

Aus dem Institut für diagnostische und interventionelle Radiologie

der Heinrich-Heine-Universität Düsseldorf

Direktor: Univ.-Prof. Dr. med. Gerald Antoch

**Prädiktion quantitativer respiratorischer Parameter aus kardialen
Echtzeit-MR-Bildern mittels maschinellen Lernens unter Verwendung
von MR-kompatibler Spirometrie**

Dissertation

Zur Erlangung des Grades eines Doktors der Medizin

der Medizinischen Fakultät der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Halima Malik

2026

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

gez.:

Dekan: Prof. Dr. med. Nikolaj Klöcker

Erstgutachter: Prof. Dr. med. Frank Pillekamp

Zweitgutachter: Prof. Dr. rer. nat. Ulrich Flögel

For the hands that carried me.

Teile dieser Arbeit wurden veröffentlicht in: Malik H, Uelwer T, Röwer LM, et al. Machine-learning-based prediction of respiratory flow and lung volume from real-time cardiac MRI using MR-compatible spirometry. *Med Phys*. 2025;52(8):e18019. doi:10.1002/mp.18019

Zusammenfassung

Die Funktion des Herzens wird wesentlich durch die Atmung beeinflusst. Diese Interaktion zu untersuchen, ist daher von großem wissenschaftlichem Interesse. Während die Echtzeit-Magnetresonanztomographie (MRT) die kontinuierliche Erfassung der kardialen Funktion ermöglicht, ist die Spirometrie der Goldstandard zur Messung der Lungenfunktion. Die Kombination beider Verfahren eignet sich daher sehr gut, die Herz-Lungen-Interaktion zu untersuchen. Besonders bei herzkranken Kindern wäre allerdings eine Kürzung der Untersuchungsdauer mit Atemmaske wünschenswert.

Ziel dieser Arbeit war die Entwicklung und Validierung eines Modells für maschinelles Lernen (ML), das nach einer kurzen Trainingsphase mit Atemmaske, die quantitativen Atemparameter aus Bildern der folgenden kardialen Echtzeit-MRT vorhersagen kann.

Bei zehn gesunden erwachsenen Probanden wurde die kardiale Kurzachsendvolumetrie mittels Echtzeit-MRT während freier Atmung in Kombination mit MR-kompatibler Spirometrie (MRkS) aufgenommen (1.5 T; 30 Bilder/s; 19 Schichten). Aus den MR-Bildern wurden mittels der iterativen Lucas-Kanade-Methode optische Flussvektoren (OFV) berechnet. Jeder Datensatz eines Probanden wurde in einen Trainingsatz (OFV der MR-Bilder und MRkS) sowie einen Testsatz (nur OFV der MR-Bilder) unterteilt. Ein ML-Modell wurde mittels Ridge Regression trainiert, um den Atemfluss des ungesehenen Testsatzes vorherzusagen. Zur Optimierung der Trainingsphase wurden die Trainingsätze bezüglich Atemmuster und Protokolldauer variiert. Die statistische Validierung der prädiktiven Güte erfolgte anhand der Mittelwerte der Determinationskoeffizienten (R^2), der relativen mittleren quadratischen Fehler (rMSE) und der Bland-Altman-Analysen der absoluten Tidalvolumendifferenz (aTVD) und der absoluten maximalen Flussdifferenz (aMFD).

Ein kurzes Trainingsprotokoll (100 s), das aus Aufnahmen jeder zweiten Schicht mit jeweils zwei normalen Atemzügen bestand, erzielte ein Optimum zwischen prädiktiver Güte und Trainingszeit. Für den Atemfluss ergaben sich Mittelwerte von $R^2 = 0,984$, rMSE = 0,015 sowie ein Bias (aMFD) von -0,01 L/s (95 %-Konfidenzintervall (KI): +0,084/-0,1 L/s). Für das Lungenvolumen lagen die Mittelwerte bei $R^2 = 0,990$, rMSE = 0,003 und der Bias (aTVD) bei 4,27 mL (95 %-KI: +33/-24 mL). Die Anwendung von trainierten Modellen auf weitere Studien bekannter oder unbekannter Probanden konnte gute qualitative Resultate erzielen.

Zusammenfassend konnte gezeigt werden, dass nach einer kurzen Trainingsphase mit MRkS ein ML-Modell quantitative respiratorische Daten aus ungesehenen MR-Bildern der restlichen Studie mit hoher Genauigkeit vorhersagen konnte und somit eine Reduktion der Akquisitionszeit mit MRkS möglich ist.

Abstract

Heart function is significantly affected by respiration, and analyzing this interaction is of great scientific interest. While real-time magnetic resonance imaging (RT-MRI) enables the continuous monitoring of cardiac function, spirometry is the gold standard for assessing lung function. Combining both techniques offers great possibilities to analyze heart-lung interaction. However, particularly in children with heart disease, shortening the examination time with a breathing mask would be beneficial.

The aim of this study was to develop and validate a machine learning (ML) model that, after a brief training phase with a breathing mask, would be able to predict quantitative respiratory parameters from subsequent cardiac RT-MRI.

For ten healthy adult subjects, cardiac short-axis volumetry was acquired using RT-MRI during free breathing in combination with magnetic resonance-compatible spirometry (MRcS) (1.5 T; 30 frames/s; 19 slices). Optical flow vectors (OFVs) were computed from the MR images using the iterative Lucas-Kanade method. Each subject's dataset was divided into a training set (OFVs of the MR images with MRcS) and a test set (OFVs from MR images only). An ML-model was trained using ridge regression to predict respiratory flow from the unseen test set and to evaluate its predictive performance. To optimize the training phase, training sets were varied with respect to breathing pattern and protocol length. Statistical validation included mean coefficients of determination (R^2) and mean relative root mean squared errors (rMSE), as well as Bland-Altman analyses of absolute tidal volume difference (aTVD) and absolute maximal flow difference (aMFD).

A shortened training protocol (100 s), consisting of acquisitions of every second slice with two normal breaths each, achieved the optimal balance between predictive quality and training time. For airflow, mean values of $R^2 = 0.984$, rMSE = 0.015, and a bias (aMFD) of -0.01 L/s (95% confidence interval (CI): +0.084/-0.1 L/s) were shown. For lung volume, mean values of $R^2 = 0.990$, rMSE = 0.003 and a bias (aTVD) of 4.27 mL (95% CI: +33/-24 mL) were observed. The application of the trained models to additional studies of known and unknown subjects yielded good qualitative results.

In summary, it was shown that after a brief training phase with MRcS, an ML-model was able to predict quantitative respiratory data from unseen MR images of the remaining study with high accuracy, thereby allowing a reduction in acquisition time with MRcS.

Abkürzungsverzeichnis

aMF	absoluter maximaler Fluss
aMFD	absolute maximale Flussdifferenz
aTV	absolutes Tidalvolumen
aTVD	absolute Tidalvolumendifferenz
AZ	Atemzüge
BA	Bland-Altman
b-SSFP	balanced Steady-State Free Precession
EKG	Elektrokardiographie
FFT	schnelle Fouriertransformation
FAST	Fast Acquired Steady-State Technique
FLASH	Fast Low Angle Shot
FOV	Field Of View
KAV	Kurzachsenstapel-Volumetrie
KI	Konfidenzintervall
LOO-KV	Leave-One-Out Kreuzvalidierung
ML	maschinelles Lernen
MR	Magnetresonanz
MRkS	Magnetresonanz-kompatible Spirometrie
MRT	Magnetresonanztomographie
MV	mittventrikuläre Schicht
OFV	optische Flussvektoren
QC _{Fluss}	Qualitätskriterium Atemfluss
QC _{Volumen}	Qualitätskriterium Lungenvolumen
R ²	Determinationskoeffizient
RF-Impuls	Radiofrequenzimpuls
rMSE	relativer mittlerer quadratischer Fehler
S	Schicht
TE	Echozeit
TR	Repetitionszeit
TTE	transthorakale Echokardiographie
WTA	Winkel zwischen unterer Bildkante und der anterioren Thoraxwand im Kurzachsenschnitt

Inhaltsverzeichnis

1	Einleitung	1
1.1	Kardiale bildgebende Diagnostik bei Kindern	1
1.2	Kardiopulmonale Hämodynamik	2
1.2.1	Einfluss der physiologischen Atmung	2
1.2.2	Einfluss durch Atemanhaltenmanöver	2
1.2.3	Einfluss durch maschinelle Beatmung	3
1.3	Stand der Forschung	4
1.3.1	Physikalische Grundlagen des MRTs	4
1.3.2	Konventionelle kardiale MRT	6
1.3.3	<i>Gating</i> -Techniken	6
1.3.4	<i>Fast</i> -Sequenzen	10
1.3.5	Echtzeit-MRT	11
1.3.6	Akquise von Ateminformation	14
1.3.7	Ateminformation aus MR-Bildern	14
1.3.8	MR-kompatible Spirometrie	15
1.4	Maschinelles Lernen	16
1.4.1	Definition und Arten	16
1.4.2	Lineare Regression und Ridge Regression	18
1.4.3	Maschinelles Lernen im medizinischen Arbeitsbereich	19
1.5	Ziele dieser Arbeit	21
2	Material und Methoden	22
2.1	Datenakquise	22
2.2	MR-kompatible Spirometrie	25
2.3	Modellarchitektur	26
2.3.1	Vorverarbeitung	26
2.3.2	Respiratorisches Modell	26
2.3.3	Prädiktion	27
2.3.4	Nachbearbeitung	28
2.4	Modellkonfiguration und -optimierung	30
2.4.1	Originale kardiale KAV-Protokoll	30
2.4.2	Optimierung Training - Atemmuster und Anzahl der Atemzüge	30
2.4.3	Optimierung Training - Anzahl an Schichten	30
2.4.4	Trainiertes Modell - Wiederverwendung auf unbekanntem Probanden	33

2.4.5	Trainiertes Modell - Wiederverwendung auf zeitlich versetzter Aufnahme	34
2.5	Evaluationsmetrik.....	35
2.5.1	Qualitative Evaluation	35
2.5.2	Quantitative Evaluation.....	35
3	Ergebnisse.....	38
3.1	Prädiktion auf einer mittventrikulären Schicht	38
3.2	Optimierung Training - Atemmuster und Anzahl der Atemzüge.....	39
3.3	Prädiktion der originalen kardialen Volumetrie	43
3.4	Optimierung Training - Anzahl der Schichten	44
3.5	Regionale Gewichtsverteilung des Modells	51
3.6	Wiederverwendung von trainierten Modellen	52
3.6.1	Testdatensatz - unbekannter Proband	52
3.6.2	Testdatensatz - zeitlich versetzte Aufnahme	59
3.7	Technische Aspekte	63
4	Diskussion	67
5	Schlussfolgerung	73

1 Einleitung

In Deutschland wurde in den Jahren 2006 bis 2007 eine umfassende Studie durchgeführt, um die Gesamtprävalenz von kongenitalen Herzerkrankungen zu erfassen. Danach weist eines von 100 Neugeborenen einen Herzfehler auf (1).

1.1 Kardiale bildgebende Diagnostik bei Kindern

Zur bildmorphologischen Darstellung des Herzens werden vorwiegend nichtinvasive Methoden ausgewählt. Von klinischer Bedeutung sind insbesondere die transthorakale Echokardiographie (TTE) und die Magnetresonanztomographie (MRT) (2).

Am häufigsten wird die TTE des Herzens mittels Ultraschalls verwendet, da sie sich durch ihre schnelle Verfügbarkeit, ihre hohe zeitliche Auflösung, ihre niedrigen Kosten und ihre unmittelbare Darstellung der Aufnahmen für den Untersucher auszeichnet (3). Die TTE zeigt Grenzen bei der Anwendung. So kann bei älteren Kindern das Schallfenster beispielsweise durch Thoraxdeformitäten oder durch Luftüberlagerung eingeschränkt sein (3,4). Außerdem ist zu beachten, dass eine gewisse Expertise des Untersuchenden erforderlich ist und eine ausreichende Compliance der Kinder benötigt wird, um präzise Bilder zu erzeugen (4).

So wird immer häufiger auch auf die ebenfalls nichtinvasive und – im Gegensatz zu CT-Untersuchungen – nicht durch ionisierende Strahlen belastende kardiale MRT zurückgegriffen. Primär wird diese Technik zur genauen Darstellung von ventrikulären Dimensionen und ihrer funktionellen Analyse genutzt (5). Die MRT zur genaueren Beurteilung des Myokards und seiner Vitalität (3) genutzt. Auch erweist sich die MRT etwa bei der Visualisierung von benachbarten anatomischen Strukturen, wie den systemischen und pulmonalen Gefäßen oder dem tracheobronchialen Baum gegenüber der TTE überlegen (6).

Trotz ihrer vielseitigen Einsatzmöglichkeiten weist die MRT jedoch auch einige Nachteile auf. Zum einen sind die Kosten deutlich höher, zum anderen kann die Bildqualität durch metallische Implantate limitiert sein (7). Des Weiteren erfordert diese Form der Untersuchung eine gute Patientenkooperation. Oft treten Probleme durch die inadäquate Umsetzung der teils komplexen Anweisungen, wie etwa Atemhalteübungen von zumeist 7-10 Sekunden, gelegentlich aber auch bis zu 20 Sekunden, oder durch Bewegung aufgrund von Angst oder langer Untersuchungsdauer (8,9,10,11) auf. Dies macht bei Kindern nicht selten eine Sedierung oder in einigen Fällen sogar eine Allgemeinanästhesie erforderlich. Hierfür müssen mehr Ressourcen aufgebracht werden (12) und die Kinder werden den allgemeinen Risiken einer Sedierung oder Narkose ausgesetzt. Insbesondere

eine Beatmung kann darüber hinaus zu erheblichen Veränderungen der Hämodynamik des kardiopulmonalen Kreislaufs führen.

1.2 Kardiopulmonale Hämodynamik

Zum Verständnis der Hämodynamik des kardiopulmonalen Kreislaufs werden grundlegende Kenntnisse der Physiologie vorausgesetzt. Veränderungen der Atmung haben sowohl in Ruhe, beim Atemanhalten als auch während externer Beatmung eine Auswirkung auf die kardiale Hämodynamik. Entscheidende Einflussgrößen sind dabei das Lungenvolumen und die intrathorakalen Druckverhältnisse, insbesondere der intrapleurale Druck (13).

1.2.1 Einfluss der physiologischen Atmung

In Ruhe führt die Inspiration zu einem größeren Druckgradienten zwischen dem Herzen und dem großen systemischen Blutkreislauf. Dies basiert auf der Abnahme des intrapleuralen Drucks und der intrathorakalen Verlagerung des Herzens. Infolgedessen kommt es zu einem erhöhten venösen Rückstrom in den rechten Vorhof. Dies wiederum bewirkt eine Zunahme des enddiastolischen rechtsventrikulären Volumens (14) und des rechtsventrikulären Schlagvolumens.

Der erhöhte transpulmonale Druck und die relative Verlagerung des Herzens in inferiorer Richtung führen zu einer erhöhten Nachlast, gegen die der linke Ventrikel arbeiten muss. Dies führt zu einer transienten Abnahme des linksventrikulären Schlagvolumens mit erhöhtem endsystolischem linksventrikulärem Volumen. Gleichzeitig wird das Blut der alveolären Gefäße aufgrund der luftgefüllten Lunge in Richtung des linken Herzens gedrückt. Dieser Rückstrom und das verbliebene endsystolische linksventrikuläre Volumen führen in den nachfolgenden Zyklen zum Angleich des linksventrikulären Schlagvolumens. Während der Expiration wird der Mechanismus umgekehrt (14).

1.2.2 Einfluss durch Atemanhaltenmanöver

Verlängerte Atemanhaltenmanöver, wie sie bei konventionellen MRT-Untersuchungen durchgeführt werden, führen zu Veränderungen der kardialen Hämodynamik. Für diese Untersuchungen wird in Endexpiration oder Endinspiration die Luft oft für 7-10 Sekunden, in manchen Situationen sogar bis zu 20 Sekunden angehalten (8,9).

Ein Anhalten der Atmung in Endinspiration führt zu einem erhöhten intrathorakalen Druck mit verringertem venösem Rückfluss in den rechten Ventrikel und resultiert schließlich in einem geringeren Schlagvolumen. Hingegen zeigt das Atemanhalten bei geringerem Lungenvolumen, wie es beispielsweise bei Manövern in Endexpiration erfolgt, einen

geringeren Einfluss auf die kardiopulmonale Hämodynamik und ähnelt stärker dem physiologischen Einfluss der Atmung (15,16,17).

Sowohl das Anhalten in Endinspiration als auch Endexpiration führen zu einem geringeren pulmonalen Blutfluss (18).

1.2.3 Einfluss durch maschinelle Beatmung

Im Gegensatz zur physiologischen Atmung erfolgt die maschinelle Beatmung nicht durch Unterdruck, sondern üblicherweise mit Hilfe eines Überdrucks. Dadurch entsteht ein kontinuierlich positiver intrathorakaler Druck, der zu einer Abnahme des venösen Rückflusses zum Herzen, einer Reduktion des rechtsventrikulären Schlagvolumens und einer Verminderung des pulmonalen Flusses führt. Daraus resultiert eine verringerte Vorlast, die nach dem Frank-Starling-Mechanismus, zu einer Abnahme des Schlagvolumens und folglich zur Reduktion des Herzzeitvolumens führt (19).

So wird deutlich, dass durch externe Beatmung die physiologische Hämodynamik des Herz-Kreislaufs beeinflusst wird. Zur Beurteilung der kontinuierlichen Analyse der Herz-Lungen-Interaktion bei kardial vorerkrankten Patienten bietet die kardiale MRT vielseitige Möglichkeiten. Dabei ist es wünschenswert, den Einfluss von Atemanhaltenmanövern oder externer Beatmung, wie sie im Rahmen einer Allgemeinanästhesie notwendig wird, zu minimieren und somit pathophysiologische Veränderungen der Herz-Lungen-Interaktion bei spontaner Atmung analysieren zu können.

Diese Fragestellung bildet die Grundlage eines breiten Forschungsgebiets. Dabei hat sich der Stand der Technik im Bereich der kardialen MRT und der gleichzeitigen quantitativen Atemmessung stark weiterentwickelt.

1.3 Stand der Forschung

1.3.1 Physikalische Grundlagen des MRTs

Die Magnetresonanztomographie beruht auf der Eigenschaft von Protonen, sich aufgrund ihres Eigendrehimpulses (Spin) und des daraus resultierenden magnetischen Dipolmoments in einem Magnetfeld in einem Gleichgewichtszustand (Equilibrium) auszurichten. Durch einen hochfrequenten Radiofrequenzimpuls (RF-Impuls) können diese Protonen angeregt und aus dem Gleichgewicht gekippt werden. Nach der Anregung kehren die Spins durch Relaxation zurück in das Equilibrium (**Abbildung 1a**) (20).

Im Tomographen besteht zunächst ein Magnetfeld B_0 , welches kontinuierlich auf den Körper des Menschen wirkt. Dadurch richten sich die magnetischen Dipolmomente der Protonen überwiegend parallel oder antiparallel zum Magnetfeld aus und es entsteht eine longitudinale Gesamtmagnetisierung. (**Abbildung 1a**). Mit Hilfe eines RF-Impulses, eines kurzzeitigen oszillierenden Feldes B_1 , werden die Spins aus dem Gleichgewicht in die Transversalebene gekippt und führen eine Präzessionsbewegung um die Achse des Magnetfeldes B_0 aus. Dabei entsteht eine transversale Magnetisierung. Das Feld wird mit der Larmorfrequenz appliziert, der Resonanzfrequenz bei der Protonen Energie absorbieren. Die präzedierenden Spins induzieren in einer externen Spule eine Wechselspannung. Dieses Signal wird als Echo bezeichnet und bildet die Grundlage der Bildinformation. Die Spins kehren nach Beenden des RF-Impulses in ihr Gleichgewicht entlang B_0 zurück. Die Zeit zwischen dem RF-Impuls bis zur Registrierung des Signals wird Echozeit (TE) genannt. Die Dauer zum nächsten ausführbaren RF-Impuls wird als Repetitionszeit (TR) definiert. Beide Zeiten sind unter anderem abhängig von der Relaxation des Protons zurück in sein Equilibrium, beschrieben durch die Zeit T_1 , als auch der Abnahme der transversalen Magnetisierung, beschrieben durch die Zeit T_2 (20).

Das Echo wird in einem abstrakten Raum hinterlegt. Die räumliche Lokalisation erfolgt in der Regel mittels ortsabhängiger Gradientenfelder in drei Dimensionen.

In der z-Richtung werden Schichten definiert, für die jeweils dann ein zweidimensionaler k-Raum angelegt wird. Die zweidimensionale räumliche Zuordnung in x- und y- Richtung innerhalb jeder Schicht erfolgt über eine frequenz- und phasenabhängige Kodierung der Spins (**Abbildung 1a**). So wird der gesamte dreidimensionale Raum schrittweise ausgefüllt. Zur eigentlichen MR-Bild Rekonstruktion wird der k-Raum anschließend anhand unterschiedlicher Methoden abgetastet und üblicherweise durch eine schnelle Fouriertransformation abgebildet (**Abbildung 1b**). Gängige Abtastmethoden sind unter anderem die kartesische, spiralförmige oder die radiale Abtastung (20,21).

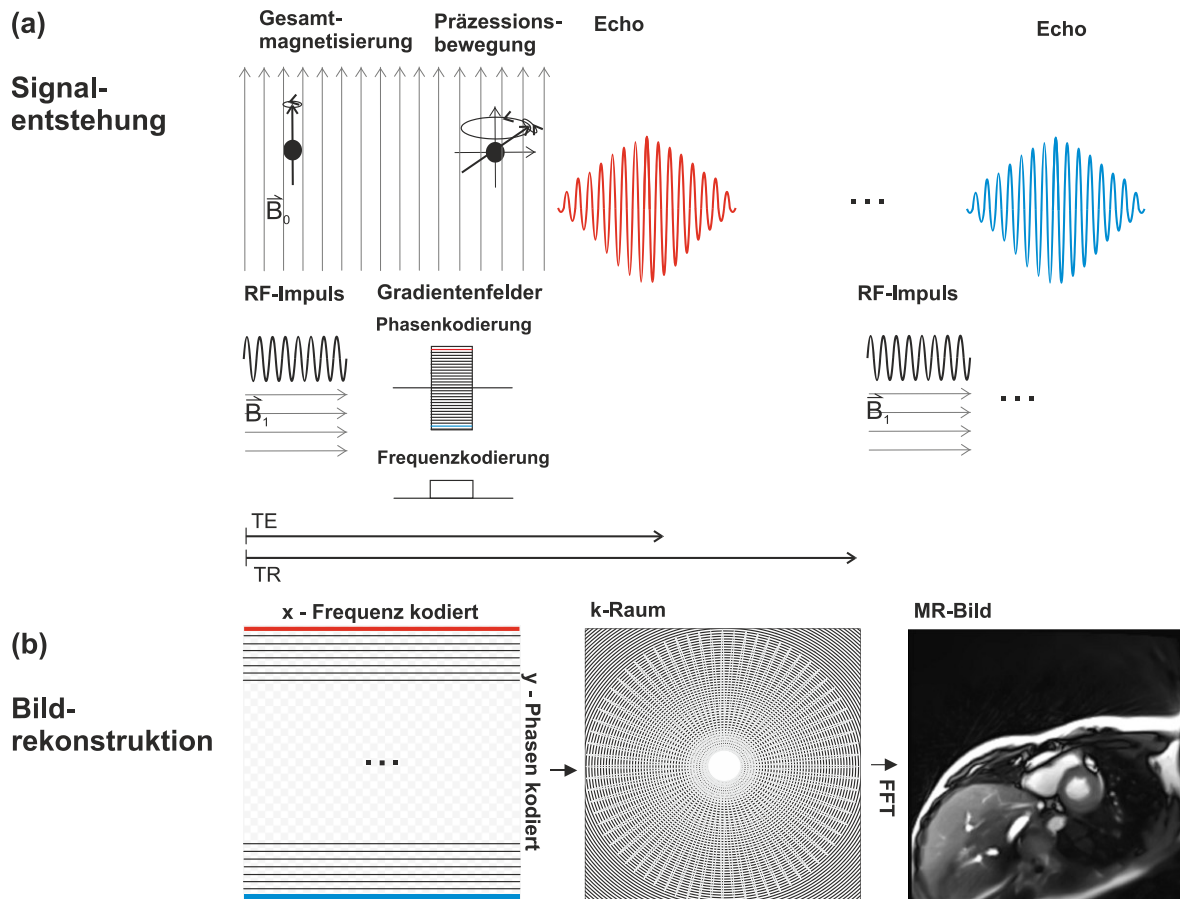


Abb. 1: Grundprinzipien MRT. Schematische Darstellung der Funktionsweise eines MRTs.

(a) Signalentstehung: Protonen werden durch das longitudinale Magnetfeld B_0 entlang dessen Achse parallel oder antiparallel ausgerichtet und tragen zu einer Gesamtmagnetisierung bei. Nach Applikation eines radiofrequenten (RF)-Impulses und ortsabhängiger Gradientenfelder werden die Spins der Protonen aus ihrer Position gekippt und führen eine Präzessionsbewegung durch. Dadurch wird ein Echo in der Empfängerspule induziert. Die Zeit zwischen RF-Impuls und Registrierung des Echos wird als Echozeit (TE) definiert. Die Zeit bis zum nächsten RF-Impuls wird durch die Repetitionszeit (TR) beschrieben (20).

(b) Bildrekonstruktion: Jedes Echo wird mittels Gradientenfeldern ortsabhängig nach Frequenz (x-Richtung) und Phase (y-Richtung) kodiert und in einem k-Raum abgelegt. Durch eine Fouriertransformation (FFT) wird das MR-Bild rekonstruiert (20). (Eigene Grafik mit modifizierten Abschnitten aus Abbildung 2 und 12 von Ridgway JP et al. (20))

1.3.2 Konventionelle kardiale MRT

Die Aufnahme bei einer kardialen MRT muss einem hohen Qualitätsstandard entsprechen. Innerhalb weniger Millisekunden muss ein Bild aufgenommen werden (20). Dafür ist ein hoher technischer Stand notwendig und verschiedene Hürden müssen überwunden werden. Die Bildakquise und die Wahl der kardialen MR-Sequenzen werden im Wesentlichen davon bestimmt, welche anatomischen Areale und welche kardiale Funktionalität dargestellt werden sollen (22).

Über den kardialen Zyklus können einzelne statische Bilder oder mehrere Bilder (*Cine imaging*) akquiriert werden. Während statische Bilder üblicherweise mittels Spin-Echo Techniken aufgenommen werden, werden bewegte (*Cine-*) Bilder überwiegend mittels Gradienten-Echo Techniken akquiriert. Im Wesentlichen unterscheiden sich beide Techniken in der Art der Echoerzeugung und der Kompensationsmechanismen von Dephasierungseffekten. In der Spin-Echo Technik wird ein Refokussierungsimpuls genutzt, um eine Dephasierung durch Spin-Spin Interaktionen oder B_0 -Inhomogenitäten zu kompensieren. Bei Gradienten-Echo Techniken fehlt dieser Impuls. Stattdessen wird das Echo durch Umkehrung des angelegten Gradientenfeldes erzeugt, wodurch die Spins teilweise rephasiert werden. Da in diesem Fall eine Dephasierung durch Spin-Spin Interaktionen und B_0 -Inhomogenitäten nicht vollständig kompensiert wird, führt dies zu einer schnelleren Signalabnahme. Dies ermöglicht eine kürzere TR und damit eine schnellere Füllung des k-Raums, wie sie für *Cine*-Bilder benötigt wird (20).

Statische Bilder zeigen hohe Kontrastverhältnisse im Bild und werden daher vorwiegend zur Beurteilung morphologischer Strukturen, wie des Herzgewebes und dessen Vitalität eingesetzt. Beim *Cine imaging* werden einzelne Bilder der Herzphasenabschnitte zusammengesetzt und stellen einen kardialen Zyklus in Bewegung dar. Das *Cine imaging* im Bereich der kardialen MRT ist der Goldstandard zur Evaluierung der kardialen Funktionsdiagnostik, wie beispielsweise der Analyse von Herzwandbewegung sowie von ventrikulären systolischen oder diastolischen Volumina (20,21).

1.3.3 Gating-Techniken

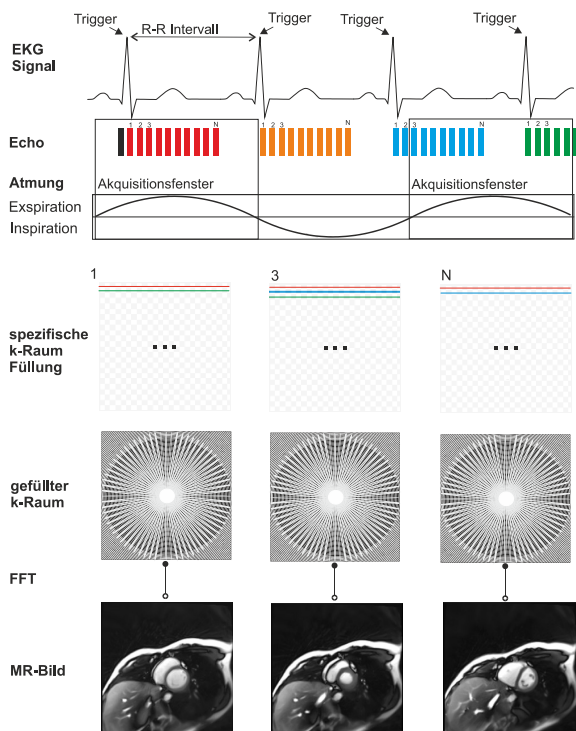
Bewegungen im thorakalen Raum aufgrund der Atmung, der Herzmechanik oder des Blutflusses stellen große Herausforderungen bei der Bildakquise dar. In rekonstruierten Bildern können Bewegungen zu Artefakten führen, die durch einen verzerrten k-Raum entstehen. Daher wurden Techniken eingeführt, die zum einen Bewegungen minimieren und zum anderen Aufnahmen ermöglichen, die wenig Zeit beanspruchen (21).

Standardmäßig werden kardiale MRT-Aufnahmen zum Herzschlag synchronisiert. Dafür werden die ausgeführten MR-Sequenzen an das R-R Intervall eines Herzzyklus angepasst. Die Anpassung an ein vorbestimmtes Intervall, auch *Gating* genannt, erfolgt entweder prospektiv oder retrospektiv und beim kardialen *Gating* mittels EKG.

Beim prospektiven EKG-*Gating* wird die R-Zacke detektiert und abhängig davon ein Trigger zum Start der Bildakquise versendet. Dabei umfasst das Aufnahmezeitfenster jeweils ca. 80 - 90% des kardialen Zyklus. Der enddiastolische Abschnitt wird dabei nicht vollständig erfasst. Jeder Herzphase wird ein korrespondierender k-Raum zugeteilt und über mehrere Herzzyklen gefüllt, sodass ein Bild einer bestimmten Herzphase stets eine Zusammensetzung aus mehreren Herzzyklen widerspiegelt (20) (**Abbildung 2a**). Der Nachteil bei dieser *Gating*-Methode ist, dass die globale ventrikuläre Ejektionsfraktion unterschätzt wird (23).

Beim retrospektiven EKG-*Gating* erfolgt eine kontinuierliche Aufnahme mit einer festen Anzahl an Bildern pro Zyklus. Die Dauer des R-R Intervalls wird über alle R-R Intervalle gemittelt. So werden die Bildsignale zeitlich den jeweiligen R-R Intervallen zugeordnet und nach Herzphasen zunächst in k-Räume Fourier-transformiert. Auch in diesem Fall entsteht ein Bild einer Herzphase aus unterschiedlichen Herzzyklen (**Abbildung 2b**) (20). Der Vorteil gegenüber dem prospektiven EKG-*Gating* ist, dass das gesamte R-R Intervall abgebildet wird (23).

(a) Prospektives Gating



(b) Retrospektives Gating

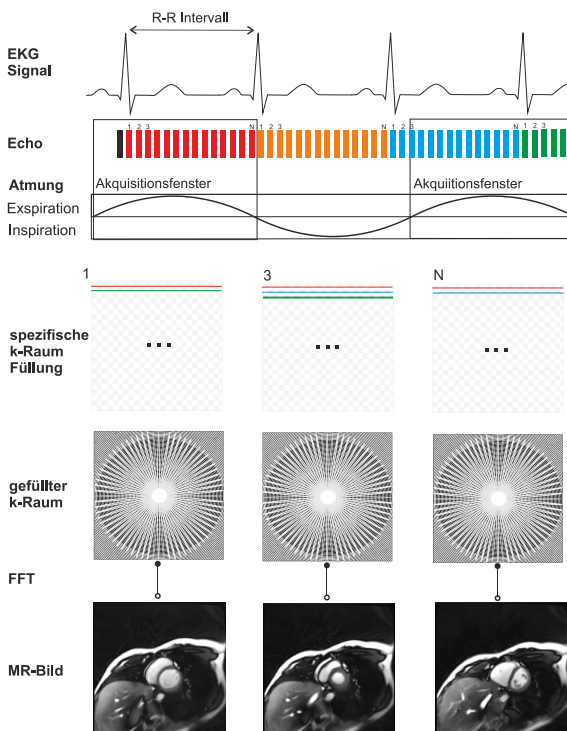


Abb. 2: Duales Gating. EKG-Gating und respiratorisches-Gating.

(a) Prospektives *EKG-Gating*: Zu Beginn jeder R-Zacke wird die Bildakquise gestartet und EKG-abhängige Echos werden aufgezeichnet. Pro R-R Intervall werden etwa 80-90% des Herzzyklus erfasst. Je Herzphase wird ein spezifischer k-Raum über mehrere Herzzyklen gefüllt.
(b) Retrospektives *EKG-Gating*. Eine durchgängige Aufnahme von einer festen Anzahl an Bildern findet statt. Die Zuteilung der Bildsignale zu den korrespondierenden Herzphasen erfolgt retrospektiv, sodass verschiedene k-Räume gefüllt werden. Die Füllung erfolgt über mehrere Herzzyklen.

Beim *dualen Gating*, dem kombinierten EKG- und respiratorischem *Gating*, wird die Bildakquise sowohl durch das EKG als auch durch die Atmung (Expiration im Vergleich zur Inspiration) begrenzt. Dies wird durch das Akquisitionsfenster verdeutlicht. Nach der k-Raum Füllung findet mit schneller Fouriertransformation (FFT) die Rekonstruktion in die einzelnen MR-Bilder statt. (Eigene Grafik mit modifizierten Anteilen aus Abbildung 25 von Ridgway JP et al. und Abbildung 31 von Biglands JD et al. (20,24)).

Zusätzliche Bewegungsartefakte können durch die Atemmechanik ausgelöst werden. In der Bildrekonstruktion kann dies zu respiratorisch bedingtem *Blurring* oder *Ghosting* führen. Unter *Ghosting* versteht man wiederholt auftretende Spiegelung des aufzunehmenden Objektes im rekonstruierten Bild. Beim *Blurring* werden Kanten unscharf dargestellt, wodurch sich ein verwaschenes Bild präsentiert (25).

Konventionelle kardiale Aufnahmen werden daher häufig mit Protokollen gefahren, in denen Atemanhaltenmanöver durchgeführt werden (21) oder respiratorische Bewegung überwacht wird (24, 26). Eine Korrektur der Bewegungsartefakte kann durch ein respiratorisches *Gating* erfolgen, wodurch die Selektion der MR-Bilder eingeschränkt wird (27,28,29).

Beim Atemanhaltenmanöver soll mehrere Sekunden lang die Luft bspw. in Endexpiration angehalten werden, um in diesem Zeitfenster die MRT-Aufnahme durchzuführen (8,29). Diese Technik ist weit verbreitet, jedoch mit einigen Nachteilen verbunden. Sie erfordert eine hohe Patientencompliance (28). Weiterhin können Atemanhaltenmanöver die Zwerchfellposition beeinflussen, vor allem wenn bei verschiedenen endexpiratorischen Phasen die Zwerchfellposition unterschiedlich ist (21,26,31,32). Letztlich führen Atemanhaltenmanöver häufig zu einer schlechteren örtlichen Auflösung, sodass die Kontrastierung des Herzens beeinträchtigt werden kann (33).

Um Aufnahmen während freier Atmung zu gewährleisten, werden Verfahren eingesetzt, die die Atembewegung mit Hilfe externer *Devices* erfassen (z.B. MR-kompatible Atemgurte) oder die MRT-Aufnahme auf bestimmte atemabhängige Intervalle prospektiv oder retrospektiv begrenzen (26).

Durch die Nutzung externer *Devices* kann hardwarebasiert die Atmung detektiert werden. Ein gängiges Verfahren ist der Einsatz eines MR-kompatiblen Atemgurts, der abdominelle oder thorakale Bewegungen erfasst und damit gut mit der Zwerchfellbewegung oder der thorakalen Ausdehnung korreliert (26,28,29). Prospektiv kann anhand von Kompression und Dekompression des Gurts bei der Atmung ein Signal erzeugt und direkt an den MR-*Scanner* übermittelt werden (33).

Ein weiterer Ansatz zur Bildstabilisierung und der Bewegungskorrektur wird durch Navigator-Echos ermöglicht.

Beim prospektiven respiratorischen-*Gating* wird in einer vorab aufgenommenen Bildsequenz ein anatomisches Areal für den Navigator definiert (**Abbildung 3a**). Der Zwerchfellübergang mit seinem starken Kontrast zur Lunge wird häufig als Navigator-Areal verwendet. Während einer bestimmten respiratorischen Phase werden Bilder aufgenommen (**Abbildung 3a**). Dafür wird das Gewebe im Navigator-Areal mit einem RF-Impuls angeregt, wodurch ein Echo emittiert wird. Dieses Signal stellt die Bewegung des Diaphragmas dar. Fällt jenes Signal in ein definiertes Akzeptanzareal, welches beispielsweise den Bereich der Endexpiration darstellt, werden die Bilddaten genutzt. Außerhalb dieses Areals werden die Bilder verworfen (**Abbildung 3b**) (24, 28, 34). Ein wesentlicher Nachteil dieses Verfahrens ist die ungewisse zeitliche Länge der Bildakquise, wenn sich die Akzeptanzrate während der Untersuchung verändert (28, 34).

Beim retrospektiven respiratorischen *Gating* werden MR-Bilder kontinuierlich aufgenommen und anschließend anhand von Atemmodellen selektiert, sodass nur Bilder innerhalb definierter Atemintervalle genutzt werden. Dieses Verfahren kann jedoch zu Überabtastung führen, da hier mehrere Atemzyklen benötigt werden, um den k-Raum ausreichend auszufüllen (35).

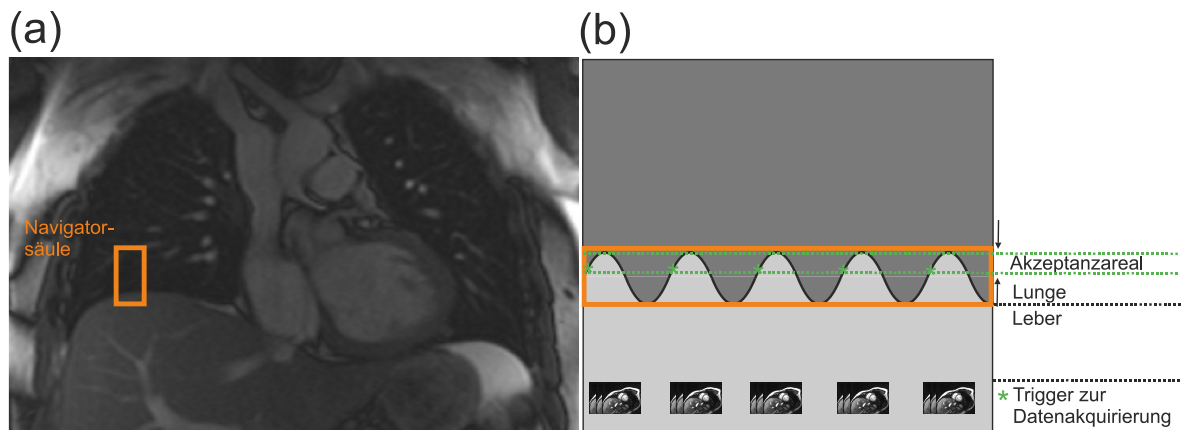


Abb. 3: Prospektives respiratorisches-Gating mit Navigator-Echos.

(a) Koronare Sicht auf den Thorax. Eine Navigatorsäule befindet sich auf Höhe der Lungen-Leber-Grenze.

(b) Im Bereich der Navigatorsäule wird das Gewebe mittel Radiofrequenzimpulsen (RF-Impuls) angeregt, um Echos zu emittieren, die den Intensitätswechsel der Kante zwischen Lunge und Zwerchfell/Leber als respiratorische Linie abbilden. Fällt diese Linie in den Bereich des Akzeptanzareals, in diesem Beispiel während der Expiration, werden Daten akquiriert, sofern keine weiteren beschränkenden Methoden die Bildaufnahme verhindern.

(Eigene Grafik mit modifizierten Anteilen aus Abbildung 31 von Biglands JD et al. (24)).

1.3.4 Fast-Sequenzen

Um Verzögerungen durch EKG- und respiratorisches-Gating zu reduzieren, wurden sogenannte *Fast Acquisition Steady State (FAST)* Sequenzen eingeführt. *FAST* Sequenzen sind schnelle Gradienten basierte Sequenzen (20). Im Rahmen dieser Arbeit werden zwei vorgestellt: die *fast low angle shot (FLASH)* und die *balancierte Steady State Free Precession (b-SSFP)* Sequenz.

Bei der *FLASH*-Sequenz wird der Kernspin in einem kleineren Winkel als herkömmlich aus dem Equilibrium gedreht. Die longitudinale Magnetisierung wird im Großteil aufrechterhalten. Es werden in kurzer Zeit viele RF-Impulse emittiert, woraus nach einer kurzen Zeit ein dynamisches Equilibrium (*Steady-State*) der Kernspins entsteht. Die Abfolge von einem Bild zum nächsten wird damit nicht durch die komplette Relaxation des Spins limitiert. Die Repetitionszeit wird dadurch deutlich verkürzt und beträgt 10-20 ms. Letztlich

führt dies zu einer schnelleren Füllung des k-Raums und damit zu einer geringeren Akquisitionsszeit (36).

Die technischen Grundlagen der *b-SSFP* Aufnahme ist vergleichbar mit der *FLASH*-Sequenz, ergänzt durch balancierte Gradientenfelder. Diese bewirken, dass die vollständige Relaxation am Ende einer Aufnahme mit einem zusätzlichen Gradientenfeld aufrecht gehalten wird, wodurch Bildeigenschaften in den nächsten k-Raum übertragen werden. Durch das gute Signal-Rausch Verhältnis werden Aufnahmen, bei denen ein hoher Anteil an bewegten Bildern vorliegt, wie beispielsweise Flussprofile, gut dargestellt. Daher wird diese Technik aufgrund ihres guten Blut-zu-Myokard-Kontrasts in der kardiovaskulären Bildgebung häufig eingesetzt (37,38).

1.3.5 Echtzeit-MRT

Die Echtzeit-MRT revolutionierte mit ihrer Einführung die Aufnahme und Analyse von Bildern, die in nahezu Echtzeit rekonstruiert werden. Zeitliche Auflösung von bis zu 20 ms pro Bild und örtliche Auflösung von 1,5-2 mm werden beschrieben (37,39). Technisch zeichnet sich das Echtzeit-MRT durch schnelle Gradienten-Echo-Sequenzen, einer effizienten k-Raum Abtastung und einer nicht linearen inversen Rekonstruktion aus. Wurde zuvor auf sequenzielle kartesische k-Raum Abtastung zurückgegriffen, werden nun radiale Abtastmethoden verwendet.

Der k-Raum ist nicht linear aufgebaut (37). Mittig befinden sich Daten mit niedrigerer örtlicher Auflösung als am Rand des k-Raums. Dies führt dazu, dass Kontraste und Bewegungen von Objekten im Bereich des mittleren k-Raums festgehalten werden, während die äußeren Bereiche des k-Raums die Kanten der MR-Bilder abbilden (40).

Bei bewegten Bildern ist es von Vorteil, den k-Raum nicht linear abzutasten. Bei der radialen Abtastung wird die Mitte des k-Raums mehrfach abgetastet, was zu einer robusteren Darstellung von Bewegung und Kontrasten führt. Letztlich wird technisch eine nicht lineare Rekonstruktion verwendet, die verschiedene Algorithmen und Filter nutzt, um Aliasing zu vermeiden und die Bildqualität zu verbessern (**Abbildung 4**) (37,39).

Pro Schicht

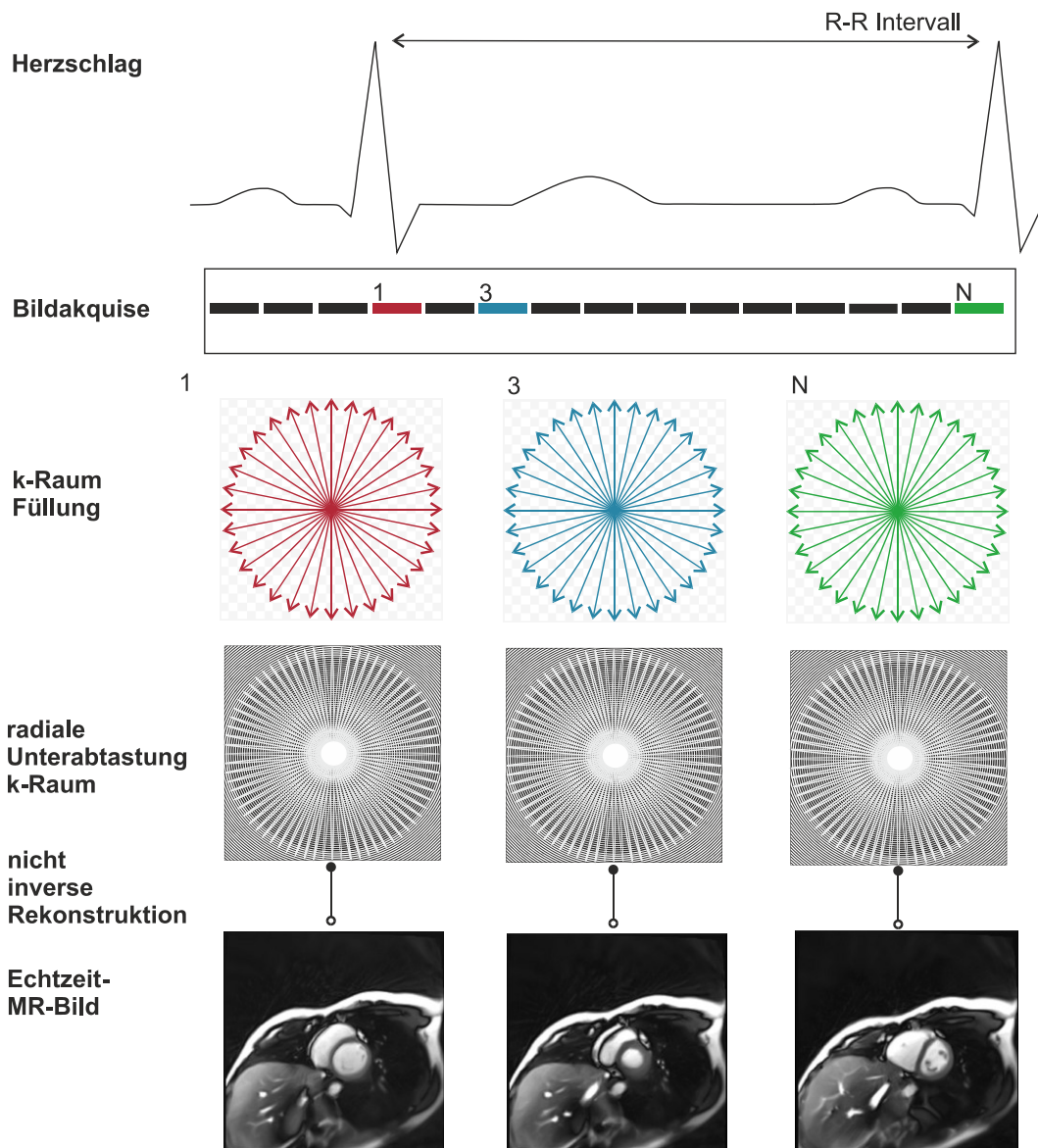
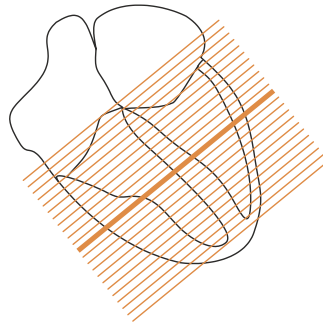


Abb. 4: Echtzeit-MRT. Die Echtzeit-MRT-Aufnahme erfolgt kontinuierlich unabhängig von respiratorischen- oder EKG-Gating. Pro Schicht des Herzens sowie abhängig von der Akquisitionszeit und der Länge des R-R Intervalls wird für jeden Herzschlag eine unterschiedliche Anzahl an Bildern aufgenommen. Der k-Raum wird nicht linear aufgebaut. Mittig befinden sich Daten niedrigerer örtlicher Auflösung als am Rand. Die Bilderstellung basiert bspw. auf radialer Unterabtastung des k-Raums, mehrfacher Abtastung von Stellen mit niedriger Resolution und der nicht linearen inversen Rekonstruktion. (Eigene Grafik)

Die Anwendung der Echtzeit-MRT findet bereits in mehreren medizinischen Bereichen statt. Bislang zeigt sich eine überwiegende Anwendung bei der Untersuchung von muskuloskelettalen, oropharyngealen und kardiovaskulären Mechanismen. In der muskuloskelettalen Forschung wird vor allem die Bewegung von Gelenken und die Interaktion mit Bändern, Muskeln und Knochen untersucht. Es wurden Studien zur Untersuchung der Knie, der Wirbelsäule, der Füße und der Hände durchgeführt. In der Forschung von oropharyngealen Mechanismen wurden beispielsweise Studien zur Sprachproduktion und zum Schluckakt durchgeführt (37).

Im Bereich der kardiovaskulären Forschung wird durch die Echtzeit-MRT die Notwendigkeit eines kardialen oder respiratorischen *Gatings* beseitigt. Hierdurch wird eine Bildgebung bei Patienten mit kardialen Rhythmusstörungen möglich (37). Der Einsatz von Atemanhaltenmanövern kann reduziert und die Indikation für eine kindlichen Sedierung bei einer MRT-Aufnahme kann relativiert werden (8, 37). Die quantitative Analyse der Herz-Lungen-Interaktion gewann durch die kardiale Echtzeit-MRT neue Relevanz. Röwer et al. konnten in Echtzeit-MRT-Aufnahmen in freier Atmung respiratorische Einflüsse auf das Herz quantifizieren (41). Des Weiteren zeigte der Vergleich zwischen kardialer Echtzeit-MRT und konventionellen kardialen *Cine*-Aufnahmen unter Atemanhaltenmanövern gute Übereinstimmungen hinsichtlich linksventrikulärer Funktionsparameter. Gleichzeitig konnte ein signifikanter Unterschied in der quantitativen Auswertung der atemabhängigen, rechtsventrikulären Funktionsparameter nachgewiesen werden. Dies verdeutlicht, dass unter Atemanhaltenmanövern in Endexpiration die physiologische Herz-Lungen-Interaktion bei freier Atmung nicht ausreichend präzise abgebildet werden kann (8). Erste Forschungsergebnisse mit kardialer Echtzeit-MRT bei Kindern mit rechtsventrikulärer Volumenbelastung konnten sich die atemabhängigen Veränderungen zunutze machen und eine Beeinträchtigung im Frank-Starling Mechanismus nachweisen (42).

1.3.6 Akquise von Ateminformation

Ateminformationen können aus unterschiedlichen Medien akquiriert werden. Zur Beurteilung der Herz-Lungen-Interaktion sind das Lungenvolumen und der intrathorakale Druck entscheidende Einflussgrößen (13). Zur quantitativen Darstellung der Lungenfunktion ist die Spirometrie weiterhin der Goldstandard (43).

Der intrathorakale Druck kann bei ruhiger Atmung annähernd durch den intrapleurale Druck beschrieben werden. Dieser lässt sich durch die Messung des ösophagealen Drucks mittels Ballonkatheter annähern (44). Die Invasivität und die eingeschränkte MR-Tauglichkeit dieses Messverfahrens verdeutlichen jedoch die Grenze bei der Untersuchung von der Herz-Lungen-Interaktion im Rahmen der Magnetresonanztomographie (13).

Daher wird im Rahmen dieser Arbeit der Fokus auf die Ermittlung der Ateminformation aus MR-Bildern und der Lungenfunktionsanalyse mittels Spirometrie gelegt.

1.3.7 Ateminformation aus MR-Bildern

Der Stand der Forschung zur Bestimmung von Ateminformationen aus thorakalen MR-Bildern verfolgt unterschiedliche Ansätze. So werden qualitative als auch quantitative Methoden zur Erfassung von Lungenfunktionsparametern untersucht.

Im Vordergrund und gleichzeitig als limitierender Faktor steht die Bildqualität der Lunge. Aufgrund des hohen Sauerstoffgehalts und der relativ niedrigen Gewebedichte und damit Protonendichte der Lunge ist die erreichbare MR-Signalintensität begrenzt. Des Weiteren ist das Lungengewebe stark inhomogen, wodurch die Signaldauer kurz und anfällig für die Bildung von Artefakten ist. Die Lunge erscheint im MR-Bild als signalschwache Entität (33,45).

Trotzdem wurden bereits mehrere Ansätze zur Detektion von Ateminformationen aus MR-Bildern erforscht.

2005 demonstrierten Voorhees et al. die Möglichkeit, aus MR-Bildern in koronarer Ansicht durch Bildsegmentierung und durch Kartierung von Vektorgeschwindigkeiten den regionalen Atemfluss und das daraus resultierende Lungenvolumen zu bestimmen (46).

Boucneau et al. stellten 2020 eine Methode vor, bei der ebenfalls regionale Veränderungen im Lungengewebe anhand der Deformation von MR-Bildern der Lunge in Form von qualitativen regionalen Fluss-Volumen-Kurven erfasst wurden. Dabei wurden sagittale und koronare Ansichten der Lunge genutzt und die Anisotropie der Lungendeformation wurde in der Berechnung inkludiert (47).

1.3.8 MR-kompatible Spirometrie

Die kombinierte Anwendung von Spirometer und MRT wurde in den 2000er-Jahren erforscht.

2004 zeigten Bankier et al., dass die Signalintensität des Lungenparenchyms mit dem Lungenvolumen einer MR-kompatible Spirometrie (MRkS) korreliert. Dafür untersuchten sie zehn Probanden und führten neben koronaren Lungen MRT-Aufnahmen MRkS Messungen durch (48).

2007 stellten Eichinger et al. eine Methode vor, mit der quantitative Lungenfunktionsparameter in einer liegenden Position mit Hilfe eines MRkS gleichzeitig zur MRT-Aufnahme aufgezeichnet werden konnten. Sie zeigten zudem, dass eine Korrelation zwischen den gemessenen MRkS-Daten und den bildmorphologischen Größenänderungen der Lunge besteht (49).

Diese Techniken eröffnen Möglichkeiten zur Beurteilung qualitativer und quantitativer respiratorischer Parameter. Mit der Echtzeit-MRT werden die technischen bildmorphologischen Rahmenbedingungen verbessert, was eine hervorragende Basis zur Weiterentwicklung dieser Methoden darstellt.

1.4 Maschinelles Lernen

1.4.1 Definition und Arten

Der Begriff maschinelles Lernen (ML) wurde von Arthur Samuel geprägt, als dieser einem Computer das Spiel Dame lehrte. Er erstellte eine Methode, mit der ein Computer die Fähigkeit des Lernens durch Erfahrung aneignet, ohne explizit programmiert zu werden (50). Neuere Definitionen erweitern diesen Ansatz um die Erkennung von „fundamentalen statistisch-rechnerisch-informatisch-theoretischen Gesetzen, die die Basis aller lernbaren Systeme vereinbart“ (51).

In der praktischen Umsetzung wird ML als Teilbereich der künstlichen Intelligenz betrachtet. Dabei wird ein Modell erstellt, das anhand einer definierten Menge an Trainingsdaten gefittet wird, um Muster und grundlegende Gesetzmäßigkeiten zu erlernen. Das trainierte Modell soll anschließend eine Generalisierung von Mustern und grundlegenden Gesetzen auf einer bisher ungesehenen Datenmenge, dem Testdatensatz, ermöglichen (52) **(Abbildung 5)**.

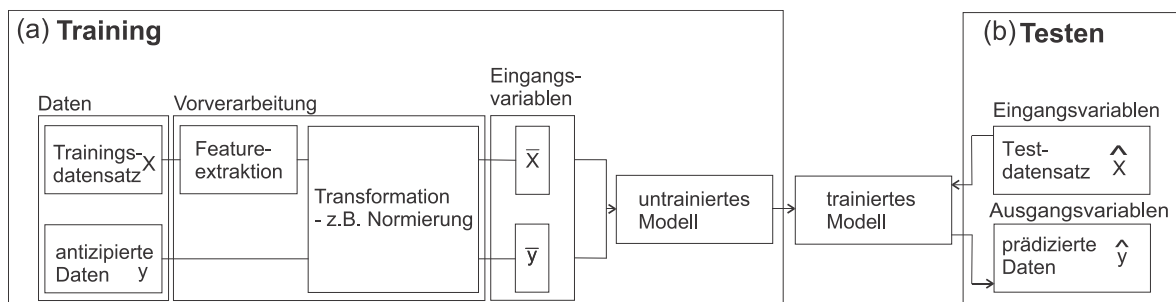


Abb. 5: Grundprinzipien des maschinellen Lernens (ML).

(a) Training: Die Daten, die zum Training verwendet werden, bestehen aus dem Trainingsdatensatz X und einem Datensatz y , der die antizipierten Daten enthält. Im Weiteren wird häufig eine Präprozessierung der Rohdaten durchgeführt, um Fehler zu bereinigen, *Features* zu extrahieren oder notwendige Transformationen der Daten, bspw. eine Normierung, durchzuführen. Die vorbereiteten Eingangsvariablen \bar{X} und \bar{y} werden dann dem untrainierten Modell übergeben, um dieses zu fitten.

(b) Testen: Das trainierte Modell soll auf einem unbekanntem Testdatensatz \hat{X} generalisieren, um die Zielwerte \hat{y} vorherzusagen. (Eigene Grafik)

Maschinelles Lernen lässt sich in drei Hauptkategorien unterteilen: *supervised learning*, *unsupervised learning* und *reinforcement learning* (51).

Beim "*supervised learning*" (überwachtes Lernen) werden dem Modell zum Training neben den zu untersuchenden Daten X auch die antizipierten Daten y übergeben. Das trainierte Modell soll anschließend auf ungesehenen Daten \hat{X} eine Prädiktion zu den dazugehörigen

Daten \hat{y} treffen (**Abbildung 5**). Die Architektur solcher Modelle ist ebenfalls sehr vielseitig. Hierzu gehören unter anderem Entscheidungsbäume, Zufallswälder, lineare Regression, logistische Regression, oder neuronale Netzwerke (51).

Beim „*unsupervised learning*“ (unüberwachten Lernen) wird dem Modell zum Training lediglich der zu untersuchende Datensatz X übergeben. Die vorherzusagenden Werte y sind zuvor unbekannt und sollen durch das Modell selbst ermittelt werden (51).

Beim „*reinforcement learning*“ wird ein Modell auf Basis eines Belohnungssystems trainiert. Das Modell kennt die antizipierten Ergebnisse zu Beginn nicht. Stattdessen erhält das Modell ein Feedback, ob ein erzieltetes Ergebnis den Zielvorgaben entspricht oder nicht. Bei einem korrekten Ergebnis wird eine Belohnung abgegeben, bei unerwünschtem Ergebnis wird keine abgegeben. Im letzten Fall wird dies vom Modell wahrgenommen und das Modell wird neu trainiert (51).

Im Rahmen dieser Arbeit wurden Modelle des überwachten Lernens mittels Linearer Regression verwendet, weswegen ausschließlich diese weiter erläutert wird.

Um die Validierung von ML-Algorithmen weiter zu verstehen, sind die Begriffe ‚Überanpassung‘ und ‚Unteranpassung‘ von zentraler Bedeutung. Unteranpassung oder *Underfitting* tritt auf, wenn ein Modell unzureichend trainiert wurde. Die Variabilität des Trainingsdatensatzes wurde nicht ausreichend erkannt, was zu einer schlechten prädiktiven Güte auf dem Testdatensatz führt (53) (**Abbildung 6a**). Überanpassung oder *Overfitting* beschreibt hingegen eine Vorhersage, die zu stark auf den trainierten Datensatz abgestimmt ist. Eine Verallgemeinerung auf einen ungesehenen Datensatz kann nur unzureichend und fehlerbehaftet ausgeführt werden (53) (**Abbildung 6b**). Um ein Gleichgewicht zwischen Überanpassung und Unteranpassung zu finden, ist ein robustes Training notwendig, bei dem sowohl das richtige Modell gewählt wird als auch geeignete Optimierungsverfahren im Training eingesetzt werden (53) (**Abbildung 6c**).

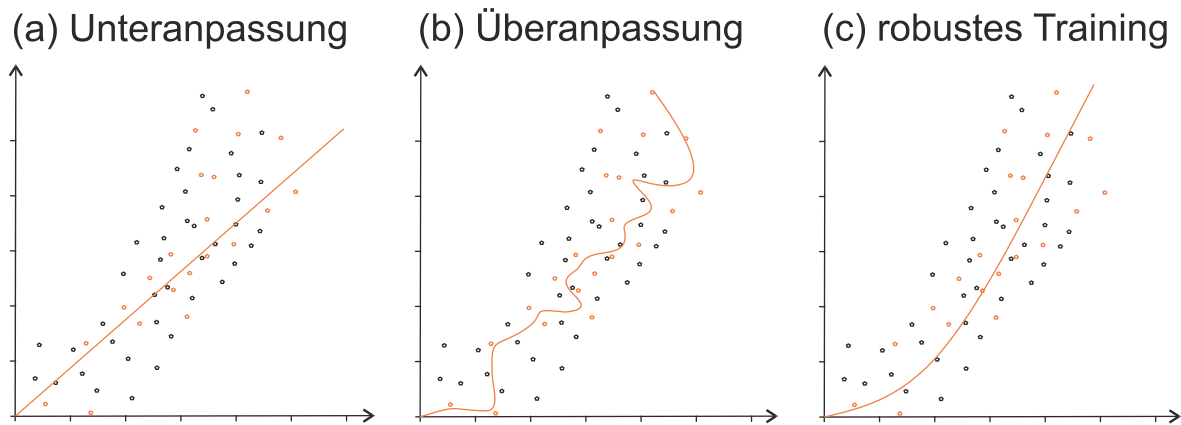


Abb. 6: Modellvalidierung. Die Validierung eines trainierten Modells hängt von (a) Unteranpassung, (b) Überanpassung oder (c) einem robusten Training ab. Die schwarzen Polygone visualisieren die Testdaten und die orangenen Polygone die Trainingsdaten. Die Kurven (orangene Linie) in (a)-(c) stellen die geschätzten Funktionen der trainierten Modelle dar. (a) Bei einer Unteranpassung findet eine unzureichende Anpassung an den Trainingsdatensatz statt. Die Generalisierung auf Testdatensätze ist nur begrenzt möglich. (b) Bei der Überanpassung findet eine zu starke Anpassung an den vorliegenden Trainingsdatensatz statt. Eine Generalisierung auf ungesehene Testdatensätze ist nur begrenzt möglich. (c) Beim robusten Training wird anhand der Trainingsdaten die unterliegende Funktion ausreichend gut geschätzt, sodass eine Generalisierung auf einem ungesehenen Testdatensatz mit geringem Fehler stattfinden kann (53). (Eigene Grafik mit modifizierten Anteilen aus Aliferis C et al. (54))

1.4.2 Lineare Regression und Ridge Regression

Die lineare Regression gehört zu einer Subkategorie des überwachten Lernens. Lineare Regressionsmodelle zeichnen sich unter anderem durch eine geringe Komplexität, hohe Interpretierbarkeit und häufig durch einen niedrigen Rechenaufwand aus (55). Bei der linearen Regression wird ein mathematisches Modell definiert, das zwischen einer unabhängigen Variable $X^{n \times m}$ und einer abhängigen Variable $y^{n \times 1}$ die vorhandene lineare Relation bestimmt. Die Relation folgt der mathematischen Funktion:

$$y = \beta X + \varepsilon$$

in der $\beta^{m \times 1}$ berechnete Gewichte des Modells und ε den nicht durch das Modell erklärbaren Fehler darstellen (56).

Die genannte mathematische Funktion wird nach den Gewichten β gelöst. Üblicherweise wird dafür die Methode der kleinsten Quadrate verwendet, indem die Summe der Quadratresiduen (SSE) minimiert wird (57):

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Bei komplexe Datensätzen mit einer geringen Menge an Daten zeigt sich häufig eine Multikollinearität, d.h. eine beinahe lineare Abhängigkeit zwischen den Variablen X , y oder β (58). Im Bereich der linearen Regression kann dies zu Überanpassung führen, insbesondere bei geringer Trainingsdatenmenge und hoher Varianz in der vorhergesagten Datenmenge (58).

Um diesem Problem entgegenzuwirken, wird Ridge Regression verwendet. Sie verbessert die Generalisierbarkeit auf den Testdaten durch Regularisierung der Gewichte β während des Trainings. Die Regularisierung erfolgt, indem ein Strafterm in die Minimierung der Summe der Quadratresiduen (SSE_{L2}) eingeführt wird:

$$SSE_{L2} = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^m \beta_j^2$$

Dieser Strafterm besteht aus der Summe der Gewichte β und einem vorab eingestellten Parameter (Hyperparameter) λ , welcher durch Kreuzvalidierung gefunden wird. Die Lösung des gesamten Systems gibt die Regressionsgerade an, bei der die Varianz der vorhergesagten Datenmenge minimiert ist (55, 59).

1.4.3 Maschinelles Lernen im medizinischen Arbeitsbereich

Maschinelles Lernen gehört zu einem rasant wachsenden Forschungsgebiet so auch in der medizinischen Forschung.

ML in der Medizin adressiert verschiedene Aspekte. Zum einen verbessert es in der Diagnostik Reliabilität, Performanz, Genauigkeit und Umsetzung und leitet Strategien zur Optimierung von Therapie ab. Zum anderen soll eine Verbesserung und Erleichterung im Bereich der Prognosestellung hervorgerufen werden (60). Darüber hinaus sollen ML-erzeugte automatisierte Prozesse eine Arbeitserleichterung schaffen. Indem beispielsweise Aufgaben übernommen werden, die von Medizinern nur eingeschränkt bewältigt werden können. Letztlich dient es als Instrument zur Verbesserung der Ausbildung künftiger Ärztinnen und Ärzte (60,61).

So vielseitig wie die Aufgabengebiete des ML sind, so breit ist auch sein Einsatzspektrum: nahezu alle Fachdisziplinen der Medizin profitieren davon, insbesondere die Radiologie, Pathologie, Onkologie oder Kardiologie. Typische Forschungsdatensätze stammen aus der medizinischen Bildgebung, der medizinischen Chemie, der Genexpressionsanalyse oder der Krebsforschung (62).

Ein Anwendungsbeispiel der Klinik und dem alltäglichen Gebrauch findet sich in der Auswertung von Arrhythmien im EKG. So gibt es bereits die Möglichkeit EKG-Ableitungen

über Smartwatches ML-basiert auszuwerten oder sowohl in implantierbaren Kardioverter-Defibrillatoren als auch in automatisierten externen Defibrillatoren lebensbedrohliche Arrhythmien zu erkennen und automatische Defibrillationen auszulösen (63).

Die Anwendung von ML in der Bildanalyse ist ebenfalls ein sehr großer Arbeitsbereich. Dabei werden ML-basierte Techniken der Segmentation, Techniken zur Extraktion von *Features* und Techniken zur Klassifizierung angewendet (62).

Auch in der kardiovaskulären Bildgebung, vorwiegend im Bereich der Echokardiographie, der Computertomographie und der Magnetresonanztomographie werden bereits ML-basierte Verfahren eingesetzt, z.B. zur Messung kardiovaskulärer Größen wie der Wandverdickung oder der volumetrischen Größen des linken Ventrikels (64).

Trotz dieser Fortschritte und weitreichenden Nutzung, gibt es einige limitierende Faktoren in der klinischen Umsetzung. Die Akquise großer Datenmenge zum Training ist zeitaufwendig. Um eine gewisse Güte in der Vorhersage zu erzielen, müssen Daten oft umfangreich vorverarbeitet werden. Darunter fällt die Minimierung von redundanter Information zur Besserung des Trainings. Im Weiteren müssen Datensätze von sensiblen personalisierten Patientendaten befreit und anonymisiert werden, damit ein adäquater Umgang in Hinblick des Datenschutzes erfolgt. Außerdem birgt die Übergabe von vorgegebenen Zielwerten im maschinellen Lernen Potenzial zum Erlernen von groben Fehlern, die beispielsweise einen diskriminierenden Charakter haben können. Wenn diese Vorhersagen zur Therapie oder zur Diagnose von Krankheiten verwendet werden, stellt sich als nächstes die Frage der Verantwortung. Wer wird bei inakkurater Therapie oder Diagnose verantwortlich gemacht, das trainierte Modell oder der Arzt, der das trainierte Modell genutzt hat (65 ,66)? Ein weiteres zentrales Problem ist das *Black Box* Prinzip. Im Allgemeinen wird der zugrunde liegende Algorithmus eines trainierten Modells nicht offengelegt, sodass die Nachvollziehbarkeit der Vorhersage teils nur stark eingeschränkt möglich ist (67). All diese Faktoren sind bei der Nutzung von maschinellem Lernen zu beachten. Dennoch zeigt die Forschung im Bereich des ML viele neue Möglichkeiten im medizinischen Umfeld, die den Arbeitsalltag der Klinik vereinfachen können.

1.5 Ziele dieser Arbeit

Die experimentelle Studie wurde zur Entwicklung und Validierung eines Modells durchgeführt, welches die Hypothesen testet, dass mittels maschinellen Lernens nach kurzer Trainingsphase mit kardialen Echtzeit-MR-Bildern kombiniert mit MR-kompatibler Spirometrie quantitative respiratorische Parameter aus ungesehenen kardialen Echtzeit-MR-Bildern

- 1.) einer mittventrikulären Schicht desselben Probanden der verbliebenen Sequenz vorhergesagt werden können.
- 2.) einer kompletten kardialen Kurzachsenstapel-Volumetrie desselben Probanden der verbliebenen Sequenz vorhergesagt werden können.
- 3.) einer kompletten kardialen Kurzachsenstapel-Volumetrie eines zuvor unbekanntem Probanden einer separaten Sequenz vorhergesagt werden können.
- 4.) einer kompletten kardialen Kurzachsenstapel-Volumetrie desselben Probanden einer zeitlich versetzt aufgenommenen Sequenz vorhergesagt werden können.

2 Material und Methoden

Die Studie wurde von der Ethikkommission der Heinrich-Heine-Universität Düsseldorf mit der Studiennummer 6176R im Jahr 2017 bewilligt. Anschließend wurden 2022 Daten von zehn freiwilligen (sechs weiblichen und vier männlichen) Probanden aufgenommen. Vor Beginn der Untersuchung fand eine Aufklärung statt, indem jeder Proband bezüglich der Funktionen der MRT, der bestehenden Risiken und der durchzuführenden Untersuchung eingewiesen wurde. Ein Informationsblatt wurde an jeden Probanden ausgehändigt. Kontraindikationen oder gesundheitliche Einschränkungen in Bezug auf die MRT wurden vor Beginn bei allen Probanden ausgeschlossen. Alle Probanden gaben ihre schriftliche Einwilligung in Form einer Einverständniserklärung zur Ausführung der Studie ab (68).

2.1 Datenakquise

Für jeden Probanden wurden Echtzeit-MR-Bilder und simultan MR-kompatible Spirometrie (MRkS) Daten akquiriert. Die Echtzeit-MR-Bilder wurden vom Sequenztyp b-SSFP, mit einem TR / TE Verhältnis von 3,7 / 1,85, einem FOV (*Field of View*) von 320-400 mm, einer planaren Auflösung von 1,6 mm x 1,6 mm und einer Schichtdicke von 8 mm aufgenommen. Die rekonstruierten Bilder waren 200 x 200 Pixel groß. Die Akquisitionszeit für jedes Bild betrug 33 ms (**Tabelle 1**). Zur MR-Aufnahme wurde ein 1,5 T *Scanner* (MAGNETOM Avanto fit Siemens Healthcare, Software Version Syngo MR VE11) mit einer 32-Kanal-*Spine-Matrix*-Spule und einer 18-Kanal-Körperspule genutzt. Das MR-Protokoll für die aufzunehmenden Sequenzen bestand aus standardisierten kardialen Lokalisationsdaten und retrospektiv aufgenommenen Zwei- und Vierkammerblick-Datensätzen. Diese Einstellungen entsprachen den Protokolleinstellungen, die in vorherigen veröffentlichten Studien der Arbeitsgruppe verwendet wurden (40,68).

In dieser Studie wurden je Proband zwei Datensätze aufgenommen, eine singuläre mittventrikuläre Schicht mit 4000 Bildern und eine kardiale Kurzachsenstapel-Volumetrie (KAV) mit 19 Schichten und jeweils 900 Bildern. Für einen Probanden wurde eine zweite zeitversetzte KAV aufgenommen (**Tabelle 1**). Jeder Freiwillige lag in Rückenlage mit den Füßen voran im MR-Scanner (**Abbildung 7a**). Jeder Proband wurde mit einem Bauchgurt und einem EKG ausgestattet, um physiologische qualitative Parameter der Atmung und des Herzens aufzunehmen (**Abbildung 7b**). Die Software ‚Siemens Signal Logging System VE11C‘ wurde dafür genutzt. Die quantitativen Daten der Atmung wurden mit Hilfe einer MRkS aufgenommen. Alle akquirierten Daten wurden im Kontrollraum registriert (68).

Sequenzparameter	Originale KAV	Mittventrikuläre Schicht	Optimiertes Training
Sequenztyp	b-SSFP		
TR/TE (ms)	3,7/1,85		
Orientation	Kurze Achse		
Flip Winkel (°)	60		
Bandweite (Hz/Pixel)	760		
FOV (mm)	320-400		
Bildmatrix (Pixel)	200 x 200		
planare Auflösung (mm x mm)	1,6 x 1,6		
Schichtdicke (mm)	8		
Bildakquisitionszeit (ms)	33		
Anzahl an Schichten	19	1	10
Schichtabstand(mm)	0	0	8
Bilder	900	4000	300
Totale Aufnahmedauer(s)	570	133	100
Atemtyp	freie Atmung	tiefe & freie Atmung	freie Atmung

Tabelle 1: MR-Parameter. Sequenzparameter, die für die Echtzeit-Magnetresonanztomographie (MRT) zur Bestimmung der originalen kardiale Kurzachsenstapel-Volumetrie (KAV) und einer mittventrikulären Schicht verwendet wurden. Zusätzlich dargestellt sind die gesetzten Parameter der resultierenden Sequenz eines optimierten Trainings nach der Methodenentwicklung. *b-SSFP*, balanced steady-state free precession; *FOV*, field of view; *TE*, Echozeit; *TR*, Repetitionszeit. (Modifizierte Tabelle aus Malik et al. (68))

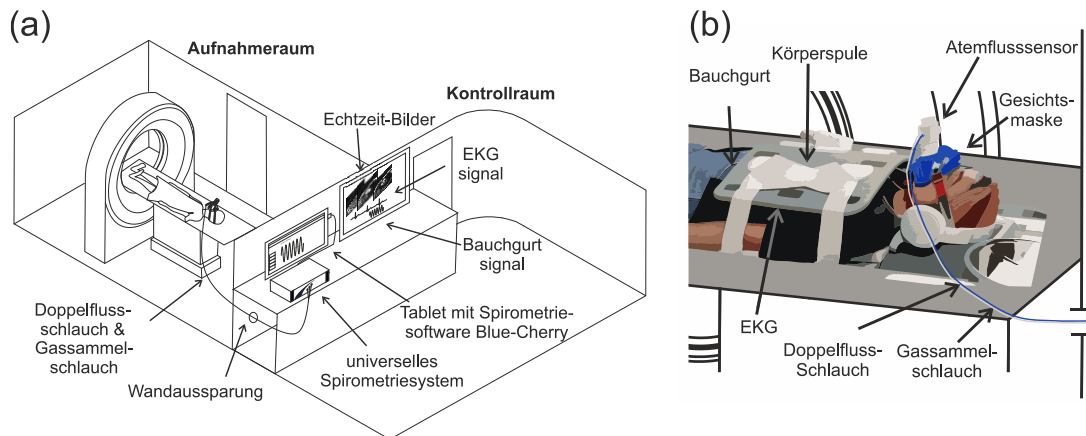


Abb.7: Aufbau und Monitoring.

(a) Darstellung des Aufnahmerraums und des dazugehörigen Kontrollraums.

(b) Physiologisches Monitoring während der Aufnahme. Der Atemfluss in mL/s wurde vom Flusssensor mittels einer Ableitung des Differentialdrucks an einer Membran gemessen. Der Differentialdruck wurde über einen doppellumigen Atemschlauch (blau) weitergeleitet. O₂/CO₂ - Messungen wurden durch den Gassammelschlauch (weiß) transportiert. Das Schlauchsystem wurde durch eine Wandaussparung geleitet und im Kontrollraum mit einem universellem Spirometriesystem verbunden. Auf dem dazugehörigen *Tablet* konnten die Atemkurven in Echtzeit über die *Software Blue Cherry* beobachtet werden, das EKG -, das Bauchgurtsignal und die Echtzeit-MR-Bilder wurden auf einem separaten Rechner verzeichnet und abgespeichert. (Modifizierte Grafik aus Malik et al. (68))



Video 1. Abnahme der Gesichtsmaske. Die Maske kann innerhalb weniger Sekunden vom Gesicht des Probanden entfernt werden. Der MR-Tisch muss dafür nicht aus seiner Position verschoben werden. (In seperater Anlage, Video aus dem Supplement von Malik et al. (68))

2.2 MR-kompatible Spirometrie

Quantitative Atemflussdaten wurden mit Hilfe einer MRkS (Geratherm Respiratory GmbH, Bad Kissingen, Deutschland) aufgezeichnet. Das MRkS bestand aus einem leichtgewichtigen Flusssensor (Ergoflow, Geratherm Respiratory GmbH, Bad Kissingen, Deutschland), welcher mit Hilfe einer Silikongesichtsmaske (COSMED Deutschland GmbH, Werneck, Deutschland, Größe S) am Probanden angebracht wurde (**Abbildung 7b**). Über eine Membran am Flusssensor wurde der differentielle Druckunterschied der Atmung abgeleitet und über einen modifizierten doppellumigen Atemschlauch (Ergostik Schlauchsatz, 6m, Geratherm Respiratory GmbH, Bad Kissingen, Deutschland) an das Spirometriesystem im Kontrollraum weitergeleitet. Ein Adapter (Ergoclip-Adapter, Geratherm Respiratory GmbH, Bad Kissingen, Deutschland) verband den doppellumigen Atemschlauch mit dem Flusssensor. Der Atemschlauch hatte ein effektives Totraumvolumen von < 96 mL. Weiter war zum Feuchtigkeitsausgleich im Schlauchset ein Permapur-Trockenschlauch (PermaPure, Lakewood, NJ, USA) mit einem hydrophobem Minifilter installiert (68).

Das Signal hatte eine Bandbreite von ± 5 L/s und wurde mit einer Abtastrate von 125 Hz und einer Genauigkeit von $\pm 3\%$ oder ± 50 mL/s aufgezeichnet (68). Die Steuerung der MRkS und der akquirierten Daten wurde mittels der *Software Blue Cherry* (Geratherm Respiratory GmbH, Bad Kissingen, Deutschland) auf einem separaten Tablet im Kontrollraum durchgeführt (**Abbildung 7a**).

Im Weiteren der Studie war die Adjustierung oder Abnahme der Silikonmaske innerhalb des MR-Scanners von Relevanz. Diese konnte innerhalb weniger Sekunden ohne Positionsänderungen des Probanden durchgeführt werden (**Video 1**) (68).

2.3 Modellarchitektur

Zur Vorhersage des quantitativen Atemfluss wurde ein Modell des MLs in Python entwickelt und mittels Ridge Regression trainiert. Dafür wurden respiratorische Daten der aufgenommenen Probanden vorverarbeitet, Echtzeit-MR-Bilder aufbereitet und in Trainings- und Testdaten unterteilt. L Daten wurden für das Training des Modells verwendet. Die übrigen T Daten wurden für das Testen genutzt. Das trainierte Modell sollte mit Hilfe der ungesehenen Testdaten die dazugehörigen respiratorischen Parameter vorhersagen. Die vorhergesagten Daten wurden nachbearbeitet (68).

2.3.1 Vorverarbeitung

Die Vorverarbeitung der Trainingsdaten erfolgte durch die Normalisierung der respiratorischen Flussmessung $f^{(1)}, \dots, f^{(L)} \in \mathbb{R}$:

$$f_{norm}^{(l)} = \frac{f^{(l)} - \mu_f}{\sigma_f} \text{ für } l = 1, \dots, L,$$

mit μ_f und σ_f als Mittelwert und Standardabweichung von $f^{(1)}, \dots, f^{(L)}$. Die Vorverarbeitung der Echtzeit-MR-Bilder des Trainings- und Testdatensatzes waren essenziell zur Erstellung des respiratorischen Modells. Dafür wurde zunächst jedes MR-Bild, bestehend aus 200 x 200 Pixeln, durch eine Maskierung manuell auf das Interessengebiet begrenzt (**Abbildung 8a**). Dadurch wurden Bildartefakte aus dem Modelltraining und der Prädiktion exkludiert. Darauf folgend fand eine Heruntertaktung der Pixel statt, wodurch nur jeder fünfte Pixel im MR-Bild zur weiteren Berechnung berücksichtigt wurde. Es resultierte pro Aufnahme eine Auflösung mit 40 x 40 Pixeln (68).

2.3.2 Respiratorisches Modell

Der zentrale Kern des respiratorischen Modells bildeten normalisierte optische Flussvektoren (OFV). Sie repräsentierten die Pixelbewegung durch Richtung und Geschwindigkeit von einem Bild zum nächsten (69). Diese wurden aus zwei aufeinanderfolgenden maskierten, abgetasteten Bildern bestimmt (**Abbildung 8b**). Die Berechnung der OFV beruhte auf der Applikation der zehn Schritten der iterativen Lucas-Kanade Methode (70,71). Dies wurde mit Hilfe der Scikit-Image Bibliothek in Python (72) implementiert. Jeder OFV wurde ausgehend von jedem heruntergerechneten Pixel des maskierten Bildes mit zwei Werten, ihrer Richtung und Geschwindigkeit, als Reihe von Tensoren $X^{(1)}, \dots, X^{(L)} \in \mathbb{R}^{40 \times 40 \times 2}$ für den Trainingsdatensatz und $X^{(1)}, \dots, X^{(T)} \in \mathbb{R}^{40 \times 40 \times 2}$ für den Testdatensatz beschrieben. Dabei beinhaltete der Trainingsdatensatz weniger als ein Drittel der Gesamtdatenmenge (68).

Das respiratorische lineare Modell ermöglichte die Prädiktion eines normalisierten Atemflusses:

$$\hat{f}_{norm}^{(l)} = \sum_{m=1}^{40} \sum_{n=1}^{40} \sum_{k=1}^2 W_{mnk} X_{mnk}^{(l)} + b.$$

Die Modellparameter $W \in \mathbb{R}^{40 \times 40 \times 2}$ und $b \in \mathbb{R}$ wurden durch das Training des Modells bestimmt (**Abbildung 8c**). Dafür wurde Ridge Regression genutzt, um die optimalen Größen zu finden. Das Optimierungsproblem wurde durch die regularisierte Methode der kleinsten Quadrate gelöst:

$$\min_{W,b} \frac{1}{L} \sum_{l=1}^L \left(\hat{f}_{norm}^{(l)} - f_{norm}^{(l)} \right)^2 + \lambda \sum_{m=1}^{40} \sum_{n=1}^{40} \sum_{k=1}^2 W_{mnk}^2.$$

Wobei der Hyperparameter $\lambda > 0$ durch die *Leave-one-out* Kreuzvalidierung (LOO-KV) im Trainingsdatensatz gesetzt wurde. Zur Umsetzung und Anpassung des Modells sowie der Kreuzvalidierung wurde Scikit-Learn in Python (73) genutzt (68).

2.3.3 Prädiktion

Die Vorhersage des Atemflusses des Testdatensatzes $\hat{f}_{norm}^{(1)}, \dots, \hat{f}_{norm}^{(T)}$ erfolgte in dem die Tensoren des Testdatensatzes $X^{(1)}, \dots, X^{(T)}$ dem trainierten Modell übergeben wurden:

$$\hat{f}_{norm}^{(t)} = \sum_{m=1}^{40} \sum_{n=1}^{40} \sum_{k=1}^2 W_{mnk} X_{mnk}^{(t)} + b$$

Abbildung (8d) (68).

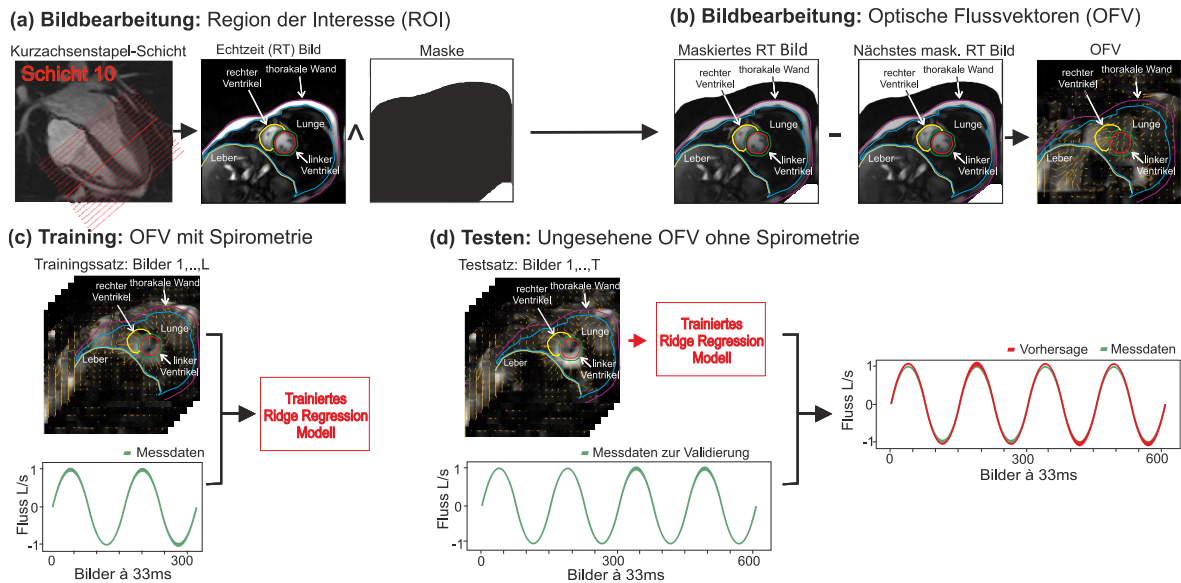


Abb. 8: Respiratorisches Modell. Aus einer kardialen Kurzsachsenstapelvolumetrie wurden Echtzeit-MR-Bilder akquiriert.

- (a) Das Interessengebiet der Atmung wurde durch eine Maske begrenzt.
- (b) Aus aufeinanderfolgenden heruntergerechneten maskierten Echtzeit-MR-Bildern des Trainingsdatensatzes und des Testdatensatzes wurden anhand der 10 Schritte der Lucas-Kanade Methode optische Flussvektoren (OFV) bestimmt.
- (c) Die OFV des Trainingsdatensatzes wurden gepaart mit den gemessenen Atemflüssen einem linearen Modell zum Training mittels Ridge Regression übergeben.
- (d) Die OFV des ungesehenen Testdatensatzes wurden dem trainierten Modell übergeben, um eine Vorhersage des respiratorischen Atemflusses durchzuführen. (Modifizierte Grafik aus Malik et al. (68))

2.3.4 Nachbearbeitung

Nach Vorhersage des Atemflusses $\hat{f}_{norm}^{(1)}, \dots, \hat{f}_{norm}^{(T)}$ wurde dieser denormalisiert

$$\hat{f}_{denorm}^{(t)} = \sigma_f \hat{f}_{norm}^{(t)} + \mu_f \text{ für } t = 1, \dots, T$$

mit μ_f und σ_f als Mittelwert und Standardabweichung von $f^{(1)}, \dots, f^{(L)}$ (68).

Im nächsten Schritt wurde der denormalisierte vorhergesagte Fluss $\hat{f}_{denorm}^{(1)}, \dots, \hat{f}_{denorm}^{(T)}$ geglättet. Dafür wurde der Savitzky-Golay Filter (74), von SciPy (75) implementiert, angewandt. Der denormalisierte, geglättete vorhergesagte Fluss wird im Weiteren durch die Folge $\hat{f}^{(1)}, \dots, \hat{f}^{(T)}$ beschrieben (68).

Aus den gemessenen Daten des Atemflusses $f^{(1)}, \dots, f^{(T)}$ und den Daten des vorhergesagten denormalisierten geglätteten Atemflusses $\hat{f}^{(1)}, \dots, \hat{f}^{(T)}$ wurden die zugehörigen Lungenvolumina $v^{(1)}, \dots, v^{(T)}$ der Messung und $\hat{v}^{(1)}, \dots, \hat{v}^{(T)}$ der Vorhersage durch

$$v^{(t)} = c \sum_{k=1}^t f^{(k)} \text{ for } t = 1, \dots, T,$$

$$\hat{v}^{(t)} = c \sum_{k=1}^t \hat{f}^{(k)} \text{ for } t = 1, \dots, T,$$

berechnet. Um die richtige Messgröße in mL zu erhalten, wurde zur Faktorisierung die Konstante $c = 10$ verwendet (68). Die berechneten Lungenvolumina zeigten eine versetzte Basislinie, weswegen diese korrigiert wurde. Die Korrektur der Basislinie erfolgte durch Subtraktion der mit Hilfe von medianer Interpolation geglätteten Minima von den unkorrigierten Lungenvolumina.

2.4 Modellkonfiguration und -optimierung

2.4.1 Originale kardiale KAV-Protokoll

Ziel der Modellkonfiguration war, dass die Prädiktion respiratorischer Daten einer originalen kardialen KAV (**Tabelle 1 und Abbildung 9a**) ermöglicht wird. Im Weiteren wurden die Modelle so optimiert, dass ein optimales Gleichgewicht zwischen der prädiktiven Güte und der Trainingsdauer mit MRkS gefunden wurde. Dafür wurden die Trainingsdatensätze hinsichtlich des Atemmusters, der Anzahl der Atemzüge und der Anzahl der genutzten Schichten variiert (68).

2.4.2 Optimierung Training - Atemmuster und Anzahl der Atemzüge

Zunächst wurde die Güte der Vorhersage des Modells auf einer mittventrikulären (MV) Schicht analysiert. Dafür wurde für jeden Probanden ein Ausgangstrainingsdatensatz, bestehend aus vier tiefen und vier normalen Atemzügen (1200 Bilder, 40s), erstellt. Der Testdatensatz bestand unveränderbar für jeden Probanden aus 18 normalen Atemzügen (2800 Bilder, 94s). Die Sequenzparameter der Datensätze sind aus **Tabelle 1** zu entnehmen (68).

Zunächst wurde der Einfluss des Atemmusters untersucht. Dafür wurde ein Modell mit vier tiefen Atemzügen (600 Bilder, 20s) und ein Modell mit vier normalen Atemzügen (600 Bilder, 20s) trainiert. Respiratorische Vorhersagen beider Modelle wurden miteinander und mit dem Ausgangsmodell verglichen und analysiert (**Abbildung 9b**) (68).

Darauffolgend wurde die Auswirkung der Anzahl der Atemzüge auf das vorhergesagte Ergebnis der Testdatensätze betrachtet. Dazu wurden die Trainingsdatensätze auf vier normale Atemzüge (600 Bilder, 20s), drei normale Atemzüge (450 Bilder, 15s), zwei normale Atemzüge (300 Bilder, 10 s), einen normalen Atemzug (150 Bilder, 5 s) sowie einen halben Atemzug (75 Bilder, 2,5 s) gekürzt und die Vorhersage auf den Testdatensätzen wurden miteinander und mit dem Ausgangsmodell verglichen und analysiert (**Abbildung 9b**) (68).

2.4.3 Optimierung Training - Anzahl an Schichten

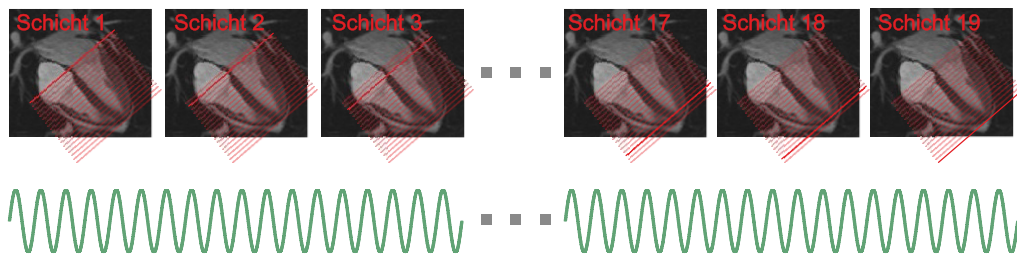
Nach der Optimierung des Atemmusters und der Anzahl an Atemzügen auf einer MV-Schicht wurde die Übertragbarkeit des Modells und die Güte der Vorhersage auf einer originalen kompletten KAV mit 19 Schichten überprüft und evaluiert (68).

Der Ausgangstrainingsdatensatz für die Prädiktion einer kompletten KAV beinhaltete zunächst 19 Schichten mit jeweils 300 Bilder und hatte somit eine komplette Dauer von 190s. Das trainierte Modell wurde als Referenzmodell genutzt. Der Testdatensatz blieb über

alle Trainingsdatensätze gleich und beinhaltete 19 Schichten mit den jeweiligen ungesehenen restlichen 600 Bilder pro Schicht und einer kompletten Dauer von 380s (68).

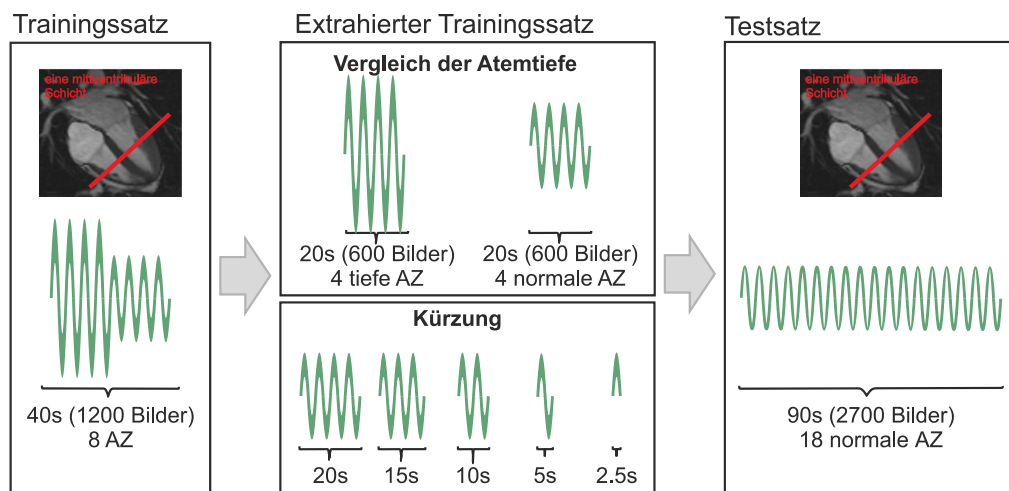
Der Ausgangstrainingsdatensatz wurde anhand der Anzahl der genutzten Schichten, in Anbetracht der technischen Umsetzbarkeit, optimiert (**Abbildung 9c**). Zur Optimierung wurden drei Trainingsdatensätze untersucht. Die in den Trainingsdaten nicht enthaltenen Testschichten wurden durch benachbarte trainierte Schichten vorhergesagt. Beim Ersten gekürzten Protokoll wurde jede zweite Schicht beim Training berücksichtigt, somit wurde der Trainingsdatensatz effektiv auf 10 Schichten (jeweils 300 Bilder/Schicht, komplette Dauer 100 s) gekürzt. Technisch wurde damit der Schichtabstand von 0 mm auf maximal 8 mm angehoben (**Tabelle 1**). Im nächsten Schritt wurde jede dritte Schicht zum Trainieren genutzt, somit wurden sieben Schichten (jeweils 300 Bilder/Schicht, komplette Dauer 70s, Schichtabstand maximal 16 mm) berücksichtigt. Im letzten Reduktionsschritt wurde jede sechste Schicht in den Trainingsdatensätzen inkludiert. Dadurch wurden vier Schichten (jeweils 300 Bilder/Schicht, komplette Dauer 40s, Schichtabstand maximal 40 mm) im Training betrachtet. Die trainierten Modelle wurde hinsichtlich ihrer Güte untereinander und gegenüber dem Referenzmodell verglichen (68).

(a) Original-Protokoll: Kurzachsen-Volumetrie mit simultaner Spirometrie



~ 10 min: (900 Bilder) / 6 Atemzüge(AZ) / Schicht (S)

(b) Protokoll Optimierung: Atemmuster und Anzahl an Atemzügen



(c) Protokoll Optimierung: Anzahl an Schichten

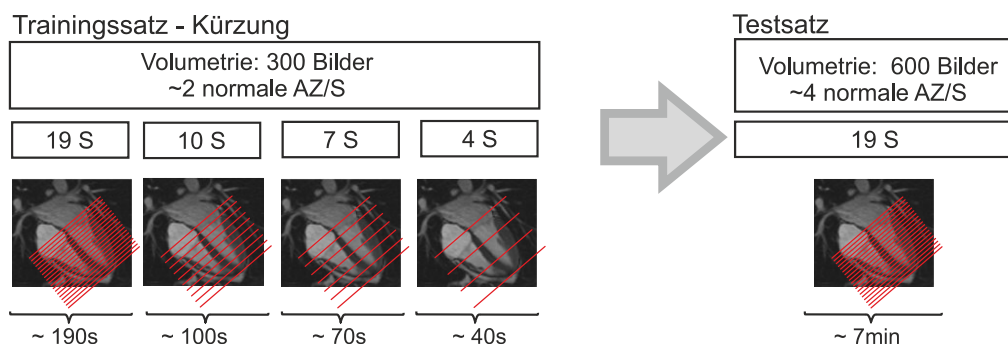


Abb. 9: Methodenentwicklung. (a) Die Entwicklung eines Trainingsdatensatzes zur Vorhersage respiratorischer Messgrößen einer kardialen Kurzachsenstapel-Volumetrie mit freier Atmung. Der Trainingsdatensatz wurde in Abhängigkeit des Atemmusters, der Anzahl der Atemzüge (AZ) (b) und der Anzahl der Schichten (c) variiert. (Modifizierte Grafik aus Malik et al. (68))

2.4.4 Trainiertes Modell - Wiederverwendung auf unbekanntem Probanden

Bislang wurde für jeden Probanden aus dem aufgenommenen KAV-Datensatz ein Trainings- und Testdatensatz erstellt, sodass für jeden Probanden ein separates Modell erzeugt wurde. Im Weiteren war die Wiederverwertbarkeit und die Übertragbarkeit bereits trainierter Modelle bekannter Probanden auf Testdatensätzen unbekannter Probanden von Interesse. Ein iteratives Verfahren zur Analyse wurde erstellt. Als Testdatensatz wurde die komplette KAV eines willkürlich selektierten Probanden (19 Schichten, jeweils 900 Bilder/Schicht, Nr.208) gewählt, der von allen Trainingsdatensätzen exkludiert wurde (68).

Zum Training wurden zunächst die Daten kompletter kardialer Volumetrie bekannter einzelner Probanden (19 Schichten, jeweils 900 Bilder, Nr. 205 und 210) selektiert und genutzt. Dabei wurden die Trainingsdatensätze bezüglich Gemeinsamkeiten analysiert, die zu einem guten prädiktiven Resultat beim unbekanntem Testdatensatz führten. Unter den Gemeinsamkeiten wurde der Einfluss von Körpermaßen, dem Körpergewicht und der Körpergröße und der Einfluss des Winkels WTA. Wir definierten den WTA anhand der Echtzeit-MR-Bilder einer Kurzachsenvolumetrie, in dem der Winkel zwischen der unteren Bildkante und der anterioren Thoraxwand oder der Parallelen dazu gemessen wurde. Der Einfluss des WTA wurde weiter vertiefend untersucht. In einer zweiten Analyse wurden die Bilder des Testdatensatzes, dessen WTA nicht mit dem WTA der Bilder des Trainingsdatensatzes (Nr. 210) übereinstimmte, rotiert. Damit die Bildorientierung und der WTA des Testdatensatzes und des Trainingsdatensatzes angeglichen wurden. Die Bildrotation wurde mit Hilfe von Scikit-Image (70) und SciPy (73) implementiert und im Bereich der Vorverarbeitung des Bildes nach der Maskierung des Interessengebietes und vor der Berechnung der OFV hinzugefügt (68).

Im Anschluss wurden Daten der bekannten Probanden unter Betrachtung der gemeinsamen Eigenschaften zusammengesetzt und zum Training eines Modells genutzt. Hierfür wurden drei komplette KAV (19 Schichten, jeweils 900 Bildern, Nr. 205, 207 und 212) mit einer Dauer von 1710 Sekunden eingesetzt (68).

In einer letzten Untersuchung wurde ein Trainingsdatensatz aus allen neun bekannten Probanden unabhängig der Gemeinsamkeiten zusammengesetzt und zum Trainieren eines Modells genutzt. Somit wurden neun KAV (19 Schichten, jeweils 900 Bildern, Nr. 205-207,209-214) mit einer Dauer von 5130 Sekunden zum Training zusammengeführt (68).

Zusätzlich wurde der Einfluss der Denormalisierung des prädizierten respiratorischen Flusses untersucht (68). Wie in 2.3.4 beschrieben wurde diese bislang durch den Mittelwert und der Standardabweichung des Atemflusses des Trainingsdatensatzes durchgeführt. Es wurde daher eine Methode für eine adaptive Denormalisierung untersucht, bei der der

Mittelwert und die Standardabweichung des respiratorischen Flusses des Testdatensatzes genutzt wurde:

$$\hat{f}_{denorm}^{(t)} = \sigma_{ft} \hat{f}_{norm}^{(t)} + \mu_{ft} \text{ für } t = 1, \dots, T$$

mit μ_{ft} und σ_{ft} als Mittelwert und Standardabweichung vom gemessen respiratorischen Fluss des Testdatensatzes $f^{(1)}, \dots, f^{(T)}$. Hier wurden iterativ alle Probanden bis auf den Testprobanden einzeln unabhängig ihrer Gemeinsamkeiten als Trainingsdatensatz selektiert. Die trainierten Modelle wurden zur Vorhersage auf dem Testdatensatz genutzt und validiert. Beide Verfahren der Denormalisierung wurden gegenübergestellt und ausgewertet.

2.4.5 Trainiertes Modell - Wiederverwendung auf zeitlich versetzter Aufnahme

Letztlich wurde untersucht, inwiefern ein trainiertes Modell bei einem zeitlich versetzt aufgenommenen Testdatensatz desselben Probanden wiederverwertbar ist. Dafür wurden für einen Probanden (Nr.205) zwei zeitlich versetzte KAV akquiriert. Für den Trainingsdatensatz wurde die optimierte KAV (10 Schichten, jeweils 300 Bildern) der ersten Aufnahme genutzt. Die zweite, zeitlich versetzte Aufnahme einer KAV (19 Schichten, jeweils 900 Bildern) wurde als Testdatensatz verwendet (68).

Die Güte der Vorhersage auf dem zweiten Testdatensatz, wurde mit der prädiktiven Güte des ersten Testdatensatzes der ersten Studie des Probanden verglichen.

Auch nach Vergleich beider Testdatensätze wurde eine Anpassung des WTA des zweiten Testdatensatzes durchgeführt, um dem WTA des Trainingsdatensatz zu entsprechen. Dafür wurden die Echtzeit-MR-Bilder des zweiten Probanden, nach der in 2.4.4 erwähnten Methode, rotiert (68).

2.5 Evaluationsmetrik

Die Datenanalyse und -auswertung wurde mit Hilfe von Python durchgeführt. Dafür wurde in Python auf Funktionen der Bibliotheken Statsmodels (74), Numpy (75) und SciPy (73) zurückgegriffen. Die Evaluation wurde auf den unbekannt Testdatensätzen $X^{(1)}, \dots, X^{(T)} \in \mathbb{R}^{40 \times 40 \times 2}$ mit den dazugehörigen gemessenen respiratorischen Flüssen $f^{(1)}, \dots, f^{(T)}$ und den vorhergesagten respiratorischen Flüssen $\hat{f}^{(1)}, \dots, \hat{f}^{(T)}$ durchgeführt (68).

2.5.1 Qualitative Evaluation

Zur qualitativen Validierung der Ergebnisse wurde die Übereinstimmung der Graphen der vorhergesagten respiratorischen Flüsse und Lungenvolumina mit den Graphen vom MRkS gemessenen Atemflüssen und Lungenvolumina verglichen. Im Weiteren erfolgten eine qualitative Visualisierung und eine Analyse der trainierten Modellgewichte für jeden Probanden. Dafür wurden die Gewichte des trainierten Modells auf ein unspezifisch selektiertes Bild eines Trainingsdatensatz mit einer Größe von 5x5 Pixeln abgebildet (68).

2.5.2 Quantitative Evaluation

Zur quantitativen Validierung wurde zunächst der Determinationskoeffizienten

$$R^2_{Fluss} = 1 - \frac{\sum_{t=1}^T (f^{(t)} - \hat{f}^{(t)})^2}{\sum_{t=1}^T (f^{(t)} - \bar{f})^2}$$

als Maß der Korrelation zwischen gemessenen Atemfluss $f^{(1)}, \dots, f^{(T)}$ und vorhergesagten Atemfluss $\hat{f}^{(1)}, \dots, \hat{f}^{(T)}$ bestimmt. Wobei $\bar{f} = \frac{1}{T} \sum_{t=1}^T f^{(t)}$ den Mittelwert der gemessenen Atemflüsse bildete. Ähnlich wurde der Determinationskoeffizient

$$R^2_{Volumen} = 1 - \frac{\sum_{t=1}^T (v^{(t)} - \hat{v}^{(t)})^2}{\sum_{t=1}^T (v^{(t)} - \bar{v})^2}$$

als Maß der Korrelation zwischen gemessenen Lungenvolumina $v^{(1)}, \dots, v^{(T)}$ und den vorhergesagten Lungenvolumina $\hat{v}^{(1)}, \dots, \hat{v}^{(T)}$ bestimmt. Wobei $\bar{v} = \frac{1}{T} \sum_{t=1}^T v^{(t)}$ den Mittelwert der gemessenen Lungenvolumina bildete (68).

Zur quantitativen Beurteilung des Fehlers wurde zusätzlich der relative mittlere quadratische Fehler

$$rMSE_{Fluss} = \frac{\frac{1}{T} \sum_{t=1}^T (f^{(t)} - \hat{f}^{(t)})^2}{\frac{1}{T} \sum_{t=1}^T f^{(t)2}}$$

zwischen dem gemessenen Atemflüssen $f^{(1)}, \dots, f^{(T)}$ und den vorhergesagten Atemflüssen $\hat{f}^{(1)}, \dots, \hat{f}^{(T)}$ ermittelt. Analog wurde der relative mittlere quadratische Fehler

$$rMSE_{Volumen} = \frac{\frac{1}{T} \sum_{t=1}^T (v^{(t)} - \hat{v}^{(t)})^2}{\frac{1}{T} \sum_{t=1}^T v^{(t)2}}$$

zwischen den korrespondierenden gemessenen Lungenvolumina $v^{(1)}, \dots, v^{(T)}$ und den vorhergesagten Lungenvolumina $\hat{v}^{(1)}, \dots, \hat{v}^{(T)}$ berechnet (68).

Für die Anwendung im klinischen Kontext wurden die Differenzen zwischen den gemessenen und vorhergesagten absoluten maximalen Atemflüssen (aMF) und die Differenzen zwischen den gemessenen und vorhergesagten absoluten Tidalvolumina (aTV) untersucht.

Für jeden Atemzug wurde ein Paar von Atemfluss-Extrema $z^{(p)} = (f^{(p)}, \hat{f}^{(p)})$ für $p = 1, \dots, P$ ausgewählt und die absolute maximale Atemflussdifferenz

$$aMFD_{z^{(p)}} = | \max f^{(p)} - \max \hat{f}^{(p)} |$$

zwischen den gemessenen Atemflüssen $f^{(1)}, \dots, f^{(P)}$ und den vorhergesagten Atemflüssen $\hat{f}^{(1)}, \dots, \hat{f}^{(P)}$ wurde kalkuliert (68).

Analog wurde für jeden Atemzug ein Paar von Lungenvolumen-Extrema $z^{(p)} = (v^{(p)}, \hat{v}^{(p)})$ für $p = 1, \dots, P$ bestimmt und damit wurde die absolute Tidalvolumendifferenz

$$aTVD_{z^{(p)}} = |TV_{v^{(p)}} - TV_{\hat{v}^{(p)}}|,$$

mit $TV_{v^{(p)}} = | \max v^{(p)} - \min v^{(p)} |$ und $TV_{\hat{v}^{(p)}} = | \max \hat{v}^{(p)} - \min \hat{v}^{(p)} |$,

zwischen den gemessenen Lungenvolumina $v^{(1)}, \dots, v^{(P)}$ und den vorhergesagten Lungenvolumina $\hat{v}^{(1)}, \dots, \hat{v}^{(P)}$ ermittelt (68).

Für jeden Probanden wurde über alle Atemzüge innerhalb eines Testdatensatzes der Mittelwert (μ), die Standardabweichung (SA) und die minimale und maximale Begrenzung von rMSE, R^2 , aTVD und aMFD ausgerechnet. Um eine allgemeine Aussage der Modellgüte über alle Probanden zu treffen, wurde der Mittelwert dieser Werte über alle Probanden errechnet. Zusätzlich wurde die Variabilität von R^2 zwischen den Probanden über alle Schichten und die Variabilität von R^2 zwischen den vorhergesagten Schichten über alle Probanden untersucht (68).

Im Weiteren wurde das Maß der Übereinstimmung von dem aTV und dem aMF von vorhergesagten und gemessenen respiratorischen Daten mit Hilfe von Bland-Altman (BA)

Diagrammen illustriert und statistisch analysiert. Die Messpunkte der Bland-Altman Graphen wurden durch zufällig ausgewählte Punktpaare aus aTV oder aMF jeder prädizierten Schicht (19 insgesamt) von jedem Probanden (10 insgesamt) kumuliert und grafisch veranschaulicht. Für jeden Optimierungsschritt, in dem die Anzahl an verwendeten Schichten im Training gekürzt wurde, wurde eine BA-Grafik erstellt. Jede Grafik stellte somit eine kumulierte Verteilung der Differenzen der aTV und aMF mit einem 95% Konfidenzintervall (KI) und einer Linie der mittleren Differenz, dem Bias, dar (68).

Zur Beurteilung der Übereinstimmung wurden Qualitätskriterien für die mittlere aMFD und die mittlere aTVD definiert. Eine gute Übereinstimmung und damit ein Qualitätskriterium (QC_{Fluss}) für die mittlere aMFD wurde definiert als eine maximale Abweichung von ± 50 mL/s bis ± 100 mL/s (76). Für die mittlere aTVD wurde als Qualitätskriterium (QC_{Volumen}), eine absolute maximale Abweichung von < 20 mL festgesetzt. Dies übertraf die Mindestanforderung, die für klinische Anwendung einer Spirometrie festgelegt wurde (100 mL für Tidalvolumina unter 1 L) (76). Bei Überschreitung dieser Grenzen des mittleren aMFD und der mittleren aTVD wurden die dazugehörigen Vorhersagen der einzelnen Probanden als quantitativ unzureichend gewertet. Darüber hinaus wurden die Anpassungen an das Modell, bei denen weniger als 9 Probanden die mittleren Qualitätskriterien erfüllen konnten, verworfen (68).

3 Ergebnisse

Die Studie wurde an 10 gesunden Probanden durchgeführt. In der Probandengruppe lag das durchschnittliche Alter bei $25,4 \pm 2,9$ Jahre (zwischen 21 und 30 Jahren), die durchschnittliche Körpergröße betrug $173,1 \pm 8,5$ cm (zwischen 157 und 185 cm) und das durchschnittliche Gewicht lag bei $67,6 \pm 11,8$ kg (zwischen 54 und 93 kg) (68).

Das Ziel der Studie war die Prädiktion von respiratorischen Atemparameter aus Echtzeit-MR-Bildern einer kardialen KAV. Es wurden Modelle auf einzelnen Schichten bekannter Probanden oder auf kompletten kardialen KAV bekannter Probanden, unbekannter Probanden oder bekannter Probanden mit zeitlich versetzten Aufnahmen trainiert und analysiert. Die prädiktive Güte dieser Modelle wurde untersucht.

3.1 Prädiktion auf einer mittventrikulären Schicht

Zunächst wurde die Vorhersage auf einer MV-Schicht geprüft. Dafür wurde als Erstes ein Ausgangsmodell auf einer MV-Schicht mit acht Atemzügen (vier tiefen und vier normalen Atemzügen) mit einer Gesamtdauer von 40 s trainiert und auf den verbliebenen 18 normalen Atemzüge derselben Schicht getestet.

Über alle Probanden konnten respiratorische Daten mit einer sehr guten prädiktiven Güte vorhergesagt werden. Graphisch ließ sich eine sehr genaue Übereinstimmung der vorhergesagten und gemessenen Atemfluss- und Lungenvolumenkurven feststellen (**Abbildung 10**). Auch die statistische Analyse bestätigte dies: Der mittlere R^2 für alle Probanden lag für den Atemfluss bei 0,986 und für das Lungenvolumen bei 0,993. Der mittlere rMSE betrug 0,017 für den Atemfluss und 0,003 für das Lungenvolumen. Die Qualitätskriterien (QC_{Volumen} and QC_{Fluss}) konnten für alle Probanden erfüllt werden. Im Mittel lag die absolute maximale Flussdifferenz (aMFD) bei 34 mL/s und die absolute Tidalvolumendifferenz (aTVD) bei 7 mL (**Tabelle 2**) (68).

Die Übertragung dieses Ausgangsmodell auf die Konstruktion eines Vorhersagemodells für eine komplette kardiale KAV mit 19 Schichten würde zu einer sehr langen Trainingszeit (>760 s) führen, die die Zeit des Testdatensatzes (570 s) überschreiten würde. Dies motivierte die weitere Modelloptimierung. Dabei wurde die prädiktive Güte von unterschiedlichen Trainingseinheiten analysiert. Untersucht wurden dabei unterschiedliche Atemmuster, die Anzahl der Atemzüge und die Anzahl der verwendeten Schichten (68).

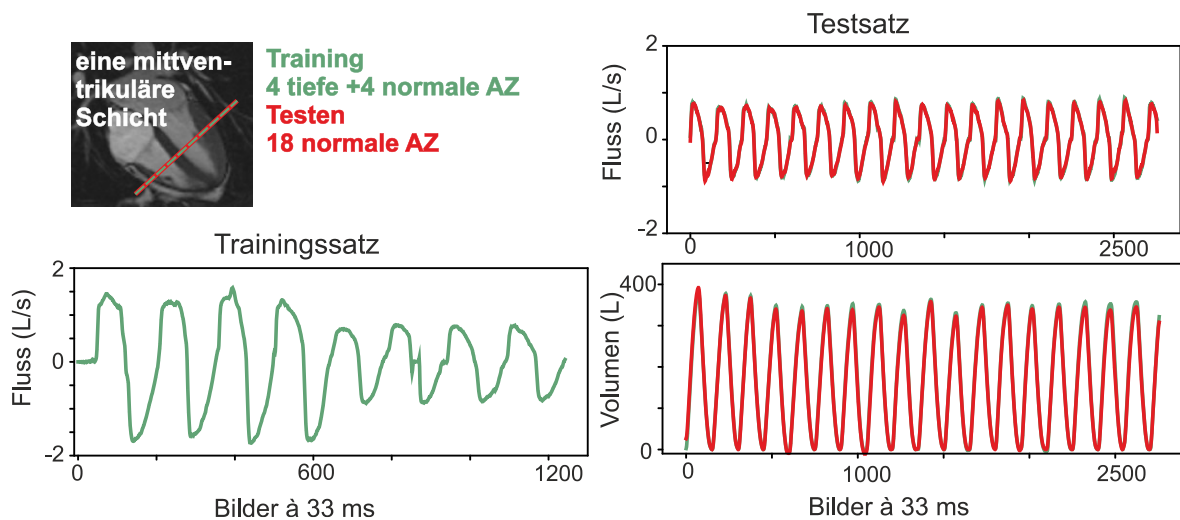


Abb. 10: Ausgangsmodell einer mittventrikulären Schicht. Training auf einer mittventrikulären (MV) Schicht mit vier tiefen und vier normalen Atemzügen (AZ) (grüne Kurve, links unten) und die Vorhersage eines Probanden vom Ausgangsmodell des respiratorischen Flusses (rote Kurve, oben rechts) und dem korrespondierenden Volumen (rote Kurven, unten rechts) auf dem Testdatensatz. Unterlegt sind die gemessenen Kurven des Testdatensatzes (grün, rechts). (Modifizierte Grafik aus Malik et al. 2025 (68))

3.2 Optimierung Training - Atemmuster und Anzahl der Atemzüge

Zur Untersuchung des Atemmusters wurden Modelle mit tiefen Atemzügen oder normalen Atemzügen trainiert und die prädiktive Güte wurde miteinander verglichen. Zur Analyse der benötigten Anzahl der genutzten Atemzüge im Training wurden Modelle mit Trainingsdaten aus vier, drei, zwei, ein und einem halben normalen Atemzug und deren prädiktive Güte wurde miteinander verglichen.

Beim Vergleich der Atemtiefe – vier tiefe oder vier normale Atemzüge – im Training zeigte sich, dass zwar qualitativ keine wesentlichen Unterschiede in der Kurvenübereinstimmung von vorhergesagten und gemessenen Kurven vorlagen (**Abbildung 11a, b**), allerdings eine quantitativ höhere Prädiktionsgüte bei den Modellen vorlag, die mit vier normalen Atemzügen trainiert wurden (**Tabelle 2**). Bei den Vorhersagen der Modelle, die mit vier tiefen Atemzügen trainiert wurden, präsentierte sich schlechtere mittlere aMFD und mittlere aTVD. Das Qualitätskriterium für das Lungenvolumen konnte bei vier Probanden nicht erfüllt werden. Zudem verschlechterten sich der mittlere rMSE und der mittlere Determinationskoeffizient sowohl für den vorhergesagten respiratorischen Fluss als auch für das vorhergesagte Lungenvolumen (**Tabelle 2**). Daraus folgend wurde in den weiteren Trainingseinheiten ausschließlich normale Atemzüge betrachtet (68).

Ein Training mit vier normalen Atemzügen zeigte qualitativ (**Abbildung 11b**) als auch quantitativ (**Tabelle 2**) ähnliche Ergebnisse, wie das Ausgangsmodell (**Abbildung 10**). Die

vorhergesagten Kurven illustrierten eine gute Übereinstimmung mit den gemessenen Kurven (**Abbildung 11b**). Die Qualitätskriterien konnten bei allen Probanden erfüllt werden (**Tabelle 2**) (68).

Weitere Kürzungen der Trainingseinheiten auf drei oder zwei normalen Atemzug gingen mit minimalen Verlusten der prädiktiven Güte im Vergleich zum Ausgangsmodell einher (**Tabelle 2**). Eine Kürzung des Trainingsdatensatzes auf einen normalen Atemzug zeigten qualitativ weiterhin gute Kurvenübereinstimmungen (**Abbildung 11c**) und quantitativ tendenziell schlechtere aber ausreichend gute prädiktive Werte für R^2 und rMSE (**Tabelle 2**). Allerdings konnte in diesem Fall das Qualitätskriterium für das Lungenvolumen bei drei Probanden nicht erfüllt werden (**Tabelle 2**) (68).

Eine Kürzung des Trainingsdatensatzes auf einen halben normalen Atemzug lediglich aus einer Inspiration ohne Expiration führte sowohl qualitativ als auch quantitativ zu deutlichen Verlusten in der Vorhersage. Die negativen Werte des respiratorischen Flusses konnten in der prädiktiven Kurve nicht abgebildet werden, weswegen auch die Maxima im prädizierten Lungenvolumen nicht berechnet werden konnten (**Abbildung 11d**). Der mittlere rMSE und der mittlere Determinationskoeffizient zeigten schlechtere Ergebnisse. Zudem konnte das Qualitätskriterium für das Lungenvolumen bei keinem Probanden und für den Atemfluss nur ungenügend erfüllt werden (**Tabelle 2**) (68).

Unter Berücksichtigung der technischen Umsetzbarkeit, der qualitativen und quantitativen Ergebnisse und um sicherzustellen, dass mindestens ein normaler Atemzug im Trainingsdatensatz enthalten ist, wurde die Reduktion auf zwei normale Atemzüge favorisiert (68).

Trainings-satz	Volumen				Fluss			
	R ² $\mu \pm SA$ [Min-Max]	rMSE $\mu \pm SA$ [Min-Max]	aTVD [mL] $\mu \pm SA$ [Min-Max]	QC _{Vol.}	R ² $\mu \pm SA$ [Min-Max]	rMSE $\mu \pm SA$ [Min-Max]	aMFD [mL/s] $\mu \pm SA$ [Min-Max]	QC _{Fluss}
4 tiefe + 4 normale	0,993 \pm 0,003 [0,989-0,997]	0,003 \pm 0,002 [0,001-0,007]	7 \pm 2 [3-11]	10/10	0,986 \pm 0,005 [0,977-0,994]	0,017 \pm 0,009 [0,006-0,026]	34 \pm 20 [13-85]	10/10
4 tiefe	0,961 \pm 0,026 [0,916-0,996]	0,009 \pm 0,008 [0,001-0,030]	16 \pm 10 [5-36]	6/10	0,961 \pm 0,034 [0,871-0,994]	0,034 \pm 0,028 [0,007-0,109]	47 \pm 20 [16-76]	10/10
4 normale	0,993 \pm 0,003 [0,987-0,997]	0,002 \pm 0,001 [0,001-0,004]	7 \pm 2 [4-11]	10/10	0,986 \pm 0,006 [0,974-0,994]	0,013 \pm 0,004 [0,005-0,019]	24 \pm 8 [14-37]	10/10
3 normale	0,986 \pm 0,015 [0,942-0,996]	0,003 \pm 0,002 [0,001-0,006]	9 \pm 3 [4-15]	10/10	0,984 \pm 0,008 [0,967-0,993]	0,014 \pm 0,005 [0,007-0,024]	31 \pm 18 [15-79]	9/10
2 normale	0,981 \pm 0,018 [0,937-0,996]	0,005 \pm 0,004 [0,001-0,009]	11 \pm 6 [6-25]	9/10	0,978 \pm 0,012 [0,954-0,989]	0,019 \pm 0,007 [0,010-0,035]	35 \pm 18 [19-86]	10/10
1 normaler	0,975 \pm 0,026 [0,969-0,996]	0,012 \pm 0,013 [0,002-0,037]	17 \pm 9 [8-36]	7/10	0,959 \pm 0,035 [0,859-0,986]	0,032 \pm 0,014 [0,014-0,057]	40 \pm 13 [23-65]	10/10
0.5 normale	-0,762 \pm 1,372 [-2,585-0,919]	0,496 \pm 0,393 [0,027-1,098]	84 \pm 52 [23-167]	0/10	-0,383 \pm 1,124 [-2,084-0,881]	0,441 \pm 0,239 [0,091-0,821]	76 \pm 39 [26-147]	8/10

Tabelle 2: Deskriptive statistische Daten zur vorhergesagten Ateminformation von Modellen, die auf einer mittventrikulären Schicht trainiert wurden. Modelltraining wurde mit verschiedenen Atemmuster und Anzahl an Atemzügen durchgeführt.

μ , Mittelwert, SA, Standardabweichung und minimale bis maximale Grenzen von R^2 , Determinationskoeffizient, $rMSE$, relativer mittlerer quadratischer Fehler, $aTVD$, absolute Tidalvolumendifferenz in mL und $aMFD$, absolute maximale Atemflussdifferenz in mL/s von allen Probanden (N=10, jeweils eine Schicht mit 18 Atemzügen). $QC_{Vol.}$, Anzahl an Probanden, die das Qualitätskriterium für das Volumen ($aTVD_s$, Mittelwert $aTVD$ für einen Probanden <20ml) erfüllen und QC_{Fluss} , Anzahl an Probanden, die das Qualitätskriterium für den Atemfluss ($aMFD_s$, Mittelwert $aMFD$ für einen Probanden <100 ml/s) erfüllen. (Modifizierte Tabelle aus dem Supplement von Malik et al. (68))

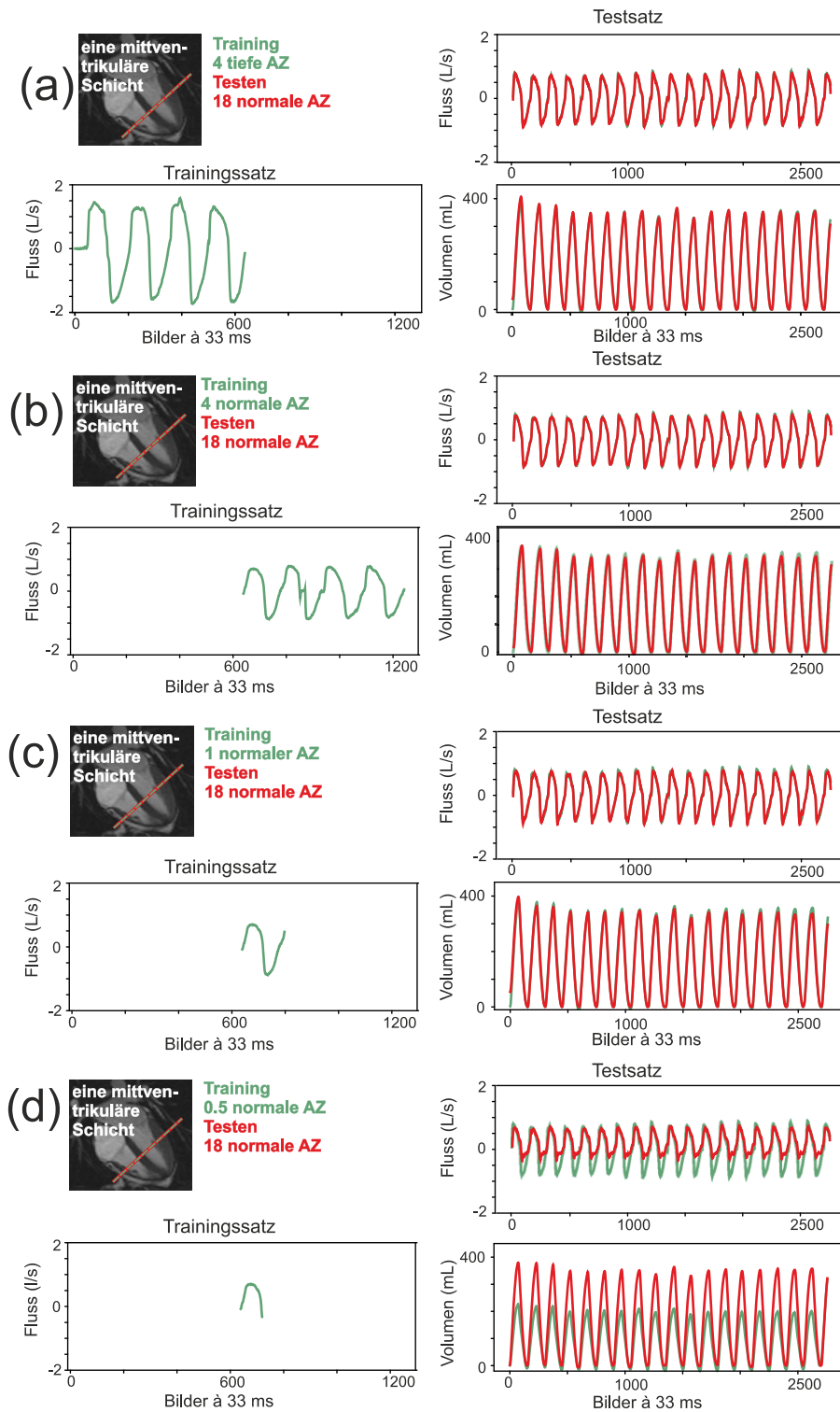


Abb. 11: Modelloptimierung Atemmuster und Anzahl an Atemzügen. Für progressiv gekürzte Trainingsdatensätze auf einer mittventrikulären Schicht: (a) vier tiefe Atemzüge (grün, links), (b) vier normale Atemzüge (grün, links), (c) ein normaler Atemzug (grün, links) und (d) ein einzelner inspiratorischer Atemzug (grün, links) wurden vorhergesagte Daten (rot, rechts) und gemessene Daten (grün, rechts) des Atemfluss (rechts) und des Lungenvolumen (rechts) verglichen. (Modifizierte Grafik aus Malik et al. (68))

3.3 Prädiktion der originalen kardialen Volumetrie

Die Vorhersage auf allen Schichten einer kompletten originalen kardialen KAV wurde analysiert. Die Trainingseinheit des Referenzmodells enthielt 19 Schichten mit jeweils 300 Bildern entsprechend ca. zwei normalen Atemzügen und zeigte eine Aufnahmedauer von 190s. Der ungesehene Testdatensatz umfasste alle 19 Schichten mit jeweils 600 Bildern entsprechend ca. vier normalen Atemzügen.

Über alle Probanden konnten mittels des Referenzmodells respiratorische Daten einer KAV mit einer sehr guten prädiktiven Güte vorhergesagt werden.

Qualitativ konnte zwischen den gemessenen und den vorhergesagten Kurven des Atemflusses als auch des Lungenvolumens eine hohe Übereinstimmung gezeigt werden (**Abbildung 12a**).

In der quantitativen Analyse konnte verdeutlicht werden, dass die Differenzen der randomisierten ausgewählten Paare aus allen Schichten und allen Probanden von aTV und aMF gleichmäßig verteilt waren. Die mittlere Differenz für die aTV lag bei 7 mL (95% KI: +40/-26 mL) und die mittlere Differenz der aMF 0,0 L/s (95% KI: +0,088/-0,94 L/s) (**Abbildung 13a**). Somit konnten kumuliert im Mittel die Qualitätskriterien QC_{Volumen} und QC_{Fluss} eingehalten werden. Weiter lagen für über alle Probanden und über allen Schichten ähnliche R^2 vor (**Abbildung 14a, Abbildung 15a**). Damit konnte unabhängig der vorherzusagenden Schicht als auch unabhängig des ausgewählten Probanden eine kontinuierlich gute Vorhersage veranschaulicht werden. Bei einzelner Betrachtung der Probanden erfüllten neun Probanden QC_{Volumen} und zehn Probanden QC_{Fluss} . Auch die mittleren R^2 und rMSE präsentierten sowohl für das Lungenvolumen als auch für den Atemfluss gute Ergebnisse (**Tabelle 3**).

3.4 Optimierung Training - Anzahl der Schichten

Zur weiteren Optimierung der Modellvorhersage wurde die Anzahl der benötigten Schichten zum Training reduziert. Unter Betrachtung der technischen Umsetzbarkeit wurde die Reduzierung auf jede zweite Schicht (insgesamt zehn Schichten, ca. 100 s Aufnahmedauer), jede dritte Schicht (insgesamt sieben Schichten, ca. 70 s Aufnahmedauer) und jede sechste Schicht (insgesamt vier Schichten, ca. 40 s Aufnahmedauer) untersucht.

Die Reduktion des Trainingsdatensatzes auf jede zweite Schicht ermöglichte weiterhin eine sehr gute Vorhersage.

Qualitativ konnte eine präzise Überlagerung der gemessenen und vorhergesagten Kurven des respiratorischen Flusses und des Lungenvolumens gezeigt werden, vergleichbar mit dem Referenzmodell (**Abbildung 12b**) (68).

Auch die quantitative Analyse der prädiktiven Güte veranschaulichte ein ähnliches Bild (**Tabelle 3**). Weiterhin konnte eine homogene Verteilung der Differenzen der randomisierten ausgewählten Paare der aMF und der aTV gezeigt werden. Auffällig war, dass die mittlere Differenz der aTV mit 4,27 mL (95% KI: +33/-24 mL) tendenziell geringer ausfiel und ein engeres KI aufwies als im Referenzmodell. Die mittlere Differenz des aMF lag bei -0,1L/s (95% KI: +0,084/-0,1 L/s) (**Abbildung 13b**) (68).

Auch war die Variabilität von R^2 zwischen den vorhergesagten Schichten über alle Probanden (**Abbildung 14b**). und zwischen allen Probanden über alle Schichten (**Abbildung 15b**) gering. Es konnte gezeigt werden, dass Testschichten, die nicht im Trainingsdatensatz enthalten waren, ebenso gut vorhergesagt werden konnten wie Testschichten, die im Trainingsdatensatz enthalten waren. Es gab keine bemerkenswerten Differenzen zu den Ergebnissen des Referenzmodells (**Abbildung 14a, b und 15a, b**) (68).

Auch die mittlere R^2 und die mittlere rMSE unterschieden sich nicht wesentlich vom Referenzmodell. Die Qualitätskriterien QC_{Volumen} und QC_{Fluss} konnten für alle Probanden erfüllt werden (**Tabelle 3**).

Im Gegensatz dazu führte die Reduktion auf jede dritte Schicht und die Reduktion auf jede sechste Schicht zu schlechteren Ergebnissen (68).

Qualitativ zeigten sich eine unzureichende Kurvenübereinstimmung, bei der die Maxima der Volumen- und Flusskurven nicht adäquat abgebildet wurden (**Abbildung 12 c, d**). Dieses Ergebnis spiegelte sich auch in der quantitativen Auswertung wider: Es konnte eine stärkere Streuung der Differenzen der zufällig ausgewählten Paare von aMF und aTV in den Bland-Altman-Diagrammen wahrgenommen werden. Besonders bei der Reduktion auf

insgesamt vier Schichten im Trainingsdatensatz zeigte sich eine erhöhte mittlere Differenz der aTV von 11,17 mL (95%-KI: +73/-51 mL) und eine mittlere Differenz der aMF von -0,01 L/s (95%-KI: +0,13/-0,15 L/s) (**Abbildung 13 c, d**).

Diese Inhomogenität tritt noch deutlicher in den Unterschieden der Vorhersage zwischen den Schichten hervor (**Abbildung 14 c, d**). Besonders sichtbar wurde, dass die Prädiktion auf Schichten, die nicht im Trainingsdatensatz erhalten waren, schlechtere Determinationskoeffizienten aufwies (68).

Die probandenspezifischen Vorhersagen zeigten ein deutlich schlechteres Ergebnis als das Ergebnis des Referenzmodells und des Modells, welches mit zehn Schichten trainiert wurde (**Abbildung 15 c, d**) (68).

Bezüglich der weiteren Güteparameter stellte sich eine abnehmende Tendenz der mittleren R^2 und eine leichte Zunahme des mittleren rMSE für den respiratorischen Fluss und das Lungenvolumen dar (**Tabelle 3**).

Letztlich konnte das Qualitätskriterium für das Lungenvolumen bei einem Probanden nach dem Training mit sieben Schichten und bei sechs Probanden nach dem Training mit vier Schichten nicht erfüllt werden (**Tabelle 3**).

Schlussfolgernd wurde ein optimiertes Training mit zehn Schichten zur Vorhersage einer kompletten kardialen KAV bevorzugt.

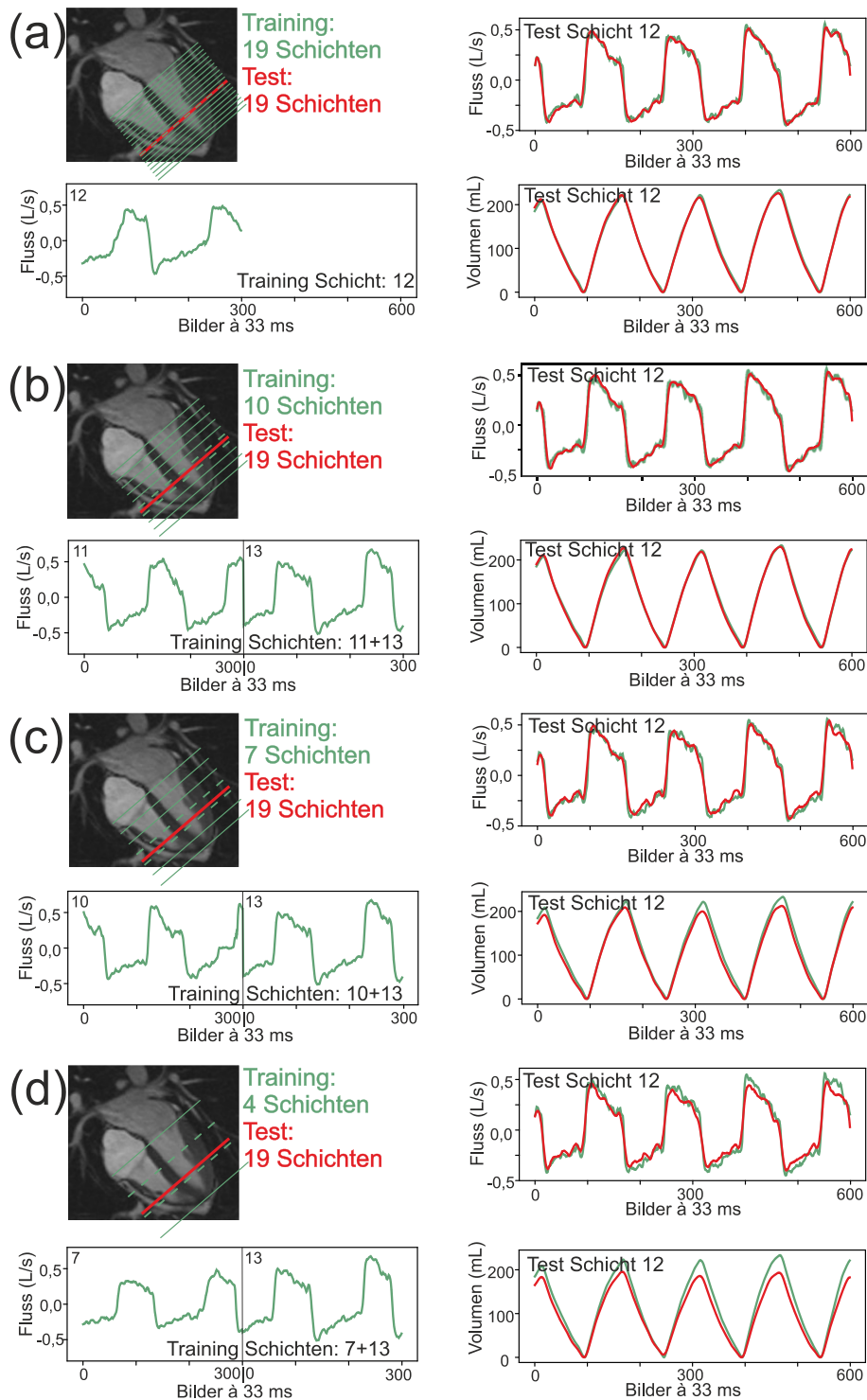


Abb. 12: Modelloptimierung Schichtreduktion. Optimierung der Trainingsdatensätze durch eine Schichtenreduktion. Exemplarisch an einer Testschicht (12, rote Linie) dargestellt. Training (grüne gestrichelte Linie/n) fand (a) in der gleichen Schicht (12) (b) in den unmittelbar benachbarten zwei Schichten (11+13), (c) in weiter voneinander entfernten Schichten (10+13), (d) in noch weiter voneinander entfernten Schichten (7+13) statt. Für eine komplette kardiale Kurzachsenstapel-Volumetrie wurden insgesamt 19 (a), 10 (b), 7 (c), 4 (d) Schichten trainiert. Respiratorische Kurven des Testens (rechte Spalte) mit vorhergesagtem (rot) und gemessenem (grün) Atemfluss (je Abschnitt, obere Grafik) und dem dazugehörigen Lungenvolumen (je Abschnitt, untere Grafik). (Modifizierte Grafik aus Malik et al. (68))

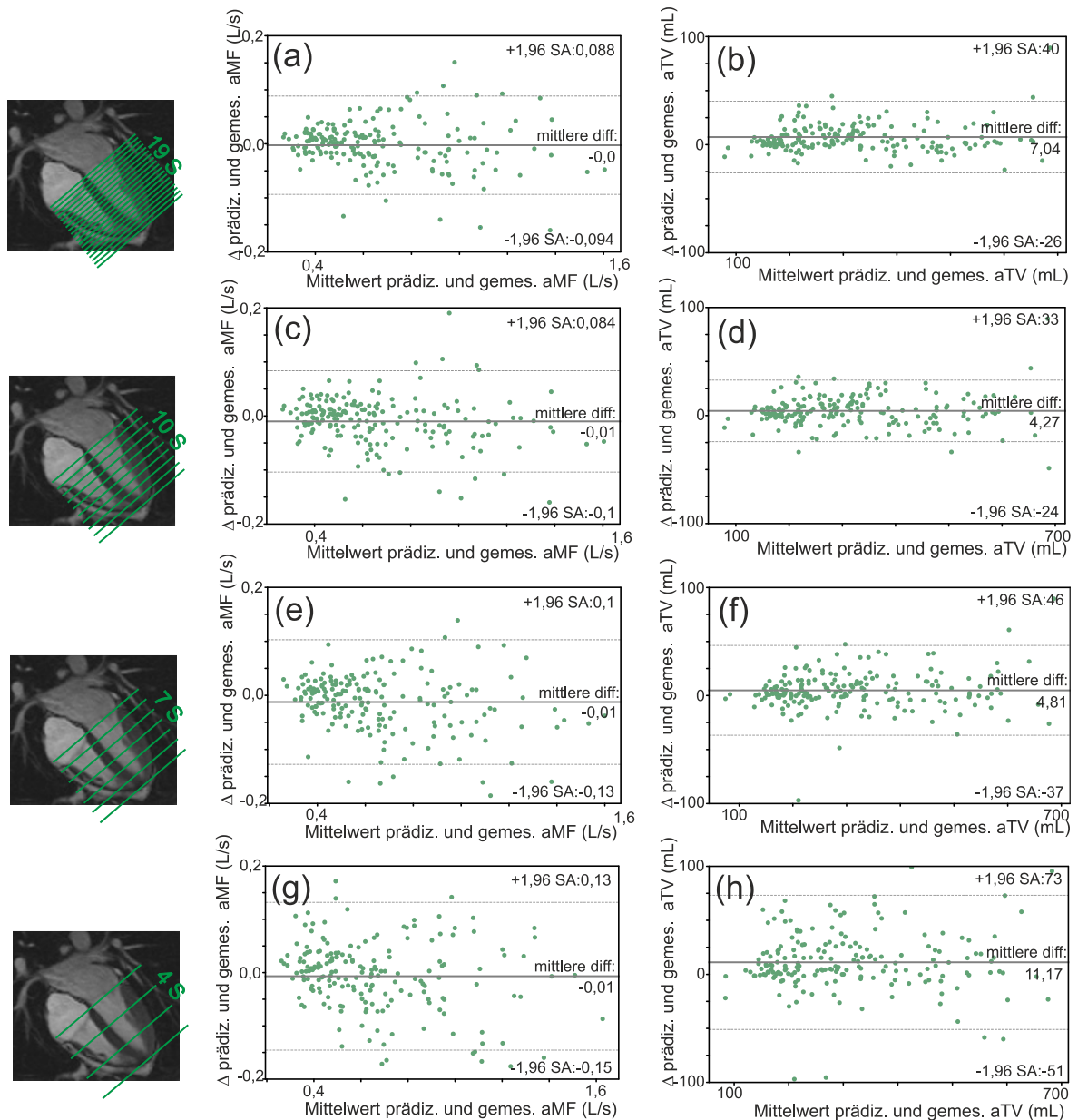


Abb. 13: Bland-Altman (BA) Modelloptimierung Schichtreduktion. Für jeden Reduktionsschritt der Trainingsschichten (linke Spalte: 19 Schichten (a), 10 Schichten (b), 7 Schichten (c) und 4 Schichten (d)) wird für jeden Probanden und jede Schicht eines Testdatensatzes ein randomisiertes Paar von Maxima für prädizierte (prädiz.) und gemessene (gemes.) Daten von absoluten maximalen Atemfluss (aMF) (mittlere Spalte) und von absoluten Tidalvolumina (aTV) (rechte Spalte) selektiert und in einer BA-Grafik illustriert. Ähnliche mittlere Differenzen (Bias) mit ähnlicher Streuung werden für ein Training mit allen 19 Schichten und ein Training mit 10 Schichten für aMF und aTV erreicht. Eine Reduktion auf sieben Schichten oder vier Schichten zum Training führt zu einer stärkeren Streuung und einem höheren Bias. (Modifizierte Grafik aus dem Supplement von Malik et al. (68))

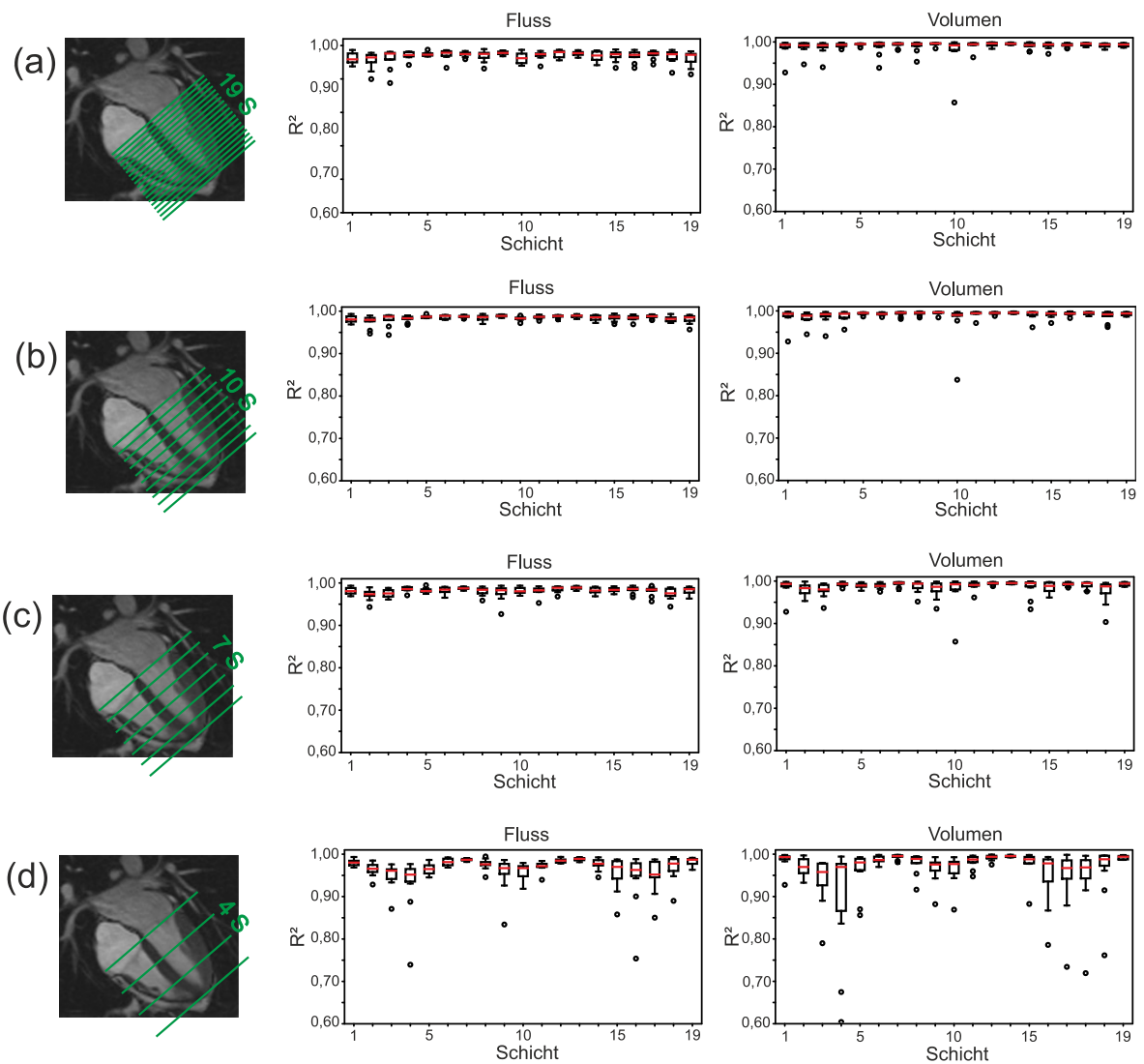


Abb. 14: Boxplot Schichtabhängigkeit. Modelltraining mit 19 Schichten (a), mit 10 Schichten (b), mit 7 Schichten (c) und 4 Schichten (d) mit jeweils 2 normalen Atemzügen pro Schicht. Und die dazugehörigen Boxplots der Determinationskoeffizienten R^2 aller Probanden für den prädizierten Atemfluss (mittlere Spalte) und das prädizierte Lungenvolumen (rechte Spalte) in Abhängigkeit der Schicht, auf welcher getestet wurde. (Modifizierte Grafik aus dem Supplement von Malik et al. (68))

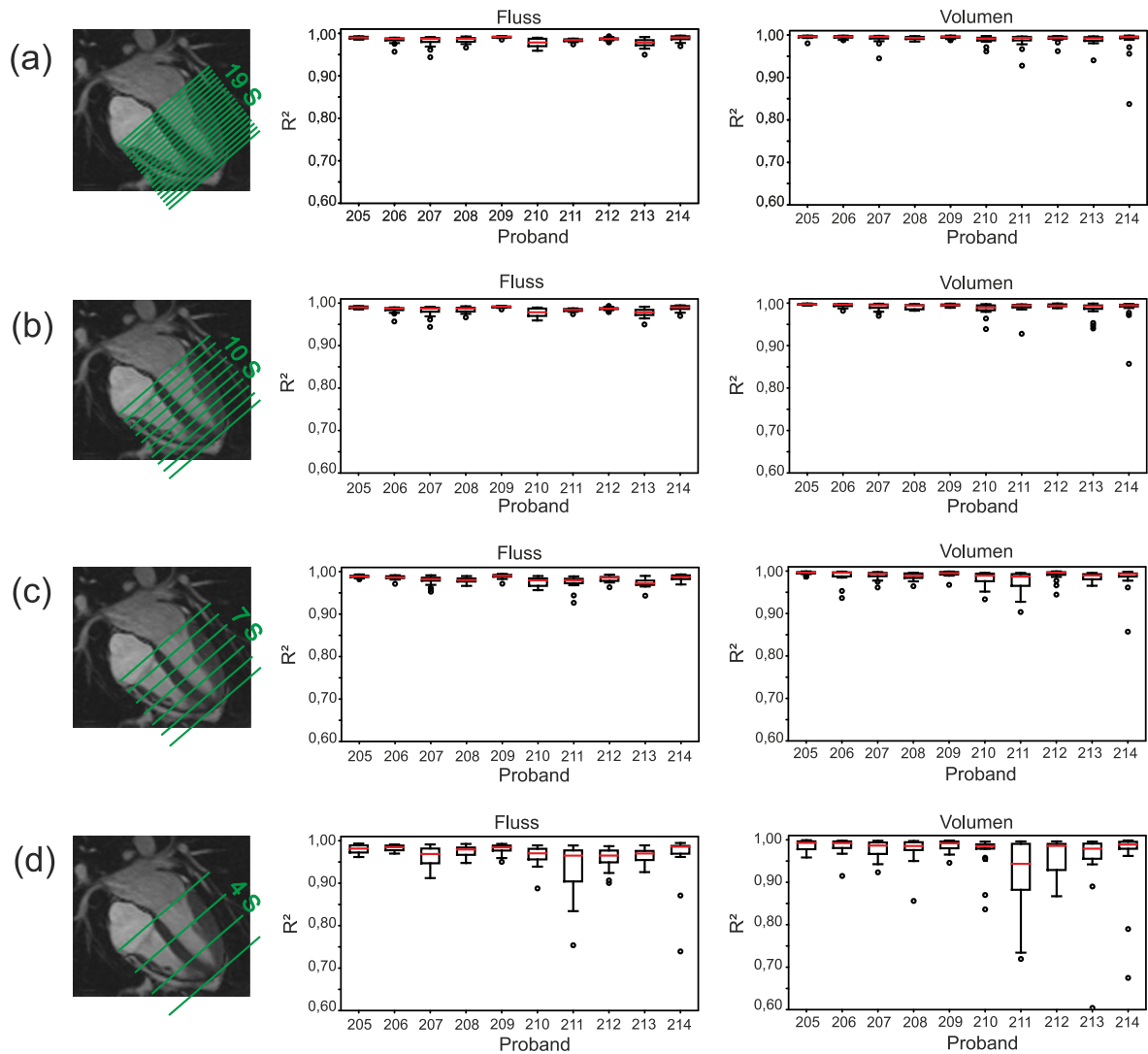


Abb. 15: Boxplot Probandenabhängigkeit. Modelltraining mit 19 Schichten (a), mit 10 Schichten (b), mit 7 Schichten (c) und 4 Schichten (d) mit jeweils 2 normalen Atemzügen pro Schicht. Und die dazugehörigen Boxplots der Determinationskoeffizienten für alle Schichten für den prädizierten Atemfluss (mittlere Spalte) und das prädizierte Lungenvolumen (rechte Spalte) in Abhängigkeit des Probanden, mit welchem trainiert und auf welchem getestet wurde. (Modifizierte Grafik aus dem Supplement von Malik et al. (68))

Trainingsatz	Volumen			Atemfluss		
Anzahl an Schichten	R^2 $\mu \pm SA$ [Min-Max]	$rMSE$ $\mu \pm SA$ [Min-Max]	$QC_{vol.}$	R^2 $\mu \pm SA$ [Min-Max]	$rMSE$ $\mu \pm SA$ [Min-Max]	QC_{Fluss}
19	0,991 \pm 0,010 [0,985-0,997]	0,003 \pm 0,001 [0,001-0,005]	9/10	0,984 \pm 0,007 [0,977-0,989]	0,015 \pm 0,004 [0,009-0,022]	10/10
10	0,990 \pm 0,009 [0,983-0,995]	0,003 \pm 0,001 [0,002-0,005]	10/10	0,984 \pm 0,006 [0,977-0,990]	0,015 \pm 0,004 [0,009-0,023]	10/10
7	0,987 \pm 0,013 [0,974-0,995]	0,004 \pm 0,002 [0,002-0,007]	9/10	0,982 \pm 0,009 [0,973-0,989]	0,018 \pm 0,005 [0,010-0,026]	10/10
4	0,967 \pm 0,039 [0,913-0,987]	0,011 \pm 0,006 [0,004-0,026]	4/10	0,967 \pm 0,023 [0,933-0,982]	0,029 \pm 0,008 [0,017-0,045]	10/10

Tabelle 3: Deskriptive statistische Daten zur vorhergesagten Ateminformation kompletter kardialer Volumetriem mit verschiedenen Trainingsätzen. Training wurde mit unterschiedlicher Anzahl an Schichten durchgeführt.

μ , Mittelwert, SA , Standardabweichung und minimale bis maximale Grenzen von R^2 , Determinationskoeffizient, $rMSE$, relativer mittlerer quadratischer Fehler von allen Probanden ($N=10$, 19 Schichten pro Probanden mit jeweils 4 Atemzügen pro Testschicht). $QC_{vol.}$, Anzahl an Probanden, die das Qualitätskriterium für das Volumen ($aTVD_s$, Mittelwert $aTVD$ für einen Probanden <20 mL) erfüllen und QC_{Fluss} , Anzahl an Probanden, die das Qualitätskriterium für den Atemfluss ($aMFD_s$, Mittelwert $aMFD$ für einen Probanden <100 mL/s) erfüllen. (Modifizierte Tabelle aus dem Supplement von Malik et al. (68))

3.5 Regionale Gewichtsverteilung des Modells

Zur Bewertung der regionalen Gewichtsverteilung der Modelle wurden diese auf der zugehörigen Schicht abgebildet. Abhängig der Schichtlokalisierung konnten unterschiedliche Verteilungen wahrgenommen werden. Auf basaler kardialer Ebene lag eine stärkere Gewichtung in der Region um das Zwerchfell und um die kaudalen Anteile des Thorax vor. In den mittventrikulären Schichten zeigte sich eine homogenere Verteilung entlang der anterioren thorakalen Wand. Letztlich wurden die stärksten Gewichte der apikalen kardialen Schnittebene im Bereich des apikalen Thorax abgebildet (**Abbildung 16**) (68).

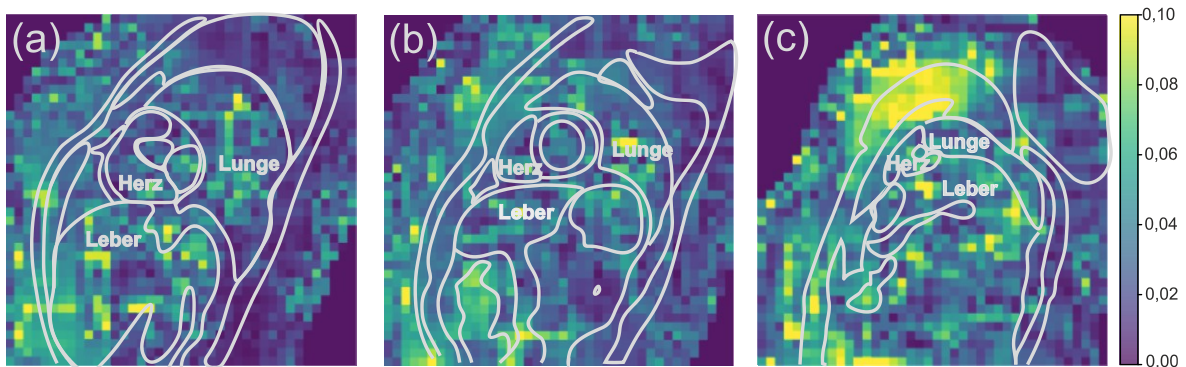


Abb. 16: Regionale Gewichtsverteilung. Beitrag der Gewichte (0: kein Beitrag, 0,1: hoher Beitrag) der optischen Flussvektoren und ihre regionale Verteilung für eine (a) basale Schicht, eine (b) mittventrikuläre Schicht und (c) eine apikale Schicht. (Modifizierte Grafik aus Malik et al. (68))

3.6 Wiederverwendung von trainierten Modellen

Nach der Modelloptimierung wurde die Wiederverwertbarkeit von trainierten Modellen untersucht. Zum einen wurde dafür die Übertragbarkeit von bereits trainierten Modellen auf Testdatensätzen von unbekanntem Probanden untersucht. Zum anderen wurde die Wiederverwendung eines trainierten Modells eines Probanden auf einem zeitlich versetzt aufgenommenen Testdatensatz desselben Probanden analysiert.

3.6.1 Testdatensatz - unbekannter Proband

Zur Validierung der Vorhersage bereits trainierter Modelle, welche auf dem Testdatensatz eines unbekanntem Probanden (Nr.208) angewandt wurden, wurde ein iteratives Verfahren entwickelt. In dem Verfahren wurde sukzessiv der Einfluss der Körpermaße, des WTA und der daraus resultierenden Bildorientierung sowie der Anzahl der Trainingsprobanden untersucht. Die **Abbildung 17** illustriert die qualitativen Ergebnisse und die **Tabelle 4** stellt die quantitativen Ergebnisse dieses Verfahrens dar.

Der Einfluss der Körpergröße und des Körpergewichts (**Tabelle 5**) ist von geringer Bedeutung. Bei Probanden mit ähnlichen Körpermaßen, aber unterschiedlichen WTA – Trainingsproband (Nr.210) und Testproband (Nr. 208) – konnte keine gute Übereinstimmung zwischen den gemessenen und vorhergesagten Daten festgestellt werden (**Abbildung 17 k**) (**Tabelle 4**) (68).

Der Einfluss eines ähnlichen WTA ist von entscheidender Bedeutung. Ein Modell, welches mit Daten des Probanden Nr. 205 (WTA: 36°) trainiert und auf den Daten des Probanden Nr.208 (WTA: 28°) getestet wurde, konnte gute Ergebnisse erzielen: Die vorhergesagten und gemessenen Kurven zeigten eine gute Überlagerung mit geringen Differenzen in den Spitzenwerten und einem wenig verrauschten Flusssignal (**Abbildung 17 l**). Auch quantitativ konnten akzeptable Ergebnisse erreicht werden: Die mittlere aTVD lag über dem definierten Qualitätskriterium für das Lungenvolumen, lag jedoch im Bereich des von Criée et al. für Spirometrie festgelegten Wertes (< 100 mL für Tidalvolumina < 1 L) (76) (**Tabelle 4**). Die mittlere aMFD verfehlte das Qualitätskriterium für den Atemfluss. Dennoch zeigte sich eine hohe Korrelation im R^2 sowie ein niedriger rMSE zwischen gemessenen und prädizierten Werten (68) (**Tabelle 4**).

In einer weiteren Untersuchung wurde die Bildorientierung zweier Probanden mit unterschiedlichen WTA angepasst. Dafür wurde der Datensatz des Testprobanden (Nr. 208) rotiert, sodass dessen WTA dem des Trainingsprobanden (Nr. 210) entsprach. Durch die Rotation konnte eine deutliche Verbesserung in qualitativer (**Abbildung 18**) als

auch quantitativer Vorhersage der Güteparameter (**Tabelle 4**) von Atemfluss und Lungenvolumen erzielt werden (68).

Des Weiteren wurde untersucht, wie sich die Kombination aus mehreren Probanden (Nr. 205 - WTA :36°, Nr. 207 - WTA: 32°, Nr. 212 - WTA: 29°) mit ähnlichen WTA in einem Trainingsdatensatz auf die prädiktive Güte auswirkt.

Im Vergleich zum Trainingsdatensatz mit nur einem Probanden (Nr. 205) wurde eine minimale Verbesserung (**Tabelle 4**) qualitativ und quantitativ für den vorhergesagten respiratorischen Fluss erkenntlich (68). Die Kurvenübereinstimmung zeigte ein geringeres Rauschverhalten mit verbesserten Spitzenabgleich zwischen dem vorhergesagten und gemessenen Atemflusses (**Abbildung 17 m**). Quantitativ konnte eine Abnahme der mittleren aMFD erkannt werden (**Tabelle 4**). Die Vorhersage des Lungenvolumens zeigte qualitativ (**Abbildung 17 m**) als auch quantitativ (**Tabelle 4**) ähnliche Ergebnisse wie beim Training mit einem Probanden, allerdings mit leicht verschlechtertem Spitzenabgleich zwischen gemessenen und vorhergesagten Kurven und einer minimalen Zunahme der mittleren aTVD.

Letztlich konnte eine reine Erhöhung der Anzahl der Trainingsprobanden ohne Beachtung der Bildorientierung zu keiner Besserung der prädiktiven Güte führen. Dafür wurde die Kombination von neun Probanden (Nr. 205-207, 209-214) als Trainingsatz verwendet und die Modellvorhersage analysiert. Sowohl qualitativ als auch quantitativ präsentierte sich eine prädiktive Verschlechterung im Vergleich zum Trainingsatz mit einem Probanden (Nr. 205), welcher basierend auf dem WTA ausgewählt wurde. Auch wenn die mittlere aMFD und der mittlere R^2 ähnliche quantitative Werte zeigten, wiesen die mittlere aTVD und der rMSE schlechtere Werte auf (**Tabelle 4**). In der Kurvenüberlagerung des vorhergesagten und gemessenen Flusses und des Lungenvolumens traten zudem stärkere Unstimmigkeiten in den Kurvenmaxima auf (**Abbildung 17 n**).

Trainingsdatensatz	Volumen			Fluss		
	R^2 $\pm SA_B$	$rMSE$ $\pm SA_B$	$aTVD$ [mL] $\mu \pm SA$	R^2 $\pm SA_B$	$rMSE$ $\pm SA_B$	$aMFD$ [mL/s] $\mu \pm SA$
210 (unterschiedliche Winkel)	0,55 \pm 0,003	0,370 \pm 0,184	311 \pm 148	-0,08 \pm 0,01	0,450 \pm 0,222	518 \pm 338
210 (ähnliche Winkel nach Rotation der Testbilder (208))	0,89 \pm 0,002	0,038 \pm 0,019	133 \pm 92	0,89 \pm 0,002	0,112 \pm 0,055	247 \pm 257
205 (ähnliche Winkel)	0,92 \pm 0,002	0,027 \pm 0,023	55 \pm 59	0,92 \pm 0,002	0,077 \pm 0,038	253 \pm 211
205, 207, 212 (ähnliche Winkel)	0,91 \pm 0,002	0,030 \pm 0,015	65 \pm 62	0,92 \pm 0,001	0,079 \pm 0,039	228 \pm 218
205-207, 209-214	0,92 \pm 0,002	0,038 \pm 0,019	126 \pm 91	0,90 \pm 0,002	0,100 \pm 0,052	217 \pm 232

Tabelle 4: Deskriptive statistische Daten zur vorhergesagten Ateminformation für einen Probanden 208 mit Trainingsdatensätzen von unbekanntem Probanden. R^2 , Determinationskoeffizient, $rMSE$, relative mittlere quadratischer Fehler und SA_B , Standardabweichung durch Bootstrapping von R^2 und $rMSE$. μ , Mittelwert und SA , Standardabweichung von $aTVD$, absolute Tidalvolumendifferenz in mL and $aMFD$, absolute maximale Atemflussdifferenz in mL/s für alle Atemzüge innerhalb einer kardialen Kurzachsen-Volumetrie (19 Schichten mit je sechs Atemzügen pro Schicht) von Proband 208. (Modifizierte Tabelle aus dem Supplement von Malik et al. (68))

Probandennummer	Gewicht [kg]	Größe [cm]
205	65	178
206	93	185
207	54	157
208	69	178
209	54	164
210	60	174
211	65	166
212	63	184
213	68	170
214	85	175

Tabelle 5: Körpermaße. Maße der Probanden. *Gewicht* in kg und *Größe* in cm.

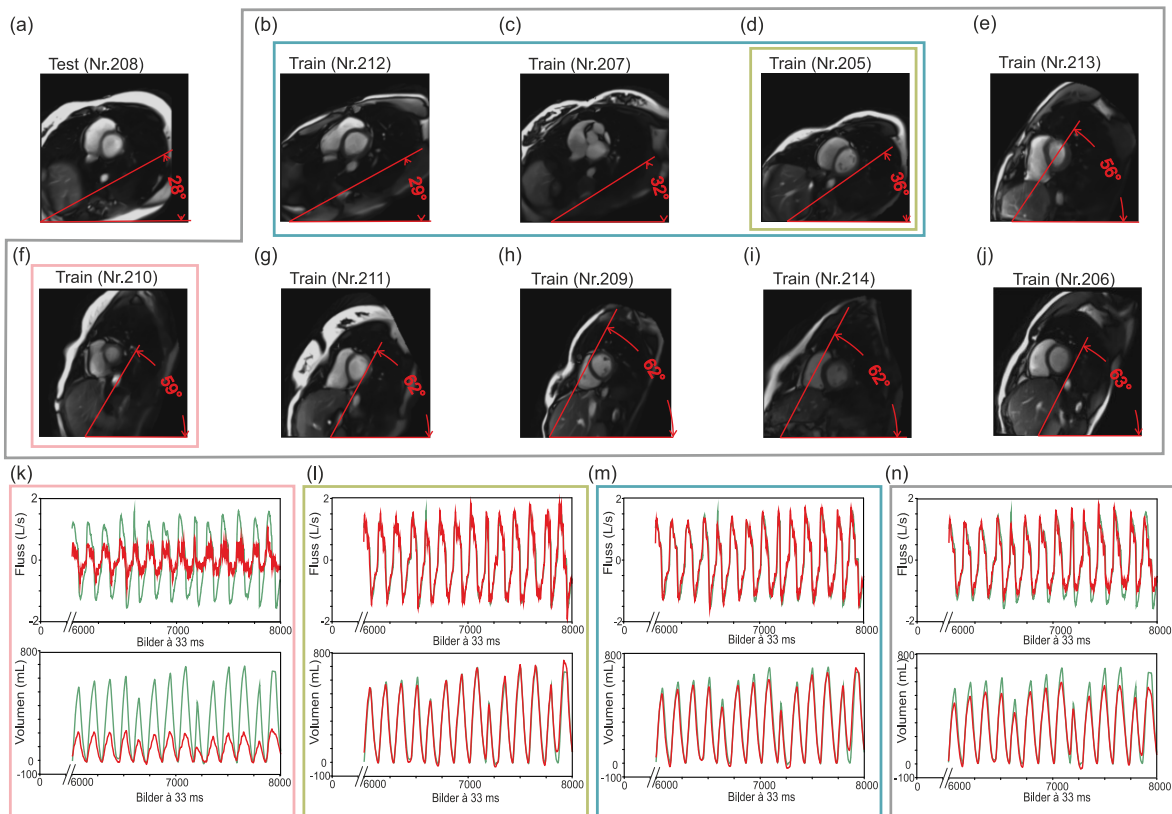


Abb. 17: Unbekannter Proband. Qualitative Evaluation der Vorhersage von respiratorischen Daten eines Testprobanden (Nr. 208) (a), der nicht im Trainingsdatensatz (b-j) enthalten war. Gemessene Daten (grüne Linien) und vorhergesagte Daten (rote Linien) für Atemfluss (vorletzte Reihe) und Lungenvolumen (letzte Reihe). Die Vorhersage des Testprobanden ist abhängig des Winkels der zwischen unterer Bildkante und der anterioren Thoraxwand im Kurzachsenschnitt (WTA) (rote Winkelangabe in (a-j)) vorhanden ist. Ein Trainingsproband mit unterschiedlichem WTA führt zu schlechter Vorhersage (k, pinke Umrandung), ein Trainingsproband (l, grüne Umrandung) oder mehrere Probanden (m, türkise Umrandung) mit ähnlichen WTA führen zu einer guten Vorhersage, alle Probanden ohne Beachtung des WTA (n, graue Umrandung) führen zu keiner Besserung der Vorhersage. (Modifizierte Grafik aus Malik et al. (68))

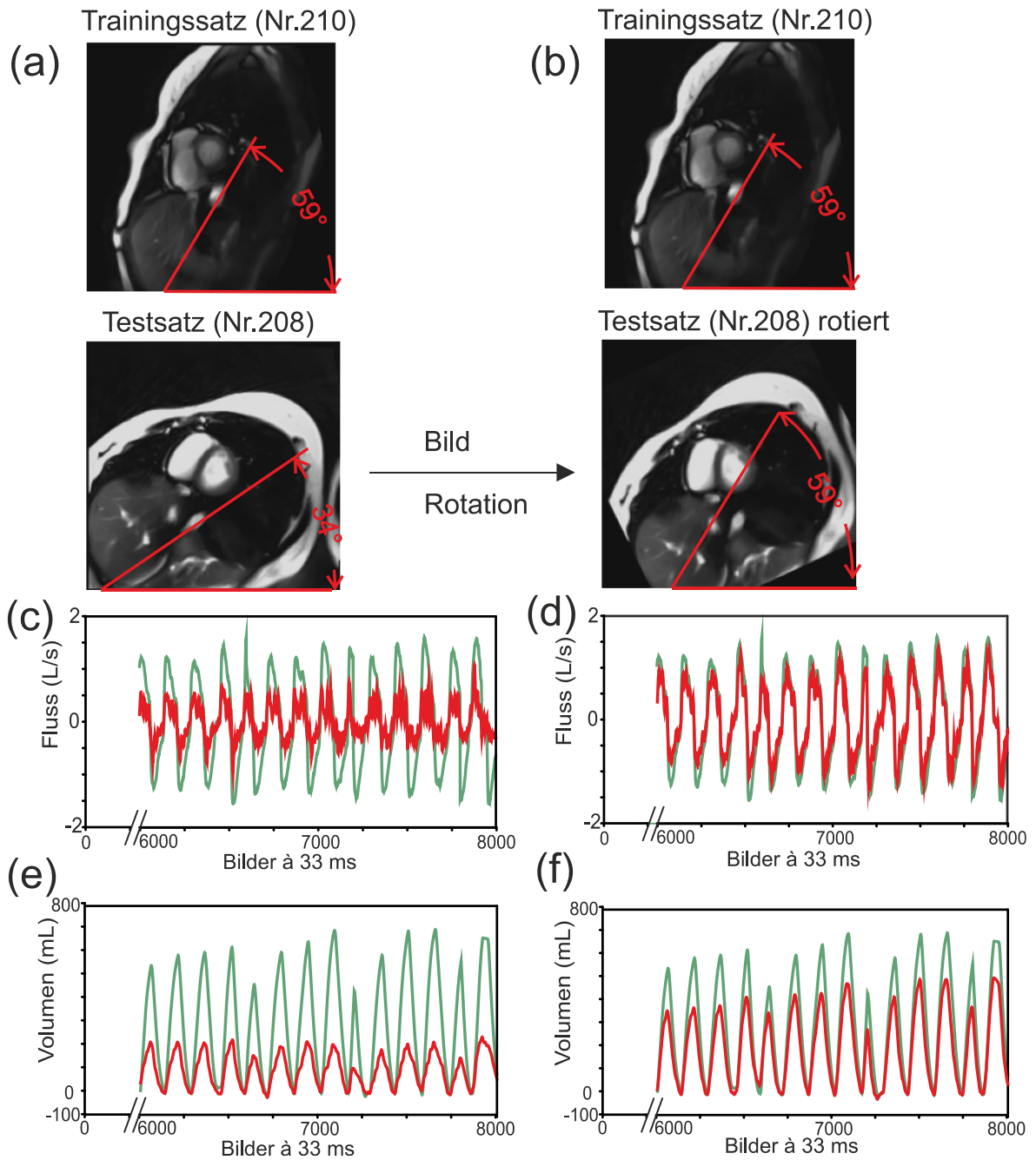


Abb. 18: Rotation des Testdatensatzes. Anpassung des Winkels der thorakalen Achse eines Testdatensatzes (Nr. 208, zweite Reihe von oben) an den des Trainingsdatensatzes (Nr. 205, obere Reihe) durch Rotation. Originaler Testdatensatz (a), rotierter Testdatensatz (b). Vorhergesagter Atemfluss (vorletzte Reihe) vor Rotation (c, rot) nach Rotation (d, rot) und vorhergesagtes Lungenvolumen (letzte Reihe) vor Rotation (e, rot) und nach Rotation (f, rot). Unterlegt ist der jeweilige gemessene Atemfluss (grün) und das gemessene Lungenvolumen (grün). (Modifizierte Grafik aus Supplement aus Malik et al. (68))

Es wurde geprüft, ob eine adaptive Denormalisierung die Vorhersage für unbekannte Probanden verbessern könnte. Dafür wurde die bisher verwendete Denormalisierung, basierend auf dem Mittelwert und der Standardabweichung des Trainingsdatensatzes, mit einer adaptierten Denormalisierung unter Verwendung des Mittelwertes und der Standardabweichung des Testdatensatzes verglichen. Die adaptive Denormalisierung führte weder zu einer eindeutigen Verbesserung noch zu einer eindeutigen Verschlechterung der Vorhersage (**Tabelle 6, Abbildung 19**) (68).

Trainings- satz	Volumen						Fluss					
	Denormalisierung mit μ und SA des Trainingsprobanden			Denormalisierung mit μ und SA des Testprobanden			Denormalisierung mit μ und SA des Trainingsprobanden			Denormalisierung mit μ und SA des Trainingsprobanden		
Probanden- nummer	R^2	rMSE	aTVD [mL] μ	R^2	rMSE	aTVD [mL] μ	R^2	rMSE	aMFD [mL/s] μ	R^2	rMSE	aMFD [mL/s] μ
205	0,92	0,027	55	0,92	0,034	64	0,92	0,077	253	0,90	0,081	291
206	0,92	0,028	67	-0,58	2,700	601	0,92	0,083	272	-6,81	1,580	1747
207	0,92	0,138	173	0,90	0,052	99	0,60	0,172	325	0,85	0,099	212
209	0,66	0,280	270	0,55	0,300	236	0,19	0,335	462	0,13	0,450	917
210	0,55	0,370	311	0,72	0,153	172	-0,08	0,450	518	0,56	0,284	228
211	0,82	0,117	181	0,18	0,864	420	0,66	0,180	278	-1,50	0,824	1249
212	0,92	0,031	78	0,40	0,594	362	0,91	0,078	219	-0,72	0,600	961
213	0,75	0,195	233	0,79	0,227	178	0,44	0,252	398	0,34	0,210	179
214	0,63	0,344	302	0,86	0,055	73	0,00	0,374	590	0,84	0,143	372

Tabelle 6: Adaptive Denormalisierung. Statistische Daten zur Vorhersage für einen Probanden (Nr. 208) welcher mit unterschiedlichen unbekanntem Probanden trainiert wurde (Nr. 205-207,209-214).

Zur Normalisierung der Trainingsdaten wurden der Mittelwert (μ) und die Standardabweichung (SA) des Atemflusses des Trainingsprobanden genutzt. Die Denormalisierung der prädizierten Testdaten wurde entweder durch μ und der SA des Atemflusses des Trainingsprobanden oder adaptiv durch μ und der SA des Atemflusses des Testprobanden.

R^2 , Determinationskoeffizient, $rMSE$, relativer mittlerer quadratischer Fehler, $aTVD [\mu]$, Mittelwert der absoluten Tidalvolumendifferenz in mL and $aMFD [\mu]$, Mittelwert der absoluten maximalen Atemflussdifferenz in mL/s. (Modifizierte Tabelle aus dem Supplement von Malik et al. (68))

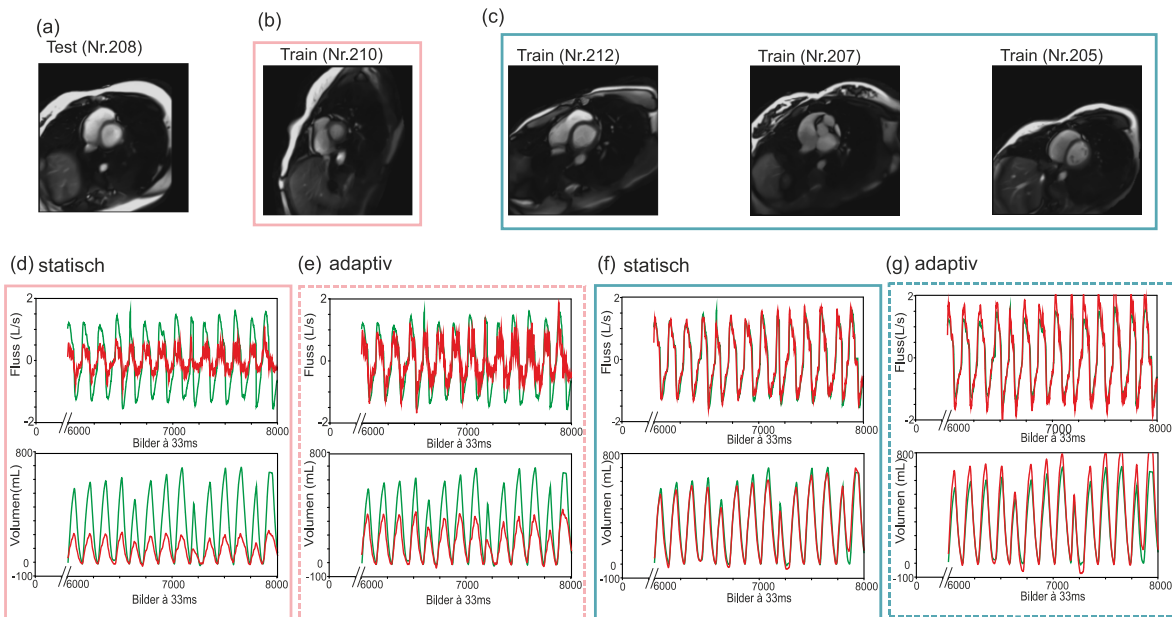


Abb.19: Adaptive Denormalisierung. Für einen Testprobanden (Nr. 208) (a) welcher mit unterschiedlichen unbekanntem Probanden trainiert wurde (Nr. 210 (b, pinke Umrandung) oder Nr. 212, Nr. 207 und Nr. 205 (c, blaue Umrandung)) wurde eine statische und adaptive Denormalisierung durchgeführt. Denormalisierung statisch: Mittelwert und Standardabweichung des Trainingsprobanden (d, pinke Umrandung) / der Trainingsprobanden (f, blaue Umrandung). Denormalisierung adaptiv: Mittelwert und die Standardabweichung des Testprobanden (e, pinke Umrandung, g, blaue Umrandung). Gemessene (grüne) und vorhergesagte (rote) Atemfluss- (vorletzte Reihe) und Lungenvolumenkurven (letzte Reihe). (Eigene Grafik)

3.6.2 Testdatensatz - zeitlich versetzte Aufnahme

Die Wiederverwendung eines trainierten Modells auf einem zeitlich versetzt aufgenommenen zweiten Testdatensatz desselben Probanden präsentierte gute Ergebnisse. Es konnten gute Übereinstimmungen der vorhergesagten und gemessenen respiratorischen Daten gefunden werden (68).

Qualitativ konnten in den basalen Schichten die Maxima von Atemfluss- und Lungenvolumenkurven gut vorhergesagt werden (**Abbildung 20 d**). Mit zunehmender Schichttiefe verschlechterte sich die Vorhersage, sodass die Spitzen der Maxima von Lungenvolumen als auch Atemfluss in den apikalen Schichten nur unzureichend abgebildet wurden (**Abbildung 20 e, f**). Die quantitative Analyse zeigte, dass die Vorhersagegüte sich hinsichtlich der mittleren aMFD, der mittleren aTVD, des mittleren rMSE und des mittleren R^2 im zweiten Testdatensatz verschlechterte (**Tabelle 7**). Die Verteilung der Differenzen zwischen prädizierten und gemessenen aTV und aMF sind inhomogener als in der ersten Studie. Mit steigenden Mittelwerten der prädizierten und gemessenen aTV nahm auch die Differenz zwischen prädizierten und gemessenen aTV zu. Zusätzlich zeigte sich ein breiteres Konfidenzintervall für die Differenzen des aMF (**Abbildung 21 c, d**) (68).

Bei Vergleich der MR-Bilder der ersten Aufnahme konnte entlang zunehmender Schichttiefe ein größer werdender WTA festgestellt werden (**Abbildung 20 a, b, c**). In der zweiten Aufnahme war die Zunahme des WTA entlang der Schichttiefe nicht so groß wie in der ersten Studie (**Abbildung 20 d, e, f**). Daraus wurde abgeleitet, dass die ungleiche Winkelzunahme ein möglicher Grund für die Verschlechterung der Prädiktion entlang der kurzen Achse sein könnte. Aufgrund dieser Beobachtung wurde auch in diesem Experiment eine Anpassung der Bildorientierung durchgeführt. Dabei wurden die einzelnen Schichten des Testdatensatzes der zweiten Aufnahme rotiert, sodass sie annähernd der Bildorientierung der einzelnen Schichten des Trainingsdatensatzes der ersten Aufnahme entsprachen (**Abbildung 20 g, h, i**) (68).

Die qualitative Evaluation präsentierte eine Verbesserung der Kurvenübereinstimmung insbesondere in den apikalen Schichten (**Abbildung 20 g, h, i**). Quantitativ konnte eine Verbesserung des mittleren R^2 , des mittleren rMSE, der mittleren aTVD und der mittleren aMFD erzielt werden (**Tabelle 7**). In der kumulativen Darstellung aller Schichten in den BA-Figuren, konnte eine homogenere Verteilung der Differenzen des aMF und insbesondere der aTV beobachtet werden. Der Anstieg der Differenzen zwischen prädizierten und gemessenen aTV konnte deutlich reduziert und die 95%-KI konnten deutlich verkleinert werden (**Abbildung 21 e, f**) (68).

Trainingsatz	Volumen			Fluss		
	R^2 $\mu \pm SA$	$rMSE$ $\mu \pm SA$	$aTVD$ [mL] $\mu \pm SA$	R^2 $\mu \pm SA$	$rMSE$ $\mu \pm SA$	$aMFD$ [mL/s] $\mu \pm SA$
Erste	0,995±0,004	0,002±0,001	12±10	0,99±0,002	0,011±0,003	65±54
Zweite	0,76±0,20	0,085±0,072	148±81	0,70±0,37	0,115±0,068	178±139
Zweite (rotiert)	0,89±0,09	0,039±0,031	98±54	0,89± 0,09	0,065±0,040	111±84

Tabelle 7: Deskriptive statistische Daten zur vorhergesagten Ateminformation für einen Probanden 205 mit zwei zeitlich versetzt aufgenommenen Datensätzen.

Die erste Aufnahme wurde in einen Trainings- und Testdatensatz geteilt. Die zweite Aufnahme fungierte als reiner Testdatensatz. Der zweite Testdatensatz wurde zweifach getestet, ohne und mit Rotation. Letztere wurde durchgeführt, um der Bildorientierung der ersten Aufnahme zu entsprechen. μ , Mittelwert, SA , Standardabweichung von R^2 , Determinationskoeffizient, $rMSE$, relativer mittlerer quadratischer Fehler, $aTVD$, absolute Tidalvolumendifferenz in mL and $aMFD$, absolute maximale Atemflussdifferenz in mL/s für die kardiale Volumetrie (19 Schichten mit 4-6 Atemzügen pro Schicht) von Proband 205. (Modifizierte Tabelle aus dem Supplement von Malik et al. (68))

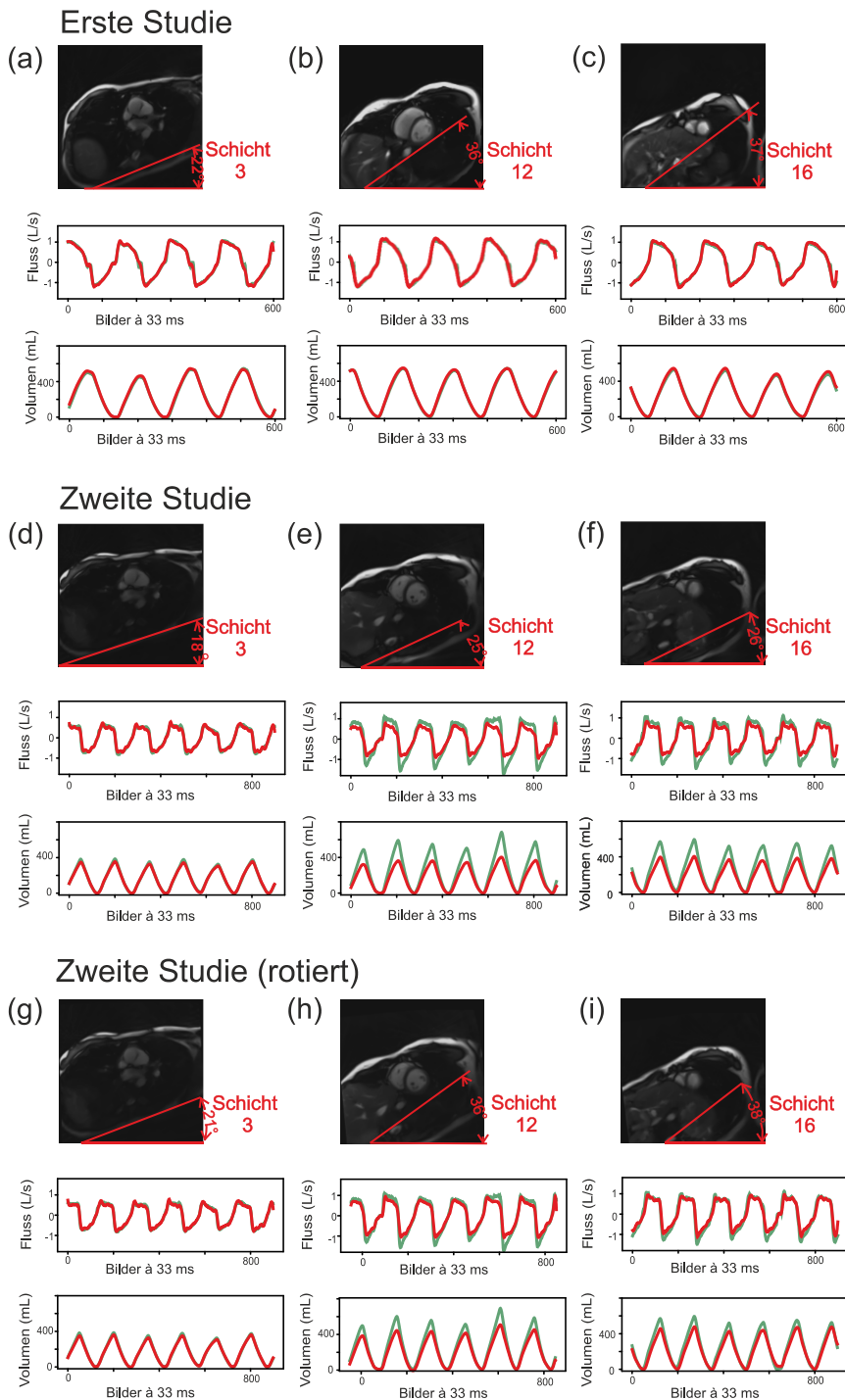
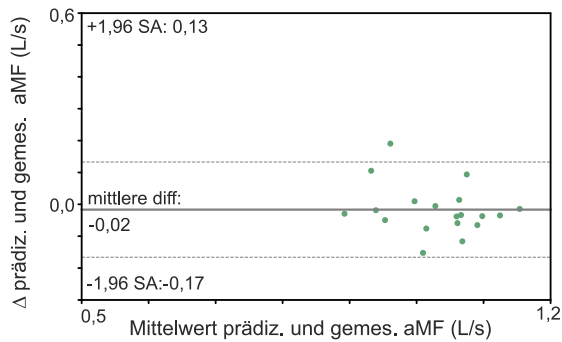
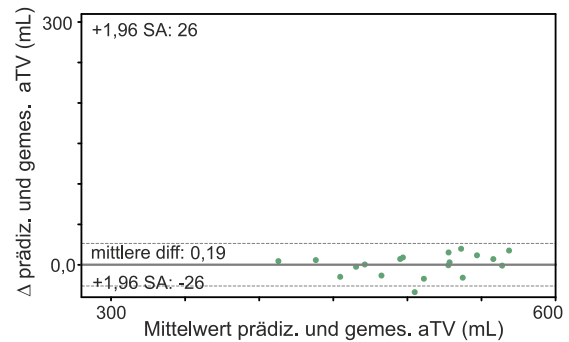


Abb. 20: Test-Retest. Vergleich der Vorhersagen eines trainierten Modells eines Probanden für den zeitlich versetzt zwei Testdatensätze (erste Studie (basal (a), mittventrikulär (b), apikal (c)), zweite Studie (basal (d), mittventrikulär (e), apikal (f)), aufgenommen wurde. Der Trainingsdatensatz (insgesamt 10 Schichten, mit jeweils 300 Bildern) entstammt aus der ersten Studie. Die zweite Studie wurde in weiterer Instanz rotiert, um dem Winkel zwischen unterer Bildkante und der anterioren Thoraxwand im Kurzachsenschnitt (WTA) des Trainingsdatensatzes zu entsprechen ((basal (g), mittventrikulär (h), apikal (j))). Die dazugehörigen vorhergesagten (rot) und gemessenen (grün) Atemflüsse (je Abschnitt, obere Reihe) und Lungenvolumen (je Abschnitt, untere Reihe). (Modifizierte Grafik aus Malik et al. (68))

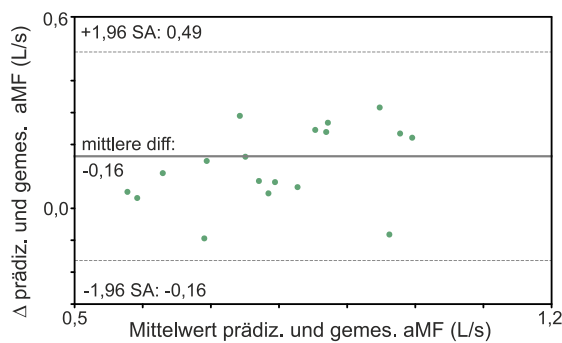
(a) Erste Studie



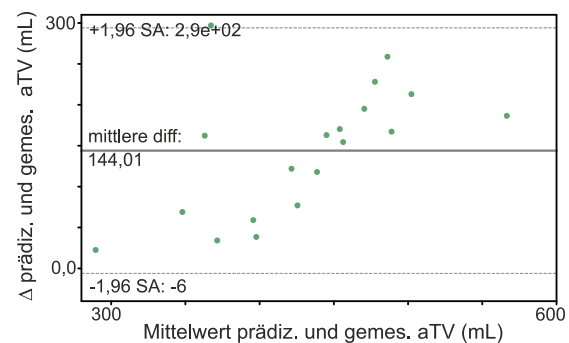
(b)



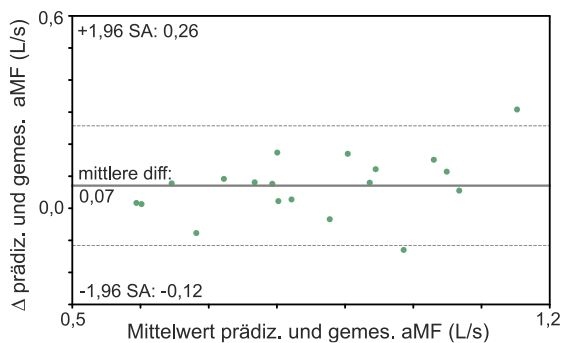
(c) Zweite Studie



(d)



(e) Zweite Studie (rotiert)



(f)

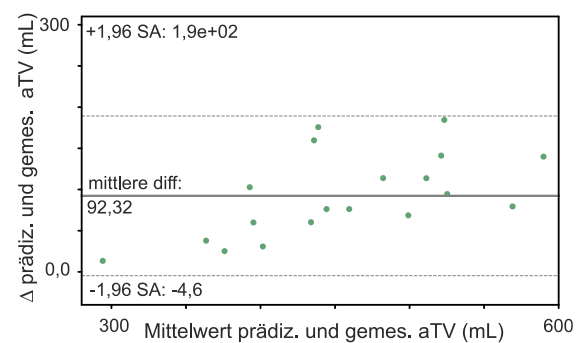


Abb. 21: Test-Retest Bland-Altman. Quantitative Evaluation für die Verwendung eines trainierten Modells (erste Studie) in zwei zeitlich verschiedenen Testdatensätze (erste Studie und zweite Studie) des Probanden Nr. 205 mittels Bland-Altman (BA) Grafiken. Für beide Aufnahmen wurde für jede Schicht eines Testdatensatzes ein randomisiertes Paar von Extrema für prädizierte (prädiz.) und gemessene (gemes.) Daten von absoluten maximalen Atemfluss (aMF) (a, c) und von absoluten Tidalvolumina (aTV) (b, d) selektiert und in einer BA-Grafik illustriert. Die BA-Grafik der zweiten Studie zeigte höhere mittlere Differenzen mit breiteren 95%- Konfidenzintervallen (KI) in aMF (b) und aTV (d). Zusätzlich konnte mit höheren Mittelwerten von prädizierten und gemessenen aTV eine größere Differenz in prädizierten und gemessenen aTV ausgemacht werden (d). Die BA-Grafik der zweiten Studie mit rotierter Bildorientierung zeigte kleinere mittlere Differenzen und geringere Breiten des KI in aMF(e) und aTV(f). (Modifizierte Grafik aus dem Supplement von Malik et al. (68))

3.7 Technische Aspekte

Aus technischer Perspektive wurden unterschiedliche Aspekte des Modells hinterleuchtet. Zum einen wurde geprüft, inwiefern eine Überanpassung die Prädiktion beeinflusst. Dafür wurde ein trainiertes Modell sowohl auf dem Trainingsdatensatz als auch auf dem zugehörigen Testdatensatz getestet. Es lag keine Überanpassung vor (**Abbildung 22**).

Im Weiteren wurden unterschiedliche Methoden zur Validierung untersucht. Dabei wurde der Hyperparameter λ , der Strafterm zur Bestimmung der Modellgewichte, automatisch mit Werten zwischen 0,01 und 1 ausgewählt. Diese Werte wurden mit manuellen Werten zwischen 0,001 bis 1000 verglichen. Die besten internen Modellbewertungen zeigten sich für den Atemfluss als auch für das Lungenvolumen im automatisch gewählten Bereich von $\lambda = 0,01$ und 1 (**Abbildung 23**).

Des Weiteren wurde die Kreuzvalidierung mittels *LOO-KV* durchgeführt. Ein Vergleich mit der *k-Fold-Kreuzvalidierung* ($k = 5$) wurde hinsichtlich der internen Modellwertung und der benötigten Trainingszeit durchgeführt. *LOO-KV* führte sowohl zu besseren internen Modellbewertungen als auch Trainingszeiten (**Tabelle 8**).

Abschließend wurde in der Nachbearbeitung die Wahl des Filters zur Beseitigung von Rauschartefakten im Atemfluss untersucht. Dabei wurden ein Savitzky-Golay Filter, ein Gauß Filter und ein Filter des gleitenden Mittelwertes verglichen. Am besten schnitt der Savitzky-Golay Filter ab: Hier konnten die Spitzen der vorhergesagten Kurven, im Vergleich zu den beiden anderen Filtern am genauesten erfasst werden, während Rauschartefakte gleichzeitig ausreichend beseitigt wurden (**Abbildung 24**).

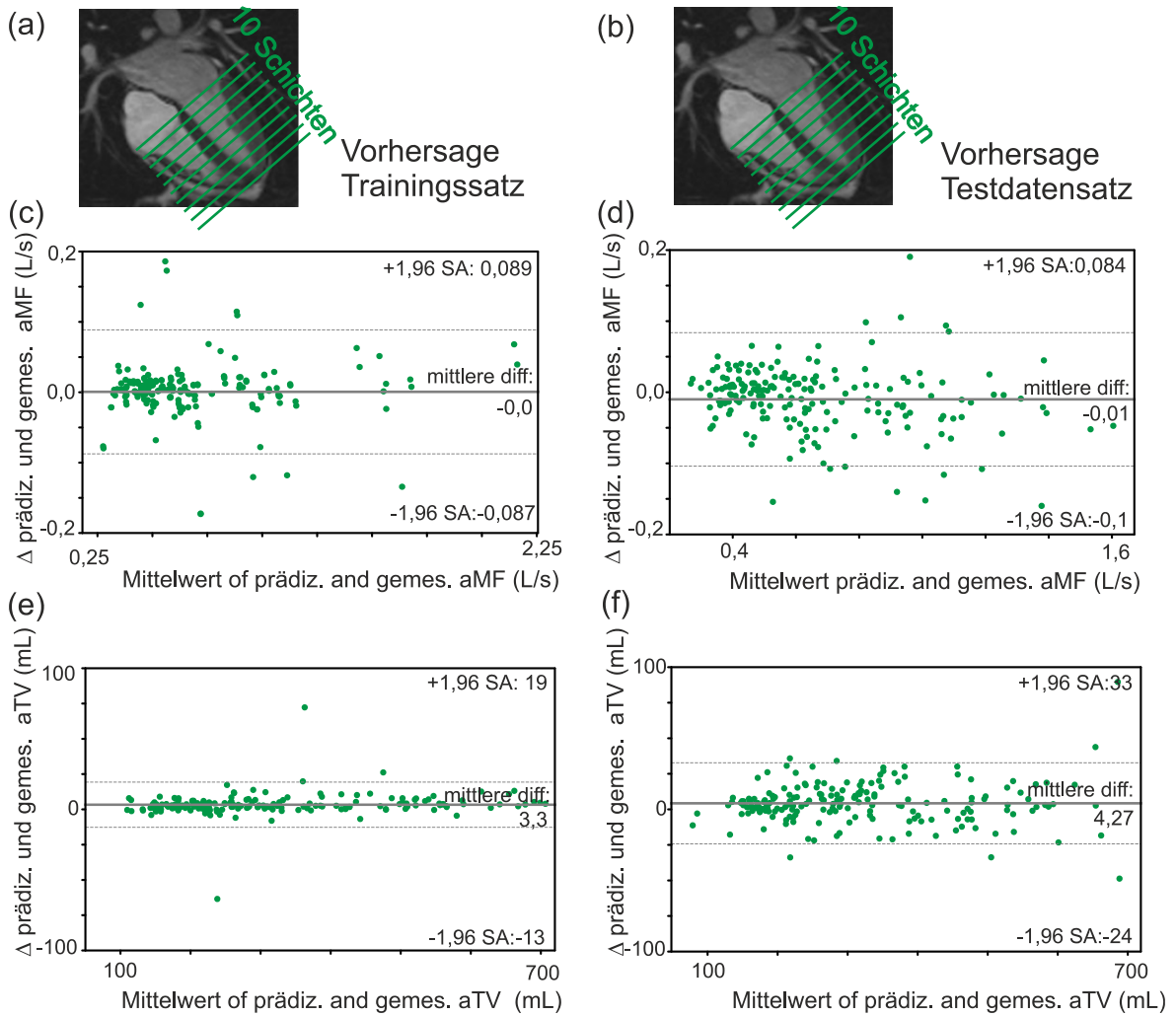


Abb. 22: Interne Modellvalidierung. Zur Überprüfung ob bei einem trainierten Modell eines Probanden (Nr. 205) eine Überanpassung vorlag. Das trainierte Modell wurde auf dem Trainingsdatensatz (linke Spalte) und dem Testdatensatz (rechte Spalte) getestet. Bland-Altman Grafik mit den Differenzen der randomisiert ausgewählten Paaren zwischen prädictierten und gemessenen absoluten Tidalvolumina (untere Reihe) und absoluten maximalen Flüssen (obere Reihe). Eine stärkere Streuung ist im Testdatensatz vorzufinden.

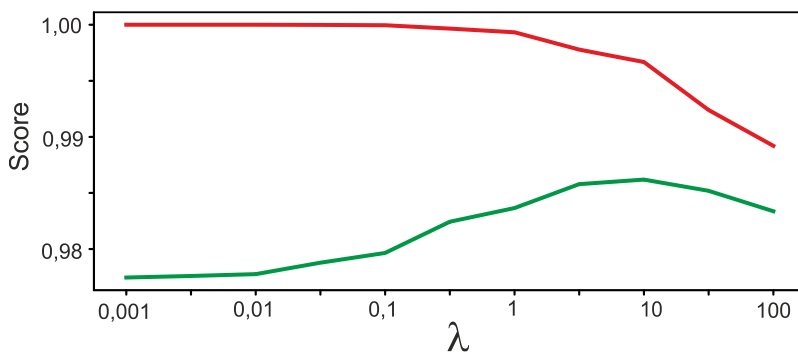


Abb. 23: Regularisierungsparameter λ : Interne Wertung (Score) eines trainierten Modells für Atemfluss (grün) und Lungenvolumen (rot) in Abhängigkeit des ausgewählten Regularisierungsparameter λ .

Trainingsatz Proband 205	Kreuzvalidierung					
Anzahl an Schichten	LOO-KV			k-Fold = 5		
	Zeit [s]	score		Zeit [s]	score	
		[Train] μ	[Test] μ		[Train] μ	[Test] μ
19	2	0,998	0,988	9	0,998	0,988
10	3	0,998	0,985	13	0,997	0,985
7	4	0,998	0,982	15	0,997	0,982
4	3	0,996	0,971	13	0,997	0,969

Tabelle 8: Evaluation der internen Modellwertung und der benötigten Rechenzeit für verschiedene Kreuzvalidierungstechniken. Für einen Probanden (Nr. 205) wurden unterschiedliche Anzahl an Schichten für das Training genutzt. Die Kreuzvalidierung fand mit *Leave-One-Out* (LOO-KV) oder *k-Fold* = 5 statt. Die interne Modellwertung (*score*) vor der Postprozessierung für den Trainingsdatensatz und den Testdatensatz und die Rechenzeit in Sekunden für das Training wurden präsentiert.

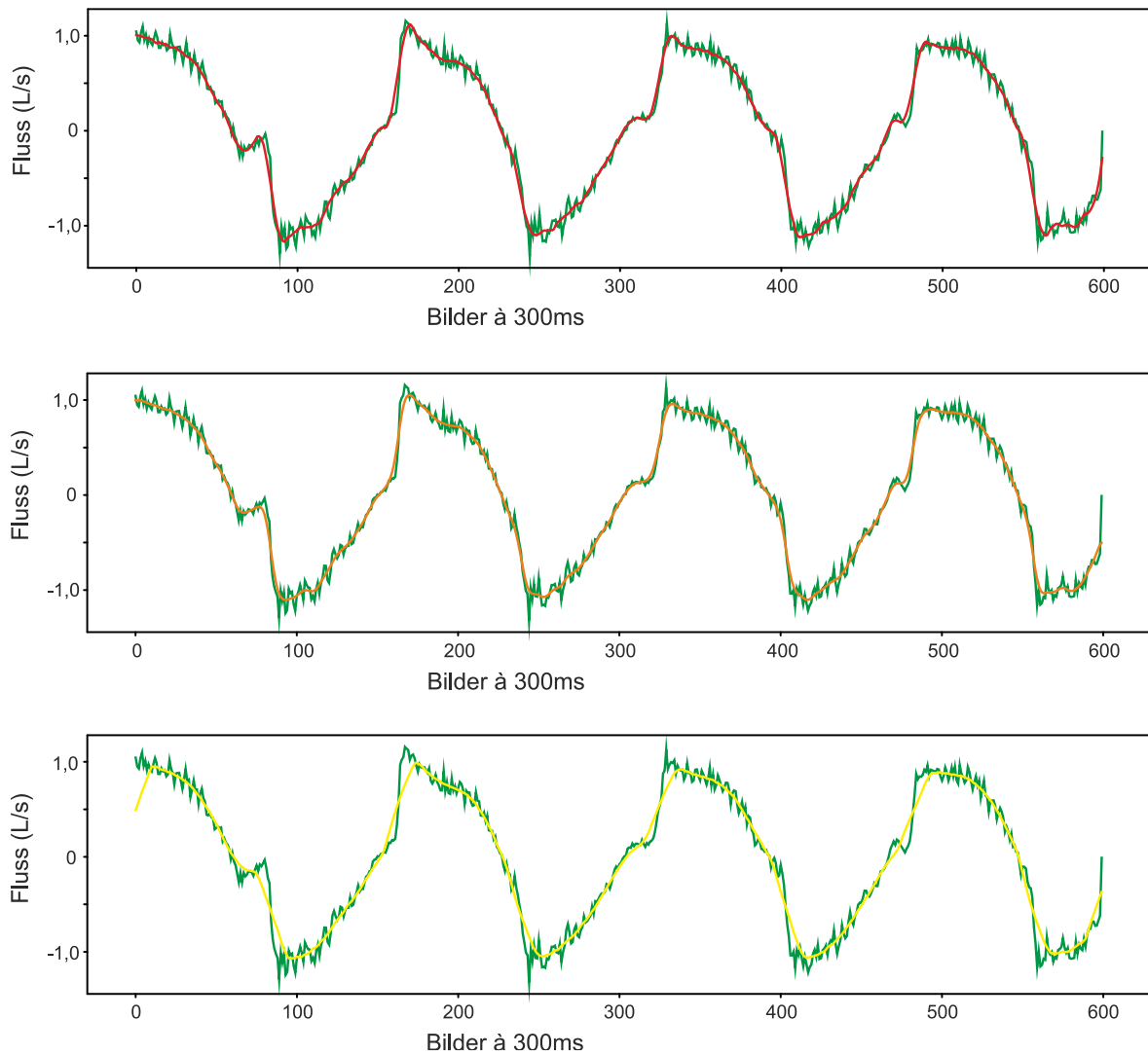


Abb. 24: Postprozessierung. Vorhersage des Atemflusses eines Probanden einer Schicht ohne Postprozessierung (grün). Filterauswahl zur Glättung des Signals (obere Zeile: Savitzky-Golay (rot), mittlere Zeile: Gauß (orange), untere Zeile: gleitender Mittelwert (gelb)).

4 Diskussion

Kardiale Echtzeit-MRT-Untersuchungen eröffnen neue Möglichkeiten zur Untersuchung herzkranker Patienten, insbesondere durch die umfassende Funktionsanalyse bei freier Atmung (79,80). Zur quantitativen Darstellung der Lungenfunktion ist die Spirometrie weiterhin der Goldstandard (43). Eine Kombination beider diagnostischer Verfahren bietet eine hervorragende Grundlage, die Herz-Lungen-Interaktion umfassend quantitativ zu analysieren. Dass die Verwendung eines MRkS während kardialer Echtzeit-MRT-Untersuchungen durchführbar und gut verträglich ist, wurde bereits gezeigt (41). Gerade bei kardial vorerkrankten Kindern spielt aber der Komfort während einer MRT-Untersuchung eine entscheidende Rolle. Da die enganliegende Maske eines MRkS bei längeren Messungen belastend sein kann (81), ist eine Kürzung der Zeit mit Spirometrie Maske wünschenswert (68).

Ziel dieser Arbeit war daher, ein ML-Modell zu entwickeln und zu validieren, welches nach einer kurzen Trainingsphase mit MRkS und einer Atemmaske, die Vorhersage quantitativer Lungenfunktionsparameter (Atemfluss und Lungenvolumen) aus nachfolgenden MR-Bildern einer kardialen KAV ermöglicht.

Zunächst wurde die Hypothese getestet, ob ein Modell, welches mit dem Atemfluss und den Echtzeit-MR-Bildern einer mittventrikulären Schicht eines bekannten Probanden trainiert wurde, aus nachfolgenden Echtzeit-MR-Bildern derselben Schicht Lungenfunktionsparameter vorhersagen kann. Die Hypothese konnte anhand unserer Qualitätskriterien bestätigt werden.

Die initiale Trainingsphase, die gute prädiktive Resultate lieferte, nutzte vier normale und vier tiefe Atemzüge dauerte aber 40 Sekunden pro Schicht. Die Übertragung dieser Trainingssequenz auf eine kompletten kardialen KAV wäre zu lang (760 Sekunden), da allein die Trainingsphase die Dauer der kompletten kardialen KAV (570 Sekunden) überschreiten würde (68). Nach Optimierung konnte, ohne einen bedeutsamen Verlust der Vorhersagegüte zu haben, die Trainingsphase auf zwei normale Atemzüge (ca. 300 Bilder, ca. 10 s) gekürzt werden. Ein Atemzug (ca. 150 Bilder, ca. 5 s) lieferte ebenso gute quantitative Ergebnisse. Die weitere Kürzung auf einen halben normalen Atemzug im Trainingsdatensatz ging mit erheblichen Verlusten in der Vorhersage einher. Diese Beobachtung erklärt sich dadurch, dass Expiration und Inspiration einer entgegengesetzten Mechanik folgen und somit die OFV einer Inspiration die fehlende Information der OFV der Expiration nicht kompensieren können. Obwohl 150 Bilder im Trainingsdatensatz eine ausreichende prädiktive Güte boten, wurden 300 MR-Bilder pro Schicht favorisiert. Der Grund dafür war, dass ein repräsentativer Trainingsdatensatz bei

automatischer Selektion der MR-Bilder mit mindestens einem Atemzug pro Schicht und ohne Verzerrung der Güte durch Atempausen gewährleistet werden sollte.

Ein Training mit normalen Atemzügen erzielte bessere Vorhersagen als ein Training mit tiefen Atemzügen. Ursächlich für dieses Ergebnis könnte dabei ein Informationsverlust sein, der bei tiefer schneller Atmung im Vergleich zu normaler Atmung von einem Bild zum nächsten Bild auftreten könnte. Relativ starke Bewegung können durch die Lucas-Kanade Methode die Bildung von schwachen OFV begünstigen (82). Zudem wies die untersuchende Testeinheit normale Atmung auf, weswegen ein Trainingsdatensatz mit demselben Atemmuster repräsentativer war (68).

Die zentrale Hypothese dieser Arbeit war, dass ein ML-Modell, welches mit MR-Bildern einer verkürzten kardialen KAV und den dazugehörigen Lungenfunktionsparameter trainiert wurde, aus nachfolgenden unbekanntem MR-Bildern derselben kompletten kardialen KAV, desselben Probanden, die dazugehörigen quantitativen Lungenfunktionsparameter vorhersagen kann. Diese Hypothese konnten anhand unserer Qualitätskriterien bestätigt werden.

Initial wurde das Modell mit 19 Schichten (Referenzmodell) mit jeweils 300 Bildern (insgesamt 190 Sekunden) und den dazugehörigen respiratorischen Daten trainiert und lieferte gute Prädiktionen.

Nach weiterer Optimierung konnte, ohne quantitativ hochwertige prädiktive Information über alle 19 Schichten zu verlieren, die Anzahl der selektierten Schichten im Training auf 10 Schichten mit jeweils 300 Bildern reduziert werden (Gesamtdauer 100 Sekunden). Bei einer Reduktion auf sieben (Gesamtdauer 70 Sekunden) oder vier Schichten (Gesamtdauer 40 Sekunden) wurde die prädiktive Güte unzureichend.

Beim Betrachten eines Trainingsdatensatzes aus jeder zweiten Schicht (gesamt 10 Schichten) konnte das beste Ergebnis erzielt werden, weil zwischen den benachbarten Trainingsschichten und der Testschicht der Abstand konstant klein blieb (maximal 8 mm). Trainingsdatensätzen aus jeder dritten Schicht (gesamt 7 Schichten, bis zu 16 mm Abstand) oder jeder sechsten Schicht (gesamt 4 Schichten, bis zu 40mm Abstand) wiesen deutlich größere und variablere Abstände auf, wodurch die prädiktive Güte verschlechtert wurde. Ursächlich für diese Beobachtung war, dass sich mit größer werdendem Abstand die MR-Bilder und damit die OFV der Trainings- und Testdatensätze stärker unterschieden. Kleinere variable Abstände waren nicht möglich, denn der gewählte Abstand zwischen den Schichten im MR-Protokoll musste aufgrund der technischen Rahmenbedingungen ein Vielfaches der Schichtdicke (8 mm) betragen und die erste und letzte Schicht mussten im Trainingsdatensatz enthalten sein.

Die Kürzung des Trainingsdatensatz auf ca. 1/6 des Gesamtdatensatzes führte trotz untypischer Aufteilung von Trainings- und Testdaten zu keiner Verschlechterung der Generalisierbarkeit. Zwar nutzt die Literatur häufig größere Trainingsdatensätze (z.B. 5/6 zu 1/6), jedoch hängt die Generalisierbarkeit und damit robuste Prädiktion ohne Unter- oder Überanpassung maßgeblich auch von der Komplexität des Modells ab (83). In unserem Modell konnte gezeigt werden, dass trotz sehr guter Determinationskoeffizienten und niedriger Trainingsfehler keine Überanpassung vorlag. Ursache dafür war der Einsatz von linearer Ridge Regression. Dieses Modell zeichnet sich durch seine geringe Komplexität und niedrige Varianz aus, wodurch auch kleine Trainingsdatensätze sehr gute Ergebnisse liefern können (84). Entscheidend war zudem, dass ein repräsentativer Trainingsdatensatz mit strenger Leave-One-Out Kreuzvalidierung sowie die automatische Selektion von Regularisierungsparameter zur besseren Abstimmung der unterliegenden Parameter vorlag (84,85).

Bei der Verteilung der Modellgewichte fiel auf, dass die bevorzugten Bildareale für das Modelltraining für alle Schichten nicht primär im Lungengewebe lokalisiert waren. Abhängig von der Schnittebene des Herzens wurden Areale im Bereich des Zwerchfells oder der äußeren Thoraxwand stärker gewichtet. Es ist denkbar, dass obwohl die MR-Signalintensität nach Bankier et al. (47) mit dem Lungenvolumen korreliert, diese im Vergleich zur Signalintensität durch die thorakale Wandbewegung oder die Zwerchfellbewegung zu niedrig und das Gewebe zu inhomogen (33,45) war, um stark gewichtete OFV zu bilden. Diese Annahme wird weiter unterstützt durch die Grundlagen der Lucas-Kanade Methode. Zum einen wird bei dieser Methode eine Konsistenz der Pixelintensität von einem zum anderen Bild benötigt und zum anderen hängen stabile optischen Flussvektoren von einer ausreichenden Textur der Region ab (82).

Unter der Annahme, dass Atemmechanik und die thorakale Anatomie von Mensch zu Mensch ähnlich sind und grundlegend durch Expansion des Thorax und dem Absenken des Zwerchfells bei Inspiration sowie der Umkehrung bei Expiration bestimmt werden, scheint die Übertragung von trainierten Modellen eines Probanden auf den nächsten sinnvoll.

Es wurde die Hypothese getestet, dass ein Modell, welches mit Echtzeit-MR-Bildern aus kardialen KAV und den dazugehörigen Lungenfunktionswerten von einzelnen oder mehreren Probanden trainiert wurde, quantitative Lungenfunktionsparameter aus unbekanntem Echtzeit-MR-Bildern einer kompletten kardialen KAV eines unbekanntem Probanden vorhersagen kann. Die Hypothese zeigte zwar qualitativ hochwertige Ergebnisse, konnte allerdings anhand der Gütekriterien quantitativ nicht bestätigt werden.

Des Weiteren wurden die gemeinsamen Faktoren, die zu qualitativ hochwertigen Ergebnissen führten, untersucht. Die Größe und das Gewicht der Probanden spielten eine

untergeordnete Rolle. Ähnliche Winkel zwischen unterer Bildkante und der anterioren Thoraxwand im Kurzachsenschnitt (WTA) im MR-Bild des Test- und Trainingsdatensatzes waren ein entscheidender Faktor für die Güte der Vorhersage. Wir konnten zeigen, dass durch Rotation und damit dem Angleichen der Bildausrichtung des Testdatensatzes an die des Trainingsdatensatzes, die prädiktive Güte sich deutlich verbesserte.

Die Grundlage des Modells bilden optische Flussvektoren, die die Pixelbewegung von einem zum nächsten Bild erfassen (69). Zuvor konnte bereits gezeigt werden, dass die Gewichtung der OFV um die Thoraxwand und das Zwerchfell am stärksten ausgeprägt war. Eine einheitliche Rotation der Bildausrichtung des gesamten Testdatensatz führte dazu, dass die Lage der Thoraxwand und des Zwerchfells besser mit der Lage des Trainingsdatensatzes übereinstimmte und somit in der Vorhersage stärker berücksichtigt werden konnten. Bei Verwendung der klassischen Lucas-Kanade Methode können große Rotationen innerhalb eines Datensatz zu systematischen Fehlern in der Bildung von OFV führen, da die Voraussetzung der lokalen Translokation verletzt wird (83). Die hier erwähnte Rotation wurde einheitlich im gesamten Datensatz durchgeführt, sodass die lokale Translokation von einem rotierten MR-Bild zum nächsten rotierten MR-Bild nicht verletzt wird. Trotzdem war keine ausreichende quantitative Vorhersage möglich. Gründe dafür können unter anderem sein, dass die Anatomie und Atmung zwischen Menschen zwar ähnlich aber nicht gleich sind. Besonders bei freier Atmung über einen längeren Zeitraum, selbst bei vorgegebener Atemfrequenz, ist eine natürliche Varianz in der Atemtiefe, im Atemfluss oder der Atemmechanik zu erwarten. Ob vermehrt aus dem Bauch oder eher flach geatmet wird oder ob Inspiration und Expiration gleichmäßig erfolgen, kann die Vorhersage zusätzlich beeinflussen.

Die Hypothese, dass ein Modell, welches mit respiratorischen Daten und den Echtzeit-MR-Bildern einer kompletten kardialen KAV eines bekannten Probanden trainiert wurde, aus Echtzeit-MR-Bildern einer zeitlich versetzt aufgenommenen kardialen KAV desselben Probanden quantitative respiratorische Daten vorhersagen kann, wurde untersucht. Die Hypothese konnte anhand der Gütekriterien nicht bestätigt werden, zeigte aber qualitativ bedeutsame Ergebnisse.

Bei genauer Betrachtung der Bildausrichtung in den Schichten von der ersten und der zweiten zeitlich versetzten Aufnahme des gleichen Probanden zeigten sich leichte Abweichungen, insbesondere im WTA entlang der Schichten. Dies deutete daraufhin, dass die Ausrichtung der kurzen Achse leicht verschoben war. Daher wurde der Testdatensatz schichtweise an den Trainingsdatensatz durch Rotation angepasst. Dadurch konnte eine deutliche Verbesserung der prädiktiven Güte erzielt werden. Die Rotation wurde kontinuierlich entlang der kurzen Achse durchgeführt, wodurch leicht verschiedene

Rotationswinkel zwischen den Schichten vorhanden waren. Der daraus entstehende Fehler ist jedoch vernachlässigbar. Denn zum einen unterstützt die klassische Lucas-Kanade Methode kleine Rotationen durch lineare Approximation (70) und zum anderen verwendet die angewandte dichte Lucas-Kanade Methode von Le Besnerais integriertes *Warping*, sodass der resultierende Fehler durch die Rotation praktisch vollständig kompensiert wird (71).

Eine Verbesserung mit ausreichender quantitativer Güte konnte vermutlich nicht erreicht werden, weil die Achse durch den kardialen kurzen Stapel anatomisch durch drei Freiheitsgrade händisch am MR-Scanner eingezeichnet wird. Einer Rotation, die nur auf einem zweidimensionalen Bild basiert, fehlt der dritte Freiheitsgrad. Perspektivisch wäre es sinnvoll, sofern technisch möglich, die Positionierung des Probanden und vor allem die genaue Einstellung der kurzen Achse zu speichern und von einer zur nächsten Aufnahme zu übernehmen.

Zur klinischen Anwendung des Modells sind vor allem Forschungsgebiete, die gezielt das quantitative kardiopulmonale Zusammenspiel untersuchen, von Interesse.

Röwer et al. untersuchten 2023 bei Kindern, die eine angeborene chronische Rechtsherzbelastung haben, den Einfluss der Atmung auf den rechten Ventrikel (42). Bislang wurde dies mit geschätzten Tidalvolumina aus Bildkorrelationen zum Zwerchfellstand durchgeführt. Die weitere Untersuchung des Herzens mit präziseren quantitativen Daten könnte die Aussagekraft der Ergebnisse verbessern. Da die Forschung an herzkranken Kindern besonders schonend erfolgen sollte, ist es vorteilhaft, Methoden zu verwenden, die möglichst belastungsarm sind. So kann bei notwendigen längeren Aufnahmen auf unsere Methode zurückgegriffen werden, und das Spirometer nach kurzer Zeit abgelegt werden. Der Komfort während der MR-Untersuchung würde dadurch maßgeblich verbessert werden (68).

Claessen et al. untersuchten 2014 bereits den Effekt der Atmung auf den rechten und linken Ventrikel während einer Echtzeit-MRT-Untersuchung und konnten zeigen, dass sich die ventrikulären Volumina atemabhängig verändern (87). Auch hier wurde zur Bestimmung der Atmung ein qualitatives Verfahren gewählt, bei dem ein Plethysmograph den Druck des Abdomens wahrgenommen hat. Eine quantitative Beurteilung der Atmung könnte hier die Forschung auch weiter verbessern.

Bei längeren Untersuchungszeiten eröffnet sich ein weiteres Anwendungsgebiet: die Untersuchung von Herz-Lungen-Interaktion unter Belastung mit Hilfe von MR-kompatibler Ergospirometrie (68). Dabei werden Gesamtbelastungszeiten von mindestens acht Minuten betrachtet (88). Bei Patientinnen mit chronisch-thromboembolischem pulmonalem

Bluthochdruck zeigen sich unter Belastung bei Inspiration Veränderungen der ventrikulären Füllung, der interventrikulären septalen Krümmung und des Schlagvolumens (89). Die belastungsabhängige unzureichende Leerung des rechten Vorhofs, dient als diagnostischer Marker von rechtsventrikulären Dysfunktionen (90). An diese Studien kann angeknüpft werden. Die Auswirkung von sportlicher Aktivität auf die Herz-Lungen-Interaktion ist ein wichtiges Forschungsgebiet, welches mit Hilfe von kardialer Echtzeit-MRTs in Kombination mit MR-kompatibler Ergospirometrie und der Vorhersage quantitativer respiratorischer Daten weiter untersucht werden könnte.

Neben der Analyse quantitativer respiratorischer Daten können diese Daten qualitativ zur Kompensation von Bewegungsartefakten durch respiratorisches *Gating* (24,26,27,28,29) in der Bildrekonstruktion von konventionellen kardialen MRTs verwendet werden. Die Anwendung bereits trainierter Modelle auf Testdaten unbekannter Probanden zeigte, dass die Unterteilung in Ein- und Ausatmung effizient vorhergesagt werden kann. Wenn also die qualitative Bemessung der Atmung ausreicht, wie beispielsweise bekannte Verfahren des *Atemmonitorings* (24,26,27,28,29), stellt dies eine neue ergänzende Methode dar (68).

Obwohl die Methode sowohl bei quantitativer oder qualitativer Bestimmung respiratorischer Parameter angewendet werden kann, gibt es wesentliche limitierende Faktoren. Der größte begrenzende Faktor ist, dass für quantitative Ergebnisse zwar die Zeit mit dem Spirometer deutlich verkürzt werden kann, ein Verzicht auf das MRkS derzeit nicht möglich ist (68). Die Güte des Modells hängt stark von der Güte der Bilder und ihrer Aufbereitung ab. Ohne Vorverarbeitung durch Selektion der Region würden Artefakte des Bildes in das Modell integriert und gewichtet werden. Ohne Nachbearbeitung würden rauschende Signale der Optical-Flow-Vektoren das eigentliche Signal verzerren. Für die Güte von trainierten Modellen, die auf unbekannte Probanden angewendet werden, kann eine Verbesserung durch Rotation der Bildebene des Testdatensatz erzielt werden. Eine reine Erhöhung der Probandenanzahl hat eine deutlich geringere Verbesserung als die Bildrotation (68). Denn die ausschlaggebende Limitierung wird durch unterschiedliche Anatomien, Ausrichtung und Größe der kurzen Achse und letztlich Unterschiede in der freien Atmung der Probanden ausgemacht. Der Aufbau des Experiments erfordert im Vergleich zum konventionellen kardialen MRT einen zusätzlichen Zeitaufwand. Letztlich kann eine Vorhersage nicht während der Echtzeit-MRT-Messung erfolgen. Bisläng wurde die Methode nur an Echtzeit-MR-Daten ausprobiert, ob eine Übertragung für konventionelle kardiale MRTs oder andere schnelle Sequenzen möglich ist, ist noch weiter zu prüfen (68).

5 Schlussfolgerung

Kardiale Echtzeit-MR-Bilder in Kombination mit MR-kompatibler Spirometrie bieten eine herausragende Möglichkeit Herz-Lungen-Interaktion quantitativ zu analysieren.

Im Rahmen dieser Arbeit konnte eine neue Methode entwickelt werden, mit der gezielt quantitative Ateminformation aus Echtzeit-MR-Bildern mittels maschinellen Lernens vorhergesagt werden kann. Dafür kamen Ridge Regression Modelle zum Einsatz, die mit kardialen Echtzeit-MR-Bildern und den zeitlich korrespondierenden respiratorischen Messdaten eines MRkS trainiert wurden.

Die Ergebnisse zeigten, dass sowohl quantitativ als auch qualitativ hochwertige respiratorische Daten aus ungesehenen kardialen Echtzeit-MR-Bildern einzelner mittventrikulärer Schichten sowie kompletter kardialer KAV desselben Probanden vorhergesagt werden können. Durch gezielte Optimierung des Modelltrainings konnte die benötigte Trainingsphase für die Vorhersage von kardialer KAV auf eine Dauer von 100 Sekunden reduziert werden. Dadurch konnte ein außerordentlich gutes Verhältnis zwischen Trainingszeit und Vorhersagegüte erreicht werden. Darüber hinaus konnten qualitativ robuste Vorhersagen auf Testdatensätzen von unbekanntem Probanden sowie auf zeitlich versetzt aufgenommenen Testdatensätzen eines bekannten Probanden erzielt werden.

Für die Zukunft bietet diese Methode vielfältige Anwendungsbereiche. Dazu gehören zum einen die quantitative Analyse der Herz-Lungen-Interaktion von kardial vorerkrankten Patienten, insbesondere bei Kindern, bei denen der Komfort durch Ablegen der MRkS-Maske gesteigert werden kann. Zum anderen ist die Methode sehr gut für den Einsatz im Rahmen von Herzbelastungsstudien mit MR-Ergospirometrie geeignet. Bei Nutzung trainierter Modelle zur Bereitstellung qualitativer Daten ganz ohne vorangehende Spirometrie könnte die Methode – ähnlich einem Bauchgurt – zur Bildstabilisierung oder als retrospektives *Gating*-Verfahren in der Bildakquise eingesetzt werden.

Literaturverzeichnis

1. Lindinger A, Schwedler G, Hense HW. Prevalence of congenital heart defects in newborns in Germany: Results of the first registration year of the PAN Study (July 2006 to June 2007). *Klin Padiatr.* 2010;222(5):321-326. doi:10.1055/s-0030-1254155
2. Rao PS. Advances in the Diagnosis and Management of Congenital Heart Disease in Children. *Children (Basel).* 2023;10(4):753. Published 2023 Apr 21. doi:10.3390/children10040753
3. Pignatelli RH, McMahon CJ, Chung T, Vick GW 3rd. Role of echocardiography versus MRI for the diagnosis of congenital heart disease. *Curr Opin Cardiol.* 2003;18(5):357-365. doi:10.1097/00001573-200309000-00006
4. Sganga D, Behera S, Beattie MJ, et al. Quality Improvement in a Pediatric Echocardiography Laboratory: A Collaborative Process. *Children (Basel).* 2022;9(12):1845. Published 2022 Nov 28. doi:10.3390/children9121845
5. Keenan NG, Pennell DJ. CMR of ventricular function. *Echocardiography.* 2007;24(2):185-193. doi:10.1111/j.1540-8175.2007.00375.x
6. Odegard KC, DiNardo JA, Tsai-Goodman B, Powell AJ, Geva T, Laussen PC. Anaesthesia considerations for cardiac MRI in infants and small children. *Paediatr Anaesth.* 2004;14(6):471-476. doi:10.1111/j.1460-9592.2004.01221.x
7. Dillman JR, Hernandez RJ. Role of CT in the evaluation of congenital cardiovascular disease in children. *AJR Am J Roentgenol.* 2009;192(5):1219-1231. doi:10.2214/AJR.09.2382
8. Röwer LM, Radke KL, Hußmann J, et al. Comparison of cardiac volumetry using real-time MRI during free-breathing with standard cine MRI during breath-hold in children. *Pediatr Radiol.* 2022;52(8):1462-1475. doi:10.1007/s00247-022-05327-5
9. Hartnell GG. Breathhold cardiac MRI and MRA. *Int J Card Imaging.* 1999;15(2):139-150. doi:10.1023/a:1006138713163
10. Ciet P, Tiddens HA, Wielopolski PA, et al. Magnetic resonance imaging in children: common problems and possible solutions for lung and airways imaging. *Pediatr Radiol.* 2015;45(13):1901-1915. doi:10.1007/s00247-015-3420-y
11. Zucker EJ. Compact pediatric cardiac magnetic resonance imaging protocols. *Pediatr Radiol.* 2023;53(7):1336-1351. doi:10.1007/s00247-022-05447-y
12. Heinrich S, Irouschek A, Ackermann A, et al. Magnetresonanztomografie bei Kindern in Sedierung--anästhesiologisches Prozedere, Zeitbedarf und Determinanten verzögerter Prozessabläufe [Anesthesiological time requirements and reasons for delayed processes in MRI examinations of children undergoing sedation or general anesthesia]. *Rofa.* 2013;185(5):467-473. doi:10.1055/s-0032-1330742

13. Berger D, Werner Moller P, Bachmann KF. Cardiopulmonary interactions-which monitoring tools to use?. *Front Physiol.* 2023;14:1234915. Published 2023 Aug 9. doi:10.3389/fphys.2023.1234915
14. Magder S. Heart-Lung interaction in spontaneous breathing subjects: the basics. *Ann Transl Med.* 2018;6(18):348. doi:10.21037/atm.2018.06.19
15. Lisanti CJ, Douglas DB. Effects of breath-hold and cardiac cycle on the MRI appearance of the aorta and inferior vena cava in t2 abdominal imaging. *AJR Am J Roentgenol.* 2009;192(5):1348-1358. doi:10.2214/AJR.08.1646
16. Ferrigno M, Ferretti G, Ellis A, et al. Cardiovascular changes during deep breath-hold dives in a pressure chamber. *J Appl Physiol (1985).* 1997;83(4):1282-1290. doi:10.1152/jappl.1997.83.4.1282
17. Sakuma H, Kawada N, Kubo H, et al. Effect of breath holding on blood flow measurement using fast velocity encoded cine MRI. *Magn Reson Med.* 2001;45(2):346-348. doi:10.1002/1522-2594(200102)45:2<346::aid-mrm1044>3.0.co;2-i
18. Ley S, Fink C, Puderbach M, et al. MRI Measurement of the hemodynamics of the pulmonary and systemic arterial circulation: influence of breathing maneuvers. *AJR Am J Roentgenol.* 2006;187(2):439-444. doi:10.2214/ajr.04.1738
19. Soni N, Williams P. Positive pressure ventilation: what is the real cost?. *Br J Anaesth.* 2008;101(4):446-457. doi:10.1093/bja/aen240
20. Ridgway JP. Cardiovascular magnetic resonance physics for clinicians: part I. *J Cardiovasc Magn Reson.* 2010;12(1):71. Published 2010 Nov 30. doi:10.1186/1532-429X-12-71
21. Menchón-Lara RM, Simmross-Wattenberg F, Casaseca-de-la-Higuera P, Martín-Fernández M, Alberola-López C. Reconstruction techniques for cardiac cine MRI. *Insights Imaging.* 2019;10(1):100. Published 2019 Sep 23. doi:10.1186/s13244-019-0754-2
22. Pfeiffer MP, Biederman RW. Cardiac MRI: A General Overview with Emphasis on Current Use and Indications. *Med Clin North Am.* 2015;99(4):849-861. doi:10.1016/j.mcna.2015.02.011
23. Sievers B, Addo M, Kirchberg S, et al. Impact of the ECG gating method on ventricular volumes and ejection fractions assessed by cardiovascular magnetic resonance imaging [published correction appears in J Cardiovasc Magn Reson. 2005;7(5):871-2]. *J Cardiovasc Magn Reson.* 2005;7(2):441-446. doi:10.1081/jcmr-200053515
24. Biglands JD, Radjenovic A, Ridgway JP. Cardiovascular magnetic resonance physics for clinicians: Part II. *J Cardiovasc Magn Reson.* 2012;14(1):66. Published 2012 Sep 20. doi:10.1186/1532-429X-14-66

25. Bertelsen A, Bernchou U, Schytte T, Brink C, Mahmood F. The effect of respiration-induced target motion on 3D magnetic resonance images used to guide radiotherapy. *Phys Imaging Radiat Oncol.* 2022;24:167-172. Published 2022 Nov 15. doi:10.1016/j.phro.2022.11.010
26. Ladrova M, Martinek R, Nedoma J, et al. Monitoring and Synchronization of Cardiac and Respiratory Traces in Magnetic Resonance Imaging: A Review. *IEEE Rev Biomed Eng.* 2022;15:200-221. doi:10.1109/RBME.2021.3055550
27. Choe KA, Smith RC, Wilkens K, Constable RT. Motion artifact in T2-weighted fast spin-echo images of the liver: effect on image contrast and reduction of artifact using respiratory triggering in normal volunteers. *J Magn Reson Imaging.* 1997;7(2):298-302. doi:10.1002/jmri.1880070207
28. McConnell MV, Khasgiwala VC, Savord BJ, et al. Comparison of respiratory suppression methods and navigator locations for MR coronary angiography. *AJR Am J Roentgenol.* 1997;168(5):1369-1375. doi:10.2214/ajr.168.5.9129447
29. Santelli C, Nezafat R, Goddu B, et al. Respiratory bellows revisited for motion compensation: preliminary experience for cardiovascular MR. *Magn Reson Med.* 2011;65(4):1097-1102. doi:10.1002/mrm.22687
30. Scott AD, Keegan J, Firmin DN. Motion in cardiovascular MR imaging. *Radiology.* 2009;250(2):331-351. doi:10.1148/radiol.2502071998
31. Holland AE, Goldfarb JW, Edelman RR. Diaphragmatic and cardiac motion during suspended breathing: preliminary experience and implications for breath-hold MR imaging. *Radiology.* 1998;209(2):483-489. doi:10.1148/radiology.209.2.9807578
32. Kolar P, Neuwirth J, Sanda J, et al. Analysis of diaphragm movement during tidal breathing and during its activation while breath holding using MRI synchronized with spirometry. *Physiol Res.* 2009;58(3):383-392. doi:10.33549/physiolres.931376
33. Puderbach M, Hintze C, Ley S, Eichinger M, Kauczor HU, Biederer J. MR imaging of the chest: a practical approach at 1.5T. *Eur J Radiol.* 2007;64(3):345-355. doi:10.1016/j.ejrad.2007.08.009
34. Ehman RL, McNamara MT, Pallack M, Hricak H, Higgins CB. Magnetic resonance imaging with respiratory gating: techniques and advantages. *AJR Am J Roentgenol.* 1984;143(6):1175-1182. doi:10.2214/ajr.143.6.1175
35. Gottwald LM, Blanken CPS, Tourais J, et al. Retrospective Camera-Based Respiratory Gating in Clinical Whole-Heart 4D Flow MRI. *J Magn Reson Imaging.* 2021;54(2):440-451. doi:10.1002/jmri.27564
36. Frahm J, Haase A, Matthaei D. Rapid NMR imaging of dynamic processes using the FLASH technique. *Magn Reson Med.* 1986;3(2):321-327. doi:10.1002/mrm.1910030217

37. Nayak KS, Lim Y, Campbell-Washburn AE, Steeden J. Real-Time Magnetic Resonance Imaging. *J Magn Reson Imaging*. 2022 Jan;55(1):81-99. doi: 10.1002/jmri.27411. Epub 2020 Dec 9. PMID: 33295674; PMCID: PMC8435094.
38. Bieri O, Scheffler K. Fundamentals of balanced steady state free precession MRI. *J Magn Reson Imaging*. 2013;38(1):2-11. doi:10.1002/jmri.24163
39. Uecker M, Zhang S, Voit D, Karaus A, Merboldt KD, Frahm J. Real-time MRI at a resolution of 20 ms. *NMR Biomed*. 2010;23(8):986-994. doi:10.1002/nbm.1585
40. Moratal D, Vallés-Luch A, Martí-Bonmatí L, Brummer M. k-Space tutorial: an MRI educational tool for a better understanding of k-space. *Biomed Imaging Interv J*. 2008;4(1):e15. doi:10.2349/bij.4.1.e15
41. Röwer LM, Uelwer T, Hußmann J, et al. Spirometry-based reconstruction of real-time cardiac MRI: Motion control and quantification of heart-lung interactions. *Magn Reson Med*. 2021;86(5):2692-2702. doi:10.1002/mrm.28892
42. Röwer LM, Radke KL, Hußmann J, et al. First experience with real-time magnetic resonance imaging-based investigation of respiratory influence on cardiac function in pediatric congenital heart disease patients with chronic right ventricular volume overload. *Pediatr Radiol*. 2023;53(13):2608-2621. doi:10.1007/s00247-023-05765-9
43. Schneider A, Gindner L, Tilemann L, et al. Diagnostic accuracy of spirometry in primary care. *BMC Pulm Med*. 2009;9:31. Published 2009 Jul 10. doi:10.1186/1471-2466-9-31
44. Akoumianaki E, Maggiore SM, Valenza F, et al. The application of esophageal pressure measurement in patients with respiratory failure. *Am J Respir Crit Care Med*. 2014;189(5):520-531. doi:10.1164/rccm.201312-2193CI
45. Wielpütz M, Kauczor HU. MRI of the lung: state of the art. *Diagn Interv Radiol*. 2012;18(4):344-353. doi:10.4261/1305-3825.DIR.5365-11.0
46. Voorhees A, An J, Berger KI, Goldring RM, Chen Q. Magnetic resonance imaging-based spirometry for regional assessment of pulmonary function. *Magn Reson Med*. 2005;54(5):1146-1154. doi:10.1002/mrm.20682
47. Boucneau T, Fernandez B, Larson P, Darrasse L, Maître X. 3D Magnetic Resonance Spirometry [published correction appears in *Sci Rep*. 2020 Sep 24;10(1):16012. doi: 10.1038/s41598-020-72363-2]. *Sci Rep*. 2020;10(1):9649. Published 2020 Jun 15. doi:10.1038/s41598-020-66202-7
48. Bankier AA, O'Donnell CR, Mai VM, et al. Impact of lung volume on MR signal intensity changes of the lung parenchyma. *J Magn Reson Imaging*. 2004;20(6):961-966. doi:10.1002/jmri.20198
49. Eichinger M, Puderbach M, Smith HJ, et al. Magnetic resonance-compatible-spirometry: principle, technical evaluation and application. *Eur Respir J*. 2007;30(5):972-979. doi:10.1183/09031936.00040607

50. Samuel A L. Some studies in machine learning using the game of checkers. *IBM Journal of research and development*. 1959; 3(3): 210-229.
51. Jordan MI, Mitchell TM. Machine learning: Trends, perspectives, and prospects. *Science*. 2015;349(6245):255-260. doi:10.1126/science.aaa8415
52. El Naqa I, Murphy MJ. What is machine learning?. In: El Naqa I, Murphy MJ, eds. *Machine Learning in Radiation Oncology: Theory and Applications*. Cham: Springer International Publishing; 2015:3-11.
53. Jabbar H, Khan R Z. Methods to avoid over-fitting and under-fitting in supervised machine learning (comparative study). *Computer Science, Communication and Instrumentation Devices*. 2015: 70(10.3850): 978-981.
54. Aliferis C, Simon G. Overfitting, underfitting and general model overconfidence and under-performance pitfalls and best practices in machine learning and AI. In: Simon GJ, Aliferis C, eds. *Artificial Intelligence and Machine Learning in Health Care and Medical Sciences*. Cham: Springer; 2024. doi:10.1007/978-3-031-39355-6_10
55. Kuhn M, Johnson K. Linear regression and its cousins. In: Kuhn M, Johnson K, eds. *Applied Predictive Modeling*. 1st ed. Springer Nature; 2013:101-139. doi:10.1007/978-1-4614-6849-3
56. James G, Witten D, Hastie T, Tibshirani R. *An Introduction to Statistical Learning: With Applications in R*. Vol 103. New York, NY: Springer; 2013.
57. Maulud D, Abdulazeez AM. A review on linear regression comprehensive in machine learning. *Journal of Applied Science and Technology Trends*. 2020;1(2):140-147. doi:10.38094/jastt1457
58. Saleh AKMdE, Arashi M, Kibria BG. Introduction to ridge regression. In: Saleh AKMdE, Arashi M, Kibria BG, eds. *Theory of Ridge Regression Estimation with Applications*. 1st ed. John Wiley & Sons; 2019:1-13. doi:10.1002/9781118644478.ch1
59. Saleh AKMdE, Arashi M, Kibria BG. Ridge regression in theory and applications. In: Saleh AKMdE, Arashi M, Kibria BG, eds. *Theory of Ridge Regression Estimation with Applications*. 1st ed. John Wiley & Sons; 2019:143-169. doi:10.1002/9781118644478.ch6
60. Shehab M, Abualigah L, Shambour Q, et al. Machine learning in medical applications: A review of state-of-the-art methods. *Comput Biol Med*. 2022;145:105458. doi:10.1016/j.combiomed.2022.105458
61. Deo RC. Machine Learning in Medicine. *Circulation*. 2015;132(20):1920-1930. doi:10.1161/CIRCULATIONAHA.115.001593
62. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42:60-88. doi:10.1016/j.media.2017.07.005

63. Brown G, Conway S, Ahmad M, et al. Role of artificial intelligence in defibrillators: a narrative review. *Open Heart*. 2022;9(2):e001976. doi:10.1136/openhrt-2022-001976
64. Knackstedt C, Bekkers SC, Schummers G, et al. Fully Automated Versus Standard Tracking of Left Ventricular Ejection Fraction and Longitudinal Strain: The FAST-EFs Multicenter Study. *J Am Coll Cardiol*. 2015;66(13):1456-1466. doi:10.1016/j.jacc.2015.07.052
65. Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. *Lancet Oncol*. 2019;20(5):e262-e273. doi:10.1016/S1470-2045(19)30149-4
66. Char DS, Shah NH, Magnus D. Implementing Machine Learning in Health Care - Addressing Ethical Challenges. *N Engl J Med*. 2018;378(11):981-983. doi:10.1056/NEJMp1714229
67. Giger ML. Machine Learning in Medical Imaging. *J Am Coll Radiol*. 2018;15(3 Pt B):512-520. doi:10.1016/j.jacr.2017.12.028
68. Malik H, Uelwer T, Röwer LM, et al. Machine-learning-based prediction of respiratory flow and lung volume from real-time cardiac MRI using MR-compatible spirometry. *Med Phys*. 2025;52(8):e18019. doi:10.1002/mp.18019
69. Horn BKP, Schunck BG. Determining optical flow. *Artif Intell*. 1981;17(1-3):185-203.
70. Lucas BD, Kanade T. An iterative image registration technique with an application to stereo vision. In: *Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI '81)*. San Francisco, CA, USA: Morgan Kaufmann Publishers; 1981:674-679.
71. Le Besnerais G, Champagnat F. Dense optical flow by iterative local window registration. In: *Proceedings of the IEEE International Conference on Image Processing*; September 2005:Vol 1, I-137. IEEE.
72. Van der Walt S, Schönberger JL, Nunez-Iglesias J, et al. scikit-image: image processing in Python. *PeerJ*. 2014;2:e453. doi:10.7717/peerj.453
73. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine learning in Python. *J Mach Learn Res*. 2011;12:2825-2830.
74. Savitzky A, Golay MJ. Smoothing and differentiation of data by simplified least squares procedures. *Anal Chem*. 1964;36(8):1627-1639.
75. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods*. 2020;17:261-272. doi:10.1038/s41592-019-0686-2
76. Seabold S, Perktold J. statsmodels: Econometric and statistical modeling with Python. In: *Proceedings of the 9th Python in Science Conference*; 2010.
77. Harris CR, Millman KJ, van der Walt SJ, et al. Array programming with NumPy. *Nature*. 2020;585:357-362. doi:10.1038/s41586-020-2649-2

78. Criée CP, Berdel D, Heise D, et al. Empfehlungen der Deutschen Atemwegsliga zur Spirometrie [Recommendations on spirometry by Deutsche Atemwegsliga]. *Pneumologie*. 2006;60(9):576-584. doi:10.1055/s-2006-944245
79. Contijoch F, Rasche V, Seiberlich N, Peters DC. The future of CMR: All-in-one vs. real-time CMR (Part 2). *J Cardiovasc Magn Reson*. 2024;26(1):100998. doi:10.1016/j.jocmr.2024.100998
80. Wang X, Uecker M, Feng L. Fast real-time cardiac MRI: a review of current techniques and future directions. *Investig Magn Reson Imaging*. 2021;25(4):252–265.
81. Casier A, Goubert L, Gebhardt WA, et al. Acceptance, well-being and goals in adolescents with chronic illness: a daily process analysis. *Psychol Health*. 2013;28(11):1337-1351. doi:10.1080/08870446.2013.809083
82. Chiba N, Kanade T. A tracker for broken and closely-spaced lines. *International Archives of Photogrammetry and Remote Sensing*. 1998;32:676-683.
83. Baker S, Matthews I. Lucas–Kanade 20 years on: A unifying framework. Part 1. *Int J Comput Vis*. 2004;56(3):221–255. doi:10.1023/B:VISI.0000011205.11775.fd
84. Sivakumar M, Parthasarathy S, Padmapriya T. Trade-off between training and testing ratio in machine learning for medical image processing. *PeerJ Comput. Sci*. 2024;10:e2245. doi:10.7717/peerj-cs.2245
85. Mehta P, Bukov M, Wang C-H, Day AGR, Richardson C, Fisher CK, Schwab DJ. A high-bias, low-variance introduction to Machine Learning for physicists. *Phys Rep*. 2019;810:1–124. doi:10.1016/j.physrep.2019.03.001
86. Ghojogh B, Crowley M. The theory behind overfitting, cross validation, regularization, bagging, and boosting: tutorial. *arXiv [preprint]*. 2019:arXiv:1905.12787.
87. Claessen G, Claus P, Delcroix M, Bogaert J, La Gerche A, Heidbuchel H. Interaction between respiration and right versus left ventricular volumes at rest and during exercise: a real-time cardiac magnetic resonance study. *Am J Physiol Heart Circ Physiol*. 2014;306(6):H816-H824. doi:10.1152/ajpheart.00752.2013
88. Herdy AH, Ritt LE, Stein R, et al. Cardiopulmonary Exercise Test: Background, Applicability and Interpretation. *Arq Bras Cardiol*. 2016;107(5):467-481. doi:10.5935/abc.20160171
89. Claeys M, Petit T, Bogaert J, et al. Dynamic aspects of ventricular interaction during exercise in HFpEF and in pre-capillary pulmonary hypertension. *ESC Heart Fail*. 2023;10(1):650-660. doi:10.1002/ehf2.14216
90. Schnell F, Claessen G, La Gerche A, et al. Atrial volume and function during exercise in health and disease. *J Cardiovasc Magn Reson*. 2017;19(1):104. Published 2017 Dec 18. doi:10.1186/s12968-017-0416-9

Danksagung

Als Erstes möchte ich von ganzem Herzen meinem Doktorvater Herrn Prof. Pillekamp danken, für die Mühe, die Geduld, die Nachsicht und die kontinuierliche Unterstützung. Auch wenn ich teils den Wald vor lauter Bäumen nicht erkannt habe, bin ich umso dankbarer, dass du den Wald nie aus den Augen verloren hast. Mit kleinen Schritten konnten wir so gemeinsam das Ziel erreichen. Ich bin an allen Problemen und den noch so vielen Hürden persönlich gewachsen und bin dafür sehr dankbar.

Als nächstes möchte ich Herrn Dr. Uelwer danken, für die kontinuierliche Unterstützung bei der Programmierung. Für meine unzähligen Fragen. Es war mir eine immense Freude so viel von dir lernen zu dürfen. Vielen Dank auch, dass du nach deiner Promotion noch weiter so tatkräftig mitgeholfen hast. Ohne dich wäre die Arbeit in diesem Sinne nicht möglich gewesen. Vielen Dank.

Als nächstes möchte ich Herrn Dr. Klee und Herrn Prof. Antoch danken, für die Möglichkeit das Echtzeit-MRT der Kinderradiologie benutzen zu dürfen. Vielen Dank für die Hilfe bei der logistischen Umsetzung meines Vorhabens.

Weiter möchte ich Lena Maria Röwer und Janina Hussmann danken - ihr seid unglaublich liebevolle Menschen. Vielen Dank, dass ich mit so viel Freude und so viel Hilfe die MRT-Aufnahmen mit euch durchführen konnte. Vielen Dank, dass ihr wirklich immer geholfen habt, wo ihr auch konntet. Vielen Dank, dass ihr mit mir endlose Stunden im MRT verbracht habt, um wieder eine Probandenaufnahme durchzuführen.

Vielen Dank an Herrn Dr. Harmeling, dafür dass Sie mir die Nutzung Ihres Rechners zur Verfügung gestellt haben, auch nachdem Sie die Universität gewechselt haben. Vielen Dank, dass Sie bei vielen neuen Denkanstößen geholfen haben.

Vielen Dank an Herrn Dr. Verde, dafür dass ich für die Statistik eine exzellente Beratung und bei allen Fragen fundierte und lehrreiche Antworten bekommen habe. Danke schön.

Vielen Dank an das gesamte Team der Radiologie, die mich bei Fragen rund ums MRT unterstützt haben.

Vielen Dank an Herrn Wittsack und Herrn Kieslich, für die hilfreichen Ideen zur logistischen Umsetzung unseres Vorhabens und der Organisation von technischem Zubehör.

Vielen Dank an all meine Freunde, meine Familie und vor allem an meine Mutter. Dafür, dass Sie mit mir durch den Schweiß und die Tränen durchgehalten haben. Mit aufmunternden Worten, mit hilfreichen Ratschlägen, mit andauerndem Zuspruch und kontinuierlicher Unterstützung. Ich werde euch für immer dankbar sein.