

Aus der Klinik für Gastroenterologie, gastrointestinale Onkologie und Endokrinologie
(Prof. Dr. med. V. Ellenrieder)
der Medizinischen Fakultät der Universität Göttingen

**Evaluation des hepato-renalen Index zur
nicht-invasiven Bestimmung des
Steatosegrades der Leber bei chronisch
Leberkranken mit der Histologie als
Referenzstandard**

INAUGURAL-DISSERTATION

zur Erlangung des Doktorgrades
der Medizinischen Fakultät der
Georg-August-Universität zu Göttingen

vorgelegt von

Julian Klemens Lasser

aus

Graz

Göttingen 2022

Dekan: Prof. Dr. W. Brück

Betreuungsausschuss

Betreuer/in: PD Dr. S. Kunsch

Ko-Betreuer/in: Prof. Dr. C. Ritter, MHBA

Prüfungskommission

Referent/in: PD Dr. S. Kunsch

Ko-Referent/in: PD Dr. Johannes Uhlig

Drittreferent/in:

Datum der mündlichen Prüfung: 04.08.2022

Hiermit erkläre ich, die Dissertation mit dem Titel " Evaluation des hepatorenenalen Index zur nicht-invasiven Bestimmung des Steatosegrades der Leber bei chronisch Leberkranken mit der Histologie als Referenzstandard " eigenständig angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Göttingen, den
.....
(Unterschrift)

Die Daten, auf denen die vorliegende Arbeit basiert, wurden teilweise publiziert:

Petzold G, **Lasser J**, Rühl J, Bremer SCB, Knoop RF, Ellenrieder V, Kunsch S, Neesse A (2020): Diagnostic accuracy of B-Mode ultrasound and Hepatorenal Index for graduation of hepatic steatosis in patients with chronic liver disease. PLOS ONE 15, e0231044

Inhaltsverzeichnis

Abbildungsverzeichnis	III
Tabellenverzeichnis	IV
Abkürzungsverzeichnis.....	V
1 Einleitung	1
1.1 Einteilung der Hepatosteatose	1
1.1.1 Alkoholische Fettlebererkrankung.....	1
1.1.2 Nicht alkoholische Fettlebererkrankung.....	2
1.1.3 Sekundäre Hepatosteatose und kryptogene Formen.....	3
1.2 Epidemiologie	4
1.2.1 Prävalenz der Hepatosteatose	4
1.2.2 Einflussfaktoren auf Hepatosteatose	6
1.2.3 Folgen der Hepatosteatose	8
1.3 Diagnostik.....	10
1.3.1 Invasive Diagnostik: Leberbiopsie.....	10
1.3.2 Nicht-invasive Diagnostik.....	13
1.4 Fragestellung	20
2 Material und Methoden	21
2.1 Primäres Auswahlkriterium: Transkutane Leberbiopsie	21
2.2 Sekundäres Auswahlkriterium: Ultraschallbilder	22
2.3 Bestimmung des hepatorenenalen Index.....	23
2.4 Erfassung der Interraterreliabilität.....	25
2.5 Erstellung der Datenbank	26
2.6 Statistische Auswertung.....	27
3 Ergebnisse.....	29
3.1 Studienkollektiv und Ablauf	29
3.2 Statistische Auswertung.....	32
3.2.1 Durchschnittsgröße der ROI-Rechtecke	32
3.2.2 Durchschnittlicher HRI und HRI-Verteilung nach Steatosegrad	32
3.2.3 Gruppenvergleich	34
3.2.4 Bestimmung von Cut-off-Werten.....	35
3.2.5 Subgruppe Personen ohne NAFLD	37
3.2.6 Einfluss von Leberfibrose auf die diagnostische Messgenauigkeit des HRI	40
3.2.7 Korrelation anderer Parameter mit dem histologisch bestimmten Steatosegrad.....	42
3.2.8 Bestimmung der Interraterreliabilität	43
4 Diskussion.....	44
5 Zusammenfassung.....	50

6	Literaturverzeichnis	52
----------	-----------------------------------	-----------

Abbildungsverzeichnis

Abbildung 1: Einteilung der Hepatosteatose und mögliche Progressionswege von NAFLD.....	2
Abbildung 2: Unterschiedliche Progressionswege von NAFLD zu HCC.	9
Abbildung 3: Diagnostischer Algorithmus bei NAFLD - Indikation zur Leberbiopsie.	10
Abbildung 4: Ultraschallbilder der Leber mit und ohne Leberverfettung.....	15
Abbildung 5: Auswahlkriterien des Studienkollektivs.	22
Abbildung 6: Ultraschallbild der Leber und der linken Niere mit Pixelstatistik und ROI-Rechtecken.	23
Abbildung 7: Platzierungskriterien der ROI-Rechtecke.....	24
Abbildung 8: Alternativplatzierungen der ROI-Rechtecke.....	25
Abbildung 9: Workflow Studiendesign.	29
Abbildung 10: Prozentuelle Verteilung der Lebererkrankungen im Patient*innenkollektiv.	30
Abbildung 11: Darstellung der Verteilung von HRI-Werten der Logiq-Kohorte und des Gesamtkollektivs als Box-Plots nach histologischem Steatosegrad.	33
Abbildung 12: ROC-Kurven für die HRI-Cut-off-Werte vom Gesamtkollektiv und der Logiq-Kohorte zur Detektion von $S \geq 1$, $S \geq 2$ und $S = 3$	35
Abbildung 13: Darstellung der Verteilung von HRI-Werten der Nicht-NAFLD-Kohorte und des Gesamtkollektivs als Box-Plots aufgeteilt nach histologischem Steatosegrad.	38
Abbildung 14: Darstellung der Verteilung von HRI-Werten der $F \leq 1$ - und der $F \geq 2$ -Kohorte als Box-Plots aufgeteilt nach histologischem Steatosegrad.....	41

Tabellenverzeichnis

Tabelle 1: Ursachen der sekundären Hepatosteatose	3
Tabelle 2: NAFLD-Prävalenz aufgeteilt nach Region	6
Tabelle 3: Geschlechterunterschied der NAFLD-Prävalenz in populationsbasierten Studien....	7
Tabelle 4: NAFLD Activity Score.	11
Tabelle 5: Fibroestaging nach Desmet und Scheuer	12
Tabelle 6: Diagnostische Genauigkeit von Ultraschall zur Hepatosteatosebestimmung.....	16
Tabelle 7: Studienvergleich der Messgenauigkeit des HRI zur Detektion einer Hepatosteatose	17
Tabelle 8: Diagnostische Genauigkeit des Controlled Attenuation Parameters (CAP)	18
Tabelle 9: Parameter des SteatoTest-1 und -2.....	19
Tabelle 10: Auflistung aller Parameter der Studiendatenbank.....	26
Tabelle 11: Epidemiologische Daten des Studienkollektivs	30
Tabelle 12: Verteilung der histologischen Steatose auf das Gesamtkollektiv und die Logiq- und Alokakohorte.....	31
Tabelle 13: Verteilung des HRI nach Steatosegrad und Studienkollektiv	32
Tabelle 14: Paarweiser Gruppenvergleich der HRI-Werte der jeweiligen histologischen Steatosegrade	34
Tabelle 15: Diagnostische Genauigkeit des HRI zur Prädiktion von Steatose im Gesamtkollektiv	36
Tabelle 16: Diagnostische Genauigkeit des HRI zur Prädiktion von Steatose in der Logiq- Kohorte	37
Tabelle 17: Verteilung der histologischen Steatose in der Nicht-NAFLD-Kohorte und im Gesamtkollektiv	37
Tabelle 18: Diagnostische Genauigkeit des HRI zur Prädiktion einer Steatose im Gesamtkollektiv und der Logiq- und Nicht-NAFLD-Kohorte.....	39
Tabelle 19: Verteilung histologischer Steatose bei Personen mit $F \geq 2$ und $F \leq 1$ im Vergleich zum Gesamtkollektiv	40
Tabelle 20: Diagnostische Genauigkeit des HRI zur Prädiktion einer Steatose im $F \leq 1$, $F \geq 2$ - und Gesamtkollektiv.....	42
Tabelle 21: Vergleich von Studien, die die Messgenauigkeit des HRI überprüften.....	46
Tabelle 22: Vergleich der diagnostischen Genauigkeit von visueller Einschätzung und HRI ...	48

Abkürzungsverzeichnis

AFLD	<i>Alcoholic Fatty Liver Disease</i> , Alkoholische Fettlebererkrankung
AIH	Autoimmunhepatitis
ALT	Alanin-Aminotransferase
AUROC	<i>Area Under Receiver Operating Characteristic</i> , Fläche unter der Grenzwertoptimierungskurve
BMI	<i>Body Mass Index</i>
CAP	<i>Controlled Attenuation Parameter</i>
CLD	<i>Chronic Liver Disease</i> , Chronische Lebererkrankung
CT	Computertomographie
HCC	Hepatozelluläres Karzinom
HRI	Hepatorenaler Index
H-MRSI	Protonenmagnetresonanzspektroskopie
ICC	<i>Intraclass Correlation Coefficient</i> , Intraklassenkorrelationskoeffizient
MRT	Magnetresonanztomographie
MRT-PDFF	<i>Proton Density Fat Fraction MRT</i>
NAFL	<i>Non-Alcoholic Fatty Liver</i> , Nicht alkoholische Fettleber
NAFLD	<i>Non-Alcoholic Fatty Liver Disease</i> , Nicht alkoholische Fettlebererkrankung
NASH	<i>Non-Alcoholic Steatohepatitis</i> , Nicht alkoholische Steatohepatitis
NPW	Negativ prädiktiver Wert
PPW	Positiv prädiktiver Wert
ROC-Kurve	<i>Receiver Operating Characteristic Curve</i> , Grenzwertoptimierungskurve
ROI	<i>Region Of Interest</i>
RR	Blutdruck nach Riva-Rocci
γ -GT	Gamma-Glutamyltranspeptidase

1 Einleitung

Diese Arbeit befasst sich mit der diagnostischen Genauigkeit des mittels *High-End*-Ultraschallgeräten ermittelten hepatorenenalen Index zur Detektion und Klassifikation von hepatischer Steatose bei Patient*innen mit chronischen Lebererkrankungen (CLD).

Hepatosteatoese wird als abnormale Akkumulation von Triglyceriden ($\geq 5\%$ der Hepatozyten) in der Leber definiert (Kleiner et al. 2005). Dem liegt mikroskopisch eine Einlagerung von Fettsäuren in die Hepatozyten zu Grunde (Tannapfel et al. 2011).

Hepatosteatoese gewinnt als Krankheitsbild immer mehr an Bedeutung. Dies beruht darauf, dass die Prävalenz stetig steigt (Estes et al. 2018; Younossi et al. 2018). In einer Übersichtsarbeit zur nicht alkoholischen Fettlebererkrankung vermuten Aurújo et al (2018), dass deren Prävalenz bei 24 – 25 % liegt. Damit gilt die nicht alkoholische Fettlebererkrankung als größte Untergruppe der CLD und nimmt somit einen hohen Stellenwert im klinischen Alltag ein (Adams et al. 2005; Younossi et al. 2018).

Diese Tatsache ist insofern bedeutsam, da sich eine Leberverfettung über eine Steatohepatitis zu einer Leberfibrose bis hin zu einem hepatozellulären Karzinom (HCC) entwickeln kann (Paul 2002; Anstee et al. 2013).

1.1 Einteilung der Hepatosteatoese

In den Amerikanischen und Deutschen Leitlinien zur Therapie der nicht alkoholischen Fettlebererkrankung wird die Hepatosteatoese in zwei Gruppen unterteilt: alkoholische Fettlebererkrankung (*Alcoholic Fatty Liver Disease*, AFLD) und nicht alkoholische Fettlebererkrankung (*Non-Alcoholic Fatty Liver Disease*, NAFLD). Davon abzugrenzen sind die sekundären Fettlebererkrankungen (Chalasanı et al. 2012; Roeb et al. 2015).

1.1.1 Alkoholische Fettlebererkrankung

Die alkoholische Fettlebererkrankung setzt einen chronischen Alkoholabusus als Ursache voraus. Der Schwellenwert, ab dem eine Fettleber durch chronischen Alkoholmissbrauch nicht mehr ausgeschlossen werden kann, wird von der Deutschen Leitlinie für die nicht alkoholische Fettlebererkrankung mit 20 g Alkohol pro Tag bei Männern und 10 g pro Tag bei Frauen definiert (Roeb et al. 2015).

In der amerikanischen Leitlinie wird der Schwellenwert anders definiert. Dieser wird in Drinks pro Woche angegeben. Pro Drink wird von einer Alkoholmenge von ca. 10 g ausgegangen. Für Männer gilt ein durchschnittlicher Alkoholkonsum von > 21 Drinks pro Woche und für Frauen > 14 Drinks als Grenzwert (Chalasanı et al. 2012).

1.1.2 Nicht alkoholische Fettlebererkrankung

Die nicht alkoholische Fettlebererkrankung wird als Leberverfettung definiert, die nicht auf einen signifikanten Alkoholkonsum oder sekundäre Ursachen zurückzuführen ist (Roeb et al. 2015).

Ihr untergeordnet sind die nicht alkoholische Fettleber (*Non-Alcoholic Fatty Liver*, NAFL), welche sich histologisch als blande, hauptsächlich makrovesikuläre Hepatosteatose darstellt und die nicht alkoholische Steatohepatitis (*Non-Alcoholic Steatohepatitis*, NASH), welche histologische Zeichen von Fibrosierung und nekroinflammatorischer Zellschädigung aufweist (Hashimoto et al. 2015). Im weiteren Krankheitsverlauf kann eine NASH über eine Leberzirrhose bis hin zu einem HCC fortschreiten (Anstee et al. 2013). Eine Übersicht über die Unterformen der Hepatosteatose und die Progressionswege einer NAFLD gibt Abbildung 1.

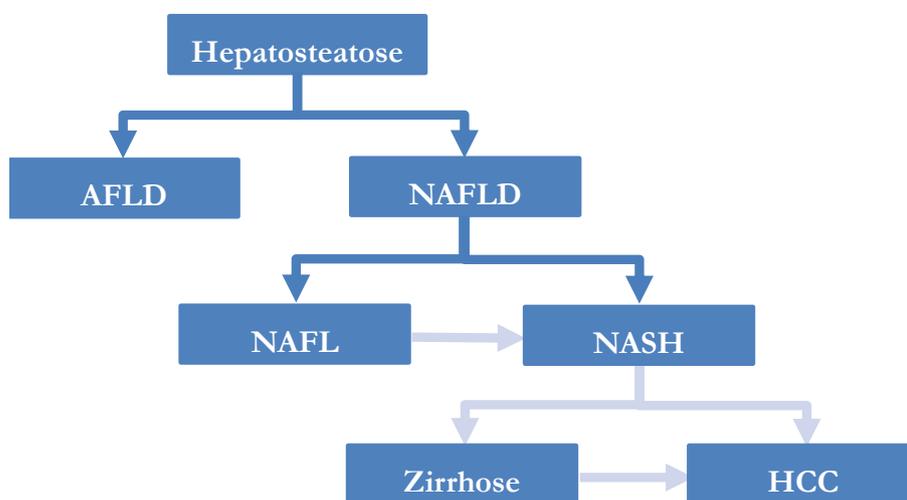


Abbildung 1: Einteilung der Hepatosteatose und mögliche Progressionswege von NAFLD. Die Einteilung zwischen einer nicht alkoholischer Fettlebererkrankung (NAFLD) und einer alkoholischer Fettlebererkrankung (AFLD) ist abhängig vom chronischem Alkoholkonsum. Eine NAFLD dient als Überbegriff für die nicht alkoholische Fettleber (NAFL) und die nicht alkoholische Steatohepatitis (NASH). Eine NAFL kann zu einer NASH, einer Leberzirrhose und einem hepatozellulären Karzinom (HCC) fortschreiten. Dunkelblaue Pfeile stellen die Subkategorien dar, hellblaue eine Progredienz (adaptiert nach Hashimoto et al. 2015; Calzadilla Bertot und Adams 2016). CC BY 4.0. Die Verwendung erfolgt mit freundlicher Genehmigung von John Wiley and Sons.

1.1.3 Sekundäre Hepatosteatose und kryptogene Formen

Als sekundäre Steatose gilt wiederum jegliche Hepatosteatose, die nicht durch einen chronischen Alkoholmissbrauch oder Teilkomponenten des metabolischen Syndroms bedingt ist, sondern durch andere Auslöser (Roeb et al. 2015). Eine Auflistung dieser Ursachen ist in Tabelle 1 zu sehen.

Tabelle 1: Ursachen der sekundären Hepatosteatose

Fettstoffwechselstörungen	Medikamenteninduziert
Abetalipoproteinämie	Amiodaron
Hypobetalipoproteinämie	Tamoxifen
Kombinierte familiäre Hyperlipidämie	HAART (highly active antiretroviral therapy)
Glukogenspeicherkrankheiten	Corticosteroide
Weber-Christian-Syndrom	Methotrexat
Lipodystrophie	Valproeate
LCAT-Störung	Mangelernährung
Cholesterinesterspeichererkrankung	Morbus Wilson
Morbus Wolman	Schäden durch Gifte in der Umwelt
Zöliakie	Parenterale Ernährung
Reye-Syndrom	Hepatitis C Infektion
Akute Schwangerschaftsfettleber	Starker Gewichtsverlust nach bariatrischen Operationen
HELLP Syndrom	

(basierend auf Chalasani et al. 2012; basierend auf Kneeman et al. 2012). Die Verwendung erfolgt mit freundlicher Genehmigung von John Wiley and Sons und SAGE Publications.

Lässt sich eine Hepatosteatose nicht sicher einer der oben genannten Ursachen zuordnen, fällt sie in die Gruppe der kryptogenen Formen (Roeb et al. 2015).

1.2 Epidemiologie

1.2.1 Prävalenz der Hepatosteatose

Zur Prävalenz der allgemeinen Hepatosteatose in der Normalbevölkerung gibt es nur wenige Publikationen. Deren Ergebnisse weisen eine starke Variabilität auf. Diese Schwankung liegt ebenso bei der besser erforschten NAFLD, einer Subgruppe der Hepatosteatose, vor und wird in der aktuellen S2k Leitlinien zur nicht alkoholischen Fettlebererkrankung diskutiert (Roeb et al. 2015). Da NAFLD die größte Untergruppe unter den Fettlebererkrankungen darstellt, kann davon ausgegangen werden, dass der starken Prävalenzvariabilität der allgemeinen Hepatosteatose die gleichen Ursachen wie der NAFLD zu Grunde liegen (Caruso et al. 2019).

Die Autoren der S2k Leitlinie führen das oben genannte Phänomen auf die unterschiedlichen diagnostischen Modalitäten der bisher durchgeführten Studien zurück. In den Studien wurde die Prävalenz in der Normalbevölkerung mit unterschiedlichen histopathologischen Protokollen, radiologischen Verfahren oder Biomarkern ermittelt. Diesem breiten Spektrum an Messverfahren liegt zu Grunde, dass der eigentliche Diagnosestandard zur Bestimmung einer NAFLD, die Leberbiopsie, aufgrund ihrer Invasivität, für populationsbasierte Studien nicht genutzt werden kann (Roeb et al. 2015). Die alternativen nicht-invasiven Diagnoseverfahren, welche den Großteil der populationsbasierten Studien ausmachen, gelten jedoch als weniger aussagekräftig und geben je nach angewandtem Verfahren eine unterschiedliche Prävalenz an (Vernon et al. 2011).

Die Prävalenz jener oben erwähnten Studien zur allgemeinen Hepatosteatose variierte zwischen 14 % und 58,3 %. Als Untersuchungsmethoden wurden Ultraschall (Nomura et al. 1988; Bellentani et al. 2000) und Protonenmagnetresonanzspektroskopie (H-MRS) genutzt (Browning et al. 2004; Szczepaniak et al. 2005).

Um die Prävalenz der Hepatosteatose besser einordnen zu können, ergibt es Sinn sich zwei weitere Studien anzusehen. Lee et al. (2007) und Choudhary et al. (2017) haben die Prävalenz von NAFLD bei Leberlebendspendern mittels Biopsie, dem diagnostischen Goldstandard, ermittelt. An der Studie von Lee et al. (2007) nahmen 589 Koreaner*innen teil. Eine Hepatosteatose wurde bei 51,4 % festgestellt. Die in Indien stattgefundene Studie von Choudhary et al. (2017) kam bei 573 Teilnehmenden auf eine Prävalenz von 49,2 %. Bei der Interpretation beider Studien muss jedoch berücksichtigt werden, dass die Prävalenz von Hepatosteatose, wie unten gezeigt, von der Ethnie des Studienkollektivs abhängig ist (Pan 2014).

Da NAFLD eine Subgruppe der Fettlebererkrankungen darstellt, muss für die Hepatosteatose von einer Prävalenz im Prozentbereich der beiden obengenannten Studien, oder darüber ausgegangen werden.

Die Annahme wird durch eine Studie von Caruso et al. (2019) gestützt. Die Autoren haben in einer prospektiven Kohortenstudie anhand von 1.445 Teilnehmenden die Prävalenz von NAFLD und AFLD ermittelt. Hierbei stellten sie für NAFLD eine Prävalenz von 41,2 % und für AFLD von 15,8 % fest. Aus den Resultaten lässt sich ebenso abschätzen, in welchem Verhältnis die Prävalenzen von NAFLD und AFLD zueinanderstehen (Caruso et al. 2019).

Eine weitere große prospektive Kohortenstudie mit 218.030 Teilnehmenden von Chang et al. (2019) kommt auf eine deutlich geringere Prävalenz von NAFLD (22 %) und AFLD (6,4 %) unter der Normalbevölkerung.

Beide Studien veranschaulichen, dass AFLD nur circa ein Viertel bis ein Fünftel aller Hepatosteatosefälle ausmacht und NAFLD die mit Abstand größte Untergruppe darstellt. Jedoch muss ebenso bei beiden Studien angemerkt werden, dass die Prävalenz anhand von nicht-invasiven Methoden ermittelt wurde.

In einer Metaanalyse fassen Vernon et al. (2011) die Ergebnisse aller epidemiologischen Studien zur Prävalenz von NAFLD zwischen 1980 und 2010 zusammen. Die Studienergebnisse variieren zwischen 6 % und 35 %. Für US-Amerikaner vermuten die Studienautoren die Prävalenz bei 30 %, für den Rest der Welt bei 20 %. Diese Studie zeichnet sich durch ihre Teilnehmendengröße aus, hat jedoch wieder die Schwäche, dass sie auf mehreren unterschiedlich zuverlässigen Messmethoden basiert.

Wie in der oben genannten Studie bereits beschrieben, ist die Prävalenz von NAFLD nicht nur von der diagnostischen Modalität, sondern auch von der analysierten Region abhängig. Younossi et al. (2016) haben in einer Metaanalyse die Daten von insgesamt 8.515.431 Personen weltweit untersucht und die NAFLD-Prävalenz für unterschiedliche Regionen bestimmt. Es zeigt sich, dass Afrika mit nur 13,48 % die geringste Prävalenz und der Mittlere Osten mit 31,79 % die höchste hat. Eine ausführliche Übersicht über die Prävalenz von NAFLD in unterschiedlichen Regionen der Welt ist in Tabelle 2 zu sehen. Auch hier müssen die Ergebnisse hinterfragt werden, da in die Metaanalyse sowohl Biopsiedaten, als auch nicht-invasive Daten miteingeflossen sind und die Autoren selbst feststellen, dass bei der Diagnose anhand von erhöhten Leberenzymen die NAFLD-Prävalenz signifikant unterschätzt wird. Die unterschiedlichen Werte je nach Region erklären sich die Autoren mit dem Zusammenhang zwischen NAFLD und der Prävalenz von metabolischen Störungen wie Adipositas.

Dieser Zusammenhang ist ebenso von Bedeutung, um einschätzen zu können, ob die Prävalenz von Leberverfettung steigt oder sinkt. Chooi et al. (2019) stellen in ihrer Studie fest, dass die weltweite Prävalenz von Übergewicht und Adipositas nach wie vor steigt. Basierend auf den Erkenntnissen von Younossi et al. (2016) ist anzunehmen, dass dies auch für NAFLD gelten wird.

Tabelle 2: NAFLD-Prävalenz aufgeteilt nach Region

Region	n	Prävalenz	95 % Konfidenzintervall (%)
Afrika	2	13,48	(5,69 - 28,69)
Asien	14	27,37	(23,29 - 31,88)
Europa	11	23,71	(16,12 - 33,45)
Mittlerer Osten	3	31,79	(13,48 - 58,23)
Nordamerika	13	24,13	(19,73 - 29,15)
Südamerika	2	30,45	(22,74 - 39,44)
Gesamt	45	25,24	(22,1 - 28,65)

n: Anzahl der in die Metaanalyse einbezogenen Studien (basierend auf Younossi et al. 2016). Die Verwendung erfolgt mit freundlicher Genehmigung von John Wiley and Sons.

1.2.2 Einflussfaktoren auf Hepatosteatose

Da sich ein Großteil der epidemiologischen Forschung mit NAFLD befasst und somit kaum Daten zur allgemeinen Hepatosteatose vorhanden sind, wird im folgenden Kapitel größtenteils auf die Epidemiologie der NAFLD eingegangen. Da NAFLD bei weitem die größte Subgruppe der Hepatosteatose darstellt (s. o.), werden deren epidemiologische Daten als teilrepräsentativ angenommen.

1.2.2.1 Alter

Vernon et al. (2011) stellen in ihrer Übersichtsarbeit eine steigende Prävalenz von NAFLD und NAFLD-assoziiierter Fibrose mit steigendem Lebensalter fest. Des Weiteren sind die Progressionswahrscheinlichkeit und Mortalitätswahrscheinlichkeit bei älteren Menschen erhöht. Es wird allerdings angemerkt, dass der genaue Zusammenhang zwischen Alter und NAFLD noch nicht geklärt ist. Die hohe Prävalenz von NAFLD und NAFLD-assoziiierter Fibrose könnte auch durch die Dauer der Erkrankung anstatt dem Alter selbst bedingt sein.

1.2.2.2 Geschlecht

Die Datenlage zur Prävalenz von NAFLD in beiden Geschlechtern hat sich geändert. Wurde in älteren Studien noch eine erhöhte Prävalenz von NAFLD bei Frauen festgestellt, zeigt sich in neueren Studien eine signifikant höhere Prävalenz bei Männern. Dies wird in Tabelle 3 veranschaulicht. Der Grund für die unterschiedlichen Ergebnisse wird darin gesehen, dass die älteren Studien nicht populationsbasiert waren und somit für Messfehler anfällig sind (Pan 2014).

Tabelle 3: Geschlechterunterschied der NAFLD-Prävalenz in populationsbasierten Studien

Referenz	Studienpopulation	n	Definition von NAFLD	Prävalenz von NAFLD (%)	
				M	F
Ruhl et al. (2003)	NHANES III (1988 - 1994)	5.724	ALT ¹	4,3	1,6
Clark et al. (2003)	NHANES III (1988 - 1994)	15.676	ALT oder AST ²	5,7	4,6
Browning et al. (2004)	Dallas Heart Study	734 ⁵	MRS ³	42	24
Ioannou et al. (2006)	NHANES (1999 - 2002)	6.823	ALT oder AST ⁴	13,4	4,5
Lazo et al. (2013)	NHANES III (1988 - 1994)	12.454	Ultraschall	20,2	15,8
Schneider et al. (2014)	NHANES III (1988 - 1994)	4.037	Ultraschall	15	10,1

¹ALT > 43 U/L; ²ALT > 40 U/L und AST > 37 U/L für Männer; ALT und AST > 31 U/L für Frauen; ³Hepatischer Triglyceridgehalt > 5,5 %; ⁴ALT > 43 U/L oder AST > 40 U/L; n: Personenanzahl; M: Männer; F: Frauen (basierend auf Pan 2014). CC BY 4.0.

Als Ursachen für die erhöhte Prävalenz von NAFLD bei Männern vermuten Pan et al. (2014) Unterschiede in der Körperfettverteilung, im Lebensstil und im Metabolismus von Sexualhormonen.

1.2.3 Folgen der Hepatosteatose

Die Folgen einer Hepatosteatose können systemisch-, oder lokal auf die Leber begrenzt sein. Hierbei muss angemerkt werden, dass eine Hepatosteatose nicht nur bei einer NAFLD und AFLD vorliegt, sondern auch sekundär bei Personen mit anderen Lebererkrankungen, wie Hepatitis B und C oder Morbus Wilson, auftreten kann (Gordon et al. 2005; Stättermayer et al. 2015). Die Folgen einer Hepatosteatose unterscheiden sich somit auch abhängig von der zugrundeliegenden oder begleitenden Erkrankung.

Als relevanteste systemische Auswirkung einer NAFLD gelten sowohl kardiovaskuläre Ereignisse, wie auch die Entwicklung eines Typ 2 Diabetes, bzw. dessen Progression (Roeb et al. 2015). Der Zusammenhang zwischen Diabetes Typ 2 und NAFLD wird als bidirektional bezeichnet, da beide meist miteinander z. B. im Rahmen des metabolischen Syndroms einhergehen und nicht zur Gänze geklärt ist, wie sie sich gegenseitig beeinflussen (Anstee et al. 2013).

Auch die genaue Beteiligung von NAFLD an kardiovaskulären Ereignissen wird kontrovers diskutiert. Als sicher gilt jedoch der Zusammenhang zwischen NAFLD und einer erhöhten Prävalenz von Herz-Kreislaufgeschehen (Anstee et al. 2013). Stepanova und Younossi (2012) ist es gelungen NAFLD als unabhängigen Risikofaktor für kardiovaskuläre Ereignisse nachzuweisen. Erstaunlich ist jedoch, dass trotz NAFLD als Grunderkrankung die Mortalität an Herz-Kreislaufgeschehen im Beobachtungszeitraum von 14 Jahren nicht signifikant gestiegen ist. Auch diese Studie wird jedoch kritisiert, da die Diagnosestellung der NAFLD anhand von Laborparametern und Ultraschall erfolgt ist (Anstee et al. 2013).

Zu den organspezifischen Folgen von NAFLD ist die Progression zur NASH zu zählen. Das Risiko, dass eine NAFLD im Verlauf eine NASH ausbildet beträgt bis zu 30 % (Calzadilla Bertot und Adams 2016). Die Wahrscheinlichkeit einer Progression ist vom Grad der Hepatosteatose abhängig. Insofern spielt der Hepatosteatosegrad im Speziellen bei der Risikoabschätzung für die Progression einer frühzeitig erkannten NAFLD zu einer fibrotischen NASH eine wichtige Rolle in der klinischen Praxis (Ajmera et al. 2018). Ebenso fördern vor allem Adipositas und Insulinresistenz eine zelluläre Entzündungsreaktion und somit die Ausbildung einer NASH auf Boden einer NAFLD (Naik et al. 2013).

NASH nimmt eine besondere Rolle ein, da diese mit steigender Erkrankungsdauer zu einer Zirrhose werden oder in ein HCC übergehen kann (Anstee et al. 2013). Die Entwicklung eines HCC muss hierbei nicht aus einer Leberzirrhose erfolgen, sondern kann auch direkt aus einer NASH oder Hepatosteatose entstehen. Die Entstehung aus einer Zirrhose ist jedoch mit 4 – 27 % am wahrscheinlichsten im Vergleich zu 0 – 2,8 % aus NASH und 0 – 0,5 % aus einer reinen Steatose. Eine direkte Entwicklung eines HCC aus einer einfachen Hepatosteatose heraus wurde nie im Menschen diagnostiziert. Die Zahlen hierfür beziehen sich auf fragliche Evidenz und stellen somit nur einen möglichen Pathomechanismus dar.

Diese Beobachtungsergebnisse beziehen sich auf einen Beobachtungszeitraum von 19,5 Jahren (Duan et al. 2014). Eine Übersichtsgrafik zu den Progressionswegen ist in Abbildung 2 zu sehen.

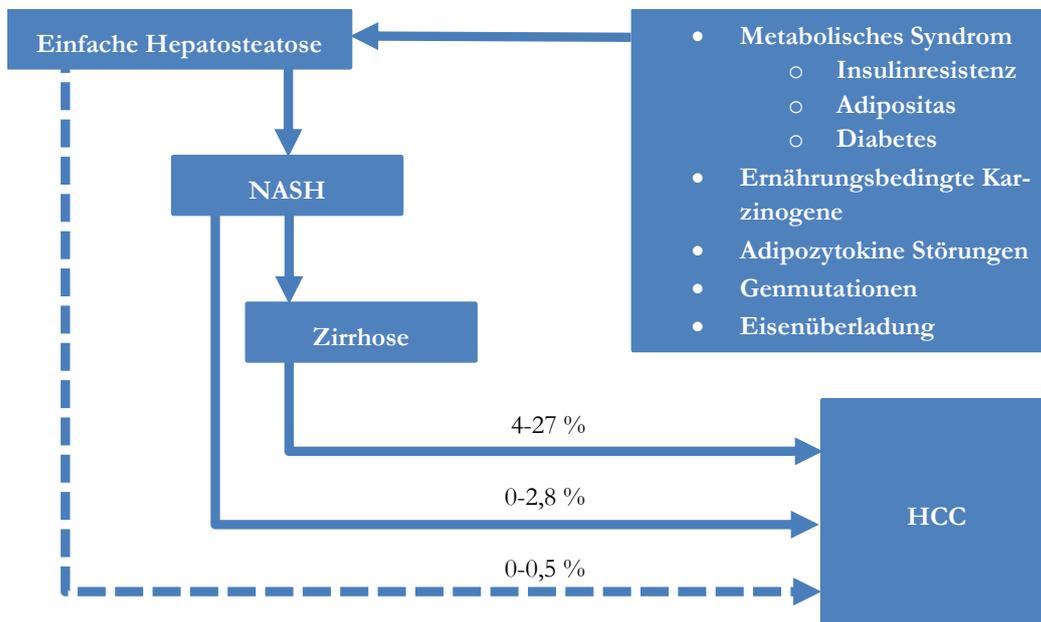


Abbildung 2: Unterschiedliche Progressionswege von NAFLD zu HCC. Die Anwesenheit eines metabolischen Syndroms, von Adipokinalterationen, Toxinexposition und Mutationen von Schlüsselgenen spielt eine Rolle in der Progression von Steatose zum HCC. Dies gilt sowohl in Anwesenheit als auch in Abwesenheit von Inflammation, Fibrose und Zirrhose. HCC wurde nie in einer einfachen Fettleber im Menschen diagnostiziert (der gepunktete Pfeil repräsentiert fragliche Evidenz). Die Prävalenz von HCC aus NAFLD und NASH beträgt 0 – 0,05 % und 0 – 2,8 % innerhalb von 19,5 Jahren. Circa 4 – 27 % der NASH-zugehörigen Zirrhosepatient*innen progredieren in dieser Zeit zu einem HCC. Durchgehende Pfeile: mit sicherer experimenteller oder epidemiologischer Evidenz; gepunktete Linie: möglicher Mechanismus. NASH: nicht alkoholische Steatohepatitis; HCC: hepatozelluläres Karzinom (basierend auf Duan et al. 2014) . Die Verwendung erfolgt mit freundlicher Genehmigung von Elsevier.

Einen relevanten Einfluss von sekundärer Hepatosteatose auf den Verlauf einer CLD stellten Chan et al. (2017) fest. Ihnen gelang es nachzuweisen, dass eine parallel zu einer HBV-Infektion bestehende Hepatosteatose ein unabhängiger Risikofaktor für die Ausbildung eines HCC ist. Ein ähnlicher Zusammenhang ist auch bei Personen mit gleichzeitig bestehender Autoimmunhepatitis (AIH) und Hepatosteatose zu sehen. Hier konnte bei Personen mit AIH und Steatose ein höheres Risiko für die Entwicklung einer Leberzirrhose und für einen klinisch schlechteren Verlauf nachgewiesen werden als bei Personen, die nur eine AIH hatten (De Luca-Johnson et al. 2016). Analog dazu wurde in einer anderen Studie ein erhöhtes Risiko bei Personen mit Alpha-1-Antitrypsinmangel und sekundärer Hepatosteatose festgestellt. Auch hier lag bei gleichzeitigem Vorhandensein beider Erkrankungen ein erhöhtes Leberzirrhoserisiko vor (Strnad et al. 2019).

Die oben genannten Studienergebnisse führten dazu, dass die Autoren der Überprüfung einer möglichen sekundären NAFLD bei der Risikoevaluation der primären Erkrankung eine wichtige Rolle zukommen lassen (De Luca-Johnson et al. 2016; Chan et al. 2017; Strnad et al. 2019).

1.3 Diagnostik

Für die Diagnose und Quantifizierung einer Hepatosteatose stehen unterschiedliche Methoden zur Verfügung. Als Goldstandard gilt die Leberbiopsie, die jedoch aufgrund ihrer Invasivität nur nach einer strengen Indikationsstellung durchgeführt werden darf. Als nicht-invasive Diagnoseverfahren stehen unterschiedliche bildgebende und auf Biomarkern basierende Methoden zur Verfügung (Roeb et al. 2015).

1.3.1 Invasive Diagnostik: Leberbiopsie

Für die Indikationsstellung einer Leberbiopsie bei Verdacht auf Hepatosteatose wird von den deutschen S2k Leitlinien für die nicht alkoholische Fettlebererkrankung ein Diagnosepfad empfohlen. Der ausführliche Entscheidungsbaum ist in Abbildung 3 zu sehen.

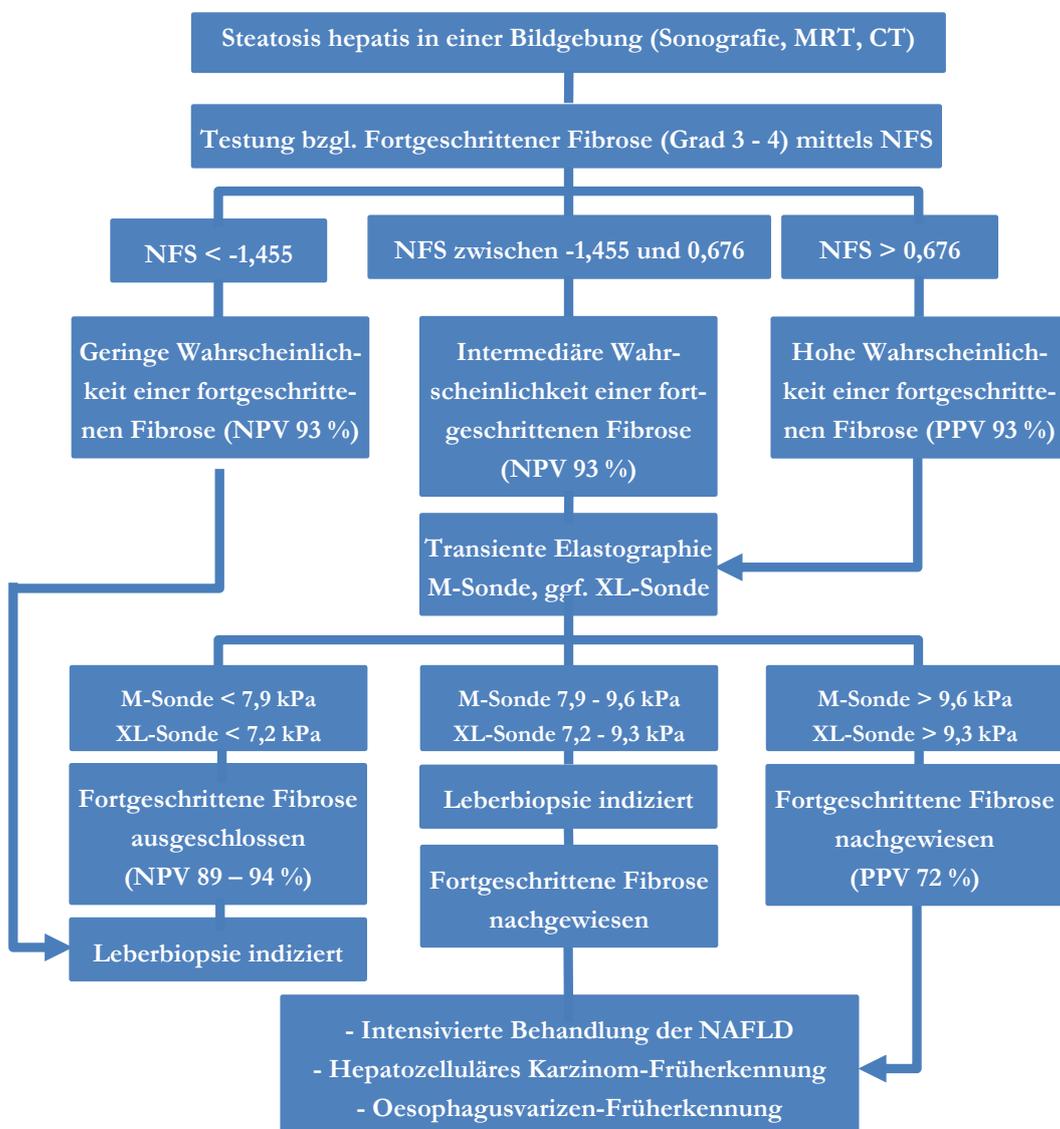


Abbildung 3: Diagnostischer Algorithmus bei NAFLD - Indikation zur Leberbiopsie. NFS: NAFLD-Fibrosis-Score; NPV: negativer prädiktiver Wert; PPV: positiver prädiktiver Wert (basierend auf Roeb et al. 2015). Die Verwendung erfolgt mit freundlicher Genehmigung von Georg Tieme Verlag KG.

Um eine Leberbiopsie in Erwägung ziehen zu können sollte, zuerst eine Hepatosteatose mit Hilfe eines bildgebenden Verfahrens (Ultraschall, MRT, CT) festgestellt werden. Im Anschluss muss der NAFLD *Fibrosis Score* (NFS) bestimmt werden. Bei einem intermediären und einem hohen Fibrosierisiko (NFS $-1,455 - 0,676$ und NFS $> 0,676$) wird eine transiente Elastographie empfohlen. Liegt der Elastographiewert im Bereich von $7,9 - 9,6$ kPa (M-Sonde), oder $7,2 - 9,3$ kPa (XL-Sonde) wird eine Leberbiopsie empfohlen (Roeb et al. 2015).

Histologisch stellt sich eine NAFLD durch makrovesikuläre oder mediovesikuläre Fetteinlagerungen in die Hepatozyten dar. Die Fettvakuolen können hierbei so groß werden, dass sie den Zellkern an den Rand drängen und die Hepatozyten eine morphologische Ähnlichkeit zu Adipozyten zeigen. Es gibt weitere mikrovesikuläre Formen, die als eher unüblich gelten und in bis zu 10 % der Fälle vorkommen. Die Fetteinlagerungen können insbesondere perizentral oder auch panazinär in Zone 3 auftreten (Bedossa 2016).

Eine Hepatosteatose kann histologisch in drei Grade unterteilt werden. Von einer physiologischen Leber spricht man bei einem Anteil von < 5 % fettenthaltenden Hepatozyten. Als Hepatosteatose Grad I bezeichnet wird eine Verfettung von $\geq 5 - 33$ % der Hepatozyten und als Grad II $> 33 - 66$ %. Eine Hepatosteatose Grad III wird ab einem Anteil von > 66 % verfetteter Hepatozyten diagnostiziert (Kleiner et al. 2005; Nalbantoglu und Brunt 2014; Bedossa 2017).

Neben der Verfettung der Leber können auch weitere Parameter histologisch ermittelt werden, um NAFLD-assoziierte Erkrankungen, oder Folgeerkrankungen wie eine NASH oder Leberfibrose zu diagnostizieren. So wird die NASH anhand der histologischen Kriterien Hepatosteatose, lobuläre Inflammation und Leberzellballonierung diagnostiziert (Bedossa 2017). Die drei Parameter können semiquantitativ zusammengefasst werden, um eine differenzierte Diagnose zu ermöglichen. Hierfür empfiehlt die aktuelle S2k Leitlinie den NAS (*NAFLD Activity Score*) (Roeb et al. 2015). Dieser wird in Tabelle 4 beschrieben.

Tabelle 4: NAFLD Activity Score. Interpretation: 0 – 2 Punkte: definitiv kein NASH, 3 – 4 Punkte: fragliche NASH, ≥ 5 Punkte: definitive NASH; ¹Herde pro Gesichtsfeld bei 200-facher Vergrößerung.

NAS	Steatose (% Fetteinlagerung in Hepatozyten)	Ballionierte Hepatozyten	Inflammation, lobulär
0	< 5 % (0)	Keine (0)	Keine (0)
3	$5 - 33$ % (1)	Wenig (1)	1 – 2 Foci pro 20-x-Feld ¹ (1)
6	$34 - 66$ % (2)	Zahlreich (2)	2 – 4 Foci pro 20-x-Feld ¹ (2)
8	> 66 % (3)	Zahlreich (2)	> 4 Foci pro 20-x-Feld ¹ (3)

(basierend auf Kleiner et al. 2005). Die Verwendung erfolgt mit freundlicher Genehmigung von John Wiley and Sons.

Bei Punktwerten ≤ 2 schließt der NAS-Score definitiv eine Steatohepatitis aus, bei 3 – 4 ist sie fraglich und bei ≥ 5 ist eine NASH definitiv gesichert.

Ein weiterer Parameter, der histologisch beurteilt werden kann, ist die Fibrose. Sie ist das Resultat von ausgeheilten Leberschäden, die durch vielfältige Ursachen wie autoimmune-, virale-, cholestatische und metabolische Erkrankungen hervorgerufen werden können. Dadurch ist Leberfibrose eine Begleiterscheinung vieler chronischer Lebererkrankungen. (Friedman 2003). Eine Leberfibrose kann im Rahmen des Stagings z. B. nach Desmet und Scheuer klassifiziert werden. Die Klassifizierung erstreckt sich hierbei von Grad 0, keine Fibrose, über milde, mäßige bis schwere Fibrose und hat die Leberzirrhose mit Grad 4 als Maximalwert. (Schirmacher et al. 2004). Die genaue Klassifizierung wird in Tabelle 5 wiedergegeben.

Tabelle 5: Fibrorestaging nach Desmet und Scheuer

Score	Ausformuliert	Histologische Merkmale
0	Keine Fibrose	Keine Faservermehrung
1	Milde/geringgradige Fibrose	Portale Faservermehrung, keine Septen
2	Mäßige/mittelgradige Fibrose	Inkomplette oder komplette portoportale Fasersepten ¹ Erhaltene Architektur
3	Schwere/hochgradige Fibrose	Septenbildende Faservermehrung mit Architekturstörung ² kein Anhalt für kompletten zirrhotischen Umbau
4	Zirrhose	Wahrscheinlicher ³ oder definitiver zirrhotischer Umbau

¹Unabhängig von Zahl und Breite der Septen, ²Zum Beispiel portozentrale Septen, Verschiebung der portalen/azinären Architektur (portal-zentralvenöser Abstand), ³Zum Beispiel ohne definitiven Nachweis vollständig bindegewebig separierter Pseudoloboli, aber aufgrund indirekter Zeichen (z. B. fragmentierte, herausgebrochene Pseudolobolie) anzunehmen (basierend auf Schirmacher et al. 2004). Die Verwendung erfolgt mit freundlicher Genehmigung von Springer Nature.

Da eine Unterscheidung zwischen NAFL und NASH nur anhand von nicht-invasiver Diagnostik nicht möglich ist, nimmt die Leberbiopsie auch hier eine wichtige Rolle ein (Roeb et al. 2015).

1.3.2 Nicht-invasive Diagnostik

Nicht-invasive Diagnostikmethoden können in bildgebende und auf Biomarkern basierende Verfahren unterteilt werden. Zu den bildgebenden Verfahren gehören Computertomographie (CT), Magnetresonanztomographie (MRT), Protonenmagnet-resonanzspektroskopie (H-MRS) und Ultraschallverfahren. Biomarkerbasierte Verfahren wie der Steatotest, *Fatty Liver Index* und die Stuhl Diagnostik versuchen anhand von Parametern aus Blut und Stuhl, die teilweise mit zusätzlichen Messwerten wie Taillenumfang, Größe, Gewicht und Geschlecht ergänzt werden, die Wahrscheinlichkeit einer NAFLD zu hervorzusagen (Roeb et al. 2015).

1.3.2.1 Computertomographie (CT)

Die Ermittlung des Hepatosteatosegrades mittels CT hat die Vorteile, dass es sich um ein untersucherunabhängiges Messverfahren handelt und Computertomographen breit verfügbar sind. Dem gegenüber steht jedoch der große Nachteil, dass eine CT-Untersuchung mit einer hohen Strahlenbelastung einhergeht (Lăpădat et al. 2017).

Aus diesem Grund wird in den aktuellen S2k-Leitlinien NAFLD von der Verwendung der CT als primäres Screeningverfahren abgeraten (Roeb et al. 2015). Eine Ausnahme stellen Patient*innen dar, die im Rahmen der Voruntersuchungen für eine Leberlebenspende untersucht werden (Iwasaki et al. 2004).

Die Diagnose und Quantifizierung von Hepatosteatose mittels CT erfolgt anhand des Leber-Milz-Index (L/S-Ratio). Hierbei werden über Leber und Milz *Regions of Interest* (ROI) gelegt und die Hounsfield-Einheiten darin ermittelt. Aus diesen beiden Werten wird der Leber-Milz-Index errechnet. Als optimaler *Cut-off*-Wert, um eine moderate Steatose von $> 30\%$ fettenthaltender Hepatozyten zu detektieren, gilt ein L/S-Wert von $> 1,1$. Hier liegt die Sensitivität bei $83,3\%$ und die Spezifität bei $81,5\%$ (Iwasaki et al. 2004).

Bei der Detektion jeglicher Leberverfettung ($> 0 - > 5\%$ fettenthaltender Hepatozyten) zeigt die CT nur noch eine Sensitivität von 50% bzw. $46,1\%$ und Spezifität von $77,2\%$ bzw. $84,4\%$ (Lee et al. 2010; Bohte et al. 2011).

1.3.2.2 Magnetresonanztomographie (MRT)

MRT ist ein strahlungsfreies Verfahren, mit dem anhand verschiedener MRT-Techniken der Anteil von Wasser und Fett in der Leber separiert wird. Als Nachteil werden momentan der hohe Zeit- sowie Kostenaufwand genannt (Roeb et al. 2015; Lăpădat et al. 2017).

Ein vielversprechendes MRT-Verfahren zur Ermittlung einer Hepatosteatose ist das *Proton Density Fat Fraction* MRT (MRT-PDFF). Es bietet den Vorteil, dass es mit der Standard-MRT-Technologie durchgeführt werden kann und eine exzellente Messgenauigkeit liefert (Grat et al. 2020). Dies zeigt auch eine aktuellere Studie, die für die Detektion von Hepatosteatose mittels MRT-PDFF eine Sensitivität von 93% und Spezifität von 94% feststellt. Bei der Differenzierung zwischen den Steatosegraden $0 - 1$ vs. $2 - 3$ und $0 - 2$ vs. 3 erreichte das MRT eine Sensitivität von 74% und Spezifität von 90% und 87% (Gu et al. 2019). Ebenso

ist das MRT in der Lage im Rahmen einer Verlaufskontrolle geringgradige Veränderungen einer Hepatosteatose im Sinne einer Zunahme mit einer Sensitivität von 57 % und Spezifität von 90 % und einer Abnahme mit 58% und 90% zu detektieren (Middleton et al. 2017).

1.3.2.3 Protonenmagnetresonanzspektroskopie (H-MRS)

H-MRS ist eine weitere nicht-invasive Untersuchungsmethode. Ihre Verwendung ist zurzeit jedoch nur auf Forschungszentren begrenzt, da sie sehr hohe Anforderungen an Hard- und Software stellt (Roeb et al. 2015).

H-MRS zeigt im Vergleich zu anderen nicht-invasiven Untersuchungsmethoden gute Ergebnisse bei der Detektion von geringgradiger Hepatosteatose ($> 0\%$; $> 0,2\%$; $> 5\%$) mit einer Sensitivität von 88,5 % und Spezifität 92,0 %) (Bohte et al. 2011).

1.3.2.4 Ultraschallverfahren

Leberultraschall stellt ein breit verfügbares, kostengünstiges und strahlungsfreies Screeningverfahren dar (Roeb et al. 2015).

Es gibt unterschiedliche Methoden, um Ultraschall für das Screening nach einer Hepatosteatose zu nutzen. Zu diesen Methoden zählt die optische Beurteilung, der hepato-renale Index und *Controlled-Attenuation-Parameter*.

Eine nicht verfettete Leber sollte die gleiche Echogenität aufweisen wie die im gleichen Bild angeschnittene benachbarte Niere. Erscheint die Leber heller als die Niere, hat sie eine erhöhte Echogenität und es kann von einer Leberverfettung ausgegangen werden (Joy et al. 2003).

Diese Eigenschaft macht man sich bei der optischen Beurteilung der Leber und dem HRI zu Nutzen (Webb et al. 2009; Lăpădat et al. 2017).

Steatosis Hepatis kann via Ultraschall in drei Grade (leicht = 1° , moderat = 2° , schwer = 3°) eingestuft werden. Ist nur eine erhöhte Echogenität der Leber im Vergleich zur Niere in Abwesenheit von anderen Auffälligkeiten festzustellen, spricht dies für eine leichte Hepatosteatose. Eine moderate Leberverfettung liegt vor, wenn eine leichte Erhöhung der Echogenität des Leberparenchyms festzustellen ist und die Wände der Portalvenenäste durch die erhöhte Echogenität verschleiert werden. Eine schwere Steatosis hepatis liegt vor, wenn eine deutliche Erhöhung der Echogenität vorliegt und die Umrisse von Diaphragma und Lebergefäßen dadurch schlecht oder nicht abgrenzbar sind und es zu einer dorsalen Schallabschwächung kommt (Saadeh et al. 2002; Lăpădat et al. 2017). Des Weiteren wird bei einer steigenden Leberverfettung eine dorsale Schallabschwächung des Leberparenchyms beobachtet, die bis zur vollständigen Nichtdarstellbarkeit des rechten unteren Leberrandes bei schwerer Hepatosteatose gehen kann (Saadeh et al. 2002; Joy et al. 2003). Beispiele für Ultraschallbilder mit und ohne Hepatosteatose sind in Abbildung 4 zu sehen.

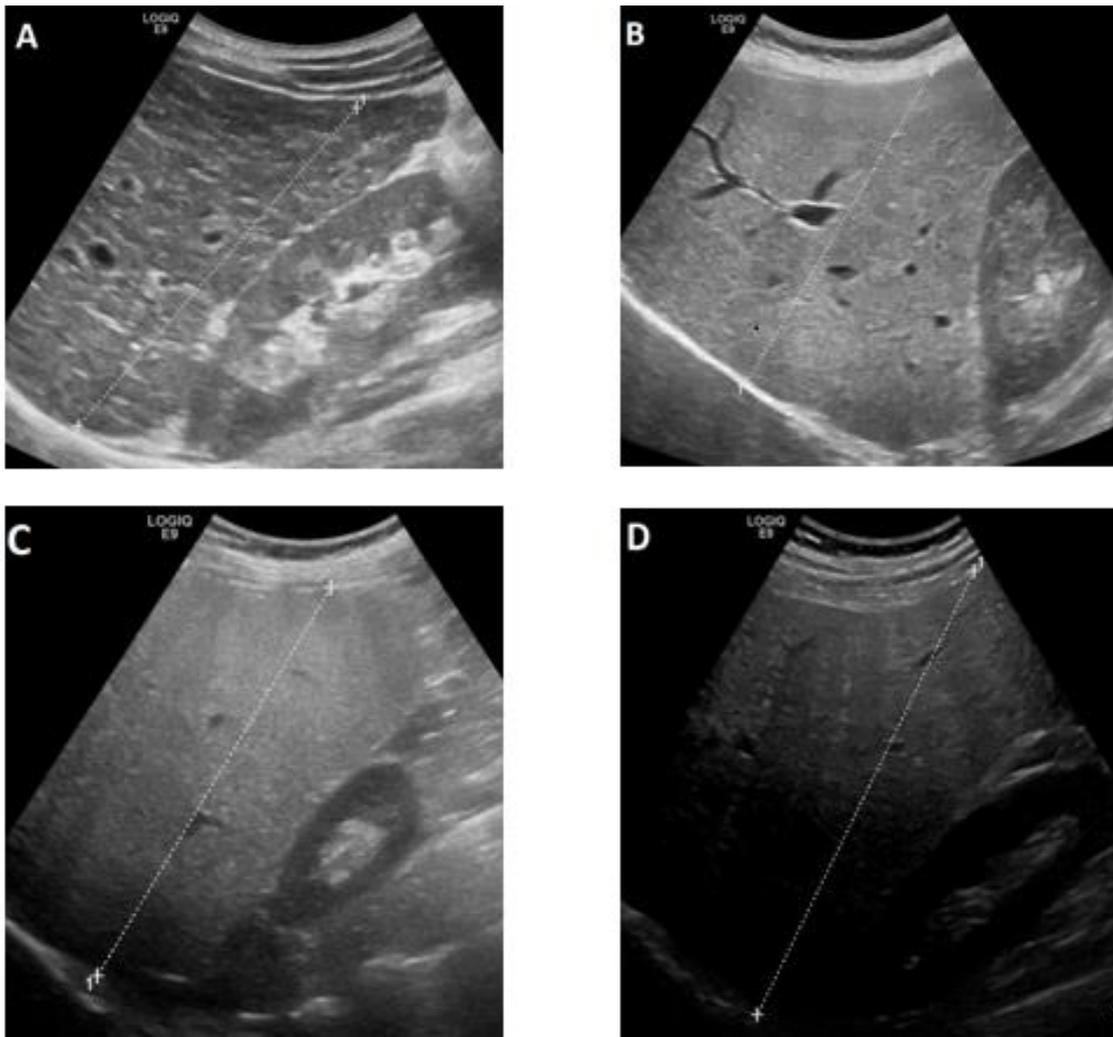


Abbildung 4: Ultraschallbilder der Leber mit und ohne Leberverfettung. A: Leber mit normwertigem Anteil an fettreichen Hepatozyten; B: Leber mit Hepatosteatose Grad 1; C: Leber mit Hepatosteatose Grad 2; D: Hepatosteatose mit Grad 3 (basierend auf Petzold et al. 2020). CC BY 4.0.

Sensitivität und Spezifität von Ultraschall in der Hepatosteatosedagnostik hängen sehr stark vom Grad der Leberverfettung ab. Bei mittlerer und schwerer Hepatosteatose zeigt Ultraschall eine gute Sensitivität (ca. 85 %) und Spezifität (ca. 94 %). Bei leichtgradiger Leberverfettung nimmt die Sensitivität jedoch stark ab (ca. 40 %). Aus diesem Grund raten die S2k Leitlinien NAFLD von einem Ausschluss einer Steatosis Hepatis durch Sonographie ab (Roeb et al. 2015).

In diesem Kontext ist eine großangelegte Metaanalyse von Bohte et al. (2011) mit 46 eingeschlossenen Studien und 4.715 Teilnehmenden von Interesse. Sie ermittelte die Sensitivität und Spezifität von Ultraschall, um eine Hepatosteatose zu erkennen und sie in eine der drei Grade einteilen zu können. Die Ergebnisse sind in Tabelle 6 aufgelistet und zeigen eine ausführlichere Aufschlüsselung der Testgenauigkeit von Ultraschall als die S2k Leitlinien NAFLD (Roeb et al. 2015).

Es muss jedoch berücksichtigt werden, dass die Steatosegrade in dieser Studie von der zurzeit üblichen Einteilung nach prozentualem Anteil von verfetteten Hepatozyten leicht abweichen (Kleiner et al. 2005; Bohte et al. 2011).

Tabelle 6: Diagnostische Genauigkeit von Ultraschall zur Hepatosteatosebestimmung

Gruppe 1: > 0 %, > 2 %, > 5 % ¹		Gruppe 3: > 25 %, > 20 %, > 22 % ¹	
n ²	Sensitivität (95 % KI)	n ²	Sensitivität (95 % KI)
19	73,3 (62,2 - 82,1)	19	85,7 (78,4 - 90,8)
n ²	Spezifität (95 % KI)	n ²	Spezifität (95 % KI)
19	84,4 (76,2 - 90,1)	19	85,2 (76,9 - 90,9)

Gruppe 3: > 10 % ¹		Gruppe 4: > 50 %, > 60 %, > 66 % ¹	
n ²	Sensitivität (95 % KI)	n ²	Sensitivität (95 % KI)
5	90,5 (79,3 - 96,0)	9	91,1 (63,0 - 98,4)
n ²	Spezifität (95 % KI)	n ²	Spezifität (95 % KI)
5	69,6 (60,0 - 77,7)	9	91,9 (74,3 - 97,8)

¹Anteil an verfetteten Hepatozyten in der Biopsie; ²Anzahl analysierter Datensätze; 95 % KI: 95 % Konfidenzintervall (basierend auf Bohte et al. 2011). CC BY-NC 2.0.

Ebenso interessant ist das 95 % Konfidenzintervall der oben genannten Studie. Es gibt eine Aussage über die Spannbreite der Einzelergebnisse aus den 46 Studien. Bemerkenswert ist, dass es sowohl bei Sensitivität als auch Spezifität um 10 bis 15 Prozentpunkte variiert (Bohte et al. 2011). Dies veranschaulicht, dass die Datenlage zur Ultraschalldiagnostik bezüglich Hepatosteatose keinesfalls eindeutig ist und die Frage gestellt werden muss, was die Ursache dieser großen Spannbreite ist.

Diese Frage stellten sich auch Hernaez et al. (2011) und befassen sich deshalb in ihrer Metaanalyse mit der Untersuchergenauigkeit. Es wurden 49 Studien zwischen 1967 und 2010 ausgewertet, die sich mit der Messgenauigkeit von Ultraschall zur Hepatosteatosediagnostik auseinandersetzten. Die Interraterreliabilität und Intraraterreliabilität wurden anhand einer Kappa Statistik ermittelt und betragen 0,44 - 1,0 und 0,54 - 0,92. Dies entspricht einer Übereinstimmung von moderat (0,41 - 0,60) bis vollkommen (0,81 - 1,00) (Strauss et al. 2007).

Eine weitere Möglichkeit zur Bestimmung des Hepatosteatosegrades mit Hilfe von Ultraschall stellt der hepatorenale Index (HRI) dar. Hierbei wird die Echogenität von Leber und Niere mit Hilfe einer Pixelstatistik für eine bestimmte Region (ROI) als Zahlenwert erfasst und aus den beiden Werten ein Index gebildet (Webb et al. 2009). Dieses Verfahren basiert

somit auf der Annahme, dass eine nicht verfettete Leber die gleiche Echogenität wie die dazugehörige Niere hat (Joy et al. 2003). Durch das Errechnen des Index ist jedoch die Abschätzung der Echogenität nicht mehr von der Einschätzung des Untersuchenden, wie bei der optischen Einschätzung des Steatosegrades, abhängig, sondern wird als objektiver Wert wiedergegeben.

Es gibt mehrere Studien, die sich mit der Messgenauigkeit des HRI befassen haben. Ein Teil der Studien stellt eine sehr gute Messgenauigkeit des HRI fest (Webb et al. 2009; Marshall et al. 2012; De Almeida e Borges et al. 2013), ein anderer zeigte jedoch widersprüchliche Ergebnisse (Wang et al. 2013; Chauhan et al. 2016). Ebenso sind mit 1,24 bis 2,01 die *Cut-off*-Werte mit deren Hilfe entschieden wird, ob eine Hepatosteatose vorliegt breit gestreut.

Die Ergebnisse der oben genannten Studien sind in Tabelle 7 aufgelistet.

Tabelle 7: Studienvergleich der Messgenauigkeit des HRI zur Detektion einer Hepatosteatose

Referenz	n ¹	Cut-off-Wert	Sensitivität	Spezifität
Webb et al. (2009)	111	1,49	100	91
Marshall et al. (2012)	101	1,28	100	54
De Almeida e Borges et al. (2013)	82	1,24	93	93
Wang et al. (2013)	175	4 ²	82	63
Chauhan et al. (2016)	45	2,01	62,5	95,2

¹n: Anzahl an in Studie eingeschlossenen Personen; ²Autoren nutzten hepatorenen Unterschied statt hepatorenen Index (basierend auf Chauhan et al. 2016). Die Verwendung erfolgt mit freundlicher Genehmigung von John Wiley and Sons.

Auffallend ist, dass die Datenerhebung des Großteils der Studien vor mehr als 10 Jahren stattgefunden hat, als noch keine *High-End*-Ultraschallgeräte zu Verfügung standen und dass die Studien teilweise nach der Ätiologie der CLD selektionierte Studienkollektive hatten. So hatten in der Studie von Wang et al. (2013) alle Personen eine Hepatitis B- und/oder eine Hepatitis C-Infektion. De Almeida e Borges et al. (2013) schlossen Personen mit oben genannten Hepatitiden aus und untersuchten nur Patient*innen mit einer NAFLD. Marshall et al. (2012), Chauhan et al. (2016) und Wang et al. (2013) inkludierten in Bezug auf die Ätiologien der CLD ein breites Patient*innenspektrum in ihre Studien, wobei Chauhan et al. Personen mit einer höhergradigen Fibrose (Ishak score ≥ 1) ausschloss (Webb et al. 2009).

Das dritte Ultraschallverfahren, das für die Hepatosteatosediagnostik genutzt werden kann, ist der *Controlled Attenuation Parameter* (CAP).

CAP ist ein Messverfahren, das nur mit geeigneten Geräten (Fibroscan, Echosens, Paris, Frankreich) im Rahmen der transienten Elastographie durchgeführt werden kann. Hierbei wird die dorsale Schallabschwächung bzw. -dämpfung quantitativ in dB/m gemessen (Roeb et al. 2015). Eine Metaanalyse aus dem Jahr 2017 analysierte die Daten von 2.735 Patient*innen und ergab Werte für Sensitivität und Spezifität, die jenen der herkömmlichen Sonographie entsprechen (Bohte et al. 2011; Karlas et al. 2017). Die Detailergebnisse sind in Tabelle 8 zu sehen.

Tabelle 8: Diagnostische Genauigkeit des Controlled Attenuation Parameters (CAP)

	Sensitivität (95 % KI)	Spezifität (95 % KI)	Optimaler Cut-off-Wert ¹ (95 % KI)
$S^2 \geq 1$	0,668 (0,660 - 0,750)	0,822 (0,761 - 0,897)	248 (237 - 261)
$S^2 \geq 2$	0,773 (0,690 - 0,838)	0,812 (0,749 - 0,879)	268 (257 - 284)
$S^2 = 3$	0,882 (0,765 - 0,956)	0,776 (0,720 - 0,821)	280 (268 - 294)

¹Wert in dB/m; ²S1 - S3: Steatosegrade nach histologischem Klassifikationsstandard; 95 % KI: 95 % Konfidenzintervall (basierend auf Karlas et al. 2017). Die Verwendung erfolgt mit freundlicher Genehmigung von Elsevier.

CAP birgt als Messverfahren jedoch einige Nachteile: CAP ist nicht breit verfügbar, da die Technik nur in Fibroscangeräten integriert ist und die Messwerte können von Kovariablen, wie Diabetes oder BMI, beeinflusst werden (Karlas et al. 2017). Ebenso ist anzumerken, dass die Messung bei einem relevanten Teil der Personen fehlschlägt. Dies zeigt eine Studie von De Lédinghen et al. (2014) in der bei 7,7 % der Messungen aufgrund von unverwertbaren Ergebnissen fehlschlugen.

1.3.2.5 Fatty Liver Index (FLI)

Eine Studiengruppe rund um Giorgio Bedogni hat anhand von leberkranken Patient*innen mehrere potenzielle Prädiktoren der Fettlebererkrankung identifiziert. Aus den einzelnen Variablen wurde ein Modell bestehend aus BMI, Hüftumfang, Triglyceriden und Gamma-Glutamyltranspeptidase (γ -GT) entwickelt, der einen Punktwert auf einer Skala von 0 bis 100 ergibt. Der sog. *Fatty Liver Index* (FLI) steht als frei nutzbares Messwerkzeug zu Verfügung und weist eine Messgenauigkeit von 0,84 % (95 % KI 0,81 – 0,87) auf, um eine Fettleber zu detektieren. Bei einem FLI von < 30 kann eine Fettleber mit einer Sensitivität von 87 % und einer *Negative Likelihood Ratio* von 2,9 ausgeschlossen und bei einem Wert von ≥ 60 mit einer Spezifität von 86 % und einer *Positive Likelihood Ratio* von 4,3 diagnostiziert werden (Bedogni et al. 2006; Roeb et al. 2015).

1.3.2.6 SteatoTest

Der SteatoTest ist ein kommerzielles Messwerkzeug, das von der Forschungsgruppe um Thierry Poynard entwickelt wurde, um Leberverfettung zu diagnostizieren (Poynard et al. 2005). Ähnlich dem FLI wurden 12 Parameter bestimmt anhand derer ein Score vergeben wird. Die Vorteile dieses Tests sind geringere Kosten im Vergleich zu Biopsie oder MRT und eine hohe Untersucherunabhängigkeit (Munteanu et al. 2008). Es gelang ein Modell zu entwickeln, das eine Hepatosteatose $\geq 5\%$ je nach Kontrollgruppe mit einer Sensitivität von 0,85 bis 1,00 und einer Spezifität von 0,83 bis 1,00 erkennt (Poynard et al. 2005). In einer 2019 veröffentlichten Publikation wurde die 2. Generation, der Steatotest -2, präsentiert. Ziel der Überarbeitung war es einen mindestens gleichwertigen Test zu entwickeln, der ohne die Parameter BMI und Bilirubin auskommt. Grund für den Ausschluss der beiden Kriterien war die Variabilität in der Erfassung von Höhe und Gewicht, sowie eine Verzerrung des Bilirubinwertes bei Hämolyse und Gilbert-Syndrom. Dies ist der Studie gelungen. Ebenso konnte ein neuer *Cut-off*-Wert zur Bestimmung einer Hepatosteatose ($> 5\%$ der Hepatozyten) mit einer Sensitivität von 79 % und einer Spezifität von 50 % bestimmt werden (Poynard et al. 2019). Als Nachteil kann genannt werden, dass für diese Untersuchung die Bestimmung vieler Blutwerte notwendig ist und dies mit einem Zeit- und Kostenaufwand einhergeht. Eine Gegenüberstellung der Parameter beider Tests ist in Tabelle 9 zu sehen.

Tabelle 9: Parameter des SteatoTest-1 und -2

Biomarker	SteatoTest-1	Steatotest-2
Alter	X	X
Geschlecht	X	X
α -2-Makroglobulin (g/l)	X	X
Haptoglobin (g/l)	X	X
Apolipoprotein A1 (g/l)	X	X
Blutzucker (mmol/l)	X	X
GGT (IU/l)	X	X
ALT (IU/l)	X	X
Gesamtcholesterin (mmol/l)	X	X
Triglyceride (mmol/l)	X	X
Gesamt Bilirubin (μ mol/l)	X	
BMI (kg/m ²)	X	

GGT: Gamma-Glutamyltransferase; ALT: Alanin-Aminotransferase; BMI: Body-Mass-Index (basierend auf Poynard et al. 2005; basierend auf Munteanu et al. 2008; basierend auf Poynard et al. 2019).

1.4 Fragestellung

Das Ziel der folgenden Arbeit war die Evaluation der diagnostischen Genauigkeit des hepato-renalen Index zur Detektion und Graduierung einer Hepatosteatose bei Patient*innen mit CLD mit der Histologie als Referenzstandard. In Abgrenzung zu bereits bestehenden Arbeiten erfolgte die Bestimmung des HRI mit *High-End*-Ultraschallgeräten und der Einschluss von Patient*innen nicht nur mit NAFLD, HBV oder HCV, sondern auch mit weiteren Lebererkrankungen. Ferner wurde der Einfluss von Leberfibrose auf die diagnostische Messgenauigkeit des hepato-renalen Index und die Korrelation weiterer Parameter mit dem histologisch bestimmten Steatosegrad untersucht. Zusätzlich erfolgte die Bestimmung der Inter-raterreliabilität.

2 Material und Methoden

Ziel dieser Studie war es, einen möglichen Zusammenhang zwischen dem mit *High-End*-Ultraschallgeräten bestimmten hepatorenen Index und dem histologisch bestimmten Steatosegrad festzustellen und zu quantifizieren.

Zur Beantwortung dieser Fragestellung wurde das Studiendesign einer retrospektiven, einfach verblindeten Studie gewählt. Es wurden Auswahlkriterien, wie unten beschrieben, festgelegt, nach denen das Patient*innenkollektiv durchsucht wurde. Wurden diese erfüllt, so wurden Hepatosteatosegrad, HRI und weitere relevante Kovariablen erhoben und in eine anonymisierte Datenbank aufgenommen.

Die statistische Auswertung erfolgte mit dem Statistikprogramm SPSS Version 26 (IBM, Armonk, NY, USA).

Die Studie wurde von der Ethikkommission der Universitätsmedizin Göttingen genehmigt (Registrierungsnummer - 28/3/18). Der Antrag auf Verzicht einer schriftlichen Einwilligung der Studienteilnehmenden wurde ebenso bewilligt. Die Studie entspricht den Vorgaben der Deklaration von Helsinki und jenen der örtlichen Gesetze.

2.1 Primäres Auswahlkriterium: Transkutane Leberbiopsie

Als primäres Aufnahmekriterium galt eine stattgehabte Leberstanzbiopsie an der Universitätsmedizin Göttingen im Zeitraum von 2.3.2016 bis 28.5.2019. Aus der Studie wurden Patient*innen ausgeschlossen, deren Leberpunktion zur Abklärung einer Raumforderung durchgeführt wurde.

Die Biopsien wurden mit einer 18-gauge semi-automatischen Biopsienadel (BioPince, Argon Medical devices, USA) entnommen. Die Probenentnahme erfolgte beim rechten Leberlappen transkostal oder subkostal und beim linken nur subkostal.

Die Proben wurden von erfahrenen Patholog*innen der Universitätsmedizin Göttingen befundet. Patient*innen mit einem Probenzylinder, der kürzer als 15 mm war, oder deren Proben nicht repräsentativ waren, wurden aus der Studie ausgeschlossen.

Die histologische Einteilung des Steatosegrades wurde durch die Patholog*innen wie folgt festgelegt: Grad 0 (< 5 % Hepatozyten mit Fetteinlagerungen), Grad I (\geq 5 % - 33 %), Grad II (> 33 % - 66 %) und Grad III (> 66 %) (Kleiner et al. 2005).

2.2 Sekundäres Auswahlkriterium: Ultraschallbilder

Es wurde sonographisches Bildmaterial von allen Patient*innen, die die Aufnahmekriterien in die Studie erfüllt haben (s. o.), durch einen erfahrenen Untersucher gesichtet.

Das Bildmaterial stammte hierbei von den *High-End*-Ultraschallgeräten Logiq E9 (GE Medical Systems, Wauwatosa, USA; Software R1.0.6) und Hitachi ALOKA (ProSound Alpha 7). Bei der Sichtung wurden nur Aufnahmen im B-Modus berücksichtigt.

Es wurden Patient*innen ausgeschlossen bei denen keine Ultraschallbilder auffindbar waren, der Zeitraum zwischen Stanzbiopsie und Abdomensonographie größer als ein Jahr war oder wenn die Aufnahmen mit *Non-High-End*-Geräten zustande gekommen waren.

Für die Bestimmung des HRI wurden Aufnahmen benötigt, die sowohl Leber als auch Niere in einem Bild darstellen. Dementsprechend wurden bei Personen, die die Auswahlkriterien bis hier hin erfüllten nach Aufnahmen, die beide Organe in einem Bild zeigten, gesucht. Im weiteren Verlauf wurden Personen aus dem Studienkollektiv ausgeschlossen, bei denen die Bildqualität der Aufnahmen inadäquat waren. Eine inadäquate Bildqualität lag vor, wenn es nicht möglich war nach unten genannten Kriterien den HRI zu bestimmen. Ebenso wurden Personen ausgeschlossen deren Biopsieprobenmaterial keine histologische Bestimmung des Steatosegrades zuließ. Eine vollständige Übersicht des Selektionsprozesses des Studienkollektivs ist in Abbildung 5 zu sehen.

Gab es bei Patient*innen mehrere sonographische Untersuchungen des Abdomens mit gleichwertigen Bildern, so wurden näher zum Biopsietermin liegende Aufnahmen bevorzugt.

Ebenso wurden Bilder, deren Schnittebene längs in der Medioklavikularline lagen gegenüber anderen präferiert, sofern die Bildqualität äquivalent war.

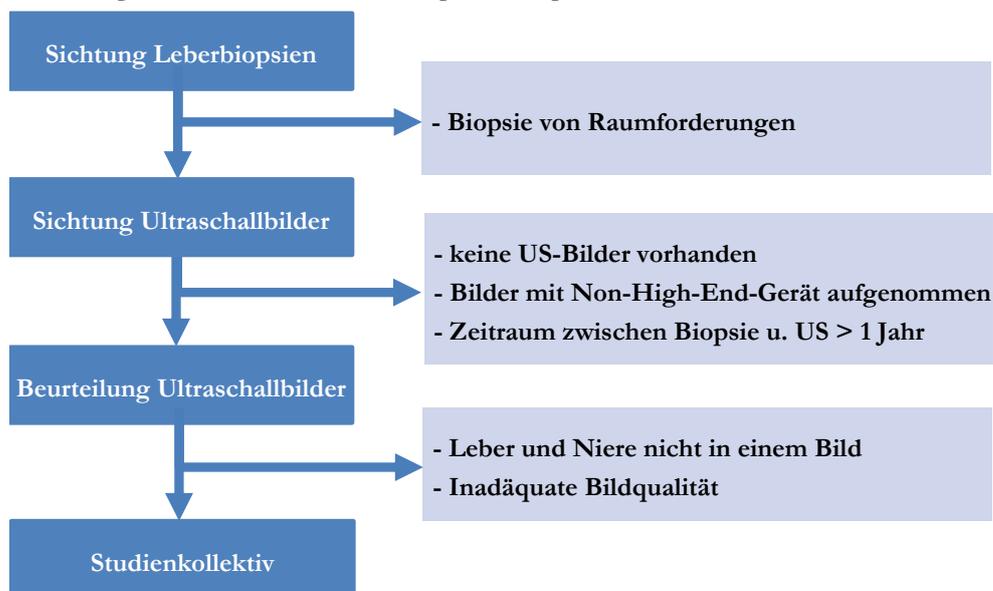


Abbildung 5: Auswahlkriterien des Studienkollektivs. Dunkelblaue Felder: Selektionsprozess des Studienkollektivs; hellblaue Felder: Ausschlusskriterien.

2.3 Bestimmung des hepatorenenalen Index

Der HRI wurde mittels PACS-Software (*Picture archiving and communication system*) Centricity healthcare web 3.0 (GE Healthcare Centricity) bestimmt. Für diesen Zweck wurde die Messfunktion „ROI-Rechteck“ genutzt. ROI-Rechtecke können beliebig im Bild gesetzt und in ihrer Größe modelliert werden. Für jedes ROI-Rechteck kann eine Pixelstatistik über die Bildpunkte innerhalb dieser definierten Fläche ausgegeben werden. Ein Beispiel ist in Abbildung 6 zu sehen.

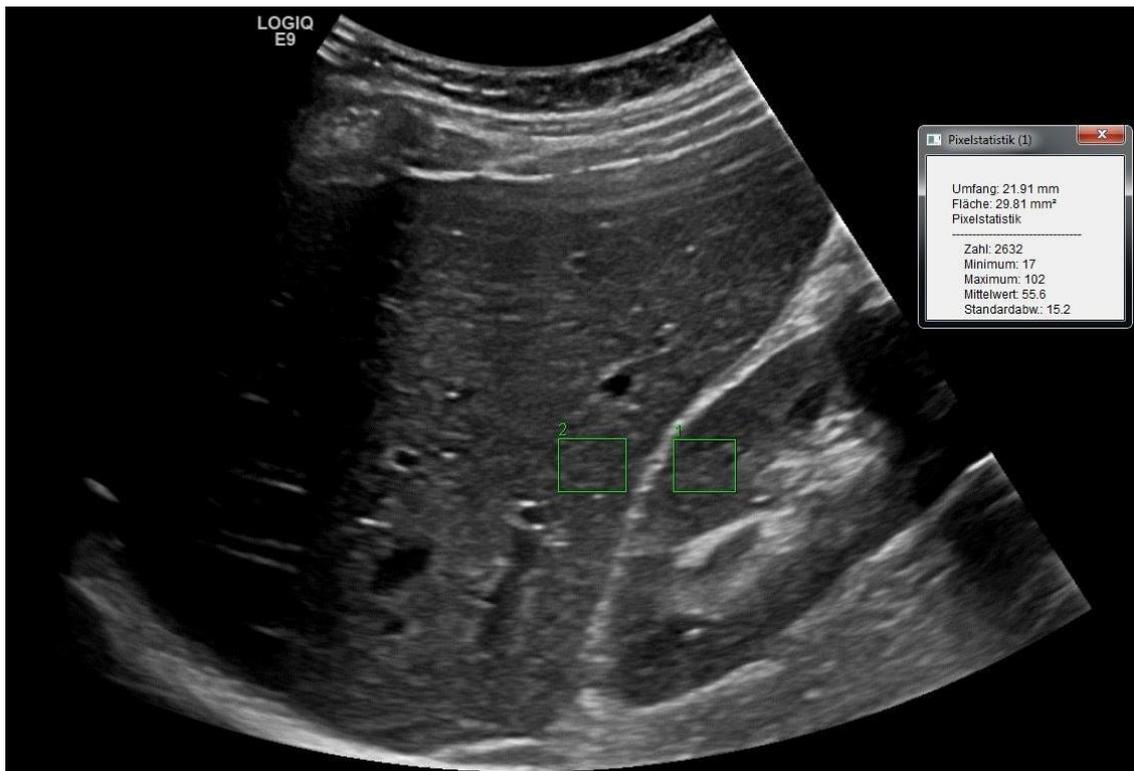


Abbildung 6: Ultraschallbild der Leber und der linken Niere mit Pixelstatistik und ROI-Rechtecken. Grüne Rechtecke: Regions of Interest.

Zur Bestimmung des HRI wurden sowohl ein ROI-Rechteck über dem Nieren- als auch eines über dem Leberparenchym gesetzt. Ein Beispiel für die Platzierung von zwei ROI-Rechtecken ist in Abbildung 6 zu sehen.

Die Platzierung der ROI-Rechtecke erfolgte nach festgelegten Regeln. Beide ROI-Rechtecke mussten mit der gleichen Entfernung (r) zur Schallsonde gesetzt werden (siehe Abbildung 7). Ebenso muss die Höhe (h) der Rechtecke gleich groß sein (siehe Abbildung 7). Die ROI-Rechtecke mussten über einem homogenen, repräsentativen Parenchymabschnitt gesetzt werden. Das Platzieren über Gefäßen musste vermieden werden. Die Distanz (d) zwischen den beiden ROI-Rechtecken sollte möglichst geringgehalten werden. Die ROI-Rechtecke mussten unter Berücksichtigung der oben genannten Regeln möglichst groß gewählt werden, um eine aussagekräftigere Pixelstatistik zu erzielen. Abbildung 6 und Abbildung 7 dienen als Beispiel für die regelrechte Platzierung der ROI-Rechtecke.

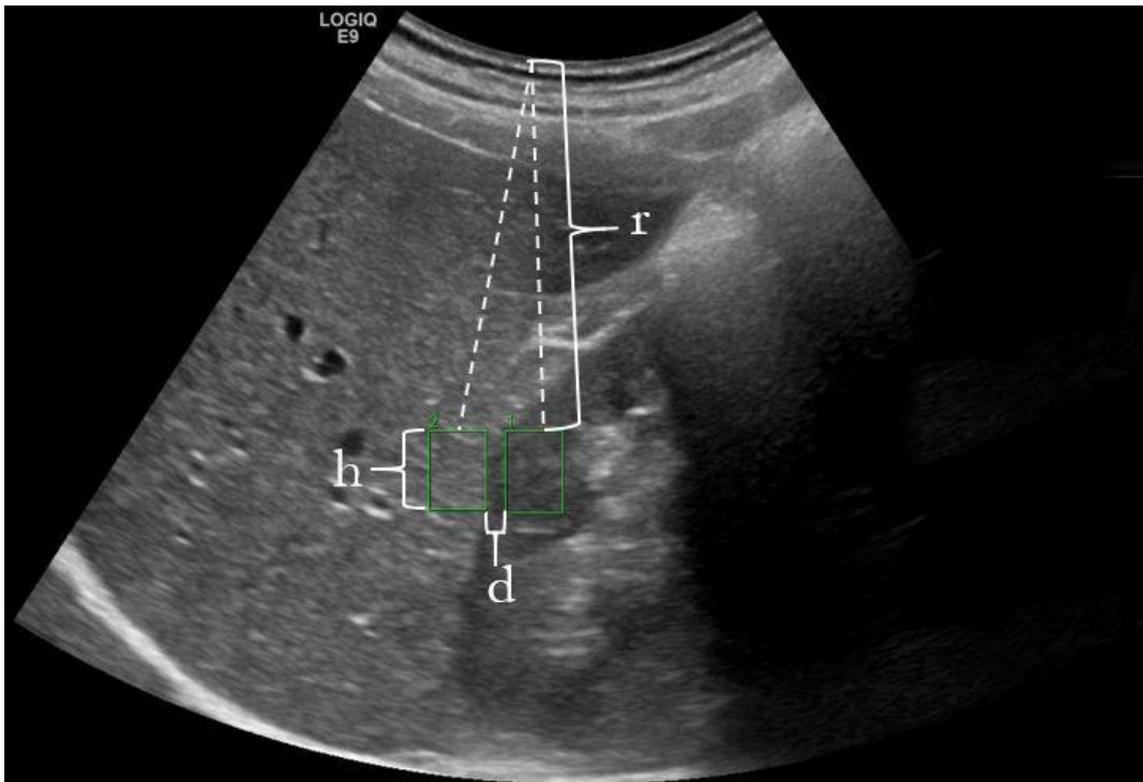


Abbildung 7: Platzierungskriterien der ROI-Rechtecke. Beide ROI-Rechtecke sollten die gleiche Entfernung zur Schallsonde und die gleiche Höhe haben, die Distanz zwischen den beiden ROI-Rechtecken sollte geringgehalten werden, die ROI-Rechtecke sollten über homogenen Parenchymabschnitten platziert werden; r: Entfernung von ROI-Rechteck zur Schallsonde, h: höhe der ROI-Rechtecke, d: Distanz zwischen ROI-Rechtecken

Für die Positionierung der ROI-Rechtecke wurden zusätzlich zu den gültigen Regeln bevorzugte Positionen definiert. Demnach sollte das ROI-Rechteck der Niere auf Höhe der Mitte des Nierenparenchyms liegen. Das ROI-Rechteck der Leber sollte auf gleicher Höhe mit der kürzesten Distanz zum Nieren-ROI-Rechteck platziert werden. Die ROI-Rechtecke in Abbildung 6 und 7 entsprechen der bevorzugten Position.

Beim Bearbeiten des Bildmaterials konnte aufgrund von Bildgegebenheiten nicht immer die bevorzugte Position berücksichtigt werden. Dies war vor allem der Fall, wenn ein inhomogener Parenchymabschnitt, Schallschatten von z. B. Rippen oder eine Messlinie im Idealbereich lagen. Infolgedessen wurde bei der Positionierung der ROI-Rechtecke mit beiden Rechtecken nach oben oder unten oder nur mit einem nach rechts oder links ausgewichen. Beispiele hierfür sind in Abbildung 8 zu sehen.

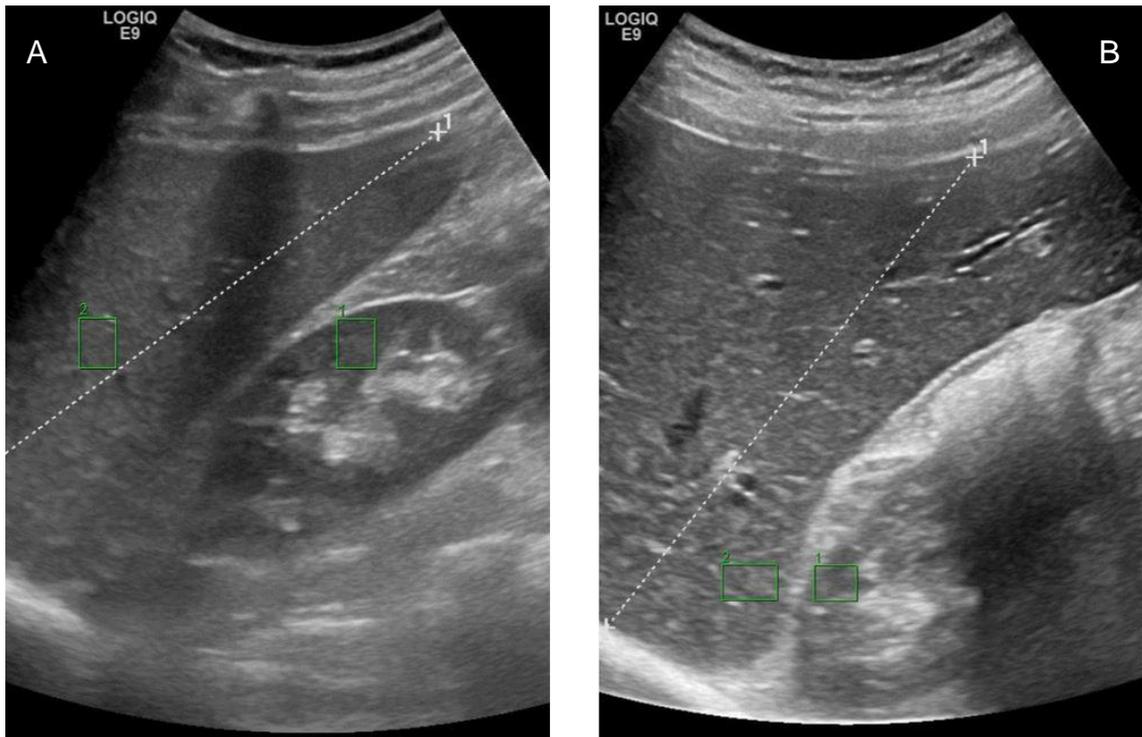


Abbildung 8: Alternativplatzierungen der ROI-Rechtecke. A: sowohl eine Messlinie als auch ein Schallschatten verhindern das Platzieren des Leber-ROI-Rechtecks in der bevorzugten Position – es wurde deshalb nach links ausgewichen; B: auf Höhe der Mitte des Nierenparenchyms liegt ein Schallschatten – es wurden deshalb beide ROI-Rechtecke weiter unten platziert.

Aus der für jedes ROI-Rechteck ausgegebenen Pixelstatistik wurden Mittelwert, Standardabweichung und Fläche in die Datenbank übernommen. Zusätzlich wurde bei jeder Person die Position der ROI-Rechtecke und die Ausgabe der Pixelstatistik mittels *Screenshots* dokumentiert und anonymisiert abgespeichert.

Anhand der erhobenen Werte wurde der HRI wie folgt berechnet:

$$\text{hepatorenenaler Index} = \frac{\text{Mittelwert Leber}}{\text{Mittelwert Niere}}$$

2.4 Erfassung der Interraterreliabilität

An 30 Personen wurden Bildauswahl und Bestimmung des HRI von zwei Untersuchern mit unterschiedlichen Erfahrungsgraden doppelt verblindet vorgenommen. Ein Untersucher hatte bis dahin mehr als 5.000 Ultraschalluntersuchungen vorgenommen, der Zweite ca. 100 unter Supervision.

2.5 Erstellung der Datenbank

Erfüllten die Patient*innen die Einschlusskriterien bei Abwesenheit von Ausschlusskriterien, so wurden neben deren HRI und histologischen Steatosegrad weitere Kovariablen aus deren Patient*innenakte mittels des krankenhausinternen Datenverwaltungsprogramms Ixserv (ix.mid Software Technologie GmbH, Köln, Deutschland) ermittelt und in die Datenbank eingepflegt. Eine vollständige Auflistung aller in die Datenbank aufgenommenen Werte ist in Tabelle 10 zu sehen.

Tabelle 10: Auflistung aller Parameter der Studiendatenbank

Bildanalyse	Biometrische Daten und Befunde
Datum Ultraschalluntersuchung	Geschlecht
Mittelwert ROI-Rechteck Leber	Alter (Jahre)
SD ROI-Rechteck Leber	Größe (cm)
Fläche ROI-Rechteck Leber	Gewicht (kg)
Mittelwert ROI-Rechteck Niere	BMI (kg/m ²)
SD ROI-Rechteck Niere	Diabetes mellitus (ja/nein)
Fläche ROI-Rechteck Niere	
Hepatorenaler Index Untersucher 1	Ätiologie der Lebererkrankung
Hepatorenaler Index Untesucher 2	
Verwendetes Ultraschallgerät	Laborwerte
Pathologiebefund	Gesamtbilirubin (mg/dl)
Datum Leberstanzbiopsie	Aspartat-Aminotransferase (U/I)
Histologischer Steatosegrad 0 - III	Alanin-Aminotransferase (U/I)
Fibrosegrad Desmet u. Scheuer (0 – 4)	Alkalische Phosphatase (U/I)
Grading Desmet u. Scheuer: (0 – 4)	Gamma-Glutamyltransferase (U/I)

ROI: Region Of Interest

2.6 Statistische Auswertung

Die Studiendaten wurden in einer Excel-Tabelle zusammengefasst. Die statistische Auswertung erfolgte mit dem Statistikprogramm SPSS Version 24 (IBM, Armonk, NY, USA). Zur Auswahl der geeigneten statistischen Verfahren wurde die Informationsseite zur Methodenberatung der Universität Zürich genutzt (Schwarz et al. 2018).

Um für die Auswertung ein einheitliches Definitionssystem einzuführen wurde der Begriff „keine Steatose“ als „S0“ bezeichnet. Die histologischen Steatosegrade I - III wurden dementsprechend „S1“, „S2“ und „S3“ genannt. Diese Definitionen werden ebenso im Kapitel „Ergebnisse“ verwendet.

Als Grundlage der Weiteren statistischen Verfahren musste ermittelt werden, ob die erhobenen HRI-Werte normalverteilt sind oder nicht. In der statistischen Analyse mittels deskriptiver Statistik stellte sich heraus, dass keine Normalverteilung vorliegt.

Um die Mittelwerte der hepatorealen Indices mit den Steatosegraden vergleichen zu können, wurde der Kruskal-Wallis-Test auf den Datensatz angewandt. Ziel war es, festzustellen, ob zwischen den HRI-Werten der histologisch bestimmten Gruppen (S0-S3) ein signifikanter Unterschied besteht. Der Test kann verwendet werden, wenn mehr als zwei unabhängige Stichprobengruppen, in diesem Fall die der Hepatosteato-segrade, vorliegen, deren Mittelwerte nicht normalverteilt und mindestens ordinalskaliert sind. Die Mittelwerte werden nach ihrer Größe sortiert und aufsteigenden Rängen zugeteilt. Die Berechnung erfolgt dann anhand der Rangsumme für die jeweilige Steatosegruppe. Da der Datensatz die Kriterien einer asymptotischen Chi-Quadrat-Verteilung erfüllt, wurde die Signifikanz anhand des entsprechenden kritischen Wertes berechnet (Schwarz et al. 2020).

Eine statistische Signifikanz wurde vom Studienteam bei $p < 0,05$ festgelegt.

Um die signifikanten Unterschiede einzelnen Gruppenpaaren, wie z. B. Hepatosteato-segrad I und II, zuordnen zu können, wurden Post-hoc-Tests in Form eines Dunn-Bonferroni-Tests durchgeführt.

Um den HRI als Messwerkzeug zur Diagnose und Einstufung einer Steatose nutzen zu können, ist die Bestimmung von *Cut-off*-Werten notwendig. Hierfür wurde zuerst definiert zwischen welchen Gruppen der Wert trennen soll. Es wurde entschieden, dass ein eigener *Cut-off*-Wert zur Differenzierung zwischen keiner Steatose und Steatose ($S \geq 1$), sowie einer für alle Steatosegrade, die größer als Grad I sind ($S \geq 2$) und zur Bestimmung einer drittgradigen Steatose ($S = 3$) ermittelt werden soll.

Sensitivität und Spezifität des HRI zur Vorhersage von Leberverfettung bzw. des Grades der Steatose wurden mit einer Grenzwertoptimierungskurve (ROC-Kurve) dargestellt. Die diagnostische Genauigkeit wurde anhand der Fläche unter dieser Kurve (AUROC) bestimmt. Hierbei kann ein Wert von 0 bis 1 erreicht werden. Ein Wert von 0,5 spricht für keinen Zusammenhang und somit eine schlechte Sensitivität und Spezifität.

Zur Ermittlung des besten *Cut-off*-Wertes des HRI um einen Hepatosteatosegrad ≥ 1 , ≥ 2 und $= 3$ zu bestimmen, wurde der Youden-Index gebildet. Hierfür wurden für jeden HRI-Wert Sensitivität und Spezifität ermittelt, zueinander addiert und der Wert 1 von der Summe abgezogen.

$$\text{Youden Index} = \text{Sensitivität} + \text{Spezifität} - 1$$

Es können Werte zwischen 0 und 1 erreicht werden. Der HRI mit dem größten Summenwert wurde als *Cut-Off*-Wert festgelegt. Sowohl die Untersuchung der Gruppenpaare als auch die Ermittlung von *Cut-Off*-Werten wurde am gesamten Datensatz als auch an Subpopulationen vorgenommen. Hierfür wurde eine gesonderte Auswertung nach Ultraschallgerättyp, Ätiologie der Lebererkrankung, BMI und Fibrosegrad nach Desmet und Scheuer vorgenommen. Grund für die Analyse der Subpopulationen war die Frage, ob sich die Messgenauigkeit durch die Verwendung eines bestimmten Ultraschallgerätes oder durch eine Vorselektion anhand anderer klinischer Parameter wie BMI, Ätiologie oder Leberfibrose steigern lässt.

Um zu überprüfen, ob es einen Zusammenhang zwischen Hepatosteatosegrad bzw. HRI und anderen Kovariablen gibt, wurde die Rangkorrelation nach Spearman auf den gesamten Datensatz angewandt. Alle Variablen, die einen p-Wert $< 0,1$ hatten, wurden zudem in ein multiples Regressionsanalysemodell integriert, um einen unabhängigen Zusammenhang zu überprüfen.

Die Interraterreliabilität wurde anhand des intraclass correlation coefficient (ICC) überprüft. Die Bewertung des ICC wurde mit „schlecht“ bei $\text{ICC} = 0,0 - 0,20$), „mäßig“ ($\text{ICC} = 0,20 - 0,40$), „moderat“ ($\text{ICC} = 0,40 - 0,75$) und „exzellent“ ($> 0,75$) definiert.

3 Ergebnisse

Im Zeitraum von Beginn der Datenerhebung im März 2016 bis zu deren Ende im Mai 2019 wurden an der Universitätsmedizin Göttingen 231 Leberbiopsien von Lebererkrankten durchgeführt. Biopsien von Raumforderungen wurden bereits vor der Datenerhebung ausgeschlossen.

3.1 Studienkollektiv und Ablauf

Nach Sichtung des Bildmaterials und Beurteilung der Biopsien wurden 157 Personen in die Datenbank aufgenommen. Bei 71 Personen erfüllten die Ultraschallbilder nicht die im Methodikteil beschriebenen Auswahlkriterien und waren somit nicht verwertbar. Drei Personen mussten aufgrund von inadäquaten Biopsieproben nachträglich ausgeschlossen werden. Der Ablauf der Teilnehmendenselektion und statistischen Auswertung ist in Abbildung 9 dargestellt.

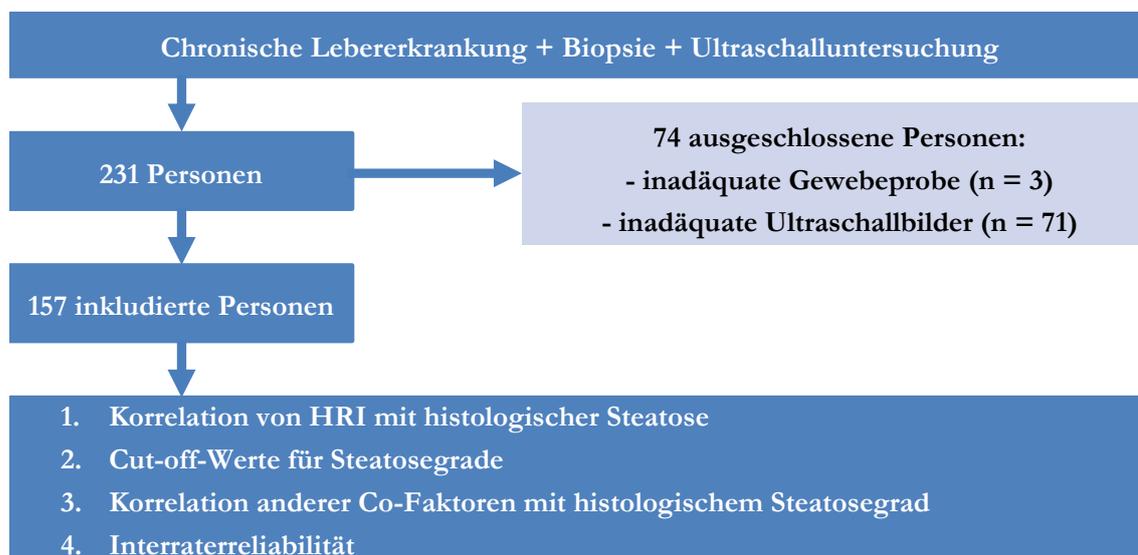


Abbildung 9: Workflow Studiendesign. Dunkelblaue Kästchen: Stationen im Studienablauf; hellblaue Kästchen: ausgeschlossene Personen

In dieser Arbeit wird der Begriff „keine Steatose“ mit „S0“ umschrieben. Die Steatosegrade I-III werden als „S1“, „S2“ und „S3“ definiert.

Von den 157 Personen waren 87 Frauen und 70 Männer. Dies entspricht einem Verhältnis von 55,4 % Frauen zu 44,6 % Männern. Die Studienteilnehmer*innen waren zwischen 18 und 85 Jahren alt. Der Mittelwert lag bei 48,2 Jahren mit einer Standardabweichung von 15,3. Der Median betrug 49 Jahre. Der Body-Mass-Index betrug im Mittel 27,7 kg/m² mit einem Minimum von 17,2 kg/m² und Maximum von 49,6 kg/m². Die Standardabweichung lag bei 5,6 kg/m² und der Median bei 27,5 kg/m². Aus den 157 Personen hatten 51 einen BMI < 25 kg/m², 65 einen BMI 25 - 30 kg/m² und 41 einen BMI > 30 kg/m². Eine Übersicht der Charakteristika des Studienkollektivs ist in Tabelle 11 zu sehen.

Tabelle 11: Epidemiologische Daten des Studienkollektivs

Kategorie	Median	Mittelwert	SD	Minimum	Maximum
Alter (Jahre)	49	48,23	± 15,29	18	85
Größe (cm)	170	171,00	± 9,38	140	192
Gewicht (kg)	79	80,87	± 17,53	43	152
BMI (kg/m ²)	27,47	27,66	± 5,55	17,2	49,63

Unter den 157 Teilnehmenden war die häufigste Ätiologie der Lebererkrankungen NAFLD mit 42-, gefolgt von AIH mit 30 Personen. Am dritthäufigsten waren alkoholische Lebererkrankung, Hepatitis C und Erkrankungen unbekannter Ätiologie mit jeweils 18 Personen. Einen kleinen Bestandteil des Patient*innenkollektivs machten die Ätiologien arzneimittelinduzierte Leberschäden mit 10, primär biliäre Cholangitis mit 6, primär sklerosierende Cholangitis mit 5, Hepatitis B mit 4, ischämische Cholangiopathie mit 3 und Alpha-1-Antitrypsinmangel mit 2 erkrankten Personen aus. Eine einzelne Person hatte eine Hämochromatose. Eine Übersicht über die Verteilung der Lebererkrankungen ist in Abbildung 10 zu sehen.

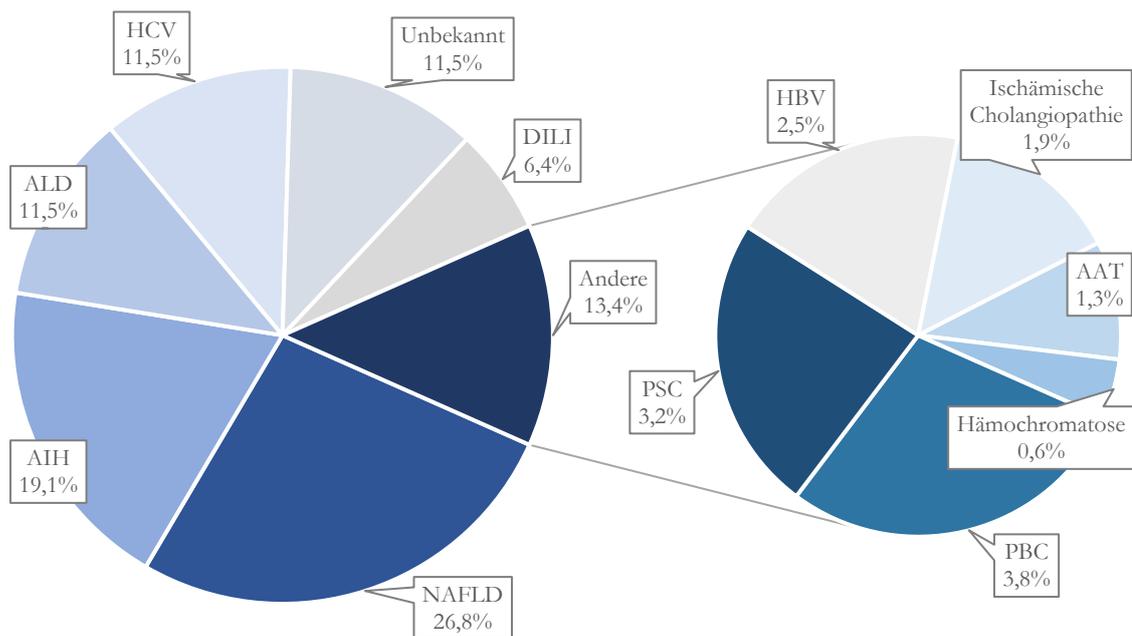


Abbildung 10: Prozentuelle Verteilung der Lebererkrankungen im Patient*innenkollektiv. NAFLD: nicht alkoholisch Fettlebererkrankung; AIH: Autoimmunhepatitis; ALD: alkoholische Lebererkrankung; HCV: Hepatitis-C-Virus; unbekannt: unbekannter Ätiologie; DILI: arzneimittelinduzierte Leberschäden; PBC: primär biliäre Cholangitis; PSC: primär sklerosierende Cholangitis; HBV: Hepatitis-B-Virus; AAT: Alpha-1-Antitrypsinmangel.

Eingeteilt nach histologisch bestimmtem Steatosegrad war die größte Gruppe im Gesamtkollektiv jene ohne Steatose mit 75 Personen. Knapp ein Viertel des Studienkollektivs, 38 Teilnehmende, hatte eine Steatose Grad I. Mit 25 Personen war die Gruppe mit Steatosegrad

II die drittgrößte und jene mit Grad III und 19 die kleinste. Demnach liegt in diesem Studienkollektiv eine Hepatosteatoseprävalenz von 52,2 % vor.

Von den 157 Ultraschalluntersuchungen wurden 121 mit dem Gerät Logiq E9 (GE Medical Systems, Wauwatosa, USA; software R1.0.6) und 36 mit Hitachi ALOKA (ProSound Alpha 7) durchgeführt.

Aus den 121 Personen, die mit dem Gerät Logiq untersucht wurden, hatten 57 Personen keine histologisch nachweisbare Steatose, 30 wiesen eine Hepatosteatose ersten Grades, 19 eine Steatose zweiten Grades und 15 eine dritten Grades.

Die Aloka-Gruppe umfasste nur 36 Personen. Von den 36 Personen hatten 18 keine Steatose, 8 eine leichte, 6 eine mittelgradige und 4 eine schwere Steatose. Die Verteilung der histologischen Steatosegrade auf des Gesamtkollektiv und die Logiq- und Alokakohorte wird in Tabelle 12 dargestellt.

Tabelle 12: Verteilung der histologischen Steatose auf das Gesamtkollektiv und die Logiq- und Alokakohorte

Kollektiv	n gesamt	n S0 ¹	n S1 ¹	n S2 ¹	n S3 ¹
Gesamt	157	75	38	25	19
Logiq ²	121	57	30	19	15
Aloka ³	36	18	8	6	4

n: Personenanzahl; ¹histologischer Steatosegrad 0 - 3; ²Ultraschallgerät Logiq E9 (GE Medical Systems, Wauwatosa, USA; software R1.0.6); ³Ultraschallgerät Hitachi ALOKA (ProSound Alpha 7).

3.2 Statistische Auswertung

Die statistische Auswertung erfolgte mit dem Statistikprogramm SPSS der Firma IBM (Armonk, NY, USA).

Da die Ultraschalluntersuchungen im Beobachtungszeitraum mit zwei unterschiedlichen *High-End*-Gerättypen durchgeführt wurden, wurde die Analyse des Datenmaterials sowohl für das gesamte Studienkollektiv als auch für jene Personengruppe, deren Bilder mit dem Gerät Logiq E9 zustande kamen, vorgenommen.

Da die Gruppe, die mit dem Ultraschallgerät ALOKA untersucht wurde, nur 36 Personen umfasste, wurde keine weitere statistische Auswertung dieses Kollektivs vorgenommen. Die geringe Stichprobenzahl lässt keinen repräsentativen Vergleich der Gerätetypen zu.

Neben der Logiq-Kohorte wurde eine weitere Subgruppe des Gesamtkollektivs gesondert statistisch analysiert. Dies waren jene Personen, die keine NAFLD als zugrundeliegende Ätiologie ihrer CLD hatten.

3.2.1 Durchschnittsgröße der ROI-Rechtecke

Bei der Platzierung der ROI-Rechtecke über Leber und Niere wurde darauf geachtet, dass eine möglichst große Fläche in einem repräsentativen Parenchymabschnitt erfasst wird. Es wurde dadurch eine durchschnittliche Größe der ROI-Rechtecke über der Leber von 18.16 mm² und über der Niere von 16.72 mm² erreicht.

3.2.2 Durchschnittlicher HRI und HRI-Verteilung nach Steatosegrad

Die Analyse des gesamten Studienkollektivs ergab für den HRI einen Mittelwert von 1,55 und eine Standardabweichung von $\pm 1,49$. Im Vergleich hierzu betragen die Werte der Logiq-Kohorte 1,65 ($\pm 1,67$).

Eine Untersuchung der HRI getrennt nach Steatosegrad und Studienkollektiv ergab in Tabelle 13 dargestellte Ergebnisse für Mittelwert und Standardabweichung.

Tabelle 13: Verteilung des HRI nach Steatosegrad und Studienkollektiv

Steatosegrad	S0 ¹		S1 ¹		S2 ¹		S3 ¹	
Gesamt	1,18	$\pm 0,23$	1,53	$\pm 1,44$	2,06	$\pm 2,62$	2,38	$\pm 1,96$
Logiq ²	1,18	$\pm 0,21$	1,60	$\pm 1,61$	2,35	$\pm 2,96$	2,65	$\pm 2,14$
HRI-Verteilung	M	SD	M	SD	M	SD	M	SD

¹histologischer Steatosegrad (0 - 3); ²Ultraschallgerät Logiq E9 (GE Medical Systems, Wauwatosa, USA; Software R1.0.6)

Ebenso wurde die Verteilung der HRI-Werte in Relation zum Steatosegrad für das Gesamtkollektiv und die Logiq-Gruppe in einer Grafik mit Hilfe von Box-Plots visualisiert. Sie können in Abbildung 11 gesehen werden.

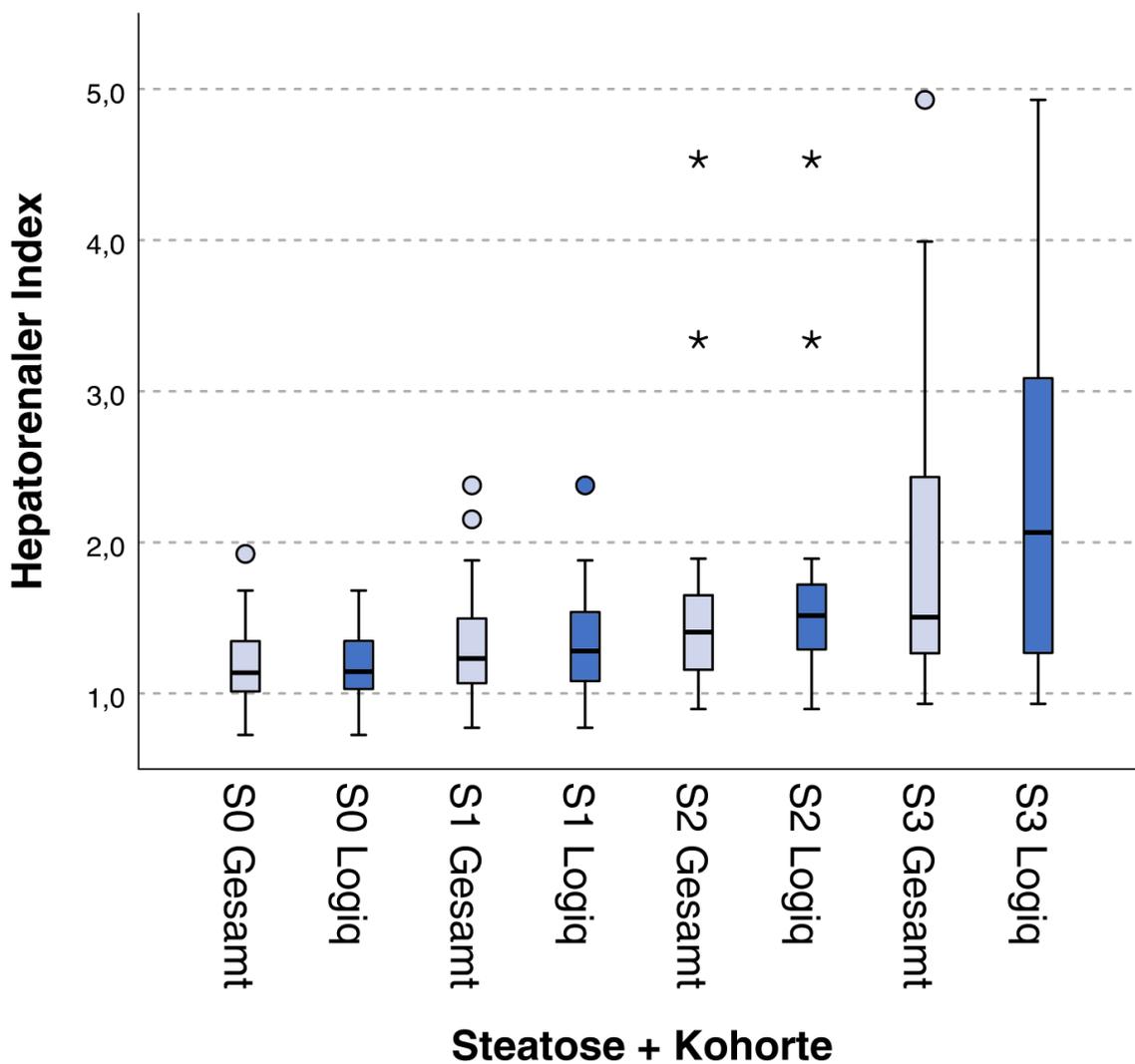


Abbildung 11: Darstellung der Verteilung von HRI-Werten der Logiq-Kohorte und des Gesamtkollektivs als Box-Plots nach histologischem Steatosegrad. Hellblaue Boxen: Verteilung HRI-Werte Gesamtkollektiv; dunkelblaue Boxen: Verteilung HRI-Werte Logiq-Kohorte; Punkte: milde Ausreißer; Sterne: extreme Ausreißer.

3.2.3 Gruppenvergleich

Die hepatorenalen Indices der histologisch bestimmten Steatosegrade S0, S1, S2 und S3 wurden paarweise mittels Kruskal-Wallis Test und Dunn-Bonferroni Post-hoc-Test auf signifikante Unterschiede untersucht. Als signifikant wurden Ergebnisse von $p < 0.05$ gewertet.

Bei der Untersuchung des gesamten Studienkollektivs wurde ein signifikanter Unterschied zwischen der Gruppe S0 gegen S3 ($p < 0,001$) und der Gruppe S0 gegen S2 ($p < 0,014$) festgestellt. Ein Trend zeichnete sich zwischen der Gruppe S1 gegen S3 ab ($p = 0,083$). Zwischen den anderen Gruppenpaaren konnten keine signifikanten Unterschiede beobachtet werden.

Die Analyse der Logiq-Kohorte zeigte ein Ergebnis mit einer höheren Signifikanz für die Paare S0 gegen S3 ($p < 0,001$) und S0 gegen S2 ($p < 0,001$) im Vergleich zum Gesamtkollektiv. Das Paar S1 gegen S3 zeigte dafür keinen eindeutigen Trend an. Die genauen Ergebnisse der Gruppenvergleiche sind in der Tabelle 14 dargestellt.

Tabelle 14: Paarweiser Gruppenvergleich der HRI-Werte der jeweiligen histologischen Steatosegrade

Steatosegruppen (histologisch)	Gesamtkollektiv	Logiq-Kohorte
S0 vs. S1	$p = 0,45$	$p = 0,471$
S0 vs. S2	$p = 0,014$	$p < 0,001$
S0 vs. S3	$p < 0,001$	$p < 0,001$
S1 vs. S2	$p = 1,0$	$p = 0,225$
S1 vs. S3	$p = 0,083$	$p = 0,132$
S2 vs. S3	$p = 1,0$	$p = 1,0$

p: Signifikanzwert

3.2.4 Bestimmung von Cut-off-Werten

Es wurde eine Grenzwertoptimierungsanalyse (ROC) mit den im Methodikteil definierten *Cut-off*-Grenzen ($S \geq 1$, $S \geq 2$, $S = 3$) als Zustandsvariablen auf den Datensatz angewandt. Die Ergebnisse für das gesamte Studienkollektiv und für die Logiq-Gruppe sind in Abbildung 12 zu sehen. Hierbei sind die Kurven beider Kollektive für die jeweilige *Cut-off*-Grenze in eine Grafik integriert und erlauben somit einen direkten Vergleich. Die Fläche (AUROC) unter der jeweiligen ROC-Kurve gibt eine Aussage über die Messgenauigkeit der *Cut-off*-Grenzen. Sie betrug im Gesamtkollektiv für die Grenze zwischen Steatose und keine Steatose ($S \geq 1$) 0,68 mit einem Konfidenzintervall von 0,60 - 0,76. Für die Detektion einer mittelgradigen Steatose oder darüber ($S \geq 2$) ergab sich eine AUROC von 0,71 (0,62 - 0,81). Die Grenze zwischen drittgradiger Steatose und den übrigen Steatosegraden ($S = 3$) erreichte eine AUROC von 0,74 (0,61 - 0,86).

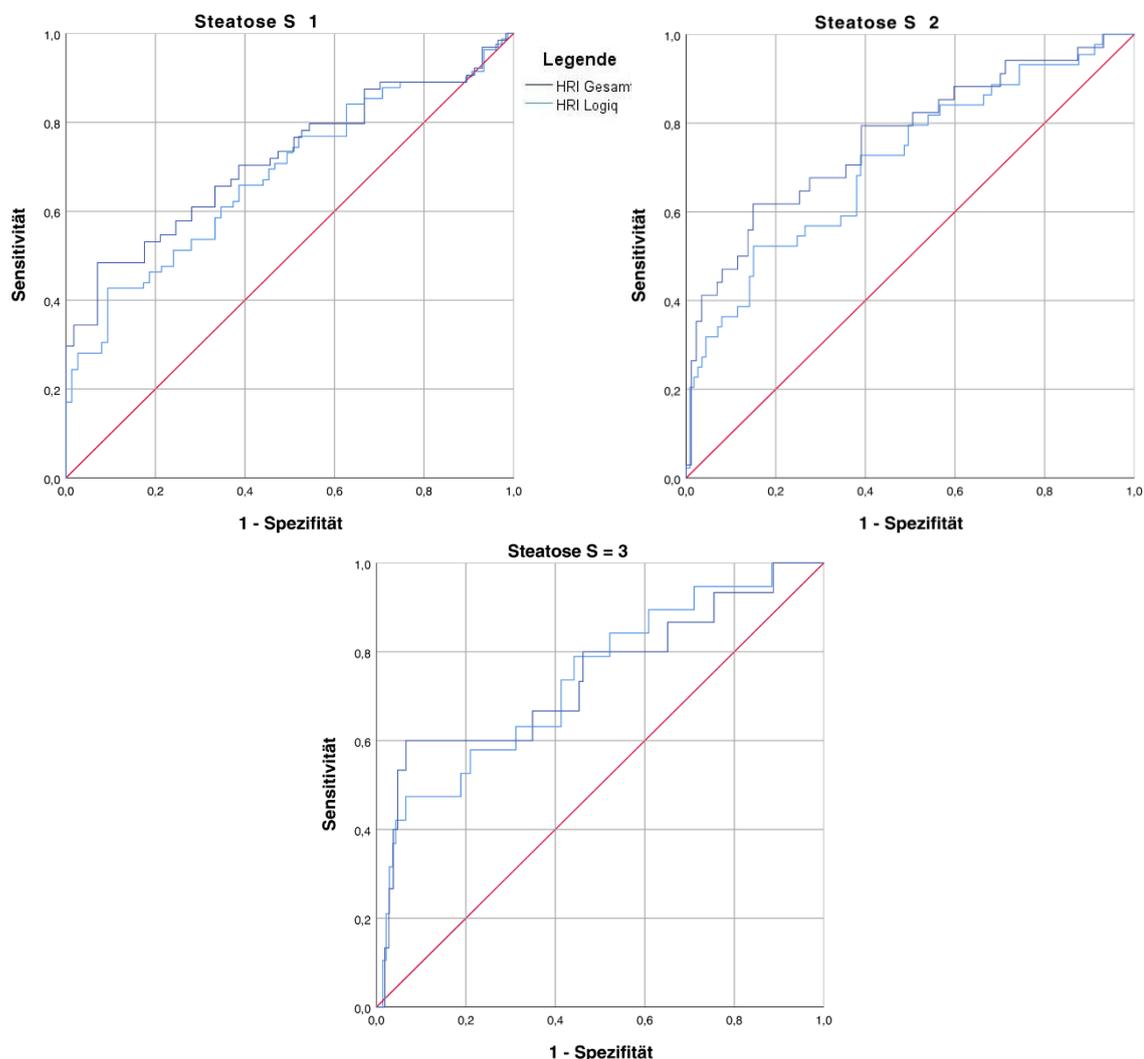


Abbildung 12: ROC-Kurven für die HRI-Cut-off-Werte vom Gesamtkollektiv und der Logiq-Kohorte zur Detektion von $S \geq 1$, $S \geq 2$ und $S = 3$. Rote Querlinie: Bezugslinie

Mit Hilfe der Grenzwertoptimierungsanalyse wurde für jeden HRI -Wert die Sensitivität und die Spezifität ermittelt, mit dem zwischen den oben beschriebenen Gruppen differenziert werden kann. Mit Hilfe des Youden-Index wurde im Anschluss der jeweils beste *Cut-off*-Wert bestimmt.

Für das Gesamtkollektiv wurden zwei *Cut-off*-Werte zur Detektion einer Hepatosteatose ermittelt. Der niedrigere *Cut-off*-Wert ist auf die Sensitivität optimiert, der höhere auf die Spezifität. Mit 85,4 % weist der *Cut-off*-Wert von 1,05 eine hohe Sensitivität auf, jedoch auf Kosten der Spezifität, welche 30,7 % beträgt. Eine hohe Spezifität mit 90,7 % erhält man bei einem *Cut-off*-Wert von 1,46. Die Sensitivität sinkt hierbei jedoch auf 42,7 %. Für den *Cut-off*-Wert von 1,46 ergeben sich dadurch ein positiv prädiktiver- und negativ prädiktiver Wert von 0,834 (PPW) und 0,591 (NPW). Der *Cut-off*-Wert von 1,12 erreicht Werte von 0,574 (PPW) und 0,658 (NPW).

Die besten *Cut-off*-Werte zur Detektion einer Steatose Grad II und darüber, sowie einer ausschließlich drittgradigen Steatose, betragen 1,48 ($S \geq 2$) und 1,79 ($S = 3$). AUROC, Sensitivität, Spezifität, PPW und NPW für alle *Cut-off*-Werte sind in der Tabelle 15 dargestellt.

Tabelle 15: Diagnostische Genauigkeit des HRI zur Prädiktion von Steatose im Gesamtkollektiv

Steatose	AUROC (95 % - KI)	Cut-off-Wert	Sens. (%)	Spez. (%)	PPW	NPW
$S \geq 1$	0,680 (0,597 - 0,763)	1,46	42,7	90,7	0,834	0,591
		1,05	85,4	30,7	0,574	0,658
$S \geq 2$	0,712 (0,619 - 0,805)	1,48	52,3	85,0	0,576	0,821
$S = 3$	0,738 (0,612 - 0,864)	1,79	47,4	93,5	0,501	0,928

AUROC: Fläche unter Grenzwertoptimierungskurve; 95 % KI: 95 % Konfidenzintervall; Sens: Sensitivität; Spez: Spezifität; PPW: positiv prädiktiver Wert; NPW: negativ prädiktiver Wert.

Bei der Analyse der Logiq-Gruppe wurde ebenso vorgegangen wie beim Gesamtkollektiv. Die AUROC für die *Cut-off*-Grenze zwischen Steatose und keine Steatose ($S \geq 1$) ergab einen Wert von 0,72 mit einem 95 % - Konfidenzintervall von 0,62 - 0,81. Die AUROC von $S \geq 2$ betrug 0,76 (0,66 - 0,87) und von $S = 3$ 0,74 (0,59 - 0,90). Somit fielen die Ergebnisse der drei *Cut-off*-Werte der Logiq-Gruppe besser aus als die *Cut-off*-Werte im Gesamtkollektiv. Dies spiegelt sich auch in den in Abbildung 13 gezeigten Vergleichen der ROC-Kurven wider.

Es wurden ebenso wie beim Gesamtkollektiv mit Hilfe des Youden-Index die besten *Cut-off*-Werte bestimmt. Die errechneten Werte waren nahezu identisch zu jenen des Gesamtkollektivs. Es zeigte sich jedoch, dass die Logiq-Gruppe analog zu den AUROC-Werten bessere Ergebnisse für Sensitivität und Spezifität erzielte. Sensitivität und Spezifität, inklusive PPW und NPW für die jeweiligen *Cut-off*-Werte, sind in Tabelle 16 erfasst.

Tabelle 16: Diagnostische Genauigkeit des HRI zur Prädiktion von Steatose in der Logiq-Kohorte

Steatose	AUROC (95 % - KI)	Cut-off-Wert	Sens. (%)	Spez. (%)	PPW	NPW
$S \geq 1$	0,715 (0,624 - 0,807)	1,46	48,4	93,0	0,886	0,616
$S \geq 2$	0,764 (0,664 - 0,865)	1,48	61,8	85,1	0,618	0,851
$S = 3$	0,741 (0,585 - 0,897)	1,79	60,0	93,4	0,563	0,943

AUROC: Fläche unter Grenzwertoptimierungskurve; 95 % KI: 95 % Konfidenzintervall; Sens: Sensitivität; Spez: Spezifität; PPW: positiv prädiktiver Wert; NPW: negativ prädiktiver Wert.

3.2.5 Subgruppe Personen ohne NAFLD

115 Personen hatten andere Lebererkrankungen als NAFLD. Im weiteren Text wird dieses Kollektiv als Nicht-NAFLD-Kohorte bezeichnet. Aus dieser Subgruppe hatten 75 keine Steatose, 24 eine Steatose ersten Grades, 11 eine zweitgradige und 5 eine drittgradige Steatose. Die Verteilung unterscheidet sich somit zu jener des Gesamtkollektivs dadurch, dass sie einen Überhang an Personen ohne Steatose und einen geringeren Anteil an Personen mit hochgradiger Steatose aufweist. Dies wird in Tabelle 17 veranschaulicht.

Tabelle 17: Verteilung der histologischen Steatose in der Nicht-NAFLD-Kohorte und im Gesamtkollektiv

Kollektiv	n gesamt	S0 ¹		S1 ¹		S2 ¹		S3 ¹	
		n	%	n	%	n	%	n	%
Nicht-NAFLD	115	75	65,2	24	20,9	11	9,6	5	4,3
Gesamt	157	75	47,8	38	24,2	25	15,9	19	12,1

n: Personenanzahl; ¹histologischer Steatosegrad 0 - 3

Die weitere statistische Analyse wurde analog zur Auswertung des Gesamt- und der Logiq-Kohorte durchgeführt.

Die HRI-Werte der Nicht-NAFLD-Kohorte lagen im Mittel bei 1,38 mit einer Standardabweichung von 1,27. Im Vergleich zum Gesamtkollektiv sind somit beide Werte kleiner.

Die Werteverteilung der hepatorenenalen Indices auf die Steatosegrade wurde als Box-Plot visualisiert und kann in Abbildung 13 gesehen werden. Vergleicht man die grafische Darstellung der Steatosegrade der nicht an NAFLD erkrankten Patient*innen mit jener der Gesamtkohorte, fällt auf, dass die hepatorenenalen Indices der Nicht-NAFLD-Kohorte in einem ähnlichen Wertebereich liegen und nicht jene deutliche ansteigende Tendenz aufweisen, wie sie die Gesamtkohorte zeigt. Insbesondere bei S3 unterscheidet sich die Verteilung der HRI-Werte beider Kohorten in Bezug auf ihren Wertebereich und dessen Ausdehnung.

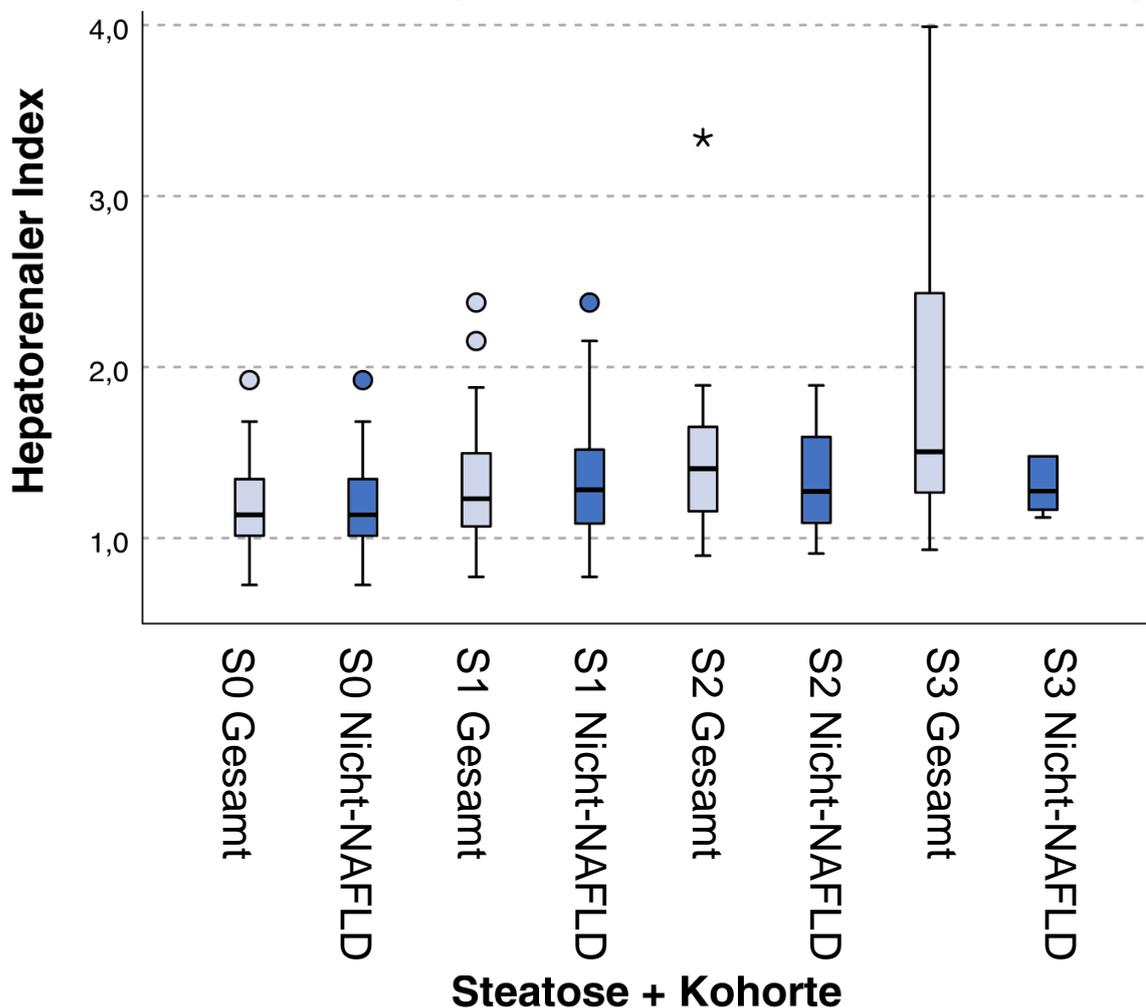


Abbildung 13: Darstellung der Verteilung von HRI-Werten der Nicht-NAFLD-Kohorte und des Gesamtkollektivs als Box-Plots aufgeteilt nach histologischem Steatosegrad. Hellblaue Boxen: Verteilung HRI-Werte Gesamtkollektiv; dunkelblaue Boxen: Verteilung HRI-Werte der Nicht-NAFLD-Kohorte; Punkte: milde Ausreißer.

Es erfolgte eine Analyse der Nicht-NAFLD-Kohorte auf signifikante Zusammenhänge zwischen den HRI-Werten und den Steatosegraden mittels Kruskal-Wallis Test. Die Untersuchung ergab einen p-Wert von 0,056 bei einem Signifikanzniveau von $p < 0,05$. Demnach liegt kein statistisch signifikanter Zusammenhang zwischen Steatosegrad und HRI in dieser Subpopulation vor, es kann jedoch eine Tendenz angenommen werden.

In der Darstellung als ROC-Kurve spiegelten sich die Erkenntnisse aus den Voruntersuchungen wider. Der ermittelte *Cut-off*-Wert zur Differenzierung zwischen Steatose und keiner Steatose ($S \geq 1$) von 1,46 ist zwar der gleiche, wie jener des Gesamtkollektivs und der Logiq-Kohorte, jedoch weist er mit 37,5 % eine deutlich schlechtere Sensitivität auf. Ein genauer Vergleich der Messgenauigkeit des HRI zur Detektion einer Hepatosteatose im Gesamtkollektiv, der Logiq- und der Nicht-NAFLD-Kohorte wird in Tabelle 18 gezeigt.

Tabelle 18: Diagnostische Genauigkeit des HRI zur Prädiktion einer Steatose im Gesamtkollektiv und der Logiq- und Nicht-NAFLD-Kohorte

Werte für $S \geq 1^1$	AUROC (95 % - KI)	Cut-off-Wert	Sens. (%)	Spez. (%)	PPW	NPW
Gesamt	0,680 (0,597 - 0,763)	1,46	42,7	90,7	0,834	0,591
Logiq	0,715 (0,624 - 0,807)	1,46	48,4	93,0	0,886	0,616
Nicht-NAFLD	0,652 (0,545 - 0,760)	1,46	37,5	90,7	0,683	0,731

¹Histologischer Steatosegrad; AUROC: Fläche unter Grenzwertoptimierungskurve; 95 % KI: 95 % Konfidenzintervall; PPW: positiv prädiktiver Wert; NPW: negativ prädiktiver Wert.

3.2.6 Einfluss von Leberfibrose auf die diagnostische Messgenauigkeit des HRI

Um den Einfluss von Leberfibrose auf die Messgenauigkeit des HRI beurteilen zu können, wurde das Gesamtkollektiv in zwei Kohorten unterteilt, die gesondert analysiert wurden. Die Unterteilung orientierte sich hierbei an dem histologischen Staging nach Desmet und Scheuer. Dementsprechend werden stellvertretend für die vier Fibrosegrade die Abkürzungen F1 – F4 und für Personen ohne Fibrose F0 verwendet.

Die erste Kohorte setzte sich aus Personen ohne oder mit leichter Fibrose zusammen ($F \leq 1$). Die Zweite bestand aus Personen mit mindestens mittelgradiger Leberfibrose ($F \geq 2$).

Aus den 157 in die Studie eingeschlossenen Teilnehmer*innen hatten 70 eine mindestens mittelgradige Leberfibrose ($F \geq 2$). Darunter hatten 29 eine zweitgradige (F2) und 23 eine drittgradige (F3) Leberfibrose. 18 Personen hatten eine Leberzirrhose (F4). Aus diesen 70 Teilnehmer*innen hatten 32 keine Steatose (S0), 18 eine leichte (S1), 12 eine moderate (S2) und 8 Personen eine schwere (S3) Hepatosteatose.

Keine (F0) oder eine geringgradige (F1) Fibrose hatten 49 bzw. 38 Personen. Demnach bestand die $F \leq 1$ -Kohorte aus 87 Teilnehmer*innen. Davon hatten 43 Personen keine Hepatosteatose (S0), 20 eine leichte (S1), 13 eine moderate (S2) und 11 Personen eine schwere (S3) Steatose. Die prozentuelle Verteilung der Steatosegrade auf die $F \geq 2$ - und $F \leq 1$ -Kohorte entspricht somit nahezu jener des Gesamtkollektivs und ist in Tabelle 19 zu sehen.

Tabelle 19: Verteilung histologischer Steatose bei Personen mit $F \geq 2$ und $F \leq 1$ im Vergleich zum Gesamtkollektiv

Kollektiv	n gesamt	S0 ¹		S1 ¹		S2 ¹		S3 ¹	
		n	%	n	%	n	%	n	%
$F^2 \geq 2$	70	32	45,7	18	25,7	12	17,1	8	11,4
$F^2 \leq 1$	87	43	49,4	20	23	13	14,9	11	12,6
Gesamt	157	75	47,8	38	24,2	25	15,9	19	12,1

¹histologischer Steatosegrad 0 – 3; ²Fibrose nach Desmet und Scheuer; n: Personenanzahl.

In der $F \geq 2$ -Kohorte lag der HRI im Mittel bei 1,56 mit einer Standardabweichung von $\pm 1,63$. Im Vergleich zum Gesamtkollektiv ($1,55 \pm 1,49$) sind Mittelwert und Standardabweichung somit größer. Analog dazu sind beide Werte der $F \leq 1$ -Kohorte mit 1,55 und $\pm 1,37$ kleiner.

Die Verteilung der HRI-Werte beider Fibrosekohorten auf die Steatosegrade ist in Abbildung 14 zu sehen. Aus Darstellungsgründen sind die beiden Kohorten mit $F \leq 1$ ($F \leq 1$) und $F \geq 2$

($F \geq 2$) abgekürzt. Es fällt auf, dass die Werteverteilung der $F \geq 2$ -Kohorte bei S0-S2 in einem sehr ähnlichen Bereich liegt und bei S3 breit gestreut ist. In der $F \leq 1$ -Kohorte sieht man, dass sich die Werte innerhalb einer Standardabweichung durchaus unterscheiden und einer steigenden Tendenz unterliegen. Betrachtet man jedoch die Mittelwerte, so fällt ein eindeutiger Unterschied nur zwischen S1 und S2 auf. Der Mittelwert von S2 liegt sogar höher als jener von S3.

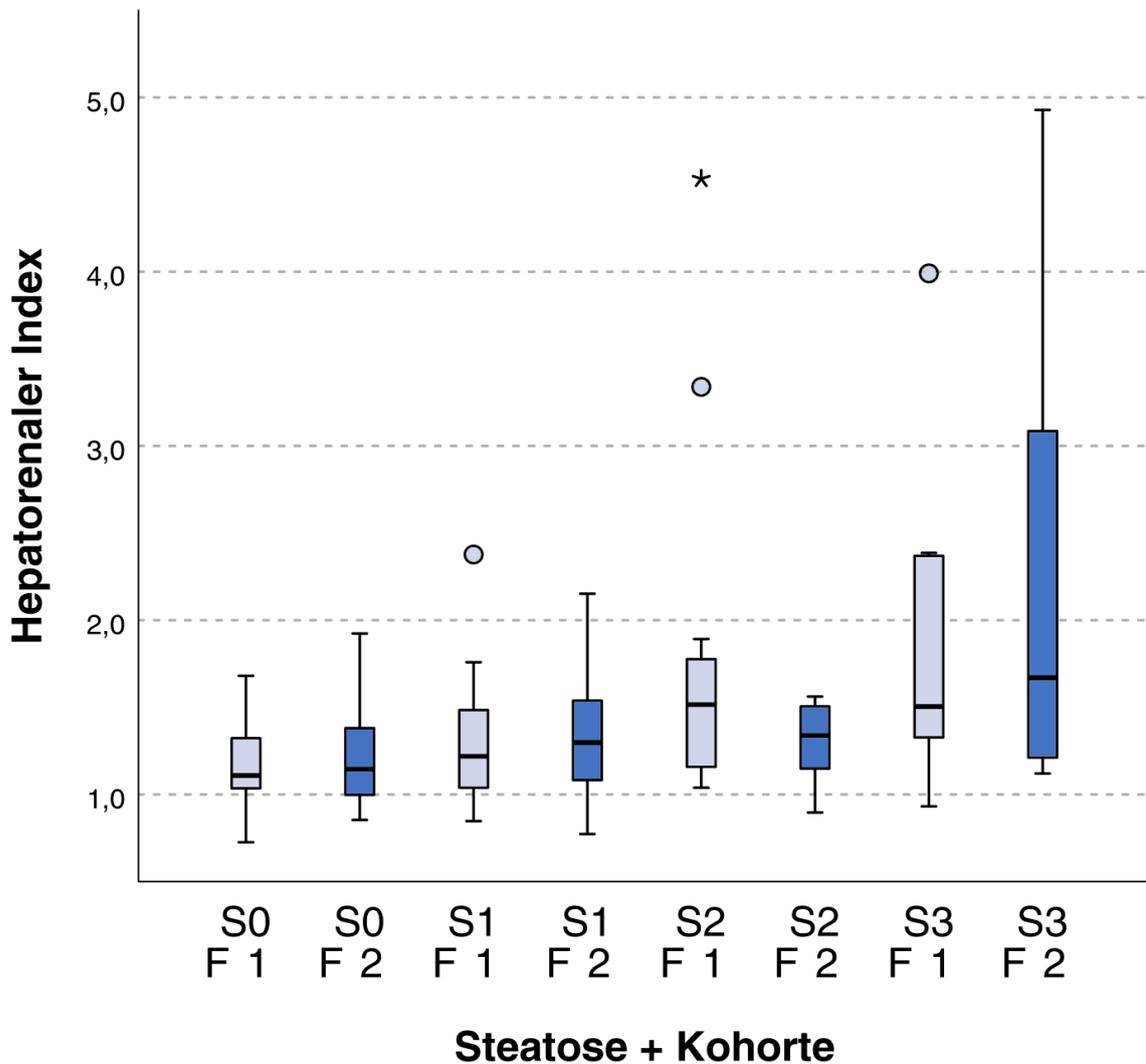


Abbildung 14: Darstellung der Verteilung von HRI-Werten der $F \leq 1$ - und der $F \geq 2$ -Kohorte als Box-Plots aufgeteilt nach histologischem Steatosegrad. Hellblaue Boxen: Verteilung HRI-Werte $F \leq 1$ -Kohorte; dunkelblaue Boxen: Verteilung HRI-Werte $F \geq 2$ -Kohorte; Punkte: milde Ausreißer.

Im Kruskal-Wallis-Test zeigte die $F \geq 2$ -Kohorte mit einem Wert von 0,079 keinen statistisch relevanten Zusammenhang zwischen HRI und Steatosegrad. In der $F \leq 1$ -Kohorte fiel der Wert mit 0,002 jedoch statistisch signifikant aus. Im darauffolgenden Post-hoc-Test konnte für die Paare S0 vs. S2 mit 0,034 und S0 vs. S3 mit 0,007 ein signifikanter Zusammenhang gezeigt werden.

Mit einer AUROC von 0,646 ist die Messgenauigkeit des HRI zur Detektion einer Hepatosteatose ($S \geq 1$) in der $F \geq 2$ -Kohorte schlechter als beim Gesamtkollektiv und allen anderen Kohorten. Der errechnete optimale *Cut-off*-Wert mit dem größten Youden-Index liegt bei einem HRI von 1,05 mit einer Sensitivität von 0,9 und Spezifität von 0,38. Der Wert unterscheidet sich hier deutlich von den anderen Gruppen, die alle bei einem HRI von 1,46 lagen.

Die $F \leq 1$ -Kohorte zeigte mit einer AUROC von 0,703 einen höheren Wert. Bei der Detektion einer Steatose ($S \geq 1$) erreicht der HRI-Grenzwert der $F \leq 1$ -Kohorte eine Sensitivität von 47,7 % und Spezifität von 97,7 %. Die Ergebnisse zu beiden Kollektiven sind in Tabelle 20 dargestellt.

Tabelle 20: Diagnostische Genauigkeit des HRI zur Prädiktion einer Steatose im $F \leq 1$ -, $F \geq 2$ - und Gesamtkollektiv

Werte für $S \geq 1$	AUROC (95 % - KI)	Cut-off-Wert	Sens. (%)	Spez. (%)	PPW	NPW
$F \geq 2$	0,646 (0,516 – 0,775)	1,05	89,5	37,5	0,630	0,750
$F \leq 1$	0,703 (0,592 – 0,815)	1,45	47,7	97,7	0,955	0,646
Gesamt	0,680 (0,597 - 0,763)	1,46	42,7	90,7	0,834	0,591

¹Histologischer Steatosegrad; ²Fibrosegrad nach Desmet und Scheuer; AUROC: Fläche unter der Grenzwertoptimierungskurve; 95 % KI: 95 % Konfidenzintervall; PPW: positiv prädiktiver Wert; NPW: negativ prädiktiver Wert.

3.2.7 Korrelation anderer Parameter mit dem histologisch bestimmten Steatosegrad

Der histologisch bestimmte Steatosegrad wurde mit weiteren Parametern auf signifikante Zusammenhänge untersucht. Hierzu zählen die Variablen HRI, Diabetes mellitus, BMI, Alter, Geschlecht, ALT, γ -GT und Bilirubin.

In der univariante Analyse der Parameter zeigte sich eine signifikant positive Assoziation zwischen HRI (Spearman's $\rho = 0,364$; $p < 0,001$), Diabetes mellitus (Spearman's $\rho = 0,264$; $p = 0,001$) und BMI (Spearman's $\rho = 0,397$; $p < 0,001$) mit dem histologisch bestimmten Steatosegrad.

Für die Variablen Alter ($p = 0,946$), Geschlecht ($p = 0,211$), ALT ($p = 0,399$), γ -GT ($p = 0,382$), Fibrose ($p = 0,828$) und Bilirubin ($p = 0,072$) konnte kein signifikanter Zusammenhang zum Steatosegrad nachgewiesen werden.

Alle Variablen, deren p -Werte in der univariaten Analyse kleiner als 0,1 waren, wurden zusätzlich mit einem multiplen Regressionsanalysemodell auf unabhängige Korrelationen untersucht. Ein signifikanter unabhängiger Zusammenhang konnte zwischen HRI ($p < 0,001$), BMI ($p = 0,004$) und dem Steatosegrad festgestellt werden.

3.2.8 Bestimmung der Interraterreliabilität

Bei 30 der 157 Personen wurde der HRI von zwei Untersuchern unabhängig voneinander bestimmt. Die Werte wurden mit Hilfe des Intraklassen-Korrelationskoeffizienten auf Übereinstimmung untersucht. Die Übereinstimmung war mit einem Intraklassenkorrelationskoeffizienten von 0,934 exzellent. Somit besteht eine hohe Interraterreliabilität.

4 Diskussion

Diese Arbeit hat sich damit befasst, ob der sonographisch bestimmte HRI geeignet ist, um eine Hepatosteatose zu erkennen und zu quantifizieren. Hierfür wurden die Daten von Personen mit einer CLD analysiert, die sich sowohl einer Leberbiopsie mit histologischer Beurteilung als auch einer sonographischen Untersuchung der Leber unterzogen hatten.

Ein zuverlässiges, nicht-invasives Diagnostiktool zur Früherkennung und Einstufung einer Hepatosteatose wäre von großer Bedeutung. Zum einen, um Patient*innen mit einer Fettlebererkrankung frühzeitig zu erkennen und anhand des Steatosegrades eine Risikostratifizierung des Fibroseerisikos vorzunehmen. Zum anderen bestehen Hinweise, dass eine zusätzlich vorliegende Hepatosteatose bei Patient*innen mit einer CLD mit einer erhöhten Mortalität einhergeht. Dies trifft vor allem auf Personen mit einer chronischen Hepatitis B Infektion oder einer Autoimmunhepatitis zu (De Luca-Johnson et al. 2016; Chan et al. 2017; Petzold et al. 2020). Aufgrund dieser Erkenntnisse wäre die Früherkennung einer zusätzlichen Hepatosteatose von großer Bedeutung für die Überwachung und Behandlung von Patient*innen mit einer CLD.

Die primäre Fragestellung war, ob ein signifikanter Zusammenhang zwischen dem histologisch ermittelten Steatosegrad und dem sonographisch bestimmten HRI festgestellt werden kann. Die Ergebnisse der Datenauswertung zeigten, dass eine unabhängige signifikante positive Korrelation gegeben ist ($p < 0,001$).

Des Weiteren befasste sich die Studie damit, ob der HRI zur Quantifizierung einer Hepatosteatose geeignet ist. Hierfür wurden unterschiedliche *Cut-off*-Werte ermittelt und deren Sensitivität, Spezifität, PPW und NPW bestimmt. Die Studie zeigte, dass eine zuverlässige Differenzierung zwischen geringgradiger und keiner Steatose nicht möglich ist. Die diagnostische Genauigkeit für die Abgrenzung einer normalen Leber von einer mittel- bis hochgradigen Steatose war sehr hoch.

Die statistische Analyse wurde ebenso für mehrere Subpopulationen durchgeführt: Es wurde eine eigene Auswertung für jene Gruppe vorgenommen, deren Leberultraschall mit dem Ultraschallgerät Logiq E9 (GE Medical Systems, Wauwatosa, USA; software R1.0.6) durchgeführt wurden. Das zweite Gerät, Hitachi ALOKA (ProSound Alpha 7), welches im Studienkollektiv vertreten war, hatte eine zu geringe Fallzahl und wurde aus diesem Grund nicht gesondert analysiert. Die Auswertung der Logiq-Gruppe im Vergleich zum Gesamtkollektiv zeigte einerseits eine höhere diagnostische Genauigkeit des HRI, andererseits aber auch eine breitere Streuung der HRI -Werte bei hochgradiger Hepatosteatose.

Eine weitere Subgruppe waren chronisch Leberkranke, die eine andere Erkrankung als NAFLD hatten. Deren Daten wurden ebenso gesondert ausgewertet. Die Verteilung der HRI -Werte auf die Steatosegrade war für alle drei Grade in einem ähnlichen Wertebereich und zeigte nicht den Trend, der im Gesamtkollektiv und bei der Logiq-Gruppe zu sehen war.

Dieses Ergebnis spiegelte sich auch in der diagnostischen Genauigkeit des HRI zur Bestimmung der Hepatosteatosegrade wider. Es wurden deutlich schlechtere Ergebnisse vor allem bei der Sensitivität erzielt. Somit eignet sich der HRI nicht gut um bei Patient*innen mit einer zusätzlichen Lebererkrankung eine Steatose zu detektieren oder zu quantifizieren.

Bemerkenswert ist, dass die Analysen des Gesamtkollektivs und beide Subpopulationen ähnliche und teilweise nahezu gleiche *Cut-off*-Werte zur Bestimmung einer Steatose und ihrer Grade lieferten.

Des Weiteren wurde der Zusammenhang zwischen dem HRI und Hepatosteatose bei Personen mit Fibrose untersucht. Hierfür wurde zwischen einer Kohorte mit einer mittelgradigen und höheren Fibrose ($F \geq 2$) und einer Kohorte mit keiner und einer geringen Fibrose ($F < 2$) unterschieden. Für die Gruppe mit mittelgradiger und höherer Fibrose konnte kein signifikanter Zusammenhang zwischen HRI und Hepatosteatose hergestellt werden. Die Kohorte aus Personen ohne oder mit geringer Fibrose zeigte ähnlich wie im Gesamtkollektiv und der Logiq-Kohorte einen signifikanten Zusammenhang zwischen HRI und Hepatosteatose. Der Zusammenhang spiegelte sich auch in der Messgenauigkeit zur Detektion einer Hepatosteatose wider und übertraf in Teilen sogar jene des Gesamtkollektivs und der Logiq-Kohorte. Demnach kann vermutet werden, dass die Anwesenheit einer signifikanten Fibrose sich negativ auf die diagnostische Genauigkeit des hepatoreanalen Index auswirkt.

Für den klinischen Alltag könnte die Spezifität des HRI-*Cut-off*-Wertes der $F < 2$ -Kohorte mit 97,7 % relevant für die Detektion einer Steatose ($S \geq 1$) sein. Es ergab sich daraus ein PPW von 0,955 und ein negativ prädiktiver Wert von 0,646. Demnach könnte man z. B. im Rahmen der Überwachung von Personen mit einer vorbekannten CLD ohne Fibrose bei Überschreiten des *Cut-off*-Wertes von 1,45 direkt Maßnahmen gegen eine Hepatosteatose setzen, ohne zuvor die Steatose histologisch sicher zu müssen.

Um den HRI auf seine Eignung als untersucherunabhängiges Diagnostiktool zu überprüfen, wurde die Interraterreliabilität anhand des ICC ermittelt. Das Ergebnis ergab eine exzellente Übereinstimmung zwischen den Untersuchungsergebnissen verschiedener Untersucher. Damit entspricht das Ergebnis jenem der vorhergegangenen Studien (Webb et al. 2009; Marshall et al. 2012).

Beim Vergleich dieser Arbeit mit bisher erschienen Publikationen fällt auf, dass die Ergebnisse in Bezug auf die Messgenauigkeit des HRI stark variieren. Dies gilt sowohl für die *Cut-off*-Werte als auch für deren Sensitivität und Spezifität. So erstreckt sich der optimale *Cut-off*-Wert zur Detektion einer leichten Hepatosteatose von 1,24 bis 2,01. Die Sensitivität schwankte zwischen 100 % und 62,5 % und die Spezifität zwischen 95,2 % und 54 %. Bemerkenswert ist hierbei, dass alle Studien die Leberbiopsie als Referenzstandard nutzten (Webb et al. 2009; Marshall et al. 2012; De Almeida e Borges et al. 2013; Wang et al. 2013; Chauhan et al. 2016). Ein direkter Vergleich der Studien ist in Tabelle 21 zu sehen.

Tabelle 21: Vergleich von Studien, die die Messgenauigkeit des HRI überprüften

Autoren und Veröffentlichungsjahr	n	Cut-off-Wert	Sensitivität	Spezifität
Webb et al. (2009)	111	1,49	100	91
Marshall et al. (2012)	101	1,28	100	54
De Almeida e Borges et al. (2013)	82	1,24	93	93
Wang et al. (2013)	175	4 ¹	82	63
Chauhan et al. (2016)	45	2,01	62,5	95,2
Aktuelle Studie	157	1,46	42,7	90,7

¹Autoren nutzten hepatorenen Unterschied statt heparorenen Index; n: Personenanzahl. (basierend auf Chauhan et al. 2016). Die Verwendung erfolgt mit freundlicher Genehmigung von John Wiley and Sons.

Die unterschiedlichen Ergebnisse der Studien in Bezug auf die Messgenauigkeit des HRI zur Detektion einer Hepatosteatose lassen vermuten, dass der HRI sehr stark vom verwendeten Ultraschallgerät und von der Ursache der Lebererkrankung abhängig ist. Diese Erklärung lässt auch diese Studie in Bezug auf die Ätiologie der Lebererkrankung zu. Um den Zusammenhang zwischen der Messgenauigkeit des HRI mit dem verwendeten Ultraschallgerät nachweisen zu können, hat die Kohortengröße der Aloka-Gruppe nicht ausgereicht. Es muss trotzdem hinterfragt werden, was den unterschiedlichen Messergebnissen des Gesamtkollektivs und jenen der Logiq-Gruppe zu Grunde liegt.

Es stellt sich die Frage, warum die *Cut-off*-Werte und die Messgenauigkeit sowohl in den vergangenen Studien als auch in der aktuellen Studie so stark variieren.

Einerseits könnte die Verwendung unterschiedlicher Ultraschallgeräte zu unterschiedlichen Studienergebnissen führen. Diese Vermutung konnte in der aktuellen Studie aufgrund der fehlenden Kohortenanalyse der Ultraschallgeräte weder widerlegt noch gestützt werden. Man konnte jedoch anhand des Vergleiches der Studienergebnisse und der Veröffentlichungsdaten schließen, dass sich die Verwendung von modernen *High-End*-Geräten nicht relevant positiv auf die Messgenauigkeit des HRI ausgewirkt hat.

Andererseits ist zu vermuten, dass unterschiedliche zugrundeliegende Lebererkrankungen einen Einfluss auf den HRI und somit auf die Genauigkeit der Detektion einer Hepatosteatose haben. Bemerkenswert ist zum Beispiel, dass die Studie von Wang et al. (2013)

nur Personen mit Hepatitis B und -C Infektion in ihrem Studienkollektiv hatte. Die Arbeit von De Almeida e Borges et al. (2013) inkludierte nur Patient*innen mit Verdacht auf eine Hepatosteatose und eine Kontrollgruppe, bei der eine Hepatosteatose als unwahrscheinlich galt. Das Studienkollektiv der aktuellen Studie setzt sich aus Patient*innen mit 12 unterschiedlichen Lebererkrankungen zusammen und kann deshalb als heterogen bezeichnet werden. Die gesonderte Analyse jener Personen, die eine andere Lebererkrankung als NAFLD hatten, zeigte deutlich schlechtere Ergebnisse als die Gesamtgruppe, weshalb ein Zusammenhang zwischen der Grunderkrankung und der diagnostischen Genauigkeit des HRI vermutet werden kann. Die Heterogenität des Studienkollektivs könnten auch zum Teil für die schlechten Ergebnisse des HRI betreffend Sensitivität und Spezifität verantwortlich sein.

Ebenso könnte die dorsale Schallabschwächung, ein Phänomen das man sich bei der visuellen Quantifizierung einer Hepatosteatose zu Nutzen macht, eine Ursache für die schlechte diagnostische Genauigkeit sein. Bei der dorsalen Schallabschwächung erscheinen distal zur Schallsonde gelegene Leberparenchymabschnitte dunkel oder sind aufgrund der vollständigen Schallauslöschung komplett schwarz. Das Phänomen steht demnach gegensätzlich zu der Annahme des hepatorenen Index, dass sich eine verfettete Leber im Ultraschall heller darstellt als eine Gesunde. Dies führt vor allem bei starker Steatose zu dem Problem, dass bereits auf Höhe der Niere eine dorsale Schallabschwächung auftritt und bei der HRI-Bestimmung über der Leber falsch niedrige Werte ermittelt werden. In diesem Zusammenhang wäre für weitere Arbeiten interessant, ob mithilfe der Schallabschwächung zwischen mehreren in festen Abständen, auf einer Geraden liegenden Messpunkten eine Hepatosteatose besser festgestellt und quantifiziert werden könnte.

In Bezug auf die beiden Einflussfaktoren Ultraschallgerät und zugrundeliegende Lebererkrankung wird deutlich, dass weitere prospektive Studien mit unterschiedlichen Ultraschallmodellen und einheitlichem Patient*innenkollektiv nötig sind.

Anhand des gleichen Patient*innenkollektivs wurde in einer Arbeit von Petzold et al. (2020) zusätzlich zur diagnostische Genauigkeit des HRI die Messgenauigkeit der visuellen Einschätzung des Hepatosteatosegrades anhand von B-Bildkriterien überprüft. Die Ergebnisse werden in Tabelle 22 gegenübergestellt. Bemerkenswert ist, dass die visuelle Einschätzung des Steatosegrades für alle drei Grade eine höhere diagnostische Genauigkeit aufweist als der hepatorenale Index. Insofern kann daraus geschlossen werden, dass der hepatorenale Index keinen Zusatznutzen bei der Diagnose und Quantifizierung einer Hepatosteatose im Vergleich zur konventionellen visuellen Bestimmung liefert.

Tabelle 22: Vergleich der diagnostischen Genauigkeit von visueller Einschätzung und HRI

		AUROC (95 % - KI)	Sens. (%)	Spez. (%)	PPW	NPW
S \geq 1 ¹	HRI ²	0,680 (0,597 - 0,763)	42,7	90,7	0,834	0,591
	Visuell ³	0,798 (0,728 – 0,868)	75,6	76,0	0,775	0,740
S \geq 2 ¹	HRI ²	0,712 (0,619 - 0,805)	52,3	85,0	0,576	0,821
	Visuell ³	0,861 (0,793 – 0,929)	90,9	64,6	0,500	0,948
S = 3	HRI ²	0,738 (0,612 - 0,864)	47,4	93,5	0,501	0,928
	Visuell ³	0,923 (0,875 – 0,971)	89,5	89,9	0,550	0,984

¹Steatosegrad; ²Einschätzung anhand des hepatorenenalen Index; ³Einschätzung anhand visueller Kriterien; AUROC: Fläche unter der Grenzwertoptimierungskurve; 95 % KI: 95 % Konfidenzintervall; PPW: positiv prädiktiver Wert; NPW: negativ prädiktiver Wert (basierend auf Petzold et al. 2020).

Diese Studie weist einige Limitationen auf. Zuallererst sei erwähnt, dass diese Arbeit aufgrund ihres retrospektiven Studiendesigns jene Schwachpunkte aufweist, die für retrospektive Studien typisch sind. Für diese Studie könnten das mögliche Übersehen von Confoundern und eine Stichprobenverzerrung durch die Selektierung des Patient*innenkollektivs als Fehlerquellen von Relevanz sein.

Ebenso wurden, wie oben angesprochen, unterschiedliche Ultraschallgeräte genutzt, die von unterschiedlichen Untersuchern bedient wurden.

Des Weiteren wurde jeweils nur eine Leberbiopsie pro Patient*in durchgeführt. In diesem Kontext kann jedoch angemerkt werden, dass eine Studie, die die Übereinstimmung der Steatosegrade zwischen paarweise entnommenen Proben von NAFLD-Patient*innen eine hohe Übereinstimmung feststellte (Ratziu et al. 2005).

Zuletzt muss auch angemerkt werden, dass die Größe von Subgruppen teilweise nicht ausreichend war, um weitere Analysen von Subpopulationen zu ermöglichen.

Eine potenzielle Schwäche dieser Arbeit, das heterogene Studienkollektiv, kann auch als Stärke gewertet werden. Die Studie umfasst ein breites Spektrum an Lebererkrankungen in unterschiedlichen Stadien. Dies entspricht auch der Verteilung, der man im klinischen Alltag begegnet. Somit werden die Realbedingungen wiedergegeben. Zu den weiteren Stärken der Studie zählt die Größe des gesamten Studienkollektivs, die die meisten der vorangegangenen Studien zu diesem Thema übertrifft.

Eine weitere Stärke ist, dass die histologische Bestimmung des Steatosegrades ausschließlich mittels Biopsie als Referenz ermittelt wurde und somit dem diagnostischen Goldstandard entspricht.

Zusammenfassend kann gesagt werden, dass es zwar einen Zusammenhang, zwischen dem sonographisch bestimmten HRI und dem histologisch bestimmten Steatosegrad gibt. Trotzdem ist der HRI nicht für die breite Bevölkerung zur Früherkennung oder Quantifizierung einer Steatose geeignet. Unabhängig davon ist die Interraterreliabilität des HRI exzellent. Die Verwendung von *High-End*-Ultraschallgeräten hat sich anscheinend nicht relevant positiv auf die Messgenauigkeit des HRI ausgewirkt.

Um den Einfluss von unterschiedlichen Ultraschallgeräten und Ätiologien der CLD auf die Messgenauigkeit des HRI beurteilen zu können benötigt es weitere groß angelegte Studien, die eine Analyse dieser Subpopulationen ermöglichen.

5 Zusammenfassung

Hepatosteatose ist eine Erkrankung der Leber, die immer mehr Bedeutung im klinischen Alltag erlangt. Bereits jetzt wird eine weltweite Prävalenz der Hepatosteatose von 25 % angenommen. Da ein direkter Zusammenhang zwischen Adipositas und NAFLD besteht und die globale Prävalenz von Übergewichtigkeit und Adipositas nach wie vor steigt, wird auch davon ausgegangen, dass die Prävalenz der Hepatosteatose weiterhin steigt. Die Folgen der Hepatosteatose sind weitreichend. Bis zu 30 % der Hepatosteatose-Erkrankten entwickeln eine Steatohepatitis, die wiederum zu einer Leberfibrose, -zirrhose bis hin zum HCC fortschreiten kann. Als Konsequenz der weiten Verbreitung von Hepatosteatose und deren Folgen kann die Tatsache, dass Leberverfettung mittlerweile die zweithäufigste Indikation für eine Lebertransplantation ist, gewertet werden.

Aufgrund dieser Bedeutung im klinischen Alltag spielt die Hepatosteatosediagnostik eine wichtige Rolle. Wegen ihrer Invasivität wird der derzeitige Goldstandard, die Stanzbiopsie, nur nach einer strengen Indikationsstellung zur Diagnostik einer Leberverfettung angewandt. Deshalb wird ein nicht-invasives Diagnostetool für das Hepatosteatosescreening dringend benötigt. Die bisherigen Diagnoseverfahren haben alle den Nachteil, dass sie entweder sehr kostenintensiv sind oder eine stark schwankende Messgenauigkeit aufweisen. Unter diesen befindet sich auch der sonographisch ermittelte HRI, dessen Messgenauigkeit in mehreren Studien untersucht wurde. Auffallend ist jedoch, dass die meisten Studien vor mehr als zehn Jahre entstanden sind. Somit wurden die Untersuchungen mit einer anderen Generation von Ultraschallgeräten als die heute verfügbaren hochauflösenden *High-End*-Versionen durchgeführt.

Aus diesem Grund untersucht diese Arbeit, ob der mit *High-End*-Ultraschallgeräten bestimmte HRI zur Detektion und Quantifizierung einer Hepatosteatose bei chronisch leberkranken Personen geeignet ist.

Für die Studie wurden im Zeitraum zwischen Januar 2016 und Mai 2019 bei 157 Personen, die sowohl eine Leberbiopsie als auch eine Ultraschalluntersuchung der Leber bekommen haben, der HRI durch einen Untersucher bestimmt und mit seinem histologisch bestimmten Steatosegrad verglichen. Hierbei wurde nach einem Zusammenhang zwischen dem HRI und dem histologisch bestimmten Steatosegrad gesucht und auf Signifikanz überprüft. Ebenso wurden *Cut-off*-Werte zur Ermittlung einer Hepatosteatose bzw. zur Quantifizierung dieser bestimmt und die Messgenauigkeit ermittelt. Des Weiteren wurde die Interraterreliabilität und der Zusammenhang mit anderen Einflussfaktoren wie Body-Mass-Index, Alter, Geschlecht und Laborparameter für den HRI überprüft.

In der univariaten Untersuchung wurde eine hochsignifikante positive Korrelation zwischen HRI und dem histologisch bestimmten Steatosegrad festgestellt ($p < 0.001$). Eine Hepatosteatoe konnte mit dem ermittelten *Cut-off*-Wert von 1,46 und einer Sensitivität von 42,7 % und Spezifität von 90,7 % bestimmt werden. In der multivariaten Analyse wurde eine Korrelation zwischen HRI ($p < 0.001$) und Body-Mass-Index ($p = 0.004$) mit dem histologisch bestimmten Steatosegrad festgestellt. Die Interraterreliabilität wurde anhand des Interklassenkoeffizienten bestimmt und mit einem Wert von 0,934 als exzellent befunden.

Die Studie zeigt, dass es einen signifikant positiven Zusammenhang zwischen dem mit *High-End*-Ultraschallgeräten bestimmten HRI und dem histologisch bestimmten Steatosegrad gibt. Die Messgenauigkeit des HRI ist jedoch trotz positivem Zusammenhang nicht ausreichend, um sich gegen andere nicht-invasive Verfahren durchzusetzen und eine sichere Bestimmung und Quantifizierung einer Hepatosteatoe zu ermöglichen. Bei Personen mit einer mäßigen oder stärkeren Fibrose, anderen Erkrankungen als NAFLD oder niedrigen BMI-Werten nimmt die Messgenauigkeit zusätzlich ab.

6 Literaturverzeichnis

- Adams LA, Lymp JF, St. Sauver J, Sanderson SO, Lindor KD, Feldstein A, Angulo P (2005): The natural history of nonalcoholic fatty liver disease: a population-based cohort study. *Gastroenterology* 129, 113–121
- Ajmera V, Park CC, Caussy C, Singh S, Hernandez C, Bettencourt R, Hooker J, Sy E, Behling C, Xu R, et al. (2018): Magnetic Resonance Imaging Proton Density Fat Fraction Associates With Progression of Fibrosis in Patients With Nonalcoholic Fatty Liver Disease. *Gastroenterology* 155, 307–310
- Anstee QM, Targher G, Day CP (2013): Progression of NAFLD to diabetes mellitus, cardiovascular disease or cirrhosis. *Nat Rev Gastroenterol Hepatol* 10, 330–344
- Araújo AR, Rosso N, Bedogni G, Tiribelli C, Bellentani S (2018): Global epidemiology of non-alcoholic fatty liver disease/non-alcoholic steatohepatitis: What we need in the future. *Liver Int* 38, 47–51
- Bedogni G, Bellentani S, Miglioli L, Masutti F, Passalacqua M, Castiglione A, Tiribelli C (2006): The Fatty Liver Index: a simple and accurate predictor of hepatic steatosis in the general population. *BMC Gastroenterol* 6, 33
- Bedossa P (2016): Histological Assessment of NAFLD. *Dig Dis Sci* 61, 1348–1355
- Bedossa P (2017): Pathology of non-alcoholic fatty liver disease. *Liver Int* 37, 85–89
- Bellentani S, Saccoccio G, Masutti F, Crocè LS, Brandi G, Sasso F, Cristanini G, Tiribelli C (2000): Prevalence of and Risk Factors for Hepatic Steatosis in Northern Italy. *Ann Intern Med* 132, 112
- Bohte AE, van Werven JR, Bipat S, Stoker J (2011): The diagnostic accuracy of US, CT, MRI and 1H-MRS for the evaluation of hepatic steatosis compared with liver biopsy: a meta-analysis. *Eur Radiol* 21, 87–97
- Browning JD, Szczepaniak LS, Dobbins R, Nuremberg P, Horton JD, Cohen JC, Grundy SM, Hobbs HH (2004): Prevalence of hepatic steatosis in an urban population in the United States: Impact of ethnicity. *Hepatology* 40, 1387–1395
- Calzadilla Bertot L, Adams L Anton (2016): The Natural Course of Non-Alcoholic Fatty Liver Disease. *Int J Mol Sci* 17, 774
- Caruso MG, Veronese N, Notarnicola M, Cisternino AM, Reddavid R, Inguaggiato R, Guerra V, Donghia R, Loggoscino A, Rotolo O, et al. (2019): Fatty liver and mortality: a cohort population study in South Italy. *BMJ Open* 9, e027379
- Chalasani N, Younossi ZM, Lavine JE, Diehl AM, Brunt EM, Cusi K, Charlton M, Sanyal AJ (2012): The diagnosis and management of non-alcoholic fatty liver disease: Practice Guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. *Hepatology* 55, 2005–2023
- Chan AWH, Wong GLH, Chan HY, Tong JHM, Yu YH, Choi PCL, Chan HLY, To KF, Wong VWS (2017): Concurrent fatty liver increases risk of hepatocellular carcinoma among patients

- with chronic hepatitis B: Steatosis and cancer in hepatitis B. *J Gastroenterol Hepatol* 32, 667–676
- Chang Y, Cho J, Cho YK, Cho A, Hong YS, Zhao D, Ahn J, Sohn CI, Shin H, Guallar E (2019): Alcoholic and Nonalcoholic Fatty Liver Disease and Incident Hospitalization for Liver and Cardiovascular Diseases. *Clin Gastroenterol Hepatol* 205–215
- Chauhan A, Sultan LR, Furth EE, Jones LP, Khungar V, Sehgal CM (2016): Diagnostic accuracy of hepatorenal index in the detection and grading of hepatic steatosis. *J Clin Ultrasound* 44, 580–586
- Chooi YC, Ding C, Magkos F (2019): The epidemiology of obesity. *Metabolism* 92, 6–10
- Choudhary NS, Saraf N, Saigal S, Duseja A, Gautam D, Lipi L, Rastogi A, Goja S, Bhangui P, Ramchandra SK, et al. (2017): Prediction of nonalcoholic fatty liver in prospective liver donors. *Clin Transplant* 31, e12890
- De Almeida e Borges VF, Diniz ALD, Cotrim HP, Rocha HLOG, Andrade NB (2013): Sonographic hepatorenal ratio: A noninvasive method to diagnose nonalcoholic steatosis. *J Clin Ultrasound* 41, 18–25
- De Lédighen V, Vergniol J, Capdepon M, Chermak F, Hiriart J-B, Cassinotto C, Merrouche W, Foucher J, Brigitte LB (2014): Controlled attenuation parameter (CAP) for the diagnosis of steatosis: A prospective study of 5323 examinations. *J Hepatol* 60, 1026–1031
- De Luca-Johnson J, Wangenstein KJ, Hanson J, Krawitt E, Wilcox R (2016): Natural History of Patients Presenting with Autoimmune Hepatitis and Coincident Nonalcoholic Fatty Liver Disease. *Dig Dis Sci* 61, 2710–2720
- Duan XY, Zhang L, Fan JG, Qiao L (2014): NAFLD leads to liver cancer: Do we have sufficient evidence? *Cancer Lett* 345, 230–234
- Estes C, Razavi H, Loomba R, Younossi Z, Sanyal AJ (2018): Modeling the epidemic of nonalcoholic fatty liver disease demonstrates an exponential increase in burden of disease. *Hepatology* 67, 123–133
- Friedman SL (2003): Liver fibrosis – from bench to bedside. *J Hepatol* 38, 38–53
- Gordon A, McLean CA, Pedersen JS, Bailey MJ, Roberts SK (2005): Hepatic steatosis in chronic hepatitis B and C: predictors, distribution and effect on fibrosis. *J Hepatol* 43, 38–44
- Grąt K, Grąt M, Rowiński O (2020): Usefulness of Different Imaging Modalities in Evaluation of Patients with Non-Alcoholic Fatty Liver Disease. *Biomedicines* 8, 298
- Gu J, Liu S, Du S, Zhang Q, Xiao J, Dong Q, Xin Y (2019): Diagnostic value of MRI-PDFF for hepatic steatosis in patients with non-alcoholic fatty liver disease: a meta-analysis. *Eur Radiol* 29, 3564–3573
- Hashimoto E, Tokushige K, Ludwig J (2015): Diagnosis and classification of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis: Current concepts and remaining challenges. *Hepatol Res* 45, 20–28

- Hernaes R, Lazo M, Bonekamp S, Kamel I, Brancati FL, Guallar E, Clark JM (2011): Diagnostic accuracy and reliability of ultrasonography for the detection of fatty liver: A meta-analysis. *Hepatology* 54, 1082–1090
- Iwasaki M, Takada Y, Hayashi M, Minamiguchi S, Haga H, Maetani Y, Fujii K, Kiuchi T, Tanaka K (2004): Noninvasive Evaluation of Graft Steatosis in Living Donor Liver Transplantation: Transplantation 78, 1501–1505
- Joy D, Thava VR, Scott BB (2003): Diagnosis of fatty liver disease: is biopsy necessary? *Eur J Gastroenterol Hepatol* 15, 539–543
- Karlas T, Petroff D, Sasso M, Fan JG, Mi YQ, De Lédinghen V, Kumar M, Lupsor-Platon M, Han KH, Cardoso AC, et al. (2017): Individual patient data meta-analysis of controlled attenuation parameter (CAP) technology for assessing steatosis. *J Hepatol* 66, 1022–1030
- Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, Ferrell LD, Liu YC, Torbenson MS, Unalp-Arida A, et al. (2005): Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* 41, 1313–1321
- Kneeman JM, Misdraji J, Corey KE (2012): Secondary causes of nonalcoholic fatty liver disease. *Ther Adv Gastroenterol* 5, 199–207
- Lăpădat AM, Jianu IR, Ungureanu BS, Florescu LM, Gheonea DI, Sovaila S, Gheonea IA (2017): Non-invasive imaging techniques in assessing non-alcoholic fatty liver disease: a current status of available methods. *J Med Life* 10, 19–26
- Lee JY, Kim KM, Lee SG, Yu E, Lim YS, Lee HC, Chung YH, Lee YS, Suh DJ (2007): Prevalence and risk factors of non-alcoholic fatty liver disease in potential living liver donors in Korea: A review of 589 consecutive liver biopsies in a single center. *J Hepatol* 47, 239–244
- Lee SS, Park SH, Kim HJ, Kim SY, Kim MY, Kim DY, Suh DJ, Kim KM, Bae MH, Lee JY, et al. (2010): Non-invasive assessment of hepatic steatosis: Prospective comparison of the accuracy of imaging examinations. *J Hepatol* 52, 579–585
- Marshall RH, Eissa M, Bluth EI, Gulotta PM, Davis NK (2012): Hepatorenal index as an accurate, simple, and effective tool in screening for steatosis. *Am J Roentgenol* 199, 997–1002
- Middleton MS, Heba ER, Hooker CA, Bashir MR, Fowler KJ, Sandrasegaran K, Brunt EM, Kleiner DE, Doo E, Van Natta ML, et al. (2017): Agreement Between Magnetic Resonance Imaging Proton Density Fat Fraction Measurements and Pathologist-Assigned Steatosis Grades of Liver Biopsies From Adults With Nonalcoholic Steatohepatitis. *Gastroenterology* 153, 753–761
- Munteanu M, Ratziu V, Morra R, Messous D, Imbert-Bismut F, Poynard T (2008): Noninvasive biomarkers for the screening of fibrosis, steatosis and steatohepatitis in patients with metabolic risk factors: FibroTest-FibroMax experience. *J Gastrointest Liver Dis* 17, 187–19
- Naik A, Košir R, Rozman D (2013): Genomic aspects of NAFLD pathogenesis. *Genomics* 102, 84–95
- Nalbantoglu ILK, Brunt EM (2014): Role of liver biopsy in nonalcoholic fatty liver disease. *World J Gastroenterol* 20, 9026–9037

- Nomura H, Kashiwagi S, Hayashi J, Kajiyama W, Tani S, Goto M (1988): Prevalence of fatty liver in a general population of Okinawa, Japan. *Jpn J Med* 27, 142–149
- Pan JJ (2014): Gender and racial differences in nonalcoholic fatty liver disease. *World J Hepatol* 6, 274–283
- Paul A (2002): Nonalcoholic fatty liver disease. *N Engl J Med* 346, 1221–1231
- Petzold G, Lasser J, Rühl J, Bremer SCB, Knoop RF, Ellenrieder V, Kunsch S, Neesse A (2020): Diagnostic accuracy of B-Mode ultrasound and Hepatorenal Index for graduation of hepatic steatosis in patients with chronic liver disease. *PLOS ONE* 15, e0231044
- Poynard T, Ratziu V, Naveau S, Thabut D, Charlotte F, Messous D, Capron D, Abella A, Massard J, Ngo Y, et al. (2005): The diagnostic value of biomarkers (SteatoTest) for the prediction of liver steatosis. *Comp Hepatol* 4, 10
- Poynard T, Peta V, Munteanu M, Charlotte F, Ngo Y, Ngo A, Perazzo H, Deckmyn O, Pais R, Mathurin P, et al. (2019): The diagnostic performance of a simplified blood test (SteatoTest-2) for the prediction of liver steatosis: *Eur J Gastroenterol Hepatol* 31, 393–402
- Ratziu V, Charlotte F, Heurtier A, Gombert S, Giral P, Bruckert E, Grimaldi A, Capron F, Poynard T (2005): Sampling Variability of Liver Biopsy in Nonalcoholic Fatty Liver Disease. *Gastroenterology* 128, 1898–1906
- Roeb E, Steffen H, Bantel H, Baumann U, Canbay A, Demir M, Drebber U, Geier A, Hampe J, Hellerbrand C, et al. (2015): S2k-Leitlinie nicht alkoholische Fettlebererkrankung. *Z Gastroenterol* 53, 668–723
- Saadah S, Younossi ZM, Remer EM, Gramlich T, Ong JP, Hurley M, Mullen KD, Cooper JN, Sheridan MJ (2002): The utility of radiological imaging in nonalcoholic fatty liver disease. *Gastroenterology* 123, 745–750
- Schirmacher P, Fleig WE, Tannapfel A, Langner C, Dries V, Terracciano L, Denk H, Dienes HP (2004): Biopsische Diagnostik der chronischen Hepatitis: Ergebnisse einer evidenzbasierten Konsensuskonferenz der Deutschen Gesellschaft für Pathologie (DGP), der Deutschen Gesellschaft für Verdauungs- und Stoffwechselkrankheiten (DGVS) und des Kompetenznetzes Hepatitis (HepNet). *Pathol* 25, 337–348
- Schwarz J, Käch W, Bruderer Enzler H (2018): Datenanalyse mit SPSS. https://www.methodenberatung.uzh.ch/de/datenanalyse_spss.html; Zugriff am 29.8.2019
- Schwarz J, Käch W, Bruderer Enzler H (2020): Rangkorrelation nach Spearman. https://www.methodenberatung.uzh.ch/de/datenanalyse_spss/zusammenhaenge/rangkorrelation.html; Zugriff am 29.4.2021
- Stättermayer AF, Traussnigg S, Dienes HP, Aigner E, Stauber R, Lackner K, Hofer H, Stift J, Wrba F, Stadlmayr A, et al. (2015): Hepatic steatosis in Wilson disease – Role of copper and PNPLA3 mutations. *J Hepatol* 63, 156–163
- Stepanova M, Younossi ZM (2012): Independent Association Between Nonalcoholic Fatty Liver Disease and Cardiovascular Disease in the US Population. *Clin Gastroenterol Hepatol* 10, 646–650

- Strauss S, Gavish E, Gottlieb P, Katsnelson L (2007): Interobserver and intraobserver variability in the sonographic assessment of fatty liver. *Am J Roentgenol* 189, W320–W323
- Strnad P, Buch S, Hamesch K, Fischer J, Rosendahl J, Schmelz R, Brueckner S, Brosch M, Heimes CV, Woditsch V, et al. (2019): Heterozygous carriage of the alpha1-antitrypsin Pi*Z variant increases the risk to develop liver cirrhosis. *Gut* 68, 1099–1107
- Szczepaniak LS, Nurenberg P, Leonard D, Browning JD, Reingold JS, Grundy S, Hobbs HH, Dobbins RL (2005): Magnetic resonance spectroscopy to measure hepatic triglyceride content: prevalence of hepatic steatosis in the general population. *Am J Physiol-Endocrinol Metab* 288, E462–E468
- Tannapfel A, Denk H, Dienes HP, Langner C, Schirmacher P, Trauner M, Flott-Rahmel B (2011): Histopathological diagnosis of non-alcoholic and alcoholic fatty liver disease. *Virchows Arch* 458, 511–523
- Vernon G, Baranova A, Younossi ZM (2011): Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. *Aliment Pharmacol Ther* 34, 274–285
- Wang JH, Hung CH, Kuo FY, Eng HL, Chen CH, Lee CM, Lu SN, Hu TH (2013): Ultrasonographic quantification of hepatic–renal echogenicity difference in hepatic steatosis diagnosis. *Dig Dis Sci* 58, 2993–3000
- Webb M, Yeshua H, Zelber-Sagi S, Santo E, Brazowski E, Halpern Z, Oren R (2009): Diagnostic value of a computerized hepatorenal index for sonographic quantification of liver steatosis. *Am J Roentgenol* 192, 909–914
- Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M (2016): Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology* 64, 73–84
- Younossi ZM, Anstee QM, Marietti M, Hardy T, Henry L, Eslam M, George J, Bugianesi E (2018): Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention. *Nat Rev Gastroenterol Hepatol* 15, 11–20

Danksagung

Ich möchte mich bei meinem Doktorvater PD Dr. Kunsch für die Gelegenheit und Förderung zu dieser Promotion bedanken. Dieser Dank gilt auch meinem Betreuer PD Dr. Golo Petzold. Durch seine gute Betreuung war diese Arbeit zwar durchaus eine Arbeit, aber niemals eine Belastung.