

Aus dem Lehrstuhl
für Epidemiologie und Präventivmedizin
Prof. Dr. Dr. Michael Leitzmann
der Fakultät für Medizin
der Universität Regensburg

Qualitätssicherung der Ultraschalluntersuchung des Abdominalfetts
im Rahmen des Data Freeze 100k der NAKO Gesundheitsstudie

Inaugural – Dissertation
zur Erlangung des Doktorgrades
der Medizin

der
Fakultät für Medizin
der Universität Regensburg

vorgelegt von
Bettina Clarissa Jobst

2021

Aus dem Lehrstuhl
für Epidemiologie und Präventivmedizin
Prof. Dr. Dr. Michael Leitzmann
der Fakultät für Medizin
der Universität Regensburg

Qualitätssicherung der Ultraschalluntersuchung des Abdominalfetts
im Rahmen des Data Freeze 100k der NAKO Gesundheitsstudie

Inaugural – Dissertation
zur Erlangung des Doktorgrades
der Medizin

der
Fakultät für Medizin
der Universität Regensburg

vorgelegt von
Bettina Clarissa Jobst

2021

Dekan: Prof. Dr. Dirk Hellwig

1. Berichterstatter: Prof. Dr. Dr. Michael Leitzmann

2. Berichterstatter: Prof. Dr. Christian Stroszczyński

Tag der mündlichen Prüfung: 01.12.2021

Inhaltsverzeichnis

1	EINLEITUNG	14
1.1	Adipositas und die Bedeutung der Körperfettverteilung	14
1.2	Methoden zur Quantifizierung des SAT und VAT	17
1.3	Qualitätssicherung in der NAKO Gesundheitsstudie	18
1.4	Ziele der Arbeit	21
2	MATERIAL UND METHODEN	23
2.1	Überarbeitung der SOP „US Abdominalfett“	23
2.2	Datenaufbereitung und -analyse des Data Freeze 100k-Datensatzes	24
2.2.1	Anforderungen und theoretischer Hintergrund	24
2.2.2	Vollständigkeit der US-Untersuchung des Abdominalfetts	25
2.2.3	Umgang mit unplausiblen Werten des SAT und VAT	28
2.2.4	Fehlende Werte des SAT und VAT	30
2.2.5	Verteilung der Werte des SAT und VAT	30
2.2.6	SZ-, Untersucher- und Zeiteffekte auf die Werte des SAT und VAT	30
2.2.7	Interne Validierung der Doppelmessung des SAT und VAT	31
2.3	Bewertung der US-Bilder	32
2.3.1	Stichprobe	33
2.3.2	Bewertung mittels Qualitätskriterien	33
2.3.3	Analyse der US-Bildbewertung	35
2.3.4	Rückmeldung an die Untersucher	35
2.4	Qualitätsbericht und Empfehlungen zur US-Messung	36
3	ERGEBNISSE	37
3.1	Überarbeitete SOP „US Abdominalfett“	37
3.2	Datenaufbereitung und –analyse des Data Freeze 100k-Datensatzes	41
3.2.1	Vollständigkeit der Untersuchung US Abdominalfett	41
3.2.2	Umgang mit unplausiblen Werten des SAT und VAT	44
3.2.3	Fehlende Werte des SAT und VAT	45
3.2.4	Verteilung der Werte des SAT und VAT	46
3.2.5	SZ-, Untersucher- und Zeiteffekte auf die Werte des SAT und VAT	49
3.2.6	Interne Validierung der Doppelmessung des SAT und VAT	64

3.3	Bewertung der US-Bilder	68
3.3.1	Stichprobe.....	68
3.3.2	Analyse der US-Bildbewertung.....	69
3.3.3	Rückmeldung an die Untersucher	76
3.4	Qualitätsbericht und Empfehlungen zur US-Messung	77
4	DISKUSSION.....	78
4.1	Überarbeitung der SOP „US Abdominalfett“	78
4.2	Datenaufbereitung und -analyse des Data Freeze 100k-Datensatzes.....	79
4.3	Bewertung der US-Bilder	86
5	ZUSAMMENFASSUNG	90
6	ANHANG	95
6.1	ANOM-Plots zu Untersuchereffekten	95
6.2	Schablonen für die US-Bildbewertung	97
6.3	Rückmeldung zur US-Messung des Abdominalfetts	97
6.4	Empfehlungen zur US-Messung des Abdominalfetts.....	99
7	LITERATURVERZEICHNIS.....	100
8	ERKLÄRUNG	109
9	DANKSAGUNG	110

Abkürzungsverzeichnis

ANOM	Analyse der Mittelwerte, <i>Analysis of Means</i>
ANOVA	Analyse der Varianz, <i>Analysis of Variance</i>
BMI	<i>Body-Mass-Index</i>
cm	Zentimeter
CRP	C-reaktives Protein
CT	Computertomographie
DALY	<i>Disability-Adjusted Life Years</i>
DEGS1	Studie zur Gesundheit Erwachsener in Deutschland des Robert Koch-Instituts
GEDA	Gesundheit in Deutschland aktuell, Studie des Robert Koch-Instituts
GEP	Gute Epidemiologische Praxis
ICC	Intraklassenkorrelationskoeffizient, <i>Intraclass Correlation Coefficient</i>
IL-6	Interleukin 6
iScan	Automatischen Optimierung von Gain und TGC
LGC	Laterale Verstärkung, <i>Lateral Gain Compensation</i>
MRT	Magnetresonanztomographie
n	Anzahl
NA	Nicht verfügbar, <i>Not Available</i>
NAKO	Nationale Kohorte
QS	Qualitätssicherung
R	Programmiersprache R
SAT	Subkutanes Fettgewebe, <i>Subcutaneous Adipose Tissue</i>
SAT1	Erster Messwert des subkutanen Fettgewebes
SAT2	Zweiter Messwert des subkutanen Fettgewebes
SD	Standardabweichung, <i>Standard Deviation</i>
SOP	Standardarbeitsanweisung, <i>Standard Operating Procedure</i>
SZ	Studienzentrum
TGC	Tiefenabhängige Verstärkung, <i>Time Gain Compensation</i>

TNF α	Tumor-Nekrose-Faktor α
US	Ultraschall
VAT	Viszerales Fettgewebe, <i>Visceral Adipose Tissue</i>
VAT1	Erster Messwert des viszeralen Fettgewebes
VAT2	Zweiter Messwert des viszeralen Fettgewebes
WHO	Weltgesundheitsorganisation, <i>World Health Organization</i>
WHR	Verhältnis von Taillen- zu Hüftumfang, <i>Waist-Hip-Ratio</i>
100k	100.000

Abbildungsverzeichnis

Abbildung 1: Studienzentren und Cluster der NAKO Gesundheitsstudie.	20
Abbildung 2: Beschriftung einer Fotografie der Messregler des US-Geräts für die überarbeitete Version der SOP „US Abdominalfett“	39
Abbildung 3: Anatomische Strukturen: Transversalschnitte durch die Rektusscheide kranial der Linea arcuata sowie durch das Abdomen auf Höhe des ersten Lendenwirbelkörpers.....	40
Abbildung 4: US-Bilder von a) SAT und b) VAT mitsamt der Messstrecken unter Optimierung der Eindringtiefe, Fokuseinstellung und Helligkeit.	40
Abbildung 5: Flowchart zur Vollständigkeit der US-Untersuchungen.....	42
Abbildung 6: Perzentilen des SAT [cm] nach Altersgruppen bei a) Männern (N = 10.075) und b) Frauen (N = 10.278).	48
Abbildung 7: Perzentilen des VAT [cm] nach Altersgruppen bei a) Männern (N = 9.960) und b) Frauen (N = 10.147).....	49
Abbildung 8: Verhältnis von SAT zu VAT nach Alter und Geschlecht. ...	49
Abbildung 9: ANOM-Plot zur Abweichung der SZ-Mittelwerte des SAT vom globalen Mittelwert 2,2 cm (N = 20.353).....	52
Abbildung 10: ANOM-Plot zur Abweichung der SZ-Mittelwerte des VAT vom globalen Mittelwert 6,7 cm (N = 20.107).....	52
Abbildung 11: ANOM-Plot zur Abweichung der Clustermittelwerte des SAT vom globalen Mittelwert 2,2 cm (N = 20.353).	54
Abbildung 12: ANOM-Plot zur Abweichung der Clustermittelwerte des VAT vom globalen Mittelwert 6,7 cm (N = 20.107).	54
Abbildung 13: ANOM-Plot zur Abweichung der Untersuchermittelwerte des SAT vom globalen Mittelwert 2,2 cm (N = 18.890).	57
Abbildung 14: ANOM-Plot zur Abweichung der Untersuchermittelwerte des VAT vom globalen Mittelwert 6,8 cm (N = 18.666).	58
Abbildung 15: Variation der Mittelwerte des SAT und VAT nach Studienmonaten, adjustiert nach Alter, Geschlecht und SZ.	60

Abbildung 16: ANOM-Plot zur Abweichung der Studienmonatsmittelwerte des SAT vom globalen Mittelwert 2,2 cm (N = 20.203).....	61
Abbildung 17: ANOM-Plot zur Abweichung der Studienmonatsmittelwerte des VAT vom globalen Mittelwert 6,7 cm (N = 19.960).....	61
Abbildung 18: Variation der Mittelwerte des SAT und VAT nach Monat, adjustiert nach Alter, Geschlecht und SZ.	62
Abbildung 19: ANOM-Plots zur Abweichung der Monatsmittelwerte des SAT vom globalen Mittelwert 2,2 cm (N = 20.203) sowie des VAT vom globalen Mittelwert 6,7 cm (N = 19.960).	63
Abbildung 20: Variation der Mittelwerte des SAT und VAT nach Uhrzeit, adjustiert nach Alter, Geschlecht und SZ.	63
Abbildung 21: ANOM-Plots zur Abweichung der Mittelwerte des SAT nach Uhrzeit vom globalen Mittelwert 2,2 cm (N = 20.217) sowie des VAT nach Uhrzeit vom globalen Mittelwert 6,7 cm (N = 19.974).	64
Abbildung 22: Bland-Altman-Plot der Doppelmessung des SAT (N = 18.883).	65
Abbildung 23: Bland-Altman-Plot der Doppelmessung des VAT (N = 18.629).	66
Abbildung 24: Flowchart zur Stichprobe für die Bewertung der US-Bilder.....	68
Abbildung 25: Fehler bei der Messung des SAT.	71
Abbildung 26: Fehler bei der Positionierung des oberen Messpunktes des VAT.	71
Abbildung 27: Fehler bei der Positionierung des unteren Messpunktes des VAT. ...	72
Abbildung 28: Mangelhafte Bildqualität bei der Messung des SAT und VAT.....	75
Abbildung 29: Beispielbilder zur Messrichtigkeit bei der Messung von a) SAT und b) VAT.	77
Abbildung 30: Beispielbilder zur Bildqualität bei der Messung von a) SAT und b) VAT.	77
Abbildung 31: ANOM-Plot zur Abweichung der Untersuchermittelwerte des SAT vom globalen Mittelwert 2,2 cm (N = 18.890) inklusive Untersuchernummern und p-Werten.	95

Abbildung 32: ANOM-Plot zur Abweichung der Untersucher Mittelwerte des VAT vom globalen Mittelwert 6,8 cm (N = 18.666) inklusive Untersuchernummern und p-Werten.	96
Abbildung 33: Schablonen für die Bewertung der formatfüllenden Darstellung der Zielstrukturen und der senkrechten Ausrichtung der Messlinie.	97
Abbildung 34: Beispielrückmeldung.	98
Abbildung 35: Empfehlungen zur US-Messung des Abdominalfetts.	99

Tabellenverzeichnis

Tabelle 1: Kodierung der Variablen Status der Untersuchung in der Erhebungssoftware und Durchführbarkeit der Untersuchung	26
Tabelle 2: Kategorien der Vollständigkeit der Untersuchung.	26
Tabelle 3: Kategorien und Suchbegriffe zur Zuordnung der Freitexteingaben der Variable Nichtdurchführbarkeit – Sonstiges zu anderen Variablen.	27
Tabelle 4: Kategorien und Suchbegriffe zur Kategorisierung der Freitexteingaben der Variable Nichtdurchführbarkeit – Sonstiges.	28
Tabelle 5: Plausibilitätsgrenzen der Messwerte des SAT und VAT.	29
Tabelle 6: Kodierung der Variable Grund für die Korrektur eines Messwerts.	29
Tabelle 7: Kriterien für die Bewertung der Messrichtigkeit und Bildqualität der US-Bilder.	34
Tabelle 8: Standardisierte Formulierung bei der Rückmeldung an die Untersucher zur Bewertung der US-Bilder	35
Tabelle 9: Vergleich der Kriterien für ein qualitativ hochwertiges US-Bild gemäß einem Sonographie-Lehrbuch mit der SOP „US Abdominalfett“	38
Tabelle 10: Vollständigkeit der US-Untersuchung des Abdominalfetts.	43
Tabelle 11: Verteilung der Gründe für nicht durchgeführte US-Untersuchungen, in 16 Fällen Mehrfachangabe von Gründen.	44
Tabelle 12: Korrekturanlässe unplausibler Messwerte.	45
Tabelle 13: Fehlende Messwerte der US-Untersuchung des Abdominalfetts nach SZ.	46
Tabelle 14: Charakteristika der analysierten Studienteilnehmer des Data Freeze 100k Datensatzes.	47
Tabelle 15: Perzentilen des SAT	47
Tabelle 16: Perzentilen des VAT.	48
Tabelle 17: Verteilung der Messwerte des SAT und VAT in Abhängigkeit des SZ, adjustiert nach Alter und Geschlecht der Studienteilnehmer.	50

Tabelle 18: Verteilung der Messwerte des SAT und VAT in Abhängigkeit des Clusters, adjustiert nach Alter und Geschlecht der Studienteilnehmer.....	53
Tabelle 19: Verteilung der Messwerte des SAT und VAT in Abhängigkeit des Untersuchers, adjustiert nach Alter und Geschlecht der Studienteilnehmer. .	55
Tabelle 20: Verteilung der Messwerte des SAT und VAT in Abhängigkeit der Zeit, adjustiert nach Alter, Geschlecht und SZ.	59
Tabelle 21: Intraobserver-Reliabilität und mittlere Differenz bei der US-Messung des SAT und VAT.	66
Tabelle 22: Charakteristika der Studienteilnehmer der Stichprobe für die Bewertung der US-Bilder.	69
Tabelle 23: Bewertung der Messrichtigkeit in Abhängigkeit des Geschlechts und Alters der Studienteilnehmer, der Menge des SAT und VAT und des SZ.	69
Tabelle 24: Logistische Regressionsmodelle zum Einfluss des Geschlechts und Alters der Studienteilnehmer, der Menge des SAT und VAT und des SZ auf die Bewertungskriterien der Messrichtigkeit.....	73
Tabelle 25: Bewertung der Bildqualität in Abhängigkeit des Geschlechts und Alters der Studienteilnehmer, der Menge des SAT und VAT und des SZ.	74
Tabelle 26: Logistische Regressionsmodelle zum Einfluss des Geschlechts und Alters der Studienteilnehmer, der Menge des SAT und VAT und des SZ auf die Bewertungskriterien der Bildqualität.....	76

1 Einleitung

1.1 Adipositas und die Bedeutung der Körperfettverteilung

Übergewicht und Adipositas werden von der Weltgesundheitsorganisation (WHO) als epidemisches Gesundheitsproblem bezeichnet. Die Prävalenz von Übergewicht und Adipositas nimmt nicht nur in den Industrieländern stetig zu. Sie betrug weltweit bei Erwachsenen im Jahr 2016 39% für Übergewicht und 13% für Adipositas (1). In Deutschland waren gemäß der Studie „Gesundheit in Deutschland aktuell“ (GEDA) des Robert Koch-Instituts im Jahr 2014/2015 36% der Befragten übergewichtig und 18% adipös (Anzahl (n) = 24.016), wobei die Prävalenzen auf Selbstangaben beruhten. In der Untersuchungsstudie DEGS1 zur Gesundheit Erwachsener in Deutschland lag die Adipositasprävalenz um das Jahr 2010 bei knapp 24%. Hier wurden Körpergewicht und Körpergröße nach einem standardisierten Verfahren gemessen (2). Anhand dieser beiden anthropometrischen Variablen wird der Body-Mass-Index (BMI) berechnet, nach welchem weltweit Übergewicht und Adipositas klassifiziert werden. Die WHO sieht für Übergewicht einen Grenzwert von 25 kg/m^2 vor, von Adipositas wird ab einem BMI von 30 kg/m^2 gesprochen (3).

Ein hoher BMI ist ein etablierter Risikofaktor für nicht übertragbare, chronische Erkrankungen. Insbesondere im Rahmen des metabolischen Syndroms prädisponiert Adipositas für metabolische und kardiovaskuläre Erkrankungen (4). Darüber hinaus sind auch Nierenerkrankungen, gastrointestinale Erkrankungen, psychische Erkrankungen, Erkrankungen des Bewegungsapparats und einige Krebserkrankungen mit Übergewicht und Adipositas assoziiert (5). Folglich spielen Übergewicht und Adipositas eine wichtige Rolle bei der Zunahme der „Disability-Adjusted Life Years“ (DALY), die den Verlust an gesunden Lebensjahren durch Erkrankungen oder Todesursachen quantifizieren (6). Der BMI ist einfach zu erfassen, differenziert jedoch weder zwischen Körperfett und fettfreier Masse noch gibt er Auskunft über die Körperfettverteilung (5). Letztere ist entscheidend für das Risiko an Folgeerkrankungen, wobei insbesondere eine abdominelle Fetteinlagerung mit einem erhöhten Risiko für Diabetes, kardiovaskuläre Erkrankungen und einige Krebserkrankungen assoziiert ist (7, 8).

Das Gesamtkörperfett wird üblicherweise in zwei Hauptkomponenten unterteilt: das subkutane Fettgewebe (SAT), das sich zwischen Haut und Muskulatur befindet und das viszerale Fettgewebe (VAT), das in den Körperhöhlen und dabei größtenteils im Omentum majus und Mesenterium in der Bauchhöhle eingelagert ist (9). Darüber hinaus finden sich ektope Fettdepots beispielsweise in Muskulatur, Leber, Niere, Herz und Pankreas. Die Körperfettverteilung hängt hauptsächlich von Alter, Geschlecht, dem Gesamtkörperfettgehalt, der Energiebilanz und genetischen Faktoren ab (10, 11). Unabhängig von der Gesamtkörperfettmasse nimmt das Verhältnis von VAT zu SAT insbesondere bei Männern mit dem Alter zu (12). Auch Sexualhormone modulieren die Körperfettverteilung, sodass bei prämenopausalen Frauen eher der gynoide, birnenförmige Fettverteilungstyp mit peripherer, subkutaner Fetteinlagerung zu finden ist. Männer und postmenopausale Frauen tendieren hingegen zu einer zentralen, intra-abdominalen Fetteinlagerung (13, 14). Diese Form der Körperfettverteilung wird auch als androider oder apfelförmiger Typ bezeichnet (15). Eine weitere wichtige Ursache für Gewichtszunahme ist Überernährung (16). In Folge einer positiven Energiebilanz werden vermehrt Triglyzeride in Adipozyten des SAT eingelagert. Dabei fungiert das SAT als metabolischer Puffer, in dem es andere Organe vor Fetteinlagerung schützt. Wenn die Kapazität des SATs erschöpft ist, akkumuliert Fett im viszeralen und ektopen Fettgewebe (17, 18). Im ektopen Fettgewebe führt die Lipidakkumulation zu strukturellen und funktionellen Schäden, was als Lipotoxizität bezeichnet wird (19). Auch bei bestimmten genetischen Konstellationen, die mit dysfunktionalem SAT einhergehen, findet die Lipidspeicherung bevorzugt im viszeralen und ektopen Fettgewebe statt (20).

Durch Hypertrophie des VAT kommt es zur Dysregulation der Sekretion und des Metabolismus der dort lokalisierten Adipozyten. Über proinflammatorische Signalwege werden Makrophagen angelockt, die das VAT infiltrieren. Sie sezernieren Zytokine wie Tumornekrosefaktor- α (TNF- α) und Interleukin-6 (IL-6), die lokal wie peripher die Insulinsensitivität verschlechtern (21). Zudem aktiviert TNF- α die Lipolyse im VAT, sodass vermehrt Triglyzeride über die Pfortader in die Leber gelangen. Diese bewirken dort eine erhöhte Produktion von Glucose und reduzieren die Degradation von Insulin. In Folge kommt es zu Hyperglykämie und Hyperinsulinämie (22). IL-6 stimuliert darüber hinaus die Synthese des C-reaktiven Proteins (CRP) in der Leber und hemmt die Adiponektin-Sekretion im Fettgewebe, sodass ein chronischer subakuter Entzündungszustand entsteht (23). Über diesen Pathomechanismus begünstigt

viszerale Adipositas die Entwicklung metabolischer Erkrankungen wie Diabetes mellitus Typ 2 und kardiovaskulärer Erkrankungen wie koronare Herzkrankheit und Schlaganfall (24).

Im Gegensatz dazu wird insbesondere vom SAT das Peptidhormon Adiponektin freigesetzt (5). Adiponektin wirkt antidiabetisch, indem es in der Skelettmuskulatur die Glucoseaufnahme und Fettsäureoxidation stimuliert und in der Leber die Gluconeogenese reduziert (25). Zudem zeigte es in Tiermodellen antiatherogene Effekte, da es die Expression endothelialer Adhäsionsmoleküle hemmt und die Mediaproliferation sowie die Schaumzellbildung unterdrückt (26). Trotz dieses potentiell protektiven Effekts korrelierte das SAT in verschiedenen Studien ebenfalls mit kardiovaskulären und metabolischen Risikofaktoren (27–29). Ein möglicher Pathomechanismus findet sich in der Beobachtung, dass bei Adipositas Stress auf das endoplasmatische Retikulum der subkutanen Adipozyten ausgeübt und dadurch weniger Adiponektin freigesetzt wird (30, 31).

Zur Abschätzung der Körperfettverteilung können der Taillenumfang oder das Verhältnis von Taillen- zu Hüftumfang („Waist Hip Ratio“, WHR) verwendet werden. Dem liegt die Annahme zugrunde, dass SAT und VAT im Bauchraum durch den Taillenumfang repräsentiert werden, während der Hüftumfang lediglich femorales und gluteales SAT misst (32). Die WHR erlaubt im Gegensatz zum BMI die Identifikation normalgewichtiger Personen mit abdomineller Fetteinlagerung, die ebenfalls ein erhöhtes Risiko für kardiovaskuläre Erkrankungen und Mortalität aufweisen (33). Dieser metabolisch kranke, normalgewichtige Phänotyp wurde erstmals im Jahr 1981 beschrieben und zeichnet sich durch Risikofaktoren wie Hyperinsulinämie und Insulinresistenz aus (34). Dahingegen konnte auch eine Subgruppe Adipöser identifiziert werden, die metabolisch gesund ist (35), wobei auch bei dieser ein erhöhtes Risiko für kardiovaskuläre Erkrankungen nachgewiesen wurde (36). Der Taillenumfang korrelierte in Studien stärker als der BMI mit dem VAT sowie dem Risiko für Folgeerkrankungen (37, 38), während der Hüftumfang erwartungsgemäß stark mit dem SAT korrelierte (38).

1.2 Methoden zur Quantifizierung des SAT und VAT

Für die genaue Differenzierung und Quantifizierung von SAT und VAT bieten sich bildgebende Verfahren an. Den Goldstandard für die Messung des VAT stellen Schnittbildverfahren wie Computertomographie (CT) und Magnetresonanztomographie (MRT) dar. Die Menge des Fettgewebes kann dabei zweidimensional unter Verwendung einer einzelnen transversalen Schnittebene oder auch dreidimensional mittels multipler Schnittbilder abgeschätzt werden (32). Einzelschnittbilder werden meist auf der Höhe des zweiten oder dritten Lendenwirbelkörpers aufgenommen (39). Die Identifizierung des Fettgewebes erfolgt über Hounsfield-Einheiten bei der CT beziehungsweise Hyperintensität des Fettgewebes im Vergleich zum Umgebungsgewebe bei verschiedenen MRT-Sequenzen (40). Aufgrund der hohen Kosten und Dauer der MRT beziehungsweise der Strahlenbelastung der CT sind diese Methoden für große epidemiologische Studien nur bedingt geeignet (32).

Eine Alternative stellt die Sonographie dar, da sie schnell, kostengünstig und für den Studienteilnehmer¹ wenig belastend durchgeführt werden kann. Fettgewebe grenzt sich dabei aufgrund seiner niedrigeren Impedanz von umliegendem Gewebe echoarm ab (41). Vorreiter auf dem Gebiet der Messung des abdominellen Fettgewebes mittels Ultraschall (US) waren Armellini et al. im Jahr 1990. Sie verwendeten als Korrelat für das SAT die Strecke zwischen Haut-Fett-Linie und Musculus rectus abdominis, für das VAT die Strecke zwischen Musculus rectus abdominis und Hinterwand der Aorta. Beide Messungen führten sie im Bereich der Linea xiphoumbilicalis kranial des Bauchnabels durch (42). Dabei erreichten sie Korrelationskoeffizienten der US- und CT-Messwerte zwischen 0,68 und 0,74 (43). Für die US-Messstrecke des SAT konnte im Weiteren eine starke Korrelation zur CT-Messstrecke gezeigt werden (44). Zwischen US und MRT war die Korrelation des SAT mit einem Korrelationskoeffizienten von 0,73 moderat, wobei das SAT durch die US-Messung tendenziell etwas überschätzt wurde (45). Für die Distanzmessung des VAT konnte in einer Metaanalyse keine signifikante Differenz zwischen US- und CT- beziehungsweise MRT-Messwerten gefunden werden (46). Die Distanzmessung mittels US korrelierte in verschiedenen Studien moderat bis stark mit der Fläche des VAT, die mittels CT oder MRT bestimmt wurde (44, 45, 47–49).

¹ Zur besseren Lesbarkeit werden personenbezogene Bezeichnungen nur in der männlichen Form aufgeführt. Dies soll jedoch keinesfalls eine Geschlechterdiskriminierung oder eine Verletzung des Gleichheitsgrundsatzes zum Ausdruck bringen.

Mehrere Studien zur Reproduzierbarkeit von US-Messwerten des SAT und VAT zeigten hohe Intraklassenkorrelationskoeffizienten über 0,9 sowohl bei der Analyse der Intra- als auch Interraterreliabilität (44, 47, 48). Trotzdem schätzten viele Autoren die US-Methode für die Messung des abdominellen Fettgewebes als stark untersucherabhängig ein. Deshalb befürworteten sie die Verwendung eines standardisierten Untersuchungsprotokolls zur Gewährleistung von Präzision und Reliabilität (24, 32). Azzi et al. fassten ihre Empfehlungen zur Messung des SAT wie folgt zusammen (50): Das SAT sollte als Strecke 2 Zentimeter (cm) kranial oder kaudal des Bauchnabels in der Medianlinie am liegenden Probanden gemessen werden. Sie empfahlen die Messung von der Haut-Fett-Linie zur Linea alba als Korrelat der Fett-Muskel-Linie. Bei der Messung sollte möglichst wenig Druck angewendet werden und die US-Sonde im 90°-Winkel zur Haut gehalten werden. Dennoch unterschied sich die Messstrecke in verschiedenen europäischen Kohortenstudien, wobei teils die Haut mitgemessen wurde, teils der Musculus rectus abdominis statt der Linea alba als Messpunkt verwendet wurde (51–53). Auch bei der Messung des VAT wurde in den letzten zwei Jahrzehnten kein einheitliches Protokoll verwendet. Der ventrale Messpunkt lag entweder am Musculus rectus abdominis oder am Peritoneum, der dorsale Messpunkt wurde an der Vorderwand der Aorta oder an der Vorderkante der Lendenwirbelsäule positioniert (51–55). Meist wurde am Ende der Ausatmung gemessen und dabei möglichst wenig Druck auf die US-Sonde ausgeübt.

1.3 Qualitätssicherung in der NAKO Gesundheitsstudie

Auch im Rahmen der NAKO Gesundheitsstudie wurden SAT und VAT mittels US gemessen, um die Körperfettverteilung zu quantifizieren. Bei der NAKO (Nationale Kohorte) Gesundheitsstudie handelte es sich um eine bevölkerungsbezogene prospektive Kohortenstudie. Ihr Ziel war die Aufklärung von Ursachen chronischer Krankheiten und deren Zusammenhang mit genetischen, Lebensstil- und Umweltfaktoren. Die Identifikation neuer Risikofaktoren und die Entwicklung von Risikovorhersagemodellen sollten zur Entwicklung personalisierter Präventionsstrategien beitragen. Darüber hinaus sollten Marker zur Früherkennung chronischer Erkrankungen identifiziert werden (56). Für die epidemiologische Forschung sollte eine umfangreiche standardisierte Daten- und Bioprobenbank aufgebaut werden, wobei großen Wert auf Kompatibilität des Studienprogramms mit anderen europäischen Kohortenstudien gelegt wurde (57).

Im Zeitraum von 2014 bis 2019 wurden für die Basisuntersuchung mehr als 205.000 Männer und Frauen im Alter von 20 bis 69 Jahren per Zufallsstichprobe von Einwohnern definierter geographischer Regionen rekrutiert (57). Diese geographischen Regionen ballten sich um die 18 beteiligten Studienzentren (SZ), die sich wiederum acht Clustern zuordnen ließen (58). Die Zufallsstichprobe wurde nach Alter und Geschlecht stratifiziert, wobei die höheren Altersgruppen zugunsten der jüngeren Altersgruppen überrepräsentiert waren (56). Bei allen Teilnehmern wurde ein etwa vierstündiges standardisiertes Untersuchungsprogramm durchgeführt (Level 1), das unter anderem ein gesundheitsbezogenes Interview, die Gewinnung von Bioproben und eine körperliche Untersuchung umfasste. Letztere beinhaltete die Messung von Körpergröße und Körpergewicht, Körperzusammensetzung mittels bioelektrischer Impedanzanalyse, Taillen- und Hüftumfang. Zusätzlich wurden Sekundär- und Registerdaten wie beispielsweise von Kranken- und Rentenversicherung sowie Krebsregistern gewonnen. Eine zufällig ausgewählte Unterstichprobe von 40.000 Personen (20%) durchlief ein umfassenderes Untersuchungsprogramm, das als Level 2 bezeichnet wurde. Hierbei wurde unter anderem das subkutane und viszerale abdominale Fettgewebe mittels US quantifiziert, was Gegenstand dieser Arbeit ist. Eine weitere Subgruppe von gut 30.000 Level-2-Studienteilnehmern erhielt in fünf SZ zusätzlich eine Ganzkörper-MRT (58). Alle Untersuchungen wurden von qualifiziertem Studienpersonal durchgeführt, das meist aus dem Bereich der Krankenpflege und medizinischen Fachanstellung stammte. Direkt im Anschluss an die Basisuntersuchung startete die Zweituntersuchung, die ein einheitliches Untersuchungsprogramm vorsah. Darüber hinaus wurde zwei bis drei Jahre nach der Basiserhebung ein schriftliches Follow-up durchgeführt (57). Ziel war die Erfassung inzidenter Erkrankungen sowie die Veränderung von Risikofaktoren, um auf die präklinische Morbidität rückzuschließen.



Abbildung 1: Studienzentren und Cluster der NAKO Gesundheitsstudie (56).

Um eine hohe Qualität der Datenerhebung in der NAKO Gesundheitsstudie zu gewährleisten, wurde ein umfassendes Qualitätsmanagementsystem etabliert (57). Als Orientierung dienten die „Leitlinien und Empfehlungen zur Sicherung von Guter Epidemiologischer Praxis (GEP)“ (59). Die interne Qualitätssicherung (QS) erfolgte durch ein zentrales Qualitätsmanagement, QS-Beauftragte der SZ sowie weitere QS-Verantwortliche. Zusätzlich fand eine externe Qualitätssicherung durch die Abteilung für Epidemiologie und Gesundheitsmonitoring des Robert-Koch-Instituts statt. Für alle Arbeitsschritte in den SZ wurden Standardarbeitsabläufe und Leitfäden entwickelt, anhand derer die Mitarbeiter geschult und zertifiziert wurden. Nach dem Train-the-Trainer-Prinzip wurden zudem einzelne Mitarbeiter in den SZ durch die Modulverantwortlichen als Trainer ausgebildet (57). Auch die Verwendung einheitlicher Untersuchungsgeräte und –materialien und deren regelmäßige Kalibrierung trug zur Standardisierung der Datenerhebung bei. Modulverantwortliche und fachspezifische

Kompetenzzentren wurden mit der Betreuung der Untersuchungsmodule und Aufbereitung der erhobenen Daten anvertraut. Von 2011 bis 2014 wurden vor Beginn der Basiserhebung in einer Prätestphase der komplette Untersuchungsablauf und die einzelnen Befragungs- und Untersuchungselemente sorgfältig getestet (57, 60). Während des Studienverlaufs wurden regelmäßig Site Visits durch geschulte und zertifizierte Mitarbeiter anderer SZ durchgeführt, um Abweichungen vom Studienprotokoll aufzudecken und Handlungsempfehlungen auszusprechen. Die erhobenen Messwerte wurden einerseits bei ihrer Eingabe in das verwendete Datenerhebungsprogramm automatisch auf Plausibilität und Vollständigkeit überprüft. Andererseits wurden durch das Datenintegrationszentrum die eingehenden Daten fortlaufend auf Vollständigkeit sowie Alters- und Geschlechtsverteilung der Teilnehmer geprüft und an das zentrale Qualitätsmanagement rückgemeldet. Darüber hinaus fand eine Kalibrierungsstudie anhand einer Zufallsstichprobe von 6000 Level-2-Studienteilnehmern aller zwölf SZ statt, wobei innerhalb von einem bis zwölf Monaten nach der Basisuntersuchung nahezu das komplette Untersuchungsprogramm wiederholt wurde. Ziel war die Quantifizierung der Messwertvariabilität, die durch zufällige Messfehler und intraindividuelle Schwankungen der Studienteilnehmer zustande kommt (57, 61).

Zur Mitte der Basisuntersuchung wurde ein Datensatz mit der ersten Hälfte der Studienteilnehmer zur Qualitätssicherung und Datenbereinigung zur Verfügung gestellt. Dieser Datensatz umfasste die Daten von gut 100.000 Studienteilnehmer und wurde dementsprechend als „Data Freeze 100k“ bezeichnet. Die Mindestanforderungen für die Datenbereinigung und Erstellung eines qualitätsgesicherten Datensatzes für wissenschaftliche Auswertungen wurden in einer Standardarbeitsanweisung (SOP) definiert (62).

1.4 Ziele der Arbeit

Im Rahmen der vorliegenden Arbeit sollte die Qualität der US-Untersuchung des Abdominalfetts in der NAKO Gesundheitsstudie zur Mitte der Basisuntersuchung beurteilt und Verbesserungsvorschläge abgeleitet werden. Dafür wurde zunächst die SOP der US-Untersuchung des Abdominalfetts nach Kriterien zur Generierung hochwertiger US-Bilder überarbeitet. Des Weiteren wurden die Messwerte der ersten Hälfte der Studienteilnehmer der Basisuntersuchung hinsichtlich Vollständigkeit, Reliabilität und Objektivität geprüft. Ein besonderes Augenmerk lag auf der Einschätzung des

Einflusses der Untersucher, der SZ und der Zeit auf die Messwerte. Der Data Freeze 100k-Datensatz wurde darüber hinaus bereinigt, um ihn für wissenschaftliche Auswertungen zur Verfügung zu stellen. Bei einer Stichprobe der gespeicherten US-Bilder der Untersucher der NAKO Gesundheitsstudie wurden Messrichtigkeit und Bildqualität bewertet, wofür Qualitätskriterien entwickelt wurde. Anhand der Bewertung wurde für jeden Untersucher eine individualisierte Rückmeldung erstellt. Die Ergebnisse der Qualitätssicherung wurden schlussendlich in einem Qualitätsbericht für die NAKO Gesundheitsstudie zusammengefasst. Wie die NAKO Gesundheitsstudie stellt auch diese Arbeit ein Kooperationsprojekt dar und wurde in enger Zusammenarbeit mit Frau Anja Sedlmeier erstellt. Sie gliedert sich inhaltlich in drei große Bereiche: die Überarbeitung der SOP, die Datenaufbereitung und -analyse sowie die Bewertung der US-Bilder.

2 Material und Methoden

2.1 Überarbeitung der SOP „US Abdominalfett“

Die US-Untersuchung des Abdominalfetts war für alle Level-2-Teilnehmer der Basisuntersuchung der NAKO Gesundheitsstudie vorgesehen. Sie umfasste die Quantifizierung des SAT und VAT als Distanzmessung auf Höhe der Taille in der Medianlinie. Die Vorgehensweise wurde in einer Vorphase der NAKO Gesundheitsstudie validiert (48) und in der SOP „US Abdominalfett“ beschrieben (63). Für die Messung wurde in allen SZ das US-Gerät iE33 sowie die dazugehörige Software und der C5-1 Breitband-Convex-Schallkopf von Philips verwendet (64). Während der US-Untersuchung lag der Teilnehmer in Rückenlage auf der Untersuchungsfläche. Der Schallkopf wurde senkrecht zur Haut ohne Kompression und unter Verwendung von US-Gel positioniert. Die Messung des SAT erfolgte am Ende der nicht forcierten Ausatmung von der Hautoberfläche bis zur Linea alba. Das VAT wurde ebenfalls am Ende der nicht forcierten Ausatmung von der Linea alba bis zur Vorderkante der Wirbelsäule gemessen. Für beide Fettkompartimente wurden zwei Messungen durchgeführt. Die Messwerte wurden in der Ergebnissoftware dokumentiert sowie die dazugehörigen Bilder gespeichert.

Vor Beginn der Zweituntersuchung der NAKO Gesundheitsstudie sollten die SOPs zur Untersuchungsdurchführung von den Modulverantwortlichen überarbeitet werden. Dafür wurde die Methodik der US-Messung in der SOP „US Abdominalfett“ (63) zunächst mit der Vorgehensweise in aktuellen Übersichtsarbeiten verglichen (50). Zur Optimierung der Bildqualität der US-Bilder wurden Empfehlungen aus einem Sonographie-Lehrbuch übernommen (41). Darüber hinaus wurden erneut US-Bilder unter Berücksichtigung aller neu etablierten Vorgehensweisen aufgenommen und als Beispiel in die SOP eingefügt. Zusätzlich wurden anatomische Abbildungen aus einem Anatomieatlas (65, 66) ausgewählt und beschriftet, um die der Messung zugrundeliegenden Strukturen zu veranschaulichen.

2.2 Datenaufbereitung und -analyse des Data Freeze 100k-Datensatzes

2.2.1 Anforderungen und theoretischer Hintergrund

Im Rahmen des Data Freeze 100.000 sollten die Daten der ersten 100.000 Teilnehmer der NAKO Gesundheitsstudie gesichtet und bereinigt werden. Die Anforderungen an die Durchführung der Datenbereinigung und Erstellung eines qualitätsgesicherten Datensatzes für wissenschaftliche Auswertungen wurden in der SOP „Vorgaben für die Datenbereinigung im Rahmen des Data Freeze 100.000“ beschrieben (62). Die standardisierte Durchführung der Datenbereinigung sollte unter anderem folgende Inhalte umfassen:

1. Vollständigkeit der Untersuchungen
2. Umgang mit unplausiblen Werten
3. Fehlende Werte
4. Verteilung der Werte
5. SZ-, Untersucher- und Zeiteffekte
6. Interne Validierung

Als Gütekriterien zur Beurteilung der Qualität der Messmethode wurden Objektivität und Reliabilität der US-Messung des Abdominalfetts betrachtet. Die Objektivität gibt an, inwieweit Messergebnisse vom Untersucher abhängen. Insbesondere systematische Messfehler führen zu einer systematischen Abweichung des Messwerts vom wahren Wert, was als Bias bezeichnet wird. Sie können durch unzureichende Schulung des Untersuchers entstehen (67–69). Die Stichprobe der NAKO Gesundheitsstudie wurde als Cluster an 18 SZ gezogen. Bei Cluster- beziehungsweise Klumpenstichproben werden statt zufälliger Individuen Gruppen der Grundgesamtheit ausgewählt, was den Vorteil der räumlichen Nähe hat. Wenn die Variabilität zwischen den Clustern größer ist als innerhalb eines Clusters, führt das jedoch zu einem höheren Standardfehler bei der Schätzung von Variablen (70). Neben dem Einfluss der Untersucher auf die Messwerte wurde deshalb auch der SZ-Effekt analysiert. Darüber hinaus können äußere Einflüsse wie die Nahrungsaufnahme zu Schwankungen der Messwerte des VAT führen (71), weshalb die Uhrzeit als deren Surrogat untersucht wurde.

Zufällige Messfehler bewirken hingegen eine Streuung um den wahren Messwert. Sie werden durch fehlende Übung und Aufmerksamkeit von Seiten des Untersuchers provoziert und beeinträchtigen die Präzision der Messung. Durch Wiederholungsmessungen können sie kontrolliert werden. Die Reproduzierbarkeit von Ergebnissen bei wiederholten Messungen unter identischen Bedingungen wird als Reliabilität bezeichnet (67–69). Die Intraobserver-Reliabilität wurde im Rahmen der internen Validierung der Messwerte geprüft. Das dritte Gütekriterium von Messmethoden, die Validität, misst die Richtigkeit einer Messmethode (67). Dafür wird die Messmethode mit dem Goldstandard verglichen, der in diesem Fall die CT- oder MRT-Messung des SAT und VAT darstellt (32). Die MRT-Daten, die bei einer Subgruppe der Level-2-Studienteilnehmer in fünf SZ erhoben wurden, lagen für diese Arbeit nicht vor. Alle Analysen wurden mit der Programmiersprache R in der Version 3.6.1 erstellt (72). Als Rückmeldung zur Datenbereinigung und -kontrolle wurde ein Qualitätsbericht verfasst.

2.2.2 Vollständigkeit der US-Untersuchung des Abdominalfetts

Der Data Freeze 100k-Datensatz umfasste neben den Studienteilnehmern der Hauptphase auch Studienteilnehmer aus der Pilotphase B und der Interimsphase, die vor der Hauptphase untersucht wurden. Für die Level-2-Studienteilnehmer der Hauptphase sollten die Anteile der Studienteilnehmer ermittelt werden, für die das Modul

1. vollständig durchgeführt und dokumentiert,
2. als nicht durchgeführt oder unvollständig abgebrochen dokumentiert
3. oder gar nicht dokumentiert wurde.

Dies sollte jeweils pro SZ und insgesamt über alle Zentren erfolgen. Für Untersuchungen, die als nicht durchgeführt oder unvollständig abgebrochen dokumentiert waren, sollte die Verteilung der Abbruchgründe dargestellt werden. Ebenfalls sollten Teilnehmer aufgelistet werden, für die keine abgeschlossene Dokumentation vorlag (62).

Anhand der Variablen *Status der Untersuchung in der Erhebungssoftware* und *Durchführbarkeit der Untersuchung* wurden die Teilnehmer den oben genannten Kategorien zugeordnet (siehe Tabelle 1).

Tabelle 1: Kodierung der Variablen Status der Untersuchung in der Erhebungssoftware und Durchführbarkeit der Untersuchung

Variable	Kodierung	Bedeutung
Status der Untersuchung in der Erhebungssoftware	-1	Aktuell nicht durchführbar
	0	Keine Daten
	1	Erstmalig oder wieder geöffnet und bearbeitet
	10	Ordnungsgemäß und vollständig abgeschlossen, keine Warnungen
	11	Ordnungsgemäß und vollständig abgeschlossen, mit Warnungen
	20	Unvollständig abgeschlossen - Abbruch durch Teilnehmer
	21	Unvollständig abgeschlossen - Abbruch durch Untersucher, technische Gründe
	22	Unvollständig abgeschlossen - Abbruch durch Untersucher, organisatorische Gründe
	23	Unvollständig abgeschlossen - Grund nicht mehr ermittelbar
	30	Durchführung nicht vorgesehen wegen fehlender Einwilligung
	31	Durchführung nicht vorgesehen aus Designgründen
	32	Durchführung nicht möglich wegen fehlender Voraussetzungen
Durchführbarkeit der Untersuchung	1	Durchführbar
	2	Nicht durchführbar

Als vollständig durchgeführt und dokumentiert (Kategorie 1, siehe Tabelle 2) wurden Teilnehmer angesehen, deren Status *ordnungsgemäß und vollständig abgeschlossen, keine Warnungen, ordnungsgemäß und vollständig abgeschlossen, mit Warnungen* oder *(erstmalig oder wieder) geöffnet und bearbeitet*, möglicherweise Daten in Arbeit und deren *Durchführbarkeit der Untersuchung* mit *ja* dokumentiert wurden. Wenn die Untersuchung als *nicht durchführbar* oder *unvollständig dokumentiert* oder die *Durchführung nicht vorgesehen* oder *nicht möglich* war oder die *Daten wieder geöffnet* wurden und die *Durchführbarkeit der Untersuchung* nicht verfügbar war, wurden die Teilnehmer der Kategorie 2 (als nicht durchgeführt oder unvollständig abgebrochen dokumentiert) zugeordnet. Nicht dokumentiert (Kategorie 3) waren all jene Teilnehmer, deren *Status der Untersuchung* nicht verfügbar war.

Tabelle 2: Kategorien der Vollständigkeit der Untersuchung mit logischer Berechnung.

Kategorie	Beschreibung	Logische Berechnung
1	Untersuchung vollständig durchgeführt und dokumentiert	$Status\ der\ Untersuchung\ in\ der\ Erhebungssoftware = 10, 11, 1\ UND\ Durchführbarkeit\ der\ Untersuchung = 1$
2	Untersuchung als nicht durchgeführt oder unvollständig abgeschlossen dokumentiert	$Status\ der\ Untersuchung\ in\ der\ Erhebungssoftware = 20, 21, 22, 23, 30, 31, 32\ ODER\ Durchführbarkeit\ der\ Untersuchung = 2\ ODER\ Status\ der\ Untersuchung\ in\ der\ Erhebungssoftware = 1\ UND\ Durchführbarkeit\ der\ Untersuchung = NA$
3	Untersuchung nicht dokumentiert	$Status\ der\ Untersuchung\ in\ der\ Erhebungssoftware = NA$

NA = nicht verfügbar.

Für die Ermittlung der Abbruchgründe der Teilnehmer in Kategorie 2 wurden die ergänzenden Variablen zur Durchführbarkeit der Untersuchung *Kontraindikation*, *technisches Problem*, *Verweigerung* und *Sonstiges* verwendet. Aus diesen Variablen konnte von den Untersuchern der Grund ausgewählt werden, weswegen die Untersuchung nicht durchführbar war. Bei *Sonstiges* konnte zusätzlich eine Freitexteingabe als Kommentar gespeichert werden. Diese Freitexteingaben wurden zunächst nach Angaben durchsucht, die die Zuordnung zu einer anderen Variablen erlaubten. Unter Verwendung von Suchbegriffen wurden die Freitexteingaben neu erstellten Kategorien zugeordnet (siehe Tabelle 3). Sofern in den Freitexteingaben Messwerte vorhanden waren, wurden diese manuell in die neu abgeleiteten Variablen der Messwerte übertragen. Außerdem wurde die Durchführbarkeit sowie die Gründe für Nichtdurchführbarkeit in neu abgeleiteten Variablen korrigiert. Wenn in der Freitexteingabe beispielsweise ein technisches Problem erwähnt wurde, wurde dies der Variable *technisches Problem* zugeordnet. Eine Schwangerschaft ab der 15. Schwangerschaftswoche wurde gemäß der SOP „US Abdominalfett“ (63) als Kontraindikation gewertet. Wenn laut Freitexteingabe der Teilnehmer die Untersuchung verweigert hatte, wurde dies als Zutreffen der Variable *Verweigerung* gewertet. Bei diesen Teilnehmern wurde die Auswahl der Kategorie *Sonstiges* revidiert.

Tabelle 3: Kategorien und Suchbegriffe zur Zuordnung der Freitexteingaben der Variable Nichtdurchführbarkeit – Sonstiges zu anderen Variablen.

Kodierung	Bedeutung	Suchbegriffe
1	SAT oder VAT	„sat“, „sub“, „vat“, „vis“
2	Technisches Problem	„ger“, „defekt“, „tech“
3	Verweigerung	„verw“
4	Schwangerschaft	„schwang“, „ssw“

SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe.

Daraufhin wurden die Freitexteingaben kategorisiert, um weitere Gründe für die Nichtdurchführbarkeit der Untersuchungen zu identifizieren (siehe Tabelle 4). All diejenigen Teilnehmer mit Freitexteingaben, die durch die Suchbegriffe keiner Kategorie zugeordnet werden konnten, wurden manuell zugeordnet. Sofern die Freitexteingabe keine der Kategorien 1 bis 5 erfüllte, wurde sie der Kategorie 6, *Sonstiges*, zugeordnet. Wenn in der Freitexteingabe angegeben wurde, dass die Untersuchung durchgeführt wurde, aber die gewünschten Strukturen im US nicht dargestellt werden konnten, wurde die Untersuchung als durchführbar gewertet.

Tabelle 4: Kategorien und Suchbegriffe zur Kategorisierung der Freitexteingaben der Variable Nichtdurchführbarkeit – Sonstiges.

Kodierung	Bedeutung	Suchbegriffe
1	<i>Zeit</i>	„zeit“
2	<i>Personal</i>	„pers“, „untersucher“, „assistent“, „studynurse“, „geschult“
3	<i>Raum</i>	„raum“
4	<i>Medizinische Gründe</i>	„bauch“
5	<i>Organisatorische Gründe</i>	„orga“, „durchg“, „nicht angeboten“
6	<i>Sonstiges</i>	

Nach der Auswertung der Freitexteingaben wurden die Teilnehmer gemäß der Vollständigkeit der Untersuchung den drei oben genannten Kategorien zugeordnet (vergleiche Tabelle 2). Dies wurde SZ-übergreifend und für jedes SZ einzeln durchgeführt. Diejenigen Teilnehmer, deren Untersuchung nicht durchführbar war, wurden gemäß den dokumentierten Gründen (*Kontraindikation, technisches Problem, Verweigerung, Sonstiges*) weiter untergliedert. Zusätzlich wurde die Kategorie *Sonstiges* gemäß der vorher durchgeführten Freitextanalyse in die Subkategorien *Zeit, Personal, Raum, medizinische Gründe, organisatorische Gründe* und *Sonstiges* aufgeteilt (vergleiche Tabelle 4).

2.2.3 Umgang mit unplausiblen Werten des SAT und VAT

Bei der Suche nach unplausiblen Werten wurden zunächst die offensichtlich unplausiblen Messwerte 999,9 und 0,000 korrigiert. 999,9 wurde bei der Untersuchung als Messwert eingetragen, wenn kein Messwert erhoben werden konnte (73). Für 999,9 und 0,000 wurden die duplizierten Variablen der Messwerte als nicht verfügbar (NA) festgelegt. Daraufhin wurden die Freitexteingaben in den Kommentaren zu den Messwerten ausgewertet und unplausible Werte korrigiert. Zusätzlich wurden die Bemerkungen nach Hinweisen auf zusätzliche Messwerte durchsucht. Bei diesen wurden die US-Bilder kontrolliert und gegebenenfalls Messwerte ausgetauscht. Die Freitextangaben in den Kommentaren zur Durchführbarkeit der Messung wurden bereits im Zuge der Analyse der Vollständigkeit der Untersuchung nach Messwerten durchsucht (vergleiche 2.2.2). Außerdem wurden diejenigen Teilnehmer überprüft, für die zwei identische Messwerte für das SAT oder VAT eingetragen wurden und eine Bemerkung vorhanden war. Sofern aus der Bemerkung hervorging, dass keine korrekte Messung möglich war, wurden die Messwerte der duplizierten Variablen als nicht verfügbar festgelegt.

Des Weiteren wurden die US-Bilder derjenigen Teilnehmer kontrolliert, deren Messwerte den weichen und harten Plausibilitätsgrenzen entsprachen, die zu Beginn der NAKO Gesundheitsstudie verwendet wurden (siehe Tabelle 5). Bei Messwerten jenseits der weichen Plausibilitätsgrenzen erschien eine Warnung in der Erhebungssoftware. Bei Unter- beziehungsweise Überschreitung der harten Plausibilitätsgrenzen konnten die Messwerte nicht gespeichert werden. Wenn der Messwert auf dem US-Bild ergab, dass statt des wahren Wertes die Plausibilitätsgrenze eingetragen wurde, wurde diese in den duplizierten Variablen durch den wahren Wert ersetzt.

Tabelle 5: Plausibilitätsgrenzen der Messwerte des SAT und VAT (63).

	Plausibilitätsgrenzen	
	Weich	Hart
SAT	0,7 cm – 3,7 cm	0,5 cm – 5,5 cm
VAT	3,0 cm – 13,0 cm	2,0 cm – 20,0 cm

SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe.

Darüber hinaus wurden Teilnehmer identifiziert, deren Differenz zwischen erstem und zweitem Messwert außerhalb des Mittelwertes der Differenz ± 4 Standardabweichungen lag (70). Der Mittelwert der Differenz wurde aus der Summe der Differenzen geteilt durch die Anzahl der Teilnehmer berechnet. Hier wurden die US-Bilder kontrolliert und falsche Messwerte aufgrund von Tippfehlern in den duplizierten Variablen korrigiert. Zuletzt wurden Extremwerte anhand ihrer Lage jenseits des Mittelwertes ± 4 Standardabweichungen bestimmt. Dafür wurden die Mittelwerte der Messwerte verwendet. Die US-Bilder zu den Extremwerten wurden kontrolliert und gegebenenfalls falsche Messwerte in den neu abgeleiteten Variablen korrigiert. Zusätzlich zur Korrektur der unplausiblen Messwerte wurde der Grund der Bearbeitung in einer eigenen Variablen dokumentiert (siehe Tabelle 6). Im Anschluss wurden die Mittelwerte der Messwertpaare neu berechnet.

Tabelle 6: Kodierung der Variable Grund für die Korrektur eines Messwerts.

Kodierung	Bedeutung
1	Messwert in <i>Kommentar zur Nichtdurchführbarkeit</i>
2	Messwert in <i>Bemerkung zur Messung</i>
3	Tippfehler (Kontrolle der US-Bilder)
4	Austausch mit 3. oder 4. Messwert aufgrund <i>Bemerkung zur Messung</i> (Kontrolle der US-Bilder)
5	Messwert gleich 999,9 oder 0,000
6	Kommentar in <i>Bemerkung zur Messung</i>
7	Messwert gleich der Plausibilitätsgrenzen (Kontrolle der US-Bilder)

US = Ultraschall.

2.2.4 Fehlende Werte des SAT und VAT

Für den Qualitätsbericht sollte eine Statistik fehlender Werte erstellt werden. Hierfür wurden lediglich diejenigen Studienteilnehmer berücksichtigt, bei denen die US-Untersuchung des Abdominalfetts dokumentiert wurde (vergleiche 2.2.2). Für jede Messung wurden die fehlenden Messwerte absolut und relativ berechnet. Dies erfolgte SZ-übergreifend und für jedes SZ separat.

2.2.5 Verteilung der Werte des SAT und VAT

Für alle Variablen sollte die Werteverteilung stratifiziert nach Alter in 10-Jahres-Gruppen und Geschlecht bestimmt und auf Plausibilität geprüft werden (62). Da die Messung des SAT und VAT als Doppelmessung durchgeführt wurde, wurden zunächst Mittelwerte aus den beiden Messwerten berechnet. Sofern nur ein Messwert vorlag, wurde dieser als Mittelwert übernommen. Die Werteverteilung wurde nach Geschlecht und Altersgruppen ermittelt und als Perzentilkurven graphisch dargestellt. Die Ergebnisse der Werteverteilung wurden bereits im Februar 2020 im Bundesgesundheitsblatt publiziert und werden im Ergebnisteil zitiert (74).

2.2.6 SZ-, Untersucher- und Zeiteffekte auf die Werte des SAT und VAT

In Ergänzung zur Verteilung der Werte (vergleiche 2.2.5) sollte nach den Vorgaben zur Datenbereinigung geprüft werden, ob es SZ-, Untersucher- und Zeiteffekte gab (62). Zum Vergleich der Messwerte der einzelnen SZ wurden nach Alter und Geschlecht adjustierte Mittelwerte für jedes SZ berechnet. Die Mittelwerte der SZ wurden zusätzlich mithilfe einer Analyse der Mittelwerte (*ANOM*) verglichen. Diese statistische Methode der Qualitätssicherung diente zum simultanen Vergleich der Mittelwerte der einzelnen SZ mit dem globalen Mittelwert aller SZ. Damit unterscheidet sie sich von einer Analyse der Varianz (*ANOVA*), bei der die Mittelwerte verschiedener Gruppen untereinander verglichen werden (75). Die Mittelwerte der SZ wurden mitsamt der Entscheidungsgrenzen in einem *ANOM*-Plot dargestellt. Dabei wurden die Entscheidungsgrenzen nach Anzahl der Studienteilnehmer pro SZ und den Kovariablen Alter und Geschlecht adjustiert. Wenn der Mittelwert eines SZ außerhalb der Entscheidungsgrenzen lag, entsprach das einer signifikanten Abweichung des Mittelwerts des SZ vom globalen Mittelwert ($p < 0,05$). Da die *ANOM* homogene Varianzen voraussetzte, wurden die Varianzen der Messwerte mit einem Levene-Test verglichen (76). Die *ANOM* wurde dann unter Annahme von Varianzinhomogenität durchgeführt (75).

Für die Ermittlung der Untersuchereffekte wurden alle Untersucher mit mindestens 50 durchgeführten Untersuchungen im Zeitraum des Data Freeze 100k betrachtet. Die Messmittelwerte der Untersucher wurden mithilfe einer *ANOM* unter Annahme von Varianzinhomogenität verglichen, wobei die Entscheidungsgrenzen nach Alter und Geschlecht der Studienteilnehmer sowie Anzahl der Studienteilnehmer pro Untersucher adjustiert wurden. Dadurch konnten Untersucher identifiziert werden, deren Messmittelwerte signifikant vom globalen Mittelwert abwichen.

Zudem wurde der Effekt des Studienmonats, des Monats und der Uhrzeit der US-Messung auf die Varianz der Messwerte betrachtet. Als Studienmonate wurden dabei die einzelnen Monate im Laufe der Hauptuntersuchung der NAKO Gesundheitsstudie bis zum Data Freeze 100k begriffen, während Monate den zwölf Monaten im Jahr entsprachen. Die nach Alter, Geschlecht und SZ adjustierten Mittelwerte wurden nach Studienmonat, Monat und Uhrzeit in Liniendiagrammen dargestellt, um Schwankungen analysieren zu können. Zusätzlich wurde mit Hilfe einer *ANOM* unter Annahme von Varianzinhomogenität die Signifikanz der Abweichungen vom globalen Messmittelwert kontrolliert, wobei die Entscheidungsgrenzen nach Alter und Geschlecht der Studienteilnehmer, SZ sowie Anzahl der Studienteilnehmer pro Monat, Studienmonat beziehungsweise Uhrzeit adjustiert wurden.

2.2.7 Interne Validierung der Doppelmessung des SAT und VAT

Da die Messwerte des SAT und VAT in Doppelmessungen unter identischen Bedingungen erhoben wurden, bestand eine Abhängigkeit zwischen erstem und zweitem Messwert. Für die interne Validierung wurden ebenfalls nur Messwerte derjenigen Untersucher betrachtet, die mindestens 50 Untersuchungen im Zeitraum des Data Freeze 100k durchgeführt hatten. Zunächst wurde die durchschnittliche Abweichung zwischen erstem und zweitem Messwert des SAT und VAT berechnet. Die SOP „US Abdominalfett“ (63) sah vor, dass die Abweichung zwischen erstem und zweitem Messwert $\leq 0,5$ cm betragen sollte. Dementsprechend wurde ermittelt, inwieweit diese Anforderung erfüllt wurde.

Um das Ausmaß sowie Muster der Diskrepanz der Doppelmessung darzustellen, wurden Bland-Altman-Plots erstellt. Sie dienen der Veranschaulichung von Abweichungen zwischen wiederholten Messungen oder zwei Messmethoden. Dafür wurden in einem Scatterplot die Mittelwerte der Messwertpaare auf der Abszisse gegen die Differenz der gepaarten Messwerte auf der Ordinate aufgetragen. Zusätzlich

wurden der Mittelwert der Differenz über alle Messwertpaare sowie zwei Standardabweichungen in beide Richtungen als horizontale Linien eingetragen. Anhand der Lage der berechneten Werte in Bezug zu diesen Linien wurden Anzahl und Anteil der Ausreißer quantifiziert. Wenn der Anteil an Ausreißern höchstens 5% betrug, wurde das als starke Übereinstimmung der Messwerte gewertet. Bei einem Anteil an Ausreißern zwischen 5% und 10% wurde die Übereinstimmung der Messwerte als gut gewertet. Darüber hinaus wurde beurteilt, ob ein Trend beobachtet werden konnte. Von einem Trend wird beispielsweise gesprochen, wenn das Ausmaß der Diskrepanz mit Zunahme der Messwertgröße steigt (69, 77).

Die Reproduzierbarkeit der Messwerte bei doppelter Messung durch einen Untersucher, auch Intraobserver-Reliabilität genannt, wurde mithilfe des Intraklassenkorrelationskoeffizienten (*ICC*) erfasst. Dieser erfasst sowohl die Korrelation von erstem und zweitem Messwert als auch systematische Unterschiede zwischen beiden Messwerten. Er misst den Anteil der Messwertvariabilität, der durch interindividuelle Varianz verursacht wurde. Die Gesamtvarianz setzt sich dabei aus der interindividuellen Varianz und der Varianz aufgrund der Untersucher durch zufällige sowie systematische Unterschiede zwischen wiederholten Messungen zusammen (69):

$$ICC = \frac{\text{variance between individuals}}{\text{variance between individuals} + \text{error variance}}$$

ICC-Werte größer 0,9 wurden als exzellente Korrelation gewertet. Der *ICC* wurde mithilfe einer *ANOVA* berechnet, wobei ein *two-way mixed effects model* angewandt wurde, das die absolute Übereinstimmung der Messwerte (*absolute agreement definition*) voraussetzt und auf eine einzelne Messung (*single rater/measurement type*) Bezug nimmt (78).

2.3 Bewertung der US-Bilder

Für die Einschätzung der Qualität der US-Messung des SAT und VAT wurde eine Stichprobe der US-Bilder anhand von bestimmten Qualitätskriterien bewertet. Daraus abgeleitet wurde für jeden Untersucher eine personalisierte Rückmeldung mit Verbesserungsvorschlägen erstellt.

2.3.1 Stichprobe

In die Bewertung der US-Bilder wurden ebenfalls lediglich die Untersucher eingeschlossen, die mindestens 50 Untersuchungen im Zeitraum des Data Freeze 100k durchgeführt hatten. Pro Untersucher wurden mindestens zehn und maximal 30 Studienteilnehmer à vier US-Bilder bewertet. Dafür wurden von Untersuchern mit weniger als 200 durchgeführten Untersuchungen pauschal 10 Studienteilnehmer per Zufallsstichprobe bestimmt. Bei Untersuchern mit 200 bis 599 durchgeführten Untersuchungen wurden 5% der Studienteilnehmer bewertet, was 10 bis 30 Studienteilnehmer entsprach. Von Untersuchern mit mindestens 600 durchgeführten Untersuchungen wurden pauschal 30 Studienteilnehmer ausgewählt. Aus den Studienteilnehmer mit vollständigen Messwerten wurde jeweils eine reproduzierbare Zufallsstichprobe der entsprechenden Größe für die Bewertung der US-Bilder gezogen. Auch die Reihenfolge für die Bewertung der Untersucher wurde mittels Zufallsstichprobe bestimmt. Die US-Bilder wurden über den zentralen Server der NAKO am Deutschen Krebsforschungszentrum im Programm Xcelera Express von Philips aufgerufen.

2.3.2 Bewertung mittels Qualitätskriterien

Für die Bewertung der US-Bilder wurden Qualitätskriterien der Messrichtigkeit und Bildqualität entwickelt (siehe Tabelle 7), wobei die Kriterien der Messrichtigkeit aus der SOP „US Abdominalfett“ abgeleitet wurden (63). Diese umfassten die korrekte Positionierung der Messpunkte sowie die senkrechte Ausrichtung der Messlinie. Als Vorlage für die Kriterien der Bildqualität dienten aus der Literatur abgeleitete Qualitätsmerkmale eines US-Bildes. Diese besagten, dass für ein gutes US-Bild die Geräteeinstellungen optimal sein sollten. Insbesondere eine schlechte Einstellung von Verstärkung und Tiefenausgleich könne die Bildqualität beeinträchtigen und dazu führen, dass Schallhindernisse übersehen werden. Verstärkung und Tiefenausgleich dienten dazu, die empfangenen Signale zu verstärken und das Gewebe schallkopfnah und –fern in gleicher Helligkeit abzubilden (41, 79). Darüber hinaus sollten die Leitstrukturen so dargestellt werden, dass der erfahrene Betrachter auch ohne Angabe der Schallkopfposition die Lage des Schnitts nachvollziehen kann. Der Fokus sollte auf Höhe der Zielstrukturen eingestellt sein, um eine hohe Auflösung zu gewährleisten. Für eine formatfüllende Aufnahme sollte das Bild so vergrößert werden, dass die Zielstruktur einige Zentimeter vom unteren Bildschirmrand entfernt zur Darstellung kommt (41).

Tabelle 7: Kriterien für die Bewertung der Messrichtigkeit und Bildqualität der US-Bilder.

Messrichtigkeit	Messlinie senkrecht	Messlinie senkrecht nach unten gelotet ($\pm 5^\circ$ Toleranzbereich)
	Messpunkte richtig gesetzt	SAT: Bauchhaut und Oberrand der Linea alba VAT: Unterrand der Linea alba und Vorderkante des Wirbelkörpers
Bildqualität	Mittlere Helligkeit	Mittlere Gesamthelligkeit mit gleicher Helligkeit an der Oberfläche und in der Tiefe
	Formatfüllende Aufnahme	SAT: Linea alba in den unteren 2/3 des Bildes oder maximal gezoomt VAT: Vorderkante des Wirbelkörpers in der unteren Bildhälfte
	Linea alba mittig	Linea alba in Bezug auf rechts und links mittig im Bild

US = Ultraschall, SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe.

Bei der Bewertung der US-Bilder wurde zunächst kontrolliert, ob die Messwerte von den US-Bildern korrekt in die Erhebungssoftware übertragen wurden. Daraufhin wurde die Bildqualität bewertet, wobei für jedes Kriterium und jedes Bild jeweils ein Punkt vergeben wurde. Die Gesamthelligkeit des US-Bildes sollte im mittleren Bereich mit gleicher Helligkeit an der Oberfläche und in der Tiefe liegen. Für grobe Fehler wie ein sehr dunkles US-Bild wurden 0 Punkte erteilt. Die Zielstrukturen sollten unter Verwendung der Zoomfunktion des US-Geräts formatfüllend aufgenommen worden sein, wobei die Eindringtiefe des Geräts minimal 5 cm und maximal 30 cm betrug. Bei der Messung des SAT wurden 0 Punkte vergeben, wenn die Linea alba im oberen Bilddrittel lag. Wenn bei der Messung des VAT die Vorderkante des Wirbelkörpers in der oberen Bildhälfte lag, wurde ebenfalls 0 Punkte erteilt. Zur Objektivierung wurden Schablonen verwendet, womit die Lage der Zielorgane bezogen auf das US-Bild einfach und sicher kontrolliert werden konnte (siehe Anhang). Mit Augenmaß wurde geprüft, ob sich die Linea alba mittig im Bild befand. Im Anschluss fand die Bewertung der Messrichtigkeit statt. Mithilfe einer Schablone wurde kontrolliert, ob die Messlinie senkrecht verlief. Bei einer Abweichung von mehr als 5° wurden 0 Punkte erteilt. Für die Positionierung jedes Messpunktes wurde ein Punkt vergeben und gegebenenfalls Abweichungen dokumentiert. Der obere Messpunkt des SAT sollte direkt an der Bauchhaut liegen, der untere Messpunkt am oberen Rand der Linea alba. Das VAT sollte vom unteren Rand der Linea alba zur Vorderkante des Wirbelkörpers gemessen worden sein. Wenn die Lage des Messpunkts nicht beurteilbar war, wurde kein Punkt vergeben. Um standardisierte Bewertungsbedingungen zu gewährleisten, wurden die US-Bilder immer an demselben Bildschirm unter Abdunkelung des Raumes und von derselben Person bewertet.

2.3.3 Analyse der US-Bildbewertung

Für jedes Kriterium der Messrichtigkeit und Bildqualität wurden die Anteile korrekt durchgeführter US-Untersuchungen in Abhängigkeit von Alter und Geschlecht der Studienteilnehmer, der gemessenen Menge des SAT beziehungsweise VAT und dem SZ berechnet. Zudem wurde analysiert, welche Fehler in der Durchführung der einzelnen Kriterien auftraten. In einem logistischen Regressionsmodell wurden die oben genannten prädiktiven Variablen auf Signifikanz getestet. Darüber hinaus wurde berechnet, welcher Anteil der Messungen des SAT und VAT nach Maßstäben der Messrichtigkeit beziehungsweise Bildqualität komplett korrekt durchgeführt wurde. Der Anteil der hinsichtlich der Messrichtigkeit komplett korrekt durchgeführten US-Bilder jedes Untersuchers wurde dem bereinigten Data Freeze 100k-Datensatz hinzugefügt, um die Einschätzung der Datenqualität bei der Verwendung der Daten im Rahmen von zukünftigen wissenschaftlichen Analysen zu ermöglichen.

2.3.4 Rückmeldung an die Untersucher

Aus der Bewertung der US-Bilder wurde für jeden Untersucher eine Rückmeldung samt Verbesserungsvorschlägen verfasst. Dafür wurde zunächst eine Vorlage erstellt, die dann für jeden Untersucher individuell ergänzt wurde. Die Bewertung der Kriterien der Messrichtigkeit und Bildqualität wurde in Schulnoten samt Farbung dargestellt und im Fließtext standardisiert formuliert (vergleiche Tabelle 8). Verbesserungsvorschläge wurden immer dann beigefügt, wenn ein Fehler in mindestens 10% der bewerteten US-Bilder aufgetreten war. Wenn ein Kriterium der Messrichtigkeit in weniger als 75% der bewerteten US-Bilder korrekt durchgeführt worden war, wurde dem jeweiligen Untersucher eine Nachschulung empfohlen. Zusätzlich wurden neue Aspekte der überarbeiteten SOP (73) wie die Verwendung des Fokus erwähnt, um die Untersucher zur weiteren Verbesserung der Messqualität anzuregen. Zur Veranschaulichung wurden Beispielbilder erstellt.

Tabelle 8: Standardisierte Formulierung bei der Rückmeldung an die Untersucher zur Bewertung der US-Bilder

Korrektur Anteil	Schulnote	Bewertung	Angabe der Häufigkeit
≥ 90%	1	Sehr gut	Immer
≥ 75% bis < 90%	2	Gut	Meistens
≥ 60% bis < 75%	3	Befriedigend	Oft
≥ 50% bis < 60%	4	Ausreichend	Gelegentlich
< 50%	5	Mangelhaft	Selten

US = Ultraschall.

2.4 Qualitätsbericht und Empfehlungen zur US-Messung

Die Ergebnisse der Datenaufbereitung und –analyse wurden im „Qualitätsbericht zur Datenbereinigung im Rahmen des Data Freeze 100.000“ des Moduls US Abdominalfett zusammengefasst. Beigefügt wurden die R-Datensätze mit den bereinigten und abgeleiteten Variablen sowie ein R-Markdown-Skript zur Datenbereinigung. In einem R-Markdown-Skript können Fließtext und Code mit den Ergebnissen der statistischen Analysen kombiniert werden, was die Anschaulichkeit und Reproduzierbarkeit verbessert (80). Als Ergänzung zur aktuellen SOP „US Abdominalfett“ (73) wurden Empfehlungen zur US-Messung mit den oben genannten Kriterien der Messrichtigkeit und Bildqualität sowie Beispielbildern aus der Rückmeldung an die Untersucher formuliert (siehe Anhang).

3 Ergebnisse

3.1 Überarbeitete SOP „US Abdominalfett“

Für die Überarbeitung der SOP „US Abdominalfett“ (63) wurde sie mit aktuellen Übersichtsarbeiten zur Methodik der US-Messung des Abdominalfetts verglichen. Die SOP erfüllte bei der Messung des SAT nahezu alle Empfehlungen von Azzi et al. (50): So wurde das SAT als Strecke und nicht als Fläche oder Volumen gemessen. Für die Messung wurde der Studienteilnehmer auf dem Rücken gelagert. Die Messposition befand sich auf Höhe der Taille und dadurch wenige Zentimeter kranial des Bauchnabels in der Medianlinie. In der SOP wurde darauf hingewiesen, dass bei der Messung möglichst wenig Druck angewendet und die US-Sonde im 90°-Winkel zur Haut gehalten werden sollte. Im Gegensatz zu anderen Studien wurde jedoch in der NAKO Gesundheitsstudie bei der Messung des SAT die Bauchhaut miteinbezogen. Bei der Messung des VAT orientierte sich die SOP an der Methodik von Schlecht et al. (48) und beschrieb eine Messung von der Linea alba zum Lendenwirbelkörper, während in anderen Studien vom Peritoneum oder der inneren Faszie der Bauchmuskeln zur Vorder- oder Hinterwand der Aorta gemessen wurde (42, 45, 49, 81, 82). Auch die Messposition unterschied sich geringfügig zwischen den Studien, wobei meist etwas kranial des Bauchnabels in der Medianlinie gemessen wurde. Viele Studien sahen ebenso wie die SOP eine Messung am Ende der Ausatmung und mit möglichst geringem Druck der US-Sonde vor.

Zur Optimierung der Bildqualität der US-Bilder wurden aus einem Sonographie-Lehrbuch (41) folgende Handlungsanweisungen abgeleitet (vergleiche Tabelle 9): Für eine hohe Auflösung des US-Bildes sollte die Eindringtiefe der US-Sonde so reguliert werden, dass die interessierende Struktur einige Zentimeter vom unteren Bildschirmrand zur Darstellung kommt. Bei der Messung des SAT sollte dafür die minimale Eindringtiefe eingestellt werden, die das SAT formatfüllend darstellt. Bei der Messung des VAT sollte der Wirbelkörper im unteren Bereich des US-Bildes abgebildet werden, um auch hier eine formatfüllende Aufnahme zu erreichen. Zusätzlich sollte der Fokus auf die Höhe der interessierenden Struktur eingestellt werden. Bei der Messung des SAT sollte die Linea alba fokussiert werden, bei der Messung des VAT die Wirbelkörpervorderkante. Die Bildqualität kann durch iScan automatisch verbessert werden (64). Mittels Gain kann die Helligkeit des US-Bildes manuell angepasst

werden. Nach Einfrieren des Bildes kann dieses unter Verwendung des Trackballs um wenige Millisekunden vorgespult werden, um eine bestmögliche Momentaufnahme für die Messung zu erhalten. Um die SOP nicht zu kompliziert zu gestalten, fanden weder die tiefenabhängige noch laterale Verstärkung Erwähnung. Auch auf die Einführung eines Linearschallkopfes für die Messung des SAT wurde verzichtet, da auch unter Verwendung der minimalen Eindringtiefe des Konvexschallkopfes von 5 cm eine gute Darstellung des SAT möglich war. Von einer speziellen Vorbereitung der Teilnehmer mit Verzicht auf größere Mahlzeiten und kohlenensäurehaltige Getränke wurde aus organisatorischen Gründen abgesehen.

Tabelle 9: Vergleich der Kriterien für ein qualitativ hochwertiges US-Bild gemäß einem Sonographie-Lehrbuch mit der SOP „US Abdominalfett“ (41, 63, 64).

Kriterium	Durchführung	Erfüllt	Teilweise/nicht erfüllt
Vorbereitung	Keine großen Mahlzeiten oder kohlenensäurehaltigen Getränke vor der Abdomensonographie		X
Position	Rückenlage des Probanden	X	
Atemkommando	Verlagerung des Querkolons nach kranial während der Ausatmung	X	
Schallkopf	Abdomensonographie: Konvexschallkopf, Darstellung der Bauchwand: Linearschallkopf		X
US-Gel	Verwendung für gute Schalleitung	X	
Schnittebene	Darstellung der rechten Anteile des Untersuchungsgebiets links im US-Bild	X	
Eindringtiefe	Formatfüllende Darstellung der Zielstrukturen		X
Fokus/Fokuszone	Einstellung auf Höhe der Zielstruktur für hohe Auflösung		X
Gain	Einstellung der Bildhelligkeit	X	
TGC/LGC	Ausgleich der tiefenabhängigen und lateralen Abschwächung des US		X
iScan	Automatischen Optimierung von Gain und TGC	X	
Freeze/Trackball	Vorspulen zur bestmöglichen Momentaufnahme		X

SOP = Standardarbeitsanweisung, US = Ultraschall, Gain = Gesamtverstärkung, TGC = tiefenabhängige Verstärkung, LGC = laterale Verstärkung, Freeze = Bild einfrieren, Trackball = Maus.

Die relevanten Messregler am US-Gerät wurden auf einer Photographie beschriftet und der SOP beigefügt (siehe Abbildung 2).

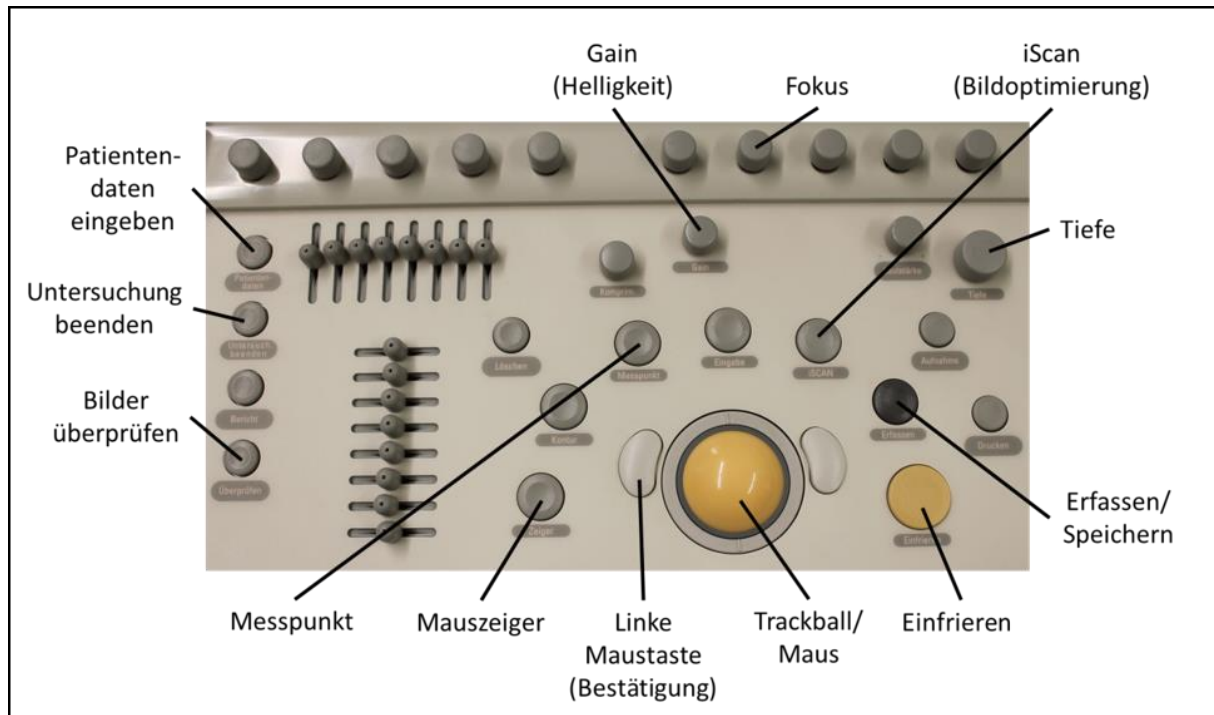


Abbildung 2: Beschriftung einer Photographie der Messregler des US-Geräts iE33 von Philips für die überarbeitete Version der SOP „US Abdominalfett“ (48, 64).
 US = Ultraschall, SOP = Standardarbeitsanweisung.

Zur Veranschaulichung der zugrundeliegenden anatomischen Strukturen wurden Bilder aus einem Anatomieatlas (65, 66) verwendet. Die für die US-Messung relevanten Strukturen wurden beschriftet und die gewünschten Messstrecken eingetragen (siehe Abbildung 3). Dabei sollte beachtet werden, dass das Bild aus dem Anatomieatlas einen Transversalschnitt durch das Abdomen auf Höhe des ersten Lendenwirbelkörpers zeigt. Diese Höhe entspricht nicht der Höhe der US-Messung im Bereich des zweiten bis dritten Lendenwirbelkörpers. Somit kam beispielsweise die Leber bei der US-Messung des VAT normalerweise nicht zur Darstellung.

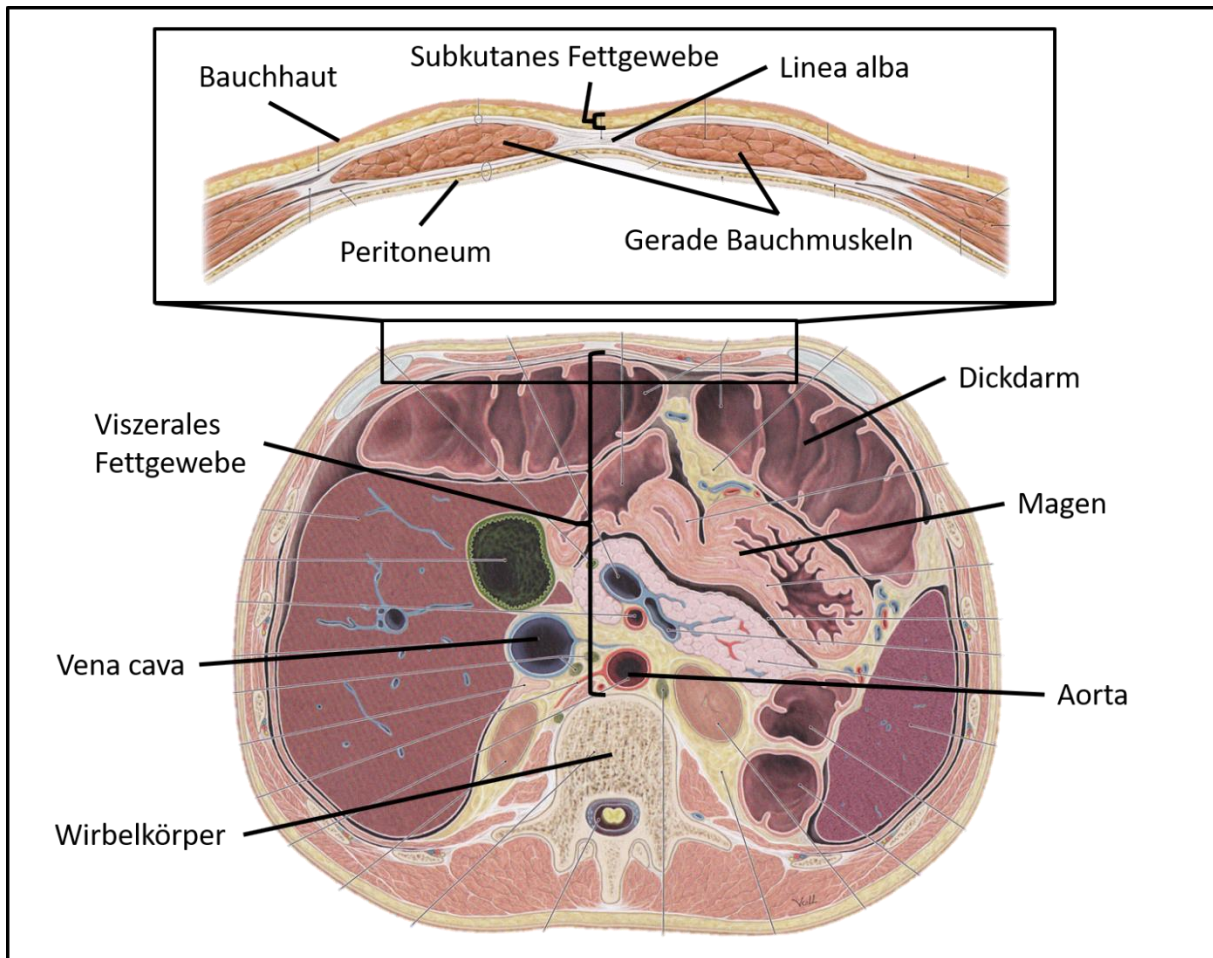


Abbildung 3: Anatomische Strukturen: Transversalschnitte durch die Rektusscheide kranial der Linea arcuata sowie durch das Abdomen auf Höhe des ersten Lendenwirbelkörpers, modifiziert nach Schünke, Schulte et al. (65, 66).

Zudem wurden US-Bilder unter Berücksichtigung aller neu etablierten Vorgehensweisen zur Bildoptimierung aufgenommen und der SOP als Beispielbilder beigefügt.

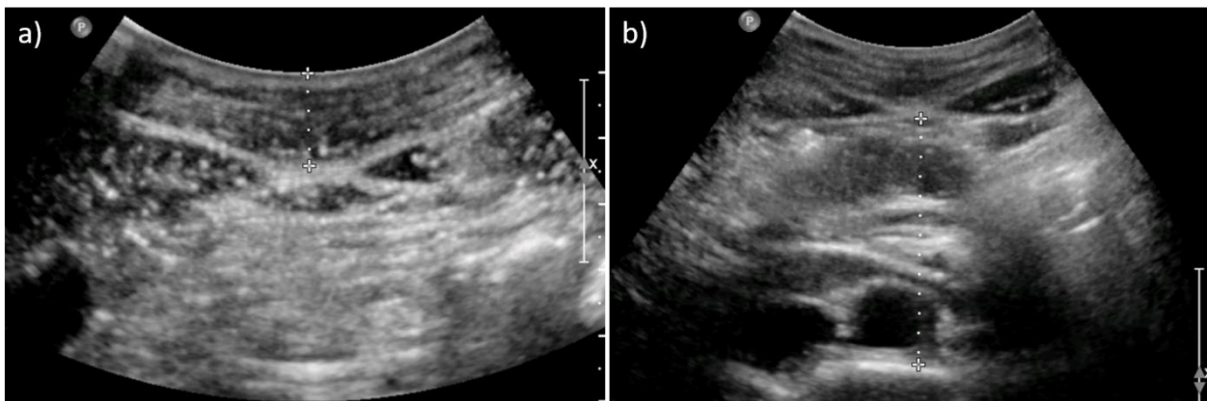


Abbildung 4: US-Bilder von a) SAT und b) VAT mitsamt der Messstrecken unter Optimierung der Eindringtiefe, Fokuseinstellung und Helligkeit.
 US = Ultraschall, SAT = subkutanes Fettgewebe, VAT = viszerale Fettgewebe.

3.2 Datenaufbereitung und –analyse des Data Freeze 100k-Datensatzes

3.2.1 Vollständigkeit der Untersuchung US Abdominalfett

Der Data Freeze 100k Datensatz umfasste 101.816 Studienteilnehmer, darunter 97.749 Studienteilnehmer der Hauptphase. Die restlichen 4.067 Studienteilnehmer wurden im Rahmen der Pilotphase B ($n = 1.012$) beziehungsweise der Interimsphase ($n = 3.055$) untersucht. 71,6% der Studienteilnehmer der Hauptphase erhielten eine Level-1-Untersuchung ($n = 69.978$), während bei 28,4% im Rahmen der Level-2-Untersuchung die US-Untersuchung des Abdominalfetts vorgesehen war ($n = 27.771$). Im Verhältnis zu Level-2-Teilnehmerzahl wurden die Anteile der Studienteilnehmer ermittelt, für die das Modul

1. vollständig durchgeführt und dokumentiert,
2. als nicht durchgeführt oder unvollständig abgebrochen dokumentiert
3. oder gar nicht dokumentiert wurde (vergleiche Abbildung 5).

In allen 18 SZ wurden insgesamt 20.390 US-Untersuchungen vollständig durchgeführt und dokumentiert, das entsprach 73,4% der vorgesehenen US-Untersuchungen (vergleiche Tabelle 10). Die meisten US-Untersuchungen wurden im SZ Augsburg durchgeführt ($n = 4.102$), die wenigsten im SZ Saarbrücken ($n = 375$). Im SZ Regensburg wurde mit 96,0% der größte Anteil der US-Untersuchungen vollständig dokumentiert und durchgeführt ($n = 667$), in Saarbrücken mit 21,8% der geringste ($n = 375$). Insgesamt wurden in allen SZ 23,8% der vorgesehenen US-Untersuchungen nicht dokumentiert und somit in der Erhebungssoftware nicht bearbeitet ($n = 6.596$). In allen SZ wurden 2,8% der US-Untersuchungen als nicht durchgeführt oder unvollständig abgebrochen dokumentiert ($n = 785$). Die wenigsten unvollständigen Untersuchungen wurden in Bremen (0,3%), die meisten in Braunschweig/Hannover (17,8%) dokumentiert.

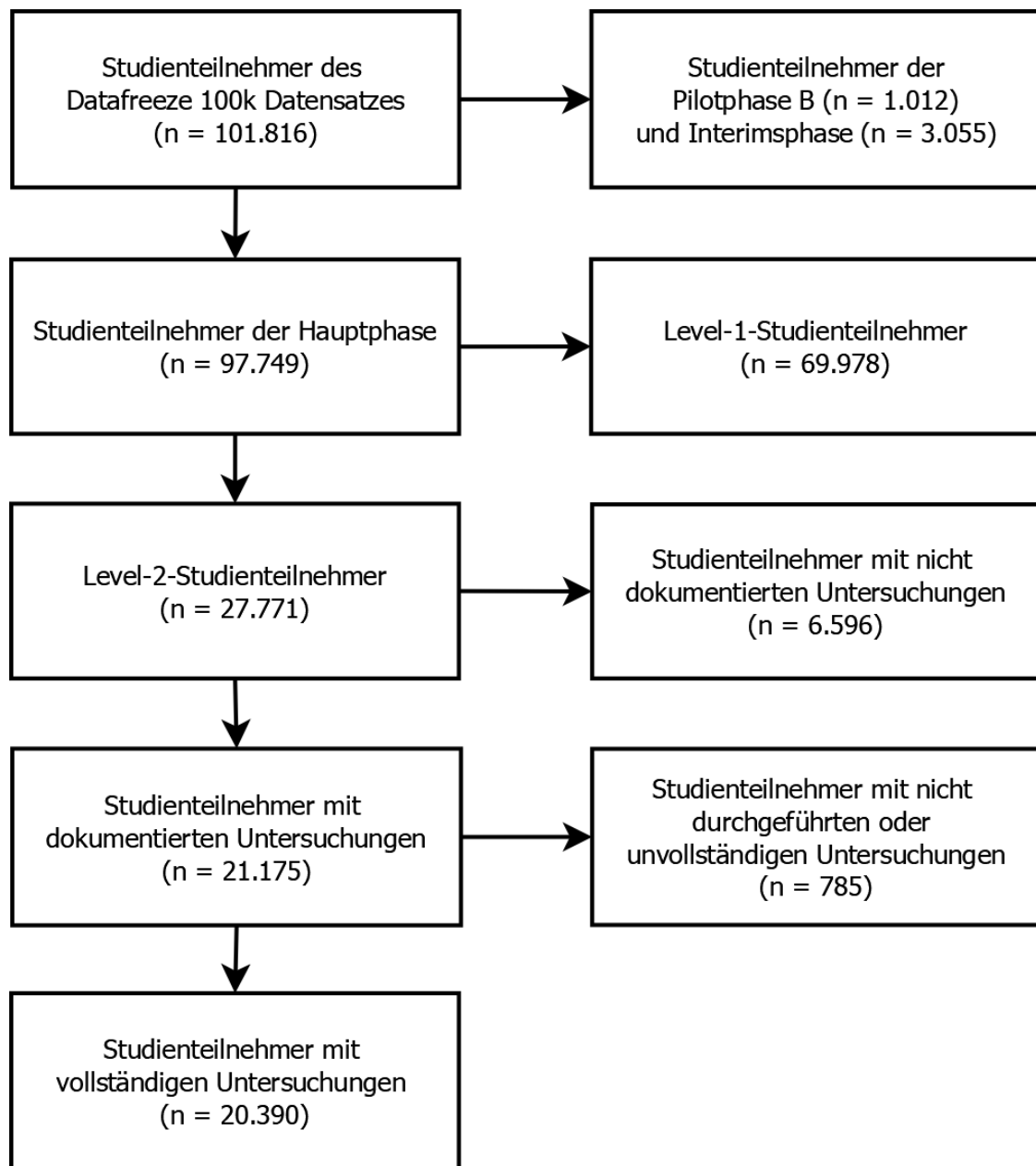


Abbildung 5: Flowchart zur Vollständigkeit der US-Untersuchungen.
 US = Ultraschall, n = Anzahl.

Tabelle 10: Vollständigkeit der US-Untersuchung des Abdominalfetts.

	Gesamt	Nicht dokumentiert		Dokumentiert						
		N	n	%	Gesamt		Vollständig		Nicht durchgeführt/unvollständig	
					n	%	n	%	n	%
Alle SZ	27771	6596	23,8	21175	76,2	20390	73,4	785	2,8	
Augsburg	4382	252	5,8	4130	94,2	4102	93,6	28	0,6	
Regensburg	695	7	1,0	688	99,0	667	96,0	21	3,0	
Heidelberg/Mannheim	2224	350	15,7	1874	84,3	1855	83,4	19	0,9	
Freiburg	956	42	4,4	914	95,6	910	95,2	4	0,4	
Saarbrücken	1717	1280	74,5	437	25,5	375	21,8	62	3,6	
Essen	2053	1523	74,2	530	25,8	495	24,1	35	1,7	
Münster	779	104	13,4	675	86,6	658	84,5	17	2,2	
Düsseldorf	697	273	39,2	424	60,8	415	59,5	9	1,3	
Halle	947	83	8,8	864	91,2	846	89,3	18	1,9	
Leipzig	1377	180	13,1	1197	86,9	1188	86,3	9	0,7	
Berlin Nord	941	32	3,4	909	96,6	830	88,2	79	8,4	
Berlin Mitte	1131	171	15,1	960	84,9	824	72,9	136	12,0	
Berlin Süd/Brandenburg	739	45	6,1	694	93,9	680	92,0	14	1,9	
Braunschweig/Hannover	1083	248	22,9	835	77,1	642	59,3	193	17,8	
Hamburg	1040	58	5,6	982	94,4	930	89,4	52	5,0	
Bremen	1183	73	6,2	1110	93,8	1106	93,5	4	0,3	
Kiel	931	349	37,5	582	62,5	569	61,1	13	1,4	
Neubrandenburg	4896	1526	31,2	3370	68,8	3298	67,4	72	1,5	

US = Ultraschall, n = Anzahl, SZ = Studienzentren.

Zusätzlich wurde ermittelt, weshalb US-Untersuchungen nicht durchgeführt oder unvollständig abgebrochen wurden (vergleiche Tabelle 11). Dafür wurden die dokumentierten Abbruchgründe „Kontraindikation“, „technisches Problem“, „Verweigerung“ und „Sonstiges“ sowie die Freitexteingaben in den Bemerkungen zur Durchführbarkeit der Messung ausgewertet. Am häufigsten wurde Personalmangel mit 30,3% ($n = 238$), gefolgt von organisatorischen Gründen mit 23,1% ($n = 181$) und technischen Problemen mit 13,5% ($n = 106$) als Grund angegeben. Dabei fanden sich Unterschiede zwischen den einzelnen SZ. Beispielsweise wurde im SZ Berlin Mitte Personalmangel in 83,8% der Fälle ($n = 114$) als Grund für nicht durchgeführte oder unvollständig abgeschlossene Untersuchungen angegeben. Im SZ Berlin Nord ($n = 63$) waren in 79,7% der Fälle organisatorische Gründe dafür verantwortlich, dass US-Untersuchungen nicht oder unvollständig durchgeführt wurden.

Tabelle 11: Verteilung der Gründe für nicht durchgeführte US-Untersuchungen, in 16 Fällen Mehrfachangabe von Gründen.

	Dok.	Nicht durchgeführt /unvollständig									
		Ges.	Pers.	Org.	Tech.	Zeit.	Verw.	Med.	Kon.	Raum.	Sonst.
	N	n	n	n	n	n	n	n	n	n	n
Alle SZ	21175	785	238	181	106	73	67	19	18	5	94
Augsburg	4130	28	0	0	0	0	3	2	6	0	20
Regensburg	688	21	1	8	0	2	4	3	3	1	1
Heidelberg/ Mannheim	1874	19	0	3	11	2	0	0	0	0	3
Freiburg	914	4	0	0	1	0	2	0	0	0	1
Saarbrücken	437	62	31	9	20	0	1	0	0	0	2
Essen	530	35	0	20	2	0	2	0	0	0	11
Münster	675	17	2	0	0	0	2	1	0	0	12
Düsseldorf	424	9	0	1	0	0	0	2	0	0	4
Halle	864	18	10	1	0	0	3	0	2	0	3
Leipzig	1197	9	1	1	2	1	2	1	0	0	2
Berlin Nord	909	79	0	63	0	19	0	0	0	0	0
Berlin Mitte	960	136	114	5	2	4	7	1	3	0	5
Berlin Süd/ Brandenburg	694	14	8	2	1	3	0	0	0	0	0
Braunschweig/ Hannover	835	193	35	55	16	23	35	4	3	4	21
Hamburg	982	52	0	2	25	18	2	0	0	0	5
Bremen	1110	4	0	0	3	0	0	0	1	0	0
Kiel	582	13	0	4	2	0	1	3	2	0	2
Neubrandenburg	3370	72	36	7	21	1	3	2	0	0	4

US = Ultraschall, Dok. = dokumentiert, Ges. = gesamt, Pers. = Personalmangel, Org. = organisatorische Gründe, Tech. = technisches Problem, Zeit. = Zeitmangel, Verw. = Verweigerung, Med. = medizinische Gründe, Kon. = Kontraindikation, Raum. = Raummangel, Sonst. = Sonstiges, n = Anzahl, SZ = Studienzentren.

3.2.2 Umgang mit unplausiblen Werten des SAT und VAT

Sowohl SAT als auch VAT wurden in Doppelmessungen bestimmt. Bei 21.175 dokumentierten Untersuchungen (vergleiche Tabelle 10) sollten jeweils 42.350 Messwerte des SAT und 42.350 Messwerte des VAT und somit insgesamt 84.700 Messwerte vorliegen. Auf vielfältige Weise wurden unplausible Werte identifiziert und insgesamt 0,9% der Messwerte korrigiert ($n = 728$). Dabei wurden 0,4% der Messwerte des SAT ($n = 160$) und 1,3% der Messwerte des VAT ($n = 568$) geändert. Der zweite VAT-Messwert (VAT2) war etwas häufiger unplausibel als der erste VAT-Messwert (VAT1). Bei der Messung des subkutanen Fettgewebes wurden erster und zweiter Messwert (SAT1 und SAT2) etwa gleich häufig korrigiert. In 0,5% der Fälle ($n = 408$) traten die offensichtlich unplausiblen Messwerte 0,000 und 999,9 auf. Aus Freitextangaben in der Bemerkung zur Messung und zur Durchführbarkeit der Messung wurden insgesamt 168 Messwerte übernommen und damit 0,2% der Messwerte korrigiert. In 0,1% der Fälle fand sich in der Bemerkung zur Messung der Kommentar, dass der angegebene Messwert unplausibel war ($n = 81$). Bei einer Differenz über 1 cm zwischen erstem und zweitem Messwert, bei Messwerten gleich den Plausibilitätsgrenzen in der Erhebungssoftware (vergleiche Tabelle 5) und bei Extremwerten

jenseits der vierten Standardabweichung wurden die US-Bilder zu den Messungen gesichtet. Dabei wurden insgesamt 0,1% der Messwerte aufgrund von Tippfehlern ($n = 48$) und zehn Messwerte gleich der Plausibilitätsgrenzen korrigiert. Wenn aus der Bemerkung zur Messung hervorging, dass mehr als zwei Messwerte erhoben wurden, wurden auch diese US-Bilder kontrolliert und insgesamt 13 Messwerte ausgetauscht. Der Grund für die Korrektur wurde in einer zusätzlichen Variablen hinterlegt (vergleiche Tabelle 6).

Tabelle 12: Korrekturanlässe unplausibler Messwerte.

	Gesamt	SAT1	SAT2	VAT1	VAT2
	N	n	n	n	n
Alle Korrekturanlässe	728	82	78	261	307
Messwert gleich 999,9 oder 0,000	408	11	12	177	208
Messwert in <i>Bemerkung zur Messung</i>	120	37	36	22	25
Kommentar „Messwert unplausibel“ in <i>Bemerkung zur Messung</i>	81	3	3	38	37
Messwert in <i>Bemerkung zur Durchführbarkeit</i>	48	17	17	7	7
Tippfehler	48	9	6	9	24
Austausch mit drittem oder viertem Messwert	13	1	1	7	4
Messwert gleich Plausibilitätsgrenze	10	4	3	1	2

SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe, n = Anzahl.

3.2.3 Fehlende Werte des SAT und VAT

Für die Statistik fehlender Werte wurden die Studienteilnehmer mit dokumentierten Untersuchungen betrachtet ($n = 21.175$, vergleiche Tabelle 10). Insgesamt fehlten 4,5% der 84.700 erwarteten Messwerte ($n = 3.830$). Bei den Messwerten des SAT fehlten je 3,9% des ersten ($n = 823$) und zweiten Messwerts ($n = 830$). Mit 5,2% fehlte VAT2 am häufigsten ($n = 1.108$), gefolgt von VAT1 mit 5,0% ($n = 1.069$). Der Anteil fehlender Werte unterschied sich zwischen den SZ (vgl. Tabelle 13). Mit 25,1% fehlten anteilig am meisten Werte im SZ Braunschweig/Hannover ($n = 839$), gefolgt von den SZ Saarbrücken mit 15,0% ($n = 262$) und Berlin Mitte mit 14,6% ($n = 559$). Am wenigsten Messwerte fehlten anteilig im SZ Augsburg mit 0,7% ($n = 119$).

Tabelle 13: Fehlende Messwerte der US-Untersuchung des Abdominalfetts nach SZ.

	Dok.	Erw.	Fehlende Messwerte									
			Gesamt		SAT1		SAT2		VAT1		VAT2	
			N	N	n	%	n	%	n	%	n	%
Alle SZ	21175	84700	3830	4,5	823	3,9	830	3,9	1069	5,0	1108	5,2
Augsburg	4130	16520	119	0,7	28	0,7	29	0,7	31	0,8	31	0,8
Regensburg	688	2752	149	5,4	25	3,6	25	3,6	48	7,0	51	7,4
Heidelberg/ Mannheim	1874	7496	157	2,1	20	1,1	20	1,1	50	2,7	67	3,6
Freiburg	914	3656	45	1,2	6	0,7	7	0,8	15	1,6	17	1,9
Saarbrücken	437	1748	262	15,0	62	14,2	62	14,2	69	15,8	69	15,8
Essen	530	2120	166	7,8	38	7,2	38	7,2	45	8,5	45	8,5
Münster	675	2700	106	3,9	20	3,0	20	3,0	33	4,9	33	4,9
Düsseldorf	424	1696	58	3,4	10	2,4	10	2,4	18	4,2	20	4,7
Halle	864	3456	78	2,3	17	2,0	17	2,0	22	2,5	22	2,5
Leipzig	1197	4788	54	1,1	11	0,9	11	0,9	16	1,3	16	1,3
Berlin Nord	909	3636	348	9,6	80	8,8	80	8,8	93	10,2	95	10,5
Berlin Mitte	960	3840	559	14,6	138	14,4	138	14,4	141	14,7	142	14,8
Berlin Süd/ Brandenburg	694	2776	90	3,2	15	2,2	17	2,4	28	4,0	30	4,3
Braunschweig/ Hannover	835	3340	839	25,1	198	23,7	199	23,8	220	26,3	222	26,6
Hamburg	982	3928	209	5,3	51	5,2	50	5,1	53	5,4	55	5,6
Bremen	1110	4440	61	1,4	8	0,7	10	0,9	20	1,8	23	2,1
Kiel	582	2328	111	4,8	16	2,7	16	2,7	39	6,7	40	6,9
Neubrandenburg	3370	13480	419	3,1	80	2,4	81	2,4	128	3,8	130	3,9

US = Ultraschall, SZ = Studienzentrum, Dok. = dokumentierte Untersuchungen, Erw. = erwartete Messwerte, SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe, n = Anzahl

3.2.4 Verteilung der Werte des SAT und VAT

Die Verteilung der US-Messwerte des Abdominalfetts aus dem Datafreeze 100k der NAKO Gesundheitsstudie wurde bereits von Fischer, Sedlmeier et al. im Februar 2020 im Bundesgesundheitsblatt publiziert (74). Die Messwerte werden im Folgenden zitiert und um einige Grafiken ergänzt. Sie basieren auf den Messwerten der Level-2-Studienteilnehmern der Hauptphase bis zur Mitte der Basiserhebung und wurden zwischen dem 01.10.2014 und 16.03.2017 erhoben. 49,5% der 20.360 Studienteilnehmer mit vorhandenen US-Messwerten waren männlich ($n = 10.078$), 50,5% waren weiblich ($n = 10.282$). Das Durchschnittsalter lag bei 51,6 Jahren und erstreckte sich von 20 Jahren bis 73 Jahren. Der Mittelwert des SAT lag bei den Männern mit $2,0 \text{ cm} \pm 0,8 \text{ cm}$ ($n = 10.075$) niedriger als bei den Frauen mit $2,3 \text{ cm} \pm 1,0 \text{ cm}$ ($n = 10.278$), wohingegen das durchschnittliche VAT bei den Männern mit $7,8 \text{ cm} \pm 2,5 \text{ cm}$ ($n = 9.960$) stärker ausgeprägt war als bei den Frauen mit $5,8 \text{ cm} \pm 2,2 \text{ cm}$ ($n = 10.147$) (vergleiche Tabelle 14). Die Werte zeigten eine leicht rechtsschiefe Verteilung (74).

Tabelle 14: Charakteristika der analysierten Studienteilnehmer des Data Freeze 100k Datensatzes (74).

	Gesamt			Männer			Frauen		
	N	Mean	SD	n	Mean	SD	n	Mean	SD
Alter [Jahre]	20360	51,6	12,1	10078	51,9	12,1	10282	51,3	12,1
SAT [cm]	20353	2,2	0,9	10075	2,0	0,8	10278	2,3	1,0
VAT [cm]	20107	6,7	2,6	9960	7,8	2,5	10147	5,8	2,2

N = Anzahl, Mean = Mittelwert, SD = Standardabweichung, SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe, cm = Zentimeter.

Die Werteverteilung wurde zusätzlich nach Geschlecht und Alter in 10-Jahres-Gruppen ermittelt (74). Bei den Frauen lag der Mittelwert des SAT bei nahezu allen Altersgruppen über dem Mittelwert der Männer (vergleiche Tabelle 15 und Abbildung 6). Den höchsten Wert des SAT erreichten Frauen über 50 Jahren mit 2,4 cm. Bei den Männern lag das Maximum des SAT bei 2,1 cm in der Altersgruppe der 30- bis 49-Jährigen und nahm bei älteren Männern wieder ab. Die Streuung der Messwerte des SAT war bei den Frauen größer als bei den Männern. Bei den Männern nahm die Streuung der Messwerte mit zunehmendem Alter ab, während sie bei den Frauen über alle Altersgruppen relativ konstant blieb.

Tabelle 15: Perzentilen des SAT (74).

Alter [Jahre]	N	SAT [cm]		SAT [cm] Perzentile						
		Mean	SD	5.	10.	25.	50.	75.	90.	95.
Männer										
Alle	10075	2,0	0,8	1,0	1,1	1,5	1,9	2,4	3,0	3,5
20-29	637	1,9	1,0	0,8	0,9	1,2	1,7	2,3	3,2	3,8
30-39	983	2,1	0,9	0,9	1,1	1,4	1,9	2,5	3,3	3,9
40-49	2256	2,1	0,8	1,0	1,2	1,5	2,0	2,5	3,2	3,7
50-59	3015	2,0	0,8	1,0	1,2	1,5	1,9	2,5	3,0	3,5
≥ 60	3184	1,9	0,7	1,0	1,1	1,4	1,8	2,3	2,8	3,2
Frauen										
Alle	10278	2,3	1,0	1,0	1,1	1,6	2,2	2,9	3,7	4,3
20-29	736	2,0	1,1	0,8	1,0	1,2	1,8	2,5	3,5	4,3
30-39	947	2,1	1,2	0,8	1,0	1,3	1,8	2,7	3,7	4,5
40-49	2479	2,3	1,1	0,9	1,1	1,5	2,1	2,9	3,7	4,3
50-59	3124	2,4	1,0	1,0	1,2	1,7	2,3	3,1	3,9	4,4
≥ 60	2992	2,4	0,9	1,1	1,3	1,7	2,3	2,9	3,6	4,1

SAT = subkutanes Fettgewebe, N = Anzahl, cm = Zentimeter, Mean = Mittelwert, SD = Standardabweichung.

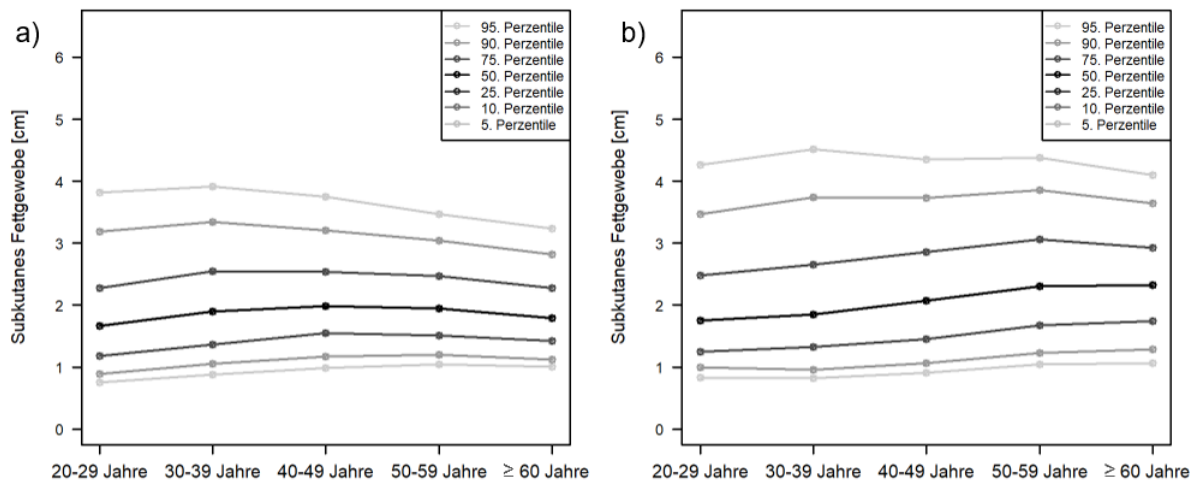


Abbildung 6: Perzentilen des SAT [cm] nach Altersgruppen bei a) Männern (N = 10.075) und b) Frauen (N = 10.278). SAT = subkutanes Fettgewebe, cm = Zentimeter, N = Anzahl.

Der Mittelwert des VAT der Männer lag in allen Altersgruppen über dem der Frauen (vergleiche Tabelle 16 und Abbildung 7). Bei beiden Geschlechtern nahm das VAT mit zunehmendem Alter kontinuierlich zu, sodass die über 60-jährigen Männer mit durchschnittlich 8,4 cm und die über 60-jährigen Frauen mit durchschnittlich 6,5 cm am meisten VAT hatten. Ebenso nahm die Streuung der Messwerte mit zunehmendem Alter zu (74).

Tabelle 16: Perzentilen des VAT (74).

Alter [Jahre]	N	Mean		VAT [cm]						
		SD	Perzentile							
				5.	10.	25.	50.	75.	90.	95.
Männer										
Alle	9960	7,8	2,5	4,2	4,7	5,9	7,5	9,3	11,1	12,4
20-29	632	5,9	1,7	3,6	4,0	4,7	5,6	6,8	7,9	9,1
30-39	977	6,6	2,0	3,9	4,3	5,2	6,5	7,7	9,2	10,3
40-49	2243	7,5	2,3	4,1	4,7	5,8	7,3	9,0	10,4	11,5
50-59	2982	8,0	2,5	4,4	5,0	6,2	7,8	9,6	11,4	12,7
≥ 60	3126	8,4	2,6	4,5	5,1	6,5	8,2	10,0	11,8	13,1
Frauen										
Alle	10147	5,8	2,2	3,0	3,4	4,2	5,3	6,9	8,9	10,0
20-29	731	4,4	1,3	2,6	3,0	3,5	4,2	5,0	6,1	6,7
30-39	934	4,7	1,7	2,6	3,0	3,6	4,4	5,6	7,0	7,9
40-49	2450	5,3	2,0	2,9	3,3	3,9	4,9	6,3	8,0	9,2
50-59	3078	6,0	2,2	3,3	3,7	4,5	5,6	7,1	9,0	10,2
≥ 60	2954	6,5	2,3	3,4	3,8	4,8	6,1	8,0	9,7	10,8

VAT = viszerales Fettgewebe, N = Anzahl, cm = Zentimeter, Mean = Mittelwert, SD = Standardabweichung.

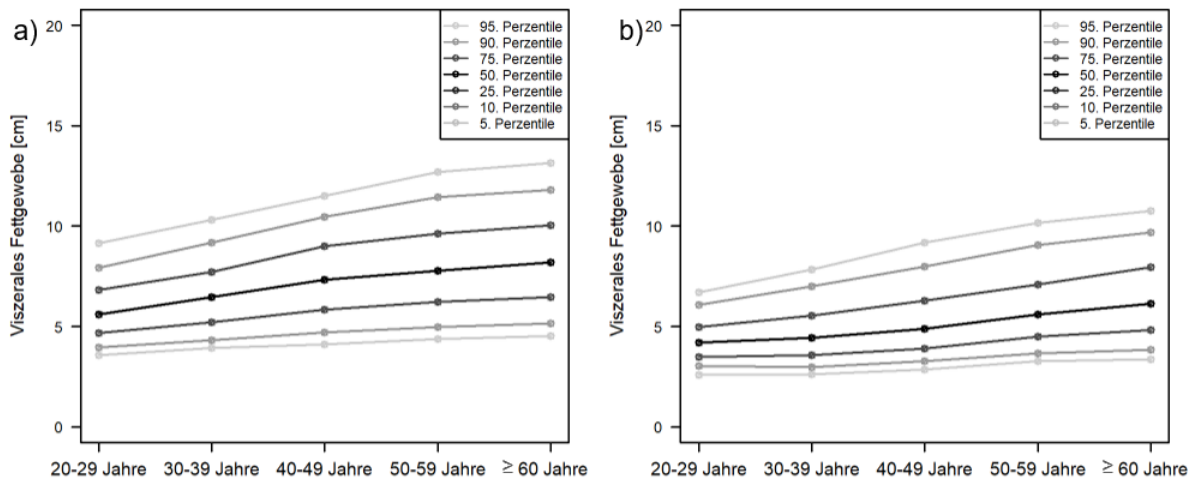


Abbildung 7: Perzentilen des VAT [cm] nach Altersgruppen bei a) Männern ($N = 9.960$) und b) Frauen ($N = 10.147$). VAT = viszerales Fettgewebe, cm = Zentimeter, N = Anzahl.

Das Verhältnis von SAT zu VAT verlagerte sich bei beiden Geschlechtern mit zunehmendem Alter auf die Seite des VAT (vergleiche Abbildung 8). Dabei wiesen die Männer aller Altersgruppen im Vergleich zu den Frauen verhältnismäßig mehr VAT als SAT auf.

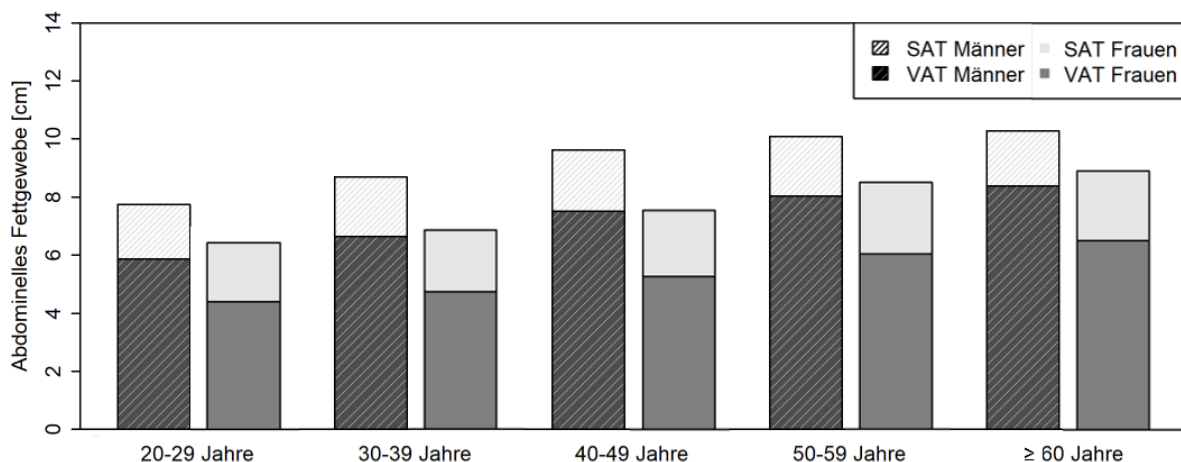


Abbildung 8: Verhältnis von SAT zu VAT nach Alter und Geschlecht. N (SAT Männer) = 10.075, N (SAT Frauen) = 10.278, N (VAT Männer) = 9.960, N (VAT Frauen) = 10.147. SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe, cm = Zentimeter, N = Anzahl.

3.2.5 SZ-, Untersucher- und Zeiteffekte auf die Werte des SAT und VAT

Neben der Verteilung der Werte wurden der Einfluss des SZ, der Untersucher und der Zeit auf die US-Messwerte des Abdominalfetts analysiert. Die US-Untersuchungen des Abdominalfetts wurden an 18 SZ durchgeführt. Der Vergleich der altersadjustierten Mittelwerte der einzelnen SZ zeigte regionale Unterschiede des SAT und VAT (siehe

Tabelle 17). Die alters- und geschlechtsadjustierten Durchschnittswerte des SAT an den einzelnen SZ variierten zwischen 1,8 cm und 2,5 cm, die des VAT zwischen 5,1 cm und 7,6 cm. Die höchsten Messwerte des SAT sowohl bei Männern als auch bei Frauen fanden sich im SZ Augsburg (Männer: 2,3 cm, Frauen: 2,8 cm). Am wenigsten SAT wurde bei beiden Geschlechtern im SZ Düsseldorf gemessen (Männer: 1,7 cm, Frauen 1,8 cm). Das VAT war bei den Männern mit 8,7 cm im SZ Essen am stärksten ausgeprägt. Bei den Frauen wurde im SZ Halle mit 6,7 cm am meisten VAT gemessen. Wie das SAT war auch das VAT im SZ Düsseldorf mit 5,8 cm bei den Männern und 4,5 cm bei den Frauen am niedrigsten.

Tabelle 17: Verteilung der Messwerte des SAT und VAT in Abhängigkeit des SZ, adjustiert nach Alter und Geschlecht der Studienteilnehmer.

	Gesamt			Männer			Frauen		
	N	Mean	SD	n	Mean	SD	n	Mean	SD
SAT [cm]									
Alle SZ	20353	2,2	0,9	10075	2,0	0,8	10278	2,3	1,0
Augsburg	4102	2,5	0,0	2099	2,3	0,0	2003	2,8	0,0
Regensburg	663	2,1	0,0	329	1,9	0,0	334	2,2	0,1
Heidelberg/Mannheim	1854	2,0	0,0	960	1,9	0,0	894	2,1	0,0
Freiburg	908	1,8	0,0	460	1,8	0,0	448	1,9	0,0
Saarbrücken	375	2,1	0,0	160	2,0	0,1	215	2,2	0,1
Essen	492	2,1	0,0	246	1,9	0,0	246	2,2	0,1
Münster	655	1,8	0,0	335	1,7	0,0	320	1,9	0,1
Düsseldorf	414	1,8	0,0	222	1,7	0,1	192	1,8	0,1
Halle	847	2,5	0,0	392	2,2	0,0	455	2,7	0,0
Leipzig	1186	2,2	0,0	558	1,9	0,0	628	2,5	0,0
Berlin Nord	829	2,0	0,0	409	1,9	0,0	420	2,1	0,0
Berlin Mitte	822	2,0	0,0	399	1,9	0,0	423	2,1	0,0
Berlin Süd/Brandenburg	679	1,9	0,0	342	1,8	0,0	337	1,9	0,1
Braunschweig/Hannover	637	1,9	0,0	308	1,8	0,0	329	1,9	0,1
Hamburg	932	1,9	0,0	462	1,9	0,0	470	1,9	0,0
Bremen	1102	2,0	0,0	458	1,9	0,0	644	2,1	0,0
Kiel	566	2,2	0,0	264	2,0	0,0	302	2,3	0,1
Neubrandenburg	3290	2,3	0,0	1672	2,1	0,0	1618	2,6	0,0
VAT [cm]									
Alle SZ	20107	6,7	2,6	9960	7,8	2,5	10147	5,8	2,2
Augsburg	4099	7,3	0,0	2097	8,6	0,0	2002	6,1	0,0
Regensburg	640	6,0	0,1	318	7,0	0,1	322	5,1	0,1
Heidelberg/Mannheim	1824	6,5	0,1	942	7,4	0,1	882	5,6	0,1
Freiburg	899	6,0	0,1	457	6,7	0,1	442	5,3	0,1
Saarbrücken	368	6,5	0,1	158	7,4	0,2	210	5,5	0,1
Essen	485	7,6	0,1	243	8,7	0,1	242	6,6	0,1
Münster	642	5,8	0,1	328	6,5	0,1	314	5,2	0,1
Düsseldorf	406	5,1	0,1	219	5,8	0,2	187	4,5	0,1
Halle	842	7,6	0,1	388	8,6	0,1	454	6,7	0,1
Leipzig	1181	7,0	0,1	555	8,1	0,1	626	6,0	0,1
Berlin Nord	816	6,7	0,1	402	7,8	0,1	414	5,7	0,1
Berlin Mitte	819	6,1	0,1	399	7,0	0,1	420	5,2	0,1
Berlin Süd/Brandenburg	666	5,9	0,1	338	6,7	0,1	328	5,1	0,1
Braunschweig/Hannover	615	6,1	0,1	302	7,0	0,1	313	5,3	0,1
Hamburg	929	5,9	0,1	459	6,7	0,1	470	5,0	0,1
Bremen	1090	6,7	0,1	455	7,8	0,1	635	5,7	0,1
Kiel	543	6,4	0,1	252	7,3	0,1	291	5,5	0,1
Neubrandenburg	3243	7,2	0,0	1648	8,3	0,1	1595	6,2	0,1

SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe, SZ = Studienzentrum, N = Anzahl, Mean = Mittelwert, SD = Standardabweichung, cm = Zentimeter.

Zur Analyse der statistischen Signifikanz der Abweichung der SZ-Mittelwerte vom globalen Mittelwert wurden *ANOM*-Plots berechnet, wofür die Daten als annähernd normalverteilt betrachtet wurden. Die Varianzen der Messwerte der einzelnen SZ waren in einem Levene-Test signifikant inhomogen ($p < 0,001$), sodass die *ANOM* unter Annahme von Varianzhomogenität durchgeführt wurde. Es zeigte sich, dass in 14 von 18 SZ die Mittelwerte des SAT signifikant vom globalen Mittelwert abwichen (siehe Abbildung 9). Der globale Mittelwert entsprach hierbei dem SZ-übergreifenden Mittelwert und lag bei 2,2 cm. Die Mittelwerte des SAT in den SZ Augsburg, Halle und Neubrandenburg waren signifikant größer als der globale Mittelwert. In den SZ Regensburg, Heidelberg/Mannheim, Freiburg, Münster, Düsseldorf, Berlin Nord, Mitte und Süd, Braunschweig/Hannover, Hamburg und Bremen war das mittlere SAT signifikant niedriger als im SZ-übergreifenden Durchschnitt. Die Nummerierung der SZ in der *ANOM* erfolgte nach Cluster (erste Ziffer) und Nummer (zweite Ziffer) (vergleiche Abbildung 1). Die Mittelwerte des VAT zeigten ebenfalls in 15 von 18 SZ signifikante Abweichungen vom globalen Mittelwert, der bei 6,7 cm lag (siehe Abbildung 10). In den SZ Augsburg, Essen, Halle, Leipzig und Neubrandenburg war das VAT signifikant größer als im globalen Mittel, während es in den SZ Regensburg, Heidelberg/Mannheim, Freiburg, Münster, Düsseldorf, Berlin Mitte und Süd, Braunschweig/Hannover, Hamburg und Kiel signifikant kleiner war. Die Entscheidungsgrenzen zur Signifikanzeinschätzung wurden nach Alter, Geschlecht und Anzahl der Studienteilnehmer je SZ adjustiert.

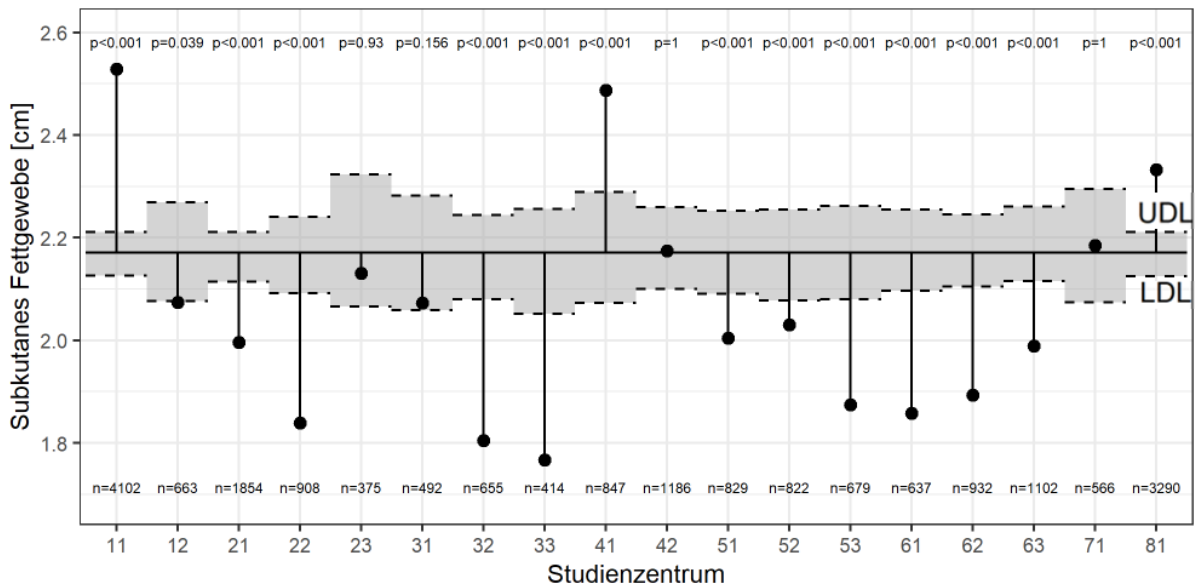


Abbildung 9: ANOM-Plot zur Abweichung der SZ-Mittelwerte des SAT vom globalen Mittelwert 2,2 cm (N = 20.353). Adjustierung der Entscheidungsgrenzen nach Alter, Geschlecht und Anzahl der Studienteilnehmer je SZ. Nummerierung der SZ: 11 = Augsburg, 12 = Regensburg, 21 = Heidelberg/Mannheim, 22 = Freiburg, 23 = Saarbrücken, 31 = Essen, 32 = Münster, 33 = Düsseldorf, 41 = Halle, 42 = Leipzig, 51 = Berlin Nord, 52 = Berlin Mitte, 53 = Berlin Süd/Brandenburg, 61 = Braunschweig/Hannover, 62 = Hamburg, 63 = Bremen, 71 = Kiel, 81 = Neubrandenburg.

ANOM = Analyse der Mittelwerte, SZ = Studienzentrum, SAT = subkutanes Fettgewebe, cm = Zentimeter, N = Anzahl, p = p-Wert, LDL = untere Entscheidungsgrenze, UDL = obere Entscheidungsgrenze.

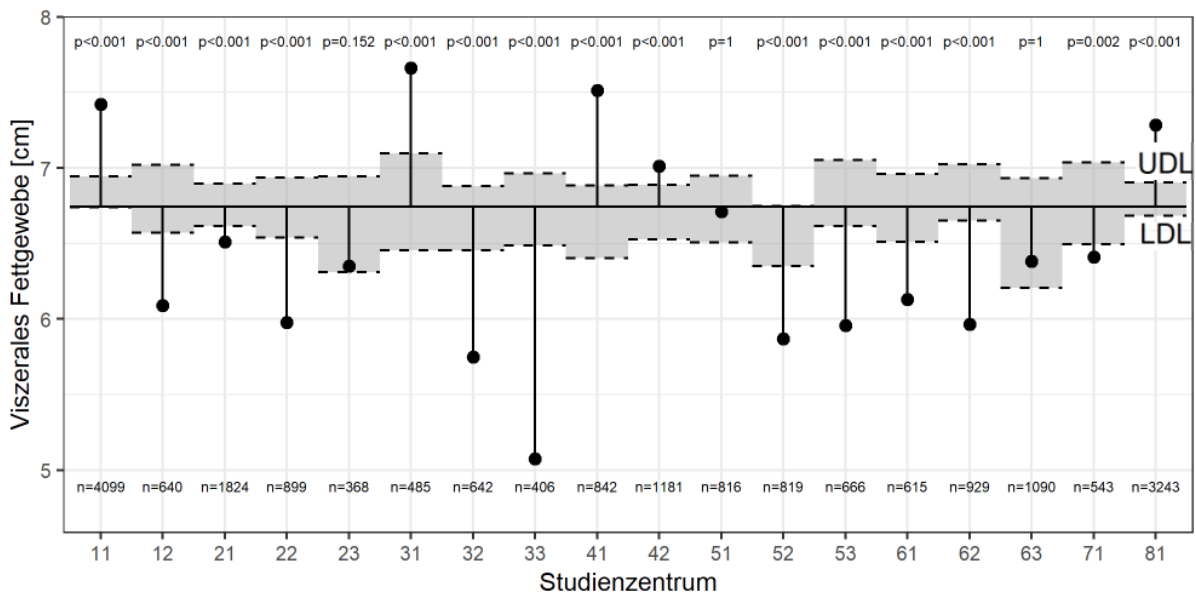


Abbildung 10: ANOM-Plot zur Abweichung der SZ-Mittelwerte des VAT vom globalen Mittelwert 6,7 cm (N = 20.107). Adjustierung der Entscheidungsgrenzen nach Alter, Geschlecht und Anzahl der Studienteilnehmer je SZ. Nummerierung der SZ: 11 = Augsburg, 12 = Regensburg, 21 = Heidelberg/Mannheim, 22 = Freiburg, 23 = Saarbrücken, 31 = Essen, 32 = Münster, 33 = Düsseldorf, 41 = Halle, 42 = Leipzig, 51 = Berlin Nord, 52 = Berlin Mitte, 53 = Berlin Süd/Brandenburg, 61 = Braunschweig/Hannover, 62 = Hamburg, 63 = Bremen, 71 = Kiel, 81 = Neubrandenburg.

ANOM = Analyse der Mittelwerte, SZ = Studienzentrum, VAT = viszerales Fettgewebe, cm = Zentimeter, N = Anzahl, p = p-Wert, LDL = untere Entscheidungsgrenze, UDL = obere Entscheidungsgrenze.

Die 18 beteiligten SZ ließen sich wiederum acht Clustern zuordnen (vergleiche Abbildung 1). Die alters- und geschlechtsadjustierten Durchschnittswerte des SAT in den acht Clustern variierten zwischen 1,9 cm und 2,5 cm, die des VAT zwischen 6,2 cm und 7,2 cm (vergleiche Tabelle 18).

Tabelle 18: Verteilung der Messwerte des SAT und VAT in Abhängigkeit des Clusters, adjustiert nach Alter und Geschlecht der Studienteilnehmer.

	Gesamt			Männer			Frauen		
	N	Mean	SD	n	Mean	SD	n	Mean	SD
SAT [cm]									
Alle Cluster	20353	2,2	0,9	10075	2,0	0,8	10278	2,3	1,0
Bayern	4765	2,5	0,0	2428	2,3	0,0	2337	2,7	0,0
Baden-Württemberg/Saarland	3137	2,0	0,0	1580	1,9	0,0	1557	2,1	0,0
Nordrhein-Westfalen	1561	1,9	0,0	803	1,8	0,0	758	2,0	0,0
Sachsen/Sachsen-Anhalt	2033	2,3	0,0	950	2,0	0,0	1083	2,6	0,0
Berlin-Brandenburg	2330	2,0	0,0	1150	1,9	0,0	1180	2,1	0,0
Niedersachsen/Hamburg/Bremen	2671	1,9	0,0	1228	1,9	0,0	1443	2,0	0,0
Schleswig-Holstein	566	2,2	0,0	264	2,0	0,0	302	2,3	0,1
Mecklenburg-Vorpommern	3290	2,3	0,0	1672	2,1	0,0	1618	2,6	0,0
VAT [cm]									
Alle Cluster	20107	6,7	2,6	9960	7,8	2,5	10147	5,8	2,2
Bayern	4739	7,1	0,0	2415	8,4	0,0	2324	5,9	0,0
Baden-Württemberg/Saarland	3091	6,3	0,0	1557	7,2	0,1	1534	5,5	0,1
Nordrhein-Westfalen	1533	6,2	0,1	790	7,0	0,1	743	5,5	0,1
Sachsen/Sachsen-Anhalt	2023	7,3	0,0	943	8,3	0,1	1080	6,3	0,1
Berlin-Brandenburg	2301	6,2	0,0	1139	7,2	0,1	1162	5,3	0,1
Niedersachsen/Hamburg/Bremen	2634	6,3	0,0	1216	7,2	0,1	1418	5,4	0,1
Schleswig-Holstein	543	6,4	0,1	252	7,3	0,1	291	5,5	0,1
Mecklenburg-Vorpommern	3243	7,2	0,0	1648	8,3	0,1	1595	6,2	0,1

SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe, SZ = Studienzentrum, N = Anzahl, Mean = Mittelwert, SD = Standardabweichung, cm = Zentimeter.

Die statistische Signifikanz der Abweichung der Clustermittelwerte vom globalen Mittelwert wurde ebenfalls mittels ANOM-Plots unter der Annahme normalverteilter Daten und Varianzinhomogenität berechnet. In sieben von acht Clustern wich der Mittelwert des SAT signifikant vom globalen Mittelwert ab. Dabei wurde in den Clustern Bayern, Sachsen/Sachsen-Anhalt und Mecklenburg-Vorpommern signifikant mehr, in den Clustern Baden-Württemberg/Saarland, Nordrhein-Westfalen, Berlin-Brandenburg und Niedersachsen/Hamburg/Bremen signifikant weniger SAT als im Durchschnitt gemessen. Lediglich im Cluster Schleswig-Holstein, das das SZ Kiel umfasst, entsprach die Menge an SAT dem globalen Mittelwert. Die Mittelwerte des VAT differierten in allen acht Clustern signifikant, wobei ebenso wie bei dem SAT in den Clustern Bayern, Sachsen/Sachsen-Anhalt und Mecklenburg-Vorpommern signifikant mehr, in den restlichen Clustern signifikant weniger VAT gemessen wurde. Im Vergleich mit der ANOM der SZ konnte beobachtet werden, dass die Tendenz der Abweichung vom globalen Mittelwert in nahezu allen SZ außer den SZ Regensburg und Essen der Tendenz der Abweichung der Cluster entsprach.

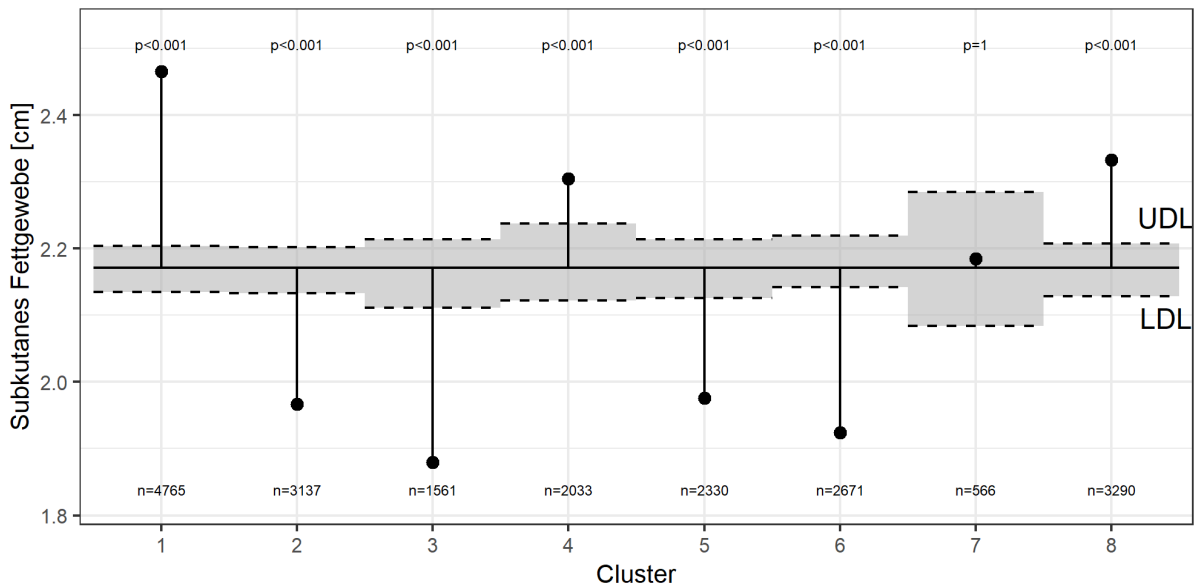


Abbildung 11: ANOM-Plot zur Abweichung der Clustermittelwerte des SAT vom globalen Mittelwert 2,2 cm (N = 20.353). Adjustierung der Entscheidungsgrenzen nach Alter, Geschlecht und Anzahl der Studienteilnehmer je Cluster. Nummerierung der Cluster: 1 = Bayern, 2 = Baden-Württemberg/Saarland, 3 = Nordrhein-Westfalen, 4 = Sachsen/Sachsen-Anhalt, 5 = Berlin-Brandenburg, 6 = Niedersachsen/Hamburg/Bremen, 7 = Schleswig-Holstein, 8 = Mecklenburg-Vorpommern. ANOM = Analyse der Mittelwerte, SAT = subkutanes Fettgewebe, cm = Zentimeter, N = Anzahl, p = p-Wert, LDL = untere Entscheidungsgrenze, UDL = obere Entscheidungsgrenze.

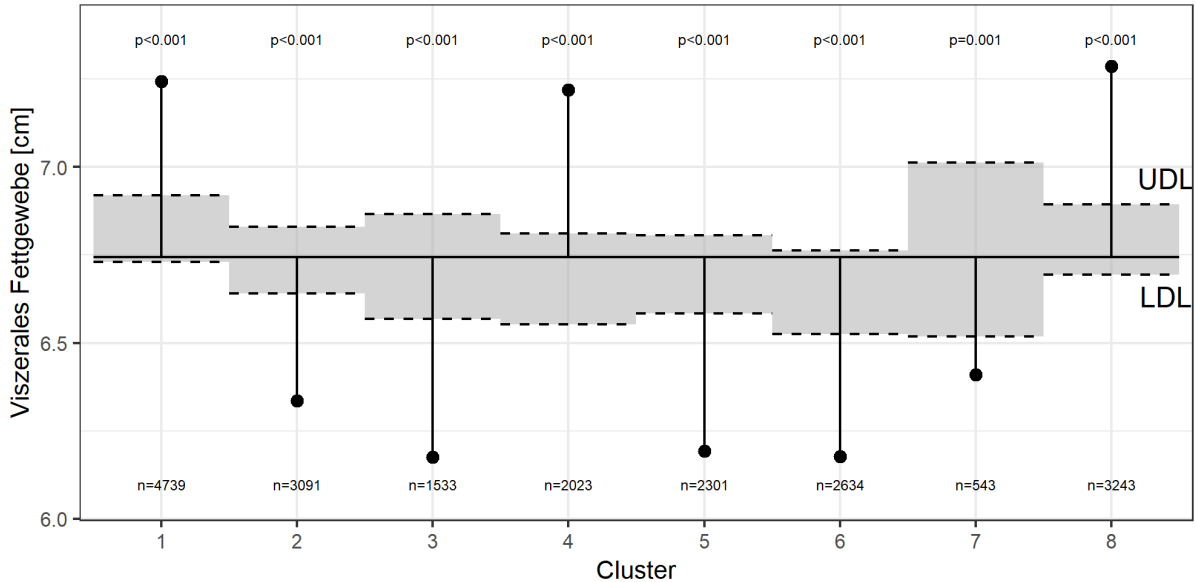


Abbildung 12: ANOM-Plot zur Abweichung der Clustermittelwerte des VAT vom globalen Mittelwert 6,7 cm (N = 20.107). Adjustierung der Entscheidungsgrenzen nach Alter, Geschlecht und Anzahl der Studienteilnehmer je Cluster. Nummerierung der Cluster: 1 = Bayern, 2 = Baden-Württemberg/Saarland, 3 = Nordrhein-Westfalen, 4 = Sachsen/Sachsen-Anhalt, 5 = Berlin-Brandenburg, 6 = Niedersachsen/Hamburg/Bremen, 7 = Schleswig-Holstein, 8 = Mecklenburg-Vorpommern. ANOM = Analyse der Mittelwerte, VAT = viszerales Fettgewebe, cm = Zentimeter, N = Anzahl, p = p-Wert, LDL = untere Entscheidungsgrenze, UDL = obere Entscheidungsgrenze.

Die US-Untersuchungen wurden an den 18 SZ von insgesamt 194 Untersuchern durchgeführt. Für die Analyse des Untersuchereffekts wurden alle jene Untersucher mit mindestens 50 Untersuchungen im Zeitraum des Data Freeze 100k betrachtet ($n = 74$). Von diesen Untersuchern waren die meisten Untersucher im SZ Heidelberg/Mannheim tätig ($n = 11$) und mit nur einem Untersucher pro SZ am wenigstens Untersucher in den SZ Saarbrücken und Essen. Diese 74 Untersucher führten insgesamt 18.922 vollständige Untersuchungen durch, was 93% aller vollständigen Untersuchungen im Data Freeze 100k-Datensatz entsprach ($n = 20.390$). Die meisten Untersuchungen wurden von einem Untersucher im SZ Augsburg durchgeführt ($n = 3.116$). Die Untersuchernummern wurden aus der Nummer des SZ und einem Pseudonym des Datenintegrationszentrums zusammengesetzt.

Tabelle 19: Verteilung der Messwerte des SAT und VAT in Abhängigkeit des Untersuchers, adjustiert nach Alter und Geschlecht der Studienteilnehmer.

	Untersucher	SAT [cm]			VAT [cm]		
		N	Mean	SD	N	Mean	SD
Alle SZ		18890	2,2	0,9	18666	6,8	2,6
Augsburg	1110002	204	2,4	0,1	203	7,4	0,2
	1110044	778	2,6	0,0	779	6,8	0,1
	1110243	3116	2,5	0,0	3113	7,4	0,0
Regensburg	1210062	323	2,2	0,0	318	6,1	0,1
	1210142	85	2,1	0,1	76	6,1	0,2
	1210277	92	2,0	0,1	84	5,7	0,2
	1210410	102	1,7	0,1	101	6,0	0,2
Heidelberg/Mannheim	2110029	65	2,0	0,1	63	5,8	0,3
	2110040	267	2,1	0,1	260	6,4	0,1
	2110049	138	2,3	0,1	133	7,6	0,2
	2110101	116	2,0	0,1	112	6,3	0,2
	2110130	95	1,8	0,1	96	5,7	0,2
	2110192	267	1,9	0,1	266	6,8	0,1
	2110209	219	2,0	0,1	219	6,2	0,1
	2110310	379	1,9	0,0	375	6,6	0,1
	2110343	110	2,2	0,1	107	5,4	0,2
	2110366	95	1,9	0,1	92	6,7	0,2
Freiburg	2110391	92	2,2	0,1	90	7,5	0,2
	2210137	58	1,9	0,1	58	6,8	0,3
	2210196	170	1,8	0,1	165	5,6	0,2
	2210240	163	1,7	0,1	163	5,6	0,2
	2210259	114	1,8	0,1	112	5,9	0,2
	2210371	157	2,1	0,1	157	6,7	0,2
	2210377	167	1,8	0,1	165	5,8	0,2
Saarbrücken	2310039	375	2,1	0,0	368	6,5	0,1
Essen	3110041	491	2,1	0,0	484	7,6	0,1
Münster	3210080	158	1,8	0,1	158	6,5	0,2
	3210128	215	1,7	0,1	216	4,7	0,1
	3210144	110	1,8	0,1	100	5,9	0,2
Düsseldorf	3310151	190	1,5	0,1	189	4,6	0,2
	3310259	98	1,8	0,1	92	5,5	0,2
	3310271	102	2,1	0,1	101	5,5	0,2
Halle	4110084	190	2,3	0,1	190	7,2	0,2
	4110115	125	2,2	0,1	121	7,7	0,2
	4110240	521	2,6	0,0	520	7,8	0,1
Leipzig	4210170	732	2,1	0,0	731	6,8	0,1
	4210228	51	1,9	0,1	49	6,7	0,3
	4210381	355	2,3	0,0	353	7,4	0,1

SZ	Untersucher	SAT [cm]			VAT [cm]		
		N	Mean	SD	N	Mean	SD
Berlin Nord	5110127	283	2,0	0,1	273	6,4	0,1
	5110319	341	1,8	0,0	339	6,2	0,1
	5110383	71	2,2	0,1	70	6,7	0,3
	5110409	134	2,4	0,1	134	8,6	0,2
Berlin Mitte	5210145	330	2,0	0,0	330	6,3	0,1
	5210174	283	2,0	0,1	283	6,2	0,1
	5210203	163	2,2	0,1	160	5,4	0,2
Berlin Süd/Brandenburg	5310003	165	2,3	0,1	162	6,2	0,2
	5310075	157	1,8	0,1	155	6,0	0,2
	5310120	108	1,9	0,1	105	6,6	0,2
	5310403	56	2,0	0,1	55	6,0	0,3
Braunschweig/Hannover	5310409	192	1,5	0,1	188	5,0	0,2
	6110196	54	1,7	0,1	51	6,4	0,3
	6110246	373	1,8	0,0	366	6,0	0,1
	6110270	104	2,0	0,1	92	6,4	0,2
Hamburg	6110287	70	2,0	0,1	70	6,0	0,3
	6210063	56	1,9	0,1	56	5,5	0,3
	6210085	59	2,5	0,1	59	7,0	0,3
	6210102	65	1,8	0,1	64	6,0	0,3
Bremen	6210173	52	1,6	0,1	53	4,9	0,3
	6310219	246	2,0	0,1	245	6,7	0,1
	6310292	254	2,0	0,1	247	6,9	0,1
	6310294	343	2,0	0,0	339	6,8	0,1
Kiel	6310345	167	1,9	0,1	168	6,6	0,2
	7110196	53	2,0	0,1	49	5,5	0,3
	7110271	192	2,4	0,1	188	7,2	0,2
	7110287	163	1,9	0,1	162	5,6	0,2
Neubrandenburg	8110077	651	2,4	0,0	615	7,3	0,1
	8110121	233	2,2	0,1	232	7,3	0,1
	8110138	212	2,8	0,1	210	6,9	0,1
	8110167	1052	2,5	0,0	1050	7,5	0,1
	8110217	89	2,1	0,1	88	7,7	0,2
	8110235	82	2,2	0,1	81	7,1	0,2
	8110326	190	2,1	0,1	190	6,1	0,2
8110364	331	2,0	0,0	329	6,9	0,1	
	8110372	431	2,2	0,0	429	7,3	0,1

SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe, cm = Zentimeter, N = Anzahl, Mean = Mittelwert, SD = Standardabweichung, SZ = Studienzentrum.

Für die Berechnung von ANOM-Plots wurden die Daten als annähernd normalverteilt betrachtet. Die Varianzen der Messwerte der einzelnen Untersucher zeigten sich in einem Levene-Test signifikant inhomogen ($p < 0,001$), sodass die ANOM unter Annahme von Varianzinhomogenität durchgeführt wurde. Die nach Alter und Geschlecht adjustierten Untersucher Mittelwerte des SAT lagen zwischen 1,5 cm und 2,8 cm. Bei 36 der 74 analysierten Untersucher wichen die Mittelwerte des SAT signifikant vom globalen Mittelwert ab. Dabei waren die Mittelwerte des SAT von 30 Untersuchern signifikant kleiner und von sechs Untersuchern signifikant größer als der globale Mittelwert. Lediglich bei einem Untersucher aus dem SZ Neubrandenburg (Untersuchernummer: 8110364) entsprach die Richtung der Abweichung vom globalen Mittelwert nicht der des SZ. Bei der Messung des VAT lagen die Untersucher-

mittelwerte zwischen 4,6 cm und 8,6 cm und es wichen 43 der 74 analysierten Untersucher signifikant vom globalen Mittelwert ab. Die Mittelwerte von 12 Untersuchern waren signifikant größer, von 31 Untersuchern signifikant kleiner als der globale Mittelwert. Bei zwei der signifikant abweichenden Untersucher (Untersuchernummern: 2110049, 8110326) entsprach die Richtung der Abweichung vom globalen Mittelwert nicht der des SZ. Auch hier wurden die Entscheidungsgrenzen nach Alter, Geschlecht und Anzahl der Studienteilnehmer je Untersucher adjustiert. Die ANOM-Plots befinden sich zudem vergrößert und einschließlich der Untersuchernummern und p-Werte im Anhang.

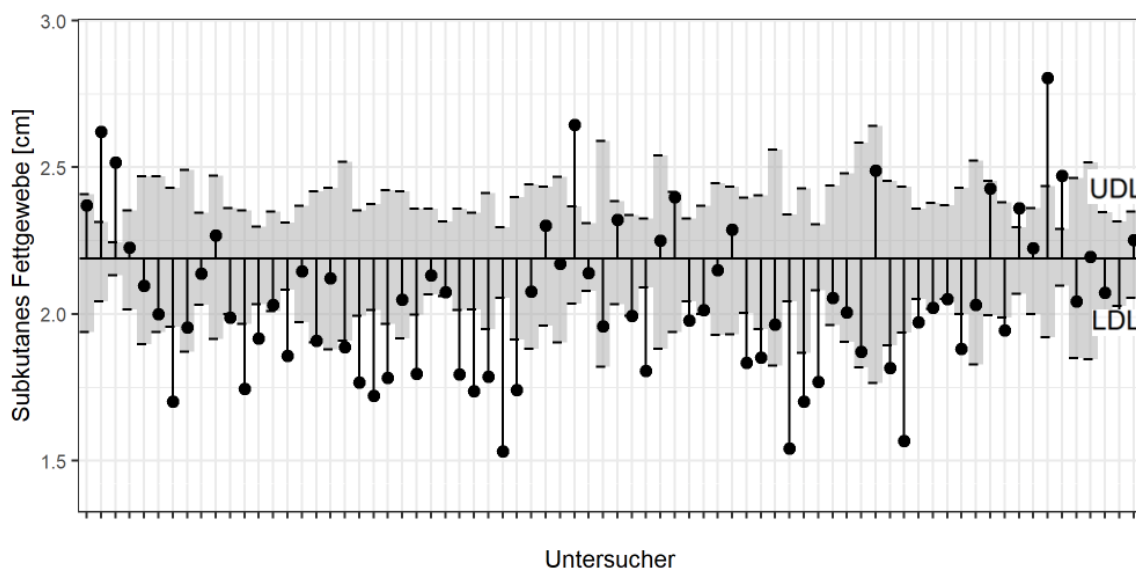


Abbildung 13: ANOM-Plot zur Abweichung der Untersuchermittelwerte des SAT vom globalen Mittelwert 2,2 cm ($N = 18.890$). Adjustierung der Entscheidungsgrenzen nach Alter, Geschlecht und Anzahl der Studienteilnehmer je Untersucher. ANOM = Analyse der Mittelwerte, SAT = subkutanes Fettgewebe, cm = Zentimeter, N = Anzahl, LDL = untere Entscheidungsgrenze, UDL = obere Entscheidungsgrenze.

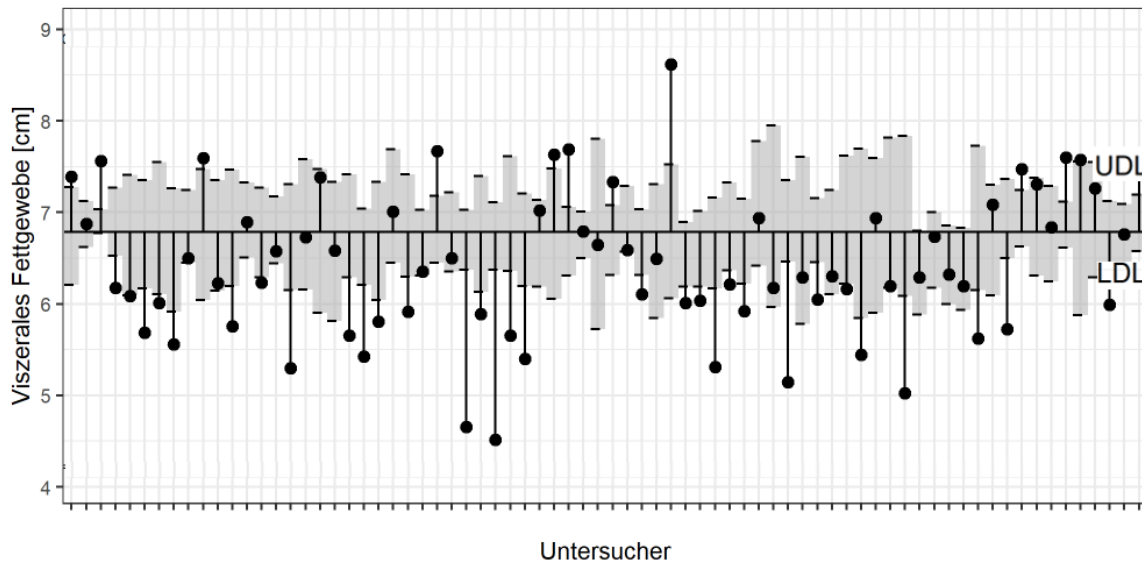


Abbildung 14: ANOM-Plot zur Abweichung der Untersuchermittelwerte des VAT vom globalen Mittelwert 6,8 cm ($N = 18.666$). Adjustierung der Entscheidungsgrenzen nach Alter, Geschlecht und Anzahl der Studienteilnehmer je Untersucher.
 ANOM = Analyse der Mittelwerte, VAT = viszerales Fettgewebe, cm = Zentimeter, N = Anzahl, LDL = untere Entscheidungsgrenze, UDL = obere Entscheidungsgrenze.

Darüber hinaus wurde der Effekt des Studienmonats, des Monats und der Uhrzeit auf die Varianz der US-Messwerte des Abdominalfetts betrachtet (vergleiche Tabelle 20). Unter Studienmonaten wurden dabei die 31 Monate der Hauptuntersuchung der NAKO Gesundheitsstudie bis zum Data Freeze 100k verstanden, während Monate den zwölf Monaten im Jahr entsprachen. Pro Studienmonat fanden durchschnittlich 652 Untersuchungen statt. Die meisten Untersuchungen wurden in den Studienmonaten Juni 2016 ($n = 1.112$) und November 2016 ($n = 1.106$) durchgeführt. Summiert fanden in den Februarmonaten ($n = 2.169$) und Novembermonaten ($n = 2.082$) der betrachteten Studienzeit am meisten US-Untersuchungen statt, wobei die Jahre 2014 und 2017 mit nur jeweils drei Monaten vertreten waren. Bei der Uhrzeit der US-Untersuchung wurden lediglich die Kernarbeitszeiten zwischen 8 Uhr morgens und 18 Uhr abends betrachtet. Die meisten US-Untersuchungen wurden zwischen 11 Uhr und 12 Uhr vormittags durchgeführt ($n = 3.756$).

Tabelle 20: Verteilung der Messwerte des SAT und VAT in Abhängigkeit der Zeit im Sinne von Studienmonat, Monat und Uhrzeit der US-Messung des Abdominalfetts, adjustiert nach Alter, Geschlecht und SZ. Fehlende Werte des Studienmonats und Monats: n = 150 (SAT) bzw. n = 147 (VAT). Fehlende Werte der Uhrzeit: n = 136 (SAT) bzw. n = 133 (VAT).

	SAT [cm]			VAT [cm]		
	N	Mean	SD	N	Mean	SD
Studienmonat	20203	2,2	0,9	19960	6,7	2,6
Oktober 2014	194	2,1	0,1	193	6,5	0,2
November 2014	259	2,1	0,1	257	6,5	0,1
Dezember 2014	217	2,0	0,1	214	6,6	0,1
Januar 2015	358	2,1	0,0	356	6,6	0,1
Februar 2015	413	2,1	0,0	410	6,8	0,1
März 2015	495	2,2	0,0	487	6,8	0,1
April 2015	443	2,2	0,0	434	6,5	0,1
Mai 2015	404	2,1	0,0	399	6,4	0,1
Juni 2015	598	2,1	0,0	594	6,3	0,1
Juli 2015	619	2,1	0,0	612	6,6	0,1
August 2015	528	2,1	0,0	520	6,6	0,1
September 2015	729	2,1	0,0	725	6,7	0,1
Oktober 2015	720	2,2	0,0	714	6,8	0,1
November 2015	717	2,2	0,0	714	6,8	0,1
Dezember 2015	493	2,1	0,0	492	6,6	0,1
Januar 2016	627	2,2	0,0	618	7,1	0,1
Februar 2016	743	2,3	0,0	734	7,0	0,1
März 2016	822	2,3	0,0	806	6,8	0,1
April 2016	909	2,2	0,0	901	6,9	0,1
Mai 2016	739	2,3	0,0	731	6,9	0,1
Juni 2016	1112	2,2	0,0	1097	6,9	0,1
Juli 2016	913	2,1	0,0	906	6,8	0,1
August 2016	783	2,2	0,0	771	6,8	0,1
September 2016	922	2,2	0,0	905	6,7	0,1
Oktober 2016	954	2,2	0,0	940	6,8	0,1
November 2016	1106	2,1	0,0	1096	6,6	0,1
Dezember 2016	736	2,1	0,0	722	6,6	0,1
Januar 2017	1002	2,1	0,0	991	7,0	0,1
Februar 2017	1013	2,2	0,0	1001	6,8	0,1
März 2017	635	2,2	0,0	620	6,9	0,1
Monat	20203	2,2	0,9	19960	6,7	2,6
Januar	1987	2,2	0,0	1965	6,9	0,0
Februar	2169	2,2	0,0	2145	6,8	0,0
März	1952	2,2	0,0	1913	6,8	0,0
April	1352	2,2	0,0	1335	6,8	0,1
Mai	1143	2,2	0,0	1130	6,7	0,1
Juni	1710	2,2	0,0	1691	6,7	0,1
Juli	1532	2,1	0,0	1518	6,7	0,1
August	1311	2,1	0,0	1291	6,7	0,1
September	1651	2,2	0,0	1630	6,7	0,1
Oktober	1868	2,2	0,0	1847	6,8	0,1
November	2082	2,1	0,0	2067	6,7	0,0
Dezember	1446	2,1	0,0	1428	6,6	0,1
Uhrzeit	20217	2,2	0,9	19974	6,7	2,6
08:00-08:59	569	2,2	0,0	561	6,7	0,1
09:00-09:59	2361	2,2	0,0	2340	6,7	0,0
10:00-10:59	3464	2,2	0,0	3436	6,7	0,0
11:00-11:59	3756	2,2	0,0	3722	6,8	0,0
12:00-12:59	3446	2,2	0,0	3405	6,7	0,0
13:00-13:59	2521	2,2	0,0	2480	6,8	0,0
14:00-14:59	1630	2,2	0,0	1604	6,9	0,1
15:00-15:59	1141	2,2	0,0	1122	6,8	0,1
16:00-16:59	647	2,1	0,0	637	6,8	0,1
17:00-17:59	464	2,2	0,0	454	6,9	0,1

SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe, US = Ultraschall, SZ = Studienzentrum, n = Anzahl, cm = Zentimeter, Mean = Mittelwert, SD = Standardabweichung.

Beim Vergleich der nach Alter, Geschlecht und SZ adjustierten Mittelwerte der US-Messwerte fiel auf, dass sie zu Beginn der Basisuntersuchung stärker um den globalen Mittelwert schwankten (siehe Abbildung 15). In der zweiten Hälfte des betrachteten Zeitraums stabilisierten sie sich. Bei der Messung des VAT konnten Peaks in den Wintermonaten der drei betrachteten Jahre beobachtet werden.

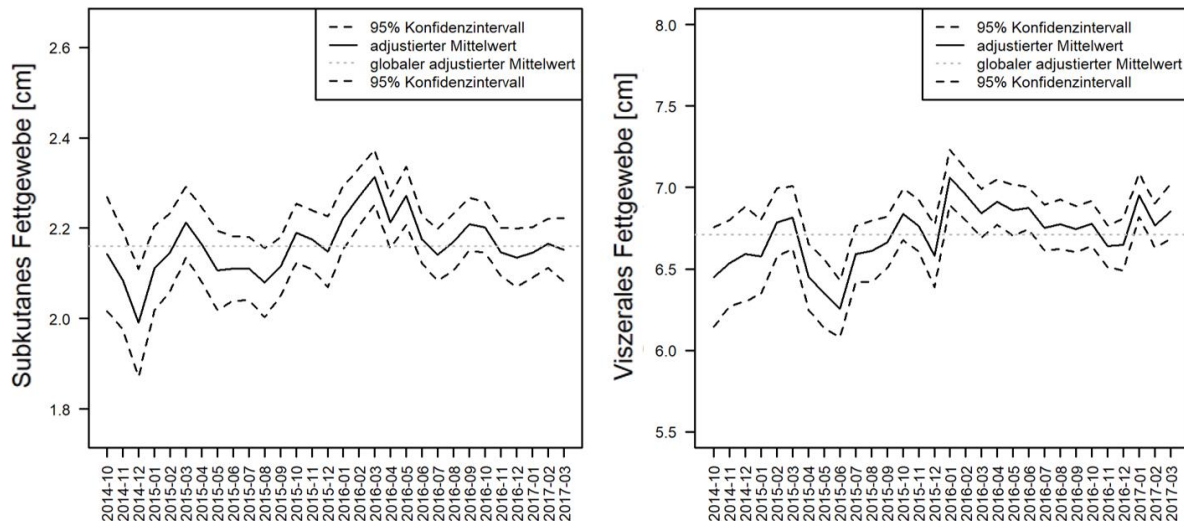


Abbildung 15: Variation der Mittelwerte des SAT und VAT nach Studienmonaten, adjustiert nach Alter, Geschlecht und SZ. Globaler Mittelwert des SAT: 2,2 cm (N = 20.203), globaler Mittelwert des VAT: 6,7 cm (N = 19.960). SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe, SZ = Studienzentrum, cm = Zentimeter, N = Anzahl.

Bei Betrachtung der ANOM-Plots zeigte sich, dass der Mittelwert des SAT im März 2016 signifikant größer war als in den restlichen Studienmonaten (siehe Abbildung 17). Bei der Messung des VAT wichen die Mittelwerte im Mai und Juni 2015 sowie im Januar 2016 signifikant vom globalen Mittelwert ab (siehe Abbildung 17). Dabei fanden sich im Mai und Juni 2015 signifikant kleinere Mittelwerte und im Januar 2016 ein signifikant größerer Mittelwert des VAT. Die Entscheidungsgrenzen zur Signifikanz-einschätzung wurden nach Alter, Geschlecht, SZ und Anzahl der Studienteilnehmer je Studienmonat adjustiert.

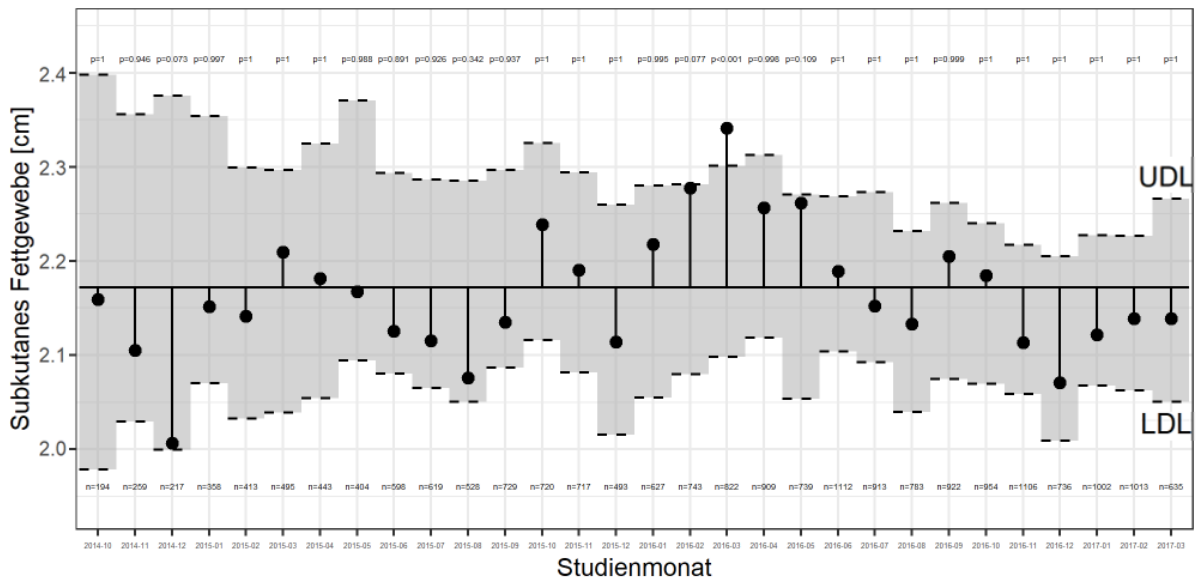


Abbildung 16: ANOM-Plot zur Abweichung der Studienmonatsmittelwerte des SAT vom globalen Mittelwert 2,2 cm (N = 20.203). Adjustierung der Entscheidungsgrenzen nach Alter, Geschlecht, SZ und Anzahl der Studienteilnehmer je Studienmonat.
 ANOM = Analyse der Mittelwerte, SAT = subkutanes Fettgewebe, cm = Zentimeter, N = Anzahl, SZ = Studienzentrum, p = p-Wert, LDL = untere Entscheidungsgrenze, UDL = obere Entscheidungsgrenze.

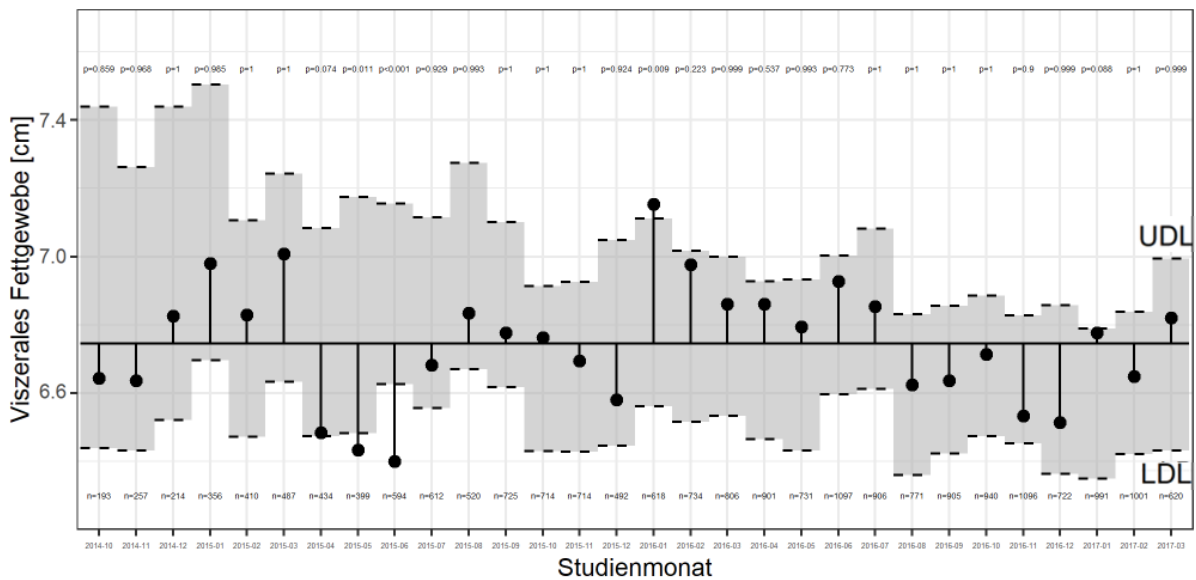


Abbildung 17: ANOM-Plot zur Abweichung der Studienmonatsmittelwerte des VAT vom globalen Mittelwert 6,7 cm (N = 19.960). Adjustierung der Entscheidungsgrenzen nach Alter, Geschlecht, SZ und Anzahl der Studienteilnehmer je Studienmonat.
 ANOM = Analyse der Mittelwerte, VAT = viszerales Fettgewebe, cm = Zentimeter, N = Anzahl, SZ = Studienzentrum, p = p-Wert, LDL = untere Entscheidungsgrenze, UDL = obere Entscheidungsgrenze.

Im Jahresverlauf variierten die US-Messwerte des SAT und VAT wenig. Bei der Messung des SAT fanden sich geringfügige Peaks im Frühjahr und Herbst, während im Winter mehr VAT gemessen wurde (siehe Abbildung 18).

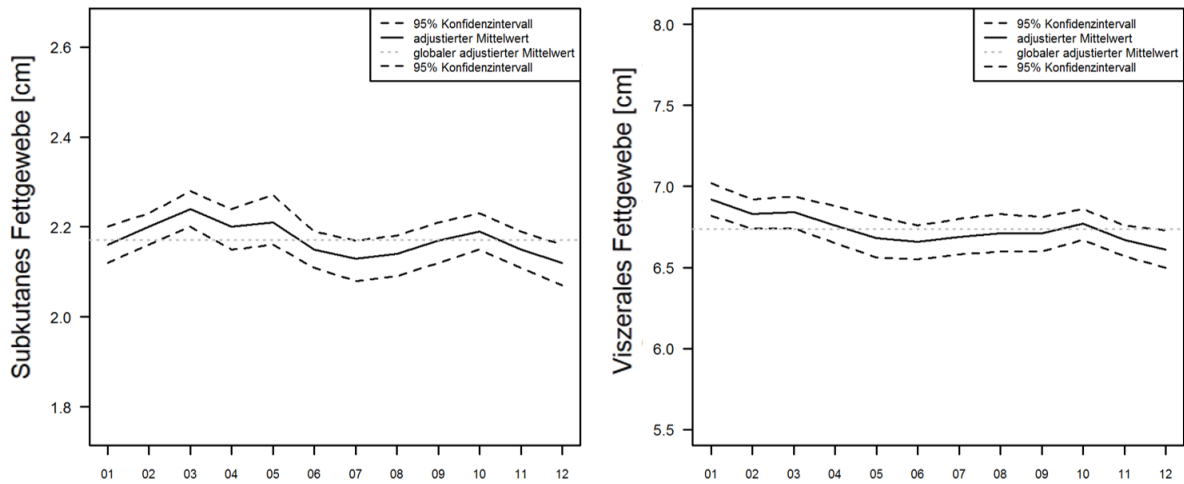


Abbildung 18: Variation der Mittelwerte des SAT und VAT nach Monat, adjustiert nach Alter, Geschlecht und SZ. Globaler Mittelwert des SAT: 2,2 cm (N = 20.203), globaler Mittelwert des VAT: 6,7 cm (N = 19.960). SAT = subkutanes Fettgewebe, VAT = viszerale Fettgewebe, SZ = Studienzentrum, cm = Zentimeter, N = Anzahl.

Im Monat März war der Mittelwert des SAT signifikant größer als der globale Mittelwert. Bei der Messung des VAT war der Mittelwert im Januar signifikant größer als der globale Mittelwert (siehe Abbildung 19). Auch hier wurden die Daten als annähernd normalverteilt betrachtet. Die Varianzen der Messwerte der Studienmonate zeigten sich in einem Levene-Test signifikant inhomogen ($p < 0,05$), sodass die ANOM unter Annahme von Varianzinhomogenität durchgeführt wurde.

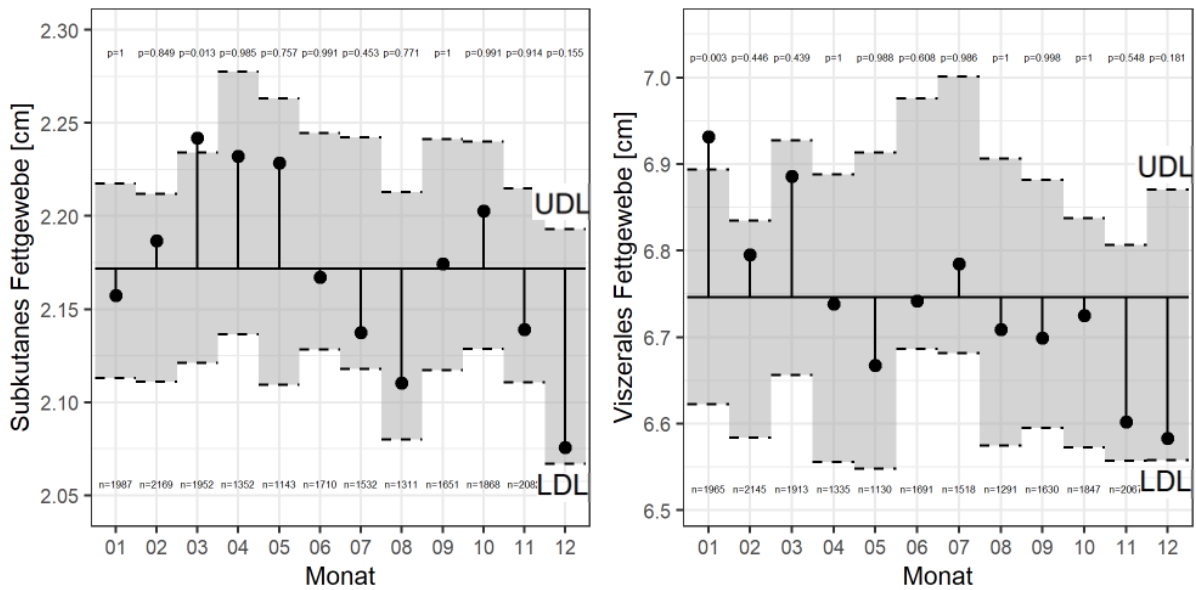


Abbildung 19: ANOM-Plots zur Abweichung der Monatsmittelwerte des SAT vom globalen Mittelwert 2,2 cm ($N = 20.203$) sowie des VAT vom globalen Mittelwert 6,7 cm ($N = 19.960$). Adjustierung der Entscheidungsgrenzen nach Alter, Geschlecht, SZ und Anzahl der Studienteilnehmer pro Monat.

ANOM = Analyse der Mittelwerte, SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe, cm = Zentimeter, N = Anzahl, SZ = Studienzentrum, p = p -Wert, LDL = untere Entscheidungsgrenze, UDL = obere Entscheidungsgrenze.

Die Messwerte des SAT verhielten sich im Tagesverlauf relativ konstant. Bei der Messung des VAT schienen die Messwerte in der zweiten Tageshälfte etwas höher zu liegen als in der ersten Tageshälfte (siehe Abbildung 20).

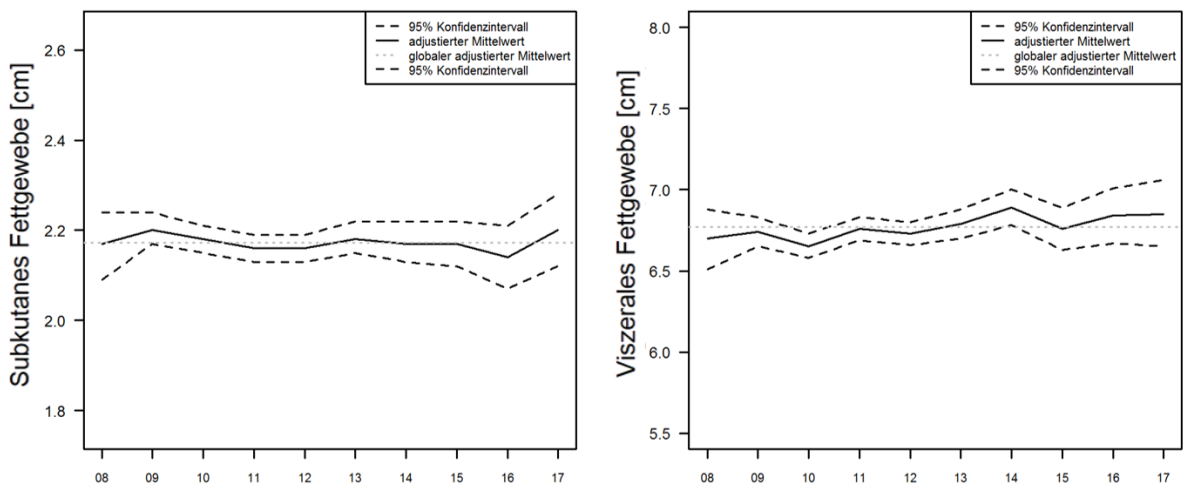


Abbildung 20: Variation der Mittelwerte des SAT und des VAT nach Uhrzeit, adjustiert nach Alter, Geschlecht und SZ. Globaler Mittelwert des SAT: 2,2 cm ($N = 20.217$), globaler Mittelwert des VAT: 6,7 cm ($N = 19.974$).

SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe, SZ = Studienzentrum, cm = Zentimeter, N = Anzahl.

In der ANOM zeigte sich, dass der Mittelwert des VAT derjenigen Studienteilnehmer, die zwischen 10:00 und 10:59 Uhr untersucht wurden, signifikant niedriger als der globale Mittelwert war. Zu den restlichen Tageszeiten wichen weder die Mittelwerte des SAT noch des VAT signifikant vom globalen Mittelwert ab (siehe Abbildung 21). Auch hier wurden die Daten als annähernd normalverteilt betrachtet. Die Varianzen der Messwerte der Studienmonate waren in einem Levene-Test signifikant inhomogen ($p < 0,01$), sodass die ANOM unter Annahme von Varianzhomogenität durchgeführt wurde.

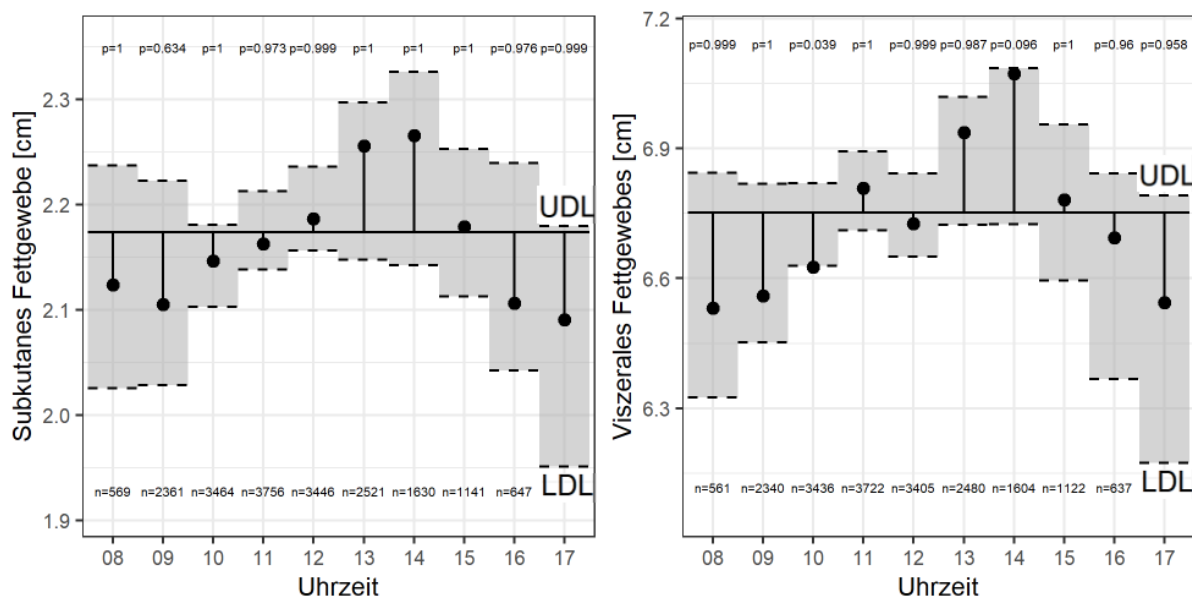


Abbildung 21: ANOM-Plots zur Abweichung der Mittelwerte des SAT nach Uhrzeit vom globalen Mittelwert 2,2 cm ($N = 20.217$) sowie des VAT nach Uhrzeit vom globalen Mittelwert 6,7 cm ($N = 19.974$). Adjustierung der Entscheidungsgrenzen nach Alter, Geschlecht, SZ und Anzahl der Studienteilnehmer je Uhrzeit. ANOM = Analyse der Mittelwerte, SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe, cm = Zentimeter, N = Anzahl, SZ = Studienzentrum, p = p -Wert, LDL = untere Entscheidungsgrenze, UDL = obere Entscheidungsgrenze.

Zusammenfassend konnte gezeigt werden, dass sowohl die SZ, Cluster als auch Untersucher einen signifikanten Effekt auf die Messwerte des SAT und VAT hatten, während die Zeit wenig Einfluss auf die Variation der Messwerte nahm.

3.2.6 Interne Validierung der Doppelmessung des SAT und VAT

Das SAT und VAT wurde in Doppelmessungen von einem Untersucher unter identischen Bedingungen gemessen, sodass eine Abhängigkeit zwischen erstem und zweitem Messwert bestand. Für die interne Validierung wurden nur Messwerte derjenigen Untersucher betrachtet, die mindestens 50 Untersuchungen im Zeitraum des Data Freeze 100k durchgeführt hatten ($n = 74$). Die durchschnittliche absolute

Abweichung zwischen erstem und zweitem Messwert des SAT betrug 0,1 cm ($n = 18.883$), während erster und zweiter Messwert des VAT um durchschnittlich 0,2 cm voneinander abwichen ($n = 18.629$). Fast alle Messwertpaare des SAT (99,7%) erfüllten die Anforderung der SOP „US Abdominalfett“ (63) und wichen weniger als 0,5 cm voneinander ab ($n = 18.829$). Von den Messwertpaaren des VAT lagen 94,4% maximal 0,5 cm voneinander entfernt ($n = 17.831$).

Um das Ausmaß sowie Muster der Diskrepanz der Doppelmessung zu analysieren, wurden Bland-Altman-Plots erstellt (77). Bei Betrachtung aller Untersucher lag der Mittelwert der Differenz zwischen erstem und zweitem Messwert des SAT bei -0,01 cm, sodass keine systematische Differenz zwischen den beiden Messungen vorhanden war (vergleiche Abbildung 22). Die Mittelwerte der Differenz der 74 Untersucher befanden sich zwischen -0,06 cm und +0,05 cm (vergleiche Tabelle 21). Im Bland-Altman-Plot der Doppelmessung des SAT verteilte sich die Punktwolke annähernd symmetrisch, sodass kein Trend beobachtet werden konnte. Das Ausmaß der Diskrepanz zwischen erstem und zweitem Messwert wurde anhand der Spannweite des Mittelwerts der Differenz abgeschätzt, welche dem Mittelwert der Differenz \pm zwei Standardabweichungen entspricht. Diese Spannweite lag bei der Messung des SAT bei 0,5 cm, wobei sich 93,2% der Messwertpaare ($n = 17.599$) innerhalb dieses Bereiches befanden. Daraus ergab sich eine gute Übereinstimmung der Messwerte des SAT.

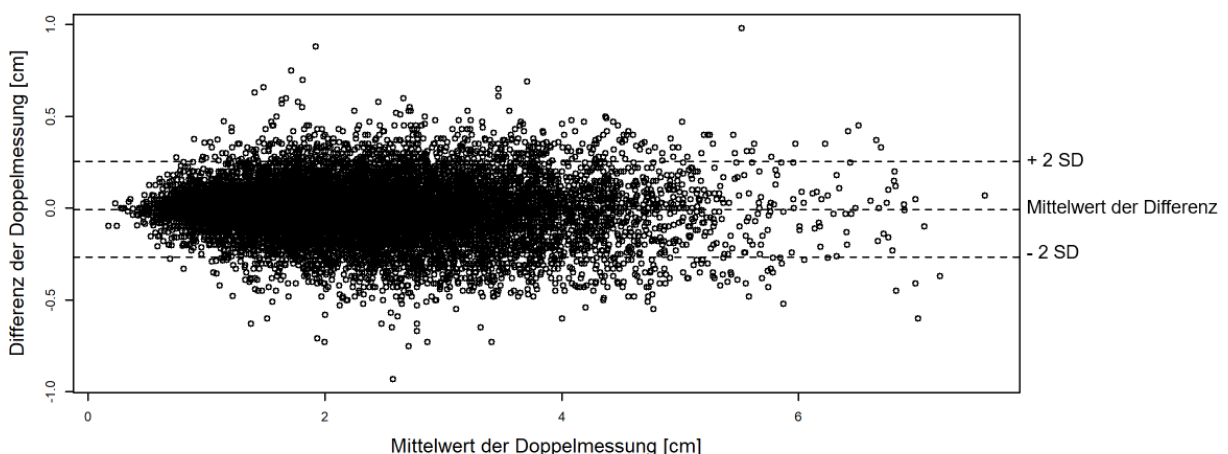


Abbildung 22: Bland-Altman-Plot der Doppelmessung des SAT ($N = 18.883$). Darstellung des Mittelwerts der Messwertpaare auf der Abszisse sowie der Differenz der Messwertpaare auf der Ordinate. Mittelwert der Differenz aller Messwertpaare = -0,01 cm, + 2 SD = 0,26 cm, - 2 SD = -0,27 cm. SAT = subkutanes Fettgewebe, N = Anzahl, SD = Standardabweichung, cm = Zentimeter.

Bei der Doppelmessung des VAT (vergleiche Abbildung 23) lag der Mittelwert der Differenz bei Betrachtung aller Untersucher mit -0,01 cm ebenfalls nahe Null, sodass auch hier keine systematische Differenz zwischen den beiden Messungen beobachtet werden konnte. Die Mittelwerte der Differenz der 74 Untersucher befanden sich zwischen -0,11 cm und +0,09 cm (vergleiche Tabelle 21). Da sich die Punktwolke im Bland-Altman-Plot auch hier annähernd symmetrisch verteilte, war kein Trend zu sehen. Die Spannweite des Mittelwerts der Differenz lag bei 1,0 cm, wobei sich 96,2% der Messwertpaare innerhalb von \pm zwei Standardabweichungen befanden ($n = 17.921$). Die Übereinstimmung der Messwerte des VAT konnte somit als sehr gut gewertet werden.

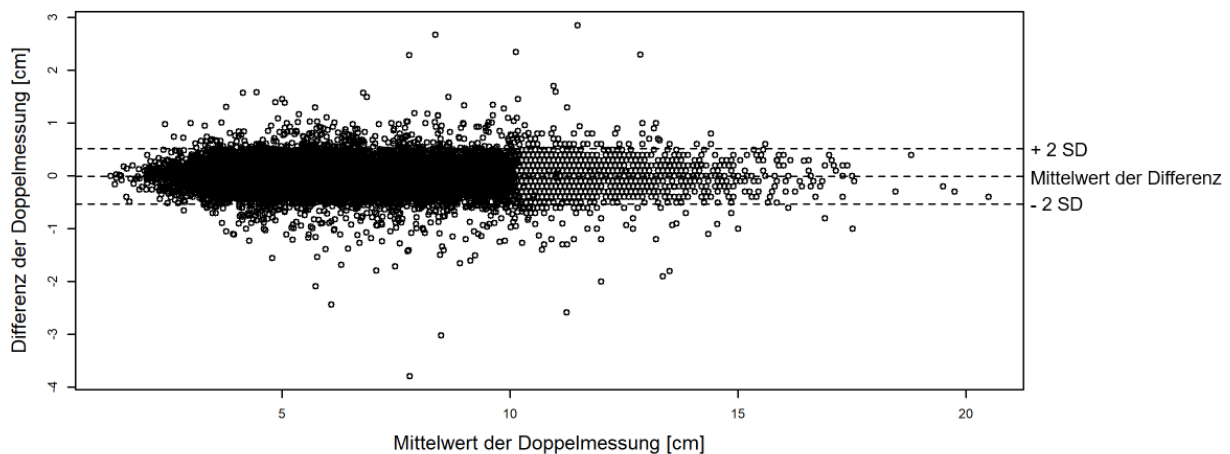


Abbildung 23: Bland-Altman-Plot der Doppelmessung des VAT ($N = 18.629$). Darstellung des Mittelwerts der Messwertpaare auf der Abszisse sowie der Differenz der Messwertpaare auf der Ordinate. Mittelwert der Differenz aller Messwertpaare = -0,01 cm, + 2 SD = 0,51 cm, - 2 SD = -0,53 cm. VAT = viszerales Fettgewebe, N = Anzahl, SD = Standardabweichung, cm = Zentimeter.

Die Intraobserver-Reliabilität wurde mithilfe des ICC gemessen. Die ICC-Werte bei der Messung des SAT lagen allesamt über 0,97, die ICC-Werte bei der Messung des VAT über 0,95 (vergleiche Tabelle 21). Somit konnte eine exzellente Reliabilität bei der US-Messung des Abdominalfetts angenommen werden.

Tabelle 21: Intraobserver-Reliabilität und mittlere Differenz bei der US-Messung des SAT und VAT.

	Untersucher	SAT			VAT		
		N	ICC	Mittlere Differenz [cm]	N	ICC	Mittlere Differenz [cm]
Alle SZ		18883	-	-0,01	18629	-	-0,01
Augsburg	1110002	204	0,98	-0,03	203	1,00	-0,02
	1110044	778	0,99	0,00	779	1,00	-0,00
	1110243	3115	0,99	-0,03	3113	1,00	-0,01
Regensburg	1210062	323	0,99	-0,04	318	0,99	-0,03
	1210142	85	0,99	0,00	74	0,99	-0,01
	1210277	92	0,98	0,04	83	0,99	-0,02
	1210410	102	0,99	-0,02	101	0,99	-0,01

SZ	Untersucher	SAT			VAT		
		N	ICC	Mitt. Diff. [cm]	N	ICC	Mitt. Diff. [cm]
Heidelberg/ Mannheim	2110029	65	0,99	-0,01	63	0,99	-0,01
	2110040	267	0,98	-0,01	260	0,99	-0,04
	2110049	138	1,00	-0,01	133	1,00	0,01
	2110101	116	1,00	0,01	111	1,00	-0,01
	2110130	95	0,99	0,01	96	0,99	0,00
	2110192	267	0,99	-0,01	266	0,99	-0,05
	2110209	219	0,98	0,04	219	0,99	-0,02
	2110310	379	0,99	-0,01	375	0,99	-0,05
	2110343	110	1,00	0,01	94	0,98	0,05
	2110366	95	0,99	0,01	92	0,99	-0,05
	2110391	92	0,97	-0,06	88	0,99	-0,04
Freiburg	2210137	58	0,99	-0,01	58	0,99	-0,06
	2210196	170	0,98	-0,01	163	0,99	-0,08
	2210240	163	0,99	-0,01	163	0,99	0,01
	2210259	114	0,99	0,03	112	0,99	-0,04
	2210371	157	0,99	0,01	157	0,99	-0,03
	2210377	166	0,98	-0,01	165	1,00	-0,03
Saarbrücken	2310039	375	0,99	0,01	368	0,99	-0,01
Essen	3110041	491	0,98	0,05	484	0,99	0,04
Münster	3210080	158	0,99	0,01	158	0,98	0,02
	3210128	215	0,98	0,01	216	0,95	0,00
	3210144	110	0,99	-0,02	100	1,00	-0,05
Düsseldorf	3310151	190	0,99	0,00	189	0,99	0,03
	3310259	98	0,99	0,02	90	0,97	-0,03
	3310271	102	0,98	0,02	101	0,99	0,00
Halle	4110084	190	0,99	-0,01	190	1,00	0,04
	4110115	125	0,99	-0,04	121	0,99	-0,06
	4110240	521	0,99	-0,01	520	1,00	0,01
Leipzig	4210170	732	0,99	-0,03	731	1,00	-0,04
	4210228	51	0,99	0,02	49	1,00	-0,06
	4210381	355	0,99	-0,00	353	1,00	-0,03
Berlin Nord	5110127	283	0,98	-0,05	271	0,99	-0,04
	5110319	341	0,99	0,01	339	1,00	-0,03
	5110383	71	0,99	-0,01	70	0,99	-0,08
	5110409	134	0,99	-0,02	134	1,00	-0,05
Berlin Mitte	5210145	330	1,00	0,02	330	1,00	-0,00
	5210174	283	1,00	0,02	283	0,99	0,00
	5210203	163	1,00	-0,01	160	1,00	-0,01
Berlin Süd/ Brandenburg	5310003	163	0,99	-0,03	162	0,99	-0,03
	5310075	157	0,99	-0,02	155	1,00	0,01
	5310120	108	0,97	-0,05	104	0,99	-0,04
	5310403	56	0,98	-0,04	55	0,99	-0,04
	5310409	192	0,99	0,01	187	0,99	-0,03
Braunschweig/ Hannover	6110196	54	0,99	0,01	50	0,99	0,09
	6110246	372	0,97	-0,01	365	0,98	0,02
	6110270	104	0,99	0,01	92	0,98	-0,03
	6110287	70	0,99	0,04	70	0,99	-0,06
Hamburg	6210063	56	0,99	-0,00	56	1,00	-0,00
	6210085	59	0,99	0,01	59	0,99	-0,11
	6210102	65	0,99	0,00	64	0,99	-0,03
	6210173	52	0,97	-0,02	53	0,99	0,02
Bremen	6310219	246	1,00	0,01	245	0,99	0,02
	6310292	254	1,00	0,00	245	1,00	0,01
	6310294	342	0,98	-0,05	339	0,99	-0,03
	6310345	167	0,99	-0,02	167	0,98	-0,05
	Kiel	7110196	53	0,99	0,02	48	0,98
7110271		192	0,99	-0,02	188	0,99	-0,03
7110287		163	0,99	0,01	162	0,98	-0,07
Neubrandenburg	8110077	651	1,00	0,03	614	1,00	0,01
	8110121	233	0,99	0,02	232	1,00	0,03
	8110138	212	0,99	-0,02	210	1,00	0,00
	8110167	1051	0,99	-0,01	1049	1,00	0,01
	8110217	89	1,00	0,02	88	1,00	0,02
	8110235	82	1,00	-0,02	80	0,99	-0,06
	8110326	190	0,99	0,03	189	0,99	0,01
	8110364	331	1,00	0,01	329	1,00	-0,00
	8110372	431	0,99	-0,01	429	0,99	-0,03

US = Ultraschall, SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe, N = Anzahl, ICC = Intraklassenkorrelationskoeffizient, cm = Zentimeter, SZ = Studienzentren, Mitt. Diff. = mittlere Differenz.

3.3 Bewertung der US-Bilder

3.3.1 Stichprobe

Für die Bewertung der US-Bilder wurde nach dem in Abbildung 24 angegebenen Schema eine Stichprobe von 981 Teilnehmern gezogen, die von insgesamt 74 Untersuchern an 18 SZ untersucht wurden. Die Stichprobe umfasste damit 4,8% der Studienteilnehmer des Data Freeze 100k mit vollständigen Untersuchungen ($n = 20.390$, vergleiche Tabelle 10).

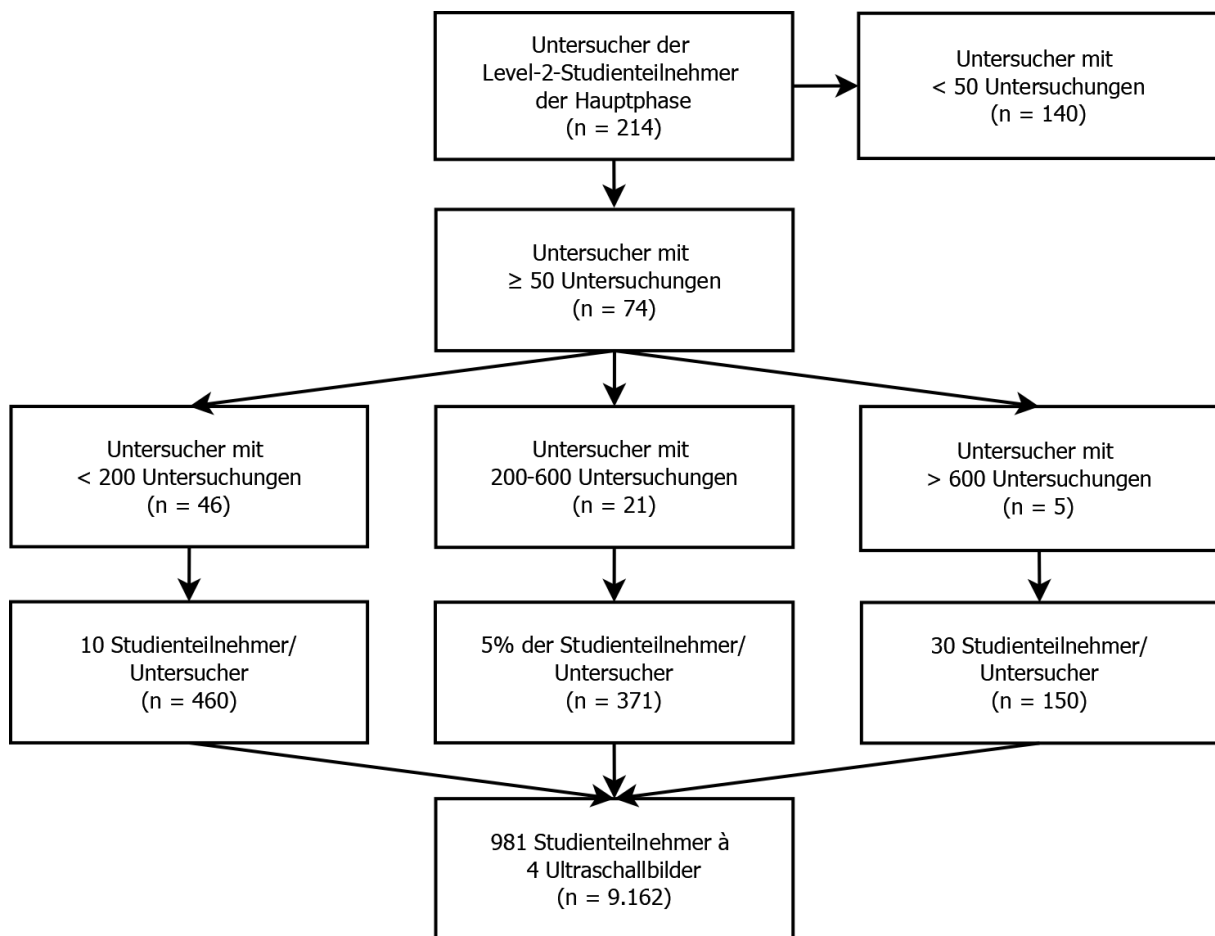


Abbildung 24: Flowchart zur Stichprobe für die Bewertung der US-Bilder. US = Ultraschall, n = Anzahl.

Die Studienteilnehmer der Stichprobe für die US-Bildbewertung wurden zwischen 06.10.2014 und 15.03.2017 untersucht. 50,2% der 981 Studienteilnehmer der Stichprobe waren männlich ($n = 492$), 49,8% waren weiblich ($n = 489$). Das Durchschnittsalter lag bei 51,2 Jahren und erstreckte sich von 21 Jahren bis 72 Jahren. Der Mittelwert des SAT lag bei den Männern der Stichprobe bei 2,0 cm, bei den Frauen bei 2,2 cm. Das durchschnittliche VAT bei den Männern der Stichprobe betrug 7,5 cm, das

der Frauen 5,6 cm (vergleiche Tabelle 22). Diese Charakteristika der Stichprobe waren denen der Studienteilnehmer des Data Freeze 100k ähnlich (vergleiche 3.2.4).

Tabelle 22: Charakteristika der Studienteilnehmer der Stichprobe für die Bewertung der US-Bilder.

	Gesamt (N = 981)		Männer (n = 492)		Frauen (n = 489)	
	Mean	SD	Mean	SD	Mean	SD
Alter [Jahre]	51,2	12,3	51,7	12,1	50,7	12,6
SAT [cm]	2,1	0,8	2,0	0,7	2,2	0,9
VAT [cm]	6,5	2,5	7,5	2,4	5,6	2,2

US = Ultraschall, N = Anzahl, Mean = Mittelwert, SD = Standardabweichung, SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe, cm = Zentimeter.

3.3.2 Analyse der US-Bildbewertung

Pro Studienteilnehmer der Stichprobe ($n = 981$) lagen jeweils zwei US-Bilder der Messung des SAT ($n = 1.962$) und zwei US-Bilder der Messung des VAT ($n = 1.962$) vor.

Tabelle 23: Bewertung der Messrichtigkeit in Abhängigkeit des Geschlechts und Alters der Studienteilnehmer, der Menge des SAT und VAT und des SZ (Anzahl der Untersucher in Klammern).

	US-Bilder	Messrichtigkeit SAT				Messrichtigkeit VAT			
		ML richtig	MP1 richtig	MP2 richtig	ML, MP1, MP2 richtig	ML richtig	MP1 richtig	MP2 richtig	ML, MP1, MP2 richtig
	N	%	%	%	%	%	%	%	%
Gesamt	1962	99,2	93,4	93,2	86,3	97,6	80,0	74,9	59,7
Geschlecht									
männlich	984	99,6	94,9	92,6	87,5	97,9	79,1	71,9	57,1
weiblich	978	98,8	91,8	93,8	85,1	97,3	81,0	78,0	62,4
Alter [Jahre]									
20 – 29	154	99,4	94,2	96,1	90,3	99,4	76,6	81,8	59,7
30 – 39	186	98,9	93,0	95,2	87,6	98,9	78,0	83,3	66,7
40 – 49	434	98,9	91,2	91,7	83,6	98,9	78,8	80,0	63,1
50 – 59	612	99,5	95,1	94,3	89,1	99,5	81,7	73,5	60,5
>60	576	99,1	93,1	91,7	83,9	99,1	80,7	68,1	54,2
SAT [cm]									
<1,5	527	98,5	92,6	91,3	83,1	-	-	-	-
≥1,5 – <2	550	99,1	93,6	93,1	86,2	-	-	-	-
≥2 – <2,5	395	99,8	92,7	93,9	86,8	-	-	-	-
≥2,5	490	99,6	94,5	94,7	89,4	-	-	-	-
VAT [cm]									
<4	296	-	-	-	-	97,3	77,4	80,7	60,8
≥4 – <6	615	-	-	-	-	96,4	79,5	76,3	61,0
≥6 – <8	551	-	-	-	-	98,7	78,6	77,1	61,3
≥8	500	-	-	-	-	98,0	83,8	67,4	55,8
SZ (74)									
Augsburg (3)	140	100,0	100,0	100,0	100,0	100,0	92,1	82,1	77,1
Regensburg (4)	92	100,0	98,9	98,9	97,8	100,0	85,9	75,0	63,0
Heidelberg/ Mannheim (11)	252	97,6	98,8	88,1	84,5	85,7	59,9	72,2	40,1
Freiburg (6)	120	99,2	98,3	94,2	92,5	100,0	84,2	60,8	51,7
Saarbrücken (1)	40	100,0	100,0	95,0	95,0	85,0	100,0	92,5	85,0
Essen (1)	50	100,0	98,0	90,0	88,0	98,0	38,0	68,0	34,0
Münster (3)	62	98,4	77,4	93,6	69,4	100,0	96,8	61,3	59,7
Düsseldorf (3)	60	93,3	93,3	93,3	85,0	100,0	68,3	73,3	60,0
Halle (3)	92	100,0	95,7	97,8	94,6	100,0	95,7	76,1	71,7
Leipzig (3)	116	100,0	97,4	93,1	90,5	100,0	92,2	81,9	75,9
Berlin Nord (4)	104	99,0	34,6	100,0	34,6	100,0	92,3	80,8	73,1

	US-Bilder	Messrichtigkeit SAT				Messrichtigkeit VAT			
		ML richtig	MP1 richtig	MP2 richtig	ML, MP1, MP2 richtig	ML richtig	MP1 richtig	MP2 richtig	ML, MP1, MP2 richtig
SZ	N	%	%	%	%	%	%	%	%
Berlin Mitte (3)	82	100,0	93,9	85,4	80,5	100,0	85,4	76,8	64,6
Berlin Süd/ Brandenburg (5)	100	100,0	94,0	93,0	87,0	100,0	73,0	65,0	51,0
Braunschweig/ Hannover (4)	104	97,1	96,2	92,3	86,5	99,0	70,2	78,9	53,9
Hamburg (4)	80	100,0	98,8	88,8	87,5	98,8	47,5	93,8	45,0
Bremen (4)	104	100,0	87,5	91,4	81,7	100,0	72,1	88,5	61,5
Kiel (3)	60	100,0	98,3	96,7	95,0	100,0	85,0	73,3	66,7
Neubrandenburg (9)	304	100,0	100,0	92,1	92,1	99,3	91,8	68,4	62,2

SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe, SZ = Studienzentrum, US = Ultraschall, ML = Messlinie, MP1 = oberer Messpunkt, MP2 = unterer Messpunkt, N = Anzahl, cm = Zentimeter.

Um auf die Qualität der US-Messung schließen zu können, wurde die Messrichtigkeit analysiert. Dafür wurde kontrolliert, ob die Messpunkte richtig gesetzt wurden und ob die Messlinie senkrecht verlief (vergleiche Tabelle 23). Sowohl bei der Messung des SAT (99,2%) als auch bei der Messung des VAT (97,6%) wurde die Messlinie fast immer senkrecht ausgerichtet. Der obere Messpunkt des SAT an der Bauchhaut wurde in 93,4% der Fälle korrekt positioniert. In 6,1% der Fälle ($n = 120$) wurde der obere Messpunkt fälschlicherweise unterhalb der Bauchhaut positioniert und damit das SAT zu kurz gemessen (siehe Abbildung 25). Der untere Messpunkt des SAT am Oberrand der Linea alba wurde in 93,2% der Fälle korrekt positioniert. In 2,1% der Fälle ($n = 42$) wurde statt zum Oberrand zum Unterrand der Linea alba gemessen, was einen zu großen Messwert ergab. Teilweise wurde die Linea alba nicht korrekt identifiziert. So wurde in 1,5% der Fälle ($n = 29$) zum Peritoneum gemessen, woraus ein zu großer Messwert resultierte. In 1,6% der Fälle ($n = 31$) wurde der untere Messpunkt an einem Reflex im SAT positioniert, weshalb zu kurz gemessen wurde. In 0,6% der Fälle ($n = 11$) wurde fälschlicherweise zum Bauchmuskel gemessen.

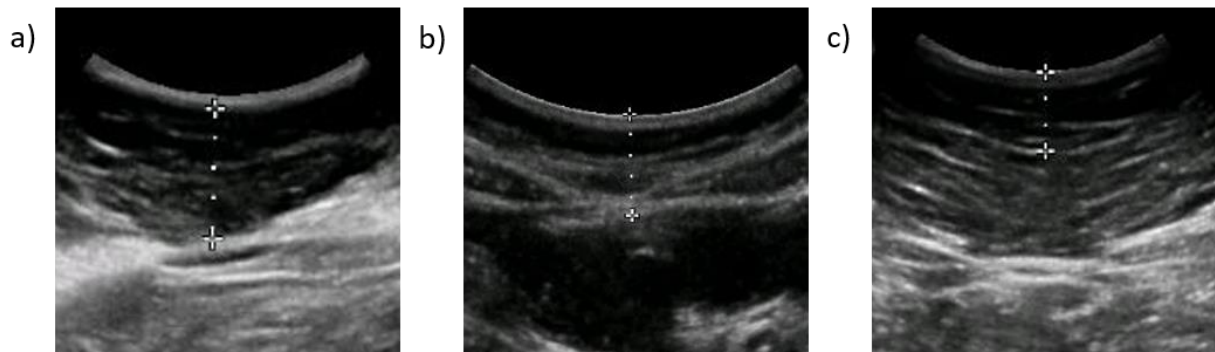


Abbildung 25: Fehler bei der Messung des SAT: a) Positionierung des oberen Messpunkts unterhalb der Bauchhaut. b) Positionierung des unteren Messpunkts unterhalb der Linea alba. c) Positionierung des unteren Messpunkts an einem Reflex im SAT.
SAT = subkutanes Fettgewebe.

Der obere Messpunkt des VAT am Unterrand der Linea alba wurde in 80,0% der Fälle korrekt positioniert. In 10,7% der Fälle ($n = 210$) wurde das VAT fälschlicherweise vom Oberrand statt vom Unterrand der Linea alba gemessen, woraus ein zu großer Messwert resultierte (siehe Abbildung 26). In 5,1% der Fälle ($n = 101$) wurde vom Peritoneum statt von der Linea alba gemessen, weshalb der Messwert zu klein war. Wenn der obere Messpunkt fälschlicherweise im Bereich des SAT positioniert wurde (2,1%, $n = 41$), wurde eine zu lange Strecke für das VAT gemessen. In weiteren 2,0% der Fälle waren andere Fehler aufgetreten ($n = 40$).

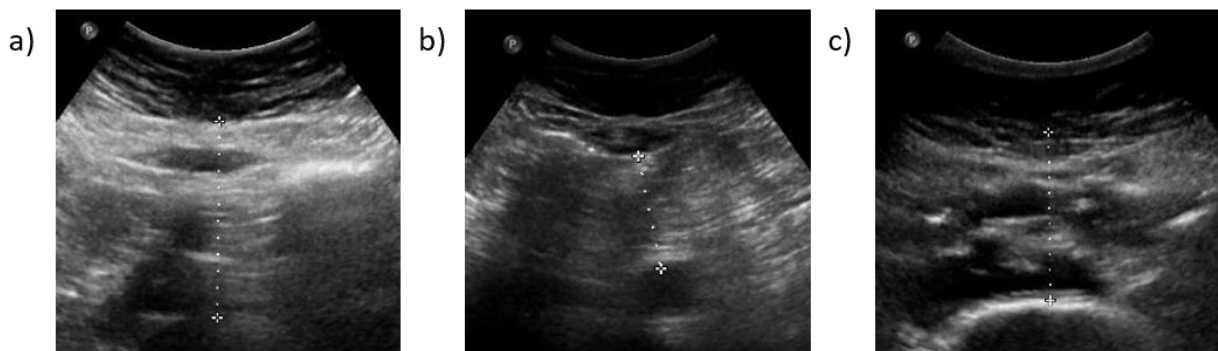


Abbildung 26: Fehler bei der Positionierung des oberen Messpunktes des VAT: a) Positionierung am Oberrand der Linea alba. b) Positionierung am Peritoneum. c) Positionierung im Bereich des SAT.
VAT = viszerales Fettgewebe, SAT = subkutanes Fettgewebe.

Die korrekte Positionierung des unteren Messpunktes des VAT an der Vorderkante des Wirbelkörpers gelang in 74,9% der Fälle. Teils war der Messpunkt falsch positioniert, sodass beispielsweise zum Ober- oder Unterrand der Aorta (6,1%, $n = 120$), zu kleineren Gefäßen oberhalb der Aorta (0,7%, $n = 14$) oder zu Darmschlingen gemessen wurde (1,2%, $n = 24$), woraus ein zu kleiner Messwert resultierte (siehe

Abbildung 27). Gelegentlich wurde auch statt zur Vorderkante mittig oder seitlich an den Wirbelkörper gemessen (2,1%, $n = 42$), weshalb der Messwert zu groß war. In 12,6% der Fälle ($n = 248$) waren die zur Orientierung dienenden Strukturen Aorta, Vena cava und der Wirbelkörper auf dem US-Bild nicht identifizierbar. Wenn dies der Fall war, wurde die Messung ebenfalls als nicht korrekt gewertet. Andere Fehler waren in 2,2% der Fälle aufgetreten ($n = 44$).

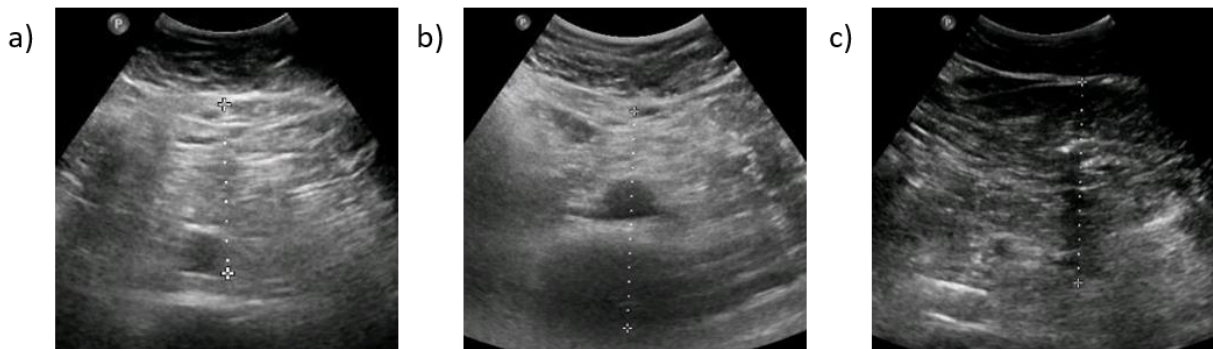


Abbildung 27: Fehler bei der Positionierung des unteren Messpunktes des VAT: a) Positionierung am Unterrand der Aorta. b) Positionierung mittig im Wirbelkörper. c) Wirbelkörper und Aorta nicht identifizierbar. VAT = viszerales Fettgewebe.

Darüber hinaus wurde berechnet, zu welchem Anteil das SAT beziehungsweise das VAT nach Maßstäben der Messrichtigkeit komplett korrekt gemessen wurde. Dieser Anteil lag bei der Messung des SAT bei 86,3% ($n = 1.693$) und bei der Messung des VAT bei 59,7% ($n = 1.172$), bezogen auf 1.962 US-Bilder. Für jeden bewerteten Untersucher wurde der Anteil komplett korrekt gemessener Studienteilnehmer dem bereinigten Data Freeze 100k-Datensatz hinzugefügt, um die Einschätzung der Datenqualität bei der Verwendung der Daten im Rahmen von zukünftigen wissenschaftlichen Analysen zu ermöglichen.

In einem logistischen Regressionsmodell wurde der Einfluss der Variablen Geschlecht und Alter der Studienteilnehmer, Menge des gemessenen SAT beziehungsweise VAT und SZ auf die Bewertung der Kriterien der Messrichtigkeit getestet (vergleiche Tabelle 24). Dabei zeigte sich, dass die Menge des SAT und das Alter der Studienteilnehmer einen signifikanten Einfluss auf die Positionierung des unteren Messpunktes bei der Messung des SAT hatte. Je mehr SAT und je jünger die Studienteilnehmer, desto eher wurde der untere Messpunkt korrekt gesetzt. Darüber hinaus wurden auch signifikant häufiger junge Studienteilnehmer sowie Studienteilnehmer mit viel SAT Fettgewebe komplett korrekt vermessen. Bei der Messung des VAT hatten sowohl das Alter der

Studienteilnehmer als auch die Menge des gemessenen VAT einen statistisch signifikanten Einfluss auf die Bewertung des unteren Messpunkts. Je älter die Studienteilnehmer und je mehr VAT, desto schlechter wurde die Positionierung des unteren Messpunkts des VAT bewertet. Dasselbe galt für komplett korrekt vermessenen Studienteilnehmer.

Tabelle 24: Logistische Regressionsmodelle zum Einfluss des Geschlechts und Alters der Studienteilnehmer, der Menge des SAT und VAT und des SZ auf die Bewertungskriterien der Messrichtigkeit.

Abhängige Variablen \ Prädiktive Variablen	Messrichtigkeit SAT				Messrichtigkeit VAT			
	ML	MP1	MP2	ML, MP1, MP2	ML	MP1	MP2	ML, MP1, MP2
	<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>
Geschlecht	0,02	0,09	0,83	0,19	0,34	0,53	0,17	0,49
Alter	0,75	0,97	0,03	0,03	0,36	0,25	0,00	0,02
SAT	0,29	0,82	0,02	0,01	-	-	-	-
VAT	-	-	-	-	0,57	0,77	0,00	0,01
SZ	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe, SZ = Studienzentrum, ML = Messlinie, MP1 = oberer Messpunkt, MP2 = unterer Messpunkt, *p* = *p*-Wert (fett gedruckt, wenn $p < 0,05$).

Das SZ, an dem die Untersuchungen durchgeführt wurden, hatte einen signifikanten Einfluss auf die Bewertung aller Kriterien der Messrichtigkeit. Das SAT wurden im SZ Berlin Nord mit 34,6% ($n = 36$) am seltensten korrekt gemessen. Bei der Sichtung der bewerteten US-Bilder in diesem SZ zeigte sich, dass alle vier bewerteten Untersucher bei der Messung des SAT den oberen Messpunkt häufig unterhalb der Bauchhaut positioniert hatten, sodass eine zu kurze Messstrecke resultierte. Dasselbe traf für zwei von drei bewerteten Untersuchern im SZ Münster zu. Im SZ Augsburg wurde dagegen das SAT bei allen 140 bewerteten US-Bildern korrekt gemessen. Bei der Messung des VAT wurden die wenigsten korrekten US-Messungen im SZ Essen mit 34,0% ($n = 17$) sowie im SZ Heidelberg/Mannheim mit 40,1% ($n = 101$) durchgeführt. In den SZ Saarbrücken und Heidelberg/Mannheim wurde die Messlinie mit 85,0% beziehungsweise 85,7% seltener senkrecht positioniert als im SZ-übergreifenden Durchschnitt mit 97,6%. Der obere Messpunkt des VAT wurde in den SZ Essen (38,0%), Hamburg (47,5%) und Heidelberg/Mannheim (59,9%) deutlich seltener korrekt positioniert als im SZ-übergreifenden Durchschnitt (80,0%). In den drei genannten SZ wurde der obere Messpunkt des VAT häufig am Oberrand der Linea alba oder am Peritoneum positioniert. Im SZ Freiburg wurde der untere Messpunkt des VAT mit 60,8% weit seltener korrekt positioniert als im Durchschnitt (74,9%). Dort waren die Zielstrukturen Wirbelkörper und Aorta bei den US-Bildern von fünf der sechs Untersucher häufig nicht sicher erkennbar. Die anteilig meisten korrekten US-

Messungen des VAT wurden im SZ Saarbrücken durchgeführt. Dort wurden 85,0% der bewerteten US-Bilder ($n = 34$) korrekt vermessen.

Tabelle 25: Bewertung der Bildqualität in Abhängigkeit des Geschlechts und Alters der Studienteilnehmer, der Menge des SAT und VAT und des SZ (Anzahl der bewerteten Untersucher in Klammern).

	US-Bilder	Bildqualität SAT				Bildqualität VAT			
		HE Richtig %	FO richtig %	LA richtig %	HE, FO, LA richtig %	HE richtig %	FO richtig %	LA richtig %	HE, FO, LA richtig %
Gesamt	N								
	1962	97,6	36,7	95,9	35,6	81,9	85,0	93,2	64,0
Geschlecht									
männlich	984	97,0	33,8	94,6	32,2	82,4	92,9	90,4	69,2
weiblich	978	98,3	39,5	97,2	39,1	81,3	77,0	96,0	58,7
Alter [Jahre]									
20 – 29	154	95,5	37,7	93,5	35,1	95,5	77,3	91,6	61,0
30 – 39	186	95,7	45,7	98,4	44,6	95,7	82,8	94,6	63,4
40 – 49	434	97,2	36,9	97,5	35,7	97,2	81,6	94,0	60,8
50 – 59	612	98,5	34,5	96,7	34,2	98,5	88,1	92,8	67,8
>60	576	98,1	35,6	93,8	34,4	98,1	87,0	92,9	63,2
SAT [cm]									
<1,5	527	94,3	28,8	93,9	26,8	-	-	-	-
≥1,5 – <2	550	98,0	31,8	95,8	31,1	-	-	-	-
≥2 – <2,5	395	98,5	33,9	96,7	32,9	-	-	-	-
≥2,5	490	100,0	52,7	97,6	52,5	-	-	-	-
VAT [cm]									
<4	296	-	-	-	-	87,8	53,4	97,0	46,3
≥4 – <6	615	-	-	-	-	85,5	82,3	92,5	65,2
≥6 – <8	551	-	-	-	-	81,1	93,1	93,5	70,1
≥8	500	-	-	-	-	74,6	98,0	91,4	66,2
SZ (74)									
Augsburg (3)	140	100,0	18,6	97,1	18,6	85,0	92,9	96,4	75,7
Regensburg (4)	92	100,0	29,4	92,4	28,3	90,2	88,0	90,2	71,7
Heidelberg/ Mannheim (11)	252	94,8	6,4	92,5	6,4	86,9	63,1	88,5	46,4
Freiburg (6)	120	100,0	3,3	97,5	3,3	85,0	88,3	97,5	73,3
Saarbrücken (1)	40	100,0	100,0	100,0	100,0	90,0	100,0	100,0	90,0
Essen (1)	50	100,0	20,0	88,0	18,0	38,0	88,0	82,0	26,0
Münster (3)	62	93,6	0,0	96,8	0,0	88,7	80,7	95,2	66,1
Düsseldorf (3)	60	98,3	0,0	98,3	0,0	85,0	65,0	96,7	58,3
Halle (3)	92	95,7	77,2	96,7	72,8	75,0	97,8	96,7	69,6
Leipzig (3)	116	100,0	43,1	98,3	41,4	83,6	98,3	95,7	77,6
Berlin Nord (4)	104	86,5	15,4	87,5	13,5	82,7	78,9	87,5	55,8
Berlin Mitte (3)	82	100,0	29,3	97,6	29,3	91,5	93,9	100,0	86,6
Berlin Süd/ Brandenburg (5)	100	98,0	2,0	92,0	2,0	82,0	86,0	88,0	59,0
Braunschweig/ Hannover (4)	104	100,0	15,4	97,1	15,4	91,4	57,7	92,3	46,2
Hamburg (4)	80	100,0	65,0	100,0	65,0	90,0	95,0	91,3	76,3
Bremen (4)	104	99,0	85,6	100,0	85,6	87,5	91,4	92,3	76,0
Kiel (3)	60	100,0	33,3	98,3	33,3	73,3	70,0	91,7	43,3
Neubrandenburg (9)	304	97,4	84,2	98,0	80,9	69,4	97,4	95,7	64,8

SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe, SZ = Studienzentrum, US = Ultraschall, HE = Helligkeit, FO = Format, LA = Linea alba, N = Anzahl, cm = Zentimeter.

Zusätzlich zur Messrichtigkeit wurde die Bildqualität bewertet, um die Präzision der Messwerte abschätzen zu können (vergleiche Tabelle 25). Die Helligkeit lag bei 97,6% der US-Bilder mit Messungen des SAT im optimalen Bereich mit mittlerer Helligkeit an

der Oberfläche und in der Tiefe, in 1,8% der Fälle ($n = 36$) war das US-Bild zu dunkel. Bei der Messung des VAT war die Helligkeit bei 81,9% der US-Bilder zufriedenstellend. In 14,7% der Fälle ($n = 288$) waren die US-Bilder bei der Messung des VAT in der Tiefe zu dunkel, was zusätzlich die Identifizierung des Wirbelkörpers auf dem US-Bild erschwerte (siehe Abbildung 28). Die formatfüllende Aufnahme der Zielstrukturen stellte eine weitere Kategorie bei der Bewertung der Bildqualität da, wenngleich sie in früheren Versionen der SOP noch nicht detailliert erklärt wurde (63). Bei der Messung des SAT wurden 36,7%, bei der Messung des VAT 85,0% der bewerteten US-Bilder formatfüllend aufgenommen. Dabei wurde sowohl bei der Messung des SAT als auch bei der Messung des VAT zu wenig gezoomt, sodass die interessierenden Strukturen zu klein zur Darstellung kamen. Als weiteres Merkmal der Bildqualität wurde die Position der Linea alba betrachtet. Diese befand sich bei 95,9% der Messungen des SAT und 93,2% der Messungen des VAT wie gewünscht mittig im Bild. Alle drei Kriterien der Bildqualität wurden bei 35,6% der US-Messungen des SAT und 64,0% der US-Messungen des VAT erfüllt.

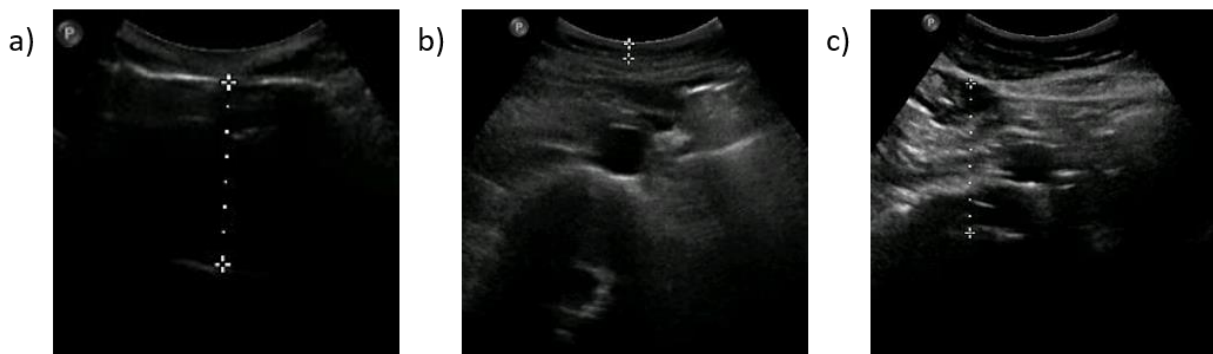


Abbildung 28: Mangelhafte Bildqualität bei der Messung des SAT und VAT: a) in der Bildtiefe zu dunkel. b) zu kleine Darstellung der Zielstrukturen SAT und Linea alba. c) Linea alba zu weit links im Bild. SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe.

Auch auf die Bewertung der Kriterien der Bildqualität wurde in einem logistischen Regressionsmodell der Einfluss verschiedener Variablen gemessen (vergleiche Tabelle 26). Auf die formatfüllende Aufnahme des SAT wirkte sich insbesondere die Menge des SAT signifikant aus. Bei Studienteilnehmern mit wenig SAT wurde das US-Bild seltener formatfüllend aufgenommen als bei Studienteilnehmern mit viel SAT. Bei der Messung des VAT gelang die Einstellung der Helligkeit bei viel Fettgewebe schlechter als bei wenig Fettgewebe. Die US-Bilder des VAT wurden bei männlichen Studienteilnehmern sowie Studienteilnehmern mit viel VAT signifikant häufiger formatfüllend aufgenommen als bei weiblichen sowie Studienteilnehmern mit wenig

VAT. Insgesamt wurde die Bildqualität der Messung des SAT bei weiblichen Studienteilnehmern, die der Messung des VAT bei männlichen Studienteilnehmern besser bewertet.

Tabelle 26: Logistische Regressionsmodelle zum Einfluss des Geschlechts und Alters der Studienteilnehmer, der Menge des SAT und VAT und des SZ auf die Bewertungskriterien der Bildqualität.

Abhängige Variablen	Bildqualität SAT				Bildqualität VAT			
	HE	FO	LA	HE, FO, LA	HE	FO	LA	HE, FO, LA
Prädiktive Variablen	<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>
Geschlecht	0,01	0,02	0,00	0,00	0,00	0,00	0,00	0,00
Alter	0,01	0,12	0,06	0,20	0,26	0,56	0,92	0,44
SAT	0,00	0,00	0,00	0,00	-	-	-	-
VAT	-	-	-	-	0,00	0,00	0,19	0,39
SZ	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe, SZ = Studienzentrum, HE = Helligkeit, FO = Format, LA = Linea alba, *p* = *p*-Wert (fett gedruckt, wenn $p < 0,05$).

Das SZ, an dem die US-Bilder aufgenommen wurden, hatte einen signifikanten Einfluss auf alle drei Kriterien der Bildqualität. Die qualitativ hochwertigsten US-Bilder wurden im SZ Saarbrücken aufgenommen, dort führte im Zeitraum des Data Freeze 100k lediglich ein Untersucher die US-Untersuchung durch. Die US-Bilder des SAT wurden in den SZ Augsburg, Heidelberg/Mannheim, Freiburg, Essen, Münster, Düsseldorf, Berlin Nord, Berlin Süd/Brandenburg und Braunschweig/Hannover mit Anteilen von $\leq 20\%$ deutlich seltener formatfüllend aufgenommen als im Durchschnitt (36,7%). Bei der Messung des VAT wurden die US-Bilder in den SZ Braunschweig/Hannover (57,7%), Heidelberg/Mannheim (63,1%) und Düsseldorf (65,0%) seltener formatfüllend aufgenommen als im SZ-übergreifenden Vergleich (85,0%). Insbesondere die US-Bilder im SZ Essen waren bei der Messung des VAT in der Tiefe zu dunkel.

3.3.3 Rückmeldung an die Untersucher

Um die Qualität der US-Messung des Abdominalfetts zu verbessern, wurde für jeden bewerteten Untersucher eine Rückmeldung basierend auf der Bewertung der US-Bilder erstellt (vergleiche Beispiel im Anhang). Diese umfasste die Bewertung der einzelnen Kategorien der Messrichtigkeit und Bildqualität in Schulnoten sowie individuelle Verbesserungsvorschläge. Sobald ein Untersucher eine der Kategorien der Messrichtigkeit bei weniger als 75% der bewerteten US-Bilder korrekt durchgeführt hatte,

wurde ihm eine Nachschulung empfohlen. Dies betraf 63,5% der bewerteten Untersucher ($n = 47$). Zur Veranschaulichung wurden folgende Beispielbilder eingefügt (vergleiche Abbildung 29 und Abbildung 30).

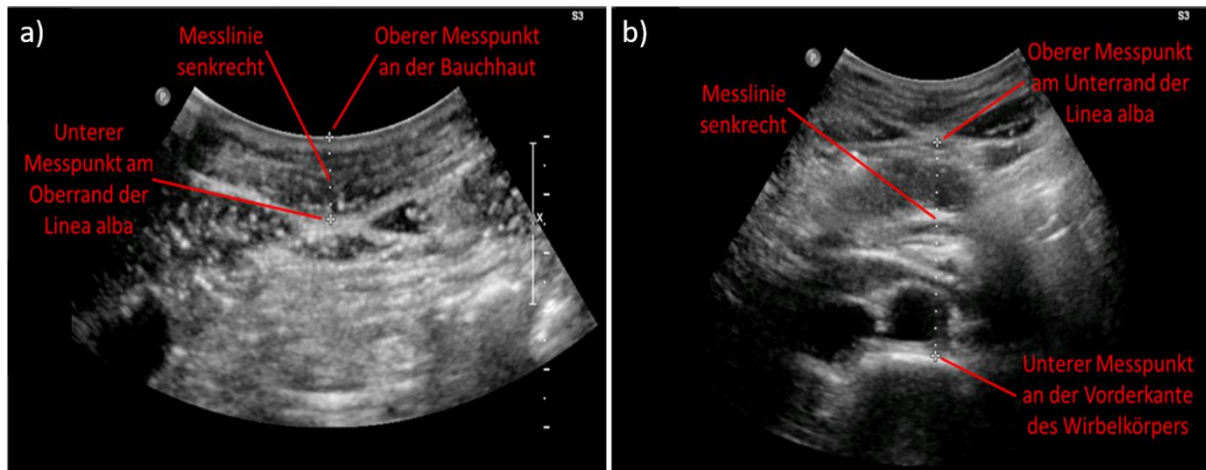


Abbildung 29: Beispielbilder zur Messrichtigkeit bei der Messung von a) SAT und b) VAT. SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe.

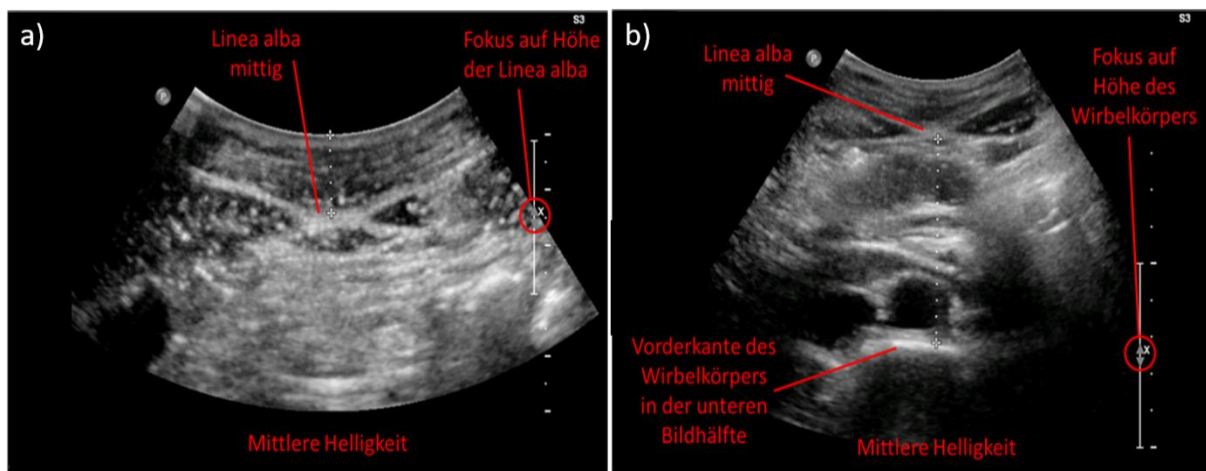


Abbildung 30: Beispielbilder zur Bildqualität bei der Messung von a) SAT und b) VAT. SAT = subkutanes Fettgewebe, VAT = viszerales Fettgewebe.

3.4 Qualitätsbericht und Empfehlungen zur US-Messung

Der Qualitätsbericht zur Datenbereinigung im Rahmen des Data Freeze 100.000 des Moduls US Abdominalfett wurde im März 2019 bei dem Integrationszentrum der NAKO Gesundheitsstudie eingereicht. Zusätzlich wurden allen SZ der NAKO Gesundheitsstudie Empfehlungen zur Durchführung der US-Messung zur Verfügung gestellt und konnten ab Januar 2019 für die Schulung der US-Untersuchung genutzt werden. Letztere befinden sich im Anhang.

4 Diskussion

4.1 Überarbeitung der SOP „US Abdominalfett“

Für die Überarbeitung der SOP „US Abdominalfett“ wurde die zweite Version der SOP mit aktuellen Übersichtsarbeiten zur Methodik der US-Messung des Abdominalfetts sowie Empfehlungen aus einem Sonographie-Lehrbuch verglichen. Dabei konnte gezeigt werden, dass die SOP bereits einige Kriterien zur Aufnahme eines qualitativ hochwertigen US-Bildes erfüllte. Im Gegensatz zu anderen Studien (50) wurde das SAT im Rahmen der NAKO Gesundheitsstudie einschließlich der Bauchhaut gemessen. Da die Dicke der Dermis am Körperstamm bis zu 0,4 cm beträgt (83), sollte dieser Unterschied bei dem Vergleich mit Messwerten anderer Studien nicht außer Acht gelassen werden. Für die Messung des VAT schien sich bisher keine standardisierte Methodik in der Literatur etabliert zu haben, sodass sich sowohl die Messposition als auch die Messpunkte in anderen Kohortenstudien unterschieden (42, 45, 49, 81, 82). Hier könnte die Vergleichbarkeit der Messwerte erschwert sein.

Zur weiteren Verbesserung der Bildqualität der US-Bilder wurde die Optimierung der Eindringtiefe sowie die Verwendung des Fokus in der überarbeiteten SOP eingeführt. Um die bestmögliche Momentaufnahme zu erhalten, wurde das Vorspulen des US-Bildes erklärt. Zudem wurde eine Fotografie der Messregler des US-Geräts sowie Abbildungen der der Messung zugrundeliegenden anatomischen Strukturen und neue Beispiel-US-Bilder hinzugefügt. Dadurch sollte die Verständlichkeit der SOP für das paramedizinische Studienpersonal weiter verbessert werden. Die Überarbeitung der SOP fand während der Basisuntersuchung der NAKO Gesundheitsstudie statt. Deshalb könnte die Änderung der Methodik bei der Aufnahme der US-Bilder in einer besseren Bildqualität, jedoch schlechteren Vergleichbarkeit der US-Messwerte resultieren. Um dessen Ausmaß zu quantifizieren, wäre ein Vergleich der Bildqualität und Messergebnisse von Stichproben vor und nach Einführung der dritten, überarbeiteten Version der SOP sinnvoll. Aus organisatorischen Gründen konnte zudem zum Zeitpunkt der Überarbeitung der SOP eine Vorbereitung des Studienteilnehmers im Sinne von Verzicht auf größere Mahlzeiten und kohlenensäurehaltige Getränke (41) nicht bewerkstelligt werden. Dies hätte durch Verminderung der Darmgasüberlagerung einen positiven Einfluss auf die Qualität der Messwerte des VAT genommen (71). Da im Rahmen der NAKO Gesundheitsstudie die US-Methode zur Messung des

Abdominalfetts erstmalig bei einer fünfstelligen Probandenzahl angewendet wurde, galt es, die Methodik anhand der gewonnenen Erfahrungen ständig zu verbessern, ohne die Vergleichbarkeit der Ergebnisse außer Acht zu lassen.

4.2 Datenaufbereitung und -analyse des Data Freeze 100k-Datensatzes

Die Datenaufbereitung und -analyse wurde anhand der Anforderungen in der SOP „Vorgaben für die Datenbereinigung im Rahmen des Data Freeze 100.000“ (62) durchgeführt. Dabei konnte gezeigt werden, dass 73,4% der vorgesehenen US-Untersuchungen des Abdominalfetts vollständig durchgeführt wurden, wobei es beträchtliche Unterschiede zwischen den 18 SZ gab. In den SZ Saarbrücken und Essen wurden lediglich 21,8% beziehungsweise 24,1% der vorgesehenen US-Untersuchungen vollständig durchgeführt, wobei jeweils nur ein Untersucher für das Gros der US-Untersuchungen verantwortlich war, während es in den anderen SZ mindestens drei Untersucher waren. Das deutet auf ein strukturelles Problem hin und lässt schlussfolgern, dass in jedem SZ mehr als ein Untersucher für die US-Untersuchung des Abdominalfetts geschult werden sollte, um die Zielzahlen zu erreichen. In den SZ mit einer Vollständigkeit von über 90% führten drei bis sechs Untersucher den Großteil der US-Untersuchungen durch. Die meisten der nicht durchgeführten US-Untersuchungen waren in der Erhebungssoftware nicht dokumentiert, weshalb der Grund für die Nichtdurchführbarkeit nicht ersichtlich war. Dies erschwerte die Ableitung von Konsequenzen mit dem Ziel, die Vollständigkeit der US-Untersuchungen zukünftig zu verbessern. Bei den dokumentierten Gründen stellten Personalmangel und organisatorische Gründe die wichtigsten Ursachen für nicht durchgeführte oder unvollständig abgebrochene Untersuchungen dar. Erwartungsgemäß selten wurden medizinische Gründe und Kontraindikationen als Grund festgehalten. Hier wären lediglich eine Schwangerschaft ab der 15. Schwangerschaftswoche und Schwierigkeiten bei der Lagerung des Studienteilnehmers in Rückenlage denkbar (63).

Bei der Prüfung der Plausibilität der US-Messwerte wurden 0,4% der Messwerte des SAT und 1,3% der Messwerte des VAT als nicht plausibel eingestuft und korrigiert. Unplausible Messwerte des VAT wurden dabei meist aufgrund der eingetragenen Werte 0,000 und 999,9 als fehlende Messwerte eingestuft. Diese offensichtlich unplausiblen Werte wurden von den Untersuchern in der Erhebungssoftware dokumentiert, wenn aufgrund von schwierigen Untersuchungsbedingungen kein Messwert erhoben

werden konnte. Bei unplausiblen Messwerten des SAT konnten hingegen häufig Messwerte aus den Kommentaren zur Messung oder Durchführbarkeit der Messung übernommen werden. Diese Messwerte lagen meist jenseits der Plausibilitätsgrenzen, wie sie zu Beginn der Hauptphase der NAKO Gesundheitsstudie in der Erhebungssoftware hinterlegt waren (vergleiche Tabelle 5) und die Dokumentation dieser Messwerte verhinderte. Gemäß den US-Bildern waren diese Messwerte jedoch plausibel. Daraus ließ sich interpretieren, dass die Plausibilitätsgrenzen des SAT zu streng gesetzt waren. Deshalb wurden die weichen Plausibilitätsgrenzen gelockert und die harten Plausibilitätsgrenzen abgeschafft, sodass auch extreme Messwerte eingetragen werden konnten (84).

Bei 21.175 dokumentierter US-Untersuchungen des Abdominalfetts sollten 84.700 Messwerte vorliegen, da sowohl SAT als auch VAT doppelt gemessen wurde. Prozentual fehlten mehr Messwerte des VAT (5,1%) als Messwerte des SAT (3,9%). Eine mögliche Erklärung wäre die teils erschwerte Darstellung und Messung des VAT aufgrund von Darmgasüberlagerungen. Der Anteil fehlender Werte erfüllte lediglich bei der Messung des SAT die Vorgabe aus der SOP zur Qualitätssicherung der US-Messung Abdominalfett und lag unter 5% (84). Zwischen den SZ gab es große Unterschiede bei dem Anteil fehlender Werte. Während im SZ Augsburg nur 0,7% der Messwerte fehlten, waren es im SZ Braunschweig/Hannover 25,1%. Hier fiel auf, dass manche SZ wie beispielsweise Braunschweig/Hannover und Berlin Mitte viele Untersuchungen unvollständig oder nicht durchgeführt hatten, während andere SZ wie Saarbrücken und Essen einen Großteil der Untersuchungen nicht dokumentiert hatten und diese deshalb nicht bei den fehlenden Messwerten erschienen. Das legt die Vermutung nahe, dass die SZ bei der Dokumentation uneinheitlich verfahren. Zukünftig wäre eine konsequente Dokumentation auch der nicht durchgeführten Untersuchungen wünschenswert, die durch eine klare Beschreibung der Vorgehensweise in einer SOP zur Dokumentation erleichtert werden könnte.

Die bereits im Bundesgesundheitsblatt publizierte Werteverteilung zeigte, dass die Männer des Data Freeze 100k der NAKO Gesundheitsstudie mit 2,0 cm weniger SAT besaßen als die Frauen mit 2,3 cm. Dahingegen war das VAT der Männer mit durchschnittlich 7,8 cm in allen Altersgruppen stärker ausgeprägt als das der Frauen mit 5,8 cm (74). Das Verhältnis von SAT zu VAT verlagerte sich bei beiden Geschlechtern mit zunehmendem Alter auf die Seite des VAT. Diese Fettumverteilung konnte auch schon

in zahlreichen Studien beobachtet werden, die SAT und VAT computertomographisch als Fläche maßen (12). Bei Frauen könnte dafür der stärkere Einfluss der Androgene bei abnehmenden Östrogenen nach der Menopause ursächlich sein (14, 85). Dabei sollte beachtet werden, dass es sich bei den vorliegenden Daten der NAKO Gesundheitsstudie nicht um longitudinale, sondern Querschnittsdaten handelte.

Anhand der Messwerte der NAKO Gesundheitsstudie können mit Vorbehalt Rückschlüsse auf die deutsche Bevölkerung gezogen werden. Die Stichprobe der NAKO Gesundheitsstudie repräsentiert die Altersstruktur der deutschen Bevölkerung aufgrund eines Selektionsbias nur bedingt, da lediglich Personen zwischen 20 und 73 Jahren eingeschlossen wurden und zudem die jüngeren Dekaden zu Gunsten der älteren Dekaden unterrepräsentiert waren. Darüber hinaus wurde die Stichprobe als Cluster im Einzugsgebiet der SZ gezogen, die wiederum nicht zufällig auf Deutschland verteilt waren. Bei der Rekrutierung der Studienteilnehmer könnte ein Freiwilligen-Bias aufgetreten sein (67), sodass diese im Vergleich zur Durchschnittsbevölkerung womöglich stärker an ihrer Gesundheit interessiert waren. Wenngleich die Studienteilnehmer für das Untersuchungsprogramm mitsamt der US-Messung des Abdominalfetts ebenfalls per Zufallsstichprobe ausgewählt wurden, könnte ein umgekehrter Healthy-Worker-Effekt entstanden sein (86): Das aufwändigere Level-2-Untersuchungsprogramm wurde vorzugsweise am Vormittag durchgeführt, weshalb die Teilnahme für vollzeitberufstätige, tendenziell gesundheitlich fittere Arbeitnehmer erschwert gewesen sein könnte. Aufschluss hierüber kann die Analyse der Sozialdaten und der Vorerkrankungen der Studienteilnehmer geben. Des Weiteren wurden die Werte des SAT und VAT in der vorliegenden Arbeit weder nach Vorerkrankungen noch nach gesundheitlichen Risikofaktoren und psychosozialen Faktoren adjustiert. Dieser Schritt wäre für die Berechnung von Referenzwerten anhand einer gesunden Untergruppe notwendig. Mithilfe der Longitudinaldaten, die im Rahmen der Folgeuntersuchungen der NAKO Gesundheitsstudie erhoben werden, könnten Grenzwerte des SAT und VAT für die Kalkulation des individuellen kardiovaskulären und metabolischen Risikos ermittelt werden. Da das abdominelle Fettgewebe den Phänotyp der Körperfettverteilung und das Risiko für Folgeerkrankungen besser erfasst, könnte es als Imaging Biomarker in Zukunft den BMI ablösen (87).

Im Vergleich mit anderen großen europäischen Studien, die eine US-Messung des Abdominalfetts durchführten, wurde in der NAKO Gesundheitsstudie weniger SAT

gemessen. Beispielsweise wurde bei italienischen, überwiegend übergewichtigen und adipösen Probanden ohne sonstige Vorerkrankungen mit vergleichbarer Altersspanne durchschnittlich 2,9 cm SAT gemessen ($n = 10.517$) (51, 53). In diesen beiden Studien lag der BMI der Teilnehmer bei knapp 30 kg/m^2 und damit über dem der Teilnehmer der NAKO Gesundheitsstudie, der sich bei $26,8 \text{ kg/m}^2$ befand (74). Bei Annahme einer Korrelation von SAT und BMI (88, 89) könnte dies der SAT-Differenz von 0,7 cm zugrunde liegen. Eine Subgruppenanalyse mit adipösen Studienteilnehmern der NAKO Gesundheitsstudie würden den Vergleich erleichtern. In einer dänischen Kohortenstudie wurde bei Probanden mit erhöhtem Risiko für Diabetes mellitus Typ 2 durchschnittlich 2,5 cm ($n = 1.342$) SAT gemessen (52). Die Frauen dieser Studie hatten durchschnittlich 0,5 cm mehr SAT als die der NAKO Gesundheitsstudie, was nicht allein durch das 15 Jahre höhere Durchschnittsalter zu erklären ist. Auch der BMI war mit $26,8 \text{ kg/m}^2$ vergleichbar. Die in der dänischen Studie postulierte Korrelation von abdominellem Fettgewebe und Diabetesrisiko könnte die entscheidende Ursache sein und sollte im Rahmen von Subgruppenanalysen der NAKO Gesundheitsstudie betrachtet werden. In den genannten italienischen und dänischen Studien wurde das SAT ebenfalls von der Haut bis zur Linea alba beziehungsweise dem Musculus rectus abdominis gemessen. Im Gegensatz dazu wurde in einer Studie zur Generierung von Referenzwerten anhand gesunder italienischer Blutspender durchschnittlich 1,8 cm SAT gemessen ($n = 250$), wobei von der Grenze zwischen Haut und subkutanem Fettgewebe zur Linea alba gemessen wurde (54). Daraus ergab sich eine kürzere Messstrecke als bei der Messung von der Hautoberfläche zur Linea alba, wie sie in der NAKO Gesundheitsstudie durchgeführt wurde. Bei einer Dicke der Dermis im Bereich des Körperstamms von bis zu 0,4 cm (83) könnte diese Abweichung der Messstrecke zu der beobachteten Differenz entscheidend beigetragen haben. Zusätzlich lag der durchschnittliche BMI dieser Probanden unter 25 kg/m^2 und damit unter dem der Studienteilnehmer der NAKO Gesundheitsstudie.

VAT wurde bei männlichen, überwiegend übergewichtig bis adipösen italienischen Probanden ähnlich viel, bei weiblichen Probanden 1,3 cm weniger gemessen als in der NAKO Gesundheitsstudie (51, 53, 55). Die Referenzwerte der gesunden, normalgewichtigen italienischen Blutspendern beider Geschlechter lagen ebenfalls etwa 40% niedriger als die Werte der NAKO Gesundheitsstudie (54). Bei allen genannten Studien wurde jedoch von der inneren Faszie des Musculus rectus abdominis zur Vorderwand der Aorta gemessen, sodass eine mindestens um den Durchmesser der Aorta kürzere

Messstrecke resultierte. Der Durchmesser der abdominellen Aorta beträgt knapp 2 cm und ist abhängig von Alter, Geschlecht und Körpergröße (90, 91). Zudem kann der Abstand von Aorta zur Wirbelvorderkante bis zu gut 1 cm betragen (41). Bei Hinzunahme dieser Differenz liegen die Messwerte des VAT der übergewichtig bis adipösen italienischen Probanden über denen der NAKO Gesundheitsstudie, was dem Verhältnis des SAT und des BMIs entspricht. Auch dänische Probanden mit erhöhtem Risiko für die Entwicklung eines Diabetes mellitus Typ 2 hatten knapp 1 cm mehr VAT (52). Hier wurde vom Peritoneum zur Lendenwirbelsäule gemessen, wodurch ebenfalls eine kürzere Strecke resultierte als bei der Messung der NAKO Gesundheitsstudie. In allen genannten Studien lassen sich die tendenziell höheren VAT-Werte vermutlich auf das ausgeprägtere Risikoprofil der Teilnehmer mit Übergewicht, Adipositas und Prädiabetes zurückführen. Auch hier wäre eine Subgruppenanalyse für den Vergleich der Messwerte interessant.

Die alters- und geschlechtsadjustierten Durchschnittswerte der 18 SZ lagen bei der Messung des SAT zwischen 1,8 cm und 2,5 cm und bei der Messung des VAT zwischen 5,1 cm und 7,6 cm. In einer ANOM konnte gezeigt werden, dass die Mehrheit der SZ signifikant vom globalen Mittelwert abwich. Folglich waren sich die Studienteilnehmer der einzelnen SZ hinsichtlich der Menge des SAT beziehungsweise VAT untereinander ähnlicher als SZ-übergreifend. Da die Tendenz der Abweichung der SZ der Tendenz der Abweichung der acht Cluster in den meisten Fällen entsprach, weist dies auf einen signifikanten Clustereffekt des SAT und VAT hin. Besonders ausgeprägt waren SAT als auch VAT in den SZ der Bundesländern Bayern, Sachsen/Sachsen-Anhalt und Mecklenburg-Vorpommern. Diese Beobachtung entsprach größtenteils der regionalen Verteilung von Adipositas aus Daten des deutschen Mikrozensus 2013 (92). Der Vergleich mit anthropometrischen Variablen der NAKO Gesundheitsstudie zeigte für den Body-Mass-Index, dass dieser ebenso in den Bundesländern Bayern, Sachsen/Sachsen-Anhalt und Mecklenburg-Vorpommern überdurchschnittlich hoch lag (74). Auch eine Gegenüberstellung mit den Werten des Taillenumfangs der NAKO Gesundheitsstudie wäre interessant, da dieser insbesondere mit dem VAT korreliert (88). Eine weitere denkbare Ursache für die signifikant voneinander abweichenden Mittelwerte des SAT und VAT an den einzelnen SZ könnten systematische Fehler wie die unterschiedliche Durchführung der Untersuchung sein, für die es bei der Bewertung der US-Bilder Hinweise gab.

Zur Kontrolle der Objektivität der US-Messung des Abdominalfetts wurde der Untersuchereffekt auf die Messwerte analysiert. Die nach Alter und Geschlecht adjustierten Untersucher Mittelwerte des SAT lagen zwischen 1,5 cm und 2,8 cm, die des VAT zwischen 4,6 cm und 8,6 cm. Damit differierten die Untersucher Mittelwerte erwartungsgemäß stärker als die Mittelwerte an den einzelnen SZ. In einer *ANOM* wichen knapp die Hälfte der analysierten Untersucher mit ihrem Mittelwert signifikant vom globalen Mittelwert des SAT ab, bei den Mittelwerten des VAT waren es knapp 60%. Auch hier könnte der oben beschriebene Clustereffekt eine tragende Rolle spielen, wobei ein Vergleich mit *ANOM*-Plots weniger untersucherabhängiger Messwerte wie dem Body-Mass-Index Aufschluss geben könnte. Dennoch wichen die Mittelwerte einzelner Untersucher in gegensätzlicher Richtung ihres SZ beziehungsweise Clusters ab, was eher durch systematische Fehler erklärbar wäre. Zur Identifikation abweichender Untersucher könnten zukünftig regelmäßig Analysen der Mittelwerte jedes einzelnen SZ durchgeführt werden und entsprechend auffallende Untersucher nachgeschult werden.

Auch für die Bewertung des Einflusses der Jahreszeit auf die US-Untersuchung des Abdominalfetts wurden Analysen der Mittelwerte durchgeführt. Im Frühjahr wurden hohe Messwerte des SAT mit einer signifikanten Abweichung im März beobachtet. Das VAT war in den Wintermonaten besonders ausgeprägt, wobei es im Januar signifikant vom globalen Mittelwert abwich. Unter Annahme einer Korrelation von Körpergewicht und abdominellem Fettgewebe (88) entsprach dies der Beobachtung einer Studie, die eine Gewichtszunahme der deutschen Bevölkerung nach den Weihnachtsfeiertagen und um Ostern nachweisen konnten (93). Dabei sollte beachtet werden, dass es sich bei den vorliegenden Daten der NAKO Gesundheitsstudie nicht um longitudinale Daten, sondern eine Querschnittsuntersuchung handelte. Die Uhrzeit der US-Messung hatte keinen signifikanten Einfluss auf die Messwerte des Abdominalfetts, jedoch konnten sowohl bei der Messung des SAT als auch bei der Messung des VAT höhere Messwerte nach Mittag festgestellt werden. Durch die Füllung der Dünndarmschlingen nach dem Mittagessen könnte das VAT größer zur Darstellung kommen, während die Nahrungsaufnahme keinen Einfluss auf die Menge des SAT haben sollte (71).

Der Einfluss von SZ, Cluster, Untersucher und Zeit auf die Messwerte des Abdominalfetts wurde mittels Analysen der Mittelwerte eingeschätzt. Diese Methode ermöglichte die simultane Identifizierung von Gruppen, deren Mittelwert signifikant vom globalen

Mittelwert abweicht (75). Die Gruppenmittelwerte mitsamt der Entscheidungsgrenzen wurden in sogenannten *ANOM*-Plots dargestellt, wobei die Entscheidungsgrenzen nach der Gruppengröße und weiteren Variablen wie dem Alter und Geschlecht der Studienteilnehmer adjustiert wurden. Die *ANOM* diente der Prüfung der Objektivität der Messung. Beim Vergleich der Untersuchermittelwerte war eine Unterscheidung zwischen Clustereffekten und systematischen Fehlern im Sinne einer mangelhaften Interraterreliabilität nicht möglich. Die Analysen der Mittelwerte wurden zudem unter der Annahme von normalverteilten Daten durchgeführt, weshalb die leicht rechtsschiefe Verteilung der Messwerte des Abdominalfetts falsch-positive Ergebnisse provoziert haben könnte. Aufgrund der großen Stichprobengröße der NAKO Gesundheitsstudie waren die Entscheidungsgrenzen der *ANOM*-Plots sehr eng, sodass Abweichungen im Millimeterbereich Signifikanz erlangen konnten. Dabei ist es fraglich, ob die statistisch signifikanten Ergebnisse auch eine praktische Relevanz haben, insbesondere weil die Differenz zwischen erstem und zweitem Messwert ebenso im Millimeterbereich lag. In der SOP wird beispielsweise ein Abweichung zwischen erstem und zweitem Messwert von bis zu 0,5 cm toleriert (73).

Die Reproduzierbarkeit der US-Messwerte des Abdominalfetts wurde anhand des *ICC* und Bland-Altman-Plots geprüft. Die *ICC*-Werte der 74 Untersucher lagen allesamt $> 0,9$ und sprachen damit für eine exzellente Reliabilität sowohl der Messwerte des SAT als auch des VAT. Dennoch bleibt zu beachten, dass wiederholte Messungen durch den gleichen Untersucher aufgrund fehlender Maskierung weniger stark zum Ausgleich von Variabilität dienen als bei der Interobserver-Variabilität. Die Untersucher konnten sich bei der zweiten Messung vermutlich an den ersten Messwert erinnern oder bemühten sich um eine sehr genaue Messung, weshalb die *ICC*-Werte möglicherweise falsch hoch sind (69). Auch in Bland-Altman-Plots wurde weder eine systematische Differenz noch ein Trend beobachtet. Hierbei wurden die Messwerte aller Untersucher in einem Plot dargestellt. Um systematische Fehler einzelnen Untersucher zu identifizieren, müsste für jeden Untersucher der NAKO Gesundheitsstudie ein Bland-Altman-Plot betrachtet werden. Da auch der *ICC* systematische Unterschiede zwischen erster und zweiter Messung erfasst (69), wurde darauf zur Gewährung der Übersichtlichkeit verzichtet. Die Diskrepanz zwischen erstem und zweitem Messwert wurde aufgrund der geringen Spannweite des Mittelwerts der Differenz als niedrig eingeschätzt. Sie betrug bei der Messung des SAT 0,5 cm und bei der Messung des VAT 1,0 cm. Die Ergebnisse der Reliabilitätsuntersuchungen

entsprachen größtenteils den Beobachtungen einer Studie zur Reproduzierbarkeit und Validität der US-Messung des abdominalen Fettgewebes, die in einer Vorphase der NAKO Gesundheitsstudie durchgeführt wurde (94). Auch dort wurden ICC-Werte $> 0,9$ und eine gute Übereinstimmung der Messwerte bei gleichem Untersuchungsablauf wie in der NAKO Gesundheitsstudie erzielt. Lediglich die Spannweite des Mittelwerts der Differenz war mit 0,4 cm bei der Messung des SAT und 0,6 cm bei der Messung des VAT noch niedriger, was vermutlich an der geringeren Stichprobengröße lag. Die Interraterreliabilität konnte aufgrund des Studiendesigns der NAKO Gesundheitsstudie nicht betrachtet werden.

Die Validität der US-Messung des Abdominalfetts könnte mittels Korrelation der US-Messwerte mit Messwerten aus der MRT geprüft werden. Bei einer Subgruppe der Level 2-Studienteilnehmer der NAKO Gesundheitsstudie wurde in fünf SZ eine MRT-Untersuchung des ganzen Körpers durchgeführt. Da die MRT als Goldstandard für die Messung des viszeralen Abdominalfetts gilt (32), wäre hier ebenfalls eine Gegenüberstellung der US- und MRT-Messwerte in einem Bland-Altman-Plot sinnvoll. Dabei könnten die MRT-Messwerte auf die Abszisse aufgetragen werden, sodass die Lage des Mittelwerts der Differenz zwischen beiden Messmethoden den Bias der US-Messung im Vergleich zum Goldstandard angeben würde (69).

4.3 Bewertung der US-Bilder

Um die Objektivität der US-Messung des Abdominalfetts in der NAKO Gesundheitsstudie genauer einschätzen zu können, wurde eine Stichprobe der US-Bilder mit Qualitätskriterien zur Messrichtigkeit und Bildqualität bewertet. Die Messung des SAT wurde bei 86,3% der bewerteten US-Bilder im Hinblick auf die Messrichtigkeit als komplett korrekt eingeschätzt. Dieses Ergebnis spricht für eine erfolgreiche Schulung der Untersucher und hohe Qualität der Messwerte. Anhand eines logistischen Regressionsmodells konnte gezeigt werden, dass insbesondere bei jungen Studienteilnehmern mit viel SAT die Messpunkte richtig gesetzt wurden. Bei älteren Studienteilnehmern und Studienteilnehmern mit wenig SAT bestanden größere Schwierigkeiten vorwiegend bei der Positionierung des unteren Messpunkts, möglicherweise aufgrund einer schwierigeren Abgrenzbarkeit der Linea alba durch Verwachsungen wie beispielsweise nach Operationen. Ein weiterer häufiger Fehler bestand darin, dass der obere Messpunkt des SAT unterhalb der Bauchhaut positioniert wurde, was in 6,1% der Fälle beobachtet wurde. Daraus resultierten eine kürzere Messstrecke und

folglich ein zu kleiner Messwert. Dieser Fehler betraf im SZ Berlin Nord gut 60% der bewerteten US-Bilder, weshalb hier ein systematischer Fehler angenommen werden kann. Möglicherweise wurde dabei die Grenze zwischen Haut und subkutanem Fettgewebe als Position des oberen Messpunktes des SAT verwendet, wie es in anderen Studienprotokollen etabliert ist (50). Auch eine Weitergabe von falschen Informationen bei der Einarbeitung der Untersucher könnte in Betracht kommen, auch wenn die Messstrecke in der SOP eindeutig beschrieben war.

Bei der Messung des VAT konnten 59,7% der bewerteten US-Bilder hinsichtlich der Messrichtigkeit als komplett korrekt eingestuft werden. Das schlechtere Abschneiden als bei der Messung des SAT liegt vermutlich an der technisch anspruchsvolleren Untersuchungsdurchführung und stärkeren Abhängigkeit von Teilnehmerfaktoren wie Darmgasüberlagerung, die die Messung zusätzlich erschweren. Insbesondere bei jungen Studienteilnehmern mit wenig VAT wurden die Messpunkte jedoch korrekt gesetzt. Der obere Messpunkt des VAT am Unterrand der Linea alba wurde in vier von fünf Fällen richtig positioniert, wobei er sich in 10,7% der Fälle am Oberrand statt am Unterrand der Linea alba befand. Daraus resultierten eine minimal längere Messstrecke sowie ein größerer Messwert. Dieser Fehler wurde besonders häufig in den SZ Essen, Hamburg und Heidelberg/Mannheim beobachtet. Der untere Messpunkt des VAT war bei drei von vier bewerteten Bildern korrekt positioniert. In 12,6% dieser Fälle waren die Zielstrukturen nicht sicher erkennbar, sodass eine Fehlmessung zu vermuten war. Daraus ließ sich ableiten, dass die richtige Darstellung des Wirbelkörpers die größten Schwierigkeiten bereitet hatte. Diese könnte durch Darmgasüberlagerungen im Querkolon erschwert sein, weshalb die Messung gemäß der SOP am Ende der Ausatmung mit Verlagerung des Querkolons nach kranial stattfinden sollte. Eine weitere Hilfe könnte die Verwendung des Trackballs sein, der das Vorspulen des Bildes um wenige Millisekunden ermöglicht, um die bestmögliche Momentaufnahme mit scharfer Darstellung der Zielstrukturen zu erhalten. Diese Methode wurde in der überarbeiteten SOP behandelt (73). Darüber hinaus sollten die Untersucher dahingehend geschult werden, dass die Dokumentation unsicherer oder falscher Messwerte für die Validität der Studie ungünstiger ist als das Fehlen einzelner Werte. Aufgrund der Häufung bestimmter Fehler an einzelnen SZ konnte eine Fehlerweitergabe innerhalb eines SZ nicht ausgeschlossen werden. Für deren Kontrolle sind regelmäßige Site Visits durch geschulte und zertifizierte Mitarbeiter anderer SZ essenziell.

Zusätzlich zur Messrichtigkeit wurde die Bildqualität der US-Bilder bewertet, um die Präzision der Messung des Abdominalfetts beurteilen zu können. Eine möglichst exakte Messung des Abdominalfetts setzt voraus, dass die Zielstrukturen formatfüllend und zentriert sowie in mittlerer Helligkeit aufgenommen werden. Die Zentrierung wurde anhand der Position der Linea alba geprüft, die sich bei über 90% der bewerteten US-Bilder mittig befand. Die formatfüllende Aufnahme der Zielstrukturen wurde in der zweiten Version der SOP „US Abdominalfett“ nicht detailliert behandelt (63), was die niedrige Erfolgsquote von lediglich 36,7% bei der Messung des SAT erklären könnte. Bei beiden Messungen wurden die Zielstrukturen zu klein dargestellt, wobei insbesondere bei der Messung des SAT zu wenig gezoomt wurde. Möglicherweise wurde die Zoom-Voreinstellung von manchen Untersuchern bei 14 cm Eindringtiefe belassen, sodass das VAT mit 85,0% deutlich häufiger formatfüllend dargestellt wurde, während insbesondere wenig SAT sehr klein zur Darstellung kam. Die Einstellung der Bildhelligkeit gelang bei nahezu allen bewerteten US-Bildern des SAT, während 14,7% der bewerteten US-Bilder des VAT in der Tiefe zu dunkel waren. Besonders bei Studienteilnehmern mit viel VAT waren die US-Bilder aufgrund der Schallabschwächung bei höherer Eindringtiefe häufig in schallkopffernen Ebenen zu dunkel. Das erschwerte die Identifikation des Wirbelkörpers, wobei die Nutzung der optimalen Eindringtiefe sowie der tiefenabhängigen Verstärkung Abhilfe verschaffen könnte. In der überarbeiteten Version der SOP wurde auf die Verwendung der iScan-Taste hingewiesen, die eine automatische Optimierung der tiefenabhängigen Verstärkung und Bildhelligkeit ermöglicht (73, 64). Bei der Bewertung der Bildqualität traten ebenso wie bei der Bewertung der Messrichtigkeit Unterschiede zwischen den 18 SZ auf. Beispielsweise wurden alle 40 bewerteten US-Bilder des SAT aus dem SZ Saarbrücken formatfüllend aufgenommen, während es in den SZ Münster und Düsseldorf bei keinem US-Bild gelang. Im SZ Essen war die Bildhelligkeit bei lediglich 38,0% der bewerteten US-Bilder des VAT korrekt eingestellt, während es im SZ-übergreifenden Durchschnitt 81,9% waren. Dieser Beobachtung könnten Unterschiede in der Schulung der US-Untersuchung an den einzelnen SZ zugrunde liegen. Durch eine detaillierte SOP und engmaschige Supervision könnte die Präzision und Objektivität der US-Untersuchung weiter erhöht werden.

Anhand der US-Bildbewertung wurde für jeden betrachteten Untersucher eine Rückmeldung erstellt, die die Bewertung der einzelnen Kategorien der Messrichtigkeit und Bildqualität sowie individuelle Verbesserungsvorschläge umfasste. Darüber hinaus

wurden Methoden aus der überarbeiteten SOP zur Verbesserung der Bildqualität und Präzision der Messung wie die Verwendung des Zooms und des Fokus beschrieben. Bei knapp zwei Drittel der bewerteten Untersucher wurde eine Nachschulung empfohlen, da mindestens eine der Kategorien der Messrichtigkeit bei über 25% der US-Bilder des jeweiligen Untersuchers als fehlerhaft bewertet wurde. Für die Definition des Nachschulungsbedarf wurde die Bewertung der Bildqualität nicht zu Rate gezogen, da diese sich einerseits vornehmlich auf die Präzision der Messwerte auswirkte und andererseits aufgrund unscharfer Beschreibung in der zweiten Version der SOP (63) nicht vorausgesetzt werden durfte. Für all diejenigen Untersucher, die erst nach dem Data Freeze 100k in der US-Untersuchung des Abdominalfetts tätig waren, wurden Empfehlungen für die korrekte und präzise Durchführung der Untersuchung mit Beschreibung der Kategorien der Messrichtigkeit und Bildqualität verfasst. Sowohl bei der Bildbewertung als auch bei der Analyse der Mittelwerte zeigte sich, dass die US-Messung auch in der NAKO Gesundheitsstudie eine stark untersucherabhängige Methode ist. Die Rückmeldung, die Nachschulung der Untersucher und die Empfehlungen sollten durch Verbesserung der Objektivität und Reliabilität eine höhere Validität der US-Messungen zur Folge haben, könnten jedoch auch die Vergleichbarkeit der Messwerte einschränken. Zur Quantifizierung dieses Effektes wäre eine Bewertung der Messrichtigkeit und Bildqualität im Verlauf sinnvoll, ebenso wie zur weiteren Qualitätskontrolle der US-Messung im Rahmen der NAKO Gesundheitsstudie. Für die Nutzung der Bildbewertung als Maßstab für die Qualität der Messwerte einzelner Untersucher wurde dem bereinigten Data Freeze 100k Datensatz der Anteil komplett korrekt durchgeführter US-Untersuchungen jedes Untersuchers als neue Variable beigefügt. Dies ermöglicht die ausschließliche Verwendung hoch qualitativer Daten einzelner Untersucher für zukünftige wissenschaftliche Auswertungen.

Eine mögliche Alternative zur manuellen Bewertung der US-Bilder wäre die automatisierte Bewertung mittels Machine Learning. Durch diese Methode der künstlichen Intelligenz wäre die Bewertung der US-Bilder zeiteffizienter und weniger vom Bewertenden abhängig. Ein weiterer Vorteil wäre, dass damit statt einer Stichprobe alle US-Bilder bewertet werden könnten. Infolgedessen wäre es möglich, lediglich korrekt gemessene US-Bilder für wissenschaftliche Auswertungen zu verwenden.

5 Zusammenfassung

Das Ziel dieser Arbeit war die Beurteilung der Qualität der US-Messung des Abdominalfetts im Rahmen der NAKO Gesundheitsstudie. Diese wurde anhand der SOP zur US-Untersuchung, der Messwerte des Data Freeze 100k Datensatzes und der Bewertung einer Stichprobe der US-Bilder eingeschätzt. Besonderes Augenmerk lag auf der Reliabilität und Objektivität der US-Messung. Zudem wurde die SOP „US Abdominalfett“ überarbeitet und der Data Freeze 100k Datensatz bereinigt. Die Ergebnisse der Qualitätssicherung wurden in einem Qualitätsbericht für die NAKO Gesundheitsstudie zusammengefasst.

Das in der SOP beschriebene Messprotokoll für die Messung von SAT und VAT wurde in einer Vorphase der NAKO Gesundheitsstudie entwickelt und validiert. Beide Messungen wurden am liegenden Probanden auf Höhe der Taille in der Medianlinie durchgeführt. Das SAT wurde von der Bauchhaut bis zur Linea alba, das VAT von der Linea alba bis zum Lendenwirbelkörper gemessen. Dieses Vorgehen entsprach dem Vorgehen anderer Studien mit Ausnahme der Messstrecken. Die Messung des SAT schloss im Gegensatz zu anderen Studien die Bauchhaut mit ein, woraus etwas größere Messwerte resultierten. Für die Messung des VAT wurden in der Literatur unterschiedliche Messpunkte verwendet, die meist kürzere Messstrecken zur Folge hatten, was die Vergleichbarkeit der Ergebnisse erschwerte. Die Beschreibung der Untersuchungsdurchführung in der SOP erfüllte einen Großteil der Kriterien für ein hochwertiges US-Bild. Lediglich die fehlende Vorbereitung der Studienteilnehmer hinsichtlich Nahrungsaufnahme könnte einen negativen Einfluss auf die Reproduzierbarkeit und Validität der US-Messwerte des VAT haben. Für eine größere Präzision der Messwerte wurden die Regulierung der Eindringtiefe für eine formatfüllende Aufnahme der Zielstrukturen und die Verwendung des Fokus bei der Überarbeitung der SOP eingeführt. Um den Untersuchern eine bessere räumliche Vorstellung der Zielstrukturen zu ermöglichen, wurden anatomische Abbildungen hinzugefügt.

Die Stichprobe der NAKO Gesundheitsstudie unterlag aufgrund der gewählten Altersstruktur mit Überrepräsentation der älteren Dekaden sowie der Erhebung in Clustern einem Selektionsbias. Aus strukturellen Gründen fehlten zum Zeitpunkt der Erstellung des Data Freeze 100k drei Viertel aller US-Untersuchungen in zwei der insgesamt 18 SZ, weshalb diese SZ in der Gesamtauswertung unterrepräsentiert waren. Die

gemessenen Werte des SAT und VAT wurden in dieser Arbeit nicht nach Vorerkrankungen sowie gesundheitlichen Risikofaktoren oder sozioökonomischen Faktoren adjustiert, weshalb sie nicht als Referenzwerte dienen können. Die Vergleichbarkeit der Messwerte mit anderen Studien ist aufgrund abweichender Messstrecken und Charakteristika der Studienpopulationen eingeschränkt. Tendenziell wurde im Rahmen der NAKO Gesundheitsstudie mit durchschnittlich 2,2 cm weniger SAT als in anderen großen europäischen Studien bei gleicher Messstrecke gemessen. Jedoch lag in den Vergleichsstudien der BMI der Teilnehmer höher oder sie wiesen ein erhöhtes Diabetesrisiko auf, welches mit dem abdominellen Fettgewebe korreliert. Im Gegensatz dazu war das VAT mit durchschnittlich 6,7 cm stärker ausgeprägt als in anderen europäischen Studien, wobei aber auch die verwendete Messstrecke um den Durchmesser der Aorta länger war. Für einen besseren Vergleich wären Subgruppenanalysen sinnvoll.

Die Qualität der US-Messung des Abdominalfetts wurde anhand der drei Gütekriterien Reliabilität, Objektivität und Validität beurteilt. Als Maß der Intraobserver-Reliabilität wurde der *ICC* berechnet, der den Anteil der interindividuellen Variabilität an der Gesamtvariabilität der Messwerte angibt. Dieser lag sowohl bei der Messung des SAT als auch bei der Messung des VAT sehr hoch, sodass lediglich ein geringer Anteil der Varianz durch systematische Unterschiede zwischen erster und zweiter Messung zustande kam. Die durchschnittliche absolute Abweichung zwischen den Messwerten der Doppelmessung war mit 0,1 cm bei der Messung des SAT und 0,2 cm bei der Messung des VAT ebenso sehr niedrig. Zur Veranschaulichung von Mustern in der Diskrepanz der Doppelmessung wurden Bland-Altman-Plots erstellt, wobei weder ein Trend noch eine systematische Differenz zwischen erstem und zweitem Messwert beobachtet werden konnte. Die Reliabilität der US-Untersuchung des Abdominalfetts im Rahmen der NAKO Gesundheitsstudie konnte dementsprechend als exzellent bewertet werden.

Für die Beurteilung der Objektivität der US-Untersuchung des Abdominalfetts wurden *ANOM* verwendet. Diese Methode dient dem simultanen Vergleich der Mittelwerte einzelner Gruppen mit dem globalen Mittelwert. Sie zeigte signifikante Unterschiede zwischen den Messwerten der Untersucher sowie der SZ, wofür zwei Ursachen denkbar wären: Einerseits könnten systematische Messfehler der Untersucher und somit eine mangelhafte Objektivität dieser Beobachtung zugrunde liegen. Andererseits

wären systematische Unterschiede zwischen den Studienteilnehmern der einzelnen SZ im Sinne eines Clustereffekts als Ursache möglich. Für einen Clustereffekt spricht die Tatsache, dass die Tendenz der Abweichung der Cluster meist der Tendenz der Abweichung ihrer zugeordneten SZ und Untersucher entsprach. Gewissheit könnte der Vergleich mit *ANOM*-Plots weniger untersucherabhängiger Messwerte der NAKO Gesundheitsstudie wie dem Body-Mass-Index geben. Darüber hinaus könnte der Einfluss der Untersucher und SZ auf die Messwerte des SAT und VAT aufgrund der großen Stichprobengröße der NAKO Gesundheitsstudie und der daraus resultierenden engen Entscheidungsgrenzen der *ANOM*-Plots überschätzt sein. Zudem lag die Abweichung der SZ beziehungsweise Untersuchern vom globalen Mittelwert ebenso wie die Abweichung zwischen erstem und zweitem Messwert im Millimeterbereich, was deren Bedeutung relativiert.

Als weiteren möglichen Einflussfaktor auf die US-Messung des Abdominalfetts wurden der Monat und die Uhrzeit der Untersuchungsdurchführung betrachtet. Im Januar wurde mehr VAT, im März mehr SAT gemessen. Unter Annahme einer Korrelation von Körpergewicht und abdominalem Fettgewebe könnte diese Beobachtung einer Gewichtszunahme an den Feiertagen entsprechen. Die Daten der NAKO Gesundheitsstudie wurden jedoch nicht longitudinal, sondern im Querschnitt erhoben. Die Uhrzeit der Untersuchungsdurchführung hatte keinen signifikanten Einfluss auf die Messwerte des SAT und VAT.

Um Messfehler bei der US-Messung des Abdominalfetts aufzudecken, wurde eine Stichprobe der US-Bilder jedes Untersuchers anhand definierter Kriterien der Messrichtigkeit und Bildqualität bewertet. Als Kriterien der Messrichtigkeit dienten die Positionierung der Messpunkte sowie die senkrechte Ausrichtung der Messlinie. Komplett korrekt gemessen wurden bei Anwendung der Kriterien der Messrichtigkeit 86% der bewerteten US-Bilder des SAT und 60% der bewerteten US-Bilder des VAT. Insbesondere bei der Messung des SAT spricht dieses Ergebnis für eine erfolgreiche Schulung der Untersucher und hohe Qualität der Messwerte. Jedoch konnten sowohl bei der Messung des SAT als auch des VAT systematische Fehler bei der Positionierung der Messpunkte beobachtet werden. So wurde beispielsweise der obere Messpunkt des SAT unterhalb der Bauchhaut positioniert oder das VAT zur Aorta statt zum Wirbelkörper gemessen. Diese Messfehler häuften sich in einzelnen SZ beziehungsweise bei einzelnen Untersuchern. In Zusammenschau mit den statistisch

signifikanten Unterschieden der Messwerte der Untersucher und SZ in den *ANOM*-Plots schränkt dies die Objektivität der Messung ein. Als Ursache der systematischen Fehler ist an erster Stelle eine Fehlerweitergabe innerhalb der SZ bei der Schulung der Untersucher zu vermuten. Besondere Übung und Konzentration waren für die Positionierung des unteren Messpunktes des VAT an der Vorderkante des Wirbelkörpers notwendig. Hier wurden bei jedem achten bewerteten US-Bild Messfehler beobachtet. Auch die Konstitution und das Alter der Studienteilnehmer hatten einen Einfluss auf die Messrichtigkeit. Diese wurde bei jungen Studienteilnehmern mit viel SAT beziehungsweise wenig VAT besser bewertet. Die Messung von wenig SAT könnte durch den stärkeren Einsatz der Zoomfunktion vereinfacht werden, wie er in der überarbeiteten Version der SOP beschrieben wird. Für die Darstellung des Wirbelkörpers kann die Bildoptimierung mittels iScan und Fokus hilfreich sein.

Neben der Messrichtigkeit wurde auch die Bildqualität der US-Bilder bewertet, um einen Eindruck der Präzision der Messwerte zu erhalten. Bei einem Großteil der bewerteten US-Bilder waren die Zielstrukturen zentriert und in mittlerer Helligkeit dargestellt. Lediglich bei Studienteilnehmern mit viel VAT waren die US-Bilder gelegentlich aufgrund von Schallabschwächung in der Tiefe zu dunkel. Das VAT wurde im Gegensatz zum SAT meist formatfüllend dargestellt. Letzteres wurde in zwei von drei bewerteten US-Bildern zu klein aufgenommen, sodass eine präzise Positionierung der Messpunkt erschwert sein könnte. Auch bei der Bewertung der Bildqualität zeigten sich deutliche Unterschiede zwischen den SZ, was auf Unterschiede in der Schulung und Übung der Untersucher zurückgeführt werden kann.

Aufgrund fehlender Dokumentationsmöglichkeit konnten weitere denkbare Einflussfaktoren auf die US-Messung des Abdominalfetts wie die Position und der Druck der US-Sonde nicht betrachtet werden. Trotzdem lässt sich anhand der dargestellten Ergebnisse schlussfolgern, dass die US-Untersuchung auch im Rahmen der NAKO Gesundheitsstudie eine stark untersucherabhängige Methode ist. Während die Messung des SAT von den Untersuchern gut durchgeführt wurde, sind bei der Messung des VAT einige Schwierigkeiten aufgetreten. Um die Objektivität der US-Messung des Abdominalfetts zu verbessern, wurde anhand der Bewertung der US-Bilder eine individuelle Rückmeldung für jeden Untersucher verfasst. Darüber hinaus wurde zwei Dritteln der Untersucher eine Nachschulung empfohlen, da bei mehr als 25% der bewerteten US-Bilder dieser Untersucher Fehler bei der Positionierung der

Messpunkte beobachtet wurden. Auch die SOP wurde mit dem Ziel überarbeitet, die Messrichtigkeit und Bildqualität der US-Messung zu fördern. Die Anwendung der dort beschriebenen Methoden sollte eine größere Präzision der Messung zur Folge haben, die sich wiederum positiv auf die Reliabilität der Messwerte auswirken könnte. Einen weiteren hohen Stellenwert nehmen eine gute Schulung und regelmäßige Supervision der Untersucher ein.

Aus der Verbesserung der Objektivität und Reliabilität der US-Messung des Abdominalfetts sollte eine Zunahme der Validität resultieren. Da die MRT den Goldstandard für die Messung des Abdominalfetts darstellt, könnte die Validität der US-Messung mittels Korrelation der US-Messwerte mit MRT-Messwerten eingeschätzt werden. MRT-Schnittbilder wurden an fünf SZ zusätzlich zu den US-Messwerten generiert und standen für die vorliegende Arbeit nicht zur Verfügung.

Diese Arbeit konnte zeigen, dass insbesondere die Messung des SAT in der NAKO Gesundheitsstudie bereits hoch präzise und reliabel durchgeführt wird. Auf vielfältige Weise wurden Fehlerquellen bei der Messung des VAT identifiziert und Techniken zur Erleichterung der Messung in die neue Version der SOP aufgenommen. Die überarbeitete SOP sowie die Bewertung der US-Bilder und das Verfassen von individuellen Rückmeldungen können zu einer weiteren Verbesserung der Untersuchungsdurchführung in den Folgeuntersuchungen entscheidend beitragen. Im weiteren Verlauf der NAKO Gesundheitsstudie sollten regelmäßig Qualitätssicherungen durchgeführt und der Effekt der getroffenen Maßnahmen zur Qualitätsförderung geprüft werden. Dafür käme auch eine automatisierte Bewertung aller US-Bilder mittels Machine Learning in Frage. Im Anschluss können Referenzwerte für das SAT und VAT anhand der Daten aller Studienteilnehmer der Basisuntersuchung erstellt werden. Dabei sollte eine Adjustierung nach Vorerkrankungen, Risikofaktoren und sozioökonomischen Faktoren erfolgen und die Repräsentativität der Studienpopulation bedacht werden. Die longitudinalen Daten der Folgeuntersuchungen ermöglichen die Risikoberechnung des abdominellen Fettgewebes für kardiovaskuläre und metabolische Erkrankungen. Gegebenenfalls stellt das abdominelle Fettgewebe in Zukunft eine Alternative zum Body-Mass-Index dar, da es den Phänotyp der Fettverteilung und das Risiko für Folgeerkrankungen besser erfasst. Dabei könnten SAT und VAT einen wichtigen Stellenwert als Imaging Biomarker einnehmen.

6 Anhang

6.1 ANOM-Plots zu Untersuchereffekten

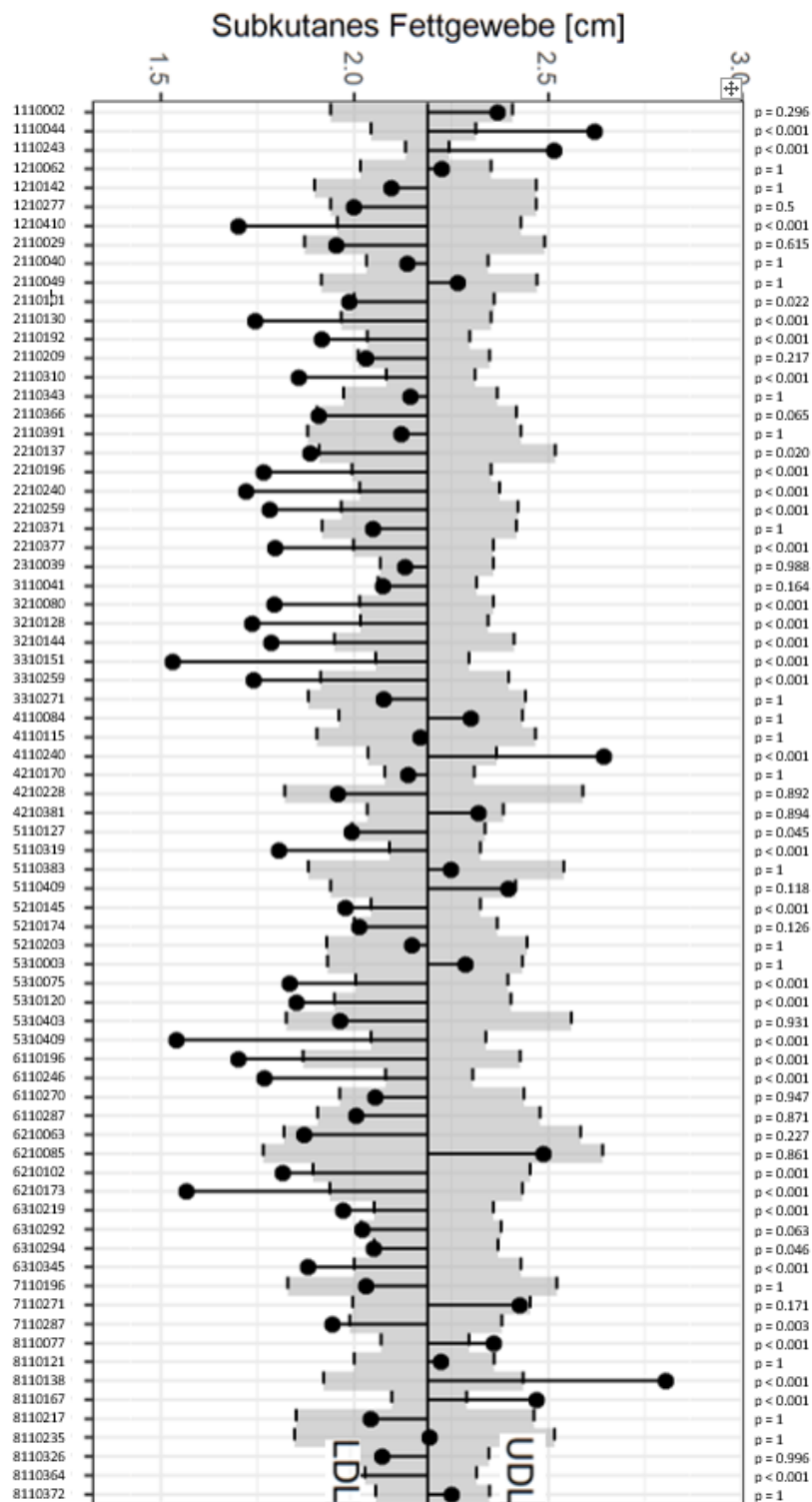


Abbildung 31: ANOM-Plot zur Abweichung der Untersuchermittelwerte des SAT vom globalen Mittelwert 2,2 cm ($N = 18.890$) inklusive Untersuchernummern und p -Werten. Adjustierung der Entscheidungsgrenzen nach Alter, Geschlecht und Anzahl der Studienteilnehmer je Untersucher.
ANOM = Analyse der Mittelwerte, SAT = subkutanes Fettgewebe, cm = Zentimeter, N = Anzahl, p = p -Wert, LDL = untere Entscheidungsgrenze, UDL = obere Entscheidungsgrenze.

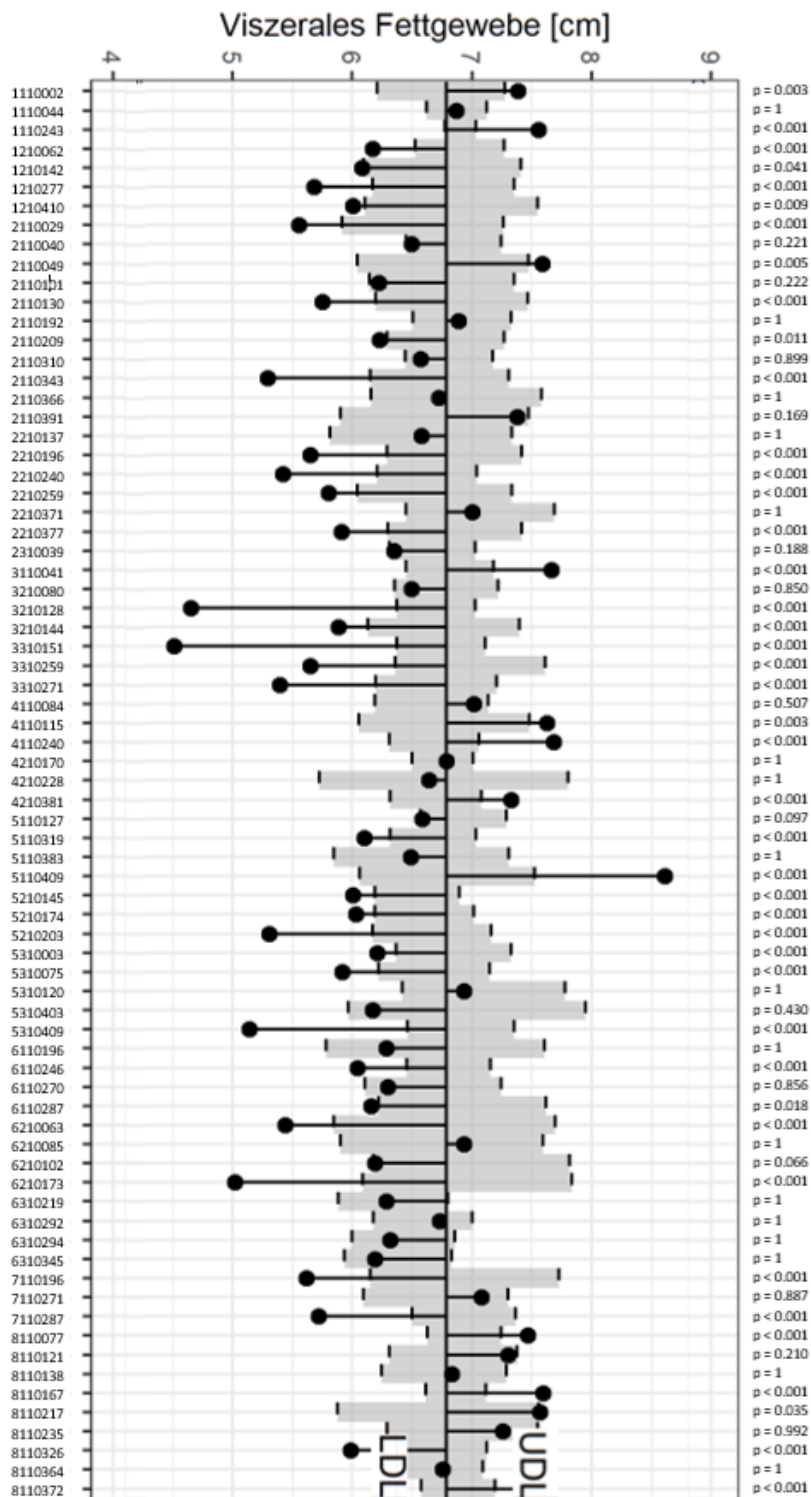


Abbildung 32: ANOM-Plot zur Abweichung der Untersuchermittelwerte des VAT vom globalen Mittelwert 6,8 cm ($N = 18.666$) inklusive Untersuchernummern und p-Werten. Adjustierung der Entscheidungsgrenzen nach Alter, Geschlecht und Anzahl der Studienteilnehmer je Untersucher. ANOM = Analyse der Mittelwerte, VAT = viszerales Fettgewebe, cm = Zentimeter, N = Anzahl, p = p-Wert, LDL = untere Entscheidungsgrenze, UDL = obere Entscheidungsgrenze.

6.2 Schablonen für die US-Bildbewertung

Für die Bewertung der Eindringtiefe des US und der Ausrichtung der Messlinie wurden Schablonen erstellt (vergleiche Abbildung 33), die für die Vollbilddarstellung der US-Bilder im Programm Xcelera Express von Philips auf einem Monitor mit 55“ Durchmesser konzipiert sind. Sie wurden auf durchsichtige Folie gedruckt und an den Monitor angelehnt.

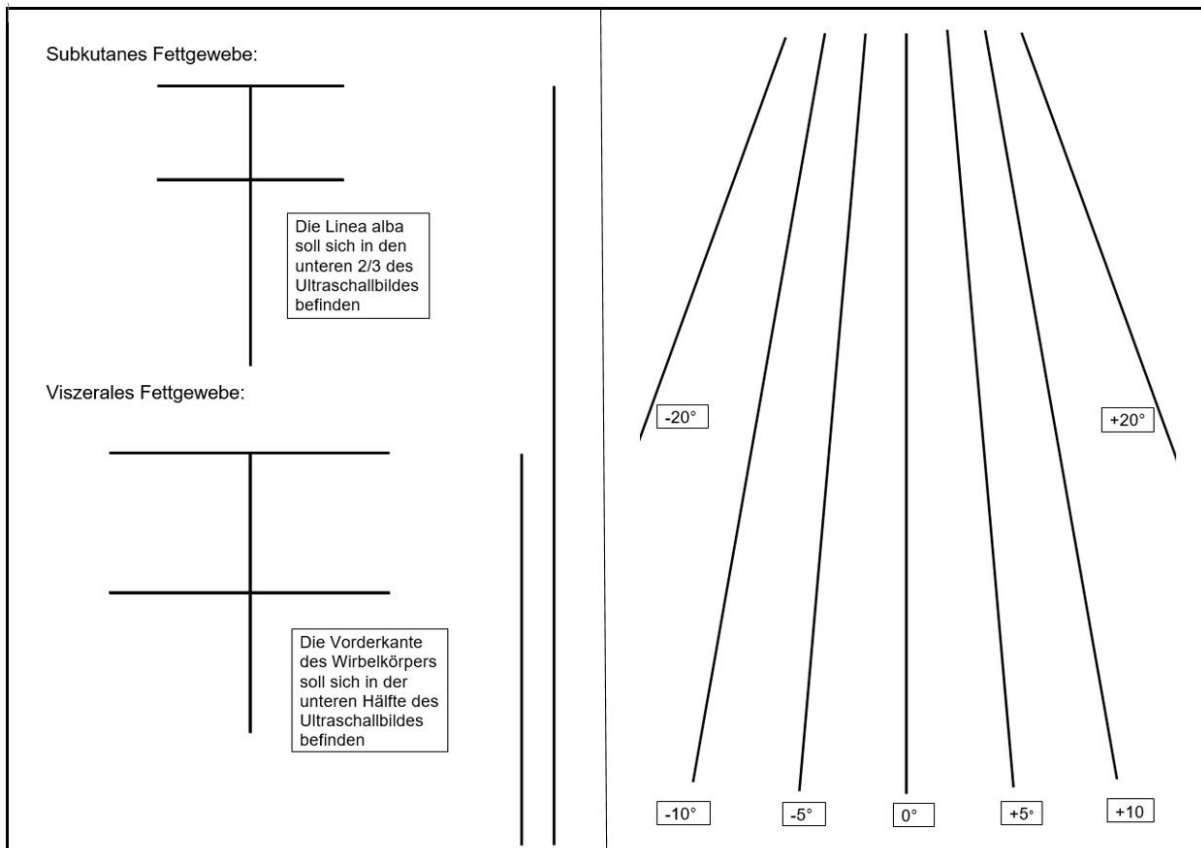


Abbildung 33: Schablonen für die Bewertung der formatfüllenden Darstellung der Zielstrukturen (links) und der senkrechten Ausrichtung der Messlinie (rechts).

6.3 Rückmeldung zur US-Messung des Abdominalfetts

Für jeden bewerteten Untersucher wurde eine Rückmeldung basierend auf der Bewertung der US-Bilder erstellt und im Januar 2019 an die Untersucher verschickt. Abbildung 34 zeigt beispielhaft eine solche Rückmeldung.

Rückmeldung an Untersucher

Im Rahmen der Qualitätssicherung des Moduls Ultraschall Abdominalfett wurden 22 von Ihnen untersuchte Probanden per Zufallsstichprobe bestimmt und die dazugehörigen Ultraschallbilder im Hinblick auf Messrichtigkeit und Bildqualität bewertet.

Tabelle 1. Bewertung der Messrichtigkeit und Bildqualität der Ultraschallmessung des subkutanen und viszeralen Fettgewebes

	Kriterien	Subkutanee Fettgewebe	Viszerale Fettgewebe
Messrichtigkeit	Messlinie senkrecht	Sehr gut	Sehr gut
	Messpunkte richtig gesetzt	Gut	Befriedigend
Bildqualität	Mittlere Gesamthelligkeit	Sehr gut	Befriedigend
	Formatfüllende Aufnahme	Gut	Sehr gut
	Linea alba mittig im Bild	Sehr gut	Sehr gut

Messrichtigkeit:

Messlinie: Die Messlinie verlief immer (88/88 Messungen) senkrecht.

Messpunkte: Sie haben den oberen Messpunkt des subkutanen Fettgewebes an der Bauchhaut immer (44/44 Messungen) korrekt positioniert. Den unteren Messpunkt des subkutanen Fettgewebes am Oberrand der Linea alba haben Sie oft (32/44 Messungen) korrekt positioniert. Den oberen Messpunkt des viszeralen Fettgewebes am Unterrand der Linea alba haben Sie oft (31/44 Messungen) korrekt positioniert. Den unteren Messpunkt des viszeralen Fettgewebes an der Vorderkante des Wirbelkörpers haben Sie gelegentlich (29/44 Messungen) korrekt positioniert.

Gelegentlich (12/44 Messungen) haben Sie den unteren Messpunkt des subkutanen Fettgewebes fälschlicherweise oberhalb der Linea alba im subkutanen Fettgewebe, am Unterrand der Linea alba, unterhalb der Linea alba oder am Peritoneum positioniert. Gelegentlich (13/44 Messungen) haben Sie den oberen Messpunkt des viszeralen Fettgewebes fälschlicherweise am Peritoneum, mittig oder am Oberrand der Linea alba oder am Oberrand des Bauchmuskels positioniert. Selten (3/44 Messungen) haben Sie den unteren Messpunkt des viszeralen Fettgewebes fälschlicherweise seitlich am Wirbelkörper positioniert. Gelegentlich (12/44 Messungen) war der untere Messpunkt des viszeralen Fettgewebes nicht beurteilbar, da der Wirbelkörper nicht eindeutig zu erkennen war.

Verbesserungsvorschläge:

Messlinie: Keine Verbesserungsvorschläge.

Messpunkte: Den unteren Messpunkt des subkutanen Fettgewebes bitte am Oberrand der Linea alba positionieren. Den oberen Messpunkt des viszeralen Fettgewebes bitte am Unterrand der Linea alba positionieren. Den unteren Messpunkt des viszeralen Fettgewebes bitte an der Vorderkante des Wirbelkörpers positionieren.

Bei Schwierigkeiten, die Zielstrukturen darzustellen, ist es im Zweifelsfall besser, keinen Wert in das NatCoEdo einzutragen als einen möglicherweise falschen Wert. Schreiben Sie dazu die 999,9 in das dazugehörige Feld.

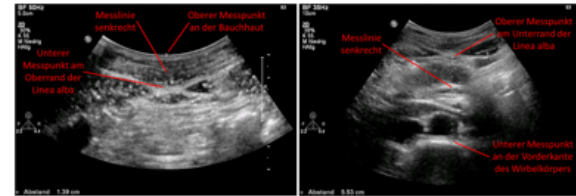


Abbildung 1: Beispielfotografen zur Messrichtigkeit. Links: Messung des subkutanen Fettgewebes. Rechts: Messung des viszeralen Fettgewebes.

Bildqualität:

Helligkeit: Die Helligkeit des Ultraschallbilds war bei der Messung des subkutanen Fettgewebes nahezu immer (42/44 Messungen) im optimalen Bereich mit gleicher Helligkeit an der Oberfläche und in der Tiefe. Bei der Messung des viszeralen Fettgewebes war das Ultraschallbild gelegentlich (14/22 Messungen) in der Tiefe zu dunkel.

Zoom: Mithilfe der Zoomfunktion des Ultraschallgeräts haben Sie immer (44/44 Messungen) so gezoomt, dass das viszerale Fettgewebe formatfüllend dargestellt wurde. Bei der Messung des subkutanen Fettgewebes haben Sie meistens (36/44 Messungen) so gezoomt, dass das subkutane Fettgewebe formatfüllend dargestellt wurde.

Linea alba: Die Linea alba befand sich nahezu immer (84/88 Messungen) mittig im Bild.

Verbesserungsvorschläge:

Helligkeit: Zur besseren Darstellung der Zielstrukturen bitte die Helligkeit des viszeralen Fettgewebes so einstellen, dass sie im mittleren Bereich mit gleicher Helligkeit an der Oberfläche und in der Tiefe liegt. Dafür können die Tasten iScan und Gain verwendet werden. Zusätzlich kann der Fokus bei der Messung des subkutanen Fettgewebes auf die Linea alba bzw. bei der Messung des viszeralen Fettgewebes auf die Vorderkante des Wirbelkörpers gelegt werden, um diese Strukturen noch schärfer darzustellen.

Zoom: Für eine genauere Messung des subkutanen Fettgewebes bitte mit dem Tieferegler stärker zoomen, sodass die Zielstrukturen formatfüllend dargestellt werden. Die Linea alba sollte bei der Messung des subkutanen Fettgewebes im mittleren oder unteren Bild Drittel liegen. Bei der Messung des viszeralen Fettgewebes sollte sich die Vorderkante des Wirbelkörpers in der unteren Bildhälfte befinden.

Linea alba: Keine Verbesserungsvorschläge.

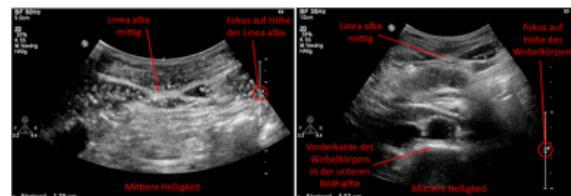


Abbildung 2: Beispielfotografen zur Bildqualität. Links: Messung des subkutanen Fettgewebes. Rechts: Messung des viszeralen Fettgewebes.

Übertragung:

Sie haben die Messwerte nahezu immer (82/88 Messwerte) korrekt in das NatCoEdo übertragen. Selten (6/88 Messwerte) gab es Übertragungsfehler.

Schlussbemerkung:

Da Sie gelegentlich Schwierigkeiten mit der Messung des viszeralen Fettgewebes hatten, wäre eine Nachschulung sinnvoll. Bitte wenden Sie sich dafür an den Trainer in Ihrem Studienzentrum, in naheliegenden Studienzentren oder an die Modulverantwortliche.

Der genaue Ablauf der Messung und einige Hilfestellungen sind in der SOP UN25-SD_V03 beschrieben.

Bei Fragen zu den Bewertungen oder zur Messung selbst können Sie sich gerne jederzeit an die Modulverantwortliche Anja Sedlmeier wenden:

Adresse: Universität Regensburg
Lehrstuhl für Epidemiologie und Präventivmedizin
Franz-Josef-Strauß-Allee 11
D-93053 Regensburg
E-Mail: anja.sedlmeier@ukr.de
Telefon: 0941-0446218

6.4 Empfehlungen zur US-Messung des Abdominalfetts

Als Ergänzung zur SOP wurden Empfehlungen zur US-Messung mit den Kriterien der Messrichtigkeit und Bildqualität sowie Beispielbildern verfasst (vergleiche Abbildung 35) und allen SZ der NAKO Gesundheitsstudie im Januar 2019 zur Verfügung gestellt.

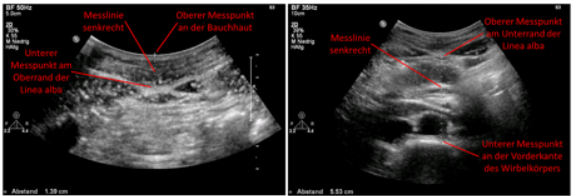
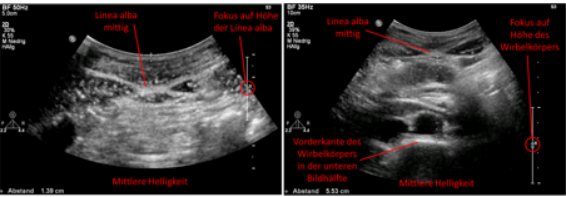
<p>Empfehlungen zur Ultraschallmessung des Abdominalfetts</p> <p>Für eine korrekte und präzise Durchführung der Ultraschallmessung des Abdominalfetts bitte ich Sie, folgende Kriterien der Messrichtigkeit und Bildqualität zu berücksichtigen.</p> <p>Kriterien der Messrichtigkeit:</p> <p>Messlinie: Für eine exakte Messung bitte ich Sie, darauf zu achten, dass die Messlinie senkrecht verläuft. Sie können sich gerne ein Blatt Papier oder ein Lineal zu Hilfe nehmen und an den Bildschirm als Orientierung anlehnen.</p> <p>Messpunkte: Den oberen Messpunkt des subkutanen Fettgewebes bitte direkt an der Bauchhaut positionieren. Den unteren Messpunkt des subkutanen Fettgewebes bitte am Oberrand der Linea alba positionieren. Den oberen Messpunkt des viszeralen Fettgewebes bitte am Unterrand der Linea alba positionieren. Den unteren Messpunkt des viszeralen Fettgewebes bitte an der Vorderkante des Wirbelkörpers positionieren.</p> <p>Bei Schwierigkeiten, die Zielstrukturen darzustellen, ist es im Zweifelsfall besser, keinen Wert in das NatCoEdc einzutragen als einen möglicherweise falschen Wert. Schreiben Sie dazu die 999,9 in das dazugehörige Feld.</p>  <p>Abbildung 1: Beispielbilder zur Messrichtigkeit. Links: Messung des subkutanen Fettgewebes. Rechts: Messung des viszeralen Fettgewebes.</p>	<p>Kriterien der Bildqualität:</p> <p>Helligkeit: Zur besseren Darstellung der Zielstrukturen bitte die Helligkeit so einstellen, dass sie im mittleren Bereich liegt mit gleicher Helligkeit an der Oberfläche und in der Tiefe. Dafür können die Tasten iScan und Gain verwendet werden. Zusätzlich kann der Fokus bei der Messung des subkutanen Fettgewebes auf die Linea alba bzw. bei der Messung des viszeralen Fettgewebes auf die Vorderkante des Wirbelkörpers gelegt werden, um diese Strukturen noch schärfer darzustellen.</p> <p>Zoom: Für eine genauere Messung bitte mit dem Tieferegler stärker zoomen, sodass die Zielstrukturen formatfüllend dargestellt werden. Die Linea alba sollte bei der Messung des subkutanen Fettgewebes im mittleren oder unteren Bilddrittel liegen. Bei der Messung des viszeralen Fettgewebes sollte sich die Vorderkante des Wirbelkörpers in der unteren Bildhälfte befinden.</p> <p>Linea alba: Um sicherzustellen, dass an der Mittellinie des Körpers gemessen wird, bitte darauf achten, dass sich die Linea alba mittig im Ultraschallbild befindet.</p>  <p>Abbildung 2: Beispielbilder zur Bildqualität. Links: Messung des subkutanen Fettgewebes. Rechts: Messung des viszeralen Fettgewebes.</p> <p>Der genaue Ablauf der Messung und einige Hilfestellungen sind in der SOP UN25-SD_V03 beschrieben. Bei Fragen zur Messung können Sie sich gerne jederzeit an die Modulverantwortliche Anja Sedlmeier wenden:</p> <p>Adresse: Universität Regensburg Lehrstuhl für Epidemiologie und Präventivmedizin Franz-Josef-Strauß-Allee 11 D-93053 Regensburg</p> <p>E-Mail: anja.sedlmeier@ukr.de Telefon: 0941-9445216</p>
---	---

Abbildung 35: Empfehlungen zur US-Messung des Abdominalfetts.
US = Ultraschall.

7 Literaturverzeichnis

1. World Health Organisation. Obesity and overweight [Fact Sheet]; 2020 [cited 2020 Dec 17]. Available from: URL: <http://www.who.int/en/news-room/fact-sheets/detail/obesity-and-overweight>.
2. Mensink GBM, Schienkiewitz A, Haftenberger M, Lampert T, Ziese T, Scheidt-Nave C. Übergewicht und Adipositas in Deutschland: Ergebnisse der Studie zur Gesundheit Erwachsener in Deutschland (DEGS1). Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2013; 56(5-6):786–94.
3. World Health Organisation. Obesity: Preventing and managing the global epidemic. Geneva: World Health Organization; 2000. (Technical report series / World Health Organization; vol 894).
4. Eckel RH, Grundy SM, Zimmet PZ. The metabolic syndrome. Lancet 2005; 365(9468):1415–28.
5. Wirth A, Hauner H, editors. Adipositas: Ätiologie, Folgekrankheiten, Diagnostik, Therapie. 4., vollst. überarb. und aktual. Aufl. Berlin: Springer; 2013.
6. Dai H, Alsallhe TA, Chalghaf N, Riccò M, Bragazzi NL, Wu J. The global burden of disease attributable to high body mass index in 195 countries and territories, 1990-2017: An analysis of the Global Burden of Disease Study. PLoS Med 2020; 17(7).
7. Després JP. Intra-abdominal obesity: an untreated risk factor for Type 2 diabetes and cardiovascular disease. J Endocrinol Invest 2006; 29(3 Suppl):77–82.
8. Ungefroren H, Gieseler F, Fliedner S, Lehnert H. Obesity and cancer. Horm Mol Biol Clin Investig 2015; 21(1):5–15.
9. Shen W, Wang Z, Punyanita M, Lei J, Sinav A, Kral JG et al. Adipose tissue quantification by imaging methods: a proposed classification. Obes Res 2003; 11(1):5–16.
10. Schleinitz D, Böttcher Y, Blüher M, Kovacs P. The genetics of fat distribution. Diabetologia 2014; 57(7):1276–86.
11. Wajchenberg BL. Subcutaneous and visceral adipose tissue: their relation to the metabolic syndrome. Endocr Rev 2000; 21(6):697–738.

12. Tchernof A, Després J-P. Pathophysiology of human visceral obesity: an update. *Physiol Rev* 2013; 93(1):359–404.
13. Goss AM, Darnell BE, Brown MA, Oster RA, Gower BA. Longitudinal associations of the endocrine environment on fat partitioning in postmenopausal women. *Obesity (Silver Spring)* 2012; 20(5):939–44.
14. Ambikairajah A, Walsh E, Tabatabaei-Jafari H, Cherbuin N. Fat mass changes during menopause: a metaanalysis. *Am J Obstet Gynecol* 2019; 221(5):393-409.
15. Krotkiewski M, Björntorp P, Sjöström L, Smith U. Impact of obesity on metabolism in men and women: Importance of regional adipose tissue distribution. *J Clin Invest* 1983; 72(3):1150–62.
16. Romieu I, Dossus L, Barquera S, Blottière HM, Franks PW, Gunter M et al. Energy balance and obesity: what are the main drivers? *Cancer Causes Control* 2017; 28(3):247–58.
17. Ibrahim MM. Subcutaneous and visceral adipose tissue: structural and functional differences. *Obes Rev* 2010; 11(1):11–8.
18. Freedland ES. Role of a critical visceral adipose tissue threshold (CVATT) in metabolic syndrome: implications for controlling dietary carbohydrates: a review. *Nutr Metab (Lond)* 2004; 1(1):12.
19. Unger RH. Lipotoxic diseases. *Annu Rev Med* 2002; 53:319–36.
20. Hammarstedt A, Gogg S, Hedjazifar S, Nerstedt A, Smith U. Impaired Adipogenesis and Dysfunctional Adipose Tissue in Human Hypertrophic Obesity. *Physiol Rev* 2018; 98(4):1911–41.
21. Fain JN, Madan AK, Hiler ML, Cheema P, Bahouth SW. Comparison of the release of adipokines by adipose tissue, adipose tissue matrix, and adipocytes from visceral and subcutaneous abdominal adipose tissues of obese humans. *Endocrinology* 2004; 145(5):2273–82.
22. Luca C de, Olefsky JM. Stressed out about obesity and insulin resistance. *Nat Med* 2006; 12(1):41-42.
23. Shoelson SE, Lee J, Goldfine AB. Inflammation and insulin resistance. *J Clin Invest* 2006; 116(7):1793–801.

24. Ponti F, Santoro A, Mercatelli D, Gasperini C, Conte M, Martucci M et al. Aging and Imaging Assessment of Body Composition: From Fat to Facts. *Front Endocrinol (Lausanne)* 2020; 10.
25. Kadowaki T, Yamauchi T, Kubota N, Hara K, Ueki K, Tobe K. Adiponectin and adiponectin receptors in insulin resistance, diabetes, and the metabolic syndrome. *J Clin Invest* 2006; 116(7):1784–92.
26. Achari AE, Jain SK. Adiponectin, a Therapeutic Target for Obesity, Diabetes, and Endothelial Dysfunction. *Int J Mol Sci* 2017; 18(6).
27. Fox CS, Massaro JM, Hoffmann U, Pou KM, Maurovich-Horvat P, Liu C-Y et al. Abdominal visceral and subcutaneous adipose tissue compartments: Association with metabolic risk factors in the Framingham Heart Study. *Circulation* 2007; 116(1):39–48.
28. Liu J, Fox CS, Hickson DA, May WD, Hairston KG, Carr JJ et al. Impact of abdominal visceral and subcutaneous adipose tissue on cardiometabolic risk factors: the Jackson Heart Study. *J Clin Endocrinol Metab* 2010; 95(12):5419–26.
29. Oka R, Miura K, Sakurai M, Nakamura K, Yagi K, Miyamoto S et al. Impacts of visceral adipose tissue and subcutaneous adipose tissue on metabolic risk factors in middle-aged Japanese. *Obesity (Silver Spring)* 2010; 18(1):153–60.
30. Torre-Villalvazo I, Bunt AE, Alemán G, Marquez-Mota CC, Diaz-Villaseñor A, Noriega LG et al. Adiponectin synthesis and secretion by subcutaneous adipose tissue is impaired during obesity by endoplasmic reticulum stress. *J Cell Biochem* 2018; 119(7):5970–84.
31. Han J, Kaufman RJ. The role of ER stress in lipid metabolism and lipotoxicity. *J Lipid Res* 2016; 57(8):1329–38.
32. Shuster A, Patlas M, Pinthus JH, Mourtzakis M. The clinical importance of visceral adiposity: a critical review of methods for visceral adipose tissue analysis. *Br J Radiol* 2012; 85(1009):1–10.
33. Sahakyan KR, Somers VK, Rodriguez-Escudero JP, Hodge DO, Carter RE, Sochor O et al. Normal-Weight Central Obesity: Implications for Total and Cardiovascular Mortality. *Ann Intern Med* 2015; 163(11):827–35.

34. Ruderman NB, Schneider SH, Berchtold P. The "metabolically-obese," normal-weight individual. *Am J Clin Nutr* 1981; 34(8):1617–21.
35. Sims EA. Are there persons who are obese, but metabolically healthy? *Metab Clin Exp* 2001; 50(12):1499–504.
36. Caleyachetty R, Thomas GN, Toulis KA, Mohammed N, Gokhale KM, Balachandran K et al. Metabolically Healthy Obese and Incident Cardiovascular Disease Events Among 3.5 Million Men and Women. *J Am Coll Cardiol* 2017; 70(12):1429–37.
37. Balkau B, Deanfield JE, Després J-P, Bassand J-P, Fox KAA, Smith SC et al. International Day for the Evaluation of Abdominal Obesity (IDEA): A study of waist circumference, cardiovascular disease, and diabetes mellitus in 168,000 primary care patients in 63 countries. *Circulation* 2007; 116(17):1942–51.
38. Neamat-Allah J, Wald D, Hüsing A, Teucher B, Wendt A, Delorme S et al. Validation of anthropometric indices of adiposity against whole-body magnetic resonance imaging - a study within the German European Prospective Investigation into Cancer and Nutrition (EPIC) cohorts. *PLoS ONE* 2014; 9(3):e91586.
39. Fang H, Berg E, Cheng X, Shen W. How to best assess abdominal obesity. *Curr Opin Clin Nutr Metab Care* 2018; 21(5):360–5.
40. Hu HH, Chen J, Shen W. Segmentation and quantification of adipose tissue by magnetic resonance imaging. *MAGMA* 2016; 29(2):259–76.
41. Delorme S, Debus J, Jenderka K-V. *Sonografie: 105 Tabellen*. 3., vollst. überarb. und erw. Aufl. Stuttgart: Thieme; 2012. (Duale Reihe).
42. Armellini F, Zamboni M, Rigo L, Todesco T, Bergamo-Andreis IA, Procacci C et al. The contribution of sonography to the measurement of intra-abdominal fat. *J Clin Ultrasound* 1990; 18(7):563–7.
43. Armellini F, Zamboni M, Robbi R, Todesco T, Rigo L, Bergamo-Andreis IA et al. Total and intra-abdominal fat measurements by ultrasound and computerized tomography. *Int J Obes Relat Metab Disord* 1993; 17(4):209–14.

44. Bazzocchi A, Filonzi G, Ponti F, Sassi C, Salizzoni E, Battista G et al. Accuracy, reproducibility and repeatability of ultrasonography in the assessment of abdominal adiposity. *Acad Radiol* 2011; 18(9):1133–43.
45. Lucia Rolfe E de, Sleigh A, Finucane FM, Brage S, Stolk RP, Cooper C et al. Ultrasound measurements of visceral and subcutaneous abdominal thickness to predict abdominal adiposity among older men and women. *Obesity (Silver Spring)* 2010; 18(3):625–31.
46. Murphy J, Bacon SL, Morais JA, Tsoukas MA, Santosa S. Intra-Abdominal Adipose Tissue Quantification by Alternative Versus Reference Methods: A Systematic Review and Meta-Analysis. *Obesity (Silver Spring)* 2019; 27(7):1115–22.
47. Stolk RP, Wink O, Zelissen PMJ, Meijer R, van Gils APG, Grobbee DE. Validity and reproducibility of ultrasonography for the measurement of intra-abdominal adipose tissue. *Int J Obes* 2001; 25(9):1346–51.
48. Schlecht I, Wiggermann P, Behrens G, Fischer B, Koch M, Freese J et al. Reproducibility and validity of ultrasound for the measurement of visceral and subcutaneous adipose tissues. *Metab Clin Exp* 2014; 63(12):1512–9.
49. Mauad FM, Chagas-Neto FA, Benedeti ACGS, Nogueira-Barbosa MH, Muglia VF, Carneiro AAO et al. Reproducibility of abdominal fat assessment by ultrasound and computed tomography. *Radiol Bras* 2017; 50(3):141–7.
50. Azzi AJ, Lafrenière A-S, Gilardino M, Hemmerling T. Ultrasonography Technique in Abdominal Subcutaneous Adipose Tissue Measurement: A Systematic Review. *J Ultrasound Med* 2019; 38(4):877–88.
51. Bertoli S, Leone A, Vignati L, Spadafranca A, Bedogni G, Vanzulli A et al. Metabolic correlates of subcutaneous and visceral abdominal fat measured by ultrasonography: a comparison with waist circumference. *Nutr J* 2016; 15(2).
52. Philipsen A, Jørgensen ME, Vistisen D, Sandbaek A, Almdal TP, Christiansen JS et al. Associations between ultrasound measures of abdominal fat distribution and indices of glucose metabolism in a population at high risk of type 2 diabetes: the ADDITION-PRO study. *PLoS ONE* 2015; 10(4):e0123062.
53. Leone A, Battezzati A, Bedogni G, Vignati L, Vanzulli A, Amicis R de et al. Sex- and Age-Related Differences in the Contribution of Ultrasound-Measured Visceral

- and Subcutaneous Abdominal Fat to Fatty Liver Index in Overweight and Obese Caucasian Adults. *Nutrients* 2019; 11(12).
54. Bazzocchi A, Ponti F, Diano D, Moio A, Albisinni U, Pasquali R et al. Abdominal adiposity by ultrasonography: a "pocket" database for reference standard in Italian people. *Prim Care Diabetes* 2014; 8(4):358–64.
55. Bertoli S, Leone A, Krakauer NY, Bedogni G, Vanzulli A, Redaelli VI et al. Association of Body Shape Index (ABSI) with cardio-metabolic risk factors: A cross-sectional study of 6081 Caucasian adults. *PLoS ONE* 2017; 12(9):e0185013.
56. Wichmann H-E, Kaaks R, Hoffmann W, Jöckel K-H, Greiser KH, Linseisen J. Die Nationale Kohorte. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 2012; 55(6-7):781–7.
57. Schipf S, Schöne G, Schmidt B, Günther K, Stübs G, Greiser KH et al. Die Basiserhebung der NAKO Gesundheitsstudie: Teilnahme an den Untersuchungsmodulen, Qualitätssicherung und Nutzung von Sekundärdaten. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 2020; 63(3):254–66.
58. German National Cohort (GNC) Consortium. The German National Cohort: aims, study design and organization. *Eur J Epidemiol* 2014; 29(5):371–82.
59. Deutsche Gesellschaft für Epidemiologie, Deutsche Gesellschaft für Medizinische Informatik Biometrie und Epidemiologie, Deutsche Gesellschaft für Sozialmedizin und Prävention, Deutsche Region der Internationalen Biometrischen Gesellschaft, Technologie- und Methodenplattform für die Vernetzte Medizinische Forschung E. V., Deutsches Netzwerk Versorgungsforschung E. V., editor. Leitlinien und Empfehlungen zur Sicherung von Guter Epidemiologischer Praxis (GEP); 2018. Available from: URL: https://www.dgepi.de/assets/Leitlinien-und-Empfehlungen/Leitlinien_fuer_Gute_Epidemiologische_Praxis_GEP_vom_September_2018.pdf [cited 2020 Dec 17].
60. Ahrens W, Greiser H, Linseisen J, Kluttig A, Schipf S, Schmidt B et al. Das Design der Machbarkeitsstudien für eine bundesweite Kohortenstudie in Deutschland: Die Pretests der Nationalen Kohorte (NAKO).

Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2014;
57(11):1246–54.

61. Frost C, White IR. The effect of measurement error in risk factors that change over time in cohort studies: do simple methods overcorrect for 'regression dilution'? *Int J Epidemiol* 2005; 34(6):1359–68.
62. Schipf S, Ostrzinski S, Melzer J. NAKO Standardarbeitsanweisung (SOP): Vorgaben für die Datenbereinigung im Rahmen des Data Freeze 100.000 [QM03-GE]; 2017.
63. Sedlmeier A, Schlecht I. NAKO Standardarbeitsanweisung (SOP): Ultraschall Abdominalfett Version 02 [UN25-SD_V02]; 2017.
64. Philips. iE33 Ultraschallsystem Benutzerhandbuch: 4535 616 23841 Rev A; 2014.
65. Schünke M, Schulte E, Schumacher U. Prometheus - LernAtlas der Anatomie: Allgemeine Anatomie und Bewegungssystem; 182 Tabellen. 2., überarb. und erw. Aufl. Stuttgart [u.a.]: Thieme; 2007.
66. Schünke M, Schulte E, Schumacher U. Prometheus - Lernatlas der Anatomie: Innere Organe; 118 Tabellen. 2., überarb. und erw. Aufl. Stuttgart [u.a.]: Thieme; 2009.
67. Weiß C. Basiswissen Medizinische Statistik. Berlin, Heidelberg: Springer Berlin Heidelberg; 2019.
68. Biemer PP, Lyberg L. Introduction to survey quality. Hoboken, N.J.: Wiley; 2003. (Wiley series in survey methodology).
69. Nieto F, Szklo M, editors. Epidemiology: Beyond the basics. 2. ed. Sudbury, Mass.: Jones and Bartlett; 2007.
70. Hedderich J, Sachs L. Angewandte Statistik: Methodensammlung mit R. 16. Auflage. Berlin, Heidelberg: Springer; 2018.
71. Bazzocchi A, Filonzi G, Ponti F, Amadori M, Sassi C, Salizzoni E et al. The role of ultrasonography in the evaluation of abdominal fat: analysis of technical and methodological issues. *Acad Radiol* 2013; 20(10):1278–85.

72. R: A Language and Environment for Statistical Computing. Version 3.6.1. Vienna, Austria: R Foundation for Statistical Computing; 2019. Available from: URL: <https://www.R-project.org/>.
73. Sedlmeier A, Schlecht I. NAKO Standardarbeitsanweisung (SOP): Ultraschall Abdominalfett Version 03 [UN25-SD_V03]; 2018.
74. Fischer B, Sedlmeier AM, Hartwig S, Schlett CL, Ahrens W, Bamberg F et al. Anthropometrische Messungen in der NAKO Gesundheitsstudie – mehr als nur Größe und Gewicht. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2020; 63(3):290–300.
75. Pallmann P, Hothorn LA. Analysis of means: a generalized approach using R. Journal of Applied Statistics 2016; 43(8):1541–60.
76. Wollschläger D. Grundlagen der Datenanalyse mit R: Eine anwendungsorientierte Einführung. 4., überarbeitete und erweiterte Auflage. Berlin: Springer Spektrum; 2017. (Statistik und ihre Anwendungen).
77. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1(8476):307–10.
78. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med 2016; 15(2):155–63.
79. Haber P. Die Dokumentationsprüfung nach der neuen Ultraschallvereinbarung: Tipps zur Verbesserung der technischen Bildqualität; 2011.
80. Xie Y, Allaire JJ, Grolemond G. R Markdown: The definitive guide. Boca Raton, London, New York: CRC Press Taylor & Francis Group a Chapman & Hall book; 2018. (Chapman & Hall/CRC the R series).
81. Gradmark AMI, Rydh A, Renström F, Lucia-Rolfe E de, Sleigh A, Nordström P et al. Computed tomography-based validation of abdominal adiposity measurements from ultrasonography, dual-energy X-ray absorptiometry and anthropometry. Br J Nutr 2010; 104(4):582–8.
82. Kuchenbecker WKH, Groen H, Pel H, Bolster JHT, Wolffenbuttel BHR, Land JA et al. Validation of the measurement of intra-abdominal fat between ultrasound and CT scan in women with obesity and infertility. Obesity (Silver Spring) 2014; 22(2):537–44.

83. Held L, Metzler G, Schaller M. Histologische Strukturen der normalen Haut. In: L. Cerroni et al. (Hg.), *Histopathologie der Haut*, Springer Reference Medizin. 61-76.
84. Sedlmeier A, Schlecht I. NAKO Standardarbeitsanweisung (SOP): Qualitätssicherung Ultraschall Abdominalfett [UN25-QS_V03]; 2018.
85. Elffers TW, Mutsert R de, Lamb HJ, Roos A de, van Willems Dijk K, Rosendaal FR et al. Body fat distribution, in particular visceral fat, is associated with cardiometabolic risk factors in obese women. *PLoS ONE* 2017; 12(9):e0185403.
86. Shah D. Healthy worker effect phenomenon. *Indian J Occup Environ Med* 2009; 13(2):77–9.
87. Neeland IJ, Lemos JA de. Time to Retire the BMI?: Evaluating Abdominal Adipose Tissue Imaging as Novel Cardiovascular Risk Biomarker. *J Am Coll Cardiol* 2016; 68(14):1522–4.
88. Ping Z, Pei X, Xia P, Chen Y, Guo R, Hu C et al. Anthropometric indices as surrogates for estimating abdominal visceral and subcutaneous adipose tissue: A meta-analysis with 16,129 participants. *Diabetes Res Clin Pract* 2018; 143:310–9.
89. Nadeem B, Bacha R, Gilani SA. Correlation of Subcutaneous Fat Measured on Ultrasound with Body Mass Index. *J Med Ultrasound* 2018; 26(4):205–9.
90. Kawel-Boehm N, Maceira A, Valsangiacomo-Buechel ER, Vogel-Claussen J, Turkbey EB, Williams R et al. Normal values for cardiovascular magnetic resonance in adults and children. *J Cardiovasc Magn Reson* 2015; 17(29).
91. Sienz M, Ignee A, Dietrich CF. Normwerte in der Abdomensonografie - Aorta, Vena cava inferior, Niere. *Z Gastroenterol* 2012; 50(3):293–315.
92. Robert Koch-Institut. Bericht Gesundheit in Deutschland 2015: Gesundheitsberichterstattung des Bundes gemeinsam getragen von RKI und Destatis. Berlin; 2015 [cited 2018 Apr 30].
93. Helander EE, Wansink B, Chieh A. Weight Gain over the Holidays in Three Countries. *N Engl J Med* 2016; 375(12):1200–2.
94. Schlecht I. Reproducibility and validity of ultrasound-assessed visceral and subcutaneous body fat and associations with inflammatory parameters and metabolites [Dissertation zur Erlangung des Doktorgrades der Biomedizinischen Wissenschaften (Dr. rer. physiol.)]. Regensburg: Universität Regensburg; 2015.

8 Erklärung

Diese Arbeit wurde von Herrn Prof. Dr. Dr. Michael Leitzmann angeregt und in Kooperation mit Frau Anja Sedlmeier erstellt.

Ich erkläre hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Insbesondere habe ich nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater oder andere Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeit erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

9 Danksagung

Mein Dank geht an allererster Stelle an Prof. Dr. Dr. Michael Leitzmann, dem ich für die Ermöglichung der Arbeit, die angenehmen Besprechungstermine und seine Geduld danken möchte.

Ebenso dankbar bin ich Anja Sedlmeier, die meine Arbeit über zweieinhalb Jahre betreute und immer ein offenes Ohr hatte. Sie ließ mir das perfekte Maß an Eigeninitiative und war trotzdem leitend im Hintergrund. Zudem las sie den Großteil meiner Arbeit Korrektur und gab mir produktives Feedback.

Ein besonderes Dankeschön geht an Sylvia Pietsch, die mir als Sekretärin des Lehrstuhls für Epidemiologie und Präventivmedizin bei technischen Problemen immer zur Seite stand.

Darüber hinaus möchte ich Andrea Weber für ihre Hilfe bei statistischen Fragestellungen danken. Mit ihr und Anja Sedlmeier hatte ich eine wunderbare Zeit im Büro, an die ich gerne zurückdenke.

Auch dem Team des NAKO-Studienzentrums in Regensburg bin ich dankbar für die schöne Zeit als Untersucherin, die nicht zuletzt den Weg zu dieser Arbeit bahnte.

Danke auch an meine Eltern, die mir das Studium und diese Arbeit ermöglichten und mich in allem unterstützten.

Zuletzt möchte ich meinem Freund Jens für seine besonnene Art danken, er ist mein Ruhepol und immer für mich da.