegende Unterschiede beider Verfahren	95
5.2.1.2.1	Datenextraktion	95
5.2.1.2.2	Konsistenzanalyse	96
5.2.1.2.3	Analyse der Belegtheit (Validierung)	99
5.2.1.2.4	Informationssynthese	100
5.2.2	Andere systematische Reviews von Leitlinien.....	100
5.3	Diskussion der Ergebnisse der Anwendung am Beispiel der	
	chronischen Herzinsuffizienz.....	101
5.3.1	Voraussetzungen für eine unverzerrte Analyse.....	102
5.3.1.1	Systematische Leitlinienrecherche	102
5.3.1.2	Auswahl der Leitlinien	102
5.3.1.3	Frameworkentwicklung.....	102
5.3.2	Prüfung der formalen methodischen Qualität	103
5.3.3	Prüfung der inhaltlichen Angemessenheit	105
5.3.3.1	Konsistenzanalyse	105
5.3.3.1.1	Typ-(3)-Konsistenzen	105
5.3.3.1.2	Typ A-Inkonsistenzen, Fallbeispiel BNP-/NT-proBNP-Test.....	106
5.3.3.2	Analyse der Belegtheit	109
5.3.3.2.1	Typ-1-Konsistenzen, Fallbeispiel ACE-Hemmer	110
5.3.3.2.2	Typ-2-Konsistenzen, Fallbeispiel Aldosteron-Antagonisten	111
5.3.3.2.3	Typ-4-Konsistenzen, Fallbeispiel Diuretikatherapie.....	112

5.3.3.2.4	Typ-B-Inkonsistenzen, Fallbeispiel 1: β -Rezeptorenblocker bei asymptomatischen Patienten ohne Infarktereignis	114
5.3.3.2.4	Typ-B-Inkonsistenzen, Fallbeispiel 2: AT ₁ -Blocker bei ACE-Hemmer-Intoleranz	114
5.4	Limitationen der vorliegenden Arbeit	115
5.4.1	Limitation der Publikationssprache der Leitlinien	115
5.4.2	Fehlende Zweitbewertung	116
5.4.3	Fehlende Validierung des Klassifikationssystems für die Konsistenzanalyse	117
5.5	Weiterer Forschungsbedarf	117
6	Schlussfolgerungen	118
D - Literatur		119
1	Berücksichtigte Leitlinien	119
2	Sonstige Literatur	122
E - Anhang		I
1	Anhänge des Methodikkapitels des Hauptdokuments	I
1.1	Recherchestrategie	I
1.2	Verzeichnis besuchter Internetadressen	I
1.3	Das AGREE-Instrument	IV
2.1	Ausgeschlossene Leitlinien	V
2.2	Weitere Ergebnisse der Bewertung der formalen methodischen Qualität berücksichtigter Leitlinien	VII
2.3	Datenextraktion aus Leitlinien	XII
2.4	Weitere Ergebnisse der Konsistenzanalyse	XVI
2.5	Ergebnis der Informationssynthese	XXIV
2.2.3.1	Therapie der systolischen Herzinsuffizienz	XXV
2.2.3.1.1	Mortalitätsreduzierende Therapie	XXV
2.6	Weitere Ergebnisse der Analyse der Belegtheit	XXIX
2.6.1	Kurzbeschreibung eines HTA-Berichts [Shekelle et al. 2003]	XXIX
2.6.2	Checkliste eines HTA-Berichts [Shekelle et al. 2003]	XXXII
2.6.3	Beispiel einer Evidenztabelle von Primärstudien	XXXIII
2.6.4	Weitere Fallbeispiele	XXXVII
2.6.4.1	Weitere konsistente Empfehlungen vom Typ (2)	XXXVII
2.6.4.2	Weitere konsistente Empfehlungen vom Typ (4)	XXXIX
2.6.4.3	Weitere Typ-B-Inkonsistenzen	XL
Danksagung und Persönliche Daten		XLIV

Verzeichnis der Abbildungen

Abb. 1.1: Modell evidenzbasierter klinischer Entscheidung	4
Abb. 1.2: Validität von Leitlinien.....	9
Abb. 1.3: Leitlinien-„Bi-cycle“ zu Entwicklung und Implementierung.....	12
Abb. 1.4: Einflüsse auf die Formulierung von Leitlinienempfehlungen	29
Abb. 3.1: Pathophysiologische Modelle der chronischen Herzinsuffizienz	34
Abb. 3.2: Standardisierter Domänen-Summenscore, Berechnungsformel	45
Abb. 4.1: Ergebnisse von systematischer Leitlinienrecherche und Ein..... schlussprüfung	53
Abb. 4.2: Frameworkentwicklung: Beispiel BNP-Test in der Erstdiagnostik	60
Abb. 5.1: Gegenüberstellung von transkontextueller Adaptation und syste..... matischem Leitlinienreview	92

Verzeichnis der Tabellen

Tab. 1.1: Gemeinsamkeiten und Unterschiede von EbM, Leitlinien und	5
HTA	5
Tab. 1.2: Anforderungen an Leitlinien für die medizinische Versorgung.....	6
Tab. 1.3: Der DEGAM-Zehnstufenplan.....	13
Tab. 1.4: Transkontextuelle Leitlinienadaptation	18
Tab. 1.5: Inadäquate methodische Vorgehensweise bei der Entwicklung..... von Leitlinien	26
Tab. 3.1: Funktionelle Stadieneinteilung der Herzinsuffizienz in der revi..... dierten NYHA-Klassifikation	37
Tab. 3.2: Stadieneinteilung der Herzinsuffizienz nach AHA und ACC	38
Tab. 3.3: Klassifikation von Konsistenzen und Inkonsistenzen	47
Tab. 3.4: Vorläufige Evidenzlevel und Empfehlungsstärken im Erstent..... wurf	49
Tab. 3.5: Strategie zur Gewinnung weiterer Evidenz.....	51
Tab. 4.1: Berücksichtigte Leitlinien: Empfehlungsbereich, Versorgungs	53
kontext.....	53
Tab. 4.2: Berücksichtigte Leitlinien: weitere Merkmale.....	55
Tab. 4.3: Beobachtete Übereinstimmung ein- und ausgeschlossener	57
Dokumente	57
Tab. 4.4: Bewertung der Qualität: absolute u. standardisierte Domänen	58
scores.....	58
Tab. 4.5: Bewertung der Qualität: Median und Streuung.....	59
Tab. 4.6: Erstdiagnostik ohne BNP-Tests.....	64
Tab. 4.7: BNP-Test.....	65
Tab. 4.8: Themenzusammenstellung für die Analyse der Belegtheit.....	66

Tab. 4.9:	Themenliste für systematische/umfassende Evidenzrecherchen ..	67
Tab. 4.10:	Anzahl und Art der in den Leitlinien zitierten Evidenzquellen	68
Tab. 4.11:	Konsistente Leitlinienempfehlungen zur Therapie mit ACE- Hemmern.....	69
Tab. 4.12:	Konsistenzanalyse der ACE-Hemmertherapie	70
Tab. 4.13:	Schlüsselempfehlungen zur Therapie mit Aldosteron- Antagonisten	76
Tab. 4.14:	Konsistenzanalyse zur Therapie mit Aldosteron-Antagonisten.....	76
Tab. 4.15:	Schlüsselempfehlungen zur Therapie mit Diuretika	78
Tab. 4.16:	Konsistenzanalyse zur Diuretikatherapie	78
Tab. 4.17:	Schlüsselempfehlungen zur Therapie mit β -Blockern	80
Tab. 4.18:	Konsistenzanalyse zur β -Rezeptorenblockertherapie	81
Tab. 4.19:	Fragestellungen mit Bedarf zur weiteren Evidenzgewinnung	84

Tab. 5.1:	Vergleich verschiedener Bewertungen der Leitlinienqualität	104
-----------	--	-----

(Tabellen im Anhang)

Tab. A.1:	Institutionen, deren Homepages bei der systematischen Leitlinienrecherche besucht wurden - Nationale Organisationen	I
Tab. A.2:	Institutionen, deren Homepages bei der systematischen Leitlinienrecherche besucht wurden - Internationale Organisationen....	II
Tab. A.3:	Das AGREE-Instrument – Domänen und Items	IV
Tab. A.4:	Ausgeschlossene Leitlinien unter Nennung des Ausschlussgrundes	V
Tab. A.5:	Ausschnitt aus einer Evidenztafel zu Leitlinienempfehlungen über ACE-Hemmer	XIII
Tab. A.6:	Nichtpharmakologische Therapie	XVI
Tab. A.7:	Digitalistherapie	XVII
Tab. A.8:	Pharmakotherapie mit Angiotensin-II-Antagonisten (AT ₁ -Blocker)	XVII
Tab. A.9:	Pharmakotherapie mit Hydralazin / ISDN	XVIII
Tab. A.10:	Therapie mit Kalziumantagonisten	XIX
Tab. A.11:	Orale Antikoagulation und Thrombozytenaggregationshemmung	XIX
Tab. A.12:	Pharmakotherapieempfehlungen zur diastolischen Herzinsuffizienz	XX
Tab. A.13:	Konsistente und inkonsistente Empfehlungen der berücksichtigten Leitlinien, geordnet nach Themen	XX
Tab. A.14:	Evidenztafel zu Primärstudien über Aldosteron-Antagonisten	XXXIV

Verzeichnis der Abkürzungen

ACC	American College of Cardiology
ACE-Hemmer	Angiotensin Converting Enzyme-Hemmer
AGREE	Apraisal of Guidelines for Research & Evaluation
AHA	American Heart Association
AHCPR	Agency for Health Care Policy and Research
AHRQ	Agency for Healthcare Research and Quality
AKDAE	Arzneimittelkommission der deutschen Ärzteschaft
ANP	Atrial Natriuretic Peptide
AT ₁ -Blocker	Synonym für Angiotensin II-Rezeptorantagonist
BNP	Brain Natriuretic Peptides
CADTH	Canadian Agency for Drugs and Technologies in Health
CCOHTA	Canadian Coordinating Office for Health Technology Assessment
CCS	Canadian Cardiovascular Society
CM	Kardiomyopathie
CRD	Centre for Reviews and Dissemination
CSANZ	Cardiac Society of Australia & New Zealand
DEGAM	Deutsche Gesellschaft für Allgemeinmedizin und Familienmedizin
DGK	Deutsche Gesellschaft für Kardiologie
DiEM	Institut für evidenzbasierte Medizin
DVA & VHA	Department of Veterans Affairs & Veterans Health Administration
ECHTA	European Collaboration for Health Technology Assessment
EF	Ejektionsfraktion
ESC	European Society of Cardiology
HI	Herzinsuffizienz
ICSI	Institute for Clinical System Improvement
INR	International Normalised Ratio
LL	Leitlinie(n)
LLGH	Leitlinien-Gruppe Hessen
LVEF	Linksventrikuläre Ejektionsfraktion
NHF/Austr.	National Heart Foundation of Australia
NHF/NZ	National Heart Foundation of New Zealand
NICE	National Institute for Clinical Excellence
NNH	Numbers Needed to Harm
NNT	Numbers Needed to Treat
NTpro-BNP	N-terminal pro-B-type natriuretic peptide

(Fortsetzung Abkürzungsverzeichnis)

NYHA	New York Heart Association
OPOT	Ontario Program for Optimal Therapeutics
RAAS	Renin-Angiotensin-Aldosteron-Systems
RCT	Randomized Controlled Trial
SLR	Systematischer Leitlinienreview
SR	Systematischer Review
UM	University of Michigan
UWH	Universität Witten / Herdecke
VES	Ventrikuläre Extrasystolie
VHF	Vorhofflimmern
VT	Ventrikuläre Tachykardie

A - Abstrakts

1 Deutschsprachiger Abstrakt

Hintergrund: Leitlinien können potentiell die gesundheitliche Versorgung verbessern. Ihre evidenzbasierte Entwicklung ist aufwendig, insbesondere bei komplexen Krankheitsbildern, wie chronischer Herzinsuffizienz. Transkontextuelle Leitlinienadaptationen sind problematisch, da kontextuell beeinflusste Empfehlungen der Quelleitlinie die Zielleitlinie verzerren können.

Forschungsgegenstand ist eine Verfahrensentwicklung zur systematischen Evaluation von Leitlinien (systematischer Leitlinienreview, SLR) und deren Erprobung am Gegenstand, der Entwicklung einer evidenzbasierten Leitlinie zur hausärztlichen Versorgung der chronischen Herzinsuffizienz.

Methode: Der SLR umfasst folgende Verfahrensschritte: 1) Systematische Recherche nach evidenzbasierten Leitlinien, 2) Auswahl von Quelleitlinien nach prädefinierten Kriterien, 3) Bewertung der methodischen Qualität (AGREE-Instrument), 4) Systematische Framework-Entwicklung (klinische Fragen als strukturbildendes Raster der Zielleitlinie), 5) Datenextraktion aus den Quelleitlinien, 6) Systematisch-vergleichende Konsistenzanalyse der Empfehlungen, 7) Informationssynthese im Leitlinien-Erstentwurf, 8) Validierung der Evidenzbezüge und Prüfung der Konkordanz zu Empfehlungen, 9) Überarbeitung des Leitlinien-Erstentwurfs und Formulierung weiteren Forschungsbedarfs.

Ergebnisse: N=16 Quelleitlinien erfüllten die Einschlusskriterien (zwei Reviewer, Übereinstimmung $\kappa=0,95$), davon zeigten fünf eine hohe methodische Qualität. Zum Framework (27 Fragen) wurden N=35 komplexe Empfehlungen identifiziert, darunter waren n=25 konsistent, n=9 inkonsistent, n=1 nicht vergleichbar (einzelne Leitlinienempfehlung). 14/25 konsistenten Empfehlungen waren konsensbasiert (Diagnostik, nicht-pharmakologische Therapie, diastolische Herzinsuffizienz), 7/25 basierten auf belastbarer Evidenz, 4/25 hatten differierende Evidenzlevel. 3/9 Inkonsistenzen waren bedeutsam (BNP-Tests, AT₁-Blocker). Für n=16 Empfehlungen wurden Evidenzbelege reevaluiert: die Mehrzahl war kongruent. Einzelne Inkongruenzen (fehlende Evidenzbelege, überoptimistische Trends, inadäquate Risikobewertung), Inkonsistenzen und konsensbasierte Empfehlungen indizierten Forschungsbedarf. Der Erstentwurf wurde nach sechs Monaten fertiggestellt.

Diskussion: Die Hauptlimitation der vorgelegten Studie ist die fehlende Zweitbegutachtung, die nur in der Auswahl der Quelleitlinien erfolgte.

Schlussfolgerung: Der systematische Leitlinienreview ist ein systematischer, valider, transparenter und effektiver Ansatz zur Entwicklung evidenzbasierter Leitlinien.

2 English Abstract

Background: Clinical guidelines potentially improve healthcare. However, the de-novo development of evidence-based guidelines requires remarkable resources – especially for complex conditions like chronic heart failure (CHF), and the adaptation of existing guidelines may be biased as a result of contextually influenced recommendations in source-guidelines.

Objective: To design a systematic guideline review (SGR) method, and to apply it to the development of an evidence-based guideline on CHF.

Methods: The SGR involves the following steps: 1) Systematic search for, and 2) Selection of eligible source guidelines according to predefined criteria, 3) Assessment of methodological quality (AGREE-Instrument), 4) Systematic definition of the clinical questions, 5) Data extraction from the source guidelines, 6) Analysis for consistency in the recommendations, 7) Synthesis of the data, 8) Validation of the evidence and assessment of recommendations' congruency, 9) Reformulation of the draft and identification of evidence gaps.

Results: A total of 16 guidelines was included (two reviewers, $\kappa=0.95$), 5/16 of high quality. Within the framework (27 questions) we identified 35 complex recommendations: 25/35 were consistent, 9/35 inconsistent, and 1/35 unratable (derived from a single guideline). Out of the 25 consistencies, $n=14$ were based on consensus (diagnostics, non-pharmacologic therapy, diastolic CHF), $n=7$ on evidence, and $n=4$ differed in grading. Major inconsistencies were found in 3/9 recommendations (Brain Natriuretic Peptides, AT_1 -Blockers). For 16 recommendations, the evidence was re-evaluated whereby the majority were congruent. A few incongruencies (lacking evidence, over-optimistic trends, inadequate risk assessment), inconsistencies and consensus based statements necessitated further research. After six months the first draft was completed.

Discussion: Main limitation to the study is the lack of a second review, which was only done in the form of the guideline selection.

Conclusion: The SGR is a completely systematic, valid, transparent and efficient approach to the development of evidence based guidelines.

B - Ausführliche Zusammenfassungen

1 Ausführliche Zusammenfassung (deutsch)

Hintergrund: Klinische Leitlinien haben das Potential, die gesundheitliche Versorgung von Patienten zu verbessern, sofern bei ihrer evidenzbasierten Entwicklung wichtige methodische Grundsätze berücksichtigt werden, die ihre Validität gewährleisten. Die Leitlinienentwicklung erfordert erhebliche Ressourcen, insbesondere wenn die Zielerkrankung durch vielfältige und komplexe Entscheidungssituationen gekennzeichnet ist, die in der Leitlinie behandelt werden sollen. Bisher wurde darum eine ‚Adaptation‘ bereits existierender Leitlinien vorgeschlagen und verschiedentlich durchgeführt. Dieses Vorgehen ist nicht unproblematisch, da kontextspezifische Empfehlungen in der Quelleitlinie zu Bias in der Zielleitlinie führen können.

Fragestellung: Ist ein „systematischer Leitlinienreview“ eine effektive Vorgehensweise in der Entwicklung einer qualitativ hochwertigen evidenzbasierten Leitlinie zur hausärztlichen Versorgung der chronischen Herzinsuffizienz?

Methoden: Für die Zielleitlinie wurden ex ante folgende *Prämissen* definiert:

- Zielerkrankung ist eine chronische Herzinsuffizienz beim Erwachsenen inkl. akuter Dekompensationen. Ausgeschlossen wurden akute Herzinsuffizienz und isolierte Rechtsherzinsuffizienz auf dem Boden eines Cor pulmonale.
- Primäre Nutzer sind *hausärztlich* tätige ÄrztInnen und deren Patienten.
- Die Leitlinie muss die Spezifika des allgemeinmedizinischen Versorgungs-Settings berücksichtigen (hohe Kontinuität und Longitudinalität, Niedrig-Prävalenz-Bereich).
- Ergebnisparameter sind v.a. Mortalität, Morbidität, Lebensqualität, Symptomverbesserung und Belastungstoleranz, weitere werden spezifisch ergänzt.

Der *systematische Leitlinienreview* umfasste folgende *Verfahrensschritte*:

1. *Systematische Leitlinienrecherche:* im März 2004 wurde eine systematische Recherche in elektronischen Datenbanken durchgeführt und durch ausführliche Handsuchen ergänzt.

2. *Auswahl von Leitlinien:* zwei voneinander unabhängige Reviewer wählten anhand prädefinierter Ein- und Ausschlusskriterien Quelleitlinien aus, die Untersucherübereinstimmung wurde bestimmt.

3. *Bewertung der methodischen Qualität:* diese wurde mit dem international validierten AGREE-Instrument bewertet, um sicherzustellen, dass sich Leitlinien von hoher methodischer Qualität unter den berücksichtigten befanden.

4. *Systematische Framework-Entwicklung*: aus den Quelleitlinien wurden klinisch relevante Fragestellungen extrahiert, auf den Versorgungskontext abgestimmt und als strukturbildendes Raster für die Zielleitlinie entwickelt.

5. *Datenextraktion*: Daten der Leitlinien wurden in Evidenztabelle extrahiert.

6. *Konsistenzanalyse*: unter Berücksichtigung der berichteten Evidenz wurden die Empfehlungen berücksichtigter Leitlinien vergleichend analysiert und in einem dafür entwickelten sechsstufigen Klassifikationsschema eingeordnet: Konsistenzen in vier Typen (je nach berichteter Evidenzgrundlage) und Inkonsistenzen in bedeutsame und weniger bedeutsame.

7. *Informationssynthese*: die Ergebnisse der Konsistenzanalyse wurden in einem vorläufigen Erstentwurf zusammengefasst.

8. *Analyse der Belegtheit der Empfehlungen („Validierung“)*: bei konsistenten Empfehlungen mit aufgeführten Evidenzbelegen und bei weniger bedeutsamen Inkonsistenzen mit Evidenzbezug wurden Reevaluationen der relevanten Primärquellen vorgenommen und auf Kongruenz zu den Empfehlungen geprüft.

9. *Überarbeitung des Erstentwurfs und Identifikation weiteren Forschungsbedarfs*: der vorläufige Erstentwurf wurde unter Berücksichtigung der Validierungsergebnisse überarbeitet. Für Bereiche mit hoher Entscheidungsunsicherheit (bedeutsame Inkonsistenzen und Konsistenzen ohne belastbare Evidenz) sowie Fragestellungen, die sich aus Inkongruenzen ergaben, wurde weiterer Forschungsbedarf formuliert.

Ergebnisse: Von den 52 in der *systematischen Leitlinienrecherche* identifizierten entsprachen 16 umfassende Leitlinien den *Ein- und Ausschlusskriterien* (Untersucherübereinstimmung $\kappa=0,95$). In der *Bewertung ihrer formalen methodischen Qualität* mit dem AGREE-Instrument zeigten fünf eine hohe Qualität. Aus den Quelleitlinien wurde ein strukturbildendes Raster (*Framework*: 27 klinisch relevante Fragen) entwickelt. Gemäß diesem Framework wurden die Daten aus den Leitlinien in 20 Evidenztabelle extrahiert.

Insgesamt wurden $N=35$ spezifische, z.T. mehrgliedrige und komplexe Empfehlungen identifiziert ($n=6$ zur Erstdiagnostik der Herzinsuffizienz, $n=4$ zur nicht-pharmakologischen Therapie, $n=19$ zur Pharmakotherapie der systolischen Herzinsuffizienz, $n=6$ zur diastolischen Herzinsuffizienz) und auf *Konsistenz* geprüft. Von diesen waren $n=25$ konsistent, $n=9$ inkonsistent, $n=1$ nicht vergleichbar (sie entstammte einer einzelnen Leitlinie). 14/25 konsistenten Empfehlungen waren konsensbasiert und betrafen insbesondere die Bereiche Diagnostik, nicht-pharmakologische Therapie und diastolische Herzinsuffizienz. 7/25 Konsistenzen basierten auf belastbarer Evidenz (Pharmakotherapie der systolischen Herzinsuffizienz), 4/25 waren inhaltlich konsistent mit unterschiedlich graduierter Evidenz (z.B. körperliches Training und Diuretikatherapie). Von neun Inkonsistenzen waren drei bedeutsam (zur diagnostischen Wertigkeit des BNP-/NT-proBNP-Tests und zum Wert der AT_1 -Blocker). Die sechs weniger bedeutsamen Inkonsistenzen betrafen insbesondere spezifische Empfehlungen

zur Pharmakotherapie (z.B. zu β -Blockern bei asymptomatischer Herzinsuffizienz nicht-ischämischer Genese) und beruhten auf einer kontroversen Bewertung vorliegender Studienevidenz durch einzelne Leitlinien. Die bis dahin gewonnenen Ergebnisse wurden im *vorläufigen Erstentwurf* zusammengefasst und diskutiert.

In der *Analyse der Belegtheit* wurden für n=16 Empfehlungen (n=11 konsistente, n=5 inkonsistente) die entscheidungsrelevanten Evidenzbelege reevaluiert, wobei die jeweils höchstwertigen 21 Arbeiten aus insgesamt mehr als 300 zitierten Evidenzbezügen ausgewählt werden konnten und um zwei neu identifizierte HTA-Berichte und eine Metaanalyse ergänzt wurden. Ausreichende Evidenzbelege fanden sich bei 5/16 Empfehlungen (z.B. zu β -Blockern bei chronischer Herzinsuffizienz, NYHA II-III), bei 8/16 Empfehlungen fehlten Belege für Teilaussagen, u.a. der Beleg für die Wirksamkeit von ACE-Hemmern bei asymptomatischen Patienten, der erst in einem neu identifizierten HTA-Bericht erbracht wurde und der fehlende Beleg der Wirksamkeit von β -Rezeptorenblockern bei asymptomatischen Patienten ohne Infarktereignis, der auch in weiteren Evidenzrecherchen nicht ermittelt werden konnte. In 3/16 Fällen wurde auf eine Reevaluation von Primärquellen verzichtet (z.B. zu Hydralazin/ISDN wegen fehlender Marktverfügbarkeit in Deutschland).

Im Ergebnis wurde der vorläufige *Erstentwurf* im Wortlaut der Empfehlungen, in Evidenzleveln und im Stärkegrad der Empfehlungen korrigiert bzw. ergänzt. Außerdem wurden auf der Basis von identifizierten Inkonsistenzen, wenig belastbaren Evidenzbelegen und Inkongruenzen spezifische Forschungsfragen für eine weitere Evidenzgewinnung formuliert, die einer formalen Priorisierung zugänglich waren (*strukturierter Forschungsbedarf*).

Diskussion: Mit der Zielsetzung, für die Neuentwicklung einer deutschsprachigen hausärztlichen Leitlinie zur Herzinsuffizienz einen Erstentwurf zu erarbeiten, wurde das Verfahren des systematischen Leitlinienreviews (SLR) neuentwickelt und erstmalig durchgeführt. Um einen der arbeits- und ressourcenaufwendigsten Schritte einer de novo-Leitlinienentwicklung (die Recherche und Bewertung der Primär-Evidenz) zu reduzieren, sollte dieses Verfahren von bereits existierenden Leitlinien ausgehen. Zudem sollte der SLR auf den Prinzipien und Arbeitsmethoden der evidenzbasierten Medizin basieren und den Anforderungen an die Validität in der Leitlinienentwicklung genügen.

In einem Vergleich mit anderen Verfahren, die auf der Auswertung existierender Leitlinien beruhen, wie z.B. der transkulturellen Adaptation nach Fervers et al. [2006], konnte gezeigt werden, dass zwar Ähnlichkeiten im Vorgehen bestehen – insbesondere in der Suche nach Leitlinien und der Bewertung von deren formaler Qualität –, dass jene Verfahren jedoch keine Prüfung der *inhaltlichen Angemessenheit* von Leitlinienempfehlungen vornehmen können und daher in der Gefahr stehen, kurzschlüssige Evidenzbewertungen oder kontextfremde Werturteile aus den ausgewerteten Leitlinien ungeprüft in den eigenen Leitlinienent-

wurf 'einzuschleppen'. Im SLR hingegen werden systematisch die Konsistenz von Empfehlungen, die Validität der Evidenzbezüge und die Kongruenz zwischen beiden geprüft, um die *Validität der Zielleitlinie* zu sichern, was am Beispiel der chronischen Herzinsuffizienz belegt wurde:

Zwar waren die in den Leitlinien identifizierten Empfehlungen mehrheitlich konsistent (etwa drei Viertel der Empfehlungen), jedoch wurde nur zu einem geringen Teil der konsistenten Empfehlungen (weniger als ein Drittel) eine belastbare Evidenzgrundlage berichtet. In der Validierung der angegebenen Evidenzbezüge wurden wiederum Inkongruenzen in Form fehlender Evidenzbelege, überoptimistischer Ergebnisse und inadäquater Risikoeinschätzungen identifiziert. Außerdem wurden bedeutsame Inkonsistenzen in den Leitlinienempfehlungen gefunden, die – gemeinsam mit Empfehlungen ohne belastbare Evidenz – für eine Neuerschließung von Evidenz vorgemerkt wurden.

Die *zentrale Limitation* der vorgelegten Arbeit besteht in einer fehlenden Zweitbegutachtung in Datenextraktion, Konsistenzanalyse und Validierung, so dass ein gewisser Untersucherbias nicht ausgeschlossen werden kann. Weitere Limitationen bestehen in möglichen kontextbezogenen Verzerrungen, die sich ergeben können, wenn Leitlinien einer bestimmten Publikationssprache den SLR dominieren – wie z.B. englischsprachige Leitlinien mit den dazugehörigen angloamerikanischen Kontexten. Auch bestehen im Rahmen der Neuentwicklung und erstmaligen Anwendung des SLR (wie auch des Klassifikationssystems für Konsistenzen / Inkonsistenzen) bislang nur begrenzte Erfahrungen. Künftige SLR-Anwendungen werden u. U. weitere Limitationen aufzeigen.

Schlussfolgerungen: Der systematische Leitlinienreview ist ein

- *zusammenhängender systematischer und transparenter* Ansatz zur Entwicklung einer evidenzbasierten Leitlinie von der Entwicklung des Frameworks bis zur Fertigstellung des Erstentwurfs, da jeder einzelne Verfahrensschritt festen Ordnungsprinzipien folgt und nachvollziehbar dokumentiert werden kann;
- *valider* Ansatz, da die in den Quelleitlinien berichteten Evidenzgrundlagen anhand einer Neubewertung der entscheidungsrelevanten Primärquellen verifiziert werden und Empfehlungen in Bezug auf ihre Kongruenz zur herrschenden Evidenzlage und zum Zielkontext kritisch überprüft werden;
- *effektiver* und *ressourcensparender* Ansatz zur de novo-Leitlinienentwicklung, insbesondere da unnötige Redundanzen in Literaturreviews vermieden werden können.

Aus diesen Gründen empfiehlt es sich dort, wo bereits qualitativ hochwertige evidenzbasierte Leitlinien zu einem Thema existieren, den systematischen Leitlinienreview als Verfahrensschritt in der de novo-Entwicklung einer Leitlinie einzusetzen.

2 Executive Summary

Background: Clinical practice guidelines have the potential to bring about a substantial improvement in healthcare, if basic methodological principles of evidence-based guideline development, which ensure their validity, are maintained. However, the de-novo development of an evidence-based guideline requires time, and considerable resources – in particular for guidelines on conditions of high complexity. Basing guideline development on the adaptation of existing guidelines was therefore suggested. This procedure, however, may lead to problems, as the adoption of context-dependent recommendations from source guidelines may cause a bias in the newly developed guideline (target guideline), if the target guideline is developed in another context.

Research question: Is the proposed 'Systematic Guideline Review' (SGR) an effective means of developing a high-quality, evidence-based guideline on chronic heart failure (CHF) for general practice?

Methods: For the guideline under development the following premises were defined a priori:

- Target condition is chronic heart failure in adults incl. exacerbations of the syndrome i.e. acute decompensations. The guideline does not address 'acute heart failure' or isolated right heart failure as a consequence of respiratory disease.
- Target users of the guideline are primarily general practitioners / family physicians and their patients.
- The guideline has to recognise the specificity of the primary care setting (continuity and longitudinality of care, low disease prevalence and the inclusion of earlier disease).
- The main outcome parameters for the assessment of effectiveness are mortality, morbidity, health-related quality of life, symptom improvement, and exercise tolerance. Further parameters are included for specific conditions.

The systematic guideline review consists of the following methodological steps:

1. *Systematic retrieval of existing guidelines:* In March 2004 a systematic search in electronic databases was conducted and complemented using extensive hand-searches.
2. *The selection of eligible guidelines* following predefined, explicit criteria: Two researchers assessed independently, whether the available guidelines fulfilled explicit criteria for inclusion/exclusion. The degree of agreement was computed.
3. *Appraisal of methodological quality:* The internationally validated AGREE-instrument was applied to ensure that high-quality guidelines were among the sample of selected source guidelines.

4. *Systematic framework development*: Clinical questions, relevant for the care of CHF-patients in the primary care setting, were defined on the basis of the source guidelines. This framework became the grid structure in the development of the target guideline.

5. *Data extraction*: Data from the source guidelines were entered (recommendations and their link to the evidence base / studies) into evidence tables.

6. *Consistency analysis*: The recommendations of the source guidelines were compared and analysed for each clinical question with respect to the underlying (reported) evidence level. The observed consistencies / inconsistencies among the recommendations were classified according to four types of consistency (depending on the quality of the evidence reported) and two types of inconsistency (major vs. minor).

7. *Information synthesis*: Based on the results of the preceding analyses we composed a first (preliminary) draft of the target guideline.

8. *Analysis of the evidence basis of recommendations ('Validation')*: In consistent or minor inconsistent recommendations, backed up by evidence, the most relevant evidence sources cited in the guidelines were re-assessed. The congruence was proved when the results of the cited studies justified the recommendations.

9. *Review of the preliminary guideline draft and identification of research gaps*: The preliminary draft of the guideline was revised with respect to the validation results. In questions of significant uncertainty (consistencies without satisfying evidence, relevant inconsistencies) the need for further research, and the clinical questions begged by the inconsistencies, were formulated.

Results: Out of 52, a total of 16 comprehensive guidelines were included as source guidelines which fulfilled the predefined criteria (interobserver agreement $\kappa=0.95$). In accordance with the AGREE-instrument 5/16 of the source guidelines were of high quality. From the source guidelines a framework of 27 clinically relevant questions was derived and the data (guideline recommendations and their evidence basis) were entered into 20 evidence tables covering these questions.

Consistency analysis: To answer these questions a total of N=35 partly complex recommendations was identified (primary diagnostics of CHF: n=6, non-pharmacologic therapy: n=4, pharmacologic therapy in systolic CHF: n=19, diastolic CHF: n=6). Among these 25/35 recommendations were consistent, 9/35 were inconsistent, one was not classifiable (it was derived from a single guideline, and no comparison was possible). Fourteen of the 25 consistent recommendations were merely based on consensus (particularly in the fields of diagnostics, non-pharmacological therapy, and diastolic CHF). Based on qualified evidence, were 7/25 recommendations (mainly in pharmacotherapy of systolic

CHF), 4/25 recommendations were consistent in content but differed in level of evidence (e.g. exercise training, benefit of diuretics). Major inconsistencies were found in 3/9 cases (diagnostic relevance of testing for brain natriuretic peptides, benefit of combination therapies with angiotensin II receptor blockers). The minor inconsistencies (6/9) concerned recommendations on specific pharmacologic questions (e.g. beta-blockers in the therapy of asymptomatic CHF with non-ischemic aetiology) and were mostly the result of a controversial appraisal of studies in single guidelines. The results were summarised and discussed in a first, preliminary draft of the guideline.

Validation: In n=16 recommendations (11 consistent, 5 inconsistent) the evidence basis which led to the evidence grading in the source guidelines, was re-analysed. By means of restriction to the most relevant sources of evidence according to the reported evidence levels in the source guidelines, 21 of more than 300 cited references were selected for re-evaluation, supplemented by two identified health technology assessments and one meta-analysis (results from the initial literature search, not mentioned in the source guidelines). A sufficient evidence basis was found for 5/16 recommendations (e.g. beta-blocker in CHF, NYHA II-III). For 8/16 recommendations, the cited evidence could not (at least in parts) support the recommendation: e.g. evidence was lacking in the cited references for the benefit of ACE-inhibitors in asymptomatic CHF patients – it was only proven in a newly identified HTA report. Another example is the use of beta-blockers for asymptomatic patients without previous myocardial infarction, for which even our additional research did not reveal any evidence. In 3/16 recommendations we refrained from a re-evaluation (e.g. due to a lack of market availability as for the fixed drug combination hydralazine plus ISDN).

In the final step, the preliminary draft was amended in 'wording', evidence levels and the grading of recommendations. Furthermore, the identified inconsistencies and the recommendations based on weak evidence led to specific questions for further research (structured research needs).

Discussion: The SGR method evolved from the need to develop a German guideline on CHF for general practice in a transparent and valid, but lean process. To save resources during the most expensive phase of de-novo guideline development (search for, retrieval and appraisal of primary evidence) the process should consider the relevant source guidelines. Moreover, the SGR method should comprise the basic concepts and means of evidence-based medicine, as well as the fundamental requirements of validity in guideline development.

Compared to other methods of guideline development that make use of existing guidelines, e.g. trans-contextual guideline adaptation [Fervers et al. 2006], there are similarities such as in the retrieval and appraisal of appropriate source guidelines. In contrast to the SGR, the other methods are unable to assess the *appropriateness* of adapted guideline recommendations. Therefore, these methods of adaptation potentially risk transferring recommendations based on

weak evidence, or containing context-dependent value judgements to the target guideline. In the SGR the consistency of recommendations, the validity of references to evidence, as well as the congruency between both are checked systematically in order to ensure the validity of the target guideline, as demonstrated in examples on chronic heart failure.

Certainly most of the recommendations in the source guidelines (about three quarters) were consistent, but less than one third of these consistencies was based on proper and sufficient evidence. Furthermore, the validation has shown considerable flaws: a lack of evidence, over-optimistic interpretation of the evidence, and inadequate estimation of risks. Major inconsistencies as well as evidence gaps call for further research resp. retrieval of primary evidence.

The main drawback to this study is that due to limited resources, data extraction and the evaluative steps were performed by only one reviewer. Therefore, a certain observer bias cannot be ruled out. Another potential limitation of this study is the limitation to the German and English languages. This may have led to a contextual bias due to the fact that within the selected guidelines, as well as in the primary studies, there is a predominance of contributions from within the Anglo-Saxon cultural context (and its values and beliefs). Last but not least, the present study was the first application of the SGR method, as well as corresponding concepts like the classification of consistency. Hence, practical experience with the method is limited, and further applications may indicate additional limitations.

Conclusions: The methodical approach of the systematic guideline review is

- a *consistent, systematic and transparent approach* to all developmental stages of an evidence-based clinical guideline, from the development of a framework of clinical questions to the preparation of a first consistent draft of the target guideline. In each of the developmental steps clear criteria and requirements are applied and the results can be reproducibly documented;
- a *valid approach*, since the reported evidence for recommendations in the source guidelines are re-assessed and the empirical justification for the recommendations examined. Implicit context-specific normative values can be detected and taken into account in the subsequent guideline development process;
- an *effective and resource saving approach* to de-novo guideline development, since redundant work (particularly in reviewing literature) can be avoided.

The SGR method is generally applicable in situations where (more than one) high-quality guideline is already available and despite limited resources de-novo development (and not a mere adaptation) is desired.

C - Hauptdokument

1 Hintergrund

1.1 Bezugsrahmen der Arbeit

In Kooperation mit der Deutschen Gesellschaft für Allgemeinmedizin und Familienmedizin (DEGAM) wurde im Institut für Allgemeinmedizin der Johann Wolfgang Goethe-Universität Frankfurt eine *evidenzbasierte Leitlinie zur hausärztlichen Versorgung der chronischen Herzinsuffizienz* entwickelt. Dieses Vorhaben wurde als Teilprojekt im Kompetenznetz ‚Herzinsuffizienz‘ durch das Bundesministerium für Bildung und Forschung (BMBF) industrieunabhängig gefördert (FKZ: 01GI0205).

Innerhalb dieses Projekts wurde von der Autorin der vorliegenden Arbeit eine in dieser Form neue Methode entwickelt – der so genannte systematische Leitlinienreview. Mit Hilfe dieser Methode wurde ein ‚systematisches Lernen‘¹ aus bereits existierenden Leitlinien und damit eine ressourcensparende Entwicklung der späteren DEGAM-Leitlinie Nr. 9 „Herzinsuffizienz“ ermöglicht, was in der vorliegenden Arbeit gezeigt wird.

Im folgenden Abschnitt wird zunächst der Leitlinienbegriff definiert, und es wird ein Bezug zwischen Leitlinien und dem Konzept der evidenzbasierten Medizin hergestellt, in dem sie verortet sind. Es werden die Ziele umrissen, die mit der Entwicklung und Implementierung von Leitlinien verfolgt werden, und die Anforderungen, denen Leitlinien gerecht werden sollen. Danach werden Grundsätze und Methoden in der Leitlinienentwicklung dargestellt, die für die vorliegende Arbeit bedeutsam sind. Im Anschluss an diese einführenden Betrachtungen werden offene Fragen, die sich gegenwärtig aus der Umsetzung von etablierten Entwicklungskonzepten ergeben, diskutiert. Dazu werden Probleme der Strukturbildung in Leitlinien, der Gewinnung und Bewertung von Evidenz sowie bei deren Graduierung aufgezeigt.

Der hohe Aufwand, der sich aus Literaturrecherche und -review ergibt, wird auch in anderen westlichen Industriestaaten als gravierendes Ressourcenproblem aufgefasst. Es wurden daher Adaptationsverfahren angeregt, bei denen in der Entwicklung auf bereits existierende Leitlinien zurückgegriffen wird. Diese Verfahren werden vorgestellt, und den spezifischen Problemen, die dabei zu beachten sind, wird in einem separaten Abschnitt zu ‚Einflüssen auf die Formulierung von Empfehlungen‘ nachgegangen – u.a. der (kontextspezifischen)

¹ Systematisches Lernen wird hier nicht als terminologischer Begriff verwendet, es soll vielmehr verdeutlicht werden, dass dieser Prozess nicht zufällig oder beliebig ist.

Normativität in Leitlinien. Diese Probleme sind für die vorgelegte Arbeit bedeutsam und begründen das hier vorgestellte Verfahren, da darin auch internationale Leitlinien mit abweichenden Kontextbezügen berücksichtigt wurden.

1.2 Leitlinien

1.2.1 **Definition und Ziele von Leitlinien in der gesundheitlichen Versorgung**

„Leitlinien sind systematisch entwickelte Empfehlungen, die Grundlagen für die gemeinsame Entscheidungsfindung von Ärzten und deren Patienten zu einer im Einzelfall sinnvollen gesundheitlichen Versorgung darstellen.“

Diese Formulierung der Deutschen Gesellschaft für Allgemeinmedizin und Familienmedizin (DEGAM) basiert auf Definitionen des *Institute of Medicine* [Field & Lohr 1990] und der *Agency for Health Care Policy and Research* [AHCPR 1995]. [Gerlach, Beyer et al. 1999]

Leitlinien sollen

- zur Verbesserung der Qualität der medizinischen Versorgung beitragen, indem sie unangemessene Versorgung und unbegründete (z.B. geographische) Varianz der Versorgungsqualität verhindern sollen [Field & Lohr 1992];
- einen effektiven Gebrauch von Ressourcen ermöglichen und damit die Kosteneffektivität verbessern [Woolf 1990];
- die Lücke zwischen dem, was Kliniker in der Routineversorgung tun und was von wissenschaftlichen Forschungsergebnissen unterstützt wird, schließen [Eccles & Mason 2001];
- den Innovationstransfer in die Praxis beschleunigen, Schnittstellen beschreiben und Handlungspfade systematisieren (Arbeitsteilung und Kooperation) [McKinlay et al. 2001].
- Die Ziele für Patienten bestehen in der Verbesserung der gesundheitlichen Outcomes. Durch patientengerechte Informationen werden jedoch noch weitere Effekte möglich, wie Informationsgewinn und Stärkung des Selbstsorgeverhaltens [Woolf et al. 1999].
- Ärzten ermöglichen Leitlinien eine verbesserte Qualität der Entscheidungsfindung [Woolf et al. 1999].

Das Ausmaß, in dem diese Ziele durch die Anwendung einer Leitlinie erreicht werden, ist zwar nur bei wenigen Leitlinien systematisch evaluiert worden. Aus den Ergebnissen dieser internationalen Evaluationsstudien [Grimshaw & Russell 1993, Thomas et al. 1998, Davis & Taylor-Vaisey 1997, Worrall et al. 1997 u.a.] konnten jedoch wichtige Grundsätze für die Leitlinienentwicklung abgelei-

tet werden. Die Einhaltung dieser Grundsätze ist nach Mulrow & Lohr [2001] notwendige Voraussetzung für die *Validität* einer Leitlinie (vgl. Abschn. 1.2.3.2), die ganz wesentlich von der Substanz und Qualität der zugrundeliegenden *Evidenz* bestimmt wird [Mulrow & Lohr 2001].

Evidenz wird in der gesundheitlichen Versorgung als empirisch begründeter Beleg definiert, basierend auf systematischen Untersuchungen im Rahmen eines rigorosen Forschungsdesigns oder auf der Grundlage von nicht-systematischen Beobachtungen (z.B. Erfahrung), die in jedem Falle auf Tatsachen und nicht auf „Offenbarungen, Träumen oder historischen Texten“ beruhen² [Eddy 2001].

Neben evidenzbasierten Leitlinien haben sich weitere Formen evidenzbasierter Entscheidungsunterstützung etabliert, die im folgenden Abschnitt beschrieben werden.

1.2.2 Leitlinien und andere Formen evidenzbasierter Entscheidungsunterstützung

Evidenzbasierte Medizin (EbM), *evidenzbasierte Leitlinien* und *Health Technology Assessment* (HTA) wurden auch als die drei konstituierenden Schlüssel-disziplinen evidenzbasierter Entscheidungsunterstützung bezeichnet [Perleth et al. 2000]. Als vierte sollte hier die „evidenzbasierte Gesundheitsversorgung“ (*Evidence-based Healthcare*, EHC) ergänzt werden.

Die *evidenzbasierte Medizin*, ein neuer Denkansatz in der medizinischen Versorgung individueller Patienten, wurde seit Mitte der 80er Jahre ausgehend von der McMasters Universität in Kanada entwickelt [Perleth & Beyer 1996] und 1996 von Daniel Sackett et al. definiert als

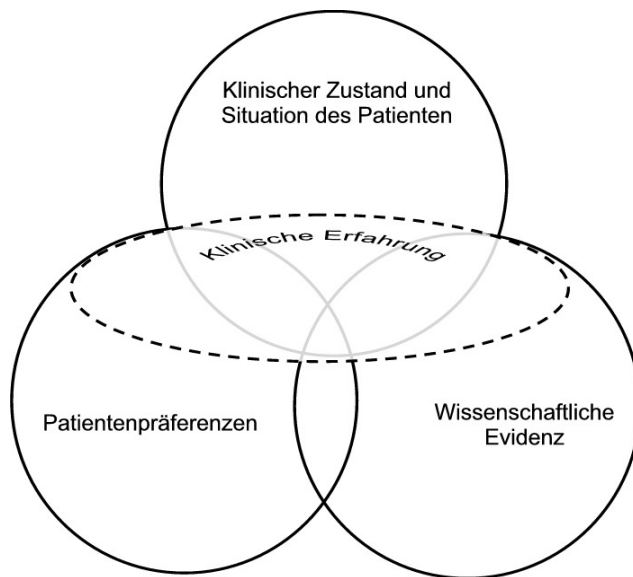
„the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients. The practice of evidence based medicine means integrating individual clinical expertise with best available external evidence from systematic research.“ [Sackett et al. 1996]

Der darin enthaltene Ansatz, dass die evidenzbasierte klinische Entscheidungsfindung auf der Basis von externer wissenschaftlicher Evidenz und individueller klinischer Erfahrung des Arztes stattfindet, wurde von Brian Haynes et al. erweitert [Haynes et al. 2002]. Das 2002 aktualisierte Modell (Abb. 1.1) berücksichtigt weitere essentielle Dimensionen des Entscheidungsprozesses: den klinischen Status und die besonderen Umstände des Patienten sowie unterschiedliche

² “In health care, evidence means empirical observations. They may be systematic observations using rigorous experimental designs or nonsystematic observations (e.g., experience), but they are all empirical observations of real events... not revelations, dreams, or ancient texts.” Übersetzung d. Verf. [Eddy 2001]

Präferenzen von Patienten, Klinikern und Kostenträgern [Eccles & Mason 2001, Haynes et al. 2002].

Abb. 1.1: Modell evidenzbasierter klinischer Entscheidung



Quelle: Raspe in Kunz et al. 2007 nach Haynes et al. 2002

Im Unterschied zur EbM, die Entscheidungen in *ausgewählten Fragestellungen* bei *einzelnen Patienten* unterstützt, greifen evidenzbasierte *Leitlinien* klinische Entscheidungsprobleme in der Versorgung individueller Patienten *systematisch* auf. Sie strukturieren komplexe Versorgungsprobleme und recherchieren nach Maßgabe des *Entscheidungsbedarfs* die jeweils beste externe Evidenz. Der Wert von Leitlinien muss daher ganz wesentlich darin gesehen werden, dass sie komplexe klinische Entscheidungen transparent machen und deswegen ggf. neben der kritischen Bewertung (dem *critical appraisal*) der verfügbaren externen Evidenz auch auf andere Techniken (z.B. Entscheidungsanalysen, Konsensustechniken) zurückgreifen [Elstein 2004, Molewijk et al. 2003, Eccles & Mason 2001, Murphy et al. 1998]. Indem sie vorhandene Evidenz kondensieren und bewerten, überwinden sie zudem zentrale Barrieren für die Umsetzung der EbM: um externe Evidenz in verschiedenen, teilweise unübersichtlichen, uneinheitlichen und / oder kostenpflichtigen Datenbanken aufzuspüren sowie deren Relevanz und Validität einzuschätzen, sind speziell trainierte Techniken der Anwender erforderlich, die darüber hinaus einen erheblichen Zeitaufwand bedeuten [Lelgemann & Donner-Banzhoff in Kunz et al. 2007].

Anders als die EbM und Leitlinien, die Entscheidungssituationen primär unter *klinischer*, d.h. *Handlungs- und Behandlungsperspektive*, ansprechen, wenden sich *Evidence-based Healthcare* und *Health Technology Assessment* an Entscheidungsträger auf Systemebene - z.B. im Gesundheitswesen. Die EHC entstand im angelsächsischen Raum auf die Frage, ob und in welcher Form die Grundsätze der klinisch orientierten EbM auf die Versorgung der gesamten Bevölkerung übertragen werden können. In Anlehnung an das Modell evidenzba-

sierter klinischer Entscheidung (s. Abb. 1.1) wurde für die EHC die Trias aus Erkenntnissen aus hochwertigen wissenschaftlichen Untersuchungen, zur Verfügung stehenden Ressourcen und „Werten“ intendiert. [Busse & Gibis in Kunz et al. 2007] Und *Health Technology Assessment* bezeichnet eine

„Methodik zur systematischen und transparenten Bewertung medizinischer Verfahren und Technologien unter medizinischen, ökonomischen, sozialen und ethischen Aspekten mit dem Ziel, die dazu gehörigen Entscheidungsprozesse mit wissenschaftlichen Erkenntnissen zu unterstützen“ [Henshall et al. 1997 in Perleth 2001].

Methodisch verbindet diese Strömungen evidenzbasierter Entscheidungsunterstützung, dass zunehmend systematische Übersichten die Grundlage dieser Entwicklungen bilden. Tabelle 1.1 zeigt Gemeinsamkeiten und Unterschiede der für die vorliegende Arbeit bedeutsamen Bereiche EbM, Leitlinien und HTA im Überblick.

Tab. 1.1: Gemeinsamkeiten und Unterschiede von EbM, Leitlinien und HTA

	EbM	Leitlinien	HTA
Anwender	Klinisch tätige Ärzte (und Patienten)	Klinisch tätige Ärzte und Patienten, (sekundäre Adressaten)	Entscheidungsträger im Gesundheitswesen, (sekundäre Adressaten)
Zielbevölkerung	Individuelle Patienten	Patientengruppen	Bevölkerung / Bevölkerungsgruppen
Anwendungskontext	Klinische Entscheidungsfindung in ausgewählten Fragestellungen	Unterstützung klinischer Entscheidungsfindung zur Verbesserung der Qualität der medizinischen Versorgung und zur Minimierung ungerechtfertigter Praxisvariation	Entscheidungsfindung bei Kostenübernahme, Investitionen, Regulation
Fragestellung	Beste klinische Untersuchungs- und Behandlungsweise beim individuellen Patienten	Beste regelhafte Entscheidungs- und Vorgehensweise bei einem klinischen Problem (mit der Möglichkeit der Abweichung im Einzelfall)	Effektivität und Effizienz sowie gesellschaftliche Folgen einer Technologie / eines medizinischen Verfahrens

(Fortsetzung Tabelle 1.1)

	EbM	Leitlinien	HTA
Evidenz- grundlage	Zumeist klinische Studien, (Systematische Übersichten, Metaanalysen)	Systematische Übersichten, Metaanalysen, klinische und epidemiologische Studien, (Entscheidungsanalysen), formale und nicht-formale Konsensustechniken	Systematische Übersichten, Metaanalysen, klinische und epidemiologische Studien, ökonomische Evaluationen, Entscheidungsanalysen, politische, ethische, soziologische Analysen
Produkt	(POEM, systematische Reviews)	Leitlinie, Behandlungspfad	HTA-Report

POEM – Patient Oriented Evidence that Matters; Quelle: modifiziert nach Perleth 2001

1.2.3 Anforderungen an Leitlinien

Seit Beginn der 90er Jahre wurde weltweit eine rasch wachsende Zahl von Leitlinien produziert. Bei einer Vielzahl dieser Leitlinien wurde die Validität aufgrund mangelnder methodischer Qualität in Zweifel gezogen [u.a. Eddy 1990, Grilli et al. 2000, Shaneyfelt et al. 1999]. Hauptkritikpunkte lagen insbesondere in inadäquaten Vorgehensweisen zur Identifikation, Evaluation und Synthese der wissenschaftlichen Evidenz sowie in den fehlenden Angaben zur Wahrung der Unabhängigkeit [Grilli et al. 2000, Shaneyfelt et al. 1999]. Daher wurden im Laufe der 90er Jahre spezifische Kriterien entwickelt, deren Berücksichtigung bei der Leitlinienentwicklung die interne und externe Validität von Leitlinienempfehlungen sowie deren praktische Umsetzbarkeit gewährleisten sollten. 1992 formulierten Field und Lohr für das *Institute of Medicine* Anforderungen an medizinische Leitlinien, die von den meisten Ländern sinngemäß übernommen wurden – u.a. 1997 für Deutschland durch die Bundesärztekammer und die Kassenärztliche Bundesvereinigung [BÄK & KBV 1997]. Tabelle 1.2 gibt einen Überblick über diese Anforderungen. Diese Kriterien flossen auch in die spätere Entwicklung von Instrumenten ein, mit deren Hilfe die Qualität von Leitlinien bewertet werden kann, die im folgenden Abschnitt dargestellt werden.

Tab. 1.2: Anforderungen an Leitlinien für die medizinische Versorgung

Validität	Anwendung der Leitlinien sollte zu den gewünschten Gesundheitsergebnissen und -ausgaben führen
Wissenschaftliche Nachweisbarkeit	Angabe von Umfang und Qualität vorhandener Nachweise und untermauerter Expertenurteile

(Fortsetzung Tabelle 1.2)

Voraussichtlicher Outcome	Angabe zu erwartender Gesundheitsergebnisse und damit verbundener Kosten im Vergleich zu alternativem Vorgehen unter Berücksichtigung von Patientenerwartungen
Reliabilität, Reproduzierbarkeit	Leitlinienentwicklung und -anwendung durch unterschiedliche Experten bzw. Anwender sollte zu konsistenten Ergebnissen führen
Praktische Anwendbarkeit	Angabe der Anwendungsfälle (Ein- bzw. Ausschlusskriterien)
Flexibilität	Identifizierung und Berücksichtigung von spezifischen bzw. allgemein erwarteten Ausnahmefällen sowie individuellen Präferenzen der Patienten
Klarheit	Unmissverständliche Terminologie sowie eindeutige, leicht nachvollziehbare Empfehlungen
Multidisziplinarität	Berücksichtigung der Sichtweisen und Probleme aller beteiligten Berufsgruppen im Entwicklungsprozess
Planmäßige Revision	Angabe von Zeitpunkten für planmäßige Aktualisierungen und Überarbeitungen
Dokumentation	Angabe von zugrundeliegenden Annahmen, wissenschaftlichen Nachweisen, Methoden der Leitlinienentwicklung sowie beteiligten Personen und Institutionen

Quelle: entnommen aus Gerlach et al. 1998

1.2.3.1 Methodische Qualität von Leitlinien

Graham et al. [2000] berücksichtigten in einem systematischen Review 13 Instrumente, die dafür konzipiert wurden, die methodische Qualität von Leitlinien zu bewerten, wovon jedoch nur zwei [Cluzeau et al. 1999, Shaneyfield et al. 1999] international validiert waren. Sie identifizierten in einer Inhaltsanalyse 44 Items und gruppieren sie thematisch in folgende Leitlinienattribute: [Graham et al. 2000]

- Methodische Härte (Validität, Reliabilität, Reproduzierbarkeit)
- Klinischer Inhalt (klinische Anwendbarkeit und Anpassungsfähigkeit)
- Prozess der Leitlinienentwicklung (Multidisziplinarität)
- Präsentation (Klarheit der Darstellung)
- Aktualität und Aktualisierung der Leitlinie (planmäßige Reviewverfahren)
- Dissemination der Leitlinie
- Umsetzbarkeit und Implikationen für die Implementierung der Leitlinie
- Evaluation des Leitlinienimpacts

Ergänzt und spezifiziert wurden diese Attribute im AGREE-Instrument, das von einer internationalen Arbeitsgruppe (*AGREE-Collaboration, Appraisal of Guidelines for REsearch & Evaluation*) entwickelt und in 11 Ländern bzw. Regionen – auch in Deutschland – validiert wurde [The AGREE 2003a]. Zusätzlich aufgenommen wurden die dort als *Domänen* bezeichneten Attribute „Geltungsbereich und Zweck“ sowie „redaktionelle Unabhängigkeit“. Die o.g. „Multidisziplinarität“ wurde durch die Beteiligung von Patienten um eine weitere Perspektive erweitert. Die Domäne zur methodologischen Exaktheit bildet in sieben einzelnen Items Bewertungskriterien für die Validität und Reliabilität der Leitlinie ab (detaillierte Beschreibung des AGREE-Instruments: siehe Abschnitt 3.2.3). Von der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) und dem Ärztlichen Zentrum für Qualität in der Medizin (ÄZQ) wurde 2005 das **DE**utsche Instrument zur methodischen **LE**itlinien-**B**ewertung (DELBI) veröffentlicht. Darin wurde das AGREE-Instrument durch ein Zusatzmodul detaillierter Fragen zur Präsentation von Methodik, Inhalt und Implementierungsstrategien ergänzt, welche für die Akzeptanz und Umsetzung von Leitlinien im deutschen Gesundheitssystem besonders relevant sind [AWMF/ÄZQ 2005].

Wegen der zentralen Bedeutung der *Validität* von Leitlinien für ihre Qualität wird diese Eigenschaft im folgenden Abschnitt vertiefend betrachtet.

1.2.3.2 Validität von Leitlinien

“...*the most important attribute of guidelines is validity*” [Mulrow & Lohr 2001] – wobei auch bei Leitlinien in *interne* und *externe* Validität unterschieden werden kann:

Die *interne Validität* ist das Ausmaß, in dem die darin enthaltenen Empfehlungen auf nachgewiesener Wirksamkeit von medizinischen Interventionen beruhen, also in der Regel auf der nachgewiesenen *efficacy*³ [Graham et al. 2000]. Sie wird ganz wesentlich dadurch bestimmt, auf welche Art und Weise die Evidenz gewonnen und evaluiert wurde, welcher Gehalt an Evidenz für die Beantwortung einer klinischen Frage gewonnen werden konnte, ob die Stärke der Empfehlung in Relation zum Evidenzgehalt formuliert wurde und ob Risiko-Nutzen-Abwägungen in die Formulierung der Empfehlungen eingeflossen sind [Mulrow & Lohr 2001].

Die *externe Validität* einer Leitlinie hingegen ist das Ausmaß, in dem ihre „Anwendung ... zu den gewünschten Gesundheitsergebnissen und -ausgaben“ führt [Gerlach et al. 1998], d.h. die Leitlinie muss zunächst verbreitet und dann auch im Praxisalltag umgesetzt werden, bevor ihre Effekte hinsichtlich gesund-

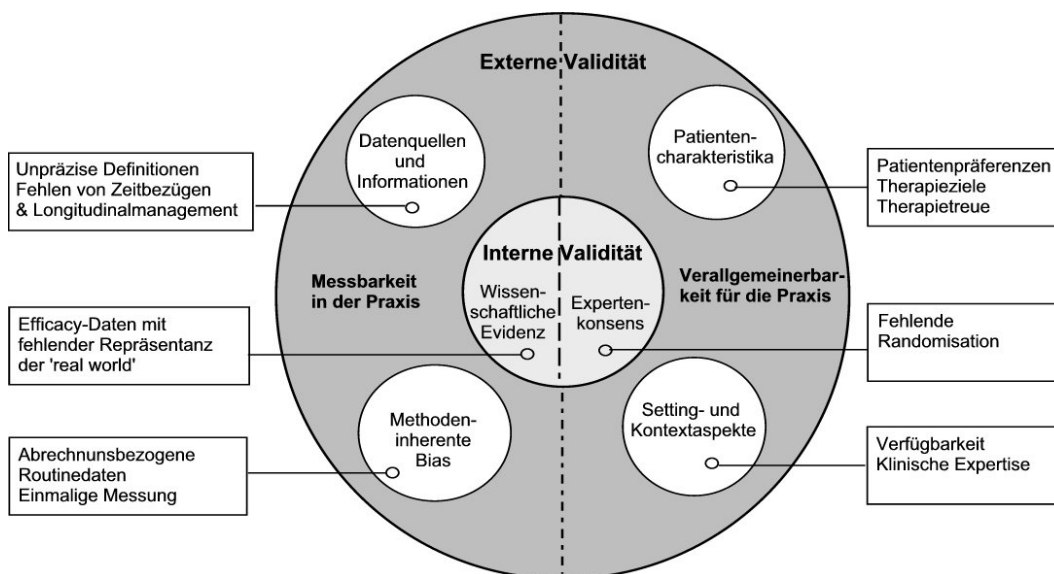
³ *Efficacy*: Beschreibt das Ausmaß der Effekte, die eine Intervention unter Idealbedingungen hervorruft (Wirksamkeit unter Idealbedingungen) [Kunz et al. 2005].

heitlicher und Kosten-Outcomes unter *effectiveness*-Bedingungen⁴ gemessen werden können.

Die Untersuchung der externen Validität einer Leitlinie ist hochgradig komplex, wie u.a. von Graham et al. [2000] für die 1996 publizierte Leitlinie der *Agency for Health Care Policy and Research* (AHCPR) zur chronischen linksventrikulären Dysfunktion gezeigt werden konnte.

Graham et al. [2000] untersuchten die externe Validität der genannten Leitlinie in der Primärversorgung und fanden eine große Zahl patientenseitiger, arztseitiger, organisations- und systemimmanenter Faktoren, welche die externe Validität der Leitlinie beeinträchtigten. In Abb. 1.2 wird eine Auswahl dieser Faktoren graphisch als „Schweizer-Käse-Modell“ dargestellt, d.h. die ermittelten Faktoren reduzieren die Validität. Die Autoren wiesen jedoch auch auf die methodischen Probleme hin, die sich durch eingeschränkte Messbarkeit in der Praxis ergeben.

Abb. 1.2: Validität von Leitlinien



Quelle: modifiziert nach Graham et al. 2000

Auch Pilote & Tager [2002] zeigten, dass exakte Methoden zur Evaluation von Leitlinien fehlen, insbesondere wenn die Effekte von Leitlinien auf die Praxisvarianz untersucht werden sollen. Sie wiesen darauf hin, dass die Messung der Praxisvarianz in erheblichem Maße fehleranfällig sei, besonders wenn aus diesen Ergebnissen Rückschlüsse auf eine Assoziation von Praxisvarianz und Gesundheitsergebnissen gezogen werden. Grund für diese Fehleranfälligkeit sei einerseits die Vielzahl von Faktoren auf Seiten von Patienten, Ärzten und im Gesundheitssystem, die zu berücksichtigen sei – andererseits die einge-

⁴ *Effectiveness*: Beschreibt das Ausmaß der Effekte, die eine Intervention unter Routinebedingungen hervorruft (Wirksamkeit unter Routinebedingungen) [Kunz et al. 2005].

schränkte Qualität von Routinedaten, die typischerweise für derartige Evaluationen herangezogen werden. [Pilote & Tager 2002]

In der vorliegenden Arbeit steht indes die *interne Validität* im Vordergrund, da hier eine Methode zur *Entwicklung* von Leitlinien dargestellt wird. Weitere Schritte, wie Verbreitung und Implementierung von Leitlinien werden davon nicht direkt betroffen.

1.2.4 Leitlinienentwicklung

Anknüpfend an das Konzept der internen Validität von Leitlinien werden zunächst methodisch-praktische Grundsätze zur Leitlinienentwicklung dargestellt, die aus den Ergebnissen internationaler Evaluationsstudien [Grimshaw & Russell 1993, Thomas et al. 1998, Davis & Taylor-Vaisey 1997, Worrall et al. 1997 u.a.] abgeleitet wurden.

1.2.4.1 Grundsätze im methodischen Vorgehen zur Entwicklung valider Leitlinien

Nach Grimshaw & Russell [1993] ist für die Entwicklung einer *validen* Leitlinie erforderlich,

- dass sie auf systematischen Evaluationen beruht,
- dass alle wichtigen Akteure in die Entwicklung einbezogen werden,
- dass eine evidenzbasierte Strategie verfolgt wird, bei der die beste verfügbare wissenschaftliche Literatur systematisch recherchiert und aufgearbeitet wird,
- dass das empfohlene Vorgehen aus der wissenschaftlichen Evidenz hergeleitet wird
- und dass der jeweilige Zusammenhang zwischen Empfehlung und zugehöriger Evidenzstufe dokumentiert wird.

1.2.4.2 Leitlinienentwicklung als mehrstufiger Prozess

Unter Berücksichtigung dieser methodischen Grundsätze konzipierte Margolis [1999] einen mehrstufigen Entwicklungsprozess mit den folgenden Phasen:

(1) *Generieren eines Initialalgorithmus*: Der initiale Algorithmus dient insbesondere dazu, die klinischen Schlüsselfragen zu formulieren, für die nachfolgend Evidenz gefunden werden soll. [Margolis 1999]

(2) *Gewinnen der Evidenz*: Zunächst empfahl Margolis den Entwicklern von Leitlinien ein einführendes Studium von Übersichtsarbeiten und Lehrbüchern, um ein möglichst umfassendes Bild von der klinischen Versorgung der Zielkondition der Leitlinie zu erhalten und um die im Initialalgorithmus definierten Versorgungsschwerpunkte kritisch auf Imbalancen zu hinterfragen. Daran anschließend sollte eine kritische Bewertung *aller relevanten* Forschungsarbeiten zum Thema inkl. einer Graduierung der Qualität der darin enthaltenen Evidenz erfolgen [Margolis 1999]. Dazu sind systematische Recherchen erforderlich, die Studien mit dem für die Fragestellung am besten geeigneten Design identifizieren sollen, insbesondere systematische Übersichten und hochwertige Primärstudien. Bei deren Fehlen sind nach dem Prinzip der Suche nach der *besten verfügbaren Evidenz* auch weniger hochwertige Studien einzubeziehen. [Kopp et al. 2007 in Kunz et al. 2007, Antes 2004, Shekelle et al. 1999 u.a.]

(3) *Durchführung eines Konsensusprozesses*: Mit Hilfe formaler Konsensustechniken, wie z.B. dem Nominalen Gruppenprozess oder der Delphi-Technik werden auf dem Boden der gewonnenen Evidenz Empfehlungen formuliert und hinsichtlich ihrer Stärke graduiert. Dieser Prozess schließt das Abwägen von Risiken und Nutzen von Maßnahmen ein und wird erschwert, wenn belastbare Evidenz fehlt oder wenn konfligierende Studienergebnisse vorliegen. [Margolis 1999]

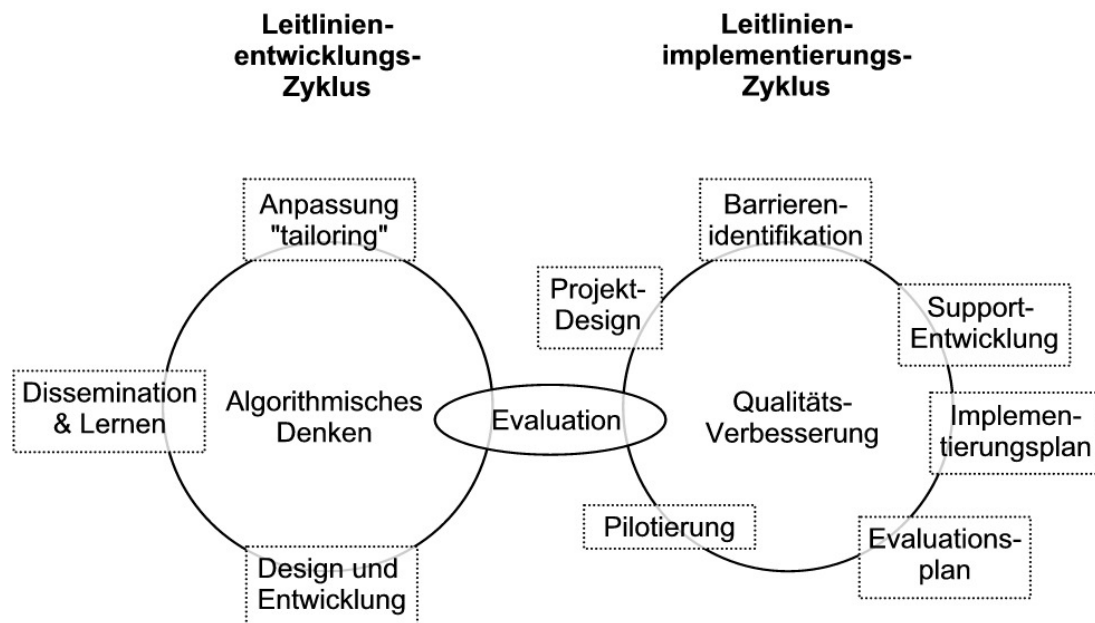
(4) *Niederschrift der ersten Revision des initialen Algorithmus*: Von den Editoren werden die im Konsensusprozess formulierten Empfehlungen inkl. der festgesetzten Empfehlungsstärken in einem ersten Leitlinienentwurf zusammengefasst, und der initiale Algorithmus wird ggf. überarbeitet [Margolis 1999].

(5) *Iterationen der Schritte 3 und 4*: Der Erstentwurf der Leitlinie inkl. dem darin enthaltenen Algorithmus wird mit der Leitlinienentwicklergruppe, bestehend aus methodischen und klinischen Experten, konsentiert und ggf. überarbeitet [Margolis 1999]. Der Prozess kann weitere iterative Überarbeitungsschritte implizieren. In internationalen Leitlinienprogrammen sind oft mehrstufige interne und externe *Peer-Review*-Verfahren und ggf. Pilotierungen der Leitlinie zur Prüfung von Umfang, interner Validität, Akzeptanz und Praktikabilität unter Anwendern vorgesehen [McCorry et al. 2003, Shekelle et al. 1999, Woolf 2000, ACC/AHA 2001, NICE 2003].

(6) *Erstellen des „penultimate draft“* (vorletzter Entwurf): Mit dieser Formulierung wies Margolis darauf hin, dass es sich bei klinischen Leitlinien um Systeme handelt, die auch nach ihrer Dissemination einem fortlaufenden Veränderungsprozess unterworfen sind, indem sie zunehmend an lokale Erfordernisse angepasst werden („*local tailoring*“). Auch unterliegen Leitlinieninhalte dem Veralten und müssen daher kontinuierlich aktualisiert werden, so dass die hier beschriebenen Schritte im Idealfall in einen kontinuierlichen Leitlinienentwicklungs-Kreislauf einmünden. [Margolis 1999]

In der Entwicklung einer klinischen Leitlinie bis zu ihrer Umsetzung in der Praxis sind weitere Schritte nötig, die eine effektive Dissemination und Implementierung beinhalten. Diese Schritte liegen außerhalb der vorliegenden Arbeit und werden daher nur kurz in der untenstehenden Graphik dargestellt, ohne vertiefend besprochen zu werden (Abb. 1.3).

Abb. 1.3: Leitlinien-„Bi-cycle“ zu Entwicklung und Implementierung



Quelle: Margolis & Cretin 1999 (Übers. d. Verf.)

In der internationalen Literatur zur Beschreibung des Vorgehens in der Leitlinienentwicklung finden sich zahlreiche Darstellungsweisen, die verschiedene Entwicklungsphasen unterschiedlich detailliert beschreiben [Kopp et al. in Kunz et al. 2007, McCrory et al. 2003, Woolf 2000, Shekelle et al. 1999]. Auch ist die Grenze zwischen dem, was unter Entwicklung und dem, was unter Implementierung subsummiert wird, nicht einheitlich definiert. Die Darstellung von Margolis [1999] wurde gewählt, da sie durch ihren hohen Abstraktionsgrad variable Ausprägungen in der Anwendung durch unterschiedliche Leitlinienentwicklergruppen zulässt – wie sie z.B. durch setting- und kontextspezifische Erfordernisse notwendig werden – und dennoch eine systematische Vorgehensweise in der Leitlinienentwicklung unterstützt.

Im folgenden Abschnitt wird das Konzept zur Entwicklung von Leitlinien der Deutschen Gesellschaft für Allgemeinmedizin und Familienmedizin (DEGAM) beschrieben, das erstmals im Jahre 1997 verabschiedet und seither fortgeschrieben wurde. Es weist inhaltliche Parallelen zur Darstellung von Margolis [1999] auf und berücksichtigt die spezifischen Erfordernisse der hausärztlichen Versorgung. Alle hausärztlichen Leitlinien, die als DEGAM-Leitlinien autorisiert werden, folgen diesem Konzept – also auch die DEGAM-Leitlinie ‚Herzinsuffizienz‘. Deshalb wird der Darstellung dieses Konzepts besonderer Raum gewidmet.

1.2.4.3 Der DEGAM-Zehnstufenplan

Leitlinien in der Primärversorgung müssen die Besonderheiten des hausärztlichen Handlungsfeldes aufgreifen, insbesondere die dort häufig anzutreffenden komplexen Entscheidungssituationen. Evidenzbasierte Leitlinien für die hausärztliche Praxis müssen nicht nur komprimiertes Wissen in praxisrelevanten Zusammenhängen liefern, sondern sollen auch Transparenz ermöglichen, die im Dialog zwischen Arzt und Patient wesentlich zur Vertrauensbildung beitragen kann. Zu diesen Zwecken wurde von der DEGAM der sogenannte „Zehnstufenplan“ zur Entwicklung, Verbreitung und Implementierung von hausärztlichen Leitlinien konzipiert. In diesem werden wesentliche Prämissen definiert, die in den jeweiligen Entwicklungsphasen bedeutsam sind. Insbesondere sind eine umfassende Suche nach Evidenz, eine sorgfältige Formulierung von Empfehlungen, ein von den Autoren unabhängiger fachinterner *Peer Review* sowie ein Praxistest vor Veröffentlichung der Leitlinie vorgesehen (siehe auch Tab. 1.3). [Gerlach, Beyer et al. 1999, Gerlach, Abholz et al. 1999]

Tab. 1.3: Der DEGAM-Zehnstufenplan

1.	Beauftragung eines koordinierenden Zentrums
2.	Konstituierung einer Arbeitsgruppe und Durchführung des Literaturreviews
3.	Erstellung eines Erstentwurfs
4.	Strukturierte Stellungnahme eines Ärztepannels (Peer Review, Allgemeinmediziner)
5.	Erstellung eines Zweitentwurfs zur Diskussion mit externen Experten (externer Review) mit anschließender erneuter Überarbeitung
6.	Drittentwurf zur Diskussion mit Fachgesellschaften und Verbänden, ggf. anschließender Viertentwurf
7.	„Praxistest“ in Qualitätszirkeln, Praxisnetzen und Einzelpraxen
8.	Erstellung eines Fünftentwurfs und formale Überprüfung
9.	Präsentation durch Autoren und Autorisierung als DEGAM-Empfehlung
10.	Erstveröffentlichung im Forum Qualität der Zeitschrift für Allgemeinmedizin

Quelle: gekürzt nach Gerlach, Beyer et al. 1999

Inzwischen existieren umfangreiche Erfahrungen und ausgereifte Konzepte zur Entwicklung evidenzbasierter Leitlinien, deren Umsetzung jedoch zahlreiche Fragen aufwerfen. Zentrale Aspekte zur Strukturbildung, Evidenzgewinnung und -graduierung sollen im Folgenden näher betrachtet werden.

1.2.5 Offene Fragen in der Leitlinienentwicklung

1.2.5.1 Leitlinienstruktur: Initialalgorithmus und Framework

Kernelemente der Leitlinienentwicklung sind die Entwicklung eines sogenannten *Initialalgorithmus* sowie die Formulierung der ‚richtigen‘ *klinischen Fragen*:

Nach Margolis [1999] werden im *Initialalgorithmus* die notwendigen Versorgungsschritte für das gesundheitliche Problem, das von der Leitlinie behandelt wird, in einem logischen Zusammenhang visualisiert. Margolis [1999] wies darauf hin, dass dieser Algorithmus den systematischen Entwicklungsansatz von Leitlinien begründet und dass Fehler bei der weiteren Leitlinienentwicklung oft darauf zurückgeführt werden können, dass wesentliche Aspekte darin nicht berücksichtigt wurden. In ihrem Ansatz stellt sie die visualisierte (dichotome) Darstellung in Form von Entscheidungsbäumen (Flussdiagrammen) in den Vordergrund, was bei hoch komplexen Gesundheitsproblemen in der primärärztlichen Versorgung nicht unproblematisch ist:

Niederstadt et al. [2001] wiesen darauf hin, dass ein Entscheidungsbaum *entweder* eine größtmögliche Vielzahl von Entscheidungssituationen und Beratungsoptionen berücksichtigt, dafür aber unübersichtlich ist und keine erkennbaren praktischen Vorteile der Anwendung mehr bietet. *Oder* er ist übersichtlich, mit einem hohen Erinnerungswert und hoher Anwendungssicherheit versehen, vernachlässigt dabei aber eine Vielzahl der tatsächlich für individuelle Patienten relevanten Entscheidungs- und Beratungsoptionen. Sie schlussfolgerten, dass Entscheidungsbäume in der Leitlinienentwicklung für die allgemeinmedizinische Versorgung komplexer Gesundheitsprobleme nur begrenzten Wert haben und plädierten für eine „auf Modellbildung basierende Leitlinienstruktur“. [Niederstadt et al. 2001, Niederstadt 2001]

Daher soll der Begriff ‚Initialalgorithmus‘ hier nicht im strengen kybernetischen Sinn (binär formalisierbar) aufgefasst werden, sondern in einem etwas weiteren Sinne: folgerichtig, schlüssig. Er erlangt als Bezugsrahmen, innerhalb dessen die klinischen Schlüsselfragen formuliert werden, jedoch umso größere Bedeutung, je komplexer das angesprochene Gesundheitsproblem ist. Gleichzeitig bleibt die Frage offen, auf welche Art und Weise ein Initialalgorithmus – bzw. eine ‚auf Modellbildung basierende Leitlinienstruktur‘ – entwickelt werden sollte, um *alle relevanten* Entscheidungsprobleme zu erfassen. Zahlreiche Leitlinienprogramme verlangen deswegen schon auf der Stufe der Problemformulierung das Zusammenwirken einer interdisziplinären Entwicklergruppe, um dieser Komplexität gerecht werden zu können [The AGREE Collaboration 2002, AWMF & ÄZQ 2005].

„Asking the right questions and ask them right“ [Oosterhuis et al. 2004] ist essentiell - sowohl für den *critical appraisal* identifizierter Evidenz [Rosenberg & Donald 1995, Richardsen et al. 1995] als auch für die praktischen Entscheidungserfordernisse, die von der Leitlinie angesprochen werden. Mit anderen Worten steht und fällt die Qualität einer Leitlinie schon damit, dass sie *die* praxisrelevanten Probleme behandelt und dass sie die für die Evidenzgewinnung erforderlichen Fragen präzise formuliert.

In Hinblick auf begrenzte Ressourcen findet bei der Entwicklung einer jeden Leitlinie ein *Trade-off*⁵ zwischen Breite und Tiefe der Fragestellung statt: je begrenzter der Anwendungsbereich definiert, je fokussierter die Arbeit ausgerichtet wird, desto profunder kann die Evidenzgewinnung erfolgen. Jede Fokussierung hat jedoch einen Verlust an Breite zur Folge, so dass in Kauf genommen wird, dass relevante klinische Fragen unbeantwortet bleiben. [Hadorn & Baker 1994] *Trade-offs* und Fokussierungsvorgänge setzen notwendigerweise voraus, dass von den Leitlinienautoren Priorisierungen vorgenommen werden, die grundsätzlich zwei Schlüsselfragen aufwerfen: wer nimmt die Priorisierungen vor? Und wie wurde sichergestellt, dass dieser Prozess nicht von Einzelnen dominiert wurde? [Ashton et al. 1999] Zur Unterstützung von Priorisierungsverfahren gibt es unterschiedliche Methoden, die in ihren Auswirkungen kontrovers beurteilt werden: häufig wird eine große Zahl von Prinzipien identifiziert, die nicht immer gut zu definieren und nicht klar zu gewichten sind und für die keine *expliziten Trade-offs* existieren (prinzipienbasierte Priorisierung). Und bei der technikbasierten Priorisierung, wie z.B. dem *Needs Assessment*⁶, ist von Nachteil, dass für einzelne Bedarfe analysiert wird und Vergleiche zwischen verschiedenen Bedarfen häufig unklar bleiben. [Ashton et al. 1999] Abgesehen von diesen offenen methodischen Fragen besteht in Leitlinien oft das Problem, dass nicht dargelegt wird, *auf welche Weise* Themen und Fragestellungen ausgewählt wurden [Rogers 2002], d.h. welche *Trade-offs* definiert wurden und worin diese Priorisierung begründet war.

In der vorliegenden Arbeit wird das strukturbildende Raster, bestehend aus einem Initialalgorithmus und den daraus abgeleiteten spezifischen klinischen Fragestellungen als das *Framework der Leitlinie* bezeichnet.

⁵ *Trade-off* beschreibt eine negative wechselseitige Abhängigkeit zweier Aspekte. Ein *Trade-off* liegt dann vor, wenn man eine Verbesserung oder Erlangung eines Aspektes nur unter Inkaufnahme der Verschlechterung oder dem Verlust des anderen Aspektes erreichen kann (Zielkonflikt). *Trade-off* bedeutet aber auch Kompromiss und Ausgleich des Zielkonfliktes. [Anonymus 2006]

⁶ *Needs Assessment* ist die Analyse der Trias: Bedarf an, Nachfrage nach und Angebot von gesundheitlicher Versorgung. Insbesondere werden Daten aus Epidemiologie (z.B. Prävalenz, Inzidenz, Mortalität), Versorgungsforschung (vorhandene Versorgungspfade und -muster, Inanspruchnahme, erkennbare Defizite) und zur Wirksamkeit von Maßnahmen der strukturierten Versorgung berücksichtigt [Wright et al. 1998].

1.2.5.2 Gewinnung von Evidenz

Diese Phase umfasst verschiedene Handlungsschritte, die teilweise sequentiell, teilweise iterativ erfolgen. Beginnend mit Literaturrecherchen zu den jeweiligen Themen des Frameworks, die je nach Fragestellung und Ressourcen systematisch, umfassend oder orientierend angelegt sein können, wird von der Entwicklergruppe eine Auswahl der identifizierten Literatur vorgenommen. Diese wird einer kritischen Bewertung ihrer internen und externen Validität unterzogen. [Breen & Feder 1999]

Dabei unterscheiden sich die Anforderungen an das prinzipielle Vorgehen zunächst nicht von denen an andere Produkte evidenzbasierter Medizin (EbM), wie etwa systematische Reviews. Im Unterschied zu anderen EbM-Produkten behandeln Leitlinien jedoch eine wesentlich größere Zahl von z.T. komplexen Einzelfragen, da sie Antworten auf die Entscheidungsnotwendigkeiten des klinischen Handelns und nicht auf eine einzelne wissenschaftliche Fragestellung bereitstellen müssen. Weiterhin wird für Leitlinien nach Maßgabe des jeweiligen Entscheidungsbedarfs recherchiert. Das bedeutet, dass für einige entscheidungsrelevante Fragen ohne identifizierbare *hochwertige* Evidenz u.U. ein Hinziehen anderer Quellen und Techniken erforderlich wird. [McIntosh 1999]

Deutlich ist, dass Qualität und Umfang der Evidenzgewinnung nicht nur von spezifischen (insbesondere methodischen) Kompetenzen abhängig sind, sondern auch ganz maßgeblich zur Ressourcenfrage für die Organisation geraten, in deren Auftrag eine Leitlinie entwickelt wird.

1.2.5.3 Graduierung von Evidenz und Empfehlungen

Mit dem Ziel, dem Nutzer von Leitlinien die Vertrauenswürdigkeit der wissenschaftlichen Ergebnisse zu vermitteln, wird die aufgefundene Evidenz graduiert – erkennbar als ‚*levels of evidence*‘. Die Angabe von Stärkegraden der Empfehlung („*grading of recommendations*“) soll hingegen ausdrücken, mit welchem Maß an Sicherheit erwartet werden kann, dass die Anwendung einer Empfehlung mehr nutzt als schadet [Atkins et al. 2004].

Auch dieser Prozess ist sehr komplex. In einer kritischen Analyse von sechs derzeit existierenden Graduierungssystemen wurden erhebliche Mängel aufgezeigt – insbesondere schlechte diskriminatorische Fähigkeiten und niedrige Reproduzierbarkeit. Nach Atkins et al. war keines der bekannten Systeme für alle Anwenderzielgruppen geeignet. [Atkins et al. 2004a]

Von der GRADE-Arbeitsgruppe wurde daraufhin eine Vorgehensweise empfohlen, in der stufenweise Bewertungen der Qualität sowie der Konsistenz vorliegender Studien (und Outcomes) in Hinblick auf deren Entscheidungsrelevanz vorgenommen werden. Sie testeten die Methode in einer Pilotstudie, indem

Evidenzprofile für jedes Outcome erstellt, quantitative Risiko-Nutzen-Bewertungen vorgenommen und abschließende Empfehlungen formuliert wurden⁷. Die Autoren schätzten dieses System selbst als klar, änderungssensitiv und einfach in der Handhabung ein. [Atkins et al. 2005]

Das GRADE-Verfahren ist jedoch noch in seiner Erprobungsphase und findet bislang (noch) keine breite Anwendung. In praktischem Gebrauch befinden sich derzeit bei den verschiedenen Leitlinienentwicklerorganisationen eine große Zahl unterschiedlicher Systeme (mit den oben geschilderten Problemen), die in der vorliegenden Arbeit für die untersuchten Leitlinien berücksichtigt werden mussten.

1.2.5.4 Hoher Ressourcenaufwand in der de novo-Entwicklung

Aus der bisherigen Darstellung lässt sich ableiten, dass die de novo-Entwicklung von qualitativ hochwertigen evidenzbasierten Leitlinien komplex und aufwendig ist. Die erforderliche Zeit wie auch die aufzuwendenden Ressourcen werden von den Entwicklergruppen häufig unterschätzt [Feder et al. 1999].

Die Kosten pro entwickelte Leitlinie bewegen sich nach Literaturangaben zwischen 25.000 £ für eine britische Asthma-Leitlinie und etwa 1 Mill. US-Dollar pro Leitlinie, die von der *Agency for Health Care Policy and Research* (AHCPR) in den 90er Jahren entwickelt wurde [American College of Cardiology's Foundation 2005, Foy et al. 2004, Royal College of Psychiatrists 1994]. Allein für Literaturrecherchen und -Review im Rahmen der Evidenzgewinnung wurde von der AHCPR damals etwa eine Viertelmillion US-Dollar pro Leitlinie aufgewendet [Shekelle et al. 2001]. Allerdings wurden nur etwa 1 bis 5% der 6.000 bis 18.000 Publikationen, die für den jeweiligen Literaturreview kritisch bewertet wurden, in der fertigen Leitlinie zitiert [American College of Cardiology's Foundation 2005].

Die AHCPR (bzw. ihre heutige Nachfolgeorganisation, die *Agency for Healthcare Research and Quality*, AHRQ) ist seit den späten 90er Jahren nicht mehr mit der Entwicklung von eigenen Leitlinien befasst, und Leitlinienautoren sehen sich weltweit (in unterschiedlichem Ausmaß) mit limitierten Ressourcen für die Leitlinienentwicklung konfrontiert.

⁷ *Evidenzprofile* enthielten Angaben über Design der vorliegenden Studien sowie deren Qualität, die Konsistenz der Ergebnisse und die Deutlichkeit der Ergebnisse, *Risiko-Nutzen-Bewertungen* wurden zusammengefasst als Netto-Nutzen / Trade-off / unklare Trade-offs / kein Netto-Nutzen; *Empfehlungen* wurden in folgenden Kategorien abgegeben *do it / probably do it / toss it / don't do it* [Atkins et al. 2005].

Daher wurde die *Adaptation international existierender Leitlinien* als Alternative zur aufwendigen de novo-Entwicklung vorgeschlagen [Europarat 2002, NZGG 2001] und verschiedentlich durchgeführt [u.a. Van Tulder et al. 2006, Graham et al. 2005, Glasier et al. 2003, Voellinger et al. 2003, MacLeod et al. 2002, Wabitsch et al. 1998]. Dabei wurden verschiedene Verfahren verwendet, die darauf abzielen, eine oder mehrere, hinsichtlich Zielerkrankung und Versorgungs-Setting geeignete, internationale Leitlinien zu identifizieren und deren formale methodische Qualität zu bewerten (z.B. mittels AGREE-Instrument [AGREE 2002]), um sie in weiteren Schritten an die nationalen Erfordernisse anzupassen [z.B. NZGG 2006, Oosterhuis et al. 2004].

Fervers et al. [2006] zeigten jedoch erst kürzlich in ihrem umfassenden Literaturreview, dass bislang kein validiertes Verfahren zur *transkontextuellen* Adaptation von Leitlinien existiert, d.h. zur Adaptation einer für *einen* (z.B. nationalen) Kontext entwickelten Leitlinie in einen *anderen* (nationalen) Kontext. Die *ADAPTE Working Group* schlug daraufhin auf der Basis dieses Reviews das in Tab. 1.4 dargestellte Vorgehen vor [Fervers et al. 2006].

Tab. 1.4: *Transkontextuelle Leitlinienadaptation*

- | |
|---|
| <ol style="list-style-type: none">(1) Definition von klinischen Fragestellungen, die von der Zielleitlinie beantwortet werden sollen (Berücksichtigung der Zielpopulation, der interessierenden Interventionen, der Zielgruppe, welche die Leitlinie nutzen soll, der klinischen Outcomes, die berücksichtigt werden sollen sowie von Versorgungskontext und -setting)(2) Recherche nach Quell-Leitlinien(3) Vergleich der Inhalte der Quell-Leitlinien hinsichtlich der von ihnen abgedeckten Fragestellungen mit den Fragestellungen der Zielleitlinie(4) Bewertung von methodischer Qualität und Kohärenz zwischen Evidenz und Empfehlungen in der/den Quell-Leitlinie(n) sowie von deren Anwendbarkeit und Akzeptanz im Zielkontext(5) Adaptation der Empfehlungen(6) Externer Review der adaptierten Leitlinie(7) Annahme, Bestätigung und Implementierung der adaptierten Leitlinie |
|---|

Quelle: nach Fervers et al. [2006]

Als problematisch muss dabei angesehen werden, dass zwar validierte Instrumente zur Bestimmung der *formalen* – insbesondere methodischen – Qualität einer Leitlinie verfügbar sind (wie z.B. das AGREE-Instrument [AGREE 2003a] oder das darauf basierende deutsche DELBI-Instrument [AWMF/ÄZQ 2005]). Diese Instrumente sind jedoch nicht geeignet, um die *inhaltliche Angemessenheit* von Leitlinienempfehlungen zu überprüfen [AWMF/ÄZQ 2005]. Eine derartige Prüfung, die von Fervers et al. [2006] ebenfalls gefordert wird, ist jedoch von zentraler Bedeutung für die Validität der Zielleitlinie.

Da für eine *Prüfung der inhaltlichen Angemessenheit* derzeit keine Verfahren oder Instrumente zur Verfügung stehen, ist es erforderlich, die Dimensionen zu erfassen, die dabei berücksichtigt werden müssen. Daher wird der Prozess der

Formulierung von Empfehlungen mit den dabei auftretenden wichtigsten Problemen im folgenden Abschnitt näher betrachtet.

1.2.5.5 Der Prozess der Formulierung von Leitlinienempfehlungen

Wie oben beschrieben wurde, wird im Prozess der Leitlinienentwicklung zunächst nach der besten verfügbaren Evidenz recherchiert. Die aufgefundene Evidenz enthält Statements darüber, welche Beobachtungen unter spezifischen Umständen in wissenschaftlichen Studien gewonnen wurden, ist also in der Essenz *deskriptiv* [McIntosh 1999]. Von der Leitlinienentwicklergruppe werden in einem *Bewertungsprozess* Leitlinienempfehlungen formuliert, in den mehrdimensionale Risiko-Nutzen-*Abwägungen* einfließen sollen [AGREE 2003a]. Diese enthalten aktive *Handlungsanweisungen*, welche die Entscheidungsfindung zwischen Arzt und Patient leiten sollen und weisen somit eine *normative* Dimension auf. Das bedeutet zunächst, dass Evidenzaussagen und Handlungsempfehlungen *nicht deckungsgleich* sein können. [McIntosh 1999]

Verschiedene Autoren betrachten diesen Prozess der Formulierung von Empfehlungen darum auch als unvermeidliche Quelle von *Subjektivität* in der Leitlinienentwicklung [McIntosh 1999, Fletcher in Margolis & Cretin 1999, Mulrow & Field 2001, Detsky 2006]. (Unkontrollierte) *Subjektivität* wird in der evidenzbasierten Medizin jedoch als wesentliche Quelle für *Bias* – also verzerrte oder verzerrende Folgerungen aus der gefundenen Evidenz – angesehen.

Unseres Erachtens greift diese ‚bias-zentrierte‘ Betrachtungsweise zu kurz, was im Folgenden vertiefend betrachtet werden soll.

1.2.5.5.1 Ableitung normativer aus deskriptiven Aussagen

Ob und wie man die Ableitung eines *normativen* Urteils (Aussage, die ein Sollen vorschreibt) aus einem *deskriptiven* Urteil (beschreibende Aussage) wissenschaftlich begründen kann, ist von jeher in der Philosophie umstritten gewesen. Insbesondere mit der Aufklärung und der Entwicklung positiver (empirisch vorgehender) Wissenschaften und der zugehörigen Wissenschaftstheorie ist die Frage noch verschärft in den Vordergrund getreten.

Erstmals wies der schottische Philosoph David Hume (1711-1776) in seinem *“Treatise of Human Nature. Book III: Of morals. Part I: Of virtue and vice in general. Sect. I: Moral distinctions not deriv’d from reason.”* darauf hin, dass es einen logischen Unterschied zwischen beschreibenden und bewertenden Sätzen gibt. Er gelangte zu dem Schluss, dass wir kein „was sein soll“ („*what ought to be*“) aus einem „was ist“ („*what is*“) herleiten können [Anonymus 2006a, Fieser 2006, Nielsen 1979]. Für Hume ist die Deduktion eines „*ought*“ aus einem „*is*“ *“altogether inconceivable”* [Cohon 2004], da es keine *logische* Regel für die

Ableitung normativer aus deskriptiven Sätzen gibt. Man spricht daher auch von einem *deskriptivistischen Fehlschluss*, wenn vom Sein auf ein Sollen geschlossen wird. [Grunwald 2004]

Wie Grunwald⁸ [2000] weiter ausführt, lässt sich einem deskriptivistischen Fehlschluss jedoch entgegenhalten, „dass es praktisch weder möglich noch sinnvoll ist, ständig das gesamte Arsenal an normativen Anteilen der Gesellschaft einer ethischen Reflexion zu unterziehen: ein reflexiver Stillstand sämtlicher gesellschaftlicher Aktivitäten wäre die Folge“. Unter Berufung auf Habermas' „Wahrheitstheorien“⁹ legt er dar, dass in *lebensweltlich* relevanten Entscheidungen *explizite* ethische Reflexionen die Ausnahme darstellen, da die weitaus meisten Entscheidungen in Form von Zweck- / Mittel-Erwägungen auf der *Handlungsebene* ohne besondere Reflexion ihrer normativen Hintergründe getroffen werden, „der normative Hintergrund dieser Entscheidungen ... in seiner Geltung zumeist nicht hinterfragt, sondern als gültig anerkannt (wird)“ [Grunwald 2000]. Die (ethische) *Diskursebene*, auf der dieser normative Hintergrund infrage gestellt werden kann, werde jedoch nur in besonderen Ausnahmefällen betreten: wenn die Moral als handlungsleitende Maxime und Regel in Individuen, Gruppen oder der Gesellschaft aufgefasst wird und die Ethik die Zusammenhänge und Geltungsbedingungen dieser Regeln in einem normativen Gefüge aufdeckt, dann sind nur *im Falle von moralischen Konflikten* (oder Ambiguitäten) ethische Reflexionen als weiterreichende Orientierung erforderlich. [Grunwald 2004]

Grunwald ordnet auch die meisten *technikrelevanten* – inkl. HTA – Entscheidungen in dem Sinne ein, dass der normative Anteil der Entscheidungsgrundlagen nicht einer eigenen Reflexion unterzogen wird, sondern dass es als für die jeweilige Situation vorgegeben angesehen und für die Entscheidung auch so akzeptiert wird und nennt sie *business as usual*-Situationen. [Grunwald 2004, Grunwald 2000]

⁸ Prof. Dr. Armin Grunwald ist Mathematiker und promovierter Physiker, habilitierte auf dem Gebiet der Philosophie und leitet seit 1999 das Institut für Technikfolgenabschätzung und Systemanalyse (ITAS) im Forschungszentrum Karlsruhe. Er befasst sich insbesondere mit Fragen zur Theorie der Technikfolgenabschätzung und zur Ethik der Technik und hat dazu bereits zahlreiche nationale und internationale Publikationen verfasst (Homepage von A. Grunwald: <http://www.itas.fzk.de/mahp/grunwald/grunwald.htm>). Der hier zitierte Aufsatz [Grunwald 2004] reflektiert den Bedarf nach und die Probleme um normative Bewertungen in der Technikfolgenabschätzung (*technology assessment*, TA) und im *Health Technology Assessment* (HTA), das sich historisch aus dem TA heraus entwickelt hat [Perleth 2001]. Dieser Aufsatz wurde ausgewählt, da seine Prämissen in engem Bezug zum Leitlinienthema stehen (vgl. Gegenüberstellung HTA und Leitlinien im Kapitel 1.2.2) und weil seine Argumentation wesentliche philosophische Fragestellungen abdeckt, die für die vorliegende Arbeit von Bedeutung sind. Für die wörtlichen Zitate wurden zwei weitere Aufsätze [Grunwald 2002, Grunwald 2000] hinzugezogen, in denen wesentliche Passagen des o. g. in deutscher Sprache dargestellt sind.

⁹ Grunwald zitiert: Habermas J (1973) Wahrheitstheorien. In: Fahrenbach H (Hrsg.): Wirklichkeit und Reflexion. Walter Schulz zum sechzigsten Geburtstag. Pfullingen, S. 211-265

„Problematisch wird es dort, wo die Grenze eines anerkannten normativen Gerüsts überschritten wird; problematisch aber vielleicht noch weit mehr ..., diese Grenze überhaupt zu erkennen, vor allem angesichts der Situation, dass sie nicht a priori festgelegt ist, sondern einerseits einem historischen Wandel unterliegt, gesellschaftlich selbst konstruiert und andererseits durch das Einbringen neuer Technik und neuer wissenschaftlicher Ergebnisse selbst verändert werden kann“ [Grunwald 2000]. Daher besteht für Grunwald die entscheidende Frage darin, „ob man sich in konkreten Entscheidungssituationen innerhalb oder außerhalb eines anerkannten normativen Gerüsts befindet und auf welche Weise diese Frage beantwortet wird“ [Grunwald 2000].

Für die Beantwortung des ersten Teils der Frage wäre es also wichtig, woran sich festmachen lässt, dass eine *business as usual*-Situation vorliegt, die keine eigene ethische Reflexion erfordert (und somit einen Sein-Sollen-Schluss erlaubt). Dafür nennt Grunwald Anforderungskriterien an den normativen Rahmen¹⁰, innerhalb dessen keine moralischen Konfliktsituationen oder Ambiguitäten vorliegen. Er macht jedoch darauf aufmerksam, dass diese genannten Kriterien immer nur – und nur dann – in Bezug auf einen spezifischen örtlich und zeitlich definierten *Entscheidungskontext* gelten können. [Grunwald 2004] Der zweite Teil der Frage, auf welche Weise mit den Wertbezügen in der Technikfolgenabschätzung umgegangen werden soll, ist noch weitgehend ungeklärt – praktiziert werden derzeit verschiedene Techniken (empirische Wertforschung, normative Ethik, Partizipation). [Grunwald 2000, Grunwald 2004]

Es kann aus dem oben gesagten für Leitlinien erstens gefolgert werden, dass eine Empfehlung nur dann *inhaltlich angemessen* sein kann, wenn sie nicht nur die empirischen Fakten, sondern darüber hinaus den dafür geltenden normativen Rahmen berücksichtigt. Da dieser normative Rahmen nur für einen spezifischen Kontext gilt, lässt sich daraus auch der strenge *Kontext-* (und *Setting-*) *bezug* von Leitlinienempfehlungen ableiten.

Zweitens folgt daraus: wenn in Leitlinien (normative) Empfehlungen auf dem Boden (deskriptiver) Evidenz formuliert werden, bedeutet das zunächst, dass hier prinzipiell ein deskriptivistischer Fehlschluss vorliegt, der allerdings unter gewissen Umständen ‚erlaubt‘ ist. Voraussetzung dafür ist, dass ein normativer Rahmen vorliegt,

- der *vollständig* ist, indem er alle erforderlichen Werturteile abdeckt,
- der *eindeutig zu interpretieren* ist,

¹⁰ Anforderungen an einen normativen Rahmen: Vollständigkeit (d.h. alle Wertimplikationen der anstehenden Entscheidung müssen hinreichend abgedeckt sein), Konsistenz (Freisein von Widersprüchen), Eindeutigkeit (hinreichende Einvernehmlichkeit unter den Beteiligten über die Interpretation des normativen Rahmens), Akzeptanz (des normativen Rahmens als Entscheidungsgrundlage bei den Beteiligten), Befolgung (des normativen Rahmens durch die Beteiligten im betreffenden Feld). [Grunwald 2000, Grunwald 2004]

- der in sich *konsistent*, d.h. frei von Widersprüchen ist,
- der von allen Beteiligten – Entwicklern wie Nutzern von Leitlinien – *akzeptiert* wird und
- der von allen Beteiligten *befolgt* wird.

Diese Anforderungen an den normativen Rahmen dürften nur in den seltensten Fällen a priori vorauszusetzen sein, etwa wenn dieser Rahmen durch die Gesetzgebung oder bspw. Regelungen zur Ausübung des ärztlichen Berufs vorgegeben wird. Hier dürfte unstrittig sein, dass die innerhalb dieses Rahmens abgegebenen Empfehlungen befolgt werden, ohne den normativen Hintergrund infrage zu stellen. Das trifft jedoch eher auf Richtlinien zu und erklärt den Imperativ der darin abgegebenen „Empfehlungen“ („Richtlinien hat man zu befolgen“ [Gerlach et al. 1998, Selbmann 2006]).

Die weitaus meisten Empfehlungen in klinischen Leitlinien überschreiten jedoch den normativen Rahmen von Gesetzgebung und berufsständischem Regelwerk, indem sie auch andere Ebenen von Werturteilen betreffen, wie im folgenden Abschnitt näher ausgeführt wird. Wenn für diese Ebenen eine Pluralität von Werturteilen angenommen werden muss, können moralische Konfliktsituationen und Ambiguitäten nicht ausgeschlossen werden, die einem weiteren *Diskurs* zugeführt werden müssen. Es ist daher notwendig, die *strittigen* Fragen und Themen zu identifizieren, die mit allen Beteiligten – z.B. im Rahmen eines formalen interdisziplinären, multiprofessionellen Konsensprozesses unter Beteiligung von Patientenvertretern – weiter bearbeitet werden müssen. Da ein derartiger Diskurs notwendigerweise unvollkommen bleiben muss – u.a., da es schon aus praktischen Erwägungen nicht möglich ist, *alle Beteiligten* in den Diskurs einzubeziehen und die dafür getroffenen *prädeliberativen Einverständnisse*¹¹ nicht per se für jeden der Teilnehmer zutreffen – lässt sich daraus weiterhin ableiten, dass Leitlinienempfehlungen keine imperativen Aussagen treffen (können) und warum im begründeten Einzelfall – z.B. bei Patientenpräferenzen, die sich außerhalb des bedingenden normativen Rahmens der Empfehlung befinden – von ihnen abgewichen werden kann und soll.

1.2.5.5.2 Ebenen der Normativität in Leitlinien

Folgt man der Aussage von Hartmann, ist die Medizin eine „*Handlungswissenschaft* eigenen Rechts“ [Hartmann 1989 in Raspe 2005]. Handlungswissenschaften – neben der Medizin auch andere, wie z.B. die klinische Psychologie – sind soziale Handlungssysteme mit je eigenen wissenschaftlichen Gegenständen, Problemen, Zielen, Normen und Mitteln. Gemeinsames Ziel von Hand-

¹¹ Als prädeliberatives Einverständnis bezeichnet Grunwald nach Gethmann die minimalen Voraussetzungen, denen alle Diskursteilnehmer zustimmen und von denen ausgehend der Diskurs entwickelt wird. [Grunwald 2004, Grunwald 2002]

lungswissenschaften ist es, den natürlichen Verlauf der Entwicklung von Menschen, die in direkten Kontakt zu ihnen kommen, in eine als positiv bewertete Richtung zu verändern. Dabei sind die intendierten Veränderungen mit persönlichen, zwischenmenschlichen und sozialen Werten verbunden. Daraus folgt, dass Handlungswissenschaften immer auch Normwissenschaften sind, deren „Feststellungen ... gleichzeitig auch Wertstellungen (sind)“. [Raspe 2005]

Das lässt sich in der Medizin bereits an der (nicht unumstrittenen) Definition von Gesundheit als „*state of complete physical, mental and social well-being*“ [WHO 2006] und den daraus abgeleiteten Grundeinschätzungen von Krankheit und Kranksein ablesen: „Krank heißt unter irgendeinem, aber keineswegs immer gleichen Gesichtspunkt schädlich, unerwünscht, minderwertig.“ [Jaspers 1973 in Raspe 2005]. Grundeinschätzungen von Krankheit als „schädlich, unerwünscht, minderwertig“ sind eindeutige Werturteile und führen überdies zu einem Handlungszwang in der Medizin, dem sogenannten „therapeutischen Imperativ“ [Katsch 1958, Hartmann 1994 in Raspe 2005]. Dieser ist so stark, „dass jeder Handlungsverzicht (der auch eine Form von Handeln ist) einer besonderen ... Begründung bedarf“ [Raspe 2005].

Wenn medizinisches *Handeln*, d.h. die klinische Tätigkeit, an Werten und Normen ausgerichtet ist, so stellt sich die Frage nach der Normativität der medizinischen *Wissenschaft*, die „die Determinanten und Konsequenzen ... (des medizinischen) Handelns ... einer rigorosen wissenschaftlichen Kontrolle unterzieht“ [Raspe 2005]. Dazu heißt es im Grundsatz 6 der Deklaration des Weltärztebundes von 2002:

„Oberstes Ziel der medizinischen Forschung am Menschen muss es sein, prophylaktische, diagnostische und therapeutische Verfahren sowie das Verständnis für die Ätiologie und Pathogenese der Krankheit zu verbessern. Selbst die am besten erprobten prophylaktischen, diagnostischen und therapeutischen Methoden müssen fortwährend durch Forschung auf ihre Effektivität, Effizienz, Verfügbarkeit und Qualität geprüft werden“ [zitiert in Raspe 2005].

Eines der zentralen Probleme in der gesundheitlichen Versorgung besteht daher darin, auf welcher Grundlage entschieden werden kann, ob eine Maßnahme – z.B. eine therapeutische Intervention – *wirkt*, um dem oben genannten Ziel der Medizin als Handlungswissenschaft zu entsprechen. Im klinischen Sinne – und darum geht es in der vorliegenden Arbeit¹² – steht dahinter das *Konzept der Effektivität*. Wie Ashcroft [2002] ausführt, kann eine Maßnahme nur in Hin-

¹² In der Frage nach der Wirksamkeit sind durchaus auch andere Dimensionen enthalten, wie etwa nach pathophysiologischen, pathobiochemischen, pharmakologischen etc. Zusammenhängen, die ähnlich wie in den nicht-medizinischen Naturwissenschaften Fragen nach der Kausalität aufwerfen oder gesundheitspolitische Dimensionen, in denen beispielsweise Fragen nach den Kosten einer Maßnahme zur Entscheidung über Ressourcenallokationen von Bedeutung sind [vgl. Ashcroft 2002].

blick auf irgendein Ziel effektiv sein, sei es intendiert, wie z.B. die Verbesserung eines gesundheitlichen Zustands durch eine medikamentöse Therapie oder aber ungewollt, wie die mit dieser Medikation einhergehenden Nebenwirkungen. Wiewohl von der *klinischen Forschung* objektive Ergebnisparameter bestimmt und gemessen werden können, um eine klinische Entscheidung auf einen rationalen Boden zu stellen, unterliegen diese jedoch persönlichen Präferenzen, wie z.B. die Frage danach, ob die Nebenwirkungen des Medikaments um der erstrebten Verbesserung des Gesundheitszustandes willen in Kauf genommen werden. Gerade diese Präferenzen sind nach Ashcroft [2002] einer empirischen Betrachtung zwar ebenfalls zugänglich, deren Ergebnisse sind für die individuelle Entscheidung für oder gegen diese Therapie jedoch irrelevant. Es gibt also kein allumfassendes und rein faktisch begründbares Effektivitätsgebot, das Konzept als solches ist untrennbar mit Werturteilen verbunden.

Das *Konzept der evidenzbasierten Medizin* (EbM) ist historisch¹³ in besonderem Maße mit dem Konzept der Effektivität verbunden und steht auch heute „im Dienst einer klinischen Zweckrationalität ..., (in der es) ihr um klinische Wirksamkeit, um klinischen Nutzen (geht)“ [Raspe 2007 in Kunz et al. 2007]. Dahinter stehen starke ethische Motive: vermeidbare Schäden von Patienten abzuwenden (*primum nil nocere*) und die gesundheitliche Versorgung zu verbessern [Ashcroft 2004]. Zu diesem Zweck stellt die EbM Methoden bereit, mit deren Hilfe externe Evidenz importiert und genutzt werden kann [Raspe 2005]. Es steht außer Zweifel, dass diese Methodik zahlreiche Fragen aufwirft, die ihr von Kritikern auch immer wieder vorgeworfen werden und die auch den erkenntnistheoretischen wie ethischen Hintergrund betreffen. Dazu zählt insbesondere die Frage, wie unterschiedliche Arten von Evidenz miteinander verglichen oder kombiniert werden können [Ashcroft 2004]. Vineis [2004] nannte dazu vier Kernprobleme:

1. das Problem der ‚gewichteten Evidenz‘, wenn mehrere Studien (z.B. multiple RCTs) mit konfliktären Ergebnissen vorliegen
2. das Problem, dass zu einer Frage bislang keine hochwertigen Studien durchgeführt wurden
3. das Problem, dass nur Studien von minderer methodischer Qualität vorliegen und deren Ergebnisse – allein oder in einer meta-analytischen Synthese – schwierig zu interpretieren sind
4. das Problem, dass zu einer Fragestellung keine hochwertigen Studien durchgeführt werden können, weil es z.B. unethisch oder unpraktikabel wäre. [Vineis 2004]

¹³ Die evidenzbasierte Medizin wird geschichtlich insbesondere auf die Arbeiten von Archie Cochrane zurückgeführt, dessen Werk der frühen 70er Jahre den Titel ‚*Effectiveness and efficiency: random reflections on health services*‘ trug und sich insbesondere mit Interventionen zweifelhafter Wirksamkeit und Sicherheit auseinandersetzte. [Ashcroft 2004]

Nach Vineis ist es daher schwierig, eine klare Grenze zwischen ‚*evidence*‘ und ‚*lack of evidence*‘ festzulegen– diese beruht immer darauf, was als Evidenz *akzeptiert* wird [Vineis 2004] und daher auch auf Werturteilen, die je nach vorliegendem normativen Rahmen voneinander abweichen können.

Die *Primärversorgung* weist darüber hinaus spezifische Probleme und bisweilen sogar immanentes Konfliktpotential zur EbM auf. Slowther und Mitarbeiter [2004] wiesen darauf hin, dass hier häufig belastbare Evidenz über die Effektivität einer Maßnahme in diesem Setting¹⁴ fehlt, dass zahlreiche Interventionen der Primärversorgung hochgradig komplex und nur schwerlich in RCTs abzubilden sind, dass die Berücksichtigung von Patientenautonomie und -präferenzen hier stärker im Vordergrund steht als in anderen Versorgungsbereichen und dass Hausärzte auch einen bedeutenden Teil von Public Health-Aufgaben, wie z.B. Impfprophylaxen wahrnehmen, deren Ziele mitunter mit individuellen Behandlungszielen konfliktieren. [Slowther et al. 2004]

Wenn also in evidenzbasierten Leitlinien Handlungsempfehlungen abgegeben werden, die sich auf die beste verfügbare Evidenz stützen und dabei Präferenzen von Patienten und Klinikern berücksichtigen, die sich schon aus Gründen der praktischen Umsetzbarkeit der Empfehlungen an den lokalen Gegebenheiten des Gesundheitssystems mit Verfügbarkeit, Zugang und Leistungsfinanzierung orientieren müssen, ohne dabei ethische Prinzipien zu vernachlässigen und sich diese Leitlinie an die primärmedizinische Versorgung wendet, müssen demnach alle oben genannten Ebenen von Normativität berücksichtigt werden. Es müssen jedoch nicht alle Ebenen in einem *Diskurs* infrage gestellt werden, sondern nur jene mit *strittigen Fragen*: da in *evidenzbasierten* Leitlinien *primär* klinisch tätige Ärztinnen und Ärzte angesprochen werden, können fundamentale ethische Prinzipien, wie z.B. das *primum nil nocere* hier ebenso im Sinne von *prädeliberativen Einverständnissen* [Grunwald 2004, Grunwald 2002] vorausgesetzt werden, wie die grundsätzliche Akzeptanz von Konzept und Methoden der EbM.

Zu der Frage, *auf welche Art und Weise* mit Normativität in der Leitlinienentwicklung oder ganz allgemein in der Nutzenbewertung umgegangen werden soll, gibt es bislang zwar einige grundsätzliche Vorstellungen und – eher abstrakte – Konzepte [Grunwald 2004, Strech 2007], jedoch keine allgemein akzeptierten und praktikablen Lösungen. Einvernehmen herrscht jedoch darüber, dass man diese Normativität zur Kenntnis nehmen muss und normative Aspekte nach Möglichkeit transparent darstellen sollte.

¹⁴ Anders als die Sekundär- oder Tertiärversorgung ist die hausärztliche Versorgung häufig mit unspezifischen Erkrankungen oder Symptomen konfrontiert, und es wird hier z.B. bei chronisch verlaufenden Krankheiten nicht nur ein Ausschnitt einer (besonders schweren, eindeutig zuzuordnenden) Krankheitsperiode versorgt, sondern auch die vom Patienten häufig kaum wahrgenommenen Frühstadien sowie die terminalen Phasen.

1.2.5.5.3 Weitere Einflüsse auf die Formulierung von Empfehlungen mit Beispielen aus der Leitlinienentwicklung

Im Anschluss an die theoretischen – und teilweise abstrakten – Betrachtungen zur Normativität in Leitlinien sollen nun anhand praktischer Beispiele normative und weitere Aspekte dargestellt werden, für die Einflüsse auf die Formulierung von Leitlinienempfehlungen gezeigt wurden. Anhand der ausgewählten Primär- und Sekundärstudien soll anschließend ein Modell entwickelt werden, das die *Prüfung der inhaltlichen Angemessenheit* von Leitlinienempfehlungen als konzeptionelles Gerüst unterstützen kann. Die folgenden Beispiele entstammen Primärstudien und Übersichtsarbeiten, die verschiedene Einflüsse auf die Formulierung von Leitlinienempfehlungen untersucht haben und hier ohne Anspruch auf Vollständigkeit diskutiert werden:

1. Noch in neueren Untersuchungen zur formalen Leitlinienqualität wurden *inadäquate methodische Vorgehensweisen* bei der Leitlinienentwicklung beobachtet, die zu inhaltlich unangemessenen Empfehlungen führen können. Häufig beobachtete Mängel sind in der nachfolgenden Tabelle (Tab. 1.5) zusammengefasst.

Tab. 1.5: *Inadäquate methodische Vorgehensweise bei der Entwicklung von Leitlinien*

Unsystematische Leitlinienentwicklung, in der Folge: <ul style="list-style-type: none">- Unterschiedliche und teilweise konfligierende Empfehlungen in 20 Leitlinien zur oralen Antikoagulation in England, Wales und Schottland [Thomson et al. 2000]- Unterschiedliche und teilweise konfligierende Empfehlungen in fünf Leitlinien und vier Konsensus-Statements zur Versorgung von Demenzerkrankungen [Beck et al. 2000]
Abweichende Interpretationen der bestehenden Evidenzlage mit unterschiedlichen und teilweise konfligierenden Empfehlungen in zwei US-amerikanischen Leitlinien zur Versorgung einer ambulant erworbenen Pneumonie [Ravago et al. 2000]
Unterschiedliche Auswahl der Evidenzquellen: <ul style="list-style-type: none">- Geringer Literatur-Overlap der Zitate mit einigen wichtigen Inkonsistenzen in den Therapieempfehlungen von Leitlinien aus 13 Ländern zum Diabetes mellitus Typ 2 [Burgers et al. 2002]- Inadäquate Berücksichtigung von systematischen Übersichtsarbeiten in einer vergleichenden bibliographischen Untersuchung von vier internationalen Leitlinien zur Raucherentwöhnung [Silagy et al. 2001]
Überrepräsentation lokaler Evidenzquellen: <ul style="list-style-type: none">- Überrepräsentation britischer Studien in einer vergleichenden Untersuchung der bibliographischen Quellen in 15 Leitlinien, entwickelt in und für Großbritannien [Grant et al. 2000]- Selektion lokaler Evidenzquellen, in der Folge unterschiedliche und teilweise konfligierende Empfehlungen in vier internationalen Leitlinien zur Therapie des arteriellen Hypertonus [Ramsey et al. 1998]

Quelle: Eigene Darstellung

2. Ein weiterer, in jüngster Zeit häufig geäußelter Kritikpunkt an Leitlinienautoren ist deren *Umgang mit potentiellen Interessenkonflikten*: eine bloße Offenlegung von finanziellen Verflechtungen mit Industrie und anderen (Profit-) Organisationen verhindert nicht deren potentiellen Einfluss auf die Formulierung von Empfehlungen durch einzelne Mitglieder der Leitlinienentwicklergruppe. Diesen ist u. U. das Ausmaß nicht bewusst, in dem sie (auch unbewusst) in ihren Entscheidungen beeinflusst sind. [Choudhry et al. 2002, Taylor & Giles 2005]
3. Andere Arbeiten gingen der Frage nach, inwiefern die *Zusammensetzung der Leitlinienentwicklergruppe* und deren *Interaktionen* untereinander die Formulierung von Empfehlungen beeinflussten. Aus der Sozialforschung bekannte Phänomene der Gruppendynamik wurden auch in Leitlinienentwicklergruppen beobachtet, wie z.B. positive Assoziationen zwischen dem Status eines Teilnehmers im Konsensusprozess und seinem Beitrag zum Entscheidungsprozess [Pagliari & Grimshaw 2002]¹⁵. Dieser Effekt wurde auch von Murphy et al. [1998] in einem HTA-Bericht gezeigt. Darüber hinaus fanden Murphy et al. Evidenz dafür, dass unterschiedlich zusammengesetzte Gruppen zu abweichenden Entscheidungen gelangen (z.B. tendieren Spezialisten signifikant häufiger dazu, Interventionen ihres Spezialgebiets zu favorisieren), dass heterogene (multidisziplinäre) Gruppen zumeist anders entscheiden als homogene Gruppen und dass die Gruppengröße einen Einfluss auf Entscheidungen in Konsensusprozessen hat [Murphy et al. 1998].
4. Darüber hinaus wurden Arbeiten identifiziert, die *soziokulturelle Einflüsse* für abweichende Empfehlungen in verschiedenen Leitlinien zu einem Thema verantwortlich machten:
 - **Christiaens et al. [2004]** befragten Autoren verschiedener europäischer Leitlinien zur unkomplizierten Zystitis und wiesen darauf hin, dass auf derselben Evidenzgrundlage *diametrale* Empfehlungen formuliert wurden: eine ‚positive‘ Darstellung der ‚besten verfügbaren Evidenz‘ führte zur Therapieempfehlung bei allen Patienten mit dieser Kondition in der einen Leitlinie, während in anderen Leitlinien eine *kritische* Diskussion der Limitationen der Evidenz herangezogen wurde, um die Option ‚keine Therapie‘ zu begründen. Christiaens et al. diskutierten kulturelle Hintergründe dafür, ob ‚keine

¹⁵ Zentrale Ergebnisse der Sozialforschung über Gruppenverhalten sind nach Pagliari & Grimshaw [2002] z.B.,

- dass Gruppen komplexe und dynamische ‚Organismen‘ sind, deren Zusammensetzung, Ziele und Sozialstruktur sich über die Zeit verändern und deren Leistung von der Interaktion ihrer Mitglieder abhängt
- dass das Verhalten der Gruppenmitglieder wie auch der gesamten Gruppe sozialen Einflüssen unterliegt, die sowohl vom Druck der Mehrheit (Konformität) als auch von einzelnen Mitgliedern (Anpassung und Gehorsam) ausgehen
- dass der Grad des sozialen Einflusses mit dem Status des Individuums oder der Gruppe positiv assoziiert ist.

Therapie' als Alternative akzeptiert wird, wie hoch das empfundene Sicherheitsbedürfnis ist, und wem die höchste Entscheidungsautorität¹⁶ zugeschrieben wird.

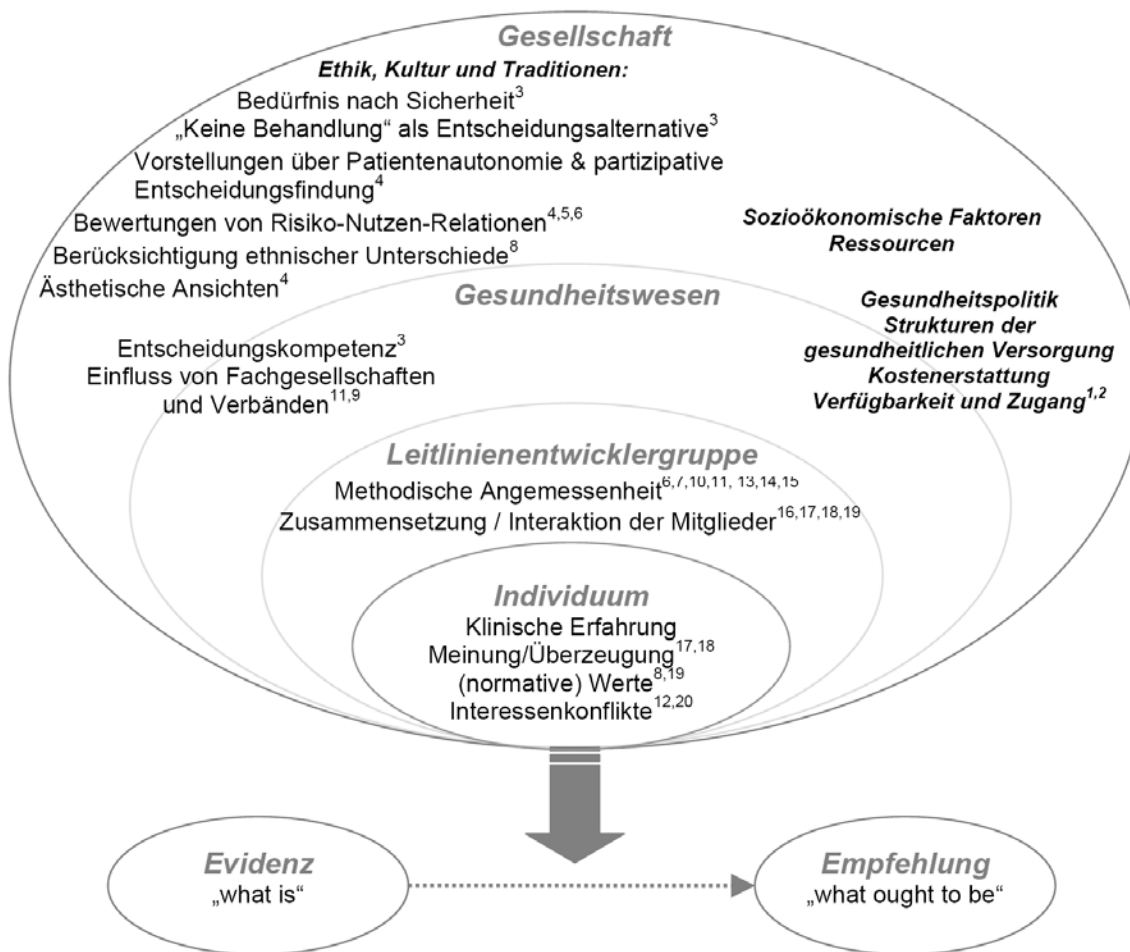
- In zwei anderen Arbeiten wurden unterschiedliche Therapieempfehlungen internationaler Hypertonieleitlinien untersucht [Ramsay et al. 1998, Fahey & Peters 1996]. **Ramsay et al. [1998]** verglichen die Indikationen zur medikamentösen Therapie bei leichter arterieller Hypertonie und berichteten, dass deutlich differierende Schwellenwerte in den untersuchten Leitlinien angegeben wurden. **Fahey & Peters [1996]** fanden in ihrer Querschnittstudie einen variablen Anteil von Patienten mit ‚kontrollierter‘ Hypertonie (17,5% bis 84,6%), je nachdem, welche der untersuchten fünf Leitlinien zugrundegelegt wurde. Ursache der abweichenden Empfehlungen war in beiden Arbeiten die unterschiedliche Risiko-Nutzen-Bewertung in den Leitlinien.
- **Eisinger et al. [1999]** verglichen die Empfehlungen aus zwei evidenzbasierten Konsensus-Statements aus den USA und Frankreich für Frauen mit einem erhöhten Brustkrebs- und Ovarialkarzinom-Risiko bei Nachweis von Mutationen der Gene BRCA 1 und 2. Beide Statements unterschieden sich erheblich in Bezug auf ihre Empfehlungen zu Brust-Selbstuntersuchung, Lebensstilfaktoren und insbesondere in der Beratung zur prophylaktischen Mastektomie und Ovariectomie. Die Autoren machten divergente Ansichten zur Patientenautonomie und partizipativen Entscheidungsfindung, zum wahrgenommenen Patienten-Selbstsorgeverhalten sowie kulturelle Unterschiede im symbolischen und ästhetischen Wert der weiblichen Brust in beiden Ländern für die Varianz der Empfehlungen verantwortlich.

1.2.5.5.2 Modell von Einflussfaktoren auf die Formulierung von Empfehlungen

Die Ergebnisse der vorgenannten und weiterer Untersuchungen zu diesem Thema sind in Abb. 1.4 schematisch dargestellt, indem die Beobachtungen verschiedenen Ebenen – vom individuellen Leitlinienautor bis zur gesamtgesellschaftlichen Ebene – zugeordnet wurden.

¹⁶ Befragte deutsche Leitlinienautoren beschrieben einen impliziten Einfluss von Fachspezialisten und deren tradiertem Verhalten, ‚ganz sicher zu gehen‘, auf den Inhalt der Empfehlungen, während holländische Autoren davon überzeugt waren, dass Hausärzte ihre Patienten schnell und effektiv behandeln. Die befragten belgischen Leitlinienautoren wiederum zielten darauf ab, in den Formulierungen ihrer Empfehlungen die Reaktionen von Patienten zu antizipieren, und ihre norwegischen Kollegen glaubten, dass evidenzbasierte Empfehlungen den Prozess partizipativer Entscheidungsfindung zwischen Arzt und Patient befördern. [Christiaens et al. 2004]

Abb. 1.4: Einflüsse auf die Formulierung von Leitlinienempfehlungen



Quellen: ¹Aldrich et al. 2003, ²DeMaeseneer & Derese 1999, ³Christiaens et al. 2004, ⁴Eisinger et al. 1999, ⁵Fahey & Peters 1996, ⁶Ramsay et al. 1998, ⁷Beck et al. 2000, ⁸Manna et al. 2003, ⁹Littlejohns et al. 1999, ¹⁰Grant et al. 2000, ¹¹Burgers et al. 2002, ¹²Choudry et al. 2002, ¹³Ravago et al. 2000, ¹⁴Silagy et al. 2001, ¹⁵Thomson et al. 1998, ¹⁶Murphy et al. 1998, ¹⁷Pagliari & Grimshaw 2002, ¹⁸Pagliari et al. 2001, ¹⁹Shekelle et al. 1999, ²⁰Taylor & Giles 2005

Die vorgenommene Zuschreibung dürfte nicht in allen Fällen unstrittig sein (z.B. ob es sich bei Wertvorstellungen von Leitlinienautoren um tradierte Wertvorstellungen ‚der Gesellschaft‘ oder um die von Individuen handelt), und die Darstellung erhebt auch keinen Anspruch auf Vollständigkeit. So wurden beispielsweise normative Einflüsse, die implizit¹⁷ die Präsentation von Evidenz und die damit zusammenhängenden Entscheidungen beeinflussen [Molewijk 2003 und 2004], nicht berücksichtigt. Die modellhafte Reduktion liefert jedoch pragmatische Ansatzpunkte für eine *Prüfung der inhaltlichen Angemessenheit*.

¹⁷ Der Begriff *implizit* wurde von Molewijk et al. [2003] als ‚nicht intendiert‘ oder ‚unabsichtlich‘ aber auch als ‚unbemerkt‘ bzw. ‚unbewusst‘ verwendet. Diese Definition wurde für die vorliegende Arbeit übernommen.

1.2.5.2.5 Schlussfolgerungen für Leitlinienentwicklung und -adaptation

Zusammenfassend kann aus dem oben Gesagten folgendes abgeleitet werden:

- Bei der Formulierung von Handlungsempfehlungen in Leitlinien entstehen nicht nur zufällige (oder systematische) *Verzerrungen*, sondern es fließen dabei zwangsläufig – auch durchaus gewollt – Wertvorstellungen, Normen oder z.B. auch ‚Sachzwänge‘, die sich aus intendierten Ressourcenallokationen ergeben, ein.
- Da es keine *logische* Regel für die Ableitung von (normativen) Leitlinienempfehlungen (die immer auch Werturteile enthalten) aus (deskriptiver) Evidenz gibt, ist erklärlich, warum Leitlinienempfehlungen zu derselben Fragestellung selbst bei der Berücksichtigung identischer Evidenz in unterschiedlichen (normativen) Kontextbezügen voneinander abweichen (können). Das gleiche gilt auch für die Leitlinienanwender – aus keiner Leitlinienempfehlung folgt *logisch*, dass sie (ungeachtet des Patienten) befolgt werden soll.
- Normativität in Leitlinien ist ein Umstand, der zur Kenntnis genommen werden muss und möglichst *transparent und explizit* dargestellt werden sollte, um den erforderlichen Diskurs zu unterstützen.

1.2.5.2.6 Implikation für die vorliegende Arbeit

Wenn also unterschiedliche Leitlinienentwicklergruppen unter Berücksichtigung identischer Evidenz zu abweichenden (kontextspezifischen) Empfehlungen gelangen – was insbesondere bei voneinander abweichenden Wertesystemen zu erwarten ist – dann darf einer *transkulturellen* Leitlinienadaptation in der oben dargestellten Vorgehensweise [Fervers et al. 2006] mit einiger Skepsis begegnet werden.

Ein Verfahren, das existierende Quelleitlinien bei der Entwicklung einer Zielleitlinie berücksichtigt, sollte unseres Erachtens in der Lage sein, die *inhaltliche Angemessenheit* von Leitlinienempfehlungen zu prüfen,

- indem Ebenen des Faktischen, basierend auf Studienergebnissen, herausgearbeitet werden und deutlich gemacht wird, wo potentiell konfligierende Werturteile einen Diskurs erforderlich machen,
- indem Evidenzbewertungen auf ihre Validität geprüft werden, wobei ein prädeliberatives Einverständnis in die fundamentalen ethischen Prinzipien ärztlichen Handelns und in das Konzept der evidenzbasierten Medizin vorausgesetzt wird.

Ein solches Verfahren wird in der vorliegenden Arbeit vorgestellt und im systematischen Leitlinienreview für die Entwicklung einer evidenzbasierten Leitlinie zur hausärztlichen Versorgung der chronischen Herzinsuffizienz als Modell angewendet.

2 Forschungsfragen

Ziel der vorliegenden Arbeit ist es, ein *systematisches, transparentes* und *effektives* Verfahren darzustellen, mit dessen Hilfe eine *ressourcensparende de novo-Entwicklung* einer qualitativ hochwertigen evidenzbasierten Leitlinie ermöglicht wird. Dazu sollen bereits existierende evidenzbasierte Leitlinien ausgewertet werden, um aus deren Erkenntnissen systematisch zu lernen. Die Methode hat zu berücksichtigen, dass Leitlinien beschreibende, aber auch bewertende Aussagen enthalten. Da Werturteile in verschiedenen Kontextbezügen nicht notwendigerweise identisch sind, muss das Verfahren normative Bezüge transparent machen, vorgenommene Evidenzbewertungen unter Anwendung von grundsätzlichen Arbeitsmethoden der evidenzbasierten Medizin auf ihre Validität prüfen und damit einen Diskurs strittiger Fragen vorbereiten. *Thema probandum* der Methode des systematischen Leitlinienreviews ist deswegen v.a., wie methodisch zuverlässig verhindert werden kann, dass in der Zielleitlinie inhaltlich unangemessene Empfehlungen formuliert werden, welche auf Verzerrungen, kurzschlüssigen Evidenzbewertungen oder unreflektierten Werturteilen der Quellleitlinien beruhen. Als Ergebnis des Vorgehens sollen ein *valider Erstentwurf* der zu entwickelnden Leitlinie sowie der *identifizierte Forschungsbedarf* zur weiteren Evidenzgewinnung *in strukturierter Form* vorliegen.

Das Verfahren, das als ‚systematischer Leitlinienreview‘ (SLR) bezeichnet wird, wird am Modell der Entwicklung einer evidenzbasierten Leitlinie für die hausärztliche Versorgung der chronischen Herzinsuffizienz vorgestellt.

Leitende Forschungsfrage:

Ist ein ‚systematischer Leitlinienreview‘ ein systematischer, transparenter und effektiver Verfahrensschritt für eine ressourcensparende Entwicklung einer qualitativ hochwertigen evidenzbasierten Leitlinie zur hausärztlichen Versorgung der chronischen Herzinsuffizienz?

1. Welche Verfahrensschritte umfasst ein systematischer Leitlinienreview, und wie sind sie zu begründen?
2. Wie wird die interne Validität der zu entwickelnden Leitlinie bei dieser Vorgehensweise sichergestellt?
3. Wie wird Transparenz für normative Bezüge hergestellt?
4. Welchen Limitationen unterliegt der hier vorgelegte SLR zur chronischen Herzinsuffizienz?

3 Methodisches Vorgehen

Die Verfahrensweise des ‚systematischen Leitlinienreviews‘ wurde ausgehend von grundsätzlichen Arbeitsmethoden der evidenzbasierten Medizin konzipiert. Im Ablauf orientiert sie sich prinzipiell an der Vorgehensweise des systematischen Reviews mit systematischer Literaturrecherche, Auswahl der identifizierten Publikationen nach a priori festgelegten Kriterien, Bewertung der methodischen Qualität der berücksichtigten Dokumente, Datenextraktion und Informationssynthese [z.B. Egger et al. 2001, Khan et al. 2001]. Da der SLR auf die systematische Evaluation von *Leitlinien* abzielt, sind davon abweichende zusätzliche Verfahrensschritte erforderlich wie:

- die Entwicklung des Frameworks (strukturbildendes Raster), da Leitlinien im Gegensatz zu systematischen Reviews eine Vielzahl von teilweise komplexen Fragestellungen zu beantworten haben,
- die Konsistenzprüfung und die Prüfung der Validität der Ergebnisse anhand von Primärquellen zur Prüfung der inhaltlichen Angemessenheit und zur Identifikation strittiger Fragen in Vorbereitung des anschließenden Diskurses.

Das Verfahren des SLR wird hier am Beispiel der Leitlinienentwicklung zur chronischen Herzinsuffizienz beschrieben und ausgeführt. Daher sind zunächst zentrale Voraussetzungen festzulegen, welche die Zielerkrankung, den Zielkontext, den dafür identifizierten Veränderungsbedarf sowie die ausgewählten Ergebnisparameter betreffen.

3.1 Prämissen

Wichtige Voraussetzungen für die vorliegende Arbeit zielen auf die Festlegung von Gültigkeitsbedingungen und Adressaten der zukünftigen Leitlinie ab. Dazu sollen folgende Fragen beantwortet werden:

- Wie ist die Zielerkrankung zu beschreiben, für welche die zukünftige Leitlinie gelten soll, und welche pathophysiologischen Modelle werden für die weitere Leitlinienentwicklung herangezogen?
- Welcher Zielkontext (Zielgruppe, Setting) soll von der Leitlinie adressiert werden?
- Welcher Veränderungsbedarf wurde aus der gegenwärtigen Versorgungssituation ermittelt, der durch die Implementierung der fertigen Leitlinie ange-regt werden soll?
- Welche Ergebnisparameter werden als Zielkriterien für die Bewertung der Evidenz ausgewählt?

3.1.1 Beschreibung und Eingrenzung der Zielerkrankung

3.1.1.1 Definition und Beschreibung der Herzinsuffizienz

3.1.1.1.1 Definition der Herzinsuffizienz (ICD-10: I50)

In einer häufig verwendeten *pathophysiologisch* orientierten Definition hat die WHO 1995 die Herzinsuffizienz als die Unfähigkeit des Herzens charakterisiert, den Organismus mit ausreichend Blut und damit genügend Sauerstoff zu versorgen, um den Stoffwechsel unter Ruhe- wie unter Belastungsbedingungen zu gewährleisten [WHO 1995 in Hoppe et al. 2005].

Neuere Forschungsergebnisse haben zu einer erweiterten Definition der Herzinsuffizienz als multisystemische Störung geführt. Durch die Dysfunktionen der Herz- und Skelettmuskulatur sowie der Nierenfunktion wird ein komplexes Muster neurohumoraler Veränderungen – u.a. eine Stimulation des sympathischen Nervensystems – kompensatorisch hervorgerufen [Jackson et al. 2000]. Diese Betrachtungsweise ist weniger hämodynamisch orientiert und erleichtert damit das Verständnis z.B. für die Wirksamkeit von β -Rezeptorenblockern bei einem reduziert kontraktionsfähigen Herzen [Jessup & Brozena 2003].

Klinisch liegt dann eine Herzinsuffizienz vor, wenn typische Symptome wie z.B. Dyspnoe, Müdigkeit (Leistungsminderung) und/oder Flüssigkeitsretention auf dem Boden einer kardialen Funktionsstörung bestehen [Hoppe et al. 2005].

3.1.1.1.2 Ätiologie

In westlichen Industrienationen sind für neun von zehn Fällen der Herzinsuffizienz eine koronare Herzerkrankung, ein arterieller Hypertonus oder eine Kombination aus beiden verantwortlich [Lip et al. 2000]; auch Vorhofflimmern (VHF) allein oder in Kombination mit anderen Ursachen ist nicht selten [McMurray & Stewart 2000].

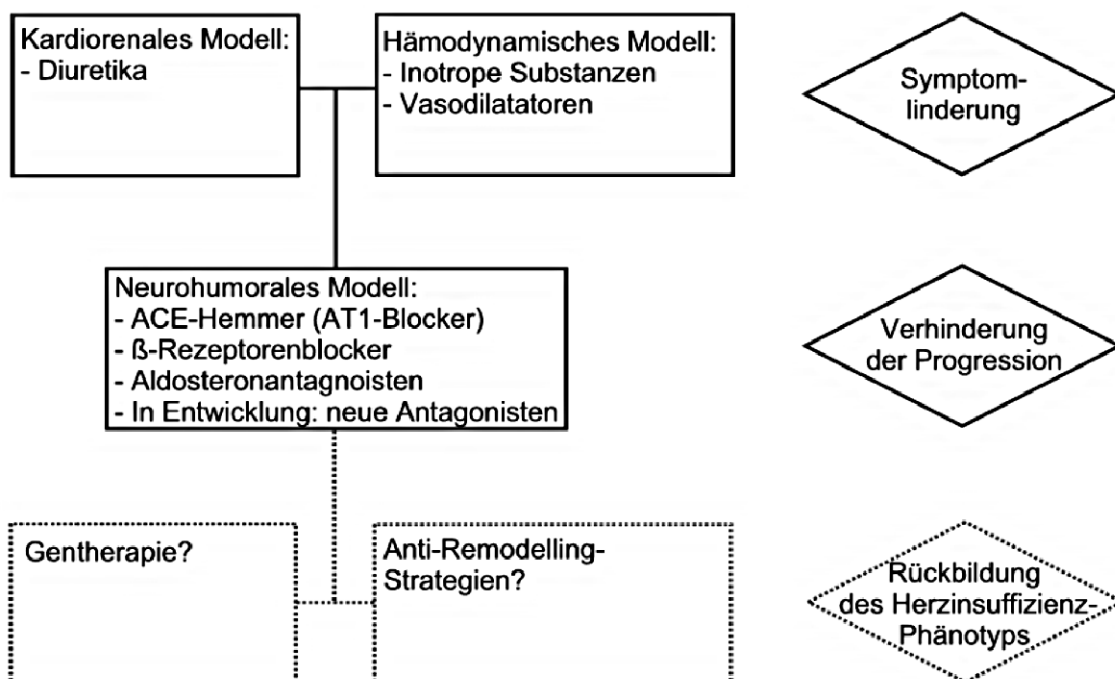
Mit Abnahme des akuten rheumatischen Fiebers sind Herzklappenerkrankungen eine zunehmend seltenere Ursache: in der Framingham-Studie waren sie für jede 50. Herzinsuffizienz bei Männern und für jede 30. bei Frauen ursächlich. Weitere eher seltene Ursachen sind Kardiomyopathien (dilatativ, hypertroph / obstruktiv, restriktiv, obliterativ), angeborene (valvuläre) Herzerkrankungen, andere Arrhythmien und Perikarderkrankungen. Unter den Kardiomyopathien sind dilatative Formen die häufigsten, insbesondere infolge alkoholtoxischer (ca. 2 bis 3% der Patienten mit chronischer Herzinsuffizienz) oder infektiöser Genese. Eine Sonderform ist das *high output failure*, bspw. verursacht durch Anämie, Thyreotoxikose oder av-Fisteln. [Lip et al. 2000]

3.1.1.1.3 Pathophysiologie

Das hochgradig komplexe Syndrom der chronischen Herzinsuffizienz ist Gegenstand aktueller Forschung, da die vielfältigen pathophysiologischen Mechanismen, die bei der Entstehung und Unterhaltung eine Rolle spielen, bislang noch nicht vollständig aufgeklärt werden konnten. Auch existiert derzeit kein allgemeingültiges Modell, das alle bekannten pathophysiologischen Mechanismen vereint. [Jessup & Brozena 2003]

Die folgende Abbildung (Abb. 3.1) veranschaulicht verschiedene Modelle, die über die Zeit entwickelt wurden und die daraus abgeleiteten Therapieansätze.

Abb. 3.1: Pathophysiologische Modelle der chronischen Herzinsuffizienz



Kardiorenales Modell: extensive Wasser- und Salzretention, verursacht durch pathologische Veränderungen des renalen Blutflusses

Hämodynamisches oder kardiozirkulatorisches Modell: reduziertes Herzzeitvolumen als Folge reduzierter Pumpleistung und erhöhten peripheren Gefäßwiderstands; Vorwärts- und Rückwärtsversagen

Neurohumorales Modell: Bedeutung des Renin-Angiotensin-Systems und des sympathischen Nervensystems für das Fortschreiten der Erkrankung, die mit den vorangegangenen Modellen nicht erklärt werden konnte

Quelle: modifiziert nach Mann 1999

Die neueren Darstellungen des *neurohumoralen Modells* gehen davon aus, dass eine Initialschädigung des Myokards eine Kaskade neurohumoraler Aktivierungen auslöst, die ihrerseits sequentiell und regional spezifische Effekte an

unterschiedlichen Zielorganen auslösen, [Davila et al. 2006, Watson et al. 2006] mit der der Organismus versucht, die kardiale Auswurfleistung zu erhalten. Von klinisch zentraler Bedeutung hinsichtlich diagnostischer oder therapeutischer Ansätze sind für die vorliegende Arbeit:

1. Die Aktivierung des sympathischen Nervensystems
2. Die Aktivierung des Renin-Angiotensin-Aldosteron-Systems (RAAS)
3. Die Aktivierung von natriuretischen Peptiden

Das **sympathische Nervensystem** – als Kompensationsmechanismus zur Aufrechterhaltung der kardialen Auswurfleistung schon früh aktiviert – verschlechtert bei chronischer Aktivierung die kardiale Funktion durch Erhöhung des peripheren Gefäßwiderstands, durch Salz- und Wasserretention im Rahmen einer RAAS-Aktivierung und durch direkte Effekte am Herzmuskel mit Absterben von Muskelzellen und Hypertrophie-Induktion sowie durch Störung der autonomen Regulation am Sinusknoten und Senkung der Herzfrequenzvariabilität [Jackson et al. 2000].

Die **Aktivierung des RAAS** manifestiert sich u.a. in einem erhöhten Anfall von Angiotensin II, einem potenten Vasokonstriktor im renalen und systemischen Kreislauf, welches die Freisetzung von Noradrenalin im sympathischen Nervensystem stimuliert, den Vagotonus hemmt und die Freisetzung von Aldosteron fördert. Aldosteron führt zu Wasser- und Salzretention und erhöht die Kaliumausscheidung [Jackson et al. 2000]. Darüber wirkt es offenbar auch direkt auf Myokard-Dysfunktion und -Fibrose sowie auf die Endotheldysfunktion [Struthers 2005].

Die **natriuretischen Peptide ANP** (*Atrial Natriuretic Peptide* – atrialer Herkunft) und **BNP** (*Brain Natriuretic Peptide* freigesetzt aus Zentralnervensystem und Ventrikelmyokard) wirken antagonistisch zu den oben genannten Systemen und führen zu erhöhter Natriumausscheidung und Vasodilatation. [Jackson et al. 2000]

Eine Reihe weiterer Mediatoren (wie z.B. Endothelin, NO-Metaboliten, Tumornekrosefaktor α , Arachidonsäurederivate und Vasopressin – die meisten von ihnen sympathomimetisch und nur wenige sympatholytisch) sowie verschiedene chemo- und barorezeptorenvermittelte Reflexe werden derzeit hinsichtlich ihrer pathophysiologischen Rolle unter der Frage künftiger neuer Therapieansätze bei der Herzinsuffizienz untersucht [Watson et al. 2006, Davila et al. 2005, Champion et al. 2003, Jackson et al. 2000]. Da sie sich noch im Experimentalstadium befinden, wird die Auswahl der pathophysiologischen Modelle für die vorliegende Arbeit auf die oben genannten beschränkt.

3.1.1.1.4 Formen der Herzinsuffizienz

Unter den zahlreichen verschiedenen Einteilungen der Formen der Herzinsuffizienz werden im Folgenden nur diejenigen aufgeführt, die für die Definition der Zielkondition bedeutsam sind:

Links- und Rechtsherzinsuffizienz: Basierend auf der Theorie eines ‚Vorwärts-‘ bzw. ‚Rückwärtsversagens‘ zeigt eine Linksherzinsuffizienz Symptome einer pulmonalen Stauung, während eine Rechtsherzinsuffizienz mit Ödemen, Ergussbildung und Leberstauung imponiert [Braunwald & Grossman 1992]. Da sich im Laufe einer Linksherzinsuffizienz häufig eine Generalisierung (mit Symptomen der Links- wie der Rechtsherzinsuffizienz) herausbildet [Braunwald & Grossman 1992], ist diese Einteilung von untergeordneter Bedeutung.

Abgegrenzt werden soll davon eine *isolierte Rechtsherzinsuffizienz bei Cor pulmonale* (Kombination aus Hypertrophie und Dilatation des rechten Ventrikels auf dem Boden einer pulmonalen Hypertension [McFadden & Braunwald 1992]). Sie wurde aus der Leitlinienthematik *ausgeschlossen*, da sich Ätiologie, Pathogenese und Behandlung grundsätzlich unterscheiden und ein anderes Vorgehen erfordern.

Akute und chronische Herzinsuffizienz: Während eine *akute* Herzinsuffizienz bei einer *vormals gesunden* Person durch ein plötzlich einsetzendes Ereignis (z.B. massiver Myokardinfarkt, akute bradykarde oder tachykarde Herzrhythmusstörung) auftritt, erfolgt die Ausbildung einer *chronischen* Herzinsuffizienz über einen längeren Zeitraum mit *typischen strukturellen Veränderungen* und Anpassungsreaktionen [Braunwald & Grossman 1992], wobei keine genaue zeitliche Abgrenzung zu definieren ist.

Die vorliegende Arbeit ist auf die Versorgung der *chronischen Herzinsuffizienz* gerichtet, wobei *akute Dekompensationen* (z.B. tachykarde Herzrhythmusstörungen, die zu einer akuten Beschwerdeverschlechterung führen) ebenfalls berücksichtigt wurden.

Systolische und diastolische Herzinsuffizienz: Bei Patienten mit linksventrikulärer Dysfunktion findet sich in etwa der Hälfte der Fälle eine *reduzierte Ejektionsfraktion* (EF)¹⁸, d.h. eine *systolische* Herzinsuffizienz¹⁹.

¹⁸ In der Literatur wird kein einheitlicher cut-off verwendet. In der vorliegenden Arbeit wurde eine systolische Dysfunktion für eine EF < 40% definiert [Hoppe et al. 2005, Jessup & Brozena 2003].

¹⁹ Prävalenzangaben zur Häufigkeit von systolischer und diastolischer Herzinsuffizienz variieren je nach untersuchter Stichprobe und verwendeten diagnostischen Kriterien. Bursi et al. [2006] fanden bei 44% der untersuchten Patienten (Durchschnittsalter 75 J) eine isolierte diastolische Herzinsuffizienz, bei 37% eine kombinierte systolische (EF < 50%) und diastolische Dysfunktion. D.h. bei ca. 80% wurde echokardiographisch eine diastolische Herzinsuf-

An eine *diastolische* Herzinsuffizienz muss bei Vorliegen typischer klinischer Symptome und Zeichen sowie weiterer technisch-apparativer Befunde der Herzinsuffizienz (z.B. im Röntgen-Thorax) und *erhaltener linksventrikulärer Ejektionsfraktion* gedacht werden. In der Echokardiographie finden sich Hinweise auf eine verminderte myokardiale Dehnbarkeit (Compliancestörung) und im Doppler erhöhte enddiastolische linksventrikuläre Füllungsdrücke. [Vasan & Levy 2000] In Linksherzkatheteruntersuchungen konnten dabei pathologische Veränderungen in der aktiven *Relaxation* und in der passiven *Stiffness* (Steifigkeit) gezeigt werden [Zile et al. 2004]. Eine diastolische Dysfunktion ist häufiger bei älteren Patienten, insbesondere bei Frauen mit systolischem Hypertonus und Myokardhypertrophie [Bursi et al. 2006, Aronnow 2006, Swedberg et al. 2005].

In der vorliegenden Arbeit werden beide Formen mit ihren teilw. voneinander abweichenden Therapieempfehlungen berücksichtigt.

Herzinsuffizienz bei Neugeborenen und Kindern: Diese Kondition wird *nicht berücksichtigt*, da ein von adulten Formen abweichendes klinisches Bild vorliegt [Braunwald & Grossman 1992] und die Versorgung einer kindlichen Herzinsuffizienz Spezialisten vorbehalten bleiben sollte.

3.1.1.1.5 Stadien der chronischen Herzinsuffizienz

Die am weitesten verbreitete Stadieneinteilung – die Klassifikation der *New York Heart Association* (NYHA) – folgt *funktionellen* Gesichtspunkten. Tabelle 3.1 zeigt diese Stadieneinteilung in ihrer revidierten Form [Hoppe et al. 2005].

Tab. 3.1: Funktionelle Stadieneinteilung der Herzinsuffizienz in der revidierten NYHA-Klassifikation

I.	Herzerkrankung ohne körperliche Limitation. Alltägliche körperliche Belastung verursacht keine inadäquate Erschöpfung, Rhythmusstörungen, Luftnot oder Angina pectoris
II.	Herzerkrankung mit leichter Einschränkung der körperlichen Leistungsfähigkeit. Keine Beschwerden in Ruhe. Alltägliche körperliche Belastung verursacht Erschöpfung, Rhythmusstörungen, Luftnot oder Angina pectoris
III.	Herzerkrankung mit höhergradiger Einschränkung der körperlichen Leistungsfähigkeit bei gewohnter Tätigkeit. Keine Beschwerden in Ruhe. Geringe körperliche Belastung verursacht Erschöpfung, Rhythmusstörungen, Luftnot oder Angina pectoris
IV.	Herzerkrankung mit Beschwerden bei allen körperlichen Aktivitäten und in Ruhe. Bettlägerigkeit

Quelle: Hoppe et al. 2005

fizienz diagnostiziert und damit deutlich mehr als die 34 bis 63% in anderen Untersuchungen an älteren Patienten beobachteten Fälle [Aronnow 2006].

2001 wurde von der *American Heart Association* (AHA) und dem *American College of Cardiology* (ACC) ergänzend eine neue Klassifikation vorgestellt, welche die Entstehung und Progredienz der Erkrankung sowie objektive Untersuchungsbefunde berücksichtigt (vgl. Tabelle 3.2) [Hoppe et al. 2005]. Mit der Absicht, eine Progression bereits in frühen Stadien des *Remodelling* zu verhindern, wurden *Risikopatienten ohne strukturelle Herzerkrankung* in Klasse A aufgenommen [Jessup & Brozena 2003].

Tab. 3.2: *Stadieneinteilung der Herzinsuffizienz nach AHA und ACC*

Stadium A	Hohes Herzinsuffizienzrisiko, da Faktoren vorliegen, die stark mit der Entstehung einer Herzinsuffizienz assoziiert sind; keine strukturelle Herzerkrankung, noch nie Herzinsuffizienzsymptome
Stadium B	Strukturelle Herzerkrankung, die eng mit der Entstehung einer Herzinsuffizienz assoziiert ist, bisher keine Herzinsuffizienzsymptome
Stadium C	Frühere oder derzeitige Herzinsuffizienzsymptome bei struktureller Herzerkrankung
Stadium D	Fortgeschrittene strukturelle Herzerkrankung und schwere Herzinsuffizienzsymptome in Ruhe trotz maximaler medikamentöser Therapie (spezielle Therapie erforderlich, z. B. Herztransplantation, i.v.-Inotropika, <i>assist device</i>)

Quelle: Hoppe et al. 2005

3.1.1.2 Eingrenzung der Zielerkrankung

Eingeschlossen wurde die chronische Herzinsuffizienz (systolisch und diastolisch) inkl. akuter Dekompensationen, *ausgeschlossen* wurden die isolierte Rechtsherzinsuffizienz auf dem Boden eines Cor pulmonale, die akute Herzinsuffizienz und die Herzinsuffizienz im Kindes- und Jugendalter.

3.1.2. Definition des Zielkontextes

Nutzer der Leitlinie: *Primäre Zielgruppe* sind einerseits hausärztlich tätige Ärztinnen und Ärzte, d.h. Allgemeinmediziner, hausärztlich tätige Internisten und praktische Ärzte, *sekundär* richtet sich die Leitlinie an die in Aus- und Weiterbildung tätigen Mediziner sowie an Entscheidungsträger im Gesundheitswesen.

Andererseits sollen durch die Entwicklung einer evidenzbasierten Patienteninformation auch Patienten direkt angesprochen werden, um notwendige Informationen zu vermitteln und damit das Selbstsorgeverhalten der Patienten zu stärken.

Setting: In der Leitlinienentwicklung müssen die Besonderheiten der hausärztlichen Versorgung berücksichtigt werden. Dazu zählen vor allem:

- Die Hausarztpraxis ist im „Niedrig-Prävalenz-Bereich“ angesiedelt und weist daher – im Zusammenhang mit dem Bayes' Theorem – Besonderheiten in Diagnostik und Therapie auf. [Gerlach 2002, Gerlach 2002b]
- Die hausärztliche ist gegenüber der spezialärztlichen Versorgung von Patienten durch ihre besonders hohe Kontinuität und Longitudinalität gekennzeichnet. Darin begründet liegen spezifisch hausärztliche Aufgaben, wie bspw. die Koordinierung verschiedener Behandlungen (und Behandler), regelmäßige Follow-up-Visiten oder etwa die Beratungsaufgaben gegenüber den Patienten und deren Angehörigen. [Haggerty et al. 2003, Reid et al. 2002, Starfield et al. 1976]

Da insbesondere bei herzinsuffizienten Patienten sowohl Hausärzte als auch ambulant und stationär tätige Spezialisten in die Versorgung – regelhaft oder im Rahmen von akuten Verschlechterungen – einbezogen werden, sind in der zu entwickelnden Leitlinie Schnittstellen zur spezialisierten ambulanten und stationären Versorgung zu definieren.

3.1.3 Veränderungsbedarf

Die Notwendigkeit der Entwicklung einer Leitlinie zur chronischen Herzinsuffizienz kann aus der Epidemiologie (Häufigkeitszunahme) und der Versorgungsqualität (Unter- und Fehlversorgung) begründet werden:

3.1.3.1 Epidemiologie

Inzidenz: In der Framingham-Kohorte stieg die Inzidenz altersabhängig von 0,2% bei 45- bis 54-Jährigen bis 4% bei 85- bis 94-Jährigen an, in einer Kohorte aus Minnesota wurden etwa 50% der inzidenten Fälle bei über 80-Jährigen beobachtet [Weir et al. 2006], Daten aus Deutschland fehlen. Im Trend über die letzten 20 Jahre zeigten vergleichende Untersuchungen für Daten von *Kaiser Permanente Northwest* (USA) bei über 70-Jährigen einen Inzidenzanstieg von 14% zwischen den 70er und den 90er Jahren, wenn für Alter und Komorbiditäten kontrolliert wurde [Barker et al. 2006].

Prävalenz: sie steigt altersabhängig ebenfalls an. Bei 50- bis 59-Jährigen wurden in Framingham 0,8% und bei über 80-Jährigen 9,1% beobachtet [Barker et al. 2006]. Für Deutschland liegen Daten zur Prävalenz der linksventrikulären *systolischen* Dysfunktion vor, die auf 2,3% geschätzt wird. In einer bevölkerungsbezogenen Stichprobe 25- bis 75-Jähriger war bei den unter 40-Jährigen einer von 66 Untersuchten davon betroffen, bei den über 60-Jährigen einer von 25. Männer hatten tendenziell häufiger eine linksventrikuläre *systolische* Dysfunktion als Frauen (2,8% vs. 1,9%). [Fischer et al. 2003]

Mortalität: Die Daten aus internationalen epidemiologischen Studien variieren erheblich. In der *Framingham Heart Study* von 2001 betrug die 1-Jahresmortalität der systolischen Herzinsuffizienz 19%, die der diastolischen Herzinsuffizienz 9%. In einer Untersuchung an Postinfarkt-Patienten betrug sie 41 – 47% für die systolische Herzinsuffizienz gegenüber 19 – 22% bei diastolischer Herzinsuffizienz. [Aronnow 2006] Cowie et al. [2000] fanden bei einer populationsbezogenen Untersuchung eine 1-Jahresmortalität von 38% bei inzidenten Fällen. Nach Statistiken des *National Heart, Lung, and Blood Institute* (NHLBI) der USA starben 80% der Männer und 70% der Frauen, die im Alter unter 65 Jahren an Herzinsuffizienz erkrankten, innerhalb von acht Jahren, und die Patienten mit Herzinsuffizienz starben 6 bis 9 mal häufiger an plötzlichem Herztod als die Allgemeinbevölkerung [ASA & AHA 2005].

Longitudinale Vergleichsuntersuchungen weisen auf einen günstigen Trend für die Sterblichkeit an Herzinsuffizienz hin: Barker et al. [2006] zeigten an einer retrospektiven Kohorte ein für Männer signifikant verbessertes Überleben im Zeitraum 1990-1994 im Vergleich zu 1970-1974, dieser Effekt war bei Frauen allerdings deutlich geringer. Ähnliche Trenddaten liegen aus den Framingham- und Rochester-Kohorten vor [Barker et al. 2006].

Aus Deutschland können nur Daten aus der allgemeinen Todesursachenstatistik herangezogen werden: die Herzinsuffizienz ist hier die dritthäufigste Todesursache mit 5,8% aller Gestorbenen, wobei diese Ursache bei Frauen (7,4%) häufiger war als bei Männern (3,9%) [Statistisches Bundesamt 2007]. Für Deutschland stiegen zwischen 2001 und 2003 die Sterbeziffern allerdings um 4% [Bruckenberger 2004].

Individuelle Prognose und Überlebenswahrscheinlichkeit: Epidemiologische Daten zeigen für die Herzinsuffizienz eine dem Kolonkarzinom vergleichbar ungünstige Prognose [Quinn et al. 2001 in NICE 2003]. Vorhersagemöglichkeiten zu individuellem Verlauf und Prognose der Herzinsuffizienz sind jedoch deutlich geringer als bei anderen Erkrankungen [Cowie 2003], da etwa die Hälfte der Sterblichkeit durch einen plötzlichen Herztod und nicht durch ein allmähliches progredientes Pumpversagen verursacht wird [Cleland et al. 2002].

In zahlreichen Studien wurden Parameter für eine reduzierte Überlebenswahrscheinlichkeit ermittelt (z.B. männliches Geschlecht, Alter, höheres NYHA-Stadium, KHK als Ursache, begleitende Komorbiditäten – v.a. Diabetes mellitus, Niereninsuffizienz, Depression, Anämie u.v.a.m.), die Studienlage ist jedoch inkonsistent [Cowburn et al. 1998, Jünger et al. 2005], und derzeit existiert kein zuverlässiger und gleichzeitig praktikabler Risikoscore²⁰.

²⁰ Risikovorhersage-Scores sind in der hausärztlichen Versorgung für das Patientengespräch sehr wertvoll. Für die Herzinsuffizienz existieren jedoch nur unpraktikable [z.B. Aaronson et al. 1997] oder nicht ausreichend validierte Scores [z.B. Bouvy et al. 2003].

Prognose der Krankheitslast: Projektionen ergaben eine erhebliche Zunahme der Krankheitslast für die nächsten Jahrzehnte. Hauptverantwortlich werden dafür allgemeine demographische Entwicklungen und verbesserte Überlebenschancen bei der Koronaren Herzkrankheit (KHK), eine der häufigsten Ursachen der Herzinsuffizienz, (sowie der Herzinsuffizienz selbst) gemacht [Stewart et al. 2003].

Morbidität: In 2004 wurden 260.801 Fälle an 3,2 Mio. Pflgetagen vollstationär behandelt [Statistisches Bundesamt 2006], wobei ein steigender Trend beschrieben wurde: zwischen 2001 und 2003 stiegen die Morbiditätsziffern um 3,3% [Bruckenberg 2004].

Kosten: Die Herzinsuffizienz verursachte im Jahre 2002 Krankheitskosten von geschätzt etwa 2,74 Mrd. € [Bruckenberg 2004], die ganz wesentlich durch stationäre Behandlungskosten verursacht wurden.

3.1.4 Versorgungssituation

Gegenwärtig wird die Mehrzahl von Patienten mit Herzinsuffizienz hausärztlich versorgt (hochgerechnet etwa 1,2 bis 1,5 Mio. Patienten in Deutschland mit durchschnittlich 7,5 Kontakten im Quartal – bei fachärztlich tätigen Internisten waren es hochgerechnet 110 – 114.000 Patienten mit durchschnittlich 4,3 Kontakten) [ZI 2006].

Die derzeitige Versorgungssituation ist verbesserungsbedürftig: nach den Ergebnissen von zwei Surveys wurden nur etwa 17% der Herzinsuffizienz-Patienten in der Primärversorgung leitliniengerecht versorgt [Cleland et al. 2002 (IMPROVEMENT), Komajda et al. 2003 (EURO-HF)]. In einer weiteren Untersuchung wurde gezeigt, dass gerade Patienten mit dem höchsten Sterblichkeitsrisiko die geringste Wahrscheinlichkeit haben, mortalitätssenkende Pharmaka zu erhalten [Lee, Tu et al. 2005].

Als vordringliche Versorgungsdefizite wurden die unzureichende Evidenzbasierung von Diagnostik und Therapie, eine unkoordinierte und fragmentierte Versorgung, ein inadäquates Monitoring sowie die fehlende Patientenaktivierung identifiziert [Gensichen et al. 2006].

3.1.4 Festlegung der Ergebnisparameter

Aus den Besonderheiten der Zielerkrankung, ihrer Epidemiologie und Versorgungssituation lassen sich folgende Outcomeparameter der Evidenzbewertung für die Leitlinie begründen:

Mortalität (bzw. Letalität), Krankheitsprogression, Morbidität (Hospitalisierungshäufigkeit und -dauer), allgemeine und krankheitsspezifische Lebensqualität, Symptomverbesserung sowie Belastungstoleranz.

Weitere Ergebnisparameter – wie z.B. Parameter der diagnostischen Wertigkeit in der Bewertung diagnostischer Tests – müssen je nach spezifischer Fragestellung präzisiert werden.

Nach der Festlegung wichtiger Prämissen für die Leitlinie zur chronischen Herzinsuffizienz wird nun die Vorgehensweise des SLR im Einzelnen dargelegt.

3.2 Verfahrensschritte im systematischen Leitlinienreview

Das methodische Vorgehen beim systematischen Leitlinienreview kann in folgende neun Einzelschritte gegliedert werden, die in den anschließenden Unterkapiteln detailliert beschrieben werden:

1. Systematische Recherche nach existierenden Leitlinien
2. Auswahl von Leitlinien gemäß prädefinierter Ein- und Ausschlusskriterien
3. Bewertung der formalen methodischen Qualität der berücksichtigten Leitlinien
4. Systematische Framework-Entwicklung
5. Datenextraktion aus den berücksichtigten Leitlinien in Evidenztabelle mit standardisiertem Berichtsformat
6. Konsistenzanalyse der Empfehlungen eingeschlossener Leitlinien unter Berücksichtigung berichteter Evidenz
7. Informationssynthese zum vorläufigen Erstentwurf
8. Analyse der Belegtheit der Empfehlungen (Validierung, Rekurs auf die Primärliteratur)
9. Überarbeitung des Erstentwurfs (zweite Informationssynthese) und Identifikation weiteren Forschungsbedarfs

Eine systematische Recherche nach Leitlinien und deren Auswahl nach prädefinierten Kriterien zielt auf einen unverzerrten Einschluss von allen relevanten Leitlinien in die vergleichende Analyse. Die Bewertung der formalen Qualität soll sicherstellen, dass sich der SLR auf eine ausreichende Zahl qualitativ hochwertiger Leitlinien stützen kann. Die systematische Framework-Entwicklung soll zunächst alle relevanten klinischen Fragestellungen erfassen und deren Auswahl für die Leitlinie nachvollziehbar abbilden.

Das wesentliche Ziel der Daten-(Evidenz-)Extraktion, der Konsistenzanalyse und der anschließenden Validierung ist es zu überprüfen, ob die Empfehlungen unterschiedlicher analysierter Leitlinien auf einer zureichenden und homogenen

Evidenzgrundlage aus Studien zu einem Problem beruhen. Unzureichende oder inhomogene Evidenz kann hierbei auf unzureichende Recherche in den Quelleitlinien, auf ein Forschungsproblem und / oder auch auf implizite Normativität von Empfehlungen hinweisen. Im Rahmen der Informationssynthese bzw. Abfassung des Erstentwurfs sind noch verstärkt die Setting-Bedingungen der Zieleitlinie (also der eigene Entwicklungskontext) sowie ggf. abweichende Setting-Einflüsse in den Quelleitlinien zu beachten.

Mit der Vorlage des Erstentwurfs ist zwar der systematische Leitlinienreview, aber natürlich nicht der Entwicklungsprozess der Zieleitlinie abgeschlossen. Weitere von den aktuellen Bewertungsmaßstäben methodisch hochwertiger Leitlinien (AGREE) geforderte Entwicklungsschritte bestehen in einem formalen Konsensprozess, im mehrstufigen externen und internen *Peer Review*-Verfahren sowie in Panel- und Praxistest, durch die jeweils multiprofessionelle, spezifisch allgemeinmedizinische und Patientenperspektiven einbezogen werden. Diese Schritte wurden auch von der hier beschriebenen DEGAM-Leitlinie Herzinsuffizienz erfolgreich durchlaufen, ihre Darstellung liegt allerdings außerhalb des Rahmens der vorliegenden Arbeit zum SLR.

Während der weiteren Entwicklungsstufen der Leitlinie bleiben die Bezüge zum kontextspezifischen Wertesystem der Leitlinienentwicklung ein fortdauerndes Problem, insbesondere dann, wenn im Konsensbildungsprozess auf unzureichender Evidenzgrundlage praktisch erforderliche Entscheidungshilfen und praktikable Handlungsempfehlungen entwickelt werden müssen. Durch den vorgängigen SLR soll allerdings sichergestellt werden, dass jeweils die 'beste verfügbare Evidenz' zugrundegelegt werden kann, dass bei Inkonsistenzen in der verfügbaren Evidenz eine offene Debatte der Empfehlungen stattfindet, und keine 'bias' aus anderen Leitlinien unreflektiert in die Zieleitlinie übernommen werden.

3.2.1 Systematische Leitlinienrecherche

Mit dem Ziel, alle existierenden Leitlinien zur chronischen Herzinsuffizienz zu identifizieren, wurde nach vorausgegangenen orientierenden Handrecherchen im März 2004 eine systematische Recherche durchgeführt. In den elektronischen Datenbanken MEDLINE, The Cochrane Library, DARE und HSTAT wurde eine Kombination aus Freitextbegriffen und kontrollierten Schlagworten verwendet (vollständige Suchstrategie siehe Anhang). Diese Recherche wurde durch ausführliche Handsuchen auf den Internetseiten von führenden Anbietern von Leitliniensammlungen (z.B. Ärztliches Zentrum für Qualität in der Medizin (ÄZQ), *National Guideline Clearinghouse* (NGC) USA), medizinischen Fachgesellschaften und deren Dachorganisationen (z.B. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF), Deutsche Gesellschaft für Kardiologie (DGK), *American College of Cardiology* (ACC)), anderen Leitlinienanbietern (z.B. *Veterans Health Administration* (VHA) USA) und in den

Referenzlisten der bereits erhaltenen Leitlinien ergänzt (vollständige Liste besuchter Internetadressen siehe Anhang, Tab. A.1 und A.2). Die erhaltenen Zitate wurden auf Relevanz geprüft, Dubletten wurden entfernt, und die Dokumente wurden im Volltext bestellt.

3.2.2 **Ein- und Ausschlusskriterien**

Um für die weitere Evaluation

- evidenzbasierte,
- hinsichtlich Zielerkrankung und -kontext für die hausärztliche Versorgung relevante,
- untereinander vergleichbare,
- sprachlich auswertbare und
- aktuelle

Leitlinien auszuwählen, wurden folgende **Einschlusskriterien** *a priori* festgelegt:

1. Die Leitlinie sollte Empfehlungen zur Primärversorgung erwachsener Patienten mit chronischer Herzinsuffizienz enthalten.
2. Die Leitlinie sollte evidenzbasiert sein. Dazu wurde gefordert, dass bei der Mehrzahl der Empfehlungen Evidenzlevel und/oder Stärke der Empfehlungen angegeben waren und dass diesen klare Quellenangaben zugrundeliegender Evidenz zugeordnet waren.
3. Die Leitlinie sollte im Jahr 2000 oder später publiziert oder aktualisiert worden sein.
4. Die Leitlinie sollte in deutscher oder englischer Sprache publiziert worden sein.

Eine Leitlinie wurde **ausgeschlossen**, wenn sie die Einschlusskriterien nicht erfüllte oder

1. wenn sie sich ausschließlich auf eine *isolierte Rechtsherzinsuffizienz* bei Cor pulmonale bezog oder
2. wenn *keine detaillierten Angaben zum methodischen Vorgehen* bei der Leitlinienentwicklung vorlagen. Bei etablierten Leitlinienprogrammen wurden mindestens generelle Angaben zur Vorgehensweise der leitlinienproduzierenden Organisation gefordert.

Wenn die Leitlinie lediglich eine *spezialisierte Fragestellung* behandelte (z.B. Therapie der Herzinsuffizienz mit implantierbaren Schrittmacher-Defibrillator-Geräten), wurde sie von der vergleichenden Analyse ausgeschlossen und als Hintergrundinformation verwendet.

Anhand dieser Kriterien wurden die in der Recherche identifizierten Dokumente durch zwei Reviewer²¹ *unabhängig voneinander* auf ihre Einschlussfähigkeit geprüft. Bei unterschiedlichen Prüfergebnissen wurde ein Konsensusurteil gebildet. Die Untersucher-Übereinstimmung wurde mittels Kappa-Statistik [Cohen 1960, McGinn et al. 2004] ermittelt.

Ausgeschlossene Leitliniendokumente werden unter Nennung des Ausschlussgrundes aufgelistet (siehe Anhang Tab. A.4).

3.2.3 Bewertung der formalen methodischen Qualität

Die formale methodische Qualität der berücksichtigten Leitlinien wurde anhand des international validierten AGREE-Instruments [AGREE 2002, AGREE 2003, AGREE 2003a] bewertet, um sicherzustellen, dass in den Review Leitlinien von hoher methodischer Qualität eingeschlossen wurden. AGREE enthält 23 einzelne Items, die in sechs 'Domänen' separate Dimensionen der Leitlinienqualität abbilden (siehe Anhang, Tab. A.3).

Jedes Item wurde mit Hilfe einer 4-stufigen-Likertskala eingestuft, die von 4 = "trifft uneingeschränkt zu" bis 1 = "trifft überhaupt nicht zu" über zwei mittlere Werte: 3 = "trifft zu" und 2 = "trifft nicht zu" reicht. Die Skala misst das Ausmaß, in dem ein Kriterium erfüllt wurde. Eine kommentierte Anleitung zum Rating ist Bestandteil des AGREE-Trainingsmanuals [AGREE 2003].

Für jede Domäne wurde ein Summenscore gebildet. Eine Standardisierung erfolgte durch Darstellung der erreichten Gesamtpunktzahl als prozentualer Anteil der maximal möglichen Punktzahl dieser Domäne (Berechnungsformel: s. Abb. 3.2) [AGREE 2002, AGREE 2003]. Die Ergebnisse der Bewertung werden in einer Gesamtübersicht tabellarisch präsentiert.

Abb. 3.2: Standardisierter Domänen-Summenscore, Berechnungsformel

$$\text{Standardisierter Domänenwert} = \frac{(\text{erreichte Punktzahl} - \text{minimal mögliche Punktzahl})}{(\text{maximal mögliche Punktzahl} - \text{minimal mögliche Punktzahl})}$$

Quelle: AWMF/ÄZQ 2005

²¹ Der Erstreview erfolgte durch die Verfasserin der vorliegenden Arbeit. Der unabhängige Zweitreview wurde von Dr. med. Dipl.-Päd. Jochen Gensichen, MPH, wissenschaftlicher Assistent und Leiter des Arbeitsbereichs Chronic Care und Versorgungsforschung des Instituts für Allgemeinmedizin der Johann Wolfgang Goethe-Universität, durchgeführt.

3.2.4 Framework-Entwicklung

Die berücksichtigten Leitlinien wurden einer vergleichenden Inhaltsanalyse unterzogen. Die von den jeweiligen Leitlinienentwicklergruppen als relevant erachteten Themen und klinischen Fragestellungen wurden identifiziert. Eine daraus resultierende Liste von Fragestellungen, die von der zu entwickelnden Leitlinie beantwortet werden sollten, wurde in der Arbeitsgruppe²² hinsichtlich ihrer Relevanz für den deutschen hausärztlichen Versorgungskontext abgestimmt. Die als relevant eingestuften Themen und Fragestellungen bildeten das Raster für die weitere Datenextraktion.

3.2.5 Datenextraktion

Zunächst wurde für jede berücksichtigte Leitlinie eine Kurzdarstellung in standardisierter Berichtsform erstellt. Gesamtziel der Leitlinie und Anwenderzielgruppe, konkrete Fragestellungen, Zielpopulation, Methodik inkl. Entwicklergruppe, Evidenzgewinnung und verwendete Systeme für Evidenzlevel und Stärken der Empfehlung (Gradings), Konsentierung, externer Begutachtungsprozess, Praxiserprobung, Gültigkeit und Zusatztools wurden dokumentiert.

Weiterhin wurde zu jeder klinischen Fragestellung des Rasters eine Evidenztabelle erstellt. Darin wurden die in den Leitlinien zur jeweiligen Fragestellung enthaltenen Empfehlungen inkl. Evidenzlevel und Grading und die berichteten Ergebnisse der Evidenzbewertung sowie die dazu zitierten Quellen (klinische Primärstudien und Übersichtsarbeiten) gesondert erfasst und dokumentiert. Evidenzlevel und Stärkegrad der Empfehlung wurden in dem jeweils verwendeten System notiert, da die Systeme untereinander nicht kompatibel waren. Sofern in den Leitlinien quantitative Risiko-Nutzen-Bewertungen (z.B. als *Number Needed to Treat* – NNT²³ – oder *Number Needed to Harm* – NNH²⁴ – oder NNT:NNH-Ratio²⁵) angegeben wurden, wurden diese gleichfalls vermerkt.

²² Neben dem bereits erwähnten Dr. med. Dipl.-Päd. Jochen Gensichen, MPH, gehörten dieser Arbeitsgruppe Dipl. Soz. Martin Beyer und Prof. Dr. med. Ferdinand M. Gerlach, MPH, an.

²³ Die NNT ist eine quantitative Risiko-Nutzen-Angabe über die Effektivität einer Therapie. Sie ist behandlungsspezifisch und errechnet sich aus dem Kehrwert der absoluten Risikoreduktion. NNT=1 beschreibt den Idealzustand, dass jeder Patient der Interventionsgruppe das Behandlungsziel erreicht, jedoch kein Patient in der Kontrollgruppe. [Moore & McQuay 2003]

²⁴ Die NNH ist eine quantitative Risiko-Nutzen-Angabe, bei der die Risikobewertung, also der Sicherheitsaspekt im Vordergrund steht. Sie gibt an, wie viele Patienten behandelt werden müssen, um ein unerwünschtes Ereignis zu beobachten. [Watt et al. 2003]

²⁵ Die NNH:NNT-Ratio ist das Verhältnis aus beiden quantitativen Risiko-Nutzen-Angaben, eine NNH:NNT-Ratio kleiner 1 würde bedeuten, dass die potentiellen Risiken den Nutzen überwiegen. [Watt et al. 2003]

3.2.6 Konsistenzanalyse

Für die Konsistenzanalyse wurden die in Evidenztabelle extrahierten Empfehlungen zu einzelnen klinischen Fragestellungen systematisch miteinander verglichen, um *Konsistenzen* und *Inkonsistenzen* darzustellen. Konnte in den Empfehlungen eine inhaltliche Übereinstimmung für eine Formulierung gefunden werden, wurden die Empfehlungen hinsichtlich dieser Formulierung als konsistent gewertet²⁶. Fehlte eine solche inhaltliche Übereinstimmung der Empfehlungen zu einer klinischen Fragestellung, lag Inkonsistenz vor.

Konsistenzen wie Inkonsistenzen wurden weiter differenziert: zu jeder *konsistenten* Empfehlung wurde die *Härte der zugrundeliegenden Evidenz* herangezogen (wie von den Autoren berichtet), um daraus unterschiedliche *Typen von Konsistenzen* zu bilden. *Inkonsistenzen* wurden danach unterschieden, ob vollständige Nicht-Übereinstimmung zwischen den Empfehlungen bestand, oder ob eine Minderheit von Leitlinien Empfehlungen enthielt, die sich von einer Mehrheit der anderen Leitlinien unterschied. Daraus wurde das folgende Klassifikationsschema hergeleitet:

Tab. 3.3: Klassifikation von Konsistenzen und Inkonsistenzen

Typ	Definition	Handlungsoptionen
Konsistenzen		
1	Empfehlungen sind konsistent in Inhalt, Evidenzlevel und Grading <u>und</u> sie basieren auf einer Vielzahl von Studienergebnissen eines hohen Evidenzlevels (z.B. zahlreiche klinische Primär- oder Sekundärstudien, wie z.B. systematische Review(s) mit/ohne darin enthaltener Metaanalyse), deren Ergebnisse eine hohe interne wie externe Validität aufweisen	Analyse der Belegtheit mindestens anhand einer Evaluation der höchstwertigen zitierten Evidenzquellen und ggf. Update-Recherchen
2	Empfehlungen sind konsistent in Inhalt, Evidenzlevel und Grading <u>und</u> sie basieren auf einzelnen oder wenigen Studien eines hohen Evidenzlevels (z.B. einzelne klinische Primärstudie), deren Ergebnisse eine hohe interne wie externe Validität aufweisen	Analyse der Belegtheit anhand der Evaluation der zitierten Evidenzquelle(n); Weitere Recherchen, insbesondere zu Sicherheitsaspekten der empfohlenen Technologie und ggf. Update-Recherchen

²⁶ Zur Erläuterung: wir erwarteten und fanden bei dem Leitlinienvergleich folgende Standardsituationen der Konsistenz: Typ 1: die Empfehlungen werden mit erstklassiger Evidenz (nach z.B. dem geläufigen Oxford-Schema Ia) begründet. Typ 2: die Empfehlungen stützen sich auf erstklassige, jedoch in der Belastbarkeit hinterfragbare (z.B. Oxford Ib) Evidenz. Typ 3: - ein typisches Leitlinienentwicklungsproblem zu hausärztlichen Vorgehensweisen – die Empfehlungen beruhen nur auf Studien weniger hochwertiger (Oxford II – III) Evidenz oder auf Expertenkonsens. Typ 4: die Evidenz wurde deutlich unterschiedlich graduiert.

(Fortsetzung Tabelle 3.3)

Typ	Definition	Handlungsoptionen
Konsistenzen		
3	Empfehlungen sind konsistent in Inhalt, Evidenzlevel und Grading <u>und</u> sie basieren auf Studien eines niedrigen Evidenzlevels (z.B. klinische Studien mit Design-bezogenen Bias oder anderen methodischen Mängeln, die ergebnisverzerrende Effekte wahrscheinlich machen) oder auf Expertenkonsens	Weitere Evidenzgewinnung ist notwendig (systematische oder umfassende Recherchen)
4	Empfehlungen sind inhaltlich konsistent, jedoch weichen Evidenzlevel u./o. Grading voneinander ab oder sind konfliktär	Analyse der Belegtheit anhand der Evaluation der zitierten Evidenzquelle(n); und ggf. Update-Recherchen
Inkonsistenzen		
A	Empfehlungen lassen keine inhaltliche Übereinstimmung erkennen, es kann kein mehrheitlicher Trend ermittelt werden.	Weitere Evidenzgewinnung ist notwendig (systematische oder umfassende Recherchen erforderlich)
B	Empfehlungen sind inhaltlich bei einer Mehrheit von eingeschlossenen Leitlinien konsistent, eine einzelne Leitlinie oder eine Minderheit von Leitlinien weicht in ihren Empfehlungen von diesem Trend ab oder deren Empfehlungen sind sogar konfliktär zur Mehrheit	Analyse der Belegtheit anhand der Evaluation der zitierten Evidenzquelle(n) zur Lösung des Konflikts und Identifikation weiteren Forschungsbedarfs und ggf. Update-Recherchen

Quelle: Eigene Darstellung

Mit dem Ziel, bestehende Ressourcen der Leitlinienentwicklung auf Forschungsfragen mit hoher Relevanz zu konzentrieren, wurden Handlungsoptionen für die Priorisierung der weiteren Evidenzgewinnung vorgeschlagen, die aus dem Klassifikationssystem von (In-)Konsistenzen abgeleitet sind:

Vorliegende *Inkonsistenzen* und / oder mangelnde Härte des Evidenzbezugs machen eine umfassende (bzw. systematische) Gewinnung neuer Evidenz erforderlich, während konsistente Empfehlungen auf der Grundlage von ‚harter‘ Evidenz – wie z.B. eine Vielzahl hochwertiger klinischer Studien – nach der ‚Validierung‘ (siehe 3.2.8) eine Beschränkung auf Update-Recherchen ermöglichen, d.h. Recherchen, die Publikationen nach Rechereschluss der eingeschlossenen Leitlinien erfassen.

3.2.7 Informationssynthese

In den vorläufigen Erstentwurf der zu entwickelnden Leitlinie wurden die als relevant identifizierten klinischen Fragen aufgenommen. Zu den Fragestellungen wurden Empfehlungen formuliert, die aus den jeweiligen Ergebnissen der Konsistenzanalyse zusammengefasst wurden. Die Empfehlungen wurden mit vorläufigen Evidenzleveln und Empfehlungsstärken verknüpft. Da vorläufige Evidenzlevel und Empfehlungsstärken ausschließlich auf der in Leitlinien berichteten (bewerteten) Evidenz beruhten, wurde ein vereinfachtes System gewählt (siehe Tab. 3.4). Die in den Leitlinien berichtete Evidenz wurde zusammengefasst und diskutiert.

Tab. 3.4: Vorläufige Evidenzlevel und Empfehlungsstärken im Erstentwurf

Evidenzlevel: IA _{X1} – Mehrere klinische Primärstudien oder mindestens eine Sekundärstudie (z.B. systematische(r) Review(s) mit/ohne darin enthaltene Metaanalysen), deren Ergebnisse von den anderen Leitlinienentwicklergruppen als valide eingestuft wurden IB _{X1} – Mindestens eine klinische Primärstudie (in der Regel ein RCT), deren Ergebnisse von den anderen Leitlinienentwicklergruppen als valide eingestuft wurden IV _{X1} – Nicht-I-Evidenz oder Expertenkonsens
Empfehlungsstärken: A _{X1} – basiert auf wissenschaftlichen Studien hoher Qualität C _{X1} – basiert auf Nicht-I-Evidenz oder Expertenkonsens

(_{X1}) – Tiefgestelltes „X“ für „Entwurf“, die folgende Ziffer bezeichnet das Stadium; Quelle: Eigene Darstellung

3.2.8 Analyse der Belegtheit (Validierung)

Ziel dieser Analyse war es, die Validität der aus den berücksichtigten Leitlinien extrahierten Evidenz anhand von Originalquellen (Primärliteratur) zu überprüfen und kritisch zu hinterfragen, ob die Evidenz eine ausreichende Grundlage zur Ableitung der Empfehlungen bereitstellte. Im positiven Fall wurden Evidenz und Empfehlung als zueinander kongruent²⁷ vermerkt. Dazu wurden vorrangig die höchstwertigen Primärquellen analysiert, bei deren Fehlen auch schrittweise Quellen von niedrigerem methodischem Wert.

²⁷ Der hier verwendete Begriff „Kongruenz“ ist als formale inhaltliche Übereinstimmung aufzufassen (eine Deckungsgleichheit ist nicht möglich, wie unter 1.2.5.5 dargestellt wurde).

3.2.8.1 Recherchen

Aus den Evidenztabelle wurden die bibliographischen Angaben der zitierten Evidenzquellen zusammengestellt und um Dubletten bereinigt. Diese wurden nach orientierender Prüfung von Titel und Abstract in Primär- und Sekundärstudien (Übersichtsarbeiten) unterteilt.

Als zusätzliche Evidenzquelle wurden *HTA-Berichte* in der initial durchgeführten systematischen (Leitlinien-)Recherche identifiziert, da diese hochwertige Evidenzquelle von keiner der berücksichtigten Leitlinien verwendet wurde.

3.2.8.2 Bewertung von Primärquellen

3.2.8.2.1 Reevaluation der in Leitlinien zitierten Sekundärstudien

Die von den berücksichtigten Leitlinien zitierten Sekundärstudien (systematische Reviews und Metaanalysen) wurden im Volltext bestellt und auf ihre Relevanz (Fragestellung, Outcomes, Studienpopulation, etc.) geprüft. Bei Vorliegen mehrerer Sekundärstudien wurden diejenigen, die nur über *Surrogatparameter* berichteten – sofern *klinische* Ergebnisparameter von Interesse aus anderen Arbeiten vorlagen – oder deren Informationsgehalt gegenüber neueren Arbeiten deutlich veraltet war, nicht erneut bewertet.

Die darin enthaltenen Daten wurden in weitere Evidenztabelle (Evidenztabelle der Sekundärstudien) extrahiert. Zur Einschätzung der internen und externen Validität wurden die Dokumente in einem *critical appraisal* auf methodische Mängel untersucht, die ergebnisverzerrende Effekte wahrscheinlich machen [z.B. Egger et al. 2001].

3.2.8.2.2 Evaluation neu identifizierter HTA-Berichte

Die relevanten HTA-Berichte wurden im Volltext bestellt. Ihre methodische Qualität wurde mit Hilfe von nicht-scorebildenden Instrumenten, die von der *German Scientific Working Group on Technology Assessment for Health Care* entwickelt wurden, bewertet [Perleth 2001, Siebert et al. 2003] und in Form von Checklisten präsentiert. Für jeden HTA-Report wurde eine Beschreibung in standardisierter Berichtsform angefertigt, in der Dokumenttyp und Bezugsrahmen, konkrete Fragestellung, Zielpopulation, Methode, Ergebnisse, Diskussion, Schlussfolgerungen der Autoren und eine abschließende Bewertung erfasst wurden.

3.2.8.2.3 *Reevaluation der in Leitlinien zitierten Primärstudien*

Wurde in den zitierten Sekundärstudien und neu identifizierten HTA-Berichten hinreichende und konsistente Evidenz gefunden, welche die Empfehlungen der zu entwickelnden Leitlinie unterstützte, wurden keine Primärstudien reevaluiert. Fehlten Belege für die Empfehlungen, wurde ggf. nach weiterer Evidenz in den zitierten Primärstudien geforscht. Dazu wurden die Daten wiederum in Evidenztabellen (Evidenztabellen der Primärstudien) übertragen, und die Studien wurden auf potentiell ergebnisverzerrende Effekte untersucht. Wurden auch in diesem Schritt keine ausreichenden Belege identifiziert, wurde Forschungsbedarf zur Gewinnung weiterer Evidenz formuliert.

3.2.9 *Überarbeitung des Erstentwurfs und Identifikation weiteren Forschungsbedarfs*

Die Ergebnisse der Validierung inkl. der neu evaluierten HTA-Berichte wurden als zusammenfassende Evidenzbewertung in den bestehenden Erstentwurf eingefügt und diskutiert.

Aus den Ergebnissen der Konsistenzanalyse und der Validierung wurde weiterer Forschungsbedarf formuliert (Tab. 3.5, vgl. auch Tabelle 3.3):

Tab. 3.5: Strategie zur Gewinnung weiterer Evidenz

Update-Recherchen zur Aktualisierung des Evidenzgehaltes: <ul style="list-style-type: none">- Prinzipiell für alle Empfehlungen
Systematische oder umfassende Recherchen: <ul style="list-style-type: none">- Konsistente Empfehlungen vom Typ (3)- Inkonsistente Empfehlungen vom Typ (A)
Fokussierte Recherchen, insbesondere zu Sicherheitsaspekten: <ul style="list-style-type: none">- Konsistente Empfehlungen vom Typ (2)
Differenziertes Vorgehen in Abhängigkeit von den Ergebnissen der Validierung: <ul style="list-style-type: none">- Konsistente Empfehlungen vom Typ (4)- Inkonsistente Empfehlungen vom Typ (B)- Empfehlungen vom Typ (1), für die in der Validierung keine ausreichenden Belege in den zitierten Evidenzquellen gefunden wurden

Quelle: eigene Darstellung

Mit diesem Vorgehen wurden differenzierte klinische Forschungsfragen identifiziert, die einem formalen Priorisierungsverfahren unterzogen werden konnten, um sie an die Ressourcen der zu entwickelnden Leitlinie anzupassen (außerhalb der vorliegenden Arbeit).

4 Ergebnisse

Dieses Kapitel folgt im Aufbau den Verfahrensschritten des systematischen Leitlinienreviews. Da eine vollständige Ergebnispräsentation den vorgegebenen Rahmen der Arbeit sprengen würde, wurde hier eine Auswahl getroffen. Ergänzende Ergebnisse werden im Anhang dargestellt (diesbezügliche Verweise finden sich im Text). Darüber hinaus können die vollständigen Ergebnisse – insbesondere zur Datenextraktion, wie Kurzberichte, Checklisten und Evidenztabelle – im Methodenreport zur Leitlinie²⁸ eingesehen werden.

Aufgrund der eingangs dargelegten Problematik der Normativität in Leitlinien wird einer ausführlichen Ergebnismitteilung zu Konsistenzanalyse und Validierung besondere Bedeutung für die vorgelegte Arbeit beigemessen. Da auch hier eine Beschränkung erforderlich wurde, wurden die Ergebnisse, bezogen auf einzelne (In-)Konsistenztypen, exemplarisch dargelegt.

4.1 Ergebnisse der systematischen Literaturrecherche

Abb. 4.1 auf der folgenden Seite zeigt die Ergebnisse der systematischen Leitlinienrecherche schematisch als Flussdiagramm. Darin sind Datenquellen, Anzahl identifizierter Zitate und das schrittweise Vorgehen bis zum Einschluss der Leitliniendokumente wiedergegeben.

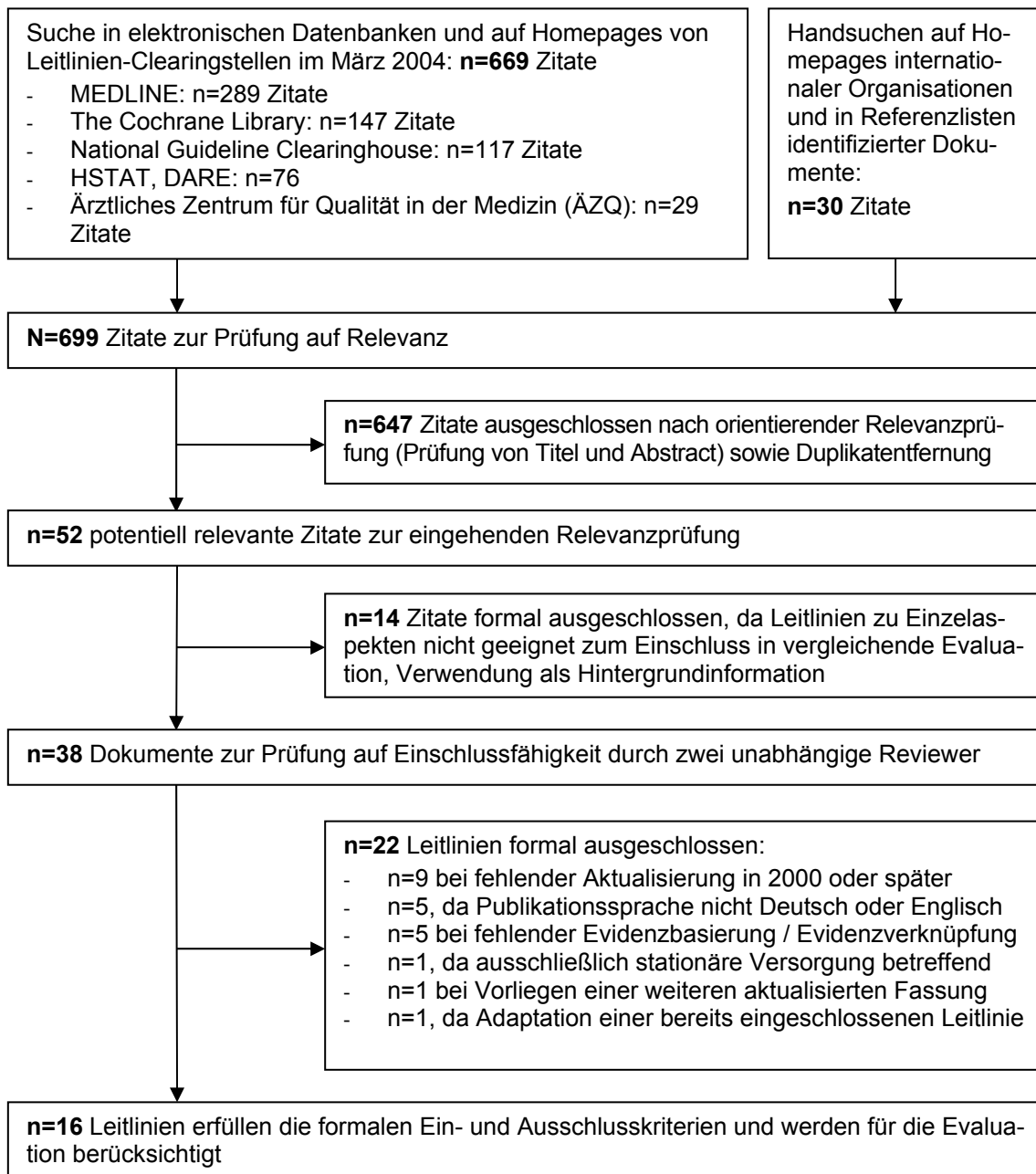
4.2 Prüfung der Einschlussfähigkeit

4.2.1 Berücksichtigte Leitlinien

Von den 52 im Volltext bestellten Leitliniendokumenten erfüllten 16 umfassende Leitlinien die Ein- und Ausschlusskriterien: vier Leitlinien [AKDAE 2001, DGK 2001, LLGH 2003, UWH 2001] und ein leitlinienähnliches Dokument (evidenzbasierte Entscheidungsgrundlage) [DieM 2003/2004] aus Deutschland, vier aus den USA [ACC/AHA 2001, DVA & VHA 2002, ICSI 2003, UM 2001], zwei aus Kanada [CCS 2002/3, OPOT 2000], zwei aus Australien bzw. Neuseeland [NHF/Austr & CSANZ 2002, NHF/NZ 2001] und je eine aus Finnland [Duodecim 2004] und Großbritannien [NICE 2003]. Eine weitere Leitlinie war für den gesamten europäischen Raum konzipiert [ESC 2002/2001]. Eine Übersicht dieser Leitlinien zeigt Tabelle 4.1. auf der folgenden Seite.

²⁸ Der Methodenreport zur Leitlinie kann unter der folgenden Adresse heruntergeladen werden: http://www.allgemeinmedizin.uni-frankfurt.de/forschung/herzinsuffizienz_internet.html

Abb. 4.1: Ergebnisse von systematischer Leitlinienrecherche und Einschlussprüfung



Tab. 4.1: Berücksichtigte Leitlinien: Empfehlungsbereich, Versorgungskontext

Kurzbezeichnung	Herkunftsland	Organisation	Versorgungskontext	Empfehlungsbereich*
ACC/AHA 2001	USA	American College of Cardiology / American Heart Association	Primär- und Sekundärversorgung	D, T
AKDAE 2001	Deutschland	Arzneimittelkommission der deutschen Ärzteschaft	Primärversorgung (sekundär auch stationäre Versorgung)	P

(Fortsetzung Tabelle 4.1)

Kurzbezeichnung	Herkunftsland	Organisation	Versorgungskontext	Empfehlungsbereich*
CCS 2001 / CCS 2002/3	Kanada	Canadian Cardiovascular Society	Primär- und Sekundärversorgung	(D), T
DGK 2001	Deutschland	Deutsche Gesellschaft für Kardiologie	Ohne nähere Angaben	T, P!
DieM 2003/2004	Deutschland	Institut für evidenzbasierte Medizin	Primärversorgung, Entscheidungsunterstützung DMP	T
Duodecim 2004	Finnland	The Finnish Medical Society Duodecim	Primärversorgung	D, T
DVA & VHA 2002	USA	Department of Veterans Affairs & Veterans Health Administration	Primär- und Sekundärversorgung bei DVA & VHA [‡]	P
ESC 2002/2001°	(Europa)	European Society of Cardiology	Ohne nähere Angaben	D, T
ICSI 2003	USA	Institute for Clinical System Improvement	Primär- und Sekundärversorgung bei Health Plans [†]	D, T
LLGH 2003	Deutschland	Leitlinien-Gruppe Hessen	Primärversorgung	(D), T
NHF/Austr. & CSANZ 2002	Australien / Neuseeland	National Heart Foundation of Australia & Cardiac Society of Australia and New Zealand	Ohne nähere Angaben	D, T
NHF/NZ 2001	Neuseeland	The National Heart Foundation of New Zealand	Ohne nähere Angaben	D, T
NICE 2003	Großbritannien	The National Collaborating Centre for Chronic Conditions / National Institute for Clinical Excellence	Primär- und Sekundärversorgung	D, T
OPOT 2000	Kanada	Ontario Program for Optimal Therapeutics	Primärversorgung in Ontario	P
UM 2001	USA	University of Michigan	Primärversorgung	D, T
UWH 2001	Deutschland	Universität Witten / Herdecke	Ohne nähere Angaben	D, T

*Empfehlungsbereich, der von der Leitlinie abgedeckt wurde: D – Diagnostik, (D) – Diagnostik enthalten, jedoch nicht schwerpunktmäßig, T – gesamte Therapie, P – Pharmakotherapie, P! – Therapieempfehlungen mit Schwerpunkt auf der Pharmakotherapie; [‡]Gesundheitsversorgung des Militärs und ehemaliger Militärangehöriger in den USA; [†]Health Plan: Krankenkassenleistungen von Arbeitgebern in den USA, häufig zusammengefasst zu großen überregionalen Organisationen.

Wie aus Tabelle 4.1 ersichtlich ist, unterschieden sich die berücksichtigten Leitlinien hinsichtlich des adressierten Versorgungskontexts, bei fünf Leitlinien fehlten explizite Angaben dazu. Der Empfehlungsbereich umfasste bei allen Leitlinien die Pharmakotherapie, bei 13 Leitlinien wurden auch Empfehlungen zur nicht-pharmakologischen Therapie abgegeben. Im Bereich Diagnostik arbeiteten nur fünf Leitlinien durchgängig mit Evidenzleveln und / oder Stärkegraden der Empfehlung [ACC/AHA 2001, Duodecim 2004, LLGH 2003, NHF/NZ 2001, NICE 2003], bei weiteren fünf Leitlinien wurden sie mindestens bei Empfehlungen zur technisch-apparativen Diagnostik angewendet [DieM 2003/2004, ICSI 2003, NHF/Austr & CSANZ 2002, UM 2001, UWH 2001]. In einer Leitlinie wurde nur ein einzelnes diagnostisches Verfahren bewertet [CCS 2002/3].

Die Leitlinien unterschieden sich auch hinsichtlich weiterer Parameter, wie der Anzahl der beteiligten Autoren, der Art und des Umfangs der Recherchen und der Zahl der bewerteten und im Text zitierten Referenzen (vgl. Tab. 4.2)²⁹.

Tab. 4.2: Berücksichtigte Leitlinien: weitere Merkmale

Kurzbezeichnung	Anzahl Autoren	Publikationszeitpunkt	Systematische Recherchen: Zeitraum / verwendete Datenbanken	Anzahl Referenzen
ACC/AHA 2001	14	Dezember 2001	Zeitraum: k. A. MEDLINE, EMBASE	573
AKDAE 2001	k. A.	Februar 2001	Keine syst. Recherche	216
CCS 2001 / CCS 2002/3	k. A.	(2001) Dez. 2003	(2001): k. A. 2002/3: MEDLINE	2001: 79; 2002/3: 42
DGK 2001	2	März 2001	Syst. Recherche 1990-2000 Datenbanken: k. A.	213
DieM 2003 / 2004	3	Juli 2003	Zeitraum: k. A. letzte Suche März 2003 in MEDLINE, Cochrane Library, „Best Evidence“	42
Duodecim 2004	k. A.	(2004)	Zeitraum: k. A. letzte Suche März 2004 in DARE, „Best Evidence“	50*
DVA & VHA 2002	14**	Dezember 2002	Updaterecherche Januar 2001 bis November 2002 MEDLINE	197
ESC 2002 / 2001°	18	September 2001	Detaillierte Angaben fehlen	196
ICSI 2003	11	Juli 2003	Detaillierte Angaben fehlen	Ca. 50

²⁹ Weitere Angaben zu den eingeschlossenen Leitlinien sind den Kurzbeschreibungen im oben genannten Methodenreport (vgl. Fußnote 25) zu entnehmen. Diese wurden in standardisierter Berichtsform erstellt.

(Fortsetzung Tabelle 4.2)

Kurzbezeichnung	Anzahl Autoren	Publikationszeitpunkt	Systematische Recherchen: Zeitraum / verwendete Datenbanken	Anzahl Referenzen
LLGH 2003	2	September 2003	Zeitraum: k. A. MEDLINE	72
NHF/Austr. & CSANZ 2002	33	(2002)	k. A.	143
NHF/NZ 2001	19	Dezember 2001	Zeitraum: k. A. letzte Suche April 2000 in MEDLINE	44
NICE 2003	15 [†]	Juli 2003	Beginn der Datenbanken bis September 2002 und Handsuchen; MEDLINE, EMBASE, CINAHL, PsycINFO, AMED, Cochrane Library, EconLit	347 ^{††}
OPOt 2000	k. A.	Herbst 2000	k. A.	24
UM 2001	6	September 2001	Systematisch von 1994 – Februar 1998 und Handsuchen bis 2001; MEDLINE	50
UWH 2001	5 + 2 [‡]	Mai 2001	k. A.	157

*Nur gelistete Dokumente erfasst, weitere Dokumente aus hinterlegten Evidenztafeln nicht vollständig erfasst, da aufgrund der interaktiven Struktur mit zahlreichen Querverlinkungen zu anderen Leitlinien-Dokumenten nicht eindeutig abgrenzbar; **Autoren sind Mitglieder der Pharmacy Benefits Management (PBM) Strategic Healthcare Group (SHG); verantwortlich für das Leitlinienprogramm bei DVA & VHA ist The Medical Advisory Panel for the Pharmacy Benefits Management Strategic Healthcare Group mit 13 Mitgliedern, die zusätzliche Unterstützung bei der LL-Entwicklung gaben; [†]Autoren waren Mitglieder der Guideline Development Group; in der Guideline Reference Group weitere 14 Experten, die an einzelnen Entwicklungsschritten der Leitlinie beteiligt waren; ^{††}Weitere Referenzen wurden in Evidenztabelle dargestellt, aber nicht im Text zitiert - sie wurden in der Referenzliste nicht erfasst; [‡]Fünf Autoren und zwei Berater

4.2.2 Ausgeschlossene Leitlinien

22 Leitlinien wurden formal ausgeschlossen, sie wurden im Anhang in Tabelle A.4 unter Nennung des Ausschlussgrundes aufgeführt. Wichtigste Ausschlussgründe waren: fehlende Evidenzbasierung (drei Leitlinien), fehlende Evidenzbezüge (zwei Leitlinien), fehlende Aktualisierung (neun Leitlinien) und Publikati-

onssprache anders als englisch oder deutsch (fünf Leitlinien). Vierzehn Leitlinien behandelten nur einzelne Aspekte zur medizinischen Versorgung herzinsuffizienter Patienten, wie z.B. zur Myokardszintigraphie. Sie waren für die vergleichende Evaluation nicht geeignet und wurden als Hintergrunddokumente verwendet.

4.2.3 Untersucherübereinstimmung

Beide Reviewer prüften unabhängig voneinander die Einschlussfähigkeit identifizierter Leitliniendokumente mit anschließender Konsentierung – die Ergebnisse zeigt Tabelle 4.3 in Form einer Kontingenztabelle.

Tab. 4.3: Beobachtete Übereinstimmung ein- und ausgeschlossener Dokumente

	B Ausschluss	B Einschluss	Summe
A Ausschluss	22	1	23
A Einschluss	0	15	15
Summe	22	16	38

A – Reviewer 1, B – Reviewer 2

Aus den Ergebnissen lässt sich die Kappa-Statistik als Maß dafür, inwieweit die beobachtete Übereinstimmung beider Reviewer ein rein zufälliges Maß an Übereinstimmung übertrifft, nach Cohen [1960] und McGinn et al. [2004] ermitteln. Danach betrug der Kappa-Wert $\kappa=0,95$ und war damit ‚almost perfect‘ [Landis & Koch 1977].

4.3 Ergebnisse der Bewertung der formalen methodischen Qualität

Die berücksichtigten Leitlinien wurden mit dem AGREE-Instrument [AGREE 2002, AGREE 2003, AGREE 2003a, AWMF & ÄZQ 2003] auf ihre methodische Qualität untersucht. Da diese Bewertung ressourcenbedingt nur durch eine Person (die Verfasserin) durchgeführt werden konnte, sind die Ergebnisse für jede eingeschlossene Leitlinie im Checklisten-Format dokumentiert (s. Methodenreport). Hier werden nur die wichtigsten Ergebnisse berichtet, im Anhang ist eine ausführliche Ergebnismitteilung zu den einzelnen Domänen aufgeführt.

Tabelle 4.4 auf der folgenden Seite zeigt die absoluten Domänenscores, die mit dem AGREE-Instrument für die berücksichtigten Leitlinien ermittelt wurden, sowie deren Standardisierung (zur Berechnung vgl. Abschnitt 3.2.3). In Tabelle 4.5 (Folgeside) werden Median und Streuung der Domänenscores dargestellt.

Tab. 4.4: Bewertung der Qualität: absolute u. standardisierte Domänenscores

Quelle (Kurzbezeichnung)	Domäne 1 (max. 12 Punkte) „Geltungsbereich und Zweck“	Domäne 2 (max. 16 Punkte) „Beteiligung von Interessengruppen“	Domäne 3 (max. 28 Punkte) „Methodologische Exaktheit der Leitlinienentwicklung“	Domäne 4 (max. 16 Punkte) „Klarheit und Gestaltung“	Domäne 5 (max. 12 Punkte) „Anwendbarkeit“	Domäne 6 (max. 8 Punkte) „Redaktionelle Unabhängigkeit“	Medianer Zielerreichungsgrad [†]
ACC/AHA 2001	8 (0,56)	11 (0,58)	18 (0,52)	16 (1,0)	7 (0,44)	4 (0,33)	0,54
AKDAE 2001	5 (0,22)	7 (0,25)	13 (0,29)	12 (0,67)	3 (0)	4 (0,33)	0,27
CCS 2001 / CCS 2002/3	6 (0,33)	8 (0,33)	18 (0,52)	10 (0,50)	3 (0)	3 (0,17)	0,33
DGK 2001	5 (0,22)	4 (0)	13 (0,29)	10 (0,50)	3 (0)	3 (0,17)	0,19
DleM 2003 / 2004	7 (0,44)	7 (0,25)	19 (0,57)	11 (0,58)	3 (0)	2 (0)	0,35
Duodecim 2004	8 (0,56)	6 (0,17)	21 (0,67)	12 (0,67)	5 (0,22)	2 (0)	0,39
DVA & VHA 2002	9 (0,67)	11 (0,58)	22 (0,71)	16 (1,0)	7 (0,44)	5 (0,50)	0,63
ESC 2002/2001a	5 (0,22)	9 (0,42)	16 (0,43)	16 (1,0)	4 (0,11)	2 (0)	0,32
ICSI 2003a	11 (0,89)	10 (0,50)	19 (0,57)	16 (1,0)	6 (0,33)	7 (0,83)	0,70
LLGH 2003	9 (0,67)	11 (0,58)	17 (0,48)	11 (0,58)	8 (0,56)	8 (1,0)	0,58
NHF/Austr. & CSANZ 2002	4 (0,11)	10 (0,50)	13 (0,29)	15 (0,92)	3 (0)	4 (0,33)	0,31
NHF/NZ 2001	9 (0,67)	5 (0,08)	17 (0,48)	15 (0,92)	5 (0,22)	2 (0)	0,35
NICE 2003	11 (0,89)	12 (0,67)	23 (0,76)	16 (1,0)	9 (0,67)	5 (0,50)	0,71
OPOt 2000	5 (0,22)	9 (0,42)	13 (0,29)	13 (0,75)	4 (0,11)	4 (0,33)	0,31
UM 2001	6 (0,33)	8 (0,33)	15 (0,38)	11 (0,58)	5 (0,22)	4 (0,33)	0,33
UWH 2001	5 (0,22)	6 (0,17)	11 (0,19)	12 (0,67)	4 (0,11)	5 (0,50)	0,21

Standardisierte Domänenscores in Klammern; [†]Ein Summenscore über alle Domänen wird nach AGREE nicht empfohlen. Zum Vergleich der Leitlinien über

alle Domänen wurde hier pro Leitlinie der „mediane Zielerreichungsgrad“ angegeben (Median der standardisierten Domänenscores aus Domäne 1 bis 6).

Tab. 4.5: Bewertung der Qualität: Median und Streuung

Quelle	Domäne 1	Domäne 2	Domäne 3	Domäne 4	Domäne 5	Domäne 6
Median	6,5 (0,39)	8,5 (0,38)	17 (0,48)	12,5 (0,71)	4,5 (0,17)	4 (0,33)
Spannweite	4-11 (0,11-0,89)	4-12 (0,0-0,67)	11-23 (0,19-0,76)	10-16 (0,5-1,0)	3-9 (0,0-0,67)	2-8 (0,0-1,0)

Unter den 16 berücksichtigten Leitlinien befanden sich fünf Leitlinien – davon eine aus Deutschland –, die einen medianen Zielerreichungsgrad von mindestens 0,5 über alle Domänen erzielten [ACC/AHA 2001, DVA & VHA 2002, ICSI 2003, LLGH 2003, NICE 2003]. In den Domänen 3 („Methodologische Exaktheit der Leitlinienentwicklung“) und 4 („Klarheit und Gestaltung“) wurde der höchste Zielerreichungsgrad über alle Leitlinien mit einem Median von 17 (0,48) in Domäne 3 resp. 12,5 (0,71) in Domäne 4 erzielt, wobei die Einzelergebnisse der Leitlinien stark streuten (Domäne 3: Spannweite 11 bis 23 (0,19 – 0,76), Domäne 4: 10 bis 16 (0,5 – 1,0)). Die größten Qualitätsunterschiede in Bezug auf die AGREE-Kriterien wurden in Domäne 6 („Redaktionelle Unabhängigkeit“) beobachtet (Spannweite 2 bis 8 (0,0 – 1,0)).

4.4 Ergebnis der Framework-Entwicklung

Aus den berücksichtigten Leitlinien wurden 92 Einzelthemen identifiziert. Angesichts der Bedeutung der Framework-Entwicklung für den gesamten weiteren Entwicklungsprozess (Stellen der 'richtigen Fragen', Extraktion der Empfehlungen etc.) wurde die Auswahl und Priorisierung der Themen von insgesamt vier Personen unserer Arbeitsgruppe mit breit gestreuten Kompetenzen und Erfahrungen durchgeführt. Von dieser Arbeitsgruppe wurden 16/92 Themen abgeschlossen:

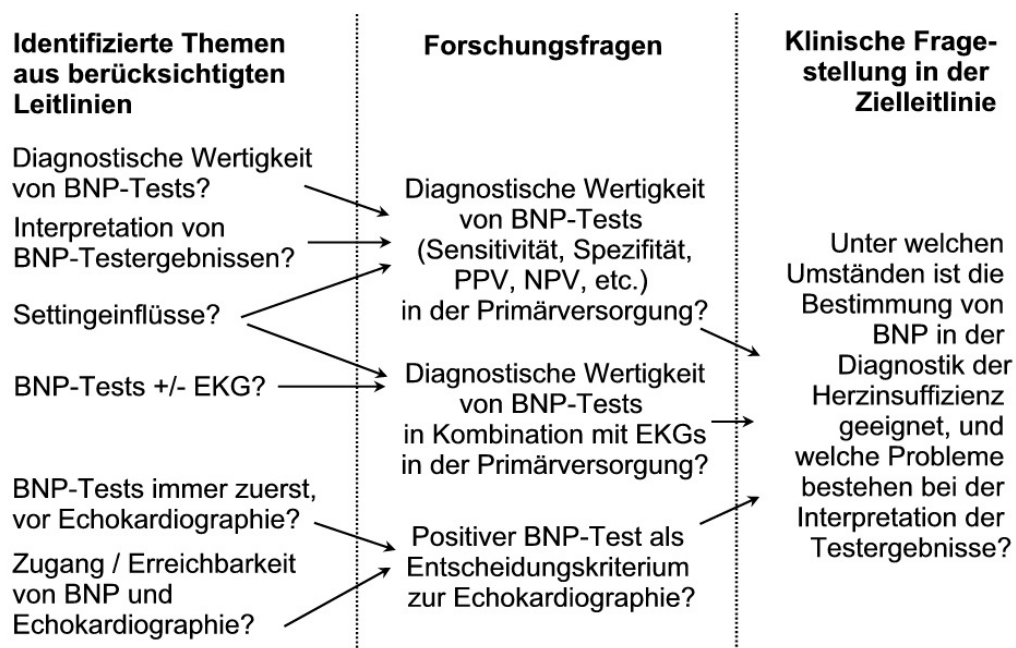
- 5/16 Themen lagen außerhalb des allgemeinmedizinischen Entscheidungsbereiches, wie z.B. die intravenöse Gabe von positiv inotropen Substanzen bei geplanter Herztransplantation.
- Bei 11/16 Themen wurde die Relevanz als weniger dringlich angesehen (z.B. die Sauerstoff-Heimtherapie als palliative Maßnahme) – die Bearbeitung der Fragen wurde für einen späteren Updateprozess der Leitlinie zurückgestellt.

Bei weiteren 17/92 Themen wurden Teilaspekte berücksichtigt, d.h. die Themenbereiche wurden stärker fokussiert. So wurden z.B. komplementärmedizinische Verfahren nur hinsichtlich möglicher Kontraindikationen zur Herzinsuffizienz

zienz sowie bezüglich potentieller Interaktionen zur evidenzbasierten Pharmakotherapie betrachtet. Eine eigenständige Evaluation ihrer Wirksamkeit bei Herzinsuffizienz wurde ebenfalls für einen späteren Aktualisierungsprozess zurückgestellt.

Die verbleibenden 59/92 Themen wurden zu 48 Themenkomplexen zusammengefasst, aus denen 27 explizite klinische Fragestellungen für die zu entwickelnde Leitlinie abgeleitet wurden. In Abbildung 4.2 wird dieser Prozess anhand eines Beispiels zum Test mit *Brain Natriuretic Peptides* (BNP) in der Erstdiagnostik der chronischen Herzinsuffizienz verdeutlicht.

Abb. 4.2: Frameworkentwicklung: Beispiel BNP-Test in der Erstdiagnostik



4.5 Ergebnisse der Datenextraktion

4.5.1 Kurzberichte der berücksichtigten Leitlinien

Zu den 16 in die systematische Evaluation eingeschlossenen Leitlinien wurden Kurzdarstellungen in standardisierter Berichtsform erstellt (siehe Methodenreport), die sich in den erfassten Merkmalen eng an die AGREE-Kriterien [AGREE 2002, AGREE 2003, AGREE 2003a] anlehnten und die hier zusammengefasst berichtet werden:

Gesamtziel der Leitlinie und Anwenderzielgruppe: Als Ziele der Leitlinien wurden klinische Outcomeparameter sowie Ergebnisparameter für Versorgungsziele benannt, wie z.B. Verbesserung der Symptomatik und Lebensqualität des Patienten, Hemmung der Krankheitsprogression, Verminderung von

Hospitalisierung und Sterblichkeit. Darüber hinaus wurden bei einer Leitlinie [ICSI 2003] auch Prozessparameter definiert (Erhöhung des Anteils der Herzinsuffizienzpatienten, deren Herzinsuffizienz echokardiographisch gesichert ist, die adäquat in ihrem Selbstsorgeverhalten geschult werden, die eine kausale Therapie erhalten, die in der Pharmakotherapie eine mortalitätssenkende Therapie mit ACE-Hemmern und Betablockern im Zieldosisbereich erhalten). Sofern eine Anwenderzielgruppe genannt wurde [ACC/AHA 2001, AKDAE 2001, CCS 2002/3, DiEM 2003/2004, Duodecim 2004, DVA & VHA 2002, ESC 2002/2001, ICSI 2003, LLGH 2003, NHF/Austr & CSANZ 2002, NICE 2003, OPOT 2000, UM 2001], wurde (auch) die Primärversorgung adressiert.

Konkrete Fragestellungen: Alle Leitlinien waren inhaltlich klar gegliedert, drei von ihnen [DVA&VHA 2002, ICSI 2003, NICE 2003] explizierten durchgängig spezifische klinische Fragestellungen bzw. -probleme.

Zielpopulation: Alle Leitlinien wurden ausschließlich für erwachsene Patienten mit chronischer Herzinsuffizienz konzipiert. Eine isolierte Rechtsherzinsuffizienz wurde nur bei drei Leitlinien [ACC/AHA 2001, DiEM 2003/2004, UWH 2001] und eine akute Herzinsuffizienz wurde bei sechs Leitlinien explizit ausgeschlossen [ACC/AHA 2001, ESC 2002/2001, NHF/NZ 2001, NICE 2003, OPOT 2000, UM 2001], ohne dass sich die Empfehlungen der übrigen Leitlinien auf diese Konditionen bezogen. In drei Leitlinien [DiEM 2003/2004, NHF/NZ 2001, UM 2001] wurde auch die diastolische Herzinsuffizienz ausgeschlossen, die übrigen Leitlinien behandelten diese Kondition.

Methodik inkl. Entwicklergruppe, Evidenzgewinnung und verwendete Systeme für Evidenzlevel und Stärken der Empfehlung: Sieben der berücksichtigten Leitlinien wurden in interdisziplinär und teilweise multiprofessionell besetzten Teams von z.T. deutlich mehr als 10 Autoren entwickelt [ACC/AHA 2001, DVA & VHA 2002, ESC 2002/2001, ICSI 2003, NHF/Austr & CSANZ 2002, NHF/NZ 2001, NICE 2003]. Zur Evidenzgewinnung wurden bei einigen der berücksichtigten Leitlinien umfassende und teilweise systematische Recherchen beschrieben (vgl. Tabelle 4.2), zur Bewertung von Evidenzleveln und Empfehlungsstärken wurden multiple, untereinander nicht kompatible Systeme verwendet.

Konsentierung: Zwei der berücksichtigten Leitlinien machten explizite Angaben zu einem Verfahren zur formalen Konsensfindung: bei der Leitlinie der *Canadian Cardiovascular Society* [CCS 2002/3] wurde über eine modifizierte Delphi-Technik mit interaktiven Anteilen berichtet, bei der Leitlinie des *National Institute of Clinical Excellence* [NICE 2003] über den Einsatz der „RAND/UCLA Appropriateness Method (RAM)“³⁰. Ergebnisse wurden zu keinem der beiden Konsensprozesse mitgeteilt.

³⁰ Die „RAND/UCLA Appropriateness Method (RAM)“ ist eine interaktive Modifikation der Delphi-Technik mit einem zweistufigen Bewertungsprozess (Rating) durch ein Expertenpanel.

Externer Begutachtungsprozess: Vier Leitlinien berichteten über ein formales mehrstufiges internes und externes Begutachtungsverfahren [ACC/AHA 2001, Duodecim 2004, DVA & VHA 2002, NICE 2003]. Bei vier Leitlinien wurden Begutachtungsverfahren durch Fachgesellschaften und externe Experten mitgeteilt [ESC 2002/2001, NHF/Austr & CSANZ 2002, NHF/NZ 2001, OPOT 2000], weitere vier Leitlinien eröffneten Diskussionsforen (z.T. internetbasiert) [CCS 2002/3, DleM 2003/2004, ICSI 2003, LLGH 2003], und eine Leitlinie nutzte ein *Peer-Review*-Verfahren bei der Publikation in einer Fachzeitschrift [UM 2001]. Die übrigen drei Leitlinien enthielten keine Angaben dazu.

Praxiserprobung: Vor ihrer Publikation wurde keine der berücksichtigten Leitlinien einem formalen Praxistest unterzogen. Eine Leitlinie berichtete über ein nach ihrer Veröffentlichung begonnenes Projekt zur Praxiserprobung und Implementierung [ACC/AHA 2001], eine Leitlinie wurde in enger Zusammenarbeit mit ‚Qualitätszirkeln Pharmakotherapie‘ entwickelt und diskutiert [LLGH 2003].

Gültigkeit: Es wurden unterschiedliche Vorgehensweisen zur Aktualisierung berichtet, darunter zwei Leitlinien mit institutionell standardisiertem Verfahren [DVA & VHA 2002, ICSI 2003]. Die meisten Leitlinien sollten regelmäßigen Reviews mit Update je nach Änderungsbedarf unterzogen werden (jährlich [ACC/AHA 2001, DleM 2003 / 2004, NHF/NZ 2001], zweijährlich [AKDAE 2001, ESC 2002/2001] usw.). Drei Leitlinien enthielten ein angekündigtes Revisionsdatum (Revision in 2002 [UWH 2001], 2004 [DGK 2001] und 2007 [NICE 2003]), das teilweise bereits überschritten war.

Zusatztools: Die Art und Zahl der für die praktische Leitlinienimplementierung entwickelten Zusatzinstrumente war sehr variabel. Vier Leitlinien waren in drei Versionen erhältlich (Langfassung, Kurzfassung und Patienteninformation) [AKDAE 2001, NHF/Austr & CSANZ 2002, NICE 2003, UWH 2001]. Bei vier Leitlinien fanden sich darüber hinaus eine Anzahl weiterer Tools, wie *Pocket-Guidelines* und Qualitätsindikatorensätze zur Evaluation der Leitlinienimplementierung [ACC/AHA 2001, DVA & VHA 2002, ICSI 2003]. Auch wurden z.B. Hilfen zur Entlassungsplanung nach stationärer Behandlung [DVA & VHA 2002], ein Instrument zur Messung der krankheitsspezifischen Lebensqualität [ACC/AHA 2001] oder Audiobänder [Duodecim 2004] angeboten. Die übrigen Leitlinien arbeiteten mit wenigen oder ohne Zusatztools.

4.5.2 Datenextraktion in Evidenztabelle

Wie unter Kapitel 4.4 beschrieben, wurden für das Framework der zu entwickelnden Leitlinie 27 spezifische klinische Fragen abgeleitet. Zu jeder dieser Fragen wurden die dazu in den berücksichtigten Leitlinien vorhandenen Daten in insgesamt 20 Evidenztabelle extrahiert (Anamnese und klinische Zeichen

Im Ergebnis liefert die RAM nicht zwangsläufig einen Konsens, sondern abschließende Bewertungsurteile (*„appropriate“*, *„inappropriate“*, *„uncertain“* oder *„necessary“*). [Fitch et al. 2001]

sowie verschiedene apparativ-technische diagnostische Tests wurden jeweils zusammenfassend extrahiert). Wegen des Umfangs dieser Tabellen wird hier nur ein Ausschnitt aus einer Evidenztabelle präsentiert (siehe Anhang, Tab. A.5; vollständige Evidenztabelle siehe Methodenreport zur Leitlinie). Nach Abschluss der Datenextraktion wurden 2 der 20 Evidenztabelle nicht näher berücksichtigt, weil sie für den hausärztlichen Entscheidungskontext in Deutschland nicht relevant waren. Sie bezogen sich auf die spezialärztliche Versorgung (Amiodaron und Dofetilid), oder die primäre Therapieindikation einer Intervention lag außerhalb der zu entwickelnden Leitlinie (Statine³¹).

4.6 Ergebnisse der Konsistenzanalyse

In der Konsistenzanalyse wurden die Empfehlungen der 16 Leitlinien, die in 18 entscheidungsrelevante Evidenztabelle extrahiert worden waren, systematisch miteinander verglichen. Daraus konnten insgesamt N=35, z.T. mehrgliedrige und komplexe Empfehlungen isoliert werden, die in der folgenden Konsistenzanalyse geprüft wurden: n=6 Empfehlungen zur Erstdiagnostik der Herzinsuffizienz, n=4 zur nicht-pharmakologischen Therapie, n=19 zur Pharmakotherapie der systolischen Herzinsuffizienz sowie n=6 zur diastolischen Herzinsuffizienz.

Zusammenfassend wurden 25 konsistente und neun inkonsistente Empfehlungen identifiziert (zur Klassifikation vgl. Tabelle 3.3):

- Konsistenzen: Typ 1, n=3; Typ 2, n=4; Typ 3, n=14; Typ 4, n=4
- Inkonsistenzen: Typ A (bedeutsam), n=3 und Typ B (weniger bedeutsam), n=6

Eine weitere Empfehlung zur Therapie der diastolischen Herzinsuffizienz wurde nur in einer einzelnen Leitlinie ausgesprochen und konnte daher nicht auf Konsistenz geprüft werden.

Tabelle 4.6 auf der nächsten Seite zeigt detaillierte Ergebnisse der Konsistenzanalyse zur Erstdiagnostik. Es handelte sich jeweils um konsistente Empfehlungen des Typs (3).

³¹ Im *Controlled Rosuvastatin Multinational Trial in Heart Failure* (CORONA) an älteren Patienten mit ischaemisch bedingter systolischer Herzinsuffizienz konnten keine Effekte von Rosuvastatin auf den kombinierten primären Endpunkt oder auf sekundäre Endpunkte wie Gesamtmortalität, koronare Ereignisse oder Tod aus kardiovaskulärer Genese gezeigt werden [Kjekshus et al. 2007]. Die Reduktion der Hospitalisierungen in dieser Studie sowie weitere, derzeit erfolgende Untersuchungen [Masoudi 2007] machen jedoch deutlich, dass diese Frage im Aktualisierungsprozess der Leitlinie erneut diskutiert werden muss und dass der SLR wichtige Hinweise für diesen Prozess liefert.

Tab. 4.6: Erstdiagnostik ohne BNP-Tests

	ACC/AHA 2001	AKDAE 2001	CCS 2001, 2002/3	DGK 2001	DieM 2003/2004	Duodecim 2004	DVA & VHA 2002	ESC 2002/2001	ICSI 2003	LLGH 2003	NHF/Austr. & CSANZ 2002	NHF/NZ 2001	NICE 2003	OPOT 2000	UM 2001	UWH 2001
Symptome und klinische Zeichen																
Evidenzlevel	C	#	Ø	#	Ø	A	#	#	#	IV	#	B	IV / IIb [†]	#	#	#
Grading	I		Ø				#		#	C			GPP			
Labor: Blutbild, Urinstatus, Serumelektrolyte, Harnstoff, Kreatinin (weitere Tests bei einzelnen LL empfohlen)																
Evidenzlevel	C	#	Ø	#	Ø	#	Ø	#	R	Ø	EO	#	Ø	Ø	Ø	C
Grading	I		Ø				Ø		Ø	C			GPP			
Ruhe-EKG (12 Ableitungen)																
Evidenzlevel	C	#	Ø	#	Ø	#	Ø	#	R	Ø	EO	#	Ø	Ø	D	B
Grading	I		Ø				Ø		Ø	B			GPP			
Röntgen-Thorax																
Evidenzlevel	C	#	Ø	#	Ø	#	Ø	#	R	Ø	EO	#	Ø	Ø	Ø	C
Grading	I		Ø				Ø		Ø	Ø			GPP			
Transthorakale Echokardiographie																
Evidenzlevel	C	#	Ø	#	I	#	Ø	#	C, R	Ø	EO	#	IIb	Ø	D	C
Grading	I		Ø				Ø		II	Ø			GPP			
Ø – Keine Angabe, da außerhalb der LL; # - Empfehlung im Text ohne Evidenzlevel & Grading; [†] - verschiedene diagnostische Studien mit teilweise inkonsistenten Ergebnissen																

In Tabelle 4.7 auf der folgenden Seite wird ein Beispiel für eine *bedeutsame Inkonsistenz* des Typs (A) vorgelegt: Empfehlungen zum BNP-Test.

Die Ergebnisse der Konsistenzanalyse zu den Typen (1), (2), (4) und (B) werden anhand ausgewählter Beispiele zur besseren Übersichtlichkeit gemeinsam mit den Ergebnissen zur Analyse der Belegtheit präsentiert (vgl. Abschnitt 4.8.2). Die übrigen Ergebnisse der Konsistenzanalyse wurden im Anhang aufgeführt (siehe Anhang, Tabellen A.6 – A.13).

Tab. 4.7: BNP-Test

	ACC/AHA 2001	AKDAE 2001	CCS 2001, 2002/3	DGK 2001	DileM 2003/2004	Duodecim 2004	DVA & VHA 2002	ESC 2002/2001	ICSI 2003	LLGH 2003	NHF/Austr. & CSANZ 2002	NHF/NZ 2001	NICE 2003	OPOT 2000	UM 2001	UWH 2001
BNP-Test zum Ausschluss einer Herzinsuffizienz (allein od. in Kombination m. EKG)																
Evidenzlevel	Ø	Ø		Ø	Ø		Ø			Ø		Ø	IIb	Ø		Ø
Grading	Ø						Ø			Ø			B			
BNP-Test bei akuter Dyspnoe und HI-Verdacht (ggf.) hinzuziehen																
Evidenzlevel	Ø	Ø	1	Ø	Ø	#	Ø	#	A, C	Ø		Ø		Ø		Ø
Grading	Ø		A				Ø		Ø	Ø						
BNP-Test für Routinevorgehen nicht empfohlen																
Evidenzlevel	Ø	Ø		Ø	Ø		Ø			Ø	#	Ø		Ø	#	Ø
Grading	Ø						Ø			Ø						
Ø – Keine Angabe, da außerhalb der LL; # - Empfehlung im Text ohne Evidenzlevel & Grading																

4.7 Ergebnisse der Informationssynthese

Unter Berücksichtigung der allgemeinen Gliederungsformate von DEGAM-Leitlinien³² entstanden auf der Basis der Frameworkentwicklung ein erster Gliederungsentwurf sowie spezifische klinische Fragen, die von der zu entwickelnden Leitlinie beantwortet werden sollten. Im Ergebnis der Konsistenzanalyse wurden Leitlinien-Empfehlungen formuliert. Die Ergebnisse der Konsistenzanalyse und der in den Leitlinien berichteten Evidenz wurden zu jeder Fragestellung unter der Überschrift „Evidenz und Rationale“ zusammengefasst und diskutiert. Auch konnten aus den berücksichtigten Leitlinien verschiedene Tools zur Unterstützung der Praxisimplementierung (z.B. Dosierungshilfen und Warnhinweise) in den Erstentwurf der Leitlinie aufgenommen werden. Der Erstentwurf wurde von der Autorin der vorliegenden Arbeit nach sechs Monaten fertiggestellt (Musterkapitel siehe Anhang).

³² Gliederungsformate von DEGAM-Leitlinien betreffen Aufbau und Nummerierung der Kapitel: 1. Einführung (inkl. Ziele und Definition, Epidemiologie, Ursachen und natürlicher Verlauf sowie Abwendbare und gefährliche Verläufe), 2. Vorgehensweise, 3. Zusammenfassung, 4. Methodik, 5. Ergebnisse, 6. Gesundheitsökonomische Aspekte, 7. Verbreitungs- und Implementierungsplan, 8. Evaluation, 9. Verantwortung, Gültigkeitsdauer und Fortschreibung, 10. Autoren, Kooperationspartner, Förderung und Konsensusverfahren, 11. Literatur.

4.8 Ergebnisse der Analyse der Belegtheit

Das Ergebnis der Konsistenzanalyse weist darauf hin, dass die Quelleitlinien in diesem systematischen Leitlinienreview nicht nur zu teilweise divergierenden oder divergent begründeten Empfehlungen kommen, sondern auch einen unterschiedlichen Gebrauch von der (prinzipiell) vorhandenen Evidenz machen. In einem Entwicklungsprozess einer qualitativ hochwertigen Leitlinie darf es jedoch nicht geschehen, sich von derartigen Prämissen abhängig zu machen. Der SLR umfasst daher als eigenständigen Arbeitsschritt die Analyse der Belegtheit von Schlüsselempfehlungen.

Da der Methode des SLR der Vorwurf entgegengehalten werden könnte, die Methode öffne der unkritischen Übernahme fremder evidenzbasierter Analyse- und Konsensergebnisse Tür und Tor, kommt diesem Arbeitsschritt eine hohe Bedeutung zu. Es werden deshalb hier einige Beispiele der Analyse der Belegtheit ausführlich dokumentiert. Der eilige Leser kann diesen Abschnitt kursorisch lesen und anschließend bei Abschnitt 4.9 fortsetzen.

Wie in Tabelle 3.4 angegeben, erfolgte eine Analyse der Belegtheit für konsistente Empfehlungen der Typen (1), (2) und (4) sowie bei Inkonsistenzen des Typs (B) anhand von Originalquellen, die in den berücksichtigten Leitlinien zitiert wurden. Eine Zusammenstellung der dabei behandelten Themen zeigt Tab. 4.8 (Themen mit ausführlicher Ergebnisdarstellung sind fett-kursiv gedruckt).

Tab. 4.8: Themenzusammenstellung für die Analyse der Belegtheit

(In-)Konsistenztyp	Themen
1	ACE-Hemmer bei allen NYHA-Klassen, β-Blocker bei NYHA I nach Infarkt, β-Blocker bei NYHA II-III
2	Aldosteron-Antagonisten bei NYHA III-IV, Digitalis bei HI + tachykardem VHF, Orale Antikoagulation bei HI mit VHF oder stattgehabtem thrombembolischen Ereignis, Kontrolle des arteriellen Hypertonus bei diastolischer Herzinsuffizienz
4	Körperliches Training, Diuretika bei NYHA II-IV, Hydralazin/ISDN als Mittel 2. Wahl, Unbedenklichkeit von langanflutenden Dihydroperidinen
B	Salz- und Flüssigkeitsrestriktion, β-Blocker bei klinisch stabilen NYHA IV, β-Blocker bei NYHA I unabhängig von Infarktereignis, AT₁-Blocker bei ACE-Hemmerintoleranz, Digitalis bei SR und NYHA III-IV, absolute Kontraindikation für Kalziumantagonisten (außer langanflutende Dihydroperidine)

Bei Konsistenzen des Typs (3) und Typ-(A)-Inkonsistenzen erscheinen umfassende bzw. systematische Recherchen zur Gewinnung neuer Evidenz erforderlich. Diese Themen wurden für ein formales Priorisierungsverfahren zur Anpassung an verfügbare Ressourcen vorbereitend gelistet (siehe Tab. 4.9).

Tab. 4.9: Themenliste für systematische/umfassende Evidenzrecherchen

<p>Konsistente Empfehlungen der berücksichtigten Leitlinien mit schwacher oder fehlender Evidenz (Typ-(3)-Konsistenzen:</p> <ul style="list-style-type: none"> - Symptome und klinische Zeichen der Herzinsuffizienz - Klinisch-chemische Erstdiagnostik - Ruhe-EKG (12 Ableitungen) bei Erstdiagnostik - Röntgen-Thorax bei Erstdiagnostik - Echokardiographie bei Erstdiagnostik - Lebensstilmodifikation - Impfeempfehlungen - Keine AT₁-Blocker anstelle von ACE-Hemmern bei ACE-Hemmer-naiven Patienten - Orale Antikoagulation bei Pat. ohne VHF oder stattgehabtem thrombembolischen Ereignis - Thrombozytenaggregationshemmung bei gleichzeitig bestehender KHK - Pharmakotherapie der diastolischen Herzinsuffizienz
<p>Komplette Inkonsistenz von Leitlinienempfehlungen (Typ-(A)-Inkonsistenz):</p> <ul style="list-style-type: none"> - BNP-/NT-proBNP-Test als diagnostischer Test - AT₁-Blocker zusätzlich zu ACE-Hemmern und β-Blockern (Triplet-Therapie) - AT₁-Blocker zusätzlich zu ACE-Hemmern bei Patienten, die keine β-Blocker erhalten können

4.8.1 Rechercheergebnisse

Die in den berücksichtigten Leitlinien zitierten Evidenzquellen wurden in der Datenextraktion in Evidenztabelle dokumentiert, zusammengefasst und um Dubletten bereinigt. Dabei wurden insgesamt 309 Publikationen identifiziert (siehe Tabelle 4.10 auf der folgenden Seite).

In der initialen Recherche wurden gegenüber den Quellleitlinien zwei zentrale Arbeiten neu identifiziert, obwohl sie z.T. im Recherchezeitraum dieser Leitlinien bereits hätten gefunden werden können³³:

- ein US-amerikanischer HTA-Bericht der *Agency for Healthcare Research and Quality* zur Pharmakotherapie bei Herzinsuffizienz und linksventrikulärer Dysfunktion [Shekelle et al. 2003],

³³ Dieser Umstand kann darauf zurückgeführt werden, dass beide Arbeiten als HTA-Berichte nicht in den gängigen elektronischen Datenbanken, wie MEDLINE oder EMBASE, gelistet sind.

- ein Kurz-HTA des Canadian Coordinating Office for Health Technology Assessment (CCOHTA)³⁴ aus deren Horizon Scanning-Programm [Ma & Boucher 2002].

Tab. 4.10: Anzahl und Art der in den Leitlinien zitierten Evidenzquellen

Thema	Zahl systematischer Übersichten	Zahl Primärpublikationen	Sonstige Dokumente
Nicht-pharmakologische Therapie der Herzinsuffizienz			
Körperliches Training	2 SRs	16 RCTs	9 Ü
Salz- und Flüssigkeitsrestriktion	0 SRs od. MA	0 RCT + 2 S	3 Ü
Pharmakotherapie der systolischen Herzinsuffizienz			
ACE-Hemmer	6 SRs	45 RCTs	
β-Blocker	9 SRs + 4 MA	40 RCTs	
AT ₁ -Blocker	1 SR + 2 MA	20 RCTs	
Aldosteron-Antagonisten	0 SRs od. MA	9 RCTs (darunter 2 Mortalitätsstudien)	
Diuretika	0 SRs + 1 MA	4 RCTs + 24 S	
Digitalis	5 SRs	11 RCTs + 25 S	1 Ü
Hydralazin/ISDN	0 SRs od. MA	6 RCTs	
Kalziumantagonisten	1 SR	11 RCTs + 16 S	1 Ü
Orale Antikoagulation	3 SRs + 1 MA	6 RCTs + 3 S	12 Ü/K
Therapie einer diastolischen Herzinsuffizienz			
Kontrolle eines arteriellen Hypertonus	0 SRs + 1 MA	2 RCTs	7 Ü
Gesamtzahl	27 SRs + 9 MA	170 RCTs + 70 S	33 Ü

SR – systematischer Review mit oder ohne darin enthaltene Metaanalysen, MA – Metaanalyse(n) von Studien, nicht auf systematischem Review beruhend, RCT – randomisierte klinische kontrollierte Studie, S – sonstige Studien: nicht-randomisiert und/oder Fallzahl N < 50, Ü – unsystematische Übersicht (traditioneller Review / Editorial), K – Konsensusempfehlungen

³⁴ Das Canadian Coordinating Office for Health Technology Assessment (CCOHTA) heißt seit April 2006 Canadian Agency for Drugs and Technologies in Health (CADTH). <http://www.cadth.ca/>

4.8.2 Ergebnisse der Bewertung von Primärquellen und HTA-Berichten

In diesem Kapitel werden die Ergebnisse an je einem ausgewählten Fallbeispiel zu jedem der Konsistenz-Typen (1), (2) und (4) sowie zum Inkonsistenz-Typ (B) präsentiert, wobei in der Darstellung folgendes Berichtsschema gewählt wurde:

- (1) Klinische Fragestellung, die von der Leitlinie beantwortet werden soll
- (2) Schlüsselempfehlungen zu dieser Fragestellung
- (3) Tabellarische Präsentation der Ergebnisse der Konsistenzanalyse
- (4) Ergebnisse der Reevaluation von Sekundärstudien, die bereits Bestandteil der berücksichtigten Leitlinien waren
- (5) Ergebnisse aus neu identifizierten HTA-Berichten, sofern vorhanden
- (6) Ergebnisse der Reevaluation von Primärstudien, die bereits Bestandteil der berücksichtigten Leitlinien waren (sofern dieser Schritt erforderlich wurde)
- (7) Zusammenfassung der Ergebnisse und weiterer Forschungsbedarf

Über das jeweils ausgewählte Fallbeispiel hinaus sind weitere Ergebnisse im Anhang zusammengefasst.

4.8.2.1 Konsistente Empfehlung, Typ (1), Fallbeispiel ACE-Hemmertherapie

- (1) **Klinische Fragestellung:** „Welche Patienten sollten einen ACE-Hemmer erhalten und welche Dosierung ist dabei sinnvoll?“
- (2) **Schlüsselempfehlungen:** die sechs in den berücksichtigten Leitlinien konsistenten Empfehlungen zur ACE-Hemmertherapie zeigt Tab. 4.11.

Tab. 4.11: Konsistente Leitlinienempfehlungen zur Therapie mit ACE-Hemmern

i	NYHA I-IV: Alle Patienten mit einer nachgewiesenen systolischen Dysfunktion und fehlenden Kontraindikationen sollten ACE-Hemmer erhalten, unabhängig davon, ob sie Symptome einer Herzinsuffizienz aufweisen, oder nicht [alle eingeschlossenen Leitlinien] $\{I_{x1} A_{x1}\}$.
ii	Ziel der ACE-Hemmertherapie ist die Prognoseverbesserung mit Senkung der Sterblichkeit und Verzögerung des Krankheitsprogresses sowie die Verbesserung der Symptomatik und der Hospitalisierungsrate bei leichter, mäßiger und schwerer Herzinsuffizienz (NYHA II-IV) [alle eingeschl. Leitlinien] $\{I_{x1} A_{x1}\}$.
iii	Bei asymptomatischen Patienten mit eingeschränkter LV-Funktion (NYHA I) sollen durch die Gabe von ACE-Hemmern die Inzidenz einer symptomatischen Herzinsuffizienz und die Hospitalisierungsrate gesenkt und die langfristige Prognose verbessert werden. Bei postinfarzieller LV-Dysfunktion soll zusätzlich die Reinfarktrate gesenkt werden [alle eingeschlossenen Leitlinien] $\{I_{x1} A_{x1}\}$.

(Fortsetzung Tabelle 4.11)

iv	Insgesamt können ACE-Hemmer die Mortalität und Morbidität der systolischen Herzinsuffizienz umso stärker verbessern, je schwerer die Herzinsuffizienz ist.
iv.a	Die Effekte sind jedoch in allen NYHA-Klassen nachweisbar [DieM 2003/2004, Duodecim 2004, DVA & VHA 2002, ICSI 2003/2004, NHF/Austr. & CSANZ 2002, NICE 2003, OPOT 2000, UM 2001, UWH 2001] {Ia _{x1} A _{x1} }.
v	ACE-Hemmer sind Mittel der ersten Wahl vor AT ₁ -Blockern, da eine Überlegenheit der Wirksamkeit dieser Substanzgruppe bislang nicht gezeigt werden konnte und die Datenlage zu ACE-Hemmern zuverlässiger ist. Ein Wechsel auf AT ₁ -Blocker sollte nur bei ACE-Hemmerhusten erwogen werden (näheres siehe Abschnitt AT ₁ -Blocker). [Alle eingeschlossenen Leitlinien] {Ia _{x1} A _{x1} }.
vi	Dosierung: Nach Möglichkeit sollten HI-Patienten auf die jeweils höchste, in Mortalitätsstudien ermittelte Zieldosis bzw. die maximal tolerierte Dosis aufgetitert werden. Für höhere Dosierungen konnten verminderte Hospitalisierungsraten gegenüber niedrigen Dosierungen gezeigt werden, der Effekt auf die Mortalität war nicht konsistent dosisabhängig [ACC/AHA 2001, CCS 2002/2003, DGK 2001, Duodecim 2004, DVA & VHA 2002, ESC 2002/2001, ICSI 2004/2003, NHF/Austr. & CSANZ 2002, NICE 2003, OPOT 2000, UM 2001] {Ib _{x1} A _{x1} }.

Vorläufige Evidenzlevel und Stärkegrade der Empfehlung in geschweiften Klammern

(3) **Konsistenzanalyse:** siehe Tabelle 4.12.

Tab. 4.12: Konsistenzanalyse der ACE-Hemmertherapie

	ACC/AHA 2001	AKDAE 2001	CCS 2001, 2002/3	DGK 2001	DieM 2003/2004	Duodecim 2004	DVA & VHA 2002	ESC 2002/2001	ICSI 2003	LLGH 2003	NHF/Austr. & CSANZ 2002	NHF/NZ 2001	NICE 2003	OPOT 2000	UM 2001	UWH 2001
ACE-Hemmer bei asymptomatischen Patienten mit oder ohne Infarktereignis																
Evidenzlevel	B	↑	1	B	I	A	B	A	A	Ib	I	A	Ia	A	A	A
Grading	I		A				A		I	A			A			
ACE-Hemmer bei asymptomatischen Patienten nach stattgehabtem Infarktereignis																
Evidenzlevel	B	↑	1	A	I	A	A	A	A	Ib	I	A	Ia	A	A	A
Grading	I		A				A		I	A			A			
ACE-Hemmer bei HI-Patienten mit NYHA II-IV																
Evidenzlevel	A	↑↑	1	A	I	A	A	A	A	Ib	I	A	Ia	A	A	A
Grading	I		A				A		I	A			A			

(4) **Reevaluation von Sekundärstudien:** Zur therapeutischen Wirksamkeit der ACE-Hemmer wurden von den Leitlinien sechs systematische Übersichtsarbeiten als Evidenzquellen angegeben. Davon wurden vier systematische Reviews mit darin enthaltenen Metaanalysen [Eccles et al. 1998, Flather et al. 2000,

Lubsen et al. 1996 und Neal et al. 2000] neu bewertet (Evidenztabelle siehe Methodenreport). Zwei weitere systematische Reviews wurden nicht näher untersucht: die bei Garg & Yusuf [1995] eingeschlossenen 32 RCTs waren auch Bestandteil einer bereits berücksichtigten Metaanalyse [Eccles et al. 1998], und die Arbeit von Mancini [2000] untersuchte Surrogat-Endpunkte.

- **Eccles et al. [1998]** untersuchten im Rahmen des *North of England Evidence Based Guideline Development Project* die Wirksamkeit verschiedener ACE-Hemmer auf die Mortalität, die Lebensqualität und Kosteneffektivität im Vergleich zu Placebo. Sie schlossen dazu 39 RCTs mit N=8.457 Patienten (davon 28 RCTs an Populationen mit niedrigem Risiko und 11 RCTs an Hochrisikokollektiven, Risikogrenzwert: Mortalität < bzw. > 15%; 81% Männer; keine Angabe von NYHA-Klassen) in ihre Metaanalyse ein und fanden unter ACE-Hemmern eine Verringerung der Gesamtmortalität, der Sterblichkeit nach stattgehabtem Myokardinfarkt sowie eine Verbesserung von Lebensqualität, Belastungstoleranz und Krankheitssymptomen. Die Effekte waren in der Hochrisikogruppe größer. Zwischen den Studien bestanden deutliche Unterschiede (eingeschlossene Population, Dauer des Follow-ups von 8 Wochen bis über 40 Monaten, etc.), statistisch konnte keine Heterogenität nachgewiesen werden.

Die methodische Qualität und Berichtsqualität waren deutlich eingeschränkt (Ein- und Ausschlusskriterien für Studien wurden nicht expliziert, die Qualität der Studien wurde nicht bewertet, detaillierte Angaben zur verwendeten Methodik fehlten, der Grenzwert zur Unterscheidung zwischen *high* vs. *low risk*-Studien wurde nicht begründet, die Metaanalyse erfolgte nach dem *fixed effect model*, die klinische Heterogenität wurde nicht diskutiert, es erfolgten keine Sensitivitätsanalysen zur Prüfung auf Robustheit der Ergebnisse). Der SR wurde öffentlich gefördert, potentielle Interessenkonflikte wurden verneint.

- **Flather et al. [2000]** führten eine Metaanalyse individueller Patientendaten (IPD) aus «Mega-Trials» (N > 1.000 Patienten) zur Wirksamkeit von ACE-Hemmern vs. Placebo für einen Follow-up-Zeitraum von mindestens 12 Monaten durch. Aus den Daten von N=12.763 Patienten (5 RCTs) konnte eine signifikante Reduktion der Gesamtmortalität sowie der Hospitalisierung unter ACE-Hemmern ermittelt werden. Weiterhin wurden in der Verumgruppe deutlich mehr unerwünschte Arzneimittelwirkungen (UAW) beobachtet, hauptsächlich Hypotension und (nach Absetzen reversible) Niereninsuffizienz. Bei Stratifikation nach Schweregrad der Einschränkung der Ejektionsfraktion (EF) wurden die stärksten Effekte bei der größten Einschränkung der EF beobachtet (< 23%), bei einer EF oberhalb von 27% waren keine statistisch signifikanten Effekte der ACE-Hemmer auf die Mortalität mehr nachweisbar. NYHA-Klassen wurden nicht berichtet.

Die methodische und Berichtsqualität waren eingeschränkt: das Einschlusskriterium N > 1000 wurde nicht begründet, ein *selection bias* ist möglich. Die methodische Qualität der eingeschlossenen Studien wurde nicht bewertet.

Die IPD-Analyse war problematisch, da in den eingeschlossenen Studien unterschiedliche Definitionen für akute Infarktereignisse und für eine ‚Hospitalisierung durch Verschlechterung der HI‘ verwendet wurden, es wurde mit dem *fixed effect model* gerechnet, die Robustheit der Ergebnisse wurde nicht in Sensitivitätsanalysen geprüft. Die Untersuchung wurde partiell industriell gefördert, Angaben zu potentiellen Interessenkonflikten der Autoren fehlten.

- **Lubsen et al. [1996]** fassten die Ergebnisse aus fünf multizentrischen RCTs aus Europa, USA und Südafrika (N=1.053 Patienten) metaanalytisch zusammen. Dabei wurden Verumgruppe (Ramipril) und Placebokontrollen verglichen, wobei die statistische Power nicht ausreichte, um statistisch signifikante Effekte auf die Mortalität aufzuzeigen, die funktionelle Kapazität (Fahrradergometrie und NYHA-Klasse) war unter Ramipril gebessert.

Die methodische und Berichtsqualität waren deutlich eingeschränkt: bibliographische Angaben zu den eingeschlossenen Studien fehlten (!), die Studienqualität wurde nicht bewertet, Ein- und Ausschlusskriterien zur Auswahl der Studien wurden nicht expliziert, Angaben zur weiteren Vorgehensweise fehlten ebenso, wie Heterogenitätsuntersuchungen. Es wurde das *fixed effect model* verwendet, Sensitivitätsanalysen zur Prüfung auf Robustheit der Ergebnisse fehlten. Die Studie wurde komplett industriell gesponsert, zu potentiellen Interessenkonflikten machten die Autoren keine Angaben.

- **Neal et al. [2000]** berichteten von einer prospektiv geplanten IPD-Metaanalyse aus 15 RCTs mit verschiedenen Untersuchungszielen, u.a. zu Effekten von ACE-Hemmern auf die inzidente Entwicklung einer Herzinsuffizienz bei hypertensiven, initial nicht herzinsuffizienten Patienten (nicht primäres Untersuchungsziel der Metaanalyse (!), n=4 RCTs (N=12.124 Patienten)). Im Vergleich zwischen ACE-Hemmern der Interventionsgruppe mit Placebo- sowie verschiedenen Verum-Kontrollgruppen (je Gruppenvergleich zwischen 2.400 und 8.100 Patienten pro Gruppe, Verum: Diuretika oder β -Blocker, Kalziumantagonisten) zeigte sich ein jeweils leicht vermindertes relatives Risiko zur Ausbildung einer Herzinsuffizienz unter ACE-Hemmern (nicht statistisch signifikant).

Die methodische und Berichtsqualität waren eingeschränkt: die Angaben zu ACE-Hemmern und Begleitmedikation waren unzureichend, es erfolgte keine Bewertung der Studienqualität, es wurde das *fixed effect model* angewendet, Sensitivitätsanalysen zur Prüfung der Robustheit der Ergebnisse fehlten. Die Untersuchung war teilweise industriell gefördert, Angaben zu potentiellen Interessenkonflikten der Autoren fehlten.

(5) Ergebnisse aus einem neu identifizierten HTA-Bericht:

- Der **US-amerikanische HTA-Bericht [Shekelle et al. 2003]** untersuchte die Wirksamkeit von ACE-Hemmern und β -Rezeptorenblockern auf die Gesamtmortalität von weiblichen und männlichen Patienten, Patienten mit und ohne Diabetes mellitus, sowie weißen und farbigen Patienten. In einer Me-

taanalyse (*random effects model*) konnte auch für asymptotische Patienten mit Herzinsuffizienz (NYHA I) ein statistisch signifikanter Überlebensvorteil durch die Einnahme von ACE-Hemmern gezeigt werden (allerdings nur für Männer), ebenso wie bei symptomatischer Herzinsuffizienz (NYHA II-IV). Bei Frauen konnte ein Überlebensvorteil durch ACE-Hemmer nur für symptomatische Stadien der Herzinsuffizienz dargestellt werden, nicht für asymptotische. Frauen waren in den betreffenden Studien jedoch deutlich unterrepräsentiert.

Die methodische und Berichtsqualität waren gut (s. Kurzbeschreibung und Checkliste im Anhang). Der HTA-Bericht wurde öffentlich gefördert, potentielle Interessenkonflikte wurden für Autoren, Reviewer etc. abgefragt, die Ergebnisse wurden jedoch nicht publiziert.

- Bei dem **kanadischen HTA-Bericht [Ma & Boucher 2002]** handelte es sich um eine Kurzbewertung der Wirksamkeit und Sicherheit von AT₁-Blockern in der Therapie der Herzinsuffizienz. Die von den Autoren aufgefundene Evidenz reichte nicht aus, um eine Empfehlung für AT₁-Blocker als gleichwertige oder als additive Therapie zu ACE-Hemmern auszusprechen. Die Autoren empfahlen, AT₁-Blocker als Therapiealternative bei ACE-Hemmer-Intoleranz einzusetzen und weitere klinische Studien durchzuführen.

Die methodische und Berichtsqualität waren für einen Kurz-HTA angemessen (Kurzbeschreibung und Checkliste s. Methodenreport). Dieser wurde öffentlich gesponsert, zu potentiellen Interessenkonflikten der Autoren fehlten die Angaben.

(6) **Reevaluation von Primärstudien:** Für die Empfehlungen zur *Dosierung* von ACE-Hemmern (vi) mussten Primärstudien reevaluiert werden, da diese Frage durch die Sekundärpublikationen nicht ausreichend beantwortet wurde. Von den insgesamt 45 Primärpublikationen wurden drei RCTs in den berücksichtigten Leitlinien als Evidenzquellen für Dosierungsempfehlungen angeführt (Evidenztabelle s. Methodenreport) [Packer et al. 1999 (ATLAS), Nanas et al. 2000, NETWORK 1998].

- **Packer et al. [1999] (ATLAS)** führten einen Multicenter-RCT zur Untersuchung von Gesamtmortalität, Hospitalisierung und kombiniertem Endpunkt aus beiden an N=3.194 überwiegend männlichen Patienten mit chronischer Herzinsuffizienz im Stadium NYHA II-IV durch. Die Interventionsgruppe wurde mit 32,5 bis 35 mg Lisinopril täglich behandelt, die Kontrollgruppe mit 2,5 bis 5,0 mg Lisinopril, der Follow-up erfolgte über mindestens drei Jahre. Ergebnisse: s.u.

Die methodische und Berichtsqualität waren eingeschränkt: es fehlten Angaben zu Randomisation, *allocation concealment*³⁵ und Verblindung, die Ergebnismitteilung der sekundären Outcomes war unvollständig, die Niedrigdosis-Gruppe wurde mit sehr (!) niedrigen Dosierungen behandelt, die deut-

³⁵ *Allocation concealment* – Geheimhaltung der Behandlungsfolge [Moher et al. 2001]

lich unter den Dosierungen vorangegangener Mortalitätsstudien lagen. Die Studie war auch für den Nachweis eines 8%-igen Mortalitätsunterschieds unterpowerd, und die klinische Relevanz eines solchen Effekts wurde nicht diskutiert. Das Sponsoring der Studie erfolgte industriell, zu potentiellen Interessenkonflikten fehlten die Angaben.

- **Nanas et al. [2000]** berichteten über eine Studie an N=248 Patienten zur Untersuchung der 1-Jahres-Gesamtmortalität und Hospitalisierung unter 20 mg Enalapril / d (Kontrollen) vs. 60 mg Enalapril / d (Intervention). Ergebnisse: s.u.

Die methodische und Berichtsqualität waren erheblich eingeschränkt: Angaben zu Randomisation und *allocation concealment* fehlten, die Untersucher waren für die Dosis nicht verblindet, die *Intention-to-treat*-Analyse erfolgte ohne nähere Angaben. Eine Fallzahlkalkulation wurde nicht berichtet, aber die Studie erscheint deutlich unterpowerd für einen zweiseitigen Test und den Endpunkt Mortalität. Die Studie war komplett industriell gesponsert, Angaben zu potentiellen Interessenkonflikten der Autoren fehlten.

- **NETWORK [1998]**: Dieser dreiarmlige Multicenter-RCT über 24 Wochen untersuchte die Mortalität, die Hospitalisierung aus kardialer Ursache und/oder die Verschlechterung der bestehenden Herzinsuffizienz (NYHA II-IV) an N=1532 Patienten (ca. 500 Patienten je Gruppe) unter 2x2,5 mg (Gruppe I) vs. 2x5 mg (Gruppe II) vs. 2x10 mg (Gruppe III) Enalapril. Ergebnisse: s.u. Die methodische und Berichtsqualität waren eingeschränkt: die Angaben zur Vorgehensweise waren stark verkürzt, die von den Autoren berichtete separate Publikation des Studienprotokolls war anhand der bibliographischen Angaben nicht auffindbar. Angaben zu Randomisation und *allocation concealment* fehlten. Das dreiarmlige Design ist schwierig zu interpretieren, der Endpunkt ‚Verschlechterung der Herzinsuffizienz-Symptomatik‘ war unscharf definiert. Die Studie wurde komplett industriell gefördert, Angaben zu potentiellen Interessenkonflikten der Autoren fehlten.

Wirksamkeit: In keiner der drei Studien konnten statistisch signifikante Mortalitätsvorteile für die Patienten der Interventionsgruppen nachgewiesen werden, wobei alle drei Studien dafür eine ungenügende Power³⁶ aufwiesen. Für die Patienten der Interventionsgruppe zeigte sich in ATLAS ein statistisch signifikanter Vorteil in einer geringeren Hospitalisierung aus kardiovaskulärer Ursache und im kombinierten Endpunkt Mortalität / kardiovaskuläre Morbidität.

Sicherheit: Bei ATLAS wurden in der Interventionsgruppe vermehrt Schwindel, Hypotension, Verschlechterung der Nierenfunktion und Hyperkaliaemie beobachtet, in der Kontrollgruppe vermehrt Husten und Verschlechterung der Herz-

³⁶ Die Fallzahlkalkulation bei ATLAS bspw. war geeignet, um einen 15%-igen Unterschied zwischen den Gruppen bei 90% Power, $\alpha=0,05$ im zweiseitigen Test zu detektieren. Die beobachtete Risikoreduktion lag bei 8% der Gesamtmortalität.

insuffizienz. Bei Nanas et al. [2000] erreichten 32,5% nach drei Monaten bzw. 45% nach 12 Monaten die angestrebte Zieldosis der Intervention.

Die Zieldosierung der „Hochdosis-“Gruppe von NETWORK entsprach den Zieldosierungen, wie sie in anderen Mortalitätsstudien zu Enalapril³⁷ verwendet wurden. Die Zieldosis lag bei ATLAS mit 35 mg / Tag deutlich über anderen Studien³⁸, ebenso bei Nanas et al. mit 60 mg Enalapril / Tag.

(7) **Zusammenfassung:** Die Reevaluation der zitierten Sekundärpublikationen erbrachte ausreichende Belege mit den in den Quelleitlinien angegebenen höchsten Evidenzleveln für die Empfehlungen (ii) und (iv). Für die Empfehlungen (i) und (iv.a) fanden sich zwar Belege für die NYHA-Klassen II-IV, *nicht* jedoch für *asymptomatische* Patienten. Diese wurden erst in dem neu identifizierten HTA-Bericht von Shekelle et al. [2003] für männliche Patienten (NYHA I) für die Empfehlungen (i), (iii) und (iv.a) mit den dort angegebenen höchsten Evidenzleveln (reserviert für systematische Reviews / Metaanalysen oder mehrere ‚Megatrials‘) gefunden. Der Kurz-HTA von Ma & Boucher [2002] ist für die vorliegende Arbeit nur von begrenztem Wert: – die Empfehlung für ACE-Hemmer als Mittel erster Wahl (v) wird untermauert, allerdings gründet sie sich auf fehlende Studienevidenz für eine Überlegenheit der AT₁-Blocker. Diese Fragestellung ist jedoch bereits für umfassende Recherchen zur Gewinnung neuer Evidenz vorgesehen (vgl. Tab. 4.9). Die Dosierungshinweise der Empfehlung (vi) wurden in mindestens einem RCT belegt.

4.8.2.2 Konsistente Empfehlungen vom Typ (2), Fallbeispiel Aldosteron-Antagonisten bei Herzinsuffizienz, NYHA III-IV

(1) **Klinische Fragestellung:** „Welche Patienten sollten einen Aldosteron-Antagonisten erhalten und welche Risiken sind dabei besonders zu beachten?“

(2) **Schlüsselempfehlungen:** die drei in den berücksichtigten Leitlinien konsistenten Empfehlungen zur Therapie mit Aldosteron-Antagonisten zeigt Tab. 4.13 auf der folgenden Seite.

(3) **Konsistenzanalyse:** siehe Tabelle 4.14 (Folgeseite).

³⁷ Sowohl bei SOLVD [SOLVD 1991, Pitt et al. 1992, Yusuf et al. 1992] als auch bei CONSENSUS [Swedberg et al. 1999] wurde Enalapril im Interventionsarm mit 2 x 10 mg / Tag dosiert.

³⁸ In anderen RCTs wurde Lisinopril auf 20 mg / Tag auftitriert, wie z.B. in Chalmers et al. [1987], Giles et al. [1990], Herlitz [1992], Bach & Zardini [1992], Zannad et al. [1992], Beller et al. [1995] – diese Studien untersuchten überwiegend funktionelle Outcomes, keine Mortalität und hatten zumeist nur kurze Follow-up-Zeiten.

Tab. 4.13: Schlüsselempfehlungen zur Therapie mit Aldosteron-Antagonisten

i	Patienten mit Herzinsuffizienz NYHA III-IV oder kürzlich stattgehabter NYHA-IV-Symptomatik, die unter optimaler Therapie (ACE-Hemmer + β -Blocker + Diuretikum +/- Digitalis) mäßig bis hochgradig symptomatisch bleiben, sollten bei erhaltener Nierenfunktion niedrig dosiert Spironolacton erhalten [alle eingeschlossenen Leitlinien {Ib _{x1} A _{x1} }].
ii	Das Therapieziel besteht in einer Verbesserung des Überlebens und der bestehenden Symptome [alle eingeschlossenen Leitlinien] {Ib _{x1} A _{x1} }.
iii	Voraussetzung: bei Therapiebeginn sollten das Serum-Kreatinin unterhalb 2,5 mg/dl und das Serum-Kalium unterhalb 5,0 mmol/l betragen! [ACC/ AHA 2001, DieM 2003, DVA & VHA 2002, ESC 2002/2001, NHF/Austr. & CSANZ 2002, NHF/NZ 2001, NICE 2003, UM 2001] {Ib _{x1} A _{x1} }.

Vorläufige Evidenzlevel und Stärkegrade der Empfehlung in geschweiften Klammern

Tab. 4.14: Konsistenzanalyse zur Therapie mit Aldosteron-Antagonisten

	ACC/AHA 2001	AKDAE 2001	CCS 2001, 2002/3	DGK 2001	DieM 2003/2004	Duodecim 2004	DVA & VHA 2002	ESC 2002/2001	ICSI 2003	LLGH 2003	NHF/Austr & CSANZ 2002	NHF/NZ 2001	NICE 2003	OPOT 2000	UM 2001	UWH 2001
Niedrig dosierte Gaben von Aldosteron-Antagonisten bei NYHA III-IV																
Evidenzlevel	B	↑	1	B	I	B	B	B	A	Ib	II	A	Ib	A	A	A
Grading	Ila		A				B		A	A			A			

(4) **Reevaluation von Sekundärstudien:** es wurden keine systematischen Reviews oder Metaanalysen identifiziert.

(5) **Ergebnisse aus neu identifizierten HTA-Berichten:** es wurden keine HTA-Berichte zu dieser Fragestellung aufgefunden.

(6) **Reevaluation von Primärstudien:** Von den berücksichtigten Leitlinien wurden neun Primärpublikationen angeführt, davon zwei RCTs mit dem primären Outcome Mortalität, die hier reevaluiert wurden [Pitt et al. (RALES) 1999, Pitt et al. (EPHESUS) 2003] (Evidenztable siehe Anhang, Tab. A.14), zusätzliche Primärstudien zu anderen Endpunkten werden bei der weiterführenden Recherche zu Sicherheitsaspekten (vgl. Handlungsoptionen, Tab. 3.3) berücksichtigt.

- **Pitt et al. (RALES) [1999]:** In diesem Multicenter-RCT wurden in 195 Zentren (!) N=1.663 Patienten mit chronischer Herzinsuffizienz im Stadium NYHA III-IV, die bereits ACE-Hemmer, β -Blocker, Diuretika (+/- Digitalis) als Basistherapie erhalten durften und die bei Einschluss ein Serum-Kreatinin \leq 2,5 mg/dl und ein Serum-Kalium \leq 5 mmol/l aufwiesen, untersucht. Die Studie wurde nach 24 Monaten vorzeitig beendet, da sich ein signifikanter

Überlebensvorteil für die Interventionsgruppe unter 25-50 mg Spironolacton gegenüber der Placebo-Kontrollgruppe ergab. Es wurden ebenfalls statistisch signifikante Vorteile für die sekundären Endpunkte Tod und/oder Hospitalisierung aus kardialer Ursache gezeigt. Unter Spironolacton wurden mittlere Anstiege des Serum-Kreatinins um 0,05-0,1 mg/dl und des Serum-Kaliums um 0,3 mmol/l im ersten Therapiejahr sowie eine relevante Zahl von Gynäkomastien / Mastodynien mit Therapieabbruch beobachtet.

Die methodische und Berichtsqualität waren eingeschränkt: die große Zahl von Zentren mit nur etwa 8 Patienten / Zentrum machen eine hohe Varianz zwischen Patienten, verwendeter Diagnostik und Therapie wahrscheinlich, Angaben zu Randomisation, *allocation concealment* und Bestimmungsmethode der Ejektionsfraktion fehlten. Die berichtete Symptomverbesserung war verzerrt (NYHA-IV-Verschlechterung nur bei Tod möglich). Die Studie war industriell gefördert, zu potentiellen Interessenkonflikten fehlten die Angaben, s.u. (EPHESUS).

- **Pitt et al. (EPHESUS) [2003]**: Dieser Multicenter-RCT verglich in 674 Zentren (!) die Wirksamkeit von 25 bis 50 mg Eplerenon / d mit Placebo hinsichtlich der primären Endpunkte Gesamtmortalität und kombiniertem Endpunkt aus Tod aus kardiovaskulärer Ursache und Hospitalisierung aus derselben Ursache. Eingeschlossen wurden insgesamt N=6.632 Patienten am 3.-14. Tag nach akutem Myokardinfarkt bei LVEF \leq 40 %, Stauungszeichen (klinisch od. radiologisch), Serum-Kreatinin \leq 2,5 mg/dl, -Kalium \leq 5 mmol/l. Der primäre Endpunkt wurde mit dem 1.012. Todesfall erreicht und wies dabei eine Follow-up-Zeit von durchschnittlich 16 Monaten auf. Die primären Endpunkte traten in der Interventionsgruppe signifikant seltener auf als in der Placebo-Kontrollgruppe. Als Begleitmedikation wurden für Interventions- und Kontrollgruppe ACE-Hemmer, β -Blocker, Diuretika, Acetylsalicylsäure und Statine berichtet. In der Verum-Gruppe traten häufiger als in der Placebo-Gruppe schwere Hyperkaliaemien auf, insbesondere bei initial erniedrigter endogener Kreatinin-Clearance.

Die methodische und Berichtsqualität waren leicht eingeschränkt: die große Zahl von Zentren: siehe RALES; die Randomisation wurde ohne nähere Angaben als *Interactive Voice Randomization System* beschrieben, Dosierungsangaben für ACE-Hemmer- und β -Blocker-Begleitmedikation fehlten. Die Studie wurde industriell gefördert, zu potentiellen Interessenkonflikten wurde berichtet, dass zwei Autoren als Berater beim Förderer arbeiteten.

(7) **Zusammenfassung**: die Empfehlung für den Einsatz von Aldosteron-Antagonisten (Spironolacton und Eplerenon) gründet sich auf je eine Mortalitätsstudie, die von allen Leitlinien zitiert wurden. In beiden RCTs konnten Überlebensvorteile sowie verminderte Hospitalisierungsraten aus kardiovaskulärer Genese für die Interventionsgruppen gezeigt werden. Die Wirksamkeit für Eplerenon wurde mit EPHESUS nur für Herzinsuffizienzpatienten nach akutem Myokardinfarkt belegt. Die Empfehlungen sind durch die reevaluierten Studien belegt.

4.8.2.3 Konsistente Empfehlung, Typ (4), Fallbeispiel Diuretikatherapie bei symptomatischer Herzinsuffizienz

- (1) **Klinische Fragestellung:** „Welche Patienten sollten Diuretika erhalten?“
- (2) **Schlüsselempfehlungen:** die vier im Wortlaut konsistenten Empfehlungen der berücksichtigten Leitlinien zeigt Tabelle 4.15.

Tab. 4.15: Schlüsselempfehlungen zur Therapie mit Diuretika

i	Alle Herzinsuffizienzpatienten, die Zeichen einer Flüssigkeitsretention aufweisen, sollten Diuretika erhalten, da Diuretika die einzige Therapieoption zur Kontrolle des Volumenhaushalts darstellen [ACC/AHA 2001, AKDAE 2001 / DGK 2001, DieM 2003, Duodecim 2004, DVA & VHA 2002, ESC 2002/2001, ICSI 2004/2003, LLGH 2003, NHF/Austr. & CSANZ 2002, NHF/NZ 2001, NICE 2003, OPOT 2000, UM 2001, UWH 200126], {Ia _{x1} A _{x1} }.
ii	Ziel der Therapie ist eine Symptombesserung mit Reduktion von pulmonalem Rückstau und / oder rechtsventrikulärem / abdominellem Rückstau mit peripheren Ödemen sowie Verbesserung der Belastungstoleranz [ebenda] {Ia _{x1} A _{x1} }.
iii	Diuretika sollten bei Herzinsuffizienz in Kombination mit mortalitätssenkenden Pharmaka verabreicht werden, um eine langfristige Stabilisierung zu erreichen [ebenda] {Ia _{x1} A _{x1} }.
iv	(Abrupte) Diuretika-Auslassversuche provozieren akute Verschlechterungen mit der Gefahr eines Lungenödems und sollten darum vermieden werden [AKDAE 2001 / DGK 2001, DieM 2003, DVA & VHA 2002] {Ia _{x1} A _{x1} }.

Vorläufige Evidenzlevel und Stärkegrade der Empfehlung in geschweiften Klammern

- (3) **Konsistenzanalyse:** siehe Tabelle 4.16.

Tab. 4.16: Konsistenzanalyse zur Diuretikatherapie

	ACC/AHA 2001	AKDAE 2001	CCS 2001, 2002/3	DGK 2001	DieM 2003/2004	Duodecim 2004	DVA & VHA 2002	ESC 2002/2001	ICSI 2003	LLGH 2003	NHF/Austr & CSANZ 2002	NHF/NZ 2001	NICE 2003	OPOT 2000	UM 2001	UWH 2001
Diuretika bei allen HI-Patienten der Klassen NYHA II-IV																
Evidenzlevel	A	↑↑	∅	A	I	∅	A	A	A	Ia	EO	B	Ib	B	D	A
Grading	I		∅				A		∅	A			C			

- (4) **Reevaluation von Sekundärstudien:** Zur therapeutischen Wirksamkeit von Diuretika bei der Herzinsuffizienztherapie wurde eine Metaanalyse [Faris et al. 2002] von den Leitlinien als Evidenzquelle zitiert (Evidenztabelle):

- **Faris et al. [2002]** berichteten über Metaanalysen von acht placebokontrollierten (inkl. RALES [Pitt et al. 1999]) und 10 aktiv kontrollierten RCTs zur Wirksamkeit von Diuretika bei Herzinsuffizienzpatienten. In den placebokontrollierten Studien wurden N=1.286 Patienten eingeschlossen (ohne RALES: N=436), in den aktiv kontrollierten Studien N=383 Patienten, wobei verwendete Diuretika, Follow-up, Begleitmedikation, Outcomes und Studiendesigns unter den Studien variierten. Der von den Autoren berichtete mortalitätsreduzierende Effekt der Diuretika beruhte auf Ergebnissen aus drei placebokontrollierten Studien (N=221) ohne RALES, zur Krankheitsprogression und Verschlechterung der bestehenden Herzinsuffizienzsymptomatik nach Absetzen der Diuretika auf vier Studien³⁹, zur Verbesserung der Belastungstoleranz unter Diuretika auf sechs Studien. Die Autoren schlussfolgerten, dass ihre Ergebnisse keine ausreichende Evidenz liefern, um einen mortalitätsreduzierenden Effekt von Diuretika anzunehmen und empfahlen den Diuretika-Einsatz auf der Basis von klinischen Erfahrungen, da zukünftig keine neuen placebokontrollierten Studien zu diesen Substanzen zu erwarten seien. Die methodische und Berichtsqualität waren erheblich eingeschränkt: die Qualität der berücksichtigten Studien wurde nicht bewertet, klinisch bestand ausgeprägte Heterogenität innerhalb und zwischen den Studien, die Autoren wählten das *fixed-effects-model*, bei fehlendem Heterogenitätsnachweis im χ^2 -Test (Power gering) wurde Homogenität angenommen. Sensitivitätsanalysen zur Prüfung auf die Robustheit der Ergebnisse erfolgten nicht, die Limitationen der Analyse wurden ungenügend diskutiert. Aber: die konservativen Schlussfolgerungen der Autoren waren angemessen. Zu Sponsoring und potentiellen Interessenkonflikten fehlten die Angaben.

(5) **Ergebnisse aus neu identifizierten HTA-Berichten:** es wurden keine HTA-Berichte identifiziert.

(6) **Reevaluation von Primärstudien:** Primärstudien wurden nicht reevaluiert.

(7) **Zusammenfassung:** es konnten Studienergebnisse identifiziert werden, welche die Empfehlungen zur Diuretikatherapie (i) bis (iv) unterstützten. Die Vergabe von Evidenzlevel und Stärkegrad der Empfehlung ist weiterhin nicht eindeutig: formal liegen mehrere randomisierte kontrollierte Studien in einer Metaanalyse vor, die methodische Qualität von eingeschlossenen Primärstudien wie auch der Metaanalyse selbst machen potentiell ergebnisverzerrende Effekte jedoch sehr wahrscheinlich.

³⁹ In den vier Studien wurde ein sogenanntes *withdrawal-design* eingesetzt, d.h. in der Interventionsgruppe wurde die Diuretikatherapie abgebrochen.

4.8.2.4 Inkonsistenz Typ (B), Fallbeispiel: β -Blockertherapie bei asymptomatischen Patienten ohne Infarktereignis

(1) **Klinische Fragestellung:** „Welche Patienten sollten einen β -Rezeptorenblocker erhalten, welche Substanzen sollten bevorzugt werden, und welche Dosierung ist anzustreben?“

(2) **Schlüsselempfehlungen:** Die sechs Empfehlungen zur β -Blockertherapie in den berücksichtigten Leitlinien zeigt Tab. 4.17, wobei hier nur die Ergebnisse bezüglich der Aussagen (i) und (iv) ausführlich dargestellt werden, da sich die Inkonsistenz auf asymptotische Patienten ohne stattgehabtes Infarktereignis bei sieben Leitlinien vs. einer Leitlinie bezieht, die diese Subpopulation explizit ausschließt [UM 2001]. Die übrigen Empfehlungen sind konsistent (Typ 1 bzw. 3) und werden am Ende kurz zusammengefasst.

Tab. 4.17: Schlüsselempfehlungen zur Therapie mit β -Blockern

I	Alle klinisch stabilen Patienten mit nachgewiesener Herzinsuffizienz der NYHA-Klassen I-IV und Fehlen von Kontraindikationen sollten mit β -Rezeptorenblockern therapiert werden [ACC/AHA 2001, CCS 2002/2003, Duodecim 2004, DVA & VHA 2002, ICSI 2003/2004, LLGH 2003, NICE 2003, UWH 2001] $\{Ia_{x1} A_{x1}\}$.
ii	Mindestvoraussetzungen für Auswahl der Patienten zur β -Blocker-Therapie: Normovolämie (!), keine Ruhedyspnoe, systolischer Blutdruck >90 mmHg, Herzfrequenz >65 /Min., keine kürzlich stattgehabte i.v.-Therapie mit positiv inotropen Substanzen, keine höhergradigen AV-Blockierungen ohne effektive Schrittmacherversorgung, keine schwere COPD [ACC/AHA 2001, DVA & VHA 2002, LLGH 2003, NHF/Austr. & CSANZ 2002, NHF/NZ 2001, NICE 2003, UM 2001] $\{IV_{x1} C_{x1}\}$.
iii	Bei Problemen mit Hypotension, Bradykardie, Flüssigkeitsretention oder COPD vor oder während der Therapieeinstellung oder bei Intoleranzerscheinungen schon unter geringen Dosierungen Kooperation mit Kardiologen oder stationäre Einstellung erwägen [DGK 2001, Duodecim 2004, DVA & VHA 2002, ESC 2002/2001, ICSI 2004/2003, NHF/NZ 2001, NICE 2003, UM 2001] $\{IV_{x1} C_{x1}\}$.
iv	Die Therapieziele der β-Blocker-Therapie bestehen langfristig in einer Verminderung von Sterblichkeit, Krankheitsprogression, Hospitalisierungsrate sowie Symptomen und einer Verbesserung des klinischen Status und der Lebensqualität [alle eingeschlossenen Leitlinien] $\{Ia_{x1} A_{x1}\}$.
V	β -Blocker offenbar ohne Klasseneffekt: Herzinsuffizienzpatienten sollten mit Bisoprolol oder Metoprolol (beide kardioselektiv) oder Carvedilol (nicht-kardioselektiv) behandelt werden. Bei Patienten mit COPD sollten kardioselektive β -Blocker bevorzugt werden. Kontraindiziert sind β -Blocker mit ISA (intrinsischer sympathomimetischer Aktivität), wie z.B. Sotalol oder Xamoterol, da sie die Sterblichkeit erhöhen. Bucindolol hatte keinen mortalitätssenkenden Effekt und zu anderen β -Blockern, wie z.B. Atenolol fehlten ausreichende Daten. [AKDAE 2001 / DGK 2001, DieM 2003, Duodecim 2004, LLGH 2003, NHF/Austr. & CSANZ 2002, NICE 2003, UM 2001] $\{Ia_{x1} A_{x1}\}$.
Vi	Dosierung: β -Blocker sollten, beginnend bei einer geringen Startdosis, sehr langsam bis zum Erreichen der Zieldosis bzw. der maximal tolerierten Dosis auftitriert werden. Die oben genannten Effekte wurden unter den Zieldosen, deren Verträglichkeit im Allgemeinen als gut bewertet wurde, beobachtet [alle eingeschlossenen Leitlinien] $\{IV_{x1} C_{x1}\}$.

(3) **Konsistenzanalyse:** siehe Tabelle 4.18.

Tab. 4.18: Konsistenzanalyse zur β -Rezeptorenblockertherapie

	ACC/AHA 2001	AKDAE 2001	CCS 2001, 2002/3	DGK 2001	DileM 2003/2004	Duodecim 2004	DVA & VHA 2002	ESC 2002/2001	ICSI 2003	LLGH 2003	NHF/Austr & CSANZ 2002	NHF/NZ 2001	NICE 2003	OPOT 2000	UM 2001	UWH 2001
β-Blocker bei asymptomatischen Patienten mit oder ohne stattgehabtes Infarkt-ereignis																
Evidenzlevel	B	Ø	Ø	Ø	Ø	A	B	Ø	A	Ø	Ø	Ø	Ia	Ø	D	A
Grading	I		A				A		I	Ø			A			
β-Blocker bei asymptomatischen Patienten nach stattgehabtem Infarkt ereignis																
Evidenzlevel	A	Ø	Ø	Ø	Ø	A	A	B	A	Ø	II	Ø	Ia	B	A	A
Grading	I		A				A		I	Ø			A			
β-Blocker bei Herzinsuffizienz-Patienten mit NYHA II-III																
Evidenzlevel	A	↑↑	1	A	I	A	A	A	A	Ia	I	A	Ia	A	A	A
Grading	I		A				A		I	A			A			
β-Blocker bei klinisch stabilen Herzinsuffizienz-Patienten mit NYHA IV																
Evidenzlevel	A	↑↑	1	A	I	A	A	A	A	Ia	II	A	Ia	C*	D**	Ø
Grading	I		A				A		I	A			A			
*Kein Therapiebeginn bei NYHA IV, bei bereits bestehender Therapie Dosis halbieren, ggf. Therapieabbruch nach Entscheidung des behandelnden Arztes; **Keine generelle Empfehlung einer β -Blockertherapie bei NYHA IV																

(4) **Reevaluation von Sekundärstudien:** Zur therapeutischen Wirksamkeit der β -Rezeptorenblocker wurden 13 systematische Übersichtsarbeiten und Metaanalysen von den Leitlinien als Belege zitiert. Davon wurden vier systematische Reviews mit darin enthaltenen Metaanalysen [Bonet et al. 2000, Bouzamondo et al. 2001, Brophy et al. 2001, Whorlow & Krum 2000] reevaluiert (Evidenztabelle siehe Methodenreport). Die übrigen neun Übersichtsarbeiten lieferten gegenüber den genannten keine zusätzlichen Informationen, da sie über Surrogatparameter berichteten [Packer et al. 2001], da sie nicht auf andere Primärstudien zurückgriffen [Doughty et al. 1997, Exner et al. 1999, Heidenreich et al. 1997, Lechat et al. 1998, Shibata et al. 2001, van Campen et al. 1998, Yancy et al. 2001] oder sich auf andere Zielerkrankungen bezogen [Freemantle et al. 1999].

- **Bonet et al. [2000]:** Dieser systematische Review im Auftrag der katalanischen HTA-Agentur evaluierte 21 RCTs (mit N=5.849 Patienten, 88,7% NYHA II-III) mit Follow-up Zeiten von 3 bis 32 Monaten (Median 6 Monate) systematisch und fasste deren Ergebnisse quantitativ zusammen. 18 nicht berücksichtigte RCTs wurden unter Nennung des Ausschlussgrundes aufge-

führt. Es wurden Interventionsgruppen (Bisoprolol, Bucindolol, Carvedilol, Metoprolol, Propranolol) mit Placebo- bzw. Standard-Kontrollgruppen hinsichtlich Mortalität, Hospitalisierung und weiterer Unterschiede, wie z.B. der Effektunterschiede zwischen vasodilatatorisch und nicht-vasodilatatorisch wirksamen β -Blockern, verglichen. Gesamtmortalität, kardiovaskuläre Mortalität und Hospitalisierungsrate wurden in den Interventionsgruppen stärker gesenkt als in den Kontrollen, wobei für die Mortalität ein stärkerer Effekt durch vasodilatatorisch wirksame β -Blocker zu beobachten war (bei Hospitalisierung diesbezüglich kein signifikanter Unterschied). Von den eingeschlossenen Studien schloss nur eine NYHA I-Patienten (30% von N=415) ein, die jedoch ausschließlich eine ischaemische Genese aufwiesen [ANZ Heart Failure Research Collaborative Group 1997].

Die methodische und Berichtsqualität waren gut (klar explizierter Methodenteil mit adäquater Recherchestrategie, Bemühungen zum Auffinden nicht-publizierter Literatur, unabhängige Bewertung der Studien durch zwei Autoren, Nennung der ausgeschlossenen Studien mit Ausschlussgrund, Wahl einer konservativen Analysemethode, Patientencharakteristika explizit beschrieben, absolute und relative Angaben zu Effektschätzern; Mangel: keine Bewertung der Studienqualität berichtet). Die Untersuchung wurde öffentlich gefördert, zu potentiellen Interessenkonflikten der Autoren wurde nichts berichtet.

- **Bouzamondo et al. [2001]** fertigten eine Metaanalyse über 16 RCTs (davon vier RCTs neu gegenüber Bonet et al.) über die Behandlungseffekte von Bisoprolol, Bucindolol, Carvedilol und Metoprolol auf Gesamtmortalität, Hospitalisierung und Therapieeffektivität hinsichtlich der Schwere der Erkrankung bei insgesamt N=14.857 Patienten an. Die bei Bonet et al. [2000] beschriebene Mortalitätsreduktion und Senkung der Hospitalisierungsrate bei NYHA II-III wurde von Bouzamondo et al. [2001] für Bisoprolol, Carvedilol und Metoprolol bestätigt, wobei ein RCT⁴⁰ unter den Studien für die statistische Heterogenität verantwortlich war und daraufhin von der Metaanalyse ausgeschlossen wurde. In einer substanzspezifischen Subgruppenanalyse wurden unter Bucindolol deutlich schlechtere Effekte beobachtet. Keine Aussage erfolgte zu Patienten mit NYHA I mit nicht-ischämischer Genese und zu NYHA IV-Patienten, da diese unterrepräsentiert waren.

Die methodische und Berichtsqualität waren erheblich eingeschränkt: die Recherchemethoden wurden nicht berichtet, ausgeschlossene Studien wurden nicht aufgeführt, die Studienqualität wurde nicht bewertet, es lagen keine Angaben zur Begleitmedikation vor. Die Autoren wählten das *fixed effects model*. Eine kritische Diskussion der Ergebnisse fehlte, aber die Autoren gelangten zu angemessenen Schlussfolgerungen. Zu Sponsoring und potentiellen Interessenkonflikten fehlten die Angaben.

⁴⁰ In der Studie erhielten die Interventionspatienten Bucindolol [The BEST Steering Cmte. 1995]

- **Brophy et al. [2001]** führten eine bayesianische Metaanalyse über 22 RCTs (N=10.135) durch, die wegen der abweichenden Methode reevaluiert wurde, obgleich diese RCTs bereits bei Bonet et al. [2000] und Bouzamondo et al. [2001] eingeschlossen waren. Die Ergebnisse der oben beschriebenen Metaanalysen wurden in dieser Arbeit erneut für NYHA II-III-Patienten bestätigt. Weiterhin ist keine Aussage über NYHA IV-Patienten (weniger als 5% der eingeschlossenen Patienten) und über NYHA I-Patienten mit nicht-ischämischer Genese möglich.

Die methodische und Berichtsqualität waren sehr gut (klar explizierter Methodenteil, die methodische Qualität der eingeschlossenen Studien wurde untersucht und als gut bewertet, Detailergebnisse wurden jedoch nicht berichtet, die Metaanalyse erfolgte im hierarchischen *random effects model*, Sensitivitätsanalysen zur Prüfung auf Robustheit der Ergebnisse wurden angefertigt). Die Untersuchung wurde öffentlich gefördert, zu potentiellen Interessenkonflikten der Autoren lagen keine Angaben vor.

- **Whorlow & Krum [2000]** untersuchten die Wirksamkeit von β -Rezeptorenblockern auf NYHA IV-Patienten in einer Metaanalyse über 18 RCTs mit N=635 Patienten (Verumgruppe: n=313 vs. Placebo n=322). Ein statistisch signifikanter Mortalitätsgewinn konnte für diese Patienten gezeigt werden, Untersuchungen zur Verträglichkeit wurden nicht vorgenommen. Die methodische und Berichtsqualität waren erheblich eingeschränkt: Recherchen waren auf MEDLINE begrenzt, eine Untersuchung auf einen *Publication Bias* erfolgte nicht, das Einschlusskriterium ‚objektiv dokumentierte Herzinsuffizienz‘ wurde nicht definiert, die Studienqualität wurde nicht bewertet, Heterogenitätstests wurden nicht berichtet, ebenso fehlten Sensitivitätstests zur Prüfung der Robustheit der Ergebnisse. Zu Sponsoring und potentiellen Interessenkonflikten der Autoren lagen keine Angaben vor.

(5) **Ergebnisse aus neu identifizierten HTA-Berichten:** In dem o.g. US-amerikanischen HTA-Bericht [Shekelle et al. 2003] wurden keine geschlechtsspezifischen Wirksamkeitsunterschiede der β -Blocker beobachtet, weitere Aussagen wurden zu Effektstärken bei weißen / nichtweißen, diabetischen / nicht-diabetischen Herzinsuffizienzpatienten getroffen. Der HTA-Bericht enthielt keine Informationen zu asymptomatischen Patienten mit nicht-ischämischer Genese.

(6) **Reevaluation von Primärstudien:** Von den Leitlinien wurden 40 Primärstudien zitiert, davon wurden 33 Studien bereits bei den o. g. Sekundärpublikationen evaluiert, sechs weitere Primärstudien waren nach Prüfung von Titel und Abstract nicht relevant für die Fragestellung zur Untersuchung asymptomatischer Patienten mit nicht-ischämischer Genese [Metra et al. (COMET) 2005, Coats (CAPRICORN) 2001, Haim et al. (BIP) 1998, MRC Working Party 1992,

Dahlof et al. (STOP-Hypertension) 1991, Lund-Johansen 1981]⁴¹, die siebte Studie konnte unter der angegebenen Bibliographie nicht aufgefunden werden.

(7) **Zusammenfassung:** in keiner der zitierten Sekundär- oder Primärpublikationen fanden sich Belege dafür, dass *asymptomatische* Patienten *ohne* Infarkt-ereignis von einer β -Blockertherapie profitieren. Auch im neu identifizierten HTA-Bericht war diesbezüglich keine Zusatzinformation enthalten, da diese Patientengruppe in klinischen Studien bislang unterrepräsentiert war. Die Belege wurden für symptomatische Patienten und asymptomatische Patienten mit ischämischer Genese in ausreichender Form erbracht (Empfehlungen i und iv), gleiches gilt bezüglich der genannten Substanzen (Empfehlung v). Die konsensbasierten Empfehlungen (ii) und (vi) ließen sich aus den Vorgehensweisen in den zitierten Mortalitätsstudien ableiten, für die Empfehlung (iii) ist i.S.v. *good medical practice* keine höhergradige Evidenz erforderlich.

4.9 Weiterer Forschungsbedarf zur Entwicklung der Leitlinie

In Tabelle 4.19 wurde der Forschungsbedarf für die weitere Leitlinienentwicklung in strukturierter Form dargestellt.

Tab. 4.19: Fragestellungen mit Bedarf zur weiteren Evidenzgewinnung

Update-Recherchen	Prinzipiell <u>alle</u> Empfehlungen zur Aktualisierung des Evidenzgehalts
Systematische oder umfassende Recherchen	Konsistente Empfehlungen vom Typ (3): <ul style="list-style-type: none"> - Erstdiagnostik (diagnostische Wertigkeit von: Symptomen und klinischen Zeichen, klinisch-chemischer und apparativ-technischer Diagnostik) - Effektivitätsnachweise einer Lebensstilmodifikation - Wirksamkeit und Sicherheit: Impfungen, (AT₁-Blocker bei ACE-Hemmer-naiven Patienten[†]), Pharmakotherapie der diastolischen Herzinsuffizienz - Indikation zu oraler Antikoagulation bei Herzinsuffizienz und Sinusrhythmus (Wirksamkeit / Sicherheit) - Klinische Relevanz von Interaktionen zwischen ACE-Hemmern und Acetylsalicylsäure
	Inkonsistente Empfehlungen vom Typ (A): <ul style="list-style-type: none"> - Diagnostische Wertigkeit von BNP-/NT-proBNP-Tests - AT₁-Blocker-Therapie (Wirksamkeit und Sicherheit von AT₁-Blockern zusätzlich zu ACE-Hemmern und β-Rezeptorenblockern (Tripel-Therapie) / anstelle von β-Rezeptorenblockern bei deren Unverträglichkeit)*

⁴¹ CAPRICORN schloss nur Herzinsuffizienz-Patienten nach Infarkt ein, BIP nur NYHA-Klasse II+III mit ischämischer Genese, COMET untersuchte symptomatische Herzinsuffizienz-Patienten, Dahlof et al. sowie die MRC Working Party untersuchten Hypertoniepatienten und die Norwegian Multicenter Study berichtete Timolol-Effekte nach Infarkt.

(Fortsetzung Tabelle 4.19)

<p>Fokussierte Recherchen</p>	<p>Konsistente Empfehlungen vom Typ (2):</p> <ul style="list-style-type: none"> - Wirksamkeit (und Sicherheit) von β-Rezeptorenblockern bei asymptomatischen Patienten ohne Infarkt ereignis - (Wirksamkeit und) Sicherheit von Aldosteron-Antagonisten bei NYHA III-IV - (Wirksamkeit und) Sicherheit von Digitalis bei Herzinsuffizienz und tachykardem Vorhofflimmern - Wirksamkeit einer Kontrolle des arteriellen Hypertonus bei diastolischer Herzinsuffizienz
<p>Differenziertes Vorgehen in Abhängigkeit von den Ergebnissen der „Validierung“</p>	<p>Konsistente Empfehlungen vom Typ (4):</p> <ul style="list-style-type: none"> - Körperliches Training: in den Trefferlisten der Initialrecherche Identifikation einer hochwertigen Sekundärpublikation, die ausreichend Belege für Effektivität lieferte → ggf. Update-Recherchen erforderlich - Diuretika: Reevaluation der Evidenz erfolgt mit ausreichender Ursachenanalyse zur inkonsistenten Vergabe von Evidenzleveln, Identifikation weiterer Mortalitätsstudien unwahrscheinlich → ggf. Update-Recherchen, ggf. andere Outcomes wählen - Hydralazin/ISDN: keine primär hausärztliche Entscheidungssituation, fehlende Marktzulassung von Kombinationspräparaten in Deutschland → z. Z. keine weiteren Recherchen erforderlich - Unbedenklichkeit von langanflutenden Dihydroperidinen → weitere Recherchen zur therapeutischen Sicherheit erforderlich <p>Inkonsistente Empfehlungen vom Typ (B):</p> <ul style="list-style-type: none"> - Salz- und Flüssigkeitsrestriktion → weitere Recherchen zu Wirksamkeit und Sicherheit bei fehlender Evidenz - β-Rezeptorenblocker bei NYHA IV: ausreichende Belege in hochwertiger Sekundärpublikation identifiziert → keine weiteren Recherchen erforderlich - AT₁-Blocker bei ACE-Hemmer-Intoleranz: Evidenz rechtfertigt Empfehlung als Mittel 2. Wahl → weitere Recherchen s.o.* - Digitalis bei NYHA III-IV und Sinusrhythmus: Warnhinweis zur Anwendung bei Frauen konnte trotz bias-anfälligem Studiendesign nicht entkräftet werden → gezielte Recherche nach prospektiver qualitativ hochwertiger Studie zur Untersuchung geschlechtsspezifischer Unterschiede in der Wirksamkeit von Digitalis erforderlich, Recherchen zur therapeutischen Sicherheit - Kontraindikation von Kalziumantagonisten auch zur kurzzeitigen Anwendung in Notfallsituationen → weitere Recherchen zur therapeutischen Sicherheit erforderlich <p>Empfehlungen vom Typ (1), für die in der Validierung keine ausreichenden Belege in den zitierten Evidenzquellen gefunden wurden:</p> <ul style="list-style-type: none"> - ACE-Hemmer: Belege für alle Empfehlungen identifiziert, neu identifizierter HTA zeigte Wirksamkeit bei asymptomatischen Patienten mit/ohne Infarkt → ggf. Update-Recherchen zur Dosierung - β-Rezeptorenblocker bei NYHA II-IV sowie I nach Infarkt ereignis: ausreichende Belege vorhanden → ggf. Update-Recherchen

Diese umfangreiche Themenliste bildete die Grundlage eines formalen Priorisierungsverfahrens⁴², bei dem diese Themen eine Rangfolge erhielten. Ziel war es, die Liste von Forschungsthemen gemäß ihrer Bedeutung für den hausärztlichen Entscheidungskontext an die bestehenden Ressourcen anzupassen.

Beginnend mit der höchsten Priorität wurden sie in der zur Verfügung stehenden Zeit „abgearbeitet“. Fragestellungen, die nicht bearbeitet werden konnten, wurden für das Aktualisierungsverfahren der Leitlinie zurückgestellt. Auch wurde dabei der Umfang bzw. die Tiefe der weiteren Evidenzgewinnung festgelegt (systematische – umfassende – orientierende Recherchen).

Daher war es möglich, in der noch verbleibenden Bearbeitungszeit die hoch priorisierten Themen ‚Bewertung des BNP-/NT-proBNP-Tests‘ als *systematische* Evaluation und ‚AT₁-Blocker‘ sowie ‚kontraindizierte Pharmaka bei Herzinsuffizienz‘ in Form von umfassenden Bewertungen vorzunehmen. Eine Reihe weiterer hausärztlich relevanter Fragen wurde orientierend bearbeitet. Andere Fragen, wie z.B. ein Thema mittlerer Priorität aus dem Bereich der komplementär- und alternativmedizinischen Verfahren (Crataegus-Extrakte in der Therapie der Herzinsuffizienz) musste mangels Ressourcen für das nächste Leitlinien-Update zurückgestellt werden⁴³.

4.10 Zusammenfassung der Ergebnisse

Von den 52 in der *systematischen Leitlinienrecherche* identifizierten Dokumenten entsprachen 16 umfassende Leitlinien den *Ein- und Ausschlusskriterien* (Untersucherübereinstimmung $\kappa=0,95$). In der *Bewertung ihrer formalen methodischen Qualität* unter Einsatz des international validierten AGREE-Instruments wurde gezeigt, dass sich unter den berücksichtigten Leitlinien solche von hoher methodischer Qualität befanden – insbesondere auch in der Domäne 3 („Methodologische Exaktheit“) – und dass die Qualität über alle Domänen des Instruments sowie im Vergleich der Leitlinien untereinander stark variierte.

In der systematischen Analyse der berücksichtigten Leitliniendokumente wurde ein *Framework* – d.h. das strukturbildende Raster der Leitlinie – aus insgesamt 27 klinisch relevanten Fragen entwickelt, die von der zu entwickelnden Leitlinie

⁴² Vorgehensweise und Ergebnisse des Priorisierungsverfahrens: siehe Methodenreport²⁵.

⁴³ Die vergleichsweise hohe Priorität dieses phytopharmakologischen Themas ist in einem Beschluss des DEGAM-Arbeitskreises ‚Leitlinien‘ vom 6. November 1999 begründet. Darin heißt es, dass „der Arbeitskreis ... mehrheitlich der Auffassung (ist), dass NHV (Naturheilverfahren) in die zu erstellenden Leitlinien Eingang finden sollen.“ Der Beschluss geht darauf zurück, dass die Verfahren „weite Verbreitung und Anwendung ... beim Publikum (und) teils ohne innere Überzeugung bei Allgemeinärzten“ hätten und dass Hausärzte in der Lage sein müssten, ihre Patienten über „Wirksamkeit und Risiken“ dieser Behandlungsmöglichkeiten aufzuklären. [Brockmann 2004]

beantwortet werden sollten. Das Framework wurde in der Arbeitsgruppe auf den hausärztlichen Versorgungskontext abgestimmt.

Zu den Fragestellungen des Frameworks wurden *Daten* aus den berücksichtigten Leitlinien in insgesamt 20 Evidenztabelle in standardisiertem Berichtsformat *extrahiert*. Im Anschluss an die Datenextraktion ergab sich, dass drei dieser Themen für den hausärztlichen Versorgungskontext in Deutschland nicht relevant waren.

Aus den entscheidungsrelevanten Evidenztabelle wurden 35 spezifische, z.T. mehrgliedrige und komplexe Empfehlungen für die *Konsistenzanalyse* isoliert (Erstdiagnostik der Herzinsuffizienz: sechs, nicht-pharmakologische Therapie: vier, Pharmakotherapie der systolischen Herzinsuffizienz: 19, diastolische Herzinsuffizienz: sechs). Zusammenfassend wurden 25 konsistente und neun inkonsistente Empfehlungen identifiziert, eine weitere Empfehlung wurde nur von einer Leitlinie abgegeben und konnte deshalb nicht auf Konsistenz geprüft werden. Anhand der Typenklassifikation der (In-)Konsistenzen ergab sich, dass drei dieser Empfehlungen auf einer übereinstimmend als höchstwertig eingestuften Evidenz beruhten (Typ 1). Bei vier konsistenten Empfehlungen wurden wenige als hochwertig bewertete Studien als Beleg angeführt (Typ 2), 14 übereinstimmende Empfehlungen folgten Expertenkonsens und bei vier inhaltlich übereinstimmenden Empfehlungen wurden unterschiedliche Wertigkeiten der zugrundeliegenden Evidenz angegeben. Bei den drei bedeutsamen Inkonsistenzen (Typ A) zur diagnostischen Wertigkeit des BNP-/NT-proBNP-Tests und zum Wert der AT₁-Blocker (als Tripeltherapie und als Substitution bei β -Rezeptorenblockerintoleranz) wurden jeweils verschiedene Evidenzgrundlagen angeführt. Die sechs weniger bedeutsamen Inkonsistenzen (Typ B) rührten von einer kontroversen Bewertung vorliegender Studienevidenz durch eine Minderheit von Leitlinienentwicklergruppen her (Ausnahme: divergierende Empfehlungen zur Salz- und Flüssigkeitsrestriktion beruhten auf unterschiedlichem Expertenkonsens).

Basierend auf dem Framework und den Ergebnissen der Konsistenzanalyse wurde in der *Informationssynthese* ein Erstentwurf der zu entwickelnden Leitlinie erstellt.

Im nächsten Schritt – der *Validierung* – wurden die zitierten Belege für insgesamt 16, teilweise komplexe Empfehlungen re-evaluiert (Konsistenzen, Typ (1): n=3, Typ (2): n=4, Typ (4): n=4; Inkonsistenztyp (B): n=5, eine weitere weniger bedeutsame Konsistenz wurde nicht re-evaluiert, da sie auf Expertenkonsens beruhte), wobei die jeweils relevanten Arbeiten (14 systematische Reviews / Metaanalysen, sechs RCTs, eine post-hoc-Subgruppenanalyse) re-evaluiert und um zwei neu identifizierte HTA-Berichte und eine kollaborative Metaanalyse ergänzt wurden. Ausreichende Evidenzbelege fanden sich bei fünf Empfehlungen (z.B. zu β -Rezeptorenblockern bei symptomatischer systolischer Herzinsuffizienz), bei weiteren acht Empfehlungen fehlten Belege für Teilaussagen, u.a.

der Beleg für die Wirksamkeit von ACE-Hemmern bei asymptomatischen Patienten, der erst in einem neu identifizierten HTA-Report erbracht wurde, und der fehlende Beleg der Wirksamkeit von β -Rezeptorenblockern bei asymptomatischen Patienten ohne Infarkt ereignis, der auch in weiteren Evidenzrecherchen nicht ermittelt werden konnte. In drei Fällen wurde auf eine Reevaluation von Primärquellen verzichtet: zum körperlichen Training, da hier eine höherwertige Evidenzquelle in der initialen Recherche identifiziert wurde, zu Hydralazin/ISDN wegen fehlender Marktverfügbarkeit in Deutschland und zur oralen Antikoagulation bei Vorhofflimmern, da das Problem nur mittelbar das Leitlinienthema berührt. Im Ergebnis wurde der Erstentwurf im Wortlaut der Empfehlungen, in Evidenzleveln und im Stärkegrad der Empfehlungen korrigiert und um die zusammengefassten Evidenzbewertungen ergänzt. Darüber hinaus entstand eine umfangreiche Liste von spezifischen Forschungsfragen, die für eine weitere Evidenzgewinnung formuliert wurden und die einer formalen Priorisierung zugänglich waren (*strukturierter Forschungsbedarf*).

5 Diskussion

Zielsetzung der vorliegenden Arbeit war es, für die Neuentwicklung einer deutschsprachigen hausärztlichen Leitlinie zur Herzinsuffizienz einen Erstentwurf zu erarbeiten, der in einem transparenten und validen Verfahren (dem systematischen Leitlinienreview, SLR) bereits existierende Leitlinien auswertet. Vorgehensweisen der einfachen Kompilation oder lokalen Adaptation von Leitlinienempfehlungen erscheinen hierbei als fehleranfällig und unbefriedigend. Die Verfahrensweise des SLRs wurde ausgehend von den grundsätzlichen Arbeitsmethoden der evidenzbasierten Medizin und der grundlegenden Anforderung an die Validität in der Leitlinienentwicklung in dieser Arbeit neuentwickelt und erstmalig durchgeführt.

Hauptzweck des SLRs ist, einen der arbeits- und ressourcenaufwendigsten Schritte einer *de novo*-Leitlinienentwicklung (die Recherche und Bewertung der Primär-Evidenz) zu reduzieren, indem bereits vorliegende Leitlinien systematisch ausgewertet werden. Zudem soll die Methode implizite Normativität in Leitlinien transparent machen und strittige Fragen für einen weiteren Diskurs vor- bzw. aufbereiten. *Thema probandum* der Methode des SLR ist es deswegen, wie methodisch zuverlässig verhindert werden kann, dass Verzerrungen, kurzschlüssige Evidenzbewertungen oder Lücken aus den ausgewerteten Leitlinien ungeprüft in den eigenen Leitlinienentwurf 'eingeschleppt' werden. Umgekehrt hatte sich gleichzeitig bei der Darstellung der Ergebnisse gezeigt, dass ein SLR sogar einen Mehrwert über den Beitrag zur Entwicklung der Leitlinie hinaus erbringen kann, indem insbesondere durch die Methoden der Validierung und Konsistenzanalyse ungeprüfte Konsensaussagen sowie aktueller Forschungsbedarf identifiziert werden können.

Das nachfolgende Diskussionskapitel folgt dem bisherigen Aufbau der Arbeit und behandelt nacheinander den Bezugsrahmen der Leitlinienentwicklung, die Methodik des SLR in der theoretischen Herleitung und anhand spezifischer Ergebnisse der Anwendung des Verfahrens in der Entwicklung zur Herzinsuffizienz-Leitlinie.

5.1 Diskussion des Bezugsrahmens

Neuentwicklungen von methodisch und qualitativ angemessenen Leitlinien, insbesondere bei komplexen Themen, sind mit erheblichem personellen, zeitlichen und finanziellen Aufwand verbunden. Besonders aufwendige Arbeitsschritte sind die umfassende und zufriedenstellende Evidenzrecherche und Studienbewertung (*critical appraisal*), und mehr noch die für Leitlinien geforderten mehrdimensionalen Risiko-Nutzen-Bewertungen [AGREE 2003, 2003a]. Diese beschränken sich nicht nur auf die klinische Effektivität als klassisches EbM-Kriterium, sondern fordern komplexere Bewertungen von Nutzen – mit der Aufächerung von Ergebnisparametern unter Einbeziehung z.B. auch der Lebens-

qualität – sowie eine Darstellung verschiedener Behandlungsalternativen. Dabei sollen inhaltliche Angemessenheit, Klarheit der Darstellung sowie praktische Umsetzbarkeit der Leitlinienempfehlungen berücksichtigt werden [AGREE 2003, 2003a].

Unvermeidlich sind daher *Trade offs* zwischen Breite und Tiefe der Evidenzbewertung einzugehen – einerseits zulasten der Anzahl von Entscheidungssituationen, die damit abgedeckt werden, andererseits zulasten der Güte der Evidenzbewertung und der Risiko-Nutzen-Bewertung, die den formulierten Empfehlungen zugrundeliegen [Hadorn & Baker 1994].

Direkte Kostenangaben für die im Rahmen unseres Reviews ausgewerteten Leitlinien ließen sich nicht ermitteln; Indikatoren können aber z.B. die von der *Agency for Healthcare Policy and Research* durchschnittlich pro Leitlinie aufgewendete 1 Mio. US-\$ [American College of Cardiology's Foundation 2005] oder die Anzahl der Autoren des *National Institute of Clinical Excellence* (15 Autoren [NICE 2003]) und des *Department of Veterans Affairs & Veterans Health Administration* (14 Autoren [DVA & VHA 2002]) darstellen.

Die bereitgestellten Fördermittel zur Entwicklung der DEGAM-Leitlinie Herzinsuffizienz dagegen umfassten nur Personalmittel für 0,5 bis 0,75 wissenschaftliche Mitarbeiter (die Verfasserin der vorliegenden Arbeit) für einen Zeitraum von drei Jahren sowie Sachmittel. Ein wesentlich höherer Aufwand wäre weder für die DEGAM noch für den Projektförderer, das BMBF, darstellbar gewesen.

Da der Erstentwurf der Leitlinie ‚Herzinsuffizienz‘ inzwischen erfolgreich durch die weiteren Stufen des DEGAM-Konzepts (insbesondere fachinternen sowie externen Review, formalen Konsentierungsprozess mit anderen Fachgesellschaften und Praxistest)⁴⁴ geführt und im September 2006 von der DEGAM auch offiziell autorisiert wurde, ohne dass gravierende Mängel des Erstentwurfs offenbar geworden sind, kann plausibel gemacht werden, dass der Entwicklungsprozess des Erstentwurfs mit der beschriebenen Methodik *effektiv* und mit den vorgegebenen Mitteln *ressourcensparend* war⁴⁵.

5.2 Diskussion der Methode

Wie im einleitenden Hintergrundkapitel dargelegt, wurden Adaptationen existierender Leitlinien als Alternative zur aufwendigen *de novo*-Entwicklung vorge-

⁴⁴ Diese Schritte liegen außerhalb der Darstellung dieser Dissertation.

⁴⁵ Die Kurzfassung, gekürzte Versionen der beiden Teile der Langfassung, die Patienteninformation sowie der komplette Methodenreport sind im Internet unter der URL http://www.allgemeinmedizin.uni-frankfurt.de/forschung/herzinsuffizienz_internet.html erhältlich. Die ungekürzte Leitlinie kann beim Omikron-Verlag unter der URL <http://www.omikronverlag.de/catalog/3009/index.html> bezogen werden.

schlagen [Europarat 2002, NZGG 2001] und verschiedentlich durchgeführt [Van Tulder et al. 2006, Graham et al. 2005, Glasier et al. 2003, Voellinger et al. 2003, MacLeod et al. 2002, Wabitsch et al. 1998].

Zu Beginn der Diskussion der Methode soll der SLR daher zunächst mit dem derzeit ambitioniertesten Verfahren zur *transkulturellen* Leitlinienadaptation von Fervers et al. [2006] vergleichend diskutiert werden, um Parallelen und grundsätzliche Unterschiede beider Verfahren darzustellen und damit die ersten drei Forschungsfragen nach der Schrittfolge des SLR und deren Begründung, nach der Sicherstellung der internen Validität sowie zur Herstellung von Transparenz über Normativität von Leitlinienempfehlungen aus methodischer Sicht zu beantworten.

5.2.1 *Schrittfolge im systematischen Leitlinienreview im Vergleich zur transkulturellen Adaptation von Leitlinien*

Bei der lokalen Adaptation von Leitlinien sollen deren Empfehlungen an lokale Versorgungskontexte angepasst werden. Häufig werden dazu internationale Leitlinien ausgewählt, wobei oft weder deren Auswahl dargestellt, noch weitere Details zum Adaptationsprozess berichtet werden. Auch werden dabei zumeist die Evidenzgrundlagen der Empfehlungen weitgehend ungeprüft übernommen [Fervers et al. 2006]. Fervers et al. [2006] schlagen darum ein Adaptationskonzept vor, dessen Gemeinsamkeiten und Unterschiede zum hier vorgestellten systematischen Leitlinienreview in der nachfolgenden Darstellung (Abb. 5.1 auf der folgenden Seite) schematisch gegenübergestellt werden, wobei gestrichelte Verbindungslinien auf vergleichbare Verfahrensschritte hinweisen, die nachfolgend zuerst diskutiert werden.

5.2.1.1 *Ähnliche Vorgehensweisen in beiden Verfahren*

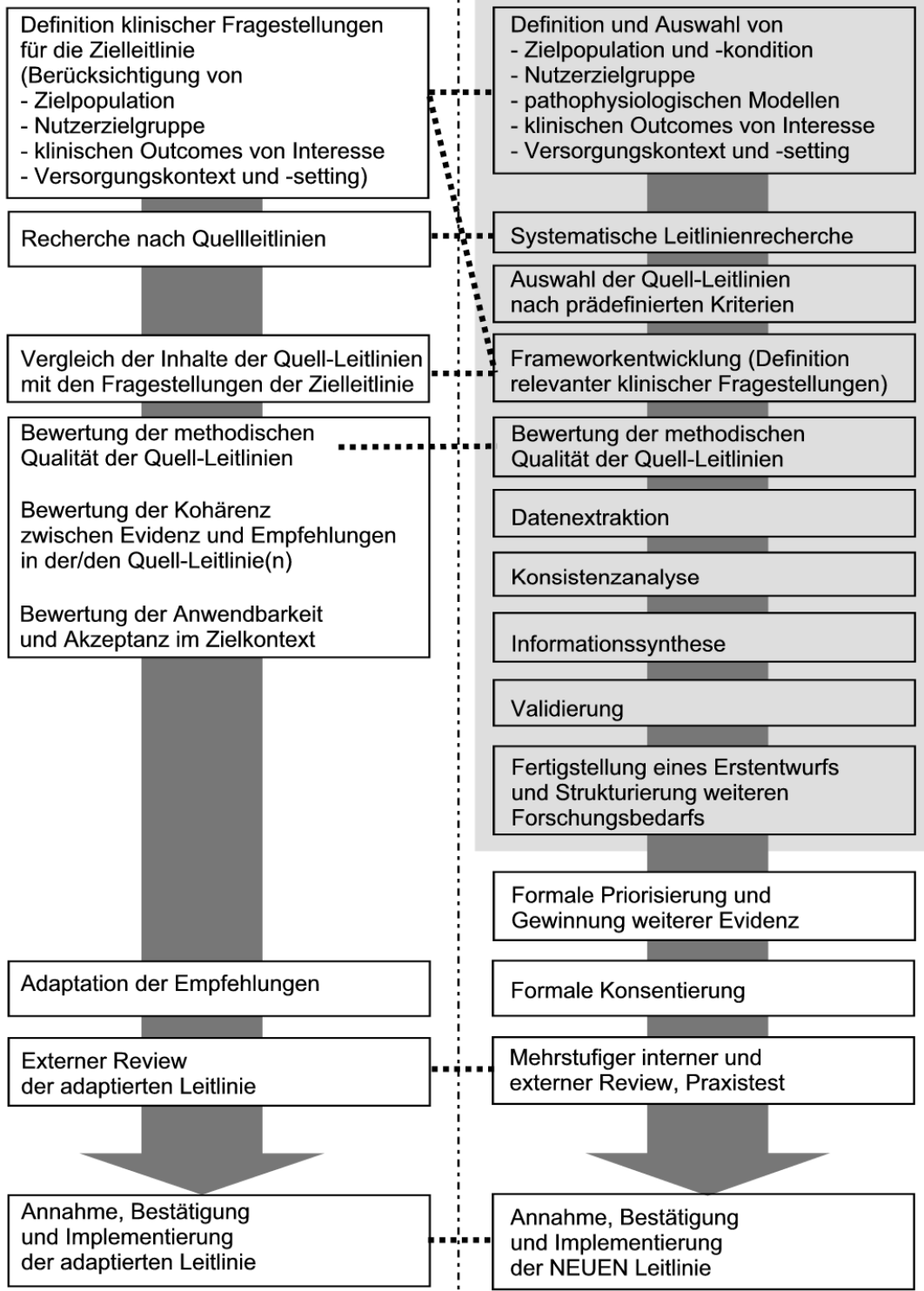
5.2.1.1.1 *Definitionsphase*

In beiden Verfahren werden zunächst wesentliche Prämissen, wie Zielpopulation und -kondition, Zielkontext mit Nutzern und Versorgungs-Setting sowie klinische Ergebnisparameter, die bei der Evidenzbewertung von Interesse sind, definiert.

Abb. 5.1: Gegenüberstellung von transkontextueller Adaptation und systematischem Leitlinienreview

**Transkontextuelle Leitlinienadaptation
[Fervers et al. 2006]**

**Systematischer Leitlinienreview
(vorliegende Arbeit, hellgrau hinterlegt)
als Bestandteil einer de novo-Entwicklung**



Abweichend von Fervers et al. [2006], die in dieser Phase eine Definition klinischer Fragestellungen *from scratch* – also als freie Konzeption ohne Vorlage – vorschlagen, um nachfolgend die Quellleitlinien dahingehend zu untersuchen, inwiefern deren Inhalte Antwort auf ihre Fragestellungen liefern, erfolgt die Konzeption klinischer Fragestellungen beim SLR als eigenständiger Prozessschritt, der Entwicklung des Frameworks (strukturbildendes Raster).

Die Framework-Entwicklung wurde als weichenstellender Prozess zur Vermeidung von Bias in der zu entwickelnden Leitlinie angesprochen [Margolis in Margolis & Cretin 1999], der sowohl die Themen auswählt, die von der Leitlinie behandelt werden sollen, als auch die Forschungsfragen formuliert, für die Evidenz gewonnen werden soll. Damit spielt dieser Schritt eine wichtige Rolle sowohl bei der Leitlinienentwicklung, als auch bei der späteren Implementierung, die ganz wesentlich davon abhängt, ob die Leitlinie die *entscheidungsrelevanten* klinischen Fragen beantwortet.

In der vorliegenden Arbeit wurden die Themen aus den berücksichtigten Leitlinien gewonnen, auf Relevanz für den Zielkontext in der internen Arbeitsgruppe geprüft und in späteren Entwicklungsstufen ggf. gemäß den Bedürfnissen breiterer Anwendergruppen (z.B. im Paneltest) sowie unter Einbeziehung weiterer Perspektiven (im Rahmen der interdisziplinären und multiprofessionellen Konsentierung unter Einschluss eines Patientenvertreters) modifiziert.

Damit wurde eine einheitliche Gestaltung des Frameworks nach klar explizierten Ordnungsprinzipien – d.h. eine Systematik – ermöglicht. Bloße Subjektivität, für die kleine Entwicklergruppen bei der Themenwahl anfällig sind [Ashton et al. 1999, Rogers 2002], wurde reduziert, da in die Entscheidung darüber, ob eine klinische Frage relevant ist, auch die Perspektiven anderer Entwicklergruppen einbezogen wurden.

Die Entwicklung des Frameworks war in allen Punkten transparent, ermöglichte den Einsatz von nachvollziehbaren Priorisierungsverfahren und wurde in weiteren Entwicklungsstufen der Leitlinie ergänzt⁴⁶.

⁴⁶ Für die DEGAM-Leitlinie ‚Herzinsuffizienz‘ wurden in Paneltest und Konsensusverfahren Themen ergänzt, die bislang von keiner der berücksichtigten Leitlinien bearbeitet wurden. Infolge des Paneltests wurden die Themen ‚akute Dekompensationen – Notfallversorgung‘, ‚Herztransplantation‘, ‚Weitere interventionelle Therapien‘ sowie ‚Palliativtherapie bei Patienten mit terminaler Herzinsuffizienz‘ in das Framework der Leitlinie aufgenommen und bearbeitet. Weitere wichtige Ergebnisse des Paneltests betrafen insbesondere die Praktikabilität der Leitlinie, wie z.B. die Aufteilung der Langfassung (Teil 1: praxisrelevanter Teil, Teil 2: ‚Evidenz und Rationale‘), die Einführung eines Stichwortverzeichnisses etc. Im Konsensusverfahren (Nominaler Gruppenprozess) wurden folgende neue bzw. inhaltlich erweiterte Themen als relevante Ergänzung des Frameworks gefordert und bearbeitet: die Frage nach einem praktikabel handhabbaren Risikoscore zur Prognoseeinschätzung beim individuellen Patienten, Fragen zur Endokarditis-Prophylaxe und zum Interaktionspotential von Digitalispräparaten.

5.2.1.1.2 Recherche nach Quelleitlinien

Auch Fervers et al. [2006] schlagen eine systematische Recherche nach allen relevanten Leitlinien zum Thema vor. Die von ihnen empfohlene Suchstrategie entspricht der in der vorliegenden Arbeit eingesetzten und ist darin begründet, dass Leitlinien nur selten in gelisteten Zeitschriften veröffentlicht werden und daher selten in elektronischen Datenbanken, wie MEDLINE, EMBASE etc. verschlagwortet werden [Feder et al. 1999, Van Tulder et al. 2006, Fervers et al. 2006]. Um dennoch *alle relevanten* Leitlinien zur chronischen Herzinsuffizienz zu identifizieren, sind aufwendige Handsuchen notwendig. Dazu wurden seitens der Verfasserin internetbasierte Quellen, wie Datenbanken von *Clearinghouses* und Homepages von internationalen Leitlinienanbietern gezielt berücksichtigt – darunter insbesondere (*Managed Care*- und andere) Organisationen im amerikanischen Gesundheitswesen mit umfangreichem eigenen Forschungs- und Entwicklungspotential, wie das *Department of Veterans Affairs & Veterans Health Administration* oder das *Institute for Clinical System Improvement*⁴⁷.

Wiederum *abweichend* von Fervers et al. [2006], die zunächst offen lassen, wie prinzipiell mit den aufgefundenen Leitlinien verfahren werden soll, und die die Auswahl der Leitlinien von deren Inhalt abhängig machen, werden die Leitlinien im SLR nach *a priori* festgelegten Kriterien, d.h. ohne vorherige Kenntnis des Inhalts ausgewählt.

Die im vornherein formulierten *Ein- und Ausschlusskriterien* sollten sicherstellen, dass in die Untersuchung Leitliniendokumente eingeschlossen wurden, die dem Zielkontext, der Zielpopulation und -kondition entsprachen und die aufgrund ihrer umfassenden Darstellung und konsequent explizierten Verknüpfungen zwischen Empfehlungen und Evidenz für das beschriebene Analyseverfahren geeignet waren. Die Auswahl wurde von zwei voneinander unabhängigen Reviewern getroffen, um einen *Selection Bias* zu vermeiden.

5.2.1.1.3 Bewertung der formalen Leitlinienqualität

Wie auch von Fervers et al. [2006] angeregt, wurden die berücksichtigten Leitlinien im SLR einer sorgfältigen Prüfung ihrer formalen, insbesondere methodischen Qualität unter Anwendung des AGREE-Instruments unterzogen. Dieses Instrument ist besonders geeignet, da es international – auch für Deutschland –

⁴⁷ Zum Recherchezeitraum (Anfang 2004) war die internationale Leitliniendatenbank des *Guideline International Network*, die im Dezember 2007 mehr als 4.800 Dokumente umfasste (URL: <http://www.g-i-n.net/index.cfm?fuseaction=about>, Zugriff vom 06.01.2008), noch im Aufbau begriffen. Zum Recherchezeitraum waren die umfassendsten Sammlungen die Leitlinienbibliothek des Ärztlichen Zentrums für Qualität in der Medizin (URL: <http://www.aezq.de/>) und die Sammlung des *National Guideline Clearinghouse* der USA (URL: <http://www.ngc.gov/>).

validiert und akzeptiert ist, sowie einfach und transparent in der Anwendung [Vlayen et al. 2005]. Anders als bei Fervers et al. [2006] zielte dieser Schritt im SLR nicht auf die abschließende Bewertung einer Leitlinie, um daraus Schlüsse für die Vertrauenswürdigkeit der Empfehlungen zu ziehen, sondern war darauf ausgelegt, sicherzustellen, dass sich unter allen berücksichtigten Leitlinien auch solche von hoher methodischer Qualität befanden. Dieser Unterschied zwischen beiden Verfahren ist bedeutsam, da die Prüfung der Angemessenheit von Leitlinienempfehlungen im SLR im Unterschied zu Fervers et al. [2006] ganz wesentlich auf deren *Inhalte* konzentriert war.

Weitere Parallelen im Vorgehen – wie Peer-Review-Verfahren und Autorisierung der Leitlinie – sind nicht mehr Gegenstand der vorliegenden Arbeit.

5.2.1.2 Grundlegende Unterschiede beider Verfahren

Fervers et al. [2006] sind sich bewusst, dass die Übertragbarkeit von Leitlinienempfehlungen ('*trans-kontextuelle*' Adaptation) ein voraussetzungsvoller Prozess ist. Ihre eigene Durchführung verbleibt aber auf der Ebene der Gewährleistung einer bloß formalen methodischen Qualität und übersieht die Kriterien *inhaltlicher Validität* von Leitlinienempfehlungen. Das ist darin zu begründen, dass sie zwar die formale methodische Qualität der Quelleitlinien bewerten, die Kohärenzprüfung zwischen Evidenz und Empfehlungen in der/den Quell-Leitlinie(n) jedoch erfolgt, ohne die von den jeweiligen Leitlinienautoren zitierten Evidenzbelege erneut zu evaluieren. Somit machen die Autoren zwar darauf aufmerksam, dass kontextuelle Faktoren einen ganz erheblichen Einfluss auf die Angemessenheit von Leitlinienempfehlungen haben können, können diesen mit ihrem Verfahren aber nicht wirksam begegnen.

Mit der Einführung der Schritte ‚Datenextraktion‘, ‚Konsistenzanalyse‘, ‚Analyse der Belegtheit‘ und der folgenden Informationssynthese geht der SLR über eine bloße Adaptation von Leitlinien weit hinaus, wie im Folgenden weiter diskutiert werden soll.

5.2.1.2.1 Datenextraktion

Die Datenextraktion hat in systematischen Reviewprozessen im wesentlichen drei Funktionen [Alderson et al. 2005]:

- (1) Sie ist eine visuelle Präsentation der eingeschlossenen Dokumente mit direkter Verbindung zu Forschungsfragen und geplanter Auswertung.
- (2) Sie fungiert als protokollarische Aufzeichnung für die Entscheidungsprozesse, die in einem Reviewverfahren ablaufen.
- (3) Sie ist die Datenquelle für weitere Analysen.

In einer systematischen Evaluation von *Leitlinien* gewinnt die Datenextraktion eine besondere Bedeutung, da hier beschreibende (und bewertende) Aussagen zur Evidenz und normative Aussagen (Handlungsempfehlungen) getrennt erfasst werden sollen, um sie in nachfolgenden Analyseschritten separat voneinander auf ihre Belegbarkeit bzw. ihre Angemessenheit im Zielkontext zu prüfen. Daher wurde für die Evidenztabellen ein Format gewählt, in dem die Ergebnisse der Evidenzbewertungen inkl. zitierter Quellen und zugeordneter Evidenzlevel sowie die darauf begründeten Handlungsempfehlungen mit den zugehörigen Stärkegraden notiert wurden.

Mit der Abfassung von Kurzberichten zu den berücksichtigten Leitlinien wurde einerseits das Ziel verfolgt, die Bewertung mit dem AGREE-Instrument nachvollziehbar zu dokumentieren, da diese nur durch die Verfasserin der vorliegenden Arbeit vorgenommen wurde. Andererseits wurden diese Daten in Zusammenschau mit der inhaltlichen Überprüfung benötigt (z.B. in Fragen des Zielkontexts und der Zielpopulation).

Dieser Verfahrensschritt war aufwendig – es entstanden 20 umfangreiche Tabellen, aber essentiell für die nachfolgenden Analyseschritte. Zudem wurde die Transparenz der Leitlinienentwicklung mit der Veröffentlichung dieser Evidenztabellen auch für künftige Nutzer erhöht.

5.2.1.2.2 Konsistenzanalyse

Wie in der vorliegenden Arbeit bereits angesprochen und von Vlayen et al. [2005] in einem systematischen Review über Instrumente zur Bewertung der formalen Leitlinienqualität gezeigt wurde, können diese keine Aussagen zu klinischem Inhalt und dessen Evidenzbasierung vornehmen [Vlayen et al. 2005]. Damit kommt der Konsistenzanalyse eine zentrale Bedeutung im systematischen Leitlinienreview zu, da sie – gemeinsam mit der Analyse der Belegtheit – eine *inhaltliche Prüfung* der eingeschlossenen Leitlinien ermöglicht.

In der Konsistenzanalyse wurden (1) Empfehlungen der berücksichtigten Leitlinien hinsichtlich ihrer *Formulierungen* und des vergebenen Stärkegrades der Empfehlung sowie (2) die zugehörigen *Evidenzbewertungen* inkl. angegebener Evidenzlevel systematisch miteinander verglichen, um *strittige Fragen* zu identifizieren. Diese können Ausdruck von inadäquatem methodischem Vorgehen, inhomogener oder wenig belastbarer Evidenz, wie auch von impliziter Normativität sein.

Wie bereits in vergleichenden Leitlinienuntersuchungen beobachtet [Campbell et al. 2006, Burgers et al. 2002, Silagy et al. 2001] und hier bestätigt wurde, besteht zwischen verschiedenen Leitlinien zum gleichen Thema trotz häufig reklamierter 'erschöpfender' Literaturrecherche häufig eine nur unzureichende Übereinstimmung in den identifizierten Evidenzquellen. Infolgedessen wurde im

Rahmen der Konsistenzanalyse auch die *Evidenzgrundlage der Empfehlungen verbreitert*, da verschiedene Evidenzbezüge zusammengeführt wurden.

Das für die Analyse entwickelte *Klassifikationsschema* ermöglichte darüber hinaus, dass Breite und Tiefe der Reevaluation bzw. Neugewinnung von Evidenz und mögliche Kompromisse zwischen diesen transparent dargestellt und begründet und unnötige Redundanzen vermieden werden können:

(1) Es wurden *konsistente Empfehlungen* identifiziert, für die in den berücksichtigten Leitlinien eine *ausreichend gute Evidenzlage* beschrieben wurde, d.h. es lagen von anderen Leitlinienautoren *konsistente Risiko-Nutzen-Bewertungen* zahlreicher hochwertiger Studien zu dieser Fragestellung vor. Die Kongruenz von Evidenz und Empfehlungen wurde nachfolgend in einer (verkürzten) Reevaluation der entscheidungsrelevanten Primärquellen (s. Validierung) geprüft. Damit wurden erhebliche Ressourcen für Recherchen und die Bewertung von Studien eingespart, die bereits (teilweise mehrfach) von anderen Leitlinienentwicklern durchgeführt wurden.

(2) Weiterhin wurden *konsistente Empfehlungen* identifiziert, die aber nur auf einer *geringen Zahl von hochwertigen Studien* beruhten. Diesem Bereich wurde eine eigene Kategorie zugewiesen, da bspw. qualitativ hochwertige RCTs zwar Informationen zur medizinischen *Wirksamkeit* von therapeutischen Interventionen liefern. Ihre Aussage zur *Sicherheit* dieser Intervention ist jedoch begrenzt, da sie zumeist keine ausreichend großen Studienpopulationen einschließen, um Risiken und deren Häufigkeit zu ermitteln [Banta & Luce 1993]⁴⁸. D.h. die Wahrscheinlichkeit, relevante Risiken aufzudecken, ist wesentlich geringer in Kategorie (2) als in Kategorie (1), in der zahlreiche Studien mit insgesamt hohen Patientenzahlen vorliegen.⁴⁹

(3) Diese Kategorie wurde definiert, da hier die *konsistenten Empfehlungen* der verschiedenen Leitlinienentwicklergruppen auf *unzureichender Evidenzgrundlage* beruhten und die Gefahr besteht, dass bloße Konsensusmeinungen ungeprüft weitergetragen werden. Diese Kategorie hatte zwei Konsequenzen in der

⁴⁸ Analoges gilt für diagnostische oder prognostische Fragestellungen, wobei dort z.T. andere Studiendesigns die hochwertigste Evidenz liefern [Sackett & Haynes 2002, Knottnerus et al. 2002].

⁴⁹ Wie aus den Ergebnissen des SLR zur chronischen Herzinsuffizienz hervorgeht, ist das Thema Sicherheit in Leitlinien deutlich ‚unterbelichtet‘. Das ist zum einen darin begründet, dass die Schwierigkeit der Bewertung von Sicherheitsaspekten zwar lange bekannt und Strategien dazu angemahnt wurden [Banta & Luce 1993], aber erst in jüngerer Zeit Evaluationskonzepte entwickelt werden [Busse et al. 2002] bzw. Arbeitsgruppen mit der Konzeptentwicklung betraut werden (z.B. die Neue Cochrane-Methoden-Gruppe: *Adverse Effects Methods Group* (AEMG), URL: <http://aemg.cochrane.org/en/index.html>). Zum anderen macht das für HTA entwickelte Konzept deutlich, dass die Bewertung von Sicherheitsaspekten ungleich ressourcenaufwendiger als die Bewertung von Wirksamkeitsaspekten ist [Busse et al. 2002].

weiteren Leitlinienentwicklung: erstens zog sie weitere Recherchen zur Gewinnung neuer Evidenz nach sich. Zweitens trug sie unter Umständen zu mehr Entscheidungssicherheit in einer konsensbasierten Expertenempfehlung bei, falls *keine* neue Evidenz gewonnen werden konnte, da die klinischen Erfahrungen und Einschätzungen anderer Autoren für die Formulierung einer neuen konsensbasierten Empfehlung berücksichtigt wurden.

(4) Die vierte Kategorie *konsistenter Empfehlungen* wurde für Empfehlungen gebildet, die zwar in der *Formulierung konsistent* waren, jedoch in den vergebenen *Evidenzleveln und Stärkegraden variierten*. Wie eine kritische Analyse von derzeit existierenden Graduierungssystemen gezeigt hatte, weisen diese eine niedrige Reproduzierbarkeit und erhebliche Mängel an diskriminatorischer Fähigkeit auf [Atkins et al. 2004]. Es war daher naheliegend, die Varianz der Graduierung zunächst als methodisches Problem zu deuten und den Versuch zu unternehmen, bestehende Konflikte anhand einer Reevaluation von Primärpublikationen auszuräumen.

A) Empfehlungen, deren *Formulierungen inhaltlich vollständig voneinander differierten*, wurden als *bedeutsame Inkonsistenzen* gewertet. Wie bereits gezeigt wurde, können diesen Inkonsistenzen vielfältige Ursachen zugrundeliegen, wie

- unterschiedliche Recherchezeiträume, bei denen jüngere Leitlinien Evidenz gewinnen konnten, die älteren noch nicht zugänglich war (insbesondere bei innovativen Technologien),
- inadäquate Methoden, infolge derer relevante Publikationen nicht erfasst oder nicht in die Bewertung eingeschlossen wurden oder deren Evidenzgehalt unangemessen bewertet wurde [Burgers et al. 2002, Silagy et al. 2001, Beck et al. 2000, Grant et al. 2000, Ravago et al. 2000, Ramsay et al. 1998, Thomson et al. 1998],
- Risiko-Nutzen-Bewertungen, die in den verschiedenen Leitlinienentwicklergruppen kontextbedingt als Ausdruck impliziter Normativität zur Formulierung unterschiedlicher Empfehlungen führten [Christiaens et al. 2004, Manna et al. 2003, Eisinger et al. 1999, Ramsay et al. 1998, Fahey et al. 1996].

Zur Lösung des Konfliktes ist im Idealfall zunächst *systematisch* gewonnene Evidenz⁵⁰ erforderlich, um eine unverzerrte Aussage über alle relevanten Studien hinweg zu erhalten – im Rahmen einer eigenen systematischen Technologiebewertung, sofern keine neueren, validen systematischen Reviews / HTA-

⁵⁰ Systematische Reviews stellen Informationen über die Effektivität einer Intervention zur Verfügung, indem sie alle relevanten Publikationen zu einer spezifischen Fragestellung identifizieren, kritisch bewerten und deren Ergebnisse qualitativ oder quantitativ (Metaanalyse) zusammenfassen, wobei die verwendeten Methoden nachvollziehbar dokumentiert werden. Oberstes Ziel ist es, möglichst unverzerrte Zusammenfassungen der besten verfügbaren Evidenz zu liefern. Damit sind sie insbesondere hilfreich bei Unsicherheit über Nutzen und Risiken von Interventionen. [Khan et al. 2001]

Berichte bei der Recherche identifiziert werden können. Auf dieser rationalen Entscheidungsgrundlage können dann – wiederum kontextspezifische – Empfehlungen formuliert und für den Nutzer transparent begründet werden.

B) Die Kategorie der weniger bedeutsamen Inkonsistenzen wurde gebildet für den Fall, dass eine *Mehrheit* von Leitlinien Empfehlungen enthielt, die inhaltlich sowie formal *konsistent* waren, zu denen sich nur einzelne bzw. eine deutliche *Minderheit* der eingeschlossenen Leitlinien *im Dissens* befanden. Ziel war zunächst eine Konfliktlösung durch Reevaluation der jeweils angeführten Evidenzquellen und ggf. eine Formulierung weiteren spezifischen Forschungsbedarfs. Auch hier wurde der anschließende Diskurs vorbereitet, indem die verschiedenen Aspekte der rationalen Entscheidungsgrundlage transparent dargestellt wurden.

5.2.1.2.3 Analyse der Belegtheit (Validierung)

Ziel dieses zweiten Teils der *inhaltlichen* Prüfung der Leitlinienempfehlungen war es, zunächst die *Validität* der von anderen Leitlinienentwicklergruppen vorgenommenen Evidenzbewertungen für Empfehlungen der Kategorien (1), (2), (4) und (B) anhand von Originalpublikationen zu verifizieren. Weiterhin wurde in diesem Analyseschritt untersucht, ob sich die jeweils zu prüfende Empfehlung kongruent zur verifizierten Evidenzaussage verhielt.

Für die Validierung ist zu erwarten, dass einerseits tatsächlich fehlerhafte Einschätzungen identifiziert werden können, wenn beispielsweise von Leitlinienautoren Aussagen zu (bedeutsamen) Subpopulationen getroffen werden, die in den zitierten Evidenzquellen nicht oder nicht ausreichend repräsentiert wurden. Andererseits ist davon auszugehen, dass Evidenzbewertungen für das hausärztliche Setting unter Umständen prinzipiell auch anders ausfallen können als für spezialärztliche / stationäre Settings. Das ist z.B. der Fall für Parameter des diagnostischen Testens, die in beiden Settings aufgrund unterschiedlicher Prävalenzen (und Vortestwahrscheinlichkeiten) eine unterschiedliche diagnostische Wertigkeit bedingen. Das kann z.B. dazu führen, dass ein Test mit guter diagnostischer Wertigkeit im stationären Bereich für den hausärztlichen Bereich unbrauchbar ist – was in einer hausärztlichen Leitlinie demzufolge zu diametralen Empfehlungen im Vgl. zu einer spezialärztlichen führt. Die hier angeführten Beispiele werden unten vertiefend diskutiert und belegt.

Um Redundanzen zu vermeiden, wurde in der Analyse von der belastbarsten Evidenz ausgegangen, die in den Leitlinien zitiert wurde, so dass zunächst die zitierten systematischen Reviews und Metaanalysen reevaluiert wurden – da diese die größte Entscheidungsrelevanz besitzen. Nur bei danach unbeantwort-

tet gebliebenen Fragen wurden weitere Primärstudien hinzugezogen⁵¹. In diese Phase wurden auch neu identifizierte HTA-Berichte (identifiziert in der initialen Leitlinienrecherche) einbezogen, da keine der berücksichtigten Leitlinien von dieser hochwertigen Evidenzquelle Gebrauch gemacht hatte.

Dieses Vorgehen ermöglichte es, die begrenzten Ressourcen auf Bereiche mit der höchsten Entscheidungs*unsicherheit* zu konzentrieren: die Kategorien (3) und (A), die in weitergehenden Recherchen zur Neugewinnung von Evidenz untersucht werden konnten.

5.2.1.2.4 Informationssynthese

Wegen der erheblichen Komplexität und des damit verbundenen Umfangs wurde bereits nach der Konsistenzanalyse eine erste Informationssynthese vorgenommen, indem die Ergebnisse der Konsistenzanalyse in einem vorläufigen ersten Entwurf zur Leitlinie zusammengefasst und diskutiert wurden. Unter Aufnahme der Ergebnisse der Validierung wurden ggf. erforderliche Korrekturen eingefügt und begründet. Dieses zweistufige Vorgehen hat sich im Nachhinein auch deshalb als vorteilhaft erwiesen, weil damit beide Verfahrensschritte klar voneinander getrennt präsentiert werden konnten, was eine hohe Transparenz zur Folge hatte.

5.2.2 Andere systematische Reviews von Leitlinien

Die hier vorgestellte Methode des SLR wurde von der Verfasserin der vorliegenden Arbeit selbstständig neuentwickelt und erstmalig angewandt. Eine nachträgliche Recherche nach anderen Arbeiten, in denen ebenfalls ein systematischer Review von Leitlinien berichtet wurde, identifizierte die folgenden:

- **Philips et al. [2004]** publizierten einen systematischen Review von Leitlinien und leitlinienähnlichen Dokumenten, die Empfehlungen zur Durchführung qualitativ hochwertiger Entscheidungsanalysen im Bereich HTA enthielten. Sie führten dazu eine systematische Recherche durch, wählten die zu analysierenden Dokumente nach a priori festgelegten Kriterien aus, bewerteten die Dokumente und synthetisierten daraus eine neue Anleitung zur Durchführung von Entscheidungsanalysen. Das Verfahren zur Bewertung der Dokumente wurde von den Autoren weder näher beschrieben noch begründet, insbesondere wurden keine Angaben zur Prüfung der inhaltlichen Angemessenheit der Empfehlungen gemacht. Es ließen sich lediglich inhalt-

⁵¹ Diesem Ansatz folgen auch andere Leitlinienentwickler in etablierten Programmen, die zunächst eine Recherche nach systematischen Reviews und Metaanalysen anstellten, bevor sie bei spezifischen Fragestellungen oder wenn keine systematischen Übersichten zu einem Thema erhältlich waren, Recherchen nach Primärstudien vornahmen, wie z.B. ACC/AHA [2001], NICE [2003] und *American College of Chest Physicians* [McCrorry et al. 2003].

liche Parallelen im Vorgehen hinsichtlich von Recherche und Selektion der Dokumente feststellen. Auch war das Ziel der Autoren nicht die Entwicklung einer klinischen Leitlinie sondern die Entwicklung eines institutionellen Standards zur Anfertigung von Entscheidungsanalysen für das *National Coordinating Centre for Health Technology Assessment*, einer HTA-Organisation des *National Health Service* in Großbritannien [Philips et al. 2004].

- In vergleichbarer Weise gingen **Larsen et al. [2007]** vor, die einen systematischen Review von Leitlinien zur Prävention von hitzebedingten Erkrankungen im Breitensport durchführten. Auch sie beschränkten sich auf eine systematische Suche, einen expliziten Ein- bzw. Ausschluss von Leitlinien und auf die Prüfung von deren formaler Qualität mittels AGREE-Instrument. Die in den Leitlinien enthaltenen Empfehlungen wurden zwar vergleichend analysiert, dann jedoch ohne Prüfung der inhaltlichen Angemessenheit und der berichteten Evidenzgrundlage synthetisiert [Larsen et al. 2007].

Beide Arbeiten unterschieden sich also grundlegend von der hier vorgestellten Methode insofern, als gerade die Prüfung der inhaltlichen Angemessenheit von Empfehlungen unterblieb.

5.3 Diskussion der Ergebnisse der Anwendung am Beispiel der chronischen Herzinsuffizienz

In diesem Abschnitt wird diskutiert, wie die (interne) Validität der zu entwickelnden Herzinsuffizienz-Leitlinie in der Vorgehensweise des SLR sichergestellt wurde (3. Forschungsfrage). Aus der Diskussion der Methode haben sich dazu drei Bereiche herauskristallisiert, die im Folgenden anhand der praktischen Anwendung des Verfahrens (u.a. in Fallbeispielen) näher betrachtet werden:

1. Verfahrensschritte, welche die *Voraussetzungen für eine möglichst unverzerrte Analyse* liefern, wie Recherche, Selektion der berücksichtigten Leitlinien und Entwicklung des Frameworks (die Datenextraktion wurde oben bereits diskutiert)
2. Verfahrensschritte zur *Prüfung der formalen methodischen Qualität* der berücksichtigten Leitlinien
3. Verfahrensschritte zur *inhaltlichen Prüfung der Leitlinienempfehlungen*, wie Konsistenzanalyse, Analyse der Belegtheit und abschließende Informationssynthese

5.3.1 Voraussetzungen für eine unverzerrte Analyse

5.3.1.1 Systematische Leitlinienrecherche

Wegen der oben geschilderten Probleme der unzureichenden Erfassung publizierter Leitlinien in elektronischen Datenbanken besteht bei systematischen Leitlinienrecherchen potentiell die Gefahr, relevante Leitlinien zu übersehen. Ein Vergleich mit dem zwischenzeitlich publizierten Leitlinien-Clearingbericht "Herzinsuffizienz" des Ärztlichen Zentrums für Qualität in der Medizin [ÄZQ 2005] zeigte jedoch, dass dort keine zusätzlichen Leitlinien identifiziert wurden, die den Suchkriterien entsprachen.

5.3.1.2 Auswahl der Leitlinien

Die Ein- und Ausschlusskriterien wurden *a priori*, d.h. vor Recherchebeginn und ohne vorherige Kenntnis von Leitlinieninhalten formuliert. In der Anwendung dieser Kriterien zeigte sich eine exzellente Übereinstimmung beider unabhängigen Untersucher in der Kappa-Statistik, so dass ein *Selection Bias* in der vorliegenden Arbeit eher unwahrscheinlich erscheint.

5.3.1.3 Frameworkentwicklung

In der Analyse der berücksichtigten Leitlinien wurden Fragen zu zahlreichen Entscheidungssituationen in der Versorgung der chronischen Herzinsuffizienz identifiziert, die in diesem Umfang mit den zur Verfügung stehenden Ressourcen nicht bearbeitet werden konnten. Da diese Leitlinien teilweise auch auf andere Zielgruppen, wie z.B. Spezialisten in ambulanter oder stationärer Versorgung gerichtet waren, konnten im ersten Schritt Themen ohne *Entscheidungsrelevanz für den hausärztlichen Versorgungsbereich* ausgeschlossen werden. Die verbleibenden Fragestellungen übertrafen jedoch weiterhin die vorhandenen Möglichkeiten, so dass im Abstimmungsprozess einige Fragen von geringerer Priorität für spätere Leitlinien-Updates zurückgestellt bzw. andere Fragen deutlich enger gefasst werden mussten. Im Ergebnis entstand ein bearbeitbares Framework aus Themen und spezifischen Forschungsfragen, das bis zur Fertigstellung der Leitlinie nur noch geringfügig verändert wurde (z.B. wurde im Konsensprozess die Frage nach einem praktikablen Risikovorhersage-Score ergänzt, die von anderen Leitlinien nicht bearbeitet wurde).

Dennoch war das Framework umfassend, was auch aus einem Vergleich mit einer US-amerikanischen Leitlinie deutlich wird: die mit etwa einer Million US-Dollar entwickelte Leitlinie der AHCPR [1996] enthielt 37 spezifische Handlungsempfehlungen [Shekelle et al. 2001], und das für die DEGAM-Leitlinie mit

deutlich geringeren Ressourcen entwickelte Framework resultierte in mehr als 44 Empfehlungen⁵².

Aktuell nicht bearbeitete Fragen wurden in einen Pool gestellt, der in späteren Aktualisierungsprozessen der Leitlinie hinsichtlich seiner Relevanz und Dringlichkeit erneut geprüft wird. Somit unterstützt die hier vorgestellte Frameworkentwicklung einen kontinuierlichen Entwicklungs- und Aktualisierungsprozess von Leitlinien und ist gleichzeitig systematisch und transparent.

5.3.2 Prüfung der formalen methodischen Qualität

Unter den berücksichtigten Leitlinien befanden sich solche mit einer guten Qualität, wobei diese zwischen den einzelnen Leitlinien, wie auch im Vergleich zwischen den Domänen stark variierte. Zur gleichen Einschätzung gelangten bereits Kulig et al. [2003] sowie später (nach Abschluss dieses Arbeitsschritts der vorliegenden Arbeit) die Fokusgruppe des Leitlinien-Clearingberichts des ÄZQ [2005]. Ein direkter Vergleich der Ergebnisse ist nur eingeschränkt möglich, da von den jeweiligen Autoren verschiedene Instrumente zur Bewertung der methodischen Qualität eingesetzt wurden, die zwar jeweils multiple Dimensionen der Leitlinienqualität abbilden, die sich hinsichtlich ihrer Items jedoch erheblich unterscheiden. Am besten war das Instrument des ÄZQ [2005] mit dem hier verwendeten AGREE-Instrument vergleichbar. Kulig et al. [2003] arbeiteten mit zwei Instrumenten, die sich von den vorgenannten stärker unterschieden. Einen näherungsweisen Vergleich mit Hilfe standardisierter Scores zeigt Tabelle 5.1 auf der folgenden Seite.

Drei Leitlinien bzw. leitlinienähnliche Dokumente wurden von keiner der anderen Arbeiten berücksichtigt [DieM 2003/2004, LLGH 2003, UWH 2001], und es ergaben sich daher 13 Vergleichsmöglichkeiten. Daraus wird ersichtlich, dass einige der in der vorliegenden Arbeit berücksichtigten Leitlinien auch vom ÄZQ [ÄZQ 2005] als qualitativ hochwertig eingeschätzt wurden. Im Trend war die Bewertung in der vorliegenden Arbeit deutlich konservativer: 10/13 Bewertungen waren schlechter als die des ÄZQ. Nur in 3/13 Leitlinien fiel die Bewertung in der vorliegenden Arbeit geringfügig (um drei bis vier Prozentpunkte) höher aus als beim ÄZQ [ACC/AHA 2001, Duodecim 2004, ICSI 2003 vs. 2004]. Eine dieser drei Leitlinien [ACC/AHA 2001] wurde auch von Kulig et al. [2003] deutlich besser bewertet als vom ÄZQ [2005], und in den beiden anderen Fällen

⁵² Den Experten des Nominalen Gruppenprozesses zur Konsentierung der DEGAM-Leitlinie Herzinsuffizienz wurde eine Liste aus 44 Items, die z.T. mehr als eine spezifische Empfehlung enthielten, zur Abstimmung vorgelegt. Da nach Abschluss des systematischen Leitlinienreviews bis zur Konsentierung nur geringfügige Veränderungen vorgenommen wurden, steht diese Anzahl für die Zahl spezifischer Empfehlungen, die zur Beantwortung des hier entwickelten Frameworks erarbeitet wurden. (Siehe auch Methodenreport zur DEGAM-Leitlinie Herzinsuffizienz, vgl. Fußnote ²⁵)

[Duodecim 2004, ICSI 2003 vs. 2004] lag eine eingeschränkte Vergleichbarkeit der Dokumente vor: vom ÄZQ wurden zwei Leitlinien zu unterschiedlichen Themen (Rechts- und Linksherzinsuffizienz) gemeinsam bewertet [Duodecim 2004, Duodecim 2002], und eine Leitlinie wurde in einer späteren Version evaluiert [ICSI 2004] – letztere unterscheidet sich jedoch nur unwesentlich von der von uns bewerteten [ICSI 2003], da der Entwicklung ein etabliertes Leitlinienprogramm zugrundelag.

Tab. 5.1: Vergleich verschiedener Bewertungen der Leitlinienqualität

	Muth Medianer Zielerreichungsgrad [†]	ÄZQ [#] Gesamtpunktwert (absolut: 0 - 40)	ÄZQ Gesamtpunktwert standardisiert ^{††}	Kulig et al. [‡] NGC - Gesamtsumme (absolut: 0 - 17)	Kulig et al. NGC - Gesamtsumme standardisiert ^{††}	Kulig et al. [‡] AQUMED – Gesamtsumme (absolut: 0 - 44)	Kulig et al. AQUMED – Gesamtsumme, standardisiert ^{††}
ACC/AHA 2001	0,48	18	0,45	15,5	0,91	22	0,50
AKDAE 2001	0,27	26	0,65	n.b.	/	n.b.	/
CCS 2001, 2002/2003	0,33	21	0,53	n.b.	/	n.b.	/
DGK 2001	0,19	16	0,40	n.b.	/	n.b.	/
DleM 2003/2004	0,35	n.b.	n.b.	n.b.	/	n.b.	/
Duodecim 2004	0,39	14	0,35	n.b.	/	n.b.	/
DVA 2002	0,63	27	0,68	n.b.	/	n.b.	/
ESC 2001/2002	0,32	19	0,48	11,5	0,68	17	0,39
ISCI 2003 vs. 2004 [*]	0,70	25	0,63	n.b.	/	n.b.	/
LLGH 2003	0,58	n.b.	n.b.	n.b.	/	n.b.	/
NHF/Austr. & CSANZ 2002	0,31	23	0,58	n.b.	/	n.b.	/
NHF/NZ 2001	0,35	23	0,58	n.b.	/	n.b.	/
NICE 2003	0,71	33	0,83	n.b.	/	n.b.	/
OPOT 2000	0,31	22	0,55	n.b.	/	n.b.	/
UM 2001	0,33	16	0,40	n.b.	/	n.b.	/
UWH 2001	0,21	n.b.	n.b.	n.b.	/	n.b.	/

[†]Werte aus [ÄZQ 2005], [‡]Werte aus [Kulig et al. 2003], ^{*}siehe Text

Zusammenfassend hat die fehlende Zweitbewertung der Leitlinienqualität nicht zu einer Überschätzung der Leitlinienqualität geführt, und es wurden – übereinstimmend zu anderen Bewertungen – qualitativ hochwertige Leitlinien in den systematischen Leitlinienreview eingeschlossen. Generell wurde eher ein kon-

servativer Trend zur Unterbewertung erkennbar, der jedoch keine weiteren Konsequenzen hatte.

5.3.3 Prüfung der inhaltlichen Angemessenheit

5.3.3.1 Konsistenzanalyse

Die Konsistenzanalyse der vorliegenden Arbeit fand überwiegend konsistente Empfehlungen zur Versorgung der chronischen Herzinsuffizienz in den berücksichtigten Leitlinien, wie bereits von Kulig et al. [2003] gezeigt wurde. Es werden nun die Ergebnisse der Konsistenzanalyse von Typ-(3)-Konsistenzen und Typ-(A)-Inkonsistenzen diskutiert, die in der weiteren Leitlinienentwicklung umfassende Recherchen zur Neugewinnung von Evidenz nach sich zogen. Die Typen (1), (2), (4) und (B) werden anhand von Fallbeispielen gemeinsam mit den Ergebnissen der Analyse der Belegtheit anschließend diskutiert.

5.3.3.1.1 Typ-(3)-Konsistenzen

Es fiel auf, dass die „Grauzonen klinischen Handelns“ [Naylor 1995], d.h. der Anteil Empfehlungen, der sich auf Expertenkonsens oder auf Primärstudien mit hoher Anfälligkeit für designbedingte Verzerrungen abstützte, beträchtlich war: insgesamt 45% aller Empfehlungen beruhten auf diesem Evidenzlevel⁵³, davon waren mehr als 90% konsistent (Typ 3) und machten damit mehr als die Hälfte aller Konsistenzen aus. Diese „Grauzonen“ betrafen insbesondere die klinische Diagnostik auf der Basis von Symptomen und klinischen Zeichen und die nicht-pharmakologische Therapie, also Bereiche, in denen in der Regel kein industrielles Sponsoring für Studien erfolgt.

Gerade diese Bereiche sind jedoch von hoher Bedeutung für die hausärztliche Versorgung: Anamnese und körperliche Untersuchung führen in fast 90% zur abschließenden Diagnose [Peterson et al. 1992], und die Vor- und Nachtestwahrscheinlichkeiten diagnostischer Tests – auch klinisch-chemischer und technisch apparativer – fallen aufgrund des niedrig-prävalenten Versorgungskontexts nach dem Bayes-Theorem deutlich geringer aus [Fagan 1975]. Auch die nicht-pharmakologische Therapie mit ihren Beratungsinhalten zur Lebensstilmodifikation, Impfprophylaxe usw. ist traditionell eng mit hausärztlicher Tätigkeit verknüpft und sollte demzufolge von einer Leitlinie für dieses Setting so evidenzbasiert wie möglich behandelt werden.

⁵³ Damit ist die Herzinsuffizienz nicht untypisch: “Proof of effectiveness is lacking for most of what is done in medicine. Clinical practice would come to a halt if practitioners could only perform tests and treatments that had been proven effective in clinical trials.” [Woolf et al 1996].

Das Fehlen von Objektivität und Evidenzbasierung sind die Hauptkritikpunkte an konsensbasierten Empfehlungen [Mulrow & Cook 1998] und berechtigen die Forderung nach umfassenden Recherchen zur Gewinnung neuer Evidenz. Ob darüber hinaus durch den SLR die Formulierung „besserer“ Konsensurteile ermöglicht wird, weil die Einschätzungen verschiedener internationaler Arbeitsgruppen systematisch analysiert und im Konsensprozess berücksichtigt werden, lässt sich nicht sicher einschätzen. Auch wenn verschiedene Leitlinienautoren scheinbar unabhängig voneinander zu konsistenten Empfehlungen gelangten, kann der Einfluss international angesehener Fachgesellschaften, der für verschiedene Leitlinien gezeigt wurde [Burgers et al. 2002, Littlejohns et al. 1999], nicht ausgeschlossen werden. So waren im vorliegenden SLR die Leitlinien der ACC/AHA und der ESC den anderen Leitlinienautoren bekannt, wurden zumeist jedoch ohne klare Bezüge zu den jeweiligen Leitlinienempfehlungen zitiert.

Nach der Diskussion von Typ (3)-Konsistenzen werden im folgenden Abschnitt die bedeutsamen Inkonsistenzen (Typ A) an einem besonders facettenreichen Fallbeispiel diskutiert – der Erstdiagnostik der Herzinsuffizienz mit Hilfe von Tests auf *Brain Natriuretic Peptides*. Andere Typ-A-Inkonsistenzen lieferten darüber hinaus keine neuen Aspekte für die Diskussion.

5.3.3.1.2 Typ A-Inkonsistenzen, Fallbeispiel BNP-/NT-proBNP-Test

Die Empfehlungen der berücksichtigten Leitlinien differierten hier vollständig – von „nicht für den Routineeinsatz empfohlen“ bis hin zu „Anwendung bei allen Patienten mit Verdacht auf Herzinsuffizienz empfohlen“. Die Diskussion wird sinnvollerweise *ex post* geführt, d.h. auf der Basis der Ergebnisse der später durchgeführten systematischen Evaluation dieser Technologie⁵⁴. Deren Ergeb-

⁵⁴ Systematische Evaluation der Tests auf *Brain Natriuretic Peptides*: Die systematische Recherche identifizierte zur diagnostischen Wertigkeit von BNP-/NT-proBNP-Tests für die Diagnostik der chronischen Herzinsuffizienz einen HTA-Bericht über 20 Primärstudien [Craig et al. 2005], einen Kurz-HTA ohne Einschluss zusätzlicher Studien [AHFMR 2005], einen systematischen Review über 20 Primärstudien [Doust et al. 2004], außerdem eine diagnostische Primärstudie [Alehagen et al. 2003], die nicht Bestandteil systematischer Bewertungen war. Weitere identifizierte Primärstudien waren bereits in den genannten Sekundärpublikationen enthalten. Die Sekundärstudien wiesen eine verkürzte Berichtsqualität auf, die methodische Qualität bei Doust et al. [2004] war gut, bei Craig et al. [2005] fielen einige Mängel auf: ein Wechsel zwischen verschiedenen Metaanalysetechniken wurde nicht begründet, Heterogenitätstests wurden nicht diskutiert, und Sensitivitätsanalysen zur Prüfung der Robustheit der Ergebnisse wurden nicht durchgeführt. Die interne Validität der in den Sekundärpublikationen eingeschlossenen Studien wurde von den Autoren als angemessen eingeschätzt. Die übergroße Mehrzahl der Primärstudien wurde an Hochrisikokollektiven durchgeführt, was die externe Validität der Ergebnisse für die Primärversorgung einschränkt. Wenige Studien im hausärztlichen Kontext zeigten in der Metaanalyse bei Doust et al. [2004] eine signifikante statistische Heterogenität – weshalb das ermittelte gewichtete Mittel auf reinem Zufall beruhen kann – und in der im hausärztlichen Setting durchgeführten Studie von Ale-

nisse führten zur Formulierung einer Negativempfehlung des Tests im Konsensusprozess der DEGAM-Leitlinie. Mögliche Ursachen für die Inkonsistenz zu diesem Thema werden nun in einer durch das Modell (vgl. Abb. 1.4) vorgegebenen Gliederung diskutiert:

Potentielle Einflussfaktoren auf die Formulierung von Empfehlungen seitens der Leitlinienentwickler(gruppe):

Methodische Angemessenheit: Leitlinien, die eine positive Empfehlung abgaben, begründeten diese auf einer Auswahl diagnostischer Studien – darunter eine Leitlinie, die ihre Empfehlung auf eine einzelne diagnostische Studie abstützte [Duodecim 2004] – eine *systematische* Evaluation wurde von keiner berichtet. Es wurden vielmehr die ermutigenden Ergebnisse zur diagnostischen Wertigkeit der Tests berichtet, ohne die oftmals schlechte methodische Qualität der Studien und designbedingte Bias zu diskutieren, die potentiell zur Überschätzung der Testwertigkeit führen können (wie z.B. Studien an Patienten der Sekundär- und Tertiärversorgung, die einen potentiellen *Spectrum Bias*⁵⁵ aufweisen, wenn ihre Ergebnisse auf die Primärversorgung übertragen werden [Knottnerus et al. 2002, Irwig et al. 2002]). Auch wurde eine zentrale Frage diagnostischen Testens außer Acht gelassen: Tests mit kontinuierlichen Testergebnissen benötigen einen Schwellenwert, ober- bzw. unterhalb dessen ein Wert als pathologisch gilt; die valide Testschwelle ist bei BNP-/NT-proBNP-Tests jedoch weiterhin unklar [Craig et al. 2005]. Darüber hinaus fehlen prognostische Studien zur Abschätzung des Effekts einer Testanwendung auf die gesundheitsbezogenen Outcomes von Patienten ebenso, wie die Implikationen dieses Tests auf die Gesundheitsversorgung [Battaglia et al. 2006]. Bereits Sackett und andere wiesen darauf hin, dass eine Testevaluation ohne Berücksichtigung des klinischen Kontexts nur begrenzten Wert hat [Sackett & Haynes 2002, Knottnerus et al. 2002].

Zusammensetzung und Interaktion – Einstellung zu Innovationen: BNP-Tests sind innovative Technologien. Banta & Luce [1993] wiesen darauf hin, dass in der frühen Diffusionsphase einer Technologie nur begrenzte Informationen über ihren Nutzen erhältlich sind, da nur Medikamente und biologisch aktive Substanzen vor ihrer Marktzulassung *systematisch* geprüft werden. Erste publizierte Studien berichten häufig ermutigende Ergebnisse und sollten daher

hagen et al. [2003] zeigte sich eine diagnostische Wertigkeit, die als ‚wasteful‘ [Linn 2005] bezeichnet werden musste. Zusammenfassend war keine ausreichende Evidenz vorhanden, um den Einsatz dieses Tests im hausärztlichen Setting zu begründen, was in zwei nachfolgend publizierten systematischen Reviews bestätigt wurde [Davenport et al. 2005, Battaglia et al. 2006] (s. Darstellung Teil 2 der Langfassung der DEGAM-LL S. 40ff. URL: http://www.allgemeinmedizin.uni-frankfurt.de/forschung/herzinsuffizienz_internet.html).

⁵⁵ *Spectrum Bias*: Die Sensitivität eines Tests wird eher überschätzt, wenn der Test an Populationen mit mehr und schwerer erkrankten Individuen entwickelt wurde, als die Zielpopulation aufweist. [Knottnerus et al. 2002, Irwig et al. 2002]

einer konservativen Bewertung unterzogen werden [Banta & Luce 1993]. Wenn Leitlinienautoren einen diagnostischen Test zu einem sehr frühen Technologiestadium evaluieren und bereits weitreichende Empfehlungen für einen *Routineeinsatz* in der Gesundheitsversorgung abgeben, handeln sie wenig konservativ. Mögliche Erklärung hierfür können Einstellungen zu Innovationen sowie die Zusammensetzung der Entwicklergruppe sein: Ärzte, die besser qualifiziert sind und sich im Gemeinwesen engagieren (beide Merkmale treffen auch für Ärzte in der Leitlinienentwicklung zu), tendieren eher dazu, ‚*early adopters*‘ zu sein [Banta & Luce 1993], und Spezialisten sind im allgemeinen weniger konservativ eingestellt als Generalisten [Banta & Luce 1993, Naylor 1995].

Potentielle Interessenkonflikte: Da nicht alle Leitlinienentwicklergruppen adäquate Informationen zur redaktionellen Unabhängigkeit bereitgestellt haben, können potentielle Interessenkonflikte – hier zu Diagnostikaherstellern – nicht ausgeschlossen werden. Eine starke Verflechtung zwischen Leitlinienentwicklern und der pharmazeutischen Industrie wurde kürzlich demonstriert [Taylor & Giles 2005, Choudhry et al. 2002], und die erhebliche Einflussnahme von Diagnostikaherstellern auf Wissenschaft und Forschung wurde bereits für andere Tests gezeigt [Broadstock 2000].

Potentielle Einflussfaktoren auf die Formulierung von Empfehlungen aus Gesundheitssystem und Gesellschaft:

Kostenerstattung: Der Umgang mit diagnostischen Tests ist in Europa uneinheitlich und nicht unproblematisch, da eine gesamteuropäische Marktzulassung ohne den Nachweis der Effektivität des Tests erfolgt [Wörz 2003]: die *In Vitro Diagnostic Directive* [IVDD 1998] fordert lediglich das Einhalten bestimmter Parameter technischer Konformität, prinzipieller Eignung, Qualität und Sicherheit. Gleichzeitig unterliegt die Kostenerstattung nationalstaatlicher Jurisdiktion [Wörz 2003], was zu unterschiedlichen Zugangsvoraussetzungen führt. In einigen europäischen Staaten werden BNP-Tests von den Kostenträgern finanziert (z.B. Großbritannien, Finnland [NICE 2003, Duodecim 2004]), in Deutschland waren sie zum Zeitpunkt der vorliegenden Untersuchung IGeL-Leistungen⁵⁶.

Verfügbarkeit und Zugang: Eine Leitlinie [Duodecim 2004] empfahl die Routinewanwendung dieses Tests in der Erstdiagnostik der Herzinsuffizienz explizit vor dem Hintergrund eines eingeschränkten Zugangs zur Echokardiographie. Bei anderen Leitlinien können Restriktionen *vermutet* werden, die den Zugang zur vergleichsweise kostenintensiven Echokardiographie begrenzen sollen, ins-

⁵⁶ IgeL – Individuelle Gesundheits-Leistungen; Nach pers. Komm. mit Dr. med. D. Sunder-Plassmann, BKK-BV, vom Oktober 2007 wurde zu Tests auf *Brain Natriuretic Peptides* unter den Vertragspartnern des Bundes-Manteltarifvertrags (Kassen, KBV) eine neue EBM-Ziffer vereinbart, die ab 01.01.2008 abrechenbar wird. Damit werden BNP-Tests dann auch in Deutschland durch die Gesetzlichen Krankenkassen finanziert.

besondere bei Leitlinien, die von Kostenträgern oder den damit in Verbindung stehenden Einrichtungen entwickelt wurden [z.B. NICE 2003].

Anhand dieses Fallbeispiels wird deutlich, dass im Zusammenhang mit bedeutenden Inkonsistenzen in Leitlinienempfehlungen eine Reihe von Aspekten diskutiert werden können, die Ausdruck mangelnder methodischer Qualität aber auch Ausdruck kontextueller Einflüsse im Sinne von impliziter Normativität sein können. Mit der einzigen Ausnahme der Duodecim-Leitlinie, in der dieser Test mit mangelndem Zugang zu höherwertigen Verfahren begründet wurde, wurden von den befürwortenden Leitlinien keine anderen Beweggründe als die Evidenzbezüge angeführt. Darunter befand sich auch die NICE-Leitlinie, die in der Bewertung der formalen methodischen Qualität – sowohl in der Bewertung durch das ÄZQ als auch durch die Verfasserin der vorliegenden Arbeit – als besonders hochwertig eingestuft wurde. Die Identifikation als *strittige Frage*, die in Zusammenschau mit normativen Aspekten im Diskurs erörtert werden sollte, wurde also nicht im Rahmen der Prüfung der formalen Qualität offenbar, sondern erst durch die Konsistenzanalyse.

5.3.3.2 Analyse der Belegtheit

Zunächst wurden die angeführten Evidenzbelege nach ihrer Härte gewichtet, so dass nur diejenigen Publikationen neu bewertet wurden, deren Ergebnisse die größte *Entscheidungsrelevanz* hatten. Dabei wurden jeweils *alle systematischen Übersichten* berücksichtigt, die *Outcomeparameter von Interesse* berichteten und einen *Zugewinn an Information* darstellten. Erst wenn in deren Ergebnissen keine ausreichenden Belege gefunden werden konnten, wurden Primärstudien des für die Fragestellung belastbarsten Designs (in der Regel RCTs) reevaluiert. Damit sollte einerseits dem geforderten systematischen Ansatz der Evidenzbewertung [Shekelle et al. 1999] Rechnung getragen werden. Andererseits konnten durch dieses Vorgehen in hohem Maße Redundanzen im Literaturreview vermieden werden: von den insgesamt mehr als 300 (vgl. Tab. 4.12)⁵⁷ zitierten und teilweise bereits mehrfach bewerteten Evidenzquellen zu den Empfehlungen der berücksichtigten Leitlinien mussten ‚nur‘ 21 Primärquellen reevaluiert werden. Die übrigen waren deshalb nicht relevant,

- weil sie entweder veraltet waren: es wurde z.B. ein systematisches Review zu ACE-Hemmern nicht erneut bewertet [z.B. Garg & Yusuf 1995], weil ein SR neueren Erscheinungsdatums dieselben und darüber hinaus zusätzliche Studien berücksichtigt hatte [z.B. Eccles et al. 1998],

⁵⁷ Die Diskrepanz zur Gesamtanzahl zitierter Referenzen in den berücksichtigten Leitlinien (s. Tab. 4.2) von z.B. 573 [ACC/AHA 2001] ist darin begründet, dass hier nur diejenigen Quellen betrachtet werden, die als Evidenzbezug für die zu untersuchenden Fragestellungen angeführt wurden.

- weil sie über Surrogatendpunkte berichteten: z.B. wurde ein SR nicht erneut bewertet, der die Effekte von ACE-Hemmern auf die endotheliale Dysfunktion untersuchte [Mancini 2000], weil andere SRs zu Gesamtmortalität und weiteren klinischen Endpunkten berichteten [u.a. Neals et al. 2000],
- weil höherwertige Evidenz vorlag: Primärstudien wurden nur dann bewertet, wenn aus der Reevaluation von SRs noch offene Fragen resultierten, wie z.B. zu ACE-Hemmer-Dosierungen [NETWORK 1998, Packer et al. 1999, Nanas et al. 2000]; 42 weitere RCTs (vgl. Tab. 4.12) wurden nicht erneut bewertet, weil sie Bestandteil der SRs [u.a. Eccles et al. 1998, Flather et al. 2000] waren.

Die erneute Bewertung verfolgte das Ziel, neben der (formalen) Überprüfung der Evidenzbelege in einem *critical appraisal* auch die Qualität der Evidenz zu untersuchen. Denn bereits Hadorn et al. [1996] wiesen darauf hin, dass die Untersuchung auf potentielle *Biases* für die Formulierung von Leitlinienempfehlungen von zentraler Bedeutung ist⁵⁸. Insgesamt fiel in der vorliegenden Arbeit auf, dass bei der Mehrzahl der neu bewerteten Sekundärpublikationen methodische Mängel auftraten, die eine Überschätzung der Effekte wahrscheinlich machen, also potentiell überoptimistische Ergebnisse produzieren, wie in den folgenden Fallbeispielen diskutiert wird.

5.3.3.2.1 Typ-1-Konsistenzen, Fallbeispiel ACE-Hemmer

Die neu bewerteten systematischen Reviews und Metaanalysen zeigten eine breit abgesicherte Evidenzbasis für die Wirksamkeit von ACE-Hemmern bei *symptomatischer* chronischer systolischer Herzinsuffizienz zur Senkung von Mortalität, Morbidität, Symptomen, Verbesserung der Belastungstoleranz und Lebensqualität. In dieser Hinsicht konnten die in den Leitlinien angeführten Evidenzbewertungen nachvollzogen werden, und die darauf gegründeten Empfehlungen waren aus den Ergebnissen ableitbar. Nicht nachvollzogen werden konnten diese Evidenzbewertungen für *asymptomatische* Patienten: keine der systematischen Übersichten und Metaanalysen hatte diese Frage untersucht, die zitierten Primärstudien zeigten zwar eine Reduktion von Symptomen und eine im Trend geringere Hospitalisierungsrate, Effekte auf die Mortalität konnten jedoch nicht gezeigt werden. Insofern war die Verallgemeinerung einiger Leitlinien, ACE-Hemmer würden positive Effekte bei *allen NYHA-Klassen* auf *alle* oben genannten Outcomeparameter zeigen, nicht zulässig. Die in diesem Zusammenhang vergebenen höchsten Evidenzlevel und Stärkegrade der Empfehlungen waren darum teilweise fehlerhaft. Für diese Subpopulation (asymptoma-

⁵⁸ "The underlying question guideline panels must ask is, 'When and how should a particular piece of evidence influence what we recommend in terms of appropriate care?' The answer should be, more or less, 'To the extent the therapeutic effect implied by the piece of evidence is likely to be a true effect, and not just an artefact or fluke.'..." [Hadorn et al. 1996]

tische chronische systolische Herzinsuffizienz) wurde jedoch in einem neu identifizierten HTA-Bericht eine Metaanalyse berichtet, die mortalitätssenkende Effekte bei Männern nachweisen konnte, Frauen waren in den untersuchten Primärstudien unterrepräsentiert. Die Reevaluation führte hier also zu einer spezifizierten Aussage darüber, für welche Subpopulationen welche Outcomeverbesserungen tatsächlich beobachtet wurden und für welche anderen Subpopulationen die Empfehlungen ggf. nur auf Analogieschlüsse begründet werden können.

Darüber hinaus führte die Reevaluation zur kritischen Auseinandersetzung mit *quantitativen* Risiko-Nutzen-Angaben, wie z.B. NNTs oder *Risk Ratios*. Shekelle et al. [2003] ermittelten eine Vielzahl verschiedener *Risk Ratios* in Abhängigkeit vom Mortalitätsrisiko der eingeschlossenen Studienpopulationen. Diese erfordern den Einsatz von Nomogrammen, um für einen Patienten mit bestimmten Risikomerkmale eine zutreffende Risikoabschätzung vorzunehmen. Damit wird die Praktikabilität solcher Angaben in klinischen Leitlinien bereits erschwert. Eine weitere grundsätzliche Überlegung führte dazu, in der Herzinsuffizienz-Leitlinie *auf quantitative Angaben weitgehend zu verzichten*: bereits Cowie et al. [2000] bemerkten gegenüber epidemiologischen Studien zur Herzinsuffizienz eine auffällig geringere Mortalität in Therapiestudien – sowohl in den Interventions- als auch in den *Kontrollgruppen*! Sie diskutierten Selektionseffekte als Ursache für diese Diskrepanz: randomisierte kontrollierte Studien schließen häufiger Patienten mit günstiger Prognose ein, welche die erste kritische Phase der Herzinsuffizienz bereits überlebt haben (*„natural survivors“*), deren Alter deutlich geringer ist als das durchschnittliche Alter von Herzinsuffizienzpatienten in der Bevölkerung unter Risiko, und die bei bestehenden Komorbiditäten zumeist ausgeschlossen werden usw. [Cowie et al. 2000]. D.h., auch wenn ACE-Hemmer-Studien konsistent positive Effekte auf Überleben, Hospitalisierung etc. zeigen konnten [Eccles et al. 1998, Flather et al. 2000, Lubsen et al. 1996, Neal et al. 2000, Shekelle et al. 2003] und damit die *Richtung der Effekte* mit großer Wahrscheinlichkeit für eine Wirksamkeit von ACE-Hemmern auf o.g. Outcomeparameter spricht, so ist doch die *Stärke der Effekte* möglicherweise überschätzt und eine valide Kalkulation von quantitativen Angaben zur Risiko-Nutzen-Bewertung derzeit nicht möglich.

5.3.3.2.2 Typ-2-Konsistenzen, Fallbeispiel Aldosteron-Antagonisten

Durch die Reevaluation wurde deutlich, dass sich die Empfehlungen aller berücksichtigten Leitlinien zum Einsatz von Aldosteron-Antagonisten für zwei Substanzen (Spironolacton und Eplerenon) auf jeweils eine Mortalitätsstudie (RALES resp. EPHESUS) bezogen, wobei die Zielpopulation bei EPHESUS nur postinfarzielle Herzinsuffizienzen einschloss. Neben methodischen Mängeln der Primärstudien (u.a. die große Zahl von Zentren, mangelnde Berichtsqualität und industrielle Förderung ohne einen plausiblen Ausschluss einer Einflussnahme durch den Förderer) fehlten zu einer umfassenden Risiko-Nutzen-Bewertung

dieser Therapie insbesondere Untersuchungen zu Sicherheitsaspekten, wie z.B. zur Gefahr einer progredienten Niereninsuffizienz und / oder Hyperkaliaemie unter der Kombinationstherapie von ACE-Hemmern und Aldosteron-Antagonisten. Dazu wurde in nachfolgenden Recherchen eine Beobachtungsstudie identifiziert [Juurlink et al. 2004]. Die unter Studienbedingungen nur selten beobachtete Komplikation einer Hyperkaliaemie stellt nach deren Ergebnis möglicherweise unter Alltagsbedingungen eine erhebliche Einschränkung der Therapiesicherheit dar. Juurlink et al. [2004] beschrieben in ihrer Zeitreihenanalyse eine deutliche Zunahme des Verschreibungsverhaltens von Spironolacton bei kanadischen Ärzten nach der Publikation von RALES. Im gleichen Zeitraum wurde nahezu eine Verdreifachung der Hyperkaliämierate beobachtet. Zeitreihenanalysen sind in hohem Maße anfällig für *Confounding* [Linden et al. 2003, Eccles et al. 2003], so dass diese Beobachtungen nicht als Kausalzusammenhang gewertet werden dürfen. Sie rechtfertigen jedoch die Forderungen nach strenger Indikation (nur Patientengruppen, die einen nachweislichen Benefit gezeigt haben), Vermeidung inadäquat hoher Dosierungen, engmaschiger Kontrolle von Nierenfunktion und Kaliumspiegeln bei Patienten, die ACE-Hemmer und Aldosteron-Antagonisten erhalten – und die besondere Vorsicht bei älteren Patienten [Juurlink et al. 2004]. Die zielgerichteten Recherchen nach Behandlungsrisiken auf der Basis der Ergebnisse des systematischen Leitlinienreviews führten hier zu praxisrelevanten Ergänzungen der Empfehlungen über die Indikationsstellung zur Therapie und das Monitoring der mit Aldosteron-Antagonisten behandelten Patienten.

5.3.3.2.3 Typ-4-Konsistenzen, Fallbeispiel Diuretikatherapie

Bei Diuretika liegt ein grundsätzliches Problem vor: sie gehörten bereits zur klinisch etablierten Therapie, bevor qualitativ hochwertige placebokontrollierte Mortalitätsstudien vorlagen. Derartige Studien sind in heutiger Zeit nicht mehr durchführbar, weil es unethisch wäre, Patienten diese Therapie vorzuenthalten, da sie als einzige Substanzen geeignet sind, den Volumenhaushalt zu kontrollieren. Somit ist aus heutiger Sicht nicht mehr zweifelsfrei zu klären, ob sie für sich allein einen mortalitätsreduzierenden und progressionshemmenden oder nur symptomatische Effekte aufweisen. Dieses Problem wurde von den meisten Leitlinien thematisiert, und eine alleinige Therapie mit Diuretika wurde daher nicht empfohlen.

In der Reevaluation der einzigen Metaanalyse zu dieser Frage [Faris et al. 2002], die aufgrund ihres Publikationsdatums nur fünf der 16 berücksichtigten Leitlinien zugänglich war und die bei drei der fünf zitiert wurde, wurde ein typisches Problem im Umgang mit Publikationen deutlich, die ein hohes Risiko für methodisch bedingte Verzerrungen aufweisen: formal lieferte die Publikation eine Metaanalyse über mehrere RCTs. Die Studien zeigten jedoch eine hohe klinische Heterogenität in nahezu allen Parametern, die nach Greenland [1987] vor der Entscheidung, ob sich quantitative Analysemethoden für eine Informati-

onssynthese eignen, geprüft werden sollten⁵⁹. Die in publizierten Metaanalysen häufig anzutreffende Argumentation, dass *keine statistische Heterogenität nachweisbar* war und darum auch keine Heterogenität anzunehmen sei, ist falsch. „Statistische Heterogenitätstests können als Warnhinweis mit hoher Spezifität aber geringer Sensitivität dienen: niedrige p-Werte indizieren, dass Heterogenität besteht und nicht ignoriert werden kann, hohe p-Werte zeigen jedoch nicht, dass Heterogenität sicher vernachlässigt werden kann.“ (Übersetzung d.Verf.⁶⁰) [Greenland 1987] In diesem Sinne wäre bei Faris et al. [2002] eine qualitative Informationssynthese vorzuziehen gewesen, zumindest wären Sensitivitätsanalysen zur Prüfung der Robustheit der Ergebnisse zu fordern. Diese erheblichen methodischen Mängel wurden bei keiner der drei Leitlinien offen diskutiert, zwei Leitlinien verwendeten den höchsten Evidenzlevel [DieM 2003, LLGH 2003], eine Leitlinie bewertete diese Evidenz mit „C“, ohne die Gründe für die offensichtliche Herabstufung zu benennen [NICE 2003].

Bei den übrigen 13 Leitlinien, welche die Arbeit von Faris et al. [2002] nicht zitiert hatten, war die Vergabe der Evidenzlevel und Stärkegrade der Empfehlungen ebenfalls inkonsistent: zwei Leitlinien vergaben den zweithöchsten Evidenzlevel (also mindestens auf hochwertiger Studienevidenz mit allenfalls geringen Mängeln beruhend), ohne eine einzige klinische Studie zu zitieren [NHF/NZ 2001, OPOT 2000], andere blieben zumindest formal konsequent, indem sie keine Literaturbewertungen vornahmen und Expertenkonsens als Evidenzbezug angaben.

Aus der inhaltlichen Prüfung im systematischen Leitlinienreview bestätigte sich für einige Leitlinien die fehlerhafte Verwendung des jeweiligen Graduierungssystems, was die Einordnung als Typ (4)-Konsistenz rechtfertigte. Die Schwierigkeit der Zuweisung von Evidenzleveln und Stärkegraden der Empfehlung wurde für den Konsensprozess der zu entwickelnden Leitlinie transparent aufbereitet.

⁵⁹ Greenland schlug vor, (1) die Spezifikation der Studienvariablen (*Outcome*, Exposition, potentielle *Confounder*, Intermediärouitcomes [Surrogate], *Effect Modifier*), (2) die Quantifizierung der Effekte (wie wurden Outcomes gemessen?), (3) die Charakteristika der Studienpopulation und (4) die Follow-up-Zeiten zu berücksichtigen. Auch wenn nicht für alle diese Variablen eine befriedigende Homogenitätsannahme zu treffen sei, so solle geprüft werden, ob die daraus erwachsenden Heterogenitäten gegenüber anderen Variationsquellen als vernachlässigbar gering eingestuft werden können. [Greenland 1987]

⁶⁰ „A statistical test of homogeneity can serve as a warning light with high specificity but low sensitivity: Small p values indicate that the heterogeneity should not be ignored, but large p values do not indicate that it can be safely ignored.“ [Greenland 1987]

5.3.3.2.4 Typ-B-Inkonsistenzen, Fallbeispiel 1: β -Rezeptorenblocker bei asymptomatischen Patienten ohne Infarktereignis

In der Reevaluation der zitierten Evidenzbezüge konnten keine Publikationen ausgemacht werden, die Belege für eine Wirksamkeit dieser Medikamente bei dieser speziellen Subpopulation erbrachten, da diese in den betreffenden Studien unterrepräsentiert waren. Die teilweise höchsten vergebenen Evidenzlevel in den Leitlinien waren demnach nicht korrekt.

Das Problem, dass interessierende Subpopulationen in klinischen Studien unterrepräsentiert sind, ist nicht selten und in der Formulierung von Leitlinienempfehlungen nicht unproblematisch. Die Vergabe von Evidenzleveln ist davon jedoch nicht betroffen und sollte so erfolgen, dass klar nachvollziehbar bleibt, für welche Populationen es Studienevidenz gibt, und bei welchen Subgruppen die Empfehlungen auf nicht systematisch untersuchter Evidenz (z.B. Expertenkonsens) beruhen. Im vorliegenden Beispiel beruhte die Empfehlung auf Expertenkonsens über plausible pathophysiologische Zusammenhänge zwischen asymptomatischer und symptomatischer Herzinsuffizienz nicht-ischämischer Genese (Analogieschluss) – was bei einzelnen Leitlinien zwar im Text expliziert wurde, sich jedoch nicht in der Vergabe der Evidenzlevel niederschlug.

Auf der Grundlage des SLR und nachfolgender Recherchen, die keine neue Studienevidenz identifizierten, formulierten die Experten im Nominalen Gruppenprozess zur DEGAM-Leitlinie keine Empfehlung für eine β -Blockertherapie bei asymptomatischer Herzinsuffizienz nicht-ischämischer Genese⁶¹ und begründeten ihren Konsensentscheid mit der Maxime des *primum nil nocere*, d.h. keinem Patienten ein potentiell nebenwirkungsträchtiges Medikament zu verordnen, für welches kein ausreichender Wirksamkeitsnachweis vorliegt.

Dieses Beispiel zeigt eindrucksvoll, dass Empfehlungen trotz identischer Evidenzbasis diametral ausfallen können und bei der Formulierung kontextuelle Faktoren – insbesondere in der Abwägung von Risiken und Nutzen und der abschließenden Einschätzung der Angemessenheit – eine bedeutsame Rolle spielen. Der normative Hintergrund für die abschließende Einschätzung wurde hier jedoch klar expliziert.

5.3.3.2.4 Typ-B-Inkonsistenzen, Fallbeispiel 2: AT_1 -Blocker bei ACE-Hemmer-Intoleranz

Die Reevaluation des systematischen Reviews mit darin enthaltener Metaanalyse zu AT_1 -Blockern wies auf ein weiteres methodisches Problem hin: die Bewertung von Metaanalyse-Ergebnissen. Einige Autoren werten die Metaanalyse nach dem *random effects model* als konservativere Methode für alle $\tau > 0$, da sie

⁶¹ Vgl. Protokoll des Nominalen Gruppenprozesses im Methodenreport (s. Fußnote²⁸).

grundsätzlich davon ausgeht, dass die Behandlungseffekte eines individuellen Patienten um einen durchschnittlichen Therapieeffekt (als Normalverteilung um einen Mittelwert mit Varianz τ) variieren und damit die sogenannte *between-study-variance* berücksichtigen [Deeks et al. 2001, Glasziou et al. 2001]. Andere Autoren lehnen metaanalytische Ansätze generell ab, wenn sich die Ergebnisse von verschiedenen Studien nicht in einem einzigen wahren Effektschätzer (*fixed effect*) kombinieren lassen und konnten zeigen, dass *random effects models* nicht grundsätzlich konservativere Effektschätzer liefern [Poole & Greenland 1999, Greenland 1994]. Für die Auswahl der Metaanalysetechnik gibt es demnach keine allgemein akzeptierten (EbM-)Regeln, wobei sich die Effektschätzer beider Methoden bei sehr kleinem τ einander annähern und *fixed effects models* im allgemeinen schmalere Konfidenzintervalle liefern [Deeks et al. 2001, Greenland 1994]. Vor diesem Hintergrund lieferte die neu bewertete Metaanalyse zur Wirksamkeit von AT₁-Blockern keine sichere Evidenzaussage, da im Vergleich zwischen AT₁-Blockern und Placebo nur im *fixed effects model* durch das schmalere Konfidenzintervall ein gerade eben signifikanter Vorteil für die Morbidität gezeigt werden konnte, nicht jedoch im *random effects model*. Diese Konstellation weist bei ausreichender statistischer Power der Analyse darauf hin, dass der *tatsächliche* Therapieeffekt noch *unterhalb* des a priori in der Analyse als *klinisch relevant* angesehenen Effekts liegt.

Spätere Recherchen zur therapeutischen Wirksamkeit von AT₁-Blockern bestätigten die unsichere Evidenzlage, als in einer kumulativen Metaanalyse gezeigt wurde, dass die mortalitätsreduzierenden Effekte von AT₁-Blockern abwechselnd nachweisbar und nicht nachweisbar waren, je nach *zeitlicher* Grenze des Publikationseinschlusses [Lee et al. 2004].

Die Ergebnisse des systematischen Leitlinienreviews und der weiteren Recherchen sowie die Hinweise, dass Kreuzreaktivität zwischen ACE-Hemmern und AT₁-Blockern hinsichtlich assoziierter Angioödeme gezeigt werden konnte [DVA & VHA 2002], führte zur Konsentierung konservativer Empfehlungen (keine AT₁-Blocker bei ACE-Hemmer-naiven Patienten, AT₁-Blocker als Mittel 2. Wahl bei ACE-Hemmer-Husten und nicht bei asymptomatischen Patienten, kein ambulanter Wechsel auf AT₁-Blocker bei ACE-Hemmer-assoziiertem Angioödem, keine Indikation zur Tripel-Therapie in der Primärversorgung etc.).

5.4 Limitationen der vorliegenden Arbeit

5.4.1 Limitation der Publikationssprache der Leitlinien

In der Recherche nach relevanten Leitlinien wurde eine Begrenzung auf deutsch- und englischsprachige Leitlinien notwendig, da für die Untersuchung von anderssprachigen Leitlinien keine Ressourcen zur Verfügung standen. Diese ist aber bezüglich potentieller Verzerrungen nicht unproblematisch: für *systematische Reviews* wurde auf die Gefahr eines *Language Bias* bei Limitation

auf bspw. englischsprachige Primärstudien hingewiesen, der potentiell überoptimistische Ergebnisse hervorrufen kann [Egger et al. 2001]. Bei einem *systematischen Leitlinienreview* kann diese sprachliche Begrenzung dazu führen, dass vermehrt auf Leitlinien aus dem Commonwealth (insbesondere Großbritannien, Kanada, Neuseeland, Australien) und den USA fokussiert wird – mit deren Kultur und Tradition, den Wertvorstellungen, Gesundheitssystemen, der Ausrichtung der Gesundheitspolitik, der Erreichbarkeit und Zugangsgerechtigkeit von Gesundheitsleistungen sowie deren Finanzierung. Die genannten normativen Aspekte beeinflussen die Formulierung von Leitlinienempfehlungen, wie im Modell (vgl. Abb. 1.4) dargestellt und oben bereits diskutiert wurde.

In der vorliegenden Arbeit wurden Leitlinien aus insgesamt sieben Ländern (Australien, Deutschland, Finnland, Großbritannien, Kanada, Neuseeland und USA) und eine gesamteuropäische Leitlinie eingeschlossen. Diese Länder unterscheiden sich hinsichtlich der Struktur ihrer Gesundheitssysteme, der Ausrichtung ihrer Gesundheitspolitik, der Erreichbarkeit und Zugangsgerechtigkeit von Gesundheitsleistungen sowie deren Finanzierung deutlich voneinander [Busse 2002, Busse & Schlette 2003, Van Doorslaer & Masseria 2004, Busse et al. 2005]. Bei einem systematischen Vergleich aller Leitlinienempfehlungen ist es also eher unwahrscheinlich, dass die Empfehlungen durch (implizite) kontextbezogene Faktoren systematisch in dieselbe Richtung beeinflusst wurden, so dass die Gefahr eines potentiellen kontextuellen Bias für die zu entwickelnde Leitlinie als gering eingeschätzt wird.

5.4.2 Fehlende Zweitbewertung

Alle Verfahrensschritte wurden mit Ausnahme der Auswahl einzuschließender Leitlinien ressourcenbedingt ohne einen zweiten unabhängigen Reviewer vollzogen und unterliegen somit der potentiellen Gefahr von untersucherbedingten Verzerrungen:

- Die Konzeption des AGREE-Instruments sieht die Bewertung der Qualität von Leitlinien durch mehrere Reviewer unabhängig voneinander vor, da die Reliabilität des Instruments mit der Anzahl der Gutachter steigt [AGREE 2003a]. Allerdings wurde für die Domäne 3 zur methodologischen Exaktheit von Leitlinien gezeigt, dass hier auch bei Einzelbewertung eine hohe Reliabilität vorliegt [AGREE 2003a]. Daher wurde die vorliegende Arbeit der zentralen Aufgabe dieses Verfahrensschritts (Sicherstellung, dass sich unter den berücksichtigten Leitlinien solche von hoher methodischer Qualität befanden) auch ohne Zweitreview gerecht. Zudem wurde im obenstehenden Vergleich zu anderen Bewertungen ein eher konservativer Bewertungstrend in der vorliegenden Arbeit gezeigt.
- Eine Datenextraktion durch mehr als einen Reviewer wurde von verschiedenen methodischen Arbeitsgruppen, wie z.B. der ECHTA *Working Group 4*

[Busse et al. 2002]⁶² oder in der „Leitlinie“ zur Anfertigung systematischer Reviews des CRD⁶³ [Khan et al. 2001] vorgeschlagen, um eine angemessene Reliabilität zu gewährleisten, auch wenn ein derartiges Vorgehen üblicherweise nicht oder nur selten umgesetzt wird. Um die Auswirkungen potentieller Fehler in der Datenextraktion auf die spätere Leitlinie zu begrenzen, wurden die Evidenztabelle allen Teilnehmern des Konsensusverfahrens vorgelegt und stehen auch den späteren Nutzern der autorisierten Leitlinie im Internet zur Einsicht. Ebenso wurde in Bezug auf Konsistenzanalyse und Analyse der Belegtheit verfahren, die denselben Limitationen unterworfen waren: alle Evidenztabelle und Ergebnisse wurden zunächst den Konsensteilnehmern und später den Nutzern der Leitlinie in Form eines ausführlichen Methodenreports zur Einsicht ins Internet gestellt. Außerdem nahm die Verfasserin nur unterstützend am Konsensusprozess teil und stimmte über die Formulierung der endgültigen Leitlinienempfehlungen nicht mit ab.

5.4.3 Fehlende Validierung des Klassifikationssystems für die Konsistenzanalyse

Eine Limitation der Konsistenzanalyse besteht darin, dass das von der Verfasserin entwickelte Klassifikationssystem der (In-)Konsistenzen bislang nicht von weiteren Anwendern getestet und validiert wurde. Demzufolge fehlen hier Daten über Reliabilität, diskriminatorische Güte und Validität dieses Instruments.

5.5 Weiterer Forschungsbedarf

Der SLR wurde im Rahmen der Entwicklung der DEGAM-Leitlinie Herzinsuffizienz neuentwickelt und erstmals erfolgreich eingesetzt. In weiteren Anwendungen werden nicht nur das Klassifikationssystem der Konsistenzen und Inkonsistenzen bezüglich Reliabilität, diskriminatorischer Güte und Validität zu testen sein, auch werden weitere Erfahrungen im Umgang mit dieser Methode notwendig sein, um ihren Wert als generisches Instrument in der Leitlinienentwicklung und ihre weiteren Limitationen zu erkennen.

Der Forschungsbedarf zur weiteren Evidenzgewinnung für die Herzinsuffizienz-Leitlinie wurde in Tabelle 4.19 dargestellt und im Zuge der weiteren Entwicklung der DEGAM-Leitlinie ‚Herzinsuffizienz‘ bereits teilweise realisiert. In diesem Zusammenhang wurde deutlich, dass auch zu einem hochaktuellen Thema, wie der chronischen Herzinsuffizienz, das Gegenstand von erheblicher Forschungstätigkeit ist, noch erhebliche Defizite in der Evidenzgrundlage bestehen.

⁶² ECHTA: *European Collaboration for Health Technology Assessment*

⁶³ CRD: *Centre for Reviews and Dissemination des National Health Service, University of York, U.K.*

6 Schlussfolgerungen

Der systematische Leitlinienreview ist ein

- *zusammenhängender, systematischer und transparenter* Ansatz zur Entwicklung einer evidenzbasierten Leitlinie von der Entwicklung des Frameworks bis zur Fertigstellung des Erstentwurfs, in dem jeder Verfahrensschritt klaren Prinzipien folgt und nachvollziehbar dokumentiert wird;
- *valider* Ansatz, da die in den Quelleitlinien berichteten Evidenzgrundlagen anhand einer Neubewertung der entscheidungsrelevanten Primärquellen verifiziert und damit Empfehlungen hinsichtlich ihrer empirischen Grundlage überprüft werden. Implizite (kontextspezifische) Normativität in Leitlinien kann aufgedeckt und für den anschließenden Diskurs im Konsensusprozess vorbereitet werden;
- *effektiver* Ansatz, wie am Beispiel der Entwicklung der DEGAM-Leitlinie ‚Herzinsuffizienz‘ gezeigt werden konnte;
- *ressourcensparender* Ansatz zur de novo-Leitlinienentwicklung, da unnötige Redundanzen in Literaturreviews vermieden werden können.

Darüber hinaus führt der SLR die Evidenzgrundlagen der analysierten Quelleitlinien zusammen und verbreitert somit die Evidenzgrundlage der zu entwickelnden Leitlinie.

Außerdem identifiziert er Bereiche von Entscheidungsunsicherheit, die einer vertiefenden Suche nach neuer Evidenz zugeführt werden können. Die Bereitstellung strukturierten Forschungsbedarfs ist einer formalen Priorisierung zugänglich, in der darüber entschieden wird, ob dieser Forschungsbedarf noch vor der Fertigstellung der Leitlinie oder während späterer Update-Prozesse realisiert werden soll. D.h. er vermittelt damit einen nachvollziehbaren Kompromiss im notwendigen *Trade-off* zwischen Breite und Tiefe der Evidenzgewinnung. Außerdem liefert er damit einen Beitrag zur späteren kontinuierlichen Aktualisierung der Leitlinie und zeigt auch außerhalb des Prozesses einer Leitlinienentwicklung Defizite und Lücken in der derzeitigen Evidenzbasis auf, die erst noch durch die Forschung geschlossen werden müssen.

Aus diesen Gründen empfiehlt es sich dort, wo bereits qualitativ hochwertige evidenzbasierte Leitlinien zu einem Thema existieren, den systematischen Leitlinienreview als Verfahrensschritt in der de novo-Entwicklung einer Leitlinie einzusetzen.

D - Literatur

1 Berücksichtigte Leitlinien

Die Dokumente wurden in alphabetischer Reihenfolge der Kurzbezeichnungen, basierend auf Organisations-Akronymen aufgeführt. Ggf. wurden mehrere Dokumente, die zur jeweiligen Leitlinie gehörten, unter einem gemeinsamen Gliederungspunkt zusammengefasst. Bei Bezug aus dem Internet wurden die Internetadressen (*Uniform Resource Locator*, URL) angegeben, der Zugriff erfolgte im ersten Quartal 2004.

1. **ACC/AHA 2001**: Hunt SA, Baker DW, Chin MH, Cinquegrani MP, Feldman AM, Francis GS, Ganiats TG, Goldstein S, Gregoratos G, Jessup ML, Noble RJ, Packer M, Silver MA, Stevenson LW (2001) ACC/AHA guidelines for the evaluation and management of chronic heart failure in the adult: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Revise the 1995 Guidelines for the Evaluation and Management of Heart Failure). American College of Cardiology Web site. URL: http://www.acc.org/clinical/guidelines/failure/hf_index.htm
2. **AKDAE 2001**: Arzneimittelkommission der deutschen Ärzteschaft (2001) Chronische Herzinsuffizienz. Empfehlungen zur Therapie der chronischen Herzinsuffizienz. Gemeinsame Empfehlung mit der Deutschen Gesellschaft für Kardiologie – Herz- und Kreislaufforschung. AVP-Sonderheft Therapieempfehlungen, 2. Aufl. URL: <http://www.akdae.de/45/Herzinsuffizienz.pdf> und http://www.akdae.de/35/10Hefte/88_Herzinsuffizienz_2001_2Auflage.pdf
3. **CCS 2001, 2002/3**: Canadian Cardiovascular Society. The 2002/3 Canadian cardiovascular society consensus guideline. URL: <http://www.ccs.ca/download/CCSHFConsUpdateDraft0103EMail.pdf>; Update von CCS 2001: The 2001 Canadian cardiovascular society consensus guideline update for the management and prevention of heart failure. Can J Cardiol 17 Suppl E: 5E-25E. URL: <http://www.ccs.ca/society/conferences/archives/2001update/2001coneng-01.asp>
4. **DGK 2001**: Hoppe UC, Erdmann E für die Kommission Klinische Kardiologie der Deutschen Gesellschaft für Kardiologie – Herz- und Kreislaufforschung (2001) Leitlinien zur Therapie der chronischen Herzinsuffizienz. Zeitschrift für Kardiologie 90(3): 218-37. URL: <http://www.dgkardio.de/Leitlinien/LeitlinienHerzinsuffizienz.pdf>
5. **DieM 2003/2004**: Kaiser T, Jennen E, Sawicki PT für die Arbeitsgruppe Praktische evidenzbasierte Medizin St. Franziskus Hospital, Köln und DieM – Institut für evidenzbasierte Medizin, Köln (2003) Entscheidungsgrundlage zur evidenzbasierten Diagnostik und Therapie bei Disease Management Programmen für Herzinsuffizienz bei systolischer linksventrikulärer Dysfunktion. 2., überarbeitete

Version 2003.

Kaiser T, Jennen E, Krones R, Franz H, Waltering A, Kleespies C, Sawicki PT. Evidenzbasierte Diagnostik und Therapie bei Disease Management Programmen. Methodik der Erstellung evidenzbasierter Entscheidungsgrundlagen. 2., überarbeitete Version 2004. URL: <http://www.di-em.de/publikationen.php?idtopic=17&la=de>

6. **Duodecim 2004:** The Finnish Medical Society Duodecim (2004) Chronic Heart Failure. URL: <http://www.ebm-guidelines.com>
7. **DVA & VHA 2002:** Department of Veterans Affairs & Veterans Health Administration (2002) The Pharmacologic Management of Chronic Heart Failure. Washington, DC: Pharmacy Benefits Management Strategic Healthcare Group and the Medical Advisory Panel, Veterans Health Administration, Department of Veterans Affairs. April 2001; Updated December 2002. PBM-MAP Publication No. 00-0015. URL: http://www.oqp.med.va.gov/cpg/CHF/CHF_Base.htm
8. **ESC 2002/2001:** Remme WJ & Swedberg K (2002) Comprehensive guidelines for the diagnosis and treatment of chronic heart failure. Task force for the diagnosis and treatment of chronic heart failure of the European Society of Cardiology. Eur J Heart Fail 4(1): 11-22
Remme WJ & Swedberg K (2001) Task Force for the Diagnosis and Treatment of Chronic Heart Failure, European Society of Cardiology: Guidelines for the diagnosis and treatment of chronic heart failure. Eur Heart J (22): 1527-60. Erratum in Eur Heart J 22: 2217-2218
Swedberg K, Cleland J, Dargie H, Drexler H, Follath F, Komajda M, Tavazzi L, Smiseth OA, Gavazzi A, Haverich A, Hoes A, Jaarsma T, Korewicki J, Levy S, Linde C, Lopez-Sendon JL, Nieminen MS, Pierard L, Remme WJ; Task Force for the Diagnosis and Treatment of Chronic Heart Failure of the European Society of Cardiology. Guidelines for the diagnosis and treatment of chronic heart failure: executive summary (update 2005): The Task Force for the Diagnosis and Treatment of Chronic Heart Failure of the European Society of Cardiology. Eur Heart J 26(11): 1115-40
Recommendations for Task Force Creation and Report Production. URL: <http://www.escardio.org/NR/rdonlyres/2BAC03CC-60A8-4676-BEF8-E3C46C8D57D7/0/Recommendations.pdf>, Pocketguidelines: URL: <http://www.escardio.org/knowledge/education/PGuidelines/>
9. **ICSI 2003:** Institute for Clinical System Improvement (2003) Health Care Guideline. Congestive Heart Failure in Adults. URL: <http://www.icsi.org/knowledge/detail.asp?catID=29&itemID=161>, Institute for Clinical System Improvement (2002) Evidence Grading System. URL: <http://www.icsi.org/knowledge/detail.asp?catID=113&itemID=619>
10. **LLGH 2003:** Leitlinien-Gruppe Hessen: Bergert FW, Conrad D, Ehrenthal K, Feßler J, Gross J, Gundermann K, Kluthe B, Krause R, Lang HW, Liesenfeld A,

Loew PG, Luther E, Pchalek R, Seffrin J, Sterzing A, Wolfring HJ, Zimmermann U (2003) Hausärztliche Leitlinie. Chronische Herzinsuffizienz. Therapie der chronischen Herzinsuffizienz. Version 2.02. Z Ärztl Fortbild Qualitätssich 97(2): 145-50. URL: <http://www.leitlinien.de/leitlinienanbieter/index/deutsch/qualitaetszirkel/index/hessen/pdf/hessenherzinsuffizienz> und <http://www.leitlinien.de/leitlinienanbieter/deutsch/pdf/hessenhireport>

11. **NHF/Austr. & CSANZ 2002:** National Heart Foundation of Australia and Cardiac Society of Australia & New Zealand (2002) Guidelines on the Contemporary Management of the Patient with Chronic Heart Failure in Australia. URL: <http://www.new.heartfoundation.com.au/downloads/cont.management.pdf>, http://www.csanz.edu.au/guidelines/practice/Chronic_Heart_Failure.pdf, http://www.heartfoundation.com.au/prof/docs/chronic_failure.pdf
12. **NHF/NZ 2001:** The National Heart Foundation of New Zealand (2001) A guideline for the management of heart failure. URL: http://www.nzgg.org.nz/guidelines/dsp_guideline_popup.cfm?guidelineCatID=32&guidelineID=26
13. **NICE 2003:** The National Collaborating Centre for Chronic Conditions (2003) Chronic heart failure. National clinical guideline for diagnosis and management in primary and secondary care. NICE Guide-line No. 5. National Institute for Clinical Excellence, ISBN: 1-84257-324-1. URL: <http://www.nice.org.uk>
14. **OPOT 2000:** Ontario Program for Optimal Therapeutics (2000) Ontario drug therapy guidelines for chronic heart failure in primary care. ISBN: 1-894706-03-X. URL: <http://www.opot.org/guidelines/chf.pdf>
15. **UM 2001:** Chavey WE II, Blaum CS, Bleske BE, Van Harrison R, Kesterson S, and Nicklas JM (2001) Guideline for the management of heart failure caused by systolic dysfunction: Part I. Guideline development, aetiology and diagnosis. American Family Physician 64(5): 769-74. URL: <http://www.aafp.org/afp/20010901/769.pdf>
Chavey WE II, Blaum CS, Bleske BE, Van Harrison R, Kesterson S, and Nicklas JM (2001) Guideline for the management of heart failure caused by systolic dysfunction: Part II. Treatment. American Family Physician 64(6): 1045-54. URL: <http://www.aafp.org/afp/20010915/1045.pdf>
16. **UWH 2001:** Universität Witten / Herdecke (2001) Leitlinie Diagnose und Therapie der chronischen Herzinsuffizienz. URL: http://www.evidence.de/Leitlinien/leitlinien-intern/Herzinsuffizienz_Start/herzinsuffizienz_start.html, http://www.evidence.de/Leitlinien/leitlinien-intern/Herzinsuffizienz_Start/Herzinsuffizienz-Patienten-LL/herzinsuffizienz-patienten-ll.html

2 Sonstige Literatur

Die Dokumente wurden in alphabetischer Reihenfolge der Erstautoren aufgeführt. Bei Bezug aus dem Internet wurden die Internetadressen (*Uniform Resource Locator*, URL) und das jeweilige Zugriffsdatum angegeben.

1. Aaronson KD, Schwartz JS, Chen TM, et al. (1997) Development and prospective validation of a clinical index to predict survival in ambulatory patients referred for cardiac transplant evaluation. *Circulation* 95(12): 2660-7
2. Adams KF, Gheorghiade M, Uretsky BF, et al. (2002) Clinical benefits of low serum digoxin concentrations in heart failure. *J Am Coll Cardiol* 39: 946-953
3. Agency for Healthcare Policy and Research (AHCPR) (1995) Using clinical practice guidelines to evaluate quality of care. U.S. Department of Health and Human Services, Volume I: Issues, AHCPR Publication No. 95-0045, p 35; Volume II: Methods, AHCPR Publication No. 95-0046 p. 117, Rockville
4. Agency for Healthcare Research and Quality (AHRQ) (2005) EPC Topic Nomination and Selection. Agency for Healthcare Research and Quality, Rockville, MD. URL: <http://www.ahrq.gov/clinic/epc/epctopicn.htm> (letzter Internet-Zugriff 03.01.2008)
5. Alberta Heritage Foundation for Medical Research (AHFMR) (2005) B-type natriuretic peptide for diagnosing congestive heart failure. TechNote 46, July 2004 (Updated January 2005). URL: <http://www.ihe.ca/documents/hta/tn46.pdf> (letzter Internet-Zugriff 03.01.2008)
6. Alderson P & Green S (eds.) (2002) Cochrane Collaboration open learning material for reviewers. Version 1.1, The Cochrane Collaboration, URL: <http://www.cochrane-net.org/openlearning/> (letzter Internet-Zugriff 03.01.2008); module 10: 3-4, module 13: 8-9
7. Alderson P, Green S, Higgins JPT (eds.) (2005) Cochrane Reviewers' Handbook 4.2.3. In: The Cochrane Library, Issue 1, 2005. Chichester, UK: John Wiley & Sons, Ltd. URL: <http://www.cochrane.org/resources/handbook/hbook.htm> (Internet-Zugriff 31.01.2005); pp 61, 116ff.
8. Aldrich R, Kemp L, Williams JS, et al. (2003) Using socioeconomic evidence in clinical practice guidelines. *BMJ* 327(7426): 1283-5
9. Alehagen U, Lindstedt G, Eriksson H, Dahlstrom U (2003) Utility of the aminoterminal fragment of pro-brain natriuretic peptide in plasma for the evaluation of cardiac dysfunction in elderly patients in primary health care. *Clin Chem* 49(8): 1337-46
10. American College of Cardiology's Foundation (2005) URL: <http://www.acc.org/clinical/bethesda/beth28/9991/convrted/> (Internet-Zugriff vom 20.08.2005)
11. American Stroke Association & American Heart Association (2005) Heart disease and stroke statistics – 2005 Update. URL: <http://www.americanheart.org>

- [/downloadable/heart/1105390918119HDSStats2005Update.pdf](#) (Internet-Zugriff vom 10.12.2005)
12. Anonymus (2006) Trade-off. Wikipedia – Die freie Enzyklopädie. URL: <http://de.wikipedia.org/wiki/Trade-off> (letzter Internet-Zugriff vom 03.01.2008)
 13. Anonymus (2006a) Humes Gesetz. Sekundärquelle in: Wikipedia – Die freie Enzyklopädie, URL: http://de.wikipedia.org/wiki/Humes_Gesetz (Internet-Zugriff vom 31.03.2006); Primäre Quelle: David Hume. Treatise of Human Nature. Book III: Of morals. Part I: Of virtue and vice in general. Sect. I: Moral distinctions not deriv'd from reason. URL: http://en.wikisource.org/wiki/Treatise_of_Human_Nature/Book_3:_Of_morals#Sect._I:_Moral_distinctions_not_deriv.E.2.80.99d_from_reason (Internet-Zugriff vom 31.03.2006)
 14. Anonymus (2004) Glossary of Terms in The Cochrane Collaboration. Version 4.2.3. URL: <http://www.cochrane.org> (Internet-Zugriff vom 06.02.2005)
 15. Antes G (2004) Die Evidenz-Basis von klinischen Leitlinien, Health Technology Assessments und Patienteninformation als Grundlage für Entscheidungen in der Medizin. Z ärztl Fortbild Qual Gesundh wes 98: 180–184
 16. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) und Ärztliches Zentrum für Qualität in der Medizin (ÄZQ) (2003) Neufassung der Checkliste „Methodische Qualität von Leitlinien“ in der Entwurfsfassung vom 23.09.2003. URL: <http://www.aezq.de> (Internet-Zugriff April 2004)
 17. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) und Ärztliches Zentrum für Qualität in der Medizin (ÄZQ) (Hrsg.) (2005) Deutsches Instrument zur methodischen Leitlinien-Bewertung (DELBI). Z ärztl Fortbild Qual (ZaeFQ) 99(8): 468-519. URL: <http://www.delbi.de> (Internet-Zugriff vom 27.08.2005)
 18. Aronow WS (2006) Epidemiology, pathophysiology, prognosis, and treatment of systolic and diastolic heart failure. Cardiol Rev 14: 108–124
 19. Asch DA & Hershey JC (1995) Why some health policies don't make sense at the bed-side. Ann Int Med 122(11): 846-850
 20. Ashcroft RE (2004) Current epistemological problems in evidence based medicine. J Med Ethics 30: 131-135
 21. Ashcroft R (2002) What is clinical effectiveness? Stud Hist Phil Biol & Biomed Sci 33: 219-233
 22. Ashton T, Cumming J, Devlin N (1999) Prioritising health and disability support services: principles, processes and problems. Report to the National Health Committee, New Zealand., URL: <http://www.nhc.govt.nz/publications/PDFs/prior.pdf> (Internet-Zugriff vom 16.03.2006)
 23. Atkins D, Best D, Briss PA, et al. for the GRADE Working Group (2004) Grading quality of evidence and strength of recommendations. BMJ 328(7454): 1490

24. Atkins D, Briss PA, Eccles M, et al. for the GRADE Working Group (2005) Systems for grading the quality of evidence and the strength of recommendations II: pilot study of a new system. *BMC Health Serv Res* 5(1): 25
25. Atkins D, Eccles M, Flottorp S, et al. for the GRADE Working Group (2004a) Systems for grading the quality of evidence and the strength of recommendations I: critical appraisal of existing approaches The GRADE Working Group. *BMC Health Serv Res* 4(1): 38
26. Australia/New Zealand Heart Failure Research Collaborative Group (1997) Randomized, placebo-controlled trial of carvedilol in patients with congestive heart failure due to ischaemic heart disease. *Lancet* 349: 375-380
27. Bach R & Zardini P (1992) Long-acting angiotensin-converting enzyme inhibition: once-daily lisinopril versus twice-daily captopril in mild-to-moderate heart failure. *Am J Cardiol.* 70(10): 70C-77C
28. Banta D & Luce B (1993) Health technology and its assessment. Oxford University Press, Oxford, pp. 35ff, 65, 70, 88.
29. Barker WH, Mullooly JP, Getchell W (2006) Changing incidence and survival for heart failure in a well-defined older population, 1970-1974 and 1990-1994. *Circulation* 113(6): 799-805
30. Battaglia M, Pewsner D, Juni P, et al. (2006) Accuracy of B-type natriuretic peptide tests to exclude congestive heart failure: systematic review of test accuracy studies. *Arch Intern Med* 166(10): 1073-80
31. Beck C, Cody M, Souder E, et al. (2000) Dementia diagnostic guidelines: methodologies, results and implementation costs. *J Am Geriatr Soc* 48: 1195–1203
32. Belardinelli R, Georgiou D, Cianci G, et al. (1999) Randomized, controlled trial of long-term moderate exercise training in chronic heart failure. *Circulation* 99:1173-82. Zitiert aus DleM 2003 (siehe oben)
33. Beller B, Bulle T, Bourge RC, et al. (1995) Lisinopril versus placebo in the treatment of heart failure: the Lisinopril Heart Failure Study Group. *J Clin Pharmacol* 35: 673-680
34. Bonet S, Agusti A, Arnau JM, et al. (2000) Beta-adrenergic blocking agents in heart failure: benefits of vasodilating and non-vasodilating agents according to patients' characteristics: a meta-analysis of clinical trials. *Arch Int Med* 160: 621–7
35. Bouvy ML, Heerdink ER, Leufkens HG, Hoes AW (2003) Predicting mortality in patients with heart failure: a pragmatic approach. *Heart* 89(6): 605-9
36. Bouzamondo A, Hulot JS, Sanchez P, et al. (2001) Beta-blocker treatment in heart failure. *Fundam Clin Pharmacol* 15: 95–109
37. Braunwald E & Grossman W (1992) Clinical aspects of heart failure. In: Braunwald E (Ed.) (1992) *Heart disease: a textbook of cardiovascular medicine*. 4th

- Edition. W.B. Saunders Company, Philadelphia. ISBN 0-7216-4408-2. Vol. 1, pp. 444-447.
38. Breen A & Feder G (1999) Where does the evidence come from? In: Hutchinson A & Baker R (Eds.) Making use of guidelines in clinical practice. Radcliffe Medical Press Ltd., Abingdon, Oxon, pp. 15-28
 39. Broadstock M (2000) Effectiveness and cost effectiveness of automated and semi-automated cervical screening devices: a systematic review of the literature. New Zealand Health Technology Assessment (NZHTA); The Clearing House for Health Outcomes and Technology Assessment. NZHTA Report 3(1). Chapter 2.9.
 40. Brockmann S (2004) Hausärztliche Leitlinien zwischen Erfahrung und ‚Eviden- ce‘. Omikron Publishing, Düsseldorf. S. 89ff
 41. Brophy JM, Joseph L, Rouleau JL (2001) Beta-blockers in congestive heart fail- ure. A Bayesian meta-analysis. Ann Int Med 134: 550–60
 42. Brucknerberger E (2004) Herzbericht 2004 mit Transplantationschirurgie. ISBN 3-00-017046-4. S. 9ff
 43. Bundesärztekammer, Kassenärztliche Bundesvereinigung (Hrsg.) Bloch RE, Lauterbach K, Oesingmann U, et al. (1997) Beurteilungskriterien für Leitlinien in der medizinischen Versorgung. Beschlüsse der Vorstände von Bundesärzte- kammer und Kassenärztlicher Bundesvereinigung. Dtsch Ärztebl 94: A2154- 2155.
 44. Burgers JS, Bailey JV, Klazinga NS, et al. for the AGREE Collaboration (2002) Inside guidelines: comparative analysis of recommendations and evidence in diabetes guide-lines from 13 countries. Diabetes Care 25(11): 1933-1939
 45. Bursi F, Weston SA, Redfield MM, et al. (2006) Systolic and diastolic heart fail- ure in the community. JAMA 296(18): 2209-16
 46. Busse R & Schlette S (eds.) (2003) Health policy developments – international trends and analyses. Bertelsmann Foundation Publishers, Guetersloh, pp 19- 36, 57-66. URL: http://www.bertelsmann-stiftung.de/bst/de/media/xcms_bst_dms_15030_15031_2.pdf (Internet-Zugriff vom 14.02.2007)
 47. Busse R (2002) Health care systems: Germany and Britain compared. Health care delivery – towards an agenda for policy learning between Britain and Ger- many. London: Anglo-German Foundation for the Study of Industrial Society. URL: <http://www.agf.org.uk/pubs/pdfs/1363Healthweb.pdf> (Internet-Zugriff vom 14.02.2007)
 48. Busse R, Orvain J, Velasco M, et al. (2002) Best practice in undertaking and reporting health technology assessments. Working Group 4 Report. Int J Tech- nol Assess Health Care 18(2): 361-422; p 375
 49. Busse R, Schlette S, Weinbrenner S (eds.) (2005) Health policy developments issue 4: Focus on access, primary care, health care organization. Bertelsmann

- Foundation Publishers, Guetersloh. URL: <http://www.wm.tu-berlin.de/fachgebiet/mig/papers/index.html> (Internet-Zugriff vom 09.03.2006)
50. Cahn M & Jones RH (1996) Personal communication, 1996. In: American College of Cardiology's Foundation. URL: <http://www.acc.org/clinical/bethesda/beth28/9991/convrted/> (Internet-Zugriff vom 20.08.2005)
 51. Campbell F, Dickinson HO, Cook JVF, et al. (2006) Methods underpinning national clinical guidelines for hypertension: describing the evidence shortfall. *BMC Health Serv Res* 6: 47
 52. Chalmers JP, West MJ, Cyran J, et al. (1987) Placebo-controlled study of lisinopril in congestive heart failure: a multicentre study. *J Cardiovasc Pharmacol* 9 Suppl 3: S89-97
 53. Champion HC, Skaf MW, Hare JM (2003) Role of nitric oxide in the pathophysiology of heart failure. *Heart Fail Rev* 8(1): 35-46
 54. Choudhry NK, Stelfox HT, Detsky AS (2002) Relationships between authors of clinical practice guidelines and the pharmaceutical industry. *JAMA* 287: 612-617
 55. Christiaens T, De Backer D, Burgers J, Baerheim A (2004) Guidelines, evidence, and cultural factors. Comparison of four European guidelines on uncomplicated cystitis. *Scand J Prim Health Care* 22: 141-145
 56. CIBIS-II Investigators and Committee (1999) The cardiac insufficiency bisoprolol study II (CIBIS-II): a randomised trial. *Lancet* 353: 9-13
 57. Cleland JGF, Chattopadhyay S, Khand A, et al. (2002) Prevalence and incidence of arrhythmias and sudden death in heart failure. *Heart Failure Reviews* 7: 229-242
 58. Cleland JGF, Cohen-Solal A, Cosin J, et al, for the IMPROVEMENT of Heart Failure Programme Committees and Investigators* and the Study Group on Diagnosis of the Working Group on Heart Failure of The European Society of Cardiology (2002) Management of heart failure in primary care (the IMPROVEMENT of Heart Failure Programme). An international survey. *Lancet* 360: 1631-9
 59. Cluzeau FA, Littlejohns P, Grimshaw JM, et al. (1999) Development and application of a generic methodology to assess the quality of clinical guidelines. *Int J Qual Health Care* 11(1): 21-8
 60. Coats AJ (2001) CAPRICORN: a story of alpha allocation and beta-blockers in left ventricular dysfunction post-MI. *Int J Cardiol* 78(2): 109-13
 61. Cohen J (1960) A coefficient of agreement for nominal scales. *Educ Psychol Meas* 20: 37-46
 62. Cohn JN & Tognoni G for the Valsartan Heart Failure Trial Investigators (Val-HeFT) (2001) A randomized trial of the angiotensin-receptor blocker valsartan in chronic heart failure. *N Engl J Med* 345: 1667-75

63. Cohn JN, Johnson G, Ziesche S et al. (1991) A comparison of enalapril with hydralazine-isosorbide dinitrate in the treatment of chronic congestive heart failure. *N Engl J Med* 325: 303-10
64. Cohon R (2004) Hume's Moral Philosophy. *Stanford Encyclopedia of Philosophy*. URL: <http://plato.stanford.edu/entries/hume-moral/> (Internet-Zugriff vom 30.09.2007)
65. Cowburn PJ, Cleland JGF, Coats AJS, Komajda M (1998) Risk stratification in heart failure. *Eur Heart J* 19: 696-710
66. Cowie MR (2003) Estimating prognosis in heart failure: time for a better approach. *Heart* 89: 587-588
67. Cowie MR, Wood DA, Coats AJS, et al. (2000) Survival of patients with a new diagnosis of heart: a population based study. *Heart* 83: 505-510
68. Craig J, Bradbury I, Cummings E, et al. (2005) The use of B-type natriuretic peptides (BNP and NT-proBNP) in the investigation of patients with suspected heart failure. NHS Quality Improvement Scotland (NHS QIS). Health Technology Assessment Re-port No. 6. ISDN 1-903961-49-1. URL: http://www.nhshealthquality.org/nhsqis/files/bnp_final_report.pdf (Internet-Zugriff vom 21.06.2005)
69. Dahlöf B, Lindholm LH, Hansson L, et al. (1991) Morbidity and mortality in the Swedish Trial in Old Patients with Hypertension (STOP-Hypertension). *Lancet* 338(8778): 1281-5
70. Dahlöf B, Pennert K, Hansson L (1992) Reversal of left ventricular hypertrophy in hypertensive patients: a meta-analysis of 109 treatment studies. *Am J Hypertens* 5: 95-110
71. Davenport C, Cheng EY, Kwok YT, et al. (2006) Assessing the diagnostic test accuracy of natriuretic peptides and ECG in the diagnosis of left ventricular systolic dysfunction: a systematic review and meta-analysis. *Br J Gen Pract* 56(522): 48-56
72. Davila DF, Nunez TJ, Odreman R, de Davila CA (2005) Mechanisms of neuro-hormonal activation in chronic congestive heart failure: pathophysiology and therapeutic implications. *Int J Cardiol* 101(3): 343-6
73. Davis DA & Taylor-Vaisey A (1997) Translating guidelines into practice. *Can Med Assoc* 157: 408-416
74. Deeks JJ (2001) Systematic reviews of evaluations of diagnostic and screening tests. In: Egger M; Smith GD; Altman DG (ed.) (2001) *Systematic reviews in health care. Meta-analysis in context*. 2nd Edition. BMJ Books, BMJ Publishing Group, BMA House, London. PP 248-282
75. Deeks JJ, Altman DG, Bradburn MJ (2001) Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In: Egger M; Smith GD; Altman DG (ed.). *Systematic reviews in health care. Meta-*

- analysis in con-text. 2nd Edition. BMJ Books, BMJ Publishing Group, BMA House, London. PP 285-312
76. DeMaeseneer J & Derese A (1999) European general practice guidelines: a step too far? *Eur J Gen Pract* 5: 86-87
 77. Detsky AS (2006) Sources of bias for authors of clinical practice guidelines. *CMAJ* 175(9): 1033
 78. Doughty RN, Rodgers A, Sharpe N, MacMahon S (1997) Effects of beta-blocker therapy on mortality in patients with heart failure: a systematic overview of randomized controlled trials. *Eur Heart J* 18: 560-5
 79. Doust JA, Glasziou PP, Pietrzak E, Dobson AJ (2004) A systematic review of the diagnostic accuracy of natriuretic peptides for heart failure. *Arch Int Med* 164: 1978-1984.
 80. Duodecim (2002) Right ventricular failure. Zitiert in: Fessler J, Gross J, Hoepf HW et al. (Autoren), Thole H (Ed.), *Ärztliches Zentrum für Qualität in der Medizin, Leitlinien-Clearingstelle (Hrsg.) (2005) Leitlinien-Clearingbericht "Herzinsuffizienz": Leitlinien-Clearingverfahren von Bundesärztekammer und Kassenärztlicher Bundesvereinigung in Kooperation mit Deutscher Krankenhausgesellschaft, Spitzenverbänden der Krankenversicherungen und gesetzlicher Rentenversicherung*. Niebuell: Verl. Videel, (AeZQ-Schriftenreihe; Bd. 18) ISBN 3-89906-909-9
 81. Eccles M & Mason J (2001) How to develop cost-conscious guidelines. *Health Technol Assess* 5(16). ISSN 1366-5278
 82. Eccles M, Freemantle N, Mason J (1998) North of England evidence based development project: guideline for angiotensin converting enzyme inhibitors in primary care management of adults with symptomatic heart failure. *BMJ* 316(7141): 1369-1375
 83. Eccles M, Grimshaw J, Campbell M, Ramsay C (2003) Research designs for studies evaluating the effectiveness of change and improvement strategies. *Qual Saf Health Care* 12(1): 47-52
 84. Eddy DM (2001) The use of evidence and cost effectiveness by the courts: how can it help improve health care? *J Health Polit Pol* 26(2): 387-408
 85. Eddy DM (1990) Practice policies – what are they? *JAMA* 263(6): 877-880
 86. Egger M, Dickersin K, Smith GD (2001) Problems and limitations in conducting systematic reviews. In: Egger M, Smith GD, Altman DG (Editors) (2001) *Systematic Reviews in Health Care. Metaanalysis in Context*. 2nd Edition. BMJ Publishing Group London, pp 60-62
 87. Eisinger F, Geller G, Burke W, Holtzman NA (1999) Cultural basis for differences between US and French clinical recommendations for women at increased risk of breast and ovarian cancer. *Lancet* 353: 919-920

88. Elstein AS (2004) On the origins and development of evidence-based medicine and medical decision making. *Inflamm Res* 53 (Suppl. 2): 184-189
89. Europarat (2002) Entwicklung einer Methodik für die Ausarbeitung von Leitlinien für optimale medizinische Praxis. Empfehlung Rec(2001)13 des Europarates und Erläuterndes Memorandum – Deutschsprachige Ausgabe. Bern (Verbindung der Schweizer Ärztinnen und Ärzte), Köln (Ärztliche Zentralstelle Qualitätssicherung), Wien (Ludwig Boltzmann Institut für Krankenhausorganisation). S. 27. URL: <http://www.leitlinien.de/informationen/pdf/europaratmethdt.pdf> (Internet-Zugriff vom 31.01.2007)
90. Exner DV, Dries DL, Waclawiw MA, et al. (1999) Beta-adrenergic blocking agent use and mortality in patients with asymptomatic and symptomatic left ventricular systolic dysfunction: a post hoc analysis of the studies of left ventricular dysfunction. *J Am Coll Cardiol* 33: 916-923
91. Fagan TJ (1975) Nomogram for Bayes's Theorem. *N Engl J Med* 293: 257
92. Fahey TP & Peters TJ (1996) What constitutes controlled hypertension? Patient based comparison of hypertension guidelines. *BMJ* 313: 93–96
93. Faris R, Flather M, Purcell H, et al. (2002) Current evidence supporting the role of diuretics in heart failure: a meta analysis of randomised controlled trials. *Int J Cardiol* 82: 149–58
94. Feder G, Eccles M, Grol R, et al. (1999) Clinical guidelines. Using clinical guidelines. *BMJ* 318: 728–730
95. Fervers B, Burgers JS, Haugh MC, et al. (2006) Adaptation of clinical guidelines: literature review and proposition for a framework and procedure. *Int J Qual Health Care* 18(3): 167-76
96. Fessler J, Gross J, Hoepf HW et al. (Autoren), Thole H (Ed.), Ärztliches Zentrum für Qualität in der Medizin (ÄZQ), Leitlinien-Clearingstelle (Hrsg.) (2005) Leitlinien-Clearingbericht "Herzinsuffizienz": Leitlinien-Clearingverfahren von Bundesärztekammer und Kassenärztlicher Bundesvereinigung in Kooperation mit Deutscher Krankenhausgesellschaft, Spitzenverbänden der Krankenkassen und gesetzlicher Rentenversicherung. Niebuell: Verl. Videel, (AeZQ-Schriftenreihe; Bd. 18) ISBN 3-89906-909-9
97. Field MJ & Lohr KN (eds) (1990) Clinical practice guidelines. Directions for a new program. National Academic Press, Washington D.C., p. 38
98. Field MJ & Lohr KN (eds) (1992) Guidelines for clinical practice. From development to use. National Academy Press, Washington D.C.
99. Fieser J (2006) David Hume (1711-1776) Moral Theory. The Internet Encyclopedia of Philosophy. URL: <http://www.iep.utm.edu/h/humemora.htm> (Zugriff vom 30.09.2007)
100. Fischer M, Baessler A, Holmer SR, et al. (2003) Epidemiologie der linksventrikulären Dysfunktion in der Allgemeinbevölkerung Deutschlands. *Ergebnisse*

echokardiographischer Untersuchungen einer großen Bevölkerungsstichprobe. Z Kardiol 92: 294-302

101. Fitch K, Bernstein S, Aguilar M (2001) The RAND/UCLA appropriateness method: user's manual. URL: http://www.rand.org/pubs/monograph_reports/MR1269/index.html (Internet-Zugriff vom 20.02.2004)
102. Flather MD, Yusuf S, Kober L, et al. (2000) Long-term ACE-inhibitor therapy in patients with heart failure of left ventricular dysfunction: a systematic overview of data from individual patients. Lancet 355: 1575-81
103. Fletcher RH (1999) Building in the evidence. In Margolis CZ & Cretin S (eds.) Implementing clinical practice guidelines. AHA-Press, Chicago, pp. 47
104. Fonarow GC, Chelimsky-Fallick C, Stevenson LW, et al. (1992) Effect of direct vasodilation with hydralazine versus angiotensin-converting enzyme inhibition with captopril on mortality in advanced heart failure: the Hy-C trial. J Am Coll Cardiol 19: 842-50
105. Foy R, Ramsay CR, Grimshaw JM, et al. (2004) The impact of guidelines on mild hypertension in pregnancy: time series analysis. BJOG 111(8): 765-70
106. Freemantle N, Cleland J, Young P, et al. (1999) Beta Blockade after myocardial infarction: systematic review and meta regression analysis. BMJ 318(7200): 1730-7
107. Garg R & Yusuf S (1995) Overview of randomized trials of angiotensin converting enzyme inhibitors on mortality and morbidity in patients with heart failure. JAMA 273: 1450-6
108. Gartlehner G, West SL, Lohr KN, et al. (2004) Assessing the need to update prevention guidelines: a comparison of two methods. Int J Qual Health Care 16(5): 399-406
109. Gensichen J, Muth C, Butzlaff M, et al. (2006) Die Zukunft ist chronisch: das Chronic Care-Modell in der deutschen Primärversorgung: Übergreifende Behandlungsprinzipien einer proaktiven Versorgung für chronische Kranke. Z ärztl Fortbild Qual Gesundh wes 100(5): 365-374
110. Gerlach FM (2002) Allgemeinmedizin in einem modernen Gesundheitssystem – Konsequenzen für Forschung, Lehre und Praxis (Teil 1). Z ärztl Fortbild Qualsich 96: 243-249
111. Gerlach FM (2002b) Allgemeinmedizin in einem modernen Gesundheitssystem – Konsequenzen für Forschung, Lehre und Praxis (Teil 2). Z ärztl Fortbild Qualsich 96: 307-311
112. Gerlach FM, Abholz HH, Berndt M, Beyer M, et al. für den DEGAM-Arbeitskreis ‚Leitlinien‘ in der Sektion Qualitätsförderung (1999) Konzept zur Entwicklung, Verbreitung, Implementierung und Evaluation von Leitlinien für die hausärztliche Praxis. Deutsche Gesellschaft für Allgemeinmedizin (DEGAM), Köln; ISBN 3-9806900-0-8

113. Gerlach FM, Beyer M, Berndt M, et al. (1999) Das DEGAM-Konzept – Entwicklung, Verbreitung, Implementierung und Evaluation von Leitlinien für die hausärztliche Praxis. *Z ärztl Fortbild Quallsich* 93: 111-120
114. Gerlach FM, Beyer M, Szecsenyi J, Fischer GC (1998) Leitlinien in Klinik und Praxis. Welche Anforderungen sollten moderne Leitlinien erfüllen? Welche Strategien zur Entwicklung, Verbreitung und Implementierung haben sich bewährt? Welchen Beitrag können Leitlinien zur Qualitätsförderung in der Medizin leisten? *Deutsches Ärzteblatt (Suppl)* 95: A-1014-A-1021
115. Giles TD, Katz R, Sullivan JM, et al. (1989) Short- and long-acting angiotensin-converting enzyme inhibitors: a randomized trial of lisinopril versus captopril in the treatment of congestive heart failure. The Multicenter Lisinopril-Captopril Congestive Heart Failure Study Group. *J Am Coll Cardiol* 13(6): 1240-7
116. Glasier A, Brechin S, Raine R, Penney G (2003) A consensus process to adapt the World Health Organization selected practice recommendations for UK use. *Contraception* 68(5): 327-33
117. Glasziou P, Irwig L, Bain C, Colditz G (2001) Systematic reviews in health care. A practical guide. Cambridge University Press, Cambridge, PP 36
118. Goldenberg MJ (2005) Evidence-based ethics? On evidence-based practice and the “empirical turn” from normative bioethics. *BMC Medical Ethics* 6: 11
119. Gottdiener JS, Reda DJ, Massie BM, et al. (1997) Effect of single-drug therapy on reduction of left ventricular mass in mild to moderate hypertension: comparison of six antihypertensive agents. *Circulation* 95: 2007-14
120. Graham ID, Calder LA, Hebert PC, et al. (2000) A comparison of clinical practice guideline appraisal instruments. *Int J Technol Assess Health Care* 16(4): 1024-38
121. Graham ID, Harrison MB, Lorimer K, et al. (2005) Adapting national and international leg ulcer practice guidelines for local use: the Ontario Leg Ulcer Community Care Protocol. *Adv Skin Wound Care* 18(6): 307-18
122. Graham RP, James PA, Cowan TM (2000) Are clinical practice guidelines valid for primary care? *J Clin Epidemiol* 53: 949-954
123. Grant J, Cottrell R, Cluzeau F, Fawcett G (2000) Evaluating “payback” on biomedical research from papers cited in clinical guidelines: applied bibliometric study. *BMJ* 320: 1107-1111
124. Greenland S (1987) Quantitative methods in the review of epidemiologic literature. *Epidemiol Rev* 9: 1-30
125. Greenland S (1994) Invited commentary: a critical look at some popular meta-analytic methods. *Am J Epidemiol* 140: 290-6
126. Grilli R, Magrini N, Penna A, et al. (2000) Practice guidelines developed by specialty societies: the need for a critical appraisal. *Lancet* 355: 103-06

127. Grimshaw J & Russell I (1993) Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet* 342: 1317-1322
128. Grunwald A (2004) The normative basis of (health) technology assessment and the role of ethical expertise. *Poiesis Prax* 2: 175-193
129. Grunwald A (2002) Normativ, aber nicht normativistisch. Bemerkungen zur Grenze zwischen Wissen, Werten und Bekehrungen als Antwort auf Christian Berg. *Technikfolgenabschätzung – Theorie und Praxis* 11(2): 155-157, URL: <http://www.itas.fzk.de/tatup/022/grun02b.pdf> (Zugriff vom 13.01.2008)
130. Grunwald A (2000) Technikfolgenabschätzung und Ethik. Eine Verhältnisbestimmung in Theorie und Praxis. Rezension. *TA-Datenbank-Nachrichten* 9(2): 98-102, URL: <http://www.itas.fzk.de/deu/tadn/tadn002/grun00b.pdf> (Zugriff vom 13.01.2008)
131. Hadorn DC & Baker D (1994) Development of the AHCPR-sponsored heart failure guideline: methodological and procedural issues. *Jt Comm J Qual Patient Saf* 20: 539-47
132. Hadorn DC, Baker D, Hodges JS, Hicks N (1996) Rating the quality of evidence for clinical practice guidelines. *J Clin Epidemiol* 49: 749-754
133. Haggerty JL, Reid RJ, Freeman GK, Starfield BH, et al. (2003) Continuity of care: a multidisciplinary review. *BMJ* 327: 1219-1221
134. Haim M, Shotan A, Boyko V, et al. (1998) Effect of beta-blocker therapy in patients with coronary artery disease in New York Heart Association classes II and III. The Bezafibrate Infarction Prevention (BIP) Study Group. *Am J Cardiol* 81(12): 1455-60
135. Haynes RB, Devereaux PJ, Guyatt GH (2002) Physicians' and patients' choices in evidence based practice. *BMJ* 324: 1350
136. Heidenreich PA, Lee TT, Massie BM (1997) Effect of beta-blockade on mortality in patients with heart failure: a meta-analysis of randomized clinical trials. *J Am Coll Cardiol* 30: 27-34
137. Henshall C, Oortwijn W, Stevens A, et al. (1997) Priority setting for health technology assessment. Theoretical considerations and practical approaches. Priority setting subgroup of the EUR-ASSESS Project. *Int J Technol Assess Health Care* 13: 144-85
138. Herlitz J (1992) Comparison of lisinopril versus digoxin for congestive heart failure during maintenance diuretic therapy. The Lisinopril-Digoxin Study Group. *Am J Cardiol* 70(10): 84C-90C
139. Hood WB, Dans A, Guyatt GH, et al. (2001) Digitalis for treatment of congestive heart failure in patients in sinus rhythm. *Cochrane Database Systematic Review*
140. Hoppe UC, Böhm M, Dietz R, et al. in Zusammenarbeit mit der Arzneimittelkommission der Deutschen Ärzteschaft (2005) Leitlinien zur Therapie der chronischen Herzinsuffizienz. *Z Kardiologie* 94: 488–509

141. Irwig L, Bossuyt P, Glasziou P, et al. (2002) Designing studies to ensure that estimates of test accuracy are transferable. *BMJ* 324 (7338): 669-71.
142. Jaarsma T, Halfens R, Huijjer-Abu Saad H, et al. (1999) Quality of life in older patients with systolic and diastolic heart failure. *Eur J Heart Fail* 1: 151-60
143. Jackson G, Gibbs CR, Davies MK, Lip GYH (2000) ABC of heart failure. *Pathophysiology*. *BMJ* 320: 167-170
144. Jaeschke R, Oxman AD, Guyatt GH (1990) To what extent do congestive heart failure patients in sinus rhythm benefit from digoxin therapy? A systematic review and meta-analysis. *Am J Med* 88: 279-86
145. Jessup M & Brozena S (2003) Heart Failure. *N Engl J Med* 348: 2007-2018
146. Jong P, Demers C, McKelvie RS, Liu PP (2002) Angiotensin receptor blockers in heart failure: meta-analysis of randomized controlled trials. *J Am Coll Cardiol* 39: 463-70
147. Junger J, Schellberg D, Muller-Tasch T, et al. (2005) Depression increasingly predicts mortality in the course of congestive heart failure. *Eur J Heart Fail* 7(2): 261-7
148. Juurlink DN, Mamdani MM, Lee DS, et al. (2004) Rates of hyperkalemia after publication of the randomized aldactone evaluation study. *N Engl J Med* 351: 543-51
149. Khan KS, ter Riet G, Glanville J, et al. (eds.) (2001) Undertaking systematic reviews of research on effectiveness CRD's guidance for those carrying out or commissioning reviews. *CRD Report Number 4 (2nd Edition)*, ISBN 1 900640 20 1; URL: <http://www.york.ac.uk/inst/crd/report4.htm> (Internet-Zugriff vom 16.11.2004); p 4 (Phase 0), p 4 (Phase 6)
150. Khand AU, Rankin AC, Kaye GC, Cleland JG (2000) Systematic review of the management of atrial fibrillation in patients with heart failure. *Eur Heart J* 21: 614-32
151. Kjekshus J, Apetrei E, Barrios V, et al. (2007) Rosuvastatin in older patients with systolic heart failure. *N Engl J Med* 357(22): 2248-61
152. Knottnerus JA, van Weel C, Muris JW (2002) Evaluation of diagnostic procedures. *BMJ* 324 (7335): 477-80
153. Komajda M, Lapuerta P, Hermans N, et al. (2005) Adherence to guidelines is a predictor of outcome in chronic heart failure: the MAHLER survey. *Eur Heart J* 26: 1653-1659
154. Komajda M, Follath F, Swedberg K, et al, for The Study Group of Diagnosis of the Working Group on Heart Failure of the European Society of Cardiology (2003) The Euro Heart Failure survey programme – a survey on the quality of care among patients with heart failure in Europe. Part 2: treatment. *Eur Heart J* 24: 464-74

155. Kulig M, Schulte E, Willich SN (2003) Comparing methodological quality and consistency of international guidelines for the management of patients with chronic heart failure. *Eur J Heart Fail* 5: 327-335
156. Kunz R, Ollenschläger G, Raspe H, et al. (Hrsg.) (2007) Lehrbuch Evidenzbasierte Medizin in Klinik und Praxis. 2. überarbeitete und erweiterte Aufl. Deutscher Ärzte-Verlag Köln, S. 16, 24, 61, 249ff, 361ff
157. Kunz R und AG Glossar des Deutschen Netzwerks evidenzbasierter Medizin (2005) Glossar zur Evidenzbasierten Medizin. URL: <http://www.ebm-netzwerk.de/grundlagen/glossar> (Internet-Zugriff vom 18.04.2005)
158. Landis JR & Koch GG (1977) The measurement of observer agreement for categorial data. *Biometrics* 33: 159
159. Larsen T, Kumar S, Grimmer K, et al. (2007) A systematic review of guidelines for the prevention of heat illness in community-based sports participants and officials. *J Sci Med Sport* 10(1): 11-26
160. Last JM (edt) (2001) A dictionary of epidemiology. 4th Ed. Edited for the International Epidemiological Association. Oxford University Press Oxford, New York, pp 14-16, 37-38
161. Lechat P, Packer M, Chalon S et al. (1998) Clinical effects of β -adrenergic blockade in chronic heart failure: a meta-analysis of double-blind, placebo-controlled, randomized trials. *Circulation* 98: 1184-91
162. Lee VC, Rhew DC, Dylan M, et al. (2004) Meta-analysis: angiotensin-receptor blockers in chronic heart failure and high-risk acute myocardial infarction. *Ann Intern Med* 141(9): 693-704
163. Lee DS, Tu JV, Juurlink DN, et al. (2005) Risk-treatment mismatch in the pharmacotherapy of heart failure. *JAMA* 294: 1240-7
164. Lee DTF, Yu DSF, Woo J, Thompson DR (2003) Health-related quality of life in patients with congestive heart failure. *Eur J Heart Fail* 7: 419-422
165. Levy D, Kenchaiah S, Larson MG, et al. (2002) Long-term trends in the incidence of and survival with heart failure. *N Engl J Med* 2002 347(18): 1397-402
166. Lijmer JG, Mol BW, Heisterkamp S, et al. (1999) Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 282(11): 1061-1066
167. Linden A, Adams JL, Roberts N (2003) Evaluating disease management program effectiveness: an introduction to time-series analysis. *Disease Management* 6: 243-255
168. Linn S (2005) New patient-oriented diagnostic test characteristics analogous to the likelihood ratios conveyed information on trustworthiness. *J Clin Epidemiol* 58(5): 450-7
169. Lip GYH & Freestone B (2003) Cardiovascular disorders. Atrial fibrillation (acute). URL: <http://www.clinicalevidence.com/cweb/conditions/cvd/0210/0210.jsp> (Internet-Zugriff vom 12.02.2005)

170. Lip GYH, Gibbs CR, Beevers DG (2000) ABC of heart failure: aetiology. *BMJ* 320: 104-107
171. Lip GYH, Kamath S, Freestone B (2003) Effects of treatments for acute atrial fibrillation. *Clinical Evidence* 7: 1-10
172. Littlejohns P, Cluzeau F, Bale R, et al. (1999) The quality and quantity of clinical practice guidelines for the management of depression in primary care in the UK. *Br J Gen Pract* 49: 205-210
173. Lloyd-Williams F, Mair FS, Leitner M (2002) Exercise training and heart failure: a systematic review of current evidence. *Br J Gen Pract* 52: 47-55
174. Lubsen J, Chadha DR, Yotof YT, Swedberg K (1996) Meta-analysis of morbidity and mortality in five exercise capacity trials evaluating ramipril in chronic congestive cardiac failure. *Am J Cardiol* 77: 1191-6
175. Lund-Johansen P (1981) The Norwegian Multicenter Study on timolol after myocardial infarction. Part II. Effect in different risk groups, causes of death, heart arrest, reinfarctions, rehospitalizations and adverse experiences. *Acta Med Scand* 651: 243-52
176. Ma J & Boucher M (2002) Heart Failure: is there a role for angiotensin II receptor blockers? *Issues Emerg Health Technol* Issue 38, URL: http://www.cadth.ca/media/pdf/167_arbs_cetap_e.pdf (letzter Internet-Zugriff vom 03.01.2008)
177. MacLeod FE, Harrison MB, Graham ID (2002) The process of developing best practice guidelines for nurses in Ontario: risk assessment and prevention of pressure ulcers. *Ostomy Wound Manage* 48(10): 30-2, 34-8
178. Mancini GB (2000) Long-term use of angiotensin-converting enzyme inhibitors to modify endothelial dysfunction: a review of clinical investigations. *Clin Invest Med* 23: 144-61
179. Mann DL (1999) Mechanisms and models in heart failure. A combinatorial approach. *Circulation* 100: 999-1008
180. Manna DR, Bruijnzeels MA, Mokkink HGA, Berg M (2003) Ethnic specific recommendations in clinical practice guidelines: a first exploratory comparison between guidelines from the USA, Canada, the UK, and the Netherlands. *Qual Saf Health Care* 12: 353-358
181. Margolis CZ (1999) Developing and constructing practical guidelines. In: Margolis CZ & Cretin S (eds.) (1999) *Implementing clinical practice guidelines*. AHA-Press, Chicago, pp. XXI, 24-33
182. Masoudi FA (2007) Statins for ischemic systolic heart failure. *N Engl J Med* 357(22): 2248-61
183. McCrory DC, Colice GL, Lewis SZ, et al. (2003) Overview of the methodology for lung cancer evidence review and guideline development. *Chest* 123: 3S-6S
184. McFadden ER & Braunwald E (1992) Cor pulmonale. In: Braunwald E (Ed.) (1992) *Heart disease: a textbook of cardiovascular medicine*. 4th Edition. W.B.

Saunders Company, Philadelphia, London, Toronto, Montreal, Sydney, Tokyo, ISBN 0-7216-4409-0. Vol. 2, p. 1581

185. McGinn T, Wyer PC, Newman TB, et al., for the Evidence-Based Medicine Teaching Tips Working Group (2004) Tips for learners of evidence-based medicine: 3. Measures of observer variability (kappa statistic). *CMAJ* 171 (11): 1369-73
186. McIntosh A (1999) Guidance: what role should consensus play in guideline development? In: Hutchinson A & Baker R (Eds.) Making use of guidelines in clinical practice. Radcliff Medical Press Ltd., Abingdon, Oxon, pp 63-81
187. McKelvie RS, Yusuf S, Pericak D, et al. (1999) Comparison of candesartan, enalapril, and their combination in congestive heart failure: randomized evaluation of strategies for left ventricular dysfunction (RESOLVD) pilot study. The RESOLVD Pilot Study Investigators. *Circulation* 100: 1056-64
188. McKinlay E, McLeod D, Dowell T, Howden-Chapman P (2001) Clinical practice guidelines. A selective literature review. Report prepared by the Wellington School Of Medicine for The New Zealand Guidelines Group Inc. URL: <http://www.nzgg.org.nz> (Internet-Zugriff vom 20.09.2004)
189. McMurray JJV, Pfeffer MA (2005) Heart failure. *Lancet* 365: 1877-89
190. McMurray JJ & Stewart S (2000) Epidemiology, aetiology, and prognosis of heart failure. *Heart* 83(5): 596-602
191. Metra M, Torp-Pedersen C, Swedberg K, et al. (2005) Influence of heart rate, blood pressure, and beta-blocker dose on outcome and the differences in outcome between carvedilol and metoprolol tartrate in patients with chronic heart failure: results from the COMET trial. *Eur Heart J* 26(21): 2259-68
192. Moher D, Schulz KF, Altman DG für die CONSORT Gruppe (2004) Das CONSORT Statement: Überarbeitete Empfehlungen zur Qualitätsverbesserung von Reports randomisierter Studien im Parallel-Design. [The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomized trials 2001 – German version]. *Dtsch Med Wochenschr* 129: T16-T20
193. Molewijk AC, Stiggelbout AM, Otten W, Dupuis HM, Kievit J (2004) Empirical data and moral theory. A plea for integrated empirical ethics. *Medicine, Health-care and Philosophy* 7: 55-69
194. Molewijk AC, Stiggelbout AM, Otten W, Dupuis HM, Kievit J (2003) Implicit normativity in evidence-based medicine: a plea for integrated empirical ethics research. *Health Care Anal* 11(1): 69-92
195. Moore & McQuay 2003: Moore A & McQuay HJ (2003) What is an NNT? 'What is...?' bulletins 1 (1). Hayward Medical Communications. URL: <http://www.evidencebased-medicine.co.uk> (Internet-Zugriff vom 18.03.2004)

196. MRC Working Party (1992) Medical Research Council trial of treatment of hypertension in older adults: principal results. *BMJ* 304(6824): 405-12
197. Mulrow CD & Cook D (1998) Systematic reviews: synthesis of best evidence for health care decisions. Philadelphia, PA: American College of Physicians; In: Gartlehner G, West SL, Lohr KN, Kahwati L, Johnson JG, Harris RP, Whitener L, Voisin CE, Sutton S (2004) Assessing the need to update prevention guidelines: a comparison of two methods. *Int J Qual Health Care* 16(5): 399-406
198. Mulrow CD & Lohr KN (2001) Proof and policy from medical research evidence. *J Health Polit Pol* 26(2): 249-266
199. Murphy MK, Black NA, Lamping DL, et al. (1998) Consensus development methods, and their use in clinical guideline development. *Health Technol Assessment* 2(3). URL: http://www.hta.nhsweb.nhs.uk/ProjectData/3_project_record_published.asp?PitId=944 (Internet-Zugriff vom 15.11.2004)
200. Nanas JN, Alexopoulos G, Anastasiou-Nana MI, et al. for the High Enalapril Dose Study Group (2000) Outcome of patients with congestive heart failure treated with standard versus high doses of enalapril: a multicenter study. *J Am Coll Cardiol* 36: 2090-5
201. Naylor CD (1995) Grey zones of clinical practice: some limits to evidence-based medicine. *Lancet* 345: 840-842
202. Neal B, MacMahon S, Chapman N, Blood P (2000) Effects of ACE inhibitors, calcium antagonists, and other blood-pressure-lowering drugs: results of prospectively designed overviews of randomised trials. Blood Pressure Lowering Treatment Trialists' Collaboration. *Lancet* 356: 1955-64
203. New Zealand Guidelines Group (2001) Handbook for the preparation of explicit evidence-based clinical practice guidelines. URL: <http://www.nzgg.org.nz> (Internet-Zugriff vom 20.09.2004)
204. New Zealand Guidelines Group (2006) Notes on the Adaptation / Synthesis of Guidelines. Last update 28.09.2006, URL: <http://www.nzgg.org.nz/index.cfm?fuseaction=evidence&fusesubaction=article&documentID=10&articleID=54> (Internet-Zugriff vom 31.01.2007)
205. Niederstadt CJ, Doering TJ, Fischer GC (2001) Allgemeinmedizinische Leitlinien-Entwicklung: Problematik der dichotomisierenden Algorithmenbildung. *Z Arztl Fortbild Qualitatssich* 95(8): 561-6
206. Niederstadt CJ (2001) Entwicklung einer Evidenz-basierten Leitlinie zur Behandlung der Harninkontinenz erwachsener Patienten in der Hausarztpraxis. Dissertation zur Erlangung des Doktorgrades der Medizin der Medizinischen Hochschule Hannover. URL: http://deposit.ddb.de/cgi-bin/dokserv?idn=965812979&dok_var=d1&dok_ext=pdf&filename=965812979.pdf (Internet-Zugriff vom 02.05.2007)
207. Nielsen K (1979) On deriving an ought from an is: A retrospective look. *Review of Metaphysics* 32(3): 487-514

208. Oosterhuis WP, Bruns DE, Watine J, et al. (2004) Evidence-based guidelines in laboratory medicine: principles and methods. *Clin Chem* 50(5): 806-818
209. Packer M, Antonopoulos GV, Berlin JA, et al (2001) Comparative effects of carvedilol and metoprolol on left ventricular ejection fraction in heart failure: results of a meta-analysis. *Am Heart J* 141: 899-907
210. Packer M, Gheorghide M, Young JB, et al. (1993) Withdrawal of digoxin from patients with chronic heart failure treated with angiotensin converting enzyme inhibitors. *N Engl J Med* 329: 1-7
211. Packer M, Poole-Wilson PA, Armstrong PW, et al. on behalf of the ATLAS Study Group (1999) Comparative effects of low and high doses of the angiotensin-converting enzyme inhibitor, lisinopril, on morbidity and mortality in chronic heart failure. *Circulation* 100: 2312-2318
212. Pagliari C & Grimshaw J (2002) Impact of group structure and process on multidisciplinary evidence-based guideline development: an observational study. *J Eval Clin Pract* 8(2): 145-53
213. Pagliari C, Grimshaw J, Eccles M (2001) The potential influence of small group processes on guideline development. *J Eval Clin Pract* 7(2): 165-73
214. Perleth M (2001) Evidenzbasierte Entscheidungsunterstützung im Gesundheitswesen. Konzepte und Methoden der systematischen Bewertung medizinischer Technologien (Health Technology Assessment) in Deutschland. Hannover: Habilitationsschrift, Medizinische Hochschule Hannover
215. Perleth M & Beyer M (1996) Evidenzbasierte Medizin, die Cochrane Collaboration und der Umgang mit medizinischer Literatur. Implikationen für die ärztliche Aus-, Fort- und Weiterbildung. *Z ärztl Fortbild* 90: 67-73
216. Perleth M, Jakubowski E, Busse R (2000) „Best Practice“ im Gesundheitswesen oder warum wir evidenzbasierte Medizin, Leitlinien und Health Technology Assessment brauchen. *Z ärztl Fortbild Quallsich* 94: 741-744
217. Peterson MC, Holbrook JH, Hales D, et al. (1992) Contributions of the history, physical examination, and laboratory investigation in making medical diagnoses. *West J Med* 156: 163-165
218. Philips Z, Ginelly L, Sculpher M, et al. (2004) Review of guidelines for good practice in decision-analytic modelling in health technology assessment. NHS R&D HTA Programme. *Health Technology Assessment* 2004; Vol. 8: Nr. 36. URL: <http://www.ncchta.org/fullmono/mon836.pdf> (Internet-Zugriff vom 15.11.2004)
219. Piepoli MF, Davos C, Francis DP, Coats AJ; ExTraMATCH Collaborative (2004) Exercise training meta-analysis of trials in patients with chronic heart failure (ExTraMATCH). *BMJ* 328(7433): 189
220. Pilote L & Tager IB (2002) Outcomes research in the development and evaluation of practice guidelines. *BMC Health Services Research* 2: 7

221. Pitt B, Poole-Wilson PA, Segal R, et al. (2000) Effect of losartan compared with captopril on mortality in patients with symptomatic heart failure: randomised trial – the losartan heart failure survival study ELITE II. *Lancet* 355: 1582-7
222. Pitt B, Remme W, Zannad F, et al. (2003) Eplerenone, a selective aldosterone blocker, in patients with left ventricular dysfunction after myocardial infarction. *N Engl J Med* 348: 1309-21
223. Pitt B, Yusuf S, for the SOLVD Investigators (1992) Study of the left ventricular dysfunction (SOLVD) subgroup results. *J Am Coll Card* 19: 215A
224. Pitt B, Zannad F, Remme WJ, et al (1999) The effect of spironolactone on morbidity and mortality in patients with severe heart failure: Randomized Aldactone Evaluation Study Investigators. *N Engl J Med* 341: 709-17
225. Pogue J & Yusuf S (1998) Overcoming the limitations of current meta-analysis of randomised controlled trials. *Lancet* 351: 47–52
226. Poole C & Greenland S (1999) Random-effects meta-analyses are not always conservative. *Am J Epidemiol* 150: 469-475
227. Quinn M, Babb P, Brock A, et al. (2001) Cancer trends in England and Wales 1950–1999. The Stationery Office, London. Zitiert in: NICE 2003 (s.o.)
228. Ramsay LE, Wallis EJ, Yeo WW, Jackson PR (1998) The rationale for differing national recommendations for the treatment of hypertension. *Am J Hypertens* 11: 79S-88S
229. Raspe H (2005) Konzept und Methoden der Evidenz-basierten Medizin: Besonderheiten, Stärken, Grenzen, Schwächen und Kritik. URL: http://www.med.uni-magde-burg.de/fme/institute/ism/Lehre/Wahlfach/Hintergrundmaterial/konzepte_ebm_raspe.pdf (Zugriff vom 02.10.2007)
230. Rathore SS, Wang Y, Krumholz HM (2002) Sex-based differences in the effect of digoxin for the treatment of heart failure. *N Engl J Med* 347: 1403-11
231. Ravago TS, Mosniam J, Alem F (2000) Evaluation of community acquired pneumonia guidelines. *J Med Sys* 24: 289-296
232. Reid RJ, Haggerty JL, McKendry R (2002) Defusing the confusion: concepts and measures of continuity of healthcare. Canadian Health service Research Foundation / Fondation canadienne de la recherche sur les services de santé. URL: http://www.chsrf.ca/final_research/commissioned_research/programs/index_e.php, Zugriff vom 30.01.2006
233. Richardson WS, Wilson MC, Nishikawa J, Hayward RSA (1995) The well-built clinical question: a key to evidence-based decisions. *ACP J Club* 123: A-12
234. Rogers WA (2002) Are guidelines ethical? Some considerations for general practice. *Br J Gen Pract* 52: 663-669
235. Rosenberg W & Donald A (1995) Evidence based medicine: an approach to clinical problem-solving. *BMJ* 310: 1122-1126

236. Rothwell PM (2005) Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. Series: Treating Individuals 2. Lancet 365: 176-186
237. Royal College of Psychiatrists, London (1994) Clinical Practice Guidelines and Their Development. Council Report CR34. URL: <http://www.rcpsych.ac.uk/publications/cr/council/cr34.pdf> (Internet-Zugriff vom 21.08.2005)
238. Sackett DL & Haynes RB (2002) Evidence base of clinical diagnosis: The architecture of diagnostic research. BMJ 324; 539-541
239. Sackett DL, Rosenberg WMC, Gray JAM, et al. (1996) Evidence based medicine: What it is and what it isn't. BMJ 312: 71-72
240. Selbmann HK (2006) Ist Konsens bei der Leitlinienerstellung Nonsense? Z ärztl Fortbild Qual Gesundh wes 100: 536-541
241. Shaneyfelt TM, Mayo-Smith MF, Rothwangl J (1999) Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer-reviewed medical literature. JAMA 281(20): 1900-5
242. Sharma D, Buyse M, Pitt B, Rucinska EJ and the Losartan Heart Failure Mortality Meta-analysis Study Group (2000) Meta-analysis of observed mortality data from all-controlled, double-blind, multiple-dose studies of losartan in heart failure. Am J Cardiol 85: 187-92
243. Shekelle P, Eccles MP, Grimshaw JM, Woolf SH (2001) When should clinical guidelines be updated? BMJ 323: 155-157
244. Shekelle P, Rich M, Morton S, et al. (2003) Pharmacologic Management of Heart Failure and Left Ventricular Systolic Dysfunction: Effect in Female, Black, and Diabetic Patients, and Cost Effectiveness. Evidence Report/Technology Assessment No. 82 (Prepared by the Southern California-RAND Evidence-based Center under Contract No 290-97-0001). AHRQ Publication No. 03-E045. Rockville, MD: Agency for Healthcare Research and Quality. URL: <http://www.ahrq.gov/> (Internet-Zugriff vom 19.03.2004)
245. Shekelle PG, Ortiz E, Rhodes S, et al. (2001) Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated? JAMA 286(12): 1461-7
246. Shekelle PG, Woolf SH, Eccles M, Grimshaw J (1999) Clinical guidelines. Developing guidelines. BMJ 318: 593-596
247. Shibata MC, Flather MD, Wang D (2001) Systematic review of the impact of beta blockers on mortality and hospital admissions in heart failure. Eur J Heart Fail 3: 351-57
248. Siebert U, Muth C, Sroczynski G, Velasco-Garrido M, et al. (Hrsg.) (2003) Dünnschichtpräparationen und computergestützte Untersuchungen von Zervixabstrichen im Rahmen der Krebsfrüherkennung. Medizinische Effektivität, gesundheitsökonomische Evaluation und systematische Entscheidungsanalyse.

- Health Technology Assessment 35. Schriftenreihe des Deutschen Instituts für medizinische Dokumentation und Information im Auftrag des Bundesministeriums für Gesundheit und Soziale Sicherung. Asgard-Verlag, Sankt Augustin. ISBN 3-537-27035-6. Titel-Nummer 270035
249. Silagy CA, Stead LF, Lancaster T (2001) Use of systematic review in clinical practice guidelines: case study of smoking cessation. *BMJ* 323: 833-836
 250. Slowther A, Ford S, Schofield T (2004) Ethics of evidence based medicine in the primary care setting. *J Med Ethics* 30: 151-154
 251. Smith GD & Egger M (2001) Going beyond the grand mean: subgroup analysis in meta-analysis of randomized trials. In: Egger M, Smith GD, Altman DG (Editors) (2001) *Systematic Reviews in Health Care. Metaanalysis in Context*. 2nd Edition. BMJ Publishing Group, London, p 143-156
 252. Starfield BH, Simborg DW, Horn SD, Yourtee SA (1976) Continuity and coordination in primary care: their achievement and utility. *Med Care* 14: 625-36
 253. Statistisches Bundesamt (2006) Statistisches Bundesamt, persönliche Kommunikation G. Vierkant v. 28.08.2006. Aus dem Krankenhaus entlassene vollstationäre Patientinnen und Patienten (einschl. Sterbe- und Stundenfälle) 2004 nach Pflegetage, Geschlecht und Altersklassen, Anzahl für ICD-10 I50
 254. Statistisches Bundesamt (2007) Sterbefälle nach den 10 häufigsten Todesursachen insgesamt und nach Geschlecht 2005. URL: <http://www.destatis.de/basis/d/gesu/gesutab20.htm> (Abfrage vom 27.01.2007)
 255. Stewart AL, Greenfield S, Hays RD, et al. (1989) Functional status and well-being of patients with chronic conditions: results from the medical outcomes study. *JAMA* 262: 907-13
 256. Stewart S, MacIntyre K, Capewell S, McMurray JJ (2003) Heart failure and the aging population: an increasing burden in the 21st century? *Heart* 89: 49-53
 257. Strech D (2007) Vier Ebenen von Werturteilen in der medizinischen Nutzevaluation. Eine Systematik zur impliziten Normativität in der Evidenz-basierten Medizin. *Z Arztl Fortbild Qualitätssich* 101(7): 473-80
 258. Struthers AD (2005) Pathophysiology of heart failure following myocardial infarction. *Heart* 91 Suppl 2: ii14-6; discussion ii31, ii43-8
 259. Swedberg K, Kjekshus J, Snapinn S (1999) Long-term survival in severe heart failure in patients treated with enalapril. Ten year follow-up of CONSENSUS I. *Eur Heart J* 20(2): 136-9
 260. Taylor DO (2003) Modern concepts in heart failure. Etiology, pathophysiology, & recent guidelines for diagnosis and treatment of heart failure patients. The Cleveland Clinic Centre for Continuing Education – Presentation 13.03.2003. URL: <http://www.clevelandclinicmeded.com/modern/case1/Taylor.pdf> (Internet-Zugriff am 02.03.2004)
 261. Taylor R & Giles J (2005) Cash interests taint drug advise. *Nature* 437: 1070-71

262. The AGREE Collaboration (2002) Checkliste zur Qualitätsbeurteilung von Leitlinien (AGREE-Instrument) – Deutschsprachige Version. Ärztliche Zentralstelle Qualitätssicherung; Köln – Verbindung der Schweizer Ärztinnen und Ärzte FMH, Bern
263. The AGREE Collaboration (2003) Appraisal of Guidelines for Research & Evaluation (AGREE) instrument training manual. URL: <http://www.agreecollaboration.org/pdf/aitraining.pdf> (Internet-Zugriff vom 10.05.2004)
264. The AGREE Collaboration (2003a) Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. *Quality and Safety in Health Care* 12(1): 18-23
265. The Atrial Fibrillation Follow-Up Investigation Of Rhythm Management (AFFIRM) Investigators (2002) A Comparison Of Rate Control And Rhythm Control In Patients With Atrial Fibrillation. *N Engl J Med* 347: 1825-33
266. The BEST Steering Cmte (1995) Design of the Beta-blocker Evaluation Survival Trial. *Am J Cardiol* 75: 1220-1223
267. The Digitalis Investigation Group (1997) The effect of digoxin on mortality and morbidity in patients with heart failure. *N Engl J Med* 336: 525-33
268. The European Parliament And The Council Of The European Union (1998) Directive 98/79/EC of the European Parliament and of the Council of 27 October 1998 on in vitro diagnostic medical devices. URL: http://europa.eu.int/eur-lex/en/consleg/main/1998/en_1998L0079_index.html (Internet-Zugriff vom 13.03.2006)
269. The NETWORK investigators (1998) Clinical outcome with enalapril in symptomatic chronic heart failure; a dose comparison. *European Heart Journal* 19: 481-489
270. The SOLVD Investigators (1991) Effect of enalapril on survival in patients with reduced left ventricular ejection fractions and congestive heart failure. *N Engl J Med* 325(5): 293-302
271. Thomas LH, McColl E, Cullum N, et al. (1998) Effect of clinical guidelines in nursing, midwifery, and the therapies: a systematic review of evaluations. *Qual Health Care* 7(4): 183-191
272. Thomson R, McElroy H, Sudlow M (1998) Guidelines on anticoagulant treatment in atrial fibrillation in Great Britain: variation in content and implications for treatment. *BMJ* 316: 509-513
273. Uretsky BF, Young JB, Shahidi FE, et al. (1993) Randomized study assessing the effect of digoxin withdrawal in patients with mild to moderate congestive heart failure: results of the PROVED trial. *J Am Coll Cardiol* 22: 955-62
274. Van Campen LCMC, Visser FC, Visser CA (1998) Ejection fraction improvement by beta-blocker treatment in patients with heart failure: an analysis of studies published in the literature. *J Cardiovasc Pharmacol* 32(suppl 1): 31-35

275. Van Doorslaer E & Masseria C (2004) OECD Health Equity Research Group. Income-related inequality in the use of medical care in 21 OECD countries. In: To-towards high-performing health systems: policy studies. OECD, Paris, p. 109-66. URL: <http://www.oecd.org> (Internet-Zugriff vom 22.02.2006)
276. Van Tulder M, Becker A, Bekkering T, et al. (2006) Chapter 3. European guidelines for the management of acute nonspecific low back pain in primary care. Eur Spine J 15 Suppl 2:S169-91; URL: http://www.backpaineurope.org/web/files/WG1_Guidelines.pdf (Internet-Zugriff vom 31.01.2007)
277. Vasan RS & Levy D (2000) Defining diastolic heart failure. A call for standardized diagnostic criteria. Circulation 101: 2118-2121
278. Vasan RS, Larson MG, Benjamin EJ, et al. (1999) Congestive heart failure in subjects with normal versus reduced left ventricular ejection fraction. J Am Coll Cardiol 33: 1948-1955
279. Vineis P (2004) Evidence based medicine and ethics a practical approach. J Med Ethics 30: 126-130
280. Vlayen J, Aertgeerts B, Hannes K, et al. (2005) A systematic review of appraisal tools for clinical practice guidelines: multiple similarities and one common deficit. Int J Qual Health Care 17(3): 235-242
281. Voellinger R, Berney A, Baumann P, et al. (2003) Major depressive disorder in the general hospital: adaptation of clinical practice guidelines. Gen Hosp Psychiatry 25(3): 185-93
282. Wabitsch R, Margolis CZ, Malkin J-E, et al. (1998) Evaluating the process of tailoring the global HIV/AIDS practice guidelines for use in developing countries. Int J Qual Health Care 10: 147-154
283. Warner JG Jr, Metzger DC, Kitzman DW, et al. (1999) Losartan improves exercise tolerance in patients with diastolic dysfunction and a hypertensive response to exercise. J Am Coll Cardiol 33: 1567-72
284. Watson AM, Hood SG, May CN (2006) Mechanisms of sympathetic activation in heart failure. Clin Exp Pharmacol Physiol 33(12): 1269-74
285. Watt E, Burrell A, Duerden M (2003) Implementing NNTs. 'What is...?' bulletins 1 (7). Hayward Medical Communications. URL: <http://www.evidencebased-medicine.co.uk> (Internet-Zugriff vom 18.03.2004)
286. Weir RA, McMurray JV, Taylor J, Brady AJ (2006) Heart failure in older patients. Br J Cardiol 13: 257-66
287. Whorlow SL & Krum H (2000) Meta-analysis of effect of beta-blocker therapy on mortality in patients with New York Heart Association class IV chronic congestive heart failure. Am J Cardiol 86: 886-9
288. Woolf SH (2000) Evidence-based medicine and practice guidelines: an overview. Cancer Control 7 (4): 362-367

289. Woolf SH (1990) Practice guidelines: a new reality in medicine. I. Recent developments. Arch Intern Med 150: 1811-1818
290. Woolf S, DiGiuseppi CG, Atkins D, Kamerow DB (1996) Developing evidence-based clinical practice guidelines: lessons learned by the US Preventive Services Task Force. Ann Rev Publ Health 17: 511-538
291. Woolf SH, Grol R, Hutchinson A, et al. (1999) Clinical guidelines. Potential benefits, limitations, and harms of clinical guidelines. BMJ 318: 527-530
292. Worrall G, Chaulk P, Freake D (1997) The effects of clinical practice guidelines on patient outcomes in primary care: a systematic review. Can Med Assoc J 156(12): 1705-1712
293. Wörtz M, Perleth M, Schöffski O, Schwartz FW (2003) Innovative Medizinprodukte im deutschen Gesundheitswesen – Wege und Verfahren der Bewertung im Hinblick auf Regelungen zur Marktzulassung und Kostenübernahme von innovativen Medizinprodukten. BVMed, Norderstedt, S. 23ff. URL: http://www.bvmed.de/linebreak4/mod/netmedia_pdf/data/innovative_medizinprodukte.pdf (Internet-Zugriff vom 03.01.2008)
294. Wright J, Williams R, Wilkinson JR (1998) Development and importance of health needs assessment. BMJ 316(7140): 1310-3
295. Yancy CW (2001) Clinical trials of β -blockers in heart failure: a class review. Am J Med 110: 7S-10S
296. Yusuf S, Nicklas JM, Timmis G, et al. (1992) Effect of enalapril on mortality and the development of heart failure in asymptomatic patients with reduced left ventricular ejection fractions. The SOLVD investigators. N Engl J Med 327(10): 685-691
297. Zannad F, van den Broek SA, Bory M (1992) Comparison of treatment with lisinopril versus enalapril for congestive heart failure. Am J Cardiol 70(10): 78C-83C
298. Zentralinstitut (ZI) der Kassenärztlichen Bundesvereinigung (KBV) (2006) Daten des ADT-Panels zur Diagnose I50 Herzinsuffizienz im Quartal IV/2005. Persönliche Kommunikation Dr. Gerhard Brenner vom 30.08.2006
299. Zile MR, Baicu CF, Gaasch WH (2004) Diastolic heart failure – abnormalities in active relaxation and passive stiffness of the left ventricle. N Engl J Med 350: 1953-1959

E - Anhang

1 Anhänge des Methodikkapitels des Hauptdokuments

1.1 Recherchestrategie

Suchstrategie für die elektronische Suche bei MEDLINE, Recherche am 11.03.2004:

Search	Most Recent Queries	Result
#10	Search #9 AND 2000:2004 [dp]	289
#9	Search #5 AND #8	457
#8	Search #6 OR #7	73409
#7	Search heart failure, congestive [mh]	47973
#6	Search "heart failure"	68245
#5	Search #1 OR #2 OR #3 OR #4	50837
#4	Search health planning guidelines/	1733
#3	Search guidelines [mh]	37857
#2	Search practice guideline [pt]	7698
#1	Search guideline [pt]	11649

1.2 Verzeichnis besuchter Internetadressen

Tab. A.1: Institutionen, deren Homepages bei der systematischen Leitlinienrecherche besucht wurden – Nationale Organisationen

Abkürzung	Name und Internetadresse der Organisation
ÄZQ	Ärztliches Zentrum für Qualität in der Medizin http://www.aezq.de Leitlinienseite: http://www.leitlinien.de/leitlinienthemmen/index/view?show=11
AWMF	Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften http://www.uni-duesseldorf.de/WWW/AWMF/gs/gs_list.htm
DAHTA@DIMDI	Deutsches Institut für Dokumentation und Information http://www.dimdi.de
DEGAM	Deutsche Gesellschaft für Allgemein- und Familienmedizin http://www.degam.de/

(Fortsetzung Tab. A.1)

Abkürzung	Name und Internetadresse der Organisation
DGIM	Deutsche Gesellschaft für Innere Medizin e.V. http://www.dgim.de/
DGK	Deutsche Gesellschaft für Kardiologie – Herz- und Kreislaufforschung http://www.dgk.org/
KBV	Kassenärztliche Bundesvereinigung, HTA-Berichte: http://www.kbv.de/hta/hta.htm

Tab. A.2: Institutionen, deren Homepages bei der systematischen Leitlinienrecherche besucht wurden – Internationale Organisationen

Abkürzung	Name und Herkunftsland der Institution, Internet-Adresse
ACC	American College of Cardiology, http://www.acc.org
AETMIS	Agence d'Evaluation des Technologies et des Modes d'Intervention en Santé (Kanada), http://www.aetmis.gouv.qc.ca/fr/
AHFMR	Alberta Heritage Foundation for Medical Research (Kanada), http://www.ahfmr.ab.ca/
AHRQ [AHCPR]	Agency for Healthcare Research and Quality, http://www.ahrq.gov (vormals Agency for Health Care Policy and Research (USA))
AMA	Alberta Medical Association (Kanada), http://www.albertadoctors.org/
CAHTA	Catalan Agency for Health Technology Assessment and Research (Spanien), http://www.aatrm.net/ang/ang.html
CCOHTA	Canadian Coordinating Office for Health Technology Assessment (Kanada), http://www.ccohta.ca/
CDHSH	Commonwealth Department of Human Services and Health (Australien), www.health.gov.au
CEDIT	Comité d'Evaluation et de Diffusion des Innovations Technologiques (Frankreich), http://cedit.aphp.fr/english/index_present.html
CTFPHC	Canadian Task Force on Preventive Health Care (Kanada), http://www.ctfphc.org/
Duodecim	Leitlinienseite von The Finnish Medical Society Duodecim (Finnland), http://www.ebm-guidelines.com/home.html
GIN	Guideline International Network, http://www.g-i-n.net
GR	Gezondheidsraad (Niederlande), http://www.gr.nl/
HHS	Unites States Department of Health and Human Services (USA), http://www.hhs.gov
ICSI	Institute for Clinical Systems Improvement (USA), http://www.icsi.org
INAHTA	International Network of Agencies for HTA (ehemalige internationale HTA-Organisation, heute HTAI – Health Technology Assessment International), http://www.inahta.org
ITA	Institut für Technikfolgen-Abschätzung (Österreich), http://www.oeaw.ac.at/ita/hta/
MSAC	Medical Services Advisory Committee (Australien), http://www.msac.gov.au/

(Fortsetzung Tab. A.2)

Abkürzung	Name und Herkunftsland der Institution, Internet-Adresse
NCCHTA	National Coordinating Centre for Health Technology Assessment (UK), http://www.hta.nhsweb.nhs.uk/
NGC	National Guideline Clearinghouse (Organisation der AHQR, USA), http://www.guideline.gov
NHG	Nederlands Huisartsen Genootschap (NL), http://nhg.artsennet.nl
NHS	National Health Services (UK), http://www.nhsdirect.nhs.uk
NHSC	National Horizon Scanning Centre (UK), http://www.publichealth.bham.ac.uk/horizon/
NHS QIS	NHS Quality Improvement Scotland (UK), http://www.nhshealthquality.org/nhsqis/nhsqis_sub_publications.jsp
NICE	National Institute for Clinical Excellence (UK), http://www.nice.org.uk/
NQMC	National Quality Measures Clearinghouse (Organisation der AHQR, s.o.), http://www.qualitymeasures.ahrq.gov
NZGG	New Zealand Guideline Group (Neuseeland), http://nzgg.org.nz
NZHTA	New Zealand Health Technology Assessment (Neuseeland), http://nzhta.chmeds.ac.nz/
SBU	The Swedish Council on Technology Assessment in Health Care (Schweden), http://www.sbu.se/www/index.asp
SMM	Senter for medisinsk metodevurdering (norwegische HTA-Organisation), http://www.sintef.no/smm/Publications/Engsmdrag/FramesetPublications.htm
SNHTA	Swiss Network of Health Technology Assessment (Schweiz), http://www.snhta.ch
TA-SWISS	Zentrum für Technikfolgenabschätzung (Schweiz), http://www.ta-swiss.ch/framesets/home-d.htm
TNO	Nederlandse Organisatie voor toegepast-natuurwetenschappelijk onderzoek, TNO (Organisation für angewandte Wissenschaft und Forschung, Niederlande), http://www.tno.nl/homepage.html
USPSTF	US Preventive Task Force (USA), http://www.ahrq.gov/clinic/uspstfix.htm
VATAP	VA Technology Assessment Program, Department of Veterans Affairs (USA), http://www.va.gov/vatap/publications.htm
WVVH	Wetenschappelijke Vereniging van Vlaamse Huisartsen (Belgien), http://www.wvvh.be
ZonMw	Netherlands Organization for Health Research and Development (NL), http://www.zonmw.nl/index.asp?s=4535

1.3 Das AGREE-Instrument

Tab. A.3: Das AGREE-Instrument – Domänen und Items

<p>Domäne 1: Geltungsbereich und Zweck:</p> <p>(1) Das Gesamtziel der Leitlinie ist differenziert beschrieben.</p> <p>(2) Die in der Leitlinie behandelten medizinischen Fragen sind differenziert beschrieben.</p> <p>(3) Die Patienten, für die die Leitlinie gelten soll, sind spezifisch beschrieben.</p>
<p>Domäne 2: Beteiligung von Interessengruppen</p> <p>(4) Die Entwicklergruppe der Leitlinie schließt Mitglieder aller relevanten Berufsgruppen ein.</p> <p>(5) Die Ansichten und Präferenzen der Patienten wurden ermittelt.</p> <p>(6) Die Anwenderzielgruppe der Leitlinie ist definiert.</p> <p>(7) Die Leitlinie wurde in einem Pilotversuch von Mitgliedern der Anwenderzielgruppe getestet.</p>
<p>Domäne 3: Methodologische Exaktheit der Leitlinienentwicklung</p> <p>(8) Bei der Suche nach der Evidenz wurden systematische Methoden angewandt.</p> <p>(9) Die Kriterien für die Auswahl der Evidenz sind klar beschrieben.</p> <p>(10) Die zur Formulierung der Empfehlung verwendeten Methoden sind klar beschrieben.</p> <p>(11) Gesundheitlicher Nutzen, Nebenwirkungen und Risiken wurden bei der Formulierung der Empfehlungen berücksichtigt.</p> <p>(12) Die Verbindungen zwischen Empfehlungen und der zugrundeliegenden Evidenz sind explizit dargestellt.</p> <p>(13) Die Leitlinie ist vor ihrer Veröffentlichung von externen Experten begutachtet worden.</p> <p>(14) Ein Verfahren zur Aktualisierung der Leitlinie wird zur Verfügung gestellt.</p>
<p>Domäne 4: Klarheit und Gestaltung</p> <p>(15) Die Empfehlungen der Leitlinie sind spezifisch und eindeutig.</p> <p>(16) Die verschiedenen Handlungsoptionen für das Versorgungsproblem sind dargestellt.</p> <p>(17) Schlüsselempfehlungen der Leitlinie sind leicht zu identifizieren.</p> <p>(18) Es existieren Instrumente bzw. Materialien, die die Anwendung der Leitlinie unterstützen.</p>
<p>Domäne 5: Anwendbarkeit</p> <p>(19) Die möglichen organisatorischen Barrieren gegenüber der Anwendung der Empfehlungen werden diskutiert.</p> <p>(20) Die durch die Anwendung der Leitlinie möglicherweise entstehenden finanziellen Auswirkungen werden berücksichtigt.</p> <p>(21) Die Leitlinie benennt wesentliche Messgrößen für das Monitoring und / oder Überprüfungskriterien.</p>
<p>Domäne 6: Redaktionelle Unabhängigkeit</p> <p>(22) Die Leitlinie ist redaktionell von der (den) finanzierenden Organisation(en) unabhängig.</p> <p>(23) Interessenkonflikte von Mitgliedern der Leitlinienentwicklungsgruppe wurden dokumentiert.</p>

Quelle: [AGREE 2002], modifiziert durch [AWMF & ÄZQ 2003]

2 Anhänge des Ergebniskapitels des Hauptdokuments

2.1 Ausgeschlossene Leitlinien

Tab. A.4: Ausgeschlossene Leitlinien unter Nennung des Ausschlussgrundes

Leitlinie (Quelle)	Ausschlussgrund
1. ABFP 2001 : American Board of Family Practice (2001) Chronic Heart Failure. http://www.familypractice.com/references/guidesframe.htm	Leitlinie ist nicht evidenzbasiert*
2. ACC 1999 : American College of Cardiology (1999) Consensus recommendations for the management of chronic heart failure. American Journal of Cardiology 83;1A-38A	Leitlinie nach 2000 nicht aktualisiert
3. AHCPR 1996 : Heart Failure: Evaluation and Care of Patients With Left-Ventricular Systolic Dysfunction. Clinical Practice Guideline Number 11. AHCPR Publication No. 94-0612. http://hstat.nlm.nih.gov/hq/Hquest/db/local.arahcpr.arclin.lvdc/screeen/TocDisplay/da/1/s/47037/action/Toc;jsessionid=9D9CAB3AC792D15C6A79A780B204AB37 Living With Heart Disease: Is It Heart Failure? Consumer Guideline Number 11. AHCPR Publication No. 94-0614. National Library of Medicine DOCLINE Information: CAT/9436359 http://hstat.nlm.nih.gov/hq/Hquest/db/local.arahcpr.aruser.lvdpscreen/TocDisplay/s/47037/action/Toc	Dokument nicht länger als Leitlinie geführt, da veraltet
4. AMDA 2002 : American Medical Directors Association (2002) Heart Failure. Clinical Practice Guidelines No. 6. http://www.amda.com/info/cpg/heartfailure.htm	Leitlinie ist nicht evidenzbasiert*
5. BCC 2001 : Brisbane Cardiac Consortium. (2001) Clinical Practice Guideline: Hospital management of congestive heart failure. http://www.health.qld.gov.au/bcc/guidelines/CHF%20HospGuidelines2%281%29.pdf , http://www.health.qld.gov.au/bcc/clinical_practice.htm	Leitlinie bezieht sich ausschließlich auf das stationäre Management von Patienten
6. Brito 2001 : Brito D (2001) [Guidelines for the diagnosis and treatment of chronic heart failure] Rev Port Cardiol 20(10): 1049-51	Leitlinie in Portugiesisch
7. GICRP 2003 : Task Force per le Attivita di Psicologia in Cardiologia Riabilitativa e Preventiva, Gruppo Italiano di Cardiologia Riabilitativa e Preventiva (2003) [Guidelines for psychology activities in cardiologic rehabilitation and prevention]. Monaldi Arch Chest Dis 60(3): 184-234	Leitlinie in Italienisch
8. HFSA 1999 : Heart Failure Society of America (1999) HFSA guidelines for management of patients with heart failure caused by left ventricular systolic dysfunction: Pharmacological approaches. http://www.hfsa.org/pdf/lvsd_heart_failure.pdf	Leitlinie nach 2000 nicht aktualisiert, Update für 2004 angekündigt
9. KPNC 2000 : Kaiser Permanente Northern California (2000) Clinical Practice Guideline For Heart Failure Due To Left-Ventricular Systolic Dysfunction. http://www.kaiserpapers.org/leftven.html	Leitlinie nicht evidenzbasiert*

(Fortsetzung Tab. A.4)

Leitlinie (Quelle)	Ausschlussgrund
<p>10. MSSM 2003: Mount Sinai School of Medicine: Galin I & Baran DA (2003) Congestive Heart Failure: Guidelines for the Primary Care Physician. The Mount Sinai Journal Of Medicine 70(4): 251-264 Und: Gomberg-Maitland M, Baran DA, Fuster V (2001) Treatment of Congestive Heart Failure. Guidelines for the Primary Care Physician and the Heart Failure Specialist. Arch Intern Med 161: 342-352</p>	<p>Leitlinie nicht evidenzbasiert*</p>
<p>11. Navazio 2000: Navazio F (2000) [New guidelines in the treatment of congestive heart failure] Clin Ter 151(3):179-81</p>	<p>Leitlinie in Italienisch</p>
<p>12. NHF/NZ 1997: National Heart Foundation of New Zealand (1997) New Zealand guidelines for the management of chronic heart failure NZ Med J 99-107</p>	<p>Leitlinie veraltet – neue Version berücksichtigt</p>
<p>13. NHG 1995. Walma EP, Bakx HCA, Besselink RAM, Hamstra PWJ, Hendrick JMA, Koote JHA, van Veelen AWC, Vink R und Geijer RMM (1995) NHG-Standaard Hartfalen. M51. http://nhg.artsennet.nl/upload/104/standaarden/M51/start.htm</p>	<p>Leitlinie nach 2000 nicht aktualisiert, nicht evid.-basiert* LL in Holländisch</p>
<p>14. Ohkusa 2003: Ohkusa T, Matsuzaki M (2003) [Guidelines for the evaluation and management of chronic heart failure] Nippon Rinsho 61(5):723-30. Und: Matsuzaki M, Okusa T (2002) [Practice guidelines for patients with chronic heart failure] Nippon Naika Gakkai Zasshi 10; 91(3):973-90.</p>	<p>Leitlinie in Japanisch</p>
<p>15. PRODIGY 2004: PRODIGY Guidance – Heart failure. http://www.prodigy.nhs.uk/guidance.asp?gt=Heart%20Failure Patient Information Leaflet Heart Failure. http://www.prodigy.nhs.uk/clinicalguidance/releasedguidance/webBrowser/pils/PL83.htm</p>	<p>Leitlinie nicht evidenzbasiert*</p>
<p>16. SAMA 1998: South African Medical Association Heart Failure Working Group. Heart failure clinical guideline. South African Medical Journal 1998;88:1135-55.</p>	<p>Leitlinie nach 2000 nicht aktualisiert</p>
<p>17. SGK 2002: Schweizerische Gesellschaft für Kardiologie. Empfehlungen zur Diagnose und Behandlung der chronischen Herzinsuffizienz. Schweizerische Ärztezeitung 2002;1233-42.</p>	<p>Keine originäre LL, basiert auf der LL der ESC 2001</p>
<p>18. SIGN 1999: Scottish Intercollegiate Guidelines Network. Diagnosis and treatment of heart failure due to left ventricular systolic dysfunction. http://www.sign.ac.uk/pdf/sign35.pdf, http://www.sign.ac.uk/pdf/grg35.pdf, http://www.sign.ac.uk/patients/index.html</p>	<p>Leitlinie nach 2000 nicht aktualisiert</p>
<p>19. Thilly 2003 : Thilly N, Zannad F, Dufay E, Juilliere Y, Briancon S; Service d'Epidemiologie et Evaluation Cliniques (UPRES EA1124), Hôpital Marin, Nancy, France. [Angiotensin-converting enzyme inhibitors in congestive heart failure: clinical practice guidelines] Therapie. 2003; 58(4): 341-9.</p>	<p>Leitlinie in Französisch</p>
<p>20. UMHS 1999: University of Michigan Health System. Heart failure: systolic dysfunction, Guidelines for clinical care. http://www.cme.med.umich.edu/pdf/guideline/heart.pdf</p>	<p>Leitlinie nach 2000 nicht aktualisiert</p>

(Fortsetzung Tab. A.4)

Leitlinie (Quelle)	Ausschlussgrund
21. VPQ 1999 : Vermont Program for Quality in Health Care. Heart failure in Vermont. http://www.vpqhc.org	Leitlinie nach 2000 nicht aktualisiert, Leitlinie vom Server entfernt
22. WHO 1995 : World Health Organization. Concise guide to the management of heart failure. A patient's guide to control of heart failure. http://www.who.int/cardiovascular_diseases/resources/publications/en/print.html	Leitlinie nach 2000 nicht aktualisiert, Leitlinie vom Server entfernt

**Fehlende Evidenzbasierung nach den Kriterien, die unter Abschnitt 3.2.2 des Hauptdokuments expliziert wurden*

2.2 Weitere Ergebnisse der Bewertung der formalen methodischen Qualität berücksichtigter Leitlinien

In Ergänzung zu Kapitel 4.3 des Hauptdokuments folgt hier die ausführliche Bewertung der berücksichtigten Leitlinien mittels AGREE-Instrument:

Domäne 1: Geltungsbereich und Zweck

Item 1 – Gesamtziel: Mit einer Ausnahme [UWH 2001] nannten die Leitlinien mindestens allgemeine Ziele zur Optimierung der medizinischen Versorgung von Patienten mit Herzinsuffizienz, wie z.B. die Bereitstellung eines praktischen Leitfadens zur Diagnostik, Beurteilung und Behandlung der Herzinsuffizienz. Die Hälfte aller Leitlinien beschrieben spezifische Ziele, wie bspw. Letalitäts-senkung, Progressionshemmung etc., eine Leitlinie [ICSI 2003] führte eingangs ein Set aus messbaren und spezifischen Zielen an.

Item 2 – Beschreibung medizinischer Fragen / Probleme: Alle Leitlinien wiesen eine klare, durch Kapitelüberschriften unterteilte Gliederung auf, spezifische Fragestellungen zu klinischen Entscheidungssituationen fehlten bei 4/12 Leitlinien (25%), wurden bei 5/12 (31%) Leitlinien zu einzelnen Schlüsselempfehlungen, bei 3/16 (19%) Leitlinien zu den meisten Schlüsselempfehlungen und bei drei Leitlinien durchgängig expliziert [DVA&VHA 2002, ICSI 2003, NICE 2003].

Item 3 – Geltungsbereich Patienten: Bei 9/16 (56%) Leitlinien wurde die Patientenzielgruppe mindestens benannt, bei 6/16 (37,5%) wurden zusätzliche Ein- und Ausschlusskriterien von Merkmalen aufgeführt, bei zwei Leitlinien (12,5%) [ACC/AHA 2001, NICE 2003] wurde darüber hinaus ein klarer Bezug zum Versorgungsbereich hergestellt.

Zusammenfassung Domäne 1: 7/16 (44%) Leitlinien erreichten standardisierte Domänenscores von 0,5 und darüber. Alle eingeschlossenen Leitlinien beschrieben Geltungsbereich und Zweck hinreichend genau, insbesondere die Zieldefinition erfolgte bei der Mehrzahl der Leitlinien sehr spezifisch. Schwächen traten bei der genauen Beschreibung der Patientenzielgruppe auf – nur wenige Leitlinien schlossen Patienten explizit aus, für welche die Leitlinienempfehlungen nicht zutreffen, wie z.B. Patienten mit isolierter Rechtsherzinsuffizienz.

Domäne 2: Beteiligung von Interessengruppen

Item 4 – Einschluss relevanter Berufsgruppen: Bei 13/16 (81%) Leitlinien wurden die Autoren mit Namen und zugehöriger Fachgruppe aufgeführt – 12 arbeiteten mit interdisziplinären Entwicklungsteams. Bei der Hälfte der Leitlinien war darüber hinaus eine Zuordnung zur Funktion bei der Leitlinienentwicklung möglich, bei einem Viertel wurde auch über die Einbeziehung nicht-ärztlichen Personals berichtet [ACC/AHA 2001, DVA & VHA 2002, ICSI 2003, NICE 2003].

Item 5 – Patientenpräferenzen: Zwei Leitlinien bezogen Patientenvertreter in den Reviewprozess ein [ACC/AHA 2001; NHF/NZ 2001], zwei weitere Leitlinien beteiligten Patientenvertreter direkt am Entwicklungsprozess [NHF/Austr. & CSANZ 2002, NICE 2003]. Bei Dreiviertel der Leitlinien (12/16) wurden keine Verfahren zur Berücksichtigung von Patientenpräferenzen berichtet.

Item 6 – Definition der Anwenderzielgruppe: Bei 13/16 (81%) Leitlinien wurden Anwenderzielgruppen mindestens benannt, bei 6/16 (37,5%) wurde auch auf sekundäre Adressaten bei Kostenträgern, Ausbildung und Forschung verwiesen. Nur bei drei Leitlinien fehlte die Nennung der Anwenderzielgruppe [DGK 2001, NHF/NZ 2001, UWH 2001].

Item 7 – Pilotierung: 5/16 Leitlinien (31%) berichteten vom Verfahren einer Pilotierung, davon explizierte eine Leitlinie die Vorgehensweise und die Ergebnisse der Auswertung [LLGH 2003], bei den übrigen fehlten Angaben dazu.

Zusammenfassung Domäne 2: In der Gesamtbewertung der Beteiligung von Interessengruppen erzielten 6/16 (38%) Leitlinien standardisierte Domänenscores von 0,5 und darüber. Während die überwiegende Mehrheit von Leitlinien in interdisziplinären Gruppen entstand und ihre Anwenderzielgruppen gut definierte, wurden die Patientenpräferenzen nur bei einem Viertel der Leitlinien nachvollziehbar berücksichtigt.

Domäne 3: Methodische Exaktheit der Leitlinienentwicklung

Item 8 – Systematische Evidenzrecherche: Bei 13/16 (81%) Leitlinien lagen mindestens Angaben vor, die auf ein Verfahren zur systematischen Evidenzgewinnung schließen ließen, drei dieser Leitlinien explizierten darüber hinaus auch verwendete Datenbanken, Suchzeiträume oder Suchstrategien, bei einer Leitlinie [DVA&VHA 2002] wurden auch die Ergebnisse der Suche als Trefferzahlen ausgewiesen. Bei 9/16 (56%) Leitlinien fehlte nur eine der erforderlichen Angaben zur Vergabe des Likert-Scores „3“ (verwendete Datenbank und Suchzeitraum und Suchstrategie), sie wurden daraufhin mit „2“ geratet. 3/16 (19%) Leitlinien machten keine Angaben zum Vorgehen bei der Recherche bzw. es wurden keine systematischen Suchen vorgenommen.

Item 9 – Auswahlkriterien: 4/16 (25%) Leitlinien berichteten formale Ein- und Ausschlusskriterien für die Berücksichtigung identifizierter Dokumente der Literaturrecherche, bei den übrigen fehlten jegliche Angaben zur Auswahl der Evidenz. In keiner der eingeschlossenen Leitlinien wurden in Bezug auf spezielle Empfehlungen ausgeschlossene Publikationen aufgeführt oder Gründe, die zu deren Ausschluss führten.

Item 10 – Konsentierung: In 12/16 (75%) Leitlinien lagen Informationen vor, dass eine Konsensbildung stattgefunden hatte, 4/16 (25%) Leitlinien machten keine Angaben. Bei 2/12 konsentierten Leitlinien wurden die formalen Konsens-techniken berichtet (Delphi-Technik bei CCS [2002/3] und *RAND/UCLA Appropriateness Method* (RAM) bei NICE [2003]).

Item 11 – Risiko-Nutzen-Darstellung: In allen berücksichtigten Leitlinien fanden sich Angaben zu potentielltem Nutzen und zu Risiken der Schlüsselempfehlungen, wobei mehr als ein Ergebnisparameter herangezogen wurde. Bei 8/16 (50%) Leitlinien wurden die Empfehlungen mindestens teilweise, bei 6/16 (37,5%) konsequent mit Maßzahlen, wie *Numbers Needed to Treat* (NNT), verknüpft.

Item 12 – Evidenzverknüpfung: Alle berücksichtigten Leitlinien enthielten überwiegend Verknüpfungen zwischen Empfehlungen und zugrundeliegender Evidenz (Einschlusskriterium für vorliegende Bewertung), mehr als zwei Drittel (11/16) durchgängig.

Item 13 – Externe Begutachtung: 11/16 (69%) Leitlinien wurden vor ihrer Publikation extern begutachtet, davon führten 10 Leitlinien die Gutachter / begutachtende(n) Organisation(en) namentlich auf und machten darüber hinaus Angaben zur Vorgehensweise. Bei zwei Leitlinien wurde auf institutionelle Standards verwiesen [DVA & VHA 2002; NICE 2003]. Die Ergebnisse des Reviewprozesses waren bei keiner Leitlinie öffentlich zugänglich – die Ergebnisdarstellung erfolgte bei NICE internetgestützt und wurde nach Publikation der fertigen Leitlinie von der Homepage entfernt. Bei den deutschen Leitlinien (5/16)

fehlten konkrete Angaben, die auf die Durchführung eines formalen externen Reviewprozesses schließen ließen. Bei DieM [2003/2004] wurde ein internetgestützter offener Feedback-Prozess berichtet, nähere Angaben dazu fehlten jedoch auch hier.

Item 14 – Gültigkeit: 12/16 (75%) Leitlinien nannten mindestens einen Zeitpunkt zur Aktualisierung, 7/12 (44%) machten ergänzende Angaben zur Verantwortlichkeit. 5/12 (31%) Leitlinien werden darüber hinaus im Rahmen etablierter Programme revidiert – auf den Homepages der Leitlinienhersteller waren detaillierte Angaben in separaten Publikationen erhältlich [DieM 2003 / 2004, Duodecim 2004, DVA & VHA 2002, ICSI 2003a, NICE 2003].

Zusammenfassung Domäne 3: 7/16 (44%) Leitlinien zeigten standardisierte Domänenscores von 0,5 und darüber. Hohe Scores wurden insbesondere bei der Risiko-Nutzen-Darstellung mit mehrdimensionalen Ergebnisparametern sowie bei der Verknüpfung von Empfehlungen mit Evidenz ermittelt, die geringsten Scores bei der Darlegung der Auswahl zugrundeliegender Evidenz.

Domäne 4: Klarheit und Gestaltung

Item 15 – Eindeutigkeit der Empfehlungen: Bezogen auf den jeweiligen Empfehlungsbereich der Leitlinie (vgl. Tab. 4.1, Kap. 4.2.1) waren alle berücksichtigten Leitlinien mindestens „überwiegend eindeutig“ in der Formulierung ihrer Empfehlungen, mehr als zwei Drittel (11/16, 69%) durchgängig eindeutig.

Item 16 – Handlungsoptionen: Wegen der heterogenen Empfehlungsbereiche – einige Leitlinien behandelten ausschließlich die Pharmakotherapie der Herzinsuffizienz, andere erfassten die gesamte Versorgung (vgl. Tab. 4.1, Kap. 4.2.1) – war eine Wertung der angegebenen Handlungsoptionen erschwert. Wurde der jeweilige Empfehlungsbereich zugrundegelegt, wurden in allen Leitlinien zumindest in einigen Bereichen verschiedene Handlungsoptionen angegeben. Bei 7/16 (48%) wurden Handlungsoptionen in den meisten Bereichen aufgezeigt. Bei weiteren 7/16 Leitlinien fand sich durchgängig eine klare Darstellung der Handlungsoptionen mit Auflistung von Entscheidungskriterien und Evidenzbezügen.

Item 17 – Identifizierung von Schlüsselempfehlungen: Alle berücksichtigten Leitlinien enthielten zumindest einige, leicht zu identifizierende Schlüsselempfehlungen, bei der Mehrzahl (14/16, 88%) waren die Schlüsselempfehlungen durchgängig im Text hervorgehoben und leicht erkennbar. Bei 9/16 (56%) Leitlinien wurden die Empfehlungen im Fließtext durch Tabellen, besonders hervorgehobene Textabschnitte sowie durch Flussdiagramme unterstützt.

Item 18 – Tools: 13/16 (81%) Leitlinien enthielten weitere Materialien, welche die Anwendung unterstützten – in der Mehrzahl graphische Aufbereitungen der

wichtigsten Handlungsoptionen als Flussdiagramme. Fünf dieser Leitlinien lieferten eine Reihe weiterer Instrumente, wie z.B. Messbogen, Kurzfassungen im Taschenformat, diverse verschiedene Patientenmaterialien etc. [ACC/AHA 2001, ESC 2002/2001a, DVA & VHA 2002, ICSI 2003, NICE 2003]. Drei Leitlinien lagen in nur einer einzigen Version und ohne unterstützende Materialien vor [DGK 2001, LLGH 2003, UM 2001].

Zusammenfassung Domäne 4: Alle 16 Leitlinien erreichten standardisierte Domänenscores von 0,5 und mehr, insbesondere waren die Empfehlungen klar und eindeutig formuliert. Die größten Unterschiede fielen hinsichtlich Art und Anzahl mitgelieferter Tools auf.

Domäne 5: Anwendbarkeit

Item 19 – Diskussion potentieller Barrieren: 2/16 (13%) Leitlinien wiesen auf mögliche Barrieren hin [ACC/AHA 2001, LLGH 2003]. Die Bedeutung einer erforderlichen Barrierenanalyse wurde bei ACC/AHA [2001] im Rahmen der Implementierung hervorgehoben und mit Evidenz unterlegt. Spezifische Barrierenanalysen zu Herzinsuffizienz-Leitlinie wurden jedoch nicht berichtet. Bei LLGH [2003] fanden sich einzelne Hinweise auf potentielle Barrieren im Leitlinienreport. Die übrigen 14 Leitlinien diskutierten dieses Thema nicht.

Item 20 – Berücksichtigung finanzieller Auswirkungen: 11/16 (69%) Leitlinien enthielten keine Angaben zu finanziellen Auswirkungen, drei Leitlinien berichteten über Tagestherapiekosten verschiedener (z.T. alternativ einsetzbarer) Pharmakotherapien [DVA & VHA 2002, OPOT 2000, UM 2001], eine Leitlinie [LLGH 2003] unternahm eine orientierende Prognoseschätzung aus Sicht der Zielgruppe (Steigerung der Verordnungen). Eine Leitlinie enthielt Betrachtungen zur Kosteneffektivität der empfohlenen Maßnahmen aus volkswirtschaftlicher Perspektive auf der Basis gesundheitsökonomischer Studien [NICE 2003].

Item 21 – Messgrößen / Indikatoren: 10/16 (62%) Leitlinien nannten klinische Messgrößen, die mindestens eine orientierende Prüfung der Versorgung erlauben. Von diesen berichteten fünf Leitlinien über einen Satz von Qualitätsindikatoren, die eine Evaluation hinsichtlich der Leitlinienanwendung ermöglichen. Vier Leitlinien publizierten diese Qualitätsindikatoren [ACC/AHA 2001, DVA & VHA 2002, ICSI 2003, NICE 2003], bei einer weiteren Leitlinie wurde im Leitlinienreport über die Vorgehensweise zur Erstellung berichtet, ohne dass der Indikatorensatz veröffentlicht wurde [LLGH 2003].

Zusammenfassung Domäne 5: In dieser Domäne zur Anwendbarkeit erzielten 2/16 (13%) Leitlinien standardisierte Domänenscores von 0,5 und mehr. 5/16 (31%) Leitlinien wurden mit standardisierten Domänenscores von 0,0 bewertet, d.h. sie wurden in allen Items mit Likert „1“ geratet. Insbesondere die Barrieren-

diskussion fehlte, und finanzielle Auswirkungen der Leitlinienempfehlungen wurden nicht diskutiert.

Domäne 6: Redaktionelle Unabhängigkeit

Item 22 – Redaktionelle Unabhängigkeit der Leitlinie: In 10/16 (63%) Leitlinien erfolgte ein expliziter Hinweis auf den finanziellen Förderer, in vier dieser Leitlinien wurden darüber hinaus auch Details über Finanzierungsmodalitäten mitgeteilt. Bei zwei Leitlinien aus dieser Gruppe wurde explizit jede inhaltliche Einflussnahme durch den Förderer verneint. Fünf Leitlinien machten keine Angaben zur Förderung. Eine Leitlinie wurde durch eine Industriespende gesponsert [NHF/NZ 2001], wobei keine Erklärung über eine fehlende inhaltliche Einflussnahme abgegeben wurde.

Item 23 – Interessenkonflikte der Leitlinienentwicklergruppe: Bei der Hälfte der Leitlinien (8/16) lagen zumindest globale Erklärungen zum Fehlen potentieller Interessenkonflikte vor, fünf davon berichteten auch den Inhalt der Abfrage, von diesen wiederum legten drei Leitlinien auch die Ergebnisse der Abfrage offen. Bei der anderen Hälfte der Leitlinien fehlten Erklärungen über potentielle Interessenkonflikte der Autoren bzw. in einem Fall war ein Autor Mitarbeiter der pharmazeutischen Industrie [NHF/NZ 2001].

Zusammenfassung Domäne 6: 5/16 (31%) Leitlinien wurden mit einem standardisierten Domänenscore von 0,5 und darüber bewertet. Bei Dreiviertel der berücksichtigten Leitlinien lagen mindestens globale Angaben zur Finanzierung der Leitlinie oder zu potentiellen Interessenkonflikten der Entwickler vor – bei einem Viertel der Leitlinien fehlten jegliche Angaben dazu. Bei einer Leitlinie kann eine direkte Einflussnahme durch den Industriesponsor via Gesamtfinanzierung der Leitlinienentwicklung und durch die Mitarbeit eines Angehörigen der pharmazeutischen Industrie als Leitlinienautor nicht ausgeschlossen werden [NHF/NZ 2001].

2.3 Datenextraktion aus Leitlinien

Die folgende Tabelle (A.5) zeigt einen Ausschnitt aus der Evidenztafel zu Evidenzbewertungen und Empfehlungen über ACE-Hemmer, in der die Daten einer Leitlinie [DVA & VHA 2002] präsentiert wurden. Diese Leitlinie war von hoher Qualität und berichtete als einzige der berücksichtigten auch eine zusammenfassende Einschätzung der Qualität der von ihr bewerteten Evidenz.

Die vollständige Evidenztafel ist Bestandteil des Methodenreports:

http://www.allgemeinmedizin.uni-frankfurt.de/forschung/herzinsuffizienz_internet.html

Tab. A.5: Ausschnitt aus einer Evidenztabelle zu Leitlinienempfehlungen über ACE-Hemmer

Ergebnisse	Evidenzlevel	Grading	Empfehlungen	Evidenzbezüge / Kommentare
ACE-Hemmer reduzieren bei Pat. mit aktuell oder kürzlich durchgemachtem Myokardinfarkt und Nachweis einer LV-Dysfunktion die Mortalität und verlangsamen die Progression zur symptomatischen Herzinsuffizienz; postinfarzielle ACE-Hemmer-Therapie kann bei Pat. mit periinfarzieller Herzinsuffizienz das Risiko für Reinfarkte und Tod senken; Pat. mit Arteriosklerose oder Diab. mell. & weiteren koronaren Risikofaktoren ohne reduzierte Ejektionsfraktion hatten unter ACE-Hemmer reduzierte Ereignisse für Tod, Myokardinfarkt und Schlaganfall	A	A (Class I ACC /AHA)	ACE-Hemmer für Patienten mit Myokardinfarkt (akut, kürzlich oder in zurückliegender Anamnese), unabhängig von Ejektionsfraktion, asymptomatisch für Herzinsuffizienz Qualität der Studien insgesamt: gut Netto-Effekte: substantiell	Zitierte Studien und Reviews: ISIS-4 1995, GISSI-3 1996, SAVE 1992, AIRE 1993, AIREX 1997, TRACE 1995, HOPE 2000, SOLVD 1991, 1992, [ACTION-HF 1999, CR] Captopril Multicenter Research Group 1983, Sharpe et al. 1984, Chalmers et al. 1987, Pflugfelder et al. 1993, Gunderson et al. 1994, FEST 1995, Lechat et al. 1993,
Signifikante Reduktion des kombinierten Endpunktes Tod und Hospitalisierung unter ACE-Hemmern bei asymptomatischen Pat. mit LV-Dysfunktion; keine signifikante Reduktion der Sterblichkeit	B	A (Class I ACC /AHA)	ACE-Hemmer für asymptomatische Pat. mit reduzierter Ejektionsfraktion, unabhängig von Infarkt-Anamnese Qualität der Studien insgesamt: gut Netto-Effekte: substantiell	

Ergebnisse	Evidenzlevel	Grading	Empfehlungen	Evidenzbezüge / Kommentare
<p>Verbesserung von Symptomen u. funktionellem Status, Senkung von Mortalität u. Hospitalisierungsrate unter ACE-Hemmern; Senkung der Rate akuter Dekompensationen: (ARR und NNT SOLVD s. DieM 2003 / 2004) Pat. mit leichter – mäßiger Herzinsuffizienz, ACE-Hemmer vs. Hydralazin / ISDN, 2,5 Jahre, Risiko Tod: ARR=5,4%, NNT=18,5 (V-HeFT) Pat. mit NYHA IV, Gesamtmortalität nach 6 Mon.: ARR=17,7%, NNT=5,7 (CONSENSUS); Uneinheitliche Evidenz für dosisabhängige Effekte auf Mortalität (signifikante Reduktion vs. kein Effekt), Hinweise für Symptomverbesserung und reduzierte Hospitalisierungsrate bei Hochdosispat. im Vgl. zu Niedrigdosispatienten.</p>	A	A (Class I ACC /AHA)	<p>Maximal tolerierte ACE-Hemmer-Dosis für Pat. mit symptomatischer Herzinsuffizienz, um Symptome und Mortalität zu verbessern und Hospitalisierungsrate zu senken Qualität der Studien insgesamt: gut Netto-Effekte: substantiell</p>	<p>Captopril-Digoxin Multicenter Research Group 1988, Cohn et al. (V-HeFTⁿ) 1991, CONSENSUS 1987, Bungard et al. 2001, ATLAS 1999, NETWORK 1998, Garg & Yusuf 1995 [SR], Nanas et al. 2000, Punzi 1993, Sharif et al. 1994, Germino et al. 1993</p>
<p>(Schlussfolgerung durch Vergleich der Ergebnisse der Mortalität der Niedrigdosisgruppe eines RCT mit der Placebogruppe eines anderen RCT)</p>	k.A.	A (ACC /AHA k.A.)	<p>Niedrigere ACE-Hemmer-Dosen reduzieren ebenfalls die Mortalität, wenn Zieldosis nicht toleriert wird. Qualität der Studien insgesamt: gut Netto-Effekte: mäßig</p>	

Ergebnisse	Evidenz-level	Grading	Empfehlungen	Evidenzbezüge / Kommentare
Unzureichende Evidenz für Pharmakotherapie der diastolischen Herzinsuffizienz, da Patienten in RCTs unterrepräsentiert, darum Einbeziehung eines Kardiologen empfohlen.	C	C (Class IIb ACC/AHA)	ACE-Hemmer bei diastolischer Herzinsuffizienz empfohlen, sofern Symptome sich unter kontrolliertem Hypertonus nicht bessern. Qualität der Studien insgesamt: schlecht Netto-Effekte: gering	Zitierte Studien und Reviews: Goldsmith & Dick 1993 [Ü], Bonow & Udelson 1992 [Ü], Weinberger 1999 [Ü], Setaro et al. 1990 (N=20), Dahloff et al. 1992 [MA zu LVH], Gottdiener et al. 1997 [RCT, LVH]
Kontraindikationen für ACE-Hemmer: Angioödem ⁶⁴ oder andere Hypersensitivitätsreaktion auf ACE-Hemmer in Anamnese, bilaterale Nierenarterienstenose, Schwangerschaft, Serum-K ⁺ >5,5 mEq/L, symptomatische Hypotension; ACE-Hemmer-Husten: alternative Ursachen abklären, versuchsweise auf Fosinopril wechseln, wenn unter anderen ACE-Hemmern aufgetreten, Ungefährlichkeit mit Pat. besprechen (Risiko-Nutzen-Abwägung); Wechselwirkungen zwischen ACE-Hemmer und Aspirin: Evidenz aus retrospektiven Analysen unzureichend, prospektive Studien abwarten	Ø	Ø	Aspirin-Gabe bei gleichzeitiger KHK weiterhin empfohlen (trotz möglicher reduzierter ACE-Hemmer-Benefits), da keine ausreichende Evidenz vorhanden	

⁶⁴ Die Inzidenz des Angioödems unter ACE-Hemmern beträgt etwa 0,1 bis 1,2% (Pylypchuk GB. ACE inhibitor versus angiotensin II blocker-induced cough and angioedema. Ann Pharmacother 1998; 32: 1060-1066. in DVA & VHA 2002). Das relative Risiko für farbige Patienten, ein Angioödem unter ACE-Hemmern zu entwickeln, beträgt gegenüber weißen Patienten 4,5 (DVA & VHA 2002). Angioödeme jedoch auch unter AT₁-Blocker-Therapie – siehe dort.

In Tab. A.5 verwendete Abkürzungen: CR – Consensus Recommendation, SR – Systematic Review, MA – Meta-analysis, RCT – Randomized Controlled Trial, Ü – traditionelle (unsystematische) Übersicht, LVH – linksventrikuläre Hypertrophie

2.4 Weitere Ergebnisse der Konsistenzanalyse

Ergänzend zum Kapitel 4.6 des Hauptdokuments werden hier weitere Ergebnisse aus den Konsistenzanalysen gezeigt.

Tab. A.6: Nichtpharmakologische Therapie

	ACC/AHA 2001	AKDAE 2001	CCS 2001, 2002/3	DGK 2001	DieM 2003/2004	Duodecim 2004	DVA & VHA 2002	ESC 2002/2001	ICSI 2003	LLGH 2003	NHF/Austr. & CSANZ 2002	NHF/NZ 2001	NICE 2003	OPOT 2000	UM 2001	UWH 2001
Lebensstilmodifikation (Rauchen, Alkohol)																
Evidenzlevel	C	#	2	B	Ø	#	Ø	#	Ø	III	EO, IV	C	†	Ø	Ø	
Grading	I		B				Ø		Ø	B			†			
Körperliches Training																
Evidenzlevel	A	#	2	A	I	C, B [†]	Ø	B	††	Ø	II	A	Ia	Ø	D	A
Grading	IIa		B				#			A			b			
Mindestens leichte Salz- und Flüssigkeitsrestriktion																
Evidenzlevel	C	#	2	B	IV	#	#	#	R	IV	IV	C	#	Ø	D	C
Grading	III		B				#		R	C			#			
Gripeschutzimpfung																
Evidenzlevel	#	Ø	Ø	Ø	Ø	Ø	#	#	Ø	IIb	EO	C	Ø	Ø	Ø	B
Grading	#		C				#		Ø	B			GPP			
Pneumokokkenimpfung																
Evidenzlevel	#	Ø	Ø	Ø	Ø	Ø	#	#	Ø	IIb	EO	C	IV	Ø	Ø	B
Grading	#		C				#		Ø	B			GPP			
Ø – Keine Angabe, da außerhalb der LL; # - Empfehlung im Text ohne Evidenzlevel & Grading; †differenzierte Evidenzlevel & Grading: Rauchen: IV, GPP; Alkohol bei alkoholtox. Kardiomyopathie: III, C; Alkohol ohne alkoholtox. Kardiomyopathie: III, GPP; Rauchen: B, Alkohol: C; †abweichendes Grading für HI-Pat. mit KHK; ††Evidenzlevel nicht eindeutig festgelegt																

Tab. A.7: Digitalistherapie

	ACC/AHA 2001	AKDAE 2001	CCS 2001, 2002/3	DGK 2001	DieM 2003/2004	Duodecim 2004	DVA & VHA 2002	ESC 2002/2001	ICSI 2003	LLGH 2003	NHF/Austr. & CSANZ 2002	NHF/NZ 2001	NICE 2003	OPOT 2000	UM 2001	UWH 2001
Digitalis bei HI-Pat. mit tachykardem Vorhofflimmern zur Kontrolle der Ruhefrequenz, additiv zur Standardtherapie																
Evidenzlevel	A	↑↑	1	B	♀	A**	A	B	A	Ib	II	B	Ib	A	A	A
Grading	I*		A				A		I	A			A			
Digitalis bei HI-Pat. NYHA III-IV und Sinusrhythmus zur Symptomverbesserung, additiv zur Standardtherapie																
Evidenzlevel	A	↑↑	1	A	♀	A**	A	∅	A	Ib♂	II	A	Ib	A	A	A
Grading	I		A				A†		I	A♂			A			
Digitalis bei asymptomatischen HI-Pat.																
Evidenzlevel	∅	↓↓	∅	D	∅	∅	∅	B††	∅	∅	∅	B††	IV††	∅	A††	A††
Grading	∅		∅				∅		∅	∅			C††			
Nur Digoxinstudien, andere Digitalispräparate nicht untersucht; *β-Blocker effektiver zur Frequenzkontrolle des chronischen VHF; ** nur bei erheblicher LV-Dilatation und deutlich reduzierter EF empfohlen; ♀ - Sterberate für Frauen in einer post-hoc-Analyse erhöht!, Hospitalisierungsrate leicht vermindert, kein Effekt auf Lebensqualität nachgewiesen; †Digoxin bei HI ohne Standardtherapie potentiell gefährlich und nicht empfohlen; ††Nur zur Frequenzkontrolle bei VHF; ♂gilt nur für Männer																

Tab. A.8: Pharmakotherapie mit Angiotensin-II-Antagonisten (AT₁-Blocker)

	ACC/AHA 2001	AKDAE 2001	CCS 2001, 2002/3	DGK 2001	DieM 2003/2004	Duodecim 2004	DVA & VHA 2002	ESC 2002/2001	ICSI 2003	LLGH 2003	NHF/Austr. & CSANZ 2002	NHF/NZ 2001	NICE 2003	OPOT 2000	UM 2001	UWH 2001
AT ₁ -Blocker als Standardtherapie bei ACE-Hemmer-Intoleranz																
Evidenzlevel	A	↑↑	2	B	!	B	A	C	A	Ib	II	C	Ia	B	A	A
Grading	Ila	/↑*	B				B		∅	A			A			
AT ₁ -Blocker zusätzlich zu ACE-Hemmern und β-Blockern (Tripel-Therapie)																
Evidenzlevel	B	∅	∅	∅	!	∅	∅	B	A	∅	∅	∅	∅	∅	∅!	∅!
Grading	Ilb		∅				∅		∅!	∅			∅!			

(Fortsetzung Tab. A.8)

	ACC/AHA 2001	AKDAE 2001	CCS 2001, 2002/3	DGK 2001	DieM 2003/2004	Duodecim 2004	DVA & VHA 2002	ESC 2002/2001	ICSI 2003	LLGH 2003	NHF/Austr. & CSANZ 2002	NHF/NZ 2001	NICE 2003	OPOT 2000	UM 2001	UWH 2001
AT₁-Blocker zusätzlich zu ACE-Hemmern bei HI-Pat., die keine β-Blocker erhalten																
Evidenzlevel	A	Ø	1**	Ø	!	Ø	B**	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø
Grading	III		A**				C**		Ø	Ø			Ø			
AT₁-Blocker anstelle ACE-Hemmern bei ACE-Hemmer-naiven Patienten																
Evidenzlevel	B	Ø	Ø	Ø	! !	Ø	B	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø!	Ø!
Grading	III		Ø				D (=!)		Ø	Ø			Ø!			
*Unterschiedliche Gradings für unterschiedliche Outcomes verwendet (Senkung der Sterblichkeit vs. Symptomverbesserung); **Empfehlung nur im Falle einer β -Blocker-Intoleranz; Ø – Keine Angabe von Evidenzlevel / Grading; ! – Potentiell schädliche Therapie, Negativempfehlung																

Tab. A.9: Pharmakotherapie mit Hydralazin / ISDN

	ACC/AHA 2001	AKDAE 2001	CCS 2001, 2002/3	DGK 2001	DieM 2003/2004	Duodecim 2004	DVA & VHA 2002	ESC 2002/2001	ICSI 2003	LLGH 2003	NHF/Austr. & CSANZ 2002	NHF/NZ 2001	NICE 2003	OPOT 2000	UM 2001	UWH 2001
Mittel 2. Wahl bei ACE-Hemmer- und AT₁-Blocker-Intoleranz																
Evidenzlevel	B	↓↓*	Ø	C*	#	Ø [†]	B	A	Ø	Ø	II	Ø	Ia	A	A	A
Grading	Ila		Ø				B ^{††}		Ø	Ø			A			
Additive Gabe, zusätzlich zur Standardtherapie																
Evidenzlevel	B	↔	Ø	Ø	#	Ø [†]	B	Ø	Ø	Ø	Ø	Ø	Ib [‡]	C	Ø	Ø
Grading	Ilb		Ø				C ^{††}						A [‡]			
*Keine gleichwertige Therapie zu ACE-Hemmern und AT ₁ -Blockern, nur als Ersatz zu erwägen; #kein Nutznachweis zu Mortalität und Morbidität erbracht; [†] Kombination mit anderen Therapeutika möglich, keine Angabe von Evidenzlevel/Grading; ^{††} bei persistierender Symptomatik unter Standardtherapie; [‡] Evidenzlage zu einem additiven Effekt unklar																

Tab. A.10: Therapie mit Kalziumantagonisten

	ACC/AHA 2001	AKDAE 2001	CCS 2001, 2002/3	DGK 2001	DileM 2003/2004	Duodecim 2004	DVA & VHA 2002	ESC 2002/2001	ICSI 2003	LLGH 2003	NHF/Austr. & CSANZ 2002	NHF/NZ 2001	NICE 2003	OPOT 2000	UM 2001	UWH 2001
Langwirksame Dihydroperidine (z.B. Amlodipin) unbedenklich																
Evidenzlevel	Ø	Ø	Ø	B	Ø	Ø [†]	Ø	A	A	Ø	II	Ø [†]	Ia	Ø	A	C ^{††}
Grading	Ø		Ø				Ø [†]		Ø [†]	Ø			A			
Kontraindikation für Nifedipin-Typ, Verapamil-Typ und Diltiazem																
Evidenzlevel	B	↓↓	Ø	D ^{**}	! !	Ø	Ø	C	A	Ø	Ø !	Ø	Ia	C	A	Ø
Grading	I [*]		Ø				Ø!		Ø	Ø			A			
*Auch kurzfristige Gaben sollen vermieden werden; **kurzfristige Gaben von Verapamil bei tachykardem VHF erlaubt; ! Potentiell schädliche Substanzgruppe; [†] Amlodipin erlaubt; ^{††} Amlodipin bei schwerer therapierefraktärer nicht-ischämischer CM																

Tab. A.11: Orale Antikoagulation und Thrombozytenaggregationshemmung

	ACC/AHA 2001	AKDAE 2001	CCS 2001, 2002/3	DGK 2001	DileM 2003/2004	Duodecim 2004	DVA & VHA 2002	ESC 2002/2001	ICSI 2003	LLGH 2003	NHF/Austr. & CSANZ 2002	NHF/NZ 2001	NICE 2003	OPOT 2000	UM 2001	UWH 2001
Orale Antikoagulation bei Pat. mit paroxysmalem oder permanentem VHF oder stattgehabtem thrombembolischen Ereignis																
Evidenzlevel	A	↑↑	2	A	Ø	D [*]	Ø	C	B/C	Ø	#	A	Ia	Ø	A	A
Grading	I		B				#		Ø	C			A			
Orale Antikoagulation bei Pat. ohne VHF oder stattgehabtem thrombembolischen Ereignis bei deutlich dilatiertem Ventrikel/deutlich reduzierter EF oder intrakardialen Thromben empfehlenswert																
Evidenzlevel	B/C	#	Ø	B	Ø	Ø	Ø	Ø	B/C	Ø	Ø	C ^{††}	IV	Ø	C	Ø
Grading	IIb		Ø			Ø	Ø!		Ø	C [†]			GPP			
Niedrig dosierte Thrombozytenaggregationshemmung bei gleichzeitiger KHK																
Evidenzlevel	B	Ø [‡]	Ø	B [‡]	Ø	D	Ø	Ø [‡]	Ø	Ø	Ø [‡]	A [‡]	Ia ^{††}	Ø	A	Ø [‡]
Grading	Ila [‡]		Ø				Ø [‡]		Ø	Ø [‡]			B			
#positive Empfehlung im Text; ! Negativempfehlung im Text; * nur bei gleichzeitiger Kardiomegalie empfohlen; **individuelle Entscheidung; [†] nur bei großen Ventrikelaneurysmen, flottierenden Wandthromben; ^{††} nur bei intraventrikulärem Thrombus; [‡] Hinweis auf potentielle Interaktion mit ACE-Hemmern; ^{††} Evidenzlevel für Interaktionen IIb																

Tab. A.12: Pharmakotherapieempfehlungen zur diastolischen Herzinsuffizienz

	ACC/AHA 2001	DVA & VHA 2002	ICI 2004/3	NICE 2003	Gesamtbe- wertung der Studienqua- lität [nach DVA & VHA 2002]	Netto- Effekte der Studien [nach DVA & VHA 2002]
Kontrolle des arteriellen Hypertonus	A/I	A/A	∅	∅	Gut	Substantiell
Diuretika zur Kontrolle des Volumenstatus, Start mit niedrigen Dosen (z.B. 20 - 40 mg Furosemid/d)	C/I	C/B	C	GPP	Schlecht	Mäßig
Frequenzkontrolle bei VHF	C/I	C/B	C*	∅	Schlecht	Mäßig
Digoxin bei Sinusrhythmus zur Verbesserung von Symptomen und Hospitalisierungsrate	C/IIb	C/C	∅	∅	Schlecht	Mäßig
β-Blocker, ACE-Hemmer, Kalziumantagonisten, AT ₁ -Blocker bei Patienten mit kontrolliertem Blutdruck, die weiterhin symptomatisch bleiben	C/IIb	C/C	C	∅	Schlecht	Gering
Nitrate bei Patienten mit diastolischer Herzinsuffizienz bei KHK	∅	C/C	∅	∅	Schlecht	Gering
*Therapie mit β-Blockern empfohlen						

Tab. A.13: Konsistente und inkonsistente Empfehlungen der berücksichtigten Leitlinien, geordnet nach Themen

Fragestellung / klinisches Problem	Anzahl konkordanter LL / Gesamtzahl LL*	Evidenzbasie- rung, wie in LL angegeben	Konsistenz- / Inkonsistenztyp**
Erstdiagnostik			
Symptome und klinische Zeichen	5/5 Positive Empfehlungen	Expertenkonsens / quasi-experimentelle Studien	3
Klinisch-chemische Erstdiagnostik (Basis-Labor)	6/6 Positive Empfehlungen	Expertenkonsens	3
Ruhe-EKG, 12 Ableitungen	7/7 Positive Empfehlungen	Expertenkonsens / quasi-experimentelle Studien	3
Röntgen-Thorax	5/5 Positive Empfehlungen	Expertenkonsens	3
Echokardiographie	7/7 Positive Empfehlungen	Expertenkonsens / quasi-experimentelle Studien	3

(Fortsetzung Tabelle A.13)

Fragestellung / klinisches Problem	Anzahl konkordanter LL / Gesamtzahl LL*	Evidenzbasie- rung, wie in LL angegeben	Konsistenz- / Inkon- sistenztyp**
Erstdiagnostik			
BNP/NT-proBNP-Test als diagnostischer Test	Komplette Inkon- sistenz	Expertenkonsens bis hochwertige Evidenz	A
Nicht-pharmakologische Therapie			
Lebensstilmodifikation (Rauchen, Alkohol)	8/8 Positive Empfeh- lungen	Expertenkonsens	3
Körperliches Training	13/13 Positive Emp- fehlungen	Expertenkonsens bis hochwertige Evidenz	4
Salz- und Flüssigkeitsre- striktion	9/10 Positive Empfeh- lungen (z.T. mit leich- ten Einschränkungen), 1/10 negative Empfeh- lung	Expertenkonsens / quasi- experimentelle Studien	B
Impfempfehlungen (In- fluenza, Pneumococcus)	6/6 Positive Empfeh- lungen	Expertenkonsens / quasi- experimentelle Studien	3
Pharmakotherapie der systolischen Herzinsuffizienz			
ACE-Hemmer in Zieldosis bei allen NYHA-Klassen	16/16 Positive Emp- fehlungen	Hochwertige Evi- denz, zahlreiche RCTs und SRs	1
β-Blocker in Zieldosis bei allen HI-Pat. mit NYHA I, mit oder ohne stattgehab- ten Myokardinfarkt	7/8 Positive Empfeh- lungen, 1/8 Negative Empfeh- lung	Hochwertige Evi- denz, wenige Studien	B
β-Blocker in Zieldosis bei allen HI-Pat. mit NYHA I und nach stattgehabtem Myokardinfarkt	11/11 Positive Emp- fehlungen	Hochwertige Evi- denz, zahlreiche RCTs und SRs [†]	1
β-Blocker bei HI-Patienten mit NYHA II-III	16/16 Positive Emp- fehlungen	Hochwertige Evi- denz, zahlreiche RCTs und SRs	1
β-Blocker bei klinisch stabilen HI-Patienten mit NYHA IV	13/15 Positive Emp- fehlungen, 2/15 Negative Emp- fehlungen	Hochwertige Evi- denz, zahlreiche RCTs	B
Niedrig dosierte Gaben von Aldosteron-Antagonisten bei NYHA III-IV	16/16 Positive Emp- fehlungen	Hochwertige Evi- denz, wenige Studien	2

(Fortsetzung Tabelle A.13)

Fragestellung / klinisches Problem	Anzahl konkordanter LL / Gesamtzahl LL*	Evidenzbasie- rung, wie in LL angegeben	Konsistenz- / Inkon- sistenztyp**
Pharmakotherapie der systolischen Herzinsuffizienz			
AT ₁ -Blocker als Standard- therapie bei ACE-Hemmer- Intoleranz	15/16 Positive Emp- fehlungen, 1/16 negative Empfeh- lung (Warnung) [‡]	Positive Empfeh- lung: hochwertige Evidenz, zahlrei- che RCTs / SRs Cave: geringer Evidenzgehalt [‡]	B
AT ₁ -Blocker zusätzlich zu ACE-Hemmern und β- Blockern (Tripel-Therapie)	2/6 Positive Empfeh- lungen, 1/6 negative Empfeh- lung (Warnung) [‡] , 3 LL rieten im Text von Tripel-Therapie ab	Wenige Studien, quasi- experimentelle Studien [‡]	A
AT ₁ -Blocker zusätzlich zu ACE-Hemmern bei HI-Pat., die wegen Kontraindikati- onen oder Intoleranz keine β-Blocker erhalten können	3/4 Positive Empfeh- lungen, 1/4 negative Empfeh- lung (Warnung) [‡] , 1 LL riet im Text ab	Unterschiedliche Evidenzgrade be- richtet	A
AT ₁ -Blocker anstelle ACE- Hemmern bei ACE- Hemmer-naiven Patienten	3/3 Negative Empfeh- lungen, 3 weitere LL rieten im Text davon ab	Zu ACE- Hemmern hoch- wertige Evidenz, Fehlen gleichwer- tiger Evidenz zu AT ₁ -Blockern	3
Diuretika bei allen HI- Patienten der Klassen NY- HA II-IV zur Kontrolle der Flüssigkeitsbalance	14/14 Positive Emp- fehlungen	Fehlen von Mor- talitätsstudien	4
Digitalis bei HI-Pat. mit tachy- kardem Vorhofflimmern zur Kontrolle der Ruhefre- quenz, additiv zur Stan- dardtherapie	15/15 Positive Emp- fehlungen	Hochwertige Evi- denz, wenige Studien	2
Digitalis bei HI-Pat. NYHA III-IV und Sinusrhythmus zur Symptomverbesse- rung, additiv zur Standard- therapie	14/14 Positive Emp- fehlungen (z.T. mit dif- ferenzierten Ein- schränkungen) 1 LL mit Warnhinweis [†]	Hochwertige Evi- denz, wenige Studien	B
Hydralazin / ISDN als Mittel 2. Wahl bei ACE-Hemmer- und AT ₁ -Blocker-Intoleranz	10/10 Positive Emp- fehlungen	Unterschiedliche Evidenzgrade be- richtet	4
Langwirksame Dihydropy- ridine (z.B. Amlodipin) un- bedenklich	7/7 Positive Empfeh- lungen	Unterschiedliche Evidenzgrade be- richtet	4

(Fortsetzung Tabelle A.13)

Fragestellung / klinisches Problem	Anzahl konkordanter LL / Gesamtzahl LL*	Evidenzbasie- rung, wie in LL angegeben	Konsistenz- / Inkonsistenztyp**
Pharmakotherapie der systolischen Herzinsuffizienz			
Bei systolischer Herzinsuffizienz absolute Kontraindikation für Kalziumantagonisten vom Nifedipin- und Verapamil-Typ sowie Diltiazem (auch kurzfristige Gaben sollen vermieden werden)	7/9 Positive Empfehlungen als absolute Kontraindikation (1 weitere LL mit Warnhinweisen im Text); 2/9 LL mit Ausnahmen bei kurzfristiger Gabe	Unterschiedliche Evidenzgrade berichtet	B
Orale Antikoagulation bei Pat. mit paroxysmalem od. permanentem VHF od. stattgehabtem thromboembolischen Ereignis	12/12 Positive Empfehlungen	Hochwertige Evidenz, wenige Studien	2
Orale Antikoagulation bei Pat. ohne VHF oder stattgehabtem thromboembolischen Ereignis bei deutlich dilatiertem Ventrikel / deutlich reduzierter EF od. intrakardialen Thromben	7/7 Positive Empfehlungen	Fehlende Studien, Expertenkonsens	3
Niedrig dosierte Thrombozytenaggregationshemmung bei gleichzeitig bestehender KHK	6/6 Positive Empfehlungen ^{††}	Schwache Evidenz / Expertenkonsens	3
Therapie der diastolischen Herzinsuffizienz (Angaben aus vier Leitlinien)			
Kontrolle des arteriellen Hypertonus	2/2 Positive Empfehlungen	Hochwertige Evidenz, wenige Studien	2
Vorsichtig Diuretika zur Kontrolle des Volumenstatus, Start mit niedrigen Dosen (z.B. 20-40 mg Furosemid / d)	4/4 Positive Empfehlungen	Wenige Studien, quasi-experimentelle Studien, Expertenkonsens	3
Frequenzkontrolle bei VHF	3/3 Positive Empfehlungen	Wenige Studien, quasi-experimentelle Studien, Expertenkonsens	3
Digoxin bei Sinusrhythmus zur Verbesserung von Symptomen und Hospitalisierungsrate	2/2 Positive Empfehlungen	Wenige Studien, quasi-experimentelle Studien, Expertenkonsens	3

(Fortsetzung Tabelle A.13)

Fragestellung / klinisches Problem	Anzahl konkordanter LL / Gesamtzahl LL*	Evidenzbasie- rung, wie in LL angegeben	Konsistenz- / Inkon- sistenztyp**
Therapie der diastolischen Herzinsuffizienz (Angaben aus vier Leitlinien)			
Vorsichtig β -Blocker, ACE- Hemmer, Kalziumantago- nisten, AT ₁ -Blocker bei Pa- tienten mit kontrolliertem Blutdruck, die weiterhin symptomatisch bleiben	3/3 Positive Empfeh- lungen	Wenige Studien, quasi- experimentelle Studien, Exper- tenkonsens	3
Nitrate bei Patienten mit diastolischer Herzinsuffi- zienz bei KHK	1/1 Positive Empfeh- lungen	Expertenkonsens	-

2.5 Ergebnis der Informationssynthese

In Ergänzung zum Kapitel 4.7 des Hauptdokuments wird hier ein Musterkapitel des vorläufigen Erstentwurfs zur DEGAM-Leitlinie Herzinsuffizienz (Bearbeitungsstand 30.06.2004) gezeigt.

2.2.3.1 Therapie der systolischen Herzinsuffizienz

2.2.3.1.1 Mortalitätsreduzierende Therapie

1. ACE-Hemmer

Welche Patienten sollten einen ACE-Hemmer erhalten und welche Dosierung ist dabei sinnvoll?

NYHA I-IV: Alle Patienten mit einer nachgewiesenen systolischen Dysfunktion und fehlenden Kontraindikationen sollten ACE-Hemmer (ACEH) erhalten, unabhängig davon, ob sie Symptome einer Herzinsuffizienz aufweisen, oder nicht [alle eingeschlossenen Leitlinien, s. Tab. 14] $\{I_{X1}, A_{X1}\}$.

Ziel der ACEH-Therapie ist die Prognoseverbesserung mit Senkung der Sterblichkeit und Verzögerung des Krankheitsprogresses sowie die Verbesserung der Symptomatik und der Hospitalisierungsrate bei leichter, mäßiger und schwerer Herzinsuffizienz (NYHA II-IV) [alle eingeschlossenen Leitlinien, s. Tab. 14] $\{I_{X1}, A_{X1}\}$.

Bei asymptomatischen Patienten mit eingeschränkter LV-Funktion (NYHA I) sollen durch die Gabe von ACEH die Inzidenz einer symptomatischen Herzinsuffizienz und die Hospitalisierungsrate gesenkt und die langfristige Prognose verbessert werden. Bei postinfarzieller LV-Dysfunktion soll zusätzlich die Reinfarktrate gesenkt werden. [alle eingeschlossenen Leitlinien, s. Tab. 14] $\{I_{X1}, A_{X1}\}$.

Insgesamt können ACEH die Mortalität und Morbidität der systolischen Herzinsuffizienz umso stärker verbessern, je schwerer die Herzinsuffizienz ist. Die Effekte sind jedoch in allen NYHA-Klassen nachweisbar [DieM 2003/2004, Duodecim 2004, DVA & VHA 2002, ICSI 2003/2004, NHF/Austr. & CSANZ 2002, NICE 2003, OPOT 2000, UM 2001, UWH 2001] $\{I_{X1}, A_{X1}\}$.

ACEH sind **Mittel der ersten Wahl** vor AT₁-Blockern, da eine Überlegenheit der Wirksamkeit dieser Substanzgruppe bislang nicht gezeigt werden konnte. Ein Wechsel auf AT₁-Blocker sollte nur bei ACEH-Intoleranz erwogen werden (näheres siehe dort). [Alle eingeschlossenen Leitlinien, s. Tabelle 14] $\{I_{X1}, A_{X1}\}$.

Dosierung: Nach Möglichkeit sollten HI-Patienten auf die jeweils höchste, in Mortalitätsstudien ermittelte Zieldosis bzw. die maximal tolerierte Dosis auftitriert werden (s. Tabelle 12). Für höhere Dosierungen konnten verminderte Hospitalisierungsraten gegenüber niedrigen Dosierungen gezeigt werden, der Effekt auf die Mortalität war nicht konsistent dosisabhängig [ACC/AHA 2001, CCS 2002/2003, DGK 2001, Duodecim 2004, DVA & VHA 2002, ESC 2002/2001, ICSI 2004/2003, NHF/Austr. & CSANZ 2002, NICE 2003, OPOT 2000, UM 2001] $\{I_{B_{X1}}, B_{X1}\}$.

Niedrigere Dosierungen reduzieren ebenfalls die Mortalität, wenn die optimale Zieldosis nicht erreicht wird [DVA & VHA 2002, NICE 2003] $\{I_{B_{X1}}, C_{X1}\}$.

ACEH-Einsatz im Gesamttherapiekonzept: $\{I_{V_{X1}}, C_{X1}\}$.

- ACEH sollten vor Initiierung einer β -Rezeptorenblocker-Therapie bis zur Zieldosis (s.o.) auftitriert werden [NICE 2003].
- Bei Normovolämie sollten sie zuerst gegeben werden, bei Hypervolämie zusammen mit Diuretika [ESC 2002/2001].
- Salzrestriktion und Diuretika erhöhen den Effekt von ACEH [Duodecim 2004].

ASS vermindert den Effekt von ACEH, es sollte bei gleichzeitig bestehender HKH jedoch weiter gegeben werden [DVA & VHA 2002].

Absolute Kontraindikationen:

Angiooedem oder andere Hypersensitivitätsreaktionen auf ACEH in der Anamnese⁶⁵; beidseitige Nierenarterienstenose; Serum-Kalium > 5,5 mmol/l; chronische Niereninsuffizienz mit GFR < 30ml/Min.; HOCM; Herzklappenstenosen⁶⁶; Schwangerschaft; symptomatische Hypotension [AKDAE 2001, DVA & VHA 2002, UWH 2001]

Empfehlungen zum praktischen Gebrauch der ACEH [NICE 2003]

1. ‚Start low – go slow‘: mit niedrigen Startdosen beginnen, Verdoppelung alle vier Wochen (nicht schneller als zweiwöchentlich); Beispiel für Titrationsschema s. Tabelle 13, Start- und Zieldosierungen verbreiteter ACEH s. Tabelle 12;
2. Zieldosis anstreben – wenn nicht möglich, höchste tolerierte Dosis (wenig ACEH ist besser als kein ACEH);
3. Ggf. Rat eines Spezialisten einholen
 - bei Patienten mit hoher Diuretikadosierung bei ACEH-Beginn (z.B. 80 mg Furosemid /d oder mehr);
 - bei Einstellungsproblemen durch Hypotonie, Hyperkaliämie oder erhöhte Nierenretentionswerte (s.u.);
4. Kontrolle von Kalium, Harnstoff, Kreatinin und Blutdruck vor jeder Dosissteigerung;
5. Information des Patienten über
 - Ziele und Benefits der Therapie
 - Dauer bis zum spürbaren Wirkeintritt (Symptombesserung) einige Wochen bis Monate!
 - Erforderliche Rückmeldung bei Eintritt unerwünschter Wirkungen (Husten, Schwindel/symptomatische Hypotension)

Zu Beginn einer ACEH-Therapie treten häufig folgende Probleme auf:

Vorsicht bei Hypotonie: ACEH führen insbesondere bei dehydrierten Patienten zu Blutdruckabfall → Überwachung für 2-6 Std. (je nach Halbwertszeit!) nach Erstgabe; leichte Hypervolämie bei Therapiestart eher von Vorteil; kein Start bei systolischem Blutdruck < 90 mmHg!

⁶⁵ Die Inzidenz des Angiooedems unter ACEH beträgt etwa 0,1 bis 1,2% [Pylypchuk 1998 in DVA & VHA 2002]. Das relative Risiko für farbige Patienten, ein Angiooedem unter ACEH zu entwickeln, beträgt gegenüber weißen Patienten 4,5 [DVA & VHA 2002]. Angiooedeme treten jedoch auch unter AT₁B-Therapie auf – siehe dort.

⁶⁶ Uneinheitlicher Umgang in Leitlinien mit z.B. Aortenstenosen: Patienten mit Aortenstenosen tolerieren nach Duodecim [2004] den vorsichtig einschleichenden Einsatz von ACEH. Rücksprache zur Therapieentscheidung mit Kardiologen empfohlen.

Häufig folgt unter ACEH-Therapie ein Kreatininanstieg um etwa 10% – 15% des Ausgangswertes [AKDAE 2001 / DGK 2001]. Bei erhöhten Ausgangswerten oder außergewöhnlich starken Anstiegen besondere Vorsicht:

Keine weitere Dosissteigerung bei Serum-Kalium >5,0 mmol/l und/oder Serum-Kreatinin >200 µmol/l oder bei Anstieg des Serum-Kreatinin um mehr als 20% des Ausgangswertes. Kontrolle und ggf. Kooperation mit Spezialisten oder stationäre Einstellung erwägen [NICE 2003, ICSI 2003, Duodecim 2004, ESC 2002/2001, UWH 2001] {IV_{x1}, C_{x1}}.

Viele Patienten berichten über Husten:

Husten bei Herzinsuffizienzpatienten häufig als ACEH-Nebenwirkung missdeutet – andere Ursachen abklären: pulmonaler Rückstau? COPD?
ACEH-Husten: nur bei etwa 5-10% der Patienten⁶⁷, trocken, verschwindet bei Absetzen und setzt wenige Tage nach erneutem Therapiestart wieder ein,
ABER: ist ungefährlich, bei Aufklärung des Patienten meist kein Absetzen der ACEH erforderlich, nur bei quälendem nächtlichem Husten Wechsel auf AT₁-Blocker erwägen. [ACC/AHA 2001, DVA & VHA 2002, NICE 2003] {IV_{x1}, C_{x1}}.

Tab. 12: Start- und Zieldosierungen verbreiteter ACEH

ACE-Hemmer	Startdosis	Zieldosis	Tagestherapiekosten
Captopril*	3x 6,25 mg/d	3x 50-100 mg/d	(Wird in späteren Entwicklungsstufen ergänzt)
Cilazapril [§]	1x 0,5 mg/d	1x 2,5 mg/d	
Enalapril*	2x 2,5 mg/d	2x 10-20 mg/d	
Fosinopril [§]	1x 10 mg/d	1x 40 mg/d	
Lisinopril*	1x 2,5-5 mg/d	1x 30-35 mg/d	
Perindopril [§]	1x 2 mg/d	1x 4 mg/d	
Quinapril [§]	1x 2,5-5 mg/d	1x 10-20 mg/d	
Ramipril*	1x 2,5 mg/d	2x5 od. 1x10 mg/d	

Quelle: NICE 2003; * Dosierungen entstammen Mortalitätsstudien, § Dosierungen entsprechen Herstellerangaben

Tab. 13: Beispiele für Titrationsschemata von ACEH (Dosierungen in mg)⁶⁸

ACEH	Start	1.-4. Woche	5.-8. Woche	9.-12. Woche	13.-16. Woche	Ziel
Captopril	3x 6,25	3x 12,5	3x 25	3x 50	3x 100	3x 50-100
Enalapril	2x 2,5	2x 5	2x 10	2x 20	2x 20	2x 10-20
Lisinopril	1x 2,5	1x 5	1x 10	1x 20	1x30	1x 30-35

⁶⁷ Bei Patienten asiatischer Herkunft etwas häufiger; Inzidenzangaben nach ACC/AHA [2001].

⁶⁸ Quelle: Leitlinienbasierte Therapieempfehlungen der Charité: <http://www.fvk-berlin.de/fvkweb/fvkindex.html>

Evidenz und Rationale

Tab. 14: Zusammenfassung der Evidenzlevel und Stärken der Empfehlung zur ACEH-Therapie in den eingeschlossenen LL (siehe Tab. 4.13, Hauptdokument)

Die Empfehlungen der 16 eingeschlossenen Leitlinien basierten auf insgesamt sechs systematischen Reviews mit darin enthaltenen Metaanalysen und 45 Primärpublikationen, davon mehr als 30 RCTs.

Gesamtbewertung der Studienqualität: gut; Nettoeffekte: substantiell [DVA & VHA 2002, Gesamtbewertung für N=27 Primärpublikationen und N=1 systematischen Review]

Ergebnisse: ACEH verbessern die Prognose bei leichter, mäßiger und schwerer Herzinsuffizienz, ebenso bei postinfarzieller und asymptomatischer Herzinsuffizienz [Duodecim 2004]. Ihre Effekte sind umso größer, je höher das Basisrisiko ist [DieM 2003]. Sie senken die Gesamtsterblichkeit, die Zahl kardial bedingter Todesfälle und die Hospitalisierungsrate in allen NYHA-Klassen [ICSI 2003/2004]. Bei NYHA II-IV werden zusätzlich der funktionelle Status und die Symptomatik gebessert und die Rate akuter Dekompensationen vermindert. Patienten mit NYHA-I und durchgemachtem Herzinfarkt profitieren ebenfalls mit einer verminderten Mortalität, einer gesenkten Reinfarktrate und einem verlangsamten Progress zu symptomatischen Stadien der Erkrankung [DVA & VHA 2002]. Patienten mit NYHA I ohne durchgemachten Herzinfarkt zeigten in klinischen Studien unter ACEH einen verminderten kombinierten Endpunkt aus Tod und Hospitalisierung, der Unterschied in der Mortalität allein war nicht signifikant [DVA & VHA 2002].

Es wird angenommen, dass ACEH einen Klasseneffekt zeigen, da sich vergleichbare Ergebnisse unter verschiedenen Substanzen nachweisen ließen. Es gibt zuminderst keine Evidenz für einen ‚besten ACEH‘, wobei jedoch nur Captopril, Enalapril, Lisinopril und Ramipril in mortalitätsbezogenen Outcomestudien evaluiert wurden [CCS 2002/2003].

Die Evidenz für eine dosisbezogene Effektivität ist nicht konsistent. Einige Studien zeigten eine signifikante Reduktion der Mortalität in höheren vs. niedrigen Dosierungen, die in anderen Studien nicht bestätigt wurden. Es liegen allerdings Hinweise dafür vor, dass sich durch hohe Dosierungen Vorteile für Hospitalisierungsrate und Symptomverbesserung erzielen lassen [DVA & VHA 2002].

2.6 Weitere Ergebnisse der Analyse der Belegtheit

In Ergänzung zu Kapitel 4.8 des Hauptdokuments werden hier die Kurzbeschreibung und Checkliste zur Prüfung der methodischen Qualität eines HTA-Berichts, eine Evidenztabelle zu Primärstudien sowie weitere Fallbeispiele präsentiert.

2.6.1 Kurzbeschreibung eines HTA-Berichts [Shekelle et al. 2003]

a) **Dokumenttyp und Bezugsrahmen:** Die *Agency for Healthcare Research and Quality* (AHRQ) ist die oberste Bundesbehörde der USA, welche die Forschung zu Qualitätsförderung, Kostendämpfung, Patientensicherheit und zu medizinischen Fehlern beaufsichtigt. Sie fördert mit ihren *Evidence-based Practice Centers* (EPC's) die Entwicklung von Evidenzreports und Technologiebewertungen. Die HTA-Berichte durchlaufen vor ihrer Publikation ein gestuftes internes und externes Reviewverfahren und sollen nach ihrer Veröffentlichung eine breite Gruppe von Adressaten (Leistungserbringer, Kostenträger, Patienten und politische Entscheidungsträger) ansprechen.

b) **Konkrete Fragestellung:** Sie entstand aus potentiellen Key Questions, die dem AHRQ vom *American College of Physicians*, der *American Society of Internal Medicine* und der *American Academy of Family Physicians* übermittelt wurden. Nach Prüfung auf Machbarkeit und Relevanz wurden sie gemeinsam mit einem Expertenpanel überarbeitet und im weiteren Bearbeitungsprozess auf die folgende konkrete Fragestellung fokussiert:

Welche Assoziationen bestehen zwischen verschiedenen Therapieansätzen (ACE-Hemmer, β -Rezeptorenblocker) und der Gesamtmortalität bei weiblichen und männlichen Patienten, Patienten mit und ohne Diabetes mellitus, farbigen und weißen Patienten mit Herzinsuffizienz?

Variieren diese Assoziationen mit Geschlecht, diabetischem Status und Rasse?
Der zweite Teil des Reports umfasst eine Kosteneffektivitätsanalyse.

c) **Zielpopulation:** In einer Anzahl von Studien wurde der mortalitätsreduzierende Benefit der Therapie mit ACE-Hemmern und β -Rezeptorenblockern bei erwachsenen Patienten mit chronischer Herzinsuffizienz gezeigt. Die meisten der eingeschlossenen Patienten waren weiß und männlich. Daraus ergab sich die Frage nach vergleichbarem Nutzen bei Subpopulationen, wie Frauen, Angehörigen anderer Rassen und Patienten mit Komorbiditäten (Diabetes mellitus).

d) **Methode: Recherche:** Initial wurde von den Untersuchern eine systematische Suche nach allen RCTs zur Wirksamkeit von ACE-Hemmern und β -Rezeptorenblockern durchgeführt, Datenbanken und Suchstrategien wurden expliziert. Da die dazu erhältlichen Publikationen nur in Einzelfällen stratifiziert zu den gesuchten Subpopulationen berichteten, musste die ursprünglich geplante Vorgehensweise geändert werden. *Ein- und Ausschlusskriterien:* Nicht explizit genannt. Einschluss von ‚Major RCTs‘ mit einer Studienpopulation $N > 1.000$ (außer CONSENSUS: eingeschlossen, weil 1. RCT zur Mortalität von ACE-Hemmern und weil von Expertenpanel ausdrücklich gewünscht). *Datenex-*

traktion: keine Angabe zur Zahl unabhängiger Reviewer, Datenextraktion erfolgte in Evidenztabelle. Patientenbezogene Daten wurden von einigen eingeschlossenen Studien ermittelt (für 2 RCTs via FDA-Approval-Verfahren, für weitere nach Feedback von angeschriebenen Autoren), weitere erforderliche Stratifikationen konnten aus inzwischen publizierten Subgruppenanalysen der Studien gewonnen werden. Die *methodische Qualität der Studien* wurde mit einem standardisierten, nicht-score-bildenden Instrument bewertet. *Datensynthese*: getrennte Metaanalysen für medizinische Wirksamkeit (Gesamtmortalität) von ACE-Hemmern und β -Rezeptorenblockern bei Männern und Frauen, Diabetikern und Nichtdiabetikern, weißen und farbigen Patienten; Verwendung des *Random-Effects-Models* wegen bestehender *between-study*-Varianz; Kalkulation von Relativem Risiko (RR) und Hazard Ratio (HR) für alle Outcomes und Ausgabe in log RR bzw. log HR sowie der rücktransformierten Form; *Hazard Ratios* wurden zusätzlich kalkuliert, da das Relative Risiko unterschiedliche Follow-up-Zeiten der Studien nicht ausreichend berücksichtigt und die eingeschlossenen Studien in diesem Punkt heterogen waren; Heterogenitätstests: χ^2 -Test; zur Untersuchung der Unterschiede der Subgruppen wurden Teststatistiken eingesetzt, in denen die *ratios of the relative risks* (RRR) und die *ratios of hazard ratios* (RHR) ermittelt wurden. Die Robustheit der Ergebnisse wurde in Sensitivitätsanalysen geprüft. Alle Methoden wurden ausreichend expliziert und ihre Auswahl begründet.

e) **Ergebnisse**: *Geschlechtsspezifische Wirksamkeit der ACE-Hemmer*. Einschluss von sieben RCTs (CONSENSUS, SAVE, SOLVD I & II, SMILE, TRACE und AIRE), davon nur für je sechs RRR und RHR kalkulierbar; Frauen: N=2.898, Männer N=11.674; Follow-up 6...42 Monate;

pooled RR Frauen=0,92 [95% CI 0,81-1,04], pooled RR Männer=0,82 [95% CI 0,74-0,90]; pooled HR Frauen=0,84 [95% CI 0,72-0,98], pooled HR Männer=0,76 [95% CI 0,66-0,87]; Unterschiede der RR und HR erklärt durch jeweils siebte Studie, die für beide Kalkulationsarten unterschiedliche Richtungen ergab; insgesamt wurde der Unterschied zwischen Männern und Frauen bei RRR statistisch signifikant;

Bei Differenzierung der Studienpopulationen in *asymptomatische und symptomatische Herzinsuffizienz*:

asymptomatische HI: pooled RR Frauen =0,96 [95% CI 0,75-1,22], pooled RR Männer=0,82 [95% CI 0,74-0,90]; symptomatische HI: pooled RR Frauen =0,90 [95% CI 0,78-1,05], pooled RR Männer=0,80 [95% CI 0,68-0,93];

Interpretation der Autoren: Männer profitieren unter ACE-Hemmertherapie bei symptomatischer und asymptomatischer HI mit reduzierter Mortalität. Frauen zeigen bei symptomatischer HI unter ACE-Hemmern einen Mortalitätsbenefit, der geringer ausfällt als bei Männern, ein Mortalitätsbenefit bei asymptomatischer HI konnte nicht gezeigt werden. Die Datenlage bei Frauen ist jedoch unzureichend und muss in weiteren Studien geprüft werden.

Wirksamkeit von ACE-Hemmern bei Diabetikern und Nichtdiabetikern: Einschluss von n=6 RCTs (CONSENSUS, SAVE, SOLVD I & II, SMILE und TRACE); Diabetiker N=2.398, Nichtdiabetiker N=10.188; Metaanalysen zeigen vergleichbare Mortalitätsreduktion bei Diabetikern und Nichtdiabetikern unter ACE-Hemmertherapie.

Geschlechtsspezifische Wirksamkeit der β -Rezeptorenblocker: Für die Kalkulation der RRs wurden n=4 RCTs (CIBIS II, COPERNICUS, MERIT-HF und US-Carvedilol) mit Frauen N=2.134, Männer N=7.885 eingeschlossen. Mortalitätsreduktion unter β -Rezeptorenblockern bei Männern und Frauen ähnlich.

Wirksamkeit von β -Rezeptorenblockern bei Diabetikern und Nichtdiabetikern: Einschluss von n=3 RCTs (CIBIS II, COPERNICUS, MERIT-HF); Diabetiker N=1883, Nichtdiabetiker N=7042; da bei CIBIS II keine HRs berechenbar, konnte nur Metaanalyse für RRs erfolgen; pooled RR Diabetiker=0,77 [95% CI 0,61-0,96], pooled RR Nichtdiabetiker=0,65 [95% CI 0,57-0,74], Unterschied zwischen beiden nicht statistisch signifikant.

Interpretation der Autoren: Diabetiker und Nichtdiabetiker profitieren bei HI unter β -Rezeptorenblockern mit reduzierter Mortalität. Der Ausmaß des Effekts bei beiden ist schwierig zu interpretieren: einerseits erscheint die relative Reduktion bei Diabetikern geringer, andererseits ist durch ihr deutlich höheres Risiko gegenüber Nichtdiabetikern mit einer größeren Reduktion des absoluten Mortalitätsrisikos zu rechnen.

f) **Diskussion:** Limitationen von Metaanalysen allgemein sind inhärente Bias der eingeschlossenen Studien, für die nicht adjustiert werden kann; darum hier nur doppelblinde, randomisierte RCTs mit Outcome Gesamtmortalität eingeschlossen, um diese Bias zu minimieren. Beschränkung auf ‚Major RCTs‘ wurde erforderlich wegen eingeschränktem Zugang zu stratifizierten Daten; Metaanalyse erfasste bei ACE-Hemmern N=14.932 eingeschlossene Patienten in n=7 RCTs, aus o.g. Gründen nicht eingeschlossene Patienten aus n=19 RCTs N=3.033; ähnlich bei β -Rezeptorenblockern: eingeschlossen N=12.776 in n=5 RCTs, nicht eingeschlossen n=19 RCTs mit N=2.938 Patienten.

g) **Schlussfolgerungen der Autoren:** Ausreichende Evidenzlage für Mortalitätsreduktion unter β -Rezeptorenblockern bei Männern und Frauen sowie bei Diabetikern und Nichtdiabetikern; keine Evidenz für einen Mortalitätsbenefit bei Frauen mit asymptomatischer Herzinsuffizienz unter ACE-Hemmertherapie, weitere Studien erforderlich; Frauen mit symptomatischer Herzinsuffizienz, Männer mit asymptomatischer oder symptomatischer Herzinsuffizienz und Diabetiker mit Herzinsuffizienz profitieren von ACE-Hemmern mit reduzierter Sterblichkeit.

h) **Abschließende Bewertung:** Umfassende Technologiebewertung für spezifische Fragestellung nach der Wirksamkeit von ACE-Hemmern und β -Rezeptorenblockern bei verschiedenen Teilpopulationen; alle Schritte wurden klar expliziert und begründet; adäquate Auswahl der Analysemethoden; übersichtliche Ergebnispräsentation; Schlussfolgerungen folgten festgestellter Evidenz. Aus Sicht d.Verf. kann ein potentieller Selection-Bias beim Studieneinschluss durch die Einflussnahme des Expertenpanels nicht ausgeschlossen werden – die generellen Aussagen zur Gesamtmortalität unter ACE-Hemmern und β -Rezeptorenblockern decken sich jedoch mit den Ergebnissen aus anderen Metaanalysen (vgl. Evidenztabellen). Da die Ergebnisse der Bewertung der methodischen Qualität der eingeschlossenen Studien nicht berichtet wurden, können keine Aussagen über sich daraus ergebende potentiell ergebnisverzerrende Effekte gemacht werden.

2.6.2 Checkliste eines HTA-Berichts [Shekelle et al. 2003]

Titel: Pharmacologic Management of Heart Failure and Left Ventricular Systolic Dysfunction: Effect in Female, Black, and Diabetic Patients, and Cost Effectiveness. Evidence Report / Technology Assessment No. 82							
Autoren: Shekelle P, Rich M, Morton S, Atkinson Col. S, et al.							
Quelle: AHRQ Publication No. 03-E045. Rockville, MD: Agency for Healthcare Research and Quality. July 2003.							
Dokumenttyp: Evidence Report/Technology Assessment							
Adressaten: Entscheidungsträger X Kliniker X Patienten Andere							
A. Fragestellung und Kontext					Ja	Nein	?
	Werden Anlass und Ziel der Publikation im Sinne einer „Policy Question“ dargestellt?				X		
	Gibt es im Rahmen des breiteren Kontext eine präzise formulierte Forschungsfrage nach der (interessierenden) Intervention?				X		
	Sind in der Publikation Angaben zu folgenden Aspekten enthalten:						
	Epidemiologie der Zielerkrankung				X		
	(Entwicklungs-)stand der Technologie				X		
	Efficacy				X		
	Effectiveness					X	
	Nebenwirkungen					X	
	Indikationen				X		
	Kontraindikationen					X	
	Praxisvariation					X	
	Versorgungsstrukturen						
	Kostengesichtspunkten				X		
	sozioökonomischem, ethischem und juristischem Impact				X		
B. Methodik der Informationsgewinnung							
	Wurden die genutzten Quellen dokumentiert?				X		
	Wurden die Suchstrategien dokumentiert?				X		
	Wurden Einschlusskriterien definiert?				X		
	Wurden Ausschlusskriterien definiert?				X		
C. Methodik der Bewertung der Dokumentation							
	Wurden Validitätskriterien berücksichtigt?				X		
	Wurde die Bewertung unabhängig von mehreren Personen durchgeführt?						X
	Sind ausgeschlossene Studien mit ihrem Ausschlussgründen dokumentiert?					X	
	Ist die Datenextraktion nachvollziehbar dokumentiert?				X		
	Erfolgte die Datenextraktion von mehreren Personen unabhängig?						X
D. Methodik der Informationssynthese							
	Es wurden quantitative Informationssynthesen durchgeführt:				X		
	Bei Metaanalyse:						

(Fortsetzung der Checkliste zum HTA-Report von Shekelle et al. 2003)

	Wurde die Auswahl des Meta-Analyse-Verfahrens begründet?	X		
	Wurden alle relevanten Outcomes berücksichtigt?	X		
	Wurden Heterogenitätstestungen durchgeführt?	X		
	Sind die Ergebnisse in einer Sensitivitätsanalyse auf Robustheit überprüft?	X		
	Es wurden qualitative Informationssynthesen durchgeführt:		X	
	Bei qualitativer Informationssynthese:			
	Ist die Informationssynthese nachvollziehbar dokumentiert?	X		
	Gibt es eine Bewertung der bestehenden Evidenz?	X		
	Es wurden zur Ergänzung der Datenlage eigene Erhebungen durchgeführt:	X		
E. Ergebnisse/Schlussfolgerungen				
	Wird die bestehende Evidenz in den Schlussfolgerungen konsequent umgesetzt?	X		
	Werden methodisch bedingte Limitationen der Aussagekraft kritisch diskutiert?	X		
	Werden Handlungsempfehlungen ausgesprochen?	X		
	Gibt es ein Grading der Empfehlungen?		X	
	Wurde die Publikation vor der Veröffentlichung einem externen Reviewverfahren unterzogen?	X		
	Ist ein „Update“ der Publikation eingeplant?		X	
F. Übertragbarkeit der internationalen/ausländischen Ergebnisse und Schlussfolgerungen				
	Bestehen Unterschiede hinsichtlich der / des:			
	Epidemiologie der Zielkondition?			X
	Entwicklungsstandes der Technologie?		X	
	Indikationsstellung?		X	
	Versorgungskontexte, -bedingungen, -prozess?	X		
	Vergütungssysteme?	X		
	Sozioökonomischen Konsequenzen?	X		
	Patienten- und Providerpräferenzen?			X
Beurteilung: Berücksichtigt: ja				

2.6.3 Beispiel einer Evidenztafel von Primärstudien

Nachfolgend wird eine beispielhafte Tabelle zu Primärstudien über die Wirksamkeit von Aldosteron-Antagonisten gezeigt. Gegenüber der im Methodenreport enthaltenen Version wurde sie geringfügig sprachlich überarbeitet.

Tab. A.14: Evidenztabelle zu Primärstudien über Aldosteronantagonisten

Quelle	Intervention	Studienpopulation, Begleittherapie	Ein- & Ausschlusskriterien ⁶⁹	Design	Outcomes	Datenanalyse	Ergebnisse	Bemerkungen	Studienqualität
Pitt et al. 1999 (RALES)	25 (-50) mg (ø 26 mg/d) Spironolacton 1x/d vs. Placebo; Vorzeitiger Abbruch nach 24 Mon. bei Nachweis der Therapieeffektivität	N=1663 (in 195 Zentren in 15 Ländern), Placebo N=841, Verum N=822; NYHA III (~70%) u. IV (~30%), EF≤35%, (ø 25%), ø-Alter 65+/12J., 73% Männer, ~50% ischaem. Genese; Basistherapie mit ACEH (ca. 95%), Schleifendiuretikum (100%), βRB (ca. 10%) und / oder Digitalis	Crea i.S. ≤2,5 mg/dl (221 μmol/l), K+ ≤5,0 mmol/l,	RCT, Multi-center, doppelblind	Prim. Outcome: Tod; Sek. Outcome: Tod aus kard. Urs., Hospitalisierung aus kard. Urs., Veränderung der NYHA-Klasse	Intention-to-treat, Kaplan-Meier-Kurven	<p><u>Gesamt mortalität:</u> RR=0,70 [95% CI 0,60 – 0,82], p<0,001; RRR=30%; §NNT=8 (24 Monate)</p> <p><u>Mortalität aus kard. Urs.:</u> RR=0,69 [95% CI 0,58 – 0,82], p<0,001 ; RRR=31% ;</p> <p><u>Hospitalisierung aus kard. Urs.:</u> RR=0,70 [95% CI 0,59 – 0,82], p<0,001; RRR=30% ;</p> <p><u>Kombin. Endpunkt Tod oder Hospitalisierung aus kard. Urs.:</u> RR=0,68 [95% CI 0,59 – 0,78], p<0,001; RRR=32% ;</p> <p><u>Veränderung der NYHA-Klassen Ver/Plac :</u> Verbesserung 41% / 33% Keine Veränderung 21% /18% Verschlechterung 38% /48%, Unterschied stat. sign. p<0,001;</p>	<p>Methode weitgehend klar expliziert; Große Zahl von Zentren (ø 8,5 Pat/Zentrum) – hohe Varianz von Pat. / Diagn./ Therapie; Fehlende Angaben zur Randomisat.-methoden u. Bestimmungsmeth. d. EF; Signifikante Zahl von Gynäkomastien / Mastodynien mit Therapieabbruch Nur NYHA III-IV-Pat. ohne Generalisierbarkeit für</p>	(Eingeschränkte Studienqualität)

⁶⁹ Hier nicht berichtet: allgemeine Ausschlusskriterien wie kausal therapierbare HI, Klappenvitien, Z.n. HTx oder HTx in Planung, lebensbedrohliche Zweiterkrankungen oder Tumorerkrankungen sowie Leberinsuffizienz

Quelle	Intervention	Studienpopulation, Begleittherapie	Ein- & Ausschlusskriterien ⁶⁹	Design	Outcomes	Datenanalyse	Ergebnisse	Bemerkungen	Studienqualität
		(ca. 75%) erlaubt					<u>Sicherheit:</u> unter Spironolacton Serum-Crea-Anstieg um 0,05 – 0,10 mg/dl, K+-Anstieg um 0,30 mmol/l im 1.Jahr, signifik. Gynäkomastie / Mastodynie (~10%)	andere NYHA-Klassen; Symptomverbesserung gebiast (NYHA-IV-Verschlechterung nur bei Tod mögl.)	
Ramires et al. 2000	25 mg Spironolacton / d (run-in mit 50 mg/d)	N=35 (Verum 19, Plac. 16), NYHA III, Männer 32/35, ischaem. Gen. 30/35	HI NYHA III bei ischaem. Gen. oder DCM, keine weiteren berichtet	RCT, nicht verblindet	VTs und VES im 24-h-Langzeit-EKG, belastungsinduzierte VTs und VES im Belastungstest	Standard-Mittelwert-Methode	(Reduktion ventrikulärer Arrhythmien unter Spironolacton)	Keine Mortalitätsstudie; kein Informationsbenefit; Keine Angaben zu Power-Kalk., Randomisationsmethoden, Studie ? nicht verblindet (k.A.), Outcomeparameter keine gesicherten Surrogate für lebensbedrohliche Arrhythmien	(Eingeschränkte method. und Berichtsqualität, ergebnisverzerrende Effekte wahrscheinlich)
Pitt et al. 2003 (EPHES US), Pitt et al.	25...50 mg Eplerenon/d + Standardth. vs. Pla-	N=6632 (Verum N=3313, Placebo N=3319),	3-14 d nach AMI, LV-EF≤40% und RGs	RCT, Multi-center (n=674 !), in-	<u>Primär:</u> Gesamtmortalität, kombinierter	Intention-to-treat, Cox-Proport.-Hazard-	<u>Primäre Endpunkte:</u> Gesamtmortalität: RR=0,85 [95% CI 0,75-0,96], p=0,008; ^{§§} NNT=43 (16 Monate)	Methoden klar expliziert, aber: für ACEH und ßRB keine Dosis berichtet, keine	(Eingeschränkte methodi-

Quelle	Intervention	Studienpopulation, Begleittherapie	Ein- & Ausschlusskriterien ⁶⁹	Design	Outcomes	Datenanalyse	Ergebnisse	Bemerkungen	Studienqualität
2001 (Studien design), Marcinia 2003 (FDA Medical Review) [§]	cebo + Standardtherapie (Dosisreduktion bei Serum-K ^{>} 5,5 mmol/l [#])	Alter: 64+/- 11J., LVEF 33+/-6% 72% Männer, Begleitmedik.: 86% ACEH, 75% β RB, 60% Diuretika, 88% ASS, 47% Statine; Anamnese: früherer AMI 27%, Diabetes 32%, HI 14%, Hypertension 60%	od. Stau im Rö.-Th. od. 3.HT (Diabetiker ohne Klinik akzeptiert); Ausschluss bei Serum-Crea>2,5 mg/dl (<220 μ mol/l), K ^{>} 5,0 mmol/l, unkontrollierte Hypotension mit systol. RR <90mmHg*	ternat. (37 Länder), Placebo-kontrolliert doppelblind, Follow-up bis 1012 Todesfälle aufgetreten (Mean-Follow-up 16 Mon.)	Endpunkt kardiovaskul. (CV) Mortalität oder CV-Hospitalisierung (Zeit bis z. ersten Auftreten eines Ereignisses) <u>Sekundär</u> CV-Mortalität, CV-Morbidität Hospitalisierung, Kombierter Endpunkt Gesamtmortalität + Hospitalisierung	Regressionsanalyse, stratifiziert nach geographischen Regionen; Time-to-event-Analyse in Kaplan-Meier-Kurven (bei Loss in Follow-up Auswertung als event);	CV-Mortalität oder – Hospitalisierung: RR=0,87 [95% CI 0,79-0,95], p=0,002; <u>Sekundäre Endpunkte:</u> Gesamtmortalität oder Hospitalisierung bei allen Urs. (kombinierter Endpunkt): RR=0,92 [95% CI 0,86-0,98]; CV-Mortalität (allein): RR=0,83 [95% CI 0,72-0,94] Hospitalisierung bei allen Urs. (allein), Hospitalisierung bei CV-Urs. (allein): keine sign. Ergebn.; <u>Unerwünschte Ereignisse (relev. für HI):</u> Schwere Hyperkaliämie (Serum-K \geq 6 mmol/l) in 5,5% der Verumgruppe vs. 3,9% in Placebogruppe (bei initial erniedrigter ECC \leq 50ml/Min. Hyperkaliämie bei 10,1% in Verumgruppe vs. 5,9% in Placebogruppe); Schwere Hypokaliämie (Serum-K<3,5 mmol/l) bei 8,4% in Verumgr. vs. 13,1% in Placebogr.; Gynaecomastie nicht häufiger unter Eplerenon	NYHA-Klass. od. Verlaufsbeschr. (EF-Bestimmung nur bei Randomisation) (pot. Biasquellen); Viele Zentren / Länder: keine Untersch. in der geograph. stratifizierten Auswertung (o. Darstellung); Einschluss nur nach AMI – Einschränkung der Generalisierbarkeit; Hyperkaliämierate unter engmaschigen Kontrollen (Studienbedingungen!) rel. niedrig, keine SAEs in Verumgruppe!; pot. Differenz zw. Efficacy & Efficiency!	sche und Berichtsqualität; Auswertung lässt eher konservativen Fehler erwarten; CAVE: Ergebnisse unter Alltagsbedingungen ggf. deutlich differierend)

2.6.4 Weitere Fallbeispiele

In Ergänzung zum Abschnitt 4.8.2 des Hauptdokuments werden hier weitere Ergebnisse der Bewertung von Primärquellen und HTA-Berichten dargestellt. Sie werden entsprechend der klassifizierten Konsistenztypen präsentiert.

Die weiteren *Typ-(1)-konsistenten* Empfehlungen zu β -Rezeptorenblockern bei NYHA II-IV und bei NYHA I nach stattgehabtem Myokardinfarkt wurden kurz zusammengefasst im Abschnitt Typ-B-Inkonsistenz zur gleichen Substanzgruppe bei asymptomatischen Herzinsuffizienzpatienten ohne Infarktereignis dargestellt (vgl. Abschnitt 4.8.2.4 des Hauptdokuments).

2.6.4.1 Weitere konsistente Empfehlungen vom Typ (2)

Digitalis bei tachykardem Vorhofflimmern zur Kontrolle der Ruhfrequenz:

^{*70}Bei Herzinsuffizienzpatienten mit chronischem Vorhofflimmern wird der Einsatz eines Digitalispräparates zur Kontrolle der Ruhfrequenz empfohlen. Zur Kontrolle der Frequenz, insbesondere unter Belastungen, sind β -Blocker jedoch besser geeignet und sollten nicht beendet werden [ACC/AHA 2001, AKDAE 2001 / DGK 2001, CCS 2002/2003, DVA & VHA 2002, ICSI 2003/2004, NHF/Austr. & CSANZ 2002, NHF/NZ 2001, NICE 2003, OPOT 2000, UM 2001, UWH 2001] $\{I_{a_{x1}} A_{x1}\}$. Die Kombination von Herzglykosiden mit einem β -Blocker ist den Einzelsubstanzen bei der Frequenzkontrolle überlegen [ESC 2005, DGK 2005]. Eine Digitalisierung führt nicht öfter als Placebo zu einer Konversion in den Sinusrhythmus [Duodecim 2004] $\{I_{a_{x1}} A_{x1}\}$.*

Aus den aufgeführten Evidenzquellen wurden ein systematischer Review mit qualitativer Informationssynthese über therapeutische Ansätze bei akutem und chronischem Vorhofflimmern (VHF) [Khand et al. 2000], eine Übersichtsarbeit zu akutem VHF [Lip & Freestone 2003] und ein RCT zur vergleichenden Wirksamkeit von Rhythmuskontrolle vs. Frequenzkontrolle bei VHF [AFFIRM 2002] reevaluiert. Khand et al. beurteilten die Evidenzlage zur Wirksamkeit verschiedener therapeutischer Ansätze bei VHF inkonsistent und ungenügend (wenige RCTs, kontrollierte nicht-randomisierte sowie unkontrollierte Studien). Für Digoxin ließ sich keine höhere Rate von Konversionen in den Sinusrhythmus nachweisen als unter Placebo. Zur Frequenzkontrolle allein war es bei akutem VHF nicht konsistent wirksam. [Lip & Freestone 2003] Bei chronischem VHF wurde keine verlässliche Wirksamkeit des Digoxins bei belastungsinduzierten Frequenzanstiegen beobachtet, Untersuchungen zur Kombination aus Digoxin und β -Blockern waren noch nicht abgeschlossen [Khand et al. 2000]. Ein Mortali-

⁷⁰ Dieser, wie auch die folgenden von Sternchen begrenzten eingerückten Textabschnitte entstammen dem Entwurf der Zielleitlinie, wie sie aus der Informationssynthese resultierten.

tätsbenefit für Herzinsuffizienzpatienten durch eine Rhythmisierung des VHF mit Konversion in den Sinusrhythmus konnte weder bei Khand et al. [2000] noch bei AFFIRM [2002] nachgewiesen werden. Alle existierenden Studien untersuchten Digoxin, keine untersuchte Digitoxin.

Die methodische und Berichtsqualität waren gut bei Khand et al. [2000], deutlich eingeschränkt bei AFFIRM [2002], und bei Lip & Freestone [2003] war sie bei nicht explizierem Methodenteil nicht zu bewerten.

Die in den berücksichtigten Leitlinien vergebenen, teilw. höchsten Evidenzlevel und Stärken der Empfehlung waren inadäquat und wurden korrigiert ($\{Ia_{x1} A_{x1}\}$ in $\{TIIa B\}$).

Orale Antikoagulation bei Herzinsuffizienz und VHF:

Bei Patienten mit Herzinsuffizienz und Vorhofflimmern oder vorausgegangenem thrombembolischem Ereignis wird eine orale Antikoagulation (INR 2 – 3) empfohlen [ACC/AHA 2001, AKDAE 2001 / DGK 2001, CCS 2002/2003, Duodecim 2004, DVA & VHA 2002, ESC 2001, ICSI 2004/2003, LLGH 2003, NHF/Austr. & CSANZ 2002, NHF/NZ 2001, NICE 2003, UM 2001, UWH 2001] $\{Ia_{x1} A_{x1}\}$.

In den berücksichtigten Leitlinien wurde darauf verwiesen, dass sich die Empfehlungen zur oralen Antikoagulation an den allgemeinen Empfehlungen bei VHF orientieren. Dieses Thema wurde nicht als originäre Fragestellung zur Herzinsuffizienz-Leitlinie angesehen, die angeführten Evidenzbelege wurden keiner erneuten Bewertung unterzogen.

Kontrolle eines arteriellen Hypertonus bei diastolischer Herzinsuffizienz:

Die Evidenzlage zur Therapie der diastolischen Herzinsuffizienz ist unzureichend, und der überwiegende Teil der Empfehlungen beruht ausschließlich auf Expertenkonsens. Es ist empfehlenswert, Patienten mit diastolischer Herzinsuffizienz in Zusammenarbeit mit einem Kardiologen zu betreuen [ACC/AHA 2001, AKDAE 2001 / DGK 2001, Duodecim 2004, DVA & VHA 2002, ICSI 2004/2003, NHF/Austr. & CSANZ 2002, NICE 2003] $\{IV_{x1} C_{x1}\}$ (empfehlenswert:) Kontrolle eines bestehenden systolischen wie diastolischen Hypertonus [nach ACC/AHA 2001, DVA / VHA 2002] $\{Ia_{x1} A_{x1}\}$.

Zur Kontrolle des arteriellen Hypertonus wurden 10 Publikationen zitiert: ein placebo-kontrollierter RCT an N=20 Patienten zum Effekt von Losartan auf Belastungstoleranz und Lebensqualität mit zweiwöchigem Follow-up [Warner et al. 1999], je ein RCT [Gottdiener et al. 1997] und eine Metaanalyse [Dahlof et al. 1992] zu Surrogat-Outcomeparametern (Reduktion der Ventrikelhypertrophie unter antihypertensiver Behandlung) und sieben traditionelle Reviews. Die Evi-

denzlage musste daher als ungenügend eingestuft werden, der hohe Evidenzlevel war nicht nachvollziehbar. Weitere Evidenzgewinnung ist erforderlich.

2.6.4.2 Weitere konsistente Empfehlungen vom Typ (4)

Körperliches Training:

Entgegen früheren Anschauungen, in denen der Patient unter der Vorstellung einer hämodynamischen Entlastung des Herzens zu körperlicher Schonung aufgefordert wurde, haben klinische Studien durchweg positive Effekte eines angepassten körperlichen Trainings gezeigt. Es kann Symptome vermindern, die Belastbarkeitsgrenzen erhöhen und die Lebensqualität verbessern [ACC/AHA 2001, AKDAE 2001, DKG 2001] $\{I_{a_{x1}} A_{x1}\}$.

Aus den Trefferlisten der initialen systematischen Leitlinienrecherche wurde eine kollaborative prospektive Metaanalyse individueller Patientendaten zur therapeutischen Wirksamkeit von körperlichem Training bei Herzinsuffizienzpatienten identifiziert [ExTraMATCH 2004]. In der Metaanalyse konnte eine signifikante Reduktion der Gesamtmortalität durch körperliches Training gezeigt werden. Das Training unter Aufsicht war in der Anwendung sicher. Auch der kombinierte Endpunkt aus Hospitalisierung und Sterblichkeit wurde positiv beeinflusst [ExTraMATCH 2004].

Methodische und Berichtsqualität waren gut. Da diese Ergebnisparameter den Outcomes der zitierten systematischen Reviews (Senkung von Hospitalisierungsrate und kardialer Mortalität, Verbesserung der Lebensqualität [Belardinelli et al. 1999 in DieM 2003/2004], Zunahme der körperlichen Leistungsfähigkeit und verbesserte Lebensqualität [Lloyd-Williams et al. 2002 in NICE 2003]) überlegen waren, wurden weder diese noch weitere Primärstudien reevaluiert. Für die angegebenen Empfehlungen und Evidenzlevel fanden sich ausreichende Belege, und die Therapieziele wurden um die Ziele Senkung von Mortalität und Hospitalisierung erweitert.

Hydralazin als Mittel 2. Wahl bei Kontraindikationen / Intoleranz von ACE-Hemmern oder AT₁-Blockern:

Hydralazin/ISDN ist bei Herzinsuffizienzpatienten mit Intoleranz bzw. Kontraindikationen gegen ACE-Hemmer und AT₁-Blocker zur Reduktion der Mortalität und Verbesserung der Belastungstoleranz einsetzbar. Diese Indikation ist in der Praxis jedoch sehr selten. Dabei ist Hydralazin/ISDN kein gleichwertiger Ersatz, da die Mortalität deutlich geringer als durch ACE-Hemmer beeinflusst wird [ACC/AHA 2001, AKDAE 2001, DKG 2001, DVA & VHA 2002, ESC 2002/2001, ICSI 2003/2004, NHF/Austr. & CSANZ 2002, NICE 2003, OPOT 2000, UM 2001, UWH 2001] $\{I_{b_{x1}} A_{x1}\}$.

Aufgrund der fehlenden Marktverfügbarkeit für die fixe Kombination Hydralazin/ISDN in Deutschland und der sekundären Rolle der Substanzen wurden die beiden von den Leitlinien evaluierten placebokontrollierten RCTs mit primärem Endpunkt Mortalität [Cohn et al. (V-HeFTⁿ) 1991, Fonarow et al. (Hy-C Trial) 1992] keiner erneuten Bewertung unterzogen.

Unbedenklichkeit von langanflutenden Dihydroperidinen:

Kalziumantagonisten sind für die primäre Therapie der Herzinsuffizienz nicht erforderlich. Patienten mit einer begleitenden arteriellen Hypertonie oder einer symptomatischen Koronaren Herzerkrankung (KHK) können langwirksame Kalziumantagonisten (z.B. Amlodipin) zur Blutdruckeinstellung bzw. zur Therapie der Angina pectoris-Beschwerden erhalten [AKDAE 2001 / DGK 2001, Duodecim 2004, DVA & VHA 2002, ESC 2002/2001, ICSI 2004/2003, NHF/Austr. & CSANZ 2002, NHF/NZ 2001, NICE 2003, UM 2001] {Ia_{x1} A_{x1}}.

Den Empfehlungen der Leitlinien lagen 11 RCTs, eine Übersichtsarbeit zu den Ergebnissen von zwei weiteren RCTs und 16 nicht-randomisierte Studien mit kleinen Fallzahlen zugrunde. Der zitierte systematische Review betraf nicht die Zielpopulation (herzinsuffiziente Patienten waren ausgeschlossen!). In Mortalitätsstudien konnte durch den Einsatz von Kalziumantagonisten bei Patienten, die bereits ACE-Hemmer erhielten, kein Überlebensvorteil gezeigt werden. Amlodipingaben blieben ohne negative Auswirkungen. Es wurden keine der zitierten Publikationen reevaluiert, da weitere Recherchen zu Sicherheitsaspekten geplant waren. Der vergebene hohe Evidenzlevel war jedoch nicht nachvollziehbar.

2.6.4.3 Weitere Typ-B-Inkonsistenzen

Digitalistherapie bei Herzinsuffizienz NYHA III-IV und Sinusrhythmus

*(i) Eine Digitalisierung ist bei Herzinsuffizienzpatienten NYHA III-IV mit Sinusrhythmus in Erwägung zu ziehen, die unter einer Therapie mit ACE-Hemmer + β -Blocker + Diuretika länger als 2-3 Monate symptomatisch bleiben oder sich unter der genannten Therapie verschlechtern, sofern keine Kontraindikationen bestehen [ACC/AHA 2001, AKDAE 2001 / DGK 2001, CCS 2002/2003, DVA & VHA 2002, ICSI 2003/2004, NHF/Austr. & CSANZ 2002, NHF/NZ 2001, NICE 2003, OPOT 2000, UM 2001, UWH 2001] {Ia_{x1} A_{x1}}.

(ii) Ziel der Therapie ist eine Verbesserung von Symptomatik, Lebensqualität sowie Belastungstoleranz und eine Senkung der Hospitalisierungshäufigkeit, Effekte auf die Mortalität wurden nicht nachgewiesen [ACC/AHA 2001, AKDAE 2001 / DGK 2001, CCS 2002/2003, DVA & VHA 2002, ICSI 2003/2004, NHF/Austr. & CSANZ 2002, NHF/NZ 2001, NICE 2003, OPOT 2000, UM 2001, UWH 2001] {Ia_{x1} A_{x1}}.

(iii) Da in verschiedenen klinischen Studien an Patienten, die gleichzeitig ACE-Hemmer und Diuretika erhielten, durch die Beendigung einer Digitalistherapie Verschlechterungen der Symptomatik, der Belastungstoleranz, der NYHA-Klasse sowie der linksventrikulären Funktion beobachtet wurden [ACC/AHA 2001, CCS 2002/2003, DVA & VHA 2002, NHF/Austr. & CSANZ 2002, NICE 2003, UWH 2001], sollten begonnene Digitalistherapien nicht ohne zwingenden Grund abgebrochen werden. Bei Beendigung sollte das klinische Monitoring des Patienten intensiviert werden {Ib_{x1} A_{x1}}.

(iv) Bei primär asymptomatischen Patienten (NYHA I) und Patienten, die unter einer Therapie mit ACE-Hemmer + β -Blocker + Diuretikum asymptomatisch geworden sind, sollte wegen potentiell proarrhythmischer Effekte keine Digitalistherapie begonnen werden [AKDAE 2001 / DGK 2001, CCS 2002/2003, DieM 2003, Duodecim 2004, DVA & VHA 2002, ICSI 2003/2004, LLGH 2003, NHF/NZ 2001, NICE 2003, OPOT 2000, UM 2001, UWH 2001] {VI_{x1} C_{x1}}.*

Von den Leitlinien wurden zur Digitalistherapie insgesamt fünf systematische Reviews und Metaanalysen angeführt [Adams et al. 2002, Hood et al. 2001, Jaeschke et al. 1990, Khand et al. 2000, Lip et al. 2002]. Zwei der genannten Arbeiten wurden re-evaluiert [Adams et al. 2002, Hood et al. 2001], zwei weitere werden hier nicht näher dargestellt, da sie Therapieeffekte bei gleichzeitigem Vorhofflimmern untersuchen [Khand et al. 2000, Lip et al. 2002], und die Publikation von Jaeschke et al. [1990] ergab keinen Informationsgewinn. Adams et al. [2002] zeigten in einer multivariaten Analyse an Patienten aus 2 RCTs im *withdrawal design*, dass sich Herzinsuffizienzsymptome nach dem Absetzen von Digitalis verschlechterten, eine Dosis- oder Spiegelabhängigkeit ließ sich nicht nachweisen. Hood et al. [2001] zeigte in einer Metaanalyse, dass Hospitalisierung und Symptomkontrolle unter Digoxin besser waren, Einflüsse auf die Mortalität konnten für Digoxin nicht nachgewiesen werden. Ein RCT [DIG 1997] lieferte dabei das größte Gewicht (verantwortlich für 88% beim Outcome Mortalität und 94% beim Outcome Hospitalisierung).

Methodische und Berichtsqualität waren bei Hood et al. [2001] gut bis leichtgradig eingeschränkt, bei Adams et al. [2002] eingeschränkt (Verwendung selten angewandter Techniken, die bei stark verkürzter Berichtsqualität nicht eindeutig nachvollziehbar waren, zudem wurden potentielle Ergebnisverzerrungen des *withdrawal*-Studiendesigns nicht diskutiert).

Zusätzlich wurde eine post-hoc Subgruppenanalyse [Rathore et al. 2002] der DIG-Studie [1997] re-evaluiert, die eine erhöhte Sterblichkeit bei Frauen unter Digitalistherapie berichtet hatte⁷¹: Die Gesamtmortalität aller Frauen war geringer als die der Männer in der Studie, ebenso die Mortalität der Frauen der Verumgruppe vs. der Männer der Verumgruppe und der Frauen der Kontrollgruppe vs. der Männer der Kontrollgruppe (sowie die Mortalität der Männer der Ve-

⁷¹ Nach Angaben der Autoren wurde von ihnen keine weitere Subgruppenanalyse an den öffentlich zugänglichen Daten der DIG-Studie vorgenommen.

rumgruppe vs. der Männer der Kontrollgruppe). Nur der Vergleich der Frauen der Verumgruppe mit den Frauen der Kontrollgruppe lieferte einen paradoxen Wirksamkeitseffekt – die Sterblichkeit war unter behandelten Frauen etwas größer als unter den Frauen der Placebogruppe. Biasanfälligkeit des Designs sowie Konstellation der Ergebnisse deuten darauf hin, dass es sich bei dem Ergebnis höchstwahrscheinlich um reinen Zufall handelt, prospektive Studien sind erforderlich. Bis zu deren Publikation kann das Ergebnis von Rathore et al. [2002] jedoch nicht ignoriert werden.

Fazit: für die Empfehlungen (i) bis (iv) konnten in den beiden re-evaluierten Sekundärpublikationen ausreichende Belege identifiziert werden. Im Ergebnis der Auswertung der Subgruppenanalyse wurde eine zusätzliche Empfehlung aufgenommen:

Bei Frauen sollte die Indikationsstellung besonders streng und unter Einbeziehung der Patientin erfolgen, da die Evidenzlage zum Netto-Nutzen (Symptomverbesserung vs. Risiko fraglich erhöhter Sterblichkeit) derzeit unklar ist [Rathore et al. 2002] {IV_{x1} C_{x1}}.

Angiotensin-II-Rezeptorblocker bei Patienten mit ACE-Hemmer-Intoleranz

*Patienten mit einer symptomatischen Herzinsuffizienz (NYHA II-IV), die keinen ACE-Hemmer tolerieren, können bei Fehlen von Kontraindikationen ersatzweise einen AT₁-Blocker erhalten [ACC/ AHA 2001, AKDAE 2001/ DGK 2001, DVA & VHA 2002, ESC 2002/2001, ICSI 2004/2003, LLGH 2003, NHF/Austr. & CSANZ 2002, NHF/NZ 2001, NICE 2003, OPOT 2000, UM 2001, UWH 2001] {Ia_{x1} A_{x1}}.

Behandlungsziele sind Verbesserung des Überlebens, Senkung der Hospitalisierungsrate und Verbesserung von Symptomen [ebenda] {Ia_{x1} A_{x1}}.*

Es wurden ein systematischer Review und zwei Metaanalysen zur Wirksamkeit der AT₁-Blocker von den berücksichtigten Leitlinien als Evidenzbelege angeführt [Jong et al. 2002, Dahloff et al. 1992, Sharma et al. 2000]. Der systematische Review wurde einer Reevaluation unterzogen [Jong et al. 2002] (Evidenz-tabelle siehe Anhang), die Metaanalysen von Dahloff et al. [1992] und Sharma et al. [2000] enthielten keine zusätzlichen Informationen, da keine anderen Primärstudien berücksichtigt wurden. Jong et al. [2002] fanden in ihrer prospektiv geplanten Metaanalyse keine signifikanten Mortalitätsunterschiede zwischen den Gruppen, die mit AT₁-Blockern vs. Placebo, AT₁-Blockern vs. ACE-Hemmern und unter AT₁-Blocker + ACE-Hemmer vs. ACE-Hemmer allein behandelt wurden, ähnliche Ergebnisse zeigten sie für die Morbidität. Nur wenn anstelle des *random effects model* mit dem *fixed effects model* gerechnet wurde, konnte ein signifikanter Morbiditätsvorteil der AT₁-Blocker vs. Placebo gezeigt werden. Nach einer post-hoc Power-Kalkulation der Autoren hätte jedoch

eine absolute Risikoreduktion von 2% in der Gesamtmortalität gezeigt werden können. Die methodische und Berichtsqualität der Publikation war gut.

Der kanadische Kurz-HTA [Ma & Boucher 2002] lieferte keine zusätzlichen Informationen. Die bei DieM 2003 zitierten RCTs, die dort den Anlass für Warnmeldungen gegen AT₁-Blocker als potentiell schädliche Substanzen gaben, wurden nicht reevaluiert, da sie bei der Metaanalyse eingeschlossen waren. Zu AT₁-Blockern ist weitere Evidenzgewinnung erforderlich.

Salz- und Flüssigkeitsrestriktion:

Die Empfehlungen gründeten sich auf Expertenkonsens und unterschieden sich im Ausmaß der Grenzwerte: während deutsche Leitlinien 2-3 g Kochsalz / Tag empfehlen, lagen die Empfehlungen angloamerikanischer Leitlinien bei 2-3 g *sodium* [=Natrium], was etwa der doppelten Menge Kochsalz entspricht. Es wurden von einzelnen Leitlinien „quasi-experimentelle“ Publikationen zitiert, keine dieser Arbeiten wurde reevaluiert. Weitere Evidenzrecherchen sind erforderlich.

β-Rezeptorenblocker bei stabilen NYHA IV-Patienten:

Siehe Abschnitt 4.8.2.4. des Hauptdokuments.

Sicherheit kurzfristiger Anwendungen von Kalziumantagonisten vom Verapamiltyp bei tachykardem VHF:

Es wurden keine Publikationen reevaluiert, da weitere Evidenzrecherchen zu Sicherheitsaspekten erforderlich sind.

Danksagung

Herrn Prof. Dr. med. Ferdinand M. Gerlach, MPH, danke ich herzlich für die Überlassung des Themas und für seine vielfältige Unterstützung der Arbeit. Auf seinen Vorarbeiten zur Entwicklung des DEGAM-Konzeptes für hausärztliche Leitlinien konnte ich aufbauen, und dank seines Gestaltungskonzeptes für das Institut erhielt ich die notwendigen Freiräume für die Erstellung meiner Promotionsarbeit.

Mein besonderer Dank gilt:

- Herrn Dr. med. Dipl.-Päd. Jochen Gensichen, MPH, der durch seine erfolgreiche Einwerbung von Fördermitteln des Bundesministeriums für Bildung und Forschung (FKZ: 01GI0205) die vorliegende Arbeit erst ermöglicht hat. Sein stetes Bemühen um Klarheit in Analyse und Aussage war mir ein fortwährender Ansporn.
- Herrn Dipl. Soz. Martin Beyer, der mir durch seine praktischen Erfahrungen im jahrelangen Umgang mit der evidenzbasierten Medizin half, mich in die Leitlinien-Thematik einzuarbeiten. Seine medizin-soziologisch geprägten gesellschaftswissenschaftlichen Ambitionen waren für mich Anreiz zur Beschäftigung mit Grenzproblemen zwischen Medizin und Gesellschaftswissenschaft. Seinen Hinweisen verdanke ich die Grundlagen zur Bearbeitung der Thematik um Normativität in Leitlinien sowie zahlreiche Anregungen zur sprachlichen Überarbeitung.
- Frau Gisela Kassner für die geduldige Durchsicht des Manuskripts und ihre Hilfe bei Layout und Formatierung der Arbeit.
- Frau Dipl.-Informationswirtin (FH) Birgit Schorsch für ihre Unterstützung bei der Literaturbeschaffung.
- Herrn Phillip Elliott für seinen Support bei der Übersetzung der Zusammenfassungen ins Englische.

Darüber hinaus bin ich meinen Kolleginnen und Kollegen, meinen Freunden und meinem Mann für ihre moralische Unterstützung, ihren Teamgeist und ihr Verständnis dankbar.

Curriculum vitae

Persönliche Daten

Name: Christiane Muth, MPH
Geboren: am 25. Mai 1963 in Berlin
Familienstand: verheiratet mit Herrn Manfred Muth

Studium

1982 – 1985 Humanmedizin an der Karl-Marx-Universität, Leipzig
1986 – 1991 Humanmedizin an der Westfälischen Wilhelms-Universität, Münster
2000 – 2001 Ergänzungsstudiengang *Public Health* an der Medizinischen Hochschule Hannover

Berufliche Tätigkeit

1991 – 1998 Ärztin im Praktikum und Weiterbildungsassistentin im Klinikum Minden in der *Medizinischen Klinik* mit den Schwerpunkten Gastroenterologie und allgemeine Innere Medizin, Kardiologie und internistische Intensivmedizin, Hämatologie und Onkologie, sowie Nephrologie und Dialyse und in der *Klinik für Rheumatologie, Endokrinologie und Physikalische Medizin*
1998 – 1999 Weiterbildungsassistentin in der Inneren Medizin im Kreis Krankenhaus Hameln und im Krankenhaus Rahden
1999 – 2000 Praxisvertretung in einer internistischen Gemeinschaftspraxis (6 Monate)
2001 Erstellung eines *Health Technology Assessment*-Berichts zur medizinischen Effektivität von Dünnschichtpräparationstechniken und computergestützten Untersuchungen von Zervixabstrichen im Rahmen der Krebsfrüherkennung (Deutsches HTA-Projekt, MHH Hannover)
2002 - 2003 Wissenschaftliche Mitarbeiterin in der Abteilung Epidemiologie, Sozialmedizin und Gesundheitssystemforschung der Medizinischen Hochschule Hannover: wissenschaftliche Koordination des ‚deutschen HTA-Projekts 1995 – 2001‘ und Methodenentwicklung zur premarket-Bewertung innovativer

Technologien (innovationsbezogenes HTA) innerhalb des Projekts Kompetenzzentrum Kardiovaskulärer Implantate (KKI)

Seit 01.01.2004 Wissenschaftliche Mitarbeiterin im Institut für Allgemeinmedizin, zunächst in der Universität Schleswig-Holstein, Campus Kiel, und seit 10.2004 an der Johann Wolfgang Goethe-Universität, Frankfurt am Main: Mitarbeit in BMBF-geförderten Projekten zur evidenzbasierten Leitlinienentwicklung, zur Entwicklung von Case Management-Konzepten in der Primärversorgung sowie zur unangemessenen Medikation bei älteren multimorbiden Patienten mit Polypharmakotherapie

Qualifikationen

1993 Ärztliche Approbation
1999 Gebietsarztanerkennung Innere Medizin
2003 Magistra Public Health;
Magisterarbeit zum Thema "Systematisches Qualitätsmanagement von Gebärmutterhalsabstrichen beim Screening auf Zervixkarzinome"

Publikationen, die in Zusammenhang mit der vorgelegten Arbeit bislang erschienen sind:

Monographien

Muth C, Gensichen J, Butzlaff M (2006) DEGAM Leitlinie Nr. 9, Herzinsuffizienz. Omikron Publishing, Düsseldorf. ISBN 3-936572-07-0

Muth C, Gensichen J, Butzlaff M (2006) DEGAM Leitlinie Nr. 9, Herzinsuffizienz. Leitlinie Langfassung Teil 2: Evidenz und Rationale. Omikron Publishing, Düsseldorf. ISBN 3-936572-10-0

Buchbeitrag in einem Sammelband

Muth C, Beyer M, Gensichen J, Gerlach FM (2007) Setting- und Kontexteinflüsse in evidenzbasierten Leitlinien. Ergebnisse und Erfahrungen aus dem Entwicklungsprozess der DEGAM-Leitlinie "Herzinsuffizienz". In: Kirch W, Badura B, Pfaff H (Hrsg.) Prävention und Versorgungsforschung, Springer Medizin Verlag, Heidelberg, S. 281-314

Artikel in Peer Reviewed Journalen

Muth C, Gensichen G, Beyer M, Gerlach F (2006) "Diagnostik und Therapie der chronischen Herzinsuffizienz. Teil 1: Grundlagen und Diagnostik." CME-Artikel. Z Allg Med 82: 31-40

Muth C, Gensichen G, Beyer M, Gerlach F (2006) "Diagnostik und Therapie der chronischen Herzinsuffizienz. Teil 2: Therapie." CME-Artikel. Z Allg Med 82: 76-90

Gensichen J, **Muth C**, Butzlaff M, Rosemann T, Raspe H, Müller de Cornejo G, Beyer M, Härter M, Müller UA, Angermann CE, Gerlach FM, Wagner E (2006) Die Zukunft ist chronisch: das Chronic Care-Modell in der deutschen Primärversorgung: Übergreifende Behandlungsprinzipien einer proaktiven Versorgung für chronische Kranke. Z ärztl Fortbild Quallsich 100(5): 365-374

Abstracts

Muth C, Gensichen J, Beyer M, Gerlach FM (2007) Setting- und Kontexteinflüsse in evidenzbasierten Leitlinien. Ergebnisse und Erfahrungen aus dem Entwicklungsprozess der DEGAM-Leitlinie Herzinsuffizienz. Vortrag auf dem 6. Deutschen Kongress für Versorgungsforschung und 2. Nationalen Präventionskongress, 24.-27.10.2007, Dresden

Muth C, Gensichen J, Beyer M, Gerlach FM (2007) A systematic guideline review as an efficient method in evidence based guideline development. Primary care management of chronic heart failure as a model. 4th G-I-N Conference 2007, 22.-

25.08.2007, Toronto, Canada. Oral presentation, abstract in: Abstractband, o.S. Im Internet erhältlich unter URL: http://www.g-i-n.net/download/files/b70_muth.pdf

Muth C, Gensichen J, Beyer M, Gerlach FM (2007) Do evidence-based guidelines require consensus? Results from the nominal group process of the DEGAM guideline on chronic heart failure. 4th G-I-N Conference 2007, 22.-25.08.2007, Toronto, Canada. Poster, abstract in: Abstractband, o.S.

Muth C, Beyer M, Gensichen J, Gerlach FM (2006) Evidence-based and normative recommendations in guidelines. Results of a systematic guideline review as the basis for the development of a German guideline on congestive heart failure. EGPRN Workshop. 19.10.-22.10.2006, Kavala, Greece. Oral presentation, abstract in: Abstractband, o.S.

Muth C, Gensichen J, Beyer M, Gerlach FM (2007) Do evidence-based guidelines require consensus? Experience from the development of the DEGAM Guideline on Congestive Heart Failure. Vortrag auf der 8. Jahrestagung des Deutschen Netzwerks Evidenzbasierte Medizin, vom 22.-24. März 2006 in Berlin

Muth C, Gensichen J, Beyer M, Sitter H, Gerlach FM (2006). DEGAM-Leitlinien formal und interdisziplinär konsentieren? Die Herzinsuffizienz-Leitlinie im Nominalen Gruppenprozess (NGP). Vortrag auf dem 40. Kongress der DEGAM, 21.09.-23.09.2006, Potsdam. In: Zeitschrift für Allgemeinmedizin 82:9-10 (Kongress-Abstracts)

Muth C, Gensichen J, Beyer M, Butzlaff M, Gerlach FM (2005). Die DEGAM Leitlinie Herzinsuffizienz: Verifikation eines systematischen Leitlinienreviews durch Evaluation der Originalquellen. Poster auf dem 39. Kongress der DEGAM, 29.09.-01.10.2005, Potsdam. In: Zeitschrift für Allgemeinmedizin (Kongress-Abstracts) 81:11

Muth C, Gensichen G, Beyer M, Gerlach F (2005) Systematischer Leitlinienreview – neue Wege bei der Entwicklung der DEGAM-Leitlinie Herzinsuffizienz. Poster auf der 6. Jahrestagung des Deutschen Netzwerkes Evidenzbasierte Medizin (DNEbM) e.V., 04.03.2005, Berlin. Abstract in: Z ärztl Fortbild Qual Gesundh wes 100:147-148

Muth C, Gensichen J, Gerlach FM (2004) Methoden und Stand der Entwicklung der DEGAM-Leitlinie "Herzinsuffizienz". Vortrag auf dem 38. Kongress der DEGAM, 23.-25.09.2004, Potsdam. In: Zeitschrift für Allgemeinmedizin (Kongress-Abstracts) 80:381

Sonstige Zeitschriftenartikel

Muth C, Gensichen J, Beyer M, Gerlach FM (2005) Herzinsuffizienz: state of the art. Was gilt in der Hausarztpraxis? Der Hausarzt 42(20):50-55

Muth C. (2005) Herzinsuffizienz-Leitlinie für die Hausarztpraxis. Damit Sie in der Info-Flut nicht untergehen. Der Hausarzt 42(20): 49

Ehrenwörtliche Erklärung

Ich erkläre ehrenwörtlich, dass ich die dem Fachbereich Medizin der Johann Wolfgang Goethe-Universität Frankfurt am Main zur Promotionsprüfung eingereichte Dissertation mit dem Titel:

„Der systematische Leitlinienreview. Eine Methode zur Entwicklung evidenzbasierter Leitlinien, dargestellt am Beispiel der Entwicklung der DE-GAM-Leitlinie 'Herzinsuffizienz'“

im Institut für Allgemeinmedizin

unter Betreuung und Anleitung von

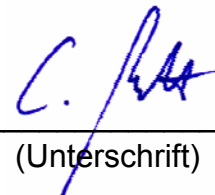
Herrn Prof. Dr. med. Ferdinand M. Gerlach, MPH

ohne sonstige Hilfe durchgeführt und bei der Abfassung keine anderen als die in der Dissertation angeführten Hilfsmittel benutzt habe.

Ich habe bisher an keiner in- oder ausländischen Universität ein Gesuch um Zulassung zur Promotion eingereicht. Die vorliegende Arbeit wurde bisher nicht als Dissertation eingereicht.

Frankfurt am Main, 07.07.2008

(Ort und Datum)



(Unterschrift)