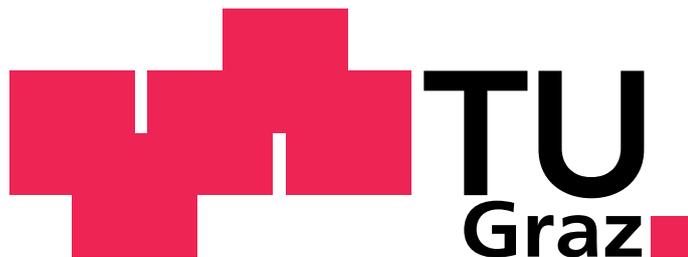


Stefan Viehhauser

Statistische Analysen zur Bewertung der Lebensqualität anhand des EQ-5D Fragebogens

Masterarbeit



Institut für Health Care Engineering
Technische Universität Graz
Kopernikusgasse 24, A-8010 Graz

Leiter: Univ.-Prof. Dr. N. Leitgeb

Betreuer: Assoc.-Prof. Dr. J. Schröttner

Begutachter: Assoc.-Prof. Dr. J. Schröttner

Graz, September, 2014

Deutsche Fassung:
Beschluss der Curricula-Kommission für Bachelor-, Master- und Diplomstudien vom 10.11.2008
Genehmigung des Senates am 1.12.2008

EIDESSTÄTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am

.....
(Unterschrift)

Englische Fassung:

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

.....
date

.....
(signature)

Danksagung

Ich möchte mich hier gerne bei all denen bedanken, die mich im Laufe meines Studiums unterstützt haben und mir zur Seite gestanden sind.

Ein ganz besonderer Dank gilt meinen Eltern. Sie haben es mir ermöglicht meinen Interessen zu folgen, und mich in meinem Tun jederzeit bekräftigt. Danke für eure Unterstützung und für den sicheren Rückhalt in jeder Lebenslage.

Weiters möchte ich mich recht herzlich bei Horst Mauschitz und Jessica Damberger für die hilfreichen Anmerkungen zur Masterarbeit bedanken.

Ein großer Dank gebührt auch meinen Freunden und Kollegen, die mich während meines Studiums begleitet haben. Weiters möchte ich mich auch bei allen Mitarbeitern des Institutes für die Unterstützung bedanken.

Zu guter Letzt möchte ich noch einen großen Dank an Veronika Mauschitz aussprechen. Auf ihren Rückhalt konnte ich jederzeit zählen, und ihre Unterstützung hat mir das Leben sehr erleichtert. Nochmals vielen Dank für alles.

Kurzfassung

Herzinsuffizienz ist eine Erkrankung, die vor allem ältere Bevölkerungsschichten betrifft. Aufgrund der demographischen Entwicklung steigt auch die Zahl der Herzinsuffizienz-Erkrankungen und die damit verbundenen Kosten. Um dem Kostenanstieg entgegenzuwirken, werden alternative Behandlungsmöglichkeiten untersucht.

Im Zuge dieser Arbeit soll eine Auswertestrategie für ein vorliegendes Studiendesign, welches sich mit dem Einsatz von Telemonitoring bei Herzinsuffizienz-Patienten beschäftigt, entwickelt werden. Um den Outcome einer Behandlungsgruppe mit dem einer Kontrollgruppe vergleichen zu können, werden die Lebensqualitätsfragebögen EQ-5D und KCCQ je dreimal zu unterschiedlichen Zeitpunkten ausgefüllt.

Hierzu wurde eine Literaturrecherche durchgeführt, um statistische Verfahren aus der Literatur zu berücksichtigen, mit dem vorliegenden Studiendesign zu vergleichen und eine Auswertestrategie zu entwickeln. Die Auswertestrategie wurde anschließend an einem Testdatensatz erprobt.

Zur Verteilungsanalyse werden der Shapiro-Wilk- und der Kolmogorov-Smirnov-Test herangezogen. Für Vergleiche zwischen zwei Gruppen kommen der T-Test bei unabhängiger Stichprobe, der Mann-Whitney-U-Test, der Chi-Quadrat-Test und der Fisher's-Exact-Test zum Einsatz. Für mehrere Gruppen werden die ANOVA und der Kruskal-Wallis-Test verwendet. Vergleiche innerhalb einer Gruppe werden mit dem T-Test bei verbundener Stichprobe bzw. dem Wilcoxon-Vorzeichenrang-Test durchgeführt.

Die entwickelte Auswertestrategie konnte am Testdatensatz erfolgreich erprobt werden, um Ergebnisse des vorliegenden Studiendesigns zu vergleichen.

Schlüsselwörter – Herzinsuffizienz, EQ-5D, KCCQ, Telemonitoring, Statistik

Abstract

Heart failure is an illness which concerns mainly people of higher age. Because of the demographic development, the number of people suffering from heart failure and the costs linked with it are constantly growing. To counteract the growing costs, alternative treatments are examined.

In the course of this work an evaluation strategy should be developed for the present study design, which deals with telemonitoring in heart-failure-patients. To be able to compare the outcome of an investigation group to that of a control group, the quality of life questionnaires EQ-5D and KCCQ are filled in three times.

In addition a literature research was done to consider the statistical procedures used in it, to compared them to the present study design and to develop an evaluation strategy. Further a data set is available to test the evaluation strategy.

The Shapiro-Wilk test and the Kolmogorov-Smirnov test are used for the distribution analysis. For comparisons between two groups the t-test for independent samples, the Mann-Whitney-U-Test, the Chi-Square test and the Fisher's-Exact test are applied. For several groups the ANOVA and the Kruskal-Wallis test are used. Comparisons within a group are carried out with the t-test for paired samples or the Wilcoxon signed-rank test.

The developed evaluation strategy could be successfully tested in a data set, to compare results of the present study design.

Keywords – Heart Failure, EQ-5D, KCCQ, Telemonitoring, Statistics

Inhaltsverzeichnis

1	Einleitung	13
1.1	Herzinsuffizienz	13
1.2	Studiendesign	15
1.3	Lebensqualitätsfragebögen	16
1.3.1	EuroQol-5Dimension	17
1.3.2	Kansans City Cardiomyopathy Questionnaire	18
2	Aufgabenstellung	19
3	Methoden	20
3.1	Literaturrecherche	20
3.2	Statistische Methoden	21
3.2.1	Deskriptive Statistik	21
3.2.2	Parametrische Tests	21
3.2.3	T-Test	22
3.2.4	ANOVA	23
3.2.5	Nichtparametrische Tests	24
3.2.6	Chi-Quadrat-Test	25
3.2.7	Shapiro-Wilk-Test	25
3.2.8	Kolmogorov-Smirnov-Test	26
3.2.9	Kruskal-Wallis-Test	26
3.2.10	Mann-Whitney-Test	26
3.2.11	Wilcoxon-Test	27
3.3	Entwicklung der Auswertestrategie	27
3.4	SPSS	28
4	Ergebnisse	29
4.1	Publikationen	29
4.2	Auswertestrategie	39
4.2.1	Rohdaten	39
4.2.2	Datenbereinigung	39
4.2.3	Deskriptive Analyse	41
4.2.4	Aufstellung und Überprüfung der Hypothese	42
4.3	Auswertung des Testdatensatzes	46
4.3.1	Testdatensatz	46

Inhaltsverzeichnis

4.3.2	Vergleich zwischen Gruppe A und B	46
4.3.3	Vergleiche innerhalb einer Gruppe	51
4.3.4	Vergleiche zwischen den NYHA-Klassen	56
4.3.5	Ergebnisse nach Datenbereinigung	57
5	Diskussion	61
5.1	Auswertestrategie	61
5.1.1	Vergleich zwischen den Gruppen	62
5.1.2	Vergleich innerhalb einer Gruppe	63
5.2	Auswertung des Testdatensatzes	64
5.3	Ergebnisse aus den Publikationen	67
6	Zusammenfassung	71
7	Literaturverzeichnis	72
8	Anhang	77

Abbildungsverzeichnis

1.1	Risikofaktoren, Komorbiditäten und Komplikationen von Herzinsuffizienz. Entnommen aus Ertl und Ruschitzka [4].	14
1.2	Das zu untersuchende Studiendesign.	16
3.1	Vorgehen bei der Entwicklung der Auswertestrategie.	27
4.1	Darstellung der EQ-5D und VAS Mittelwerte bezogen auf Befragungszeitpunkt und Erkrankung. (CHD – coronary heart disease; EQ-5D – EuroQol 5-dimensional instrument; HF – heart failure; VAS – Visual Analogue Scale). Entnommen aus Spiraki et al. [20]	30
4.2	Kaplan Meier Diagramm von Krankenhauseinweisungen und Todesfällen. Entnommen aus Scherr et al. [25]	33
4.3	Vorgehen bei der Auswertung des Testdatensatzes.	39
4.4	Vergleich zwischen den Gruppen.	43
4.5	Vergleich zwischen Zeitpunkten innerhalb einer Gruppe.	44
4.6	Vergleich zwischen NYHA-Klassen/Altersgruppen innerhalb einer Gruppe.	45
4.7	Klinische Zusammenfassung des KCCQ für die Ausgangswerte aller Probanden. Ein Vergleich zwischen Gruppe A und B.	48
4.8	Subkategorie Symptome des KCCQ für die Ausgangswerte aller Probanden. Ein Vergleich zwischen Gruppe A und B.	49
4.9	Subkategorie Lebensqualität des KCCQ für die Ausgangswerte aller Probanden. Ein Vergleich zwischen Gruppe A und B.	49
4.10	Klinische Zusammenfassung des KCCQ für alle Probanden, die die Fragebögen für alle vorgeesehenen Zeitpunkte ausgefüllt haben.	53
4.11	Verlauf der Mediane der „Klinischen Zusammenfassung“ des KCCQ für Gruppe A und B.	54
4.12	Verlauf der EQ-5D-Index-UK-Mediane für Gruppe A und B.	54
4.13	Verlauf der EQ-5D-VAS-Mediane für Gruppe A und B.	55
4.14	Verlauf der EQ-5D-Index-UK-Werte der einzelnen Probanden für Gruppe A.	55
4.15	EQ-5D-Index-UK-Werte in Abhängigkeit der NYHA-Klassen zu den unterschiedlichen Zeitpunkten für Gruppe A.	56
4.16	EQ-5D-Index-UK-Werte in Abhängigkeit der NYHA-Klassen zu den unterschiedlichen Zeitpunkten für Gruppe B.	57

Abbildungsverzeichnis

4.17	Gegenüberstellung zweier vergleichbarer Subkategorien des KCCQ und EQ-5D-Fragebogens bei ein und derselben Person (Blau: EQ-5D, Grün: KCCQ).	58
4.18	„Klinische Zusammenfassung“ des KCCQ. Vergleich zwischen Gruppe A und B zu Studienbeginn.	59
4.19	EQ-5D-Index-UK-Werte für die NYHA-Klassen 1 bis 3 zu Studienbeginn. Ohne Trennung zwischen Gruppe A und B.	60
4.20	EQ-5D-Index-UK-Werte für die 3 unterschiedlichen Altersgruppen zu Studienbeginn. Ohne Trennung zwischen Gruppe A und B.	60
8.1	Verlauf der EQ-5D-Index-UK-Werte der einzelnen Probanden für Gruppe B.	83
8.2	Verlauf der der „Klinischen Zusammenfassung“ des KCCQ der einzelnen Probanden für Gruppe A.	83
8.3	Verlauf der der „Klinischen Zusammenfassung“ des KCCQ der einzelnen Probanden für Gruppe B.	84
8.4	Verlauf der EQ-5D-VAS-Werte der einzelnen Probanden für Gruppe A. .	84
8.5	Verlauf der EQ-5D-VAS-Werte der einzelnen Probanden für Gruppe B. .	85

Tabellenverzeichnis

3.1	Keywords für die Literaturrecherche.	20
3.2	Sucheingabe in Pubmed inkl. der relevanten Treffer und der Treffer insgesamt.	20
4.1	Darstellung der EQ-5D Mittelwerte der einzelnen Gruppen.((a) Schlaganfall, (b) Herzinsuffizienz, (c) Pflegende). Entnommen aus Pinto et al. [21]	31
4.2	Vergleich der SF-36 Werte der Telemedizin-Gruppe zu Beginn und am Ende der Studie. Entnommen aus Myers et al.[28]	35
4.3	Gegenüberstellung der Dimensionen des EQ-5D-Fragebogens mit vergleichbaren Fragen des KCCQ.	40
4.4	Die Skalenniveaus der wichtigsten Parameter.	42
4.5	Vergleich der Ausgangswerte (Studien-Zeitpunkt 1) des Testdatensatzes. Angaben: MW=Mittelwert, SD=Standardabweichung, Median oder Anzahl (Prozent).	47
4.6	Vergleich der Ausgangswerte aller Probanden des Testdatensatzes, die die Fragebögen für alle vorgesehenen Zeitpunkte ausgefüllt haben. Angaben: MW=Mittelwert, SD= Standardabweichung, Median oder Anzahl (Prozent).	50
4.7	Vergleich zwischen Studien-Zeitpunkt 1 und 3 bei Gruppe A, für alle Probanden, die die Fragebögen für alle vorgesehenen Zeitpunkte ausgefüllt haben. Angaben: MW=Mittelwert, SD=Standardabweichung, Median. . .	51
4.8	Vergleich zwischen Studien-Zeitpunkt 1 und 3 bei Gruppe B, für alle Probanden, die die Fragebögen für alle vorgesehenen Zeitpunkte ausgefüllt haben. Angaben: MW=Mittelwert, SD=Standardabweichung, Median. . .	52
8.1	Darstellung der Anzahl der Probanden zu den 3 Unterschiedlichen Zeitpunkten der Studie.	77
8.2	p-Werte des Shapiro-Wilk-Tests auf Normalverteilung, für alle Probanden.	78
8.3	Vergleich der Werte zu Studien-Zeitpunkt 2 aller Probanden des Testdatensatzes, die die Fragebögen für alle vorgesehenen Zeitpunkte ausgefüllt haben. Angaben: Mittelwert \pm Standardabweichung / Median.	79
8.4	Vergleich der Werte zu Studien-Zeitpunkt 3 aller Probanden des Testdatensatzes, die die Fragebögen für alle vorgesehenen Zeitpunkte ausgefüllt haben. Angaben: Mittelwert \pm Standardabweichung / Median.	79

Tabellenverzeichnis

8.5	Vergleich der Werte zwischen Zeitpunkt 1 und Zeitpunkt 2 bei Gruppe A, für alle Probanden des Testdatensatzes, die die Fragebögen für alle vorgesehenen Zeitpunkte ausgefüllt haben. Angaben: Mittelwert \pm Standardabweichung / Median.	80
8.6	Vergleich der Werte zwischen Zeitpunkt 1 und Zeitpunkt 2 bei Gruppe B, für alle Probanden des Testdatensatzes, die die Fragebögen für alle vorgesehenen Zeitpunkte ausgefüllt haben. Angaben: Mittelwert \pm Standardabweichung / Median.	80
8.7	Vergleich der Werte zwischen Zeitpunkt 2 und Zeitpunkt 3 bei Gruppe A, für alle Probanden des Testdatensatzes, die die Fragebögen für alle vorgesehenen Zeitpunkte ausgefüllt haben. Angaben: Mittelwert \pm Standardabweichung / Median.	81
8.8	Vergleich der Werte zwischen Zeitpunkt 2 und Zeitpunkt 3 bei Gruppe B, für alle Probanden des Testdatensatzes, die die Fragebögen für alle vorgesehenen Zeitpunkte ausgefüllt haben. Angaben: Mittelwert \pm Standardabweichung / Median.	81
8.9	Darstellung der Anzahl der Probanden zu den 3 Unterschiedlichen Zeitpunkten nach Datenbereinigung.	82
8.10	Vergleich der Werte zwischen den Gruppen für alle Probanden zu Zeitpunkt 1 nach Datenbereinigung. Angaben: Mittelwert \pm Standardabweichung / Median.	82

Abkürzungsverzeichnis

ACE	Angiotensin Converting Enzyme
ANOVA	Analysis of Variance
BRUM-CHF	Birmingham Rehabilitation Uptake Maximisation study for patients with Congestive Heart Failure
CHD	Coronary Heart Disease
CIED's	Cardiovascular Implantable Electronic Devices
EKG	Elektrokardiogramm
EQ-5D-3L	EuroQol – 5 Dimension – 3 Level
EQ-5D-5L	EuroQol – 5 Dimension – 5 Level
FAIR-HF	Ferric carboxymaltose Assessment in patients with Iron deficiency and chronic Heart Failure
HF	Heart Failure
IBM	International Business Machines Corporation
ID	Identifikation
INTENSE-HF	Integrated Telemonitoring and Nurse Support Evaluation in Heart Failure
KCCQ	Kansas City Cardiomyopathy Questionnaire
MLHFQ	Minnesota Living With Heart Failure Questionnaire
MOBITEL	Mobile Telemonitoring in Heart Failure Patients
NYHA	New York Heart Association
PASW	Predictive Analysis Software
Q-Q	Quantile-Quantile
SF-36	Short Form - 36
SPSS	Statistical Package for the Social Sciences
TELEBIL	Telemonitoring Bilbao
TELE-HF	Telemonitoring to Improve Heart Failure Outcomes
TEMA-HF	Telemonitoring in the Management of Heart Failure
TIM-HF	Telemedical Interventional Monitoring in Heart Failure
UK	United Kingdom
VAS	Visual Analogue Scale

Allgemeine Orientierungshinweise

Zugunsten einer besseren Lesbarkeit wurde in dieser Arbeit auf eine gendergerechte Formulierung verzichtet. Alle männlichen Formulierungen, wie z.B. Proband oder Patient, sollen daher sinngemäß auch weibliche miteinschließen und sind somit sowohl in weiblicher als auch in männlicher Form aufzufassen.

Mit dem Ausdruck EQ-5D ist in dieser Arbeit, mit Ausnahme des Kapitels „Publikationen“, die Version EQ-5D-5L zu verstehen. Beim Kapitel „Publikationen“ und den daraus resultierenden Erläuterungen bezieht sich EQ-5D auf die Version EQ-5D-3L.

1 Einleitung

Das Gesundheitswesen in Österreich steht heutzutage vor einigen großen Herausforderungen. Eine davon betrifft die steigenden Kosten, die unter anderem durch die demographische Entwicklung mitbegründet werden. Durch die höhere Lebenserwartung treten auch Erkrankungen häufiger auf, die mit dem Alter einhergehen, wie zum Beispiel Herzinsuffizienz. Damit wachsen auch die Kosten, die für Behandlungen aufgewendet werden. Um diese zu reduzieren, werden Behandlungsmethoden optimiert. [1], [2]

Bezogen auf die Herzinsuffizienz wird vermehrt der Einsatz von Telemonitoring-Systemen untersucht. Für den Vergleich des Outcomes von Behandlungsmethoden werden unter anderem auch Lebensqualitätsfragebögen verwendet.

Die vorliegende Arbeit befasst sich mit dem Vergleich zweier Behandlungsmethoden bei Herzinsuffizienz unter der Verwendung von Lebensqualitätsfragebögen und der damit verbundenen statistischen Auswertung.

1.1 Herzinsuffizienz

Herzinsuffizienz ist eine Erkrankung, bei der das Herz nicht mehr in der Lage ist, eine ausreichende Förderleistung zu erbringen, die den Anforderungen des täglichen Lebens entspricht. Diese unzureichende Funktion des Herzens kann unterschiedlich eingeteilt werden. Zum Beispiel in eine der 4 Gruppen der New York Heart Association (NYHA) nach Schweregrad der Erkrankung, wobei der Schweregrad von Gruppe 1 bis 4 steigt. Weitere Einteilungen sind z.B. nach betroffenem Herzabschnitt oder nach Verlauf, ob chronisch oder akut. Die Ursachen einer Herzinsuffizienz sind sehr unterschiedlich und reichen von Herzinfarkt über angeborene oder erworbene Herzfehler und arterielle oder pulmonale Hypertonie bis zu Myokarditis. Die Symptome sind abhängig von der Schwere und Art der Erkrankung. Typische Symptome sind z.B. Ödeme, Herzvergrößerung, Verminderung der Blutversorgung, Zyanose usw. [3]

Es gibt unterschiedlichste Risikofaktoren, die die Entwicklung einer Herzinsuffizienz begünstigen. Aufgrund der Komorbiditäten und der unterschiedlichen Komplikationen ist Herzinsuffizienz eine sehr komplexe Erkrankung. Abbildung 1.1 zeigt einen kurzen Überblick über die Risikofaktoren, Komorbiditäten und Komplikationen von Herzinsuffizienz. [4]

Herzinsuffizienz ist heutzutage eine der am häufigsten festgestellten internistischen Erkrankungen. Ein Großteil der Herzinsuffizienz-Patienten ist über 65 Jahre alt. Aufgrund

1 Einleitung

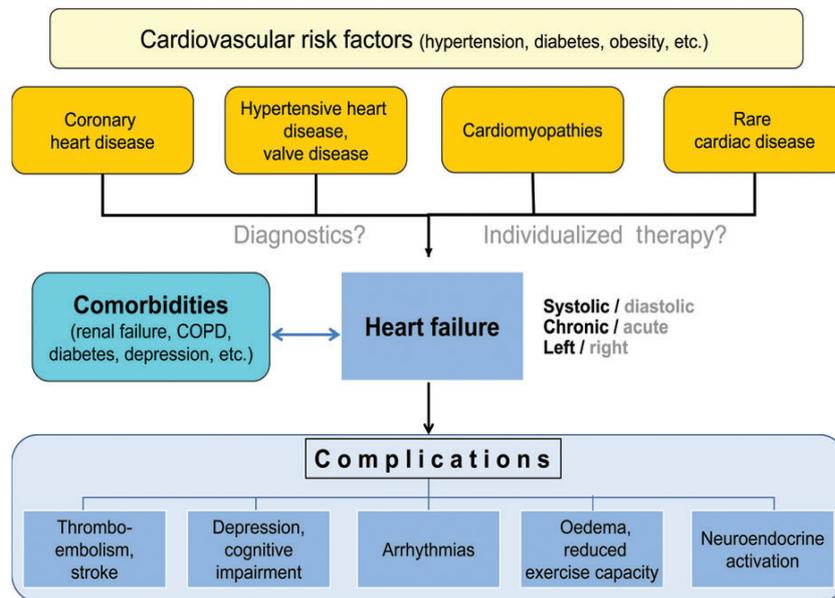


Abbildung 1.1: Risikofaktoren, Komorbiditäten und Komplikationen von Herzinsuffizienz. Entnommen aus Ertl und Ruschitzka [4].

dieser Tatsache und der immer älter werdenden Bevölkerung wird die Anzahl der Herzinsuffizienz-Erkrankungen zukünftig noch weiter steigen. Dadurch wachsen auch die Kosten für die Behandlung von Herzinsuffizienz. Es wird daher nötig sein, kosteneffektive Präventions- und Therapiemaßnahmen zu entwickeln. [2], [5]

Herzinsuffizienz-Patienten erfahren eine starke Beeinträchtigung der Lebensqualität. Deren Wichtigkeit wird durch eine Studie von Rector et al. unterstrichen. Darin haben sich mehr als die Hälfte der Patienten für eine Verbesserung der Lebensqualität entschieden, auch wenn dies eine Verkürzung des Lebens zur Folge hat. [6]

Die Lebensqualität von Herzinsuffizienz-Patienten ist vor allem in physischer, psychischer und sozialer Hinsicht beeinträchtigt. Der Grad der Beeinträchtigung ist dabei von der Schwere der Erkrankung abhängig. Juenger et al. haben den Beeinträchtigungsgrad in Abhängigkeit der NYHA-Klassen untersucht und diesbezüglich einen linearen Trend festgestellt. [7]

Die physische Beeinträchtigung geht oft so weit, dass Patienten weder ihrer Arbeit noch ihrer Freizeitaktivitäten nachgehen können. Weiters folgen häufige Krankenhausaufenthalte, und soziale Gewohnheiten müssen geändert werden. Dadurch sinkt das Selbstwertgefühl, und Patienten können in Depression verfallen. Deshalb ist es sehr wichtig, dass solche Patienten die richtige Behandlung und ausreichende Unterstützung erhalten.

Das primäre Ziel der Behandlung von Herzinsuffizienz-Patienten ist es, die Symptome zu minimieren, um somit die Lebensqualität zu verbessern. Deshalb sollten sich Studien nicht nur mit der Messung der Lebensqualität, sondern auch mit der Verbesserung von Behandlungsmethoden beschäftigen. [8]

Diagnostiziert wird die Herzinsuffizienz unter anderem mittels EKG, Thorax-Röntgen und Echokardiographie. So können erste Hinweise auf die Erkrankung, Herzvergrößerungen und Herzfehler festgestellt werden. Nach der Minimierung der Symptome wird, wenn möglich, die Ursache durch Operationen oder Medikamente beseitigt. Häufig eingesetzte Medikamente sind ACE-Hemmer, β -Blocker und Diuretika. [9]

1.2 Studiendesign

Das zu untersuchende Studiendesign beruht auf der laufenden Studie INTE-NSE-HF. Sie beschäftigt sich mit der Effektivität einer telemonitorischen Überwachung von Herzinsuffizienz-Patienten und ist eine prospektive, randomisierte, kontrollierte, multizentrische Studie. [10]

Telemedizinische Betreuung von Herzinsuffizienz-Patienten soll helfen, rechtzeitig gegenzusteuern, um eine Verschlechterung der Erkrankung zu verhindern. Durch die optimierte Behandlung und verbesserte Überwachung der Patienten soll eine erneute Aufnahme im Krankenhaus vermieden werden. Um die Effektivität der telemedizinischen Betreuung bewerten zu können, wird der Behandlungsgruppe eine Kontrollgruppe gegenübergestellt. Die Kontrollgruppe erfährt dabei eine normale Behandlung, wie sie zurzeit für Herzinsuffizienz-Patienten üblich ist, kombiniert mit pflegerischen Hausbesuchen. Die Behandlungsgruppe erhält noch zusätzlich eine telemedizinische Betreuung. Das Telemonitoring-System besteht aus Körperwaage, Blutdruckmessgerät, Mobiltelefon und Symbolkarte. Mittels „Near Field Communication“ werden die Daten an das Mobiltelefon und von dort an eine Datenzentrale weitergeleitet. Eine automatische Analysesoftware wertet die Daten aus und stellt dem Arzt Behandlungsvorschläge zur Verfügung.

Zusätzlich zu den Daten, die mit Hilfe von Telemonitoring aufgezeichnet werden, werden noch Daten bezüglich der Lebensqualität erhoben. Zu diesem Zweck füllen sowohl die Patienten der Behandlungsgruppe als auch die Patienten der Kontrollgruppe zwei Fragebögen im Verlauf der Studie je dreimal aus. Einmal werden die Fragebögen zu Beginn der Studie ausgefüllt, einmal nach 6 Monaten und ein letztes Mal nach 12 Monaten. Bei den Fragebögen handelt es sich um den Kansas City Cardiomyopathy Questionnaire (KCCQ) und den EuroQol-5-Dimension-Fragebogen (EQ-5D).

Das Diagramm 1.2 zeigt einen Überblick des Studiendesigns.

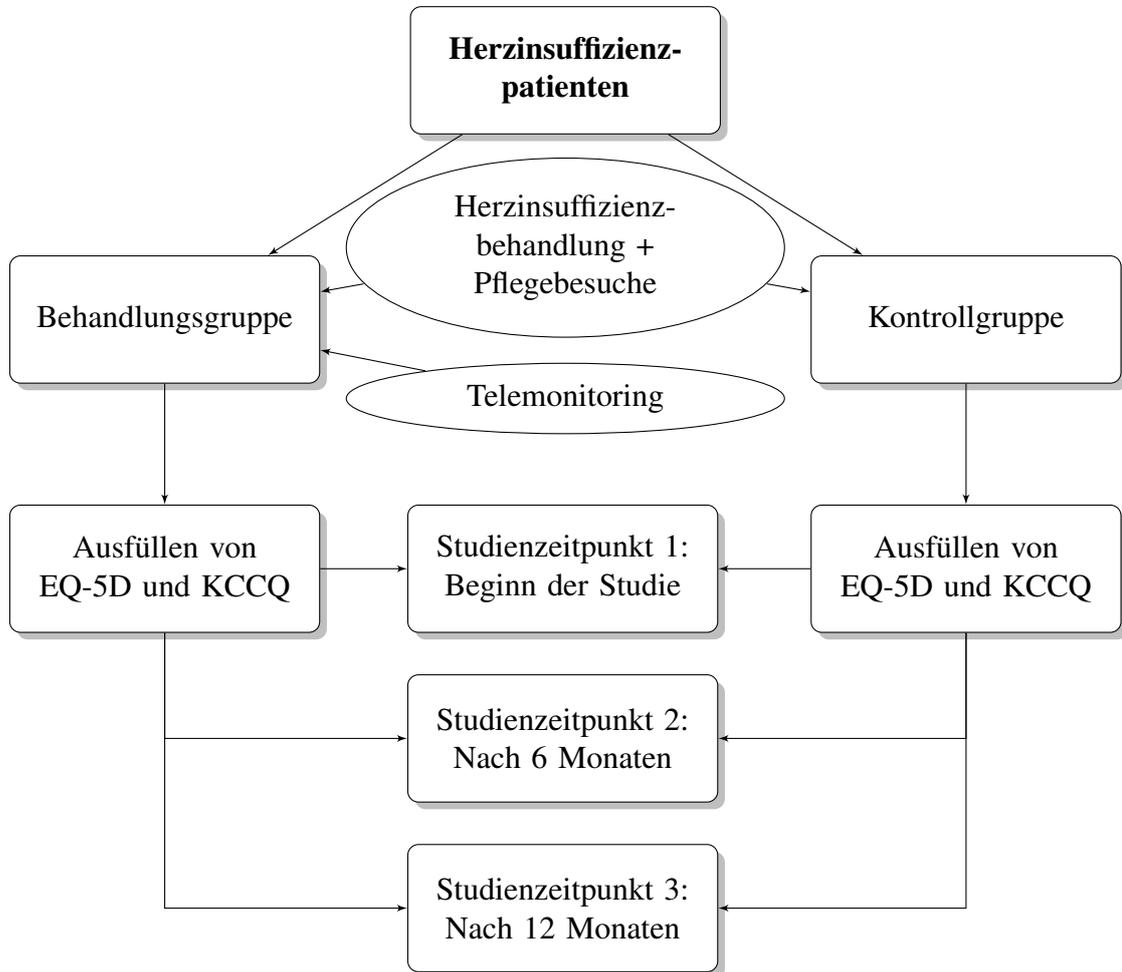


Abbildung 1.2: Das zu untersuchende Studiendesign.

1.3 Lebensqualitätsfragebögen

Bei den Lebensqualitätsfragebögen wird zwischen generischen und krankheitsspezifischen Fragebögen unterschieden. Im Gegensatz zu den krankheitsspezifischen Fragebögen zielen die generischen nicht auf eine bestimmte Erkrankung ab, sondern beziehen sich auf den allgemeinen Gesundheitszustand.

Das vorliegende Studiendesign verwendet zwei Lebensqualitätsfragebögen. Einerseits den EQ-5D als generisches Instrument zur Ermittlung der Lebensqualität, und andererseits den KCCQ als krankheitsspezifischen Fragebogen.

1.3.1 EuroQol-5Dimension

Der EQ-5D-Fragebogen wurde er von der EuroQol Gruppe entwickelt. Diese wurde 1987 gegründet und ist ein Netzwerk aus Forschern, die sich mit der Messung des Gesundheitszustandes beschäftigen.

Der EQ-5D-5L Fragebogen stellt die neueste Version dar. EQ steht dabei für die EuroQol Gruppe, 5D für die 5 unterschiedlichen Dimensionen und 5L für die 5 Ebenen (Level) der Gesundheit. Die Vorgängerversion EQ-5D-3L hatte nur 3 Ebenen.

Die 5 Dimensionen sind:

1. Beweglichkeit / Mobilität
2. Für sich selbst sorgen
3. Alltägliche Tätigkeiten
4. Schmerzen / körperliche Beschwerden
5. Angst / Niedergeschlagenheit

Diese 5 Dimensionen unterteilen sich weiter in 5 Ebenen. Damit gibt der Patient an, wie stark er in einer Dimension beeinträchtigt ist und welche Probleme er hat. Die 5 Ebenen sind:

1. keine Probleme
2. leichte Probleme
3. mäßige Probleme
4. große Probleme
5. extreme Probleme

Aus der Kombination von Dimension und Ebene ergibt sich ein Gesundheitsprofil, welches mithilfe von Value Sets in einen Index-Wert umgerechnet werden kann. Da die Value Sets für den EQ-5D-5L Fragebogen noch in der Entwicklung sind, kann mittels Cross-Walk-Tabellen über die Value Sets des EQ-5D-3L-Fragebogens der Index-Wert berechnet werden.

Weiters beinhaltet der EQ-5D-Fragebogen auch eine visuelle Analog-Skala. Diese VAS ist eine vertikale Linie mit einer Länge von 20 cm. Unterteilt wird diese Linie gleichmäßig in die Zahlen 0 bis 100, wobei 100 dem besten und 0 dem schlechtesten Gesundheitszustand entspricht. Der Patient kann dabei mit einem X seinen momentanen Gesundheitszustand vermerken und die entsprechende Zahl in ein Kästchen daneben eintragen. [11], [12]

Dyer et al. bestätigen in ihrem Review-Artikel die Validität und Zuverlässigkeit des EQ-5D-Fragebogens. Sie bezeichnen ihn auch als nützliches Werkzeug, um die Auswirkungen von kardiovaskulären Erkrankungen auf Patienten festzustellen. [13]

1.3.2 Kansans City Cardiomyopathy Questionnaire

Der KCCQ ist ein krankheitsspezifischer Fragebogen, welcher zur Ermittlung der Lebensqualität bei chronischer Herzinsuffizienz entwickelt wurde.

Er besteht aus 15 Fragen, wobei sich Frage 1 aus 6 Unterfragen und Frage 15 aus 4 Unterfragen zusammensetzen. Dabei sind je Frage 5-6 Antwortmöglichkeiten gegeben, wobei die Antwortmöglichkeit 1 immer dem schlechtesten Gesundheitszustand entspricht. Aus diesen insgesamt 23 Antworten lassen sich 6 Subskalen generieren, indem die zusammengefassten Antworten in einen Wertebereich von 0 bis 100 umgerechnet werden.

Diese Subskalen sind:

1. Physische Einschränkungen
2. Symptome
3. Symptomstabilität
4. Selbstwirksamkeit
5. Lebensqualität
6. Soziale Einschränkungen

Die Subskala „Symptome“ kann noch in die Subskalen „Symptommhäufigkeit“ und „Symptomschwere“ unterteilt werden. Bei der Subskala „Lebensqualität“ handelt es sich um die psychische Lebensqualität.

Weiters können die Subskalen noch in zwei Summenskalen zusammengefasst werden. Die Summenskala „Funktioneller Status“ umfasst die „Physischen Einschränkungen“ und „Symptome“. Die Summenskala „Klinische Zusammenfassung“ besteht aus allen Subskalen mit Ausnahme der „Symptomstabilität“ und „Selbstwirksamkeit“.

Die Sub- und Summenskalen bestehen aus einen Wertebereich von 0 bis 100. 0 stellt dabei den schlechtesten und 100 den besten Gesundheitszustand dar.

Der KCCQ versucht möglichst alle Bereiche der Gesundheit abzudecken, die für die chronische Herzinsuffizienz relevant sind. Weiters wurden die Validität und die Zuverlässigkeit des KCCQ bei der Ermittlung der Lebensqualität von Patienten mit chronischer Herzinsuffizienz bestätigt. [14], [15]

Neubauer kommt in ihrer Arbeit zu dem Schluss, dass sich der KCCQ auch im Vergleich mit anderen krankheitsspezifischen Fragebögen, z.B. dem MLHFQ, behauptet [16].

2 Aufgabenstellung

Ziel dieser Arbeit war es, eine statistische Auswertestrategie für die Daten des beschriebenen Studiendesigns zu entwickeln. Dazu sollten wichtige Publikationen und die verwendeten Verfahren in der Literatur berücksichtigt werden. Statistische Analysen zur Bewertung der Lebensqualität von Herzinsuffizienzpatienten standen dabei im Mittelpunkt.

Am Beginn dieser Arbeit erfolgte eine Einarbeitung in das Thema Herzinsuffizienz und eine Auseinandersetzung mit den Lebensqualitätsfragebögen KCCQ und EQ-5D. Eine Einarbeitung in die Statistik und die statistischen Verfahren wurde ebenfalls durchgeführt. Als Vorbereitung auf die Literaturrecherche sollte eine Suchstrategie mit Keywords ausgearbeitet werden.

In der Literaturrecherche war vor allem auf die statistischen Auswerteverfahren von vergleichbaren Studien mit der Anwendung eines Lebensqualitätsfragebogens zu achten. Dabei sollte nach relevanten Literaturstellen recherchiert werden und die Vergleichbarkeit des Probandenkollektivs hinsichtlich des zu untersuchenden Studiendesigns berücksichtigt werden.

Die Ergebnisse aus der Literaturrecherche waren entsprechend darzustellen, auszuwerten und zu interpretieren. Bei der Entwicklung der Auswertestrategie und der Auswahl von geeigneten Verfahren sollten die allgemeinen Voraussetzungen (z.B. Skalenniveau und Verteilung), Verfahren aus der Literatur und die Datenerhebung des beschriebenen Studiendesigns beachtet werden.

Die entwickelte Auswertestrategie war dann mit der Software SPSS zu testen, und diesbezüglich sollte eine Arbeitsanweisung erstellt werden. Zur Anwendung der Auswertestrategie wurde ein Testdatensatz herangezogen.

3 Methoden

3.1 Literaturrecherche

Aufbauend auf die Literaturrecherchen von Neubauer [16] und Oblinger [17] wurden weitere Keywords hinzugefügt und die Suche nach aktuellen Literaturstellen verfeinert. Die Hauptsuche nach relevanter Literatur erfolgte in Pubmed. Die iterative Suche wurde in Google durchgeführt. DIMDI, TUGraz Library Search, Google-Books und Google-Scholar lieferten zusätzlich keine relevanten Treffer. Tabelle 3.1 zeigt die verwendeten Keywords für die Literatursuche, die dann für die Sucheingabe kombiniert wurden.

<u>Keywords</u>
heart failure
herzinsuffizienz
eq 5d
kccq
<u>telemonitoring</u>

Tabelle 3.1: Keywords für die Literaturrecherche.

Die Suchergebnisse in Pubmed wurden so sortiert, dass die aktuellsten Publikationen mit der höchsten Trefferanzahl des Suchbegriffes an erster Stelle gereiht wurden. Bei der Auswahl der Literatur wurde zuerst der Titel betrachtet. Wies der Titel auf die gesuchte Thematik hin bzw. kam darin der Suchbegriff vor, wurde der Abstract eingesehen. Behandelte der Abstract die gesuchte Thematik, wurde die Publikation zur Verwendung für die vorliegende Arbeit genauer durchgearbeitet. Ausgehend von der Primärliteratur wurde dann mittels Schneeballprinzip eine iterative Literatursuche fortgesetzt, um relevante Sekundärliteratur zu finden.

Tabelle 3.2 zeigt die relevanten Treffer und die Treffer insgesamt, bei der Suche in Pubmed. Durch die iterative Suche wurden 11 weitere Artikel gefunden.

<u>Sucheingabe</u>	<u>Treffer</u>	<u>Relevant</u>
telemonitoring heart failure	212	11
heart failure eq 5d	60	5
heart failure kccq	84	2
herzinsuffizienz	18	1

Tabelle 3.2: Sucheingabe in Pubmed inkl. der relevanten Treffer und der Treffer insgesamt.

3.2 Statistische Methoden

3.2.1 Deskriptive Statistik

Die deskriptive Statistik dient der Aufbereitung von Datenmaterial. Dadurch werden die Daten übersichtlicher und es können Voraussetzungen für weitere statistische Verfahren ermittelt werden.

Eine wichtige Merkmalsausprägung ist dabei das Skalenniveau. Dabei wird grob zwischen Nominalskala, Ordinalskala und metrischer Skala unterschieden. Die Nominalskala besteht aus Merkmalen, die nicht in eine Rangfolge gebracht werden können. Dies ist bei Merkmalen der Ordinalskala durchaus der Fall. Merkmale einer Nominalskala sind zum Beispiel Augenfarben oder Berufsgruppen. Die Schulnoten stellen dagegen ordinale Merkmale dar. Der Unterschied zur metrischen Skala besteht darin, dass die Abstände zwischen Merkmalen der Ordinalskala nicht interpretierbar sind. Bei Merkmalen der metrischen Skala lässt sich der Abstand interpretieren. So kann beim Vergleich des Körpergewichtes zweier Personen sehr wohl festgestellt werden, dass eine 70kg schwere Person um 10kg leichter ist als eine Person, die 80kg wiegt.

Weiters ist für die Wahl des statistischen Verfahrens auch entscheidend, ob die Stichproben verbunden oder unabhängig sind. Eine unabhängige Stichprobe ergibt sich zum Beispiel, wenn der Gesundheitszustand von zwei unterschiedlichen Personengruppen ermittelt wird. Wird der Gesundheitszustand einer Personengruppe zu zwei unterschiedlichen Zeitpunkten gemessen, dann handelt es sich um eine verbundene Stichprobe.

Einige statistische Tests setzen die Normalverteilung voraus. Diese kann mit unterschiedlichen Verfahren überprüft werden. Häufig verwendet werden dabei der Shapiro-Wilk-Test und der Kolmogorov-Smirnov-Test. Zur unterstützenden Analyse ist es sinnvoll, auch graphische Methoden heranzuziehen. So können auch Q-Q-Plots, Histogramme und Boxplots Auskünfte über die vorliegende Verteilung geben.

Abhängig von den Merkmalsausprägungen der Daten können diese dann mittels Mittelwert \pm Standardabweichung, Anzahl und Prozentsatz, Interquartilsbereich oder Median dargestellt werden.

Weiters können auch graphische Methoden zum Einsatz kommen. Diese machen es leichter, einen Überblick zu erhalten, als es bei reiner Betrachtung von Zahlen möglich ist. Für die vorliegende Arbeit wurden vor allem Boxplots, Liniendiagramme, Histogramme und Q-Q-Plots verwendet, um die Daten graphisch aufzubereiten.

Auf der Grundlage der deskriptiven Statistik werden dann die statistischen Tests für die weitere Analyse der Daten gewählt. [18]

3.2.2 Parametrische Tests

Parametrische Tests gehen davon aus, dass die vorliegende Stichprobe einer vorgegebenen Wahrscheinlichkeitsverteilung entspringt. Folgend werden zwei parametrische Verfahren beschrieben, die sich mit Mittelwertvergleichen beschäftigen.

Durch den Vergleich von Mittelwerten können Unterschiede in den einzelnen Fallgruppen

einer Stichprobe festgestellt werden. Es kann somit untersucht werden, ob Unterschiede auftreten oder nicht. Um von einer Stichprobe auf die Grundgesamtheit zu schließen, bedarf es weiterer Methoden. Signifikanztests können, mit einer gewissen Wahrscheinlichkeit, Aufschluss darüber geben, ob eine Stichprobe repräsentativ für eine Grundgesamtheit ist. Es gibt mehrere Signifikanztests, die für Mittelwertvergleiche geeignet sind. [18], [19] Beispiele für solche Signifikanztests sind:

- T-Test
- ANOVA

3.2.3 T-Test

Der T-Test dient dazu, um mit Hilfe von Mittelwerten Zusammenhänge zwischen einer Stichprobe und der Grundgesamtheit feststellen zu können. So kann überprüft werden, ob der Unterschied zweier Mittelwerte von Stichproben auch für die Grundgesamtheit eine Gültigkeit besitzt. Es gibt noch weitere Möglichkeiten den T-Test anzuwenden. Zum Beispiel dient er auch der Untersuchung von Beobachtungen innerhalb einer Stichprobe, um dabei festzustellen, ob ein vorgegebener Wert durch den Mittelwert eines Eigenschaftsmerkmals der Grundgesamtheit über- oder unterschritten wird. Für den Vergleich von zwei Mittelwerten wird zwischen unabhängiger und verbundener Stichprobe unterschieden. Verbundene Stichproben liegen vor, wenn die Werte für den Vergleich von denselben Objekten stammen. Im Gegensatz dazu steht eine unabhängige Stichprobe für den Vergleich von Werten, die von unterschiedlichen Objekten stammen.

Wie bei der Anwendung vieler Tests, müssen die Daten auch für die Anwendung des T-Tests gewisse Anforderungen erfüllen. Die Stichproben müssen ein Teil der Grundgesamtheit sein und dabei zufällig ausgewählt werden. Weiters müssen die verwendeten Variablen mindestens intervallskaliert sein. Eine Anwendung des T-Tests bei ordinalskalierten Variablen ist somit nicht berechtigt. Eine weitere Voraussetzung ist, dass die Variablen normalverteilt sein müssen. Der T-Test ist allerdings sehr robust im Bezug auf diese Annahme. Das bedeutet, dass auch bei keiner perfekten Normalverteilung brauchbare Ergebnisse mit Hilfe des T-Tests gewonnen werden können. Auch der Standardfehler soll Beachtung finden, da dieser Auskunft über die Schätzung liefert.

Für die Auswertung bei unabhängigen Stichproben muss beachtet werden, ob diese eine gleiche oder ungleiche Varianz besitzen. Der Levene-Test dient dazu, dies zu überprüfen. Im Zweifelsfall sollte die Annahme der ungleichen Varianz verwendet werden, um die Null-hypothese zu untersuchen. Der t-Wert für unabhängige Stichproben errechnet sich mit der Formel (3.1):

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}} \quad (3.1)$$

\bar{X} steht für die Stichprobenmittelwerte, S für empirische Varianz und N für die Anzahl der Objekte. Für die Berechnung bei gleichen Varianzen werden aus den empirischen Varianzen eine gepoolte Varianz berechnet und diese dann in die Formel (3.1) eingesetzt. Mit Hilfe der Irrtumswahrscheinlichkeit können die Hypothesen überprüft werden und die Ermittlung eines Konfidenzintervalls dient dazu, etwaige Mittelwertdifferenzen einzugrenzen. [19]

Der T-Test bei verbundener Stichprobe untersucht Mittelwerte, die nicht gänzlich unabhängig voneinander sind. Solche Mittelwerte erhält man z.B. aus dem Vergleich von zwei unterschiedlichen Variablen, die von derselben Objektgruppe stammen. Der T-Test bei verbundener Stichprobe findet ebenfalls Anwendung, wenn ein Eigenschaftsmerkmal einer Fallgruppe zu zwei unterschiedlichen Zeitpunkten betrachtet wird. Formel (3.2) dient zur Berechnung des t-Werts für den Vergleich von Mittelwerten bei verbundener Stichprobe.

$$t = \frac{\bar{D}}{\sqrt{\frac{S_D^2}{N}}} \quad (3.2)$$

\bar{D} steht für die Differenz der beiden Mittelwerte, S_D bezeichnet die Standardabweichung der Differenzen und N ist die Anzahl der Objekte. [19]

3.2.4 ANOVA

ANOVA steht für Analysis of Variance und ist eine Methode, um Mittelwertvergleiche durchzuführen. Im Gegensatz zum T-Test, der dem Vergleich von zwei Mittelwerten dient, können mit Hilfe der ANOVA (Varianzanalyse) mehrere Mittelwerte verglichen werden. Auch für die Durchführung der Varianzanalyse müssen die Daten bestimmte Anforderungen erfüllen. Da die Varianzanalyse wie der T-Test zum Mittelwertvergleich eingesetzt werden, ähneln sich die Anforderungen der Daten für beide Tests. So müssen auch die Variablen für die Varianzanalyse intervallskaliert sein. Eine weitere Voraussetzung ist, dass die Daten aus der Zufallsstichprobe von einer Grundgesamtheit stammen, die eine Normalverteilung aufweist. Für die Varianzanalyse liegt auch die Annahme vor, dass Varianzgleichheit zwischen den Objektgruppen besteht.

Zur Durchführung einer Varianzanalyse benötigt man einerseits die Streuung der betrachteten Variablen innerhalb einer Objektgruppe und andererseits die Streuung zwischen den Fallgruppen. Diese beiden Werte werden benötigt, um zu untersuchen, ob sich Mittelwertunterschiede zufällig ergeben haben oder nicht. Ist z.B. die Streuung innerhalb einer Gruppe klein und die Streuung zwischen den Gruppen groß, dann sind Mittelwertunterschiede wahrscheinlich nicht zufällig entstanden.

Die Formel (3.3) zeigt die Berechnung der Quadratsummen innerhalb der Objektgruppen.

$$QSI = \sum_{i=1}^k (N_i - 1) * S_i^2 \quad (3.3)$$

Die Variable k steht für die Anzahl der Objektgruppen, N_i bezeichnet die Fälle und S_i^2 die Varianz der Werte innerhalb der i-ten Objektgruppe.

Um die Quadratsummen zwischen den Gruppen zu berechnen, wird die Formel (3.4) verwendet.

$$QSZ = \sum_{i=1}^k N_i * (\bar{X}_i - \bar{X}) \quad (3.4)$$

Die Variablen k und N_i sind analog zu Formel (3.3). \bar{X}_i bezeichnet den Mittelwert einer Objektgruppe und \bar{X} den Mittelwert der gesamten Stichprobe.

Um mittels Varianzanalyse Rückschlüsse auf die Grundgesamtheit zu erhalten, werden die Werte aus Formel (3.3) und (3.4) miteinander verglichen. Dabei wird unter Berücksichtigung der Freiheitsgrade der Quotient von QSI und QSZ gebildet. Die Freiheitsgrade lauten k-1 für QSZ und N-k für QSI, N und k stehen analog zu Formel (3.3) für die Anzahl der Fälle und Objektgruppen.

Auf diese Art und Weise lässt sich der F-Wert berechnen, in Formel (3.5) ist die Berechnung dargestellt.

$$F = \frac{\frac{QSZ}{k-1}}{\frac{QSI}{N-k}} \quad (3.5)$$

Die Verteilung der Maßzahl F ist bekannt. Dieser Umstand wird ausgenutzt, um die Signifikanz zu ermitteln, welche angibt, wie wahrscheinlich der errechnete F-Wert angenommen wird, wenn in der Grundgesamtheit identische Mittelwerte der einzelnen Gruppen vorherrschen.

Wenn das Ergebnis ein Zurückweisen der Nullhypothese (gleich hohe Mittelwerte zwischen den Gruppen in der Grundgesamtheit) zulässt, müssen aber dennoch nicht alle Mittelwerte unterschiedlich sein. Um zu untersuchen, welche Mittelwerte voneinander abweichen, müssen weitere Tests durchgeführt werden. Eine Möglichkeit wäre, die Stichproben mit einem T-Test paarweise miteinander zu vergleichen. Dies würde jedoch bei jeder Anwendung des T-Tests eine weitere Irrtumswahrscheinlichkeit bedeuten. Diese Gefahr lässt sich verringern, wenn für die Post-Hoc-Analyse entsprechende Verfahren angewendet werden.

Ein dafür häufig verwendeter Test ist der Scheffe's Test. Dieser hat als Basis die F-Verteilung und lässt auf dieser Grundlage multiple Vergleichstests zu. [19]

3.2.5 Nichtparametrische Tests

Wie bei allen Tests, die auf stichprobentheoretischen Grundlagen basieren, versuchen auch nichtparametrische Tests Rückschlüsse von Stichprobenbeobachtungen auf die Grundge-

samtheit zuzulassen. Der große Unterschied zwischen parametrischen und nichtparametrischen Tests liegt darin, dass für nichtparametrische Tests die Voraussetzungen der zu untersuchenden Daten geringer sind als für parametrische Tests. Dies gilt sowohl für die Verteilung in der Grundgesamtheit als auch für die Anforderungen an das Skalenniveau. Somit werden nichtparametrische Tests vor allem dann eingesetzt, wenn die Anforderungen von parametrischen Tests nicht erfüllt sind. Da die Voraussetzungen nicht so streng sind, können auch nur weniger starke Hypothesen überprüft werden, wodurch die Aussagekraft der Ergebnisse vermindert wird.

Nachstehend werden ein paar ausgewählte Tests erläutert. [19]

3.2.6 Chi-Quadrat-Test

Der Chi-Quadrat-Test ist ein Signifikanztest. Er dient dazu festzustellen, ob aus einer Stichprobenbeobachtung auf einen Zusammenhang in der Grundgesamtheit geschlossen werden kann. So kann festgestellt werden, ob kategoriale Variablen abhängig oder unabhängig von einander sind. Grundlage für den Chi-Quadrat-Test sind die erwarteten Häufigkeiten, die mit Hilfe einer Kreuztabelle gebildet werden. Im Gegensatz zu einer Häufigkeitstabelle kann mit einer Kreuztabelle die gemeinsame Häufigkeitsverteilung für zwei Variablen abgebildet werden. Dabei wird neben der beobachteten auch die erwartete Häufigkeitsverteilung dargestellt.

Die Formel (3.6) dient zur Berechnung des Chi-Quadrat-Wertes.

$$\chi^2 = \sum_{i=1}^n \sum_{j=1}^m \frac{\sqrt{n_{ij} - \hat{a}_{ij}}}{\hat{a}_{ij}} \quad (3.6)$$

Dabei steht n für die Anzahl der Zeilen und m für die Anzahl der Spalten. n_{ij} steht für die beobachtete Häufigkeit und \hat{a}_{ij} für die erwartete Häufigkeit, wobei i die Zeile und j die Spalte bezeichnet.

Um mit Hilfe des Chi-Quadrat-Tests festzustellen, ob ein Zusammenhang zwischen den Variablen besteht oder nicht, müssen einige Voraussetzungen erfüllt sein. Eine Voraussetzung ist, dass die erwartete Häufigkeit in weniger als 20% der Felder der Kreuztabelle unter 5 liegt. Einschränkungen können sich auch bei geringer Feldzahl der Kreuztabelle ergeben. Für solche Fälle mit geringen erwarteten Häufigkeiten liefert der Fisher's-Exact-Test die zuverlässigsten Ergebnisse. [19]

3.2.7 Shapiro-Wilk-Test

Der Shapiro-Wilk-Test wurde zur Überprüfung eines Datensatzes auf Normalverteilung entwickelt. Ursprünglich wurde er für Stichproben von $n < 50$ angedacht. Durch Erweiterungen kann er jetzt für Stichproben von $n < 5000$ eingesetzt werden. Zur Überprüfung auf Normalverteilung vergleicht der Shapiro-Wilk-Test zwei Varianzen. Eine Varianz entspricht der tatsächlichen Varianz der Stichprobe. Die zweite Varianz wird geschätzt. Sie sieht so aus, wie eine Varianz aussehen würde, wenn die Stichprobe einer normalverteilten

Grundgesamtheit entspringe. Je ähnlicher sich diese beiden Varianzen sind, umso wahrscheinlicher liegt eine Normalverteilung vor. [19]

3.2.8 Kolmogorov-Smirnov-Test

Mit Hilfe des Kolmogorov-Smirnov-Tests lässt sich die zugrunde liegende Verteilung von Werten einer Variable prüfen. Dieser Test dient oft als Grundlage zur Ermittlung von Voraussetzungen für weiterführende Testverfahren. Es ist anzumerken, dass selbst dann, wenn durch den Kolmogorov-Smirnov-Test mit hoher Wahrscheinlichkeit eine Normalverteilung ausgeschlossen wird, eine geringe Irrtumswahrscheinlichkeit besteht. Dies kann bedeuten, dass die Werte in der Grundgesamtheit dennoch annähernd eine Normalverteilung aufweisen, was als Voraussetzung für einige Testverfahren genügt. Mittels Histogramm und Normalverteilungskurve kann das z.B. überprüft werden. [19]

3.2.9 Kruskal-Wallis-Test

Der Kruskal-Wallis-Test ist das nichtparametrische Pendant zur ANOVA. Mithilfe des Kruskal-Wallis-Tests ist der Vergleich von mehreren unabhängigen Stichproben, die nicht die Voraussetzungen für die Durchführung einer ANOVA erfüllen, möglich.

Der Kruskal-Wallis-Test bildet für den Vergleich Rangfolgen für alle Stichproben. Anschließend werden die mittleren Rangzahlen der Stichproben untereinander verglichen. Dann wird die Wahrscheinlichkeit ausgegeben, mit der eine Abweichung der mittleren Rangzahlen zwischen den Gruppen durch Zufall entstanden ist. So kann überprüft werden, ob die Stichproben der gleichen Grundgesamtheit entspringen oder nicht. [19]

3.2.10 Mann-Whitney-Test

Der Mann-Whitney-Test ist für zwei unabhängige Stichproben geeignet. Im Gegensatz zum T-Test sind für den Mann-Whitney-Test ordinalskalierte Variablen ausreichend.

Beim Mann-Whitney-Test werden die Werte zweier Stichproben zusammengefasst. Für die in aufsteigender Reihenfolge angeordneten Werte werden Ränge vergeben. Anschließend werden die Rangsummen gebildet und für beide Stichproben getrennt betrachtet. Ein Vergleich der Rangsummen ergibt einen ersten Aufschluss darüber, ob die beiden Gruppen derselben Grundgesamtheit entspringen oder nicht. Allerdings kann es dabei auch leicht zu Fehlinterpretationen kommen. Wenn z.B. die Ränge der Werte von Gruppe eins sowohl im hohen als auch im niedrigen Bereich angesiedelt sind und die Ränge der Werte von Gruppe zwei sich im mittleren Bereich befinden, dann würden die Rangsummen der beiden Stichproben ähnlich sein, obwohl sie einer unterschiedlichen Grundgesamtheit entspringen. Um diesem Problem entgegenzuwirken, werden die Häufigkeiten gezählt, wie oft ein Wert der ersten Stichprobe einem Wert der zweiten Stichprobe vorangeht und umgekehrt. Sowohl die geringere der beiden Häufigkeiten als auch die geringere Rangsumme werden herangezogen, um die Signifikanz zur Bewertung der Nullhypothese zu errechnen. [19]

3.2.11 Wilcoxon-Test

Der Wilcoxon-Test eignet sich zur Untersuchung zweier verbundener Stichproben. Dabei dient er der Feststellung, ob die beiden Stichproben einer gleich verteilten Grundgesamtheit entstammen.

Beim Wilcoxon-Test werden die Rangsummen der positiven und negativen Differenzen zwischen den Wertepaaren gebildet. Weiters dienen sowohl die Anzahl der positiven und negativen Differenzen als auch die mittleren Ränge als Grundlage für einen Signifikanztest. Mit Hilfe des so genannten Z-Wertes wird die Signifikanz zur Bewertung der Nullhypothese berechnet. [19]

3.3 Entwicklung der Auswertestrategie

Die Abbildung 3.1 zeigt einen Überblick über das Vorgehen bei der Entwicklung der Auswertestrategie.

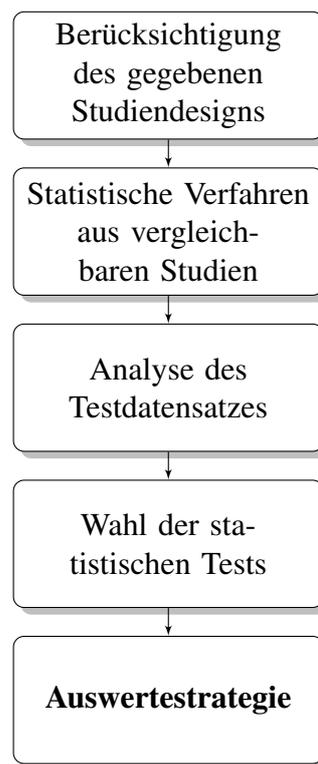


Abbildung 3.1: Vorgehen bei der Entwicklung der Auswertestrategie.

Unter Berücksichtigung des gegebenen Studiendesigns erfolgte zu Beginn eine Einarbeitung in die vorhandene Literatur. Dabei wurden statistische Verfahren aus vergleichbaren Studien festgehalten. Danach erfolgte eine Analyse des vorliegenden Testdatensatzes. Weiters wurde überlegt, wie sich eine Behandlungsmethode auf den Behandlungserfolg aus-

wirken kann und welche unterschiedlichen Unterteilungen des Probandenkollektivs sinnvoll erscheinen.

Auf dieser Grundlage wurde dann die Auswertestrategie entwickelt sowie die dazugehörigen statistischen Verfahren gewählt.

3.4 SPSS

Für statistische Auswertungen wurde die Software SPSS herangezogen. SPSS ist ein umfangreiches Programm, welches zur Datenanalyse in der Statistik eingesetzt wird. Aufgebaut ist das Programm modulartig, wobei das Basismodul recht umfangreich ist. SPSS erleichtert die Verwaltung und Analyse von Daten. Eine gute Möglichkeit Arbeitsabläufe zu speichern bietet die Syntax. Mithilfe der Syntax können Arbeitsabläufe automatisiert werden. [19]

In der vorliegenden Arbeit wurden die SPSS Versionen 18 (PASW Statistics 18) und 20 (IBM SPSS Statistics 20) verwendet.

4 Ergebnisse

4.1 Publikationen

Bei der Literaturrecherche wurden Publikationen gesucht, die Gemeinsamkeiten mit dem zu untersuchenden Studiendesign (siehe Kapitel 1.2) aufweisen. Die im Folgenden dargestellten Ergebnisse sollen einen kurzen Überblick über diese Studien und die dabei verwendeten statistischen Verfahren ermöglichen.

Spiraki et al. haben sich mit der Ermittlung der Lebensqualität von Herzinsuffizienz-Patienten und Patienten mit koronarer Herzerkrankung befasst. Verwendete Messinstrumente waren der SF-36 und der EQ-5D-Fragebogen.

Insgesamt standen für die Studie 153 Patienten zur Verfügung. Die erste Befragung wurde mit 49 Herzinsuffizienz-Patienten und 104 Patienten, die an koronarer Herzerkrankung leiden, durchgeführt. Auf Grund von Todesfällen und Verlegungen standen für die zweite Befragung von Koronaren-Herzerkrankungs-Patienten noch 98 Personen und für die dritte Befragung noch 94 Personen zur Verfügung. Für die zweite und dritte Befragung von Herzinsuffizienz-Patienten waren aus denselben Gründen noch 47 Personen verfügbar. Der Männeranteil bei den Koronaren-Herzerkrankungs-Patienten war 60% und der Anteil von Patienten über 65 Jahren betrug 76%. Bei den Herzinsuffizienz-Patienten betrug der Männeranteil 57% und der Anteil von Patienten über 65 Jahren war 86%.

Zur statistischen Auswertung wurden Deskriptive-, Korrelations- und Varianzanalysen sowie Mittelwertvergleiche durchgeführt. Um zu vergleichen, ob eine Korrelation zwischen den beiden Fragebögen (SF36 und EQ-5D) besteht, wurden Kontingenztafeln erstellt.

In der Abbildung 4.1 sind die Mittelwerte der EQ-5D und VAS Ergebnisse dargestellt. Aufgegliedert sind die Säulen einerseits nach Erkrankung (CHD-Koronare Herzerkrankung, HF-Herzinsuffizienz) und andererseits nach Befragungszeitpunkt (Phase A bis C). Bei der Aufnahme im Krankenhaus wurde die erste Befragung (Phase A) durchgeführt. Die zweite Befragung (Phase B) fand bei der Entlassung aus dem Krankenhaus statt und die Angaben zur dritten Befragung (Phase C) wurden einen Monat nach Entlassung aus dem Krankenhaus aufgenommen. Eine statistisch nicht signifikante Verbesserung der Lebensqualität durch die Behandlung ist erkennbar. Weiters ist ersichtlich, dass Herzinsuffizienz-Patienten eine im Durchschnitt 10% schlechtere Lebensqualität aufweisen als Koronare-Herzerkrankungs-Patienten. [20]

4 Ergebnisse

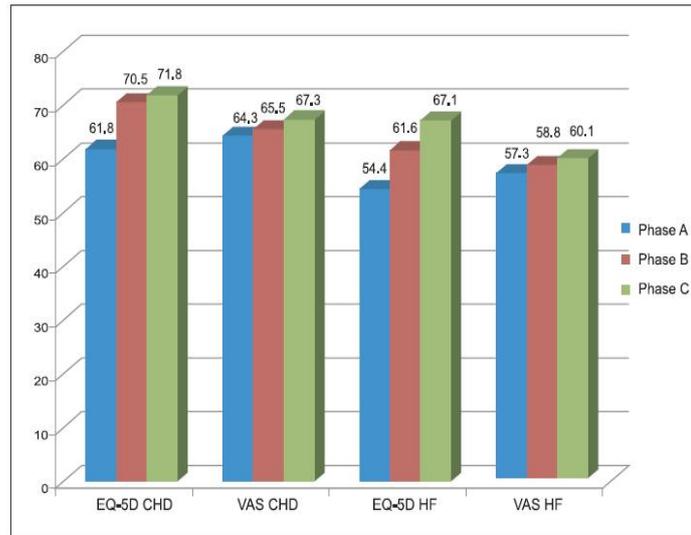


Abbildung 4.1: Darstellung der EQ-5D und VAS Mittelwerte bezogen auf Befragungszeitpunkt und Erkrankung. (CHD – coronary heart disease; EQ-5D – EuroQol 5-dimensional instrument; HF – heart failure; VAS – Visual Analogue Scale). Entnommen aus Spiraki et al. [20]

Auch Pinto et al. haben sich mit der Ermittlung der Lebensqualität beschäftigt. In ihrer Studie haben sie diesbezüglich die Lebensqualität von Herzinsuffizienz-Patienten, Schlaganfall-Patienten und Pflegern verglichen. Zur Ermittlung der Lebensqualität wurde der EQ-5D-Fragebogen herangezogen.

Um die Daten auswerten zu können, wurde eine Reihe von unterschiedlichen statistischen Verfahren angewendet. Die ANOVA wurde zum Vergleich von kontinuierlichen Variablen zwischen den Gruppen verwendet und Scheffè's Test für Post-Hoc-Vergleiche. Kategorische Variablen wurden mit Hilfe des Chi-Quadrat-Tests verglichen und für Vergleiche zwischen den Skalen wurde die Pearson-Korrelation durchgeführt. Als statistisch signifikant wurde ein p-Wert $<0,05$ angesehen.

Das Probandenkollektiv bestand aus 196 Personen. Dabei waren 67 Personen Schlaganfall-Patienten, 62 Herzinsuffizienz-Patienten und 67 Pfleger. Tabelle 4.1 zeigt die Mittelwerte der Ergebnisse des EQ-5D-Fragebogens. Der durchschnittliche Lebensqualitätswert liegt bei allen drei Gruppen unter 0,70, wobei die Gruppe der Schlaganfall-Patienten mit 0,52 den schlechtesten Wert hat.

4 Ergebnisse

Groups	EQ-5D, mean (SD)	P-value
Stroke (a)	0.52 (0.36)	
CHF (b)	0.69 (0.28)	a/b = 0.010 b/c = 0.812
Caregivers (c)	0.65 (0.24)	a/c = 0.049
ANOVA P-value (DF)	0.006 (190)	

Tabelle 4.1: Darstellung der EQ-5D Mittelwerte der einzelnen Gruppen. ((a) Schlaganfall, (b) Herzinsuffizienz, (c) Pflegende). Entnommen aus Pinto et al. [21]

Pinto et al. haben auch einen Vergleich zwischen den Gruppen für die 5 Dimensionen des EQ-5D-Fragebogens durchgeführt. Kein statistisch signifikanter Unterschied konnte zwischen Schlaganfall-, Herzinsuffizienz-Patienten und Pflegern in den Dimensionen „Schmerz“ und „Ängste/Niedergeschlagenheit“ festgestellt werden. 25% mehr Herzinsuffizienz-Patienten zeigten Beschwerden bezüglich „Mobilität“ und „Alltägliche Tätigkeiten“ verglichen mit den Pflegern, verglichen mit den Schlaganfall-Patienten waren es aber 30% weniger. Weiters zeigte sich, dass 40% mehr Schlaganfall-Patienten Probleme haben für sich selbst zu sorgen als Personen aus den anderen beiden Gruppen. [21]

Comin-Colet et al. haben die Auswirkung von intravenös verabreichten Eisenpräparaten auf die Lebensqualität von Herzinsuffizienz-Patienten mit Eisenmangel untersucht. Die dabei verwendeten Daten stammten von der FAIR-HF-Studie (Ferric carboxymaltose Assessment in patients with IRon deficiency and chronic Heart Failure).

Das Probandenkollektiv bestand aus 459 Herzinsuffizienz-Patienten. 304 Personen wurden mit Eisenpräparaten therapiert und 155 Personen wurde eine Placebo-Injektion verabreicht. Von der Eisenpräparat-Gruppe haben 26 Personen (8,6%) die Studie nicht beendet, 21 Personen sind ausgetreten und 5 Personen sind vorzeitig verstorben. Von der Placebo-Gruppe sind 4 Personen vor Beendigung der Studie gestorben und 16 Personen sind ausgetreten, was eine Gesamtzahl von 20 Personen (12,9%) ergibt. Die demografischen Patientendaten der beiden Gruppen stimmen sehr gut überein. So liegt z.B. die Altersverteilung und Anzahl der Frauen bei der Eisenpräparat-Gruppe bei 67.8 ± 10.3 bzw. 52.3% und bei der Placebo-Gruppe bei 67.4 ± 11.1 bzw. 54.8%.

Zur Ermittlung der Lebensqualität wurden der KCCQ und der EQ-5D-Fragebogen verwendet. Diese beiden Fragebögen wurden von den Herzinsuffizienz-Patienten einmal zu Beginn der Studie und je einmal nach 4, 12 und 24 Wochen ausgefüllt.

Für die statistischen Analysen der Daten wurde ein Signifikanzniveau von 0.05 angenommen. Um die kontinuierlichen Variablen zwischen der Eisenpräparat- und der Placebo-Gruppe vergleichen zu können, wurde eine Kovarianzanalyse mit Messwiederholung durchgeführt. Unterschiede in der Verteilung der beiden Gruppen wurden mittels Polytomous-Regression mit Messwiederholung untersucht.

Die Eisenpräparat- und Placebo-Gruppe weisen zu Beginn der Studie keine signifikant un-

terschiedlichen Lebensqualitätswerte auf. Bei der VAS hatten beide Gruppen einen Mittelwert von 54 ± 1 . Bei der Eisenpräparat-Gruppe betrug der Mittelwert für den Index-Wert ($\times 100$) 68 ± 1 und bei der Placebo-Gruppe 69 ± 1 , die „Klinische Zusammenfassung“ wurde mit 52 ± 1 bzw. 53 ± 1 bewertet. Die Lebensqualitätswerte der Eisenpräparat-Gruppe verbesserten sich im Laufe der Studie um durchschnittlich 28%, jene der Placebo-Gruppe um durchschnittlich 11%. Beim Index-Wert des EQ-5D-Fragebogens zeigte sich bei der Placebo-Gruppe von Woche 4 auf Woche 12 eine Verschlechterung um 4,6%. [22], [23]

Jolly et al. haben in ihrer Studie die Auswirkung von Heimtraining auf die Lebensqualität und den Gesundheitszustand von Herzinsuffizienz-Patienten untersucht. Insgesamt wurden 169 Patienten in die Studie inkludiert. 84 Personen umfasste die Trainings-Gruppe und 85 Personen die Kontroll-Gruppe. Die Patientendaten der beiden Gruppen waren sehr ähnlich. Leichte Unterschiede gab es auf das Alter bezogen. Das Durchschnittsalter der Trainings-Gruppe war um 4,1 Jahre geringer als jenes der Kontroll-Gruppe. Beide Gruppen setzten sich aus Patienten der NYHA-Klasse 1 bis 3 zusammen und der Prozentsatz an Männern lag jeweils bei über 70%. Jede Gruppe erhielt Unterstützung von Diplomkrankenschwestern, aber nur die Trainings-Gruppe ein Trainingsprogramm für zu Hause. Aufgrund von vorzeitiger Beendigung der Studie und Todesfällen waren nach 6 Monaten noch Daten von 161 Patienten verfügbar und nach 12 Monaten noch von 157 Patienten.

Um weitere Daten zu erhalten, mussten die Patienten im Laufe der Studie noch Fragebögen ausfüllen. So wurden z.B. der MLHFQ und der EQ-5D-Fragebogen angewendet. Diese Fragebögen wurden zu drei Zeitpunkten ausgefüllt. Zweimal wurden die Fragebögen persönlich am Krankenhausgelände aufgenommen, je einmal zu Beginn der Studie und nach 6 Monaten, und die dritte Abfrage der Fragebögen wurde nach 12 Monaten per Post erhoben.

Um statistische Unterschiede zwischen den Gruppen festzustellen, wurde die „Linear-Regression der kleinsten Quadrate“ verwendet. Die Cox-Regressions-Analyse wurde zum Vergleich der Zeit bis zum ersten klinischen Ereignis durchgeführt. Zur Beurteilung der statistischen Signifikanz wurde ein Signifikanzniveau von 0,05 angenommen.

Der EQ-5D-Index-Wert betrug zu Beginn der Studie bei der Trainingsgruppe 0,675 und bei der Kontrollgruppe 0,696. Im Laufe der Studie konnte kein signifikanter Unterschied zwischen Trainings- und Kontrollgruppe festgestellt werden. Die Lebensqualitätswerte unterlagen dabei auch keiner signifikanten Änderung. [24]

Das Paper von Scherr et al. befasst sich mit der MOBITEL-Studie (Mobile Telemonitoring in Heart Failure Patients). Ziel dieser Studie war es, die Auswirkung von Telemonitoring auf den Behandlungserfolg von Herzinsuffizienz-Patienten zu untersuchen.

Für die Studie standen insgesamt 120 Probanden zur Verfügung, 66 in der Telemonitoring-Gruppe und 54 in der Kontroll-Gruppe. Von den 66 Probanden in der Telemonitoring-Gruppe war es 12 Probanden nicht möglich, die Geräte zur telemedizinischen Überwachung zu bedienen. Daher wurden die weiteren Auswertungen mit den verbleibenden 54 Personen der Telemonitoring-Gruppe durchgeführt. Die Telemonitoring-Gruppe setzte sich aus

40 männlichen und 14 weiblichen Probanden zusammen, die ein Durchschnittsalter von 65 Jahren hatten. Die Kontroll-Gruppe bestand aus 39 männlichen und 15 weiblichen Personen mit einem Durchschnittsalter von 67 Jahren. Beide Gruppen umfassten Patienten der NYHA-Klasse 2 bis 4. Es waren keine signifikanten Unterschiede zwischen den Ausgangsdaten der beiden Gruppen gegeben.

Zur Untersuchung der kardiovaskulären Mortalität und der erneuten Einweisung in ein Krankenhaus wurden der Logrank-Test und der Kaplan-Meier-Schätzer verwendet. Um normalverteilte Daten auswerten zu können, wurde der T-Test angewandt. Zum Vergleich von nominal verteilten Daten wurde der Chi-Quadrat-Test durchgeführt. Bei nicht normalverteilten Werten wurde der Mann-Whitney-U-Test zum Vergleich von unabhängigen Stichproben verwendet, und der Wilcoxon-Vorzeichenrang-Test für abhängige Stichproben. Als statistisch signifikant wurde ein p-Wert von <0.05 angesehen.

Die Ergebnisse zeigen, dass es eine Verbesserung der durchschnittlichen NYHA-Klasse von 3 auf 2 bei den Telemonitoring-Patienten gab. Dies stellte einen signifikanten Unterschied (p-Wert $< 0,001$) im Vergleich zur Kontroll-Gruppe und zu den Ausgangswerten der Telemonitoring-Gruppe dar. Eine nicht signifikante Besserung der Auswurffraktion des Herzens war in beiden Gruppen ersichtlich. Insgesamt gab es bei der Telemonitoring-Gruppe 8 erneute Krankenhauseinweisungen und bei der Kontroll-Gruppe 17, dies ist in Abbildung 4.2 als Kaplan Meier Diagramm dargestellt.

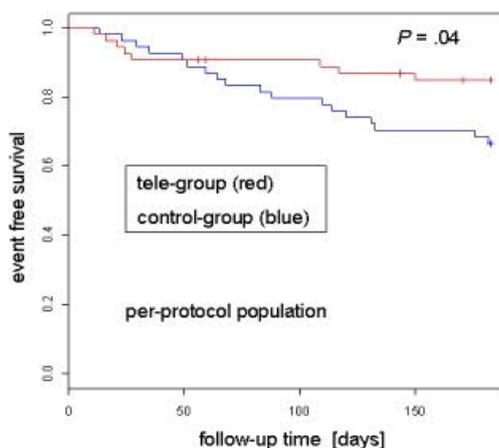


Abbildung 4.2: Kaplan Meier Diagramm von Krankenhauseinweisungen und Todesfällen. Entnommen aus Scherr et al. [25]

Es zeigte sich auch ein statistisch signifikanter Unterschied in der Krankenhausaufenthaltsdauer (p-Wert = 0.04) beim Vergleich der beiden Gruppen. Die Telemonitoring-Gruppe war durchschnittlich 3,5 Tage kürzer im Krankenhaus, als die Kontroll-Gruppe. [25]

Koehler et al. haben sich in ihrer Studie mit dem Einfluss von Telemonitoring auf die Mortalität von ambulant behandelten Herzinsuffizienz-Patienten beschäftigt. Insgesamt nah-

men 710 Herzinsuffizienz-Patienten an der Studie teil. 356 Patienten erfuhren eine normale Behandlung und 354 Patienten wurden zusätzlich noch mit Telemonitoring überwacht. Beide Gruppen weisen ähnliche Ausgangswerte auf. Das Probandenkollektiv bestand aus Patienten der NYHA-Klasse 2 und 3. Der Männeranteil war sowohl in der Telemonitoring- als auch in der Normalpflege-Gruppe über 80% und das Durchschnittsalter betrug in beiden Gruppen 66,9 Jahre.

Die Patienten verblieben dabei mindestens 12 Monate in der Studie. Zusätzlich zu den Krankenhausaufenthalten/-einweisungen und Todesfällen wurde auch die Lebensqualität mittels Short Form 36 festgehalten.

Kaplan-Meier-Diagramme wurden verwendet um die Zeit bis zu einem Vorfall gegenüber zu stellen. Diese unterschiedlichen Zeitpunktkurven der zwei Gruppen wurden anschließend mit einem Log-Rank-Test verglichen. Das Cox-Modell mit proportionalem Ausfallrisiko wurde verwendet um einen diesbezüglichen Einfluss der telemedizinischen Überwachung relativ zur normal behandelten Gruppe zu untersuchen. Korrelationsanalysen und Cox-Regression wurden durchgeführt, um den Einfluss der Übertragungshäufigkeit von telemedizinischen Daten auf die Quantität der Ergebnisse bzw. auf die Mortalität feststellen zu können. Der T-Test wurde zum Vergleich der beiden Gruppen angewandt. NYHA-Klassen Vergleiche wurden mittels Mann-Whitney-U-Tests gemacht und Lebensqualitätswertuntersuchungen wurden mit der Varianzanalyse (ANOVA) durchgeführt. Ein Signifikanzniveau von 0,05 wurde für die Analysen verwendet.

Die Anzahl der Todesfälle betrug in der Normalpflege-Gruppe 55 und in der Telemonitoring-Gruppe 54 Personen. Die Telemonitoring-Gruppe zeigte bessere physische Werte im SF-36 Test als die Normalpflege-Gruppe. Der Mittelwert betrug für die Telemonitoring-Gruppe $54,3 \pm 1,2$ und für die Normalpflege-Gruppe $49,9 \pm 1,2$ nach 12 Monaten und nach 24 Monaten $53,8 \pm 1,4$ bzw. $51,7 \pm 1,4$. Ein Vergleich der Aufenthaltstage im Krankenhaus aufgrund von Herzinsuffizienz ergibt einen Durchschnitt von $5,3 \pm 18,1$ für die Telemonitoring-Gruppe und $4,9 \pm 13,2$ für die Normalpflege-Gruppe. [26], [27]

Myers et al. haben in ihrer Studie die Effektivität von telemonitorischer Heimüberwachung bei Herzinsuffizienz-Patienten begutachtet. Dabei haben sie die Telemonitoring-Gruppe mit einer Normalpflege-Gruppe verglichen. Sowohl die Telemonitoring-Gruppe als auch die Normalpflege-Gruppe umfassten 83 Personen. Die Normalpflege-Gruppe war mit einem Durchschnittsalter von 82,5 Jahren etwas älter als die Telemonitoring-Gruppe mit durchschnittlich 79 Jahren. In beiden Gruppen lag der Frauenanteil über 50%. Eine Lebensqualitätsbefragung mittels SF-36 Fragebogen wurde bei der Telemonitoring-Gruppe sowohl zu Beginn als auch am Ende des Beobachtungszeitraums von 2 Monaten durchgeführt.

Um einen Vergleich von normalverteilten kontinuierlichen Variablen zwischen den beiden Gruppen durchzuführen, wurde der T-Test angewendet und für kategoriale Variablen der Chi-Quadrat-Test. Das Linear-Regressions-Modell wurde verwendet um die Ausgangsvariablen auf Ausgewogenheit zu überprüfen. Zur Feststellung von Veränderungen bezüglich der Lebensqualität wurde der T-Test bei verbundener Stichprobe durchgeführt.

Ebenso wurde der Wilcoxon-Vorzeichenrang-Test angewendet.

Die Ergebnisse zeigen 0,02 Krankenhausaufenthalte pro Patient und pro Jahr in der Telemonitoring-Gruppe und 0,03 Krankenhausaufenthalte in der Normalpflege-Gruppe. Bezüglich der Heimbesuche standen 5,8 pro Monat in der Telemonitoring-Gruppe 8,2 in der Normalpflege-Gruppe gegenüber. Bei den Lebensqualitätsuntersuchungen der Telemonitoring-Gruppe konnten Verbesserungen in allen Bereichen festgestellt werden. Die deutlichsten Verbesserungen zeigten sich bezüglich der körperlichen Schmerzen und der Vitalität mit einer p -Wert von $< 0,005$. Einen genaueren Überblick über die Daten des Short Form 36 liefert Tabelle 4.2. [28]

	<i>Entry</i>	<i>Exit</i>	<i>p Value^a</i>
Physical functioning	25.8	31.6	.07
Role physical	27.4	35.2	.05
Bodily pain	39.5	29.3	.00
General health	46.1	49.8	.18
Vitality	39.9	48.8	.00
Social function	44.8	57.3	.01
Role emotional	57.5	65.1	.21
Mental health (MCS)	68.7	75.8	.01

a. p values are from paired t tests.

Tabelle 4.2: Vergleich der SF-36 Werte der Telemedizin-Gruppe zu Beginn und am Ende der Studie. Entnommen aus Myers et al. [28]

Chaudhry et al. haben sich in ihrer Studie auch mit dem Behandlungserfolg von Telemonitoring bei Herzinsuffizienz-Patienten beschäftigt. Insgesamt wurden 5069 Patienten für die Studie untersucht. Nach dem Ausscheiden von 2442 ungeeigneten Patienten und dem Ablehnen einer Teilnahme an der Studie von 974 Patienten wurden die verbliebenen 1653 Patienten in die Studie aufgenommen. 827 Patienten wurden der Normalpflege-Gruppe zugewiesen und 826 Patienten erhielten zusätzlich eine telemonitorische Überwachung. Die demographischen Daten der beiden Gruppen waren sehr ähnlich. In beiden Gruppen betrug der Altersmedian 61 Jahre und der Frauenanteil knapp über 40%. Untersucht wurden Patienten aller NYHA-Klassen. Die Studiendauer betrug 6 Monate. Als Endpunkte wurden unter anderem Wiederaufnahme im Krankenhaus und Tod festgelegt.

Zum Testen der primären Hypothese und zum Untersuchen von dichotomen Variablen wurde der Chi-Quadrat-Test durchgeführt. Für Ereigniszeitanalysen wurde die Kaplan-Meier-Funktion berechnet. Weiters wurde die Cox-Regression zur Analyse der Überlebenszeit angewendet. Um Abweichungen in den Ausgangscharakteristiken zu entdecken, wurde ein logistisches Regressions-Modell angewendet. Dieses Modell nimmt das primäre Ergebnis als abhängige Variable und die unterschiedlichen Ausgangswerte der beiden Gruppen als

unabhängige Variable.

Der Endpunkt (Wiederaufnahme/Tod) wurde von 432 Herzinsuffizienz-Patienten der Telemonitoring-Gruppe und von 426 Herzinsuffizienz-Patienten der Normalpflege-Gruppe erreicht. Eine Wiederaufnahme, aufgrund der Herzinsuffizienz, in ein Krankenhaus war für 227 Personen aus der Telemonitoring-Gruppe und für 223 Personen aus der Normalpflege-Gruppe notwendig. [29], [30]

Seto et al. haben sich mit einem Telemonitoring-System für Herzinsuffizienz-Patienten beschäftigt, welches Mobiltelefone als Basis hat. Ziel war es, den Effekt des Systems auf das Herzinsuffizienz-Management und den Behandlungserfolg zu untersuchen. Insgesamt wurden 100 Patienten in die Studie inkludiert, je 50 in der Telemonitoring- und der Normalpflege-Gruppe. Das Durchschnittsalter war $55,1 \pm 13,7$ Jahre und $52,3 \pm 13,7$ Jahre bei der Telemonitoring- bzw. bei der Normalpflege-Gruppe. Der Frauenanteil lag in beiden Gruppen unter 25%.

Die Dauer der Studie betrug 6 Monate. Als Outcomes wurden neben Wiederaufnahmen im Krankenhaus, Mortalität und Anzahl der Nächte im Krankenhaus auch die Lebensqualitätswerte ermittelt.

Auf Normalverteilung wurden die Daten mit Hilfe von Kolmogorov-Smirnov- und Shapiro-Wilk-Test geprüft. Analysen zwischen den Gruppen wurden mit T-Test bei unabhängigen Stichproben und Mann-Whitney-U-Test durchgeführt. Um Ausgangsdaten mit Daten während der Studie zu vergleichen, wurde der T-Test bei verbundener Stichprobe bzw. der Wilcoxon-Vorzeichenrang-Test angewendet. Der jeweils angewendete Test wurde in Abhängigkeit der vorliegenden Verteilung gewählt.

Bezüglich der Wiederaufnahmen im Krankenhaus, Mortalität und Anzahl der Nächte im Krankenhaus konnten keine signifikanten Unterschiede zwischen der Telemonitoring- und der Normalpflege-Gruppe festgestellt werden. Die NYHA-Klasse und die LVEF zeigten bei beiden Gruppen eine signifikante Verbesserung. Bei der Lebensqualität, ermittelt mit dem MLHFQ, zeigte lediglich die Telemonitoring-Gruppe eine signifikante Verbesserung, und das sowohl bei der physischen als auch bei der emotionalen Lebensqualität. [31]

Boyne et al. haben in ihrer Studie untersucht, ob Telemonitoring den Behandlungserfolg bei Herzinsuffizienz-Patienten verbessert. Die Studie umfasste 382 Patienten, 197 in der Telemonitoring-Gruppe und 185 in der Normalpflege-Gruppe. Mit Ausnahme eines häufigeren Vorhofflimmerns in der Telemonitoring-Gruppe konnten bei den Ausgangsdaten keine signifikanten Unterschiede zwischen den zwei Gruppen festgestellt werden. Das Durchschnittsalter lag bei beiden Gruppen über 70 Jahren und der Männeranteil war durchschnittlich 18% höher als jener der Frauen. Das Probandenkollektiv umfasste Personen von NYHA-Klasse 2 bis 4.

Die Zeit bis zur ersten Aufnahme im Krankenhaus wurde als primäres Ergebnis festgelegt. Analysiert wurde dies mittels der Kaplan-Meier-Funktion. Der Shapiro-Wilk-Test wurde durchgeführt um Daten bezüglich ihrer Verteilung zu testen. Bei normalverteilten Daten wurde der Mittelwert und die Standardabweichung angegeben, anderenfalls der Score-

Range. Je nach Verteilung wurde zur Überprüfung eines Unterschiedes in den Ausgangsdaten entweder der T-Test oder der Mann-Whitney-U-Test verwendet. Unterschiede zwischen kategorischen Ausgangsdaten wurden mit Hilfe des Chi-Quadrat-Tests analysiert. Die Mortalität wurde mit der Cox-Regression überprüft. Weiters wurde der kombinierte Endpunkt, aus Einweisungen aufgrund von Herzinsuffizienz und Tod durch jegliche Ursachen, mit der Kaplan-Meier-Funktion untersucht. Alle weiteren sekundären Ergebnisse, wie zum Beispiel die Wiedereinweisung ins Krankenhaus oder die Krankenhausaufenthaltsdauer, wurden mit dem Mann-Whitney-U-Test analysiert.

Signifikante Unterschiede zwischen den Gruppen konnten nur im Bezug auf die Pflegebesuche festgestellt werden. Dabei zeigt sich, dass die Telemonitoring-Gruppe 54 Pflegebesuche weniger benötigte als die Normalpflege-Gruppe. Die mittlere Zeit bis zur ersten Krankenseinweisung betrug 163 Tage bei der Telemonitoring-Gruppe und 139 Tage bei der Normalpflege-Gruppe. Dabei wurden 18 Personen und 25 Personen bei der Telemonitoring- bzw. bei der Normalpflege-Gruppe eingewiesen. Dieser Unterschied war jedoch statistisch nicht signifikant. [32]

Martin-Lesende et al. haben die Auswirkungen von Telemonitoring auf die Behandlung von Patienten mit Herzinsuffizienz und/oder chronischer Lungenerkrankung untersucht. Dazu wurde die Telemonitoring-Gruppe mit einer Kontroll-Gruppe verglichen. Der Beobachtungszeitraum für die Patienten betrug ein Jahr. Zu Beginn der Studie wurden 28 Patienten der Telemonitoring-Gruppe und 30 Patienten der Normalpflege-Gruppe zugewiesen, wobei nur 21 bzw. 22 Patienten die 12-monatige Beobachtungszeit absolvierten. Das Durchschnittsalter lag bei beiden Gruppen knapp über 80 Jahren.

Quantitative Variablen wurden mittels T-Test oder Mann-Whitney-U-Test analysiert. Für Vergleiche zwischen qualitativen Variablen wurden der Chi-Quadrat-Test oder der Fisher's-Exact-Test verwendet. Weiters wurden auch ein logistisches Regressionsmodell und ein generelles lineares Modell für dichotom-kategorische bzw. quantitative Variablen angewandt. Zur Analyse weiterer Daten wurden auch Multilevel-Analysen durchgeführt.

Ein statistisch signifikanter Unterschied konnte nur bei den Krankenseinweisungen insgesamt festgestellt werden (p -Wert = 0,033). Dabei wurden 12 Patienten der Telemonitoring-Gruppe und 19 Patienten der Normalpflege-Gruppe mindestens einmal im Laufe der Studie in ein Krankenhaus eingewiesen. [33], [34]

Dendale et al. haben in ihrer Studie überprüft, ob Telemonitoring bei Herzinsuffizienz-Patienten die Mortalität und die Anzahl der erneuten Krankenseinweisungen reduzieren kann oder nicht. Dazu haben sie 160 Herzinsuffizienz-Patienten über den Zeitraum von 6 Monaten beobachtet. Die Patienten wurden dabei gleichermaßen einer Telemonitoring- bzw. Normalpflege-Gruppe zugewiesen. Bezüglich der demographischen Ausgangsdaten der beiden Gruppen waren keine signifikanten Unterschiede gegeben. Beide Gruppen besitzen einen Männeranteil von über 60% und weisen ein Durchschnittsalter von ca. 75 Jahren auf.

Um Vergleiche zwischen den Gruppen durchzuführen, wurden für kategorische Varia-

blen der Chi-Quadrat-Test und für kontinuierliche Variablen die ANOVA herangezogen. Die Kaplan-Meier-Methode wurde für Überlebenszeitanalysen verwendet. Weiters wurden Hazard-Ratios berechnet und ein Signifikanzniveau von 0,05 angenommen.

Bezüglich der Mortalität gibt es einen statistisch signifikanten Unterschied zwischen den Gruppen. In der Telemonitoring-Gruppe starben 4 Patienten, in der Normalpflege-Gruppe 14 Patienten. Ein weiterer signifikanter Unterschied besteht in den verlorenen Tagen aufgrund von Tod, Dialyse und Krankenhausaufenthalten wegen Herz- oder Nieren-Erkrankung. Diesbezüglich weist die Telemonitoring-Gruppe einen Durchschnitt von 13,1 Tagen pro Patient auf und die Normalpflege-Gruppe einen Durchschnitt von 30,2 Tagen. [35]

Dar et al. haben in ihrer Studie auch den Einfluss von Telemonitoring auf die Behandlung von Herzinsuffizienz-Patienten untersucht. 182 Patienten wurden in die Studie inkludiert, je 91 wurden sowohl der Telemonitoring- als auch der Normalpflege-Gruppe zugewiesen. Diese Patienten wurden in einem Zeitraum von 6 Monaten beobachtet. Die demographischen Daten der beiden Gruppen zu Beginn der Studie weisen keine signifikanten Unterschiede auf. Das Durchschnittsalter betrug für beide Gruppen ca. 70 Jahre.

Die Poisson-Regression wurde zum Vergleich der Tage, die lebend und nicht im Krankenhaus verbracht wurden, herangezogen. Die Zeit bis zum Auftreten eines Ereignisses wurde mittels Kaplan-Meier-Diagramm dargestellt und mittels Cox-Regressionsanalyse verglichen.

Die Ergebnisse zeigten eine signifikante Reduktion der Notfalleinweisungen aufgrund von Herzinsuffizienz bei der Telemonitoring-Gruppe. Abgesehen davon zeigten sich keine signifikanten Unterschiede zwischen Telemonitoring- und Normalpflege-Gruppe. Auch bei den ermittelten Lebensqualitätswerten mittels EQ-5D-Fragebogen und MLHFQ konnten keine signifikanten Unterschiede zwischen den beiden Gruppen festgestellt werden. [36]

4.2 Auswertestrategie

Einen kurzen Überblick über das Vorgehen bei der Auswertung des Testdatensatzes liefert Abbildung 4.3. Sie zeigt den groben Ablauf der wichtigsten Punkte bei der Analyse der Daten. Diese Punkte werden im Folgenden genauer erläutert.

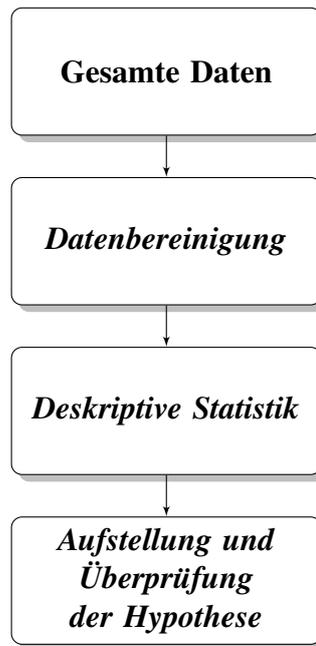


Abbildung 4.3: Vorgehen bei der Auswertung des Testdatensatzes.

4.2.1 Rohdaten

Am Beginn stand die Sichtung der Rohdaten. Sowohl beim KCCQ als auch beim EQ-5D-Fragebogen waren nur die Ergebnisse der einzelnen Fragen gegeben. Für weitere Analysen wurden aus den Ergebnissen der einzelnen Fragen die Index-Werte bzw. Subskalen berechnet.

Einen detaillierten Einblick in das genaue Vorgehen bei der Berechnung der Index-Werte und Subskalen liefert die Masterarbeit von Frau Oblinger [17].

4.2.2 Datenbereinigung

Die Rohdaten enthalten auch Widersprüche bei der Beantwortung vergleichbarer Fragen des KCCQ und des EQ-5D-Fragebogens. Tabelle 4.3 zeigt eine Gegenüberstellung der Dimensionen des EQ-5D-Fragebogens mit vergleichbaren Fragen des KCCQ.

Diese Fragen des KCCQ und des EQ-5D-Fragebogens aus Tabelle 4.3 wurden teilweise unterschiedlich bewertet. So wurde zum Beispiel beim EQ-5D-Fragebogen angegeben,

4 Ergebnisse

EQ-5D-5L Dimensionen	KCCQ Fragen
Beweglichkeit / Mobilität	Beeinträchtigung durch Atemnot oder Ermüdung bei der Ausführung folgender Tätigkeit: 1.3 Ca. 100 bis 200m auf ebener Strecke gehen
Für sich selbst sorgen	Beeinträchtigung durch Atemnot oder Ermüdung bei der Ausführung folgender Tätigkeit: 1.1 Sich selbst ankleiden 1.2 Duschen / Baden
Alltägliche Tätigkeiten	Beeinträchtigung durch Atemnot oder Ermüdung bei der Ausführung folgender Tätigkeit: 1.4 Garten- oder Hausarbeit, Einkaufstasche tragen
Schmerzen / körperliche Beschwerden	4. Beschwerlichkeit durch Schwellung der Füße, Knöchel oder Beine während der letzten 2 Wochen 6. Beschwerlichkeit durch Ermüdung während der letzten 2 Wochen 8. Beschwerlichkeit durch Atemnot während der letzten 2 Wochen
Angst / Niedergeschlagenheit	14. Häufigkeit der Depressivität oder Entmutigung durch die Herzinsuffizienz während der letzten 2 Wochen

Tabelle 4.3: Gegenüberstellung der Dimensionen des EQ-5D-Fragebogens mit vergleichbaren Fragen des KCCQ.

dass der Patient keine Schmerzen bzw. körperliche Beschwerden hat, dagegen behauptete derselbe Patient beim KCCQ, dass die Schwellungen der Füße, Knöchel oder Beine während der letzten 2 Wochen extrem beschwerlich waren. Diese widersprüchlichen Daten wurden bei der weiteren Analyse ausgeschlossen. Noch zu erwähnen ist, dass der Wert 1 beim EQ-5D-Fragebogen den besten und beim KCCQ den schlechtesten Gesundheitszustand darstellt. Die folgenden Beantwortungsmöglichkeiten stellen Widersprüche dar, welche für die weitere Datenauswertung berücksichtigt werden.

- (I) Beweglichkeit / Mobilität = 1 und KCCQ 1.3 = 1 od. 2
 Beweglichkeit / Mobilität = 2 und KCCQ 1.3 = 1 od. 2
 Beweglichkeit / Mobilität = 4 und KCCQ 1.3 = 4 od. 5
 Beweglichkeit / Mobilität = 5 und KCCQ 1.3 = 4 od. 5
- (II) Für sich selbst sorgen = 1 und KCCQ 1.1 = 1 od. 2
 Für sich selbst sorgen = 2 und KCCQ 1.1 = 1 od. 2
 Für sich selbst sorgen = 4 und KCCQ 1.1 = 4 od. 5
 Für sich selbst sorgen = 5 und KCCQ 1.1 = 4 od. 5
 bzw.
 Für sich selbst sorgen = 1 und KCCQ 1.2 = 1 od. 2

- Für sich selbst sorgen = 2 und KCCQ 1.2 = 1 od. 2
Für sich selbst sorgen = 4 und KCCQ 1.2 = 4 od. 5
Für sich selbst sorgen = 5 und KCCQ 1.2 = 4 od. 5
- (III) Alltägliche Tätigkeiten = 1 und KCCQ 1.4 = 1 od. 2
Alltägliche Tätigkeiten = 2 und KCCQ 1.4 = 1 od. 2
Alltägliche Tätigkeiten = 4 und KCCQ 1.4 = 4 od. 5
Alltägliche Tätigkeiten = 5 und KCCQ 1.4 = 4 od. 5
- (IV) Schmerzen / körperliche Beschwerden = 1 und KCCQ 4 = 1 od. 2
Schmerzen / körperliche Beschwerden = 2 und KCCQ 4 = 1 od. 2
bzw.
Schmerzen / körperliche Beschwerden = 1 und KCCQ 6 = 1 od. 2
Schmerzen / körperliche Beschwerden = 2 und KCCQ 6 = 1 od. 2
bzw.
Schmerzen / körperliche Beschwerden = 1 und KCCQ 8 = 1 od. 2
Schmerzen / körperliche Beschwerden = 2 und KCCQ 8 = 1 od. 2
- (V) Angst / Niedergeschlagenheit = 1 und KCCQ 14 = 1 od. 2
Angst / Niedergeschlagenheit = 2 und KCCQ 14 = 1 od. 2
Angst / Niedergeschlagenheit = 4 und KCCQ 14 = 4 od. 5
Angst / Niedergeschlagenheit = 5 und KCCQ 14 = 4 od. 5

4.2.3 Deskriptive Analyse

Nach der Datenbereinigung wurden mittels deskriptiver Statistik die Merkmalsausprägungen analysiert. Die deskriptive Analyse der Daten diente als Grundlage zur Erstellung und Überprüfung von Hypothesen.

Wichtige Voraussetzungen für die Auswahl der statistischen Verfahren sind unter anderem das Wissen über die Art der Stichprobe, das gegebene Skalenniveau und die vorhandene Verteilung. Bei den Stichproben wurde zwischen abhängiger und unabhängiger Stichprobe unterschieden. Weiters wurden die Daten in Nominalskala, Ordinalskala und metrische Skala aufgeteilt. Für die Analyse der nominal- und ordinalskalierten Variablen wurden nichtparametrische Testverfahren angewendet. Bei der Analyse der metrisch skalierten Variablen wurden die Testverfahren in Abhängigkeit der vorliegenden Verteilung gewählt.

Tabelle 4.4 liefert die Angaben über die Skalenniveaus der wichtigsten Parameter.

Parameter	Skalenniveau
Alter	metrische Skala
Geschlecht	Nominalskala
NYHA Klasse	Ordinalskala
EQ5D:	
Dimensionen	Ordinalskala
Index Werte	metrische Skala
VAS	metrische Skala
KCCQ:	
Fragen	Ordinalskala
Subskalen	metrische Skala

Tabelle 4.4: Die Skalenniveaus der wichtigsten Parameter.

Bei allen metrisch skalierten Variablen wurde überprüft, ob Normalverteilung vorlag oder nicht. Dies erfolgte unter anderem mithilfe des Shapiro-Wilk-Tests. Der Shapiro-Wilk-Test überprüft dabei die Nullhypothese: „Es liegt Normalverteilung vor.“ War die Irrtumswahrscheinlichkeit größergleich 5%, wurden für die Analyse parametrische Testverfahren verwendet. Bei einer geringeren Irrtumswahrscheinlichkeit wurde auf nichtparametrische Testverfahren zurückgegriffen.

Weiters wurden für die Verteilungsanalyse auch graphische Hilfsmittel verwendet. So wurden auch Q-Q-Plots und Histogramme analysiert und deren Informationen bezüglich der Verteilung berücksichtigt.

4.2.4 Aufstellung und Überprüfung der Hypothese

Nach der deskriptiven Analyse der Daten wurde eine Nullhypothese erstellt. Die Nullhypothese lautete:

- H_0 : Beide Stichproben entspringen derselben Grundgesamtheit.

Durch die Überprüfung der Nullhypothese wurde festgestellt, ob sich eine Untersuchungsgruppe signifikant von der anderen unterscheidet, dadurch konnten auch Rückschlüsse auf die Behandlungsmethoden gezogen werden. Das Vorgehen bei der Auswahl der statistischen Verfahren wird mithilfe der folgenden Flussdiagramme genauer beschrieben. Zusätzlich zu den statistischen Tests wurden auch graphische Analysen durchgeführt. Boxplots und Liniendiagramme wurden zur graphischen Gegenüberstellung von Werten und Verläufen verwendet. Das Signifikanzniveau wurde mit 0,05 angenommen.

Vergleich zwischen den Gruppen

In Abbildung 4.4 ist das Flussdiagramm für die Auswertung bei Vergleichen zwischen den zwei Untersuchungsgruppen dargestellt.

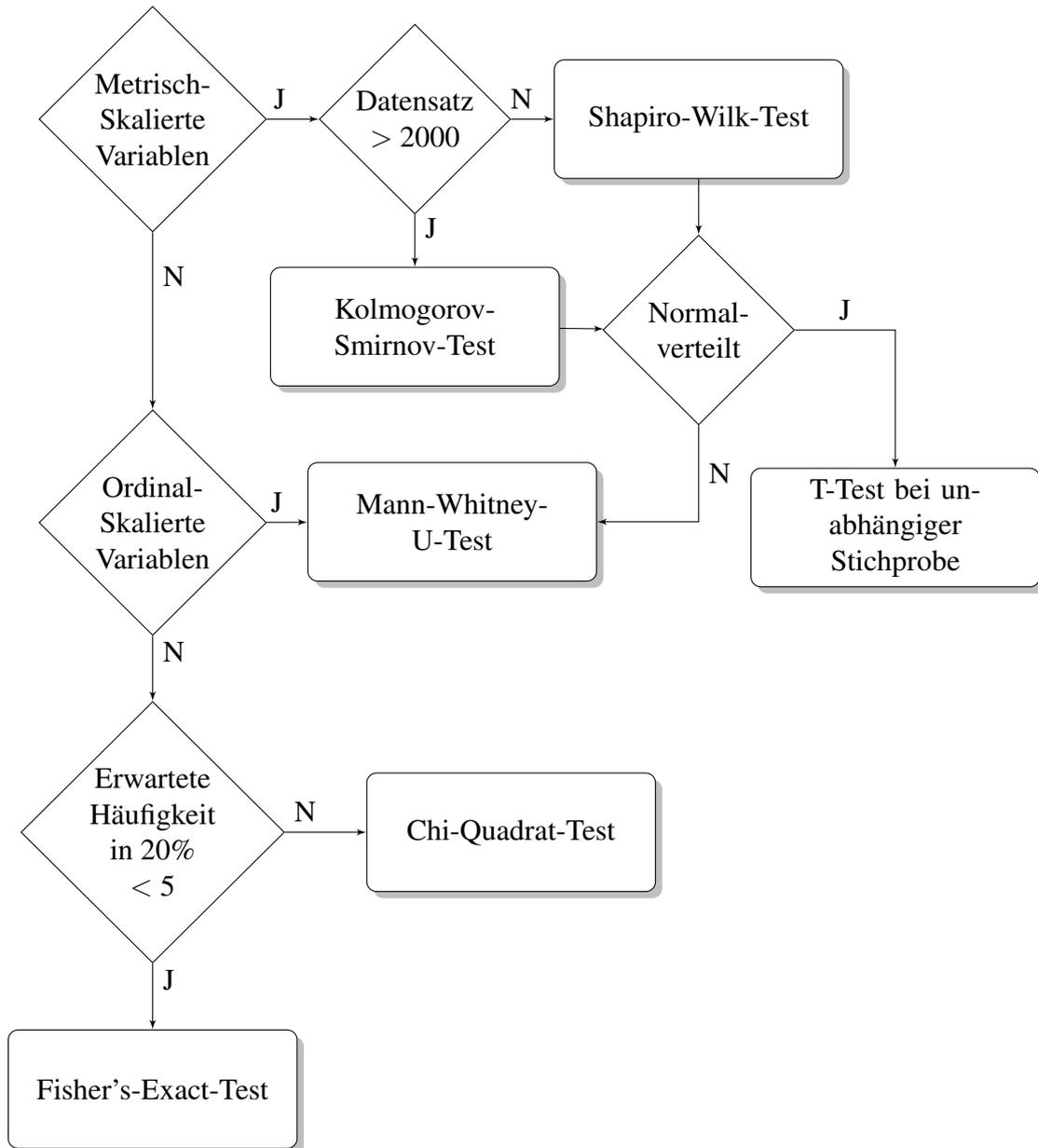


Abbildung 4.4: Vergleich zwischen den Gruppen.

Ausgehend vom Skalenniveau der Variablen werden die statistischen Verfahren gewählt. Handelt es sich um metrisch-skalierte Variablen wird zuerst auf Normalverteilung überprüft. Bei einem Datensatz, der größer ist als $N=2000$, wird der Kolmogorov-Smirnov-Test angewendet, anderenfalls der Shapiro-Wilk-Test. Liegt Normalverteilung vor, wird der T-Test

bei unabhängiger Stichprobe durchgeführt. Weisen die Daten keine Normalverteilung auf, kommt der Mann-Whitney-U-Test zum Einsatz. Dieser Test wird auch für ordinalskalierte Variablen angewendet. Handelt es sich weder um metrisch-skalierte noch um ordinalskalierte Variablen, dann besteht auch die Möglichkeit, dass nominalskalierte Variablen vorliegen. In diesem Fall wird noch überprüft, ob bei einer Kreuztabelle die erwarteten Häufigkeiten in 20% der Fälle unter 5 liegen. Ist dies nicht der Fall wird der Chi-Quadrat-Test durchgeführt, andernfalls der Fisher's-Exact-Test.

Vergleich innerhalb einer Gruppe

Die Flussdiagramme für die Auswertung bei Vergleichen innerhalb einer Untersuchungsgruppe sind in den Abbildungen 4.5 und 4.6 dargestellt. Genau wie bei Abbildung 4.4 beschrieben, wird auch hier bei metrisch-skalierten Variablen auf Normalverteilung überprüft.

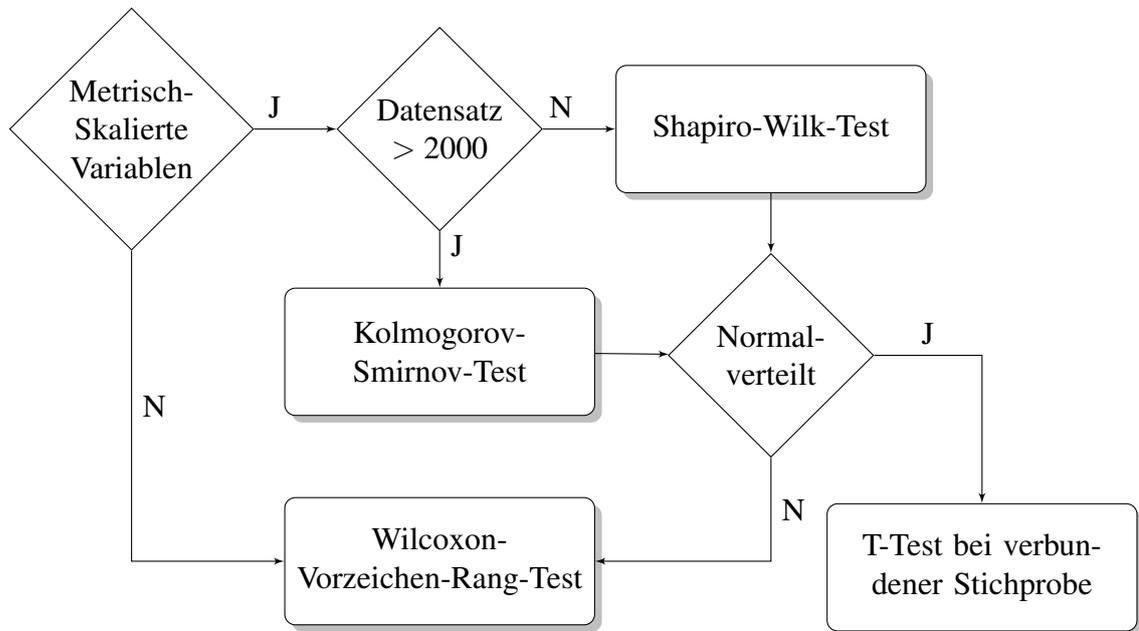


Abbildung 4.5: Vergleich zwischen Zeitpunkten innerhalb einer Gruppe.

Die Abbildung 4.5 zeigt das Vorgehen beim Vergleich zwischen Zeitpunkten innerhalb einer Gruppe. So wird zum Beispiel der EQ-5D-Index-Wert zu Beginn der Studie mit dem Wert nach 6 Monaten bei ein und derselben Untersuchungsgruppe verglichen.

Bei Normalverteilung der Daten wird der T-Test bei verbundener Stichprobe durchgeführt. Handelt es sich weder um metrisch-skalierte Variablen noch um normalverteilte Daten, wird der Wilcoxon-Vorzeichenrang-Test angewendet.

Das Flussdiagramm in Abbildung 4.6 stellt den Ablauf beim Vergleich zwischen Altersgruppen bzw. NYHA-Klassen innerhalb einer Untersuchungsgruppe dar. Beispielsweise für den Vergleich des EQ-5D-Index-Wertes zwischen NYHA-Klasse I, II und III zu Studien-Zeitpunkt 1.

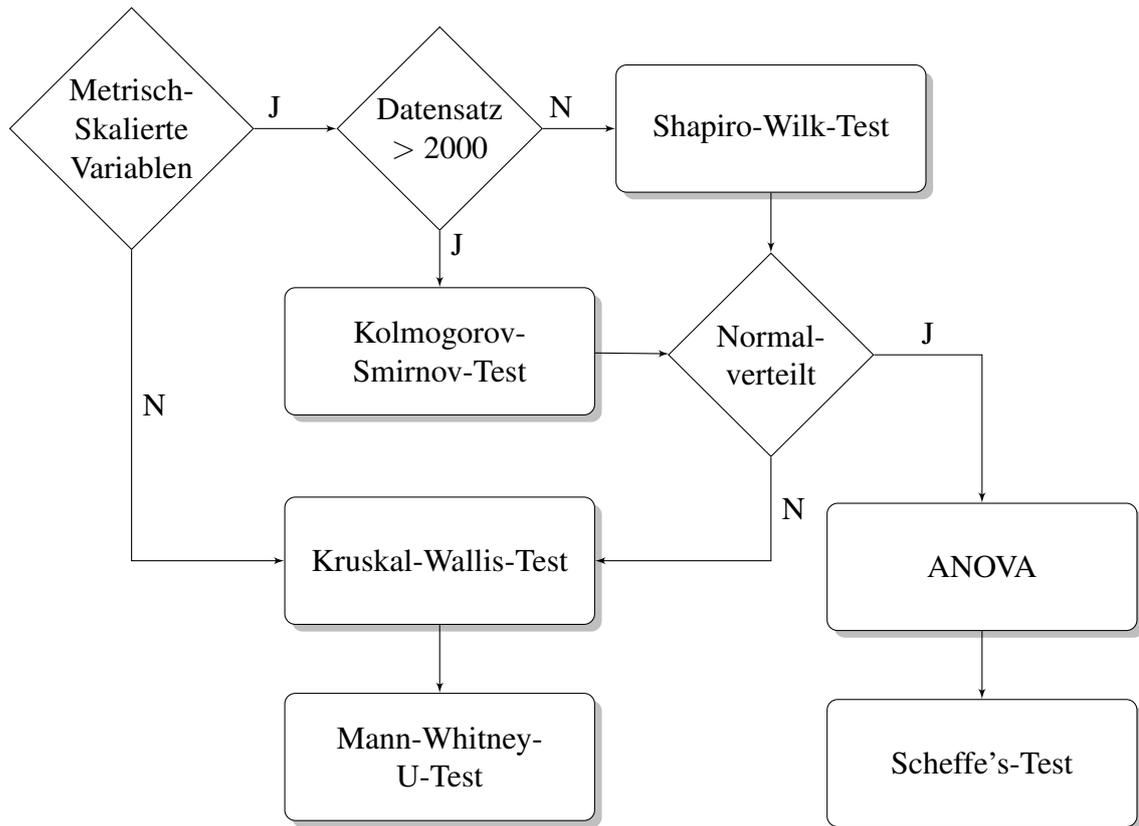


Abbildung 4.6: Vergleich zwischen NYHA-Klassen/Altersgruppen innerhalb einer Gruppe.

Handelt es sich um normalverteilte Daten, wird die ANOVA für Vergleiche herangezogen. Zeigt sich dabei ein statistisch signifikanter Unterschied, wird dieser mittels Scheffe's-Test als Post-Hoc-Vergleich genauer bestimmt. Liegt keine Normalverteilung vor, wird der Kruskal-Wallis-Test durchgeführt. Dieser wird auch für nicht metrisch-skalierte Variablen angewendet. Ergibt sich dabei ein statistisch signifikanter Unterschied, kommt für Post-Hoc-Vergleiche der Mann-Whitney-U-Test zum Einsatz.

4.3 Auswertung des Testdatensatzes

Dieses Kapitel zeigt Ergebnisse von der Auswertung des Testdatensatzes. Die dargestellten Tabellen und Abbildungen stammen von der Anwendung der entwickelten Auswertestrategie. Da die Datenauswertung unter Blind-Bedingungen (dem Auswerter war die Gruppenzuteilung unbekannt) erfolgte, konnte nicht unterschieden werden, bei welcher Gruppe es sich um die Behandlungs- bzw. Kontrollgruppe handelt. Die Gruppen wurden daher lediglich mit den beiden Großbuchstaben A und B gekennzeichnet.

4.3.1 Testdatensatz

Der Testdatensatz besteht aus 35 Patienten, von diesen wurden 18 der Gruppe A und 17 der Gruppe B zugewiesen. Alle 35 Patienten haben die Lebensqualitätsfragebögen (EQ-5D und KCCQ) mindestens einmal ausgefüllt. Von 22 Patienten lagen die Ergebnisse der Lebensqualitätsfragebögen nach 6 Monaten (Studien-Zeitpunkt 2) vor. 10 Patienten befanden sich dabei in Gruppe A und 12 in Gruppe B. Insgesamt 13 Patienten haben die Fragebögen ein drittes mal, nach 12 Monaten (Studien-Zeitpunkt 3), ausgefüllt. 7 Patienten waren dabei der Gruppe A zugeteilt und 6 der Gruppe B.

4.3.2 Vergleich zwischen Gruppe A und B

Tabelle 4.5 zeigt einen Vergleich der Ausgangsdaten zwischen Gruppe A und B. Hierfür wurden alle Probanden, die die Fragebögen mindestens einmal ausgefüllt haben, herangezogen. Statistisch signifikante Unterschiede gibt es beim Vergleich beider Gruppen im Bezug auf Ergebnisse des KCCQ.

Beide Gruppen bestehen aus Personen mit einem Durchschnittsalter von knapp über 70 Jahren. Der Frauenanteil beträgt 17% bei Gruppe A bzw. 18% bei Gruppe B. Bei den NYHA-Klassen gibt es keinen signifikanten Unterschied zwischen beiden Gruppen. Bezüglich der NYHA-Klassen ist noch zu erwähnen, dass bei Gruppe A für 4 Patienten und bei Gruppe B für einen Patienten keine Zuordnung zu einer NYHA-Klasse angegeben war. Bei den Lebensqualitätsvergleichen mit Hilfe des EQ-5D-Fragebogen zeigen sich ebenfalls keine signifikanten Unterschiede zwischen den Gruppen. Einzig bei den Ergebnissen des KCCQ sind statistisch signifikante Unterschiede vorhanden. So zeigt sich bei der „Klinischen Zusammenfassung“, dass die Gruppe B mit $68,2 \pm 13,7$ einen signifikant besseren Wert hat als Gruppe A mit $54,6 \pm 13,4$ (p-Wert = 0,006).

4 Ergebnisse

Parameter	Gruppe A	Gruppe B	p-Wert	Test
Alter				
MW	71	74		
SD	11	8		
Median	71,5	72	0,467	Mann-Whitney
Geschlecht				
Frauen	3(17)	3(18)	0,642	Fisher's-Exact
Männer	15(83)	14(82)		
NYHA Klasse				
I	5(36)	5(31)	0,843	Mann-Whitney
II	4(28)	5(31)		
III	5(36)	6(38)		
IV	0 (0)	0 (0)		
EQ5D				
UK-Index				
MW	0,728	0,807		
SD	0,252	0,145		
Median	0,796	0,809	0,596	Mann-Whitney
D-Index				
MW	0,834	0,878		
SD	0,196	0,1		
Median	0,905	0,9	0,987	Mann-Whitney
Euro-Index				
MW	0,717	0,789		
SD	0,225	0,15		
Median	0,781	0,752	0,407	Mann-Whitney
VAS				
MW	65,83	68,82		
SD	17,84	19,33		
Median	70	60	0,906	Mann-Whitney
KCCQ:				
Funktionaler Status				
MW	47,6	60,5	0,045	T-Test
SD	16,8	19,9		
Median	51	60,4		
Klinische Zusammenfassung				
MW	54,6	68,2	0,006	T-Test
SD	13,4	13,7		
Median	57,6	65,1		

Tabelle 4.5: Vergleich der Ausgangswerte (Studien-Zeitpunkt 1) des Testdatensatzes. Angaben: MW=Mittelwert, SD=Standardabweichung, Median oder Anzahl (Prozent).

Abbildung 4.7 zeigt einen Boxplot-Vergleich der „Klinischen Zusammenfassung“ zwischen Gruppe A und B. Im Boxplot aus Abbildung 4.7 ist ersichtlich, dass der Median bei Gruppe B im Vergleich zur Gruppe A um 10% höher angesiedelt ist. Ebenso liegen die mittleren 50% der Daten von Gruppe B über denen der Gruppe A.

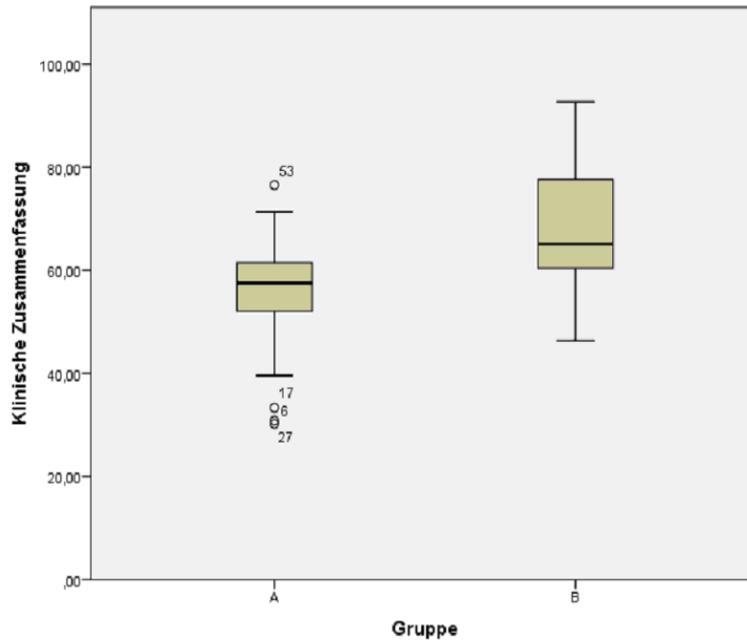


Abbildung 4.7: Klinische Zusammenfassung des KCCQ für die Ausgangswerte aller Probanden. Ein Vergleich zwischen Gruppe A und B.

In Abbildung 4.8 und 4.9 sind die Subkategorien „Symptome“ und „Lebensqualität“ des KCCQ dargestellt. Diese beiden Subkategorien weisen als einzige statistisch signifikante Unterschiede zwischen den Gruppen auf (Symptome: p-Wert = 0,015, Lebensqualität: p-Wert = 0,001) und spiegeln die „Klinische Zusammenfassung“ wider, in die beide Subkategorien miteinfließen. Das heißt auch, in diesen beiden Fällen liegt der Median von Gruppe B signifikant höher als jener von Gruppe A.

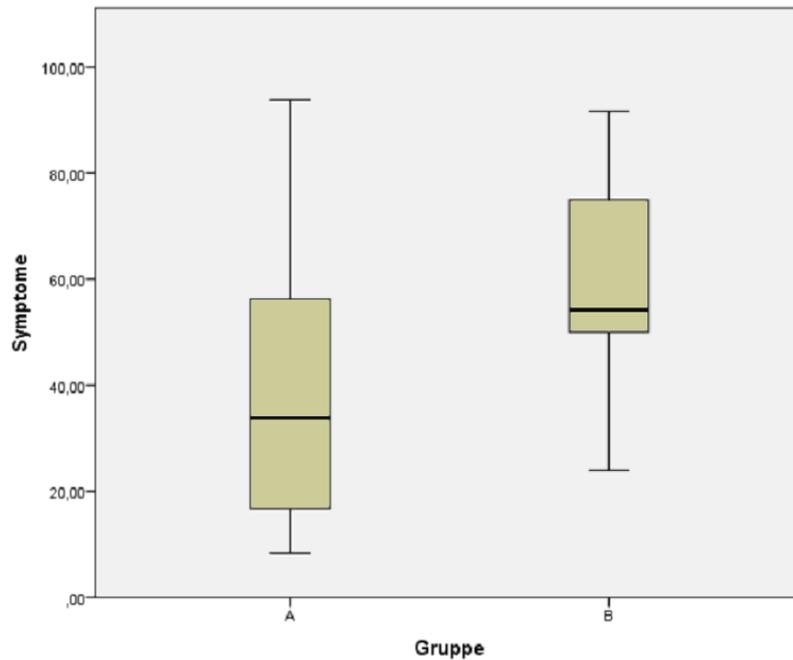


Abbildung 4.8: Subkategorie Symptome des KCCQ für die Ausgangswerte aller Probanden. Ein Vergleich zwischen Gruppe A und B.

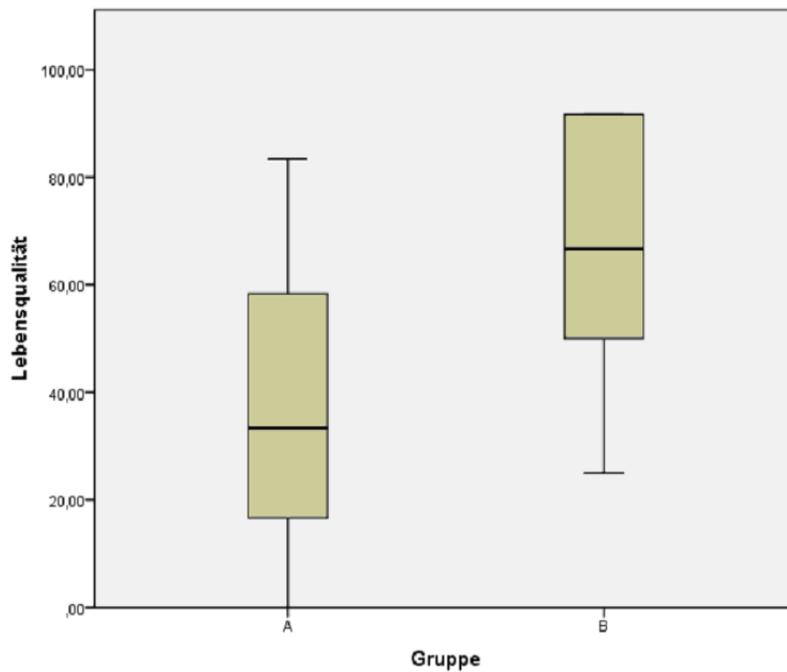


Abbildung 4.9: Subkategorie Lebensqualität des KCCQ für die Ausgangswerte aller Probanden. Ein Vergleich zwischen Gruppe A und B.

4 Ergebnisse

Tabelle 4.6 zeigt einen Vergleich der Ausgangsdaten zwischen den Gruppen für alle Probanden, die die Fragebögen dreimal ausgefüllt haben.

Parameter	Gruppe A	Gruppe B	p-Wert	Test
Alter				
MW	70	75		
SD	7	10		
Median	71	75	0,565	Mann-Whitney
Geschlecht			0,731	Fisher's-Exact
Frauen	1(14)	1(17)		
Männer	6(86)	5(83)		
NYHA Klasse			0,762	Mann-Whitney
I	3(43)	1(17)		
II	1(14)	3(50)		
III	3(43)	2(33)		
IV	0 (0)	0 (0)		
EQ5D				
UK-Index				
MW	0,66	0,747	0,599	T-Test
SD	0,361	0,174		
Median	0,74	0,753		
D-Index				
MW	0,762	0,852		
SD	0,277	0,124		
Median	0,845	0,866	0,885	Mann-Whitney
Euro-Index				
MW	0,667	0,719	0,724	T-Test
SD	0,317	0,178		
Median	0,694	0,721		
VAS				
MW	59,29	63,33		
SD	19,67	20,41		
Median	60	52,5	1,000	Mann-Whitney
KCCQ:				
Funktionaler Status				
MW	48,29	56,08	0,478	T-Test
SD	16,92	21,31		
Median	51	53,4		
Klinische Zusammenfassung				
MW	54,13	63,02	0,356	T-Test
SD	15,27	18,06		
Median	52,9	56,4		

Tabelle 4.6: Vergleich der Ausgangswerte aller Probanden des Testdatensatzes, die die Fragebögen für alle vorgesehenen Zeitpunkte ausgefüllt haben. Angaben: MW=Mittelwert, SD= Standardabweichung, Median oder Anzahl (Prozent).

Aus den Ergebnissen von Tabelle 4.6 lässt sich erkennen, dass es keine statistisch signifikanten Unterschiede zwischen den beiden Gruppen gibt. Dies setzt sich auch bei der Betrachtung von Zeitpunkt 2 und 3 fort. Hier gibt es ebenfalls keine statistisch signifikanten Unterschiede zwischen den Gruppen (siehe Anhang, Tabelle 8.3 und 8.4).

4.3.3 Vergleiche innerhalb einer Gruppe

Für diese Vergleiche wurden jene Probanden herangezogen, von denen Ergebnisse der Lebensqualitätsfragebögen nach 12 Monaten (Studien-Zeitpunkt 3) vorlagen.

Die Tabellen 4.7 und 4.8 zeigen einen Vergleich zwischen Studien-Zeitpunkt 1 und 3 innerhalb von Gruppe A und B.

Parameter	Zeitpunkt 1	Zeitpunkt 3	p-Wert	Test
EQ5D				
UK-Index				
MW	0,66	0,578		
SD	0,361	0,499		
Median	0,74	0,778	0,612	Wilcoxon
D-Index				
MW	0,762	0,692		
SD	0,277	0,402		
Median	0,845	0,91	0,398	Wilcoxon
Euro-Index				
MW	0,667	0,638	0,781	T-Test
SD	0,317	0,378		
Median	0,694	0,727		
VAS				
MW	59,29	64,14	0,639	T-Test
SD	19,67	25,5		
Median	60	50		
KCCQ:				
Funktionaler Status				
MW	48,29	66,59	0,263	T-Test
SD	16,92	36,62		
Median	51	74,5		
Klinische Zusammenfassung				
MW	54,13	77,12	0,051	T-Test
SD	15,27	24,51		
Median	52,9	83,1		

Tabelle 4.7: Vergleich zwischen Studien-Zeitpunkt 1 und 3 bei Gruppe A, für alle Probanden, die die Fragebögen für alle vorgesehenen Zeitpunkte ausgefüllt haben. Angaben: MW=Mittelwert, SD=Standardabweichung, Median.

4 Ergebnisse

Parameter	Zeitpunkt 1	Zeitpunkt 3	p-Wert	Test
EQ5D				
UK-Index				
MW	0,747	0,645	0,149	T-Test
SD	0,174	0,227		
Median	0,753	0,632		
D-Index				
MW	0,852	0,779	0,205	T-Test
SD	0,124	0,159		
Median	0,866	0,774		
Euro-Index				
MW	0,719	0,634	0,187	T-Test
SD	0,178	0,226		
Median	0,721	0,608		
VAS				
MW	63,33	62,5	0,914	Wilcoxon
SD	20,41	17,23		
Median	52,5	62,5		
KCCQ:				
Funktionaler Status				
MW	56,08	60,85	0,523	T-Test
SD	21,31	21,78		
Median	53,4	54,7		
Klinische Zusammenfassung				
MW	63,02	69,75	0,251	T-Test
SD	18,06	15,29		
Median	56,4	67,7		

Tabelle 4.8: Vergleich zwischen Studien-Zeitpunkt 1 und 3 bei Gruppe B, für alle Probanden, die die Fragebögen für alle vorgesehenen Zeitpunkte ausgefüllt haben. Angaben: MW=Mittelwert, SD=Standardabweichung, Median.

Bei Betrachtung von Tabelle 4.7 und 4.8 ist ersichtlich, dass es keine signifikanten Unterschiede innerhalb der Gruppen gibt. Bei der „Klinischen Zusammenfassung“ aus Gruppe A zeigt sich mit einem p-Wert von 0,051 eine Tendenz. Dabei gibt es von Studien-Zeitpunkt 1 auf Studien-Zeitpunkt 3 eine Verbesserung von durchschnittlich 54,13 auf 77,12.

Abbildung 4.10 zeigt die „Klinische Zusammenfassung“ des KCCQ für jene Probanden, die die Fragebögen dreimal ausgefüllt haben.

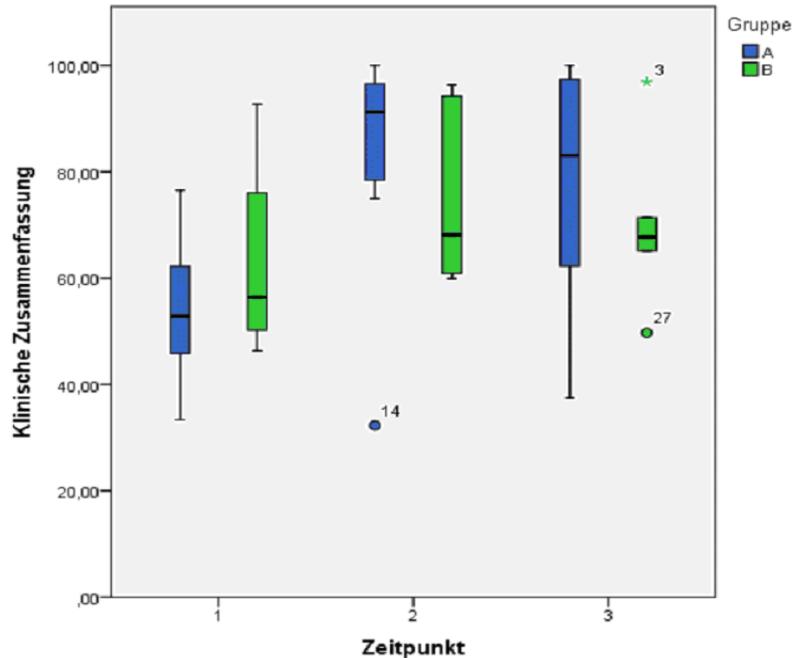


Abbildung 4.10: Klinische Zusammenfassung des KCCQ für alle Probanden, die die Fragebögen für alle vorgeesehenen Zeitpunkte ausgefüllt haben.

Die Boxplots aus Abbildung 4.10 ermöglichen einen graphischen Vergleich sowohl zwischen den Gruppen als auch zwischen den Zeitpunkten innerhalb der Gruppen. Die Verbesserung von Studien-Zeitpunkt 1 auf 2 bei Gruppe A ist statistisch signifikant (p-Wert = 0,028). Bei Gruppe B tritt zwischen Studien-Zeitpunkt 1 und 2 ebenfalls eine Verbesserung auf, jedoch nicht in diesem Ausmaß (p-Wert = 0,079). Ein genauer Überblick über die Vergleiche zwischen den Zeitpunkten innerhalb der Gruppen ist im Anhang gegeben (Tabelle 8.5, 8.6, 8.7, und 8.8).

Die Liniendiagramme aus Abbildung 4.11, 4.12 und 4.13 stellen den Verlauf des Medians beider Gruppen gegenüber. In Abbildung 4.11 zeigt sich, dass beide Gruppen einen ähnlichen Verlauf haben. Wobei die Verbesserung der Werte der „Klinischen Zusammenfassung“ von Studien-Zeitpunkt 1 auf Studien-Zeitpunkt 2 bei Gruppe A signifikanter ausgeprägt ist als bei Gruppe B. Unterschiede in den Verläufen beider Gruppen sind in den Abbildungen 4.12 und 4.13 ersichtlich. Diese Abbildungen stellen die Mediane der VAS und des Index-UK-Wertes zu den 3 unterschiedlichen Studien-Zeitpunkten dar. Gruppe B besitzt in beiden Fällen einen konstanten Verlauf. Gruppe A dagegen zeigt zuerst eine Ver-

4 Ergebnisse

besserung von Studien-Zeitpunkt 1 auf Studien-Zeitpunkt 2 gefolgt von einer Verschlechterung von Studien-Zeitpunkt 2 auf Studien-Zeitpunkt 3.

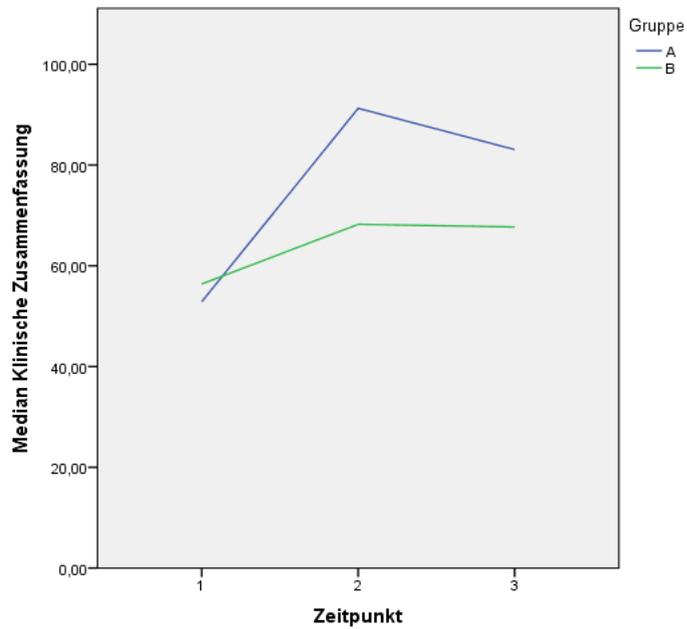


Abbildung 4.11: Verlauf der Mediane der „Klinischen Zusammenfassung“ des KCCQ für Gruppe A und B.

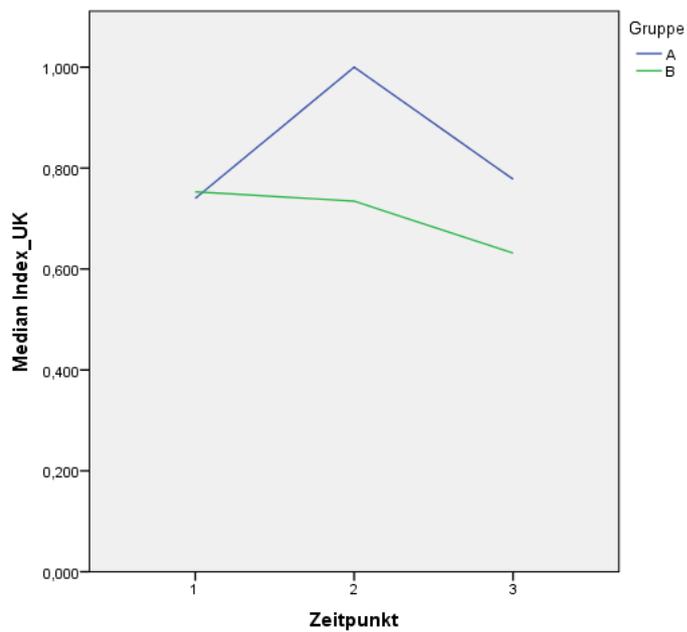


Abbildung 4.12: Verlauf der EQ-5D-Index-UK-Mediane für Gruppe A und B.

4 Ergebnisse

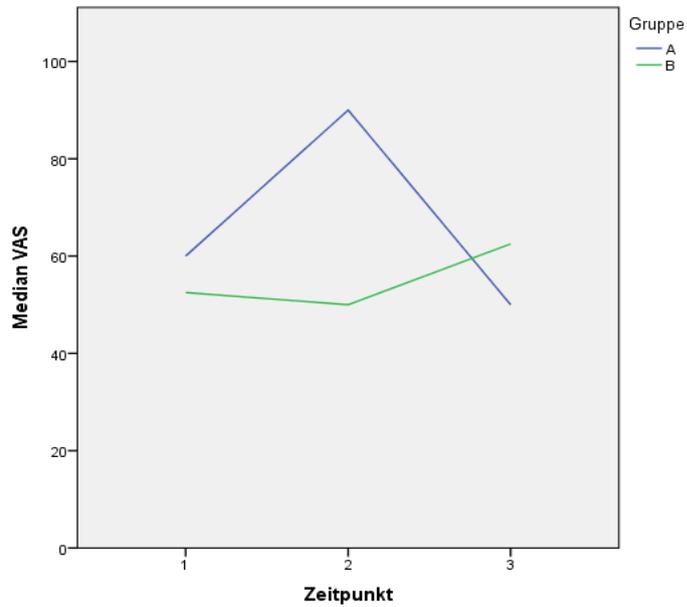


Abbildung 4.13: Verlauf der EQ-5D-VAS-Mediane für Gruppe A und B.

Der Verlauf der EQ-5D-Index-UK-Werte für alle Probanden der Gruppe A ist in Abbildung 4.14 dargestellt. Die Werte streuen über den gesamten Bereich. 3 von 7 Probanden verbessern sich von Studien-Zeitpunkt 1 auf Studien-Zeitpunkt 2 und 5 von 7 Probanden haben zu Studien-Zeitpunkt 3 schlechtere Werte als zu Studien-Zeitpunkt 1.

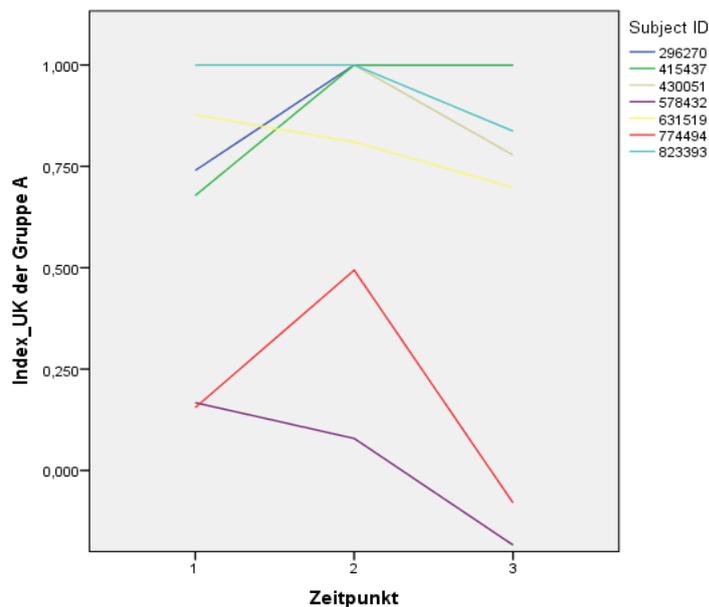


Abbildung 4.14: Verlauf der EQ-5D-Index-UK-Werte der einzelnen Probanden für Gruppe A.

4.3.4 Vergleiche zwischen den NYHA-Klassen

Abbildungen 4.15 und 4.16 stellen die EQ-5D-Index-UK-Werte für die NYHA-Klassen 1 bis 3 zu den unterschiedlichen Zeitpunkten für Gruppe A und B dar. Auch in diesem Falle sind keine signifikanten Unterschiede zwischen Gruppe A und B ersichtlich. NYHA-Klasse 1 zeigt dabei bessere Lebensqualitätswerte als NYHA-Klasse 2 und 3. Die Werte bei NYHA-Klasse 2 von Gruppe A und bei NYHA-Klasse 1 von Gruppe B stammen von einer einzelnen Person.

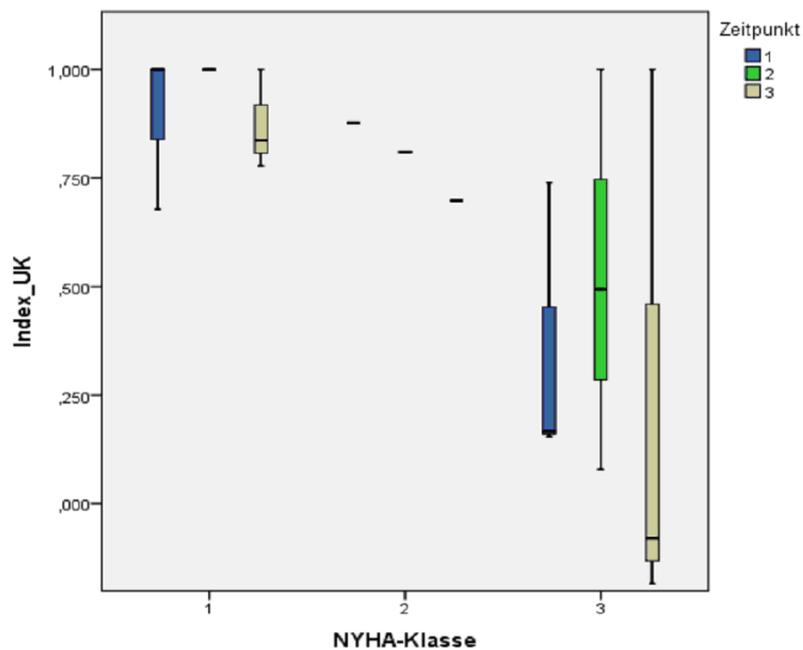


Abbildung 4.15: EQ-5D-Index-UK-Werte in Abhängigkeit der NYHA-Klassen zu den unterschiedlichen Zeitpunkten für Gruppe A.

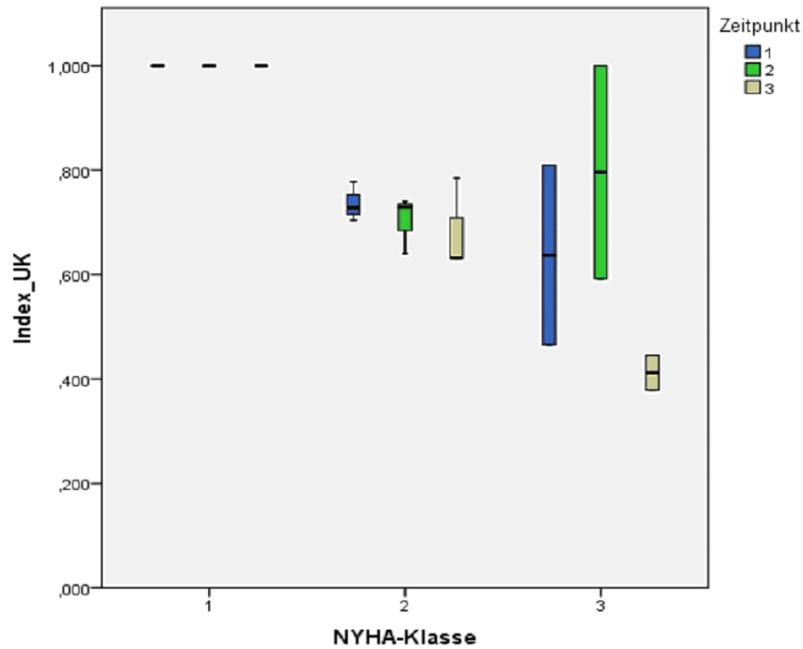


Abbildung 4.16: EQ-5D-Index-UK-Werte in Abhängigkeit der NYHA-Klassen zu den unterschiedlichen Zeitpunkten für Gruppe B.

4.3.5 Ergebnisse nach Datenbereinigung

Der Testdatensatz umfasste widersprüchliche Ergebnisse. So wurden zum Beispiel vergleichbare Subkategorien des KCCQ und EQ-5D-Fragebogens unterschiedlich bewertet. Abbildung 4.17 zeigt die Gegenüberstellung zweier vergleichbarer Subkategorien bei ein und derselben Person. Männlich, NYHA Klasse 1, Alter 76 Jahre aus Gruppe A. Hier wurde der Gesundheitszustand zum Studien-Zeitpunkt 1 sowohl bei der Subkategorie des EQ-5D-Fragebogens als auch bei der Subkategorie des KCCQ mit 1 bewertet. Bei der Subkategorie des EQ-5D-Fragebogens steht der Wert 1 für: „Ich habe keine Schmerzen oder Beschwerden“, dagegen steht der Wert 1 bei der Subkategorie des KCCQ für: „Die Schwellungen der Füße, Knöchel oder Beine waren während der letzten 2 Wochen extrem beschwerlich.“

4 Ergebnisse

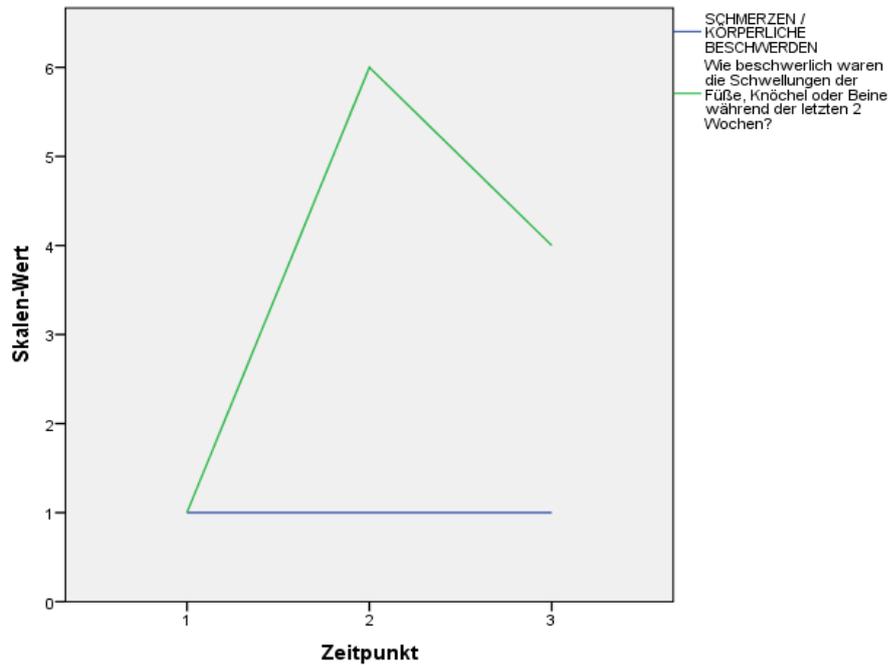


Abbildung 4.17: Gegenüberstellung zweier vergleichbarer Subkategorien des KCCQ und EQ-5D-Fragebogens bei ein und derselben Person (Blau: EQ-5D, Grün: KCCQ).

In Abbildung 4.17 steht der Skalen-Wert für die unterschiedlichen Gesundheitszustände. Beim KCCQ steht in diesem Fall der Wert 6 für den bestmöglichen Gesundheitszustand, welcher beim EQ-5D-Fragebogen dem Wert 1 entspricht. Beim EQ-5D-Fragebogen geht die Skala von 1 bis 5.

Würde man alle widersprüchlichen Daten entfernen, verbleiben noch 3 Patienten in Gruppe A und 8 Patienten in Gruppe B, die die Fragebögen zu Studienbeginn ausgefüllt haben. Ein Vergleich zwischen den Gruppen ergab keine statistisch signifikanten Unterschiede (Anhang, Tabelle 8.9 und 8.10).

Abbildung 4.18 liefert die Darstellung eines Boxplot-Vergleichs der „Klinischen Zusammenfassung“ des KCCQ zwischen Gruppe A und B. Hier ist zu erkennen, dass die Mediane beider Gruppen nicht signifikant voneinander abweichen.

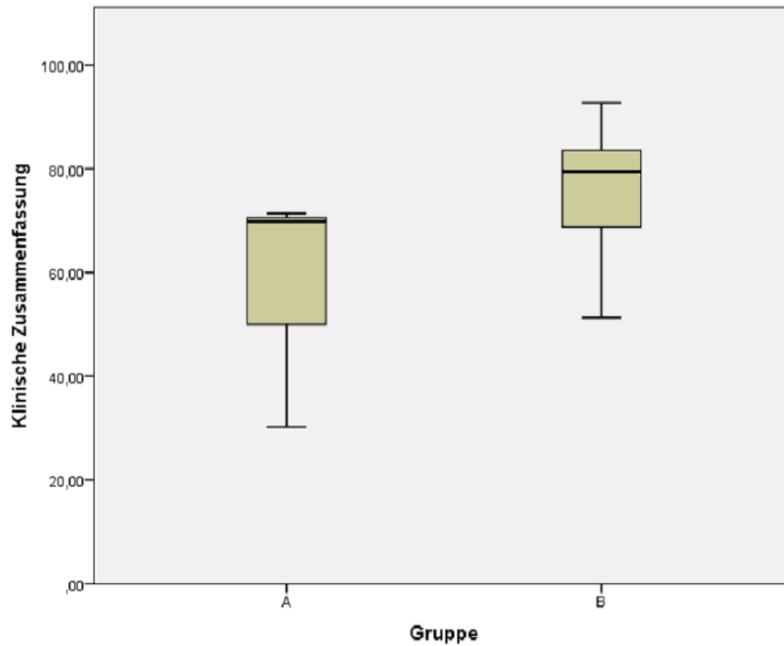


Abbildung 4.18: „Klinische Zusammenfassung“ des KCCQ. Vergleich zwischen Gruppe A und B zu Studienbeginn.

In Abbildung 4.19 ist eine Gegenüberstellung der Boxplots des EQ-5D-Index-UK-Wertes für die NYHA-Klassen 1 bis 3 zu Studienbeginn ersichtlich. Dabei wurde keine Trennung zwischen Gruppe A und B vorgenommen. Je 3 Probanden haben NYHA-Klasse 1 oder NYHA-Klasse 2 und 4 Probanden haben NYHA-Klasse 3. Es zeigt sich, dass NYHA-Klasse 1 bessere Werte als NYHA-Klasse 2 und NYHA-Klasse 2 bessere Werte als NYHA-Klasse 3 aufweisen.

Abbildung 4.20 stellt die Boxplots des EQ-5D-Index-UK-Wertes für die 3 unterschiedlichen Altersgruppen zu Studienbeginn dar (Altersgruppe 1: bis 70 Jahre, Altersgruppe 2: zwischen 70 und 80 Jahre, Altersgruppe 3: ab 80 Jahre). Dabei wurde keine Trennung zwischen Gruppe A und B vorgenommen. Altersgruppe 1 umfasst dabei 4 Personen, 2 Personen gehören der Altersgruppe 2 an und in der Altersgruppe 3 befinden sich 5 Personen. Zwischen den Altersgruppen sind hier keine signifikanten Unterschiede ersichtlich.

4 Ergebnisse

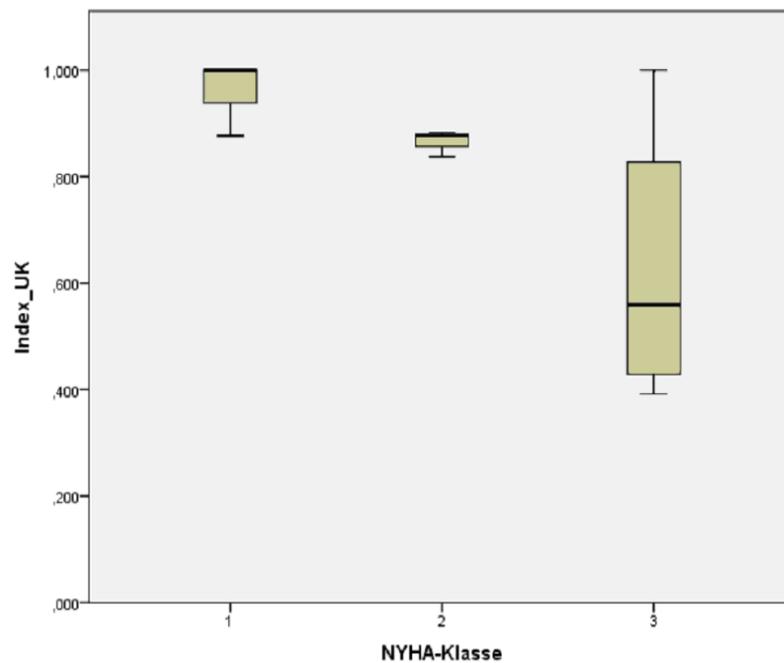


Abbildung 4.19: EQ-5D-Index-UK-Werte für die NYHA-Klassen 1 bis 3 zu Studienbeginn. Ohne Trennung zwischen Gruppe A und B.

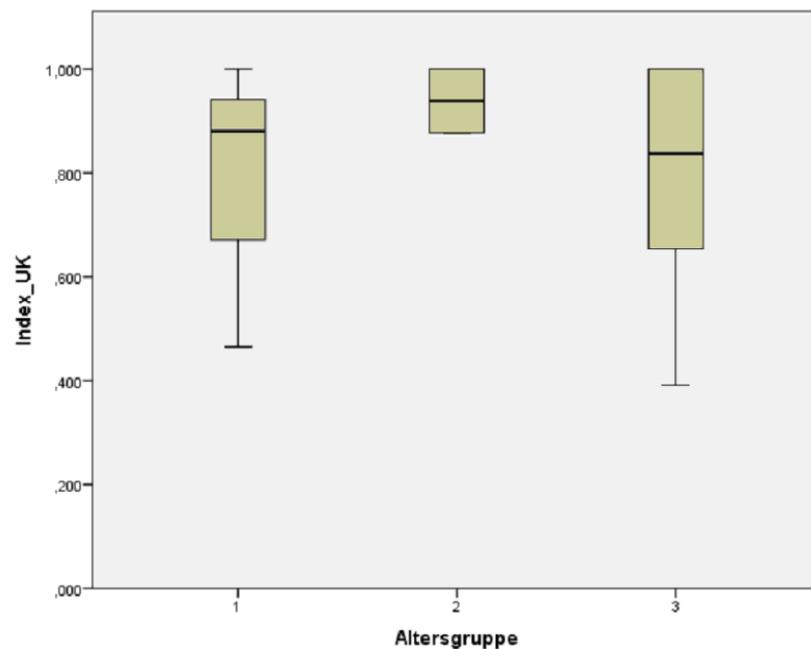


Abbildung 4.20: EQ-5D-Index-UK-Werte für die 3 unterschiedlichen Altersgruppen zu Studienbeginn. Ohne Trennung zwischen Gruppe A und B.

5 Diskussion

5.1 Auswertestrategie

Die Entwicklung der Auswertestrategie erfolgte unter Berücksichtigung der verwendeten statistischen Verfahren in den Publikationen. Für die Auswahl der statistischen Tests war die Übereinstimmung des Studiendesigns der Publikationen mit dem zu untersuchenden Studiendesign ausschlaggebend. Weiters wurde berücksichtigt für welche Vergleiche die statistischen Verfahren herangezogen wurden.

Für das zu untersuchende Studiendesign wurden sowohl der KCCQ als auch der EQ-5D-Fragebogen verwendet. Da es dabei vergleichbare Fragen gibt, wurde eine Analyse dieser Fragen in die Auswertestrategie mit eingeplant. Somit konnten etwaige Widersprüche bei der Beantwortung aufgezeigt und für die weitere Analyse der Daten berücksichtigt werden.

Die deskriptive Analyse der Daten stellt einen wichtigen Punkt dar, um aufbauend darauf die in Frage kommenden statistischen Verfahren zu wählen.

Detailliert beschrieben wird die deskriptive Analyse in keiner der gefundenen Publikationen. Einzig Spiraki et al. [20] geben an eine deskriptive Analyse durchgeführt zu haben. Sie gehen aber nicht genauer auf diesen Punkt der Datenanalyse ein.

Um die vorliegende Verteilung zu überprüfen, wurde wegen seiner relativ guten Teststärke der Shapiro-Wilk-Test verwendet [37]. Ursprünglich wurde der Shapiro-Wilk-Test für einen Stichprobenumfang von $3 < n < 50$ entwickelt [38]. Daher hat der Shapiro-Wilk-Test auch bei kleineren Stichproben eine relativ hohe Teststärke. Erst später wurde er für einen Stichprobenumfang von $n < 2000$ erweitert [39] und schlussendlich noch für $n < 5000$ [40]. Bei der Durchführung des Shapiro-Wilk-Tests in SPSS über die deskriptive Statistik wird auch das Ergebnis des Kolmogorov-Smirnov-Tests angegeben. Der Kolmogorov-Smirnov-Test wird ab einem Stichprobenumfang von $n = 2000$ berücksichtigt.

Seto et al. [31] und Boyne et al. [32] verwenden in ihren Studien zur Überprüfung der Verteilung ebenfalls den Shapiro-Wilk-Test. Bei Seto et al. kommt zusätzlich noch der Kolmogorov-Smirnov-Test zum Einsatz. Bei allen anderen Publikationen wird auf die Überprüfung der Verteilung nicht genauer eingegangen.

Zusätzlich wurde bei der Auswertung des Testdatensatzes die Verteilung auch graphisch mittels Q-Q-Plots und Histogrammen überprüft. Dies stellt eine weitere Absicherung dar, um eventuellen Fehlinterpretationen aus Ergebnissen von statistischen Tests entgegenzuwirken.

Aufgrund der Tatsache, dass sich ein Signifikanzniveau von 0,05 als Standard etabliert hat, wurde auch das Signifikanzniveau für die Auswertestrategie so gewählt. Auch bei Betrachtung der Publikationen zeigt sich, dass in 85% ein Signifikanzniveau von 0,05 angenommen wurde. Von den restlichen 15% gibt es diesbezüglich keine Angaben.

5.1.1 Vergleich zwischen den Gruppen

Das Flussdiagramm aus Abbildung 4.4 stellt den Ablauf bei der Auswertung für all die Daten dar, bei denen es sich um unabhängige Stichprobenvergleiche zwischen zwei Gruppen handelt. Dies schließt sowohl Vergleiche von Alter, Geschlecht und NYHA-Klasse als auch von EQ-5D und KCCQ-Werten mit ein.

Zum Vergleich von metrisch-skalierten Variablen wird in Abhängigkeit der Verteilung entweder der T-Test oder der Mann-Whitney-U-Test angewendet. Eine Übersicht über die Skalenniveaus der unterschiedlichen Parameter ist in Tabelle 4.4 gegeben. Für ordinalskalierte Variablen wird der Mann-Whitney-U-Test verwendet. Diese beiden statistischen Verfahren haben sich im Laufe der Recherche als zuverlässige und häufig verwendete Verfahren erwiesen.

Die MOBITEL-Studie [25] beschäftigt sich ebenfalls mit der Auswirkung von Telemonitoring auf den Behandlungserfolg bei Herzinsuffizienz-Patienten. Dabei wird, wie beim zu untersuchenden Studiendesign, eine Behandlungsgruppe mit einer Kontrollgruppe verglichen. Auch Scherr et al. verwenden dabei zur Auswertung normalverteilter Daten den T-Test. Ebenso wird zum Vergleich von nicht normalverteilten Daten bei unabhängiger Stichprobe der Mann-Whitney-U-Test angewendet.

Insgesamt zeigt sich, dass bei 54% aller betrachteten Studien der T-Test bzw. der Mann-Whitney-U-Test durchgeführt werden. Bei den restlichen Studien wurde teilweise kein Vergleich zwischen zwei Gruppen gemacht bzw. war bei 31% der betrachteten Studien das Studiendesign anders ausgelegt.

Grundsätzlich ist bei der Auswertestrategie für den Vergleich von nominalskalierten Variablen der Chi-Quadrat-Test vorgesehen. Auch Myers et al. [28] verwenden in ihrer Studie den Chi-Quadrat-Test für den Vergleich von nominalskalierten Variablen. Myers et al. befassen sich ebenfalls mit dem Telemonitoring von Herzinsuffizienz-Patienten und ihr Studiendesign weist Ähnlichkeiten mit dem zu untersuchenden Studiendesign auf.

Auch der Chi-Quadrat-Test wird in 54% der betrachteten Publikationen durchgeführt. Eine Einschränkung des Chi-Quadrat-Tests ist, dass die erwartete Häufigkeit nicht in mehr als 20% der Felder einer Kreuztabelle unter 5 liegen darf. Ist dies der Fall und es zeigen sich somit geringe erwartete Häufigkeiten, wird bei der Auswertestrategie der Fisher's-Exact-Test angewendet. Dieser liefert dabei die zuverlässigsten Ergebnisse [19].

Martin-Lesende et al. [34] erwähnen als einzige die Möglichkeit anstelle des Chi-Quadrat-Tests den Fisher's-Exact-Test bei geringen erwarteten Häufigkeiten anzuwenden.

5.1.2 Vergleich innerhalb einer Gruppe

Das Flussdiagramm aus Abbildung 4.5 stellt den Ablauf bei der Auswertung für all die Daten dar, bei denen es sich um verbundene Stichprobenvergleiche innerhalb einer Gruppe handelt. Dies wird zum Beispiel angewendet, um Vergleiche zwischen den EQ-5D-Werten zu unterschiedlichen Studien-Zeitpunkten innerhalb einer Gruppe durchzuführen.

In der Auswertestrategie ist für Vergleiche innerhalb einer Gruppe der T-Test bei verbundener Stichprobe für normalverteilte Daten vorgesehen. Für den Vergleich bei nicht normalverteilten Daten kommt in diesem Fall der Wilcoxon-Vorzeichenrang-Test zum Einsatz. Bei allen betrachteten Publikationen, die Vergleiche bei verbundener Stichprobe durchführen, wird der T-Test bzw. der Wilcoxon-Vorzeichenrang-Test angewendet. Seto et al. [31], die sich in ihrer Studie auch mit Telemonitoring bei Herzinsuffizienz-Patienten befassen, verwenden ebenso den T-Test und den Wilcoxon-Vorzeichenrang-Test für Vergleiche bei verbundener Stichprobe.

Die beiden verwendeten Verfahren ermöglichen Vergleiche zwischen zwei Variablen. Um mehrere Variablen gleichzeitig vergleichen zu können, müsste in diesem Fall der Friedman-Test bzw. die Varianzanalyse mit Messwiederholung angewendet werden. Im zu untersuchenden Studiendesign sind drei verschiedene Messzeitpunkte vorgesehen. Der Testdatensatz beinhaltet auch Patienten, die die Fragebögen zweimal ausgefüllt haben. Um eine größere Flexibilität zu erhalten, wurden daher für die Auswertestrategie der Wilcoxon-Vorzeichenrang-Test und der T-Test herangezogen. Somit besteht auch die Möglichkeit EQ-5D-Werte zwischen Studien-Zeitpunkt 1 und 3 zu vergleichen, ohne dabei auf Post-Hoc-Tests angewiesen zu sein.

Das Flussdiagramm aus Abbildung 4.6 stellt den Ablauf bei der Auswertung zwischen NYHA-Klassen bzw. Altersgruppen dar. Dabei handelt es sich zwar um Vergleiche innerhalb einer Gruppe, da die Personen in den NYHA-Klassen und Altersgruppen aber nicht dieselben sind, stellen sie sehr wohl unabhängige Stichproben dar.

Um diese Vergleiche bei normalverteilten Daten durchzuführen, ist in der Auswertestrategie die ANOVA mit Scheffe's-Test für Post-Hoc-Vergleiche vorgesehen. Für die Auswertung von nicht normalverteilten Daten wird in diesem Fall der Kruskal-Wallis-Test verwendet. Um dabei Post-Hoc-Vergleiche anstellen zu können, muss der Mann-Whitney-U-Test herangezogen werden.

Pinto et al. [21] vergleichen in ihrer Studie drei unterschiedliche Behandlungsgruppen miteinander. Dazu kommt auch bei ihnen die ANOVA mit Scheffe's-Test für Post-Hoc-Vergleiche zum Einsatz. Insgesamt wird bei 31% der betrachteten Publikationen die ANOVA für Vergleiche zwischen mehreren Gruppen verwendet. Bei den restlichen 69% der Publikationen werden solche Vergleiche nicht durchgeführt.

Der Kruskal-Wallis-Test wird in keiner der betrachteten Publikationen erwähnt. Ein Grund dafür könnte sein, dass bei diesen Studien für solche Vergleiche nur normalverteilte Daten vorliegen.

5.2 Auswertung des Testdatensatzes

Das Hauptziel der vorliegenden Arbeit war die „Entwicklung einer Auswertestrategie“, die dann an vorhandenen Daten eines Testdatensatzes angewendet wurde. Die Ergebnisse dieser Auswertung sollen im Folgenden genauer erläutert und diskutiert werden.

Der Vergleich der Ausgangsdaten aus Tabelle 4.5 zeigt einen statistisch signifikanten Unterschied zwischen den Gruppen bei Ergebnissen des KCCQ. Abgesehen von diesem Unterschied gibt es keine signifikante Abweichung zwischen den beiden Gruppen. Der statistisch signifikante Unterschied bei der „Klinischen Zusammenfassung“ ist hauptsächlich durch die Subkategorien „Symptome“ und „Lebensqualität“ begründet. In den Abbildungen 4.8 und 4.9 ist ersichtlich, dass die Werte der Gruppen beinahe über den gesamten Bereich streuen. Die Subkategorie „Lebensqualität“ des KCCQ bezieht sich rein auf die psychische Lebensqualität und kann somit nicht mit der Lebensqualität, die durch den EQ-5D-Fragebogen ermittelt wird, gleichgesetzt werden.

Beim Vergleich der Geschlechterverteilung in Tabelle 4.5 sollte der Chi-Quadrat-Test zur Anwendung kommen. Da sich in beiden Gruppen jeweils 3 Frauen befanden und dadurch die erwartete Häufigkeit in 50% der Felder einer Kreuztabelle < 5 war, wurde der Fisher's-Exakt-Test verwendet.

Bei Betrachtung der Ausgangswerte aller Probanden, die die Fragebögen bereits dreimal ausgefüllt haben, fällt auf, dass es keine statistisch signifikanten Unterschiede zwischen Gruppe A und B gibt. Ein Grund für die Nicht-Signifikanz, ersichtlich in Tabelle 4.6, könnte die geringe Anzahl an Probanden sein. Bei Gruppe A gibt es erst 7 und bei Gruppe B erst 6 Personen, die die Fragebögen bereits dreimal ausgefüllt haben. Je geringer die Anzahl an Probanden, umso mehr steigt die Wahrscheinlichkeit, dass die ermittelten Messgrößen durch Zufall zustande gekommen sind.

In Abbildung 4.10 liegen die Werte der „Klinischen Zusammenfassung“ für beide Gruppen zu Zeitpunkt 1 (Studienbeginn) näher beieinander als in Abbildung 4.7. Der Median verschiebt sich bei Gruppe B nach unten, ebenso wie das 25% Perzentil. Diese Unterschiede zwischen den Boxplots könnten auch durch die geringere Anzahl an Probanden zustande kommen.

Ein ähnlicher Verlauf der Werte beider Gruppen ist in Abbildung 4.11 ersichtlich. Allerdings zeigt sich hier, dass Gruppe A eine Verbesserung zum Zeitpunkt 2 aufweist, die doppelt so stark ist wie jene bei Gruppe B. Aufgrund der unterschiedlichen Ausgangswerte, eventueller Widersprüche in den Daten und vor allem der geringen Anzahl an Probanden lässt sich nicht ermitteln, ob die Verbesserung durch die Behandlungsmethode begründet ist oder nicht.

Einen unterschiedlichen Verlauf der VAS-Werte beider Gruppen zeigt Abbildung 4.13. In diesem Falle weisen beide Gruppen zu den Zeitpunkten 1 und 3 keine signifikanten Unterschiede auf. Bei Zeitpunkt 2 gibt es einen zwar nicht signifikanten, aber doch deutlichen

Unterschied zwischen Gruppe A und B, wobei Gruppe A die besseren Werte aufweist. Ein Grund für diesen Unterschied könnte sein, dass der Umgang mit dem Telemonitoring-System für Gruppe B zu Beginn als Belastung empfunden wurde, sollte Gruppe B die Behandlungsgruppe sein. Handelt es sich aber bei Gruppe A um die Behandlungsgruppe, könnte das Telemonitoring-System zu Beginn ein Sicherheitsgefühl vermittelt und dies zur empfundenen Verbesserung des Gesundheitszustandes geführt haben. Eine Möglichkeit, warum dieser Unterschied beim KCCQ nicht auftritt, wäre eventuell, dass die Werte der VAS sehr stark das persönliche Empfinden widerspiegeln [16], da sie über eine einzelne Skala ermittelt werden und nicht, wie zum Beispiel die Werte des KCCQ, über einen differenzierten Fragebogen.

Abbildung 4.12 zeigt ebenfalls einen unterschiedlichen Verlauf zwischen den beiden Gruppen. In diesem Falle liefert jedoch die Anwendung einer Medianbildung einen falschen Eindruck über die Verläufe beider Gruppen, da zu jedem Zeitpunkt der mittlere Wert betrachtet wird, welcher den Verlauf der Mehrzahl der Probanden, zumindest bei Gruppe A, nicht widerspiegelt. Dies lässt sich in Abbildung 4.14 genauer erkennen. Die Mehrzahl an Probanden der Gruppe A weisen einen Medianverlauf wie die Gruppe B aus Abbildung 4.12 auf.

Die Verläufe aus den Abbildungen 4.11 und 4.13 stellen die Mehrzahl der Probanden gut dar. Im Anhang in den Abbildungen 8.1, 8.2, 8.3, 8.4 und 8.5 sind die restlichen Verläufe der einzelnen Patienten gegenübergestellt.

Bei der Auswertung wurde auch die Möglichkeit berücksichtigt, dass sich das Telemonitoring nur bei einer bestimmten Altersgruppe bzw. NYHA-Klasse positiv/negativ auswirkt. Da die Anzahl der Probanden sehr gering ist, konnte die Auswertestrategie nicht getestet werden. Bei Betrachtung der Abbildungen 4.15 und 4.16 scheint es so, als würde es keine signifikanten Unterschiede zwischen Gruppe A und B geben.

Eine weitere Einschränkung des Datensatzes gibt es durch widersprüchliche Beantwortung vergleichbarer Subkategorien des KCCQ und EQ-5D-Fragebogens. Ein Beispiel einer solchen widersprüchlichen Beantwortung ist in Abbildung 4.17 ersichtlich. Hier wurden zum Zeitpunkt 1 sowohl die Subkategorie des KCCQ als auch jene des EQ-5D-Fragebogens mit 1 bewertet. Allerdings stellt der Wert 1 bei der Subkategorie des KCCQ den schlechtesten Gesundheitszustand dar, dagegen steht er bei der Subkategorie des EQ-5D-Fragebogens für den besten Gesundheitszustand. Es ist widersprüchlich, wenn der Patient angibt, keine Schmerzen oder Beschwerden zu haben, zum gleichen Zeitpunkt aber behauptet, die Schwellungen der Füße, Knöchel oder Beine seien in den letzten 2 Wochen extrem beschwerlich gewesen. Da der Patient der NYHA-Klasse 1 angehört, lässt sich vermuten, dass die Beantwortung der Subkategorie des EQ-5D-Fragebogens dem tatsächlichen Gesundheitszustand entspricht.

Würde man alle widersprüchlichen Daten entfernen, verbleiben nur noch 3 Patienten in Gruppe A und 8 Patienten in Gruppe B, welche die Fragebögen zu Studienbeginn bereits ausgefüllt haben. Dies würde bedeuten, dass bei 69% der Probanden eine widersprüchliche

Beantwortung vorliegt. Durch die Verringerung der Daten ist es nicht mehr möglich, statistisch aussagekräftige Schlüsse zu ziehen.

Bei der Betrachtung der vorhandenen NYHA-Klassen in Abbildung 4.19 zeigt die NYHA-Klasse 1 den besten Gesundheitszustand, der sich zur NYHA-Klasse 3 hin verschlechtert. Dies entspricht auch der Erwartung, dass die Personen mit der leichtesten Stufe der Erkrankung den besten Gesundheitswert aufweisen.

Aufgrund der Tatsache, dass sich die NYHA-Klassen relativ gleichmäßig auf die 3 unterschiedlichen Altersgruppen aufteilen, lässt sich auch kein Trend bei der Betrachtung der Altersgruppen in Abbildung 4.20 erkennen.

Grundsätzlich muss festgehalten werden, dass die Anzahl der Probanden im Testdatensatz zu gering ist, um endgültige Aussagen darüber machen zu können, ob sich eine Gruppe signifikant von der anderen unterscheidet und somit eine Behandlungsmethode einen Vorteil gegenüber der anderen hat.

Die Auswertestrategie konnte mit einer Ausnahme am Testdatensatz erprobt werden. Ein Vergleich zwischen NYHA-Klassen bzw. Altersgruppen war aufgrund des kleinen Probandenkollektivs nicht möglich.

Der Testdatensatz wurde auch auf widersprüchliche Antworten zwischen KCCQ und EQ-5D-Fragebogen überprüft. Da bei 69% der Probanden widersprüchliche Beantwortungen vorliegen, lässt sich bei Betrachtung der Gesamtdaten nicht erkennen, ob Unterschiede zwischen den Gruppen aufgrund der Behandlungsmethode entstanden sind oder lediglich durch die widersprüchliche Beantwortung der Fragen.

Bei zukünftigen Studien mit ähnlichem Studiendesign sollte darauf geachtet werden, dass sich die Probanden bewusst sind, dass der Wert 1 beim EQ-5D-Fragebogen den besten und beim KCCQ den schlechtesten Gesundheitszustand darstellt. Um solche Irrtümer auszuschließen, steht zwar bei jeder Frage genau beschrieben, welcher Wert welche Bedeutung besitzt, anscheinend ist dies aber nicht ausreichend.

Ein weiterer Ansatz zur Vermeidung solcher Unklarheiten wäre eine Modulation der Fragebögen. So könnte zum Beispiel die Wertigkeit bei einem Fragebogen gedreht werden, sodass die Werte bei beiden Fragebögen demselben Gesundheitszustand entsprechen.

Ein weiterer Grund für die unterschiedliche Beantwortung vergleichbarer Fragen des KCCQ und des EQ-5D Fragebogens könnte im betrachteten Zeitraum liegen. Mithilfe des EQ-5D Fragebogens wird der momentane Gesundheitszustand ermittelt, dagegen fragt der KCCQ nach dem Gesundheitszustand der letzten beiden Wochen. So könnte der Patient durchaus vor 2 Wochen noch extreme körperliche Beschwerden gehabt haben, diese könnten aber im Augenblick der Befragung verschwunden sein.

Solche widersprüchlichen Antworten sollten möglicherweise bei der Befragung direkt mit dem Patienten besprochen werden. Dadurch könnte der Grund vermerkt bzw. der Irrtum behoben werden, noch bevor die Daten endgültig abgespeichert werden.

5.3 Ergebnisse aus den Publikationen

Der Altersdurchschnitt der Herzinsuffizienz-Patienten ist bei fast allen Studien über 65 Jahren angesiedelt. Dies entspricht auch dem Probandenkollektiv des Testdatensatzes. Weiters sind in fast allen Studien mehr Männer vorhanden als Frauen, da die Zahl der Männer, die an Herzinsuffizienz leiden, etwas höher ist als jene der Frauen. Dies spiegelt sich auch im Testdatensatz wider. Das zu untersuchende Studiendesign umfasst keine Patienten der NYHA-Klasse IV. Dies ist auch bei vielen vergleichbaren Studien der Fall. Ein Grund dafür könnte sein, dass die Lebenserwartung mit der NYHA-Klasse drastisch sinkt und es daher weniger Patienten der NYHA-Klasse IV gibt. In manchen Studien wurden Patienten der NYHA-Klasse IV auch bewusst ausgeschlossen, da aufgrund der Lebenserwartung ein Verbleib bis zum Ende der Studie weniger wahrscheinlich ist als bei Patienten einer niedrigeren NYHA-Klasse.

In der Abbildung 4.1 von Spiraki et al. ist gut ersichtlich, dass die Lebensqualität zu jedem Zeitpunkt steigt. Dies ist bei den Ergebnissen aus dem Testdatensatz nicht der Fall. Allerdings liegen die 3 Zeitpunkte bei Spiraki et al. innerhalb eines Monats und beim zu untersuchenden Studiendesign innerhalb eines Jahres. Spiraki et al. erwähnen aber auch, dass die gesundheitsbezogene Lebensqualität eine gute Entscheidungsgrundlage darstellt, um zwischen unterschiedlichen Behandlungsmethoden zu wählen. [20]

Die Ergebnisse von Pinto et al. zeigen, dass Herzinsuffizienz-Patienten zwar Probleme mit der Mobilität und dem Ausüben alltäglicher Tätigkeiten haben, aber dennoch gut in der Lage sind, für sich selbst zu sorgen [21]. Auch beim Testdatensatz wird die Dimension „Für sich selbst sorgen“ des EQ-5D-Fragebogens besser bewertet als die beiden anderen Dimensionen.

Comin-Colet et al. haben auch den KCCQ und den EQ-5D-Fragebogen eingesetzt. Ebenso wie bei den Ergebnissen aus dem Testdatensatz ist auch hier der Index-Wert des EQ-5D-Fragebogens um durchschnittlich 20% besserer als jener der VAS und der „Klinischen Zusammenfassung“ des KCCQ. Die Aufteilung der Probanden zu den Gruppen weicht bei Comin-Colet et al. stark ab, was einen limitierenden Faktor darstellt. Comin-Colet et al. behaupten auch, dass die Auswirkungen einer Behandlungsmethode auf die Lebensqualität auch von der Schwere der Erkrankung abhängig sei. So ist es z.B. leichter, eine Verbesserung der Lebensqualität bei Patienten mit der NYHA-Klasse III zu erzielen als bei Patienten mit der NYHA-Klasse II. Der Grund dafür ist, dass Patienten mit der NYHA-Klasse II ohnehin schon eine bessere Lebensqualität haben und es daher schwerer ist, Verbesserungen zu erzielen, als bei Patienten mit der NYHA-Klasse III. [22]

Jolly et al. haben in ihrer Studie die Auswirkung eines Trainingsprogramms auf den Gesundheitszustand von Herzinsuffizienz-Patienten untersucht. Sie konnten aber diesbezüglich keine Verbesserung gegenüber einer Kontroll-Gruppe feststellen. Davies et al. sind in ihrem Review-Artikel auf ähnliche Ergebnisse gestoßen. Eine Verbesserung konnte nur mit einem sehr hohen Trainingsaufwand erzielt werden. Ein Grund dafür, dass Training den Gesundheitszustand von Herzinsuffizienz-Patienten nicht wesentlich verbessert, ist möglicherweise, dass die Behandlungsmethoden durch Beta-Blocker usw. schon sehr gut

auf die Patienten abgestimmt wurden. Eine Einschränkung der Studie von Jolly et al. ist die Selbstüberwachung des Trainings der Patienten. Es war also nicht möglich voll zu gewährleisten, dass der Patient das Trainingsprogramm ordnungsgemäß vollendet hat. [24], [41]

Die MOBITEL-Studie hat sich ebenso wie das zu untersuchende Studiendesign mit dem Telemonitoring von Herzinsuffizienz-Patienten beschäftigt. Die zwei größten Limitationen der MOBITEL-Studie sind die geringe Anzahl an Probanden und die noch nicht ausgereifte Technik des Telemonitorings. Einen entscheidenden Faktor stellte das richtige Einstellen der Alarm-Parameter dar. Laut Scherr et al. sollten sich zukünftige Studien in dieser Richtung auch mit automatischen Datenanalyse-Algorithmen und einer Verbesserung der Nutzerfreundlichkeit des Telemonitoring-Systems beschäftigen. [25]

Inglis et al. kommen zum Schluss, dass Telemonitoring bei Herzinsuffizienz-Patienten die Lebensqualität verbessert, Kosten, Mortalität und erneute Krankenhausaufenthalte reduziert [42].

Allerdings können nicht alle Studien einen derartig positiven Output bestätigen [29], [26]. Die Ergebnisse von Koehler et al. legen nahe, dass durch telemedizinische Überwachung die Mortalität nicht reduziert werden kann. Allerdings war die Teststärke zu gering, um Unterschiede zwischen den zwei Gruppen der Studie zu detektieren. Eine weitere Einschränkung stellt die fehlende Information über die zurückgewiesenen Patienten dar. Es besteht die Möglichkeit, dass durch die Ausschlusskriterien Patienten zurückgewiesen wurden, bei denen eine telemedizinische Überwachung einen Benefit gebracht hätte. Aufgrund dieser Tatsache und der geringen Teststärke kann das Potential einer telemedizinischen Überwachung als zusätzliche Methode nicht gänzlich ausgeschlossen werden. Weiters stellen sie fest, dass möglicherweise nicht alle Herzinsuffizienz-Patienten gleich gut von einer telemedizinischen Überwachung profitieren und in zukünftigen Studien darauf geachtet werden soll. [26]

Auch Chaudhry et al. kommen zu dem Schluss, dass telemedizinische Überwachung keine Verbesserung bezüglich der Todesrate mit sich bringt. Weiters ist auch laut den Ergebnissen keine Reduzierung der Wiederaufnahmen im Krankenhaus erkenntlich. Trotz der guten methodischen Qualität und des großen Probandenkollektivs gibt es dennoch Einschränkungen der Studie. So könnten weitere Informationen für die behandelnden Personen bezüglich der Patienten die Effektivität der Intervention verbessert haben. Auch die Art der Befragung nimmt einen Einfluss auf die Ergebnisse. Chaudhry et al. haben sich einer automatischen Befragung bedient. Persönlicher Kontakt zu den Patienten bei der Befragung bezüglich des Gesundheitsbefindens usw. könnte andere Ergebnisse liefern. Die Ergebnisse legen weiters nahe, dass eine genaue und individuelle Beurteilung des Krankheitsmanagementsystems vor der Anwendung notwendig ist. Noch erwähnenswert ist die Anzahl von 14% der für die Telemedizin-Gruppe vorgesehenen Personen, die das System nie benutzt haben. [29]

Clark et al. weisen auch darauf hin, dass die Effektivität komplexer Anwendungen, wie z.B. die telemedizinische Überwachung von Herzinsuffizienz-Patienten, sehr stark von der Art und Weise und vom Zusammenhang, in dem sie verwendet werden, abhängt [43].

Eine weitere Möglichkeit, die nicht außer Acht gelassen werden sollte, ist die Einbeziehung von kardiovaskulär implantierbaren elektronischen Geräten in die telemedizinische Überwachung. Da die Anzahl der Patienten steigt, denen elektronische Geräte, wie z.B. Herzschrittmacher, implantiert wurden, können diese auch herangezogen werden, um wertvolle Informationen über Erkrankungen zu liefern. Diese Informationen würden dann z.B. Herzinsuffizienz-Patienten bei der Behandlung zugute kommen. [44]

Myers et al. sind in ihrer Studie zu dem Schluss gekommen, dass telemedizinische Überwachung von Herzinsuffizienz-Patienten die pflegerischen Heimbesuche reduziert, ohne dass die reduzierten Heimbesuche dabei Auswirkungen auf die Krankenhausaufenthalte und Notaufnahmen haben. Insgesamt errechneten sie durch die Reduzierung der pflegerischen Heimbesuche, Kostenersparnisse bei der Telemedizin-Gruppe gegenüber der Normalpflege-Gruppe. Sie beschreiben die Kostenreduktion als abhängig vom Studiendesign, der Komplexität der telemedizinischen Überwachung und weiterer studienabhängiger Variablen. Die Verbesserungen in den Lebensqualitätswerten bei der Telemedizin-Gruppe lassen allerdings keinen Rückschluss zu, ob diese Verbesserungen durch die telemedizinische Überwachung oder durch die Behandlung an sich begründet sind. Um derartige Rückschlüsse tätigen zu können, wäre ein Vergleich zur Normalpflege-Gruppe notwendig. Eine weitere Auffälligkeit waren die erhöhten Hausbesuche bei der Telemedizin-Gruppe in den ersten Wochen der Studie. Diese Hausbesuche waren allerdings nicht pflegerischer Natur, sondern begründet durch Fragen bezüglich der telemedizinischen Gerätschaften. [28]

Seto et al. stellen fest, dass ein großer Vorteil des Telemonitorings darin besteht, dass die Medikation individuell sehr gut angepasst werden kann. Weiters achten Patienten mehr auf negative Veränderungen (z.B. Gewichtszunahme) und können diesen selbst gegensteuern. Einen positiven Effekt auf den Behandlungserfolg könnte auch eine Echt-Zeit-Instruktion mit sich bringen. Wenn zum Beispiel bei einem Patienten Gewichtserhöhung festgestellt wird, soll dieser sofortige Handlungsanweisungen erhalten und nicht erst am Ende des Tages oder der Woche. Eine Einschränkung der Studie könnte darin liegen, dass Patienten im Winter mit der Studie begonnen haben und sie im Sommer abschlossen. Zu diesem Zeitpunkt kann eine Verbesserung der Gesundheit durch einen Saison-Effekt entstehen. Eine weitere Einschränkung war die geringe Anzahl an Patienten. [31]

Boyne et al. schließen aus ihren Ergebnissen, dass die Telemonitoring-Gruppe signifikant weniger Pflege-Besuche benötigt als die Normalpflege-Gruppe. Abgesehen davon gibt es keine signifikanten Unterschiede zwischen den zwei Gruppen. Laut Boyne et al. könnten diese Ergebnisse von der geringeren Aussagekraft der Studie und dem hohen Behandlungsniveau abhängen. Durch die optimale Behandlung beider Gruppen können signifikante Unterschiede nur schwer entstehen. Weiters könnte ein größeres Probandenkollektiv andere und aussagekräftigere Ergebnisse liefern. [32]

Martin-Lesende et al. belegen in ihrer Studie die Anwendbarkeit eines Telemonitoringsystems bei älteren Patienten mit eingeschränkten Computerkenntnissen. Die Ergebnisse zeigen eine positive Auswirkung von Telemonitoring. Die Anzahl der Krankenhauseinweisungen wurde dadurch minimiert. Die Aussagekraft der Studie wird allerdings durch das kleine Probandenkollektiv verringert. [34]

Dendale et al. weisen darauf hin, dass einige ähnliche Studien nur das Körpergewicht telemonitorisch überwachten und diese Information unter Umständen zu wenig sei, um richtige Behandlungsschritte einzuleiten. Dendale et al. haben deswegen auch die Herzrate und den Blutdruck in die Studie inkludiert. Sie konnten eine Verringerung der Todesfälle und der Krankenhausaufenthalte bei der Telemonitoring-Gruppe feststellen. [35]

Dar et al. zeigen, dass Telemonitoring die Früherkennung einer Verschlechterung der Symptome ermöglicht und somit einer Notfallseinweisung entgegenwirkt. Neben Herzinsuffizienz-Symptomen wurden durch das Telemonitoringsystem auch Symptome anderer Erkrankungen, wie z.B. Atemwegserkrankungen, erkannt. Bei der Überwachung von älteren Patienten mit mehreren Komorbiditäten ist dies ein positiver Nebeneffekt. In den Ergebnissen von Dar et al. wird ersichtlich, dass Telemonitoring ähnliche Resultate liefert wie Normalpflege. Dies begründen sie in der hohen Qualität der Normalpflege. Weiters behaupten sie, dass Telemonitoring möglicherweise die Anzahl der Patienten, die ein Arzt behandeln kann, erhöht. [36]

Bei den meisten Studien wird als primäre Einschränkung die geringe Anzahl an Probanden angegeben. Dies erschwert es, statistisch signifikante Ergebnisse festzustellen, und verringert die Aussagekraft der Studie.

6 Zusammenfassung

Die Recherche hat ergeben, dass die am häufigsten verwendeten statistischen Verfahren der T-Test und der Mann-Whitney-U-Test waren, welche in 54% aller betrachteten Publikationen angewendet wurden. Diese beiden Tests wurden auch bei der Auswertestrategie für Vergleiche zwischen zwei Gruppen durchgeführt.

Der T-Test bei verbundener Stichprobe wurde für Vergleiche innerhalb einer Gruppe eingesetzt. Für nicht normalverteilte Daten kam dabei der Wilcoxon-Vorzeichenrang-Test zum Einsatz, welcher unter anderem auch bei Myers et al. [28] angewendet wurde.

Aufgrund der geringen erwarteten Häufigkeiten aus dem Testdatensatz konnte der Chi-Quadrat-Test nicht eingesetzt werden. Stattdessen wurde, wie bei Martin-Lesende et al. [34], der Fisher's-Exact-Test durchgeführt.

Zur Überprüfung auf Normalverteilung der Daten waren der Shapiro-Wilk- und der Kolmogorov-Smirnov-Test vorgesehen. Da der Testdatensatz aus 35 Probanden bestand, wurde der Shapiro-Wilk-Test zur Verteilungsanalyse herangezogen.

Eine Anwendung von ANOVA bzw. Kruskal-Wallis-Test für den Vergleich zwischen Altersgruppen bzw. NYHA-Klassen war wegen des kleinen Probandenkollektivs nicht möglich. Jedoch sollten diese Tests ab einer Gruppengröße von mindestens 3 Personen jedenfalls zur Anwendung kommen. Unter anderem verwenden Pinto et al. [21] die ANOVA für Vergleiche zwischen mehreren Gruppen und bestätigen, dass sich dieses statistische Verfahren für solche Vergleiche eignet.

Es zeigte sich jedoch auch, dass es Widersprüche bei der Beantwortung von KCCQ und EQ-5D-Fragebogen gab. Bei einem Studiendesign mit der gemeinsamen Verwendung dieser Fragebögen, sollte Augenmerk auf das richtige Ausfüllen der Fragebögen gelegt werden.

Insgesamt zeigte die Erprobung der Auswertestrategie am Testdatensatz, dass diese zur Überprüfung der aufgestellten Hypothesen geeignet ist.

7 Literaturverzeichnis

- [1] Statistik Austria: Jahrbuch der Gesundheitsstatistik. Wien. 2013
- [2] Krumholz, H.M., Baker, D.W., Ashton, C.M., Dunbar, S.B., Friesinger, G.C., Havranek, E.P., Hlatky, M.A., Konstam, M., Ordin, D.L., Pina, I.L., Pitt, B., Spertus, J.A.: Evaluating quality of care for patients with heart failure. *Circulation*. 2000, 101: 122-140
- [3] Pschyrembel, W., Dornblüth, O.: *Pschyrembel Klinisches Wörterbuch*. 260. Aufl. Berlin - New York: de Gruyter Verlag (2004)
- [4] Ertl, G., Ruschitzka, F.: The Year in Cardiology 2013: heart failure. *European Heart Journal*. 2014, 34: 470-473
- [5] Szucs, T.D.: Gesundheitsökonomische Aspekte der chronischen Herzinsuffizienz. Teil 1: Krankheitslast und ökonomische Bewertung. *Schweizerische Ärztezeitung*. 2003, 84(46): 2431-2441
- [6] Rector, T.S., Tschumperlin, L.K., Kubo S.H., Bank, A., Francis, G.S., McDonald, K.M., Keeler, C.A., Silver, M.A.: Use of the living with heart failure questionnaire to ascertain patients' perspectives on improvement in quality of life versus risk of drug-induced death. *J Cardiac Failure*. 1995, 1: 201-206
- [7] Juenger, J., Schellberg, D., Kraemer, S., Haunstetter, A., Zugck, C., Herzog, W., Haass, M.: Health related quality of life in patients with congestive heart failure: comparison with other chronic diseases and relation to functional variables. *Heart*. 2002, 87: 235-241
- [8] Polikandrioti, M.: Health failure and health related quality of life. *Health Science Journal*. 2008, 2(3): 119-120
- [9] Menche, N.: *Pflege Heute*. 4. Aufl. München. Urban & Fischer Verlag, Elsevier GmbH (2007)
- [10] Ludwig Boltzmann Gesellschaft: Herzinsuffizienz: Telemedizinische Betreuung von Hochrisiko-Patienten soll Behandlung und Prognose verbessern. 2013, online im Internet: <http://www.lbg.ac.at/en/node/2495> [Abgerufen: 26. Jänner 2014]
- [11] Oemar, M., Janssen, B.: EQ-5D-5L User Guide: Basic information on how to use the EQ-5D-5L instrument (Version 2.0). EuroQol Group 2013

- [12] Herdman, M., Gudex, C., Lloyd, A., Janssen, M.F., Kind, P., Parkin, D., Bonser, G., Badiu, X.: Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res.* 2011, 20: 1727-1736
- [13] Dyer, M.T.D., Goldsmith, K.A., Sharples, L.S., Buxton, M.J.: A review of health utilities using the EQ-5D in studies of cardiovascular disease. *Health and Quality of Life Outcomes.* 2010, 8:13
- [14] Green, C.P., Porter, C.B., Bresnahan, D.R., Spertus, J.A.: Development and Evaluation of the Kansas City Cardiomyopathy Questionnaire: A New Health Status Measure for Heart Failure. *Journal of the American College of Cardiology.* 2000, 35(5): 1245-1255
- [15] Faller, H., Steinbüchel, T., Schowalter, M., Spertus, J.A., Störk, S., Angermann, C.E.: Der Kansas City Cardiomyopathy Questionnaire (KCCQ) – ein neues krankheitsspezifisches Messinstrument zur Erfassung der Lebensqualität bei chronischer Herzinsuffizienz. *Psychother Psych Med.* 2005, 55: 200-208
- [16] Neubauer, S.: *Lebensqualität von Patienten mit Herzinsuffizienz.* Graz, 2013
- [17] Oblinger, M.L.: *Indirekte Verfahren zur Bewertung der Lebensqualität am Beispiel des EQ-5D Fragebogens.* Graz, 2013
- [18] Hartung, J., Elpelt, B., Klöser, K.H.: *Statistik: Lehr- und Handbuch der angewandten Statistik.* 15. Aufl. München: Oldenbourg Wissenschaftsverlag GmbH (2009)
- [19] Brosius, F.: *SPSS 16: Das mitp-Standardwerk.* 1. Aufl. Heidelberg: mitp Verlag (2008)
- [20] Spiraki, C., Kaitelidou, D., Papakonstantinou, V., Prezerakos, P., Maniadakis, N.: Health-related quality of life measurement in patients admitted with coronary heart disease and heart failure to a cardiology department of a secondary urban hospital in Greece. *Hellenic J Cardiol.* 2008, 49: 241-247
- [21] Pinto, E.B., Maso, I., Pereira, J.L.B., Fukuda, T.G., Seixas, J.C., Menezes, D.F., Cincura, C., Neville, I.S., Jesus, P.A.P., Oliveira-Filho, J.: Differential aspects of stroke and congestive heart failure in quality of life reduction: A case series with three comparison groups. *Health and Quality of Life Outcomes.* 2011, 9(65): 1-5
- [22] Comin-Colet, J., Lainscak, M., Dickstein, K., Filippatos, G.S., Johnson, P., Lüscher, T.F., Mori, C., Willenheimer, R., Ponikowski, P., Anker, S.D.: The effect of intravenous ferric carboxymaltose on health-related quality of life in patients with chronic heart failure and iron deficiency: A subanalysis of the FAIR-HF study. *European Heart Journal.* 2013, 34: 30-38

- [23] Anker, S.D., Comin-Colet, J., Filippatos, G., Willenheimer, R., Dickstein, K., Drexler, H., Lüscher, T.F., Bart, B., Banasiak, W., Niegowska, J., Kirwan, B.A., Mori, C., von Eisenhart Rothe, B., Pocock, S.J., Poole-Wilson, P.A., Ponikowski, P.: Ferric carboxymaltose in patients with heart failure and iron deficiency. *New England Journal of Medicine*. 2009, 361: 2436–2448
- [24] Jolly, K., Taylor, R.S., Lip, G.Y.H., Davies, M., Davis, R., Mant, J., Singh, S., Greenfield, S., Ingram, J., Stubley, J., Bryan, S., Stevens, A.: A randomized trial of the addition of home-based exercise to specialist heart failure nurse care: the Birmingham Rehabilitation Uptake Maximisation study for patients with Congestive Heart Failure (BRUM-CHF) study. *European Journal of Heart Failure*, 2009, 11: 205–213
- [25] Scherr, D., Kastner, P., Kollmann, A., Hallas, A., Auer, J., Krappinger, H., Schuchlenz, H., Stark, G., Grander, W., Jakl, G., Schreier, G., Fruhwald, F.M.: Effect of Home-Based Telemonitoring Using Mobile Phone Technology on the Outcome of Heart Failure Patients After an Episode of Acute Decompensation: Randomized Controlled Trial. *J Med Internet Res*. 2009, 11(3): e34
- [26] Koehler, F., Winkler, S., Schieber, M., Sechtem, U., Stangl, S., Böhm, M., Boll, H., Baumann, G., Honold, M., Koehler, K., Gelbrich, G., Kirwan, B.A., and Anker, S.D.: Impact of Remote Telemedical Management on Mortality and Hospitalizations in Ambulatory Patients With Chronic Heart Failure: The Telemedical Interventional Monitoring in Heart Failure Study. *Circulation*. 2011, 123: 1873-1880
- [27] Koehler, F., Winkler, S., Schieber, M., Sechtem, U., Stangl, K., Böhm, M., Boll, H., Kim, S.S., Koehler, K., Lücke, S., Honold, M., Heinze, P., Schweizer, T., Braecklein, M., Kirwan, B.A., Gelbrich, G., Anker, S.D.: Telemedical Interventional Monitoring in Heart Failure (TIM-HF), a randomized, controlled intervention trial investigating the impact of telemedicine on mortality in ambulatory patients with heart failure: study design. *European Journal of Heart Failure*. 2010, 12: 1354–1362
- [28] Myers, S., Grant, R.W., Lugn, N.E., Holbert, B., Kvedar, J.C.: Impact of Home-Based Monitoring on the Care of Patients With Congestive Heart Failure. *Home Health Care Management & Practice*. 2006, 18: 444-451
- [29] Chaudhry, S.I., Mattera, J.A., Curtis, J.P., Spertus, J.A., Herrin, J., Lin, Z., Phillips, C.O., Hodshon, B.V., Cooper, L.S., Krumholz, H.M.: Telemonitoring in patients with heart failure. *New England Journal of Medicine*. 2010, 363: 2301–2309
- [30] Chaudhry, S.I., Barton, B., Mattera, J.A., Spertus, J., Krumholz, H.M.: Randomized trial of Telemonitoring to Improve Heart Failure Outcomes (Tele-HF): study design. *J Card Fail*. 2007,13: 709-714
- [31] Seto, E., Leonard, K.J., Cafazzo, J.A., Barnsley, J., Masino, C., Ross, H.J.: Mobile Phone-Based Telemonitoring for Heart Failure Management: A Randomized Controlled Trial. *J Med Internet Res*. 2012, 14(1): e31

- [32] Boyne, J.J.J., Vrijhoef, H.J.M., Crijns, H.J.G.M., De Weerd, G., Kragten, J., Gorgels, A.P.M.: Tailored telemonitoring in patients with heart failure: results of a multicentre randomized controlled trial. *European Journal of Heart Failure*. 2012, 14: 791–801
- [33] Martín-Lesende, I., Orruño, E., Cairo, C., Bilbao, A., Asua, J., Romo, M.I., Vergara, I., Bayón, J.C., Abad, R., Reviriego, E., Larrañaga, J.: Assessment of a primary care-based telemonitoring intervention for home care patients with heart failure and chronic lung disease. The TELBIL study. *BMC Health Services Research*. 2011, 11: 56
- [34] Martín-Lesende, I., Orruño, E., Bilbao, A., Vergara, I., Cairo, C., Bayón, J.C., Reviriego, E., Romo, M.I., Larrañaga, J., Asua, J., Abad, R., Recalde, E.: Impact of telemonitoring home care patients with heart failure or chronic lung disease from primary care on healthcare resource use (the TELBIL study randomised controlled trial). *BMC Health Services Research*. 2013, 13: 118
- [35] Dendale, P., De Keulenaer, G., Troisfontaines, P., Weytjens, C., Mullens, W., Elegeert, I., Ector, B., Houbrechts, M., Willekens, K., Hansen, D.: Effect of a telemonitoring-facilitated collaboration between general practitioner and heart failure clinic on mortality and rehospitalization rates in severe heart failure: the TEMA-HF 1 (Telemonitoring in the Management of Heart Failure) study. *European Journal of Heart Failure*. 2012, 14: 333-340
- [36] Dar, O., Riley, J., Chapman, C., Dubrey, S.W., Morris, S., Rosen, S.D., Roughton, M., Cowie, M.R.: A randomized trial of home telemonitoring in a typical elderly heart failure population in North West London: results of the Home-HF study. *European Journal of Heart Failure*. 2009, 11: 319–325
- [37] D’Agostino, R.B., Belanger, A., D’Agostino, R.B.Jr.: A Suggestion for Using Powerful and Informative Tests of Normality. *The American Statistician*. 1990, 44(4): 316-321
- [38] Shapiro, S.S., Wilk, M.B.: An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*. 1965, 52(3/4): 591-611
- [39] Seier, E.: Comparison of Tests for Univariate Normality. East Tennessee State University. 2002
- [40] Rahman, M.M., Govindarajulu, Z.: A modification of the test of Shapiro and Wilk for normality. *Journal of Applied Statistics*. 1997, 24(2): 219-236
- [41] Davies, E.J., Moxham, T., Rees, K., Singh, S., Coats, A.J.S., Ebrahim, S., Lough, F., Taylor, R.S.: Exercise based rehabilitation for heart failure. *Cochrane Database of Systematic Reviews* 2010, Issue 4. Art. No.: CD003331. DOI: 10.1002/14651858.CD003-331.pub3.
- [42] Inglis, S.C., Clark, R.A., McAlister, F.A., Ball, J., Lewinter, C., Cullington, D., Stewart, S., Cleland, J.G.F.: Structured telephone support or telemonitoring programmes

for patients with chronic heart failure. Cochrane Database of Systematic Reviews 2010, Issue 8. Art. No.: CD007228. DOI: 10.1002/14651858.CD007228.pub2.

- [43] Clark, A.M., Savard, L.A., Thompson, D.R.: What is the strength of evidence for heart failure disease-management programs? *J Am Coll Cardiol.* 2009, 54: 397-401
- [44] Wilkoff, B.L., Auricchio, A., Brugada, J., Cowie, M., Ellenbogen, K.A., Gillis, A.M., Hayes, D.L., Howlett, J.G., Kautzner, J., Love, C.J., Morgan, J.M., Priori, S.G., Reynolds, D.W., Schoenfeld, M.H, Vardas, P.E.: HRS/EHRA Expert Consensus on the Monitoring of Cardiovascular Implantable Electronic Devices (CIEDs): Description of Techniques, Indications, Personnel, Frequency and Ethical Considerations. *Europace.* 2008, 10: 707–725

8 Anhang

Parameter	Zeitpunkt 1		Zeitpunkt 2		Zeitpunkt 3	
	Gruppe A	Gruppe B	Gruppe A	Gruppe B	Gruppe A	Gruppe B
Gesamt	18	17	10	12	7	6
Alter	18	17				
Geschlecht						
Frauen	3	3	2	2	1	1
Männer	15	14	8	10	6	5
NYHA Klasse	14	16	10	12	7	6
I	5	5	5	4	3	1
II	4	5	1	5	1	3
III	5	6	4	3	3	2
IV	0	0	0	0	0	0
EQ5D	18	17	10	12	7	6
VAS	18	17	10	12	7	6
KCCQ	18	17	10	12	7	6

Tabelle 8.1: Darstellung der Anzahl der Probanden zu den 3 Unterschiedlichen Zeitpunkten der Studie.

8 Anhang

Parameter	Zeitpunkt 1	Zeitpunkt 2	Zeitpunkt 3
<u>Probanden bis Zeitpunkt 1:</u>			
Alter			
Gruppe A	0,017	-	-
Gruppe B	0,722	-	-
EQ5D (UK/D/Euro)			
Gruppe A	0,004 / 0,000 / 0,031	-	-
Gruppe B	0,161 / 0,164 / 0,064	-	-
VAS			
Gruppe A	0,143	-	-
Gruppe B	0,006	-	-
KCCQ (CSS/OSS)			
Gruppe A	0,418 / 0,110	-	-
Gruppe B	0,274 / 0,611	-	-
<u>Probanden bis Zeitpunkt 2:</u>			
Alter			
Gruppe A	0,090	-	-
Gruppe B	0,901	-	-
EQ5D (UK/D/Euro)			
Gruppe A	0,061 / 0,017 / 0,121	0,013 / 0,006 / 0,016	-
Gruppe B	0,491 / 0,092 / 0,410	0,016 / 0,015 / 0,013	-
VAS			
Gruppe A	0,063	0,081	-
Gruppe B	0,010	0,028	-
KCCQ (CSS/OSS)			
Gruppe A	0,788 / 0,759	0,025 / 0,054	-
Gruppe B	0,218 / 0,538	0,257 / 0,029	-
<u>Probanden bis Zeitpunkt 3:</u>			
Alter			
Gruppe A	0,012	-	-
Gruppe B	0,927	-	-
EQ5D (UK/D/Euro)			
Gruppe A	0,083 / 0,030 / 0,194	0,010 / 0,004 / 0,013	0,032 / 0,010 / 0,098
Gruppe B	0,728 / 0,519 / 0,426	0,162 / 0,367 / 0,133	0,770 / 0,881 / 0,494
VAS			
Gruppe A	0,460	0,005	0,349
Gruppe B	0,021	0,137	0,957
KCCQ (CSS/OSS)			
Gruppe A	0,613 / 0,732	0,045 / 0,027	0,089 / 0,245
Gruppe B	0,418 / 0,280	0,308 / 0,081	0,382 / 0,267

Tabelle 8.2: p-Werte des Shapiro-Wilk-Tests auf Normalverteilung, für alle Probanden.

Parameter	Gruppe A	Gruppe B	p-Wert
EQ5D			
UK-Index	0,769±0,357 / 1,0	0,784 ±0,177 / 0,735	0,652
D-Index	0,839±0,271 / 1,0	0,88±0,102 / 0,859	0,652
Euro-Index	0,776±0,358 / 1,0	0,76±0,194 / 0,694	0,652
VAS	78,57±22,49 / 90	59,16±19,08 / 50	0,109
KCCQ:			
Funktionaler Status	78,15±29,12 / 94	69,62±25,31 / 63,5	0,508
Klinische Zusammenfassung	81,93±23,73 / 91,3	74,65±16,52 / 68,2	0,366

Tabelle 8.3: Vergleich der Werte zu Studien-Zeitpunkt 2 aller Probanden des Testdatensatzes, die die Fragebögen für alle vorgesehenen Zeitpunkte ausgefüllt haben. Angaben: Mittelwert ± Standardabweichung / Median.

Parameter	Gruppe A	Gruppe B	p-Wert
EQ5D			
UK-Index	0,578±0,499 / 0,778	0,645±0,227 / 0,632	0,667
D-Index	0,692±0,402 / 0,91	0,779±0,159 / 0,774	0,667
Euro-Index	0,638±0,378 / 0,727	0,634±0,226 / 0,608	0,982
VAS	64,14±25,5 / 50	62,5±17,23 / 62,5	0,896
KCCQ:			
Funktionaler Status	66,59±36,62 / 74,5	60,85±21,78 / 54,7	0,744
Klinische Zusammenfassung	77,12±24,51 / 83,1	69,75±15,29 / 67,7	0,538

Tabelle 8.4: Vergleich der Werte zu Studien-Zeitpunkt 3 aller Probanden des Testdatensatzes, die die Fragebögen für alle vorgesehenen Zeitpunkte ausgefüllt haben. Angaben: Mittelwert ± Standardabweichung / Median.

Parameter	Zeitpunkt 1	Zeitpunkt 2	p-Wert
EQ5D			
UK-Index	0,66±0,361 / 0,74	0,769±0,357 / 1,0	0,225
D-Index	0,762±0,277 / 0,845	0,839±0,271 / 1,0	0,144
Euro-Index	0,667±0,317 / 0,694	0,776±0,358 / 1,0	0,225
VAS	59,29±19,67 / 60	78,57±22,49 / 90	0,034
KCCQ:			
Funktionaler Status	48,29 ± 16,92 / 51	78,15±29,12 / 94	0,063
Klinische Zusammenfassung	54,13 ± 15,27 / 52,9	81,93±23,73 / 91,3	0,028

Tabelle 8.5: Vergleich der Werte zwischen Zeitpunkt 1 und Zeitpunkt 2 bei Gruppe A, für alle Probanden des Testdatensatzes, die die Fragebögen für alle vorgesehenen Zeitpunkte ausgefüllt haben. Angaben: Mittelwert ± Standardabweichung / Median.

Parameter	Zeitpunkt 1	Zeitpunkt 2	p-Wert
EQ5D			
UK-Index	0,747±0,174 / 0,753	0,784 ±0,177 / 0,735	0,422
D-Index	0,852±0,124 / 0,866	0,88±0,102 / 0,859	0,394
Euro-Index	0,719±0,178 / 0,721	0,76±0,194 / 0,694	0,448
VAS	63,33±20,41 / 52,5	59,16±19,08 / 50	0,465
KCCQ:			
Funktionaler Status	56,08 ± 21,31 / 53,4	69,62±25,31 / 63,5	0,107
Klinische Zusammenfassung	63,02 ± 18,06 / 56,4	74,65±16,52 / 68,2	0,079

Tabelle 8.6: Vergleich der Werte zwischen Zeitpunkt 1 und Zeitpunkt 2 bei Gruppe B, für alle Probanden des Testdatensatzes, die die Fragebögen für alle vorgesehenen Zeitpunkte ausgefüllt haben. Angaben: Mittelwert ± Standardabweichung / Median.

Parameter	Zeitpunkt 2	Zeitpunkt 3	p-Wert
EQ5D			
UK-Index	0,769±0,357 / 1,0	0,578±0,499 / 0,778	0,043
D-Index	0,839±0,271 / 1,0	0,692±0,402 / 0,91	0,043
Euro-Index	0,776±0,358 / 1,0	0,638±0,378 / 0,727	0,043
VAS	78,57±22,49 / 90	64,14±25,5 / 50	0,172
KCCQ:			
Funktionaler Status	78,15±29,12 / 94	66,59±36,62 / 74,5	0,463
Klinische Zusammenfassung	81,93±23,73 / 91,3	77,12±24,51 / 83,1	1,000

Tabelle 8.7: Vergleich der Werte zwischen Zeitpunkt 2 und Zeitpunkt 3 bei Gruppe A, für alle Probanden des Testdatensatzes, die die Fragebögen für alle vorgesehenen Zeitpunkte ausgefüllt haben. Angaben: Mittelwert ± Standardabweichung / Median.

Parameter	Zeitpunkt 2	Zeitpunkt 3	p-Wert
EQ5D			
UK-Index	0,784 ±0,177 / 0,735	0,645±0,227 / 0,632	0,191
D-Index	0,88±0,102 / 0,859	0,779±0,159 / 0,774	0,157
Euro-Index	0,76±0,194 / 0,694	0,634±0,226 / 0,608	0,251
VAS	59,16±19,08 / 50	62,5±17,23 / 62,5	0,444
KCCQ:			
Funktionaler Status	69,62±25,31 / 63,5	60,85±21,78 / 54,7	0,295
Klinische Zusammenfassung	74,65±16,52 / 68,2	69,75±15,29 / 67,7	0,403

Tabelle 8.8: Vergleich der Werte zwischen Zeitpunkt 2 und Zeitpunkt 3 bei Gruppe B, für alle Probanden des Testdatensatzes, die die Fragebögen für alle vorgesehenen Zeitpunkte ausgefüllt haben. Angaben: Mittelwert ± Standardabweichung / Median.

Parameter	Zeitpunkt 1		Zeitpunkt 2		Zeitpunkt 3	
	Gruppe A	Gruppe B	Gruppe A	Gruppe B	Gruppe A	Gruppe B
Gesamt	3	8	1	4	1	2
Alter	3	8				
Geschlecht						
Frauen	1	1	0	0	0	0
Männer	2	7	1	4	1	2
NYHA Klasse	3	7	1	4	1	2
I	0	3	0	2	0	1
II	2	1	1	1	1	0
III	1	3	0	1	0	1
IV	0	0	0	0	0	0
EQ5D	3	8	1	4	1	2
VAS	3	8	1	4	1	2
KCCQ	3	8	1	4	1	2

Tabelle 8.9: Darstellung der Anzahl der Probanden zu den 3 Unterschiedlichen Zeitpunkten nach Datenbereinigung.

Parameter	Gruppe A	Gruppe B	p-Wert
EQ5D			
UK-Index	0,717±0,282 / 0,877	0,854±0,199 / 0,939	0,346
D-Index	0,828±0,23 / 0,918	0,901±0,139 / 0,959	0,346
Euro-Index	0,694±0,26 / 0,841	0,846±0,294 / 0,921	0,346
VAS	66,67±15,28 / 70	73,75±22,64 / 65	0,754
KCCQ:			
Funktionaler Status	48,26 ±32,88 / 64,6	72,85±20,9 / 78,1	0,166
Klinische Zusammenfassung	57,12 ± 23,32 / 69,8	75,94±13,04 / 79,4	0,114

Tabelle 8.10: Vergleich der Werte zwischen den Gruppen für alle Probanden zu Zeitpunkt 1 nach Datenbereinigung. Angaben: Mittelwert ± Standardabweichung / Median.

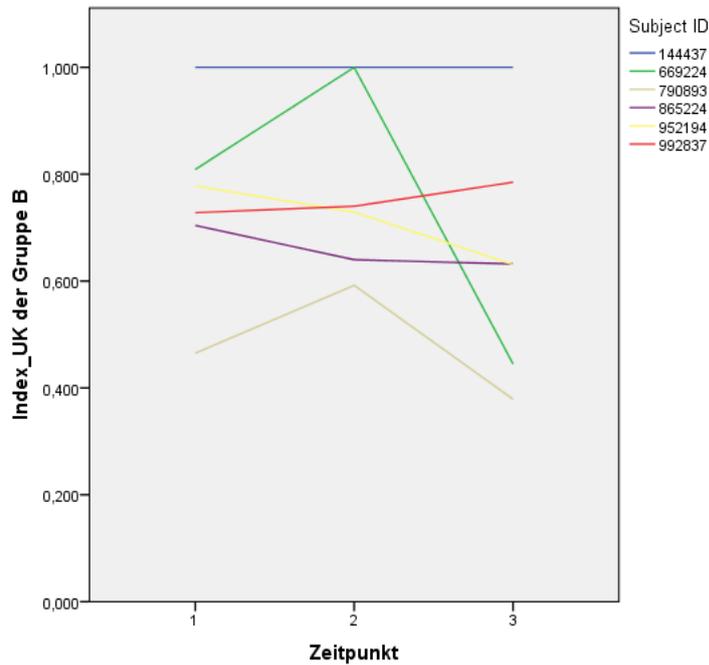


Abbildung 8.1: Verlauf der EQ-5D-Index-UK-Werte der einzelnen Probanden für Gruppe B.

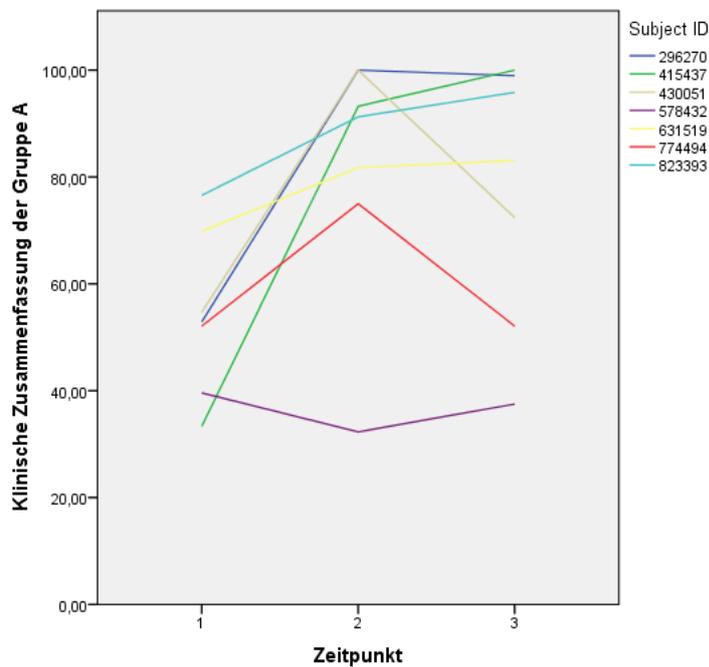


Abbildung 8.2: Verlauf der der „Klinischen Zusammenfassung“ des KCCQ der einzelnen Probanden für Gruppe A.

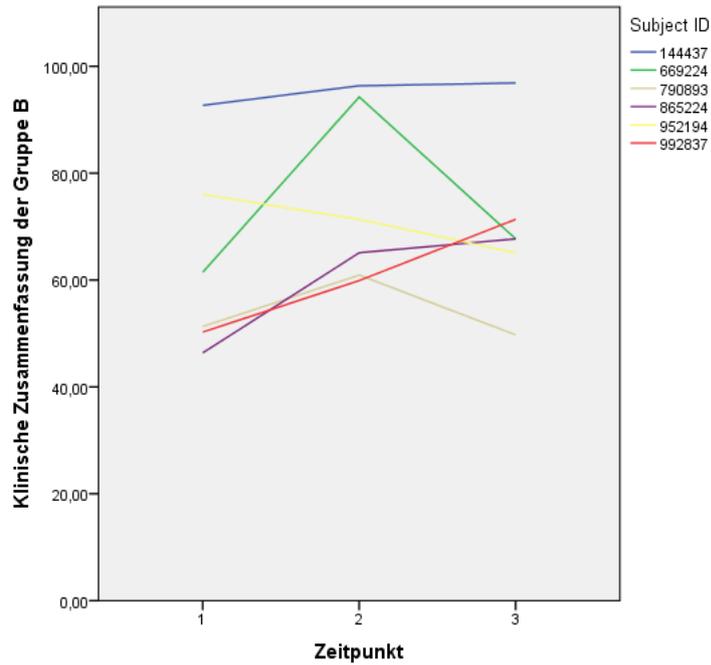


Abbildung 8.3: Verlauf der der „Klinischen Zusammenfassung“ des KCCQ der einzelnen Probanden für Gruppe B.

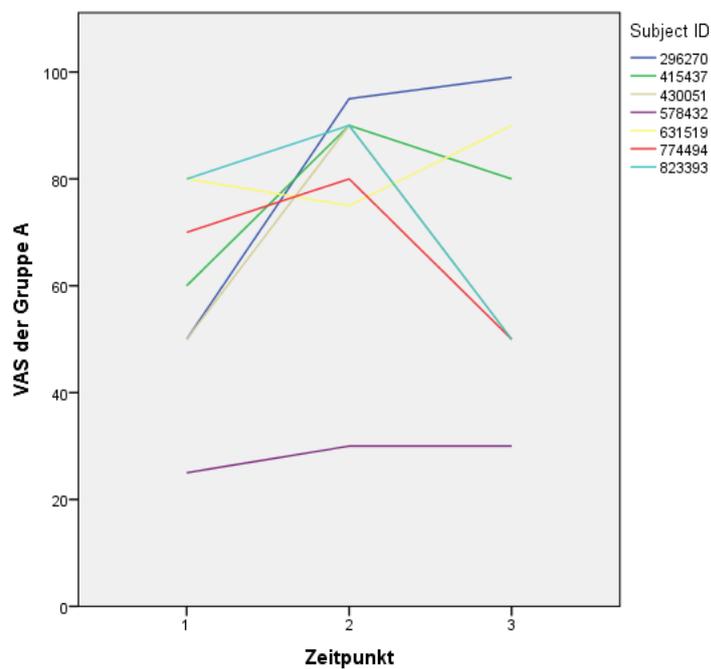


Abbildung 8.4: Verlauf der EQ-5D-VAS-Werte der einzelnen Probanden für Gruppe A.

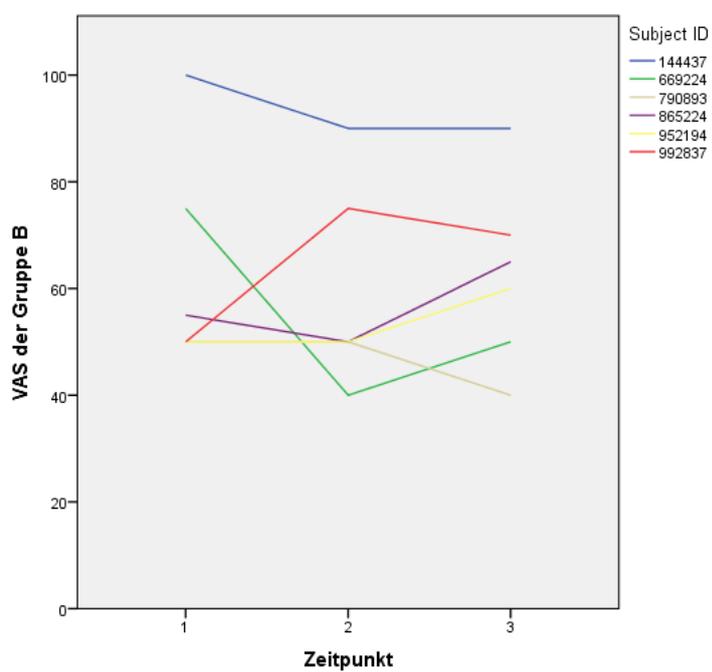


Abbildung 8.5: Verlauf der EQ-5D-VAS-Werte der einzelnen Probanden für Gruppe B.

8 Anhang

* %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Alle Beispiele sind auch für weitere Unterteilungen anwendbar (z.B. Altersgruppen, NYHA-Klassen)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

* %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Filterung von widersprüchlichen Daten
Widersprüchliche Daten werden entfernt
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

USE ALL.
SELECT IF ((BEWEG = 1 & (KCCQ1_K13 = 3 | KCCQ1_K13 = 4 | KCCQ1_K13 = 5)) | (BEWEG = 2 &
(KCCQ1_K13 = 3 | KCCQ1_K13 = 4 | KCCQ1_K13 = 5)) | (BEWEG = 3) | (BEWEG = 4 & (KCCQ1_K13 = 3
| KCCQ1_K13 = 2 | KCCQ1_K13 = 1)) | (BEWEG = 5 & (KCCQ1_K13 = 3 |
KCCQ1_K13 = 2 | KCCQ1_K13 = 1))).
EXECUTE.

USE ALL.
SELECT IF ((SORGEN = 1 & (KCCQ1_K11 = 3 | KCCQ1_K11 = 4 | KCCQ1_K11 = 5)) | (SORGEN = 2 &
(KCCQ1_K11 = 3 | KCCQ1_K11 = 4 | KCCQ1_K11 = 5)) | (SORGEN = 3) | (SORGEN = 4 & (KCCQ1_K11 =
3 | KCCQ1_K11 = 2 | KCCQ1_K11 = 1)) | (SORGEN = 5 & (KCCQ1_K11 =
3 | KCCQ1_K11 = 2 | KCCQ1_K11 = 1))).
EXECUTE.

USE ALL.
SELECT IF ((SORGEN = 1 & (KCCQ1_K12 = 3 | KCCQ1_K12 = 4 | KCCQ1_K12 = 5)) | (SORGEN = 2 &
(KCCQ1_K12 = 3 | KCCQ1_K12 = 4 | KCCQ1_K12 = 5)) | (SORGEN = 3) | (SORGEN = 4 & (KCCQ1_K12 =
3 | KCCQ1_K12 = 2 | KCCQ1_K12 = 1)) | (SORGEN = 5 & (KCCQ1_K12 =
3 | KCCQ1_K12 = 2 | KCCQ1_K12 = 1))).
EXECUTE.

USE ALL.
SELECT IF ((ALLTAG = 1 & (KCCQ1_K14 = 3 | KCCQ1_K14 = 4 | KCCQ1_K14 = 5)) | (ALLTAG = 2 &
(KCCQ1_K14 = 3 | KCCQ1_K14 = 4 | KCCQ1_K14 = 5)) | (ALLTAG = 3) | (ALLTAG = 4 & (KCCQ1_K14 = 3
| KCCQ1_K14 = 2 | KCCQ1_K14 = 1)) | (ALLTAG = 5 & (KCCQ1_K14 =
3 | KCCQ1_K14 = 2 | KCCQ1_K14 = 1))).
EXECUTE.

USE ALL.
SELECT IF ((SCHMERZ = 1 & (KCCQ4 = 3 | KCCQ4 = 4 | KCCQ4 = 5 | KCCQ4 = 6)) | (SCHMERZ = 2 &
(KCCQ4 = 3 | KCCQ4 = 4 | KCCQ4 = 5 | KCCQ4 = 6)) | (SCHMERZ = 3) | (SCHMERZ = 4) | (SCHMERZ =
5)).
EXECUTE.

USE ALL.
SELECT IF ((SCHMERZ = 1 & (KCCQ6 = 3 | KCCQ6 = 4 | KCCQ6 = 5 | KCCQ6 = 6)) | (SCHMERZ = 2 &
(KCCQ6 = 3 | KCCQ6 = 4 | KCCQ6 = 5 | KCCQ6 = 6)) | (SCHMERZ = 3) | (SCHMERZ = 4) | (SCHMERZ =
5)).
EXECUTE.

8 Anhang

USE ALL.

```
SELECT IF ((SCHMERZ = 1 & (KCCQ8 = 3 | KCCQ8 = 4 | KCCQ8 = 5 | KCCQ8 = 6)) | (SCHMERZ = 2 &
(KCCQ8 = 3 | KCCQ8 = 4 | KCCQ8 = 5 | KCCQ8 = 6)) | (SCHMERZ = 3) | (SCHMERZ = 4) | (SCHMERZ =
5)).
```

EXECUTE.

USE ALL.

```
SELECT IF ((ANGST = 1 & (KCCQ14 = 3 | KCCQ14 = 4 | KCCQ14 = 5)) | (ANGST = 2 & (KCCQ14 = 3 |
KCCQ14 = 4 | KCCQ14 = 5)) | (ANGST = 3) | (ANGST = 4 & (KCCQ14 = 3 | KCCQ14 = 2 | KCCQ14 = 1)) |
(ANGST = 5 & (KCCQ14 = 3 | KCCQ14 = 2 | KCCQ14 = 1))).
```

EXECUTE.

* %%%%%%%%%%
Boxplots für Vergleiche zwischen Gruppen und zwischen Zeitpunkten
VAS als Beispiel
%%%%%%%%%

* Chart Builder.

```
GGRAPH
  /GRAPHDATASET NAME="graphdataset" VARIABLES=Zeitpunkt GESUNDHEIT RM_GROUP_E1_C16
MISSING=LISTWISE REPORTMISSING=NO
  /GRAPHSPEC SOURCE=INLINE.
BEGIN GPL
SOURCE: s=userSource(id("graphdataset"))
DATA: Zeitpunkt=col(source(s), name("Zeitpunkt"), unit.category())
DATA: GESUNDHEIT=col(source(s), name("GESUNDHEIT"))
DATA: RM_GROUP_E1_C16=col(source(s), name("RM_GROUP_E1_C16"), unit.category())
DATA: id=col(source(s), name("$CASENUM"), unit.category())
COORD: rect(dim(1,2), cluster(3,0))
GUIDE: axis(dim(3), label("Zeitpunkt"))
GUIDE: axis(dim(2), label("VAS"))
GUIDE: legend(aesthetic(aesthetic.color), label("Gruppe"))
SCALE: linear(dim(2), include(0))
ELEMENT: schema(position(bin.quantile.letter(RM_GROUP_E1_C16*GESUNDHEIT*Zeitpunkt)),
color(RM_GROUP_E1_C16), label(id))
END GPL.
```

* %%%%%%%%%%
Boxplots für Vergleiche zwischen Gruppen
Filter für den gewünschten Zeitpunkt muss noch eingestellt werden
VAS als Beispiel
%%%%%%%%%

* Chart Builder.

```
GGRAPH
  /GRAPHDATASET NAME="graphdataset" VARIABLES=RM_GROUP_E1_C16 GESUNDHEIT
MISSING=LISTWISE REPORTMISSING=NO
  /GRAPHSPEC SOURCE=INLINE.
BEGIN GPL
SOURCE: s=userSource(id("graphdataset"))
DATA: RM_GROUP_E1_C16=col(source(s), name("RM_GROUP_E1_C16"), unit.category())
```

8 Anhang

```
DATA: GESUNDHEIT=col(source(s), name("GESUNDHEIT"))
DATA: id=col(source(s), name("$CASENUM"), unit.category())
GUIDE: axis(dim(1), label("Gruppe"))
GUIDE: axis(dim(2), label("VAS"))
SCALE: linear(dim(2), include(0))
ELEMENT: schema(position(bin.quantile.letter(RM_GROUP_E1_C16*GESUNDHEIT)), label(id))
END GPL.
```

* %%%%%%%%%%%
Liniendiagramme Medianvergleich zwischen Gruppen und Zwischen Zeitpunkten
VAS als Beispiel
%%%%%%%%%%

* Chart Builder.

```
GGRAPH
  /GRAPHDATASET NAME="graphdataset" VARIABLES=Zeitpunkt
MEDIAN(GESUNDHEIT)[name="MEDIAN_GESUNDHEIT"] RM_GROUP_E1_C16 MISSING=LISTWISE
REPORTMISSING=NO
  /GRAPHSPEC SOURCE=INLINE.
BEGIN GPL
SOURCE: s=userSource(id("graphdataset"))
DATA: Zeitpunkt=col(source(s), name("Zeitpunkt"), unit.category())
DATA: MEAN_GESUNDHEIT=col(source(s), name("MEDIAN_GESUNDHEIT"))
DATA: RM_GROUP_E1_C16=col(source(s), name("RM_GROUP_E1_C16"), unit.category())
GUIDE: axis(dim(1), label("Zeitpunkt"))
GUIDE: axis(dim(2), label("Median VAS"))
GUIDE: legend(aesthetic(aesthetic.color.interior), label("Gruppe"))
SCALE: linear(dim(2), include(0))
ELEMENT: line(position(Zeitpunkt*MEDIAN_GESUNDHEIT), color.interior(RM_GROUP_E1_C16),
missing.wings())
END GPL.
```

* %%%%%%%%%%%
Liniendiagramme Mittelwertvergleich zwischen Gruppen und Zwischen Zeitpunkten
VAS als Beispiel
%%%%%%%%%%

* Chart Builder.

```
GGRAPH
  /GRAPHDATASET NAME="graphdataset" VARIABLES=Zeitpunkt
MEAN(GESUNDHEIT)[name="MEAN_GESUNDHEIT"] RM_GROUP_E1_C16 MISSING=LISTWISE
REPORTMISSING=NO
  /GRAPHSPEC SOURCE=INLINE.
BEGIN GPL
SOURCE: s=userSource(id("graphdataset"))
DATA: Zeitpunkt=col(source(s), name("Zeitpunkt"), unit.category())
DATA: MEAN_GESUNDHEIT=col(source(s), name("MEAN_GESUNDHEIT"))
DATA: RM_GROUP_E1_C16=col(source(s), name("RM_GROUP_E1_C16"), unit.category())
GUIDE: axis(dim(1), label("Zeitpunkt"))
GUIDE: axis(dim(2), label("Mean VAS"))
GUIDE: legend(aesthetic(aesthetic.color.interior), label("Gruppe"))
```

8 Anhang

```
SCALE: linear(dim(2), include(0))
ELEMENT: line(position(Zeitpunkt*MEAN_GESUNDHEIT), color.interior(RM_GROUP_E1_C16),
missing.wings())
END GPL.
```

* %%%%%%%%%%
Linienendiagramme alle Probanden einzeln
VAS als Beispiel
%%%%%%%%%

```
* Chart Builder.
GGRAPH
  /GRAPHDATASET NAME="graphdataset" VARIABLES=Zeitpunkt
  MEAN(GESUNDHEIT)[name="MEAN_GESUNDHEIT"] SID_EQ5D MISSING=LISTWISE
  REPORTMISSING=NO
  /GRAPHSPEC SOURCE=INLINE.
BEGIN GPL
  SOURCE: s=userSource(id("graphdataset"))
  DATA: Zeitpunkt=col(source(s), name("Zeitpunkt"), unit.category())
  DATA: MEAN_GESUNDHEIT=col(source(s), name("MEAN_GESUNDHEIT"))
  DATA: SID_EQ5D=col(source(s), name("SID_EQ5D"), unit.category())
  GUIDE: axis(dim(1), label("Zeitpunkt"))
  GUIDE: axis(dim(2), label("VAS"))
  GUIDE: legend(aesthetic(aesthetic.color.interior), label("Subject ID"))
  SCALE: linear(dim(2), include(0))
  ELEMENT: line(position(Zeitpunkt*MEAN_GESUNDHEIT), color.interior(SID_EQ5D), missing.wings())
END GPL.
```

* %%%%%%%%%%
Deskriptive Analyse von metrisch skalierten Variablen
inklusive Verteilungsanalyse
Gruppe A zu Zeitpunkt 1 als Beispiel
%%%%%%%%%

```
USE ALL.
COMPUTE filter_$=(RM_GROUP_E1_C16 = 1 & Zeitpunkt = 1).
VARIABLE LABELS filter_$ 'RM_GROUP_E1_C16 = 1 & Zeitpunkt = 1 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
EXAMINE VARIABLES=Age_E1 GESUNDHEIT Index_UK Index_D Index_Euro PL Symptome LQ SL
Sym_Stab Selbst_W CSS OSS
  /PLOT HISTOGRAM NPLOT
  /STATISTICS DESCRIPTIVES
  /INTERVAL 95
  /MISSING LISTWISE
  /NOTOTAL.
```

8 Anhang

* %%%%%%%%%%%%%
Deskriptive Analyse von ordinalskalierten Variablen
Gruppe A zu Zeitpunkt 1 als Beispiel
%%%%%%%%%

```
USE ALL.  
COMPUTE filter_$=(RM_GROUP_E1_C16 = 1 & Zeitpunkt = 1).  
VARIABLE LABELS filter_$ 'RM_GROUP_E1_C16 = 1 & Zeitpunkt = 1 (FILTER)'.  
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.  
FORMATS filter_$ (f1.0).  
FILTER BY filter_$.  
EXECUTE.  
FREQUENCIES VARIABLES=NYHA_CLASS_E3_C26 BEWEG SORGEN ALLTAG SCHMERZ ANGST  
KCCQ1_K11 KCCQ1_K12 KCCQ1_K13 KCCQ1_K14 KCCQ1_K15 KCCQ1_K16 KCCQ2 KCCQ3 KCCQ4  
KCCQ5 KCCQ6 KCCQ7 KCCQ8 KCCQ9 KCCQ10 KCCQ11 KCCQ12 KCCQ13 KCCQ14 KCCQ15_K151  
KCCQ15_K152 KCCQ15_K153  
KCCQ15_K154  
/STATISTICS=MEDIAN  
/HISTOGRAM  
/ORDER=ANALYSIS.
```

* %%%%%%%%%%%%%
Vergleich zwischen Zeitpunkten innerhalb einer Gruppe
für alle 3 Studien-Zeitpunkte
%%%%%%%%%

```
NPAR TESTS  
/WILCOXON=GESUNDHEIT1_B GESUNDHEIT1_B GESUNDHEIT2_B GESUNDHEIT1_A  
GESUNDHEIT1_A GESUNDHEIT2_A Index_UK1_B Index_UK1_B Index_UK2_B Index_UK1_A  
Index_UK1_A Index_UK2_A Index_D1_B Index_D1_B Index_D2_B Index_D1_A Index_D1_A  
Index_D2_A Index_Euro1_B  
Index_Euro1_B Index_Euro2_B Index_Euro1_A Index_Euro1_A Index_Euro2_A CSS1_B CSS1_B  
CSS2_B CSS1_A CSS1_A CSS2_A OSS1_B OSS1_B OSS2_B OSS1_A OSS1_A OSS2_A WITH  
GESUNDHEIT3_B GESUNDHEIT2_B GESUNDHEIT3_B GESUNDHEIT3_A GESUNDHEIT2_A  
GESUNDHEIT3_A Index_UK3_B  
Index_UK2_B Index_UK3_B Index_UK3_A Index_UK2_A Index_UK3_A Index_D3_B Index_D2_B  
Index_D3_B Index_D3_A Index_D2_A Index_D3_A Index_Euro3_B Index_Euro2_B Index_Euro3_B  
Index_Euro3_A Index_Euro2_A Index_Euro3_A CSS3_B CSS2_B CSS3_B CSS3_A CSS2_A CSS3_A  
OSS3_B OSS2_B OSS3_B OSS3_A OSS2_A OSS3_A (PAIRED)  
/MISSING ANALYSIS.
```

```
T-TEST PAIRS=GESUNDHEIT1_B GESUNDHEIT1_B GESUNDHEIT2_B GESUNDHEIT1_A GESUNDHEIT1_A  
GESUNDHEIT2_A Index_UK1_B Index_UK1_B Index_UK2_B Index_UK1_A Index_UK1_A Index_UK2_A  
Index_D1_B Index_D1_B Index_D2_B Index_D1_A Index_D1_A Index_D2_A Index_Euro1_B  
Index_Euro1_B Index_Euro2_B Index_Euro1_A Index_Euro1_A Index_Euro2_A CSS1_B CSS1_B  
CSS2_B CSS1_A CSS1_A CSS2_A OSS1_B OSS1_B OSS2_B OSS1_A OSS1_A OSS2_A WITH  
GESUNDHEIT3_B GESUNDHEIT2_B GESUNDHEIT3_B GESUNDHEIT3_A GESUNDHEIT2_A  
GESUNDHEIT3_A Index_UK3_B  
Index_UK2_B Index_UK3_B Index_UK3_A Index_UK2_A Index_UK3_A Index_D3_B Index_D2_B  
Index_D3_B Index_D3_A Index_D2_A Index_D3_A Index_Euro3_B Index_Euro2_B Index_Euro3_B  
Index_Euro3_A Index_Euro2_A Index_Euro3_A CSS3_B CSS2_B CSS3_B CSS3_A CSS2_A CSS3_A
```

8 Anhang

```
OSS3_B OSS2_B OSS3_B OSS3_A OSS2_A OSS3_A (PAIRED)
/CRITERIA=C1(.9500)
/MISSING=ANALYSIS.
```

* %%%%%%%%%%
Vergleich zwischen Zeitpunkten innerhalb einer Gruppe
für 2 Studien-Zeitpunkte
%%%%%%%%%

```
NPART TESTS
/WILCOXON=GESUNDHEIT1_B GESUNDHEIT1_A Index_UK1_B Index_UK1_A Index_D1_B
Index_D1_A Index_Euro1_B Index_Euro1_A CSS1_B CSS1_A OSS1_B OSS1_A WITH GESUNDHEIT2_B
GESUNDHEIT2_A Index_UK2_B Index_UK2_A Index_D2_B Index_D2_A Index_Euro2_B
Index_Euro2_A CSS2_B
CSS2_A OSS2_B OSS2_A (PAIRED)
/MISSING ANALYSIS.
```

```
T-TEST PAIRS=GESUNDHEIT1_B GESUNDHEIT1_A Index_UK1_B Index_UK1_A Index_D1_B
Index_D1_A Index_Euro1_B Index_Euro1_A CSS1_B CSS1_A OSS1_B OSS1_A WITH GESUNDHEIT2_B
GESUNDHEIT2_A Index_UK2_B Index_UK2_A Index_D2_B Index_D2_A Index_Euro2_B
Index_Euro2_A
CSS2_B CSS2_A OSS2_B OSS2_A (PAIRED)
/CRITERIA=C1(.9500)
/MISSING=ANALYSIS.
```

* %%%%%%%%%%
Vergleiche zwischen Gruppe 1(A) und Gruppe 2(B)
Zeitpunkt 1 als Beispiel
%%%%%%%%%

```
USE ALL.
COMPUTE filter_$=(Zeitpunkt = 1).
VARIABLE LABELS filter_$ 'Zeitpunkt = 1 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
```

```
EXECUTE.
CROSTABS
/TABLES=Sex BY RM_GROUP_E1_C16
/FORMAT=AVALUE TABLES
/STATISTICS=CHISQ
/CELLS=COUNT
/COUNT ROUND CELL.
```

```
NPART TESTS
/M-W= Age_E1 NYHA_CLASS_E3_C26 BEWEG SORGEN ALLTAG SCHMERZ ANGST GESUNDHEIT
Index_UK Index_D Index_Euro PL Symptome LQ SL Sym_Stab Selbst_W CSS OSS BY
RM_GROUP_E1_C16(1 2)
/MISSING ANALYSIS.
```

```
T-TEST GROUPS=RM_GROUP_E1_C16(1 2)
```

8 Anhang

```
/MISSING=ANALYSIS
/VARIABLES=Age_E1 GESUNDHEIT Index_UK Index_D Index_Euro PL Symptome LQ SL Sym_Stab
Selbst_W CSS OSS
/CRITERIA=CI(.95).
```

```
* %%%%%%%%%%
Vergleich zwischen den Altersgruppen je Gruppe und Zeitpunkt
Gruppe A zu Zeitpunkt 1 als Beispiel
%%%%%%%%%
```

```
USE ALL.
COMPUTE filter_$=(RM_GROUP_E1_C16 = 1 & Zeitpunkt = 1).
VARIABLE LABELS filter_$ 'RM_GROUP_E1_C16 = 1 & Zeitpunkt = 1 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

```
NPART TESTS
/K-W=Age_E1 BEWEG SORGEN ALLTAG SCHMERZ ANGST GESUNDHEIT Index_UK Index_D
Index_Euro PL Symptome LQ SL Sym_Stab Selbst_W CSS OSS BY Altersgruppe(1 3)
/MISSING ANALYSIS.
```

```
NPART TESTS
/M-W= Age_E1 BEWEG SORGEN ALLTAG SCHMERZ ANGST GESUNDHEIT Index_UK Index_D
Index_Euro PL Symptome LQ SL Sym_Stab Selbst_W CSS OSS BY Altersgruppe(1 2)
/MISSING ANALYSIS.
```

```
NPART TESTS
/M-W= Age_E1 BEWEG SORGEN ALLTAG SCHMERZ ANGST GESUNDHEIT Index_UK Index_D
Index_Euro PL Symptome LQ SL Sym_Stab Selbst_W CSS OSS BY Altersgruppe(1 3)
/MISSING ANALYSIS.
```

```
NPART TESTS
/M-W= Age_E1 BEWEG SORGEN ALLTAG SCHMERZ ANGST GESUNDHEIT Index_UK Index_D
Index_Euro PL Symptome LQ SL Sym_Stab Selbst_W CSS OSS BY Altersgruppe(2 3)
/MISSING ANALYSIS.
```

```
ONEWAY Age_E1 GESUNDHEIT Index_UK Index_D Index_Euro PL Symptome LQ SL Sym_Stab
Selbst_W CSS OSS BY Altersgruppe
/MISSING ANALYSIS
/POSTHOC=SCHEFFE ALPHA(0.05).
```

8 Anhang

* %%%%%%%%%%
Vergleich zwischen den NYHA-Klassen je Gruppe und Zeitpunkt
Gruppe A zu Zeitpunkt 1 als Beispiel
%%%%%%%%%

```
USE ALL.  
COMPUTE filter_$=(RM_GROUP_E1_C16 = 1 & Zeitpunkt = 1).  
VARIABLE LABELS filter_$ 'RM_GROUP_E1_C16 = 1 & Zeitpunkt = 1 (FILTER)'.  
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.  
FORMATS filter_$ (f1.0).  
FILTER BY filter_$.  
EXECUTE.
```

```
NPART TESTS  
/K-W=Age_E1 BEWEG SORGEN ALLTAG SCHMERZ ANGST GESUNDHEIT Index_UK Index_D  
Index_Euro PL Symptome LQ SL Sym_Stab Selbst_W CSS OSS BY NYHA_CLASS_E3_C26(1 3)  
/MISSING ANALYSIS.
```

```
NPART TESTS  
/M-W= Age_E1 BEWEG SORGEN ALLTAG SCHMERZ ANGST GESUNDHEIT Index_UK Index_D  
Index_Euro PL Symptome LQ SL Sym_Stab Selbst_W CSS OSS BY NYHA_CLASS_E3_C26(1 2)  
/MISSING ANALYSIS.
```

```
NPART TESTS  
/M-W= Age_E1 BEWEG SORGEN ALLTAG SCHMERZ ANGST GESUNDHEIT Index_UK Index_D  
Index_Euro PL Symptome LQ SL Sym_Stab Selbst_W CSS OSS BY NYHA_CLASS_E3_C26(1 3)  
/MISSING ANALYSIS.
```

```
NPART TESTS  
/M-W= Age_E1 BEWEG SORGEN ALLTAG SCHMERZ ANGST GESUNDHEIT Index_UK Index_D  
Index_Euro PL Symptome LQ SL Sym_Stab Selbst_W CSS OSS BY NYHA_CLASS_E3_C26(2 3)  
/MISSING ANALYSIS.
```

```
ONEWAY Age_E1 GESUNDHEIT Index_UK Index_D Index_Euro PL Symptome LQ SL Sym_Stab  
Selbst_W CSS OSS BY NYHA_CLASS_E3_C26  
/MISSING ANALYSIS  
/POSTHOC=SCHEFFE ALPHA(0.05).
```