

Aus dem Institut für Klinische Chemie  
(Direktor: Prof. Dr. med. Ralf Junker)  
im Universitätsklinikum Schleswig-Holstein, Campus Kiel  
an der Christian-Albrechts-Universität zu Kiel

**Referenzwertbestimmung endokrinologischer Parameter mittels  
indirekter Methode auf Basis laborinterner Messwerte und ihre  
Relevanz im klinischen Alltag**

Inauguraldissertation  
zur  
Erlangung der Doktorwürde der Medizin  
der Medizinischen Fakultät  
der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Jessica Sommer

aus Eitorf

Kiel 2021

1. Berichterstatter: Prof. Dr. Ralf Junker, Institut für Klinische Chemie

2. Berichterstatter: Prof. Dr. Wolfgang Lieb, Institut für Epidemiologie

Tag der mündlichen Prüfung: 07.07.2023

Zum Druck genehmigt: Kiel, den 21.03.2023

gez.: PD Dr. Frank Leypoldt

(Vorsitzender der Prüfungskommission)

<b>Inhaltsverzeichnis</b>	
<b>Abbildungsverzeichnis</b>	<b>III</b>
<b>Tabellenverzeichnis</b>	<b>V</b>
<b>Abkürzungsverzeichnis</b>	<b>VI</b>
<b>1. Einleitung</b>	<b>1</b>
1.1 Definitionen	1
1.2 Aktueller Stand und Problematik der Referenzwertbestimmung	2
1.3 Entwicklung der indirekten Methode zur Berechnung von Referenzwerten	4
1.4 Endokrinologische Laborparameter	5
1.5 Zielsetzung der Arbeit	6
<b>2. Material und Methoden</b>	<b>7</b>
2.1 Daten	7
2.2 Datenvorbereitung	11
2.3 Referenzwertbestimmung	11
2.4 Datennachbereitung	15
<b>3. Ergebnisse</b>	<b>16</b>
3.1 Datenvorbereitung	16
3.1.1 Gerätetestistik	16
3.1.2 Äquivalenztest	16
3.1.3 Nachkommastellen	19
3.2 Ermittlung der Referenzbereiche	21
3.2.1. Altersklasse 18 - 100 Jahre	21
3.2.2 Weitere Eingrenzung der Altersklassen	30
<b>4. Diskussion</b>	<b>40</b>
4.1 Referenzwertbestimmung mittels RLE: Bisherige Anwendungsgebiete	40
4.2 Stärken und Schwächen der Methode	41

4.3	Diskussion der Ergebnisse unter Bezugnahme auf die klinische Anwendbarkeit sowie Vergleich mit bestehenden Referenzgrenzen	44
4.4	Vitamin D: Schlussfolgerungen zu notwendigen Merkmalen eines RLE-geeigneten Laborparameters	50
4.5	Vitamin D: Einflussfaktoren auf Referenzgrenzen und zukünftige Berücksichtigungen	51
4.6	Ausblick: Weitere Anwendungsmöglichkeiten der Methode	52
<b>5.</b>	<b>Zusammenfassung</b>	<b>54</b>
<b>6.</b>	<b>Literatur</b>	<b>56</b>
<b>7.</b>	<b>Anhang</b>	<b>62</b>

## Abbildungsverzeichnis

Abb. 1: Übersicht der Standorte der einsendenden Kliniken	7
Abb. 2: a.: Äquivalenztest der häufigsten Messgeräte für Calcium b.: Äquivalenztest der häufigsten Messgeräte für Phosphat	18
Abb. 3: Äquivalenztest der häufigsten Messgeräte für Parathormon	19
Abb. 4: Histogramm der Nachkommastellen für die Messwerte von fT4	20
Abb. 5: a.: Histogramm der Nachkommastellen für die Messwerte von TSH b.: Histogramm nach der Rundung der Nachkommastellen auf eine Nachkommastelle für die Messwerte von TSH	20
Abb. 6: a.: Altersabhängiger Verlauf der 50.-Perzentile für Renin für Männer von 80-100 Jahren b.: Zeitabhängiger Median für Renin für Männer von 80-100 Jahren	21
Abb. 7: Geschätzte Verteilung der nicht-pathologischen, pathologischen Werte und für alle Daten für Renin für Männer von 18-100 Jahren	22
Abb. 8: Geschätzte Verteilung der nicht-pathologischen, pathologischen Werte und für alle Daten für Aldosteron für beide Geschlechter von 18-100 Jahren	22
Abb. 9: Altersabhängiger Verlauf der 50.-Perzentile für Calcium für beide Geschlechter von 80-100 Jahren	23
Abb. 10: Zeitabhängiger Median für Calcium für beide Geschlechter von 80-100 Jahren	23
Abb. 11: Geschätzte Verteilung der nicht-pathologischen, pathologischen Werte und für alle Daten für Calcium für beide Geschlechter von 18-100 Jahren	24
Abb. 12: a.: 50.-Perzentile für Phosphat für beide Geschlechter von 18 – 100 Jahren b.: Geschätzte Verteilung der nicht-pathologischen, pathologischen Werte und für alle Daten für Phosphat für beide Geschlechter von 18-100 Jahren	24
Abb. 13: a.: Altersabhängiger Verlauf der 50.-Perzentile für fT3 für beide Geschlechter von 18-100 Jahren b.: Zeitabhängiger Median für fT3 für beide Geschlechter von 80-100 Jahren	25
Abb. 14: Geschätzte Verteilung der nicht-pathologischen, pathologischen Werte und für alle Daten für fT3 für beide Geschlechter von 18-100 Jahren	25
Abb. 15: a.: Altersabhängiger Verlauf der 50.-Perzentile für fT4 für Frauen von 18-100 Jahren b.: Zeitabhängiger Median für fT4 für Männer von 80-100 Jahren	26
Abb. 16: a.: Zeitabhängiger Median für fT4 für beide Geschlechter von 80-100 Jahren b.: Geschätzte Verteilung der nicht-pathologischen, pathologischen Werte und für alle Daten für fT4 für Männer von 18-100 Jahren	26
Abb. 17: a.: Altersabhängiger Verlauf der 50.-Perzentile für TSH für beide Geschlechter von 18-100 Jahren. b.: Zeitabhängiger Median für TSH für beide Geschlechter von 80-100 Jahren	27
Abb. 18: Geschätzte Verteilung der nicht-pathologischen, pathologischen Werte und für alle Daten für TSH für beide Geschlechter von 18-100 Jahren	27

Abb. 19: a.: Altersabhängiger Verlauf der 50.-Perzentile für Parathormon für beide Geschlechter von 18-100 Jahren b.: Zeitabhängiger Median für Parathormon für beide Geschlechter von 80-100 Jahren	28
Abb. 20: Geschätzte Verteilung der nicht-pathologischen, pathologischen Werte und für alle Daten für Parathormon für Männer von 18-100 Jahren	28
Abb. 21: a.: Altersabhängiger Verlauf der 50.-Perzentile für Vitamin D für beide Geschlechter von 18-100 Jahren b.: Zeitabhängiger Median für Vitamin D für beide Geschlechter von 80-100 Jahren	29
Abb. 22: Geschätzte Verteilung für alle Daten für Vitamin D für beide Geschlechter von 18-100 Jahren	29
Abb. 23: a.: Geschätzte Verteilung der nicht-pathologischen, pathologischen Werte und für alle Daten für Calcium für beide Geschlechter von 18-80 Jahren b.: Geschätzte Verteilung der nicht-pathologischen, pathologischen Werte und für alle Daten für Phosphat für beide Geschlechter von 18-80 Jahren	31
Abb. 24: Geschätzte Verteilung der nicht-pathologischen, pathologischen Werte und für alle Daten für fT4 für beide Geschlechter von 18-80 Jahren	35
Abb. 25: Geschätzte Verteilung der nicht-pathologischen, pathologischen Werte und für alle Daten für Parathormon für beide Geschlechter von 18-80 Jahren für Messgerät 1	36
Abb. 26: Geschätzte Verteilung der nicht-pathologischen, pathologischen Werte und für alle Daten für Parathormon für beide Geschlechter von 18-80 Jahren für Messgerät 2	37
Abb. 27: Geschätzte Verteilung der nicht-pathologischen Werte und für alle Daten für Sommermesswerte von Vitamin D für beide Geschlechter von 18-100 Jahren	38
Abb. 28: Geschätzte Verteilung der nicht-pathologischen, pathologischen Werte und für alle Daten für Wintermesswerte von Vitamin D für beide Geschlechter von 18-80 Jahren	39

## Tabellenverzeichnis

Tab. 1: Anzahl der Messwerte pro Parameter	9
Tab. 2: Eingesetzte Messmodelle pro Parameter	9
Tab. 3: Laborparameter und zugeordnete Messgeräte sowie Details zum Messverfahren	10
Tab. 4: Erste Ergebnistabelle für die Messwerte von Calcium für die Altersklasse 18-80 Jahren	13
Tab. 5: Zweite Ergebnistabelle für die Messwerte von Calcium für die Altersklasse 18-80 Jahre, aufgeteilt nach Geschlecht und Monat	14
Tab. 6: Dritte Ergebnistabelle für die Messwerte von Calcium für die Altersklasse 18-80 Jahre	14
Tab. 7: Erster Äquivalenztest der Geräte für Phosphat mit einem Epsilon-Wert von 10%	18
Tab. 8: Referenzintervalle verschiedener Parameter für die Altersklasse 18-100 Jahre	30
Tab. 9: Referenzintervall für Calcium für die Altersklasse 18-80 Jahre	32
Tab. 10: Saisonale Referenzintervalle für Calcium für Sommer und Winter	32
Tab. 11: Referenzintervall für Phosphat für die Altersklasse 18-80 Jahre	33
Tab. 12: Saisonale Referenzintervalle für Phosphat für Sommer und Winter	33
Tab. 13: Gerätespezifische Referenzintervalle für Phosphat für die Altersklasse 18-80 Jahre	34
Tab. 14: Referenzintervalle für TSH	34
Tab. 15: Referenzintervalle für fT3	35
Tab. 16: Referenzintervalle für fT4	36
Tab. 17: Allgemeine und gerätespezifische Referenzintervalle für Parathormon von 18-80 Jahren	37
Tab. 18: Saisonale Referenzintervalle für Vitamin D für Sommer und Winter	39
Tab. 19: Gegenüberstellung der Referenzintervalle des RLE (Altersklasse 18 – 80 Jahre), des Laborsystems des UKSH, der Hersteller Cobas und Dimension und der Literatur für Calcium, Phosphat, Parathormon, fT3, fT4 und TSH	45

## Abkürzungsverzeichnis

<b>Abb.</b>	Abbildung
<b>AK</b>	Antikörper
<b>Aldo.</b>	Aldosteron
<b>Ca</b>	Calcium
<b>C-RIDL</b>	Committee on Reference Intervals and Decision Limits
<b>DGKL</b>	Deutsche Gesellschaft für Klinische Chemie und Laboratoriumsmedizin
<b>eGFR</b>	<i>estimated GFR</i> , geschätzte glomeruäre Filtrationsrate
<b>F</b>	Frauen
<b>G</b>	Geschlecht
<b>IFCC</b>	International Federation of Clinical Chemistry
<b>IU</b>	<i>international unit</i> , internationale Einheit
<b>K</b>	Kiel
<b>KI</b>	Konfidenzintervall
<b>L</b>	Lübeck
<b>M</b>	Männer
<b>N</b>	Anzahl
<b>pD</b>	<i>permissible difference</i> , zulässige Differenz
<b>Ph</b>	Phosphat
<b>PTH</b>	Parathormon
<b>RLE</b>	<i>Reference Limit Estimator</i>
<b>RI</b>	Referenzintervall
<b>Tab.</b>	Tabelle
<b>TSH</b>	Thyroidea-stimulierendes Hormon
<b>T3, fT3</b>	Trijodthyronin, freies Trijodthyronin
<b>T4, fT4</b>	Thyroxin, freies Thyroxin
<b>UKSH</b>	Universitätsklinikum Schleswig-Holstein
<b>Vit. D</b>	Vitamin D

## **1. Einleitung**

In der Medizin haben Referenzbereiche von Messwerten einen wichtigen Stellenwert. Sie dienen als Richtwerte in der Beurteilung der Laborwerte, womit diese als pathologisch oder nicht-pathologisch eingeordnet werden. Dies trägt bei zur Diagnosefindung, zum Screening von Erkrankungen, zur Beurteilung des Therapieansprechens und des Krankheitsverlaufs. Somit sind Referenzwerte unerlässlich im klinischen Alltag, um diagnostische und therapeutische Entscheidungen auf wissenschaftlicher Grundlage zu treffen (1, 2).

Aufgrund der Bedeutsamkeit hinsichtlich medizinischer Entscheidungen muss die Bestimmung von Referenzgrenzen entsprechend definierter Vorgaben erfolgen. Die empfohlene Vorgehensweise ist, dass Referenzgrenzen über eine direkte Methode von jedem Labor selbst ermittelt werden (3). Bei dieser direkten Methode wird eine Mindestanzahl von 120 gesunden Personen aus der Population, für die die Referenzgrenzen ermittelt werden, selektiert. Die Proben werden verarbeitet und ausgewertet, um aus den Messwerten Referenzgrenzen zu berechnen. Dies stellt einen sehr großen Aufwand für Labore dar. Falls das Vorgehen aus wirtschaftlichen oder logistischen Gründen nicht möglich ist, werden oftmals Referenzgrenzen von externen Quellen übernommen. Für diese transferierten Referenzgrenzen muss eine Validierung erfolgen (3). Nur so wird gewährleistet, dass die angewandten Referenzgrenzen sowohl zu den Analysemethoden als auch zu der Patientenpopulation passen. Doch auch diese Validierung wird häufig nicht vorgenommen (4).

Neben dieser aufwändigen direkten Vorgehensweise zur Bestimmung von Referenzgrenzen wurde daher auch die indirekte Methode entwickelt, welche sich die großen Datenmengen, die in Laborinformationssystemen gespeichert sind, zu Nutze macht (5–7). Diese alternative Methode ist in mehrfacher Hinsicht von Vorteil. Sie ist kostengünstig, da keine Probanden rekrutiert und untersucht werden müssen und auch der Prozess der Probensammlung entfällt. Dementsprechend ist die Methode auch logistisch weniger aufwändig (8). Zudem wird die Referenzwertermittlung für seltene Laborparameter (z. B. aus Biopsien, Liquor) oder schwierige Patientengruppen (wie etwa Frühgeborene) erleichtert.

### **1.1 Definitionen**

Um die Bedeutung einiger Begriffe zu klären, die im Folgenden verwendet werden, sei hier auf einige Definitionen hingewiesen:

Eine Referenzpopulation umfasst alle möglichen Referenzindividuen. Diese liefern Referenzwerte, um einen Vergleich mit den Messwerten von Patienten zu ermöglichen. Die Referenzindividuen werden entsprechend definierter Kriterien ausgewählt, damit die Referenzpopulation bestmöglich wiedergespiegelt wird (9). Das Referenzintervall (oder Referenzbereich genannt) wird so bestimmt, dass es 95 Prozent der Messwerte umfasst. Es wird begrenzt durch die obere und untere Referenzgrenze. Die untere Referenzgrenze entspricht dabei der 2,5. Perzentile und die obere Referenzgrenze der 97,5. Perzentile der Messwerte der Referenzindividuen. Somit liegen jeweils 2,5 Prozent der Messwerte unterhalb bzw. oberhalb der Referenzgrenzen (10). Insgesamt 5 Prozent der Messwerte befinden sich außerhalb des Referenzbereiches und sind damit größer oder kleiner als der Mittelwert zu- bzw. abzüglich zwei Standardabweichungen, wenn die Messwerte normalverteilt sind.

## **1.2 Aktueller Stand und Problematik der Referenzwertbestimmung**

Der aktuelle Goldstandard besteht darin, dass jedes Labor für jeden angebotenen Test die Referenzbereiche selbst ermittelt. Das heißt, dass für jeden Analyten und jede Stratifizierungsgruppe (z. B. nach Alter oder Geschlecht) eine eigene gesunde Gruppe von Referenzindividuen zusammengestellt wird, die dem behandelten Patientengut möglichst ähnlich ist und mindestens 120 Individuen umfassen soll (10). Diese Selektion kann entweder a priori, also prospektiv oder a posteriori, also retrospektiv, erfolgen. Bei der prospektiven Auswahl der Referenzindividuen werden Ein- und Ausschlusskriterien auf eine große Population angewandt und so die Stichprobe zusammengestellt. Die retrospektive Selektion hingegen nutzt eine bestehende Stichprobe und wendet darauf die Auswahlkriterien an (11). Dafür müssen die Ein- und Ausschlusskriterien sorgfältig definiert werden. Ausgeschlossen werden in der Regel Personen mit Erkrankungen, unter Medikation, Alkohol- und Nikotinkonsum sowie besonderen physiologischen Zuständen (z. B. Schwangerschaft). Weitere Merkmale wie Alter, Geschlecht, Ethnie, Region und genetische Aspekte werden gegebenenfalls einbezogen (11).

Mögliche Probanden werden mittels Anamneseerhebung, körperlicher Untersuchung sowie Fragebögen untersucht und es wird eine schriftliche Einwilligung eingeholt (3). Bei Einschluss sollten in Vorbereitung auf die Probensammlung einige Faktoren, die die Probanden betreffen, standardisiert werden (u. a. Ernährung, Nahrungskarenz, Ruhephase vor Probensammlung, körperliche Aktivität und Medikation) (12). Auch sollten einheitliche Maßgaben befolgt werden bei der Probensammlung (z.B. Zeit, Körperhaltung, Materialien) und der weiteren

Probenverarbeitung (z. B. Lagerung, Transport) (12). Die Probenanalyse erfolgt mittels einer qualitätskontrollierten Methode mit geringer Messunsicherheit und -ungenauigkeit (13). Die daraus ermittelten Daten werden gesichtet hinsichtlich der Verteilung, Ausreißern und weiteren Besonderheiten sowie gegebenenfalls modifiziert. Nun werden die Referenzgrenzen berechnet. Zur Anwendung kommt in der Regel eine nicht-parametrische Methode, welcher keine Annahmen über die Verteilung der Messwerte zugrunde liegen. Ist die Verteilung der Messwerte bekannt und handelt es sich um eine Gauss-Normalverteilung oder wird die Verteilung in eine solche umgewandelt, so kann eine parametrische Methode zum Einsatz kommen (10).

Alle Schritte von der Vorbereitung bis zur Probenauswertung müssen dokumentiert werden, um die Transparenz und Reproduzierbarkeit des Vorgehens zu gewährleisten. Der aufgezeigte Prozess muss theoretisch für jeden Laborparameter, für den das Labor einen Test anbietet, durchgeführt werden. Entsprechend der Anzahl der verfügbaren Tests müssten somit Tausende Probanden rekrutiert werden (14). Diese direkte Methode der Bestimmung von Referenzgrenzen erweist sich somit als sehr schwer zu realisieren. Aufgrund des großen organisatorischen als auch finanziellen Aufwands wird die Empfehlung in der Praxis kaum umgesetzt und Referenzwerte meist von Geräteherstellern oder aus Literaturquellen übernommen (4).

Eine Möglichkeit, den aufwändigen Weg der Referenzwertbestimmung zu vermeiden, ist, bereits ermittelte Referenzgrenzen zu übernehmen. Die Grenzen können von den Herstellern der Analysengeräte selbst stammen, von anderen Laboren übernommen werden oder es handelt sich um etablierte Referenzwerte aus der Literatur. Um diese Referenzgrenzen auf das Patientenkollektiv des jeweiligen Labors und die verwendeten Geräte anwenden zu können, sollte eine Validierung erfolgen. Dafür sollte eine kleine Gruppe von 20 Referenzindividuen, die dem behandelten Patientenkollektiv entstammt, zusammengestellt und die resultierenden Messwerte mit den Referenzgrenzen verglichen werden (3). Wenn mehr als 10% der Messwerte außerhalb der Grenzen liegen, wird eine zweite Referenzgruppe von 20 Personen untersucht. Bei über 10% abweichenden Messwerten, muss eine Ursachensuche stattfinden (3). Oftmals erfolgt der Transfer der Referenzwerte jedoch ohne die nötige Validierung (4), woraus sich Schwierigkeiten ergeben können. Denn dafür müssten vergleichbare prä- und analytische Vorgehensweisen sowie vergleichbare Patientenpopulationen gegeben sein (15). Unterschiede darin können dazu führen, dass die angewandten Referenzgrenzen für das Patientenkollektiv unpassend sind. Den in der Literatur veröffentlichten Referenzgrenzen mangelt es teilweise an

Informationen und auch methodische Anforderungen zur Referenzwertermittlung werden oft nicht erreicht (16). Dabei sind Angaben bezüglich der Auswahl, Größe und Zusammensetzung der Referenzpopulation (wie etwa Alter, Geschlecht und Ethnie) dringend nötig, um die Übertragbarkeit beurteilen zu können. Auch die Referenzgrenzen, die die Gerätehersteller herausgeben, finden Anwendung in Laboren. Dieser Transfer ist ebenso kritisch zu betrachten. Meist handelt es sich um kleine, nordamerikanische Referenzpopulationen mit fehlenden Angaben zum Gesundheitsstatus (17).

### **1.3 Entwicklung der indirekten Methode zur Berechnung von Referenzwerten**

Die direkte Methode zur Referenzwertbestimmung ist die empfohlene und zu bevorzugende Option, doch aufgrund des großen Aufwands des Verfahrens, wurden schon früh Versuche unternommen, auf indirektem Wege Referenzbereiche zu erstellen (5). Verwendet werden dabei Daten von Messwerten aus Krankenhäusern, die in Laborinformationssystemen gespeichert sind und in großem Umfang zur Verfügung stehen. Ihre Quantität nimmt dabei stetig zu, auch bedingt durch die zunehmende Digitalisierung des Gesundheitswesens. Die Datensätze umfassen sowohl nicht-pathologische als auch pathologische Messwerte. Bei der statistischen Verarbeitung dieser Daten sollten jedoch keine parametrischen oder nicht-parametrischen Techniken wie bei der direkten Bestimmung von Referenzgrenzen eingesetzt werden, da diese stark von Ausreißern und pathologischen Messwerten beeinflusst werden (18). Daher erfordern solch retrospektive Daten andere statistische Ansätze. Hoffmann entwickelte ein manuelles grafisches Vorgehen, das auf Wahrscheinlichkeitspapieren basiert und wobei Daten mit einer Normalverteilung verglichen werden (19). Problematisch ist daran, dass die Schätzung der Referenzgrenzen verzerrt wird, sobald eine zweite und dritte Subgruppe, welche außerhalb des Referenzbereiches liegen, sehr groß sind (20). Eine weitere grafische Methode wurde von Bhattacharya entwickelt (21). Dabei wird die gemischte Verteilung in die Verteilungen ihrer pathologischen und nicht-pathologischen Messwerte aufgelöst. Es wird angenommen, dass diese Subgruppen einer Gauß'schen Normalverteilung folgen. Es wurden verschiedene Abwandlungen vorgenommen und die Methode zur Ermittlung von Referenzgrenzen eingesetzt (15). Es zeigte sich, dass überlappende Verteilungen voneinander getrennt werden konnten und vereinzelte Messwerte, die nicht der Normalverteilung entsprachen, einen geringeren Einfluss hatten (20). Doch das Vorgehen nach Bhattacharya hat auch einige Schwächen. Einerseits ist es abhängig vom Nutzer und seiner Auswahl der Klassengröße und -anzahl (18), andererseits ist die Parameterschätzung der Verteilung ungenau

und auch erfolgt keine Überprüfung der geschätzten Verteilung hinsichtlich ihrer Übereinstimmung mit den tatsächlichen Daten (22). Eine neue indirekte Vorgehensweise zur Referenzwertbestimmung stellt das Programm Reference Limit Estimator (RLE) dar. Dieses wurde von F. Arzideh und der Arbeitsgruppe Richtwerte der Deutschen Gesellschaft für Klinische Chemie und Laboratoriumsmedizin (DGKL) 2007 entwickelt und zur Verfügung gestellt (23). Der RLE macht sich ebenfalls die ständig anfallenden laborinternen Messwerte von Patienten zunutze und ermittelt aus dem gemischten Datensatz Referenzgrenzen (zur Funktionsweise des RLE s. Abschnitt 2.3 Referenzwertbestimmung). Die Arbeitsgruppe C-RIDL (Committee on Reference Intervals and Decision Limits) der IFCC (International Federation of Clinical Chemistry) unterstützt die Anwendung indirekter Methoden zur Beurteilung von Referenzgrenzen und zur Schätzung neuer Referenzgrenzen (18).

## 1.4 Endokrinologische Laborparameter

In dieser Arbeit werden Referenzgrenzen für einige endokrinologische Laborparameter bestimmt. Die Endokrinologie beschäftigt sich mit den Hormonen und ihren Wirkmechanismen. Hormone entstammen den verschiedensten Drüsen, dazu gehören unter anderem die Schilddrüse, die Nebenschilddrüsen und die Hypophyse (24). Parathormon (PTH) entstammt den Nebenschilddrüsen und dient als Regulator des Knochenstoffwechsels. Für diesen Parameter werden Referenzgrenzen in dieser Arbeit erstellt. Seine Hauptwirkung besteht darin, die Calciumwerte im Blut zu erhöhen. PTH steigert die Calciummobilisation aus dem Knochen und die Calciumreabsorption in der Niere. Gleichzeitig wird renal die Rückresorption von Phosphat gehemmt. Auch Calcitriol wird vermehrt gebildet (25). Daher werden auch Vitamin D, Calcium und Phosphat in diese Studie einbezogen und Berechnungen zur Bestimmung von Referenzgrenzen vorgenommen. Der Calciumhaushalt wird durch PTH, 1,25-OH-Vitamin D (Calcitriol) und Calcitonin hormonell gesteuert. Die Phosphathomöostase wird vorwiegend über PTH und Vitamin D reguliert (26). Vitamin D, genauer gesagt die biologisch aktive Form Calcitriol, steigert die Calciumaufnahme intestinal und moduliert die Phosphatresorption als auch die Knochenmineralisation. Mit steigendem Calciumspiegel fällt die PTH-Konzentration. Außerdem hat es positive Effekte u. a. auf die Glucosetoleranz, das Herz-Kreislauf-System und das Immunsystem. Die Vorstufe von Vitamin D wird aus der Nahrung aufgenommen oder vom Körper gebildet. Zunächst wird 7-Hydroxycholesterol im Darm aufgenommen. Als 7-Dehydrocholesterol wird es in der Haut über UV-Licht zu Cholecalciferol (Vitamin D<sub>3</sub>) umgewandelt. Durch eine Hydroxilierung entsteht in der Leber 25-OH-Vitamin D (Calcidiol). In der Niere kommt es zur zweiten Hydroxilierung, wodurch

1,25-OH-Vitamin D (Calcitriol) gebildet wird. Die häufigste Form im Blut stellt 25-OH-Vitamin D dar, welches bei der Bestimmung der Vitamin D-Konzentration gemessen wird (25). Auch in den folgenden Berechnungen für den Laborparameter „Vitamin D“ handelt es sich um 25-OH-Vitamin D.

Des Weiteren wurden Referenzgrenzen für die Schilddrüsenhormone (fT3 und fT4) und TSH bestimmt. Schilddrüsenhormone fördern die Entwicklung der verschiedenen Körpersysteme (z. B. Skelett, Nervensystem) und steuern den Stoffwechsel. TSH bewirkt an der Schilddrüse die Freisetzung von Thyroxin (Tetrajodthyronin, T4) und Trijodthyronin (T3), welche in freier oder gebundener Form auftreten. Physiologisch wirksam ist nur die freie Form (fT3 bzw. fT4) (25).

Einzelheiten zu den Messungen der Laborparameter, wie etwa eingesetzte Geräte und das Messprinzip, lassen sich der Tabelle 1 entnehmen.

## 1.5 Zielsetzung der Arbeit

Mittels des RLE werden Referenzgrenzen für Erwachsene für verschiedene Parameter aus dem erweiterten endokrinologischen Formenkreis (Calcium, Phosphat, TSH, fT3, fT4, Vitamin D, Parathormon, Renin, Aldosteron) berechnet. Dabei kommen Messwerte von Patienten aus dem Zentrallabor des Universitätsklinikums Schleswig-Holstein (UKSH) zum Einsatz. Somit werden mit einer indirekten Methode retrospektiv Referenzgrenzen errechnet. Dabei werden auch geschlechts- und altersspezifische Referenzgrenzen ermittelt. Weitere Stratifizierungen anhand anderer Kriterien werden gegebenenfalls vorgenommen, wenn sich diese im Verlauf als sinnvoll erweisen.

Die Anwendungsmöglichkeit der Methode wird untersucht und die klinische Relevanz der ermittelten Referenzgrenzen betrachtet. Die Ergebnisse werden einem Vergleich mit etablierten Referenzbereichen unterzogen, um die Fragestellung zu beantworten, ob sich eine sinnvolle Alternative zur aufwändigen Referenzwertbestimmung bietet. Ursachen für Unterschiede zwischen den etablierten und den berechneten Referenzgrenzen werden diskutiert und Möglichkeiten zur weiteren Optimierung des Vorgehens werden erörtert.

## 2. Material und Methoden

### 2.1 Daten

Zur Ermittlung der Referenzgrenzen für Calcium, Phosphat, TSH, fT3, fT4, Parathormon, Renin, Aldosteron und Vitamin D werden insgesamt ca. 1,5 Millionen Messwerte inklusive Patientendaten des Zentrallabors des Universitätsklinikums Schleswig-Holstein (UKSH) verwendet. Das UKSH mit seinen beiden Standorten Kiel und Lübeck ist ein Maximalversorger und umfasst 2.500 Betten. Jährlich werden etwa 450.000 Patienten behandelt. Das Zentrallabor bedient zudem weitere Krankenhäuser im Großraum von Norddeutschland, sodass die Daten auch aus Hamburg, Schleswig, Schönhagen, Damp, Elmshorn, Pinneberg, Wedel, Timmendorfer Strand, Lehmrade, Wismar, Gifhorn, Hildesheim, Lengerich, Bad Gandersheim, Northeim, Bad Herzberg/Osterode und Göttingen stammen (s. Abb. 1).



**Abbildung 1:** Übersicht der Standorte der einsendenden Kliniken. Ausgenommen sind, bis auf Hildesheim und Gifhorn, die Kliniken im grauen Oval. Stand: 20.06.2017.

Der Datensatz, der für diese Studie aus dem Laborinformationssystem extrahiert wurde, enthält neben den Messergebnissen als Zahlenwert mit bis zu drei Nachkommastellen weitere Angaben: die Art der Behandlung (angegeben als Buchstabe a, v, s, h, n für ambulant, vorstationär, stationär, halbstationär, nachstationär), das Alter (in ganzen Jahren), das Geschlecht (angegeben als F und M für Frauen bzw. Männer), das Einschleusdatum (als dd.mm.yyyy, dabei d = Tag, m = Monat, y = Jahr), die Einschleuszeit (als hh:mm:ss; wobei h = Stunden, m = Minuten und s = Sekunden), die Analysennummer (als sechsstelliger

Zahlenwert), die Einheit (bei Calcium, Phosphat: mmol/l; Aldosteron, Renin: ng/l; TSH: mIU/l; fT3, fT4: pmol/l; Vitamin D: µg/l; bei Parathormon: pg/ml), der Referenzbereich (als Zahlenintervall mit oberer und unterer Referenzgrenze), Probenart (Calcium, Phosphat, TSH, fT3, fT4: Heparinplasma; Renin, Parathormon: EDTA-Plasma; Aldosteron, Vitamin D: Serum), die Gerätenummer (als fünfstelliger Zahlenwert), Anzahl der Messungen pro Fallnummer (als Zahlenwert) und der Angabe der wievielen Messung in dem Fall (Zahlenwert). Auch vermerkt ist, wenn es sich um einen gynäkologischen oder intensivmedizinischen Fall handelt. Bei Parathormon werden zusätzlich Daten von neurologischen und urologischen Fällen gekennzeichnet.

Ausgeschlossen werden Wiederholungsmessungen, da es sich dabei wahrscheinlich um Werte von Patienten handelt, die stärker erkrankt sind, sowie Messwerte, die im Rahmen von Studien erhoben wurden. Letztere entsprechen nicht dem durchschnittlichen Krankenhauspatientenkollektiv. Messwerte von gynäkologischen Stationen und Intensivstationen werden ausgeschlossen, da dort ein überdurchschnittlich hoher Anteil pathologischer Werte zu erwarten ist, was die Berechnung der Referenzbereiche beeinflussen könnte. Zum einen hat die Schwangerschaft einen deutlichen Einfluss auf die laborchemischen Parameter, zum anderen der kritische Gesundheitszustand, in dem sich Patienten der Intensivstation befinden. Bei Parathormon wurden außerdem urologische Messwerte ausgeschlossen. Bei urologischen Fällen sind oftmals Patienten mit Niereninsuffizienz vertreten, was bewirkt, dass die renale Syntheseleistung von aktivem Vitamin D eingeschränkt ist. Erhöhte Parathormon-Konzentrationen sind die Folge.

Die Messwerte, die im Rahmen regulärer Untersuchungen erhoben wurden, entstammen dem Zeitraum März 2014 bis Dezember 2016. Initial sind 1.582.503 Werte verfügbar, nach den genannten Ausschlüssen und der Altersbegrenzung auf 18 bis 100 Jahre sind es 1.467.580. Besonders Calcium weist einen sehr hohen Datenumfang mit fast 700.000 Messwerten auf, während für Renin und Aldosteron jeweils nur etwa 3.000 Messwerte verfügbar sind (s. Tab. 1).

**Tabelle 1:** Anzahl der Messwerte pro Parameter. Ausgeschlossen wurden Messwerte von gynäkologischen und intensivmedizinischen Fällen sowie Wiederholungsmessungen. Bei Parathormon zudem Ausschluss urologischer Fälle. Altersbegrenzung 18-100 Jahre. N = Anzahl.

Parameter	N absolut	N nach Ausschluss
Calcium	670.001	629.642
Phosphat	296.285	270.496
TSH	317.023	301.115
fT3	118.303	109.918
fT4	121.714	112.453
Parathormon	25.193	18.548
Vitamin D	27.980	21.104
Renin	3.205	1.739
Aldosteron	2.799	2.565

Für jeden Analyten kommen für die Durchführung der Messungen ein bis drei Gerätemodelle zum Einsatz, welche teilweise in mehrfacher Ausführung vorhanden sind (s. Tab. 2). Weiterführende Informationen zu den jeweils eingesetzten Messgeräten und den Messverfahren können Tabelle 3 entnommen werden.

**Tabelle 2:** Eingesetzte Messmodelle pro Parameter.

Parameter	Messmodelle
Calcium	Cobas 8000, Cobas 8100
Phosphat	Cobas 8000, Cobas 8100, Dimension
TSH	Cobas 8000, Cobas 8100
fT3	Cobas 8000, Cobas 8100
fT4	Cobas 8000, Cobas 8100
Parathormon	Cobas 8000, Cobas 8100
Vitamin D	HPLC
Renin	BEP 2000
Aldosteron	BEP 2000

**Tabelle 3:** Laborparameter und zugeordnete Messgeräte sowie Details zum Messverfahren. AK = Antikörper.

Laborparameter	Messgerät (Hersteller)	Probenvolumen	Materialien	Testprinzip	Einheit
<b>Calcium</b>	Cobas 8000 (Roche)	2 µl	Puffer, o-Kresolphthalein-Komplexon und 8-Hydrochinolin	Komplexbildung (Methode nach Schwarzenbach); photometrische Messung	mmol/l
<b>Phosphat</b>	Cobas 8000 (Roche), Dimension (Siemens)	2,5 µl bzw. 3 µl	Schwefelsäure, Ammoniummolybdat, <i>Dimension zusätzlich: P-</i> Methylaminophenolsulfat (PMAPS) und Bisulfit	Komplexbildung; photometrische Messung	mmol/l
<b>Parathormon</b>	Cobas 8000 (Roche)	50 µl	biotinylierter monoklonaler PTH-spezifischer AK, Ruthenium-Komplex markierter monoklonaler PTH-spezifischer AK und Streptavidin-Mikropartikel	Sandwichkomplexbildung; unter Spannung Chemilumineszenzemission, Messung mittels Photomultiplier	pg/ml
<b>Vitamin D</b>	Cobas 8000 (Roche)	15 µl	Dithiothreitol, Natriumhydroxid, Ruthenium-markiertem Vitamin D-Bindeprotein, Streptavidin-Mikropartikel, biotinyliertes Vitamin D (25-OH)	Kompetitionsprinzip; unter Spannung Chemilumineszenzemission, Messung mittels Photomultiplier	µg/l
<b>TSH</b>	Cobas 8000 (Roche)	50 µl	Biotinylierter TSH-spezifischer AK, Ruthenium-markierter TSH-spezifischer AK, Streptavidin-Mikropartikel, Biotin-haltige Festphase	Sandwichkomplexbildung; unter Spannung Chemilumineszenzemission, Messung mittels Photomultiplier	mIU/l
<b>fT3, fT4</b>	Cobas 8000 (Roche)	15 µl	Ruthenium-markierter, T4-spezifischer AK, biotinyliertes T4, Streptavidin-Mikropartikel, Biotin-haltige Festphase	Kompetitionsprinzip; unter Spannung Chemilumineszenzemission, Messung mittels Photomultiplier	pmol/l

## **2.2 Datenvorbereitung**

Zur Vorbereitung und Analyse der Daten werden folgende R-Programme, bereitgestellt vom Institut für Klinische Chemie, Kiel, verwendet: Gerätetestistik, Äquivalenztest und Nachkommastellentest.

Das Programm Gerätetestistik filtert jeden Datensatz nach Gerätenummer, Frequenz und Gerätenamen. Der restltierenden Tabelle ist so zu entnehmen, wie viele verschiedene Geräte pro Analyt eingesetzt wurden und die Anzahl der Messungen pro Gerät.

Mit dem Programm Äquivalenztest wird überprüft, ob die verschiedenen Geräte, die für den gleichen Analyten Messungen vorgenommen haben, als äquivalent und somit vergleichbar betrachtet werden können. Dabei werden die Geräte untersucht, mit denen am häufigsten Messungen stattgefunden haben. Das Programm wendet den TOST (two one-sided-t-tests = zwei einseitige t-Tests) (27) an. Dies ist ein statistischer Hypothesentest. Dabei werden eine obere und eine untere Äquivalenzgrenze festgelegt. Es werden zwei Nullhypotesen aufgestellt, die vom Vorliegen eines Effektes, also einer relevanten Differenz zwischen den beiden zu vergleichenden Verfahren (z. B. Methoden oder Medikamente), ausgehen. Liegt der Effekt innerhalb der Äquivalenzgrenzen, werden die Nullhypotesen abgelehnt und die zwei Verfahren als äquivalent angesehen (28). Die Äquivalenzgrenzen (Epsilon) werden zunächst auf 10% festgelegt und die Konfidenzintervalle berechnet. Es werden jeweils zwei Geräte gegenübergestellt und als äquivalent bzw. nicht äquivalent bewertet.

Grafisch dargestellt wird die Verteilung der Nachkommastellen mit dem Programm Nachkommastellentest (s. Abschnitt 3.1.2 Äquivalenztest, Abb. 2). So kann beurteilt werden, ob sich die signifikanten Stellen eines Parameters unterscheiden. Sie entstehen dadurch, dass Labore die Messwerte auf eine bestimmte Weise runden, um sie mit gleicher Zifferanzahl darzustellen. Das führt im Datensatz zu unterschiedlicher Anzahl an Nachkommastellen je nach Größenordnung des Analyten, was dazu führt, dass sich die Verteilung zu stark von einer Normalverteilung unterscheidet. Bei Vorliegen unterschiedlicher signifikanter Stellen in einem Datensatz wird der komplette Datensatz eines Analyten auf zwei oder eine Nachkommastelle oder auf ganze Zahlen gerundet. Die weiteren Berechnungen werden dann mit den gerundeten Daten fortgeführt.

## **2.3 Referenzwertbestimmung**

Die Referenzbereiche werden mit dem Reference Limit Estimator (RLE) von Arzideh et al. der DGKL berechnet (23). Dieses Programm ist eine Microsoft Excel-Anwendung, die auf das

Statistikprogramm R zugreift. Es wird Version 20151017 verwendet, mittlerweile sind neuere Versionen verfügbar. Für R (Version 2.7.2) werden entsprechend den Installationsempfehlungen weitere Pakete heruntergeladen (geoR, msm, mgcv, Sp, splancs, Random Fields, mvtnorm, expm, mass, nlme, survival).

Der RLE geht mittels eines bimodalen, retrospektiven Verfahrens vor. Zunächst werden die Labordaten in Excel überführt, wobei die Messergebnisse mit Alter, Geschlecht, Einsender und einer anonymen Patientenidentifikation enthalten sein müssen. Nun werden die Daten von dem Tool eingelesen. Bei Wiederholungsmessungen an einem Patienten wird nur der erste Messwert berücksichtigt. Langfristige Drifteffekte werden geprüft. Der RLE führt für den gemischten Datensatz (enthält nicht-pathologische und pathologische Werte) eine nicht-parametrische, geglättete Kerndichteschätzung durch (29). Es wird angenommen, dass die nicht-pathologischen Werte den Großteil darstellen und eine Power-Normalverteilung sie modellieren kann. Letztere eignet sich gut für gestreute Daten (30). Falls keine Normalverteilung vorliegt, erfolgt eine logarithmische oder Box-Cox-Transformation (31). Bei den pathologischen Werten handelt es sich meist um Extremwerte (besonders hoch oder niedrig), welche nicht-parametrisch verteilt sind (29). Der überwiegende Teil dieser Werte lässt sich über zwei Trennungspunkte abschneiden. Um diese Trennungspunkte zu finden, wird der Datensatz an verschiedenen Stellen abgeschnitten (trunkiert). Für den verbleibenden, zentralen Bereich des Datensatzes wird die Parameterschätzung einer Power-Normalverteilung durchgeführt, was mittels der Maximum-Likelihood-Methode geschieht. Es wird die geschätzte Verteilung der nicht-pathologischen Werte generiert und die Güte der Anpassung dieser innerhalb und außerhalb des trunkierten Abschnitts erhoben. Dafür kommt die Kolmogorow-Smirnow-Statistik zum Einsatz. Ergibt diese niedrige Werte, so kann von einer guten Anpassung ausgegangen werden. Die niedrigsten Werte ergeben somit die besten Trennungspunkte. Danach erfolgt die Berechnung der Referenzgrenzen, also der 2,5. und 97,5. Perzentile der Verteilung der nicht-pathologischen Werte. Die Berechnung von Konfidenzintervallen folgt einer parametrischen Methode und wenn keinerlei Überschneidung dieser vorliegt, werden die verglichenen Referenzgrenzen als signifikant verschieden angesehen (23).

Die Ermittlungen erfolgen in drei Schritten. Im ersten Schritt wird die Altersspanne festgelegt. Es wird eine Ergebnistabelle generiert, die Auskunft gibt über die Anzahl der Messwerte, den Modalwert, den Durchschnitt, den Median, die Standardabweichung, den Minimal- und Maximalwert. Diese Werte werden für Männern und Frauen gemeinsam sowie nach

Geschlechtern getrennt angegeben (s. beispielhaft Tabelle 4). Außerdem werden das Histogramm mit Wahrscheinlichkeitsdichtefunktion und die 50. Perzentile in Abhängigkeit vom Alter grafisch dargestellt.

**Tabelle 4:** Erste Ergebnistabelle für die Messwerte von Calcium für die Altersklasse 18-80 Jahre. Mode = Modalwert, Mean = Durchschnittswert, Std = Standardabweichung, Min = Minimalwert, Max = Maximalwert, F = Frauen, M = Männer.

<b>Geschlecht</b>	<b>Alter</b>	<b>Anzahl</b>	<b>Mode</b>	<b>Mean</b>	<b>Median</b>	<b>Std</b>	<b>Min</b>	<b>Max</b>
F, M	18 - 80	556550	2.33	2.29	2.3	0.17	0.28	4.94
F	18 - 80	251724	2.36	2.3	2.31	0.17	0.43	4.61
M	18 - 80	304706	2.33	2.28	2.29	0.17	0.28	4.94

Im nächsten Schritt werden die 0,05.-, 0,25.-, 0,50.-, 0,75.- und 0,95.-Quantilen, aufgeteilt nach Geschlecht und Monat, berechnet (s. Tabelle 5). Grafisch dargestellt wird der Median in Abhängigkeit von der Zeit mit Konfidenzintervall. Auch die anderen Perzentilen können sich grafisch darstellen lassen. Damit sollen die Daten optisch und statistisch auf einen zeitabhängigen Drifteffekt untersucht werden, welcher bei Unstetigkeiten des Medians wahrscheinlich ist.

Im letzten Schritt werden die Referenzgrenzen berechnet. Einer Tabelle lassen sich Geschlecht, Alter, Anzahl der Messwerte, Modalwert, p (geschätzter Anteil nicht-pathologischen Werten an dem Datensatz), mu, sigma, lambda, Testwert, kritischer Wert, p-Wert, die obere und untere Referenzgrenze und die obere und untere Grenze des 90%-Konfidenzintervalls entnehmen (s. Tabelle 6). Mu, Sigma und lambda sind geschätzte Werte einer Power-Normalverteilung für die Verteilung der nicht-pathologischen Daten. Der Testwert dient der Überprüfung, ob bei der transformierten, trunkierten Stichprobe eine Normalverteilung vorliegt. Der kritische Wert darf vom Testwert nicht überschritten werden, da sonst die Hypothese, dass die Daten im trunkierten Bereich nach Transformation einer Normalverteilung folgen, abgelehnt werden muss. Der p-Wert stellt wiederum die Irrtumswahrscheinlichkeit dar. Ist er größer als der festgelegte Wert von 5%, wird angenommen, dass die Daten nach Transformation einer Normalverteilung folgen. Diese und weitere Erläuterungen der Ergebnisse sind im Handbuch zum RLE zu finden (32).

**Tabelle 5:** Zweite Ergebnistabelle für die Messwerte von Calcium (in mmol/l) für die Altersklasse 18-80 Jahre, aufgeteilt nach Geschlecht und Monat. F = Frauen, M = Männer.

		März 2014	April 2014	Mai 2014	Juni 2014
<b>Anzahl</b>	F	2566	3414	3435	3580
	M	2770	3963	4185	3976
	F, M	5336	7379	7626	7558
<b>0.05-Quantile</b>	F	1.99	1.97	1.96	1.96
	M	1.97	1.98	1.94	1.97
	F, M	1.98	1.97	1.95	1.96
<b>0.25-Quantile</b>	F	2.18	2.17	2.17	2.16
	M	2.14	2.16	2.15	2.14
	F, M	2.16	2.16	2.16	2.15
<b>0.50-Quantile</b>	F	2.29	2.28	2.27	2.27
	M	2.27	2.27	2.26	2.26
	F, M	2.28	2.28	2.27	2.27
<b>0.75-Quantile</b>	F	2.39	2.38	2.37	2.37
	M	2.37	2.365	2.36	2.36
	F, M	2.38	2.37	2.36	2.36
<b>0.95-Quantile</b>	F	2.54	2.52	2.5	2.49
	M	2.52	2.5	2.5	2.48
	F, M	2.53	2.51	2.5	2.49

**Tabelle 6:** Dritte Ergebnistabelle für die Messwerte von Calcium (in mmol/l) für die Altersklasse 18-80 Jahre. Mode = Modalwert, RG = Referenzgrenze, F = Frauen, M = Männer.

Geschlecht	Anzahl	Mode	P	mu	sigma	lambda	untere RG	obere RG
F, M	556550	2,33	0,8452	1,32536	0,11954	1	2,09	2,56
F	304706	2,34	0,8475	1,33405	0,1159	1	2,11	2,56
M	251724	2,32	0,839	1,31894	0,12191	1	2,08	2,56

Bei der Durchführung der oben genannten Schritte wurde initial die Altersspanne 18 bis 100 Jahre gewählt, dies wurde später auf 18 bis 80 Jahre angepasst. Anschließend wurde in kleinere Altersabschnitte unterteilt. Es wurden Referenzgrenzen für alle Patienten gemeinsam, sowie getrennt nach Geschlecht jeweils für Männer und für Frauen berechnet.

## 2.4 Datennachbereitung

Im Anschluss wird für jeden Analyten geprüft, ob getrennte Referenzintervalle für Männer und Frauen nötig sind oder ein gemeinsames Intervall für beide Geschlechter ausreichend ist. Auch geprüft wird, ob eine Unterteilung der Daten in engere Altersgruppen sinnvoll ist oder ob Referenzgrenzen für eine große Altersspanne ausreichend sind.

Dazu werden mittels der Excel-Anwendung „Zulässige Messunsicherheit“ der Arbeitsgruppe Richtwerte der DGKL (Deutsche Gesellschaft für Klinische Chemie und Laboratoriumsmedizin) die zulässigen Unterschiede in den Referenzgrenzen der unterteilten Population ermittelt (33). Die ermittelten Referenzgrenzen der Altersgruppe 18 bis 80 Jahre für beide Geschlechter und der zugehörige Median werden verwendet, um damit die zulässige Differenz pD (= permissible difference) für die obere und die untere Referenzgrenze zu errechnen. Eine Auflistung der angewandten Formeln findet sich im Anhang (s. Abb. 33). Es ergibt sich somit ein pD-Bereich jeweils für die obere und untere Referenzgrenze. Liegen die ermittelten oberen und unteren Referenzgrenzen aufgeteilt nach Geschlecht oder in kleinere Altersgruppen außerhalb dieser pD-Bereiche, zeugt dies von einem großen Unterschied in den verglichenen Referenzbereichen und eine entsprechende Aufteilung kann erwogen werden.

### **3. Ergebnisse**

#### **3.1 Datenvorbereitung**

##### **3.1.1 Gerätestatistik**

Zur Messung der Blutparameter kommen unterschiedliche Geräte zum Einsatz. Um sicherzustellen, dass diese vergleichbar und als äquivalent einzustufen sind, wird der Äquivalenztest angewandt. Zuvor liefert das R-Programm „Gerätestatistik“ die Information, welche Messgeräte pro Analyt wie oft eingesetzt wurden. Es wird verwendet, um danach beim Äquivalenztest die häufigsten Messgeräte miteinander vergleichen zu können. Zwar handelt es sich dabei in den meisten Fällen um das gleiche Modell, den Cobas8000 von Roche, aber um unterschiedliche Geräte. In Kiel und Lübeck gibt es jeweils zwei Geräte (K1, K2 und L1, L2). Jedes Gerät wiederum besitzt verschiedene Module. Die Messwerte von Calcium beispielsweise stammen von 13 verschiedenen Modulen. Die fünf häufigsten Module haben für Calcium jeweils zwischen 25.460 und 190.023 Messwerte erhoben. Damit stellen sie 669.092 von insgesamt 670.001 Werten, was 99,86% entspricht. Die restlichen Geräte tragen somit nur einen sehr geringen Anteil von 0,14% an Daten bei und sind zu vernachlässigen (s. Anhang Tab. 29).

##### **3.1.2 Äquivalenztest**

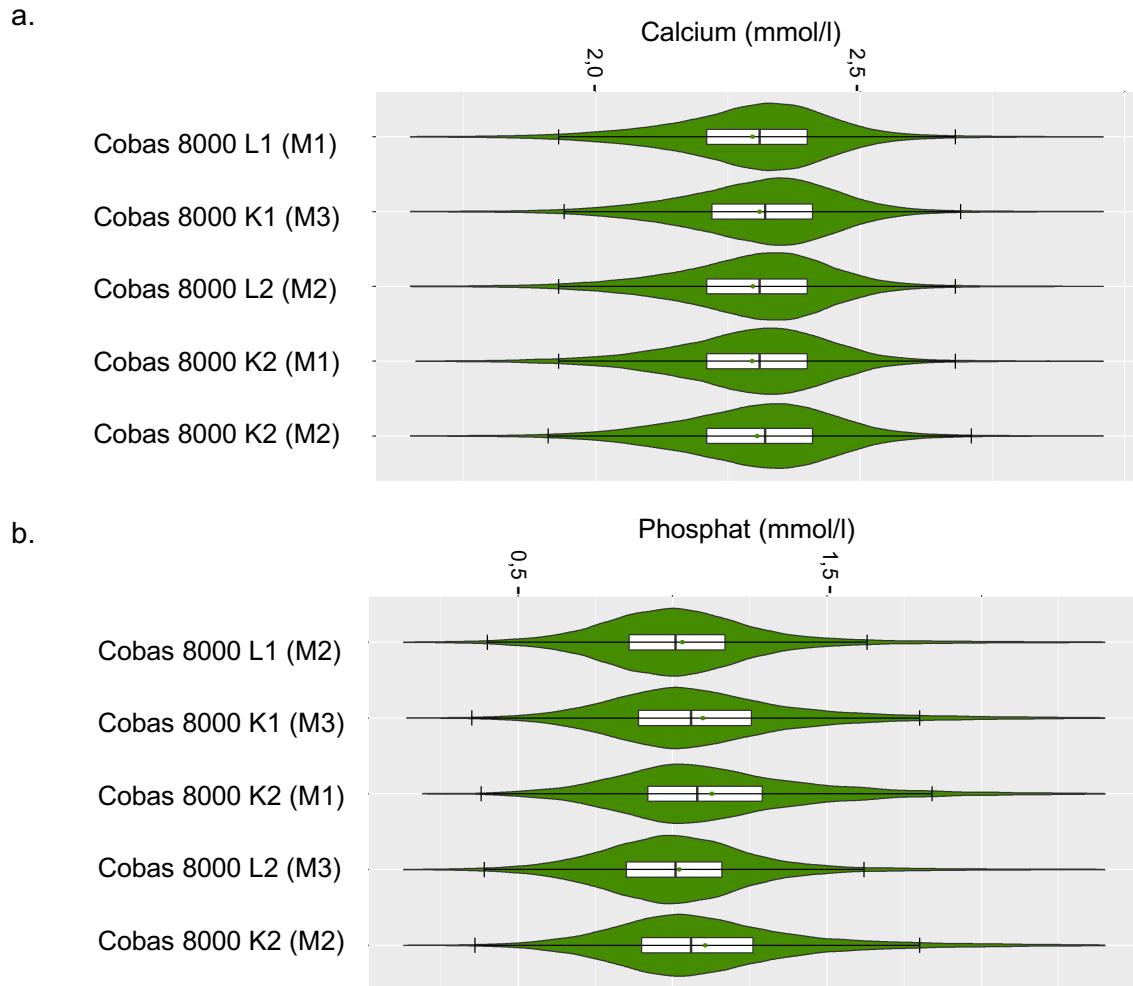
Im nächsten Schritt, dem Äquivalenztest, werden die häufigsten Geräte miteinander verglichen. Jeweils zwei Geräte stehen sich im Vergleich gegenüber, sodass sich bei fünf verschiedenen Geräten zehn Vergleiche ergeben. Zunächst wird ein Epsilon-Wert von 10% gewählt. Dieser Wert gibt den maximalen akzeptablen Unterschied zwischen zwei zu vergleichenden Geräten an. Es ergeben sich aus dem +/- Epsilon-Wert die Äquivalenzgrenzen, welche das Epsilonintervall von -0,1 – 0,1 definieren. Gleichzeitig werden zwei Nullhypthesen aufgestellt, die einen relevanten Unterschied zwischen den beiden Geräten erwarten. Es wird ein Konfidenzintervall berechnet, welches mit einer Wahrscheinlichkeit von 95 % zur Ablehnung der Nullhypthesen führt. Das Konfidenzintervall wiederum muss im Epsilonintervall liegen. Ist dies der Fall, ist die Differenz der Geräte geringer als der Epsilon-Wert und die Geräte werden als äquivalent bewertet.

Neben den Ergebnistabellen werden Grafiken generiert, welche mittels Box-Plots die Verteilung der Messwerte darstellen. Jede Box wird durch das obere und untere Quartil (25. bzw. 75. Perzentil) begrenzt und beinhaltet somit 50% der Messwerte. Der Abstand

zwischen oberem und unterem Quartil ist der Interquartilenabstand, der Strich innerhalb der Box stellt den Median dar. Die Antennen repräsentieren Daten, welche nicht innerhalb der Box liegen und sind auf den 1,5-fachen Wert des Interquartilenabstandes begrenzt. Durch die Ergänzung der Box-Plots um Violinen-Plots kann die ganze Verteilung der Messwerte eines Messgerätes durch Dichtediagramme visualisiert werden.

In der grafischen Gegenüberstellung der häufigsten Messgeräte von Calcium stellen sich die Box- und Violinen-Plots sehr ähnlich dar (s. Abb. 2a). Es zeigen sich Interquartilenabstände und Dichtediagramme, die sehr ähnliche Wertebereiche abdecken. Dabei weist Modul 2 des Gerätes Cobas 8000 K2 den weitesten Interquartilenabstand auf. Insgesamt scheinen die verschiedenen Messgeräte vergleichbare Messwerte zu liefern, was durch die Berechnungen bestätigt wird. Die Geräte werden bei einem Epsilon-Wert von 10% als äquivalent eingestuft (s. Anhang, Tab. 30).

Im Vergleich der Geräte, welche Messungen von Phosphat vorgenommen haben, zeigen sich schon grafisch größere Unterschiede. Die Boxen variieren in ihrer Höhe und in der Abdeckung ihres Messwertebereiches stärker (siehe Abb. 2b). Die Messgeräte sind wahrscheinlich nicht als äquivalent einzuordnen. Auch in den Berechnungen liegen nicht bei allen Vergleichen die Konfidenzintervalle in den entsprechenden Epsilonintervallen. Die zehn Vergleiche liefern zwei Mal ein nicht äquivalentes Ergebnis (s. Tab. 7). Erst bei einem Epsilonwert von 15% würden schließlich alle Geräte als vergleichbar bewertet werden. Aus diesem Grund werden bei Phosphat neben der standardmäßigen Referenzwertberechnung unter Einschluss aller Geräte noch einmal gerätebezogene Referenzgrenzen bestimmt.

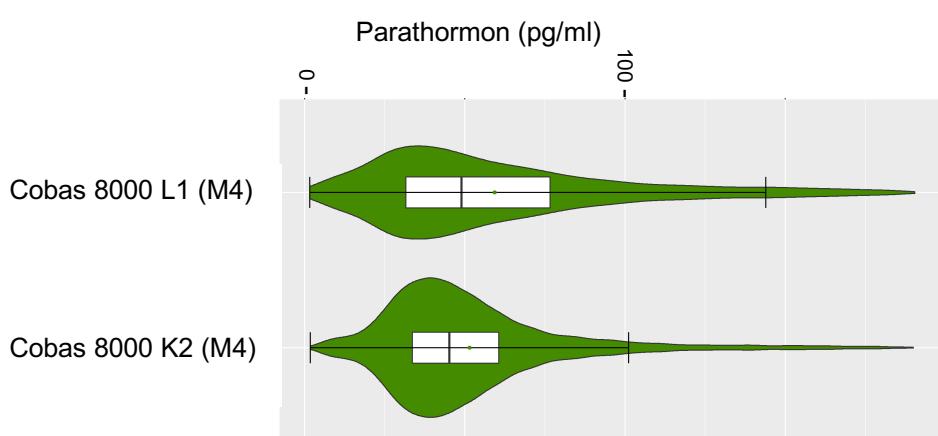


**Abbildung 2:** a.: Äquivalenztest der häufigsten Messgeräte für Calcium. K = Kiel, L = Lübeck, M = Modul.  
 b.: Äquivalenztest der häufigsten Messgeräte für Phosphat. K = Kiel, L = Lübeck, M = Modul.

**Tabelle 7:** Erster Äquivalenztest der Geräte für Phosphat mit einem Epsilon-Wert von 10%. äq.= äquivalent, n. äq.= nicht äquivalent, K = Kiel, L = Lübeck, M = Modul.

Gerät 1	Gerät 2	Konfidenzintervall	Ergebnis
Cobas 8000 K2 (M2)	Cobas 8000 K1 (M3)	[0,003 – 0,011]	äq.
Cobas 8000 K2 (M2)	Cobas 8000 L1 (M2)	[0,071 – 0,078]	äq.
Cobas 8000 K2 (M2)	Cobas 8000 L2 (M3)	[0,08 – 0,088]	äq.
Cobas 8000 K2 (M2)	Cobas 8000 K2 (M1)	[(- 0,03) – (- 0,015)]	äq.
Cobas 8000 K1 (M3)	Cobas 8000 L1 (M2)	[0,063 – 0,071]	äq.
Cobas 8000 K1 (M3)	Cobas 8000 L2 (M3)	[0,072 – 0,081]	äq.
Cobas 8000 K1 (M3)	Cobas 8000 K2 (M1)	[(- 0,037) – (- 0,022)]	äq.
Cobas 8000 L1 (M2)	Cobas 8000 L2 (M3)	[0,006 – 0,013]	äq.
Cobas 8000 L1 (M2)	Cobas 8000 K2 (M1)	[(- 0,104) – (- 0,089)]	n. äq.
Cobas 8000 L2 (M3)	Cobas 8000 K2 (M1)	[(- 0,114) – (- 0,099)]	n. äq.

Für fT3 ergibt sich beim Vergleich der fünf häufigsten Geräte in allen Fällen eine Bewertung der Geräte als äquivalent. Genauso verhält es sich bei der Prüfung der Geräte für fT4 und TSH (s. Anhang Abb. 29, Tab. 31-33). Bei Parathormon werden die zwei häufigsten Geräte verglichen und schon bei der grafischen Gegenüberstellung stellen sich die Violin- und Box-Plots sehr unterschiedlich dar (s. Abb. 3). Die Geräte sind nicht äquivalent (s. Anhang Tab. 34). Daher werden auch hier die Referenzgrenzen zum einen mit Messwerten bestimmt, die von beiden Geräten stammen, und zum anderen gerätespezifische Referenzgrenzen errechnet. Für Vitamin D, Renin und Aldosteron ist der Äquivalenztest nicht nötig, da jeweils an nur einem Gerät die Mehrheit der Messungen durchgeführt wurden.

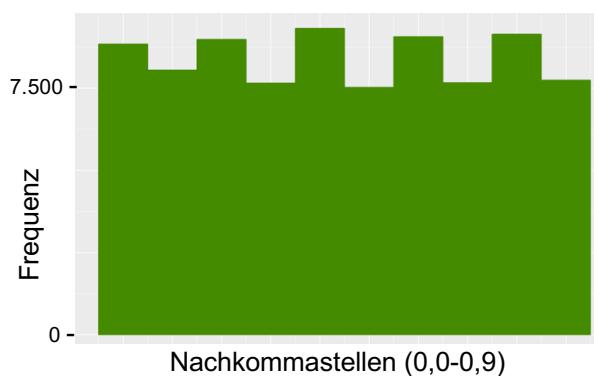


**Abbildung 3:** Äquivalenztest der häufigsten Messgeräte für Parathormon. K = Kiel, L = Lübeck, M = Modul.

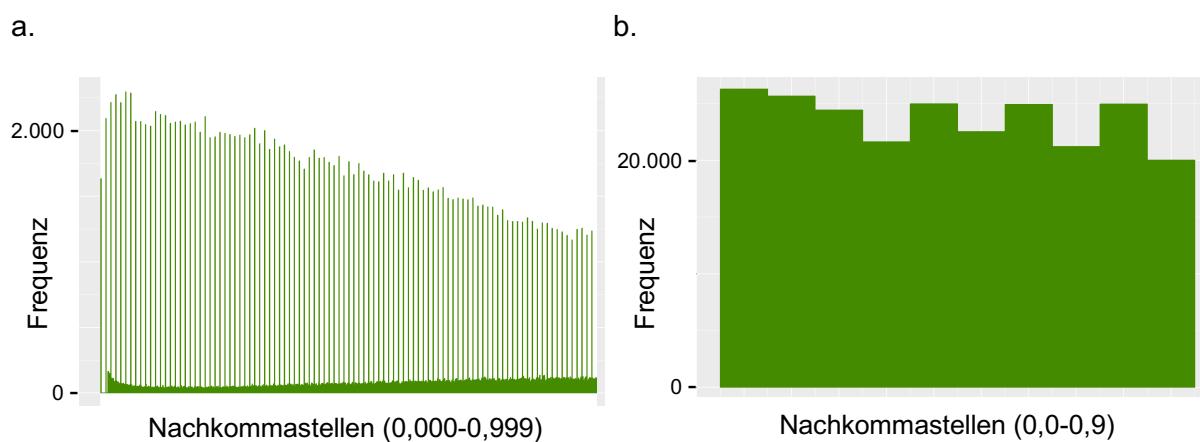
### 3.1.3 Nachkommastellen

Eine Prüfung der Nachkommastellen wird durchgeführt, um die Nachkommastellen der Messwerte für jeden Parameter zu beurteilen und zu überprüfen, ob signifikante Stellen die Verteilung der Messwerte beeinflussen. Bei Calcium und Phosphat stellten sich keine Unregelmäßigkeiten in den Nachkommastellen dar. Die Verteilungen spiegeln die schmalen physiologischen Referenzbereiche der beiden Laborparameter wider (für Calcium nach Thomas: 2,15 – 2,58 mmol/l; für Phosphat nach Thomas: 0,84 – 1,45 mmol/l; s. Anhang, Abb. 30, 31a). Erst bei den späteren Berechnungen präsentierten die Dichtefunktionen beider Laborparameter einige öfter auftretende Messwerte (s. Abb. 11, 12b), sodass eine Rundung der Werte auf zwei Nachkommastellen vorgenommen wird. Damit wurde das Phänomen behoben (s. Abb. 23).

Bei fT4 zeigen sich ungleichverteilte Nachkommastellen mit höherer Häufigkeit der geraden Nachkommastellen (s. Abb. 4). Grund dafür sind aber nicht signifikante Stellen, sondern das statistische Runden der Geräte. Auch für TSH zeigen sich Auffälligkeiten in der Grafik (s. Abb. 5a). Ursächlich dafür sind die verschiedenen Messgeräte, die jeweils eine unterschiedliche Anzahl an Nachkommastellen angeben. Daher wird eine Rundung der Daten auf eine Nachkommastelle vorgenommen, wobei sich das Phänomen des statistischen Rundens in Form von häufigeren geraden Nachkommastellen zeigt (s. Abb. 5b). Auch in den Grafiken für Parathormon und Vitamin D ist dieses sichtbar, während das Histogramm von fT3 unauffällig ist (s. Anhang, Abb. 31b, 32).



**Abbildung 4:** Histogramm der Nachkommastellen für die Messwerte von fT4.



**Abbildung 5:** a.: Histogramm der Nachkommastellen von 0,000 bis 0,999 für die Messwerte von TSH.  
 b.: Histogramm nach der Rundung der Nachkommastellen auf eine Nachkommastelle (0,0 bis 0,9) für die Messwerte von TSH.

## 3.2 Ermittlung der Referenzbereiche

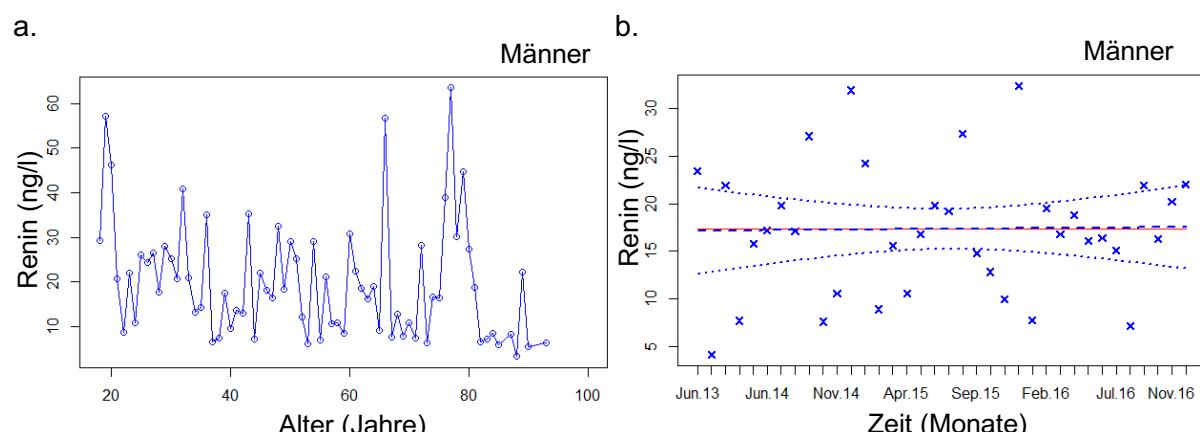
Es werden zunächst bei allen Parametern Referenzgrenzen für Messwerte von Patienten von 18 bis 100 Jahren bestimmt. Dazu gehören auch geschlechtsbezogene Referenzgrenzen. Bei Inspektion der Grafiken zeigt sich eine Streuung der Messwerte im höheren Lebensalter, da deutlich weniger Messwerte verfügbar sind. Daher wird eine Begrenzung der Altersspanne von 18 bis 80 Jahre vorgenommen. Danach werden zudem Referenzgrenzen für kleinere Altersgruppen gebildet. Bei Calcium und Vitamin D erfolgt zusätzlich eine saisonale Aufteilung und bei Phosphat sowie Parathormon eine gerätebezogene Referenzwertermittlung.

Zuletzt wird das Programm „Zulässige Messunsicherheit“ der DGKL angewandt, um zu prüfen, ob die Unterteilung der Gesamtdatenmenge nach Alter sowie Geschlecht zur Erstellung spezifischer Referenzgrenzen nötig ist. Dabei wird aus den Referenzgrenzen für beide Geschlechter für die Altersspanne 18 bis 80 Jahre die zulässige Differenz (= permissible difference, pD) errechnet. Die Referenzgrenzen für eine kleinere Altersspanne bzw. für nur ein Geschlecht werden nun mit den Intervallen der pD verglichen. Liegen die Referenzgrenzen in den pD-Intervallen, ist ein allgemeiner Referenzbereich ausreichend. Liegen sie außerhalb der pD-Intervalle, ist der Nutzen getrennter Referenzbereiche zu prüfen.

### 3.2.1. Altersklasse 18 - 100 Jahre

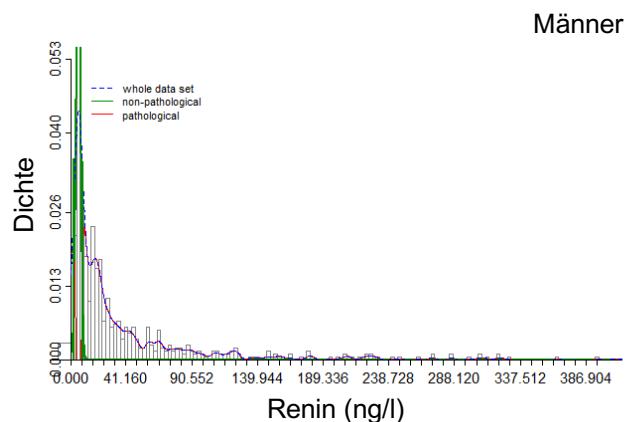
#### 3.2.1.1 Renin und Aldosteron

Bei den Berechnungen der Referenzgrenzen für Renin und Aldosteron für die Altersklasse 18 bis 100 Jahre zeigt sich bei Betrachtung der alters- und des zeitabhängigen Mediane eine starke Streuung der Daten. Bei Renin zeigt sich dies am stärksten bei den Messwerten der Männer, daher sind sie hier beispielhaft aufgeführt (s. Abb. 6).



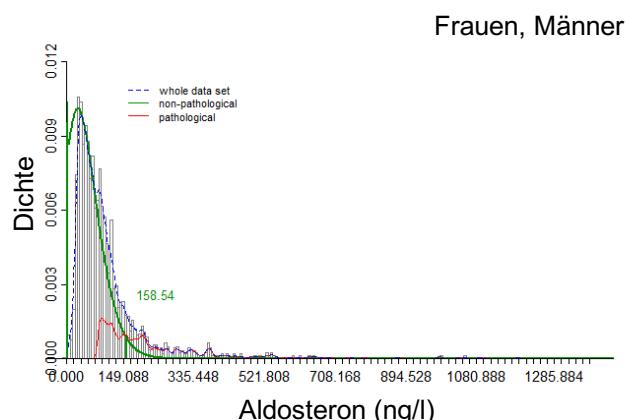
**Abbildung 6:** a.: Altersabhängiger Verlauf der 50.-Perzentile für Renin für Männer von 80-100 Jahren. b.: Zeitabhängiger Median für Renin für Männer von 80-100 Jahren (blau gestrichelt). KI=Konfidenzintervall (blau gepunktet). Median (rot).

Das Histogramm zeigt, dass die Daten keine Normalverteilung aufweisen (s. Abb. 7). Bei der Schätzung der Verteilungen können die pathologischen Werte nicht klar abgegrenzt werden. Dies ist dem geringen Umfang an Messwerten geschuldet. Für Renin sind zunächst 2164 Messwerte verfügbar. Nach Ausschluss von gynäkologischen Patienten und Patienten von Intensivstationen bleiben noch 1739 Werte, wovon 697 auf die Männer und 1042 Werte auf die Frauen entfallen.



**Abbildung 7:** Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Renin für Männer von 18-100 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

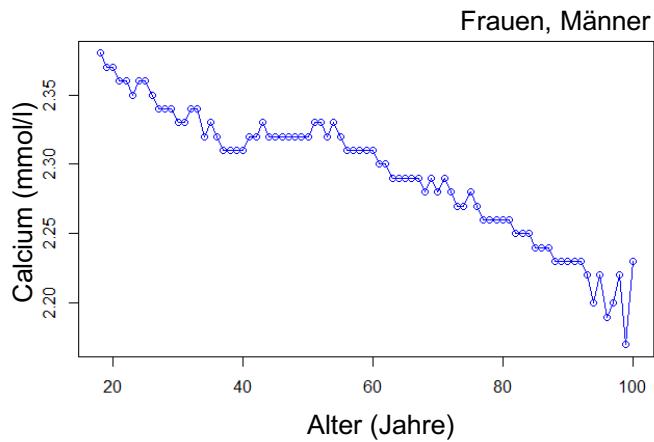
Bei Aldosteron werden ähnliche Effekte sichtbar, wenngleich hier die Abgrenzung der pathologischen von der nicht-pathologischen Verteilung besser funktioniert (s. Abb. 8). Für Aldosteron liegen initial 2736 Messwerte vor, nach den oben genannten Ausschlüssen 2565 Messwerte. Für den RLE wird eine Mindestanzahl von 2000 Werten, besser 4000 Werte empfohlen (15). Referenzgrenzen werden zwar berechnet, diese können jedoch aufgrund der wenigen Messwerte nicht als sinnvoll eingestuft werden und sollten in der Praxis nicht angewendet werden.



**Abbildung 8:** Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Aldosteron für beide Geschlechter von 18-100 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

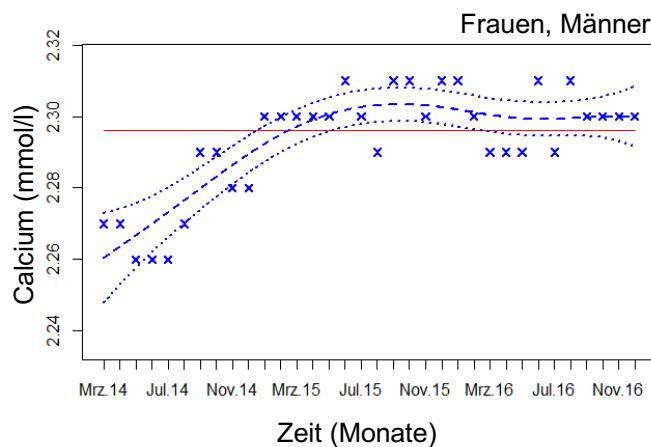
### 3.2.1.2 Calcium und Phosphat

Für Calcium zeigt sich bei der grafischen Auswertung für Messwerte von Patienten von 18 bis 100 Jahren ein gleichmäßiger, mit zunehmendem Alter relativ konstant abfallender Median der Messwerte. Im Altersbereich von 40 bis 50 Jahren zeigt sich ein Plateau, die Messwerte stellen sich stabil dar. Im Alter von 90 bis 100 Jahren präsentiert sich der Median schwankend, was auf das Vorhandensein nur weniger Messwerte in diesem hohen Lebensalter zurückzuführen ist (s. Abb. 9).



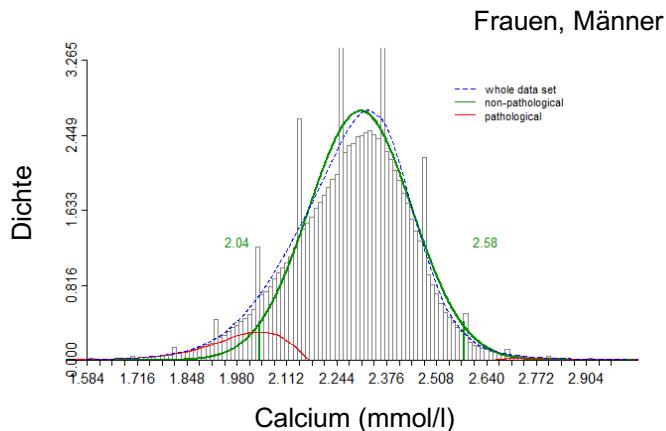
**Abbildung 9:** Altersabhängiger Verlauf der 50.-Perzentile für Calcium für beide Geschlechter von 80-100 Jahren.

Die Grafik mit dem zeitabhängigen Median zeigt für die ersten Monate der Messungen im Zeitraum März bis November 2014 leicht niedrigere Messwerte. Danach kommt es zu einem leichten Anstieg und der zeitabhängige Median stabilisiert sich bei etwa 2,30 mmol/l. Dies ist auf einen initial geringeren Datenumfang zurückzuführen (s. Abb. 10).



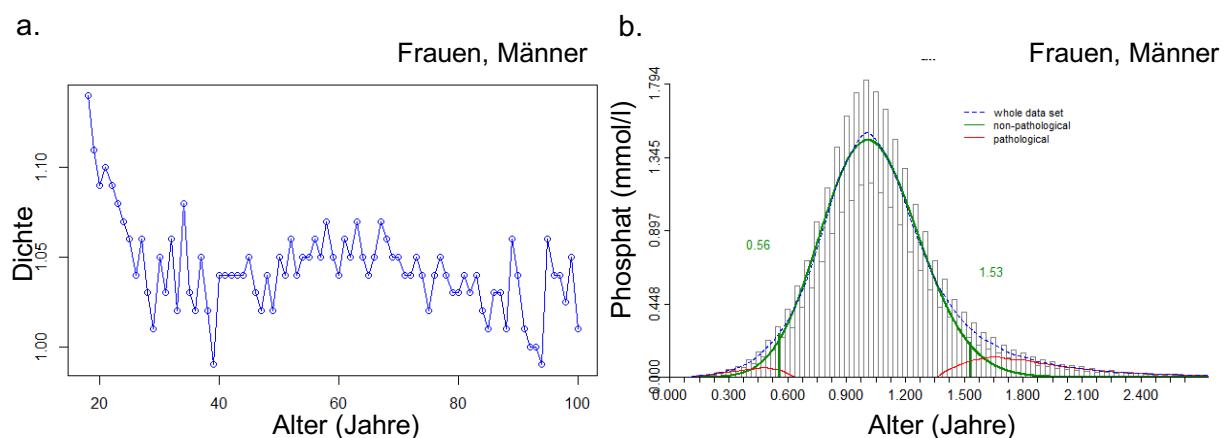
**Abbildung 10:** Zeitabhängiger Median für Calcium für beide Geschlechter von 80-100 Jahren (blau gestrichelt). KI=Konfidenzintervall (blau gepunktet). Median (rot).

Das Histogramm lässt eine Normalverteilung der nicht-pathologischen Daten annehmen. Im Bereich der niedrigeren Messwerte stellt sich der Kurvenverlauf flacher dar, dies lässt eine Subgruppe pathologischer Messwerte vermuten. Zudem zeigen sich regelmäßig einige häufiger vorkommende Messwerte (s. Abb. 11).



**Abbildung 11:** Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Calcium für beide Geschlechter von 18-100 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

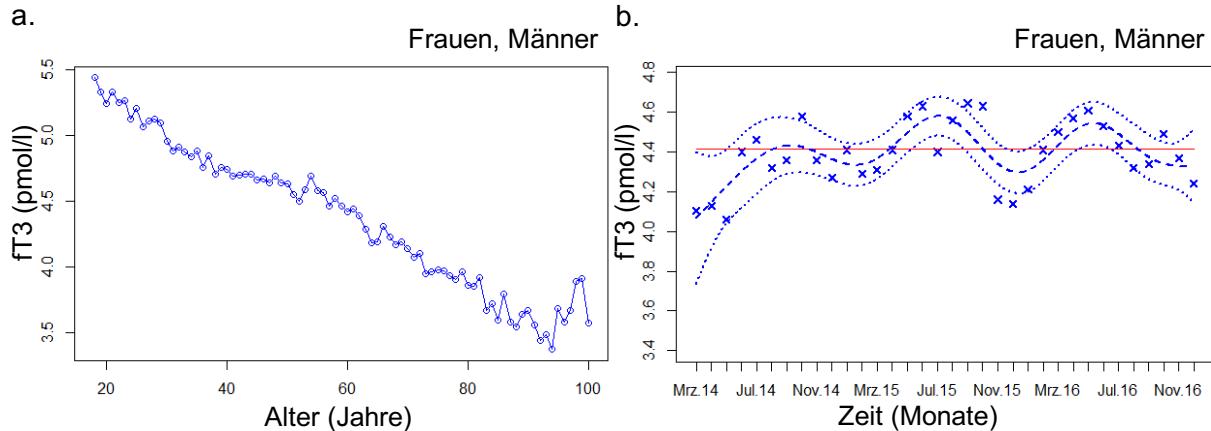
Die grafischen Darstellungen für die Messwerte von Phosphat erweisen sich als ähnlich zu denen von Calcium. Als Unterschied lässt sich der altersabhängige Median der Daten nennen, der keinen vergleichbaren kontinuierlichen altersabhängigen Abfall zeigt. Bis auf einen stärkeren Abfall zwischen 20 und 30 Jahren verläuft er ansonsten relativ stabil (s. Abb. 12a). Im Histogramm stellen sich die Messwerte normalverteilt dar, zudem werden ebenfalls häufiger auftretende Messwerte sichtbar (s. Abb. 12b).



**Abbildung 12:** a.: 50.-Perzentile für Phosphat für beide Geschlechter von 18 – 100 Jahren.  
b.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen(rot) Werte und für alle Daten (blau) für Phosphat für beide Geschlechter von 18-100 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

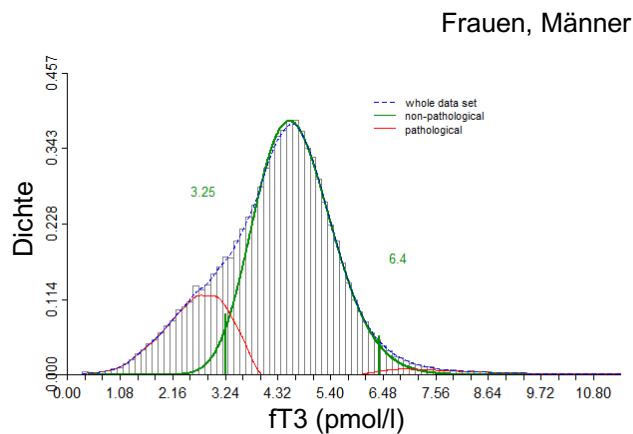
### 3.2.1.3 TSH, fT3 und fT4

Die grafischen Darstellungen der Messwerte für fT3 liefern für den altersabhängigen Median der Messwerte für beide Geschlechter gemeinsam eine mit dem Alter abfallende Kurve (s. Abb. 13a). Der Median über die Zeit zeigt einen wellenförmig schwankenden Verlauf mit höheren Messwerten im Juli, August und niedrigeren im Januar (s. Abb. 13b).



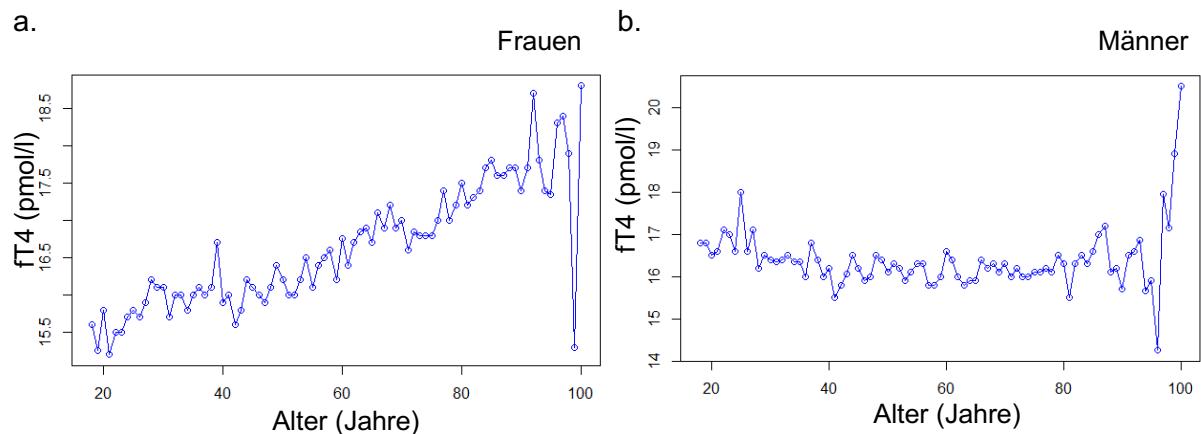
**Abbildung 13:** a.: Altersabhängiger Verlauf der 50.-Perzentile für fT3 für beide Geschlechter von 18-100 Jahren.  
b.: Zeitabhängiger Median für fT3 für beide Geschlechter von 80-100 Jahren (blau gestrichelt). KI=Konfidenzintervall (blau gepunktet). Median (rot).

Die Dichtefunktion der Messwerte für fT3 stellt eine zweite große Subgruppe im Bereich der niedrigeren Messwerte dar (s. Abb. 14).



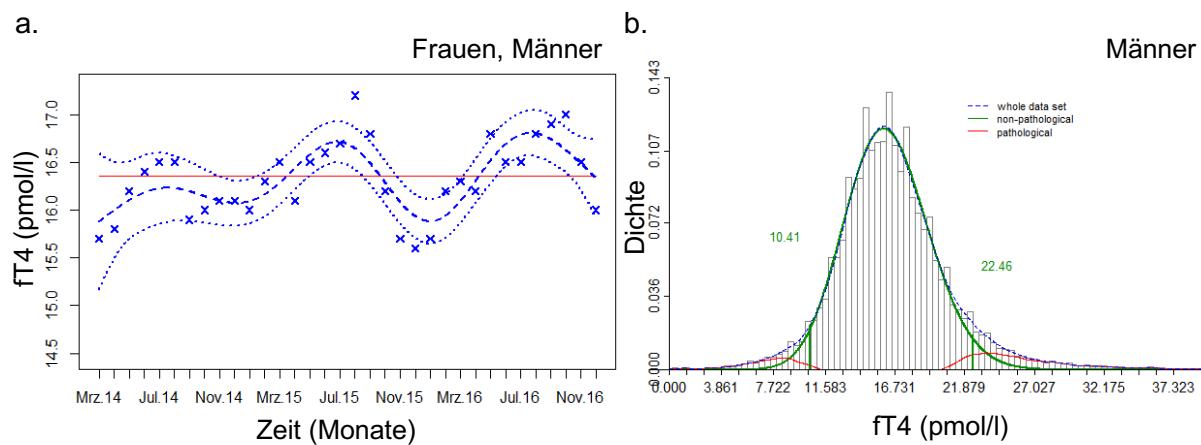
**Abbildung 14:** Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für fT3 für beide Geschlechter von 18-100 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

Bei fT4 hingegen zeigen sich die altersabhängigen Mediane der Messwerte für Frauen leicht ansteigend mit zunehmendem Alter. Für die Messwerte der Männer stellt sich der Median initial leicht fallend und danach relativ stabil dar (s. Abb. 15).



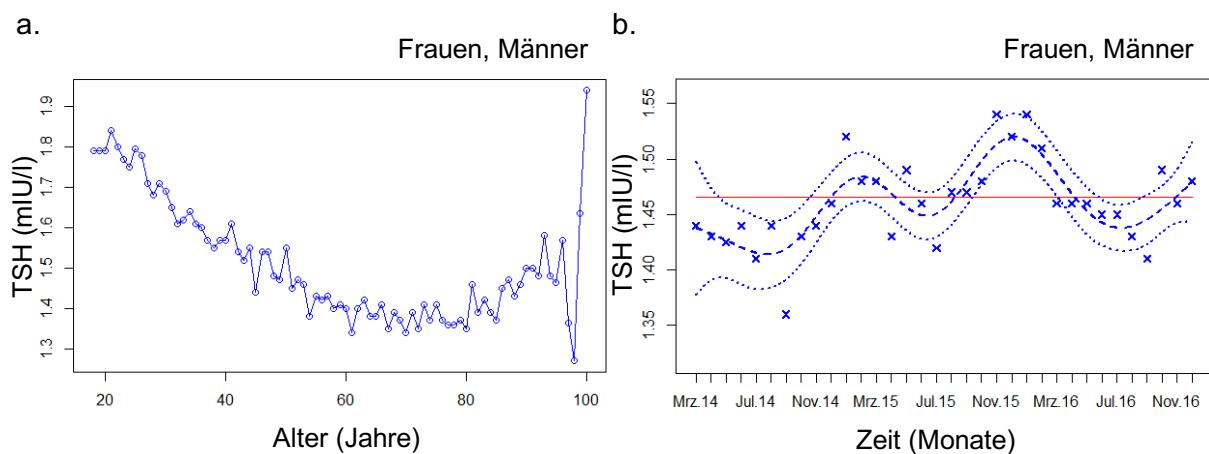
**Abbildung 15:** a.: Altersabhängiger Verlauf der 50.-Perzentile für fT4 für Frauen von 18-100 Jahren.  
b.: Zeitabhängiger Median für fT4 für Männer von 80-100 Jahren (blau gestrichelt). KI=Konfidenzintervall (blau gepunktet). Median (rot).

Der zeitabhängige Median stellt sich schwankend dar mit höheren Messwerten in den Sommermonaten, ähnlich wie bei den Daten von fT3 (s. Abb. 16a). Die Dichtefunktion zeigt einige häufiger auftretende Messwerte, hier beispielhaft in der Grafik für die Messwerte der Männer sichtbar (s. Abb. 16b).

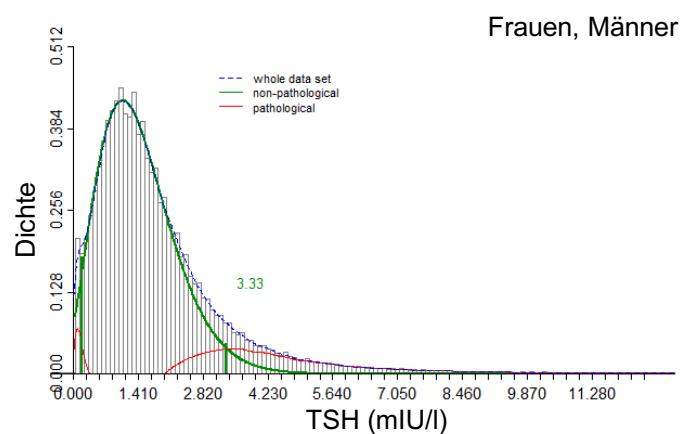


**Abbildung 16:** a.: Zeitabhängiger Median für fT4 für beide Geschlechter von 80-100 Jahren (blau gestrichelt). KI=Konfidenzintervall (blau gepunktet). Median (rot).  
b.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für fT4 für Männer von 18-100 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

Die 50. Perzentilen von TSH für Männer, Frauen und für beide Geschlechter gemeinsam fallen mit dem Alter ab bis etwa zum 60. Lebensjahr. Dann stabilisieren sich die Messwerte zwischen 1,3 und 1,4 mIU/l, um zwischen dem 80. und 100. Lebensjahr erneut anzusteigen (s. Abb. 17a). Der Median in Zeitabhängigkeit stellt sich wie bei fT3 und fT4 wellenförmig schwankend dar. Allerdings werden bei TSH im Juli und August niedrigere Ergebnisse gemessen, während im Januar und Februar hohe Messwerte dominieren (s. Abb. 17b). Die Schätzung der Messwertverteilung gelingt und zeigt eine Abflachung des Kurvenlaufes im Bereich der höheren Werte (s. Abb. 18).



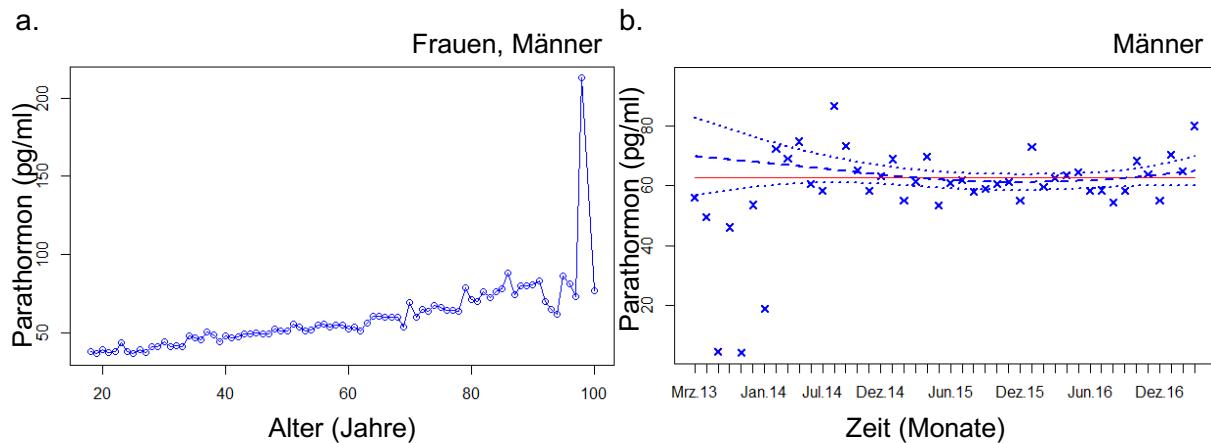
**Abbildung 17:** a.: Altersabhängiger Verlauf der 50.-Perzentile für TSH für beide Geschlechter von 18-100 Jahren.  
b.: Zeitabhängiger Median für TSH für beide Geschlechter von 80-100 Jahren (blau gestrichelt). KI=Konfidenzintervall (blau gepunktet). Median (rot).



**Abbildung 18:** Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für TSH für beide Geschlechter von 18-100 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

### 3.2.1.4 Parathormon

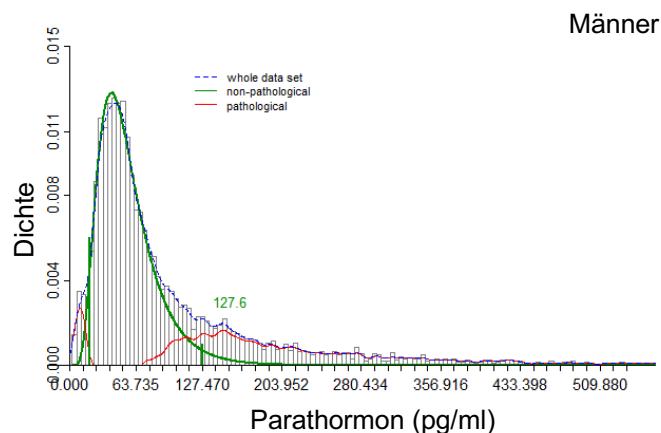
Der altersabhängige Median beschreibt für Parathormon einen sukssessive ansteigenden Verlauf von 40 bis etwa 80 pg/ml im 80. Lebensjahr, danach präsentiert er sich unsteter (s. Abb. 19a). Der Median über die Zeit stellt sich für beide Geschlechter gemeinsam relativ stabil dar, während die Grafik der Männer auf gestreutere Messwerte hindeutet (s. Abb. 19b).



**Abbildung 19:** a.: Altersabhängiger Verlauf der 50.-Perzentile für Parathormon für beide Geschlechter von 18-100 Jahren.

b.: Zeitabhängiger Median für Parathormon für Männer von 80-100 Jahren (blau gestrichelt). KI=Konfidenzintervall (blau gepunktet). Median (rot).

Die Schätzung der Verteilungen der nicht-pathologischen und pathologischen Werte gelingt. Es zeigt sich eine leicht linksgipflige Verteilung. Gerade im höheren Wertebereich verlaufen die Kurven der Messwerteverteilung flacher, was auf einen größeren Anteil pathologischer Werte schließen lässt, dies zeigt sich gut in der Verteilung der Messwerte der Männer (s. Abb. 20).

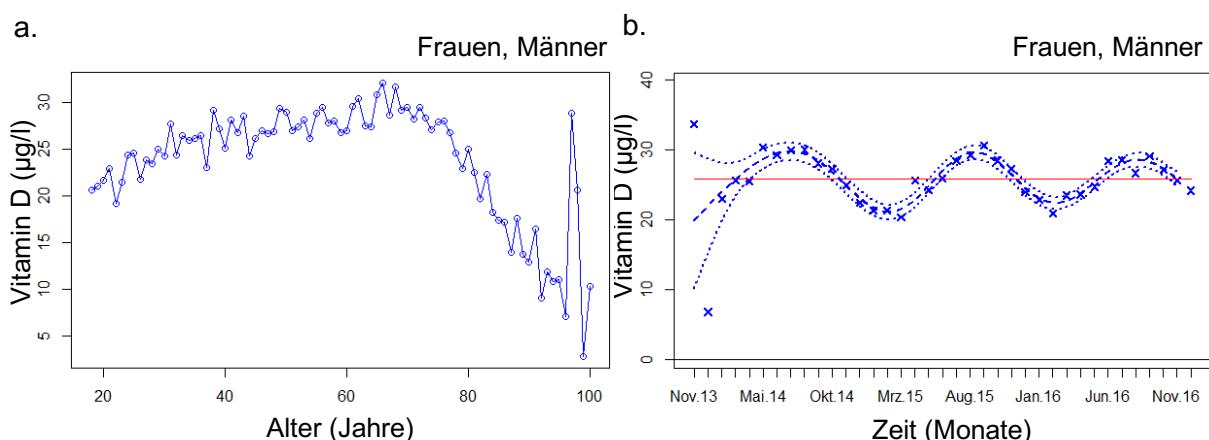


**Abbildung 20:** Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für TSH für Männer von 18-100 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

### 3.2.1.5 Vitamin D

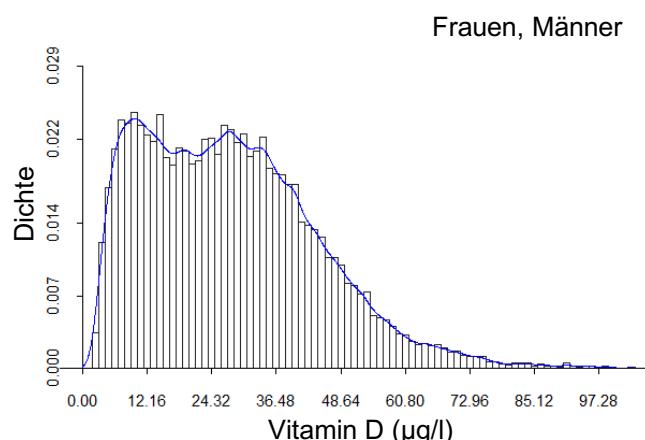
Bei dem im Folgenden als „Vitamin D“ bezeichneten Laborparameter handelt es sich um die Form 25-OH-Vitamin D.

Bei Vitamin D steigt der altersabhängige Median leicht mit Höchstwerten von bis zu 35 µg/l zwischen dem 60. und 70. Lebensjahr. Danach kommt es zu einem Abfall der Messwerte (s. Abb. 21a). Der Median erscheint über den Zeitraum der Probenentnahmen wellenförmig, wobei der Schwankungsbereich zwischen 20 und 30 µg/l liegt. In den Sommermonaten zeigen sich höhere, in den Wintermonaten niedrigere Messwerte (s. Abb. 21b).



**Abbildung 21:** a.: Altersabhängiger Verlauf der 50.-Perzentile für Vitamin D für beide Geschlechter von 18-100 Jahren.  
b.: Zeitabhängiger Median für Vitamin D für beide Geschlechter von 80-100 Jahren (blau gestrichelt). KI=Konfidenzintervall (blau gepunktet). Median (rot).

Die Dichtefunktion stellt keine Normalverteilung dar, sondern präsentiert sich zweigipflig, was auf zwei große Subgruppen schließen lässt (s. Abb. 22). Die Verteilungen der pathologischen sowie der nicht-pathologischen Werte können nicht voneinander abgegrenzt und somit keine Referenzgrenzen generiert werden. Insgesamt liegen 21.104 Daten für Vitamin D vor.



**Abbildung 22:** Geschätzte Verteilung für alle Daten (blau) für Vitamin D für beide Geschlechter von 18-100 Jahren.

### 3.2.1.6 Referenzintervalle der Parameter für das Alter 18-100 Jahre

Es ließen sich für sechs der neun Laborparameter mittels des RLE Referenzgrenzen berechnen (s. Tabelle 8).

**Tabelle 8:** Referenzintervalle verschiedener Parameter für die Altersklasse 18-100 Jahre. G = Geschlecht, RI = Referenzintervall, N = Anzahl der Messwerte, Frauen = Frauen, M = Männer, PTH = Parathormon.

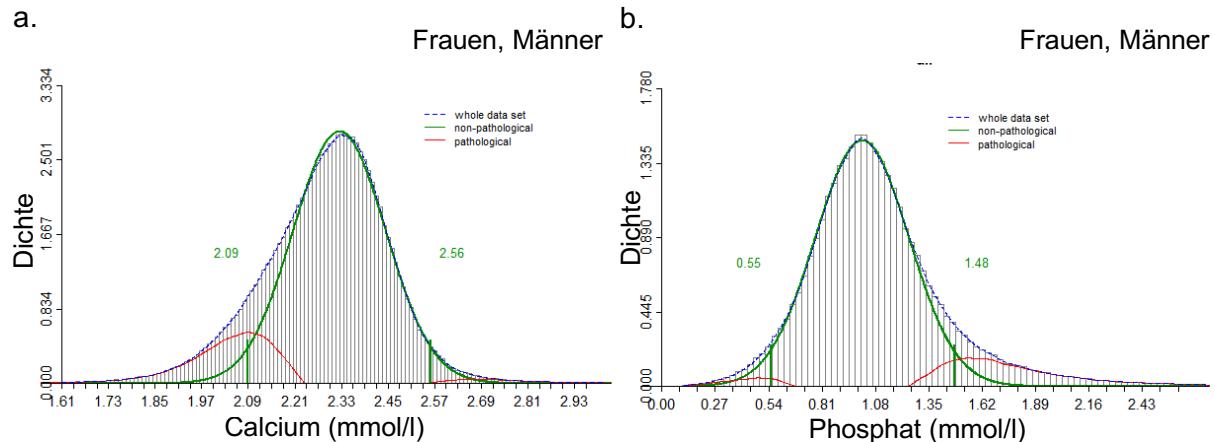
Parameter	G	RI	N	Analyt	G	RI	N
Calcium (mmol/l)	F, M	2,04 – 2,58	629.642	TSH (mIU/l)	F, M	0,18 – 3,33	301.115
	F	2,06 – 2,57	291.340		F	0,13 – 3,27	151.369
	M	2,01 – 2,58	338.176		M	0,28 – 4	149.596
Phosphat (mmol/l)	F, M	0,56 – 1,53	270.496	fT3 (pmol/l)	F, M	3,25 – 6,4	109.918
	F	0,57 – 1,49	116.634		F	3,41 – 6,28	62.107
	M	0,57 – 1,59	153.832		M	3,15 – 6,54	47.748
PTH (pg/ml)	F, M	18,31 – 89,76	17.101	fT4 (pmol/l)	F, M	10,63 – 22,92	112.453
	F	18,92 – 86,52	11.317		F	10,81 – 23,2	63.728
	M	18,35 – 119,79	5.784		M	10,41 – 22,46	48.651,

### 3.2.2 Weitere Eingrenzung der Altersklassen

Bei der visuellen Inspektion der 50. Perzentile fällt stets auf, dass diese besonders in höherem Lebensalter starke Schwankungen aufweist (s. Abb. 9, 12a, 13a, 15, 17a, 19a, 21a). Im höheren Lebensalter liegen naturgemäß weniger Daten vor und es kommt zu einer stärkeren Streuung, was die Genauigkeit der Referenzgrenzen beeinträchtigen kann. Aus diesem Grund wird nachfolgend im ersten Schritt das Altersspektrum auf 18 bis 80 Jahre eingegrenzt. Je nach Datenumfang wurden noch kleinere Altersgruppen gebildet. Bei Phosphat und Calcium mit besonders umfangreichen Datenmengen wurden jahreszeitspezifische Referenzgrenzen erstellt. Für Renin und Aldosteron erfolgen aufgrund der zu geringen Menge an Messwerten keine weiteren Berechnungen.

### 3.2.2.1 Calcium und Phosphat

Für Calcium und Phosphat haben sich bei den ersten Auswertungen für die Altersklasse von 18 bis 100 Jahren einige Unregelmäßigkeiten in den Verteilungen gezeigt (s. Abb. 11, 12b). Die beiden Datensätze werden daher auf zwei Nachkommastellen gerundet und das Phänomen damit behoben (s. Abb. 23).



**Abbildung 23:** a.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Calcium für beide Geschlechter von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.  
b.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Phosphat für beide Geschlechter von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

Die folgenden Berechnungen erfolgen alle mit den gerundeten Datensätzen. Des Weiteren werden die Daten anhand des Zeitpunkts der Einschleusung im Labor in Sommer und Winter unterteilt. Die Sommermonate umfassen Mai bis Oktober 2015 sowie Mai bis Oktober 2016. Die Wintermonate November und Dezember 2014, Januar bis April 2015, November und Dezember 2015 sowie Januar bis April 2016. Neben der alle Daten umfassenden Altersgruppe 18 bis 80 Jahre werden spezifischere Altersklassen in 10 Jahres-Schritten gebildet (s. Tab. 9, 10). Die pD-Grenzen werden für die Altersklasse 18 bis 80 Jahre ermittelt und die Referenzgrenzen für Subgruppen auf Übereinstimmung mit diesen Wertebereichen geprüft. Für die untere Referenzgrenze ergeben sich folgende pD-Grenzen: 2,03 – 2,15. Für die obere Referenzgrenze: 2,49 – 2,63. Bei den in der Tabelle grau hinterlegten Referenzgrenzen sollte erwogen werden, für diese Subgruppen eigene Referenzgrenzen zu berechnen. Bei der saisonalen Aufteilung liegen vor allem die unteren Referenzgrenzen der Gruppe Sommer als auch Winter für die Altersstufe 18 bis 27 Jahre außerhalb der zulässigen Grenzen.

**Tabelle 9:** Referenzintervall für Calcium (mmol/l) für die Altersklasse 18-80 Jahre. Alter in Jahren. G = Geschlecht, RI = Referenzintervall, N = Anzahl der Messwerte, F = Frauen, M = Männer.

Alter	G	RI	N
18 – 80	F, M	2,09 – 2,56	556.550
	F	2,11 – 2,56	251.724
	M	2,08 – 2,56	304.706

**Tabelle 10:** Saisonale Referenzintervalle für Calcium (mmol/l) für Sommer und Winter. Grau hinterlegt sind Werte außerhalb der pD-Grenzen. Alter in Jahren. G = Geschlecht, RI = Referenzintervall, N = Anzahl der Messwerte, F = Frauen, M = Männer.

Alter	G	RI Sommer	N	RI Winter	N
18 – 80	F, M	2,1 – 2,56	231.089	2,09 – 2,56	236.196
	F	2,13 – 2,56	103.875	2,11 – 2,56	106.466
	M	2,09 – 2,56	127.180	2,07 – 2,56	129.665
18 – 27	F, M	2,18 - 2,58	13.739	2,16 – 2,57	14.710
	F	2,17 - 2,55	7.175	2,15 – 2,55	7.778
	M	2,21 – 2,6	6.555	2,18 – 2,6	6.922
28 – 37	F, M	2,16 – 2,55	15.594	2,13 – 2,55	15.648
	F	2,14 – 2,53	8.564	2,12 – 2,53	8.972
	M	2,18 – 2,58	7.027	2,21 – 2,57	6.666
38 – 47	F, M	2,12 - 2,56	22.046	2,12 – 2,55	23.110
	F	2,12 - 2,54	11.135	2,12 – 2,53	11.654
	M	2,12 - 2,57	10.908	2,12 – 2,57	11.447
48 – 57	F, M	2,13 – 2,56	41.713	2,13 – 2,56	42.215
	F	2,15 – 2,57	19.029	2,13 – 2,57	19.329
	M	2,11 – 2,56	22.675	2,12 – 2,55	22.877
58 – 67	F, M	2,09 – 2,57	52.685	2,1 – 2,56	53.582
	F	2,12 - 2,58	22.329	2,11 – 2,58	23.112
	M	2,08 - 2,55	30.352	2,1 – 2,55	30.460
68 – 77	F, M	2,05 – 2,56	66.920	2.05 - 2.55	68.503
	F	2,13 – 2,56	27.574	2.07 - 2.57	27.630
	M	2,04 – 2,54	39.341	2.03 - 2.54	40.858

Für Phosphat werden für die Altersgruppe 18 bis 80 Jahre Referenzgrenzen berechnet. Für die saisonalen Datensätze werden Altersklassen in 20 Jahres-Schritten herangezogen. Aufgrund des geringeren Datenumfangs als bei Calcium werden größere Altersgruppen für Phosphat gewählt (s. Tab. 11, 12).

Da beim Äquivalenztest der fünf häufigsten Messgeräte für Phosphat einige Geräte als nicht äquivalent und somit nicht vergleichbar eingestuft wurden, werden für Phosphat zudem Referenzgrenzen spezifisch für die einzelnen Geräte bestimmt (s. Tab. 13).

Es ergeben sich für die Altersgruppe 18 bis 80 Jahre folgende pD-Grenzen: 0,51 - 0,59 für die untere und 1,39 - 1,57 für die obere Referenzgrenze. Vor allem die unteren Referenzgrenzen für die Gruppen „Sommer“ und „Winter“ der Altersklasse 18 bis 37 Jahre liegen außerhalb der zulässigen Differenz. Bei den gerätespezifischen Referenzgrenzen zeigen die Geräte Cobas 8000 K2 (M1) und Cobas 8000 K2 (M2) die größten Abweichungen.

**Tabelle 11:** Referenzintervall für Phosphat (mmol/l) für die Altersklasse 18-80 Jahre. Alter in Jahren. G = Geschlecht, RI = Referenzintervall, N = Anzahl der Messwerte, F = Frauen, M = Männer.

Alter	G	RI	N
18 – 80	F, M	0,55 – 1,48	234.242
	F	0,58 – 1,48	97.893
	M	0,55 – 1,5	136.319

**Tabelle 12:** Saisonale Referenzintervalle für Phosphat (mmol/l) für Sommer und Winter. Grau hinterlegt sind Werte außerhalb der pD-Grenzen. Alter in Jahren. RI = Referenzintervall, N = Anzahl der Messwerte, F = Frauen, M = Männer.

Alter	G	RI Sommer	N	RI Winter	N
18 – 80	F, M	0,57 – 1,53	92.360	0,57 – 1,51	100.480
	F	0,58 – 1,5	38.080	0,58 – 1,48	41.768
	M	0,57 – 1,58	54.270	0,56 – 1,48	58.695
18 – 37	F, M	0,61 – 1,5	10.813	0,63 – 1,47	11.778
	F	0,64 – 1,47	5.700	0,66 – 1,45	6.532
	M	0,56 – 1,55	5.112	0,59 – 1,51	5.245
38 – 57	F, M	0,56 – 1,52	23.099	0,54 – 1,49	25.360
	F	0,59 – 1,5	9.950	0,56 – 1,48	10.905
	M	0,56 – 1,58	13.146	0,54 – 1,52	14.446
58 – 77	F, M	0,56 – 1,55	50.131	0,56 – 1,51	54.599
	F	0,57 – 1,52	19.101	0,57 – 1,5	20.693
	M	0,57 – 1,61	31.024	0,59 – 1,57	33.900

**Tabelle 13:** Gerätespezifische Referenzintervalle für Phosphat (mmol/l) für die Altersklasse 18-80 Jahre. Grau hinterlegt sind Werte außerhalb der pD-Grenzen. Alter in Jahren. G = Geschlecht, RI = Referenzintervall, N = Anzahl der Messwerte, F = Frauen, M = Männer, K = Kiel, L = Lübeck, M = Modul.

Gerät	G	RI	N	Gerät	RI	N
Cobas 8000	F, M	0,58 – 1,59	49.700	Cobas 8000	0,56 – 1,45	32.442
K1 (M3)	F	0,58 – 1,53	18.601	L2 (M3)	0,59 – 1,45	14.851
	M	0,59 – 1,66	31.099		0,57 – 1,49	17.586
Cobas 8000	F, M	0,62 – 1,59	8.910	Cobas 8000	0,63 – 1,47	86.368
K2 (M1)	F	0,62 – 1,49	3.604	L1 (M2)	0,66 – 1,45	38.521
	M	0,64 – 1,79	5.305		0,59 – 1,51	47.829
Cobas 8000	F, M	0,61 – 1,6	54.920			
K2 (M2)	F	0,6 – 1,53	21.465			
	M	0,61 – 1,67	33.449			

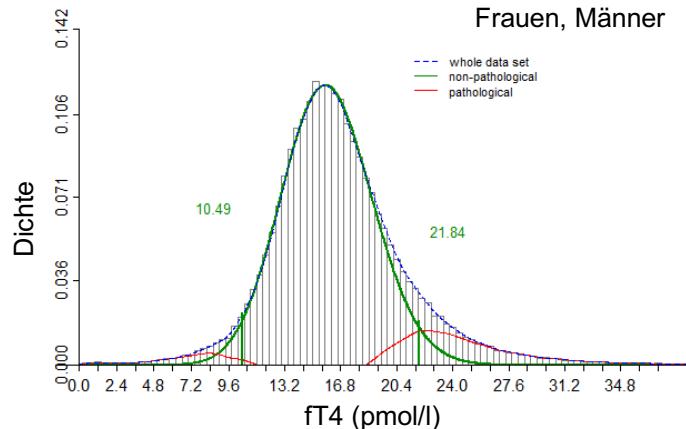
### 3.2.2.2 TSH, fT3 und fT4

Für TSH werden Referenzgrenzen für die Spanne 18 bis 80 Jahre erhoben, sowie für 10 Jahresabschnitte (s. Tab. 14). Für die Altersstufe 18 bis 80 Jahre ergeben sich folgende pD-Grenzen: 0,15 - 0,23 für die untere und 3,03 - 3,69 für die obere Referenzgrenze. Wie der Tabelle zu entnehmen ist, liegt ein Großteil der Referenzgrenzen außerhalb dieser Grenzwerte.

**Tabelle 14:** Referenzintervalle für TSH (mIU/). Grau hinterlegt sind Werte außerhalb der pD-Grenzen. Alter in Jahren. G = Geschlecht. RI = Referenzintervall, N = Anzahl der Messwerte, F = Frauen, M = Männer.

Alter	G	RI	N	Alter	RI	N
18 – 80	F, M	0,19 – 3,36	268.001			
	F	0,14 – 3,29	132.764			
	M	0,28 – 3,99	135.089			
18 – 27	F, M	0,51 – 4,63	20.728	48 – 57	0,24 – 3,71	49.400
	F	0,47 – 4,14	11.923		0,15 – 3,62	24.611
	M	0,53 – 4,82	8.774		0,31 – 3,69	24.777
28 – 37	F, M	0,4 – 3,71	23.768	58 – 67	0,14 – 3,22	55.579
	F	0,36 – 3,52	14.512		0,09 – 3,18	25.408
	M	0,49 – 4,61	9.212		0,18 – 3,11	30.160
38 – 47	F, M	0,31 – 3,5	29.228	68 – 77	0,13 – 3,37	69.728
	F	0,27 – 3,41	16.363		0,1 – 3,5	30.774
	M	0,35 – 3,69	12.837		0,15 – 3,34	38.936

Bei fT3 und fT4 wird die Spanne der Altersklassen auf 15 Jahre begrenzt. Bei fT4 haben sich signifikante Stellen im Nachkommastellentest und Auffälligkeiten im Histogramm ergeben (s. Abb. 4, 16b), sodass der Datensatz gerundet wird. Dies kann das Phänomen beheben (s. Abb. 24). Folgende Berechnungen werden mit dem gerundeten Datensatz durchgeführt.



**Abbildung 24:** Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für fT4 für beide Geschlechter von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

Die pD-Grenzen für fT3 werden erhoben: 3,18 - 3,56 für die untere und 6,15 - 6,79 für die obere Referenzgrenze. Hier zeigen sich für alle Referenzgrenzen der einzelnen Altersspannen, mit Ausnahme der Altersgruppe 48 – 62 Jahre, nicht mehr zulässige Abweichungen (s. Tab. 15).

**Tabelle 15:** Referenzintervalle für fT3 (pmol/l). Grau hinterlegt sind Werte außerhalb der pD-Grenzen. Alter in Jahren. G = Geschlecht. RI = Referenzintervall, N = Anzahl der Messwerte, F = Frauen, M = Männer.

Alter	G	RI	N	Alter	RI	N
18 – 80	F, M	3,37 – 6,47	100.692			
	F	3,43 – 6,29	56.603			
	M	3,39 – 6,64	44.027			
18 – 32	F, M	3,9 – 7,05	12.420	48 – 62	3,45 – 6,35	29.645
	F	3,79 – 6,7	7.878		3,45 – 6,2	16.834
	M	4,21 – 7,14	4.513		3,55 – 6,42	12.807
33 – 47	F, M	3,06 – 6,52	16.475	63-77	2,84 – 5,94	35.811
	F	3,15 – 6,16	10.863		2,82 – 5,94	17.772
	M	2,87 – 8,06	5.594		2,93 – 5,99	18.029

Für fT4 ergeben sich folgende pD-Bereiche: 9,86 - 11,12 für die untere und 20,71 - 22,97 für die obere Referenzgrenze. Damit befinden sich nur drei untere Referenzgrenzen knapp außerhalb der pD-Grenzen (s. Tab. 16).

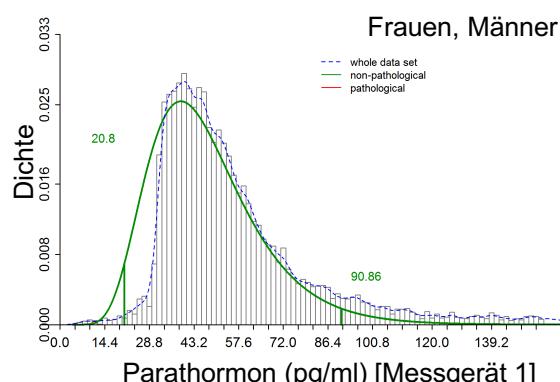
**Tabelle 16:** Referenzintervalle für fT4 (pmol/l). Grau hinterlegt sind Werte außerhalb der pD-Grenzen. Alter in Jahren. G = Geschlecht, RI = Referenzintervall, N = Anzahl der Messwerte, F = Frauen, M = Männer.

Alter	G	RI	N	Alter	RI	N
18 – 80	F, M	10,49 – 21,84	103170			
	F	10,81 – 22,57	58193			
	M	10,3 – 21,72	44904			
18 – 32	F, M	11,3 – 22,04	12.971	48 – 62	10,22 – 21,22	30.422
	F	11,25 – 21,68	8.227		10,32 – 20,98	17.323
	M	11,39 – 21,63	4.713		10,01 – 21,68	13.090
33 – 47	F, M	10,75 – 21,93	17.017	63-77	10,33 – 22,23	36.363
	F	10,76 – 21,68	11.241		10,34 – 22,24	18.115
	M	10,72 – 22,42	5.756		10,44 – 22,47	18.236

### 3.2.2.3 Parathormon

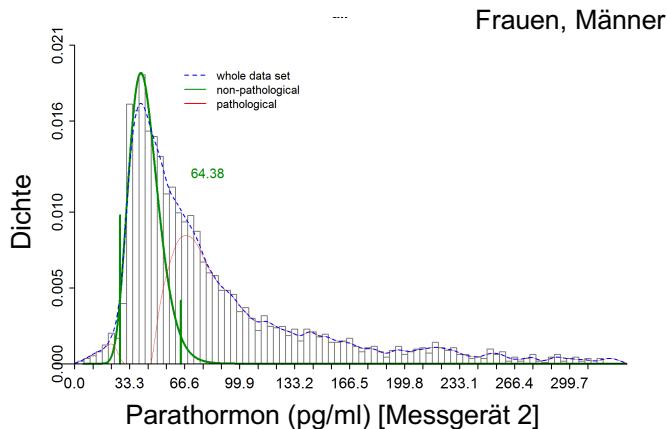
Für die Messwerte von Parathormon werden Referenzgrenzen für Patienten im Alter von 18 bis 80 Jahren erstellt. Dies geschieht zudem aufgeteilt nach verwendetem Messgerät. Bei den Berechnungen für beide Geräte zusammen zeigt sich erneut die linksgipflige Verteilung nahe der y-Achse (s. Anhang Abb. 42d-f).

Für Gerät 1 zeigt sich eine asymmetrische Verteilung der Messwerte, im Bereich niedrigerer Werte verläuft sie sehr steil, im Bereich höherer Messwerte zeigt sich eine langsam abfallende Kurve. Dies ist auf das Runden des Messgeräts zurückzuführen. Für den Datensatz beider Geschlechter gelingt die Schätzung der Verteilung der nicht-pathologischen Werte nicht (s. Abb. 25), für den Datensatz der Männer bzw. Frauen allein funktionierte dies ansatzweise (s. Anhang Abb. 43b, c). Die Referenzgrenzen der Männer wurden zwar berechnet, aufgrund des geringen Datensatzes von 2.748 Werten sollten sie jedoch nicht zur Anwendung kommen.



**Abbildung 25:** Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Parathormon für beide Geschlechter von 18-80 Jahren für Messgerät 1 (Cobas 8000 K2 (M4)). Grüne Zahlen geben die untere und obere RG an. K = Kiel, L = Lübeck, M = Modul.

Für Gerät 2 erfolgt die Schätzung der Verteilungen der nicht-pathologischen und pathologischen Werte. Im höheren Wertebereich wird ein größerer Anteil letzterer sichtbar verglichen mit Gerät 1. Referenzgrenzen werden berechnet, welche aufgrund des mäßigen Datenumfangs von insgesamt 3.065 Messwerten jedoch nicht zum praktischen Einsatz geeignet sind (s. Abb. 26 und Anhang Abb. 43e, f).



**Abbildung 26:** Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Parathormon für beide Geschlechter von 18-80 Jahren Messgerät 2 (Cobas 8000 L1 (M4)). Grüne Zahlen geben die untere und obere RG an. K = Kiel, L = Lübeck, M = Modul.

Es werden folgende pD-Grenzen auf Basis der Referenzgrenzen für beide Geräte errechnet: 17,59 – 21,89 für die untere und 99,25 – 115,61 für die obere Referenzgrenze. Dabei fällt ein Großteil der Referenzgrenzen für das Gerät 1 (Cobas 8000 K2 (M4)) aus diesen Bereichen, beim Gerät 2 (Cobas 8000 L1 (M4)) sind es alle errechneten Grenzwerte (s. Tab. 17).

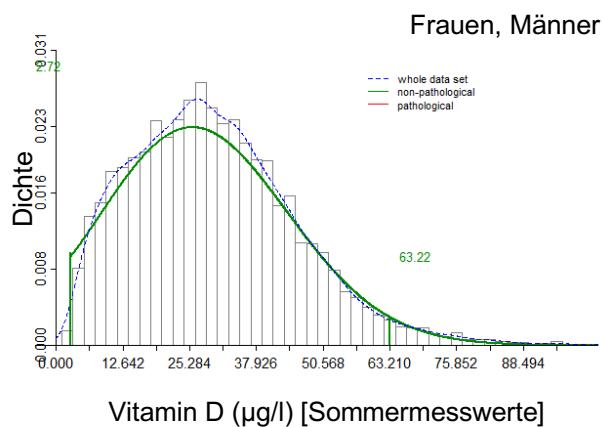
**Tabelle 17:** Allgemeine und gerätespezifische Referenzintervalle für Parathormon für die Altersklasse von 18-80 Jahren. Grau hinterlegt sind Werte außerhalb der pD-Grenzen. G = Geschlecht, RI = Referenzintervall, N = Anzahl der Messwerte, F = Frauen, M = Männer.

Gerät	G	RI	N	Gerät	G	RI	N
beide	F, M	19,74 – 107,43	22.084	Geräte	F	20,13 – 99,47	13.392
	M	18,93 – 125,72	8.691				
Cobas	F, M	20,8 – 90,86	9.228	Cobas	F, M	27,43 – 64,38	3.065
	F	22,82 – 87,82	6.480				
	M	20,82 – 97,11	2.748				
	(M4)						

### 3.2.2.4 Vitamin D

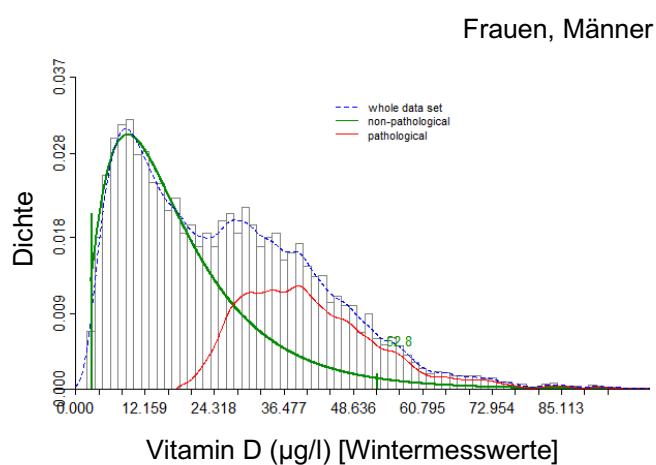
Aufgrund des auffälligen Verlaufs des Medians für die Messwerte von Vitamin D der Altersgruppe 18 bis 100 Jahre, wird eine Abhängigkeit der Ergebnisse von der Jahreszeit vermutet. Es erfolgt neben der Altersbegrenzung auf 18 bis 80 Jahre somit auch eine saisonale Aufteilung der Messwerte. Je nach Zeitpunkt der Probenentnahme werden sie der Gruppe Sommer bzw. Winter zugeordnet. Der Bereich „Sommer“ umfasst die Zeiträume Mai bis Oktober 2015 und Mai bis Oktober 2016. In den Bereich „Winter“ fallen die Monate November und Dezember 2014 und Januar bis April 2015 sowie November, Dezember 2015 und Januar bis April 2016.

Die Gruppe „Sommer“ umfasst 8.526 Messwerte. Der zeitabhängige Median der Sommermesswerte ergibt nun einen schmaleren Schwankungsbereich verglichen mit dem zeitabhängigen Median aller Messwerte. Er liegt nun bei 24 bis 31 µg/l; höhere Werte werden dabei in den Septembermonaten, niedrigere Werte im Mai gemessen (s. Anhang Abb. 44d-f). Die Dichtefunktion erscheint wieder zweigipflig, wenn auch weniger stark ausgeprägt als vor der Aufteilung. Bei der abschließenden Schätzung der beiden Verteilungen, gelingt es nur ansatzweise für die geschlechtsspezifischen Datensätze, die Verteilung der pathologischen Werte abzugrenzen und darzustellen (s. Anhang Abb. 45b, c). Bei dem gemeinsamen Datensatz für Männer und Frauen ist dies nicht erfolgreich (s. Abb. 27).



**Abbildung 27:** Geschätzte Verteilung der nicht-pathologischen (grün) Werte und für alle Daten (blau) für Sommermesswerte von Vitamin D für beide Geschlechter von 18-100 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

Für die Berechnungen der Gruppe „Winter“ sind 8.436 Ergebnisse verfügbar. Der Median ergibt Ergebnisse zwischen 20 und 27 µg/l, wobei Tiefstwerte im Februar und Höchstwerte in den Monaten April und November verzeichnet werden (s. Anhang Abb. 46a-c). Auch hier zeigt die Kerndichtefunktion eine zweigipflige Verteilung, diesmal deutlich stärker ausgeprägt. Die Verteilungen der pathologischen und nicht-pathologischen Messwerte können für den gemeinsamen Datensatz der Männer und Frauen und für die Daten der Männer geschätzt werden (s. Abb. 28; s. Anhang Abb. 46f), aber nicht für die Daten der Frauen (s. Anhang Abb. 46e). Eine Berechnung von Referenzgrenzen gelingt somit nicht für diesen Datensatz.



**Abbildung 28:** Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Wintermesswerte von Vitamin D für beide Geschlechter von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

**Tabelle 18:** Saisonale Referenzintervalle für Vitamin D (µg/l) für Sommer und Winter. Alter in Jahren. G = Geschlecht, RI = Referenzintervall, N = Anzahl der Messwerte, F = Frauen, M = Männer.

Alter	G	RI Sommer	N	RI Winter	N
18 – 80	F, M	2,72 – 63,22	7.838	2,89 – 52,8	7.941
	F	3,07 – 60,46	5.073		5.017
	M	2,38 – 60,05	2.765	3,72 – 55,33	2.923

## **4. Diskussion**

Mittels eines indirekten Vorgehens werden Referenzgrenzen für Erwachsene für verschiedene Laborparameter bestimmt. Die R-basierte Excel-Anwendung „RLE“ der DGKL und mehr als 1,5 Millionen Messwerte des Zentrallabors des UKSH werden dafür eingesetzt. Für Calcium, Phosphat, Parathormon, TSH, fT3 und fT4 können Grenzwerte ermittelt werden, bei Renin, Aldosteron und Vitamin D sind die Berechnungen nicht erfolgreich. Als Altersspanne der Patienten werden zunächst die Jahre 18 bis 100 und 18 bis 80 gewählt. Je nach Datenumfang werden weitere, kleinere Altersgruppen gebildet. Im Folgenden werden kurz die bisherigen Anwendungen skizziert und anschließend Stärken und Schwächen der Methode aufgezeigt. Die ermittelten Referenzgrenzen werden einem Vergleich mit üblichen Referenzbereichen aus der Literatur (hier: Labor und Diagnose von Thomas), von den Geräteherstellern und aus dem Laborsystem des UKSH unterzogen. Dabei werden ihre klinische Anwendbarkeit und die nötigen Voraussetzungen der Laborparameter diskutiert. Es wird erörtert, welche Faktoren besonders für Vitamin D dazu beitragen können, mit dem RLE möglichst valide Referenzgrenzen zu generieren und somit die diagnostische Effizienz zu erhöhen. Abschließend wird ein Ausblick auf mögliche Weiterentwicklungen dieser Berechnungsmethode gegeben.

### **4.1 Referenzwertbestimmung mittels RLE: Bisherige Anwendungsgebiete**

Um Referenzgrenzen zu bestimmen, wird das Programm Reference Limit Estimator (RLE) angewandt. Es ermöglicht, die umfangreiche Menge an Messwerten, die in Laborinformationssystemen gespeichert sind, auszuwerten und aus ihr Referenzgrenzen abzuleiten. Dazu separiert es die gemischten Datensätze, die sowohl pathologische als auch nicht-pathologische Werte enthalten, gemäß ihrer beiden Verteilungen. Erstmals wurde das Programm 2007 angewandt (23). Dabei wurden Referenzgrenzen für verschiedene Aktivitätskonzentrationen von Enzymen bestimmt, unter anderem für Alaninaminotransferase und Aspartataminotransferase, und deren Aufteilung entsprechend Geschlecht und Alter untersucht. Die Daten stammten aus fünf verschiedenen Laboren. Insgesamt stimmten die berechneten Referenzgrenzen gut mit den bekannten Referenzgrenzen überein, wobei aber signifikante Unterschiede zwischen den Laboren beobachtet wurden. Daher wurde die Empfehlung, laborinterne Referenzgrenzen zu bestimmen, hervorgehoben. Solche vereinfachen zudem die interne Qualitätssicherung (34). Wenig später wurden Referenzgrenzen für Kreatinin mittels der Datensätze von 14 deutschen und italienischen Zentrallaboren bestimmt. Im Vergleich mit Referenzgrenzen, die auf direkte Weise ermittelt wurden, zeigte

sich, dass die berechneten unteren Referenzgrenzen meist niedriger waren als die Referenzwerte, die von den Geräteherstellern angegeben werden. Dies wurde aber als klinisch nicht relevant eingestuft. Insgesamt wurde eine hohe Übereinstimmung zwischen den Referenzgrenzen, die auf indirektem und direktem Wege generiert wurden, vorgefunden. Wieder bestätigten sich die Unterschiede zwischen den Laboren (30). Demzufolge werden bei unseren Berechnungen nur Datensätze des Zentrallabors des UKSH verwendet, um laborinterne Grenzwerte zu erstellen, die bestmöglich zu den Analysemethoden und dem Patientenkollektiv passen. Im Jahr 2011 wurde die Methode auf TSH angewandt. 13 deutsche und italienische Labore trugen ihre Daten bei und erneut konnten plausible Referenzgrenzen errechnet werden, wobei eine Zunahme der oberen Referenzgrenze und der Breite des Referenzintervalls mit ansteigendem Alter beobachtet wurde (35). Auch in der Erstellung von pädiatrischen, kontinuierlichen Referenzgrenzen (29, 36) und von Referenzgrenzen für hämatologische Parameter für Erwachsene (37) wurde das Vorgehen schon erfolgreich eingesetzt. Unter anderem wurde darin die vorherige Präselektion betont, um eine große Gruppe nicht-pathologischer Werte zu erhalten, gerade für Leukozyten und Erythrozyten (37). Dies wurde in den erfolgten Berechnungen berücksichtigt, indem intensivmedizinische und gynäkologische Patienten ausgeschlossen wurden.

## 4.2 Stärken und Schwächen der Methode

Die indirekte Vorgehensweise mit dem RLE ist aus mehreren Gründen vorteilhaft gegenüber der direkten Bestimmung von Referenzgrenzen. Zum einen ist sie viel weniger zeitintensiv, da die langwierige Phase der Zusammenstellung des Probandenkollektivs entfällt. Es ist keine Rekrutierung von Referenzindividuen nötig, die mit klinischen Untersuchungen, Fragebögen und Interviews selektiert werden. Auch die Phase der Probengewinnung mit anschließender Probenanalyse entfällt (s. Einleitung, 1.2). Zum anderen entstehen durch sie kaum zusätzliche Kosten. Messwerte von Laborparametern fallen tagtäglich in Krankenhäusern bei der Behandlung und Versorgung der Patienten an. Darüber hinaus ist die Umsetzbarkeit der Methode mit einem deutlich geringeren Aufwand verbunden. Ein Labormitarbeiter, der mit Excel und R vertraut ist, kann sich in das Programm einarbeiten und Grenzwerte errechnen. Die Anwendung der generierten Referenzgrenzen im Krankenhausalltag wiederum sollte unter Berücksichtigung der klinischen Relevanz mit Fachpersonal diskutiert werden.

Außerdem ermöglicht die sehr große Anzahl an Messwerten in Kombination mit einer sorgfältigen Auswahl der eingeschlossenen Messwerte und Überprüfung der generierten mit etablierten Referenzbereichen, verlässliche Referenzgrenzen zu erstellen. Die direkte

Bestimmung birgt die Gefahr, dass aufgrund des enormen Aufwands nur kleine Referenzgruppen gebildet und dementsprechend wenig Proben gesammelt werden können und wenige Messwerte zur Verfügung stehen. Dieser Mangel an Daten kann zu ungenauen Referenzbereichen führen (38).

Die berechneten Grenzwerte sind zudem spezifisch für das behandelte Patientenkollektiv, da sie aus den Patientendaten abgeleitet werden. Andere Verfahren können unter Umständen zu weniger passenden Grenzwerten führen, wenn die Referenzpopulation nicht repräsentativ ist für das Patientenkollektiv. Bei der direkten Bestimmung mit mindestens 120 gesunden Probanden müssen diese entsprechend definierter Kriterien selektiert werden (unter anderem Ausschluss von Systemerkrankungen, Medikamenteneinnahme und besonderen physiologischen Zuständen wie Schwangerschaft) (11). In der Realität ist dies jedoch oft nicht umsetzbar, da keine große Gruppe gesunder Probanden verfügbar ist, sodass Labore Referenzgruppen aus ihrem Arbeitsumfeld, privaten Organisationen oder Bluttransfusionszentren erstellen (39).

Oftmals werden Referenzgrenzen auch gar nicht erst selbst ermittelt, sondern von anderen Laboren übernommen. Dieser Transfer birgt die Gefahr, dass die Referenzpopulationen, präanalytische und analytische Verfahren nicht übereinstimmen oder betreffende Informationen unzugänglich sind, was eine Validierung erschwert (15). Auch die mit den Geräteanleitungen herausgegebenen Referenzgrenzen der Hersteller wiederum enthalten oft wenig weiterführende Informationen zur Referenzpopulation. Dadurch bleiben Aspekte der genauen Altersverteilung, Nationalität und anderer Einschluss- bzw. Ausschlussfaktoren unbekannt. Wenn eine Vergleichbarkeit der Populationen und Analysesysteme nicht gegeben ist, sollten Referenzgrenzen nicht transferiert und angewendet werden, um klinische Entscheidungen zu treffen. Demgegenüber ist die Vorgehensweise zur Bestimmung von Referenzgrenzen mittels des RLE im Vorteil. Diese führt zu Grenzwerten, die repräsentativ für das behandelte Patientenkollektiv sind.

Außerdem stellt die Verfügbarkeit von Messwerten seltener Parameter einen Vorteil dar. Die Erhebung der Messwerte erfolgt von Patienten und ist daher stets medizinisch indiziert. Im Gegensatz zur direkten Methode zur Bestimmung von Referenzgrenzen ergeben sich dadurch keine ethischen Schwierigkeiten (18). Somit ermöglicht die Methode die Errechnung von Referenzgrenzen für seltene Parameter, wie beispielsweise von Frühgeborenen (40).

Doch den gespeicherten Daten mangelt es an klinischen Angaben zu den Patienten, die einen Effekt auf die Laborparameter haben können. Dazu gehören etwa Medikation, vorliegenden

Komorbiditäten (gerade mit steigendem Lebensalter), BMI, Konsumverhalten (Alkohol, Nikotin), Ethnie, körperlicher Aktivität und präanalytischen Aspekte (zB. nüchtern, Körperposition). Bei der Erstellung von pädiatrischen Referenzgrenzen wären Angaben zum Pubertätsstadium nach Tanner hilfreich, da die Pubertät wesentliche Auswirkungen auf die Konzentrationen der Geschlechts- und Wachstumshormone hat (40). Diese weitergehende Informationen würden es ermöglichen, die Daten besser zu unterteilen (15), also bestimmte Patientengruppen auszuschließen oder gezielt für diese Referenzgrenzen zu erstellen. So könnten eventuelle Effekte auf die Referenzbereiche untersucht und möglichst spezifische Grenzwerte ermittelt werden. Im Hinblick auf Vitamin D wäre ein Vermerk bezüglich einer bestehenden Therapie wegen eines Mangels von Interesse, womit sich Patienten unter Vitamin D-Substitution ausschließen ließen. Schätzungen gehen davon aus, dass etwa eine Milliarde Menschen einen Vitamin D-Mangel aufweisen (41). Die Patienten mit medikamentöser Zufuhr lieferten in den erfolgten Berechnungen vermutlich einen großen Anteil an Messwerten. Das erschwerte es dem RLE, die Verteilung der pathologischen Werte von der nicht-pathologischen Werte abzugrenzen.

Eine weitere Schwierigkeit der Methode ist die Berechnung auf Grundlage der gemischten Datensätze mit einem gewissen Anteil pathologischer Werte. Diese können von schwer kranken Patienten stammen und es ist fraglich, inwiefern dieser Anteil die Referenzgrenzen beeinflusst. Zierk et al. sind dem nachgegangen, indem Datensätze mit steigendem Anteil pathologischer Messwerte erstellt und aus diesen Referenzgrenzen für hämatologische Parameter ermittelt wurden. Dafür wurden vier verschiedene Patientenkohorten zusammengestellt, deren Messwerte zur Berechnung der Referenzgrenzen verwendet wurden. Dabei stellte sich für die meisten untersuchten Parameter heraus, dass die Referenzgrenzen mit zunehmendem Anteil pathologischer Messwerte relativ stabil blieben. Beim vierten Patientenkollektiv war von der größten Anzahl pathologischer Werte auszugehen, da es sich um onkologische und intensivmedizinische Patienten sowie um Patienten unter Bestrahlung handelte. Dadurch wurden die ermittelten Referenzgrenzen stärker beeinflusst (Zunahme der oberen Referenzgrenzen bei Leukozyten, Verbreiterung des Referenzintervalls bei Erythrozytenindices), woraufhin ein Ausschluss dieser Patientengruppe vorgenommen wurde. In simulierten Datensätzen zeigte sich, dass ein Anteil der pathologischen Werte von 15-20% oder mehr einen relevanten Effekt auf die Grenzwerte von problematischen Parametern haben kann. Letzteres sind Parameter, bei denen sich die Verteilung der nicht-pathologischen und pathologischen Werte stark überschneiden und ihre Standardabweichungen sehr ähnlich sind (37). In den durchgeführten Berechnungen für endokrinologische Parameter wurden

Patientenkollektive mit sehr kritischem Gesundheitszustand von vornherein ausgeschlossen, weshalb ein relevanter Effekt auf die Referenzgrenzen sehr unwahrscheinlich ist.

### **4.3 Diskussion der Ergebnisse unter Bezugnahme auf die klinische Anwendbarkeit sowie Vergleich mit bestehenden Referenzgrenzen**

Bei den Resultaten für Calcium konnten einige Trends in den generierten Grafiken beobachtet werden. In der 50.-Perzentile zeigte sich ein altersabhängiger Abfall der Calcium-Konzentration, welcher als Ausdruck der veränderten Calcium-Hämostase im Alter zu werten ist (42). Dieser Abfall ist vor allem durch einen Vitamin D-Mangel bedingt. Dadurch kommt es zu einer Abnahme der intestinalen Calciumabsorption und auch die renale Vitamin D-Synthese sinkt (43). Im Median über die Zeit ließ sich ein anfangs langsam ansteigender Kurvenverlauf beobachten, der sich nach etwa 9 Monaten stabilisierte. Begründen lässt sich das Phänomen mit der geringeren Anzahl an Messwerten für den Beginn des Analysezeitraums aufgrund einer Umstellung des Laborinformationssystems des UKSH.

Referenzgrenzen für Calcium bedürfen keiner geschlechtsspezifischen Aufteilung. Die untere Referenzgrenze unterscheidet sich minimal (Frauen 18-80 J.: 2,11 – 2,56 mmol/l; Männer: 2,08 – 2,56 mmol/l). Eine saisonale Unterteilung der Referenzgrenzen für Calcium ist generell nicht indiziert, die Grenzwerte für beide Geschlechter von 18 bis 80 Jahren sind fast identisch (Sommer: 2,1 – 2,56; Winter: 2,09 – 2,56). Die größeren Abweichungen von den zulässigen Grenzen in Form der höheren unteren Referenzgrenzen in den Altersgruppen 18 bis 27 und 28 bis 37 Jahre spiegeln vielmehr die Altersabhängigkeit wider. Eine entsprechend engere Unterteilung der Altersstufen erscheint jedoch nicht sinnvoll, da die klinische Relevanz dieses Effekts als gering einzustufen ist.

Im Vergleich der ermittelten Referenzgrenzen mit den Referenzgrenzen aus der Literatur, dem Laborsystem des UKSH und den Geräteherstellern Roche und Siemens besteht eine gute Übereinstimmung (s. Tab. 19).

**Tabelle 19:** Gegenüberstellung der Referenzintervalle des RLE (Altersklasse 18 – 80 Jahre), des Laborsystems des UKSH, der Hersteller Cobas und Dimension und der Literatur für Calcium, Phosphat, Parathormon, fT3, fT4 und TSH. G = Geschlecht, N = Anzahl der Messwerte, die in die Berechnungen des RLE eingegangen sind. PTH = Parathormon, F = Frauen, M = Männer. Thomas = Thomas L. (2016): Labor und Diagnose. Elektronische Auflage, TH-Books-Verlags-Gesellschaft, Frankfurt/Main.

Parameter	G	N (RLE)	RLE	Labor	Cobas	Dimension	Thomas
<b>Calcium</b> <b>(mmol/l)</b>	F, M	556.550	2,09 – 2,56	2,15 – 2,5	2,15 – 2,5	2,12 – 2,52	2,15 – 2,58
	F	251.724	2,11 – 2,56				
	M	304.706	2,08 – 2,56				
<b>Phosphat</b> <b>(mmol/l)</b>	F, M	234.242	0,55 – 1,48	0,87 – 1,45	0,81 – 1,45	0,81 – 1,58	0,84 – 1,45
	F	97.893	0,58 – 1,48				
	M	136.319	0,55 – 1,5				
<b>PTH</b> <b>(pg/ml)</b>	F, M	22.084	19,74 – 107,43	15 – 65	15 – 65		15 – 65
	F	13.392	20,13 – 99,47				
	M	8.691	18,93 – 125,72				
<b>fT3</b> <b>(pmol/l)</b>	F, M	100.692	3,37 – 6,47		3,1 – 6,8	3,4 – 6,1	3,9 – 6,7
	F	56.603	3,43 – 6,29	3,71 – 6,7			
	M	44.027	3,39 – 6,64	4,09 – 6,65			
<b>fT4</b> <b>(pmol/l)</b>	F, M	103.170	10,49 – 21,84		12 – 22	9,8 – 18,8	12,7 – 20,8
	F	58.193	10,81 – 22,57	12,3 – 20,2			
	M	44.904	10,3 – 21,72	13,1 – 21,3			
<b>TSH</b> <b>(mIU/l)</b>	F, M	268.001	0,19 – 3,36			0,358 – 3,74	0,4 – 4,2
	F	132.764	0,14 – 3,29	0,3 – 3,94	0,3 – 3,94		0,3 – 3,6
	M	135.089	0,28 – 3,99	0,3 – 3,18	0,3 – 3,18		

Die Abweichungen zwischen den ermittelten und den herkömmlichen Grenzwerten liegen bei 0,02 bis maximal 0,11 mmol/l (im Vergleich der Referenzgrenzen von 18 – 80 Jahren mit der Literatur bzw. im Vergleich der Referenzgrenzen von 18 – 60 Jahren mit der Herstellerangabe).

Auch für Phosphat sind keine geschlechtsspezifischen oder saisonalen Referenzbereiche notwendig, da diese keine großen Abweichungen zeigen. Die Aufteilung nach eingesetzten Analysegeräten zeigt besonders für zwei (Cobas 8000 K2 (M1), Cobas 8000 K2 (M2)) der fünf Geräte höhere obere und untere Grenzwerte verglichen mit dem allgemeinen, nicht nach Geräten aufgeteilten Referenzbereich (allgemein: 0,55 – 1,48 mmol/l; Cobas 8000 K2 (M1): 0,62 – 1,59 mmol/l; Cobas 8000 K2 (M2): 0,61 – 1,6 mmol/l). Diese Geräte erheben tendenziell höhere Messwerte. Die unteren Referenzgrenzen weichen um 0,07 bzw. 0,06 mmol/l und die oberen um 0,11 bzw. 0,12 ab von den geräteunspezifischen Grenzwerten. Hier erscheint ein gemeinsamer Referenzbereich für alle Geräte ausreichend, da die klinische Relevanz dieser leicht höheren Referenzgrenzen insgesamt gering ist. Bei Patienten mit normaler Nierenfunktion würden leicht elevierte Phosphatkonzentrationen den klinisch tätigen Arzt nicht zu einer Therapie veranlassen, da der Körper eine akute Hyperphosphatämie schnell ausgleicht (44). Demgegenüber wird von Malberti et al. betont, dass eine chronische

Hyperphosphatämie bevorzugt bei chronischer Niereninsuffizienz auftritt und aufgrund von Folgeschäden therapiert werden sollte. Die Arbeitsgruppe der National Kidney Foundation NKF-K/DOQI hat für die Stadien der Nierenerkrankung 3-4 den Bereich 0,87-1,49 mmol/l empfohlen, bei Dialysepatienten werden mit 1,13–1,78 mmol/l höhere Phosphatkonzentrationen toleriert (45). Für diese bestimmte Patientengruppe könnte also die Anwendung gerätespezifischer Referenzgrenzen überdacht werden, um die Messwerte bestmöglich einordnen zu können.

Im Vergleich mit den bereits etablierten Referenzgrenzen für Phosphat fällt auf, dass die oberen Referenzgrenzen beinahe übereinstimmen. Die generierten unteren Referenzgrenzen zeigen demgegenüber allerdings stärkere Abweichungen. Sie sind deutlich niedriger, mit dem größten Unterschied zur Referenzgrenze des Laborsystems (Quelle: L. Thomas, Labor und Diagnose, 8. Auflage 2012; RLE: 0,55 mmol/l; System: 0,87 mmol). Es stellt sich somit die Frage, ob eine klinische Anwendung der errechneten Referenzgrenzen ratsam ist. Messwerte, die sonst als hypophosphatämisch eingestuft würden, würden in den Referenzbereich fallen. Dabei ist festzustellen, dass für moderate Hypophosphatämien (0,32 - 0,80 mmol/l) abgesehen von beatmeten Patienten kaum Hinweise für bedeutsame medizinische Folgeerscheinungen belegt sind. Intubierte Patienten zeigten verbesserte Beatmungsparameter nach Phosphatgabe (46). Amanzadeh et al. heben hervor, dass schwere Hypophosphatämien (<0,32 mmol/l) behandlungsbedürftig sind gerade bei kritisch kranken Patienten, intubierten Patienten und bei entsprechender klinischer Symptomatik. Demzufolge erscheint der Einsatz der Referenzgrenzen als Bezugsgrößen in der Klinik möglich. Festzuhalten ist, dass bei Intensivpatienten schon frühzeitiger niedrigere Messwerte innerhalb des Referenzbereichs therapiert werden könnten und besonders für Patienten mit Niereninsuffizienz sollte die klinische Relevanz von gerätespezifischen Referenzgrenzen weiter untersucht werden.

Im Verlauf der 50.-Perzentile bezogen auf das Alter wurde für TSH ein Abfall der gemessenen Konzentrationen festgestellt (18-27 J.: 0,51 – 4,63 mIU/l; 68 – 77 J.: 0,13 – 3,37 mIU/l). Dieses Phänomen deckt sich mit Erkenntnissen aus der Studie von Clerico et al. Sie bediente sich ebenfalls des indirekten Vorgehens zur Bestimmung von Referenzgrenzen für TSH. Es wurden 146.801 Messwerte aus den Laborsystemen von vier verschiedenen Krankenhäusern in Italien untersucht (47). In einer Studie durchgeführt im Nordosten Deutschlands zur Erhebung von Referenzgrenzen für TSH, fT3 und fT4 auf direkte Weise zeigte sich ebenfalls eine altersabhängige Abnahme der TSH-Level (48).

Für die fT3-Konzentration wurde ebenfalls eine negative Korrelation zum Alter festgestellt, was sich vorher bereits in anderen Studien gezeigt hatte (48, 49). Für fT4 wurde ein altersabhängiger Anstieg der Konzentrationen beobachtet. Es handelt sich um einen leichten Anstieg der 50. Perzentile von etwa 16 auf 17 mIU/l zwischen 18 und 80 Jahren, gerade im Hinblick auf den breiten Referenzbereich für fT4. Diese Beobachtung steht im Gegensatz zu Kratzsch et al. Diese fanden mit zunehmendem Alter abfallende ft4-Werte. In ihren Untersuchungen wurden auf direkte Weise Referenzgrenzen ermittelt. 453 Proben von gesunden Personen aus Deutschland wurden analysiert. Völzke et al. fanden ein leichtes Absinken der fT4-Konzentration mit dem Alter bei den Männern, bei den Frauen einen Anstieg. Im Median über die Zeit zeigten fT3 und fT4 im Sommer ihre höchsten Konzentrationen, im Winter die niedrigsten. Für TSH war dies umgekehrt der Fall (Median TSH 11/2015, 01/2016: 1,54 mIU/l; Median TSH 09/2016: 1,41 mIU/l), was sich aus dem negativen Feedback zwischen TSH und fT3 bzw. fT4 erklärt. Die signifikante Abhängigkeit von TSH von der Jahreszeit wurde bereits gezeigt (50). Dabei wurden ebenfalls Höchstwerte für TSH im Winter vorgefunden.

Für TSH zeigten sich Unterschiede in den geschlechtsspezifischen Referenzgrenzen, für Männer ist die untere und vor allem die obere Referenzgrenze deutlich höher. Hier sollten getrennte Referenzbereiche nach Geschlecht verwendet werden. Bei der Unterteilung der Daten in Altersklassen von 10 Jahren zeigt die jüngste Altersgruppe die höchsten Referenzgrenzen, es bestätigt sich der altersabhängige Abfall der TSH-Konzentration. Arzideh et al. hingegen fanden höhere obere Referenzgrenzen für Frauen, auch war der Referenzbereich breiter als für Männer (35). fT3 und fT4 zeigen Referenzgrenzen, die keine Unterscheidung nach Geschlecht notwendig erscheinen lassen. Die Altersgruppen in 15 Jahresabschnitten zeigten für fT3 erneut den altersbedingten Konzentrationsabfall. Für TSH und fT3 könnte man also die Anwendung altersadaptierter Referenzgrenzen in Betracht ziehen. Bei fT4 ist dies nicht notwendig.

Stellt man die ermittelten Grenzen für TSH den etablierten gegenüber, fällt auf, dass die unteren als auch oberen Referenzgrenzen stets niedriger sind. Außerdem zeigt sich ein gegenteiliges Verhalten der oberen Grenzwerte für Männer bzw. Frauen. Mit der indirekten Methode wurde eine niedrigere obere Referenzgrenze für Frauen (3,29 mIU/l) ermittelt, für Männer war diese höher (3,99 mIU/l). Die Referenzgrenzen von Cobas sind jedoch für Frauen höher (3,94 mIU/l) und für Männer niedriger (3,18 mIU/l). Der größte Unterschied ergibt sich mit 0,84 mIU/l für die obere Referenzgrenze der Männer im Vergleich mit den Angaben aus der Literatur.

In der Studie von Arzideh et al. betrug die durchschnittliche untere bzw. obere Referenzgrenze für beide Geschlechter gemeinsam 0,31 mU/L bzw. 3,53 mU/L. Damit waren die Werte mit den bisher publizierten durchschnittlichen Grenzwerten für TSH gut übereinstimmend. (0,4 mU/L - 3,50 mU/L) (35). Die nun berechneten Referenzgrenzen von 0,19 und 3,36 mU/l für beide Geschlechter weisen damit zwar niedrigere Konzentrationen auf, fügen sich aber insgesamt in das sehr variable Bild der Referenzbereiche für TSH ein. Die Abweichungen von TSH zu den etablierten Grenzwerten sind größer, als dies bei Calcium und Phosphat der Fall war, dabei ist aber anzumerken, dass TSH auch einen breiteren Referenzbereich aufweist. Ohnehin würde eine Therapie hinsichtlich einer Hyper- oder Hypothyreose nicht aufgrund eines einzelnen TSH-Messwertes erfolgen. Weiteren Messungen wie die der fT4- und Gesamt-T3-Konzentration erfolgen simultan oder anschließend (51). Erst in der Zusammenschau dieser Ergebnisse mit der klinischen Symptomatik wird eine Entscheidung getroffen.

Bei fT3 zeigt sich eine gute Übereinstimmung mit den bestehenden Referenzgrenzen, am höchsten ist diese mit den Angaben von Cobas (RLE: 3,37- 6,47; Cobas: 3,1- 6,8 pmol/l). Der größte, aber angesichts der Breite des Referenzbereichs vertretbare Unterschied besteht mit 0,7 pmol/l bei der unteren Referenzgrenze der Männer zur entsprechenden Grenze im Laborsystem (Quelle: Herstellerangaben).

In der Gegenüberstellung der Referenzgrenzen für fT4 sind größere Abweichungen zu bemerken, so zum Beispiel die höhere obere Referenzgrenze für beide Geschlechter als die entsprechende Grenze vom Hersteller Siemens (21,84 vs. 18,8 pmol/l). Insgesamt ist die Übereinstimmung dennoch recht hoch, gerade der zentralen Bereiche der Referenzintervalle. Die Differenzen mit den Referenzgrenzen sind annehmbar, da fT4 den größten physiologischen Schwankungsbereich aller untersuchten Parameter besitzt. Zukünftig wäre bei den Schilddrüsenhormonen ein Vermerk des BMIs im Laborsystem erstrebenswert, da gezeigt wurde, dass die TSH-Level positiv mit diesem korreliert sind (52). In einer Studie von Bjoro et al. führte ein Ausschluss von Patienten mit positiven Thyreoperoxidase-Antikörpern zum Abfall der oberen Referenzgrenze für TSH (53). Auch zeigte sich, dass die Alters- und Geschlechtsdifferenzen des Medians und der oberen Referenzgrenze für TSH unter diesem Ausschluss nicht mehr vorlagen. Somit könnte der Hinweis bezüglich gemessener Antikörper helfen, um die variablen Referenzgrenzen der Schilddrüsenhormone weiter zu spezifizieren. Außerdem ist für die Schilddrüsenhormone der Jod-Status einer Population von Bedeutung, da gezeigt werden konnte, dass die regionale Versorgung mit Jod die Referenzgrenzen beeinflusst. Es wurden Referenzgrenzen in einer vormals jodunterversorgten Region gefunden, die sich

unterschieden von Referenzgrenzen aus einem Gebiet mit ausreichender Jod-Versorgung (48). Dies unterstreicht die Notwendigkeit der Erstellung von regions- bzw. populationsspezifischen Referenzgrenzen für die Parameter der Schilddrüsenfunktion.

Für Parathormon wurde ein leichter Konzentrationsanstieg mit dem Alter beobachtet, was sich bereits in anderen Studien gezeigt hatte (54, 55). Es könnten geschlechtsspezifische Referenzgrenzen erwogen werden. Die obere Referenzgrenze der Frauen liegt deutlich unterhalb der Männer. Auch eine Stratifikation hinsichtlich der Messgeräte erscheint sinnvoll. Die beiden eingesetzten Geräte hatten sich bereits im Äquivalenztest als sehr unterschiedlich präsentiert und auch die gerätespezifischen Referenzgrenzen zeigen, dass Messwerte, die mit Gerät 1 bestimmt wurden, einen breiteren Referenzbereich aufweisen als Messwerte von Gerät 2. Diese Tendenz muss jedoch unter Berücksichtigung der sehr geringen Datenmenge von Gerät 2 betrachtet werden.

Im Vergleich der ermittelten Referenzgrenzen mit den gängigen Referenzgrenzen zeigen sich große Abweichungen besonders der oberen Referenzgrenze. Diese weicht um 34,47 bis 60,72 pg/ml ab (Frauen bzw. Männer). Daraus ergibt sich ein sehr breiter Referenzbereich für Parathormon. Ursache dafür könnte die Rolle dieses Laborparameters im klinischen Kontext sein, da seine Messung nicht routinemäßig erfolgt, sondern bei bestimmten Indikationen, wie etwa bei chronischer Niereninsuffizienz, Verdacht auf Hyperparathyreodismus, Hyperkalzämie, Vitamin-D-Mangel und intraoperativ bei Eingriffen an Neben-/Schilddrüsen (25). Dadurch ist es wahrscheinlich, dass vorwiegend Messwerte erkrankter Personen im Datensatz von Parathormon vertreten sind. Gerade der Vitamin D-Status sollte berücksichtigt werden, da er in gegensinniger Beziehung zum Parathormon-Level steht (55). Auch die Nierenfunktion spielt eine wichtige Rolle, denn es zeigten sich bei reduzierter eGFR erhöhte Parathormon-Konzentrationen (56). Dies liegt an der reduzierten renalen Umwandlung der Vitamin D-Vorstufe in aktives Vitamin D und der daraus resultierenden niedrigen Calcium-Konzentration, näher beschrieben in Abschnitt 1.4. Aufgrund dieser Einflussfaktoren auf die Parathormon-Konzentration, sollten Referenzgrenzen von Nieren-gesunden Patienten mit ausreichender Vitamin D-Versorgung bestimmt werden (57).

In den erfolgten Berechnungen wurden Patienten von nephrologischen Stationen ausgeschlossen, dennoch können Messwerte von Nieren-erkrankten Personen damit nicht vollständig erfasst worden sein. Schließlich können sie auch auf anderen Stationen in Behandlung gewesen sein. Des Weiteren konnte kein Abgleich mit dem Vitamin D-Status vorgenommen werden, um Patienten mit zu niedrigen Konzentrationen auszuschließen. Diese

Faktoren und die Tatsache, dass es sich bei Parathormon um einen Spezial-Parameter handelt, der nicht standardmäßig erfasst wird, müssen zu einem großen Anteil pathologischer Werte beigetragen haben. Dies hat die Abgrenzung der Verteilungen der nicht-pathologischen von den pathologischen Werten erschwert. Auch die unterschiedlichen Messgeräte haben die Ermittlung von plausiblen Referenzgrenzen weiter beeinträchtigt sowie der mäßige Datenumfang. Die mittels RLE bestimmten Referenzgrenzen sind somit ungeeignet, um in der klinischen Praxis eingesetzt zu werden. Weitere klinische Informationen sind nötig, um die Daten besser präselektieren zu können und somit plausible Referenzgrenzen zu erhalten.

#### **4.4 Vitamin D: Schlussfolgerungen zu notwendigen Merkmalen eines RLE-geeigneten Laborparameters**

Für den Datensatz der Vitamin D-Messwerte war die Anwendung des RLE aus verschiedenen Gründen erschwert. Einerseits war der Datensatz an Messwerten nicht sehr umfangreich. Zunächst lagen 27.688 Messwerte vor, nach Eingrenzung auf 18 bis 100 Jahre und Ausschluss bestimmter Patientengruppen 21.104 Werte. Dies leitet über zum nächsten Grund: der starken saisonalen Schwankungen der Vitamin D-Konzentrationen. Bei der visuellen Inspektion des Medians über die Zeit wurde dies offensichtlich und daher eine Unterteilung der Messwerte von 18 bis 80 Jahre in Werte, die in Sommer- bzw. Wintermonaten erhoben wurden, vorgenommen. Auch dies trug zum schrumpfenden Umfang an Daten weiter bei (Sommer: 7.838 Messwerte für beide Geschlechter, 5.073 für Frauen, 2.765 für Männer; Winter: 7.941 Messwerte für beide Geschlechter, 5.017 für Frauen, 2.923 für Männer). Der RLE benötigt für die Berechnungen eine Mindestanzahl von 2000 Messwerten, empfohlen werden 4000 (23). Des Weiteren weist der Datensatz für Vitamin D keine Normalverteilung auf (s. Abb. 22). Es stellen sich deutlich zwei große Subgruppen dar, wodurch der RLE nicht in der Lage ist, die sonst häufigeren nicht-pathologischen Messwerte von den selteneren pathologischen Werten abzugrenzen. Die ungewöhnliche Verteilung führt zu der Annahme, dass zwei Populationen der nicht-pathologischen Werte bestehen. Dabei stellt die erste Subgruppe das Patientenkollektiv dar, das einen niedrigeren Vitamin D-Spiegel aufweist. Die zweite Subgruppe setzt sich aus Patienten mit höheren Vitamin D-Konzentrationen zusammen. Vermutlich handelt es dabei um Patienten, die Vitamin D substituieren. Letztlich zeigen diese genannten Problematiken wichtige Merkmale auf, die ein Laborparameter erfüllen sollte, um mithilfe der Methode des RLE Referenzgrenzen zu gewinnen: er sollte normalverteilt sein und häufig angefordert werden in der klinischen Praxis, um einen umfangreichen Datensatz zu gewährleisten, und nicht durch Substitution oder Medikation beeinflusst werden.

## **4.5 Vitamin D: Einflussfaktoren auf Referenzgrenzen und zukünftige Berücksichtigungen**

Die Prävalenz des Vitamin D-Mangels mit Werten unter 50 nmol/l ist hoch in Deutschland: 57% der Männer und 58% der Frauen sind betroffen. Gerade Frauen im höheren Lebensalter sind betroffen (75% der Frauen von 65 bis 79 Jahren) (58). Hintz Peter et al. nutzten für ihre Erhebungen die Daten von 2267 Frauen und 1763 Männern im Alter von 18 bis 79 Jahren, die im Rahmen des German National Health Interview and Examination Survey 1998 (GNHIES) gesammelt wurden. Dabei stellte sich heraus, dass 1.5% der Männer und 3.8% der Frauen Vitamin D-Präparate zusetzten. Bezogen auf unsere Daten kann aber von einem deutlich höheren Anteil ausgegangen werden, da in den letzten Jahren verstärkt Aufmerksamkeit auf die Bedeutung von Vitamin D gerichtet wurde und die Verteilung dies nahelegt.

Ursachen für einen Mangel an Vitamin D sind vielfältig. Da die Umwandlung in aktives Vitamin D UV-abhängig in der Haut geschieht, ist beispielsweise eine unzureichende Sonnenexposition ein Risikofaktor. Dies konnte für bestimmte Berufsgruppen wie etwa Schichtarbeiter und Pflegekräfte, aber auch generell für Arbeiter, die vornehmlich drinnen arbeiten, gezeigt werden. Bei ihnen besteht großteils ein Vitamin D-Mangel (80% der Schichtarbeiter, 77% der drinnen Arbeitenden, 72% der Studierenden aus dem Gesundheitswesen) (59). Dies führt zu der Überlegung, ob unter den klinischen Angaben ein Vermerk des Berufs hilfreich wäre, um gezielt Messwerte von Patienten dieser Berufsgruppen von der Referenzwertbestimmung auszuschließen. Sollte ihr Anteil an den verfügbaren Messwerten besonders groß sein, wäre das sinnvoll.

Bei der kutanen Vitamin D-Synthese ist auch die Hautpigmentierung zu beachten. In Kreta wurden so sehr niedrige Vitamin D-Spiegel gefunden. Durchschnittlich  $19.48 \pm 9.51$  (ng/ml+SD) für Männer und  $18.01 \pm 9.01$  (ng/ml+SD) für Frauen bei einem Grenzwert von 20 ng/ml. 21% der Einwohner waren zudem von einem starken Vitamin D-Mangel betroffen mit Werten unter 10 ng/ml. Dies wurde neben Ernährungsgewohnheiten auf die starke Hautpigmentierung zurückgeführt (60, 61). Dies lässt die zusätzliche Angabe der Ethnie im Laborinformationssystem zur Verknüpfung mit den Messwerten sinnvoll erscheinen bei der Berechnung von Referenzgrenzen für Vitamin D, da in Deutschland ein steigender Anteil von Immigranten und Geflüchteten vertreten ist.

Die ungenügende Zufuhr von Vitamin D über die Nahrung, stellt einen weiteren Einflussfaktor dar. Eine empfohlene Vitamin D-Einnahme von 5 mg/Tag wird von den meisten Männern (80.9%) und Frauen (88.5%) nicht erreicht, sie liegt durchschnittlich bei etwa der Hälfte des

Wertes (58). Dies kann einen Mangel bestärken, wenngleich aber die Hauptquelle von Vitamin D die Sonnenexposition darstellt (62). Ein Hinweis zum BMI des Patienten im Laborsystem könnte überdacht werden, da dieser sich gegenläufig zur Vitamin D-Konzentration verhält. Dies ist darin begründet, dass sich Vitamin D in Körperfett ablagert. Auch der ungesündere Lebensstil von Patienten mit hohem BMI, der eventuell durch einen geringeren Aufenthalt im Freien und eine unausgewogene Ernährung geprägt ist, kann den Mangel bestärken (63). Es konnte gezeigt werden, dass bei Patienten mit einem  $\text{BMI} \geq 40$  32% der Frauen und 46% der Männer von einem Vitamin D-Mangel betroffen sind.

Die Festlegung von festen Referenzbereichen für Vitamin D ist aufgrund der vielen Einflussgrößen, von denen im vorigen einige wichtige genannt wurden, problematisch. Weitere Faktoren spielen eine Rolle, wie etwa Bekleidung, Verwendung von Sonnenschutzmitteln, Breitengrad des Wohnorts und Vitamin D Substitution (61). Ferrari et al. schlagen daher vor, Referenzgrenzen für Vitamin D mittels eines Algorithmus unter Berücksichtigung nicht nur von Alter und Geschlecht, sondern auch der Jahreszeit, der UV-Exposition, des BMI, der Ethnie, des Messverfahrens und einer gegebenenfalls bestehenden Vitamin D Substitution. Die Berücksichtigung dieser Faktoren ließe sich mittels des RLEs umsetzen, wenn es zu einer Verknüpfung der Messwerte mit den entsprechenden weitergehenden klinischen Angaben kommt. Hinsichtlich der durchgeföhrten und erfolglosen Berechnungen würde in Zukunft vorrangig die Angabe über eine medikamentöse Zufuhr von Vitamin D zusammen mit einem größeren Datensatz (generiert z. B. durch einen Einschluss von Messwerten gesammelt über einen längeren Zeitraum) die Ermittlung von Referenzgrenzen unter Verwendung des RLE wahrscheinlich ermöglichen.

#### **4.6 Ausblick: Weitere Anwendungsmöglichkeiten der Methode**

Die durchgeföhrten Berechnungen von Referenzgrenzen für verschiedene endokrinologische Parameter haben gezeigt, dass mit der angewandten indirekten Methode Referenzgrenzen ermittelt werden konnten, die denen sehr ähnlich sind, die mit der direkten Methode bestimmt wurden. Dabei wurden besonders für Calcium und Phosphat sehr hohe Übereinstimmungen gefunden. Dies liegt in physiologischen (annähernde Normalverteilung der Messwerte) und strukturellen Gegebenheiten (große Datenmengen aufgrund hoher Anforderungsfrequenz) begründet. Für die Schilddrüsenhormone ergaben sich größere Abweichungen. Bei Parathormon ergab sich ein zu breiter Referenzbereich bedingt wahrscheinlich durch einen zu großen Anteil pathologischer Werte. Die errechneten Referenzgrenzen erscheinen daher nicht plausibel. Grenzwerte für Vitamin D, Renin und Aldosteron konnten nicht ermittelt werden.

Zukünftig wäre die Vernetzung weiterer klinischer Angaben mit den Messwerten im Laborsystem erstrebenswert. Dies ist ganz konkret für Vitamin D der Fall, wo eine zweigipflige Verteilung gefunden wurde. Dies legt die Vermutung nahe, dass die Gruppe der nicht-pathologischen Messwerte sich aus zwei Gruppen zusammensetzt: Patienten, die Vitamin D substituieren und Patienten, die dies nicht tun. Informationen bezüglich der Sonnenexposition, der Ethnie und des BMI könnten ebenfalls bei der Referenzwertbestimmung zukünftig berücksichtigt werden und ihr Einfluss auf die resultierenden Referenzgrenzen untersucht werden. Auch um relevante Referenzgrenzen für Parathormon errechnen zu können, müssten klinische Angaben einbezogen werden. Hier wären Hinweise bezüglich Niereneinschränkungen und Vitamin D-Status hilfreich, um den Anteil pathologischer Messwerte an dem Gesamtdatensatz weiter zu reduzieren.

Insgesamt bietet sich mit dieser indirekten Methode eine Alternative zur konventionellen direkten Erstellung von Referenzgrenzen. Sie ist deutlich weniger kosten- und zeitintensiv und bietet Laboren die Möglichkeit, laborinterne Referenzgrenzen zu erstellen. Voraussetzung ist ein umfangreicher Datenpool. Für selten angeforderte Parameter oder kleine Labore kann der Zusammenschluss mit anderen Laboren erwogen werden, wenngleich dann wiederum Aspekte der abweichenden analytischen Verfahren und Patientengruppen bedacht werden müssen. Spezielle Parameter bedürfen allerdings einer genaueren Präselektion anhand klinischer Informationen.

## **5. Zusammenfassung**

Medizinische Referenzgrenzen sollten von jedem Labor für jeden angebotenen Test anhand einer Gruppe gesunder Personen selbst bestimmt werden. Diese Empfehlung zur direkten Bestimmung stellt sich als aufwändig und kaum realisierbar dar, weshalb ihr kaum entsprochen wird. Labore behelfen sich, indem sie bestehende Referenzgrenzen übernehmen, welche aus der Literatur, von den Geräteherstellern oder anderen Laboren stammen. Diese Strategie ist nicht unkritisch, da oftmals keine Validierung erfolgt und die angewandten Grenzen unpassend sein können hinsichtlich der Patientenpopulation, der präanalytischen Verfahren und der Messgeräte. Daher bietet sich die indirekte Methode als Alternative an. Diese verwendet Patientendaten, welche in Laborinformationssystemen gespeichert sind, um aus der gemischten Verteilung von pathologischen und nicht-pathologischen Werten Referenzgrenzen zu bestimmen. Diese passen somit zu den laboreigenen Analysemethoden, sowie zu dem behandelten Patientenkollektiv.

Es wurden Referenzgrenzen für Erwachsene für verschiedene endokrinologische Parameter (Calcium, Phosphat, TSH, fT3, fT4, 25-OH-Vitamin D, Parathormon, Renin, Aldosteron) berechnet. Dafür wurde eine indirekte Methode, der von Arzideh entwickelte Reference Limit Estimator (RLE), auf die Messwerte von Patienten aus dem Zentrallabor des Universitätsklinikums Schleswig-Holstein (UKSH) angewandt. Diese ermöglichte es, die Verteilung der nicht-pathologischen Werte von der Verteilung der pathologischen Werte abzugrenzen und aus den nicht-pathologischen Daten retrospektiv Referenzgrenzen für Erwachsene zu ermitteln. Eine weitere Aufteilung nach Geschlecht und Alter wurde vorgenommen. Es wurden Altersklassen zwischen 18 und 80 Jahren gebildet und in einigen Fällen erfolgte eine saisonale oder gerätebezogene Stratifizierung.

Die berechneten Referenzgrenzen wurden einem Vergleich mit etablierten Referenzgrenzen unterzogen. Dabei zeigte sich eine hohe Übereinstimmung, gerade für Calcium. Bei Phosphat ergab sich eine niedrigere untere Referenzgrenze. Für die Schilddrüsenhormone ergaben sich leichte Abweichungen zu den etablierten Referenzgrenzen, aber trotzdem ähnliche Grenzwerte gerade im Hinblick auf den breiten physiologischen Schwankungsbereich dieser Laborparameter. Für Parathormon zeigte sich eine deutlich höhere obere Referenzgrenze, besonders für die Männer. Insgesamt ergab sich ein sehr breiter Referenzbereich, der nicht plausibel erscheint. Referenzgrenzen für Vitamin D konnten nicht ermittelt werden, da die Daten keiner Normalverteilung folgten, ebensowenig für Renin und Aldosteron aufgrund des zu geringen Datenumfangs.

Insgesamt handelt es sich beim RLE um eine sinnvolle Vorgehensweise, um klinisch relevante Referenzgrenzen besonders für Routineparameter mit Normalverteilung auf unkompliziertem Wege zu erhalten. Es bietet sich somit eine vorteilhafte Alternative zur konventionellen Bestimmung von Referenzgrenzen. Es erübrigen sich die finanziellen, zeitlichen und arbeitsaufwändigen Schwierigkeiten, da das Programm in die Laborsoftware integriert werden kann und auf Basis der Patientendaten Berechnungen vornimmt. Gerade für Routineparameter ist es sehr gut geeignet. Allerdings sind für schwierigere Laborparameter, die keine Normalverteilung aufweisen oder deren Datensatz zu viel pathologische Werte enthält, weitere Schritte nötig. Das Vorgehen könnte zukünftig ausgebaut werden durch die Kombination mit klinischen Informationen. Dabei wäre besonders der Vermerk der Medikation, bestehende Krankheiten und Mangelzustände von Interesse.

## **6. Literatur**

1. Sunderman FW. Current concepts of "normal values," "reference values," and "discrimination values," in clinical chemistry. *Clin Chem* 1975; 21(13):1873–7.
2. Gräsbeck R. Reference values: Relevant controls for clinical decision-making. *European Journal of Haematology* 1988; 40(1):1–6. doi: 10.1111/j.1600-0609.1988.tb00788.x.
3. Clinical and Laboratory Standards Institute. Defining, establishing, and verifying reference intervals in the clinical laboratory: Approved guideline [CLSI document EP28-A3c]. 3. ed. Wayne, PA; 2008. (Documents).
4. Leitner-Ferenc V, Atamaniuk J, Jansen-Skoupy S, Stöckelmeier B, Grohs K, Födinger M. CLSI-Based Validation of Manufacturer-Derived Reference Intervals on the Cobas 8000 Platform. *Lab Med* 2017; 48(2):e30-e35. doi: 10.1093/labmed/lmx020.
5. Neumann G. Determination of normal ranges from routine laboratory data; 1968.
6. Bock BJ, Dolan CT, Miller GC, Fitter WF, Hartsell BD, Crowson AN et al. The Data Warehouse as a Foundation for Population-Based Reference Intervals. *Am J Clin Pathol* 2003; 120(5):662–70. doi: 10.1309/W8J85AG4WDG6JGJ9.
7. Concorde D, Geffré A, Braun JP, Trumel C. A new approach for the determination of reference intervals from hospital-based data. *Clin Chim Acta* 2009; 405(1-2):43–8. doi: 10.1016/j.cca.2009.03.057.
8. Solberg HE. Using a hospitalized population to establish reference intervals: pros and cons. *Clin Chem* 1994; 40(12):2205–6.
9. Solberg HE. International Federation of Clinical Chemistry (IFCC), Scientific Committee, Clinical Section, Expert Panel on Theory of Reference Values, and International Committee for Standardization in Haematology (ICSH), Standing Committee on Reference Values. Approved Recommendation (1986) on the theory of reference values. Part 1. The concept of reference values. *J Clin Chem Clin Biochem* 1987; 25(5):337–42.
10. Solberg HE. Approved recommendation (1987) on the theory of reference values. Part 5. Statistical treatment of collected reference values. Determination of reference limits. *Clinica Chimica Acta* 1987; 170(2-3):S13-S32. doi: 10.1016/0009-8981(87)90151-3.
11. PetitClerc C, Solberg HE. Approved recommendation (1987) on the theory of reference values. Part 2. Selection of individuals for the production of reference values. *Clinica Chimica Acta* 1987; 170(2-3):S1-S11. doi: 10.1016/0009-8981(87)90150-1.

12. Solberg HE, PetitClerc C. Approved recommendation (1988) on the theory of reference values. Part 3. Preparation of individuals and collection of specimens for the production of reference values. *Clinica Chimica Acta* 1988; 177(3):S3-S11. doi: 10.1016/0009-8981(88)90074-5.
13. Solberg HE, Stamm D. IFCC recommendation: The theory of reference values. Part 4. Control of analytical variation in the production, transfer and application of reference values. *J Automat Chem* 1991; 13(5):231–4. doi: 10.1155/S146392469100038X.
14. Hoffmann G, Lichtinghagen R, Wosniok W. Simple estimation of reference intervals from routine laboratory data. *LaboratoriumsMedizin* 2016; 39(6):1033. doi: 10.1515/labmed-2015-0104.
15. Haeckel R, Wosniok W, Arzideh F. A plea for intra-laboratory reference limits. Part 1. General considerations and concepts for determination. *Clin Chem Lab Med* 2007; 45(8):1033–42. doi: 10.1515/CCLM.2007.249.
16. Siest G, Henny J, Grasbeck R, Wilding P, Petitclerc C, Queralto JM et al. The theory of reference values: An unfinished symphony. *Clin Chem Lab Med* 2013; 51(1):47–64. doi: 10.1515/cclm-2012-0682.
17. Koerbin G, Sikaris KA, Jones GRD, Ryan J, Reed M, Tate J. Evidence-based approach to harmonised reference intervals. *Clin Chim Acta* 2014; 432:99–107. doi: 10.1016/j.cca.2013.10.021.
18. Jones GRD, Haeckel R, Loh TP, Sikaris K, Streichert T, Katayev A et al. Indirect methods for reference interval determination - review and recommendations. *Clin Chem Lab Med* 2018; 57(1):20–9. doi: 10.1515/cclm-2018-0073.
19. Hoffmann RG. Statistics in the Practice of Medicine. *JAMA* 1963; 185(11):864. doi: 10.1001/jama.1963.03060110068020.
20. Gindler EM. Calculation of normal ranges by methods used for resolution of overlapping Gaussian distributions. *Clin Chem* 1970; 16(2):124–8.
21. Bhattacharya CG. A Simple Method of Resolution of a Distribution into Gaussian Components. *Biometrics* 1967; 23(1):115. doi: 10.2307/2528285.
22. Farhad Arzideh. Estimation of Medical Reference Limits by Truncated Gaussian and Truncated Power Normal Distributions.

23. Arzideh F, Wosniok W, Gurr E, Hinsch W, Schumann G, Weinstock N et al. A plea for intra-laboratory reference limits. Part 2. A bimodal retrospective concept for determining reference limits from intra-laboratory databases demonstrated by catalytic activity concentrations of enzymes. *Clin Chem Lab Med* 2007; 45(8):1043–57. doi: 10.1515/CCLM.2007.250.
24. Hofmann W, Aufenanger J, Hoffmann G. *Klinikhandbuch Labordiagnostische Pfade: Einführung ; Screening ; Stufendiagnostik.* 1. Aufl. s.l.: Walter de Gruyter GmbH Co.KG; 2012. Verfügbar unter: <http://www.degruyter.com/doi/book/10.1515/9783110228731>.
25. Thomas L. *Clinical Laboratory Diagnostics /Labor und Diagnose.* Elektronische Auflage. Frankfurt am Main: Th-Books; 2016.
26. Dörner K. *Klinische Chemie und Hämatologie.* 8., überarbeitete Auflage. Stuttgart, New York: Georg Thieme Verlag; 2013. (Taschenlehrbuch).
27. Schuirmann DJ. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *J Pharmacokinet Biopharm* 1987; 15(6):657–80.
28. Lakens D. Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-Analyses. *Soc Psychol Personal Sci* 2017; 8(4):355–62. doi: 10.1177/1948550617697177.
29. Zierk J, Arzideh F, Haeckel R, Rascher W, Rauh M, Metzler M. Indirect determination of pediatric blood count reference intervals. *Clin Chem Lab Med* 2013; 51(4):863–72. doi: 10.1515/cclm-2012-0684.
30. Arzideh F, Wosniok W, Haeckel R. Reference limits of plasma and serum creatinine concentrations from intra-laboratory data bases of several German and Italian medical centres: Comparison between direct and indirect procedures. *Clin Chim Acta* 2010; 411(3-4):215–21. doi: 10.1016/j.cca.2009.11.006.
31. G. E. P. Box, D. R. Cox. An analysis of transformations; 1964.
32. Dr. Farhad Arzideh. Reference Limit Estimator: Programm zur Abschätzung von Referenzgrenzen aus den Analysendaten klinischchemischer Laboratorien [Software Version 20151017] Oktober 2015.
33. Haeckel R, Wosniok W, Arzideh F. Equivalence limits of reference intervals for partitioning of population data. Relevant differences of reference limits. *LaboratoriumsMedizin* 2016; 40(3):891. doi: 10.1515/labmed-2016-0002.

34. Haeckel R, Gurr E, Hoff T, Guid obotwg. Bias, its minimization or circumvention to simplify internal quality assurance. *LaboratoriumsMedizin* 2016; 40(4):390. doi: 10.1515/labmed-2016-0036.
35. Arzideh F, Wosniok W, Haeckel R. Indirect reference intervals of plasma and serum thyrotropin (TSH) concentrations from intra-laboratory data bases from several German and Italian medical centres. *Clin Chem Lab Med* 2011; 49(4):659–64. doi: 10.1515/CCLM.2011.114.
36. Zierk J, Arzideh F, Rechenauer T, Haeckel R, Rascher W, Metzler M et al. Age- and sex-specific dynamics in 22 hematologic and biochemical analytes from birth to adolescence. *Clin Chem* 2015; 61(7):964–73. doi: 10.1373/clinchem.2015.239731.
37. Zierk J, Arzideh F, Haeckel R, Rauh M, Metzler M, Ganslandt T et al. Indirect determination of hematology reference intervals in adult patients on Beckman Coulter UniCell DxH 800 and Abbott CELL-DYN Sapphire devices. *Clin Chem Lab Med* 2018. doi: 10.1515/cclm-2018-0771.
38. Henny J, Vassault A, Boursier G, Vukasovic I, Mesko Brguljan P, Lohmander M et al. Recommendation for the review of biological reference intervals in medical laboratories. *Clin Chem Lab Med* 2016; 54(12):1893–900. doi: 10.1515/cclm-2016-0793.
39. Henny J, PetitClerc C, Fuentes-Arderiu X, Petersen PH, Queraltó JM, Schiele F et al. Need for revisiting the concept of reference values. *Clin Chem Lab Med* 2000; 38(7):589–95. doi: 10.1515/CCLM.2000.085.
40. Kohse KP. National and international initiatives and approaches for the establishment of reference intervals in pediatric laboratory medicine. *LaboratoriumsMedizin* 2015; 39(4):808. doi: 10.1515/labmed-2015-0055.
41. Holick MF. Vitamin D deficiency. *N Engl J Med* 2007; 357(3):266–81. doi: 10.1056/NEJMra070553.
42. Squier TC, Bigelow DJ. Protein oxidation and age-dependent alterations in calcium homeostasis. *Front Biosci* 2000; 5:D504-26.
43. Veldurthy V, Wei R, Oz L, Dhawan P, Jeon YH, Christakos S. Vitamin D, calcium homeostasis and aging. *Bone Res* 2016; 4:16041. doi: 10.1038/boneres.2016.41.
44. Malberti F. Hyperphosphataemia: Treatment options. *Drugs* 2013; 73(7):673–88. doi: 10.1007/s40265-013-0054-y.

45. National Kidney Foundation. K/DOQI clinical practice guidelines for bone metabolism and disease in chronic kidney disease. *Am J Kidney Dis* 2003; 42(4 Suppl 3):S1-201.
46. Amanzadeh J, Reilly RF. Hypophosphatemia: An evidence-based approach to its clinical consequences and management. *Nat Clin Pract Nephrol* 2006; 2(3):136–48. doi: 10.1038/ncpneph0124.
47. Clerico A, Trenti T, Aloe R, Dittadi R, Rizzardi S, Migliardi M et al. A multicenter study for the evaluation of the reference interval for TSH in Italy (ELAS TSH Italian Study). *Clin Chem Lab Med* 2018; 57(2):259–67. doi: 10.1515/cclm-2018-0541.
48. Völzke H, Alte D, Kohlmann T, Lüdemann J, Nauck M, John U et al. Reference intervals of serum thyroid function tests in a previously iodine-deficient area. *Thyroid* 2005; 15(3):279–85. doi: 10.1089/thy.2005.15.279.
49. Kratzsch J, Fiedler GM, Leichtle A, Brügel M, Buchbinder S, Otto L et al. New reference intervals for thyrotropin and thyroid hormones based on National Academy of Clinical Biochemistry criteria and regular ultrasonography of the thyroid. *Clin Chem* 2005; 51(8):1480–6. doi: 10.1373/clinchem.2004.047399.
50. Wang D, Cheng X, Yu S, Qiu L, Lian X, Guo X et al. Data mining: Seasonal and temperature fluctuations in thyroid-stimulating hormone. *Clin Biochem* 2018; 60:59–63. doi: 10.1016/j.clinbiochem.2018.08.008.
51. Kravets I. Hyperthyroidism: Diagnosis and Treatment. *Am Fam Physician* 2016; 93(5):363–70.
52. Biondi B. Thyroid and obesity: an intriguing relationship. *J Clin Endocrinol Metab* 2010; 95(8):3614–7. doi: 10.1210/jc.2010-1245.
53. Bjoro T, Holmen J, Krüger O, Midthjell K, Hunstad K, Schreiner T et al. Prevalence of thyroid disease, thyroid dysfunction and thyroid peroxidase antibodies in a large, unselected population. The Health Study of Nord-Trøndelag (HUNT). *Eur J Endocrinol* 2000; 143(5):639–47.
54. Farrell C-JL, Nguyen L, Carter AC. Parathyroid hormone: Data mining for age-related reference intervals in adults. *Clin Endocrinol (Oxf)* 2018; 88(2):311–7. doi: 10.1111/cen.13486.

55. Björkman M, Sorva A, Tilvis R. Responses of parathyroid hormone to vitamin D supplementation: a systematic review of clinical trials. *Arch Gerontol Geriatr* 2009; 48(2):160–6. doi: 10.1016/j.archger.2007.12.005.
56. Patel S, Hyer S, Barron J. Glomerular filtration rate is a major determinant of the relationship between 25-hydroxyvitamin D and parathyroid hormone. *Calcif Tissue Int* 2007; 80(4):221–6. doi: 10.1007/s00223-007-9001-9.
57. Souberbielle J-C, Brazier F, Piketty M-L, Cormier C, Minisola S, Cavalier E. How the reference values for serum parathyroid hormone concentration are (or should be) established? *J Endocrinol Invest* 2017; 40(3):241–56. doi: 10.1007/s40618-016-0553-2.
58. Hintz Peter B, Mensink GBM, Thierfelder W, Müller MJ, Scheidt-Nave C. Vitamin D status and health correlates among German adults. *Eur J Clin Nutr* 2008; 62(9):1079–89. doi: 10.1038/sj.ejcn.1602825.
59. Daniel Sowah, Xiangning Fan, Liz Dennett, Reidar Hagtvedt, Sebastian Straube. Vitamin D levels and deficiency with different occupations: a systematic review.
60. Katrinaki M, Kampa M, Margioris A, Castanas E, Malliaraki N. Vitamin D levels in a large Mediterranean cohort: reconsidering normal cut-off values. *Hormones (Athens)* 2016; 15(2):205–23. doi: 10.14310/horm.2002.1674.
61. Ferrari D, Lombardi G, Banfi G. Concerning the vitamin D reference range: pre-analytical and analytical variability of vitamin D measurement. *Biochem Med (Zagreb)* 2017; 27(3):30501. doi: 10.11613/BM.2017.030501.
62. Holick MF, Binkley NC, Bischoff-Ferrari HA, Gordon CM, Hanley DA, Heaney RP et al. Evaluation, treatment, and prevention of vitamin D deficiency: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab* 2011; 96(7):1911–30. doi: 10.1210/jc.2011-0385.
63. Wortsman J, Matsuoka LY, Chen TC, Lu Z, Holick MF. Decreased bioavailability of vitamin D in obesity. *Am J Clin Nutr* 2000; 72(3):690–3. Verfügbare unter: <https://academic.oup.com/ajcn/article-pdf/72/3/690/23960661/690.pdf>.

## 7. Anhang

Tabelle 20: Gerätetestatistik für Calcium.  
 N = Anzahl, Ca = Calcium, K = Kiel,  
 L = Lübeck, M = Modul.

N (Ca)	Gerätename
116	Cobas 8000 K1 (M1)
28	Cobas 8000 K1 (M2)
167408	Cobas 8000 K1 (M3)
25460	Cobas 8000 K2 (M1)
171251	Cobas 8000 K2 (M2)
403	Cobas 8100 L
145	Cobas 8000 L2 (M1)
114950	Cobas 8000 L2 (M2)
2	Cobas 8000 L2 (M4)
190023	Cobas 8000 L1 (M1)
210	Cobas 8000 L1 (M2)
1	Cobas 8000 L1 (M3)
4	Cobas 8000 L1 (M4)

Tabelle 22: Gerätetestatistik für TSH.  
 N = Anzahl, K = Kiel, L = Lübeck, M = Modul.

N (TSH)	Gerätename
38	Cobas 8000 K1 (M1)
37	Cobas 8000 K1 (M2)
70640	Cobas 8000 K1 (M4)
41591	Cobas 8000 K2 (M3)
44690	Cobas 8000 K2 (M4)
267	Cobas 8100 L
191	Cobas 8000 L2 (M1)
37551	Cobas 8000 L2 (M5)
18483	Cobas 8000 L2 (M4)
103535	Cobas 8000 L1 (M3)

Tabelle 21: Gerätetestatistik für Phosphat.  
 N = Anzahl, Ph = Phosphat, K = Kiel,  
 L = Lübeck, M = Modul.

N (Ph)	Gerätename
37	Cobas 8000 K1 (M1)
55	Cobas 8000 K1 (M2)
62759	Cobas 8000 K1 (M3)
11774	Cobas 8000 K2 (M1)
72773	Cobas 8000 K2 (M2)
179	Cobas 8100 L
82	Cobas 8000 L2 (M1)
40147	Cobas 8000 L2 (M3)
2	Cobas 8000 L2 (M4)
105586	Cobas 8000 L1 (M2)
125	Cobas 8000 L1 (M4)
2766	Dimension

Tabelle 23: Gerätetestatistik für fT3.  
 N = Anzahl, K = Kiel, L = Lübeck, M = Modul.

N (fT3)	Gerätename
16	Cobas 8000 K1 (M1)
10	Cobas 8000 K1 (M2)
31661	Cobas 8000 K1 (M4)
72	Cobas 8000 K1 (M5)
21321	Cobas 8000 K2 (M3)
20614	Cobas 8000 K2 (M4)
103	Cobas 8100 L
250	Cobas 8000 L2 (M1)
282	Cobas 8000 L2 (M5)
4472	Cobas 8000 L2 (M6)
39502	Cobas 8000 L1 (M4)

Tabelle 24: Gerätetestatistik für fT4.  
N = Anzahl, K = Kiel, L = Lübeck, M = Modul.

<b>N (fT4)</b>	<b>Gerätename</b>
17	Cobas 8000 K1 (M1)
5	Cobas 8000 K1 (M2)
35	Cobas 8000 K1 (M4)
32083	Cobas 8000 K1 (M5)
20981	Cobas 8000 K2 (M3)
21716	Cobas 8000 K2 (M4)
110	Cobas 8100 L
30	Cobas 8000 L2 (M1)
4759	Cobas 8000 L2 (M5)
302	Cobas 8000 L2 (M6)
240	Cobas 8000 L2 (M4)
41436	Cobas 8000 L1 (M3)

Tabelle 25: Gerätetestatistik für Parathormon.  
N = Anzahl, PTH = Parathormon, K = Kiel, L = Lübeck, M = Modul.

<b>N (PTH)</b>	<b>Gerätename</b>
14	Cobas 8000 K1 (M1)
59	Cobas 8000 K1 (M4)
14691	Cobas 8000 K2 (M4)
115	Cobas 8100 L
17	Cobas 8000 L2 (M1)
97	Cobas 8000 L2 (M5)
9	Cobas 8000 L2 (M6)
118	Cobas 8000 L2 (M4)
10073	Cobas 8000 L1 (M4)

Tabelle 26: Gerätetestatistik für Renin. N = Anzahl.

<b>N (Renin)</b>	<b>Gerätename</b>
3205	BEP2000 Lübeck

Tabelle 27: Gerätetestatistik für Aldosteron. N = Anzahl, Aldo. = Aldosteron.

<b>N (Aldo.)</b>	<b>Gerätename</b>
2799	BEP2000 Lübeck

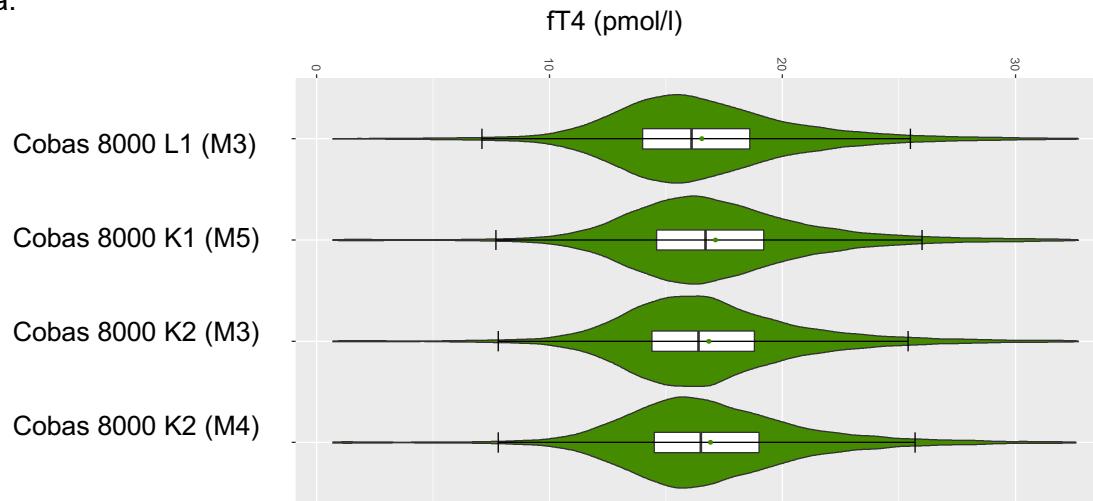
Tabelle 28: Gerätetestatistik für Vitamin D. N = Anzahl, Vit. D = Vitamin D.

<b>N (Vit. D)</b>	<b>Gerätename</b>
27980	HPLC Lübeck

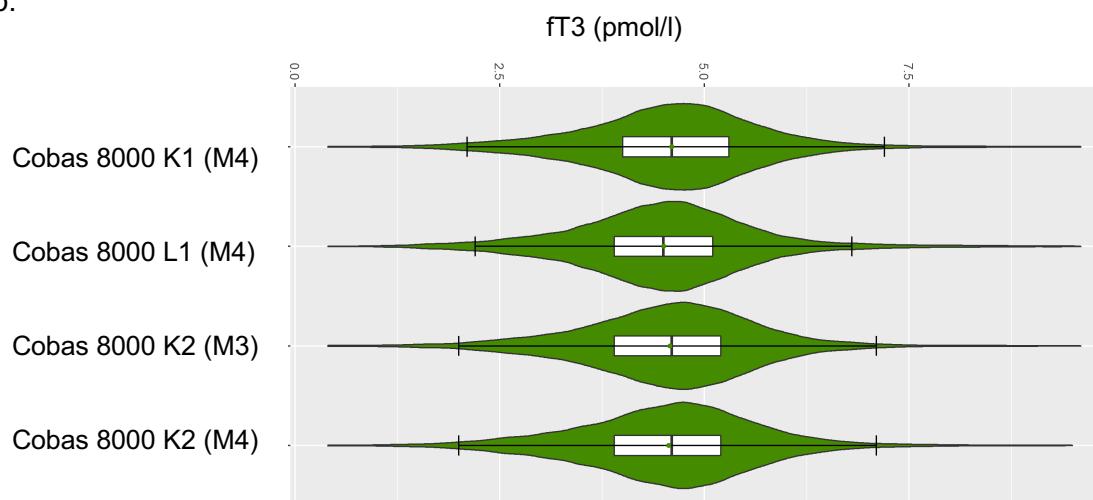
Tabelle 29: An häufigsten eingesetzte Messgeräte pro Parameter. N\* = Anzahl der Messungen der häufigsten Geräte. N\*\* = Gesamtzahl der Messungen. Ca = Calcium, Ph = Phosphat, PTH= Parathormon, Vit. D = Vitamin D, Aldo. = Aldosteron, K = Kiel, L = Lübeck, M = Modul.

Parameter	N* (N*/N**)					
<b>Ca</b>	Cobas 8000 K1 (M3)	Cobas 8000 K2 (M1)	Cobas 8000 K2 (M2)	Cobas 8000 L1 (M1)	Cobas 8000 L2 (M2)	669.092 (99,86%)
<b>Ph</b>	Cobas 8000 K1 (M3)	Cobas 8000 K2 (M1)	Cobas 8000 K2 (M2)	Cobas 8000 L1 (M2)	Cobas 8000 L2 (M3)	293.039 (98,90%)
<b>PTH</b>	Cobas 8000 K2 <del>A444</del>	Cobas 8000 L1 <del>A444</del>				24.764 (98,30%)
<b>Vit. D</b>	HPLC L					27.963 (99,94%)
<b>TSH</b>	Cobas 8000 K1 (M4)	Cobas 8000 K2 (M3)	Cobas 8000 K2 (M4)	Cobas 8000 L2 (M5)	Cobas 8000 L2 (M4)	316.490 (99,83%)
<b>fT3</b>	Cobas 8000 K1 (M4)	Cobas 8000 K2 (M3)	Cobas 8000 K2 (M4)	Cobas 8000 L2 (M6)	Cobas 8000 L1 (M4)	117.570 (99,38%)
<b>fT4</b>	Cobas 8000 K1 (M5)	Cobas 8000 K2 (M3)	Cobas 8000 K2 (M4)	Cobas 8000 L2 (M5)	Cobas 8000 L1 (M3)	120.975 (99,39%)
<b>Renin</b>	BEP2000 L					3.193 (99,63%)
<b>Aldo.</b>	BEP2000 L					2.798 (99,96%)

a.



b.



c.

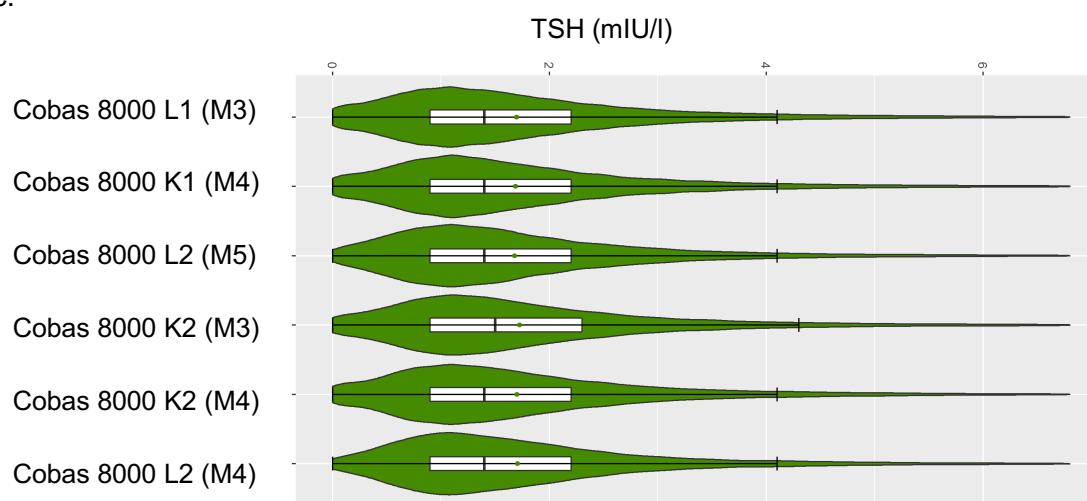


Abbildung 29: a.: Äquivalenztest der häufigsten Messgeräte für fT4. K = Kiel, L = Lübeck, M = Modul.

b.: Äquivalenztest der häufigsten Messgeräte für fT3. K = Kiel, L = Lübeck, M = Modul.

c.: Äquivalenztest der häufigsten Messgeräte für TSH. K = Kiel, L = Lübeck, M = Modul.

Tabelle 30: Äquivalenztest der Geräte für Calcium mit einem Epsilon-Wert von 10%. KI = Konfidenzintervall, äq.= äquivalent, n. äq.=nicht äquivalent, K = Kiel, L = Lübeck, M = Modul.

<b>Gerät 1</b>	<b>Gerät 2</b>	<b>KI</b>	<b>Ergebnis</b>
Cobas 8000 K2 (M2)	Cobas 8000 K1 (M3)	[(- 0,006) – (- 0,004)]	äq.
Cobas 8000 K2 (M2)	Cobas 8000 L1 (M1)	[0,008 – 0,01]	äq.
Cobas 8000 K2 (M2)	Cobas 8000 K2 (M1)	[0,007 – 0,011]	äq.
Cobas 8000 K2 (M2)	Cobas 8000 L2 (M2)	[0,007 – 0,009]	äq.
Cobas 8000 K1 (M3)	Cobas 8000 L1 (M1)	[0,012 – 0,014]	äq.
Cobas 8000 K1 (M3)	Cobas 8000 K2 (M1)	[0,012 – 0,016]	äq.
Cobas 8000 K1 (M3)	Cobas 8000 L2 (M2)	[0,012 – 0,014]	äq.
Cobas 8000 L1 (M1)	Cobas 8000 K2 (M1)	[(- 0,001) – 0,003]	äq.
Cobas 8000 L1 (M1)	Cobas 8000 L2 (M2)	[(- 0,002) – 0,001]	äq.
Cobas 8000 K2 (M1)	Cobas 8000 L2 (M2)	[(- 0,003) – 0,001]	äq.

Tabelle 31: Äquivalenztest der Geräte für TSH mit einem Epsilon-Wert von 10%. KI = Konfidenzintervall, äq.= äquivalent, n. äq.= nicht äquivalent, K = Kiel, L = Lübeck, M = Modul.

<b>Gerät 1</b>	<b>Gerät 2</b>	<b>KI</b>	<b>Ergebnis</b>
Cobas 8000 K2 (M4)	Cobas 8000 K1 (M4)	[(- 0,002) – 0,026]	äq.
Cobas 8000 K2 (M4)	Cobas 8000 K2 (M3)	[(- 0,04) – (- 0,08)]	äq.
Cobas 8000 K2 (M4)	Cobas 8000 L1 (M3)	[(- 0,009) – 0,017]	äq.
Cobas 8000 K2 (M4)	Cobas 8000 L2 (M4)	[(- 0,026) – 0,013]	äq.
Cobas 8000 K2 (M4)	Cobas 8000 L2 (M5)	[0,006 – 0,038]	äq.
Cobas 8000 K1 (M4)	Cobas 8000 K2 (M3)	[(- 0,05) – (- 0,022)]	äq.
Cobas 8000 K1 (M4)	Cobas 8000 L1 (M3)	[(- 0,019) – 0,003]	äq.
Cobas 8000 K1 (M4)	Cobas 8000 L2 (M4)	[(- 0,036) – 0]	äq.
Cobas 8000 K1 (M4)	Cobas 8000 L2 (M5)	[(- 0,004) – 0,024]	äq.
Cobas 8000 K2 (M3)	Cobas 8000 L1 (M3)	[0,015 – 0,041]	äq.
Cobas 8000 K2 (M3)	Cobas 8000 L2 (M4)	[(- 0,002) – 0,037]	äq.
Cobas 8000 K2 (M3)	Cobas 8000 L2 (M5)	[0,03 – 0,062]	äq.
Cobas 8000 L1 (M3)	Cobas 8000 L2 (M4)	[(- 0,028) – 0,007]	äq.
Cobas 8000 L1 (M3)	Cobas 8000 L2 (M5)	[0,005 – 0,031]	äq.
Cobas 8000 L2 (M4)	Cobas 8000 L2 (M5)	[0,009 – 0,048]	äq.

Tabelle 32: Äquivalenztest der Geräte für fT3 mit einem Epsilon-Wert von 10%. KI = Konfidenzintervall, äq.= äquivalent, n. äq.=nicht äquivalent, K = Kiel, L = Lübeck, M = Modul.

<b>Gerät 1</b>	<b>Gerät 2</b>	<b>KI</b>	<b>Ergebnis</b>
Cobas 8000 K1 (M4)	Cobas 8000 K2 (M3)	[(- 0,002) – 0,036]	äq.
Cobas 8000 K1 (M4)	Cobas 8000 K2 (M4)	[0,015 – 0,054]	äq.
Cobas 8000 K1 (M4)	Cobas 8000 L1 (M4)	[0,08 – 0,113]	äq.
Cobas 8000 K2 (M3)	Cobas 8000 K2 (M4)	[(- 0,004) – 0,039]	äq.
Cobas 8000 K2 (M3)	Cobas 8000 L1 (M4)	[0,061 – 0,098]	äq.
Cobas 8000 K2 (M4)	Cobas 8000 L1 (M4)	[0,043 – 0,081]	äq.

Tabelle 33: Äquivalenztest der Geräte für fT4 mit einem Epsilon-Wert von 10%. KI = Konfidenzintervall, äq.= äquivalent, n. äq.= nicht äquivalent, K = Kiel, L = Lübeck, M = Modul.

Gerät 1	Gerät 2	KI	Ergebnis
Cobas 8000 K1 (M5)	Cobas 8000 K2 (M3)	[0,217 – 0,358]	äq.
Cobas 8000 K1 (M5)	Cobas 8000 K2 (M4)	[0,156 – 0,295]	äq.
Cobas 8000 K1 (M5)	Cobas 8000 L1 (M3)	[0,527 – 0,642]	äq.
Cobas 8000 K2 (M3)	Cobas 8000 K2 (M4)	[(- 0,14) – 0,016]	äq.
Cobas 8000 K2 (M3)	Cobas 8000 L1 (M3)	[0,23 – 0,364]	äq.
Cobas 8000 K2 (M4)	Cobas 8000 L1 (M3)	[0,293 – 0,425]	äq.

Tabelle 34: Äquivalenztest der Geräte für Parathormon mit einem Epsilon-Wert von 10%. KI = Konfidenzintervall, äq.= äquivalent, n. äq.= nicht äquivalent, K = Kiel, L = Lübeck, M = Modul.

Gerät 1	Gerät 2	KI	Ergebnis
Cobas 8000 K2 (M4)	Cobas 8000 L1 (M4)	[(-10.494) – (-8.288)]	n. äq.

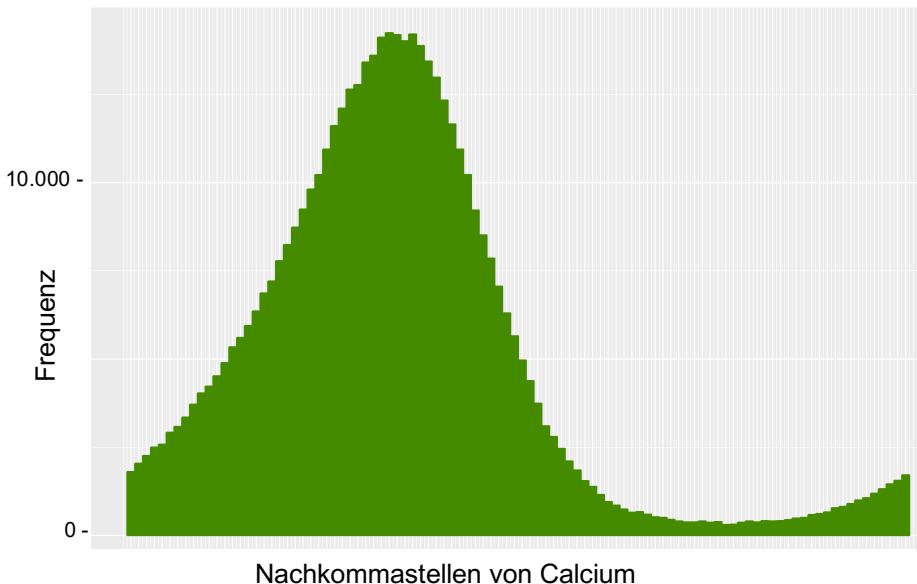
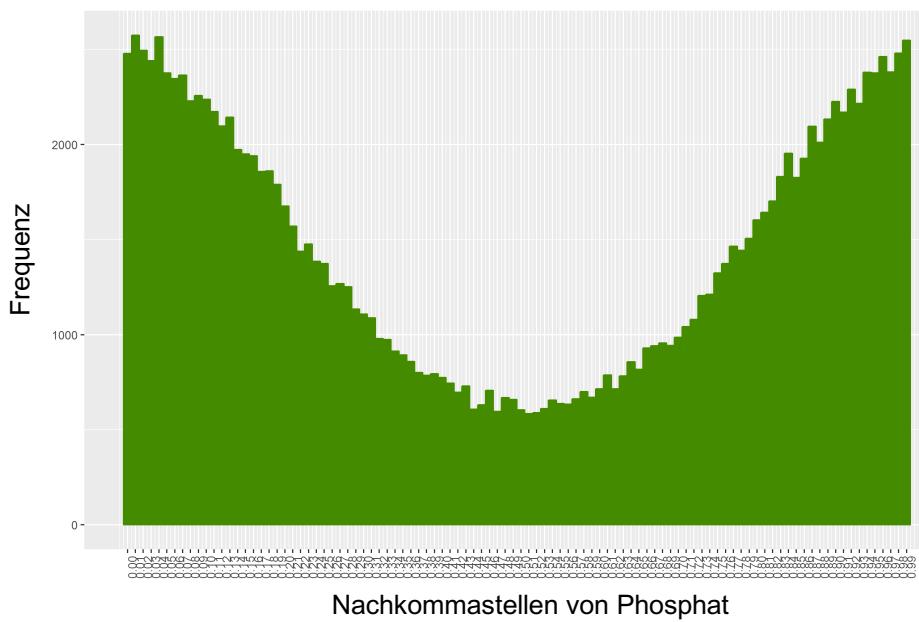


Abbildung 30: Histogramm der Nachkommastellen (0,00 bis 0,99) für die Messwerte von Calcium.

a.



b.

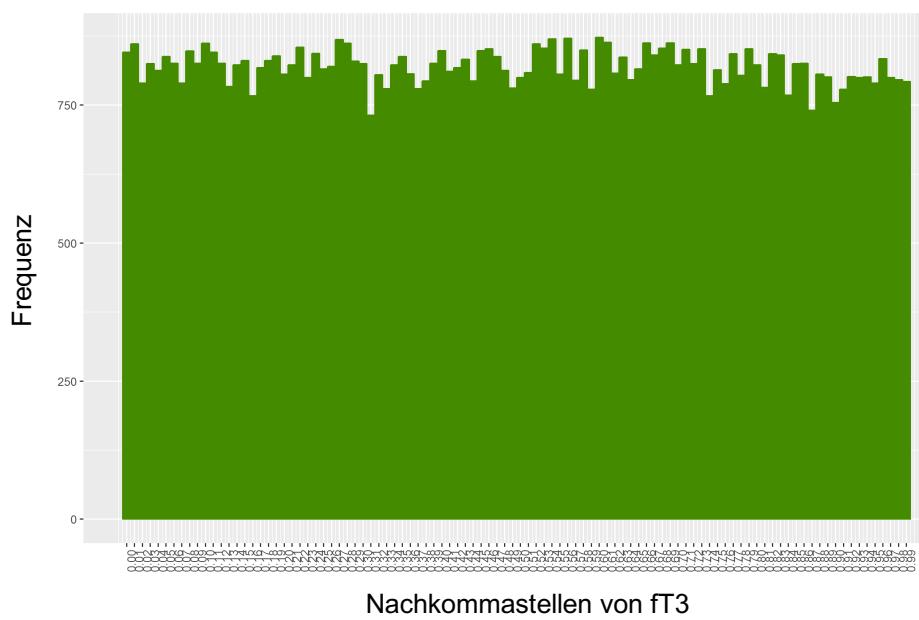


Abbildung 31: a.: Histogramm der Nachkommastellen (0,00-0,99) für die Messwerte von Phosphat.  
b.: Histogramm der Nachkommastellen (0,00-0,99) für die Messwerte von fT3.

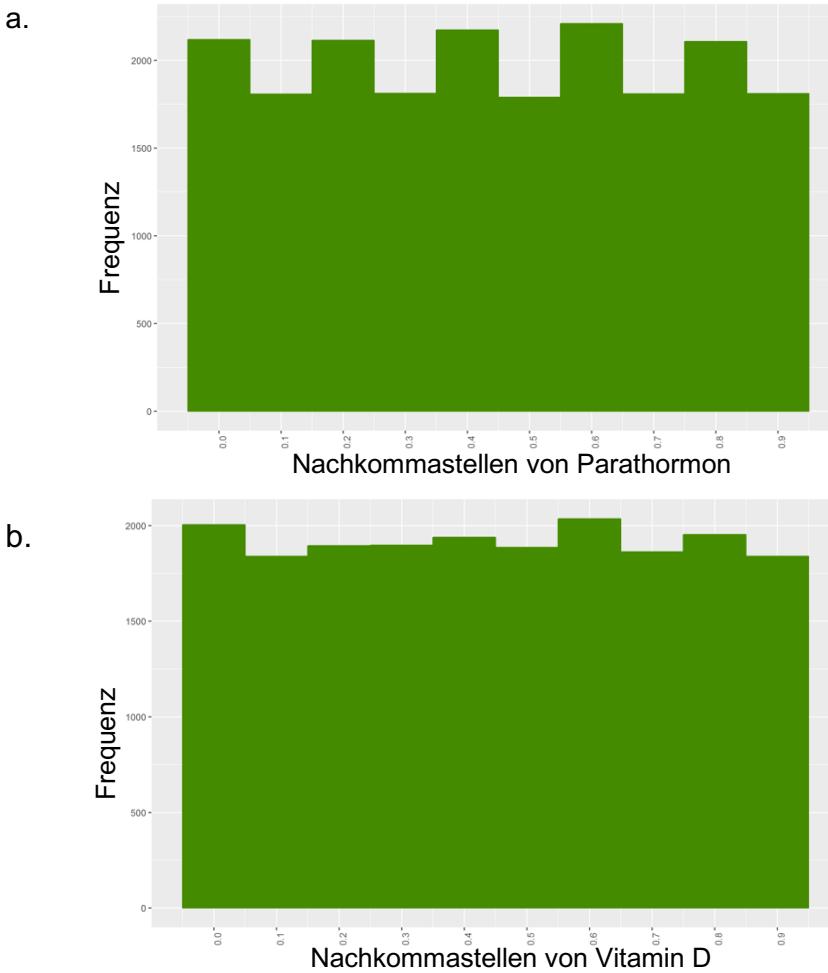


Abbildung 32: a.: Histogramm der Nachkommastellen für die Messwerte von Parathormon.  
b.: Histogramm der Nachkommastellen für die Messwerte von Vitamin D.

#### Estimation of permissible variation

The permissible analytical variation  $pCV_A$  and the empirical (biological variation)  $CV_E$  are derived as recently described [9].  $CV_E$  is only considered as a surrogate for the biological variation because it also includes the analytical variance. It is derived from the upper and lower reference limits. In the case of a normal distribution  $CV_E = (RL_2 - RL_1)/3.92$ . Because the distribution of most quantities is not known, a log-normal distribution may be assumed [18], and  $CV_E$  becomes  $CV_E^*$ .

Assuming log-normally distributed data, standard deviation and median on the logarithmic scale  $s_{E,ln}$  and  $Med_{ln}$  can be calculated by the following equation

$$\begin{aligned} s_{E,ln} &= (\ln RL_2 - \ln RL_1) / 3.92 \\ Med_{ln} &= (\ln RL_1 + \ln RL_2) / 2 \end{aligned} \quad (5)$$

On the logarithmic scale, mean and median are identical. The arithmetic mean and standard deviation on the linear scale are according to Ref. [19–21]:

$$\begin{aligned} Mean &= \exp(Med_{ln} + 0.5 s_{E,ln}^2) \\ s_{E,lin} &= Med_{ln} \cdot \{\exp(s_{E,ln}^2) - 1\}^{0.5} \end{aligned} \quad (6)$$

From these,  $CV_E^*$  valid on the linear scale can be calculated by equations 7 and 8

$$CV_E^* \cdot 100 \cdot s_{E,lin} / Mean = 100 \cdot Mean \cdot \{\exp(s_{E,ln}^2) - 1\}^{0.5} / Mean \quad (7)$$

$$CV_E^* = 100 \cdot \{\exp(s_{E,ln}^2) - 1\}^{0.5} \quad (8)$$

The permissible  $CV_A$  at the median on the linear scale [ $Med = \exp(Med_{ln}) = (RL_1 \cdot RL_2)^{0.5}$ ] can be obtained from  $CV_E^*$  by equation 9 as explained in Ref.[9]:

$$pCV_A = (CV_E^* - 0.25)^{0.5} \quad (9)$$

From this, the permissible analytical standard deviation at the median is

$$ps_{A,Med} = pCV_A \cdot 0.01 \cdot Med \quad (10)$$

The analytical standard deviation is assumed to increase linearly with the measured value  $x_i$  and can be calculated at any  $x_i$  by equations 11 and 12:

$$s_{A,Xi} = a \cdot x_i + b \quad (11)$$

The slope  $a$  in equation 11 is

$$a = (ps_{A,Med} - 0.2 \cdot ps_{A,Med}) / Med$$

and the intercept  $b$  is

$$b = 0.2 \cdot ps_{A,Med}$$

The permissible standard deviation at  $x_i$  is

$$ps_{A,Xi} = [(ps_{A,Med} - 0.2 \cdot ps_{A,Med}) / Med] \cdot x_i + 0.2 \cdot ps_{A,Med} \quad (12)$$

This can be expressed as coefficient of variation:

$$pCV_{A,Xi} = ps_{A,Xi} \cdot 100 / x_i \quad (13)$$

The intercept  $b$  in equation 11 is related to the detection limit. Because the detection limit is usually unknown for the matrix of human materials, 20% of the  $s_A$  at the median is applied here as a substitute.

Abbildung 33: Formeln zur Berechnung der zulässigen Differenz aus Haeckel et al (2016). (33)

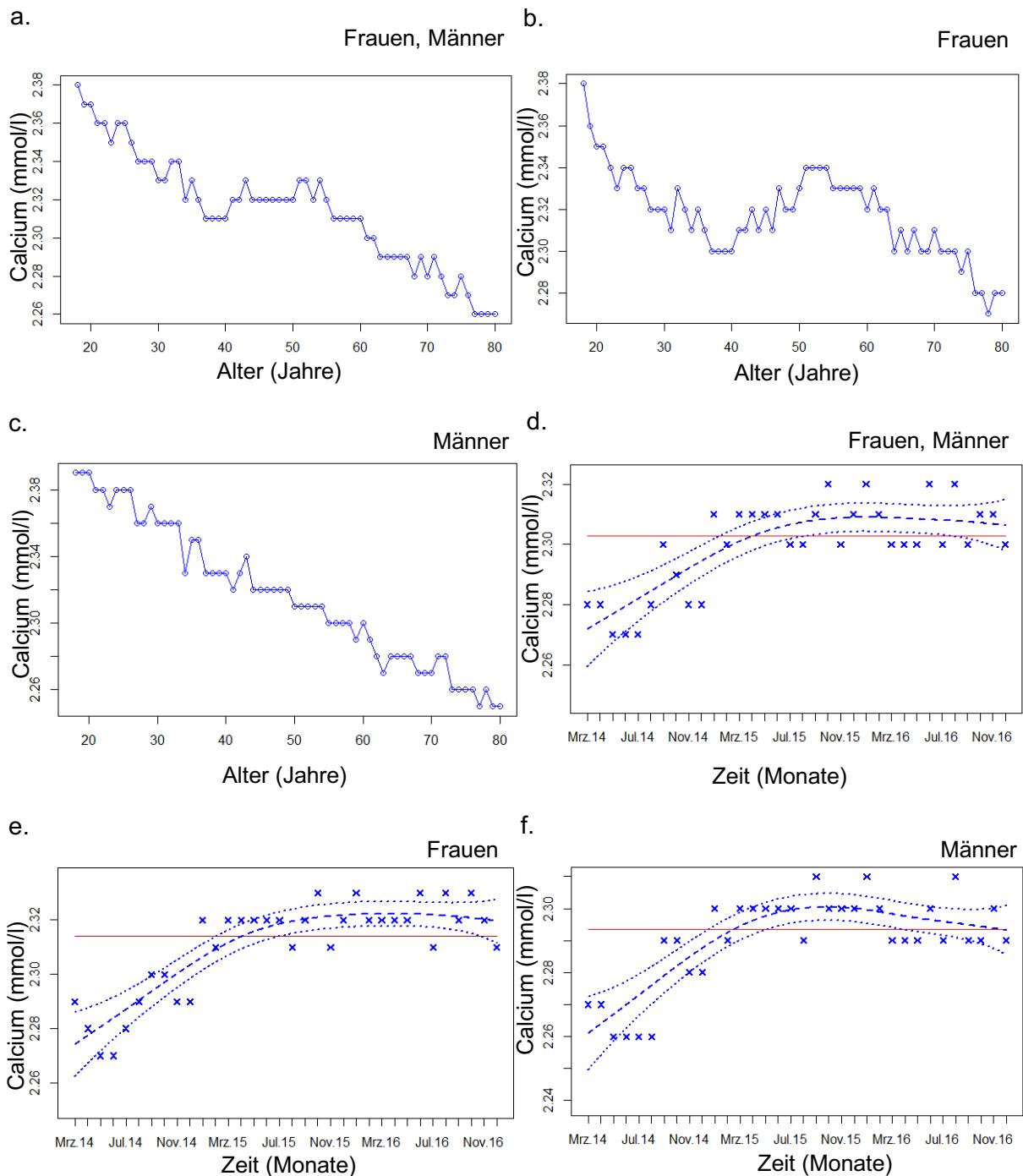


Abbildung 34: a.: Altersabhängiger Verlauf der 50.-Perzentile für Calcium für beide Geschlechter von 18-80 Jahren.

b.: Altersabhängiger Verlauf der 50.-Perzentile für Calcium für Frauen von 18-80 Jahren.

c.: Altersabhängiger Verlauf der 50.-Perzentile für Calcium für Männer von 18-80 Jahren.

d.: Zeitabhängiger Median für Calcium für beide Geschlechter von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).

e.: Zeitabhängiger Median für Calcium für Frauen von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).

f.: Zeitabhängiger Median für Calcium für Männer von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).

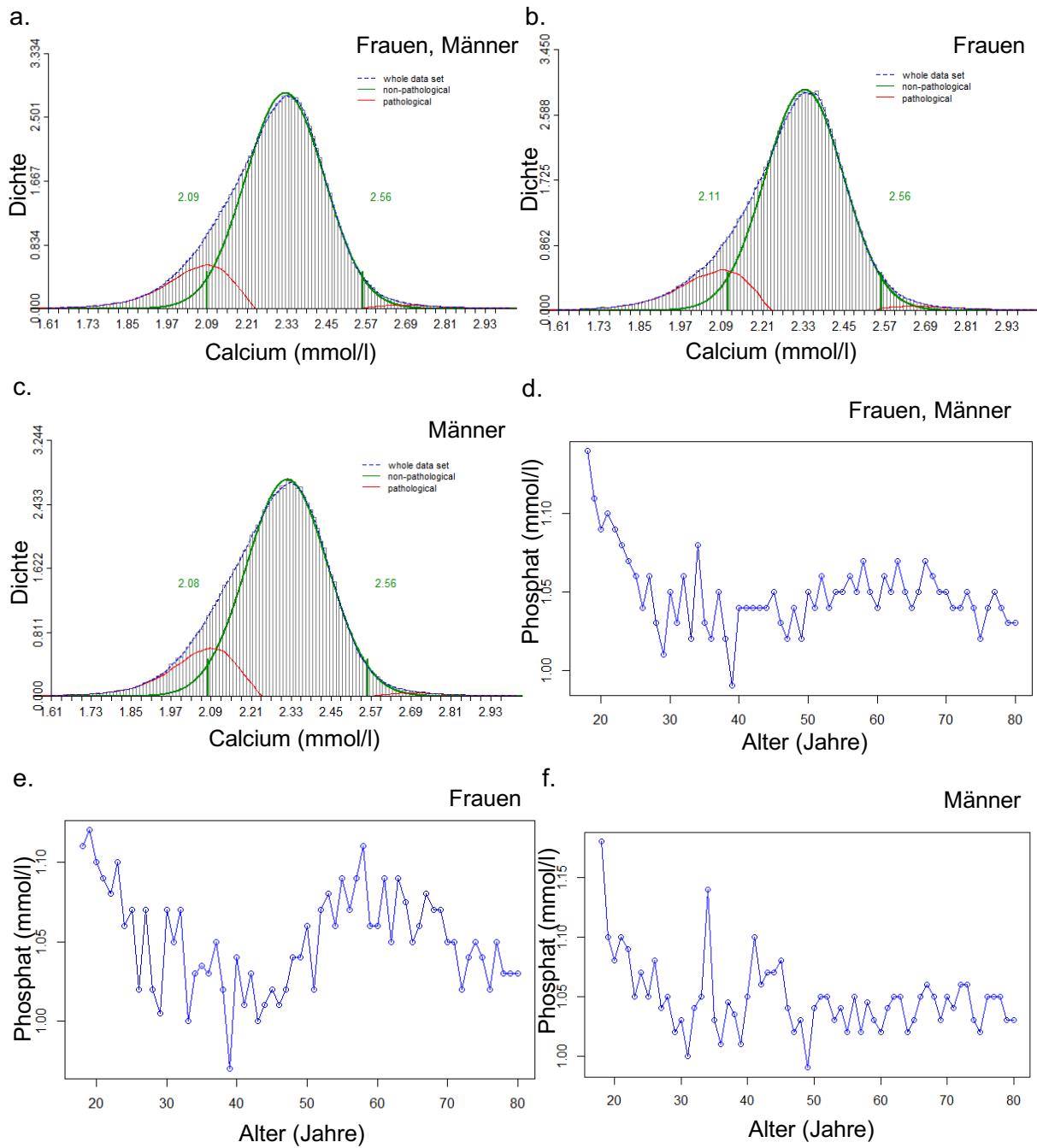


Abbildung 35: a.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Calcium für beide Geschlechter von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

b.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Calcium für Frauen von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

c.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Calcium für Männer von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

d.: Altersabhängiger Verlauf der 50.-Perzentile für Phosphat für beide Geschlechter von 18-80 Jahren.

e.: Altersabhängiger er Verlauf der 50.-Perzentile für Phosphat für Frauen von 18-80 Jahren.

f.: Altersabhängiger er Verlauf der 50.-Perzentile für Phosphat für Männer von 18-80 Jahren.

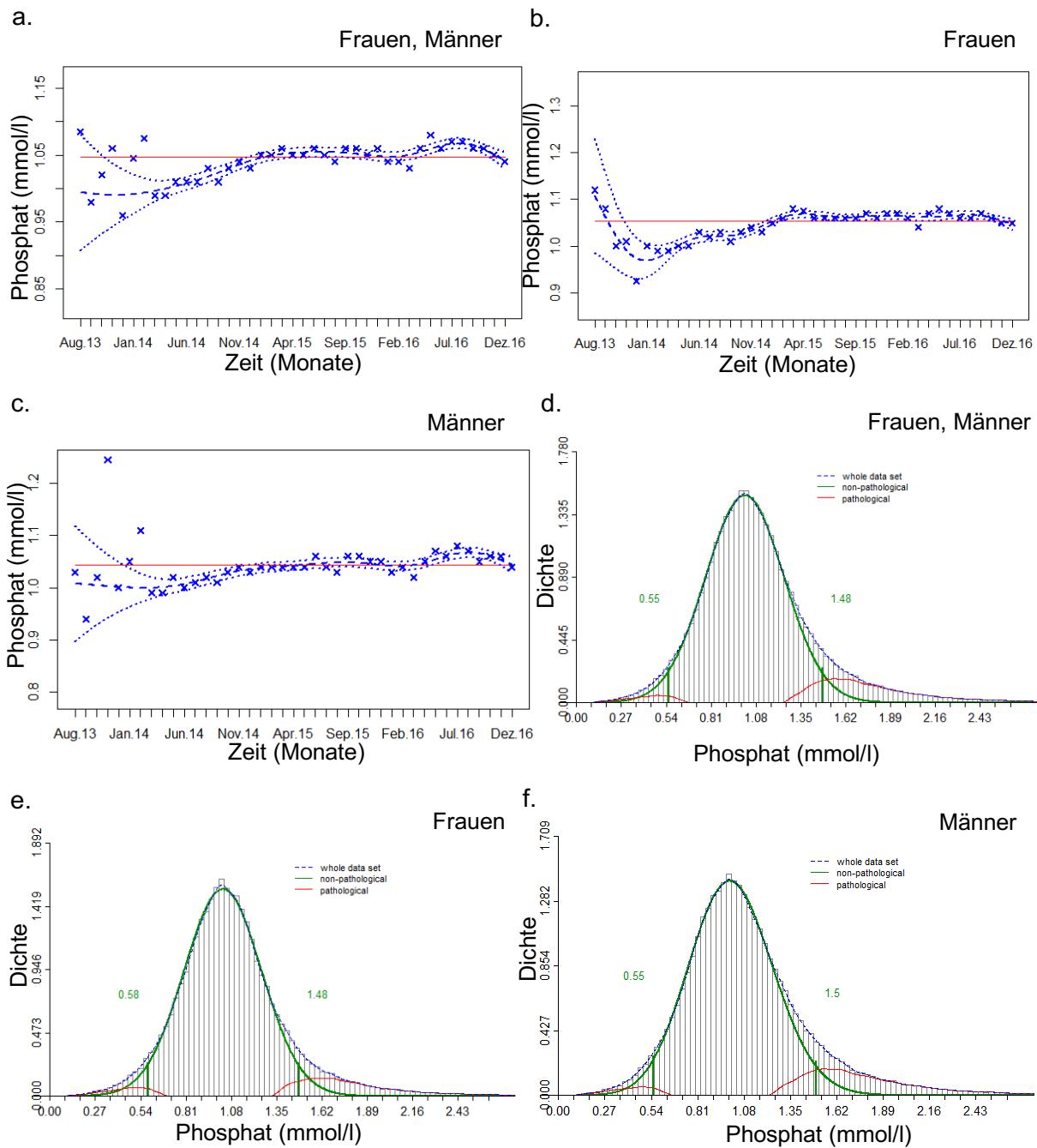


Abbildung 36: a.: Zeitabhängiger Median für Phosphat für beide Geschlechter von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).

b.: Zeitabhängiger Median für Phosphat für Frauen von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).

c.: Zeitabhängiger Median für Phosphat für Männer von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).

d.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Phosphat für beide Geschlechter von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

e.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Phosphat für Frauen von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

f.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Phosphat für Männer von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

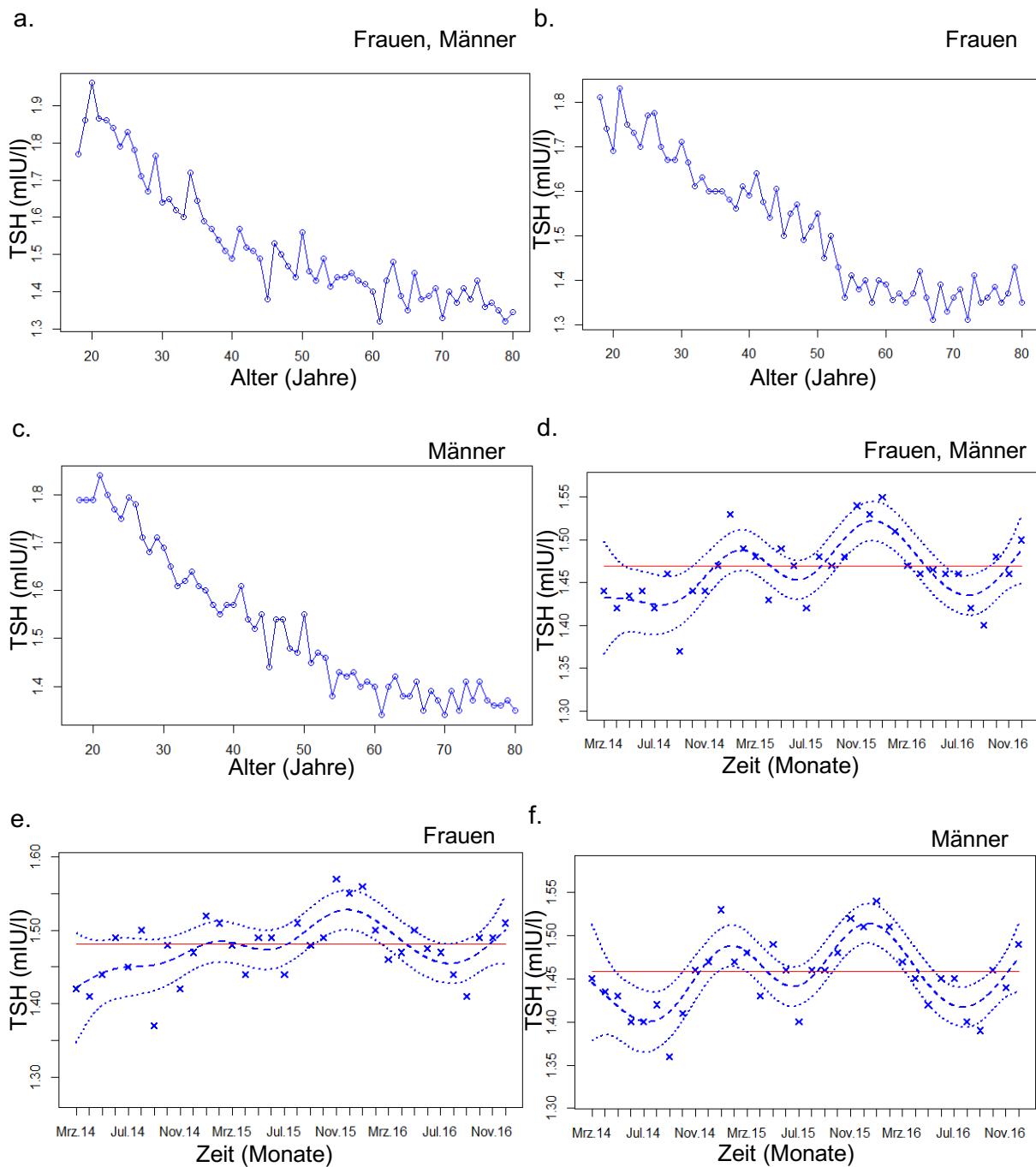


Abbildung 37: a.: Altersabhängiger Verlauf der 50.-Perzentile für TSH für beide Geschlechter von 18-80 Jahren.

b.: Altersabhängiger Verlauf der 50.-Perzentile für TSH für Frauen von 18-80 Jahren.

c.: Altersabhängiger Verlauf der 50.-Perzentile für TSH für Männer von 18-80 Jahren.

d.: Zeitabhängiger Median für TSH für beide Geschlechter von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).

e.: Zeitabhängiger Median für TSH für Frauen von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).

f.: Zeitabhängiger Median für TSH für Männer von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).

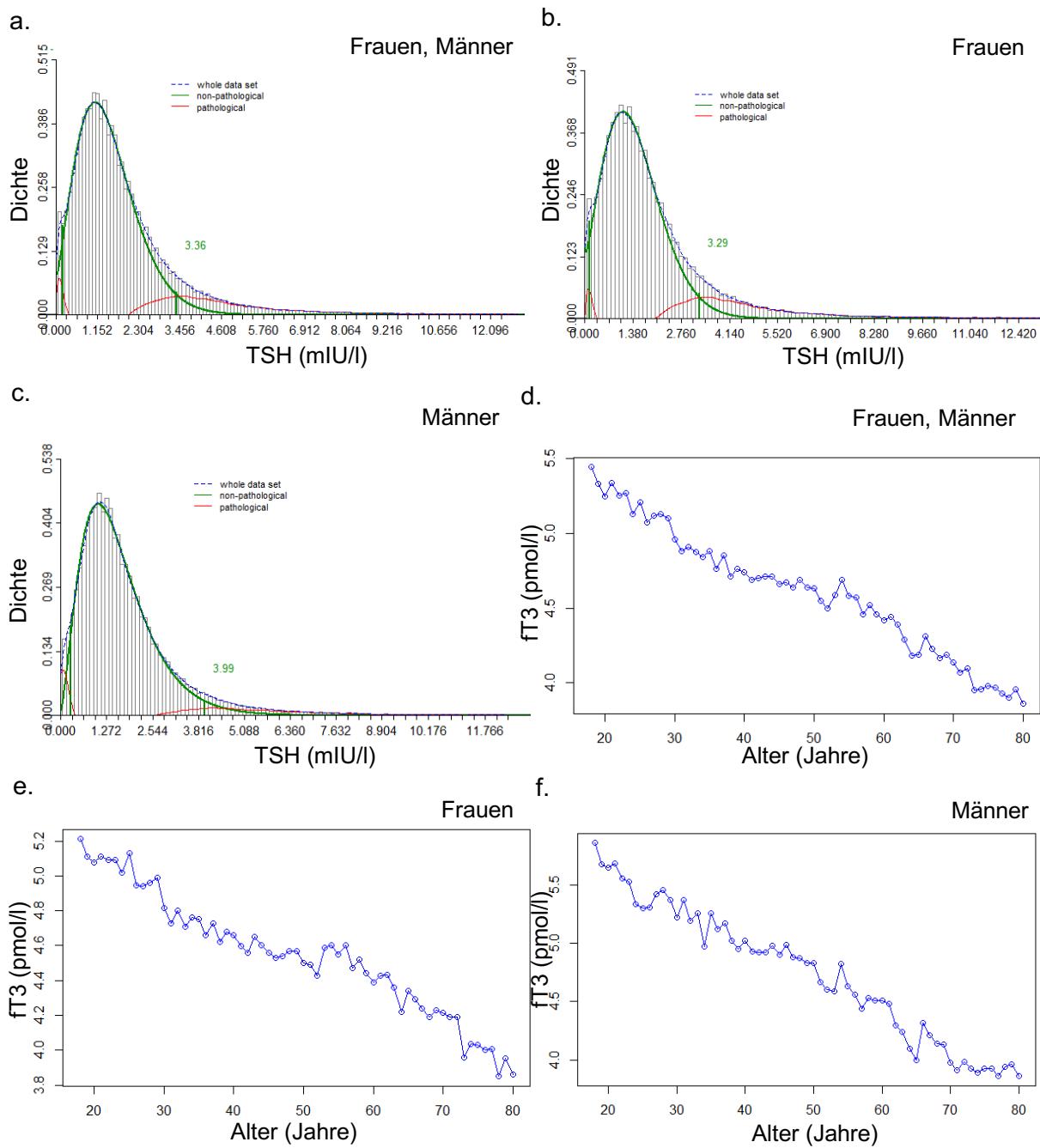


Abbildung 38: a.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für TSH für beide Geschlechter von 18-80 Jahren. Grüne Zahlen geben die obere Referenzgrenze an.

b.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für TSH für Frauen von 18-80 Jahren. Grüne Zahlen geben die obere Referenzgrenze an.

c.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für TSH für Männer von 18-80 Jahren. Grüne Zahlen geben die obere Referenzgrenze an.

d.: Altersabhängiger Verlauf der 50.-Perzentile für fT3 für beide Geschlechter von 18-80 Jahren.

e.: Altersabhängiger Verlauf der 50.-Perzentile für fT3 für Frauen von 18-80 Jahren.

f.: Altersabhängiger Verlauf der 50.-Perzentile für fT3 für Männer von 18-80 Jahren.

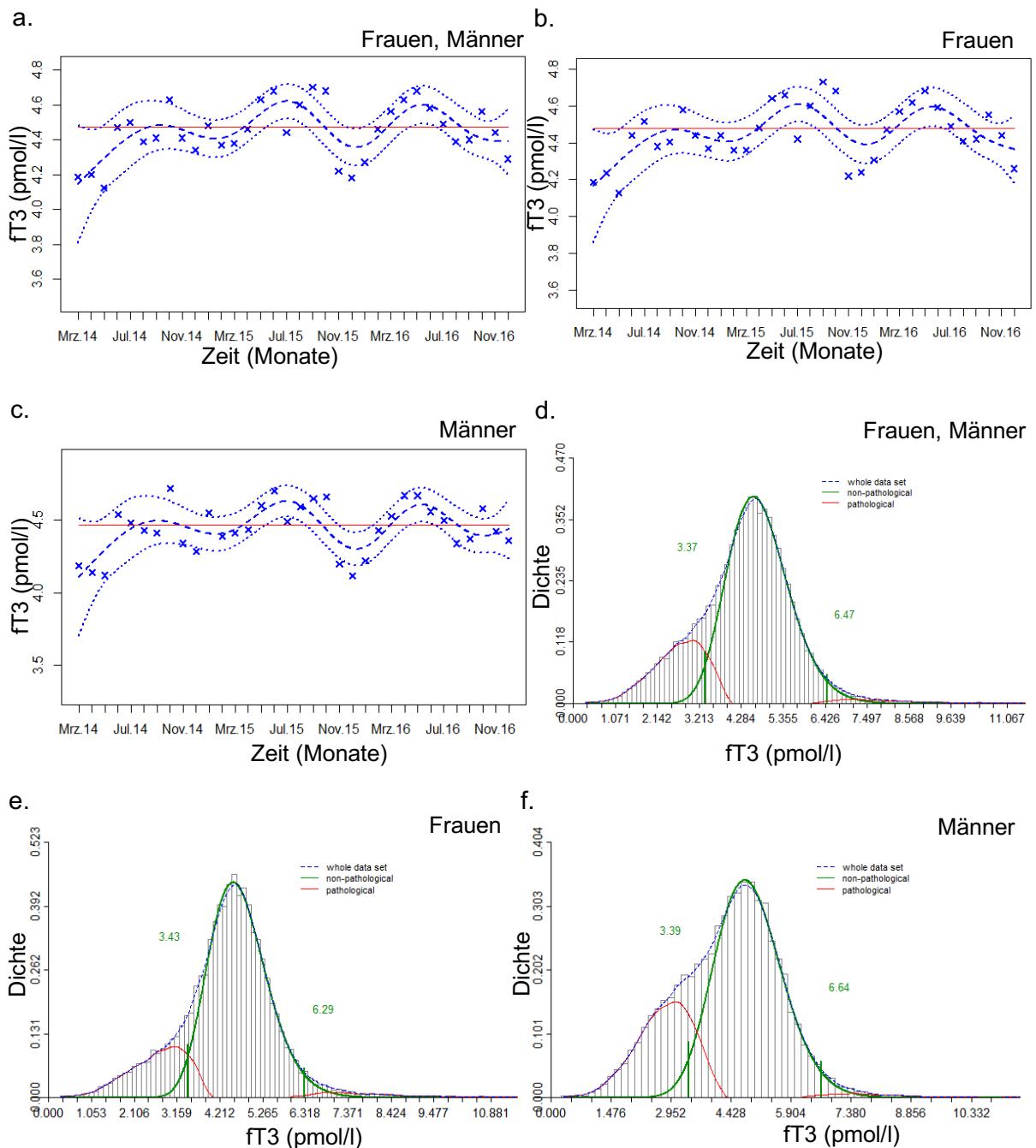


Abbildung 39: a.: Zeitabhängiger Median für fT3 für beide Geschlechter von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).

b.: Zeitabhängiger Median für fT3 für Frauen von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).

c.: Zeitabhängiger Median für fT3 für Männer von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).

d.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für fT3 für beide Geschlechter von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

e.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für fT3 für Frauen von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

f.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für fT3 für beide Geschlechter und für Frauen bzw. Männer von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

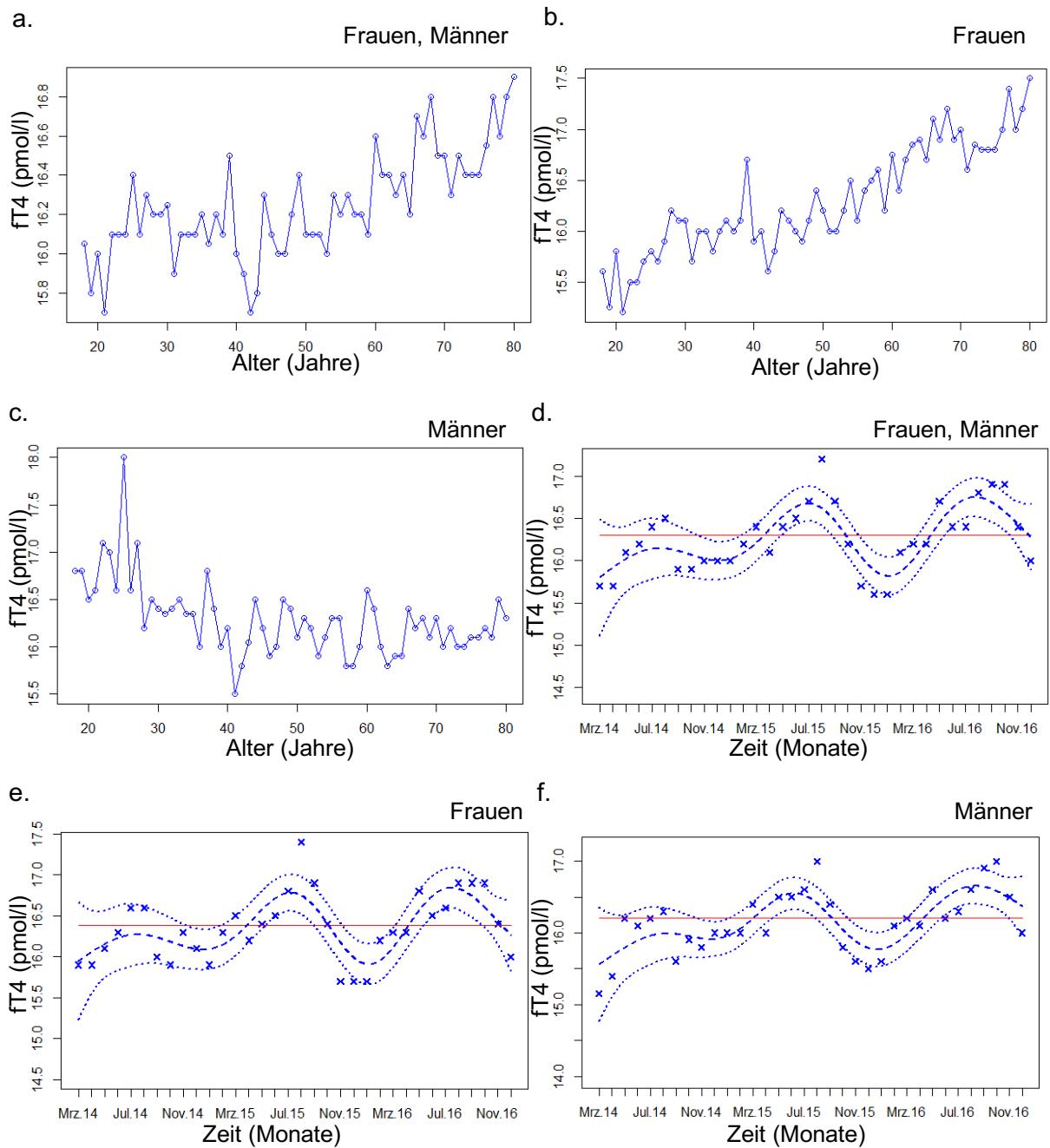


Abbildung 40: a.: Altersabhängiger Verlauf der 50.-Perzentile für fT4 für beide Geschlechter von 18-80 Jahren.

b.: Altersabhängiger Verlauf der 50.-Perzentile für fT4 für Frauen von 18-80 Jahren.

c.: Altersabhängiger Verlauf der 50.-Perzentile für fT4 für Männer von 18-80 Jahren.

d.: Zeitabhängiger Median für fT4 für beide Geschlechter von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).

e.: Zeitabhängiger Median für fT4 für Frauen von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).

f.: Zeitabhängiger Median für fT4 für Männer von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).

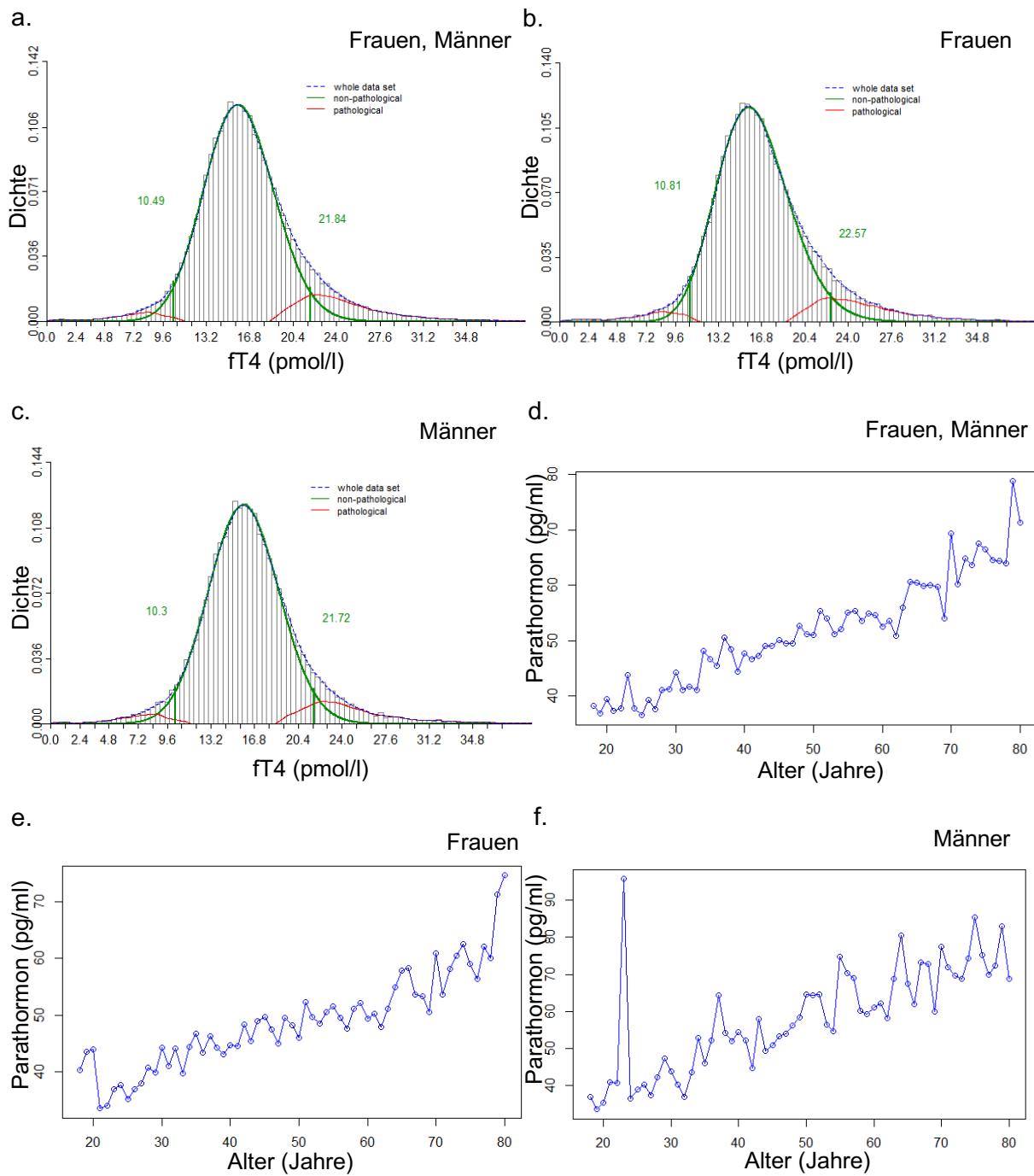


Abbildung 41: a.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für fT4 für beide Geschlechter von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

b.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für fT4 für Frauen von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

c.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für fT4 für Männer von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an.

d.: Altersabhängiger Verlauf der 50.-Perzentile für Parathormon für beide Geschlechter von 18-80 Jahren.

e.: Altersabhängiger Verlauf der 50.-Perzentile für Parathormon für Frauen von 18-80 Jahren.

f.: Altersabhängiger Verlauf der 50.-Perzentile für Parathormon für Männer von 18-80 Jahren.

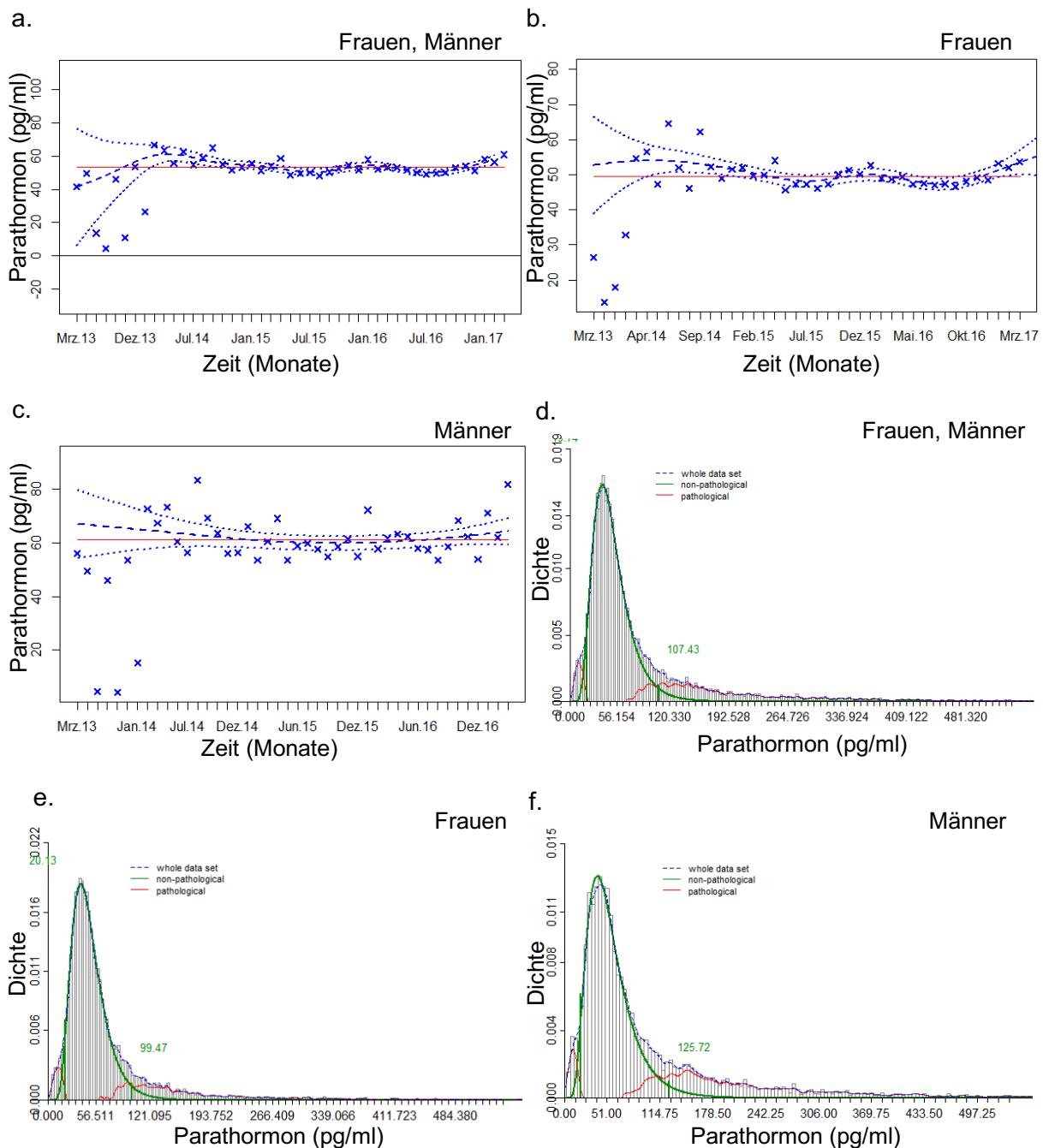


Abbildung 42: a.: Zeitabhängiger Median für Parathormon für beide Geschlechter von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).

b.: Zeitabhängiger Median für Parathormon für Frauen von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).

c.: Zeitabhängiger Median für Parathormon für Männer von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).

d.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Parathormon für beide Geschlechter von 18-80 Jahren. Grüne Zahlen geben die obere Referenzgrenze an.

e.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Parathormon für Frauen von 18-80 Jahren. Grüne Zahlen geben die obere Referenzgrenze an.

f.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Parathormon für Männer von 18-80 Jahren. Grüne Zahlen geben die obere Referenzgrenze an.

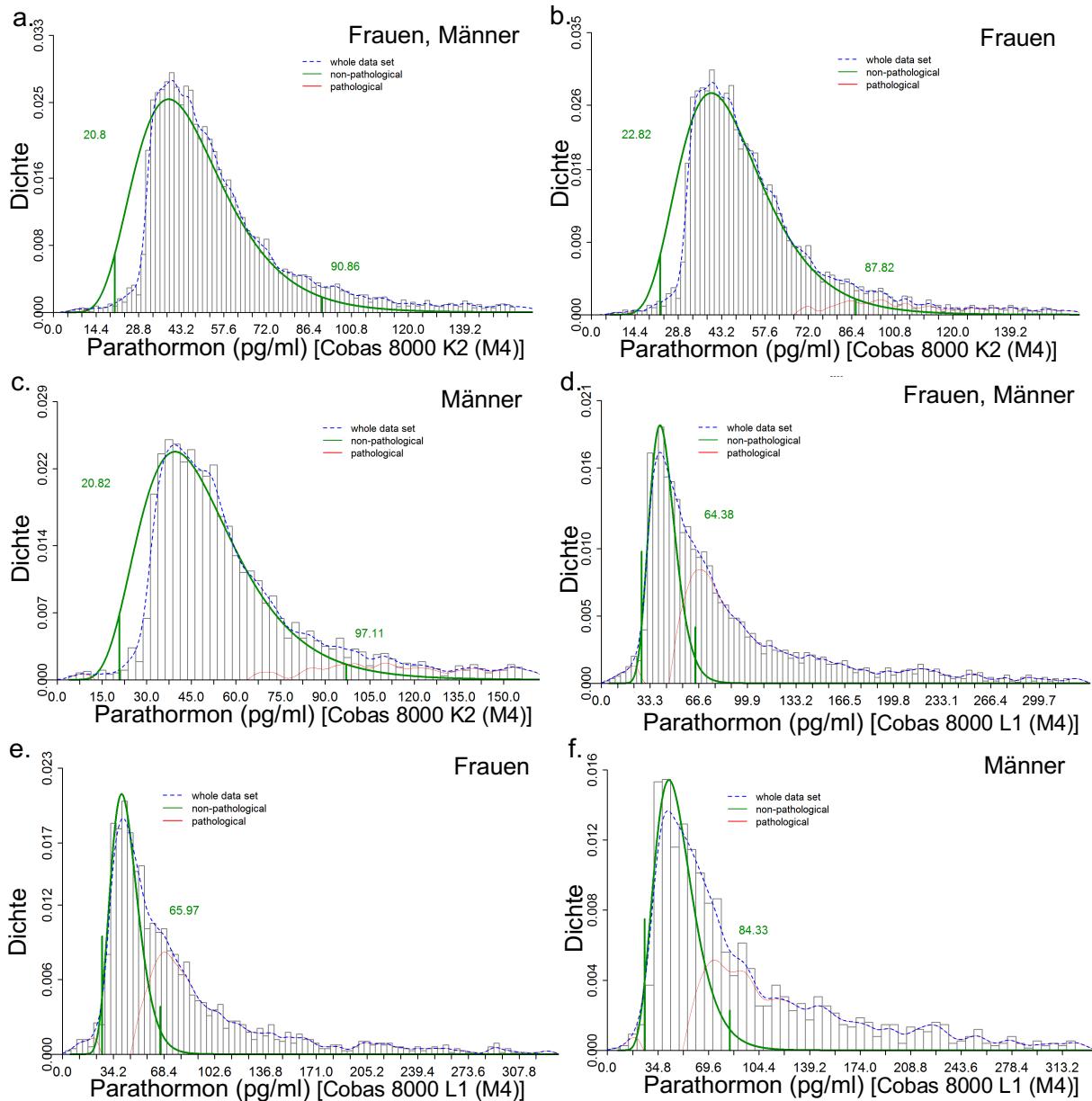


Abbildung 43: a.: Gerätespezifische geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Parathormon für beide Geschlechter von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an. K = Kiel, L = Lübeck, M = Modul.

b.: Gerätespezifische geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Parathormon für Frauen von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an. K = Kiel, L = Lübeck, M = Modul.

c.: Gerätespezifische geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Parathormon für Männer von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an. K = Kiel, L = Lübeck, M = Modul.

d.: Gerätespezifische geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Parathormon für beide Geschlechter von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an. K = Kiel, L = Lübeck, M = Modul.

e.: Gerätespezifische geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Parathormon für Frauen von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an. K = Kiel, L = Lübeck, M = Modul.

f.: Gerätespezifische geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Parathormon für Männer von 18-80 Jahren. Grüne Zahlen geben die untere und obere Referenzgrenze an. K = Kiel, L = Lübeck, M = Modul.

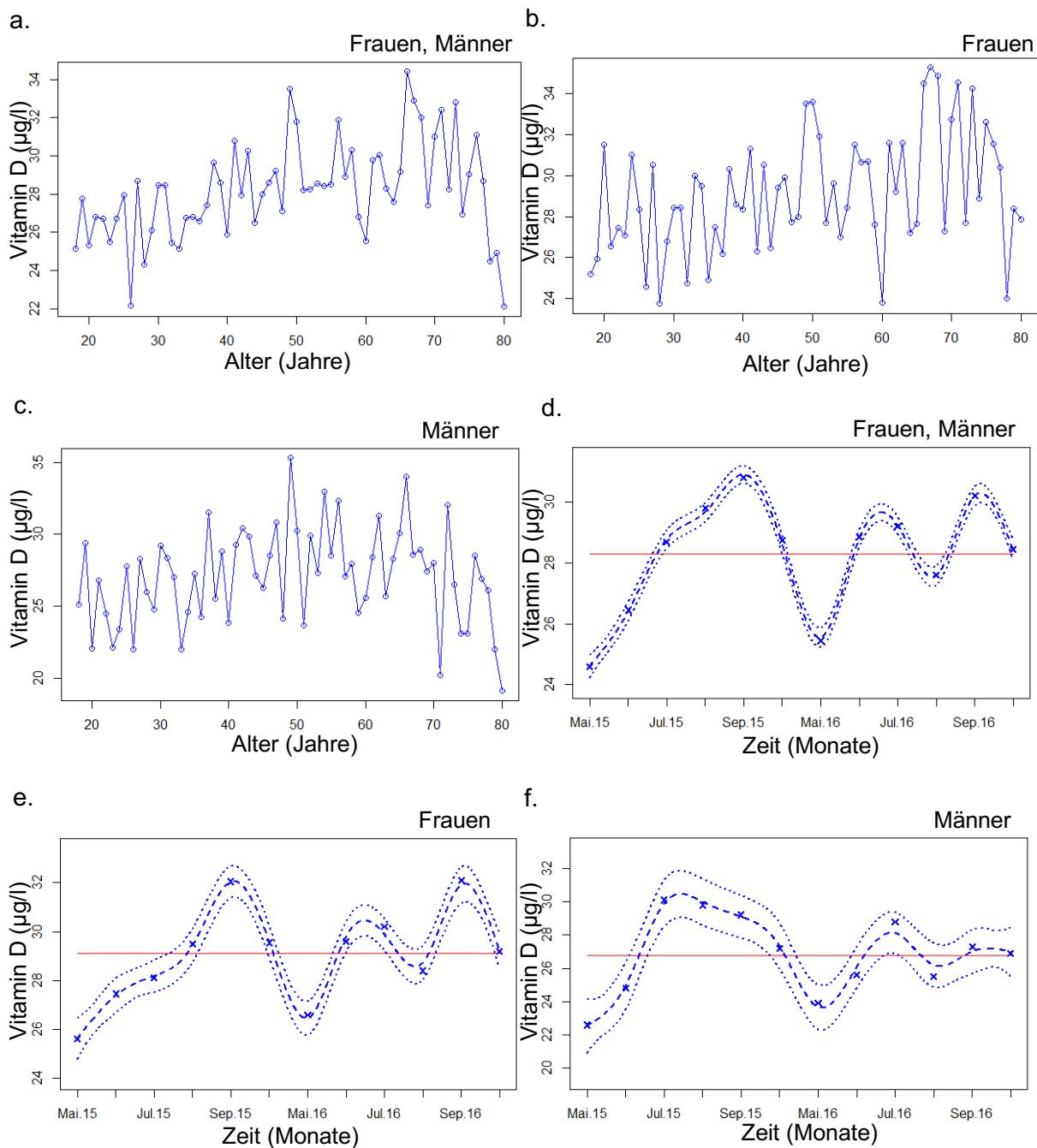


Abbildung 44: a.: Altersabhängiger Verlauf der 50.-Perzentile für Sommermesswerte von Vitamin D für beide Geschlechter von 18-80 Jahren.  
b.: Altersabhängiger Verlauf der 50.-Perzentile für Sommermesswerte von Vitamin D für Frauen von 18-80 Jahren.  
c.: Altersabhängiger Verlauf der 50.-Perzentile für Sommermesswerte von Vitamin D für Männer von 18-80 Jahren.  
d.: Zeitabhängiger Median für Sommermesswerte von Vitamin D für beide Geschlechter von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).  
e.: Zeitabhängiger Median für Sommermesswerte von Vitamin D für Frauen von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).  
f.: Zeitabhängiger Median für Sommermesswerte von Vitamin D für Männer von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).

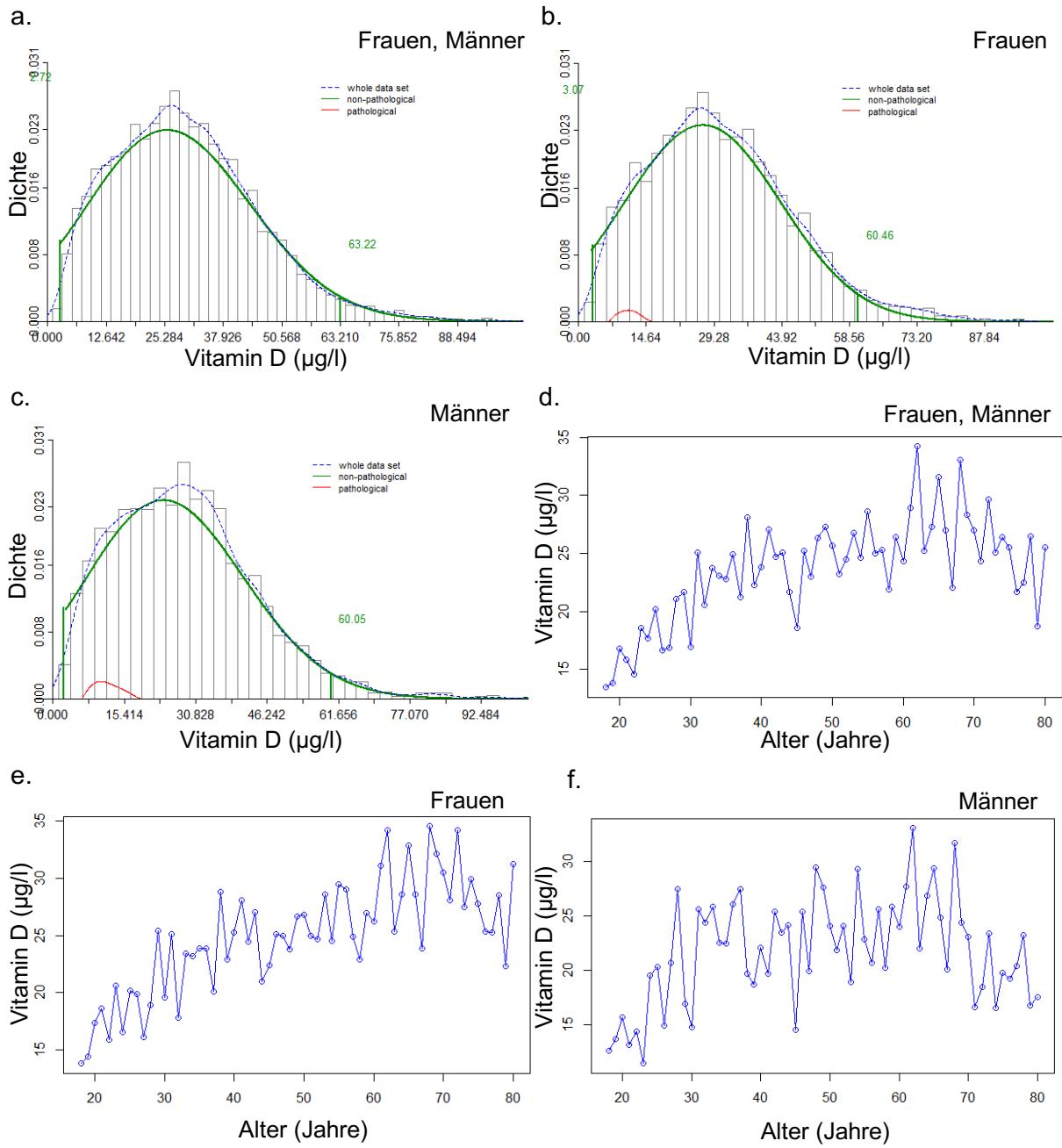


Abbildung 45: a.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Sommermesswerte von Vitamin D für beide Geschlechter von 18-80 Jahren. Grüne Zahlen geben die obere Referenzgrenze an.

b.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Sommermesswerte von Vitamin D für Frauen von 18-80 Jahren. Grüne Zahlen geben die obere Referenzgrenze an.

c.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Sommermesswerte von Vitamin D für Männer von 18-80 Jahren. Grüne Zahlen geben die obere Referenzgrenze an.

Abbildung d.: Altersabhängiger Verlauf der 50.-Perzentile für Wintermesswerte von Vitamin D für beide Geschlechter von 18-80 Jahren.

e.: Altersabhängiger Verlauf der 50.-Perzentile für Wintermesswerte von Vitamin D für Frauen von 18-80 Jahren.

f.: Altersabhängiger Verlauf der 50.-Perzentile für Wintermesswerte von Vitamin D für Männer von 18-80 Jahren.

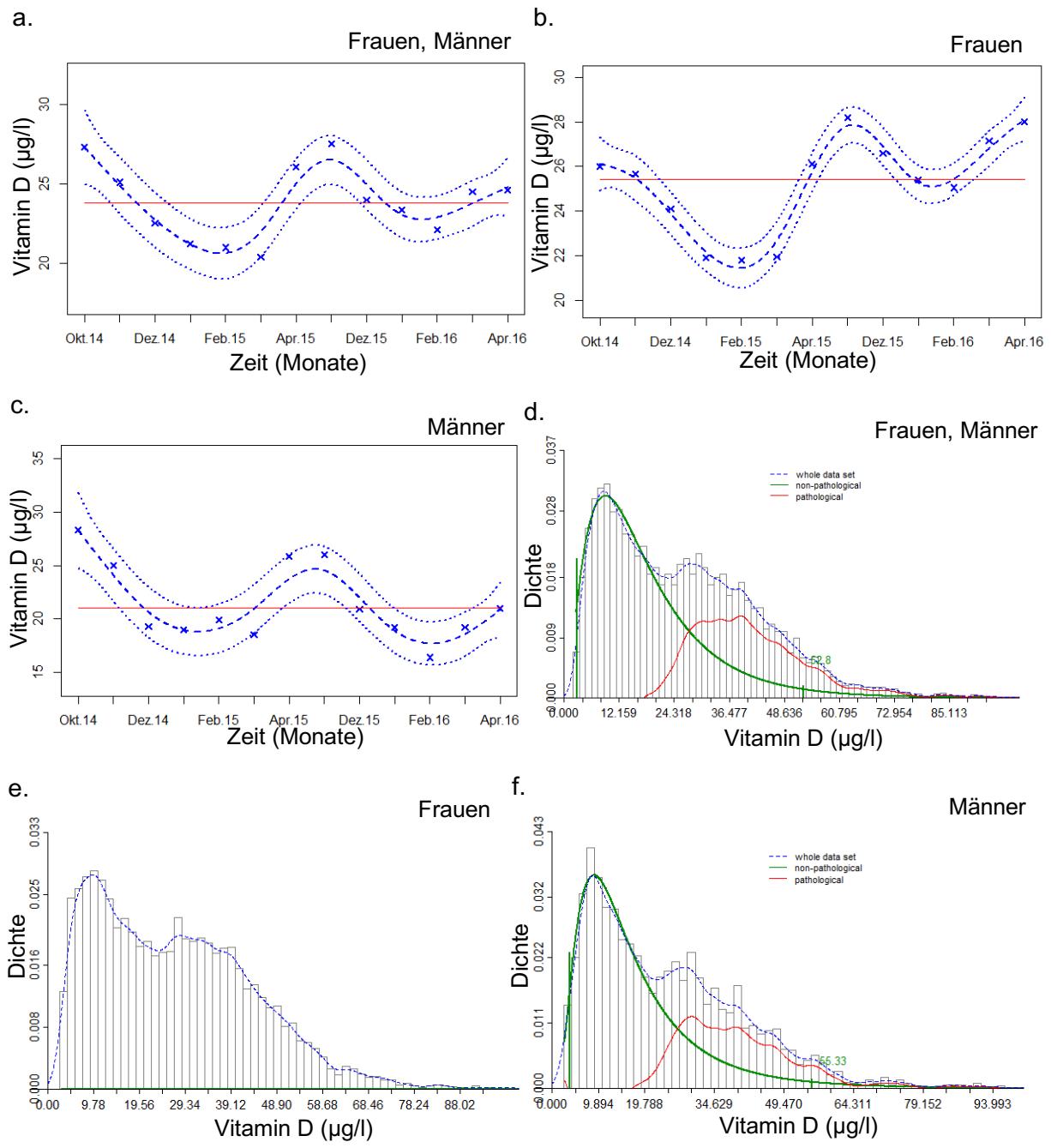


Abbildung 46: a.: Zeitabhängiger Median für Wintermesswerte von Vitamin D für beide Geschlechter von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).  
b.: Zeitabhängiger Median für Wintermesswerte von Vitamin D für Frauen von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).  
c.: Zeitabhängiger Median für Wintermesswerte von Vitamin D für Männer von 18-80 Jahren (blau gestrichelt). KI = Konfidenzintervall (blau gepunktet). Median (rot).  
d.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Wintermesswerte von Vitamin D für beide Geschlechter von 18-80 Jahren.  
e.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Wintermesswerte von Vitamin D für Frauen von 18-80 Jahren.  
f.: Geschätzte Verteilung der nicht-pathologischen (grün), pathologischen (rot) Werte und für alle Daten (blau) für Wintermesswerte von Vitamin D für Männer von 18-80 Jahren.

## **Danksagung**

Ich möchte mich an dieser Stelle bei Herrn Prof. Dr. med. Ralf Junker für die freundliche Überlassung des Themas und Bereitstellung der umfangreichen Datensätze bedanken. Die regelmäßigen Besprechungen und konstruktiven Anregungen waren ein wertvoller Bestandteil in der Bearbeitung der Fragestellung.

Einen großen Dank möchte ich an Frau Dr. rer. nat. Antje Torge für ihre stete Betreuung aussprechen. Die Bereitstellung ihrer entwickelten Methoden und ihre mathematischen Kenntnisse haben wesentlich zum Voranschreiten der Arbeit beigetragen.

Desweiteren möchte ich Herrn Dr. rer. nat. Farhad Arzideh für seine Rückmeldungen und weiterführenden Hinweise danken.

Meiner Familie und meinem Ehemann sei gedankt für ihren fortwährenden Zuspruch und ihre Geduld während der gesamten Bearbeitungsdauer und meiner Patentante für das Korrekturlesen.

Und abschließend: Soli Deo Gloria.