

Aus der  
Klinik und Poliklinik für Diagnostische und Interventionelle Radiologie  
der Universitätsmedizin der Johannes-Gutenberg-Universität Mainz

**"Myokardiale Strain-Analyse in der kardialen MRT –  
Evaluation von Referenzwerten und Reproduzierbarkeit  
von drei Software-Lösungen“**

Inauguraldissertation  
zur Erlangung des Doktorgrades der  
Medizin  
der Universitätsmedizin  
der Johannes-Gutenberg-Universität Mainz

Vorgelegt von

Marcel Maximilian Larisch  
aus Reinbek

Mainz, 2024

Wissenschaftlicher Vorstand: Univ.-Prof. Dr. Hansjörg Schild

Tag der Promotion: 17.12.2024

# Inhaltsverzeichnis

<b>Inhaltsverzeichnis .....</b>	<b>I</b>
<b>Abkürzungsverzeichnis .....</b>	<b>III</b>
<b>Abbildungsverzeichnis .....</b>	<b>IV</b>
<b>Tabellenverzeichnis .....</b>	<b>V</b>
<b>1. Einleitung .....</b>	<b>1</b>
<b>2. Literaturdiskussion .....</b>	<b>3</b>
<b>2.1 Grundlagen der Magnetresonanztomografie .....</b>	<b>3</b>
<b>2.2 Kardiale Magnetresonanztomografie .....</b>	<b>6</b>
2.2.1 Sequenzen und Schnittebenen der CINE Bildgebung .....	6
2.2.2 CINE Bildgebung .....	6
<b>2.3 Messmethoden der Myokardbewegung .....</b>	<b>8</b>
2.3.1 Tissue Tagging .....	8
2.3.2 Feature Tracking .....	9
<b>2.4 Kardiale Funktionsparameter .....</b>	<b>12</b>
2.4.1 Quantitative Funktionsparameter .....	12
2.4.2 Strain .....	13
2.4.3 Referenzwerte Strain .....	16
2.4.4 Intra- und Interobserver-Variabilität .....	18
2.4.5 Anbieterabhängige Variabilität (Intervendor-Variabilität) .....	19
<b>3. Material und Methoden .....</b>	<b>22</b>
<b>3.1 Probandenkollektiv .....</b>	<b>22</b>
<b>3.2 Kardiale Magnetresonanztomographie .....</b>	<b>22</b>
<b>3.3 CMR Analyse .....</b>	<b>23</b>
3.3.1 Cvi42 .....	24
3.3.2 SuiteHeart .....	25
3.3.3 TrufiStrain .....	25
<b>3.4 Statistische Auswertung .....</b>	<b>29</b>
<b>4. Ergebnisse .....</b>	<b>31</b>
<b>4.1 Gesunde Freiwillige .....</b>	<b>31</b>
<b>4.2 Globale Strainparameter im Softwarevergleich .....</b>	<b>33</b>
<b>4.3 Übereinstimmung der Softwareprogramme .....</b>	<b>34</b>
<b>4.4 Referenzwerte Globale Strainparameter .....</b>	<b>36</b>
<b>4.5 Untersucherabhängige Reproduzierbarkeit .....</b>	<b>37</b>
4.5.1 GRS .....	38
4.5.2 GCS .....	39
4.5.3 GLS .....	40
<b>4.6 Intrarater Reproduzierbarkeit für GCS, GLS, GRS .....</b>	<b>41</b>
<b>5. Diskussion .....</b>	<b>43</b>
<b>5.1 Unterschiede bei der linksventrikulären Volumetrie .....</b>	<b>44</b>
<b>5.2 Unterschiede linksventrikulärer globaler Strainparameter .....</b>	<b>46</b>

5.3	Anbieterabhängige Unterschiede im Bezug zur aktuellen Studienlage .....	48
5.4	Inter- und Intrarater Reproduzierbarkeit im Bezug zur aktuellen Studienlage	52
5.5	Klinische Anwendung und zukünftige Entwicklungen.....	54
5.6	Limitationen .....	56
6.	Zusammenfassung.....	59
	Literaturverzeichnis .....	61
	Danksagung .....	69

## Abkürzungsverzeichnis

CMR	Kardiale Magnetresonanztomografie
CV	Coefficient of Variation
DRA	Deformable Registration Algorithm
EDV(i)	Enddiastolisches Volumen (Index)
EF	Ejektionsfraktion
ESV(i)	Endsystolischer Volumen (Index)
FT	Feature Tracking
KÖF	Körperoberfläche
LoA	Limits of Agreement
LVGCS	Left ventricular Global Circumferential Strain
LVGLS	Left ventricular Global Longitudinal Strain
LVGRS	Left ventricular Global Radial Strain
LVM(i)	Left ventricular Massindex
MaiCo-Studie	Mainzer-Kohorten- Studie
rmANOVA	Einfaktorielle Varianzanalyse mit Messwiederholung
SSFP	Steady State free precession
SV	Schlagvolumen
ST	Strain
STE	Speckle Tracking Echokardiografie
T1T2T3-Studie	Studie zur Ermittlung von Mapping-Referenzwerten bei Herzgesunden
T1	T1 Zeit Abbau der Quermagnetisierung
T2	T2 Zeit Aufbau der Längsmagnetisierung
TE	Echozeit
TR	Repetitionszeit
TT	Tissue Tagging
R1	Untersucher 1
R2	Untersucher 2

## Abbildungsverzeichnis

Abbildung 1: Durch HF-Impuls abgelenkter Spin ( $\alpha$ =Flipwinkel) .....	3
Abbildung 2: Phasen- und Frequenzkodierung im K-Raum .....	5
Abbildung 3:CINE vs. Tissue Tagging im Vierkammerblick (links Diastole, rechts Systole).....	8
Abbildung 4: Definition Lagrange'scher und Euler'scher Strain (21) .....	14
Abbildung 5: Strainrichtungen in Diastole und Systole .....	14
Abbildung 6: Cvi42 Bearbeitungsfenster Strainberechnung.....	25
Abbildung 7: TrufiStrain Volumenberechnung, Basisschicht durch LAX definiert (61). .....	26
Abbildung 8: TrufiStrain- Konturbearbeitung (epi- und endokardiale Konturen).....	27
Abbildung 9: Einzeichnung der epi- und endokardialen Konturen in der Langachse Polygon Tool (TrufiStrain).....	27
Abbildung 10: Cvi42 Bearbeitungsfenster Strain.....	28
Abbildung 11: Boxplots EF, EDVi, ESVi (Kurzachse) in der Gegenüberstellung Cvi42, SuiteHeart, TrufiStrain .....	31
Abbildung 12: Boxplot GRS, GCS, GLS (R1) im Softwarevergleich .....	33
Abbildung 13: Bland-Altman-Plot GRS im Softwarevergleich (R1; Mittelwert) .....	35
Abbildung 14: Bland-Altman-Plot: GCS im Softwarevergleich (R1; gepoolter Mittelwert) .....	35
Abbildung 15: Bland-Altman-Plot: GLS im Softwarevergleich (R1; gepoolter Mittelwert) .....	36
Abbildung 16: Bland-Altman Plot GRS Interrater .....	38
Abbildung 17: Boxplot GRS Interrater .....	38
Abbildung 18: Bland-Altman Plot GCS Interrater .....	39
Abbildung 19: Boxplot GRS Interrater Reliabilität.....	39
Abbildung 20: Bland-Altman Plot GLS Interrater Reliabilität .....	40
Abbildung 21: Boxplot GLS Interrater Reliabilität .....	40

## Tabellenverzeichnis

Tabelle 1: Strainreferenzwerte auf Basis von 295 gesunden Männern und 301 gesunden Frauen aus Boehm et. al. (29) .....	17
Tabelle 2: Bildparameter der Maico-MR-Studie und der T1T2T3-Studie .....	23
Tabelle 3: Ausgangsmerkmale und Vergleich der Volumetrie.....	32
Tabelle 4: Globale Strainparameter Untersucher 1 im Softwarevergleich.....	33
Tabelle 5: Übereinstimmung der Softwareanbieter .....	34
Tabelle 6: Referenzwerte Globale Strainparameter nach Software und nach Geschlecht.....	36
Tabelle 7: Untersucherabhängige Reproduzierbarkeit für GRS, GCS, GLS .....	37
Tabelle 8: Intrarater Reproduzierbarkeit für GCS, GLS, GRS.....	41

## 1. Einleitung

Die kardiale Magnetresonanztomografie (CMR) ist eine vielseitig einsetzbare, nicht invasive und auf nicht-ionisierender Strahlung basierende Bildgebungstechnik, die sich in den vergangenen Jahren zum Referenzstandard für die Berechnung von links- und rechtsventrikulären (LV/RV) Volumina und ihrer Funktion sowie der myokardialen Gewebecharakterisierung entwickelt hat. Das Ziel intensiver Forschungsvorhaben beinhaltet zuletzte neue quantitative Parameter zu entwickeln, die bereits Informationen in der subklinischen Phase einer Erkrankung bieten. Insbesondere der myokardialen Strain-Analyse wird hierbei ein großes Potential zugeschrieben (1, 2). Im klinischen Alltag wichtige Parameter wie z.B. die Ejektionsfraktion (EF) liefern kaum Hinweise auf regionale Kontraktilitätsstörungen. Die Sensitivität der kardialen Strainanalyse zur Identifikation von regionalen Deformationsstörungen ist der EF überlegen (3). Es konnte nachgewiesen werden, dass die Verminderung des globalen longitudinalen Strains bei Patienten mit dilatativer oder ischämischer Kardiomyopathie bevorstehende kardiale Ereignisse prognostizieren kann (4).

Der Globale Strain wurde erstmalig in den 80er Jahren etabliert und stellt die prozentuale Längenänderung des Myokards in einer spezifischen Richtung dar. Zunächst wurde dieser mittels eigener Bildgebungssequenzen (CMR-Tagging) durchgeführt. Heutzutage hat sich die Methode des Feature Trackings durchgesetzt. Hierbei handelt es sich um eine reine Postprocessing-Methode, die es erlaubt, in der Routine eingesetzte CINE-Bildgebungssequenzen im Anschluss weiter zu analysieren (1).

In den vergangenen Jahren wurden von diversen Softwareherstellern Feature Tracking Postprocessing-Softwarelösungen zur Berechnung von Globalen Strainparametern entwickelt (5).

Somit stellt sich die Frage der Vergleichbarkeit unter den einzelnen Softwarelösungen. Schuster et al. zeigten bereits 2015 in einem Vergleich zweier Softwareanbieter signifikante Unterschiede zwischen den Ergebnissen der Strain-Berechnungen. Weiterhin konnte gezeigt werden, dass die Reproduzierbarkeit der Daten ebenfalls softwareabhängig ist (6). Untermauert wurden diese Ergebnisse von Barreiro-Perez et al., die die Ergebnisse von vier Softwareanbietern miteinander verglichen und dabei feststellten, dass insbesondere der globale radiale Strain (GRS) und der globale longitudinale Strain (GLS) zwischen verschiedenen Anbietern nicht vergleichbar ist.

Um die Strain-Analyse in die klinische Anwendung zu bringen, müssten auf Software- und Anwendererebene Standardisierungen erfolgen (7). Bis zu einer möglichen Standardisierung sind Untersuchungen von softwareabhängigen Unterschieden bei der Generierung von Strainwerten ein wichtiger Aspekt für eine zukünftige klinische Anwendbarkeit.

Ziel dieser Dissertation ist es deshalb, die MRT-basierte Strain-Analyse an Hand von drei verschiedenen Softwarelösungen zu vergleichen und Referenzwerte in einem gesunden Probandenkollektiv für diese zu erstellen. Des Weiteren soll der Einfluss der Software auf die Anwenderreproduzierbarkeit der jeweiligen Programme untersucht werden. Die Ergebnisse dieser Dissertation könnten als Basis für die Beurteilung kardialer Erkrankungen genutzt werden.

## 2. Literaturdiskussion

### 2.1 Grundlagen der Magnetresonanztomografie

Im folgenden Abschnitt soll eine Einführung in die Magnetresonanztomografie vermittelt werden.

Die Magnetresonanztomografie beruht auf dem grundlegenden Prinzip, dass bewegte Ladung einen Strom und ein eigenes Magnetfeld um sich erzeugt. Insbesondere die charakteristischen Eigenschaften des Wasserstoffatoms sind von zentraler Bedeutung. Es ist das am häufigsten vorkommende Atom im menschlichen Körper und besteht im Vergleich zu allen anderen Elementen aus nur einem Proton.

Jedes Proton besitzt eine Drehbewegung um die eigene Achse, die als Spin bezeichnet wird und jeweils in eine Richtung (Vektor) wirkt. In Ruhe – unter Einfluss des Erdmagnetfeldes – sind die einzelnen Vektoren willkürlich zueinander orientiert (8). Werden diese Protonen nun einem äußeren Magnetfeld ausgesetzt - wie z.B. der menschliche Körper in der MRT - so richten sich diese innerhalb dieses Magnetfeldes neu aus. Zusätzlich zu der Drehrichtung des Spins liegt zu jeder Zeit eine Bewegung um eine orthogonale Achse vor. Es handelt sich um eine Kreiselbewegung, die als Präzession bezeichnet wird und mit einer speziellen Geschwindigkeit – der Larmorfrequenz - sich um die eigene Achse dreht. Die Größe der Larmorfrequenz ist proportional zum außen angelegten Magnetfeld. Die Mehrheit der Vektoren richtet sich parallel aus, ein kleiner Anteil kann aber auch den energetisch ungünstigeren Zustand - antiparallel zum Magnetfeld - einnehmen. Durch das Überwiegen der parallel ausgerichteten Spins bildet sich nun ein eigenes Magnetfeld aus, das entlang des statisch angelegten Magnetfeldes verläuft und als Längsmagnetisierung bezeichnet wird (8). Mithilfe der Larmorfrequenz ist es nun möglich, durch einen Hochfrequenzimpuls der gleichen Frequenz die Spins durch Resonanz aus dem Gleichgewicht zu bringen. Der Winkel  $\alpha$  zeigt die Ablenkung der Spins.

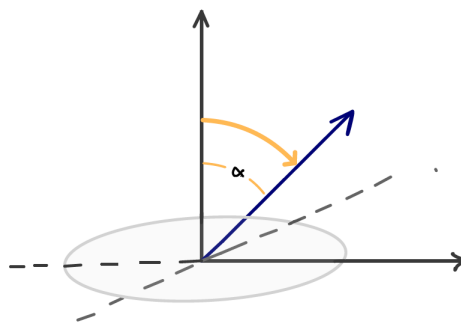


Abbildung 1: Durch HF-Impuls abgelenkter Spin (alpha=Flipwinkel)

Dieser Winkel ist abhängig von der Stärke des Hochfrequenzimpulses und wird als Flipwinkel bezeichnet.

Mit der Auslösung eines Hochfrequenzimpulses entstehen zwei Effekte. Zum einen richten sich zunehmend gleich viele Protonen parallel wie antiparallel aus, womit die Längsmagnetisierung verschwindet. Zum anderen präzedieren die Spins zunehmend synchron. Sie gehen in Phasenkohärenz, womit es zur Ablenkung der Spins in xy-Richtung kommt und sich eine Quermagnetisierung ausbildet. In Abhängigkeit des Hochfrequenzimpulses kann es auch zu einer Inversion der Längsmagnetisierung kommen.

Die Quermagnetisierung ist nicht starr, sondern bewegt sich mit den präzedierenden Protonen um die Achse des starren Magnetfeldes. Dies ist relevant, da über die Änderung dieses Magnetvektors mithilfe einer Spule ein Signal mit der Larmorfrequenz gemessen werden kann (8).

Mit Ausschalten des Hochfrequenzimpulses kehren alle Spins in den Ausgangszustand zurück. Hierdurch baut sich die Längsmagnetisierung wieder auf und die Quermagnetisierung bildet sich zurück.

Die T1-Relaxation bezeichnet den Prozess des Aufbaus der Längsmagnetisierung und wird durch die Konstante T1 beschrieben. Die Konstante T1 beschreibt die Zeit, nach der 63% der Längsmagnetisierung wieder hergestellt sind. In Abhängigkeit vom biologischen Gewebe beträgt die T1-Relaxation 300-2000ms. Je größer der Anteil an Wasser im Gewebe desto länger beträgt die T1-Zeit.

Nach Abschalten des Hochfrequenzimpulses beginnt die Dephasierung der Spins und es kommt zum Zerfall der Quermagnetisierung. Dies wird als T2-Relaxation bezeichnet. Die Zeitspanne, bis die Quermagnetisierung auf 37 % ihres Ausgangswerts abgebaut ist, ist wesentlich kürzer und wird als T2-Konstante bezeichnet. Die T2 ist wesentlich kürzer als die T1 und beträgt ca. 30 - 150 ms. T1 und T2 Zeit sind voneinander unabhängig. Sie ist neben der gewebespezifischen Anzahl der Wasserstoffatome des jeweiligen Gewebes auch von der angewandten Feldstärke des MRT abhängig und bestimmt somit entscheidend den Kontrast im MRT-Bild (9).

Da das starre Magnetfeld nur in der Theorie homogen ist, kommt es durch Spin-Spin Interaktionen zu einer schnelleren Dephasierung und zu einer Verkürzung der T2 Relaxation. Um dem entgegenzuwirken, kann nach einer Zeit x ein erneuter 180° Impuls gesendet werden, der zu einer Rephasierung der Protonen führt. Damit kann die Auswirkung störender Magnetfeldinhomogenitäten vermindert werden. Durch die Rephasierung wird die Signalqualität verbessert (9).

Die Zeit, die zwischen zwei wiederholten Anregungen liegt, wird als Repetitionszeit (TR) und die Zeit zwischen der Anregung und Registrierung des Signals wird als Echozeit (TE) benannt. Durch die Wahl von TR und TE wird maßgeblich gesteuert, wie sich ein spezifisches Gewebe in einer Sequenz abbildet. Hierdurch können Bilder generiert werden deren Kontrast entweder durch die unterschiedlichen Relaxationszeiten T1 und T2 bestimmt wird oder durch die jeweilige Protonendichte (10).

Eine weitere Möglichkeit zur Anregung der Spins sind sogenannte Gradientenechosequenzen. Entlang der Strecke des Hauptmagnetfeldes wird durch ein Gradientenmagnetfeld ein inhomogenes Magnetfeld erzeugt. Durch kurzzeitige Änderungen dieses Gradienten, kommt es zur Anregung und Rephasierung der Spins (8, 10).

Die große Herausforderung bei der MRT-Bildgebung besteht nun darin, präzise Schichten anregen zu können und diese Bildinformation einer genauen Lokalisation zuzuordnen. Hierfür werden magnetische Gradientenfelder entlang der x,y- und z-Achse erzeugt. Diese überlagern das statische Magnetfeld. Somit wird neben dem Phasenkodiergradient in x-Richtung ein Frequenzkodiergradient in y- Richtung und ein Schichtauslesegradient in Z- Richtung geschaltet. Hierdurch wird eine ortsspezifische Resonanzfrequenz erzeugt. Mit einem entsprechenden Hochfrequenzimpuls der gleichen Frequenz kann selektiv eine Stelle einer Schicht angeregt werden.

Für die Bildgebung wird ein mathematischer Raum, der k-Raum, benutzt. Der k-Raum ist eine Anordnung der Rohdaten in einer Matrix. Auf der x-Achse ist die Frequenzkodierung, auf der Y-Achse die Phasenkodierung eingetragen (9).

Diese räumliche Position erlaubt es ein zweidimensionales Bild zu generieren (10, 11,12).

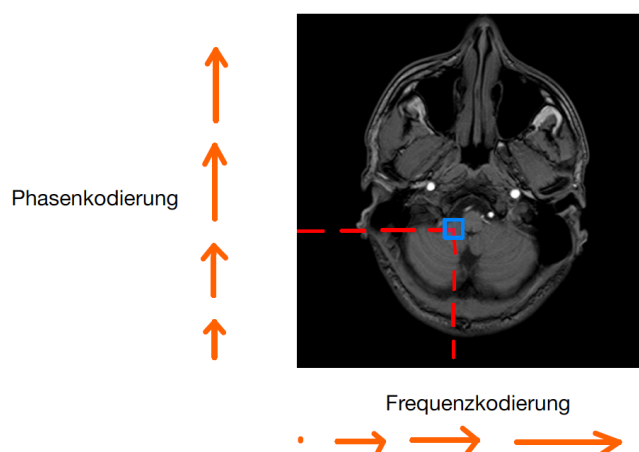


Abbildung 2: Phasen- und Frequenzkodierung im K-Raum

## 2.2 Kardiale Magnetresonanztomografie

### 2.2.1 Sequenzen und Schnittebenen der CINE Bildgebung

Die kardiale Magnetresonanztomografie ermöglicht nicht-invasiv und ohne ionisierende Strahlung Informationen über die ventrikuläre Funktion, myokardiale Perfusion sowie eine myokardiale Gewebscharakterisierung und erlaubt zudem auch angiographische Darstellungen. Um Informationen über die Pumpfunktion zu erhalten, wird die CINE Bildgebung verwendet. Die standardisierten kardialen Schnittebenen für den linken Ventrikel bestehen aus Langachsendarstellungen (Zwei-, Drei- und Vierkammerblick) sowie den Kurzachsendarstellungen. Der Zweikammerblick umfasst die Schnittebene vom Apex bis zum linken Vorhof und ermöglicht die Beurteilung der Mitralklappe sowie der Funktion des linken Herzens. Der Vierkammerblick ermöglicht die Beurteilung von linkem und rechtem Vorhof und beiden Ventrikeln, sowie der Mitralklappe und Trikuspidalklappe. Der LV-Dreikammerblick dient der Visualisierung von Mitralklappe und linksventrikulärem Ausflusstrakt mit der Aortenklappe.

Die Kurzachse wird orthogonal zu zwei Langachsenschnitten generiert. Je nach Größe des Herzens kann die Anzahl der Kurzachsenschnitte angepasst werden. Hiermit kann die Verdickung des Myokards des linken Ventrikels in der Systole von Apex bis Basis beurteilt werden, außerdem dient die Kurzachse der Berechnung der linksventrikulären quantitativen Funktionsparameter. Als Minimum werden in den Kurzachsen die Herzbasis, Mitte und der Apex abgebildet (13). Da sich diese Dissertation mit der Bestimmung von volumetrischen und Funktionsparametern mittels Feature Trackings aus CINE-SSFP Sequenzen befasst, wird im nächsten Abschnitt dezidiert auf die CINE Bildgebung und deren Postprocessing eingegangen.

### 2.2.2 CINE Bildgebung

Bei der CINE-Bildgebung handelt es sich um dynamische Bildsequenzen, die Bildinformationen über den gesamten Herzzyklus generieren (14). Insbesondere Herzfunktionsparameter wie z.B. Schlagvolumen, Ejektionsfraktion, enddiastolisches und endsystolisches Volumen und Masse können hiermit im Anschluss bestimmt werden. Gleichzeitig ist sie für die Beurteilung von regionalen Wandbewegungsstörungen z.B. nach Myokardinfarkt geeignet (15, 16). Weitere Indikationen sind angeborene Herzfehler oder Kardiomyopathien wie z.B. die hypertrophe Kardiomyopathie.

CINE-Bildgebung wird erzeugt, indem wiederholend der gewünschte Bereich einer Schicht zu unterschiedlichen Zeitpunkten dargestellt wird. Parallel hierzu erfolgt die Aufzeichnung eines EKG's, weshalb die einzelnen Bilder dem Herzzyklus zugeordnet werden können. Unterschiedliche K-Räume werden je Sequenz und Phase des Herzzyklus erstellt und es benötigt üblicherweise 25-30 R-R Intervalle bis ein Herzzyklus abgebildet werden kann. In neueren CINE Sequenzen sind zum Teil nur noch 10 R-R Intervalle notwendig. Die CINE Bildgebung stellt somit die Zusammenfassung mehrerer Ausschnitte diverser Herzschläge dar.

Um ein Bild zu erzeugen, das möglichst von der Eigenbewegung des Herzens unbeeinflusst ist, muss dieses in Millisekunden erstellt werden. Heutzutage werden in der Regel Steady-state free precession (SSFP) Bildgebungssequenzen benutzt. Sie zeichnen sich durch sehr kurze Repetitionszeiten und Echozeiten aus (12). Zwischen zwei Radiofrequenzimpulsen kommt es zum einen zum Abfall der transversalen Magnetisierung (Free induction decay), aber auch zu einem generierten Echo vor dem erneuten Radiofrequenzimpuls, womit ein kontinuierliches Signal mit einer variierenden Amplitude entsteht. FISP, TrueFISP oder PSIF sind unterschiedliche SSFP-Sequenzen. Ein großer Störfaktor für das MRT sind Bewegungsartefakte. Die Hauptbewegungsfaktoren sind die thorakale Bewegung durch In- und Expiration. Mit In- und Expiration verändert sich auch die Lage des Herzens im Thorax. Die Atmung muss somit während der Aufnahme angehalten werden und ist dementsprechend ein limitierender Faktor für die mögliche Aufnahmezeit.

## 2.3 Messmethoden der Myokardbewegung

Neben der alleinigen Analyse der CINE-Sequenz, z.B. mittels des relative wall motion Thickening zur regionalen Wandbewegungsanalyse wurden weitere Verfahren zur Analyse der Myokardbewegung entwickelt. Man unterscheidet hierbei die sogenannte „Tailored image acquisition“, wie Tissue Tagging oder Phase Velocity Mapping von Postprocessing-Methoden, wie die des Feature Trackings (17).

### 2.3.1 Tissue Tagging

Die Tissue Tagging (TT) Sequenz wurde 1988 von Zerhouni et al. entwickelt und basiert auf dem Prinzip, dass mittels Radiofrequenzimpulsen Störungen der Magnetisierung des Gewebes erzeugt werden. Es entstehen vertikale Linien (line tagging) beziehungsweise ein Gitter (grid tagging) bei einer Überlagerung von zwei orthogonal zueinander stehenden Linien (18). Die Tags verformen sich mit der Kontraktion des Herzes genauso wie das umliegende Gewebe, sodass Wandbewegungsstörungen visuell erkannt werden können. Außerdem können mit Hilfe von Postprocessing Softwarelösungen die Berechnung von quantitativen Funktionsparameter, wie z.B. die Ejektionsfraktion (EF) ermittelt werden (19).

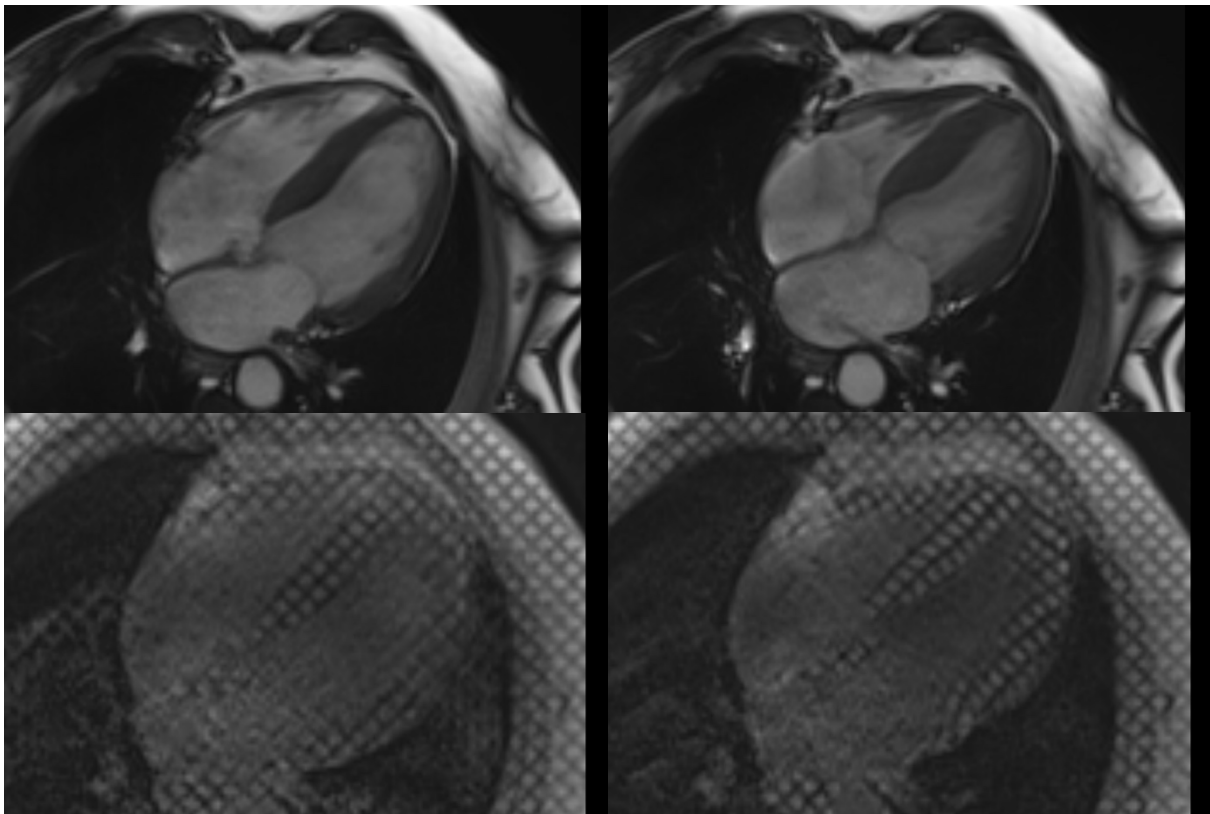


Abbildung 3: CINE vs. Tissue Tagging im Vierkammerblick (links Diastole, rechts Systole)

Die in den letzten Jahren etablierte HARP-Technologie ist die am weitesten verbreitete Optical Flow Methode. Sie basiert auf dem Prinzip, dass in jedem Bildabschnitt Pixel gleicher Helligkeit einander zugeordnet und schließlich als ein Ort innerhalb eines Gewebes definiert wird. Damit kann die Bewegung des Ortes über die Zeit verfolgt werden (17).

Obwohl das TT in der Literatur oft als Referenz deklariert wird, birgt diese Methode einige Nachteile. Diese sind insbesondere die räumliche Auflösung, als auch die so genannten „Fading Tags“. Nahe der Enddiastole beginnen die Tag-Linien zu verblassen, sodass der Kontrast zum umliegenden Gewebe abnimmt und die Differenzierung erschwert wird (20). Ferner ist ein Problem des TT, dass das Signal zu Rausch Verhältnis gering ist, wodurch die Grenze zwischen Blut und Myokard verschwimmt. Es kann somit zu einer Unterschätzung des Strains insbesondere bei einer schnelleren Herzfrequenz kommen (21). Nicht zu vernachlässigen ist die Störanfälligkeit durch Bewegungsartefakte, insbesondere durch thorakale Bewegungen wie In- und Expiration. Die Folge all dieser Aspekte ist, dass sich das TT bisher nicht in der klinischen Praxis etablieren konnte und primär lediglich zu Forschungszwecken verwendet wird (22).

### 2.3.2 Feature Tracking

Im Vergleich zum Tissue Tagging handelt es sich beim Feature Tracking (FT) um eine reine Postprocessing-Methode. Beim FT werden mittels speziellen Softwarepaketen Standard SSFP-CINE Bildgebungssequenzen im Nachhinein analysiert. Der Vorteil ist, dass keine zusätzliche Akquisition aufwendiger Sequenzen notwendig wird. Dies ist zeit- und kosteneffizient.

Das FT als Methode basiert auf dem Prinzip des Optical Flow. Hier werden entlang einer vom Untersucher definierten Grenze durch das Softwareprogramm kleine Quadrate eingezeichnet. Die Software sucht im Anschluss in aufeinander folgenden Bildern einer Sequenz nach Pixeln gleicher Helligkeit und ordnet diese einander zu. So kann die Verformung des Myokards im Vergleich zum TT über den kompletten Herzzyklus erfasst werden, da die Konturen hier, anders als beim Tagging (Fading Tags), nicht verblassen (17, 23).

In der Klinik ist die Speckle Tracking Echokardiografie (STE) etabliert (23). Sie stellt die am weitesten verbreitete Methode dar, um myokardiale Deformation zu beurteilen, und war das erste Verfahren, das Tissue Tracking in der kardialen Bildgebung

ermöglichte (24). Das Prinzip basiert darauf, dass sogenannte Speckles – das sind stabile akustische Marker in 2D Bildern, die durch die positive Interferenz von Ultraschallwellen entstehen – nachverfolgt werden können. Im Vergleich zum CMR FT hat STE eine bessere zeitliche Auflösung. Sie hat ferner den Nachteil, dass einige Anteile des Myokards weniger exakt gemessen werden können. Speckles, die näher am Ultraschallkopf lokalisiert sind, lassen eine Nachverfolgung leichter zu als ferner gelegene. Stark abhängig ist diese Methode von der Bildqualität, die allein aus anatomischen Gründen patientenabhängig variieren kann. Das führt zu einem Verlust an Genauigkeit und einer geringeren Reproduzierbarkeit (25).

Postprocessing Softwarelösungen, die für das Feature Tracking in der CMR als auch für das Speckle-Tracking in der Echokardiografie Einsatz finden, werden durch Bilderkennungsprogramme unterstützt. Diese führen eine semiautomatische Konturerkennung durch, sodass Endo- und Epikard in der Enddiastole beziehungsweise Endsystole eingezeichnet werden.

Papillarmuskeln und Trabekel sollten gemäß Konvention entweder dem Blutvolumen oder der Myokardmasse zugerechnet werden und mit den je Modus entsprechenden Referenzwerten verglichen werden(26).

Der Algorithmus folgt dabei einer Hierarchie, bei der entweder initial die Einwärts- oder die Auswärtsbewegung des Endokards entlang der Hohlraum-Gewebe-Grenzfläche verfolgt wird (27).

Im Gegensatz zum Speckle Tracking wird durch die SSFP Sequenzen ein guter Kontrast zwischen Myokard und Hohlraum erzeugt, jedoch sind CINE Aufnahmen im Vergleich eine Rekonstruktion aus vielen Herzschlägen. Dem zu Folge hat die CINE Bildgebung eine geringere zeitliche Auflösung (23).

Ein weiterer limitierender Faktor sind Artefakte, die dadurch entstehen, dass aufgrund der Verkürzung der Ventrikel in der Systole Strukturen aus der Schnittbildebene sich heraus- oder hineinbewegen. Dies ist relevant für die anschließende Analyse. Für die Generierung von Funktionsparametern muss sich der Untersucher dieses Problems bewusst sein und insbesondere in der basisnahen Kurzachse die betreffenden Schnitte aus der Untersuchung ausschließen, die nicht von Diastole bis Systole dargestellt werden können. Ansonsten kann es zu einer Verfälschung der kardialen Funktionsparameter kommen (23).

Nichtsdestotrotz hat sich insbesondere aufgrund der einfacheren Handhabbarkeit Feature Tracking zu einer vielversprechenden quantitativen Postprocessing Technik entwickelt, die zunehmend Anwendung findet (18).

Da in dieser Dissertation der Fokus auf dem Vergleich der Berechnung der Funktionsparameter liegt, wird im folgenden Kapitel auf die Definition und Berechnung sowie auf die Referenzwerte näher eingegangen.

## 2.4 Kardiale Funktionsparameter

### 2.4.1 Quantitative Funktionsparameter

Ein globaler Parameter für die Beurteilung der kardialen Funktion stellt die Ejektionsfraktion (EF) dar. Diese ist der Anteil des Schlagvolumens am maximalen Füllungsvolumen zum Zeitpunkt der Enddiastole (EDV). Das Schlagvolumen (SV) stellt die Differenz der Volumina zum Zeitpunkt der Enddiastole und der Endsystole dar (28)

$$(EF = \frac{SV}{EDV} \times 100)$$

In einer Metanalyse aus dem Jahr 2020 bestehend aus 12 Studien mit insgesamt über 800 gesunden Probanden wurden Referenzwerte zur Definition von gesunden Probanden erstellt. Als gesund wurde darin für Männer eine mittlere EF von  $63\% \pm 6\%$  definiert, während für Frauen eine mittlere EF von  $66\% \pm 7\%$  deklariert wird (29).

Das EDV ist abhängig von Alter, Geschlecht, Größe und Gewicht. Um die Vergleichbarkeit zu ermöglichen, wird der EDVi angegeben, der das auf die Körperoberfläche standardisierte EDV darstellt. Grundsätzlich ist bei der Betrachtung von Referenzwerten darauf zu achten, ob es sich um Volumina handelt, die die Papillarmuskulatur mit beinhalten oder ausschließen. Der EDVi nimmt mit dem Alter ab und liegt beim gesunden Mann im Alter von 16-83 Jahren bei ca.  $77 \text{ ml/m}^2 \pm 15 \text{ ml/m}^2$ , bei Frauen im gleichen Alter bei ca.  $69 \text{ ml/m}^2 \pm 12 \text{ ml/m}^2$ . Der ESVi liegt beim gesunden Mann bei  $29 \text{ ml/m}^2 \pm 9 \text{ ml/m}^2$  und bei Frauen bei  $24 \text{ ml/m}^2 \pm 7 \text{ ml/m}^2$  (29).

Daneben ist die linksventrikuläre Masse (LVM) ein wichtiger Parameter. Sie ist eng verbunden mit der Körpergröße, abhängig vom Geschlecht und nimmt mit dem Alter zu. Wichtig ist, dass sie ein unabhängiger Indikator kardialer Morbidität und eine wichtige Größe zur Therapieevaluation darstellt. Sie kann den kumulativen Effekt linksventrikulärer Hypertonie, Diabetes, chronischer Niereninsuffizienz oder anderer Stoffwechselerkrankungen anzeigen (30). In der bereits o.g. veröffentlichten Referenzwertübersicht von Kawel-Boehm et. al. (2020) zeigten Männer einen mittleren LVMi von  $56 \text{ g/m}^2 \pm 10 \text{ g/m}^2$  und Frauen von  $45 \text{ g/m}^2 \pm 7 \text{ g/m}^2$  (29). In Validierungsstudien konnte bewiesen werden, dass eine hohe (Inter-Studien-) Reproduzierbarkeit bei der Messung von Volumen und Funktion vorliegt. So hat sich die CMR mittlerweile zum Referenzstandard zur Berechnung von den zuvor beschriebenen Parametern entwickelt (31).

Allerdings stellen Messwerte wie die Ejektionsfraktion nur einen Indikator für die globale Herzfunktion dar und können nicht zur Beurteilung einer Ventrikeldysfunktion in sehr frühen Stadien herangezogen werden (18). Eine Reduktion der EF stellt sich meist erst in den späteren Stadien eines Krankheitsverlaufs ein, gleichzeitig hat sie eine geringe Sensitivität in der Erfassung regionaler Myokardschäden (32). Daher ist die Entwicklung eines Parameters, der eine hohe Sensitivität für die Erkennung einer kardialen Dysfunktion im Frühstadium ermöglicht von großem Interesse. Dem Strain wird dieses Potential zugeschrieben.

#### 2.4.2 Strain

Im Rahmen der Myokarddeformationsbildgebung konnte in den letzten Jahren gezeigt werden, dass kontraktile Dysfunktion bei vielen Erkrankungen durch neue Parameter wie den Strain früher erkannt werden könnten (4, 17, 33). In Abgrenzung zur EF und anderen Methoden wie dem regional wall thickening ermöglicht dieser, weitere Informationen über die Beschaffenheit des Myokards auf nicht invasivem Wege zu erhalten(34).

Der myokardiale Strain (ST) stellt die prozentuale Längenänderung des Myokards in einer Richtung dar. In der Enddiastole handelt es sich um die initiale Länge ( $L_0$ ), in der Endsystole um die maximale Länge ( $L_1$ ) des Myokards. Er ist ein Indikator für die myokardiale Deformation während des Herzzyklus.

$$\text{Myocardialer Strain (\%)} = \frac{L_1 - L_0}{L_0}$$

Es gibt zwei vorherrschende Definitionen, den Lagrange'schen und den Euler'schen ST, wobei in der kardiologischen Bildgebung der Lagrange'sche ST verwendet wird. Die Strainberechnung nach Lagrange nimmt die Summe der Längenänderungen und setzt sie in Bezug zur Ausgangslänge des Myokards  $L_0$ . Der Euler'sche Strain berechnet sich aus der Summe von  $n$  Längenänderungsanteilen ( $\Delta$ ) an  $L_n$ . Dadurch berechnet der Euler'sche Strain im Vergleich höhere absolute Werte (23, 35) .

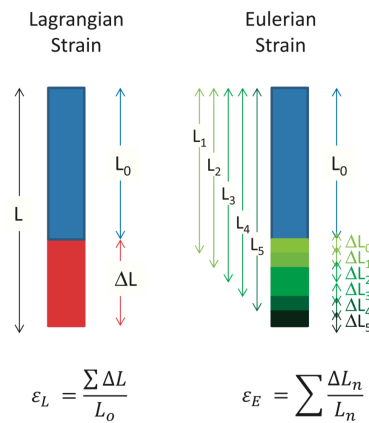


Abbildung 4: Definition Lagrange'scher und Euler'scher Strain (21)

Die Bewegung des Herzens ist komplex, da es sich um eine Verkürzung, Rotation und Verdickung des Herzmuskels handelt. In Abhängigkeit von seiner definierten Richtung spricht man vom longitudinalen Strain (LV GLS) entlang des Längsschnitts durch das Myokard, vom circumferentiellen Strain (LV GCS) entlang des Kreises und vom radialen Strain (LV GRS) bei Betrachtung der Dicke des Myokards (23,36).

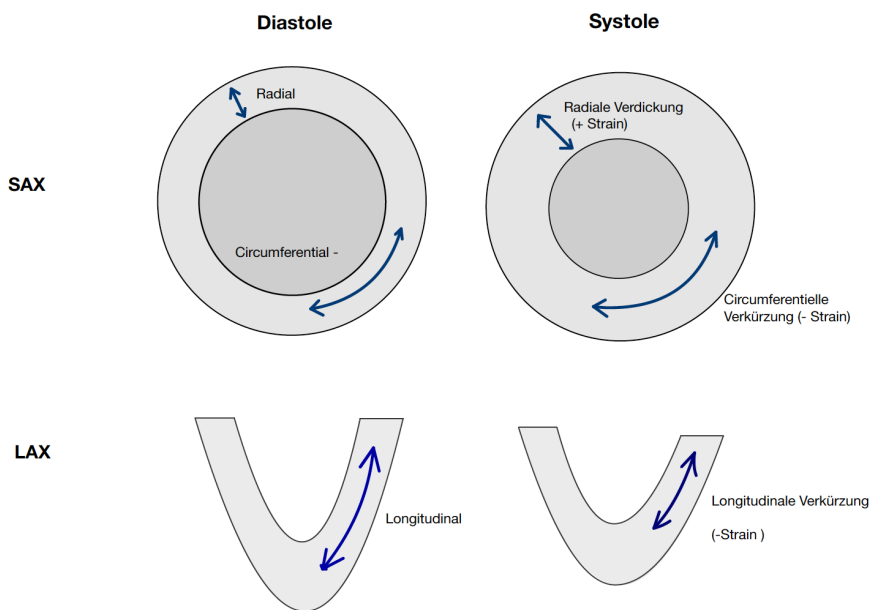


Abbildung 5: Strainrichtungen in Diastole und Systole

Grundsätzlich ist der Strain abhängig von Vor- und Nachlast sowie von der lastunabhängigen Kontraktilität und der Myozytenbeschaffenheit (32).

In unterschiedlichen Studien konnte nachgewiesen werden, dass der globale Strain abhängig vom Geschlecht ist. Insbesondere der GLS sowie der GCS waren bei Frauen größer (mehr negativ). Uneins ist die aktuelle Studienlage, ob ein Alterseffekt des Strains vorhanden ist. Einige Studien kamen zu dem Ergebnis einer Zunahme des GCS, während andere keinen Alterseffekt beim GCS oder beim GLS feststellen konnten (37-39).

Die klinische Anwendung insbesondere das prognostische Potential gegenüber etablierten Parametern wie der Ejektionsfraktion wurde intensiv untersucht. So konnten subklinische globale systolische Funktionsschwächen in frühen Stadien der hypertensiven Kardiopathie durch Veränderungen der globalen diastolischen Strainrate nachgewiesen werden (40). In einer weiteren Studie wurde ein ergänzender Einsatz zum Dobutamin Stress CMR Feature Tracking bei der Untersuchung der ischämischen Kardiopathie untersucht. Hier konnten Veränderungen des GCS festgestellt werden, selbst da, wo weder die LV-EF noch optisch erkennbare Wandbewegungsstörungen auffällig waren (41). Bei Patienten mit STEMI korrelieren die Strainparameter mit Infarktgebiet und Narbengröße. Bei Patienten mit chronischer ischämischer Kardiopathie konnten vitale von abgestorbenen Infarktgebieten unterschieden werden (35). Es konnte nachgewiesen werden, dass der GLS ein unabhängiger Wert zur Einschätzung des Myokardschadens nach ST-Hebungsinfarkt ist und in Kombination mit der Ejektionsfraktion eine Prognose für bevorstehende kardiovaskuläre Ereignisse (Tod, Reinfarkt, Herzinsuffizienz innerhalb eines Jahres nach Infarkt) zulässt (4). Weiterhin scheint der Strain eine wertvolle Ergänzung bei der Differenzierung von Patienten mit Verdacht auf eine akute Myokarditis und erhaltener Ejektionsfraktion zu sein (42). Fisher et. al konnten einen Zusammenhang zwischen dem gemessenen GLS und zukünftigen kardiovaskulären Ereignissen (wie zum Beispiel ventrikuläre Tachykardien > 30s oder Hospitalisierungsrate aufgrund von Herzinsuffizienz) nach Myokarditis nachweisen (43).

Trotz des großen klinischen Potentials ist der Einsatz globaler Strainparameter derzeit noch nicht in der klinischen Routine angekommen und wird bisher nur begrenzt zum Beispiel bei HFrEF Patient\*innen, zur Feststellung von myokardialer Ischämie in Stresstests oder auch als Verlaufparameter bei Patienten mit Morbus Fabry eingesetzt (32, 35). Die Studienlage ist zum Teil kontrovers und die Frage nach Referenzwerten gesunder bzw. kardial nicht vorbelasteter Patienten sind zum jetzigen Zeitpunkt nicht abschließend geklärt. Weitere Unsicherheit besteht bei der Vergleichbarkeit von Strainwerten, die von unterschiedlichen Softwareprogrammen erzeugt wurden. In Summe stellen somit das Thema Standardisierung im Bereich der Generierung von Strainwerten oder Korrekturfaktoren aktuelle Herausforderungen dar. Im folgenden Abschnitt wird die aktuelle Studienlage zum Thema Referenzwerte näher beleuchtet.

### 2.4.3 Referenzwerte Strain

Um unauffällige von pathologischen Befunden unterscheiden zu können, benötigt es Referenzwerte. Die Anzahl der Studien zu diesem Thema ist groß, allerdings die Anzahl der Probanden zu gering, um absolute Referenzwerte formulieren zu können. Daher haben Vo et al. 2018 mittels einer Metaanalyse globale Strainwerte bei gesunden Probanden erstellt. Zusammengefasst wurden die Ergebnisse von insgesamt 659 kardial gesunden Probanden aus insgesamt 17 Studien (3 Studien größer 100 Probanden), wovon die zwei größten aus 145 und 150 Probanden bestanden. Hierbei wurden folgende Normwerte für die globalen Strainparameter über alle Studien ermittelt:

LV GLS -20,1 % (95% CI: -20,9% bis -19,3%)

LV GCS -23,0% (95% CI: -24,3% bis -21,7 %)

LV GRS 34,1 % (95% CI: 28,5% bis 39,7%)(44).

LV GLS und LV GCS waren darunter die konkordantesten Parameter. Der GRS zeigte die größte Heterogenität und im Gegensatz hierzu der GLS die größte Homogenität (44). Die Werte geben eine gute Orientierung, wie weitere Studien zeigten. Allerdings zeigt sich bei der detaillierten Betrachtung der Ergebnisse der hier eingeschlossenen Studien, dass die globalen Strainparameter der beiden größten Studien voneinander abweichen. In Anbetracht aller 17 Studien stellt sich ein inhomogenes Bild dar. Die Heterogenität von LV GCS und LV GRS wurde zum einen auf die kleinen Studienpopulationen zurückgeführt, zum anderen auf die Art der Generierung der Strainwerte. So variierte zum Beispiel die Anzahl der Schichten, die zur Berechnung herangezogenen wurden (44). Dass dies einen Einfluss hat, wiesen 2021 Lim et al. nach. So unterschieden sich die Strainparameter, wenn sie aus drei oder allen verfügbaren Schichten generiert wurden. Insbesondere der LV GCS und der LV GRS sind davon abhängig (45).

Ein weiterer Grund für die Inhomogenität des LV GRS liegt in der hohen Abhängigkeit der Ergebnisse vom Untersucher (Interobserver-Variabilität). Der Einfluss unterschiedlich eingezeichneter Endo- und Epikardgrenzen auf den LV GRS fällt im Vergleich zu den beiden anderen Strainrichtungen aufgrund der kurzen Strecke (der Myokarddicke) groß aus (44).

Neben den technischen Faktoren konnte ein Einfluss des Geschlechts oder des Alters von Vo et. al. nicht festgestellt werden. Andere Studienergebnisse wie die von Peng et. al und weitere konnten hingegen geschlechtsabhängige Unterschiede nachweisen.

So berichtet Peng et. al. höhere Werte für Frauen beim LV GLS und LV GRS (37, 38, 46, 47).

Seitens der Autoren wurde ein Einfluss der ethnischen Herkunft der untersuchten Studienpopulation vermutet. In einer von Peng et al. 2018 veröffentlichten Studie konnte dies nicht nachgewiesen werden. Der LV GCS und LV GLS dieser gesunden chinesischen Kohorte entsprachen in etwa den Ergebnissen der zuvor zitierten Metaanalyse. Ausgenommen hiervon war der LV GRS, der stark von den Ergebnissen abweicht, da Peng et al. die Papillarmuskulatur zum Myokard zuordnete. Abschließend ist der Einfluss der Bevölkerungsgruppe zum jetzigen Zeitpunkt aber noch nicht geklärt.

Zuletzt veröffentlichten Kawel-Boehm et. al. eine aktualisierte Zusammenfassung von Referenzwerten, die ebenfalls auf Basis von hier aufgeführten Studien wie Peng, Liu und Andre et al. basierte. Sie erstellten softwareübergreifende Normwerte für den globalen Strain, definierten erstmals Werte des LV GCS und LV GLS über -14% als krank. Die Ergebnisse sind in der folgenden Tabelle zu sehen (29).

Parameter	Männer			Frauen		
	Mittelwert	SD	LL-UL	Mittelwert	SD	LL-UL
GCS (%)	-20,9	3,2	-27,2 – 14,6	-22,7	3,3	-29,2 – 16,2
GLS (%)	-19,4	3,3	-26,1 – 12,7	-21,4	3,6	-28,7 – 14,2

SD: Standardabweichung, LL-UL: Unter-Obergrenze

Tabelle 1: Strainreferenzwerte auf Basis von 295 gesunden Männern und 301 gesunden Frauen aus Boehm et. al. (29)

Im Unterschied zu den zuvor diskutierten Studien fällt auf, dass sowohl der GLS als auch der GCS niedriger eingeschätzt werden. Dementsprechend liegt eine große Bandbreite gesund klassifizierter Strainwerte vor, insbesondere wenn die oberen und unteren Grenzen (LL-UL) betrachtet werden.

Die hier aufgeführten Kritikpunkte zeigen, dass Standards noch nicht ausreichend etabliert sind. Aspekte wie die Anzahl der herangezogenen Schichten oder die Frage nach softwareübergreifender Vergleichbarkeit beeinflussen die Ergebnisse, sodass weiterhin eine Ungewissheit hinsichtlich der Übertragbarkeit und Interpretation der Daten herrscht. Neben den oben genannten Unsicherheiten ist auch der menschliche Einfluss auf die Berechnungen von Relevanz, weshalb diesbezüglich nachfolgend die aktuelle Studienlage diskutiert wird.

#### 2.4.4 Intra- und Interobserver-Variabilität

Die Intraobserver Variabilität beschreibt, ob die Ergebnisse einer Messmethode abweichen, sofern sie von einem Untersucher zu zwei unterschiedlichen Zeitpunkten gemessen wurden. Die Interobserver Variabilität beschreibt, wie abhängig ein Untersuchungsverfahren von der Person des Untersuchers ist. Sie wird bestimmt durch die Ergebnisse einer Untersuchung die von zwei oder mehr unterschiedlichen Untersuchern generiert werden. Beide Kennzahlen geben Hinweise über die Güte und Verlässlichkeit eines Untersuchungsverfahrens.

Grundsätzlich zeigen sich gemäß der aktuellen Studienlage für die globalen Strainparameter GCS und GLS eine niedrige Variabilität, wie mehrere Studien belegen (37, 38, 48). Der GRS zeigte hingegen die höchste Variabilität. Dass segmentale Strainparameter indes eine hohe Inter- und Intraobserver-Variabilität vorweisen, wurde wiederholt bestätigt. Somit sollten diese aktuell nur als ergänzende Information betrachtet werden (49, 50).

Neben den Varianzen der Strainparameter an sich wurde durch Studien belegt, dass Intra- und Interobserver Variabilität von der eingesetzten Postprocessing Software abhängig sind. So ergaben sich für Cvi42 (Version 5.5) gemäß Intraklassenkorrelations-Koeffizienten (ICC) nach Konvention exzellente ICC Ergebnisse ( $>0,95$ ) und einen Variationskoeffizienten (CV) von  $> 10\%$ , wobei MEDViso in der Studie von Barreiro Perez et. al. ICCs  $> 0,99$  und CV  $< 5,2\%$  zeigte. Die Unterschiede von MEDViso zu den anderen Programmen wie TomTec fielen noch deutlicher aus. Als Ursache ist der von MEDViso verwendete Non Rigid Algorithmus zu nennen, der im Vergleich allein die Bearbeitung der Enddiastole vorsieht. Die Anzahl der Variablen, die das Ergebnis beeinflussen, nimmt hiermit ab und ermöglicht höhere Übereinstimmungen (51, 52).

Neben der Software ist die Erfahrung des Untersuchers ein Einflussfaktor auf die Inter- und Intraobserver-Variabilität. In einer 2019 veröffentlichten Studie wurde nachgewiesen, dass gezieltes Training die Reproduzierbarkeit und die Datengenauigkeit erhöht. Hierbei wurde eine Gruppe von gesunden und kranken Probanden durch zwei Untersucher mittels der drei Softwareplattformen (TomTec, Medis und Cvi42 Circle) vor und nach einem Training untersucht. Der LV GCS und der LV GLS waren am wenigsten von der Erfahrung des Untersuchers abhängig. Beide Werte konnten jedoch durch das Training von gut auf exzellent gemäß ICC gesteigert werden. Die Intraobserver Variabilität konnte ebenfalls verbessert werden (53).

Dies verdeutlicht zum einen den Einfluss von Erfahrung oder Training auf die Ergebnisse, lässt zum anderen aber auch vermuten, dass durch semiautomatische Konturerkennung bereits ohne Erfahrung schon gute Ergebnisse erzielt werden könnten.

Dass Bilderkennungssoftware und insbesondere ihre Güte der semiautomatischen Konturzeichnung ein wichtiger Einflussfaktor sind, konnte in einer 2021 veröffentlichten Studie nachgewiesen werden. Vollautomatisch generierte Berechnungen von quantitativen Funktionsparametern wie EF, EDV und ESV sind demnach vergleichbar mit denen eines sehr erfahrenen Experten. Ein unerfahrener Untersucher war dem Computerprogramm unterlegen (54). Zwar liegen zu Strainwerten derzeit keine vergleichbaren Studien vor, allerdings sind sie mit dem EDV und ESV über die endokardiale Kontur verbunden. Dies legt nahe, dass Reproduzierbarkeit zunehmend durch verbesserte Algorithmen erhöht werden kann. Gleichzeitig bedeutet dies aber auch, dass lediglich Ergebnisse zur Inter- und Intrarater Reliabilität nur bei gleicher Softwareversion miteinander zu vergleichen sind.

In einer Studie wurden die drei Softwarepakete (Cvi42 5.10.3, SuiteHeart Neosoft 5.0.0 und DRA TrufiStrain 2.1) hinsichtlich des GLS untersucht, wobei alle drei Pakete ebenfalls hervorragende Intra- und Interrater Reproduzierbarkeit nachwiesen. Der ICC des GLS, der aus Bilddaten der Schnittebenen des Zwei- und Vierkammerblicks generiert wurde, variierte zwischen 0.86 - 1.0 (SuiteHeart ICC 0.96 – 1.00, Cvi42 ICC 0.86 – 0.99, DRA-TrufiStrain ICC 0.88 – 1.00) (55).

Hinsichtlich der Inter- und Intraobserver Variabilität zeigen die zuletzt veröffentlichten Studien gute bis exzellente Ergebnisse. Aufgrund der häufigen Anzahl an Updates der Softwarepakete muss beachtet werden, dass die jeweiligen ICC-Werte nicht direkt miteinander vergleichbar sind. Der Automatisierungsgrad und insbesondere die Fortschritte der Konturerkennung nehmen zu, sodass auch die Reproduzierbarkeit Verbesserungen zeigt und Unterschiede in der Anwendererfahrung ausgeglichen werden können (54). Die Frage nach der Verbesserung durch den Automatisierungsgrad stellt somit eine relevante Untersuchungsfrage dar.

#### 2.4.5 Anbieterabhängige Variabilität (Intervendor-Variabilität)

Die Frage, inwiefern Softwareprogramme einen Einfluss auf die Quantifizierung von Funktionsparametern des linken Ventrikels haben, untersuchten erstmals Messali et al. im Jahr 2009. Sie verglichen die Übereinstimmung von zwei Softwareprogrammen (View Forum, Version 4.2 und Argus, Version Va60C) bei der Berechnung

linksventrikulärer Volumina und EF. Hierbei konnten keine signifikanten Unterschiede zwischen den zwei Softwareprogrammen festgestellt werden. Im Jahr 2019 griff eine andere Studie erneut diese Frage auf und verglich dafür erstmals die drei Softwareprogramme Cvi42, Argus und Medis miteinander. Bei der Betrachtung der LV EF, EDV, ESV und LVM konnten Sie genauso wie Messali et al. keine Unterschiede feststellen (56, 57).

Ob dies auf globale Strainparameter, die mittels unterschiedlicher Postprocessing Software generiert wurden, übertragbar ist, wurde erstmals durch Schuster-Stahnke et al. 2015 untersucht. Eine kleine Patientengruppe aus 10 Probanden wurde mit den Programmen TomTec und Cvi42 Circle gemessen. Zwar zeigte der LV GCS im direkten Vergleich mit dem LV GRS die konkordantesten Ergebnisse. Jedoch generierte Cvi42 signifikant niedrigere LV GCS Werte als TomTec. Somit stellten sie fest, dass grundsätzlich die Vergleichbarkeit von Strainparametern unterschiedlicher Programme nicht gegeben sei (6).

Drei Jahre später wurde die Untersuchungsfrage von Barreiro-Perez et. al erneut aufgegriffen. Es wurde der linksventrikuläre Globale Strain aller drei Richtungen mit nun vier kommerziell erwerbbaaren Softwareprogrammen (TomTec, Medis, Circle Cvi42 und MEDViso) einer heterogenen Gruppe aus 45 Patienten und zusätzlich noch die Inter- und Intraobserver Variabilität bestimmt. Das Patientenkollektiv bestand aus 15 gesunden und 30 kardial vorerkrankten Patienten. Die vier Programme unterschieden sich zum Teil deutlich im Arbeitsaufwand, wobei Cvi42 die größte Anzahl an Konturkorrekturen benötigte. Sie belegten die Beobachtung von Schuster et al., dass die Programme signifikante Unterschiede insbesondere bei LV GLS und LV GRS zeigten. Der LV GCS schien am besten vergleichbar zu sein. Hinsichtlich der Intra- und Interobserver Reproduzierbarkeit von LV GLS und LV GCS schnitten alle Softwarepakete gut oder exzellent ab ( $ICC > 0,75$ ). Im Detail zeigte MEDViso die geringste Variabilität, was wie zuvor beschrieben mit dem verwendeten Algorithmus verbunden sein kann. Auffallend ist, dass der LV GRS die geringste Reproduzierbarkeit aufwies (51, 58).

Eine im Jahr 2021 vorgestellte Studie untersuchte die Sensitivität zur Erkennung einer kardialen Pathologie in Abhängigkeit von der gewählten Software. Außerdem wurde betrachtet, inwiefern die Krankheitsschwere differenziert werden kann. Die Krankheitsschwere wurde anhand der Hospitalisierungsrate oder des eingetretenen Todes nach der Untersuchung definiert. Hierfür untersuchten sie allein den GLS bei 45 HFpEF Patienten mit 26 Personen einer gesunden Kontrollgruppe durch drei

Softwarepakete (SuiteHeart 5.0.0, Neosoft Inc., Cvi42 5.10.3 und DRA TrufiStrain 2.1 Siemens Healthineers). Der GCS wurde nicht berichtet, da im Vergleich zum GLS die Sensitivität zur Differenzierung zwischen HfpEF Patienten und der Kontrollgruppe dem GLS unterlag. Bei allen drei Softwarepaketen lag ein signifikanter Unterschied zwischen Kontroll und HfpEF- Gruppe vor. Allein SuiteHeart konnte die deutlich schwerer Erkrankten von den leicht Erkrankten unterscheiden. Insbesondere der 4-Kammer GLS von SuiteHeart war signifikant assoziiert mit dem Outcome des Patienten (55). Während alle drei Programme Gesund von Krank unterscheiden konnten, waren die absoluten Werte der drei Softwareprogramme, wie zuvor beschrieben, nicht softwareübergreifend miteinander vergleichbar (exemplarisch hierzu der GLS der gesunden Kontrollgruppe 4Chamber View SuiteHeart  $-19,4 \pm 3,1$ ; Cvi42  $-17,1 \pm 3,0$ ; DRA-TrufiStrain  $-15,6 \pm 3,0$ ) (55).

Dies lässt daraus schließen, dass die untersuchte Software die generierten Strainwerte maßgeblich beeinflusst und der Einfluss des Untersuchers auf das Ergebnis geringer zu werten ist. Diese These wurde in einer weiteren Studie untersucht. Für 10 zufällig ausgewählte Patienten wurde der GLS aus manuellen vs. automatisch generierten Konturen durch SuiteHeart und Cvi42 berechnet. Hierbei konnte kein Unterschied festgestellt werden. Daraus schlussfolgerten Zhang et al., dass die Leistungsunterschiede der Programme durch die Tracking Algorithmen und nicht durch die Kontursegmentierung bestimmt werden (55).

Alle bisher veröffentlichten Studien zum Vergleich von Softwarepaketen konnten zeigen, dass die Vergleichbarkeit der Parameter nur eingeschränkt möglich ist, wobei der Vergleich des GRS nicht sinnvoll zu sein scheint. Unterschiede in der Kontursegmentierung, die Unterstützung durch automatische Konturerkennung als auch zum Beispiel die Definition von Enddiastole und Endsystole können einen Einfluss auf die Vergleichbarkeit der Parameter haben. Weiterhin unterscheiden sich diverse Softwareprogramme in den verwendeten Algorithmen. Die Prinzipien der Algorithmen wie zum Beispiel Optical Flow, non-rigid algorithm (MEDViso) oder auch der deformable registration algorithm von TrufiStrain haben einen Einfluss auf die absoluten Werte der globalen Strainparameter und auch auf deren Reproduzierbarkeit.

### **3. Material und Methoden**

#### **3.1 Probandenkollektiv**

Das untersuchte Probandenkollektiv stammt aus zwei an der Universitätsmedizin Mainz durchgeführten Studien: der MaiCo-MR-Studie aus dem Jahr 2018 und der Studie zur Ermittlung von Mapping-Referenzwerten bei Herzgesunden aus dem Jahr 2016 folgend „T1T2T3-Studie“ genannt. Die Untersuchung der Proband\*innen wurden im Vorfeld durch die Ethikkommission der Landesärztekammer Rheinland-Pfalz genehmigt (T1T2T3-Studie: (Referenznummer 837.196.13 (8881-F) MaiCo-MR-Studie 837.196.13 (8881-F)). Ausschließlich gesunde Proband\*innen wurden in die Studien eingeschlossen. Ausschlusskriterien waren kardiale Risikofaktoren wie Nikotinabusus oder Diabetes mellitus sowie kardiale Vorerkrankungen in der Anamnese. Aus der MaiCo-MR-Studie wurden fünf Proband\*innen ausgeschlossen. Zwei aufgrund einer schlechten Bildqualität und drei weitere aufgrund unvollständiger bzw. abgebrochener CINE-Sequenzen. Somit wurden 65 der 70 Proband\*innen der MaiCo-MR-Studie ausgewertet. Zur Schaffung einer möglichst repräsentativen gesunden Kohorte wurde das Kollektiv durch 33 Proband\*innen der T1T2T3-Studie erweitert. Bei SuiteHeart mussten zwei Probanden und bei TrufiStrain musste ein weiterer Proband aufgrund von Softwarefehlern bei der Strainberechnung ausgeschlossen werden (siehe Tabelle 3). Da Referenzwerte für die jeweiligen Programme erstellt werden sollten, wurde festgelegt alle Werte in die Auswertung mit einzubeziehen. Für die Intrarater Reliabilität wurde nach der Erstmessung eine zufällige Auswahl von 12 der 98 Proband\*innen durch Untersucher 1 (R1) erneut in allen drei Programmen ausgewertet.

#### **3.2 Kardiale Magnetresonanztomographie**

Die T1T2T3-Studie wie auch die MaiCo-MR-Studie wurden jeweils an einem 3 Tesla MRT der Firma Siemens Healthineers durchgeführt (MaiCo-MR-Studie mit einem Magnetom Prisma®; T1T2T3-Studie mit einem Magnetom Skyra®).

Ventral wurde das Signal von einer 18-Kanal Body-Spule (Body 18®, Siemens Healthineers) und dorsal von im Untersuchungstisch integrierten Body-Spine-Spulenelementen gemessen. Die CINE-Bildgebung erfolgte mittels einer SSFP-Sequenz. Zur Triggerung diente ein kompatibles MRT-EKG-System sodass die erzeugten Bilddaten retrospektiv R-Zacken zugeordnet werden konnten. Im Rahmen weiterer Untersuchungen der beiden Studien wurden jedem Patienten

gewichtsadaptiert (0,1 mmol/kgKG) Kontrastmittel (Gadovist®) appliziert. Für die CINE-Bildgebung war dies nicht notwendig.

Ein Herzzyklus besteht aus 25 akquirierten Phasen pro Zyklus mit einer zeitlichen Auflösung von 25 – 40 ms.

---

**Bildparameter der MaiCo-MR- Studie und der T1T2T3-Studie**

---

	Maio-MR-Studie	T1T2T3-Studie
Sequenztyp	bSSFP Cine	bSSFP Cine
EKG-Triggerung	retrospektiv	retrospektiv
Temporal Resolution (ms)	38,8	39,8
Repetition Time TR (ms)	37,7	39,2
Echo time TE (ms)	1,42	1,45
Field of View (mm)	360	360
Berechnete Herzphasen	25	25
Number of Segmentations	12	12
Slice thickness (mm)	8.0	8,0
Flip angle (°)	60	46
Voxel size (mm <sup>3</sup> )	1.5 x 1.5 x 8.0	1,5 x 1,5 x 8,0
Acceleration factor	3 (GRAPPA)	3 (GRAPPA)

bSSFP, balanced steady-state free precession; EKG Elektrokardiogramm; GRAPPA, Gene Realized Autocalibrating Partial Parallel Acquisition

---

Tabelle 2: Bildparameter der MaiCo-MR-Studie und der T1T2T3-Studie

### 3.3 CMR Analyse

Die CMR-Datensätze wurden mit drei Postprocessing CMR-FT-Software-Paketen untersucht. Es handelte sich hierbei um: Cvi42 Circle® Version 5.9.3 (Circle Vascular Imaging, Canada), SuiteHeart Version 4.0.6 (Neosoft LLC, Pewaukee, Wi, USA) und TrufiStrain 2.0 (Siemens Healthcare, Medical Imaging Technologies, Princeton, NJ, USA). Die erzeugten Bilddatensätze aus der MaiCo-MR-Studie sowie der T1T2T3-Studie wurden importiert. Hierfür wurden die SSFP CINE-Sequenzen der Kurzachsen (SAX) sowie die Langachsendarstellungen des Zwei- und Vierkammerblicks ausgewählt. Zwei geschulte Untersucher analysierten die Bilddaten unter Berücksichtigung der Richtlinien der Society for Cardiovascular Magnetic Resonance (SCMR), um die Interrater Reliabilität bestimmen zu können (59). Bei den Programmen SuiteHeart und TrufiStrain mussten aufgrund von Softwarefehlern sowie fehlenden Vergleichswerten für die Interrater Reliabilität Probanden aus der Auswertung ausgeschlossen werden. Final in die Auswertung der Interrater Reliabilität sind bei

SuiteHeart 90 von 98 Probanden und bei TrufiStrain 95 von 98 Probanden in die Auswertung eingeflossen. Um die Stichprobe möglichst groß zu halten, wurde festgelegt, die jeweilige Anzahl der Probanden in die Auswertung mit einzubeziehen. Nach einer Pause von zwei Monaten und zur Reduktion eines möglichen Recall-Bias erfolgte die erneute Messung durch R1 zur Bestimmung der Intrarater- Reliabilität. Zur Berechnung der volumetrischen Daten wie z.B. EF, EDV und Myokardmasse wurden die Kurzachsen ausgewählt. Alter, Geschlecht, Größe und Gewicht der Patienten wurden den radiologischen Anforderungsbögen entnommen und aus Größe und Gewicht wurde die Körperoberfläche (KÖF) berechnet. Zur besseren Vergleichbarkeit wurden die klinischen Daten wie EDV und LVM auf die KÖF bezogen. Es bildet sich der Index EDVi ( $EDV/KÖF$ ) und LVMI ( $LVM/KÖF$ ).

### 3.3.1 Cvi42

Für die Durchführung der Berechnung mit Cvi42 wurden zunächst semiautomatisch endokardiale Konturen eingezeichnet. Die vorgezeichneten Konturen wurden von den Untersuchern hinsichtlich Exaktheit kontrolliert und bei Bedarf angepasst. Durch die Ausbesserung der Konturen wurden bei Bedarf Endsystole (ES) und/oder die Enddiastole (ED) neu definiert. Hierfür wurde die Phase ausgewählt, welche visuell das geringste (Endsystole) beziehungsweise das größte Blutvolumen (Enddiastole) zeigt (26).

Zu den Zeitpunkten der ED und ES wurden daraufhin die epikardialen Konturen eingezeichnet. Hierbei wurde als erste basisnahe Schicht in der Kurzachse diejenige gewählt, die mindestens 50% Myokard zeigte, bzw. diejenige Schicht ausgewählt, die in der Langachse dem Linksventrikulären Ausflusstrakt am nächsten liegt. Alle Schichten, die hiervon ventrikelfern lagen, wurden von der Berechnung ausgenommen. Ausbesserungen der Konturen waren hauptsächlich basisnah oder auch am Apex erforderlich. Apexnahe Schichten wurden häufig fälschlicherweise dem Magen zugeordnet und erforderten Korrekturen.

Durch den Algorithmus wurden die Insertionspunkte des rechten Ventrikels eingezeichnet. Diese markieren den Teil der linksventrikulären Kurzachse, an dem der rechte Ventrikel beginnt und dienen der Segmentierung, die insbesondere für die Generierung regionaler Strainparameter notwendig ist. Daraufhin wurde das Volumen durch die Summierung der endokardialen Konturen aller Kurzachsen berechnet.

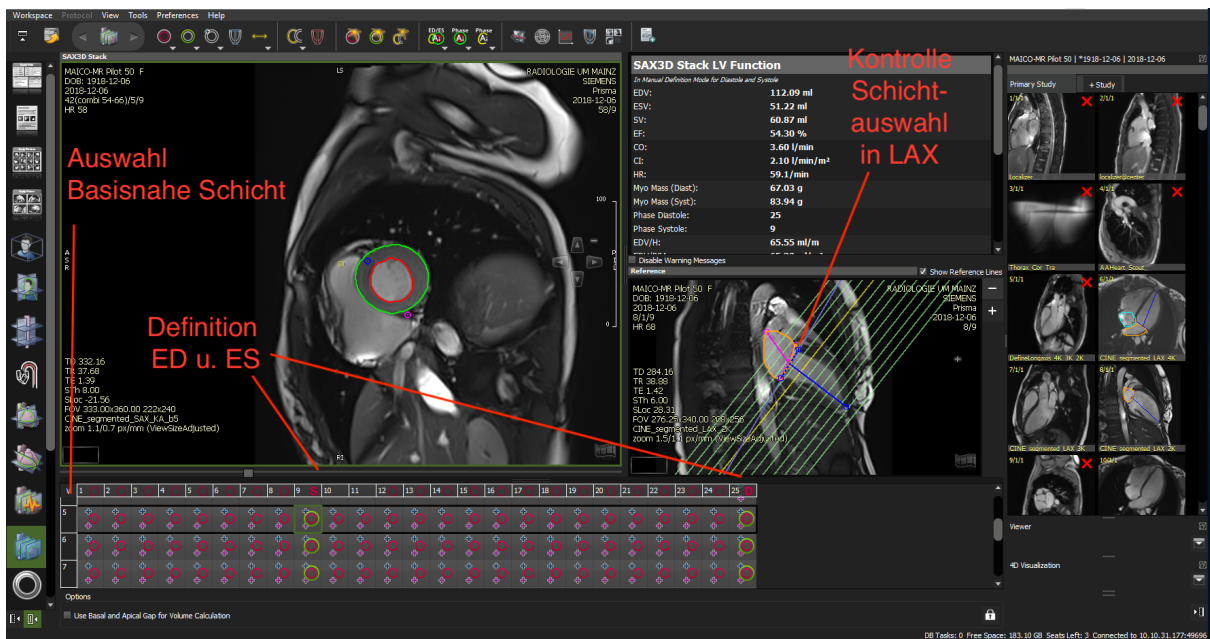


Abbildung 6: Cvi42 Bearbeitungsfenster Strainberechnung

### 3.3.2 SuiteHeart

Ähnlich dem Vorgehen bei Cvi42 erfolgte bei SuiteHeart semiautomatisch die Generierung der epi- als auch der endokardialen Konturen in der SAX. Gemäß Manual wurden im „CINE“ – Modus die Konturen kontrolliert und ausgebessert. Danach überprüfte der Untersucher die vom Programm definierten ED und ES und passte diese gegebenenfalls an (60). Es erfolgte die Definition der basalen Schicht mit der Einzeichnung der basal gelegenen Linie. Im Anschluss erfolgte die automatische Segmentierung der Langachse mit anschließender Kontrolle.

### 3.3.3 TrufiStrain

Im Gegensatz zu den beiden Programmen SuiteHeart und Cvi42 gab es einige Unterschiede. TrufiStrain basiert auf einem DR-Algorithmus (Deformation registration algorithm). Durch die Konturen wird eine Fläche definiert, in der auf Pixel-Ebene die Bewegung des Ventrikels verfolgt wird (48). Bei TrufiStrain wurden sowohl die epi- als auch endokardialen Konturen allein in der ED eingezeichnet. Die Fortführung der Konturen über die Phasen des Herzzyklus wurde von der Software übernommen. Eine Bearbeitung der semiautomatischen Konturen ist bei diesem Programm ausschließlich in der Phase der ED möglich. Die Konturkorrektur erfolgt rein manuell.

Bei der Verwendung der Korrekturwerkzeuge wurde in der Regel auf „Polygon Drawing“ oder „Circle Drawing“ zurückgegriffen. Eine Auswahl der ED oder ES war nicht möglich.

Ein Ausschluss der basisnahen Schicht aufgrund einer Überschneidung mit dem linksventrikulären Ausflusstrakt wie bei den beiden zuvor beschriebenen Programmen war sehr selten notwendig. In den vorhandenen Langachsen erkennt der Algorithmus automatisch die Mitralklappenebene und ordnet daraufhin den Anteil der Kurzachse, der unterhalb der Mitralklappe liegt, dem Volumen zu (61). Die in der folgenden Abbildung grün eingezeichneten Flächen demonstriert, die zur Volumenberechnung einberechneten Kurzachsenschnitte.

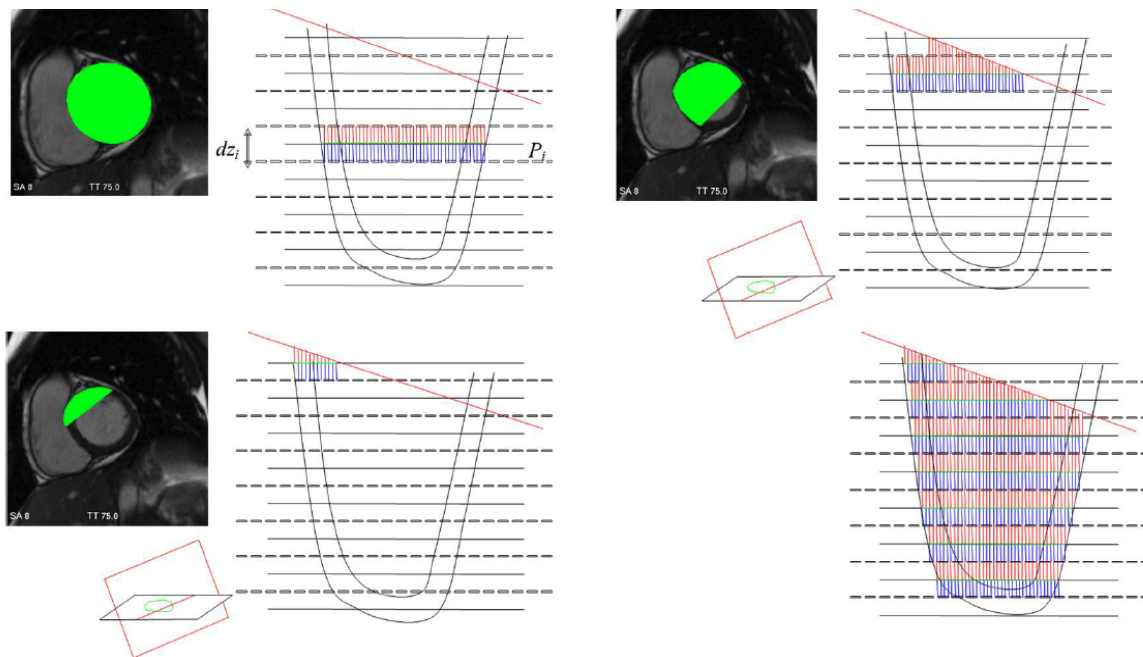


Abbildung 7: TruFiStrain Volumenberechnung, Basisschicht durch LAX definiert (61).

Für die Definition der Langachsen gab es keinen Tracking-Algorithmus, somit konnten endo- und epikardiale Konturen ausschließlich manuell eingezeichnet werden. Dies erfolgte analog zur Kurzachse ebenfalls nur in der ED. Die Fortführung über die weiteren Phasen erfolgte durch den Algorithmus. Eine Festlegung von ED oder ES ist auch in der Langachse nicht möglich.

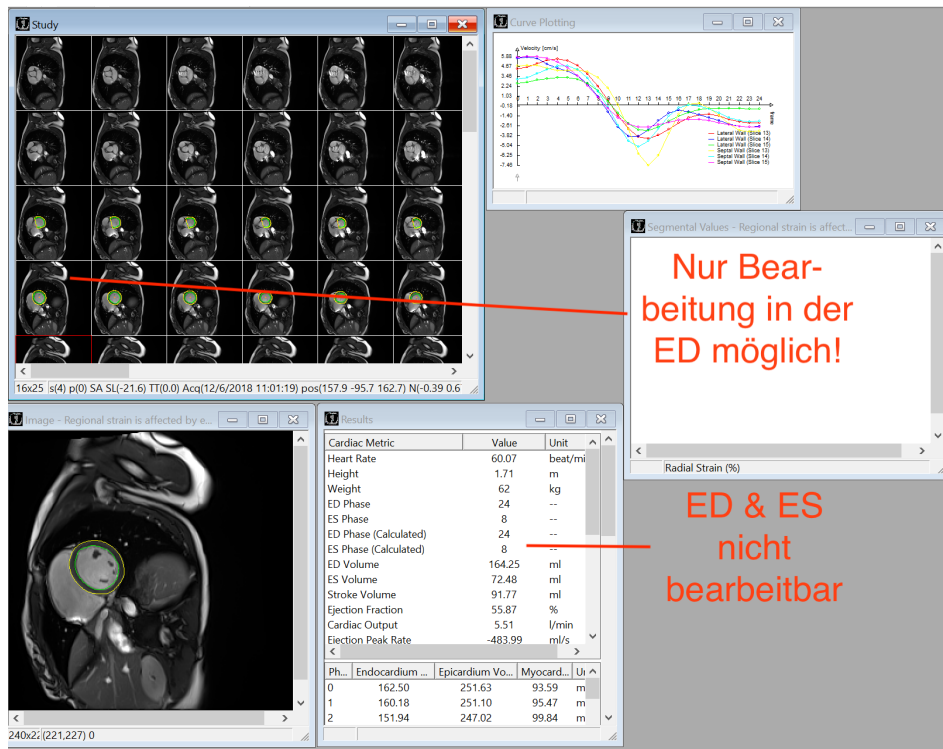


Abbildung 8: TruFiStrain- Konturbearbeitung (epi- und endokardiale Konturen)

Für die Einzeichnung der Konturen in der Langachse konnte ausschließlich das Polygon-Tool verwendet werden, das lediglich die Einzeichnung vieler gerader Linien zur Einzeichnung der endo- und epikardialen Konturen zu ließ. Bei der Verwendung der anderen Werkzeuge, die eine deutlich genauere Einzeichnung der Konturen ermöglichte (wie bei den anderen Programmen üblich), wurden die eingezeichneten Konturen durch das Programm nicht weiter prozessiert.

Siehe die folgende Abbildung:

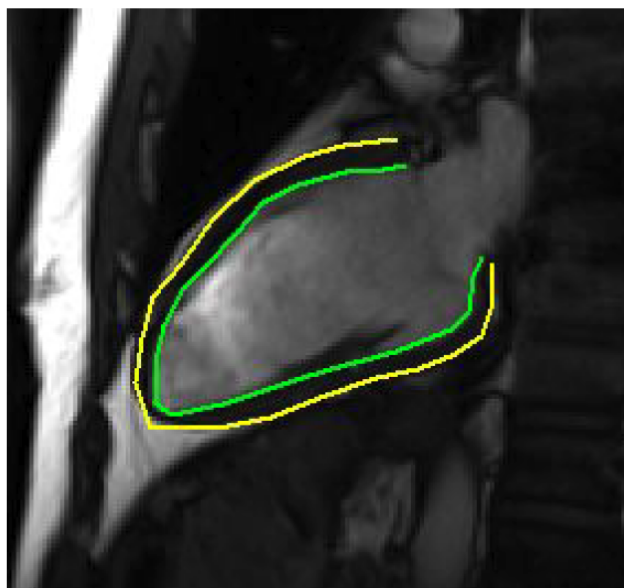


Abbildung 9: Einzeichnung der epi- und endokardialen Konturen in der Langachse Polygon Tool (TruFiStrain)

Für alle drei Programme erfolgte die Berechnung der LVM in ED und ES automatisch. Durch die Subtraktion der epi- von der endokardialen Kontur in ED und ES sowie unter Berücksichtigung der Schichtdicke und der Anzahl der Schichten wurde das daraus berechnete Volumen mit dem spezifischen Myokardgewicht von 1,05 g/ml multipliziert. Genauso wie bei der Berechnung der Volumina wurde die Papillarmuskulatur dem Ventrikelvolumen zugeordnet und somit von der Berechnung des Myokardgewichts ausgenommen. Nach dieser Konvention sind die meisten Normwerte definiert.

Die Berechnung der GCS und GRS erfolgte aus den Kurzachsenschnitten. Alle Schichten, die einen vollständig geschlossenen Ring in der ED und ES zeigten, wurden in die Berechnung mit einbezogen. Ein Anschnitt des linksventrikulären Ausflusstrakts führte zum Ausschluss der Schicht. Diese Auswahl erfolgte durch den Algorithmus. Für die Berechnung des GLS wurden die beiden langen Achsen des Zwei- und Vierkammerblicks verwendet. Hierbei war insbesondere bei Cvi42 darauf zu achten, dass die Definition der ED und ES in beiden Langachsen zum gleichen Zeitpunkt wie in den Kurzachsen gewählt wurden.



Abbildung 10: Cvi42 Bearbeitungsfenster Strain

Dies war bei SuiteHeart und TruFiStrain nicht notwendig, da durch den Algorithmus eine vollständige Einzeichnung der epi- und endokardialen Konturen über den gesamten Herzzyklus erfolgte und ein Abgleich automatisch erfolgte.

### 3.4 Statistische Auswertung

Für die statistische Auswertung wurde das Programm IBM SPSS Statistics for Windows, Version 27.0. (Armonk, NY: IBM Corp) verwendet. Die Normalverteilung der Daten wurde mittels eines Histogramms und eines Shapiro-Wilk-Tests überprüft. Hierbei zeigte sich, dass die folgenden Parameter nicht einer Normalverteilung unterlagen (Cvi42: GRS, SuiteHeart: ESVi, EDVi; TrufiStrain: EDVi, ESVi, GRS). In den aufgeführten Tabellen werden daher die Mediane verwendet. Sie wurden als Median  $\pm$  Interquartils-abstand (Q3-Q1) angegeben. Da diese von den Mittelwerten nur geringfügige Unterschiede zeigten, zeigte sich keine Auswirkung auf die statistische Auswertung, weshalb fortlaufend die Mittelwerte als Bezugsgröße verwendet wurden.

LV GCS, LV GLS und LV GRS wurden mit ihren jeweiligen Prozentzahlen in Klammern dargestellt.

Normalverteilte kontinuierliche Variablen wurden als Mittelwert  $\pm$  Standardabweichung angegeben. Als globale Strainparameter wurde der Peak Global Strain der jeweiligen Richtung für Cvi42, TrufiStrain und SuiteHeart berichtet. Für den globalen radialen Strain berichtet SuiteHeart zwei Werte: den Peak und den Mid radial Strain. Bei SuiteHeart wurde der Mid radial Strain verwendet, da dieser scheinbar den Definitionen der beiden anderen Softwareprogrammen am ehesten folgt.

Unterschiede der klinischen Daten sowie der globale Strain zwischen den drei Anbietern wurden mit einer einfaktoriellen Varianzanalyse mit Messwiederholung (rmANOVA) sowie einem post-hoc ungepaarten T-Test mit Bonferroni-Korrektur analysiert. Als ein lokales Signifikanzniveau wurde ein p- Wert von  $< 0,05$  festgelegt. Für paarweise Vergleiche wie die LVMI wurde ein gepaarter T-Test durchgeführt. Für die Definition von Referenzwerten wurde ein Mittelwert aus den Ergebnissen beider Untersucher je Strainparameter und Programm gebildet. Zur Prüfung auf geschlechtsabhängigen Einfluss wurde ein Mann-Whitney-U-Test für unabhängige Stichproben durchgeführt.

Für den Vergleich der globalen Strains der jeweiligen Softwareprogramme wurden Bland-Altman-Diagramme erstellt. Da bisher keine Postprocessing Software als Goldstandard festgelegt wurde und es bisher keinen definierten Normwert für die Strainwerte gibt, bildet der Mittelwert der drei Anbieter die hier gewählte Referenz. Die Differenz des jeweiligen Strainwertes zum gepoolten Mittelwert wurde mit 95 % Konfidenzintervall dagegen aufgetragen. Der Übereinstimmungsbereich (Limits of

Agreement) sind als Mittelwert der Differenz  $\pm 1,96 \times SD$  definiert. Hierbei wurden die Daten eines Untersuchers R1 herangezogen und nicht ein aus R1 und R2 gebildeter Mittelwert gegenüber dem gepoolten Mittelwert der drei Anbieter. Ein durch Nivellierung erzeugter Bias sollte hierdurch vermieden werden.

Die Interobserver-Variabilität wurde ebenfalls visuell mit Bland-Altman-Diagrammen mit 95% Konfidenzintervall ausgewertet. Hierfür wurde der Mittelwert aus den Ergebnissen beider Untersucher (R1 und R2) gebildet und die Differenz von R1 zu R2 dagegen aufgetragen.

Für die Ermittlung der übergeordneten Übereinstimmung der drei Softwareprogramme wurde ein „Two way mixed effects model, single rater, consistency“ ICC verwendet. Um die Übereinstimmung der beiden Untersucher wie auch die Intrarater-Reliabilität zu ermitteln, wurde ein „Two way random effects model, single rater, absolute agreement“ ICC verwendet. Im Unterschied zur Interrater-Reliabilität wurde die Intrarater-Reliabilität mit einer Stichprobe von 12 Probanden berechnet.

Die Ergebnisse des ICCs wurden wie folgt kategorisiert: ICC  $< 0,5$  schlecht;  $0,5 < \text{ICC} < 0,75$  moderat;  $0,75 < \text{ICC} < 0,9$  gut und  $\text{ICC} \geq 0,9$  exzellent (58).

Der Variationskoeffizient (CV) zur Beurteilung der Untersucher-Übereinstimmung wurde mittels der Standardabweichung der Differenz zum Mittelwert (R1; R2) in Bezug zum Mittelwert (R1; R2) ermittelt.

## 4. Ergebnisse

### 4.1 Gesunde Freiwillige

Das Durchschnittsalter der 98 freiwilligen Proband\*innen lag bei  $37 \pm 14$  Jahren mit einer Spannweite von 21 bis 72 Jahren. Da die MaiCo-MR-Studie mehrheitlich aus einem jungen Probandenkollektiv zusammengestellt war, sind die 20- bis 35-Jährigen stärker repräsentiert. Das Geschlechterverhältnis ist mit 47 Frauen (47%) und 51 Männern (52%) annähernd gleichverteilt. Weitere CMR basierte Werte für LV Volumina sind Tabelle 3: Ausgangsmerkmale und Vergleich der Volumetrie zu entnehmen. Alle hier präsentierten Daten entstammen Kurzachsen-Schnitten.

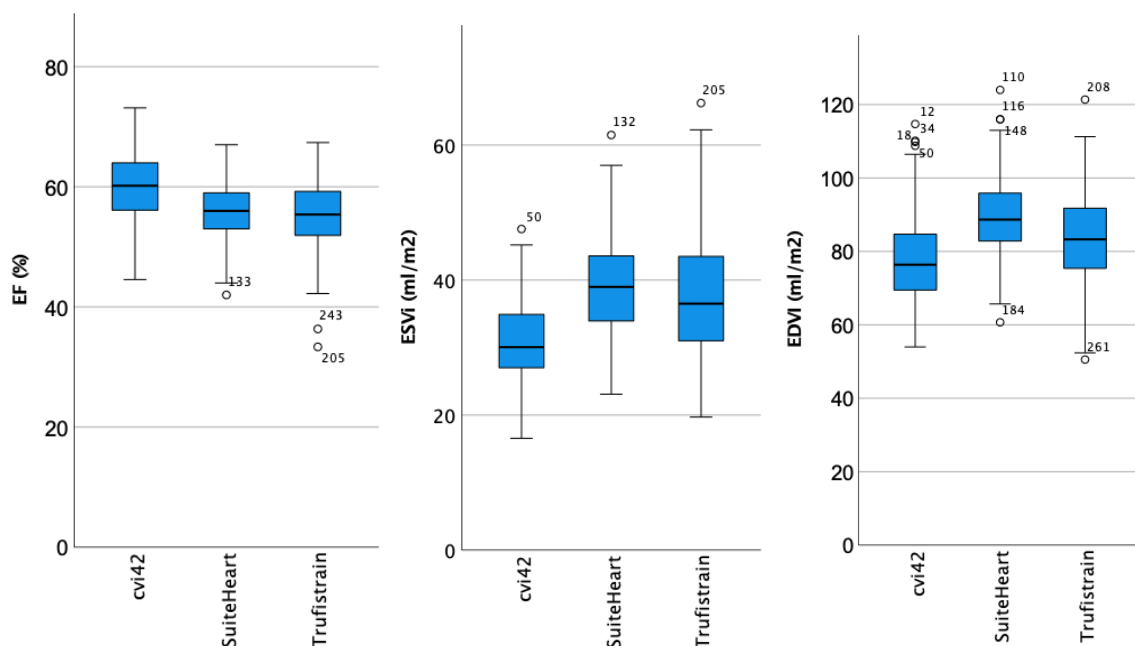


Abbildung 11: Boxplots EF, EDVi, ESVi (Kurzachse) in der Gegenüberstellung Cvi42, SuiteHeart, TrufiStrain

In der Gegenüberstellung der Ergebnisse der Volumetrieberechnung zeigten sich signifikante Unterschiede. So berechnete Cvi42 im Vergleich zu TrufiStrain und SuiteHeart eine signifikant höhere mittlere EF, während im direkten Vergleich zwischen TrufiStrain und SuiteHeart nach der posthoc Bonferroni-Korrektur kein Unterschied vorlag ( $p = 0,531$ ).

Passend zu den Ergebnissen der EF wurde mittels Cvi42 für den ESVi ein signifikant niedrigerer Mittelwert im Vergleich zu SuiteHeart und TrufiStrain berechnet ( $p < 0,01$ ), während die Mittelwerte der beiden zuletzt genannten keinen signifikanten Unterschied ergaben. Bei der Berechnung des EDVi variierten alle drei Programme signifikant voneinander, wobei Cvi42 den niedrigsten Mittelwert ergab. Den Boxplots kann entnommen werden, dass alle drei Programme trotz einer gesunden Kohorte

vereinzelt EDVI's > 100 ml/m<sup>2</sup> ausweisen. Weiterhin zeigten sich insbesondere bei TrufiStrain zwei Ausreißer mit einer EF < 40 %, die entweder ein hohes ESVI oder ein erhöhtes EDVi hatten. Diese Werte entsprechen laut Definition einer kranken Person. Da in den beiden anderen Programmen für diese Proband\*innen im direkten Vergleich nicht die gleichen Werte ermittelt wurden, ist davon auszugehen, dass es sich um Messfehler handelt und nicht um erkrankte Proband\*innen. Der LVMI zeigt einen signifikanten Unterschied zwischen den Programmen Cvi42 und TrufiStrain.

<b>Ausgangsmerkmale und Vergleich der Volumetrie (Untersucher 1)</b>				
	<b>Cvi42</b>	<b>SuiteHeart</b>	<b>DRA-TrufiStrain</b>	<b>P – Wert</b>
<b>Anzahl</b>	98 (98)	96 (98)	97 (98)	
<b>Alter (Jahre)</b>	30,6; 27	30,0 ;24	30,8; 29	
<b>(Männer)</b>	50	49	50	
<b>KÖF (m<sup>2</sup>)</b>	1,84; 0,3	1,84; 0,3	1,84; 0,3	
<b>EF (%)</b>	60,1; 7,97	56,0; 6,0	55,4; 7,4 <sup>b</sup>	<,001
<b>EDVi (ml/m<sup>2</sup>)</b>	76,4; 15,7	88,7; 13 <sup>a</sup>	83,3; 17 <sup>a</sup>	<,001
<b>ESVi (ml/m<sup>2</sup>)</b>	30,1; 8,1	39,0; 9,7	36,5; 12,7 <sup>b</sup>	<,001
<b>LVMI (g/m<sup>2</sup>)</b>	55,9; 14,9	/	61,8; 17,2	<,001 <sup>c</sup>
Daten zeigen Median; Interquartilsabstand (Q75-Q25)				
p-Wert der ANOVA Analyse zwischen den drei Gruppen				
<sup>a</sup> statistisch signifikanter Unterschied zwischen Cvi42 und SuiteHeart und DRATrufiStrain				
<sup>b</sup> statistisch kein signifikanter Unterschied zwischen SuiteHeart und TrufiStrain				
<sup>c</sup> p-Wert eines T-Tests für abhängige Stichproben				

Tabelle 3: Ausgangsmerkmale und Vergleich der Volumetrie

## 4.2 Globale Strainparameter im Softwarevergleich

Im folgenden Abschnitt werden die globalen Strainparameter im Softwarevergleich auf signifikante Unterschiede untersucht.

Globale Strainparameter Untersucher 1					
	Cvi42	SuiteHeart	DRA-TrufiStrain	Mittelwert	p- Wert
GRS (%)	31,7; 7,8	32,5; 7,3 <sup>a</sup>	29,1; 5,2	31,3; 6,0	<,001
GCS (%)	-18,3; 3,2	-17,2; 2,2	-15,8; 2,1	-17,4; 2,4	<,001
GLS (%)	-17,3; 3,9	-18,6; 2,9	-14,4; 2,9	-16,8; 2,8	<,001

Daten zeigen Median; Interquartilsabstand (Q75-Q25) von Untersucher 1  
p- Wert der rmANOVA zwischen den drei Gruppen  
<sup>a</sup>im post-hoc Vergleich zwischen Cvi42 und SuiteHeart zeigte sich kein signifikanter Unterschied

Tabelle 4: Globale Strainparameter Untersucher 1 im Softwarevergleich

Im Gruppenvergleich durch die rmANOVA zeigte sich ein signifikanter Unterschied. In den im Anschluss durchgeführten post-hoc Tests zeigte sich ein signifikanter Unterschied zwischen TrufiStrain und SuiteHeart bzw. Cvi42 ( $p < 0,001$ ). Zwischen Cvi42 und SuiteHeart gab es hingegen keinen signifikanten Unterschied ( $p = 0,092$ ).

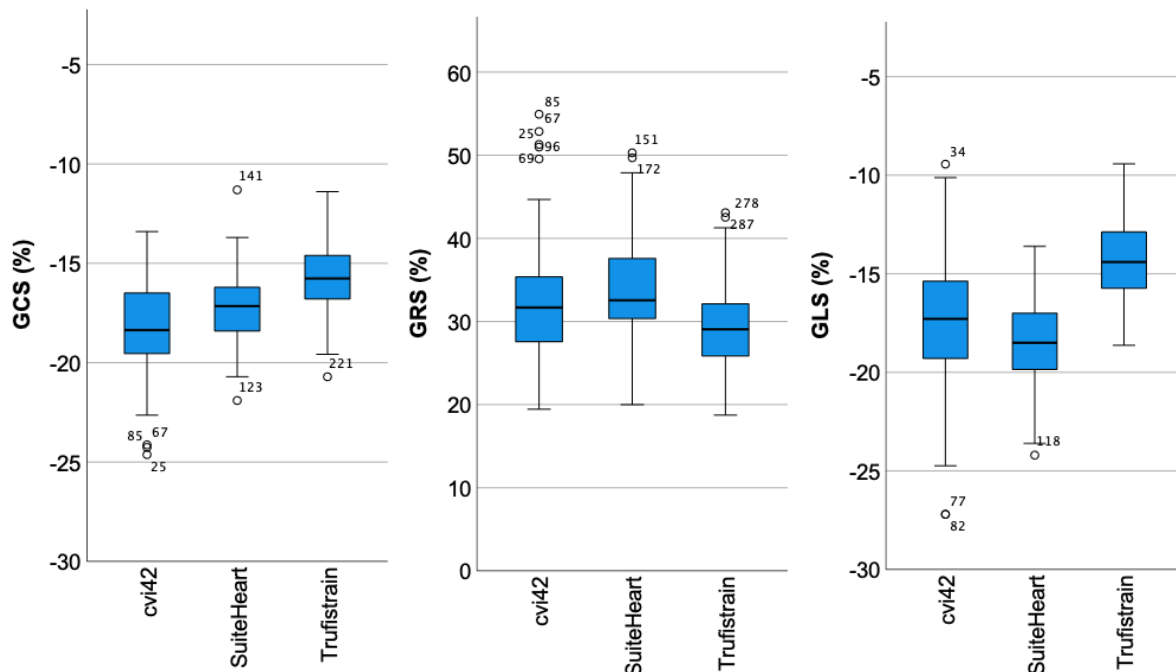


Abbildung 12: Boxplot GRS, GCS, GLS (R1) im Softwarevergleich

Im Unterschied zum GRS konnten bei der Auswertung des GCS signifikante Unterschiede zwischen allen drei Softwareprogrammen nachgewiesen werden ( $p = 0,001$ ), wobei Cvi42 den niedrigsten absoluten Mittelwert von  $-18,19 \% \pm 2,26 \%$  und

TrufiStrain den höchsten GCS mit einem Mittelwert von  $-15,68 \% \pm 1,70\%$  verzeichnete.

Für den GLS zeigte die post-hoc Analyse erneut einen signifikanten Unterschied zwischen allen drei Programmen. SuiteHeart berechnete den niedrigsten Mittelwert von  $-18,53\% \pm 2,21\%$  und TrufiStrain den höchsten Mittelwert von  $-14,29\% \pm 3,69\%$ . Die graphische Darstellung ist den beigefügten Boxplots zu entnehmen. Die hier dargestellten Ausreißer befinden sich alle im 1,5-fachen Interquartilsabstand und können daher in der Auswertung verbleiben.

Zusammenfassend ergibt sich aus der durchgeführten Auswertung, dass die Größe der Globalen Strainparameter abhängig vom ausgewählten Programm sind.

### 4.3 Übereinstimmung der Softwareprogramme

Im softwareübergreifenden Vergleich der ermittelten Parameter zeigten sich gemäß der ICC's eine gute Übereinstimmung des EDVI's (ICC = 0,860) und des GCS (ICC= 0,802), während die EF (ICC = 0,665) und der GLS (ICC = 0,631) nur eine moderate Übereinstimmung belegten. Die Ergebnisse kann der nachfolgenden Tabelle entnommen werden.

Übereinstimmung der Softwareprogramme					
Parameter	Cvi42	SuiteHeart	TrufiStrain	Gepoolter Mittelwert	ICC
EF (%)	$60,2 \pm 6,0$	$55,9 \pm 5,0$	$55,1 \pm 6,1$	$57,1 \pm 4,4$	0,665
Diff	3,2	-1,2	-2,0		
LoA	6,0	5,6	3,6		
EDVi (ml/m <sup>2</sup> )	$77,9 \pm 12,7$	$89,8 \pm 11,5$	$84,2 \pm 13,3$	$83,9 \pm 11,1$	0,860
Diff	-6,0	-5,9	0,3		
LoA	11,8	9,3	10,6		
LVGRS (%)	$32,4 \pm 7,3$	$34,1 \pm 5,9$	$29,4 \pm 5,2$	$32,0 \pm 5,2$	0,786
Diff	0,5	2,0	-2,6		
LoA	7,7	5,7	6,5		
LVGCS (%)	$-18,2 \pm 2,3$	$-17,1 \pm 1,7$	$-15,7 \pm 1,7$	$-17,0 \pm 1,6$	0,802
Diff	-1,2	-0,1	1,3		
LoA	2,2	1,9	1,9		
LV GLS (%)	$-17,4 \pm 3,3$	$-18,5 \pm 2,2$	$-14,3 \pm 3,7$	$-16,8 \pm 2,0$	0,631
Diff	-0,8	-1,7	2,5		
LoA	3,8	3,2	2,3		
Daten: Mittelwert $\pm$ Standardabweichung Untersucher 1; ICC = Intraclass Correlation Coefficient; Gepoolter Mittelwert: Mittelwert über alle Softwareprogramme; Diff: Differenz des Programms zum gepoolten Mittelwert					

Tabelle 5: Übereinstimmung der Softwareanbieter

Zum Vergleich der globalen Strainparameter durch die drei untersuchten Softwarepakete wurden Bland-Altman-Diagramme erstellt. In allen drei Softwareprogrammen zeichnete sich der LV GRS durch die größte Schwankungsbreite aus. Ein signifikanter Bias konnte in den Bland-Altman-Diagrammen für den LV GRS zwischen den

unterschiedlichen Anbietern nicht nachgewiesen werden (siehe Abbildung 15). Während Cvi42 nah am Mittelwert der drei Programme lag, zeigte SuiteHeart eine Überschätzung und TrufiStrain eine Unterschätzung des GRS gegenüber dem Mittel. Die Schwankungsbereiche um den Mittelwert waren in allen drei Programmen vergleichbar.

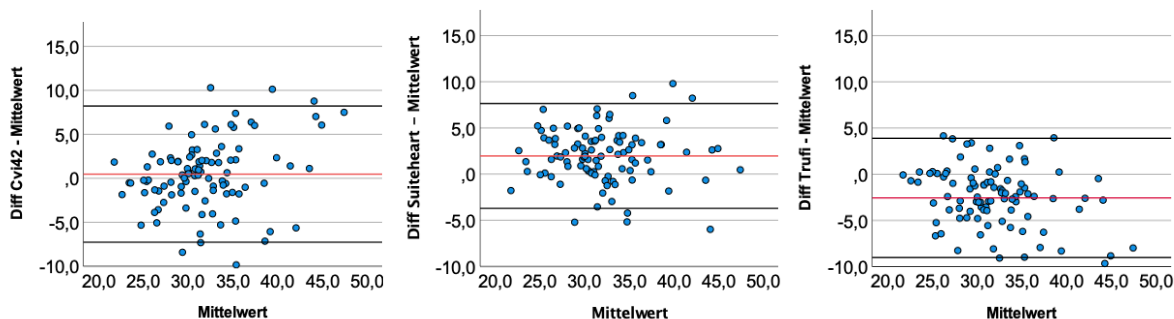


Abbildung 13: Bland-Altman-Plot GRS im Softwarevergleich (R1; Mittelwert)

Bei der Auswertung des GCS zeigte SuiteHeart die größte Übereinstimmung mit dem Mittelwert. Es konnte kein systematischer Bias zur Referenz - dem gepoolten Mittelwert der drei Programme - festgestellt werden. Cvi42 kalkuliert den GCS grundsätzlich niedriger als SuiteHeart, während TrufiStrain diesen signifikant höher berechnet. Der Bland-Altman-Plot von Cvi42 zeigte einen Trend und nahm mit größeren GCS-Werten zu. Die Mittelwerte und Standardabweichungen der Differenzen beim GCS können der obenstehenden Tabelle 5 entnommen werden.

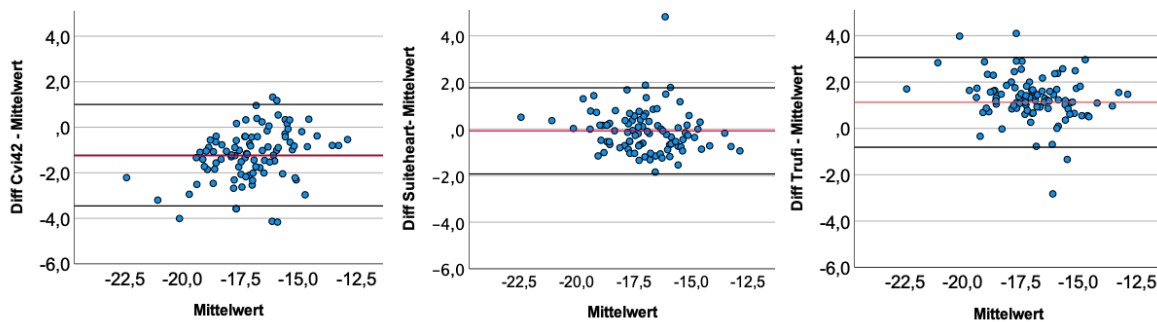


Abbildung 14: Bland-Altman-Plot: GCS im Softwarevergleich (R1; gepoolter Mittelwert)

Im direkten Softwarevergleich konnte mittels der Bland-Altman-Plots, für den durch TrufiStrain berechneten LV GLS ein signifikanter Bias gegenüber dem Mittelwert aller drei Programme nachgewiesen werden. Durch TrufiStrain erfolgt im Vergleich eine signifikante Überschätzung in Referenz zu den untersuchten Programmen. Cvi42 und SuiteHeart zeigten keinen signifikanten Bias. Neben TrufiStrain wurde durch SuiteHeart der LV GLS gegenüber der hier gewählten Referenz unterschätzt. Im Vergleich aller drei Programme stellt sich bei der Betrachtung von Cvi42 das breiteste Konfidenzintervall dar. Ein Trend konnte nicht festgestellt werden.

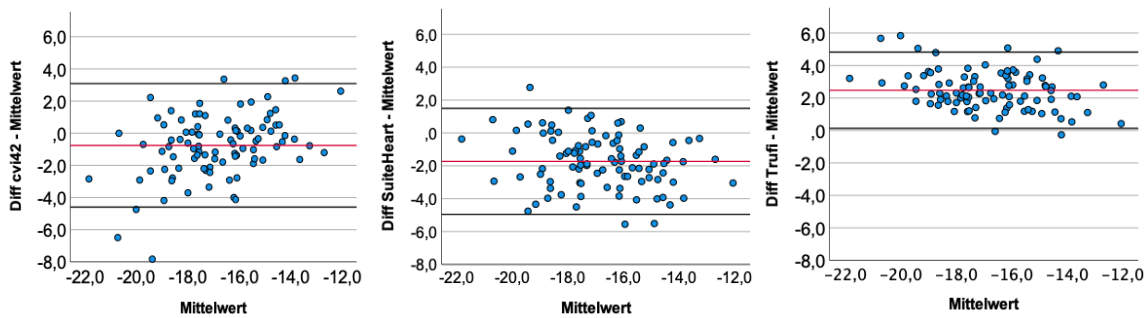


Abbildung 15: Bland-Altman-Plot: GLS im Softwarevergleich (R1; gepoolter Mittelwert)

Im abschließenden Vergleich aller drei Strainparameter bezogen auf die Softwarelösungen stellt sich der GCS als der stabilste Parameter mit dem geringsten Schwankungsbereich dar.

#### 4.4 Referenzwerte Globale Strainparameter

Referenzwerte Globale Strainparameter						
	Cvi42	p-Wert	SuiteHeart	p-Wert	DRA-TrufiStrain	p-Wert
GRS (%)	31,7 (19,2; 44,1)		33,9 (22,7; 45,2)		29,3 (19,1; 39,6)	
W	33,3 (20,9; 45,6)	0,01	35,6 (23,8; 47,4)	0,01	30,8 (21,5; 40,1)	0,03
M	30,1 (18,3; 41,9)		32,3 (22,5; 42,1)		28,0 (17,6; 38,4)	
GCS (%)	-18,0 (-22,1; -13,9)		-17,1 (-20,5; -13,7)		-15,6 (-18,9; -12,3)	
W	-18,7 (-22,5; -14,8)	<0,01	-17,5 (-20,7; -14,3)	0,018	-16,3 (-19,2; -13,4)	0,001
M	-17,4 (-21,3; -13,4)		-16,7 (-20,0; -13,3)		-14,9 (-18,0; -11,8)	
GLS (%)	-17,2 (-23,0; -11,5)		-18,5 (-22,7; -14,4)		-14,1 (-10,3; -17,8)	
W	-18,1 (-23,1; -13,1)	<0,01	-19,3 (-23,5; -15,2)	<0,001	-15,0 (-12,0; -18,0)	0,001
M	-16,4 (-22,5; -10,4)		-17,8 (-21,4; -14,2)		-13,2 (-16,7; -9,7)	

Daten zeigen Mittelwert ( $\pm$  2 fache Standardabweichung) von Untersucher 1 und Untersucher 2  
w= weiblich, m=männlich  
p-Wert nach Mann-Whitney-U-Test bei unabhängigen Stichproben zwischen männlich und weiblich  
Signifikanzniveau <0,05  
\*alle Globalen Strainparameter unterscheiden sich nach Geschlecht signifikant voneinander

Tabelle 6: Referenzwerte Globale Strainparameter nach Software und nach Geschlecht

Anhand der Messungen von Untersucher 1 und Untersucher 2 konnten nach Software getrennte Referenzwerte für Globale Strainparameter erstellt werden. Nach Geschlecht getrennt, zeigen sich signifikante Unterschiede über alle Strainparameter und alle Softwareprogramme. Bei Frauen wurden durch die jeweiligen Softwareprogramme für den GRS höhere (mehr positive) Werte gemessen. Für den GCS und GLS zeigen sich niedrigere Werte als bei Männern. Dies gilt für alle Softwareprogramme gleichermaßen.

## 4.5 Untersucherabhängige Reproduzierbarkeit

Der Vergleich der untersucherabhängigen Reproduzierbarkeit (Interrater Reliabilität) legt dar, dass Auswertungen mit Cvi42 niedrigere Übereinstimmungen erzielten. Es wurden moderate (GLS: ICC= 0,753) bis gute Reliabilitäten (GRS: ICC = 0,809) nachgewiesen. Hingegen ergaben TrufiStrain und SuiteHeart eine gute bis exzellente Reliabilität. SuiteHeart zeigte beim GCS (ICC=0,897) und GLS (ICC= 0,878) eine gute und beim GRS sogar eine exzellente Reliabilität (ICC=0,981) . TrufiStrain hingegen legte in allen drei Strainparametern eine gute bis exzellente Reliabilität dar. Der ICC und das jeweilige 95%ige Konfidenzintervall kann der folgenden Tabelle entnommen werden.

Untersucherabhängige Reproduzierbarkeit für GCS, GLS, GRS								
Strain	Software	R1	R2	ICC	95% KI	CV (%)	Bland-Altman	
							Diff	LoA
GRS (%)	Cvi42	32,4 ± 7,3	31,2 ± 6,0	0,809	[0,713-0,873]	16,4	1,2	10,4
	SuiteHeart	34,0 ± 5,8	34,0 ± 5,7	0,981	[0,970- 0,987]	4,5	-0,4	2,9
	TrufiStrain	29,4 ± 5,2	29,0 ± 5,5	0,951	[0,927-0,968]	7,9	0,1	2,0
GCS (%)	Cvi42	-18,2 ± 2,2	-17,8 ± 2,2	0,798	[0,697-0,865]	10,1	-0,4	3,5
	SuiteHeart	-17,1 ± 1,7	-17,1 ± 1,7	0,897	[0,845 -0,932]	6,1	0,1	2,0
	TrufiStrain	-15,7 ± 1,7	-15,5 ± 1,7	0,928	[0,888-0,953]	5,5	-0,2	1,8
GLS (%)	Cvi42	-17,4 ± 3,3	-17,5 ± 2,7	0,753	[0,631-0,835]	15,7	0,1	5,3
	SuiteHeart	-18,6 ± 2,2	-18,6 ± 2,2	0,878	[0,816-0,920]	7,8	0,0	2,9
	TrufiStrain	-14,3 ± 1,9	-13,8 ± 1,9	0,940	[0,852-0,970]	5,8	0,5	1,6
R1 und R2 zeigen Mittelwerte ± Standardabweichung CV auf Basis gepoolter Mittelwerte R1= Untersucher 1; R2 = Untersucher 2; ICC= Intraclasscorrelation Coefficient, CV= Variationskoeffizient, LoA = Limits of Agreement								

Tabelle 7: Untersucherabhängige Reproduzierbarkeit für GRS, GCS, GLS

## 4.5.1 GRS

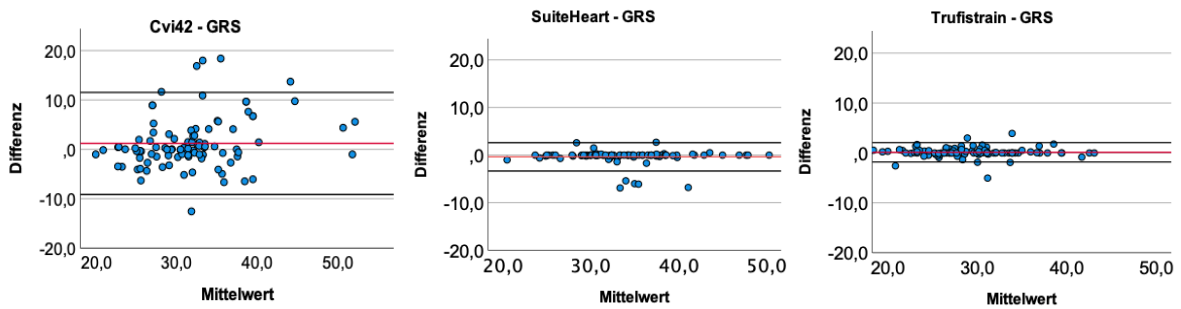


Abbildung 16: Bland-Altman Plot GRS Interrater

In den Bland-Altman-Plots konnte kein signifikanter Bias für die Berechnung des GRS durch die beiden Untersucher im Programmvergleich nachgewiesen werden. SuiteHeart und TrufiStrain weisen eine hohe Übereinstimmung auf (Daten siehe Tabelle 7). 95% der Werte der beiden Untersucher weichen nicht weiter als 2% voneinander ab. Die Übereinstimmung zwischen R1 und R2 ist in beiden Programmen exzellent ( $ICC > 0,950$ ). Wie den Bland-Altman-Plots zu entnehmen ist, zeigt Cvi42 im Vergleich hingegen das breiteste Konfidenzintervall. Gemäß ICC handelt es sich um eine gute Übereinstimmung ( $ICC = 0,809$ ). Diese Beobachtung unterstreicht auch das CV, das bei SuiteHeart und TrufiStrain (4,5%; 7,9%) deutlich niedriger war als bei Cvi42 mit 16,4%.

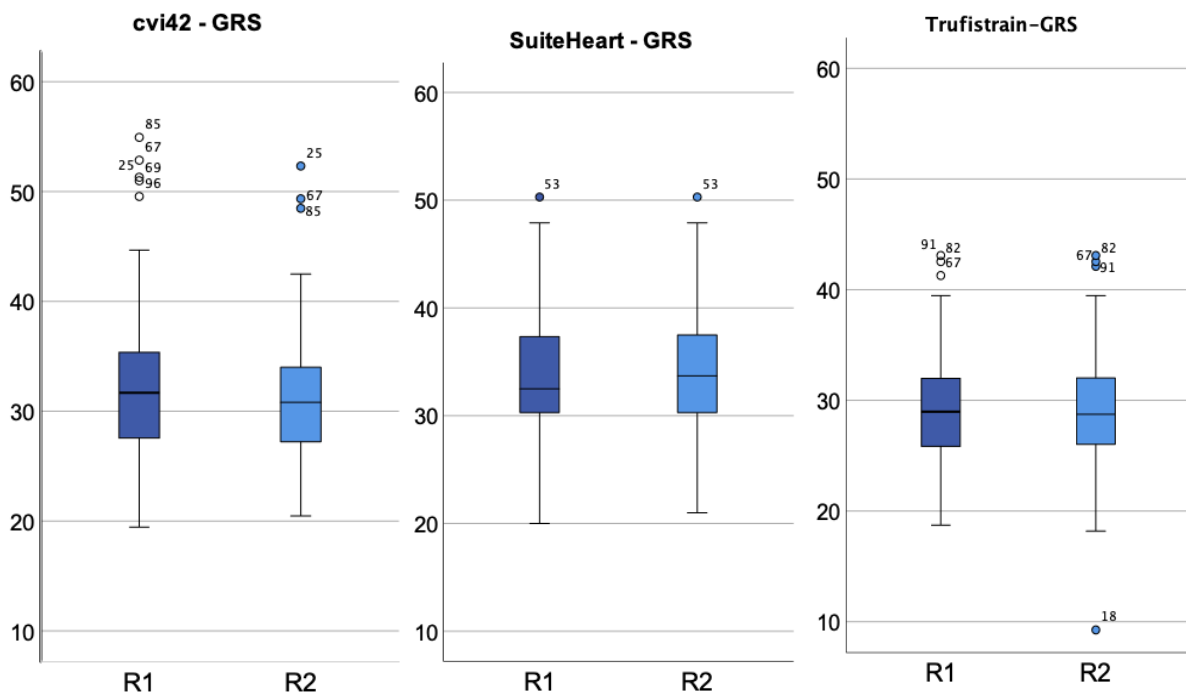


Abbildung 17: Boxplot GRS Interrater

Den Boxplots sind zu entnehmen, dass einige Ausreißer vorliegen. Alle befinden sich innerhalb des 1,5-fachen Konfidenzintervalls und werden als leichte Ausreißer

deklariert. Patient Nr.18 bei TrufiStrain zeigt einen Ausreißer, der einem Messfehler entsprechen muss, da kein Korrelat bei R1 vorliegt.

#### 4.5.2 GCS

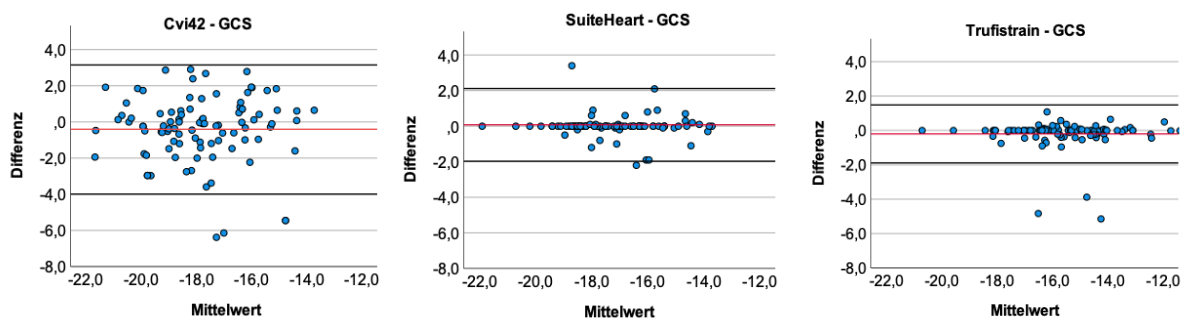


Abbildung 18: Bland-Altman Plot GCS Interrater

Keines der Programme wies beim GCS einen signifikanten Bias auf. Die Mittelwerte der Differenzen lagen nahe 0 und insbesondere SuiteHeart und TrufiStrain ergaben eine hohe Übereinstimmung. Ähnlich der Beobachtung des GRS gab es bei Cvi42 das breiteste Konfidenzintervall, sodass Cvi42 mit einer guten Übereinstimmung ( $ICC = 0,798$ ) gegenüber SuiteHeart ( $ICC = 0,987$ ) und TrufiStrain ( $ICC = 0,928$ ) im programmübergreifenden Vergleich die geringste Übereinstimmung bei der Ermittlung des GCS präsentierte. Die Abweichungen der Differenzen sind in allen Programmen gleichverteilt, sodass kein Trend erkannt werden konnte.

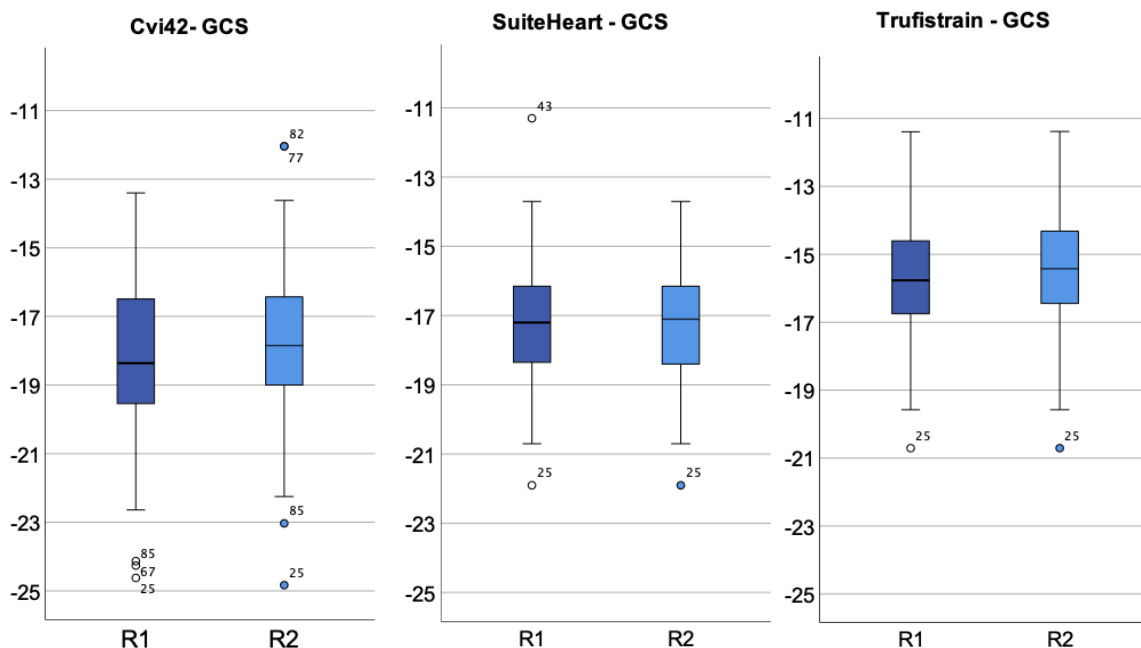


Abbildung 19: Boxplot GRS Interrater Reliabilität

### 4.5.3 GLS



Abbildung 20: Bland-Altman Plot GLS Interrater Reliabilität

Beim GLS zeigten alle drei Programme keinen signifikanten Bias, wie den Daten aus Tabelle 7 zu entnehmen ist. Alle Differenzen vom Mittelwert lagen nahe 0. Erneut wies Cvi42 analog zu den Ergebnissen des GRS und GCS den breitesten Übereinstimmungsbereich (Limits of Agreement) auf, während TrufiStrain den geringsten zeigte. Dies erklärt auch die Unterschiede, die in den ICC Werten zu erkennen sind. Ein Trend in der Datenverteilung konnte bei allen drei Programmen nicht festgestellt werden. Auch der Variationskoeffizient von Cvi42 wies den höchsten Wert mit 15,7% auf, während TrufiStrain hingegen den niedrigsten mit 5,8% ergab.

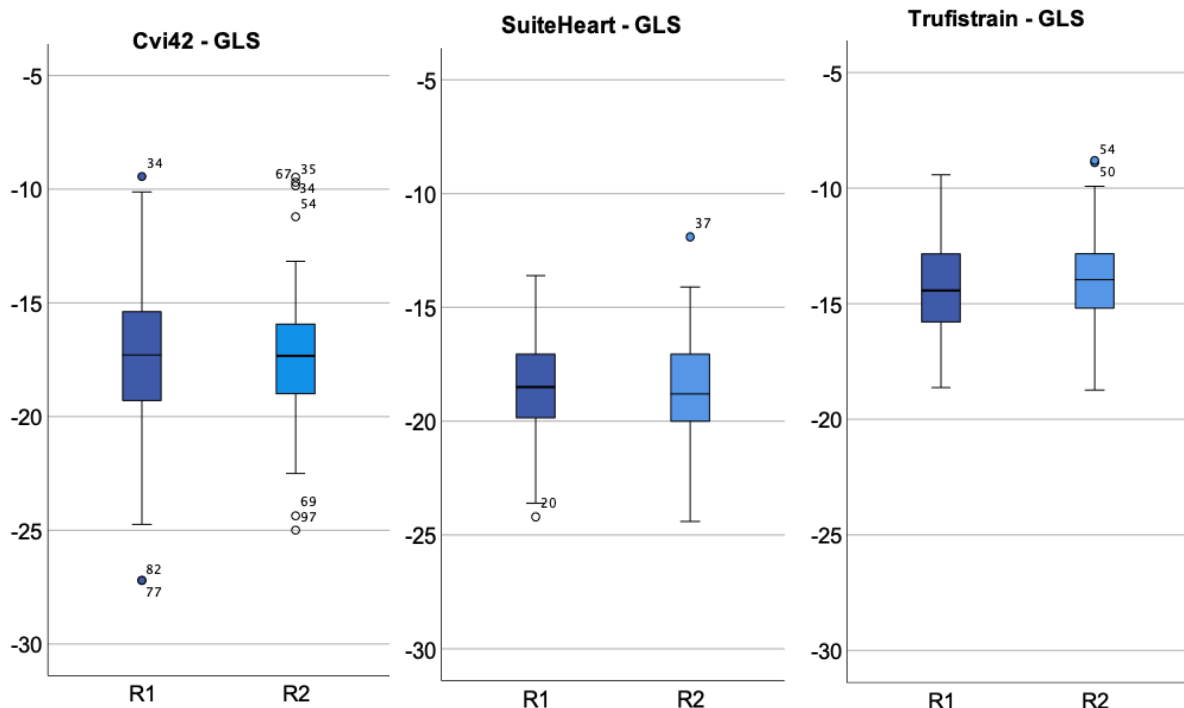


Abbildung 21: Boxplot GLS Interrater Reliabilität

An den Boxplots ist zu erkennen, dass Cvi42 die meisten Ausreißer zeigte. So wies auch der Variationskoeffizient bei Cvi42 mit 10,1% die größte Schwankungsbreite im Vergleich zu SuiteHeart (6,1%) und TrufiStrain (5,5 %) auf.

## 4.6 Intrarater Reproduzierbarkeit für GCS, GLS, GRS

Intrarater Reproduzierbarkeit für GRS, GCS und GLS								
Strain	Software	R1	R1_M2	ICC	95% KI	CV (%)	Bland-Altman	
							Diff	LoA
GRS (%)	Cvi42	29,1 ± 5,2	29,5 ± 3,4	0,905	[0,673-0,973]	9,2	0,4	5,3
	SuiteHeart	34,1 ± 4,2	34,2 ± 4,1	0,999	[0,996-0,999]	0,6	-0,1	0,5
	TrufiStrain	30,3 ± 2,7	29,9 ± 2,8	0,967	[0,888-0,991]	3,1	0,4	1,9
GCS (%)	Cvi42	-17,6 ± 2,1	-17,9 ± 1,3	0,894	[0,643-0,969]	6,3	0,3	2,2
	SuiteHeart	-16,7 ± 1,4	-16,8 ± 1,5	0,990	[0,968-0,997]	1,4	0,1	0,6
	TrufiStrain	-16,2 ± 1,8	-15,8 ± 1,6	0,795	[0,582-0,929]	8,6	-0,5	2,7
GLS (%)	Cvi42	-15,3 ± 3,3	-15,9 ± 2,4	0,886	[0,624-0,967]	11,5	0,7	3,5
	SuiteHeart	-19,2 ± 2,2	-19,5 ± 2,0	0,976	[0,907-0,993]	3,0	0,3	1,2
	TrufiStrain	-14,8 ± 1,9	-14,3 ± 1,9	0,966	[0,911-0,986]	3,9	-0,5	0,9

R1 = Untersucher 1; R1M2= Untersucher 1 Messung 2; ICC= Intraclass Correlation Coefficient; CV= Variationskoeffizient, Diff= Differenz zum Mittelwert; KI= Konfidenzintervall; LoA= Spanne des Konfidenzintervalls  
 CV auf Basis gepoolter Mittelwerte  
 Daten zeigen Mittelwert ± Standardabweichung

Tabelle 8: Intrarater Reproduzierbarkeit für GCS, GLS, GRS

Für alle drei Programme konnte beim GRS im Mittel eine exzellente Intrarater Reliabilität gezeigt werden. Im Detail wies insbesondere Cvi42 eine sehr breite Spanne des ICC- Konfidenzintervalls auf, was auch dem Variationskoeffizienten entnommen werden kann. Dieser lag bei Cvi42 bei 9,2%. Im Vergleich dazu gab es einen sehr kleinen CV bei SuiteHeart mit 0,6%. Der GCS zeigte in Summe ebenfalls eine „gute bis exzellente“ Reliabilität. Auch hier ergab sich erneut eine große Spannweite bei den einzelnen Programmen, wobei SuiteHeart die höchste Übereinstimmung mit der ersten Messung und den geringsten Variationskoeffizienten (0,6%) hatte. Anders als bei den Ergebnissen von GRS und GLS zeigte TrufiStrain beim GCS das breiteste ICC-Konfidenzintervall.

Die Intrarater Reproduzierbarkeit stellte sich beim GLS über alle Programme erneut als „gut bis exzellent“ dar. SuiteHeart zeigte die höchste Übereinstimmung und das geringste Konfidenzintervall mit der ersten Messung, während Cvi42 im Vergleich nur mit „gut“ (ICC 0,886) abschneidet.

Im Vergleich zur Interrater Reliabilität wies erwartungsgemäß die Intrarater Reliabilität bei allen Programmen eine deutlich höhere Übereinstimmung auf. SuiteHeart konnte in allen drei Strainwerten die höchste Übereinstimmung verzeichnen.

Analog zu den Ergebnissen der Interrater Reliabilität ergaben die Berechnungen durch Cvi42 die geringste Übereinstimmung und das breiteste Konfidenzintervall.

## 5. Diskussion

Die hier dokumentierten Ergebnisse zeigen, dass die Auswahl des Softwareprogrammes einen Einfluss auf die CINE basierten quantitativen Parameter der CMR hat. Sowohl die Ergebnisse der LV Volumetrie als auch die gemessenen LV globalen Strainparameter variieren in Abhängigkeit von dem ausgewählten Programm. Die Inter- und Intra-Rater-Reliabilität zeigte über alle Programme hinweg eine gute bis exzellente Übereinstimmung. Erwartungsgemäß ergab die Intra-Rater-Reliabilität eine im Vergleich noch höhere Übereinstimmung.

Der kardialen Strainanalyse könnte in Zukunft zunehmend eine klinische Bedeutung zur Diagnose von kardiovaskulären Erkrankungen im Frühstadium und damit auch zur Initiierung von therapeutischen Maßnahmen beitragen. Das größte Potential wird in der höheren Sensitivität gegenüber etablierten globalen Parametern wie der EF gesehen. Erste klinische Einsatzbereiche sind beispielsweise die frühzeitige Diagnose Chemotherapie induzierter Kardiotoxizität sowie deren Monitoring (62).

Aber auch ein Einsatz sowohl bei ischämischen oder entzündlichen Herzerkrankungen als auch zum Monitoring von Kardiomyopathien kann erwogen werden. Begrenzt wird das Potential derzeit durch eine fehlende Standardisierung, fehlende Referenzwerte und eine begrenzte Anzahl von Studien, die Aussagen zur Prognose der Erkrankung zulassen (63).

Da durch die demographische Entwicklung kardiovaskuläre Erkrankungen zunehmen werden, wird die klinische Bedeutung der CMR in der Zukunft weiter steigen. Ferner ist es von großer Bedeutung, einen Parameter zur Verfügung zu haben, der der Früherkennung von kardiovaskulären Erkrankungen dient. Dies erklärt, weshalb stetig neue CMR FT-Postprocessing Softwarelösungen auf dem Markt verfügbar sind. Es ist notwendig, diese Lösungen hinsichtlich der Reproduzierbarkeit und Vergleichbarkeit mit bereits etablierten Anbietern zu evaluieren. Zu den Programmen SuiteHeart und TrufiStrain sowie Cvi42 waren zum Zeitpunkt der Studie nur wenige Studien veröffentlicht, die eine große Kohorte gesunder Probanden untersucht. Daher wurden in dieser Studie die CINE SSFP Sequenzen von 98 gesunden Probanden mit den drei zuvor genannten Programmen retrospektiv untersucht. Ein Fokus wurde hierbei auch auf die Ermittlung von Referenzwerten gelegt.

Die Ergebnisse dieser Studie zeigten sich konkordant zu vorherigen Studien von zum Beispiel Messali, Zange und Barreiro-Pérez et al. und bestätigten, dass ein softwareübergreifender Vergleich der Globalen Strainparameter derzeit nicht möglich ist (51, 56, 57).

Die Ergebnisse dieser Arbeit zeigen, dass sich die Softwareprogramme in der Interrater Reliabilität zum Teil deutlich unterscheiden. Dies hat einen großen Einfluss auf die Reproduzierbarkeit der Messergebnisse. CMR Strain Techniken wie Optical Flow und Deformable Registration Algorithmen unterscheiden sich in der Güte der Reproduzierbarkeit und können Abweichungen in der Strainkalkulation bedingen. Die Güte der automatischen Konturerkennung birgt großes Verbesserungspotential.

### **5.1 Unterschiede bei der linksventrikulären Volumetrie**

Die Berechnung der linksventrikulären Volumetrie und Funktion belegte signifikante Abweichungen zwischen den drei Programmen. Die EF wurde durch Cvi42 im Mittel um ca. 5% höher bestimmt als bei der Berechnung durch SuiteHeart und TrufiStrain. In Referenz zu den von Kawel-Boehm et. al. definierten Referenzwerten zeigte sich, dass die durch Cvi42 ermittelte EF innerhalb der einfachen Standardabweichung lag. SuiteHeart und TrufiStrain unterschätzen jedoch die von Kawel-Boehm et.al. und lagen unterhalb des hier definierten Referenzbereichs. Die Studien anhand derer Kawel-Boehm et. al. die Referenzwerte der Volumetrie ermittelt haben, wurden mit Cvi42 und View Forum von Philips Medical und Argus Function von Siemens Medical bestimmt. Nicht alle darin enthaltenen Studien haben ihre Ergebnisse mit semiautomatischen Konturzeichnungen erzielt und SuiteHeart sowie TrufiStrain waren nicht Bestandteil (29).

Analog zur EF entsprachen das von Cvi42 berechnete EDVI sowie ESVI am ehesten dem von Kawel-Boehm et. al definierten Referenzbereich. TrufiStrain und SuiteHeart wiesen um etwa 5 bis 10 ml/m<sup>2</sup> höhere Werte auf als Cvi42 (29). Zum Teil zeigten sich auch EDVi-Werte, die nicht einem gesunden Kollektiv entsprechen sollten. Bei fehlendem Korrelat in den anderen Softwareprogrammen, ist in diesen Fällen ein untersucherabhängiger Effekt in Betracht zu ziehen. Die Definition des Endokards, der Basalebene, der Zeitpunkt der Enddiastole und Endsystole sind hierbei relevante Variablen.

Der LVMI zeigt ebenfalls einen signifikanten Unterschied zwischen Cvi42 und TrufiStrain, auch lagen die durch Cvi42 ermittelten Werte innerhalb der einfachen

Standardabweichung der Vergleichskohorte. TrufiStrain hingegen wich mit  $60 \text{ g/m}^2$  im Durchschnitt deutlich von der oben genannten Vergleichskohorte ab.

Soweit dem Autor bekannt ist, liegt derzeit keine Studie vor, in der die Volumenberechnung dieser drei Programme miteinander verglichen wurde.

Entgegen den Ergebnissen von zuvor veröffentlichten Studien konnte die vorliegende Arbeit zeigen, dass die volumetrischen Basisdaten von dem jeweiligen Programm abhängen und sich zum Teil signifikant voneinander unterscheiden (56, 57).

Mögliche Ursachen können der Algorithmus, die Güte der semiautomatischen Konturerkennung, als auch die Anwendung des Programms sein. Der manuelle Nachbearbeitungsaufwand durch Cvi42 war im Vergleich höher und erklärt auch die im Vergleich geringere Inter- und Intra-Rater-Reliabilität. Diese Beobachtung bestätigte eine bereits veröffentlichte Studie (51). In einer weiteren Studie konnte nachgewiesen werden, dass Korrekturen bei SuiteHeart seltener notwendig waren. Bei der Berechnung biventrikulärer Volumina zeigten die Werte einer voll automatischen Konturzeichnung gegenüber denen eines erfahrenen Untersuchers kaum Unterschiede und waren hochgradig reproduzierbar (53). Dies legt einen softwareabhängigen Effekt bei SuiteHeart nahe. Gegebenenfalls wurden im Vergleich endokardiale Konturen bei Cvi42 genauer angepasst oder basisnahe Kurzachsen mit einem Myokardanteil von weniger als 50% häufiger ausgeschlossen. Auch die Wahl von ED und ES sind als Einflussfaktoren auf die Parameter nicht auszuschließen.

Im Gegensatz zu Cvi42 und SuiteHeart übernimmt TrufiStrain die Definition der basisnahen Schicht anhand eines Algorithmus, der diese auf Basis der Langachse auswählt. Gegebenenfalls wurden durch TrufiStrain basisnahe Schichten mit einem Myokardanteil von unter 50% häufiger hinzugezählt, weshalb sich insbesondere der EDVi und der LVMI vergrößerte. Da die ES bei TrufiStrain durch den Untersucher nicht bearbeitet werden kann, führt dies zu einem höheren ESVi. Daraus folgend verringert sich auch die EF.

Weiterhin zeigte sich, dass sowohl der EDVi als auch der ESVi über alle Programme im Vergleich zum Referenzkollektiv von Kawel-Boehm et. al. höher gemessen wurden. In Betracht zu ziehen ist, dass in dieser Studie vorherrschend ein junges Kollektiv vertreten war, weshalb ein im Schnitt höheres EDV erklärbar wäre, da dieses mit dem Alter abnimmt.

Da die durch Cvi42 gemessenen Ergebnisse eher den Referenzwerten entsprechen, aber eine geringere Übereinstimmung zu den mit TrufiStrain und SuiteHeart gemessenen Werten zeigt, sind künftig weiterführende Untersuchungen erforderlich,

die die linksventrikuläre Volumetrie und Funktion auf anbieterabhängige Vergleichbarkeit untersuchen. Auch sollte untersucht werden, ob sich die Volumenberechnung bei TrufiStrain unterscheidet, sofern eine Langachse als Referenzpunkt fehlt.

## **5.2 Unterschiede linksventrikulärer globaler Strainparameter**

In zahlreichen Studien konnte gezeigt werden, dass aktuell kein eindeutiger Konsens zu Referenzwerten einer gesunden Kohorte besteht. Die Software, demographische Variablen (Geschlecht), der Untersucher und auch hämodynamische Faktoren, wie zum Beispiel Herzfrequenz und Blutdruck haben Einfluss auf die Ergebnisse (17, 21, 44).

Die in dieser Studie ermittelten Referenzwerte der gesunden Freiwilligen sind in allen drei globalen Strainrichtungen im Mittel um ca. 3 bis 4 % geringer ausgefallen als die publizierten Referenzwerte (29). Trotzdem liegen alle drei Parameter in den hier aufgeführten Konfidenzintervallen der Metaanalyse von Kawel-Boehm et.al. Auch in Relation zu anderen Studien wie die von Vo et. al. als auch von Liu et. al. liegen die gemessenen Werte innerhalb der einfachen Standardabweichung. Somit entspricht das hier untersuchte Kollektiv dem als gesund definierten Bereich (29, 44, 64).

Unabhängig davon lag TrufiStrain an der unteren Grenze des als gesund definierten Bereiches. Gerade in Anbetracht, der absoluten Abweichungen von größer als 1,5% beim LV GCS und größer als 3% beim LV GLS durch TrufiStrain handelt es sich um klinisch relevante, signifikante Abweichungen.

Dass die in dieser Studie gemessenen Referenzwerte geringer ausfallen als die von Kawel-Boehm et. al. definierten Referenzwerte, kann auf deren verwendeten Programme zurückgeführt werden. Die Ergebnisse der sieben darin enthaltenen Studien wurden mit vier unterschiedlichen Methoden/Softwareprogrammen berechnet (HARP- Method, Cvi42, TomTec und QStrain). Allein Cvi42 wurde auch in dieser Studie verwendet, TrufiStrain oder SuiteHeart waren nicht Teil der Studie. Dass die Auswahl der Software die Strainwerte beeinflusst, stellten bereits Barreiro-Pérez et. al. fest. Sie wiesen nach, dass sich alle vier genannten Programme bei der Berechnung des GLS und des GRS signifikant unterschieden. Im direkten Vergleich erzielte insbesondere TomTec höhere Werte als Cvi42. Allein der GCS unterschied sich nicht (51). Weiterhin variierten die Ergebnisse der darin enthaltenen Studien ebenfalls und zeigten zum Teil Abweichungen.

Jede Software unterliegt regelmäßigen Updates. Insbesondere Tracking-Algorithmen werden ständig verbessert, wodurch Ergebnisse beeinflusst werden können.

Da zur Einordnung der berechneten Parameter in unserer Auswertung eine Referenz notwendig ist, diese jedoch nicht vorliegt, wurde als Referenz der Mittelwert der drei Programme herangezogen. Dieser liegt im unteren Bereich der von Kawel-Boehm et al. definierten Referenzgruppe. Da die Mittelwerte der drei Programme sich zum Teil signifikant voneinander unterscheiden und insbesondere TruFiStrain für GCS und GLS im Verhältnis sehr hohe Strainwerte generierte, stellt dies eine Schwäche des Mittelwertes als Referenz dar.

Weiterhin legen die Hersteller bis auf die Grundprinzipien die Berechnung der Strainparameter nicht offen, somit können Unterschiede in der Berechnungsgrundlage nicht direkt erörtert werden. Dadurch, dass in der Literatur oftmals eine hohe Inter- und Intra-Rater-Reliabilität je Softwareprogramm festgestellt werden kann, legt nahe, dass der Algorithmus einen großen Einfluss auf die Generierung der Strainparameter hat.

Strainparameter können auch dahingehend variieren, wie exakt endokardiale Konturen eingezeichnet und gemessen werden. Die Organisation der Kardiomyozyten innerhalb des Myokards haben unterschiedlich großen Anteil an der Gesamtkontraktion. So erzeugen insbesondere die endokardialen gegenüber den mitt- und epikardialen Anteilen die größten absoluten Streckenveränderungen. Mit der Exaktheit der endokardialen Kontur variieren auch die generierten Strainparameter zum Teil sehr stark. So definierten Amzulescu et al. die Ermittlung der endokardialen Kontur durch eine Software als den wahrscheinlich wichtigsten Einflussfaktor auf die Vergleichbarkeit von Strainparametern. Diese softwarespezifische Information sind jedoch nicht offengelegt und können demnach auch nicht überprüft werden (21).

Weiterhin sind demographische Einflussfaktoren zu betrachten. Dass die Abweichungen zu den von Kawel-Boehm definierten Referenzwerten auf Alterseffekte zurückzuführen sind, kann zu diesem Zeitpunkt nicht postuliert werden, da die aktuelle Studienlage bzgl. altersabhängiger Effekte auf globale Strainparameter derzeit uneindeutig ist (29).

In der Auswertung nach Geschlecht ist die Studienlage zum Teil uneins, während Vo et. al. keinen geschlechtsabhängigen Einfluss sahen, wurde durch Augustine et al. ein geschlechtsabhängiger Einfluss auf den GRS und GLS festgestellt. Zuletzt wurden einige Studien veröffentlicht, die einen Einfluss des Geschlechts auf alle globalen

Strainparameter nachweisen konnten (29, 38). Hierin wurden bei weiblichen Patienten für den GRS größere, für GCS und GLS kleinere Werte gemessen. Die Ergebnisse dieser Studie gliedern sich in diese Erkenntnisse ein und konnten ebenfalls einen signifikanten geschlechtsabhängigen Einfluss nachweisen (siehe Referenzwerte 4.4). Die in dieser Studie generierten Strainparameter entsprechen einem gesunden Kollektiv. Es zeigt sich, dass ein sehr breites Intervall als gesunde Normwerte definiert werden kann. Weiterhin lässt sich nachweisen, dass insbesondere softwareübergreifende Vergleiche derzeit noch nicht möglich sind. Daraus folgert, dass bei Verlaufskontrollen immer das gleiche Softwareprogramm verwendet werden sollte. Es ist zu prüfen, ob softwareabhängige Korrekturfaktoren softwareübergreifende Vergleiche ermöglichen könnten. Die Definition einer Referenzmethode im Sinne eines Goldstandards könnte helfen Referenzwerte zu klassifizieren. Die Einführung von Standards mit dem Ziel, verlässliche Grenzwerte zu ermitteln, die softwareübergreifend gelten, stellt eine der Hauptherausforderungen eines künftigen zentralen Forschungsvorhabens dar.

### **5.3 Anbieterabhängige Unterschiede im Bezug zur aktuellen Studienlage**

In dieser Studie konnten signifikante Unterschiede für die Berechnung von globalen myokardialen Strainparametern zwischen allen drei Softwareanbietern nachgewiesen werden. Einzig beim GRS lag zwischen den von Cvi42 und SuiteHeart ermittelten Werten kein signifikanter Unterschied vor. Der GRS zeigte das breiteste Konfidenzintervall. In den Bland-Altman-Diagrammen wurden vereinzelte Ausreißer gegenüber dem gemitteltem Strainwert festgestellt. Einzig bei dem durch TrufiStrain berechneten GLS zeigte sich ein signifikanter Bias.

Trotz der Unterschiede ergab sich in der Berechnung des ICC für den GRS (ICC= 0,786) und insbesondere für den GCS (ICC = 0,802) eine gute Übereinstimmung. Passend zu den Bland-Altman-Diagrammen zeigten die Werte des GLS eine moderate Übereinstimmung (ICC = 0,631). Die Unterschiede der Werte des GCS zwischen Cvi42 und SuiteHeart waren zwar signifikant, aber in der Praxis durchaus hinnehmbar. Für GRS, GCS und GLS zeigte TrufiStrain die niedrigsten berechneten Strainparameter, während die Ergebnisse der EF und des EDVi mit den Ergebnissen von SuiteHeart vergleichbar waren.

Dass es zwischen unterschiedlichen Softwareanbietern signifikante Unterschiede bei der Berechnung globaler Strainparameter gibt, stellten bereits Schuster et al. fest, was

von Barreiro-Pérez et al. bestätigt wurde (6, 51). Aus diesen Studien ging hervor, dass sich der GCS als der robusteste Parameter im Hinblick auf softwareabhängige Unterschiede zeigte. Diese Beobachtungen konnten im Rahmen dieser Studie ebenfalls festgestellt werden. Der ICC des GCS zeigte die höchste Übereinstimmung. In den beiden zitierten Studien fiel auf, dass jeweils Cvi42 die größte Variabilität bei den Ergebnissen aufwies und der Nachbearbeitungsaufwand am höchsten war (51, 53). Auch diese Beobachtungen konnten in dieser Studie gemacht werden. Die automatische Kontursegmentierung sowohl in Kurz- als auch Langachse erforderte im Vergleich eine häufigere manuelle Bearbeitung durch den Untersucher. Insbesondere in Basis und Spitze waren Korrekturen erforderlich, da beispielsweise Anteile des Magens dem Herzen zugeordnet wurden.

Anders zu bewerten sind die Ergebnisse des GRS, die im Rahmen dieser Studie ermittelt wurden. Die bisher untersuchten Softwareprogramme zum Beispiel in der Studie von Barreiro-Pérez et. al. zeigten die größten Diskrepanzen bei der Berechnung des GRS. Dies wurde insbesondere dadurch erklärt, dass bereits kleinste Abweichungen in der Konturzeichnung aufgrund der kurzen Strecke des Myokards einen großen Einfluss auf das Gesamtergebnis des GRS haben. Die Autoren vermuteten, dass einige Softwareanbieter mit der Berechnung des GRS besser zurechtkommen als andere. So schnitt der Algorithmus von Medis und Tomtec in diesem Teilgebiet schlechter ab als der von Cvi42 und Medviso (51). In dieser Studie konnte gemäß ICC eine mäßig bis gute Übereinstimmung auch beim GRS bei den hier untersuchten Softwareprogrammen nachgewiesen werden. Dies könnte auf die Güte der semiautomatischen Konturerkennung zurückgeführt werden. So war bspw. der Nachbearbeitungsaufwand bei SuiteHeart und TrufiStrain in den Kurzachsen sehr gering. Dies spiegelt sich auch in der später noch diskutierten Interrater Reliabilität wider. Zudem könnten die drei Programme ähnlicher in der Berechnung des GRS sein, weshalb geringere Unterschiede auftraten. Dass es Unterschiede in der Art der Berechnung bzw. der Definition des GRS unter den Programmen gibt, ist allein auch darin zu erkennen, dass sich die von SuiteHeart berechneten zwei Globalen Radialen Strain- Werte, der Peak- und der Mid- Global Radial Strain, von den von Cvi42 und TrufiStrain unterscheiden. Es ist davon auszugehen, dass eine andere Berechnungsgrundlage für die jeweiligen Strainwerte bei SuiteHeart vorliegt.

Bei der Betrachtung des GLS fiel auf, dass dieser unter den globalen Strain Parametern die geringsten Übereinstimmungen unter den drei Programmen aufwies. Da in der Literatur neben dem GCS der GLS softwareübergreifend als der Parameter

mit der höchsten Übereinstimmung definiert wird, ist dies bemerkenswert. Am ehesten ist der geringe ICC auf den durch TrufiStrain erzeugten deutlich höheren Mittelwert des GLS zurückzuführen. Auffallend ist, dass der durch TrufiStrain ermittelte GLS mit einem Mittelwert von ca. -14,3 % sehr nah an der von Kawel-Boehm et. al. definierten Grenze zum Pathologischen liegt.

Die in dieser Studie vorgestellten Beobachtungen reihen sich ein in die Ergebnisse einer aus dem Jahr 2021 stammenden Studie von Zhang et.al. Sie ist die einzige Studie die den GLS von SuiteHeart, TrufiStrain und Cvi42 miteinander vergleicht und demonstriert, wie beinahe deckungsgleich die Ergebnisse der gesunden Subkohorte mit dieser sind. Ergebnisse zum GCS und GRS liegen nicht vor. Der durch TrufiStrain berechnete GLS bildet sowohl im Zwei- als auch im Vierkammerblick stets die höchsten Werte ab (55).

Die von Liu et. al und Zhang et al. veröffentlichten Ergebnisse decken sich bis auf die Werte des GRS mit denen dieser Studie (48, 55).

Die Gründe der beschriebenen Diskrepanz in allen Strainparametern sind zu klären. Die im Vergleich fehlende automatische Konturerkennung in der Langachse, die eingeschränkte Nutzbarkeit von Bearbeitungswerkzeugen, das kleine Bearbeitungsfenster sowie die fehlende Bearbeitung in der Endsystole sind wichtige den GLS beeinflussende Variablen. Die automatisch eingezeichneten Konturen der Kurzachsen, waren durch TrufiStrain in der Regel sehr gut definiert und erforderten selten Ausbesserungen in der ED. Bei der Notwendigkeit einer Korrektur, stellte sich diese wie in der Langachse auch als unpräzise dar. Wäre der Unterschied der Messwerte hauptsächlich untersucherabhängig, so müsste sich dies auch in der Interrater Reliabilität widerspiegeln, was durch die Ergebnisse dieser Studie nicht nachgewiesen werden konnte. Dieser Frage gingen Zhang et al. auch in ihrer Studie nach und verglichen die Ergebnisse durch manuelle vs. automatisch generierte Konturen in der Kurzachse. Sie konnten keine signifikanten Unterschiede feststellen, weshalb geschlussfolgert wurde, dass die softwareabhängigen Unterschiede somit auf die Algorithmen zurückzuführen seien (55).

Daneben ist der DR- Algorithmus zu nennen, der neben der Anwendung auch Unterschiede in den Prinzipien der Strainberechnung aufweist. Hierbei wird im Gegensatz zum Optical Flow (Cvi42 und SuiteHeart) jeder Pixel des Myokards verfolgt. Diese können in Schichten eingeteilt und somit zum Beispiel endo-, epi- und myokardiale Strains erzeugen. Der Optical Flow verfolgt ausschließlich Pixel entlang der epi- und endokardialen Konturen (65). So konnte bereits nachgewiesen werden,

dass mit dem DRA Ansatz sowohl in segmentalen als auch in Spitzenstrains (Peaksystolic strains) akkuratere und reproduzierbarere Ergebnisse mit diesem Ansatz erzielt werden (66). Es ist davon auszugehen, dass der von TrufiStrain verwendete DR-Algorithmus einen größeren Einfluss auf die Berechnung der Strainparameter hat. Somit wäre in einer weiteren Studie zu untersuchen, ob eine feste Korrelation zwischen den durch TrufiStrain berechneten Parametern zu denen von Cvi42 oder SuiteHeart vorliegt.

Als weitere mögliche Einflussvariable auf die Ergebnisse des Intraclass Correlation Coefficient ist auch das Studiendesign zu betrachten. Die Ergebnisse einer Studie verwiesen darauf, dass selbst eine geringe untersucherabhängige Fehlervariabilität nicht automatisch einen hohen ICC garantiert. Denn eine geringe Untersuchungsobjektvariabilität, in unserem Fall - die Untersuchung einer gesunden Kohorte- produziert ebenfalls geringere ICC's (67). Somit kann vermutet werden, dass die Untersuchung eines gesunden Kollektivs gegenüber einer Stichprobe aus Gesunden und Kranken höhere ICC's generieren könnte.

Wichtig im Gesamtzusammenhang ist auch zu bemerken, dass zum Zeitpunkt der Studie die myokardiale Strainanalyse von TrufiStrain lediglich zu Forschungszwecken und als nicht-kommerzieller Prototyp verfügbar ist.

Damit die genauen Gründe analysiert werden könnten, müssten die Algorithmen der Strainberechnungen offengelegt werden; allerdings hat nur TrufiStrain in seinem Manual die Grundzüge seiner Strainkalkulation beschrieben (68). Dies unterstreicht in Summe die Bedeutung der Vereinheitlichung der Kalkulation der globalen Strainparameter. Ähnlich wie dies bereits im Bereich der STE erfolgt ist, bei der eine Arbeitsgruppe der EACVI und der ASE eine Standardisierungsinitiative in Kooperation mit der Industrie ein besseres Intervendor Agreement erarbeitet hat, ist dies auch für das CMR – FT vonnöten(25). Hilfreich wäre es auch, den Softwareherstellern standardisierte Bilddaten zur Verfügung zu stellen, mit deren Hilfe eine Kalibrierung der Algorithmen anhand validierter Ergebnisse möglich wäre.

## **5.4 Inter- und Intrarater Reproduzierbarkeit im Bezug zur aktuellen Studienlage**

Bei der Untersuchung der Inter- und Intrarater Reproduzierbarkeit zeigten sich softwareübergreifend gute bis exzellente Ergebnisse gemäß ICC. Insbesondere TrufiStrain wies stets eine exzellente Kongruenz der Ergebnisse aus. Dies ordnet sich in die Erkenntnisse der zu diesem Punkt vorherrschenden Studienlage ein. Viele Studien bestätigten, dass grundsätzlich durch CMR-FT für den GCS und den GLS eine sehr gute bis exzellente Übereinstimmung vorlag (38, 48, 50, 69).

Dass der ICC des GRS bei TrufiStrain und SuiteHeart im Vergleich zu den anderen Strainparametern eine sehr hohe Übereinstimmung erzielte, ist hingegen neu. Barreiro-Pérez et. al. bestätigten die allgemeinen Beobachtungen, stellte aber fest, dass explizit die Berechnung des GRS durch die verschiedenen Softwareanbieter unterschiedlich gut absolviert wird (51). Der Grund dafür, könnte darin liegen, dass SuiteHeart und TrufiStrain einen robusteren Tissue-Tracking-Algorithmus für die semiautomatische Konturzeichnung aufweisen. Damit fällt der individuelle manuelle Input durch Konturkorrekturen geringer aus.

In Studien konnte nachgewiesen werden, dass die Interrater Reproduzierbarkeit sich verringert, je größer der manuelle Aufwand bei der Konturzeichnung ist. Weitere Einflussfaktoren auf die vom Untersucher abhängigen Unterschiede bei der Strainberechnung seien Expertise, Training bzw. Schulung und Grad der Automatisierung (21, 53).

Neben den untersucherbezogenen Faktoren hat der jeweilige Algorithmus für die Berechnung der Strainparameter einen großen Einfluss auf die Reproduzierbarkeit.

So wiesen Barreiro-Pérez et. al. nach, dass Medviso mit einem Non-rigid-Algorithmus die höchste Übereinstimmung in allen drei Strainparametern erzeugte(51).

Da bisher wenige Studien zu TrufiStrain und SuiteHeart veröffentlicht wurden, liegen aktuell wenige Vergleichswerte vor. Die vorliegenden Vergleichswerte, die z.B. in der Studie von Lamacie et. al. veröffentlicht wurden, untersuchten TrufiStrain im Vergleich zu konventionellen CMR-FT-Methoden sowie zu STE und bestätigten die höchste Interrater Reproduzierbarkeit. Der CV war deutlich geringer und der ICC war höher im Vergleich zu konventioneller CMR-FT-Software (65). Liu et. al. untersuchten 2017 den DRA-TrufiStrain an einer Gruppe von 130 chinesischen gesunden Probanden und zeigten ebenfalls durchweg die höchsten Übereinstimmungswerte(48).

Dass in dieser Studie der GLS bei TrufiStrain die höchste Übereinstimmung zeigte, ist auch insofern bemerkenswert, als dass für die Langachse keinerlei semiautomatische

Konturerkennung vorhanden ist und die Langachse ausschließlich manuell eingezeichnet worden ist. Dies lässt vermuten, dass der Einfluss des Untersuchers bei TrufiStrain doch geringer zu sein scheint als der des Algorithmus.

Die einzige Studie, die SuiteHeart, TrufiStrain und Cvi42 miteinander vergleicht, wurde von Zhang et. al veröffentlicht und sie untersuchten dabei lediglich die Übereinstimmung des GLS. Cvi42 zeigte auch hier die geringsten Übereinstimmungswerte. Die höchsten Übereinstimmungen erzielte SuiteHeart mit einem ICC von 0,96-1,00 und TrufiStrain schnitt mit einem ICC von 0,88-1,00 etwas schlechter ab.

Die Ergebnisse dieser Studie zur Intrarater Reliabilität ordnen sich hier konkordant ein. Über alle Studien hinweg weist die Intrarater Reliabilität höhere Werte auf als die Interrater Reliabilität. Dadurch, dass sich in der Regel das Vorgehen eines Untersuchers nicht wesentlich unterscheidet und die gleichen Anforderungen an die Konturgüte herangezogen werden, liegen geringere Prozessunterschiede vor, und folglich besteht eine höhere Übereinstimmung zwischen den Ergebnissen eines Untersuchers.

Nichtsdestotrotz sind an den Ergebnissen Unterschiede zwischen den jeweiligen Programmen erkennbar, die auch bereits bei der Interrater Reliabilität zu beobachten waren. Während SuiteHeart die höchsten Übereinstimmungen zeigten, wies Cvi42 unter den drei Programmen die geringste Reproduzierbarkeit auf. Barreiro-Pérez et.al. oder auch Schuster et al. zeigten exzellente Übereinstimmungswerte, aber im Vergleich relativ hohe Variationskoeffizienten (6, 51). Dass Cvi42 selbst in der Intrarater-Reliabilität geringere Übereinstimmungswerte erzielte, zeigt, dass die Bedienung des Programms mehr Variationsmöglichkeiten zulässt. Die automatisch generierte Konturgüte ist im Vergleich geringer und erfordert vom Untersucher wiederum mehr Korrekturen. Die Variabilität der Ergebnisse nimmt somit zu. Auch der Schulungsbedarf bei Cvi42 ist höher, um dadurch die Kongruenz der Ergebnisse zu verbessern.

Die in dieser Studie gemachten Ergebnisse bestätigen die Resultate anderer Studien. Die Güte der Tracking Algorithmen scheinen neben anderen Faktoren maßgeblich Einfluss auf die jeweiligen Ergebnisse zu haben. Wichtig hierbei ist, dass Variationen in diesem Punkt auch durch neuere Softwareupdates und somit Veränderungen der Tracking Algorithmen als Einfluss nicht außer Acht gelassen werden sollten. So können Updates zu Verbesserungen der Konturgüte maßgeblich Einfluss auf die

berechneten Parameter haben. Um die Ergebnisse dieser Arbeit weiter in den aktuellen Forschungskontext setzen zu können, wären die Ergebnisse von Zhang et. al zu GCS und GRS von Interesse, die bisher aber nicht veröffentlicht sind.

Die Ergebnisse der Interobserver Reproduzierbarkeit zeigen, dass im Grunde die Übereinstimmung der Werte zwischen den einzelnen Programmen gut bis exzellent ist. Die Werte zeigen, dass der Effekt des Untersuchers auf die Ergebnisse in einigen Programmen höher ist als in anderen und dass der Grad der Automatisierung bzw. die Genauigkeit der automatischen Konturerkennung maßgeblich einen Einfluss auf die Reproduzierbarkeit von Ergebnissen hat. Dies wiederum bedeutet, dass Programme wie Cvi42 einen höheren Schulungsbedarf benötigen, um Standardisierung zu erzielen und somit individuelle Unterschiede auf die Generierung der Strainwerte zu reduzieren.

Weitere Untersuchungen sind notwendig, die insbesondere die Qualität des Deformable Registration Algorithmus aufgrund seiner vermeintlich sehr hohen Reliabilität unter Beweis stellen.

## **5.5 Klinische Anwendung und zukünftige Entwicklungen**

Durch Verbesserung von Standardisierung und softwareübergreifender Vergleichbarkeit kann die klinische Anwendung der kardialen Strainanalyse an Bedeutung gewinnen. In vielen Bereichen konnten Studien Vorteile in der Früherkennung kardialer Erkrankungen nachweisen und eine Vorzugsstellung gegenüber etablierten Parametern belegen. Liu et al. stellten Veränderungen aller drei Globalen Strainparameter bei Patienten mit hypertensiver Kardiopathie fest, noch bevor die LV EF vermindert war. So führte der vermehrte Einbau von Kollagenfasern zu einer Verminderung des GLS und die Zunahme des LVMI zeigte eine Assoziation mit dem GRS und der GCS (70). Bei Patienten mit Myokarditis gelingt mittels des GCS der Nachweis einer noch subklinischen LV Dysfunktion bei noch bestehender LV EF (42). Weiterhin korreliert der GLS mit der Größe des Ödems und der GCS-Veränderung (43).

Bereits 2014 konnte bei Patienten mit einer ischämischen Herzerkrankung unter mittlerer- und Hochdosis Dobutamin Stress signifikante Veränderungen des GCS zwischen stenotischen und normalen Segmenten festgestellt werden. Weiterhin zeigte sich eine Überlegenheit gegenüber einer rein visuellen Beurteilung der Wandbewegung unter Dobutamin Stress(41).

In einer zuletzt veröffentlichten Studie wurden Patienten mit akutem Brustschmerz mittels einer neuen Aufnahmetechnik Fast Scenc (fSCENC) untersucht. Hier zeigte sich der GCS als ein guter Marker, um Patienten mit einer obstruktiven KHK zu identifizieren und dies ohne die Verwendung von Laborparametern wie Troponin, EKG oder Stresstests mittels Dobutamingaben (71). Weiterhin konnte nachgewiesen werden, dass der durch Feature Tracking generierte GCS signifikant mit der Infarktgröße korrelierte (72).

Während im Bereich der Chemotherapie induzierten Kardiotoxizität der durch STE generierte GLS bereits in Leitlinien der European Association of Cardiology oder der American Society of Echocardiography aufgenommen wurde, ist davon auszugehen, dass in Zukunft die durch CMR generierten Strainparameter diese ergänzen können. So zeigten bereits Studien bei Patienten mit Anthracyclin basierten Chemotherapieprotokollen eine Reduktion der LV EF als auch der LV GCS(73). Die Ergebnisse können insbesondere der therapeutischen Entscheidungsfindung während der Therapie helfen.

Im Bereich der ischämischen Kardiopathie, der dilatativen Kardiomyopathie und der Herzinsuffizienz können zukünftige Ereignisse durch Veränderungen des GLS vorhergesagt werden (35, 74). Im Rahmen der dilatativen Kardiomyopathie zeigte sich der GLS als ein prognostischer Marker hinsichtlich kardialer Mortalität.

Klinische Anwendungsgebiete können demnach in der Früherkennung, der Prognose und somit auch der therapeutischen Entscheidungsfindung liegen. Bisher existiert jedoch nur eine geringe Anzahl von Studien, die Referenzwerte für die therapeutische Entscheidungsfindung festlegen. Dies wird unter anderem dadurch bedingt, dass kein Konsens hinsichtlich normaler Referenzwerte existiert. Sie sind abhängig von demographischen und hämodynamischen Faktoren, Software, oder auch der zeitlichen und räumlichen Auflösung. Somit sind in Zukunft Studien mit großen Fallkohorten notwendig (17, 44). Mit Weiterentwicklung der künstlichen Intelligenz wird ein großes Potential in der Verbesserung der Konturqualität und eine Beschleunigung des Bearbeitungsprozesses gesehen (35). Dies wird weitere Vorteile für die klinische Anwendung bieten.

Zuvor wird jedoch die Frage nach der Standardisierung der Kalkulation zu klären sein. So sollten Unterschiede zwischen softwareabhängigen Kalkulationsmechanismen wie Optical Flow oder DRA ausgeglichen werden, um übergreifende Vergleiche zu ermöglichen.

## 5.6 Limitationen

Die Daten dieser Studie basieren auf der MaiCo-Studie aus dem Jahr 2018 und einer retrospektiven Auswahl der T1T2T3-Studie aus dem Jahr 2016.

Da die MaiCo-Studie eine sehr junge Kohorte darstellte, wurde eine Auswahl von Probanden aus der T1T2T3-Studie integriert, um eine Diversifizierung der Altersstruktur zu erreichen. Wichtig dabei war, dass eine repräsentative Verteilung der Geschlechter erzeugt wurde. Bisher ist die Studienlage zu altersspezifischen Effekten auf die Strainparameter nicht eindeutig. Während Taylor et. al. 2015 feststellten, dass der GCS mit dem Alter zunimmt und keinerlei Alterseffekte für GRS und GLS vorlagen, kamen Peng et. al. 2018 zu dem Ergebnis, dass der GCS mit dem Alter ab- und der GRS zunahm. Eine andere Studie aus dem Jahr 2019 konnte keinen Zusammenhang zwischen Alter und Strainwerten feststellen (37-39). Daraus ist zu schließen, dass zum jetzigen Zeitpunkt kein Hinweis dafür vorliegt, dass die Altersstruktur einen maßgeblichen Effekt auf die Ergebnisse hat, sodass es noch weitere Studien in weit größeren Kollektiven benötigt, die diese Frage abschließend klären.

Obwohl beide Studien (MaiCo und T1T2T3) mittels eines 3T-Gerätes durchgeführt wurden, lagen Unterschiede in der Bildqualität vor. Dies ist auf die Untersuchungsplanung zurückzuführen. Die MaiCo-Studie wurde mit einer Cardiac-Dot-Engine von Siemens durchgeführt. Dabei handelt es sich um eine semiautomatisierte Planungshilfe für alle kardialen Schnittebenen und Funktionsaufnahmen. Die T1T2T3-Studie wurde hingegen klassisch manuell durch eine\*n Medizinische-Technische\*n Radiologische\*n Assistenten\*in durchgeführt. Im Vergleich zeigte sich eine heterogenere Aufnahmequalität, insbesondere im Bereich der Langachse. Dies erforderte in dem Kollektiv der T1T2T3-Studie einen größeren Nachbearbeitungsaufwand. Eine Beeinflussung der Strainparameter hierdurch ist nicht sicher auszuschließen. Da zum Beispiel die Verkürzung des Herzens im Rahmen der Systole entlang der Langachse bei versetztem Anschnitt ggf. kürzer ausfallen könnte. Dass Bilddaten unterschiedlicher Güte ausgewertet werden müssen, ist in der Praxis gängig, inwiefern diese neben dem Untersucher und dem jeweiligen Algorithmus die Auswertungsergebnisse beeinflussen, ist eine relevante Forschungsfrage, die bereits Amzulescu et. al aufzeigten und in weiteren Untersuchungen nachverfolgt werden sollte (21).

Um die Ergebnisse etwa von EF und EDVi einordnen zu können, wurde Bezug auf die von Kawel-Boehm et. al. publizierten Werte genommen (29). Der Vergleich hat Schwächen, da Unterschiede bei den verwendeten Softwareprogrammen vorliegen. Eine weitere Limitation liegt darin, dass die Untersuchung auf links ventrikuläre globale Strainparameter beschränkt wurde. Da einige Studien den Nachweis erbrachten, dass die Übereinstimmung für segmentale Strainparameter deutlich geringer ist, wurde in dieser Studie der Fokus auf globale Strainparameter gelegt (35). Eine Aussage über die segmentalen Strainparameter kann somit nicht getroffen werden. Insbesondere aber die Tatsache, dass TruFiStrain selbst auf segmentaler Ebene eine hohe Retest-Reliabilität zugeschrieben wird, wäre eine Überprüfung dieser von Interesse (66). Eine weitere Limitation der Arbeit ist, dass Strain Rates nicht untersucht wurden. Einige Publikationen haben gezeigt, dass diese von Relevanz sein können. So konnten Hagdorn et. al. an Patienten mit Zustand nach Korrektur einer Fallot'schen Tetralogie nachweisen, dass Strain Rates das Auftreten von ventrikulären Tachykardien voraussagen können (75). Da der Fokus in dieser Arbeit auf die globalen Strainparameter und insbesondere das Thema softwareabhängige Vergleichbarkeit gelegt wurde, stellen die Strain Rates ein weiteres wichtiges Forschungsdesiderat dar. Sie sollten in zukünftigen Studien auf softwareabhängige Unterschiede untersucht werden.

Da sich die Arbeit mit dem übergeordneten Thema Vergleichbarkeit und Übertragbarkeit anbieterabhängiger Unterschiede beschäftigt, stellt sich immer auch die Frage danach, welche Software die Ergebnisse erzeugt, die der Realität am nächsten kommen. In vielen Studien wird darauf verwiesen, dass Tissue Tagging der aktuelle Goldstandard bei der Erzeugung von Strainparametern sei. Es handelt sich um eine bereits validierte Methode, die als Richtgröße der hier erzeugten Ergebnisse dienen könnte. Leider lagen keine Tagging-Sequenzen vor, die für eine Einordnung zur Verfügung standen. Um dies auszugleichen, wurde zum Beispiel in den Bland – Altman - Diagrammen der Mittelwert aus allen drei Programmen gebildet und die Differenzen der jeweiligen Programme dagegen aufgetragen. Der Mittelwert als Referenzgröße stellt eine methodische Schwäche dar, die dazu führt, dass Abweichungen unterrepräsentiert in der statistischen Auswertung dargestellt werden.

In Zusammenschau wäre für Feature Tracking eine etablierte in-vitro- sowie in vivo Validierungsmethode ähnlich der Sonomicrometry für die STE wichtig, wie es z.B.

Amundsen et. al im Jahr 2006 durchgeführt hatten (76). Da eine solche Methode nicht zur Verfügung steht, konnte keine Aussage darüber getroffen werden, welches Programm den wahren Wert am ehesten abbildet. Die Definition eines „Goldstandards“ ist eine maßgebliche Herausforderung, die für eine Weiterentwicklung notwendig ist.

## 6. Zusammenfassung

Das Ziel dieser Studie war es, die Vergleichbarkeit und Reproduzierbarkeit von drei FT-Softwareprogrammen (Cvi42, SuiteHeart und TrufiStrain) bei der Berechnung der globalen Strainparameter zu evaluieren und Referenzwerte für eine gesunde Kohorte zu definieren. Die Studienpopulation bestand aus einem Kollektiv von 98 gesunden Freiwilligen im Alter von 21 bis 72 Jahren, die ein 3T-MRT erhielten. Die Kohorte setzte sich aus zwei Studien zusammen: 1. Der Mainzer-Kohorten-Pilot-Studie (MaiCo-Studie) aus dem Jahr 2018 und 2. Einer retrospektiven Datenanalyse einer Auswahl der Studienpopulation der T1T2T3-Studie aus dem Jahr 2016. Aus den SSFP-CINE Sequenzen wurden sowohl volumetrische Daten, wie EF, EDV, ESV, LVM, als auch globale Strains ermittelt.

Besonders bei dieser Studie war, dass mit SuiteHeart und TrufiStrain zwei neue Feature-Tracking Softwareprogramme untersucht wurden, für die zum Zeitpunkt der Durchführung wenige oder keine Vergleichsparameter vorlagen.

Die Hauptergebnisse der Studie sind:

- (1) Die durch die drei Programme untersuchten klinischen etablierten Funktionsparameter (wie z.B. die EF) unterscheiden sich signifikant voneinander. Die von SuiteHeart und TrufiStrain ermittelten Werte weichen am stärksten von externen Referenzkollektiven ab.
- (2) Die ermittelten globalen Strainparameter sind abhängig von dem jeweiligen Softwareprogramm und nicht gegenseitig übertragbar. Die geringsten Unterschiede zeigten sich beim GCS und GRS, die höchsten beim GLS.
- (3) Die durch TrufiStrain berechneten globalen Strainparameter weichen am stärksten von dem bisher publizierten allgemeinen MRT-FT-Strain Referenzwerten ab.
- (4) Globale Strainparameter sind signifikant abhängig vom Geschlecht. Frauen haben signifikant unterschiedliche Strainparameter als Männer.
- (5) Die Inter- und Intrarater Reproduzierbarkeiten sind softwareübergreifend sehr gut.

Die in dieser Studie erzielten Ergebnisse sind konkordant mit der aktuell publizierten Studienlage und betonen, dass derzeit softwareübergreifende Vergleiche von Strainparametern nicht sinnvoll sind. Für sich genommen erzielen die Programme aber gute bis exzellente reproduzierbare Ergebnisse. Es bestätigte sich, dass die Güte semiautomatisierter Konturen die Reproduzierbarkeit beeinflusst.

Insbesondere zu den Programmen SuiteHeart und TrufiStrain wären weitere Studien von Interesse, die sowohl die Volumetrie als auch die Strainkalkulation evaluieren. Weitere Arbeiten, die auch die Untersuchung segmentaler Strainparameter insbesondere auch mit Hilfe eines DR- Algorithmus untersuchen, wären notwendig.

Um den Strain als Parameter zur Früherkennung oder auch für therapeutische Entscheidungen in der klinischen Praxis weiter zu etablieren, sind weitere Studien mit großen Kohorten relevant, die zum einen die Definition gesunder Referenzwerte zulassen, aber auch die Definition von pathologischen Werten ermöglichen. Hierfür wären Arbeiten vonnöten, die Standardisierungen auf Ebene des Kalkulationsalgorithmus herausarbeiten, um die Vergleichbarkeit zwischen Softwareherstellern zu ermöglichen. Mithilfe eines Validierungsprozesses im Sinne eines „Goldstandards“, der durch jedes neue Softwareprogramm durchlaufen werden müsste, könnten Ergebnisse objektiver werden und insbesondere eine erfolgreiche klinische Anwendung möglich werden. Anderenfalls sollten softwareabhängige Referenzwerte – wie die in dieser Arbeit für ein lokales Kollektiv etablierten – durch die Hersteller herausgegeben werden.

## Literaturverzeichnis

1. Schuster A, Hor KN, Kowallick JT, Beerbaum P, Kutty S. Cardiovascular Magnetic Resonance Myocardial Feature Tracking: Concepts and Clinical Applications. *Circ Cardiovasc Imaging*. 2016;9(4):e004077.
2. Vasquez M, Nagel E. Clinical indications for cardiovascular magnetic resonance. *Heart*. 2019;105(22):1755-62.
3. Dandel M, Lehmkuhl H, Knosalla C, Suram lashvili N, Hetzer R. Strain and strain rate imaging by echocardiography - basic concepts and clinical applicability. *Curr Cardiol Rev*. 2009;5(2):133-48.
4. Eitel I, Stiermaier T, Lange T, Rommel KP, Koschalka A, Kowallick JT, et al. Cardiac Magnetic Resonance Myocardial Feature Tracking for Optimized Prediction of Cardiovascular Events Following Myocardial Infarction. *JACC Cardiovasc Imaging*. 2018;11(10):1433-44.
5. Morton G, Schuster A, Jogiya R, Kutty S, Beerbaum P, Nagel E. Inter-study reproducibility of cardiovascular magnetic resonance myocardial feature tracking. *J Cardiovasc Magn Reson*. 2012;14:43.
6. Schuster A, Stahnke VC, Unterberg-Buchwald C, Kowallick JT, Lamata P, Steinmetz M, et al. Cardiovascular magnetic resonance feature-tracking assessment of myocardial mechanics: Intervendor agreement and considerations regarding reproducibility. *Clin Radiol*. 2015;70(9):989-98.
7. Claus P, Omar AMS, Pedrizzetti G, Sengupta PP, Nagel E. Tissue Tracking Technology for Assessing Cardiac Mechanics: Principles, Normal Values, and Clinical Applications. *JACC Cardiovasc Imaging*. 2015;8(12):1444-60.
8. Schild PDmHH. MRI made easy Bayer Vital GmbH; 1997.
9. Hendrix A. Magnete, Spins und Resonanzen - Eine Einführung in die Grundlagen der Magnetresonantomographie. In: Solutions SM, editor. Siemens Medical Solutions2003.
10. Bern RU. Aufbau eines MRT 2015 [Available from: <https://vetsuisse.com/vet-impl/lernmodule/htmls/carpet.html?radiosurfvet>|radgeneral|mrtomography|mrtconstruct
11. Mihaljevic. CT- und MRT Atlas [eBook]2009.
12. Allen D. Elster MF. QUESTIONS AND ANSWERS IN MRI Mallinckrodt Institute of Radiology  
Washington University School of Medicine  
St. Louis, MO2019 [Available from: <http://mriquestions.com/index.html>].

13. Russo V, Lovato L, Ligabue G. Cardiac MRI: technical basis. *Radiol Med*. 2020;125(11):1040-55.
14. Pfeiffer MP, Biederman RW. Cardiac MRI: A General Overview with Emphasis on Current Use and Indications. *Med Clin North Am*. 2015;99(4):849-61.
15. Frangi AF, Niessen WJ, Viergever MA. Three-dimensional modeling for functional analysis of cardiac images: a review. *IEEE Trans Med Imaging*. 2001;20(1):2-25.
16. Sakuma H, Takeda K, Higgins CB. Fast magnetic resonance imaging of the heart. *Eur J Radiol*. 1999;29(2):101-13.
17. Scatteia A, Baritussio A, Bucciarelli-Ducci C. Strain imaging using cardiac magnetic resonance. *Heart Fail Rev*. 2017;22(4):465-76.
18. Almutairi HM, Boubertakh R, Miquel ME, Petersen SE. Myocardial deformation assessment using cardiovascular magnetic resonance-feature tracking technique. *Br J Radiol*. 2017;90(1080):20170072.
19. Gotte MJ, Germans T, Russel IK, Zwanenburg JJ, Marcus JT, van Rossum AC, et al. Myocardial strain and torsion quantified by cardiovascular magnetic resonance tissue tagging: studies in normal and impaired left ventricular function. *J Am Coll Cardiol*. 2006;48(10):2002-11.
20. Shehata ML, Cheng S, Osman NF, Bluemke DA, Lima JA. Myocardial tissue tagging with cardiovascular magnetic resonance. *J Cardiovasc Magn Reson*. 2009;11:55.
21. Amzulescu MS, De Craene M, Langet H, Pasquet A, Vancraeynest D, Pouleur AC, et al. Myocardial strain imaging: review of general principles, validation, and sources of discrepancies. *Eur Heart J Cardiovasc Imaging*. 2019;20(6):605-19.
22. van Everdingen WM, Zweerink A, Nijveldt R, Salden OAE, Meine M, Maass AH, et al. Comparison of strain imaging techniques in CRT candidates: CMR tagging, CMR feature tracking and speckle tracking echocardiography. *Int J Cardiovasc Imaging*. 2018;34(3):443-56.
23. Pedrizzetti G, Claus P, Kilner PJ, Nagel E. Principles of cardiovascular magnetic resonance feature tracking and echocardiographic speckle tracking for informed clinical use. *J Cardiovasc Magn Reson*. 2016;18(1):51.
24. Leitman M, Lysyansky P, Sidenko S, Shir V, Peleg E, Binenbaum M, et al. Two-dimensional strain-a novel software for real-time quantitative echocardiographic assessment of myocardial function. *J Am Soc Echocardiogr*. 2004;17(10):1021-9.

25. Voigt JU, Pedrizzetti G, Lysyansky P, Marwick TH, Houle H, Baumann R, et al. Definitions for a common standard for 2D speckle tracking echocardiography: consensus document of the EACVI/ASE/Industry Task Force to standardize deformation imaging. *J Am Soc Echocardiogr*. 2015;28(2):183-93.
26. Schulz-Menger J, Bluemke DA, Bremerich J, Flamm SD, Fogel MA, Friedrich MG, et al. Standardized image interpretation and post-processing in cardiovascular magnetic resonance - 2020 update : Society for Cardiovascular Magnetic Resonance (SCMR): Board of Trustees Task Force on Standardized Post-Processing. *J Cardiovasc Magn Reson*. 2020;22(1):19.
27. Hor KN, Baumann R, Pedrizzetti G, Tonti G, Gottliebson WM, Taylor M, et al. Magnetic resonance derived myocardial strain assessment using feature tracking. *J Vis Exp*. 2011(48).
28. Schmidt RF, Lang F, Heckmann M. *Physiologie des Menschen: Mit Pathophysiologie*: Springer Berlin Heidelberg; 2010.
29. Kawel-Boehm N, Hetzel SJ, Ambale-Venkatesh B, Captur G, Francois CJ, Jerosch-Herold M, et al. Reference ranges ("normal values") for cardiovascular magnetic resonance (CMR) in adults and children: 2020 update. *J Cardiovasc Magn Reson*. 2020;22(1):87.
30. T.G. P. Left ventricular Mass. *Encyclopedia of Stress* 2007;2nd:577-8.
31. Puntmann VO, Valbuena S, Hinojar R, Petersen SE, Greenwood JP, Kramer CM, et al. Society for Cardiovascular Magnetic Resonance (SCMR) expert consensus for CMR imaging endpoints in clinical research: part I - analytical validation and clinical qualification. *J Cardiovasc Magn Reson*. 2018;20(1):67.
32. Reichek N. Myocardial Strain: Still a Long Way to Go. *Circ Cardiovasc Imaging*. 2017;10(11).
33. Sardana M, Konda P, Hashmath Z, Oldland G, Gaddam S, Miller R, et al. Usefulness of Left Ventricular Strain by Cardiac Magnetic Resonance Feature-Tracking to Predict Cardiovascular Events in Patients With and Without Heart Failure. *Am J Cardiol*. 2019;123(8):1301-8.
34. Schuster A, Backhaus SJ, Stiermaier T, Eitel I. Prognostic utility of global longitudinal strain in myocardial infarction. *World J Cardiol*. 2018;10(5):35-7.
35. Xu J, Yang W, Zhao S, Lu M. State-of-the-art myocardial strain by CMR feature tracking: clinical applications and future perspectives. *Eur Radiol*. 2022.

36. Chong A, MacLaren G, Chen R, Connelly KA. Perioperative applications of deformation (myocardial strain) imaging with speckle-tracking echocardiography. *J Cardiothorac Vasc Anesth.* 2014;28(1):128-40.
37. Taylor RJ, Moody WE, Umar F, Edwards NC, Taylor TJ, Stegemann B, et al. Myocardial strain measurement with feature-tracking cardiovascular magnetic resonance: normal values. *Eur Heart J Cardiovasc Imaging.* 2015;16(8):871-81.
38. Peng J, Zhao X, Zhao L, Fan Z, Wang Z, Chen H, et al. Normal Values of Myocardial Deformation Assessed by Cardiovascular Magnetic Resonance Feature Tracking in a Healthy Chinese Population: A Multicenter Study. *Front Physiol.* 2018;9:1181.
39. Mangion K, Burke NMM, McComb C, Carrick D, Woodward R, Berry C. Feature-tracking myocardial strain in healthy adults- a magnetic resonance study at 3.0 tesla. *Sci Rep.* 2019;9(1):3239.
40. Song Y, Li L, Chen X, Shao X, Lu M, Cheng J, et al. Early Left Ventricular Diastolic Dysfunction and Abnormal Left Ventricular-left Atrial Coupling in Asymptomatic Patients With Hypertension: A Cardiovascular Magnetic Resonance Feature Tracking Study. *J Thorac Imaging.* 2022;37(1):26-33.
41. Schneeweis C, Qiu J, Schnackenburg B, Berger A, Kelle S, Fleck E, et al. Value of additional strain analysis with feature tracking in dobutamine stress cardiovascular magnetic resonance for detecting coronary artery disease. *J Cardiovasc Magn Reson.* 2014;16:72.
42. Chen X, Hu H, Pan J, Shu J, Hu Y, Yu R. Performance of cardiovascular magnetic resonance strain in patients with acute myocarditis. *Cardiovasc Diagn Ther.* 2020;10(4):725-37.
43. Fischer K, Obrist SJ, Erne SA, Stark AW, Marggraf M, Kaneko K, et al. Feature Tracking Myocardial Strain Incrementally Improves Prognostication in Myocarditis Beyond Traditional CMR Imaging Features. *JACC Cardiovasc Imaging.* 2020;13(9):1891-901.
44. Vo HQ, Marwick TH, Negishi K. MRI-Derived Myocardial Strain Measures in Normal Subjects. *JACC Cardiovasc Imaging.* 2018;11(2 Pt 1):196-205.
45. Lim C, Blaszczyk E, Riazzy L, Wiesemann S, Schuler J, von Knobelsdorff-Brenkenhoff F, et al. Quantification of myocardial strain assessed by cardiovascular magnetic resonance feature tracking in healthy subjects-influence of segmentation and analysis software. *Eur Radiol.* 2021;31(6):3962-72.

46. Augustine D, Lewandowski AJ, Lazdam M, Rai A, Francis J, Myerson S, et al. Global and regional left ventricular myocardial deformation measures by magnetic resonance feature tracking in healthy volunteers: comparison with tagging and relevance of gender. *J Cardiovasc Magn Reson*. 2013;15:8.
47. Lawton JS, Cupps BP, Knutsen AK, Ma N, Brady BD, Reynolds LM, et al. Magnetic resonance imaging detects significant sex differences in human myocardial strain. *Biomed Eng Online*. 2011;10:76.
48. Liu H, Yang D, Wan K, Luo Y, Sun JY, Zhang TJ, et al. Distribution pattern of left-ventricular myocardial strain analyzed by a cine MRI based deformation registration algorithm in healthy Chinese volunteers. *Sci Rep*. 2017;7:45314.
49. Andre F, Steen H, Matheis P, Westkott M, Breuninger K, Sander Y, et al. Age- and gender-related normal left ventricular deformation assessed by cardiovascular magnetic resonance feature tracking. *J Cardiovasc Magn Reson*. 2015;17:25.
50. Bucius P, Erley J, Tanacli R, Zieschang V, Giusca S, Korosoglou G, et al. Comparison of feature tracking, fast-SENC, and myocardial tagging for global and segmental left ventricular strain. *ESC Heart Fail*. 2020;7(2):523-32.
51. Barreiro-Perez M, Curione D, Symons R, Claus P, Voigt JU, Bogaert J. Left ventricular global myocardial strain assessment comparing the reproducibility of four commercially available CMR-feature tracking algorithms. *Eur Radiol*. 2018;28(12):5137-47.
52. Morais P, Marchi A, Bogaert JA, Dresselaers T, Heyde B, D'Hooge J, et al. Cardiovascular magnetic resonance myocardial feature tracking using a non-rigid, elastic image registration algorithm: assessment of variability in a real-life clinical setting. *J Cardiovasc Magn Reson*. 2017;19(1):24.
53. Backhaus SJ, Metschies G, Billing M, Kowallick JT, Gertz RJ, Lapinskas T, et al. Cardiovascular magnetic resonance imaging feature tracking: Impact of training on observer performance and reproducibility. *PLoS One*. 2019;14(1):e0210127.
54. Backhaus SJ, Schuster A, Lange T, Stehning C, Billing M, Lotz J, et al. Impact of fully automated assessment on interstudy reproducibility of biventricular volumes and function in cardiac magnetic resonance imaging. *Sci Rep*. 2021;11(1):11648.
55. Zhang Y, Mui D, Chirinos JA, Zamani P, Ferrari VA, Chen Y, et al. Comparing cardiovascular magnetic resonance strain software packages by their abilities to discriminate outcomes in patients with heart failure with preserved ejection fraction. *J Cardiovasc Magn Reson*. 2021;23(1):55.

56. Zange L, Muehlberg F, Blaszczyk E, Schwenke S, Traber J, Funk S, et al. Quantification in cardiovascular magnetic resonance: agreement of software from three different vendors on assessment of left ventricular function, 2D flow and parametric mapping. *J Cardiovasc Magn Reson*. 2019;21(1):12.
57. Messalli G, Palumbo A, Maffei E, Martini C, Seitun S, Aldrovandi A, et al. Assessment of left ventricular volumes with cardiac MRI: comparison between two semiautomated quantitative software packages. *Radiologia Medica*. 2009;114(5):718-27.
58. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*. 2016;15(2):155-63.
59. Schulz-Menger J, Bluemke DA, Bremerich J, Flamm SD, Fogel MA, Friedrich MG, et al. Standardized image interpretation and post processing in cardiovascular magnetic resonance: Society for Cardiovascular Magnetic Resonance (SCMR) board of trustees task force on standardized post processing. *J Cardiovasc Magn Reson*. 2013;15:35.
60. NeoSoft L. Instructions for Use, suiteHEART Software, cMRI Analysis Software [Manufacturer Handbook ]2018.
61. Jolly MP. Trufi Strain Demonstrator User Manual [User Manual]. Siemens.
62. Houbois CP, Nolan M, Somerset E, Shalmon T, Esmaeilzadeh M, Lamacie MM, et al. Serial Cardiovascular Magnetic Resonance Strain Measurements to Identify Cardiotoxicity in Breast Cancer: Comparison With Echocardiography. *JACC Cardiovasc Imaging*. 2021;14(5):962-74.
63. Jeong D, Gladish G, Chitiboi T, Fradley MG, Gage KL, Schiebler ML. MRI in cardio-oncology: A review of cardiac complications in oncologic care. *J Magn Reson Imaging*. 2019;50(5):1349-66.
64. Liu B, Dardeer AM, Moody WE, Hayer MK, Baig S, Price AM, et al. Reference ranges for three-dimensional feature tracking cardiac magnetic resonance: comparison with two-dimensional methodology and relevance of age and gender. *Int J Cardiovasc Imaging*. 2018;34(5):761-75.
65. Lamacie MM, Thavendiranathan P, Hanneman K, Greiser A, Jolly MP, Ward R, et al. Quantification of global myocardial function by cine MRI deformable registration-based analysis: Comparison with MR feature tracking and speckle-tracking echocardiography. *Eur Radiol*. 2017;27(4):1404-15.
66. Wang J, Li W, Sun J, Liu H, Kang Y, Yang D, et al. Improved segmental myocardial strain reproducibility using deformable registration algorithms compared

with feature tracking cardiac MRI and speckle tracking echocardiography. *J Magn Reson Imaging*. 2018;48(2):404-14.

67. Mehta S, Bastero-Caballero RF, Sun Y, Zhu R, Murphy DK, Hardas B, et al. Performance of intraclass correlation coefficient (ICC) as a reliability index under various distributions in scale reliability studies. *Stat Med*. 2018;37(18):2734-52.

68. Jolly MP. Trufi Strain Demonstrator User Manual [User Manual]. Siemens.

69. Obokata M, Nagata Y, Wu VC, Kado Y, Kurabayashi M, Otsuji Y, et al. Direct comparison of cardiac magnetic resonance feature tracking and 2D/3D echocardiography speckle tracking for evaluation of global left ventricular strain. *Eur Heart J Cardiovasc Imaging*. 2016;17(5):525-32.

70. Liu H, Wang J, Pan Y, Ge Y, Guo Z, Zhao S. Early and Quantitative Assessment of Myocardial Deformation in Essential Hypertension Patients by Using Cardiovascular Magnetic Resonance Feature Tracking. *Sci Rep*. 2020;10(1):3582.

71. Riffel JH, Siry D, Salatzki J, Andre F, Ochs M, Weberling LD, et al. Feasibility of fast cardiovascular magnetic resonance strain imaging in patients presenting with acute chest pain. *PLoS One*. 2021;16(5):e0251040.

72. Khan JN, Singh A, Nazir SA, Kanagala P, Gershlick AH, McCann GP. Comparison of cardiovascular magnetic resonance feature tracking and tagging for the assessment of left ventricular systolic strain in acute myocardial infarction. *Eur J Radiol*. 2015;84(5):840-8.

73. Drafts BC, Twomley KM, D'Agostino R, Jr., Lawrence J, Avis N, Ellis LR, et al. Low to moderate dose anthracycline-based chemotherapy is associated with early noninvasive imaging evidence of subclinical cardiovascular disease. *JACC Cardiovasc Imaging*. 2013;6(8):877-85.

74. Romano S, Romer B, Evans K, Trybula M, Shenoy C, Kwong RY, et al. Prognostic Implications of Blunted Feature-Tracking Global Longitudinal Strain During Vasodilator Cardiovascular Magnetic Resonance Stress Imaging. *JACC Cardiovasc Imaging*. 2020;13(1 Pt 1):58-65.

75. Hagdorn QAJ, Vos JDL, Beurskens NEG, Gorter TM, Meyer SL, van Melle JP, et al. CMR feature tracking left ventricular strain-rate predicts ventricular tachyarrhythmia, but not deterioration of ventricular function in patients with repaired tetralogy of Fallot. *Int J Cardiol*. 2019;295:1-6.

76. Amundsen BH, Helle-Valle T, Edvardsen T, Torp H, Crosby J, Lyseggen E, et al. Noninvasive myocardial strain measurement by speckle tracking echocardiography:

validation against sonomicrometry and tagged magnetic resonance imaging. J Am Coll Cardiol. 2006;47(4):789-93.

## **Danksagung**

Mein Dank für die vielfältige Unterstützung möchte ich in erster Linie meiner Frau und meiner Familie aussprechen, die mich auf dem Weg durch das Studium und die Dissertation begleitet haben.

Ein besonderer Dank gilt auch meinen Betreuern die mir stets für fachliche Fragen zur Seite standen und mich bei der Erarbeitung und Einführung in das wissenschaftliche Arbeiten unterstützt haben.

# Lebenslauf

