

Aus der Poliklinik für Kieferorthopädie  
der Ludwig-Maximilians-Universität München  
Direktorin: Prof. Dr. Andrea Wichelhaus

**Validierung der Messgenauigkeit von Aufnahmen  
des Parodontalspaltes  
an digitalen Volumentomografien**

Dissertation  
zum Erwerb des Doktorgrades der Zahnmedizin  
an der Medizinischen Fakultät der  
Ludwig-Maximilians-Universität zu München

vorgelegt von  
Maximilian Florentin Berkel  
aus Münster

2018

**Mit Genehmigung der Medizinischen Fakultät der  
Universität München**

Berichterstatterin: Prof. Dr. med. dent. Andrea Wichelhaus

Mitberichterstatter: Prof. Dr. med. Dr. med. dent. Matthias Folwaczny  
Priv. Doz. Dr. med. Harald Kramer

Mitbetreuung durch: Dr. rer. nat. Sebastian Stapfner

Dekan: Prof. Dr. med. dent. Reinhard Hickel

Tag der mündlichen Prüfung: 14.06.2018

Meiner Familie gewidmet

## Inhaltsverzeichnis

1	Einleitung.....	6
2	Literaturübersicht.....	7
2.1	Aufbau des Zahnhalteapparates.....	9
2.2	Kieferorthopädische Zahnbewegung.....	11
2.3	Funktionsweise der Computertomografie.....	12
2.4	Funktionsweise der digitalen Volumentomografie.....	15
2.5	Radiologische 3D-Diagnostik in der Kieferorthopädie.....	17
2.6	Fragestellung.....	18
3	Material und Methoden.....	19
3.1	Versuchsobjekte und Präparation.....	19
3.2	Durchführung der Aufnahmen.....	21
3.3	Vorbereitung der Messungen.....	22
3.4	Durchführung der Messungen.....	29
3.5	Statistische Verfahren.....	31
3.5.1	Statistische Methoden und Tests.....	33
3.5.2	Intrabeurteilerdiskrepanz.....	35
3.5.3	Interbeurteilerdiskrepanz.....	37
3.5.4	Methodenvergleich.....	38
4	Ergebnisse.....	39
4.1	Intrabeurteilerdiskrepanz.....	39
4.1.1	Ergebnisse der deskriptiven Statistik.....	39
4.1.2	Ergebnisse des Technical Error of Measurement.....	46
4.1.3	Ergebnisse des nichtparametrischen Tests.....	47
4.2	Interbeurteilerdiskrepanz.....	50
4.2.1	Ergebnisse der deskriptiven und induktiven Statistik.....	50
4.2.2	Ergebnisse des Technical Error of Measurement.....	62
4.3	Methodenvergleich.....	63
4.3.1	Ergebnisse der deskriptiven und induktiven Statistik.....	63
4.3.2	Methodenvergleich anhand des Technical Error of Measurement.....	70
4.3.3	Methodenvergleich anhand von Punktwolken und Regression.....	71
5	Diskussion.....	74
5.1	Diskussion der Methodik.....	74
5.2	Beurteilung der Proben.....	75
5.3	Diskussion der Ergebnisse der Intrabeurteilerdiskrepanz.....	78

5.4	Diskussion der Ergebnisse der Interbeurteilerdiskrepanz .....	79
5.5	Diskussion der Ergebnisse des Methodenvergleichs .....	80
6	Praktischer Bezug und Schlussfolgerung.....	84
7	Zusammenfassung.....	85
8	Literaturverzeichnis .....	87
9	Danksagung.....	91

---

# 1 Einleitung

Die dreidimensionale Darstellung anatomischer Strukturen ist aus der heutigen Medizin und Zahnmedizin kaum noch wegzudenken. Fast alle Bereiche der Mund-, Kiefer- und Gesichtsheilkunde profitieren von der modernen 3D-Diagnostik.

Das bildgebende Verfahren in Form von radiologischen Aufnahmen stellt einen wichtigen Schritt in der kieferorthopädischen Diagnostik dar. Durch die technische Weiterentwicklung der digitalen dreidimensionalen Röntgentechnik gewinnt die digitale Volumentomografie (DVT) immer größere Bedeutung in der Kieferorthopädie.

Die DVT bietet den Vorteil, dass im Gegensatz zu konventionellen Röntgentechniken wie dem Orthopantomogramm (OPG) oder der Fernröntgenseitenaufnahme (FRS), anatomische Strukturen in räumlichen Bezug gestellt werden. Da die kieferorthopädische Zahnbewegung in alle Richtungen möglich ist, ist eine dreidimensionale Bildgebung zur Therapieplanung und Verlaufskontrolle wünschenswert.

Bei der kieferorthopädischen Zahnbewegung spielt der Parodontalspalt eine übergeordnete Rolle, weswegen detaillierte Informationen über dessen Anatomie essentiell und wissenschaftlich relevant für die Auswirkung von Kräften auf diesen sind. Um eine genaue Aussage über die Anatomie des Parodontalspalts zu treffen, bedarf es einer präzisen dreidimensionalen Bildgebung mit einer ethisch vertretbaren Strahlenbelastung.

Ziel dieser Untersuchung war es, herauszufinden, wie exakt und verlässlich sich kleinste anatomische Strukturen an DVT-Aufnahmen vermessen lassen. Als zu vermessene Struktur wurde der Parodontalspalt gewählt. Zusätzlich sollte ermittelt werden, wie sich Messungen des Parodontalspalt mittels der Methoden Fotografie und Mikrocomputertomografie ( $\mu$ CT) von Messungen an DVT-Aufnahmen unterscheiden.

## 2 Literaturübersicht

Die Anforderungen an moderne Röntgentechniken steigen stetig. Es soll eine möglichst gut zu beurteilende Darstellung mit guter Schärfe und Kontrast bei niedriger Strahlendosis erreicht werden. Um diese Anforderungen zu untersuchen, wurden in etlichen Studien die verschiedenen Röntgentechniken miteinander verglichen.

Besonders das digitale Volumentomogramm (engl.: „Cone-Beam-CT“ (DVT)) und das „Mehrschicht-CT“ (MSCT) spielten dabei eine große Rolle. 2005 verglichen Holberg et al. die Abbildungsqualität von Dental CT („Light Speed Ultra“ der Firma General Electric Company) und DVT (DVT 9000 der Firma QR) miteinander. Sie erhielten das Ergebnis, dass „[Metallartefakte bei den DVT-Aufnahmen im Gegensatz zum Dental-CT kaum und nur sehr abgeschwächt auftraten], während Störungen der Bildqualität durch Bewegungsartefakte ausschließlich bei der DVT zu sehen waren. Die Darstellungsqualität der dentalen und umgebenden knöchernen Strukturen war beim Dental-CT im Durchschnitt deutlich besser als bei der DVT. Während bei DVT-Aufnahmen der Parodontalspalt bei 86 % der Zähne nicht oder nur sehr schlecht zu beurteilen war, lag diese Quote bei den Dental-CT-Aufnahmen bei 20 % (...)“ (Holberg et al. 2005). Das 2005 verwendete DVT verfügte allerdings über ein geringeres Auflösungsvermögen als heutige Geräte. Hashimoto et al. verglichen 2007 das DVT mit dem CT anhand einer getrockneten Maxilla, indem sie anatomische Strukturen wie Schmelz, Pulpenkavum und andere darstellten. Sie klassifizierten die Beurteilbarkeit subjektiv von eins bis fünf. Es stellte sich heraus, dass DVT-Aufnahmen besser zu beurteilen waren. 2008 untersuchten Loubele et al. die Genauigkeit vom DVT im Vergleich mit MSCT. Sie stellten fest, dass die Genauigkeit beider Methoden bei unter einem Millimeter lag und es keinen signifikanten Unterschied bezüglich der Messgenauigkeit zwischen DVT und MSCT gab (Loubele et al. 2008). Im selben Jahr stellten Neugebauer et al. Versuche an, ob das DVT bei der Diagnostik der Lagebeziehung von Weisheitszähnen zum Nervus Alveolaris Inferior Vorteile gegenüber einem Orthopantomogramm (OPG) habe. Vor allem Informationen über die horizontale Position stellte das DVT subjektiv besser dar, als das OPG. Im selben Jahr untersuchten Suomalainen et al. die Messgenauigkeit am MSCT, verglichen mit einem DVT für die Implantatplanung, wenn man die Strahlendosis des MSCT reduziert. Trotz der reduzierten Dosis konnten sie keinen großen Verlust an Messgenauigkeit, verglichen mit der DVT, erkennen. Allerdings erhielten sie mit sinkender Strahlungsdosis mehr insuffiziente

Darstellungen (Suomalainen et al. 2008). 2011 verglichen Kamburoglu et al. verschiedene DVT Gerätetypen in Bezug auf diagnostische Beurteilung der gelieferten Darstellungen. Sie stellten heraus, dass die Art des Gerätes, die Voxelgröße und auch die Leistung des Beurteilers eine Rolle für die Abbildungsqualität und Beurteilbarkeit spielen (Kamburoglu et al. 2011). Asieh Zamani Naser et al. bezogen sich 2013 auf die Planung von Implantaten und den Vergleich der Genauigkeit der Messungen im DVT und MSCT. Sie haben definierte Strecken an der Mandibula vermessen und verglichen die Messungen, die sie im MSCT und im DVT erhielten und stellten dabei fest, dass es einen signifikanten Unterschied von 2.17 % relativ zum MSCT gab (Naser and Mehr 2013). Andere Vergleiche zwischen MSCT, DVT und weiteren Techniken wurden angestellt bzgl. der Strahlendosen der verschiedenen bildgebenden Verfahren. Einheitlich zeigte sich, dass die Computertomografie eine deutlich höhere Strahlenbelastung bedeutet, als das DVT (Coppenrath et al. 2008; Schnelle 2001; Schulze et al. 2004). Etliche Studien weisen auf die Vorteile des DVT gegenüber dem konventionellen CT hin.

Eine genauere radiologische Darstellung kann durch die Verwendung der Mikrocomputertomografie ( $\mu$ CT) erreicht werden. Diese Technik wird vor allem in der Industrie zur Untersuchung von Werkstoffen herangezogen, findet aber auch immer mehr Anwendung in der medizinischen Forschung. 2009 untersuchten Gonzales et al. mit Hilfe des Mikro CTs die Veränderung der Breite des Parodontalspalts und die Achsenverschiebung an Zähnen von Ratten nach kieferorthopädischer Kraftapplikation. Aus den im  $\mu$ CT gewonnenen Daten wurde ein Feniten-Elemente-Modell hergestellt. Sie nutzten also die  $\mu$ CT Technik für die präzise digitale Darstellung kleiner Strukturen (Gonzales et al. 2009). Idiyatullin et al. machten sich ebenfalls das hohe Auflösungsvermögen des  $\mu$ CTs bei ihrer Untersuchung zur Rissdarstellung in Zähnen zu nutze. Hierbei verwendeten sie das  $\mu$ CT als Referenz um einen Vergleich zwischen Magnetresonanztomografien und DVT-Aufnahmen bzgl. der Beurteilung der Risse herzustellen.

Leider konnte der Vergleich zwischen  $\mu$ CT und DVT nicht beurteilt werden, da die DVT-Aufnahmen aufgrund metallischer Restaurationen verzerrt waren (Idiyatullin et al. 2016).



---

## 2.1 Aufbau des Zahnhalteapparates

Der Zahnhalteapparat (Parodontium) besteht aus der Gingiva, dem Desmodont, dem Wurzelzement und der Zahnalveole. Die Gingiva besteht aus Epithel und Bindegewebe. Sie umgibt Zahn und Knochen und haftet ihnen an (Heidemann 1997). Es wird unterschieden zwischen „freier marginaler Gingiva“, „befestigter Gingiva“ und „interdentaler Gingiva“ ((Renggli et al. 1984) (S. 1 ff.)). Die gedachte Linie an der Grenze zwischen befestigter und beweglicher Gingiva wird als „mukogingivale Linie“ bezeichnet. Das Desmodont, auch parodontales Ligament (PDL) genannt, besteht aus festem Bindegewebe, das die Zahnwurzel umhüllt. Es ist zwischen Zement und Alveolarknochen lokalisiert und stellt die Verbindung zwischen Zahnwurzel und Knochen her (Heidemann 1997). Die Breite des PDLs wird bei H.P. Mueller mit 0,16-0,24 mm am Apex, 0,12-0,17 mm an der Mitte der Wurzel und 0,12-0,17 mm am Rande des Alveolarknochens angegeben (Mueller 2004). Das Bindegewebe des Desmodonts stellt sich als derbfaserige Kollagenfibrillen dar. Mehrere dieser Fibrillen bilden in Gruppen gleicher Verlaufsrichtung dann das Kollagenfaserbündel. Aufgrund der Verlaufsform, dem Ansatz und dessen Ursprung, unterscheidet man folgende Bündel: „gingivale Faserbündel“, „transseptale Faserbündel“, „zirkuläre Faserbündel“ und „zementoalveoläre Faserbündel“. Die Bündel, welche von Zement zu Knochen verlaufen, nennt man „Sharpey-Fasern“, die weiter in horizontal verlaufende, schräg verlaufende, sowie apikale und interradiäre Faserbündel unterteilt werden (Mueller 2004; Renggli et al. 1984). Ebenfalls im Desmodont vorhanden sind Zellen wie Fibroblasten, Zementoblasten, Zellen des Immunsystems und andere. Das Desmodont ist gut vaskularisiert und nerval versorgt (Mueller 2004). Das Wurzelzement ist eine dünne mineralisierte Bindegewebsschicht, die die Zahnwurzel umhüllt. Unterschieden wird histologisch in zelluläres, azelluläres, fibrilläres und afibrilläres Zement (Heidemann 1997). Aufgrund von Zementoblasten kann Zement stetig nachgebildet werden (Renggli et al. 1984). Die Aufgabe des Zements liegt hauptsächlich in der Befestigung der Desmodontalfasern an der Wurzeloberfläche (Mueller 2004).

Der Alveolarknochen besteht aus einer dünnen Knochenlamelle (Alveolen-Innenkortikalis), in die die Sharpey-Fasern eingelassen sind (Renggli et al. 1984). Der Aufbau des Parodontiums und speziell die Befestigung des Zahnes mit kollagenen Fasern bewirken, dass der Zahn nicht starr mit dem Knochen verwachsen ist. Vielmehr ergibt sich daraus eine leichte

physiologische Beweglichkeit. Die Aufgaben des Parodontium erstrecken sich von der offensichtlichen Verankerung des Zahnes im Knochen, der Anpassung an funktionelle und topografische Veränderung über die Abwehr gegen Infektionen, bis hin zur Empfindung von Druck und Schmerz (Mueller 2004). Abbildung 1 zeigt ein schematisches Schnittbild des Zahnhalteapparates. Aus Übersichtsgründen wurden zusätzlich Schmelz, Dentin und Pulpa dargestellt.

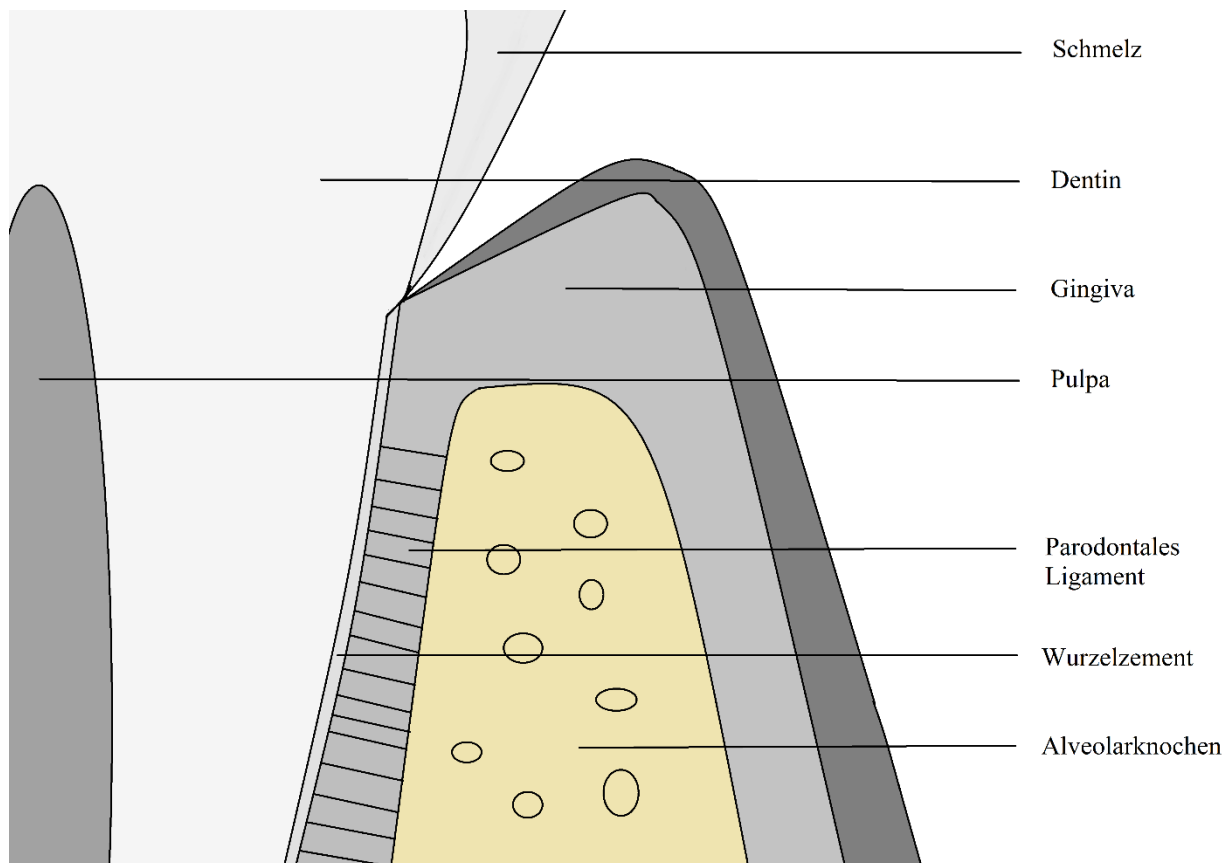


Abbildung 1: Schematisches Schnittbild des Zahnhalteapparates.

## 2.2 Kieferorthopädische Zahnbewegung

Der Parodontalspalt sowie Zellen des Alveolarknochens sind ständigen Kräften ausgesetzt. Diese können durch die physiologische Kaubewegung, aber auch durch pathologische Parafunktionen wie Bruxismus begründet sein. Diese Kräfte resultieren in Umbauprozessen des PDLs und des Alveolarknochens (Beertsen et al. 1997).

Diesen Vorgang macht sich die Kieferorthopädie zu Nutze, indem gezielt Kräfte eingesetzt werden, um gewünschte Umbauprozesse des PDLs zu induzieren. Nach Applizierung von Kräften durch kieferorthopädische Geräte wird in der initialen Phase der Zahn in der Alveole bewegt und eine sehr schnelle Stellungsänderung erreicht. Es entstehen Zonen, in denen der Zahn gegen den Alveolarknochen gedrückt wird (Druckzonen), sowie Zugzonen, in denen ein durch die Sharpey-Fasern vermittelter Zug auf den Knochen ausgeübt wird. Durch diesen Druck bzw. Zug an den Zellen des PDLs wird bereits eine zelluläre Reaktion hervorgerufen.

Diese Phase dauert, je nach Literaturquelle, ca. ein bis vier Tage. An diese initiale Zahnbewegung schließt sich eine Ruhephase an, die durch eine Minderdurchblutung (Hyalinisierung) des PDLs in den Druckzonen gekennzeichnet ist (Goz et al. 1992). In dieser Phase findet keine Zahnbewegung statt. Durch Entfernung nekrotisierten oder resorbierten Gewebes durch Makrophagen und Osteoklasten in der Druckzone wird die Zahnbewegung wieder aufgenommen. In der Zugzone beginnen Osteoblasten eine neue Knochenmatrix zu bilden.

Ca. 30 Tage nach Beginn der Krafteinwirkung kommt es zu einer Beschleunigungs- und Linearphase, in der der größte Anteil der kieferorthopädischen Zahnbewegung durch stetige Umbauprozesse stattfindet (Burstone 1962; Pilon et al. 1996).

### 2.3 Funktionsweise der Computertomografie

Die Computertomografie ist ein computergestütztes bildgebendes Verfahren in der Radiologie. Es macht sich das stoffabhängige Durchdringungsvermögen von Röntgenstrahlung zu Nutze. Dieses ist vor allem begründet durch Absorption und Streuung. Die an einem Detektor ankommende Strahlung hat abhängig von den Schwächungseigenschaften des durchdrungenen Gewebes an Intensität verloren. Diese Schwächung wird von einem Detektor gemessen und digital verarbeitet.

1972 realisierte Godfrey N. Hounsfield mit der Firma EMI den ersten CT Scanner (Hounsfield 1973). Vor diesem Jahre wurde die Rekonstruktion noch mit aufwendigen mathematischen Formeln analog errechnet. Durch den Fortschritt in der Computertechnologie wurde es 1972 möglich, die Computertomografie praktisch zu nutzen. Das Grundprinzip war damals schon das gleiche wie heute. Ausschlaggebend für die dreidimensionale Darstellung ist die digitale Aufnahmetechnik aus verschiedenen Richtungen. Die erste Generation (Rotation-Translation des Nadelstrahls) der Computertomografen verfügte über eine Röntgenquelle, die einen einzelnen Nadelstrahl aussendete, der von einem synchron rotierenden gegenüberliegendem Sensor erfasst wurde. Sensor und Röntgenquelle mussten bei dieser Generation allerdings pro Projektionswinkel linear und synchron manuell verschoben werden, was diese Methode sehr zeitaufwendig machte. In einem Radius von 180 Grad wurden pro Grad Daten gesammelt und computertechnisch verarbeitet.

In der zweiten Generation (Mehr-Detektor-Rotations-Translations-Scanner) wurde der sogenannte Fächerstrahl eingeführt, der allerdings so klein eingeblendet war, dass eine lineare Verschiebung weiterhin stattfinden musste. Dieses Problem konnte man in der dritten Generation schließlich lösen, indem man den Fächer größer wählte und somit eine lineare Verschiebung unnötig machte, was eine erhebliche Zeitersparnis bedeutete. Die dritte Generation (Rotation-Rotation in Einzelschichten) unterscheidet sich nur durch einen beweglichen Detektor von der vierten Generation, der bei den Geräten der vierten Generation im Sinne eines Detektorringes starr befestigt ist. Durch die Einführung der Schleifringtechnologie, wodurch es möglich wurde, dass die Abtasteinheit kontinuierlich rotiert, wurde die CT Technik auch im klinischen Leben für Aufnahmen des Thorax und des Abdomens zeiteffektiv praktikabel (Buzug 2004).

In dieser Untersuchung wurde ein  $\mu$ CT, welches, ähnlich wie bei der DVT-Technik, über ein kegelförmiges Strahlenbündel und einen Flächendetektor verfügt, verwendet ([www.scanco.ch](http://www.scanco.ch)). Das  $\mu$ CT rückt in der medizinischen Forschung immer mehr in den Vordergrund, da es ein nicht-invasives, sehr genaues Diagnostikinstrument darstellt (De Clerck et al. 2004). De Clerck et al. untersuchten beispielsweise Lungentumore bei Mäusen und zeigten damit, dass das  $\mu$ CT im Gegensatz zum DVT auch Weichgewebe darstellen kann.

Das besondere am  $\mu$ CT sind die Röntgenröhre und das Detektorelement. Speziell die Fokusgröße und die Größe der Detektorelemente, die das Auflösungsvermögen bestimmen. Dieses liegt nach Angaben der Firma ScyScan bei beispielsweise 10  $\mu$ m. Typischerweise ist das Messfeld mit 2  $\text{cm}^3$  sehr klein, was derzeit den hauptsächlichen Indikationsbereich begründet, nämlich der Material- und Mikrostrukturprüfung (Buzug 2004).

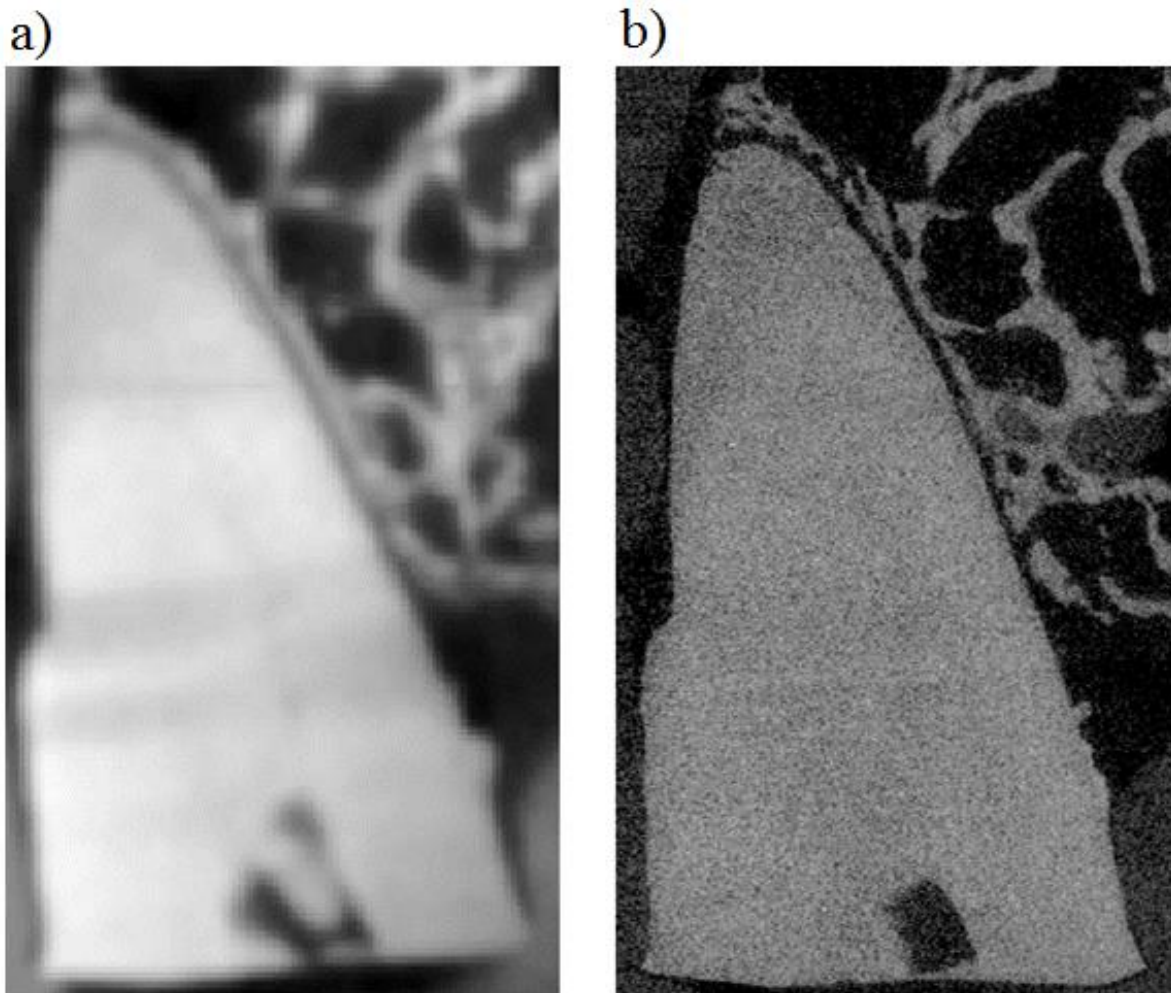


Abbildung 2: DVT(a)- und  $\mu$ CT-Aufnahme(b) eines Zahnes im Vertikalschnitt.

Abbildung 2 zeigt die Aufnahmen eines Zahnes im Vertikalschnitt, welche zum einen mit einem digitalen Volumentomografen (a) und zum anderen mit einem  $\mu$ CT (b) durchgeführt wurden. Die Unterschiede in Auflösung und Kontrast sind deutlich. Besonders die Trabekelstrukturen und der Parodontalspalt werden mit dem  $\mu$ CT wesentlich deutlicher dargestellt.

## 2.4 Funktionsweise der digitalen Volumentomografie

Vor ca. 20 Jahren entwickelt (Arai et al. 1999), stellt die digitale Volumentomografie (DVT) eine heutzutage weitverbreitete radiologische Aufnahmetechnik zur Darstellung knöcherner Strukturen am menschlichen Schädel in der Zahnmedizin dar. Die Aufnahme läuft ähnlich wie beim Orthopantomogramm (OPG) ab. Die Sender-Detektoreinheit umkreist zusammen mit der gegenüberliegenden Röntgenquelle den Kopf des Patienten und erzeugt in definierten Winkelintervallen Einzelprojektionen. Das Rotationszentrum stellt die „Region des besonderen Interesses“ (ROI, engl.: „Region Of Interest“) dar. Diese ROI ist im späteren Bildvolumen das Zentrum. Bei den meisten DVT-Geräten lässt sich das Sichtfeld eingrenzen (s. Abbildung 3). Dies hat den Vorteil einer Reduzierung der Strahlenbelastung (White and Pharoah 2014). Bei dieser sogenannten „Cone-Beam-Technik“ wird ein kegelförmiges Strahlenbündel erzeugt und auf einem Flächendetektor mit CCD Technologie (CCD = charge-couple-device) erfasst. Hieraus entstehen die Rohdaten.

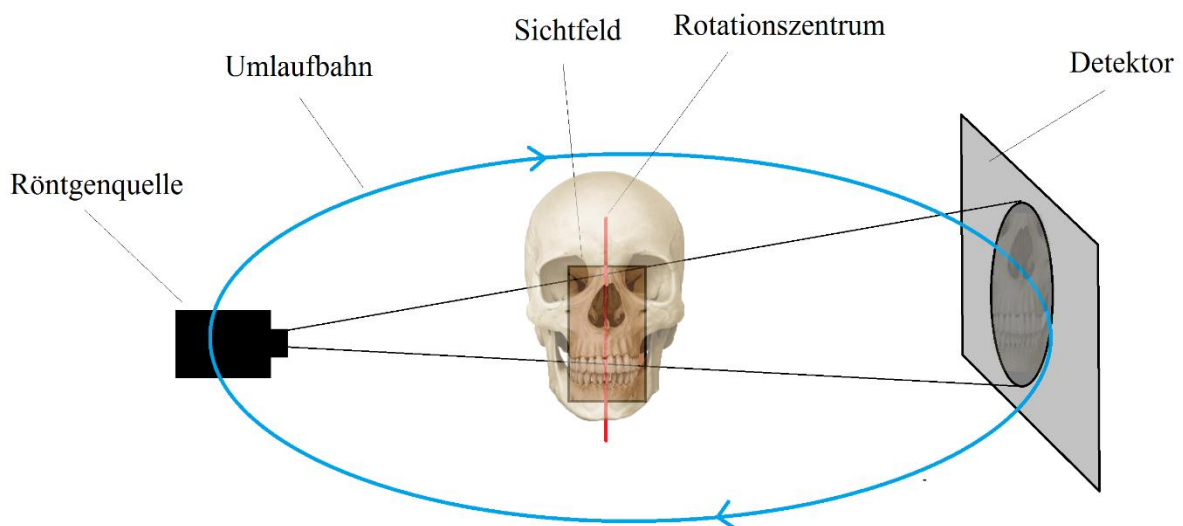


Abbildung 3: Schematische Abbildung der Funktionsweise der digitalen Volumentomografie.

Die Rohdaten werden mittels eines Computeralgorithmus in Querschnittsbilder verarbeitet. Das Grundelement ist das Pixel. Die Graustufung dieser Pixel korreliert mit der Strahlenintensität, die auf den Detektor trifft. Hieraus entstehen mehrere hundert zweidimensionale Einzelbilder, aus denen anschließend das Bildvolumen erstellt wird. Das Bildvolumen besteht aus dreidimensionalen Pixeln, den sogenannten Voxeln (White and Pharoah 2014).

Laut Scarfe (Scarfe et al. 2006) kann sich, unter anderem durch die Minimierung des Sichtfelds, bei dem „Cone-Beam-Verfahren“ eine deutliche Reduktion der Strahlenexposition gegenüber dem konventionellem CT ergeben (Cohnen et al. 2002). Nach Herstellerangaben eines Kodak Carestream CS 9300 beträgt die Scanzeit 12-28 (+/- 10 %) Sekunden, was ein Siebtel der Scanzeit eines konventionellen medizinisches CTs ist (Mohan et al. 2011). Kurze Aufnahmezeiten reduzieren Bewegungsartefakte. Die Verarbeitung der Daten liegt abhängig vom Computersystem bei unter zwei Minuten und liegt somit weit unter dem Zeitaufwand einer Computertomografie.

Eine allgemeine tabellarische Gegenüberstellung von DVT und CT bietet die Tabelle 1.

*Tabelle 1: Gegenüberstellung der Merkmale der DVT- und CT-Technik.*

	<b><i>DVT</i></b>	<b><i>CT</i></b>
Strahlendosis:	- niedrig	- hoch
Aufnahmezeit:	- kurz	- lang
Strahlengeometrie:	- kegelförmig	- fächerförmig
Datenverarbeitung:	- Datensatz als Volumenbild	- Datensatz als Einzelschichten
Funktionsweise:	- typisch eine 360° Rotation	- mehrere Umkreisungen
Detektorenart:	- nur ein Flachdetektor	- mehrere zirkuläre Detektoren



---

## 2.5 Radiologische 3D-Diagnostik in der Kieferorthopädie

Die seit 2013 amtierende Präsidentin der deutschen Gesellschaft für Kieferorthopädie, (DGKFO) Frau Prof. Dr. Ursula Hirschfelder, veröffentlichte 2008 eine Stellungnahme zur radiologischen 3D-Diagnostik in der Kieferorthopädie. Darin stellte sie Vorteile, Indikation und Indikationseinschränkung für die 3D-Diagnostik übersichtlich dar. Als Vorteile seien die dreidimensionale Erfassung von räumlichen Beziehungen, die Möglichkeit aus den vorhandenen Datensätzen transaxiale Primärschnittbilder, Panoramaschichtaufnahmen und virtuelle Fernröntgenseitenbilder ohne erneute Strahlenexposition zu generieren (Kumar et al. 2008), sowie die Weiter-, bzw. Neuverarbeitung vorhandener Volumendatensätze mittels geeigneter Auswertungssoftware, zu nennen. Als Indikation nennt sie folgende acht Punkte:

1. „Diagnostik von Anomalien des Zahnbestandes, insbesondere differentialdiagnostische Beurteilung der anatomischen Form und der Topographie noch nicht durchgebrochener überzähliger Zahnanlagen (Mesiodentes, syndromgebundene Hyperdontie) (Bernhard Böhm and Hirschfelder 2000).
2. Diagnostik von Anomalien und Dysplasien der Zahnwurzeln einschließlich von konventionell röntgenologisch nicht erkennbaren Wurzelresorptionen.
3. Differentialdiagnostische Bewertung von Zahndurchbruchstörungen (Kapila et al. 2011).
4. Ermittlung der Topographie retinierter und verlagertes Zahnkeime.
5. Beurteilung pathologischer Knochenstrukturen (z.B. Zysten, Odontome, entzündliche und tumorbedingte Läsionen).
6. 3D-Differentialdiagnostik von komplexen angeborenen oder erworbenen kraniofazialen Fehlbildungen sowie von Gesichts- und Kieferasymmetrien.
7. 3D-Differentialdiagnostik der Spaltmorphologie bei uni- und bilateralen Lippen-, Kiefer- und Gaumenspalten einschließlich der Planung und Kontrolle der Kieferspaltosteoplastik.
8. Darstellung des peridental Knochenangebots zur prognostischen Bewertung geplanter Zahnbewegungen (z.B. im parodontal vorgeschädigtem Gebiss, z.B. zur Planung von Zahnbewegungen spaltangrenzender Zähne bei Patienten mit Lippen-Kiefer-Gaumenspalten) (Romero-Delmastro et al. 2014)“(Hirschfelder 2008).

Eine Einschränkung der Indikation stellt neben den hohen Anschaffungskosten vor allem die höhere Strahlenbelastung (Silva et al. 2008) im Vergleich zur konventionellen Röntgendiagnostik, sowie die mangelnde Darstellung von Weichgewebe, dar. Daher gilt auch für den Einsatz von 3D-Röntgen der Grundsatz der strengen und zurückhaltenden Indikationsstellung (Hirschfelder 2008).

## 2.6 Fragestellung

In der vorliegenden Studie wurden die Messungen der Breite des Parodontalspalt an DVT-Aufnahmen mit einer Voxelgröße von 90  $\mu\text{m}$  vermessen werden kann.

Der Vergleich mit hochauflösenden Fotografien und  $\mu\text{CT}$ -Aufnahmen diene als Referenz bezüglich der Akkuranz der Messungen.

Zudem sollte eine eventuelle Abhängigkeit der Messungen von Beurteiler und Zeitpunkt der Messung überprüft werden.

Hieraus ergaben sich folgende Fragestellungen:

Wie unterscheiden sich die Messungen des Parodontalspalts an DVT-Aufnahmen, verglichen mit äquivalenten Messungen an Fotografien und  $\mu\text{CT}$ -Aufnahmen?

Vermessen verschiedene Beurteiler dieselben Messstellen methodenabhängig different und wie reproduzierbar sind diese Messungen?

### **3 Material und Methoden**

Im Folgenden werden die verwendeten Materialien, die Herstellung der Versuchsobjekte und die angewandten Methoden genauer erläutert. Zusätzlich wird das statistische Verfahren beschrieben.

#### **3.1 Versuchsobjekte und Präparation**

Für diese Untersuchung wurden möglichst nicht beeinträchtigte Humanpräparate verwendet. Diese wurden einer in Formalin konservierten menschlichen Leiche, die der Pathologie der Ludwig-Maximilians-Universität München für Forschungszwecke zur Verfügung stand, entnommen. Zähne mit noch umgebenden Knochen wurden aus der Mandibula gesägt. Die meist mit metallischen Restaurationen versorgten Zahnkronen wurden entfernt, da diese zu Artefaktbildung im DVT und  $\mu$ CT führen könnten (Zimmermann. et al. 2014). Die „Knochen/Zahn-Blöcke“ wurden entlang der Zahnachse zersägt, um den Parodontalspalt darzustellen und offen zu legen. Nach Beurteilung der Qualität und Verwendbarkeit der Proben ergab sich die Probenanzahl fünf.

Aufgrund der Limitierung der Größe der Proben durch den Mikrocomputertomografen wurden alle Bereiche des biologischen Materials entfernt, die für diese Untersuchung nicht von Interesse waren. Um die fünf Proben zu fixieren, wurde das Polysiloxan „Optosil P Plus“ der Firma Heraeus verwendet. Die zur Orientierung und Kalibrierung dienlichen Keramiksplitter wurden zunächst aus Wachs modelliert und anschließend als Lithiumdisilikatkeramik („IPS e.max press“) gepresst und geschliffen. Die Zahnblöcke wurden mit der Schnittseite, zusammen mit den Keramiksplintern, plan auf eine Glasplatte gelegt. Anschließend wurde das Optosil im plastischen Zustand über Zahn und Splitter auf die Glasplatte gedrückt und abgewartet, bis das Optosil ausgehärtet war. Dadurch wurde eine glatte Oberfläche geschaffen. Nach Ablösen von der Glassplatte wurden rechteckige, ca. 7 mm mal 30 mm mal 20 mm (H/B/T) große Blöcke ausgeschnitten.

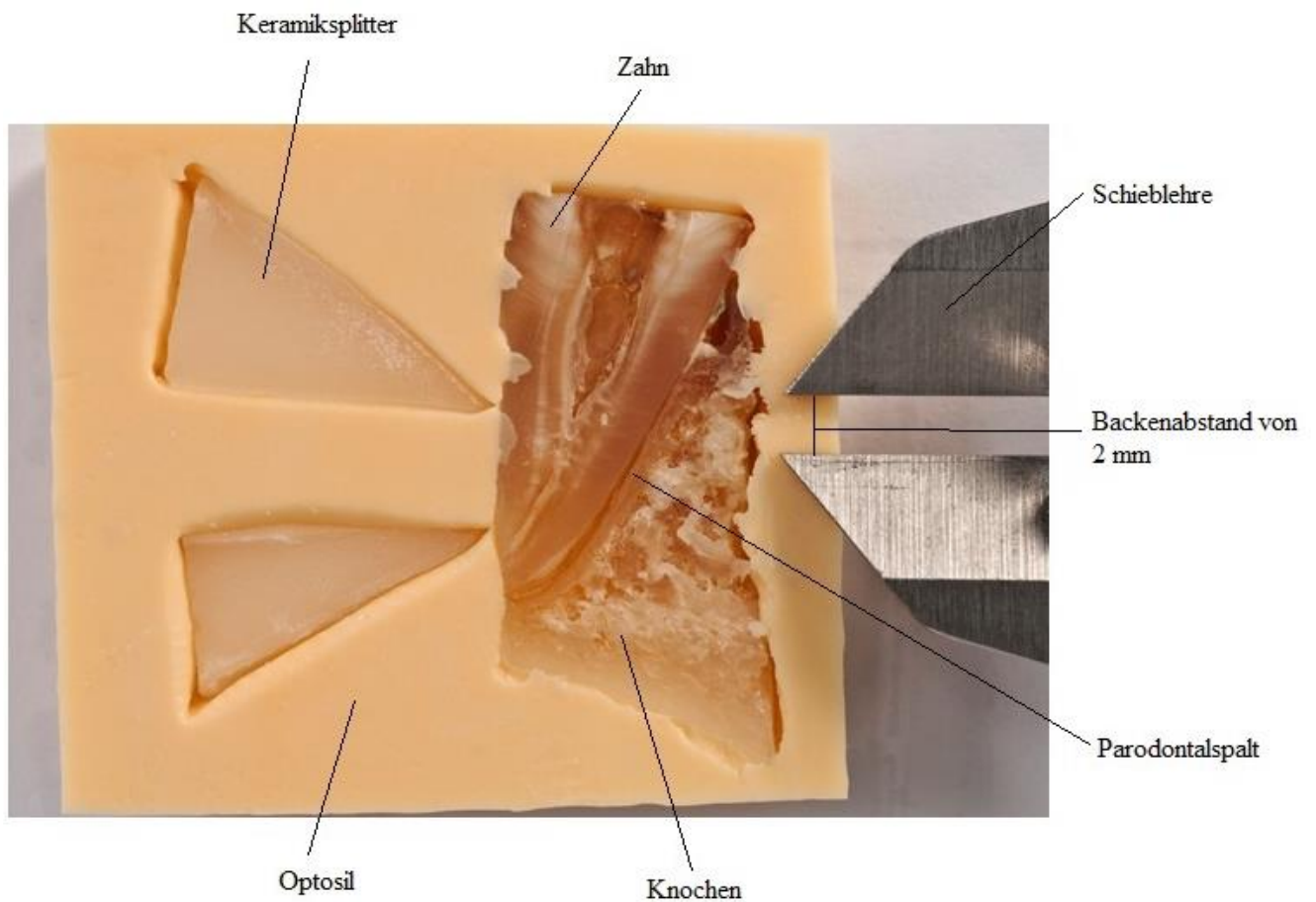


Abbildung 4: Fotografie einer Probe mit Schieblehre.

Abbildung 4 zeigt die Fotografie einer fertiggestellten Probe. Der Quader besteht aus in Optosil gedrückten Keramiksplittern und dem „Zahn/Knochen-Block“. Die abgebildeten Backen der Schieblehre haben einen exakten Abstand von zwei Millimetern und dienen der Kalibrierung der Messungen (s. Kapitel 3.2).

### **3.2 Durchführung der Aufnahmen**

Hochauflösende Fotos der Proben wurden mit der Spiegelreflexkamera „Nikon BM-8“ mit einem „Sigma Foto 70mm F2,8 EX DG Makro“ Objektiv erzeugt. Die Kamera wurde senkrecht zu den Proben in einem Abstand von 20 cm ausgerichtet. Als Referenz wurde eine digitale Schieblehre mit einem Backenabstand von exakt zwei Millimetern mit auf dem Foto abgelichtet (s. Abbildung 4). Die erzeugten Fotos wurden in „Adobe Photoshop Version: 14.0“ importiert und durch Anhebung des Kontrasts und der Schärfe optimiert.

Für die Mikrocomputertomografie wurde das „µCT 40 von Scanco Medical“ der Zahnklinik München der Ludwig-Maximilians-Universität München verwendet. Die Schichtdicke betrug 18 µm. Die Aufnahmen wurden bei 70000 Volt und 114 µA durchgeführt. Als Modus wurde „Cone-beam“ gewählt.

Die digitalen Volumentomografien wurden mittels eines „Kodak Carestream 9300“, ebenfalls aus der Abteilung der Zahnerhaltung und Parodontologie der Zahnklinik München, durchgeführt. Dieses verfügt über ein Auflösungsvermögen von 90 µm und einem TFT Sensortyp. Die Aufnahmen wurden mit einer Spannung von 70 KV und einer Stromstärke von 80 mA bei kleinstem Sichtfeld (5 cm mal 5 cm) durchgeführt. Diese Werte ergaben die besten Ergebnisse bezüglich Kontrast und Schärfe. Die Proben wurden zentral positioniert, um ein möglichst genaues Ergebnis zu erzielen.

### 3.3 Vorbereitung der Messungen

Die bei der Mikrocomputertomografie und digitalen Volumentomografie entstandenen Rohdaten wurden zur Findung der korrekten Ebene in adäquate Programme importiert. Für die Festlegung der Messebene bei den  $\mu$ CT-Aufnahmen wurde ImageJ (ImageJ 1.48, Wayne Rasband, National Institute of Health, Bethesda, USA), unter Zuhilfenahme des von Prof. Dr. Karl-Heinz Kunzelmann entwickelten Plugins "KHKs Scanco microCT ISQ FileReader", verwendet. Die DVT-Aufnahmen wurden mit dem DICOM-Viewer „Osirix 4.1.2 32 bit der Firma Pixmeo SARL“ bearbeitet. Beide Programme stellten die Proben als dreidimensionale Körper dar, die nach Belieben als Schnittbilder dargestellt wurden. Die Schichten hatten eine Dicke von 90  $\mu$ m im DVT bzw. 18  $\mu$ m im  $\mu$ CT. Für diese Untersuchung war allerdings nur die oberste Schicht, also die Sägeebene, von Interesse.

Daher mussten die Proben durch Messung und Anpassung der Winkelverschiebungen so ausgerichtet werden, dass die Sägeebene senkrecht zur Y-Achse und parallel zur XZ-Ebene lag. Die XZ-Ebene entsprach der vom Programm dargestellten Schicht. Entlang der Y-Achse erfolgte dann eine Parallelverschiebung des Quaders bis zur XZ-Ebene. Dieser Vorgang ist in Abbildung 5 auf Seite 23 illustriert. Die gelben Quader stellen vereinfacht den Messkörper dar, wobei die hellere gelbe Fläche als Sägeeben zu interpretieren ist.

In Abbildung 5a liegt der gelbe Quader nicht senkrecht zur Y-Achse und die gelbe Oberfläche nicht parallel zur XZ-Ebene. Das Programm zeigte somit einen Schnitt quer durch den Zahn, und nicht die gewünschte Sägeebene. In Abbildung 5b hingegen haben die beschriebenen Rotationen und Verschiebungen bereits stattgefunden und die Oberfläche des Quaders ist senkrecht zur Y-Achse und parallel zur XZ-Ebene. Das Programm stellte nun die gewünschte Sägeebene dar.

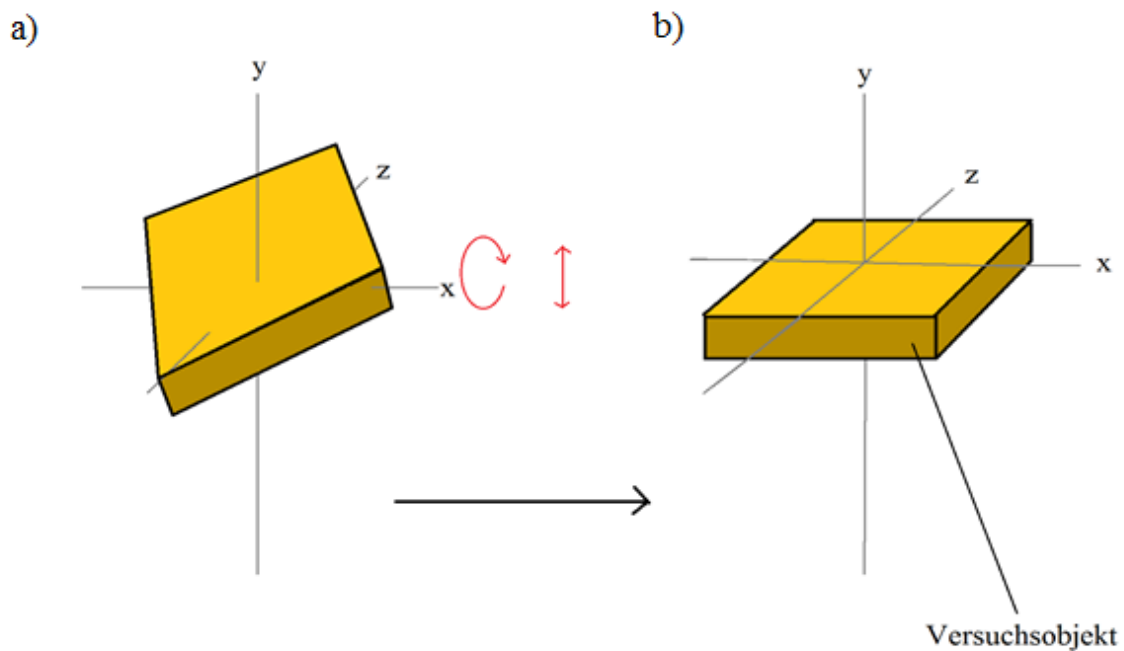
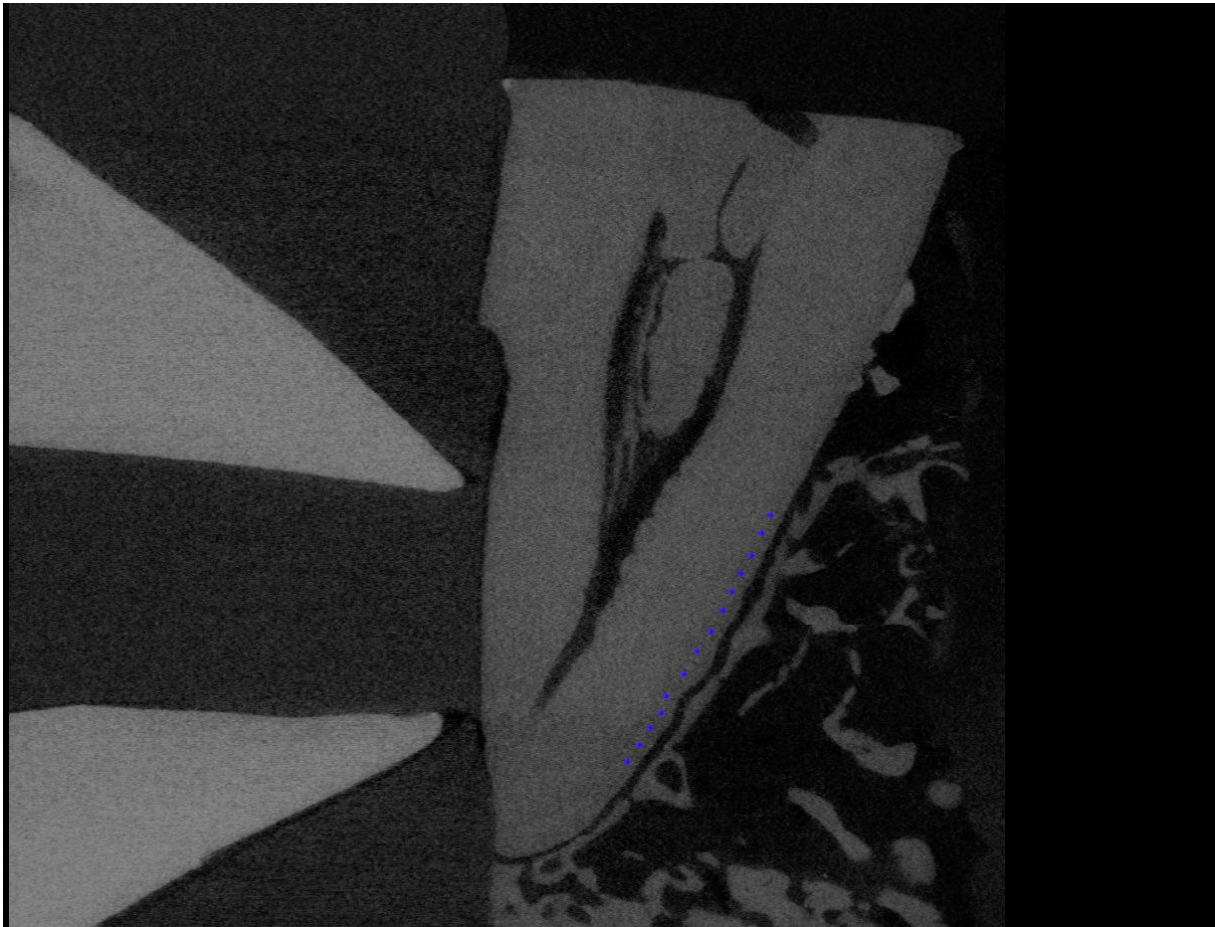


Abbildung 5: Schematische Illustration der Ausrichtung der Sägebene in ImageJ und Osirix.

Die entstandene Ansicht der Probe stellte das zu vermessende „Bild“ dar und wurde als Tiff-Datei exportiert. Dieses Prozedere musste für jede Probe sowohl für die DTV-Aufnahmen, als auch für die  $\mu$ CT-Aufnahmen durchgeführt werden.

Um vergleichbare Daten zu erhalten war es notwendig, die Messstellen bei jeder Bildgebung identisch zu wählen. Dazu wurde zunächst das entstandene Schnittbild der  $\mu$ CT-Aufnahme in Photoshop importiert. Entlang des Parodontalspalts wurden zehn gut zu messende Stellen von koronal nach apikal durch blaue Punkte definiert. Das entstandene, mit Messstellen versehene Schnittbild, in Abbildung 6 zu sehen, stellte das Referenzbild dar.



*Abbildung 6: Mit Messpunkten versehenes  $\mu$ CT-Schnittbild.*

Das in Abbildung 6 dargestellte  $\mu$ CT-Schnittbild zeigt einen sich im Knochen befindlichen Zahn. Links im Bild sind zwei Keramiksplitter zu erkennen. Entlang des Parodontalspalts sind mehrere blaue Punkte zu erkennen, die die späteren Messstellen markierten.



Die entsprechende Fotografie wurde anschließend mit dem Referenzbild überlagert. Durch Winkelmessungen und mit Hilfe der Keramiksplitter wurde die Fotografie rotiert und skaliert und somit die Überlagerung der beiden Bilder optimiert. Die Transparenz der Fotografie wurde kurzfristig erhöht, um die Messstellen des Referenzbildes, die als farbige Punkte hindurchschimmerten, auf die Fotografie zu übertragen. Abbildung 7 zeigt die beschriebene Überlagerung der  $\mu$ CT-Aufnahme mit der Fotografie. Die Transparenz der Fotografie liegt bei 50 %, sodass die blauen Messpunkte der  $\mu$ CT-Aufnahme erkennbar waren und auf die Fotografie übertragen werden konnten.



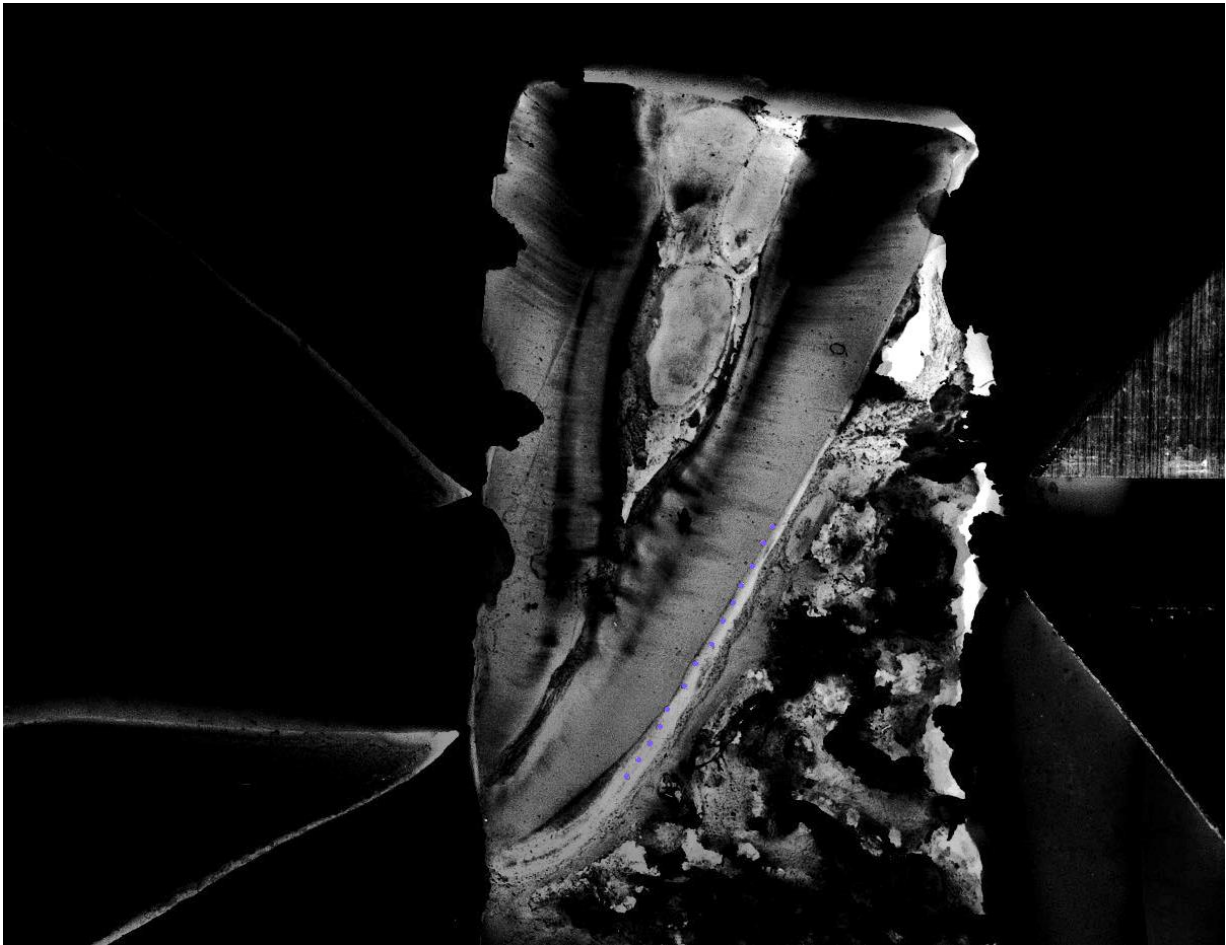
Abbildung 7: Überlagerung der Fotografie (Transparenz 50 %) mit dem Referenzbild.

Nach Zurücksetzen der Transparenz der überlagerten Fotografie auf 0, waren die Punkte deutlich auf der Fotografie zu erkennen und an denselben Stellen wie beim Referenzbild. Die Fotografie mit den übertragenen Messpunkten ist in zu sehen.



*Abbildung 8: Fotografie der Probe 4 mit Messpunkten.*

Die Darstellung des Parodontalspaltes in den Fotografien wurde durch Invertierung und Anpassung des Kontrasts, sowie durch die schwarz/weiß-Darstellung deutlich verbessert, wie die folgende Abbildung verdeutlicht.



*Abbildung 9: Invertierte Fotografie der Probe 4.*

Das Verfahren der Messpunktfestlegung wurde auf gleiche Art und Weise mit dem entsprechenden DVT-Schnittbild durchgeführt. Somit war sichergestellt, dass identische Stellen in allen drei Bildgebungen zur Vermessung verfügbar waren.

In Abbildung 10 wird vergrößert die apikale Region der Probe 4 in allen drei Bildgebungen vergleichend dargestellt. Abbildung 10a zeigt die Fotografie mit Beschriftung der anatomischen Strukturen. Darunter, in Abbildung 10b, wird dieselbe Region als  $\mu$ CT-Schnittbild und in Abbildung 10c im DVT-Schnittbild gezeigt. Die Unterschiede in Schärfe und Kontrast sind deutlich.

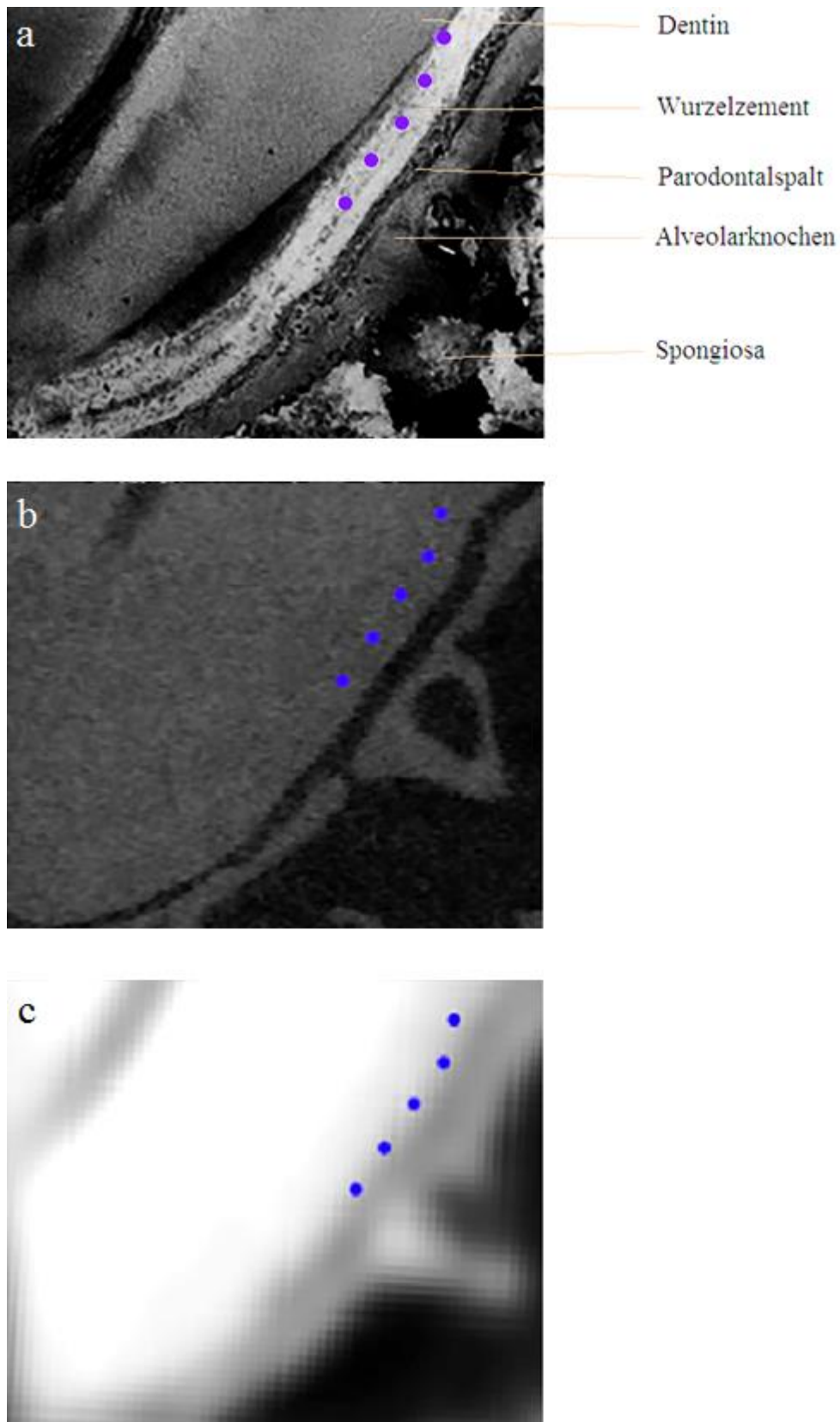


Abbildung 10: Nebeneinanderstellung der selben Strukturen im Vergleich zwischen Fotografie (a),  $\mu$ CT (b) und DVT (c).

Zur Vermessung des PDLs wurde das JAVA basierte Bildbearbeitungsprogramm „ImageJ“ (ImageJ 1.48, Wayne Rasband, National Institute of Health, Bethesda, USA) verwendet. Zunächst wurde das hochauflösende Foto einer Probe in ImageJ importiert. Beim Foto wurde eine Kalibrierung durchgeführt. Dafür wurde die mit auf dem Foto abgebildete digitale Schieblehre verwendet. Die Backen waren auf exakt zwei Millimeter eingestellt. Der Abstand der Backen auf dem Foto wurde gemessen und anschließend mit der in ImageJ enthaltenen Funktion „Set Scale“ mit dem tatsächlichen Maß von zwei Millimetern angepasst. Diese vorbereitende Maßnahme musste für jede Probe individuell wiederholt werden. Die Kalibrierung der  $\mu$ CT- und DVT-Aufnahmen erfolgte ebenfalls mit der „Set-Scale“-Funktion, indem die Pixelgröße von 18  $\mu$ m für  $\mu$ CT-Aufnahmen bzw. 90  $\mu$ m für DVT-Aufnahmen eingegeben wurde. Zur Kontrolle wurde die Kalibrierung überprüft, indem eine leicht zu messende Größe ( z. B. Zahngesamtlänge oder Keramiksplitterlänge) am Foto, DVT und am  $\mu$ CT gemessen und verglichen wurde. Nachdem sichergestellt war, dass die Kalibrierung erfolgreich war, konnten die Messungen beginnen. Zunächst wurde die zu messende Stelle optisch vergrößert, um das PDL möglichst genau erkennen zu können. Dann wurde die Breite des Parodontalspalts mithilfe der Linienfunktion vermessen und durch die Funktion „measure“ als Zahl in Mikrometer ausgegeben.

### **3.4 Durchführung der Messungen**

In dieser Studie haben drei radiologisch geschulte Personen die Messungen unabhängig voneinander durchgeführt. Die Messprozedur wurde den Personen erklärt, gezeigt und niedergeschrieben. Die Messungen erfolgten in zwei Messreihen. Jede Person hat jede Messstelle einer jeden Probe vermessen, und zwar Messstelle 1 bis 10 mit fünf Wiederholungen jeweils am Foto, DVT und  $\mu$ CT (MR1). Die gleiche Messreihe wurde nach zwei Monaten noch einmal wiederholt (MR2). (Eine Übersicht über die Verteilung der Messungen über die Beurteiler, Methoden und Proben kann der Abbildung 11 entnommen werden). Die ermittelten Werte wurden jeweils in eine für jeden Beurteiler eigene Exceltabelle eingefügt. Nach Abschluss der Messungen wurden die Tabellen zusammengeführt und eine Gesamttabelle erstellt.

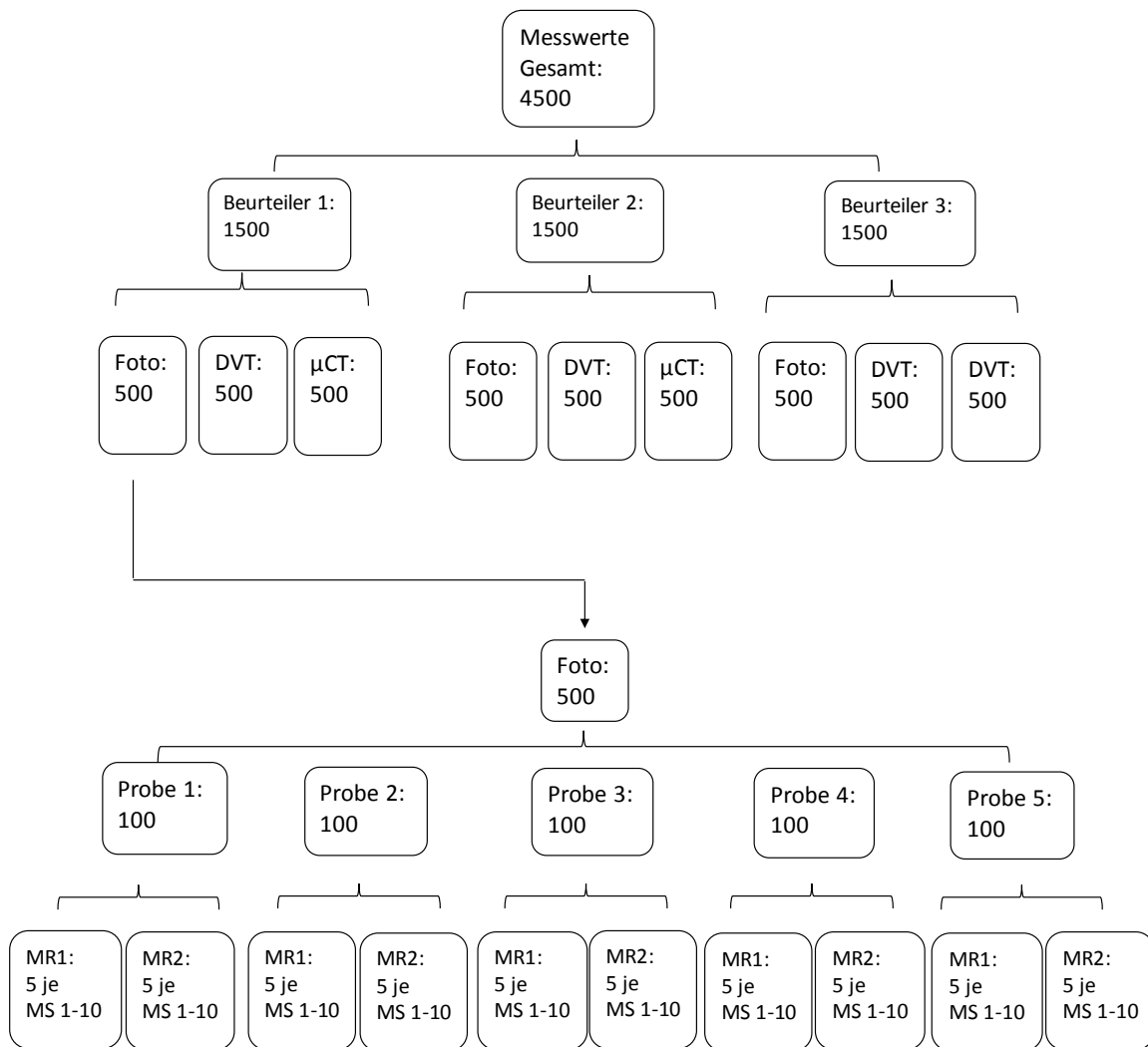


Abbildung 11: Übersicht über die Verteilung der Messwerte.

Abbildung 11 zeigt die Verteilung der Messwerte über Beurteiler, Probe und Methode. „MR1“ und „MR2“ beschreiben jeweils die Messreihen, in denen pro Probe die Messstellen eins bis zehn, fünfmal vermessen wurden. Der Abschnitt der Abbildung unterhalb des Pfeiles stellt einen Auszug aus der Gesamtaufteilung dar.

### 3.5 Statistische Verfahren

Die statistische Analyse kann in drei Teilbereiche unterteilt werden.

1. Intrabeurteilerdiskrepanz
2. Interbeurteilerdiskrepanz
3. Methodenvergleich

Die Intrabeurteilerdiskrepanz umfasst den Vergleich zwischen beiden Messreihen eines Beurteilers. Geprüft wurde, ob die Messwerte beider Messreihen aus einer Grundgesamtheit stammten. Sollte sich dies bestätigen, wurden beide Messreihen zusammengeführt.

Im zweiten Teilbereich, der Interbeurteilerdiskrepanz, wurden die zusammengeführten Messreihen der Beurteiler miteinander verglichen und wieder überprüft, ob auch diese aus einer Grundgesamtheit stammten und somit zusammengelegt werden konnten.

Teilbereich eins und zwei sind als vorbereitende Maßnahmen für den Methodenvergleich zu betrachten, bei dem die Messungen der Methoden miteinander verglichen wurden.

Eine Veranschaulichung des statistischen Verfahrens zeigt die Abbildung 12.

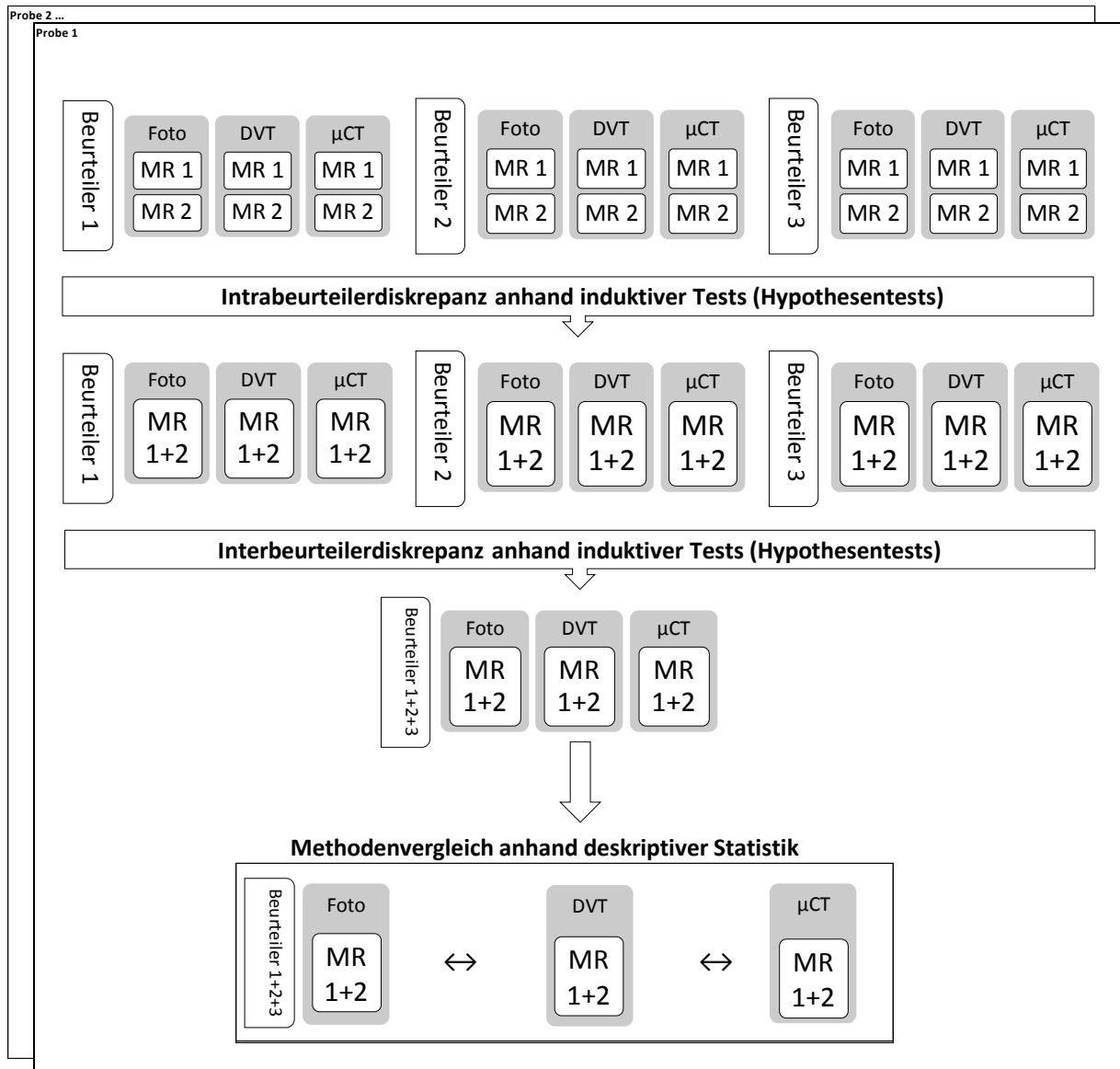


Abbildung 12: Statistisches Verfahren. „MR“ gibt die Messreihe an.

Abbildung 12 zeigt grafisch die Fusion der beiden Messreihen (MR) durch statistische Tests und die Vereinigung der Messreihen der einzelnen Beurteiler sowie die dazu verwendeten Verfahren bis hin zum Methodenvergleich.



### 3.5.1 Statistische Methoden und Tests

Dieses Kapitel beschreibt detailliert die verwendeten statistischen Methoden und Tests. Zusätzlich wird erläutert, in welchen Teilbereichen sie Anwendung fanden.

#### Deskriptive Statistik:

Die deskriptive Statistik wird tabellarisch und exemplarisch als Boxplots (Median, Maximum, Minimum, unteres/oberes Quartil und Whisker) und Balkendiagramme dargestellt. Sie umfasst, abhängig von dem jeweiligen Teilbereich, die Angabe der Probe und Messstelle (MS), die Anzahl der Messungen (N), den Mittelwert (Mittelw.), die Standardabweichung (S.D.), das Minimum (Min.), das Maximum (Max.), sowie den Median. Die deskriptive Statistik wurde für alle Teilbereiche errechnet.

#### Technischer Messfehler:

Zur Beurteilung des absoluten und relativen Fehlers wurde für alle Teilbereiche der „Technische Messfehler“ (engl. „Technical Error of Measurement“ (TEM)) inklusive des „relativen TEM“ (rTEM) errechnet.

Dieser spielt eine wichtige Rolle bei der Betrachtung der Genauigkeit der Messungen in der Anthropometrie (Lewis 1999). Weinberg und Scott stellten bereits 2005 heraus, dass die Intrabeurteilerdiskrepanz zu beträchtlichen statistischen Fehlern führen kann (Mueller. and Martorell. 1988; Weinberg et al. 2005).

Der TEM bietet ein gutes Maß, diesen Fehler zu beschreiben. Er zeigt die Qualität und Reproduzierbarkeit der Messungen und bietet nicht nur die Möglichkeit, die Messungen eines einzelnen Beurteilers der beiden Messreihen zu vergleichen, sondern auch den Vergleich zwischen verschiedenen Beurteilern sowie zwischen verschiedenen Methoden (Perini. et al. 2005).

Der TEM wird mathematisch wie folgt beschrieben:

$$TEM(\mu m) = \sqrt{\frac{\sum_{n=1}^N \left[ \left( \sum_{k=1}^K M_k(n)^2 \right) - \frac{\left( \sum_{k=1}^K M_k(n) \right)^2}{K} \right]}{N(K-1)}} \quad (1)$$

Hierbei entspricht „N“ der Anzahl der Messungen und „n“ der Messstelle. „K“ beschreibt die totale Anzahl an Beurteilern oder Methoden, je nachdem welcher Vergleich angestrebt ist, während „k“ den Beurteiler bzw. die Methode beschreibt.

Aus dem TEM lässt sich der relative TEM errechnen:

$$rTEM(\%) = \left( \frac{TEM}{\frac{1}{K} \sum_{k=1}^K \left( \frac{1}{N} \sum_{n=1}^N M_k(n) \right)} \right) \quad (2)$$

Die Berechnung erfolgte anhand einer solchen Beispieltabelle:

Tabelle 2: Rechentabelle des TEM und rTEM für die Intrabeurteilerdiskrepanz.

Beurteiler	Messreihe	Methode	Probe	MS1 ( $\mu m$ )	MS2 ( $\mu m$ )	MS3 ( $\mu m$ )	MS4 ( $\mu m$ )	MS5 ( $\mu m$ )	MS6 ( $\mu m$ )	MS7 ( $\mu m$ )	MS8 ( $\mu m$ )	MS9 ( $\mu m$ )	MS10 ( $\mu m$ )	Mittelw. ( $\mu m$ )	
1	1	1	1	164	161	148	150	124	146	108	117	181	179	147,8	
1	1	1	1	178	172	158	160	132	118	97	105	168	173	146,1	
1	1	1	1	162	161	152	151	129	145	104	110	174	198	148,6	
1	1	1	1	174	159	157	147	127	132	116	103	150	175	144	
1	1	1	1	165	151	151	155	123	133	112	111	159	159	141,9	
1	2	1	1	159	165	150	144	111	150	100	131	174	221	150,5	
1	2	1	1	183	177	148	142	144	120	99	110	159	250	153,2	
1	2	1	1	154	159	159	161	136	151	105	108	155	241	152,9	
1	2	1	1	183	174	155	148	142	136	112	130	171	221	157,2	
1	2	1	1	155	155	162	160	126	145	108	105	166	233	151,5	
SUM((M_k(n))^2)				282325	267644	237376	230860	168292	190600	112923	128594	275421	429632		
(SUM(M_k(n)))^2				2812329	2669956	2371600	2304324	1674436	1893376	1125721	1276900	2745649	4202500		
SUM(M^2)-SUM(M)^2/K				1092,1	648,4	216	427,6	848,4	1262,4	350,9	904	856,1	9382		
TEM( $\mu m$ )															13,3
rTEM(%)															8,9

Tabelle 2 zeigt die Berechnung des TEM und des rTEM der Probe 1 des Beurteilers 1 für die Intrabeurteilerdiskrepanz. Für den Interbeurteilervergleich sowie den Methodenvergleich wurden die Parameter entsprechend angepasst.

### Induktive Statistik:

Da der im Vorfeld angewandte Kolmogorow-Smirnow-Lilliefors-Test und der Shapiro-Wilk-Test ergaben, dass einige Variablen nicht einer Normalverteilung entstammten (Shapiro. and Wilk. 1965), kamen zwei nichtparametrische Tests zur Anwendung. Zum einen der Wilcoxon-Mann-Whitney-Test für die Intrabeurteilerdiskrepanz und zum anderen der Kruskal-Wallis-Test für die Interbeurteilerdiskrepanz und den Methodenvergleich.

Beide Tests bedienen sich des Verfahrens der Rangierung der Messwerte und ermöglichen die Überprüfung von vorgegebenen Nullhypothesen (Hypothesentests). Der Kruskal-Wallis-Test bietet den Vorteil, dass mehr als zwei Gruppen von Messwerten auf Signifikanz ( $p$ ) überprüft werden können (William. et al. 1952).

Das Signifikanzniveau wurde bei beiden Tests mit  $\alpha = 0,05$  festgelegt. Hieraus ergaben sich folgende Signifikanzlevels:

- $p > 0,05$ : Nullhypothese wird beibehalten
- $p < 0,05$ : Nullhypothese wird verworfen

Die Nullhypothesen wurden abhängig von dem entsprechenden Teilbereich definiert und überprüft.

### **3.5.2 Intrabeurteilerdiskrepanz**

Ziel der Intra- und Interbeurteilerdiskrepanz war es, die Messstellen zu isolieren, die entweder von einem Beurteiler zu zwei Zeitpunkten signifikant unterschiedlich vermessen wurden, oder diese, die alle drei Beurteiler different vermessen haben. Diese Messstellen wurden als fehlerbehaftet beurteilt und somit in die deskriptive Statistik des Methodenvergleichs nicht mit einbezogen.

Die Intrabeurteilerdiskrepanz beschreibt die Messdiskrepanz eines Beurteilers zwischen den beiden Messreihen. Jeder Beurteiler führte zwei Messreihen durch (MR1 und MR2). Die Anzahl an Messungen pro Messstelle einer Messreihe war  $N = 5$  (s. Abbildung 12, S. 32).

Die Intrabeurteilerdiskrepanz setzte sich aus drei Verfahren zusammen, die sich alle der gesamten Datenbasis bedienen (s. Abbildung 13).

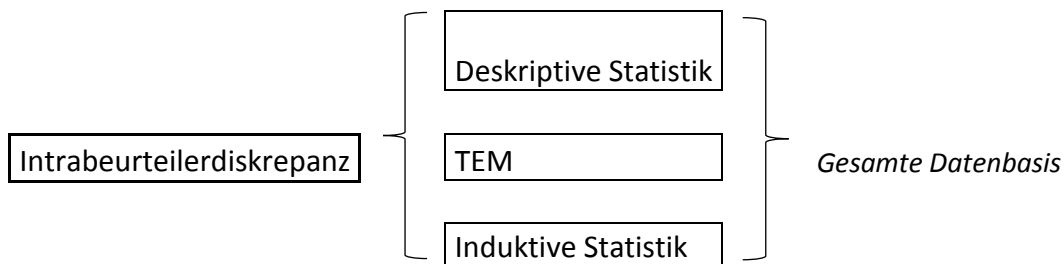


Abbildung 13: Statistischer Aufbau der Intrabeurteilerdiskrepanz.

Nach Erstellen einer deskriptiven Statistik mit Mittelwert und Standardabweichung und der Berechnung der technischen Messfehler (TEM und rTEM), wurde mit dem Wilcoxon-Mann-Whitney-Test die Nullhypothese überprüft, ob die fünf Messwerte der MR1 und die äquivalenten fünf der MR2 eines Beurteilers einer Grundgesamtheit entstammten.

Das Signifikanzniveau wurde auf  $\alpha = 0,05$  festgesetzt.

Sollte die Nullhypothese bestätigt werden (Signifikanz  $p > 0,05$ ), wurden MR1 und MR2 zusammengeführt und es ergab sich die „Messreihe Gesamt“ mit  $N = 10$ .

Ergab der Test, dass sich die Werte der beiden Messreihen signifikant unterschieden (Signifikanz  $p < 0,05$ ), wurde diese Messstelle für einen Interbeurteiler- und Methodenvergleich nicht mehr in Betracht gezogen.

### 3.5.3 Interbeurteilerdiskrepanz

Die Interbeurteilerdiskrepanz beschreibt den Vergleich zwischen den „Messreihen Gesamt“ der drei Beurteilern. Dieser Vergleich setzte sich wiederum aus drei Verfahren mit unterschiedlichen Datenbasen zusammen (s. Abbildung 14):

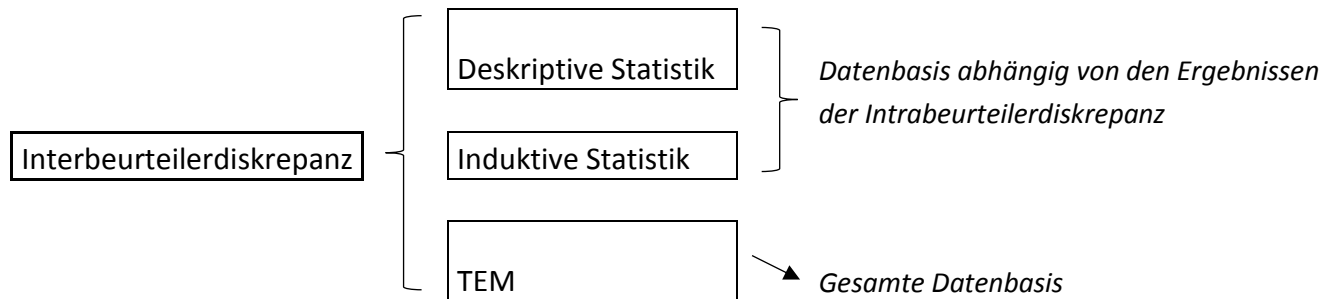


Abbildung 14: Statistischer Aufbau der Interbeurteilerdiskrepanz.

Die deskriptive Statistik umfasst Mittelwerte und Standardabweichungen und wird auszugsweise im Ergebnisteil dargestellt. TEM und rTEM wurden zur Berechnung des Fehlers zwischen den Werten der Messreihen aller drei Beurteiler herangezogen.

Die induktive Statistik wurde unter Anwendung des Kruskal-Wallis-Tests mit einem Signifikanzniveau von  $\alpha = 0,05$  durchgeführt. Die Messreihen der Beurteiler pro Messstelle, mit der Anzahl an zu vergleichenden Werten pro Messreihe eines Beurteilers  $N = 10$  („Messreihe Gesamt“), wurden auf Signifikanzen überprüft. Die Nullhypothese wurde aufgestellt, ob die Messreihen der drei Beurteiler einer Grundgesamtheit entstammten.

Es ergaben sich drei Ergebnisvarianten:

Drei Messreihen gehören einer Grundgesamtheit an:	$N = 30$
Zwei Messreihen gehören einer Grundgesamtheit an:	$N = 20$
Keine gemeinsame Grundgesamtheit:	$N = 0$

Sollten alle drei Messreihen einer Grundgesamtheit entstammen, so wurde die Anzahl an Messwerten für die geprüfte Messreihe auf  $N = 30$  erhöht, bei zwei Messreihen entsprechend auf  $N = 20$ . Sollte keine Grundgesamtheit bestehen, entfiel diese Messstelle für die deskriptive und induktive Statistik des Methodenvergleichs.

### 3.5.4 Methodenvergleich

Der Methodenvergleich erfolgte mittels zweier Datenbasen. Die Datenbasis für die deskriptive und induktive Statistik beruhte auf den Ergebnissen der Intra- und Interbeurteilerdiskrepanz. Es wurden die Werte einer Messstelle mit der Anzahl  $N = 30$  oder  $N = 20$  der drei Methoden anhand von Mittelwerten, Standardabweichungen, Minimum, Maximum und Median verglichen. Mit dem Kruskal-Wallis-Test wurde die Nullhypothese auf Signifikanzen getestet, ob die Messwerte der drei Methoden einer Grundgesamtheit entstammten.

TEM und rTEM wurden unabhängig von der Intra- und Interbeurteilerdiskrepanz durchgeführt und umfassten alle Messstellen mit  $N = 30$ .

Ebenfalls unabhängig von dem Intra- und Interbeurteilervergleich erfolgte ein Vergleich der Methoden anhand Mittelwerten der Messwerte pro „Messreihe Gesamt“ mittels Punktwolken mit Trendlinien und Bestimmtheitsmaßen.

Eine Übersicht über die statistischen Verfahren und deren Datenbasen bietet Abbildung 15:

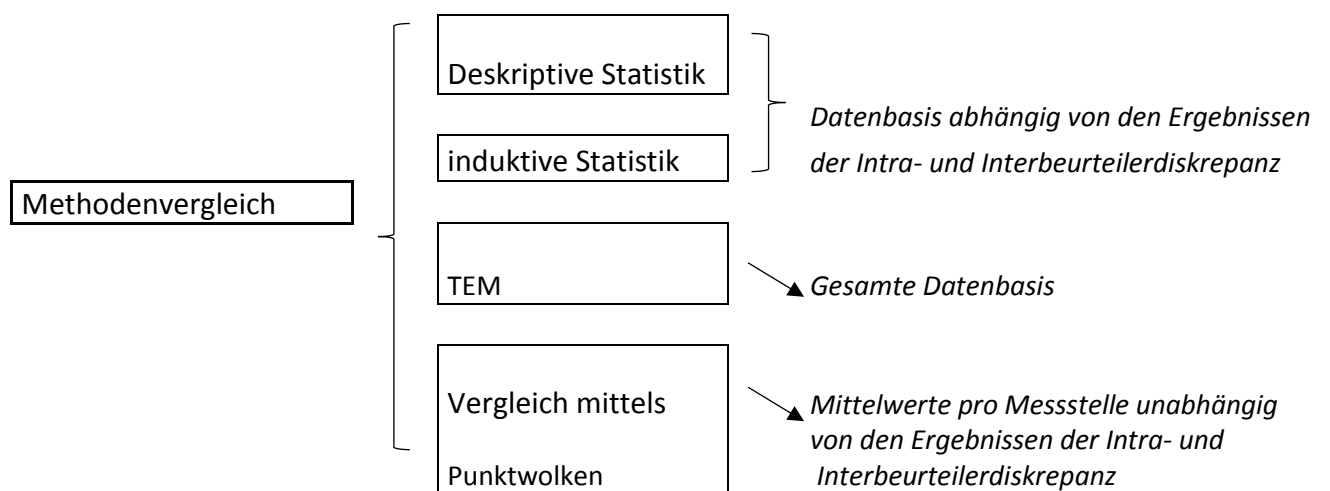


Abbildung 15: Statistischer Aufbau des Methodenvergleichs.

## **4 Ergebnisse**

Das vierte Kapitel beschreibt die Ergebnisse der Intrabeurteilerdiskrepanz, der Interbeurteilerdiskrepanz sowie des Methodenvergleichs.

### **4.1 Intrabeurteilerdiskrepanz**

Dargestellt und beschrieben werden im Folgenden die Ergebnisse der Intrabeurteilerdiskrepanz zunächst anhand der deskriptiven Statistik, gefolgt von den Ergebnissen des technischen Messfehlers (TEM) und des induktiven Tests.

#### **4.1.1 Ergebnisse der deskriptiven Statistik**

Tabelle 3 zeigt tabellarisch die Ergebnisse der deskriptiven Statistik der Intrabeurteilerdiskrepanz. Verglichen werden jeweils die Messungen von Messreihe eins mit denen von Messreihe zwei. Exemplarisch werden drei Balkendiagramme ausgewählter Messstellen gezeigt und erläutert. Die X-Achse gibt dabei die Messstellen an, während auf der Y-Achse die PDL-Breite in  $\mu\text{m}$  aufgetragen ist.

Tabelle 3: Deskriptive Statistik der Intrabeurteilerdiskrepanz.

Probe	MS	Beurteiler	Foto				DVT				μCT			
			Messreihe 1		Messreihe 2		Messreihe 1		Messreihe 2		Messreihe 1		Messreihe 2	
			Mittelw.	S.D.	Mittelw.	S.D.	Mittelw.	S.D.	Mittelw.	S.D.	Mittelw.	S.D.	Mittelw.	S.D.
			(μm)	(μm)	(μm)	(μm)	(μm)	(μm)	(μm)	(μm)	(μm)	(μm)	(μm)	(μm)
1	1	1	169	7,0	167	14,9	143	10,4	153	6,8	192	8,6	189	9,0
		2	173	8,2	175	8,6	145	11,9	152	6,0	206	11,6	172	9,7
		3	173	8,6	169	3,7	151	4,6	152	6,3	197	7,6	195	17,1
	2	1	161	7,5	166	9,4	166	5,4	119	11,9	152	13,0	149	12,6
		2	164	6,3	162	7,7	170	10,9	151	4,9	157	7,6	157	4,7
		3	162	11,6	161	11,7	169	8,9	153	5,6	164	5,8	168	7,1
	3	1	153	4,2	155	5,9	164	7,6	140	7,0	156	7,0	175	3,6
		2	154	7,7	162	3,4	172	10,8	160	5,5	160	6,4	160	2,7
		3	153	12,8	152	11,0	161	12,3	159	5,9	169	2,5	170	5,6
	4	1	153	5,0	151	8,9	142	5,4	128	15,1	115	14,3	122	5,8
		2	153	4,6	152	4,6	145	5,1	141	5,5	133	5,8	122	4,0
		3	161	7,9	163	5,5	142	5,0	146	6,8	143	2,9	148	5,0
	5	1	127	3,7	132	13,6	168	8,3	97	7,4	127	11,9	120	11,7
		2	127	4,3	140	5,7	161	8,0	168	5,0	134	5,4	121	6,7
		3	132	5,0	134	6,2	174	8,9	171	6,8	148	3,5	149	3,1
	6	1	135	11,4	140	12,9	131	10,5	122	9,7	146	9,3	132	7,8
		2	138	5,9	148	7,4	162	7,8	149	3,3	154	7,4	141	11,3
		3	148	8,4	149	5,8	133	13,6	153	6,1	143	9,2	145	5,9
	7	1	107	7,3	105	5,4	121	8,3	95	6,5	115	11,9	119	11,6
		2	107	9,4	106	10,6	164	6,3	152	6,8	133	5,4	128	4,8
		3	100	9,8	101	5,9	129	6,7	150	3,9	121	7,4	121	4,5
	8	1	109	5,5	117	12,6	156	7,1	127	7,5	85	8,6	67	10,7
		2	92	6,2	94	3,7	150	7,6	153	5,9	77	4,9	86	7,1
		3	111	10,9	112	10,1	159	9,6	152	8,0	92	5,8	95	5,0
	9	1	166	12,2	165	8,0	138	6,5	98	3,6	138	11,3	124	8,3
		2	161	9,2	171	10,3	154	6,0	123	6,1	139	4,1	112	6,6
		3	151	10,0	151	9,0	144	4,7	119	6,3	134	5,1	131	2,3
	10	1	177	14,0	233	12,7	219	26,2	210	28,7	289	9,9	268	11,3
		2	181	6,2	236	11,7	243	18,6	226	19,7	287	14,4	247	42,8
		3	190	17,4	198	12,1	228	23,4	251	13,4	299	8,1	304	7,2
2	1	1	196	4,2	189	12,8	185	10,5	182	8,2	175	14,8	179	4,3
		2	193	4,6	188	4,9	178	10,5	178	8,5	187	5,3	186	7,2
		3	206	10,5	206	6,0	202	16,8	201	12,1	185	7,2	180	5,1
	2	1	206	10,2	188	13,1	195	15,7	191	7,1	181	8,7	166	11,4
		2	197	15,6	190	16,8	194	9,2	189	11,9	185	7,2	173	5,5
		3	194	8,4	187	5,8	207	13,8	209	8,3	180	6,3	180	7,4
	3	1	176	15,2	197	17,2	197	12,1	197	12,9	179	7,3	157	7,8
		2	175	27,7	178	24,8	182	6,1	191	10,2	159	7,2	159	5,1
		3	191	9,8	190	8,4	237	12,5	231	14,3	208	16,1	202	9,8
	4	1	168	15,4	161	3,6	172	18,5	160	10,1	176	9,6	174	12,9
		2	164	8,9	169	6,9	184	9,7	169	8,3	180	8,0	174	5,0
		3	205	9,0	203	6,8	211	9,4	213	20,3	192	21,2	195	8,7
	5	1	139	9,5	155	12,9	180	11,3	216	29,7	132	3,5	140	3,7
		2	135	13,7	135	8,5	155	7,8	172	12,6	145	5,8	144	6,5
		3	131	4,4	131	5,0	238	9,1	234	11,4	191	10,2	190	8,6
	6	1	172	14,3	170	15,8	195	8,8	150	15,6	186	10,8	193	9,3
		2	183	6,2	171	7,9	206	6,5	156	11,4	207	4,5	199	9,7
		3	150	9,3	152	6,5	231	9,7	230	12,5	218	7,4	212	6,9
	7	1	94	4,3	110	14,1	129	13,8	158	6,4	122	9,1	152	7,3
		2	107	27,1	119	6,6	140	8,6	147	8,5	135	4,3	143	4,0
		3	138	5,0	141	5,1	219	9,0	217	9,6	158	9,3	160	8,6



Ergebnisse

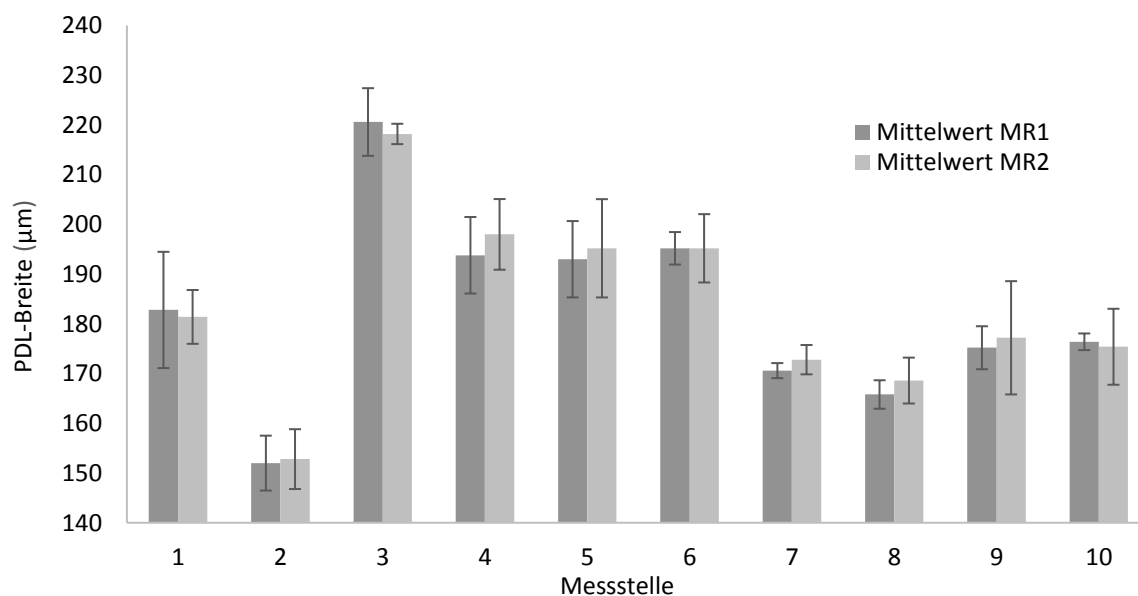
		8	1	201	7,4	200	12,7	158	22,0	190	16,6	205	5,1	217	10,6
			2	194	10,9	194	6,7	210	7,9	206	5,3	210	6,8	204	12,8
			3	216	7,4	209	6,3	230	7,6	229	8,9	230	14,5	215	4,6
		9	1	178	8,1	171	16,7	193	9,2	202	10,8	198	7,0	193	8,8
			2	185	9,9	176	12,7	189	7,5	190	6,8	180	7,2	185	8,5
			3	210	6,0	206	5,1	244	13,0	247	19,7	211	10,1	210	6,6
		10	1	169	5,8	174	7,3	213	6,8	251	9,8	179	7,3	197	10,4
			2	174	4,1	179	8,2	173	8,1	177	8,3	170	9,1	178	12,1
			3	305	22,4	301	5,3	284	11,2	285	12,3	188	9,3	193	7,7
	3	1	1	207	12,2	214	10,0	152	4,0	144	7,0	153	17,6	156	8,6
			2	215	9,8	214	8,9	175	10,5	156	7,9	155	4,3	159	8,8
			3	227	3,3	229	5,8	194	8,8	197	10,2	151	7,0	149	6,7
		2	1	295	20,6	293	8,3	215	18,4	198	21,5	230	12,0	244	16,1
			2	302	5,9	293	8,1	234	9,7	201	13,4	223	15,6	236	11,1
			3	303	7,0	303	6,3	252	6,7	257	7,5	237	5,6	246	4,7
		3	1	212	8,8	205	10,0	183	38,4	171	17,5	201	9,9	204	7,7
			2	222	9,4	212	12,8	168	5,6	223	5,5	202	13,2	181	12,0
			3	216	8,7	208	7,4	172	4,9	175	9,9	194	10,1	203	4,7
		4	1	170	5,7	142	6,9	167	14,3	167	7,9	183	9,9	179	12,5
			2	149	3,8	150	5,0	167	7,2	155	6,3	187	5,5	176	13,1
			3	142	3,9	144	11,6	177	7,8	169	8,4	186	5,6	186	4,0
		5	1	129	17,3	135	19,3	141	8,8	137	6,6	141	6,8	148	7,1
			2	122	4,9	132	5,4	144	3,9	130	4,0	141	5,7	143	6,1
			3	123	4,0	121	6,2	153	8,6	157	10,7	139	4,8	140	4,6
		6	1	200	18,7	191	14,0	109	7,2	149	7,5	124	7,9	127	12,5
			2	192	4,5	194	5,8	120	6,7	129	8,6	123	8,4	117	8,3
			3	195	4,1	194	5,9	143	5,5	142	6,7	123	5,7	119	4,7
		7	1	122	11,9	131	5,3	97	4,3	111	6,6	108	8,1	126	6,5
			2	158	2,8	158	5,3	104	7,3	109	7,7	103	4,9	116	12,2
			3	115	3,8	111	3,6	113	8,5	106	5,5	114	4,1	114	5,5
		8	1	121	8,8	107	3,9	76	4,4	99	10,2	97	7,2	107	7,8
			2	105	4,7	103	4,2	75	9,2	89	8,0	114	3,8	119	4,6
			3	99	4,8	99	2,9	94	4,5	95	1,9	98	3,6	100	5,5
		9	1	62	23,7	47	1,6	92	9,3	66	6,0	82	7,4	78	5,3
			2	45	3,5	47	3,2	53	5,3	45	4,2	76	7,0	77	8,0
			3	56	3,2	58	3,5	104	10,6	106	10,7	88	3,6	87	5,2
		10	1	48	4,8	41	2,4	38	6,2	37	4,8	69	8,1	70	8,6
			2	45	3,4	45	4,2	51	2,5	42	3,7	71	4,0	70	5,9
			3	48	2,3	47	3,1	61	9,7	44	15,4	78	1,9	82	3,8
	4	1	1	193	11,9	203	8,3	214	20,4	244	17,0	198	9,3	192	12,2
			2	190	10,1	191	2,5	211	15,0	227	10,6	192	7,9	191	5,1
			3	183	11,7	181	5,4	207	11,1	208	5,7	187	8,3	198	10,9
		2	1	148	9,8	150	9,3	333	11,0	373	32,2	255	8,8	268	23,0
			2	151	7,9	152	4,7	330	12,6	321	9,9	255	8,4	252	6,5
			3	152	5,5	153	6,0	250	9,0	238	14,3	245	6,4	246	7,1
		3	1	212	9,2	188	18,3	236	22,9	271	23,6	249	13,0	301	13,4
			2	208	10,5	195	11,5	248	55,3	238	8,6	251	14,0	267	11,5
			3	221	6,8	218	2,0	272	16,0	283	23,6	262	5,3	258	6,5
		4	1	185	6,8	188	10,3	205	16,5	217	19,0	184	9,2	176	11,5
			2	188	7,7	192	12,4	207	13,4	209	9,9	182	8,5	184	4,8
			3	194	7,7	198	7,1	182	14,2	182	6,0	191	5,1	195	6,7
		5	1	187	11,0	178	9,9	180	10,5	187	16,4	178	8,9	166	12,3
			2	186	10,1	186	8,5	177	10,2	190	9,6	176	9,2	177	5,2
			3	193	7,7	195	9,9	188	9,4	182	6,5	186	2,8	187	3,8
		6	1	235	14,8	244	9,8	278	12,4	235	17,0	248	10,9	258	9,9
			2	239	14,7	238	9,9	278	14,2	250	11,7	242	12,3	246	10,1
			3	195	3,3	195	6,9	257	24,9	252	6,5	246	9,3	242	12,8
		7	1	156	10,1	164	7,1	188	15,3	188	9,6	147	12,4	144	9,3
			2	159	10,8	164	5,5	195	12,5	187	8,5	150	10,6	146	6,0
			3	171	1,5	173	2,9	197	11,8	190	11,6	148	9,0	155	9,4
		8	1	166	16,8	171	13,3	206	13,0	199	2,7	180	14,6	183	5,1
			2	162	5,7	164	5,8	204	8,0	204	7,8	178	12,8	192	10,7
			3	166	2,9	169	4,6	203	10,6	203	11,9	182	8,7	187	6,8

9	1	168	8,6	160	6,5	182	13,4	188	21,2	205	9,1	203	19,0
	2	172	7,8	167	6,9	178	6,5	179	3,8	200	7,7	198	6,1
	3	175	4,3	177	11,4	180	14,9	176	12,8	198	5,2	196	6,8
10	1	191	12,5	221	31,5	205	11,8	210	14,3	185	8,3	186	13,2
	2	191	13,1	201	9,9	206	13,3	211	10,2	184	13,1	185	12,6
	3	176	1,7	175	7,6	203	12,8	200	9,8	181	4,6	179	4,5
5	1	97	5,8	89	8,4	197	16,1	164	16,1	144	9,6	161	10,1
	2	97	9,4	97	8,2	186	8,1	183	6,4	148	5,0	154	5,6
	3	74	9,4	77	7,3	242	19,2	220	48,2	163	7,9	166	5,2
2	1	143	1,6	88	9,3	373	23,3	192	12,8	224	6,9	194	33,2
	2	145	3,8	83	6,2	174	8,0	180	9,7	172	6,5	169	6,0
	3	134	9,5	135	3,7	370	7,7	373	12,1	193	7,8	196	6,2
3	1	100	9,8	100	8,3	185	23,8	160	19,4	145	10,0	151	8,1
	2	97	8,6	96	7,1	176	12,4	165	4,3	145	5,9	148	7,2
	3	75	7,8	77	5,9	252	7,0	259	8,9	141	10,8	142	5,2
4	1	102	4,9	92	6,7	171	35,6	172	19,8	181	9,6	171	9,7
	2	101	5,4	92	10,7	166	28,2	119	53,9	172	8,0	169	4,5
	3	96	7,5	96	5,1	220	6,1	228	2,7	207	5,6	207	2,3
5	1	71	3,6	86	7,5	172	12,7	155	4,7	169	7,4	145	6,4
	2	71	3,0	86	7,7	168	7,8	168	6,6	74	3,9	144	5,2
	3	71	2,6	73	2,9	167	17,4	170	16,4	163	3,1	166	2,3
6	1	81	7,5	70	4,0	193	19,1	201	26,6	125	9,9	125	8,6
	2	71	3,0	71	2,9	182	5,9	181	7,9	65	4,1	130	8,4
	3	70	2,3	73	3,1	177	14,8	190	10,9	112	5,9	114	3,6
7	1	94	4,0	111	8,9	214	8,4	236	26,7	93	9,3	83	13,3
	2	101	9,1	94	10,4	179	9,0	184	4,2	74	4,2	82	5,2
	3	149	9,2	154	8,4	201	11,5	199	14,6	92	5,2	88	2,4
8	1	133	4,6	118	4,9	181	21,2	260	22,2	118	9,3	124	19,6
	2	112	7,6	118	6,2	178	19,8	190	18,5	121	8,0	125	7,1
	3	116	4,4	110	5,9	186	15,6	189	15,6	118	4,7	118	4,0
9	1	62	8,4	88	5,1	121	7,2	162	13,2	118	8,2	121	7,4
	2	66	4,2	75	9,6	151	7,2	164	5,2	137	3,4	151	4,5
	3	92	5,1	94	2,9	160	4,9	162	6,1	132	4,5	129	3,9
10	1	72	16,6	100	9,3	123	12,9	179	15,7	125	10,8	149	13,9
	2	76	8,3	99	10,5	134	4,8	158	6,9	136	8,7	141	5,7
	3	96	10,3	91	3,3	154	7,6	158	5,7	128	5,5	131	2,2

Detailliert dargestellt wird der Vergleich der Messreihen der Probe 4 des Beurteilers 3 am Foto anhand von Mittelwerten und Standardabweichungen. Tabelle 4 und Balkendiagramm 1 zeigen, dass die Mittelwerte sich höchstens um 4,2  $\mu\text{m}$  unterscheiden (MS 4). Die maximale Standardabweichung liegt bei 11,69  $\mu\text{m}$  (MS 1).

Tabelle 4: Intrabeurteilerdiskrepanz der Vermessung des PDLs, Probe 4 des Beurteilers 3 am Foto.

MS	MR1		MR2	
	Mittelw. ( $\mu\text{m}$ )	S.D. ( $\mu\text{m}$ )	Mittelw. ( $\mu\text{m}$ )	S.D. ( $\mu\text{m}$ )
1	182,8	11,69	181,4	5,41
2	152	5,52	152,8	6,02
3	220,6	6,80	218,2	2,05
4	193,8	7,69	198	7,11
5	193	7,68	195,2	9,88
6	195,2	3,27	195,2	6,87
7	170,6	1,52	172,8	2,95
8	165,8	2,86	168,6	4,62
9	175,2	4,32	177,2	11,39
10	176,4	1,67	175,4	7,64

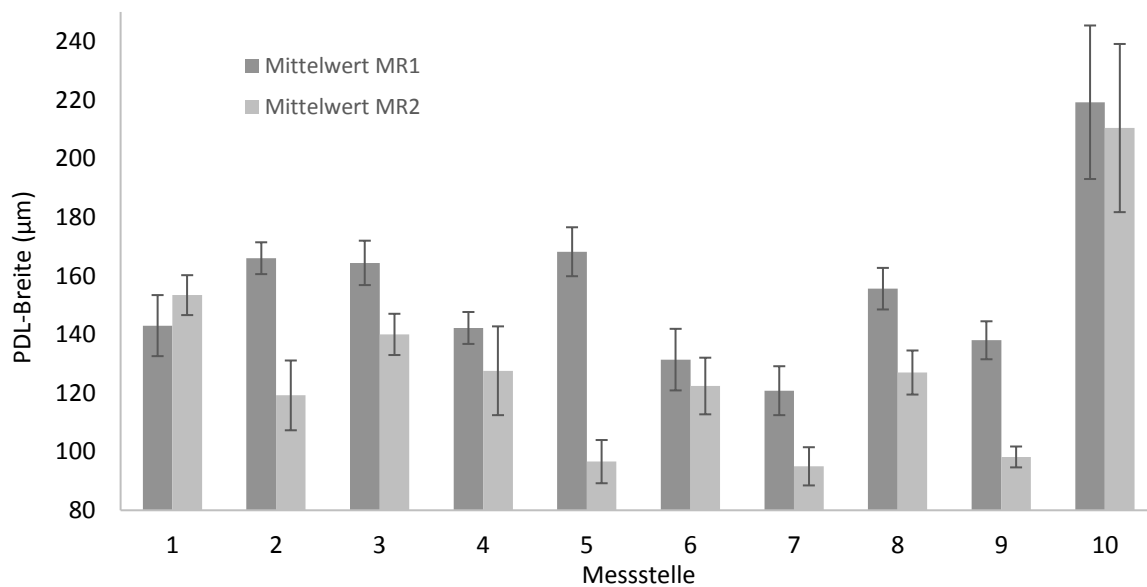


Balkendiagramm 1: Balkendiagramm der Messungen des PDLs der Probe 4 des Beurteilers 3 am Foto.

Die Probe 1, vermessen von Beurteiler 1 am DVT, weist große Differenzen zwischen den beiden Messreihen auf. Besonders die Messstelle 5 ergab einen Unterschied von 71,6  $\mu\text{m}$ . Die Standardabweichungen lagen, abhängig von der Messstelle, bei minimal 3,56  $\mu\text{m}$  (MS 9) bis maximal 28,68  $\mu\text{m}$  (MS 10). Siehe Tabelle 5 und Balkendiagramm 2.

Tabelle 5: Intrabeurteilerdiskrepanz der Vermessung des PDLs, Probe 1 des Beurteilers 1 am DVT.

MS	MR1		MR2	
	Mittelw. ( $\mu\text{m}$ )	S.D. ( $\mu\text{m}$ )	Mittelw. ( $\mu\text{m}$ )	S.D. ( $\mu\text{m}$ )
1	143	10,42	153,4	6,80
2	166	5,43	119,2	11,90
3	164,4	7,57	140	7,04
4	142,2	5,45	127,6	15,14
5	168,2	8,35	96,6	7,37
6	131,4	10,50	122,4	9,66
7	120,8	8,35	95	6,52
8	155,6	7,09	127	7,52
9	138	6,48	98,2	3,56
10	219,2	26,19	210,4	28,68

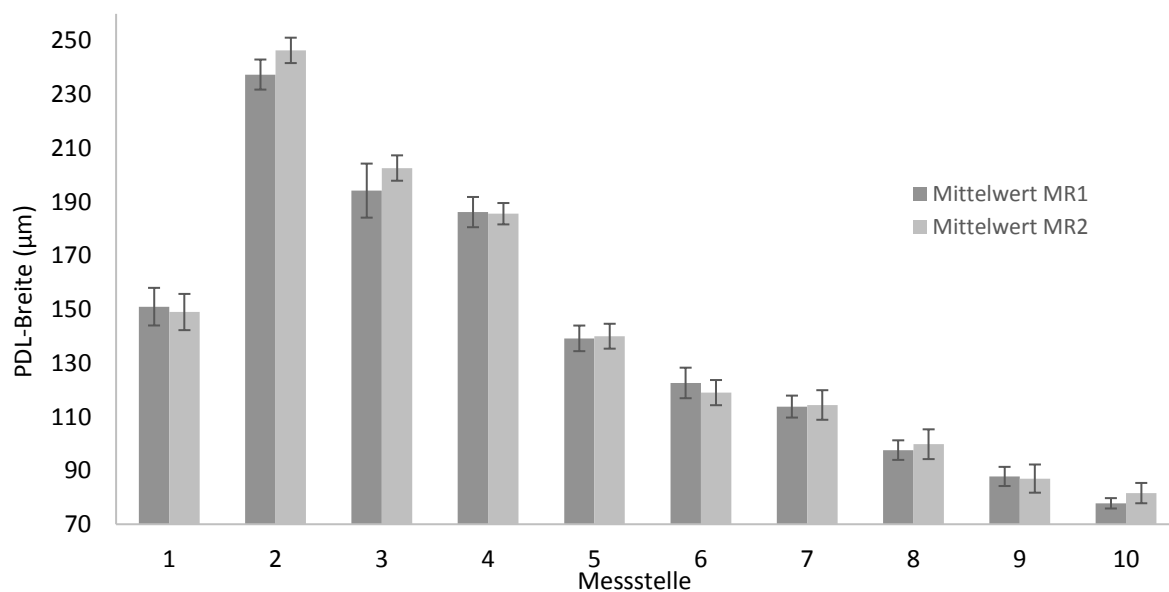


Balkendiagramm 2: Balkendiagramm der Messungen des PDLs der Probe 1 des Beurteilers 1 am DVT.

Die Tabelle 6 und das Balkendiagramm 3 beschreiben den Vergleich der Messreihen der Probe 3 des Beurteilers 3 am  $\mu$ CT. Die maximale Differenz zwischen beiden Messreihen wurde bei der Messstelle 2 mit 9,0  $\mu$ m, die niedrigste Standardabweichung bei der Messstelle 10 mit 1,92  $\mu$ m festgestellt.

Tabelle 6: Intrabeurteilerdiskrepanz der Vermessung des PDLs, Probe 3 des Beurteilers 3 am  $\mu$ CT.

MS	MR1		MR2	
	Mittelw.	S.D.	Mittelw.	S.D.
	( $\mu$ m)	( $\mu$ m)	( $\mu$ m)	( $\mu$ m)
1	151	7,00	149	6,75
2	237,4	5,59	246,4	4,72
3	194,2	10,06	202,6	4,72
4	186,2	5,63	185,6	3,97
5	139,2	4,76	140	4,64
6	122,6	5,68	119	4,69
7	113,8	4,09	114,4	5,50
8	97,6	3,65	99,8	5,54
9	87,8	3,56	87	5,24
10	77,8	1,92	81,6	3,78



Balkendiagramm 3: Balkendiagramm der Messungen des PDLs der Probe 3 des Beurteilers 3 am  $\mu$ CT.

#### 4.1.2 Ergebnisse des Technical Error of Measurement

Einen Überblick über die Ergebnisse der Berechnungen des technischen Messfehlers bietet die Tabelle 7. Gelb hinterlegt sind alle Werte des TEM größer oder genau 15  $\mu\text{m}$ .

Die Berechnungen des TEM, ausgehend von den Werten, die bei der Fotografie gemessen wurden, ergaben einen minimalen Wert von 6  $\mu\text{m}$  (Probe 4 und 5 des Beurteilers 3). Der größte Fehler wurde mit 14  $\mu\text{m}$  ebenfalls bei diesen Proben, aber von Beurteiler 1 erzeugt.

Der TEM für die Messung am DVT ergab Werte von minimal 9  $\mu\text{m}$  (Probe 3, Beurteiler 3) bis maximal 40  $\mu\text{m}$  (Probe 5, Beurteiler 1).

Die Berechnung des TEM mit den am  $\mu\text{CT}$  gemessenen Werten ergab einen minimalen TEM von 5  $\mu\text{m}$  (Probe 5, Beurteiler 3) und einen maximalen Wert von 17  $\mu\text{m}$  (Probe 5, Beurteiler 2).

Tabelle 7: TEM- und rTEM- Übersicht der Intrabeurteilerdiskrepanz.

Probe	Beurteiler	Foto		DVT		$\mu\text{CT}$	
		TEM ( $\mu\text{m}$ )	rTEM (%)	TEM ( $\mu\text{m}$ )	rTEM (%)	TEM ( $\mu\text{m}$ )	rTEM (%)
1	1	13	8,9	13	8,9	12	7,9
	2	12	8,1	12	7,2	15	10,2
	3	9	6,3	12	7,5	7	4,1
2	1	13	7,5	13	7,5	11	6,6
	2	13	7,9	13	7,0	8	4,5
	3	8	4,3	12	5,2	10	5,1
3	1	13	8,8	16	12,5	10	7,3
	2	7	4,3	14	10,8	10	7,0
	3	5	3,6	9	6,1	6	3,9
4	1	14	7,7	21	9,1	15	7,3
	2	9	5,0	17	7,4	10	4,8
	3	6	3,4	13	6,0	7	3,7
5	1	14	15,1	40	21,1	15	10,2
	2	14	14,6	18	10,7	17	13,0
	3	6	6,5	16	7,3	5	3,6
Gesamt		11,1	7,2	17,3	10,0	11,1	6,8

Bei Betrachtung des technischen Messfehlers für jeden Beurteiler individuell, aber alle Proben mit einbezogen, erhält man das Ergebnis, dass Beurteiler 3 bei allen drei Methoden den kleinsten technischen Messfehler erzeugte, während Beurteiler 1 den größten Fehler von 23  $\mu\text{m}$  (gelb hervorgehoben) am DVT zu verzeichnen hatte. Dies wird in Tabelle 8 verdeutlicht.

Tabelle 8: Technische Messfehler der Beurteiler unter Einbeziehung aller Proben.

Beurteiler	Foto		DVT		$\mu\text{CT}$	
	TEM ( $\mu\text{m}$ )	rTEM (%)	TEM ( $\mu\text{m}$ )	rTEM (%)	TEM ( $\mu\text{m}$ )	rTEM (%)
1	14	9,1	23	13,3	13	7,8
2	11	7,5	15	8,6	13	7,8
3	7	4,7	12	6,5	7	4,2

#### 4.1.3 Ergebnisse des nichtparametrischen Tests

Ausgegangen wurde von der Nullhypothese, dass die jeweils fünf Messwerte pro Messstelle einer Probe der Messreihe 1 aus einer Grundgesamtheit mit den äquivalenten fünf Messungen der Messreihe 2 stammten. In Tabelle 10 werden die einzelnen exakten Signifikanzen angezeigt, da die Anzahl an getesteten Werten kleiner als 30 war.

Bei einem p-Wert unter 0,05 war die Nullhypothese abzulehnen.

Zur Veranschaulichung wurden in der Tabelle 10 die Messstellen farblich hervorgehoben, bei denen die Nullhypothese abzulehnen war.

Bei der Probe 1, Beurteiler 2, lag der p-Wert am Foto an den Messstellen 5 und 10 bei unter 0,05. Am DVT hingegen wurde für deutlich mehr Messstellen ein p-Wert von unter 0,05 errechnet, vor allem Messstellen 2, 7 und 9 wurden von allen Beurteilern in der Messreihe 1 different zur Messreihe 2 vermessen. Die p-Werte der Berechnungen für die Werte des  $\mu$ CTs liegen achtmal unter 0,05.

Die Probe 2 wurde am Foto von allen Beurteilern an allen Messstellen ohne signifikante Unterschiede vermessen. Am DVT wurden fünf Stellen different vermessen und am  $\mu$ CT insgesamt sieben.

Die Probe 3 wurde am Foto viermal unterschiedlich vermessen. Am DVT hingegen ergab sich bei elf Messstellen ein signifikanter Unterschied zwischen den Messungen der Messreihen. Sieben dieser Stellen betrafen Beurteiler 2.

Am  $\mu$ CT wurden drei Messstellen unterschiedlich vermessen.

Die Probe 4 wurde am Foto und am  $\mu$ CT nur einmal verschieden gemessen. Am DVT war dies dreimal der Fall.

Die Probe 5 wurde am Foto in elf Fällen unterschiedlich vermessen. Am DVT kam das neunmal vor und am  $\mu$ CT siebenmal. Dies ist die einzige Probe, bei der am Foto öfter unterschiedlich vermessen wurde als beim DVT.

Bei Betrachtung der Ergebnisse des nichtparametrischen Tests auf Signifikanzen stellt man fest, dass die Signifikanzen in unterschiedlicher Häufung, abhängig von der Methode, bei unter  $p = 0,05$  lagen.

Dies spiegelt die Tabelle 9 wider, in der gezeigt wird, dass die angenommene Nullhypothese bei der Methode DVT mit 28,7 % mehr als doppelt so häufig abgelehnt wurde, als beim Foto mit 12,7 %. Dieser Wert beträgt für die Methode  $\mu$ CT 17,3 %.

Tabelle 9: Häufigkeit der Ablehnung der Nullhypothese abhängig von der Methode.

	<b>Foto</b>	<b>DVT</b>	<b><math>\mu</math>CT</b>
Gesamtanzahl der Fälle	150	150	150
Nullhypothese trifft zu (absolut)	131	107	127
Nullhypothese trifft nicht zu (absolut)	19	43	26
Nullhypothese trifft nicht zu (relativ in %)	12,7	28,7	17,3



Tabelle 10: P-Werte des Hypothesentests der Intrabeurteilerdiskrepanz. Rot hervorgehoben sind

P-Werte unter 0,05.

MS	Beurteiler	Probe 1			Probe 2			Probe 3			Probe 4			Probe 5		
		Foto	DVT	µCT	Foto	DVT	µCT	Foto	DVT	µCT	Foto	DVT	µCT	Foto	DVT	µCT
1	1	0,69	0,222	0,69	0,151	0,69	1	0,421	0,095	1	0,095	0,032	0,548	0,151	0,032	0,032
	2	0,421	0,548	0,008	0,151	1	0,69	0,69	0,032	0,421	1	0,056	0,841	1	0,841	0,095
	3	0,841	1	0,548	0,841	1	0,222	0,421	0,69	0,69	1	1	0,151	0,69	0,548	0,841
2	1	0,421	0,008	1	0,056	0,841	0,056	0,222	0,222	0,222	0,548	0,095	0,222	0,008	0,008	0,095
	2	0,841	0,008	1	0,548	0,548	0,032	0,095	0,008	0,222	1	0,151	0,548	0,008	0,421	0,421
	3	1	0,016	0,222	0,151	0,841	1	0,841	0,31	0,032	0,841	0,151	0,841	1	0,69	0,421
3	1	0,69	0,008	0,008	0,151	0,841	0,008	0,31	0,841	0,69	0,032	0,056	0,008	0,841	0,095	0,69
	2	0,095	0,056	1	0,69	0,151	1	0,222	0,008	0,056	0,095	1	0,151	0,841	0,222	0,421
	3	0,841	1	0,841	0,841	0,548	0,69	0,222	1	0,151	0,548	0,421	0,31	0,69	0,31	0,841
4	1	0,69	0,095	0,841	0,222	0,151	0,841	0,008	1	0,548	0,69	0,421	0,222	0,032	0,841	0,151
	2	0,841	0,421	0,032	0,31	0,056	0,31	0,841	0,032	0,222	0,841	0,69	0,548	0,222	0,548	0,548
	3	0,841	0,421	0,222	1	0,69	0,548	0,69	0,222	0,69	0,421	0,841	0,151	1	0,032	1
5	1	0,421	0,008	0,31	0,056	0,056	0,016	1	0,548	0,31	0,222	0,548	0,151	0,016	0,016	0,008
	2	0,008	0,151	0,016	1	0,032	0,69	0,016	0,008	0,548	1	0,095	0,548	0,008	0,841	0,008
	3	0,548	0,421	0,69	1	0,548	1	0,421	0,548	0,841	1	0,421	0,548	0,421	0,841	0,222
6	1	0,31	0,095	0,032	0,548	0,008	0,222	0,31	0,008	0,421	0,31	0,008	0,222	0,032	0,69	1
	2	0,095	0,008	0,095	0,056	0,008	0,151	0,69	0,151	0,421	1	0,016	1	1	0,841	0,008
	3	0,69	0,016	0,841	0,69	0,841	0,222	0,841	0,841	0,548	0,69	0,548	0,841	0,095	0,222	0,69
7	1	0,69	0,008	0,69	0,056	0,008	0,008	0,222	0,016	0,008	0,222	0,16	0,008	0,008	0,222	0,31
	2	0,841	0,032	0,151	0,151	0,31	0,016	0,841	0,151	0,151	0,548	0,222	0,548	0,421	0,222	0,032
	3	0,69	0,008	0,841	0,548	0,841	0,841	0,151	0,222	1	0,222	0,421	0,222	0,222	0,69	0,222
8	1	0,548	0,008	0,032	1	0,056	0,095	0,008	0,008	0,032	0,548	0,421	0,69	0,008	0,008	1
	2	0,841	0,548	0,095	1	0,31	0,69	0,421	0,056	0,151	0,69	0,841	0,222	0,31	0,222	0,421
	3	0,69	0,222	0,69	0,222	1	0,032	0,841	0,69	0,421	0,31	1	0,421	0,095	0,421	0,841
9	1	0,841	0,008	0,032	0,841	0,421	0,222	1	0,008	0,548	0,222	0,421	0,841	0,008	0,008	0,69
	2	0,151	0,008	0,008	0,31	1	0,31	0,421	0,032	0,841	0,31	0,841	0,69	0,056	0,016	0,008
	3	0,841	0,008	0,31	0,31	0,421	0,548	0,421	0,841	1	0,841	0,548	0,841	0,548	0,421	0,31
10	1	0,008	0,31	0,056	0,31	0,008	0,032	0,016	0,841	0,841	0,056	0,69	0,841	0,016	0,008	0,032
	2	0,008	0,151	0,095	0,222	0,421	0,151	1	0,008	0,841	0,222	0,69	1	0,008	0,008	0,31
	3	0,421	0,151	0,548	0,69	0,841	0,31	0,841	0,056	0,056	0,69	0,548	0,548	0,548	0,548	0,421

---

## 4.2 Interbeurteilerdiskrepanz

Dargestellt und beschrieben werden im Folgenden die Ergebnisse der Interbeurteilerdiskrepanz zunächst anhand der deskriptiven und induktiven Statistik, gefolgt von den Ergebnissen des technischen Messfehlers (TEM).

### 4.2.1 Ergebnisse der deskriptiven und induktiven Statistik

Die Ergebnisse der deskriptiven und induktiven Statistik werden zusammenhängend präsentiert. Die Interbeurteilerdiskrepanz gibt Auskunft darüber, welche Messstellen bei welcher Methode indifferent von den Beurteilern vermessen wurden. Dies zeigen sowohl die Mittelwerte der deskriptiven Statistik als auch die Tests auf Signifikanzen.

Es zeigte sich, dass es bei der Methode Foto vier Messstellen gab, die von allen drei Beurteilern signifikant unterschiedlich vermessen wurden. Bei 25 Messstellen haben zwei Beurteiler Messwerte gemessen, die sich nicht signifikant unterschieden und bei 14 Messstellen entstammten die Messwerte aller Beurteiler einer Grundgesamtheit. Sieben Messstellen konnten nicht beurteilt werden. Dies konnte aus zwei Gründen der Fall gewesen sein:

1. Es lagen aufgrund der Intra-beurteilerdiskrepanz nur Messwerte eines einzigen Beurteilers vor.
2. Die Ergebnisse des Kruskal-Wallis-Tests deuteten auf einen Fehler hin, da beispielsweise Beurteiler 1 und 3 sowie 2 und 3 indifferent gemessen haben, Beurteiler 1 und 2 aber nicht.

Die Anzahl der Messstellen, die keine Grundgesamtheit inne hatten, betrug beim DVT und  $\mu$ CT jeweils zwölf. Dass zwei Beurteiler Messwerte aus einer gemeinsamen Grundgesamtheit gemessen haben, kam bei der Methoden DVT 19 Mal und  $\mu$ CT 18 Mal vor. 14 Messstellen wurden mit der Methode  $\mu$ CT, 9 mit der Methode DVT, von allen drei Beurteilern ohne signifikanten Unterschied vermessen.

Die Anzahl an nicht zu beurteilenden Messstellen lag beim DVT bei zehn und beim  $\mu$ CT bei sechs.

Eine Übersicht über diese Verteilung wird in Tabelle 11 angeboten. Zur detaillierten Veranschaulichung werden die Vergleiche der Beurteiler zusätzlich anhand von Boxplots dargestellt.

Tabelle 11: Übersicht der Ergebnisse der induktiven Statistik der Interbeurteilerdiskrepanz.

		Foto	DVT	$\mu$ CT
Probe 1	Alle Messreihen gehören einer Grundgesamtheit an	5	3	0
	Zwei Messreihen gehören einer Grundgesamtheit an	4	2	5
	Keine Grundgesamtheit	0	1	4
Probe 2	Alle Messreihen gehören einer Grundgesamtheit an	2	0	1
	Zwei Messreihen gehören einer Grundgesamtheit an	8	6	4
	Keine Grundgesamtheit	0	3	3
Probe 3	Alle Messreihen gehören einer Grundgesamtheit an	3	0	5
	Zwei Messreihen gehören einer Grundgesamtheit an	5	3	3
	Keine Grundgesamtheit	1	6	1
Probe 4	Alle Messreihen gehören einer Grundgesamtheit an	4	5	6
	Zwei Messreihen gehören einer Grundgesamtheit an	3	3	3
	Keine Grundgesamtheit	1	0	0
Probe 5	Alle Messreihen gehören einer Grundgesamtheit an	0	1	2
	Zwei Messreihen gehören einer Grundgesamtheit an	5	5	3
	Keine Grundgesamtheit	2	2	4
Gesamt	Alle Messreihen gehören einer Grundgesamtheit an	14	9	14
	Zwei Messreihen gehören einer Grundgesamtheit an	25	19	18
	Keine Grundgesamtheit	4	12	12
	Nicht zu beurteilende Messstellen	7	10	6

### Boxplots

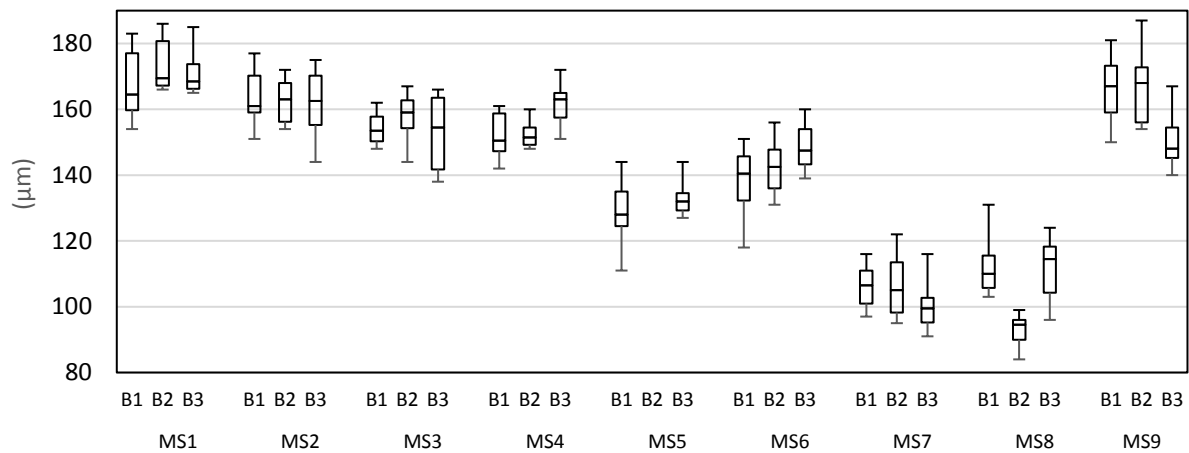
Es ergaben sich drei, bzw. abhängig von den Ergebnissen der Intrabeurteilerdiskrepanz, zwei Boxplots je Messstelle (ein Boxplot pro Beurteiler), die gruppiert nebeneinander zum Vergleich stehen.

Messstellen, die aufgrund der Ergebnisse der Intrabeurteilerdiskrepanz nicht mehr zum Beurteilervergleich zu verwenden waren (nur ein Beurteiler hat diese Messstelle in MR1 und MR2 ohne signifikante Differenz vermessen), werden nicht dargestellt. „B1-B3“ bezeichnet die Beurteiler 1 bis Beurteiler 3. „MS“ gibt die Messstelle an.

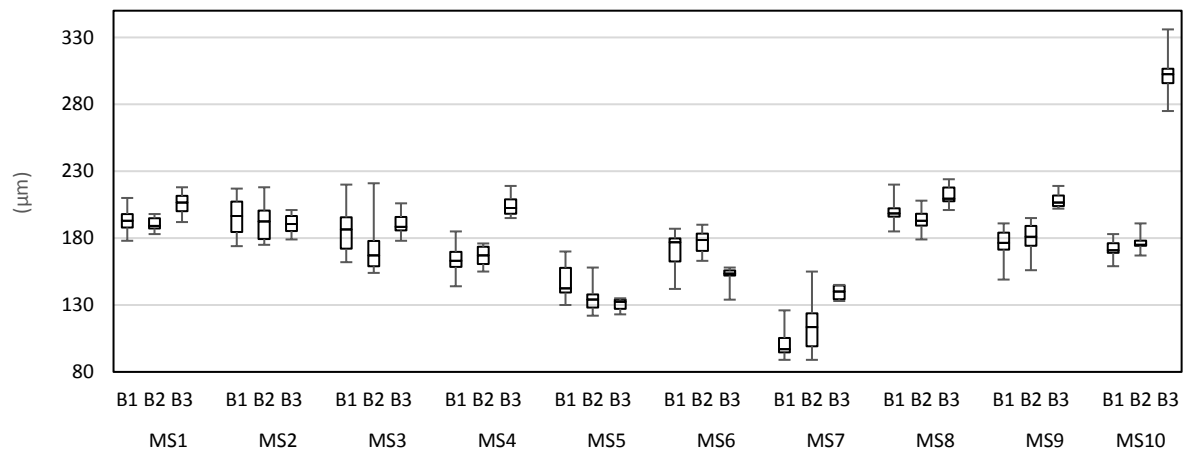
### Boxplots der Methode „Foto“:

Bei Betrachtung der Boxplots der am Foto gemessenen Werte fällt die Messstelle 10 der Probe 2 auf. An dieser Stelle maß Beurteiler 3 einen um ca. 100  $\mu\text{m}$  größeren Wert als die Beurteiler 1 und 2. Um ca. 50  $\mu\text{m}$  kleiner hingegen maß Beurteiler 3 die Messstelle 6 der Probe 4. Bei der Probe 5 standen, aufgrund der Ergebnisse der Intra-beurteilerdiskrepanz, nur sieben Messstellen zum Vergleich zur Verfügung. Lediglich die Messstelle 3 ließ einen Vergleich aller Beurteiler zu.

#### *Probe 1*

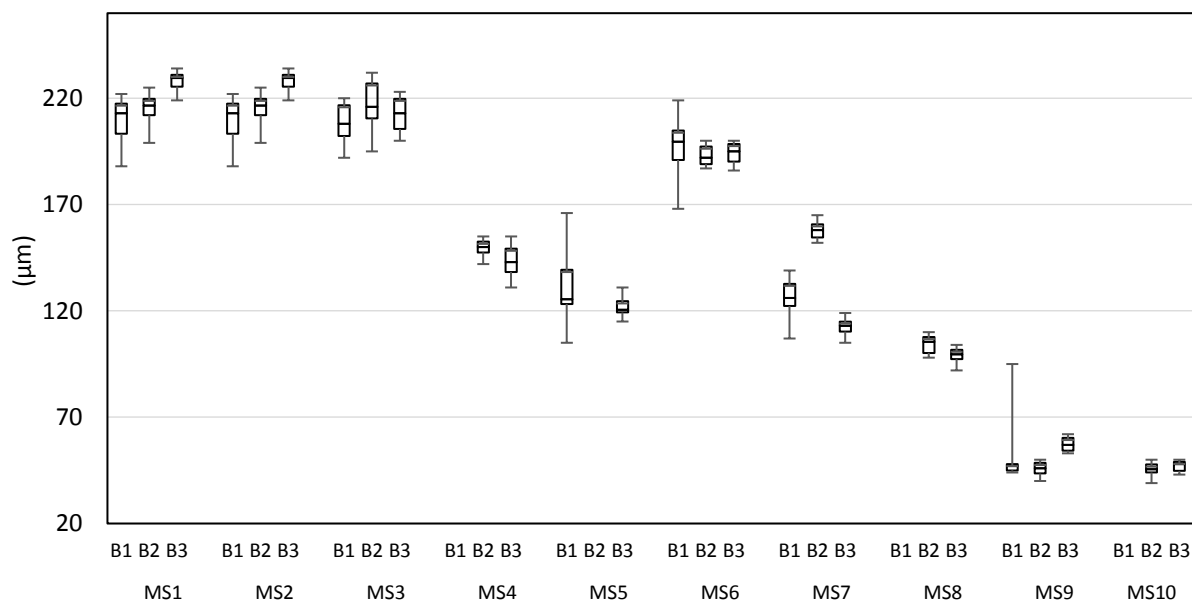


#### *Probe 2*

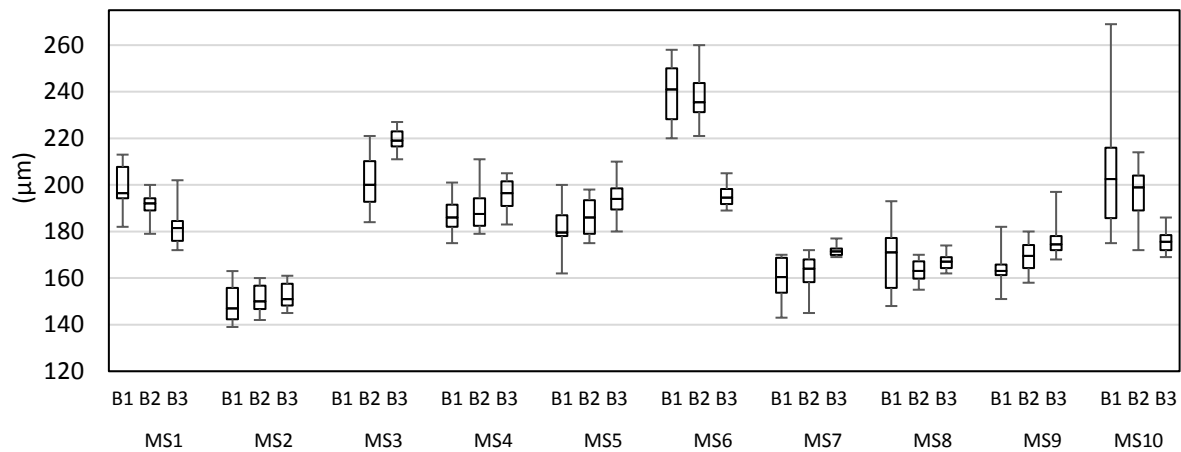


*Boxplots 1: Boxplots der Proben 1 und 2 der Interbeurteilerdiskrepanz der Methode „Foto“.*

Probe 3

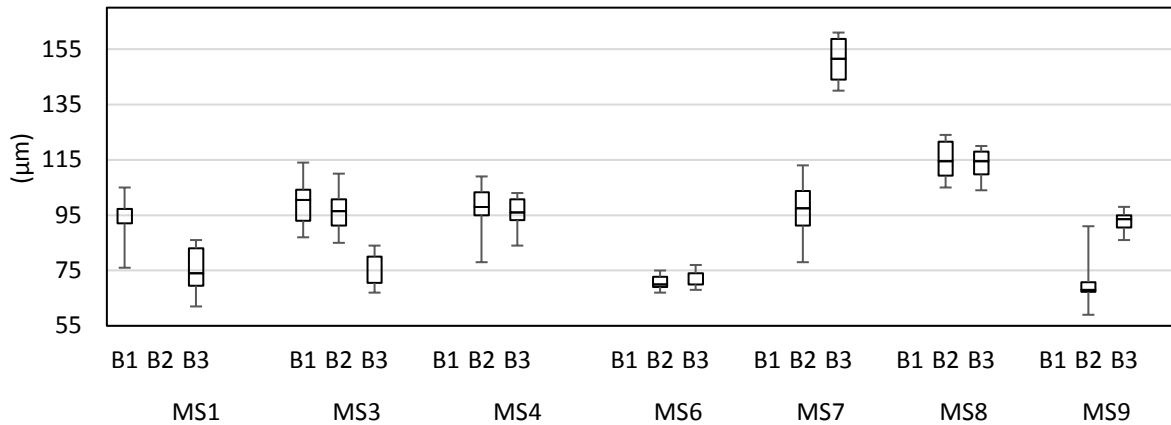


Probe 4



Boxplots 2: Boxplots der Proben 3 und 4 der Interbeurteilerdiskrepanz der Methode „Foto“.

## Probe 5

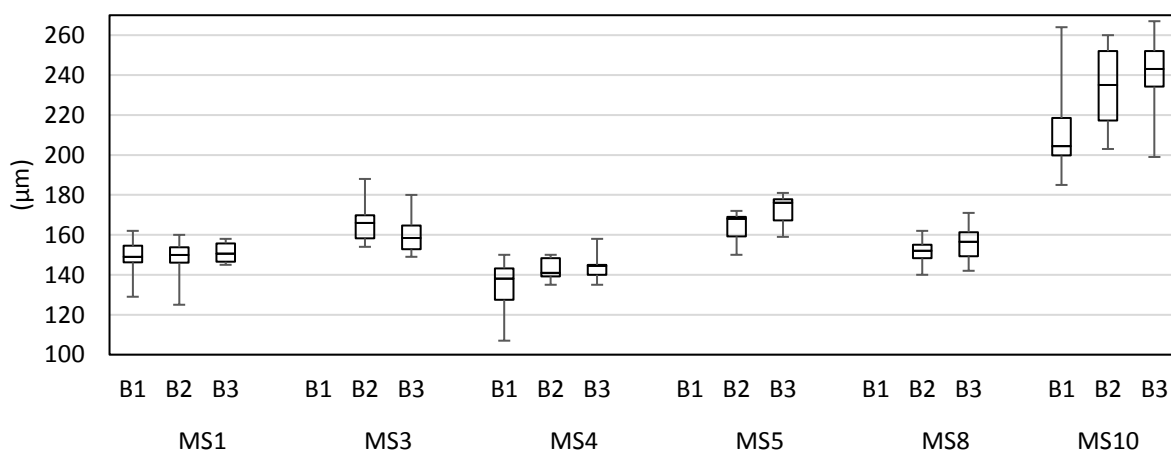


Boxplots 3: Boxplots der Probe 5 der Interbeurteilerdiskrepanz der Methode „Foto“.

Boxplots der Methode „DVT“:

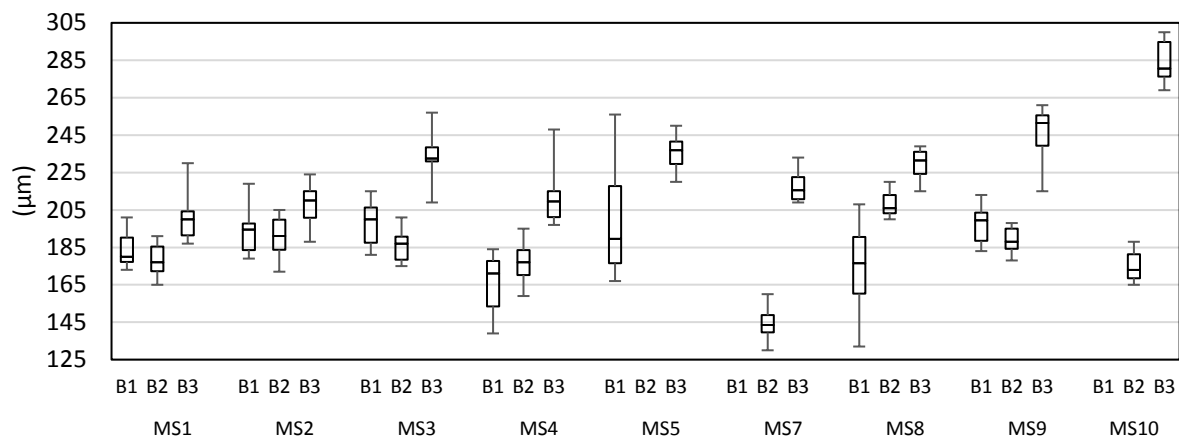
Die Boxplots der Vermessung der Messstelle 10 der Probe 2 zeigen, dass die Messungen des Beurteilers 3 um ca. +100  $\mu\text{m}$  von denen des Beurteilers 2 abwichen. Die Messungen des Beurteilers 1 standen bei dieser Messstelle nicht zum Vergleich. Auch die Messstelle 2 der Probe 5 wurde von Beurteiler 3 abweichend vermessen. Die Abweichung von den Messungen des Beurteilers 2 lag bei ca. +200  $\mu\text{m}$ .

## Probe 1

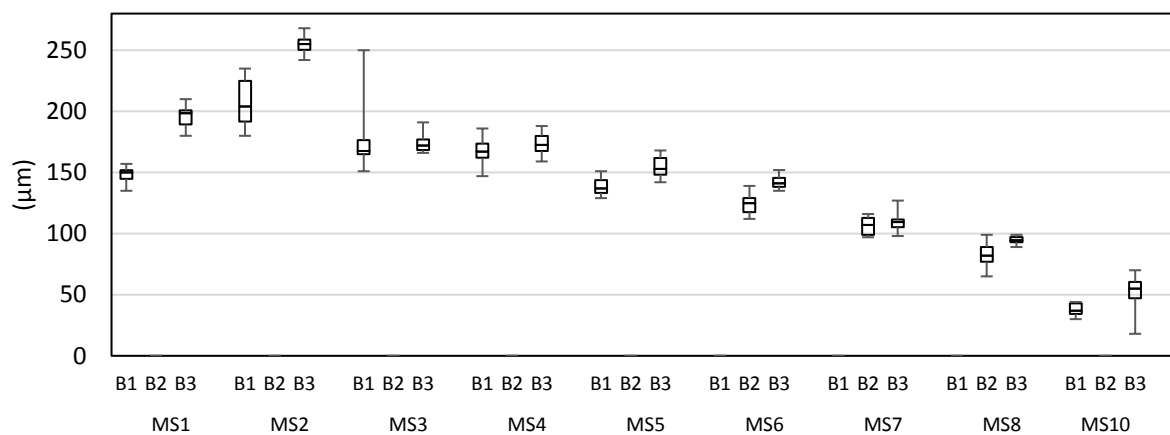


Boxplots 4: Boxplots der Probe 1 der Interbeurteilerdiskrepanz der Methode „DVT“.

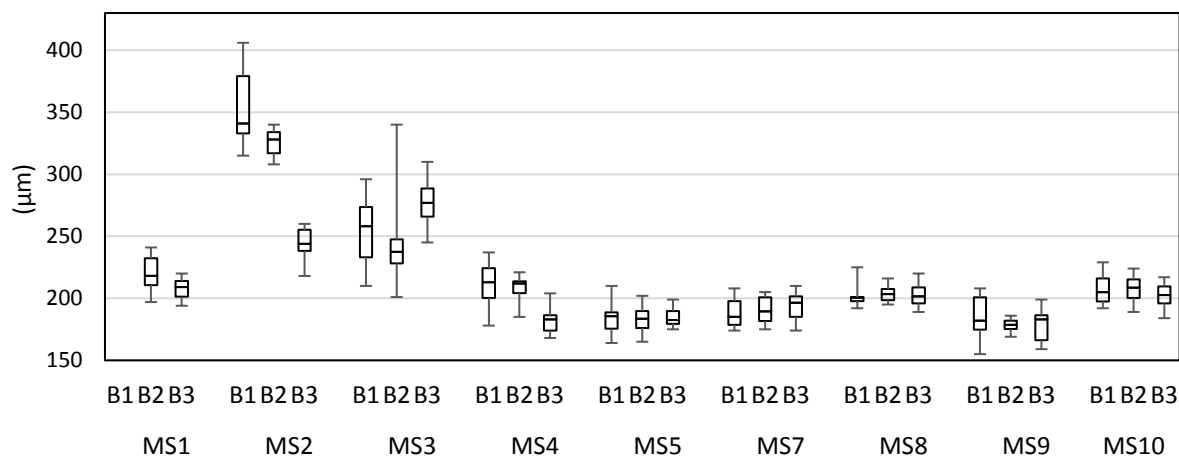
Probe 2



Probe 3

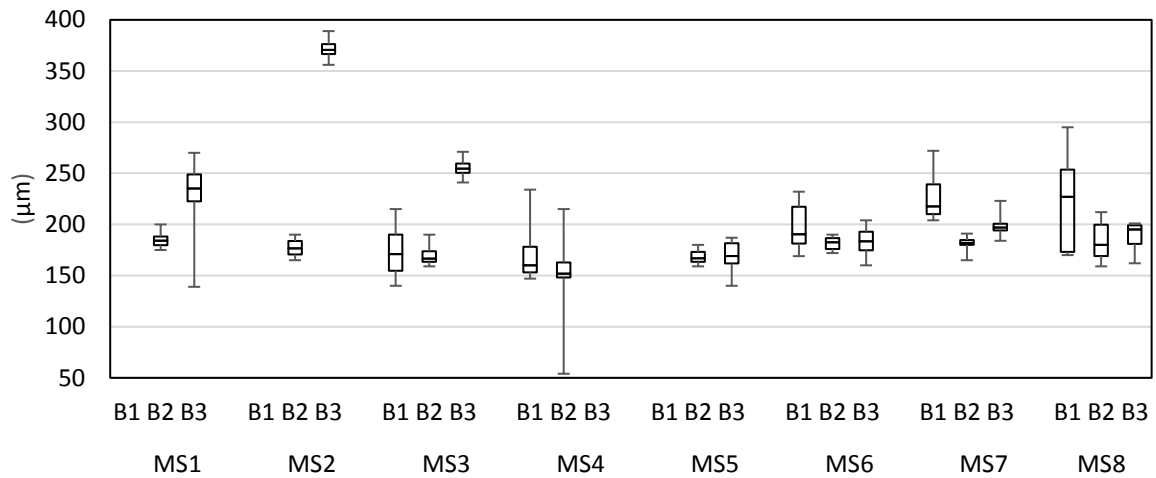


Probe 4



Boxplots 5: Boxplots der Proben 2, 3 und 4 der Interbeurteilerdiskrepanz der Methode „DVT.“

## Probe 5

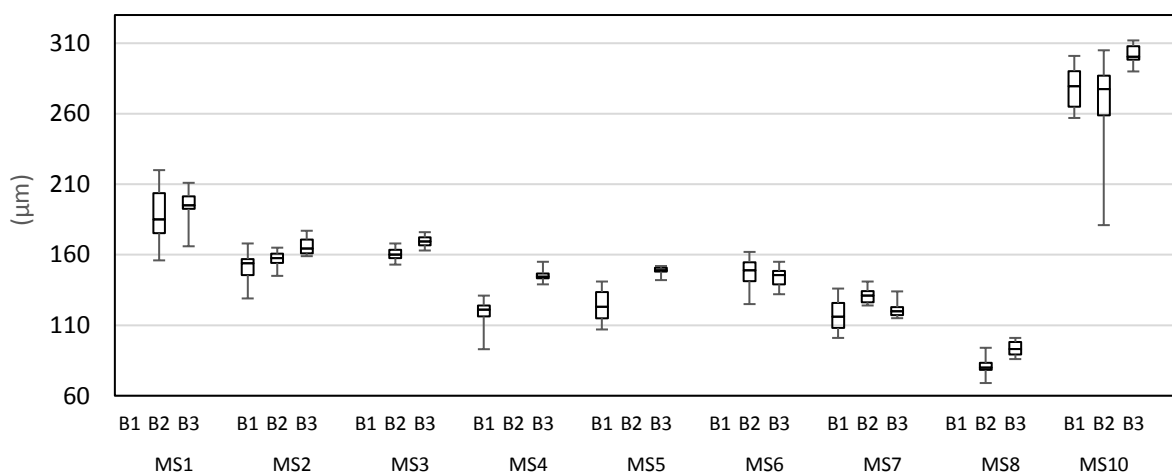


Boxplots 6: Boxplots der Probe 5 der Interbeurteilerdiskrepanz der Methode „DVT“.

Boxplots der Methode „ $\mu\text{CT}$ “:

Besonders zu betrachten sind die Boxplots der Probe 2. Zwei Messstellen wurden jeweils von Beurteiler 3 um ca. 50  $\mu\text{m}$  größer vermessen als von Beurteiler 2. Dies waren die Messstellen 3 und 5.

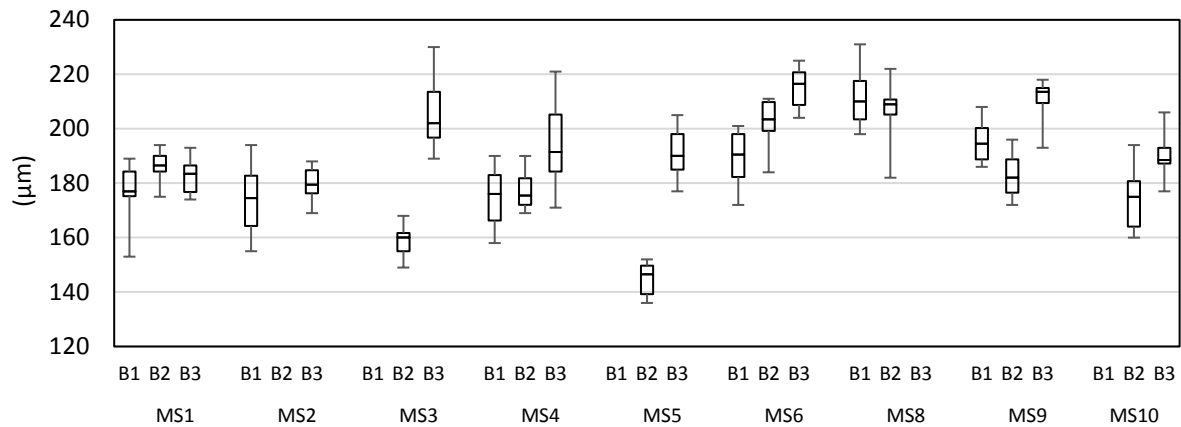
## Probe 1



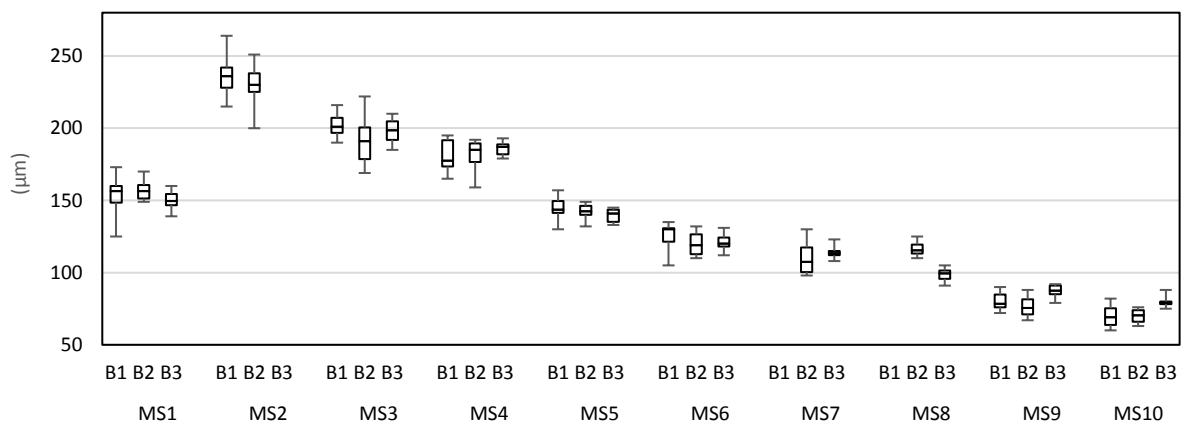
Boxplots 7: Boxplots der Probe 1 der Interbeurteilerdiskrepanz der Methode „ $\mu\text{CT}$ “.



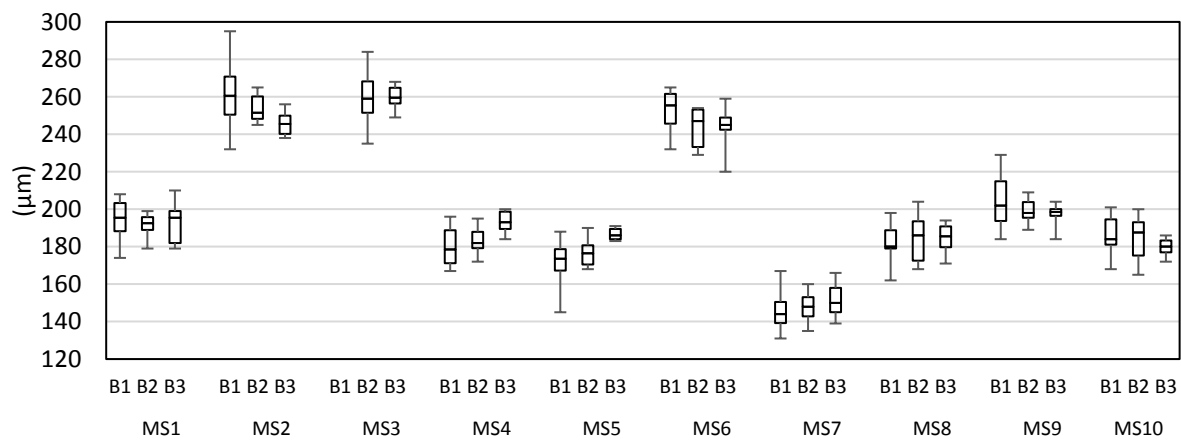
Probe 2



Probe 3

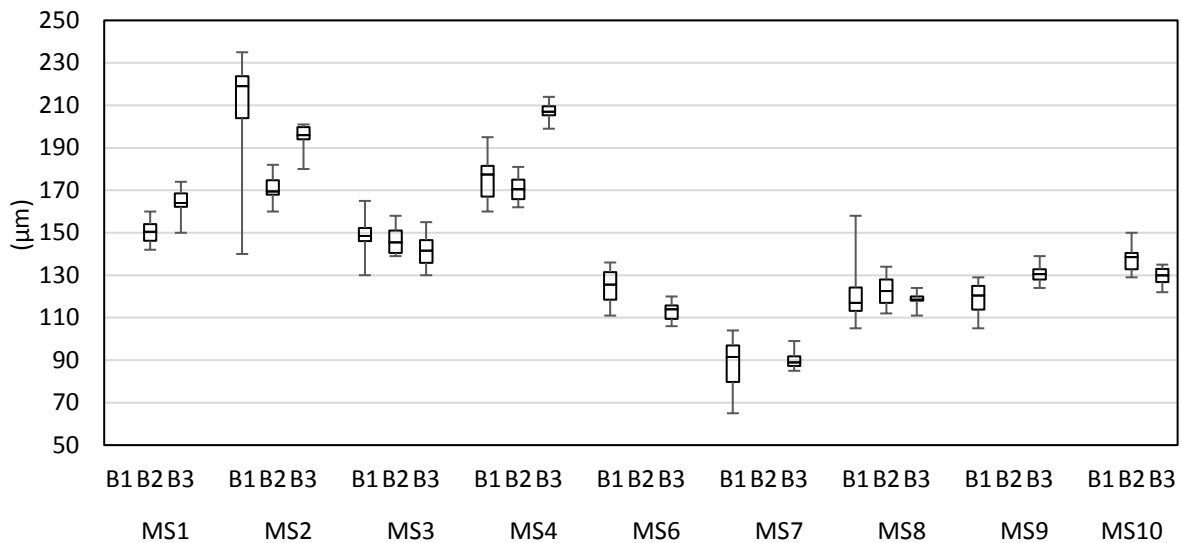


Probe 4



Boxplots 8: Boxplots der Proben 2,3 und 4 der Interbeurteilerdiskrepanz der Methode „ $\mu\text{CT}$ “.

## Probe 5



Boxplots 9: Boxplots der Probe 5 der Interbeurteilerdiskrepanz der Methode „µCT“.

Auszugsweise gezeigt werden folgend die deskriptiven Statistiken mit zugehörigen Balkendiagrammen der Interbeurteilerdiskrepanz. Rosa hinterlegt in den Tabellen sind die Messstellen, die aufgrund der Intra-beurteilerdiskrepanz nicht mehr zum Vergleich in Betracht gezogen wurden. Nebeneinander gestellt sind die Mittelwerte und die dazugehörigen Standardabweichungen (S.D.) der einzelnen Messstellen (MS).

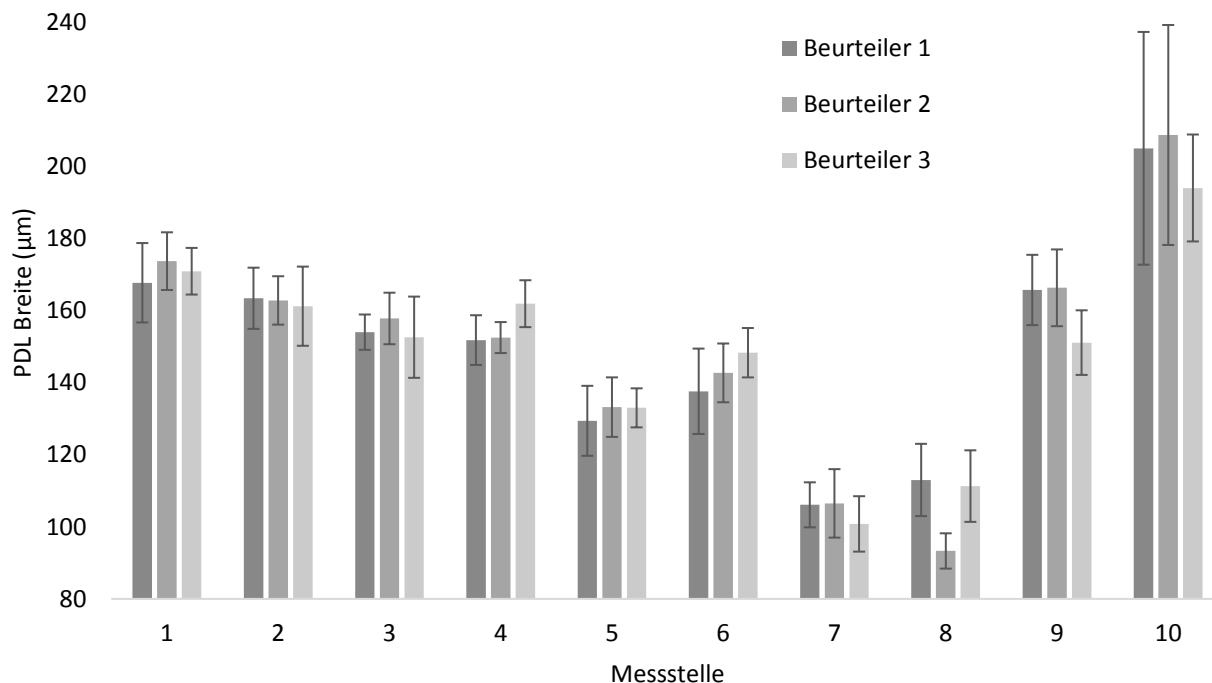
Die Tabelle 12 auf Seite 59 (*Deskriptive Statistik der Vermessung des PDLs der Probe 1 am Foto, Interbeurteilerdiskrepanz*) zeigt, dass die Standardabweichungen der Messungen der Beurteiler 1 (32,3 µm) und Beurteiler 2 (30,5 µm) an der Messstelle 10 deutlich höher waren als die Standardabweichungen an den restlichen Messstellen (maximal 11,8 µm).

Die Probe 3, am DVT vermessen (Tabelle 13, S. 60) ergab, dass Beurteiler 3 neun der zehn Messstellen, verglichen mit den anderen beiden Beurteilern, am größten vermessen hat. Nur die Messstelle 3 wurde von ihm am kleinsten vermessen.

Tabelle 14 auf Seite 61 (*Deskriptive Statistik der Vermessung des PDLs der Probe 4 am µCT, Interbeurteilerdiskrepanz*) zeigt, dass die Standardabweichungen bei maximal 17,8 µm lagen. Lediglich die Messstelle 3, vermessen von Beurteiler 1, wich stark ab. Hier lag eine Standardabweichung von 30,1 µm vor.

Tabelle 12: Deskriptive Statistik der Vermessung des PDLs der Probe 1 am Foto, Interbeurteilerdiskrepanz.

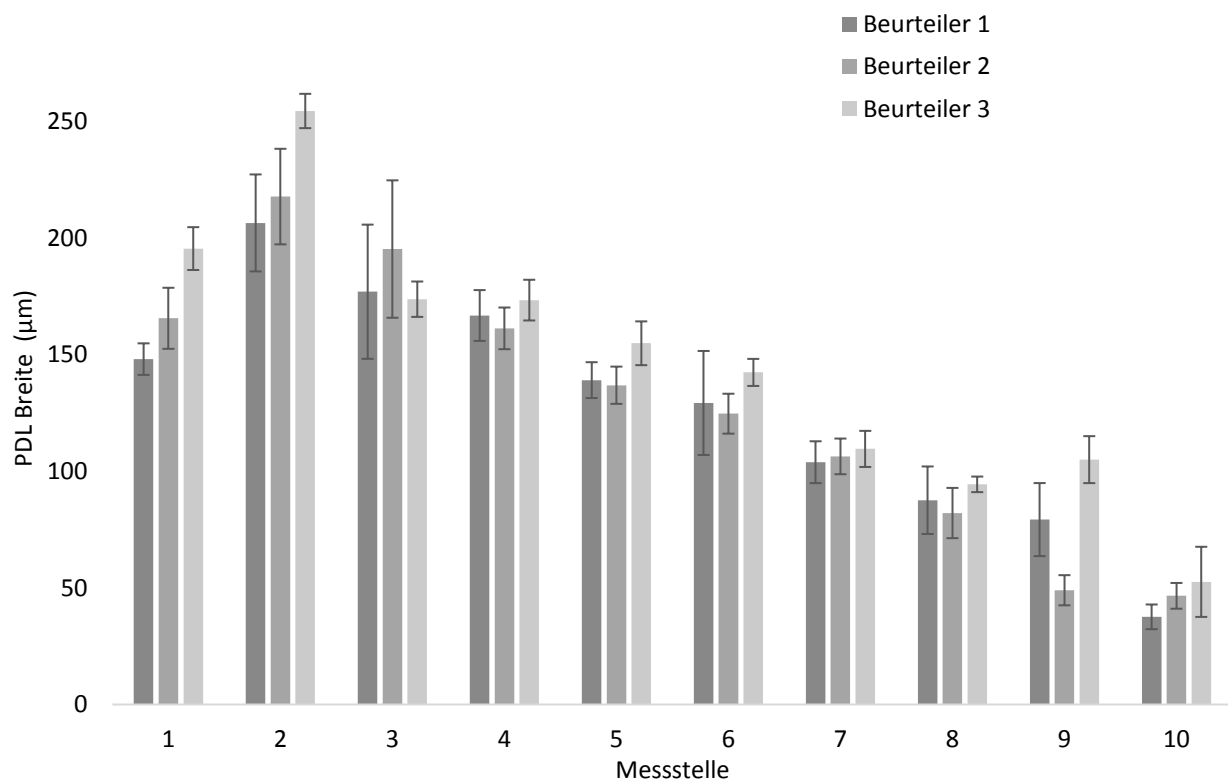
MS	Beurteiler 1		Beurteiler 2		Beurteiler 3	
	Mittelw.	S.D.	Mittelw.	S.D.	Mittelw.	S.D.
	( $\mu\text{m}$ )	( $\mu\text{m}$ )	( $\mu\text{m}$ )	( $\mu\text{m}$ )	( $\mu\text{m}$ )	( $\mu\text{m}$ )
1	168	11,0	174	8,0	171	6,5
2	163	8,5	163	6,7	161	11,0
3	154	4,9	158	7,1	153	11,3
4	152	6,9	153	4,3	162	6,5
5	129	9,7	133	8,3	133	5,4
6	138	11,8	143	8,2	148	6,8
7	106	6,2	107	9,5	101	7,7
8	113	10,0	93	4,9	111	9,9
9	166	9,8	166	10,6	151	9,0
10	205	32,3	209	30,5	194	14,8



Balkendiagramm 4: Interbeurteilervergleich der Vermessung des PDLs der Probe 1 am Foto.

Tabelle 13: Deskriptive Statistik der Vermessung des PDLs der Probe 3 am DVT, Interbeurteilerdiskrepanz.

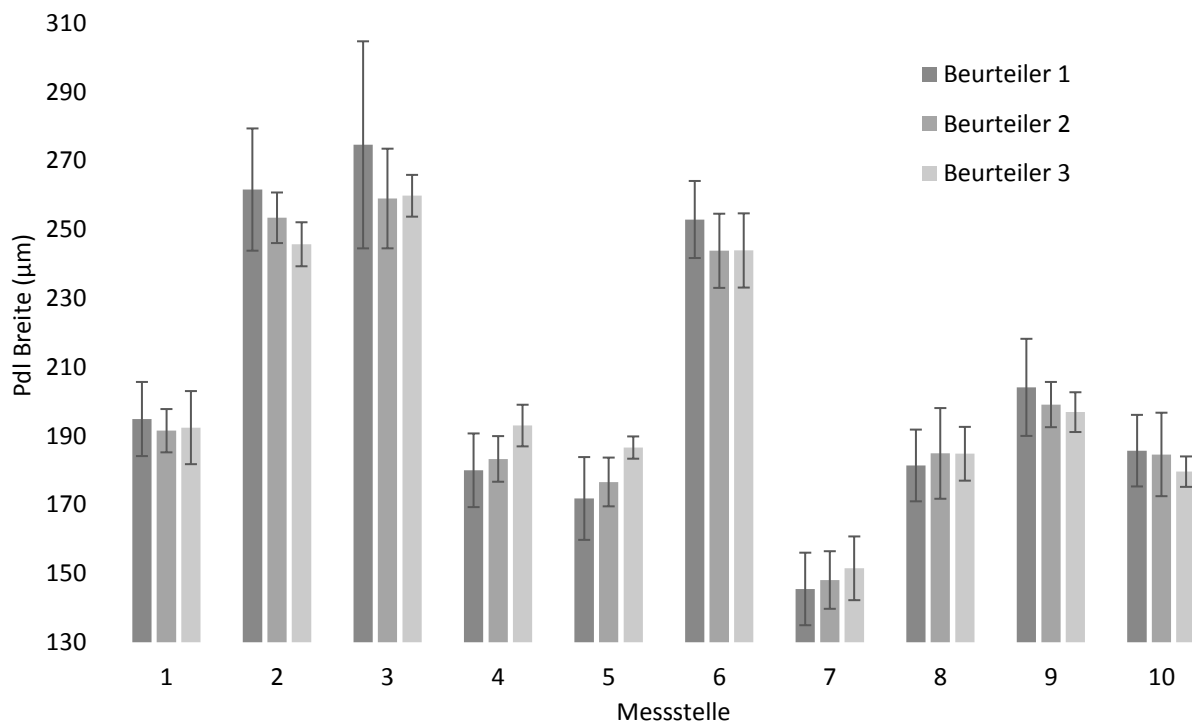
MS	Beurteiler 1		Beurteiler 2		Beurteiler 3	
	Mittelw.	S.D.	Mittelw.	S.D.	Mittelw.	S.D.
	( $\mu\text{m}$ )	( $\mu\text{m}$ )	( $\mu\text{m}$ )	( $\mu\text{m}$ )	( $\mu\text{m}$ )	( $\mu\text{m}$ )
1	148	6,8	166	13,1	196	9,2
2	207	20,8	218	20,5	255	7,4
3	177	28,8	195	29,5	174	7,6
4	167	10,9	161	8,9	173	8,7
5	139	7,7	137	8,0	155	9,4
6	129	22,3	125	8,6	142	5,8
7	104	9,0	106	7,6	110	7,7
8	88	14,5	82	10,8	94	3,3
9	79	15,7	49	6,5	105	10,1
10	38	5,3	47	5,5	53	15,0



Balkendiagramm 5: Interbeurteilervergleich der Vermessung des PDLs der Probe 3 am DVT.

Tabelle 14: Deskriptive Statistik der Vermessung des PDLs der Probe 4 am  $\mu$ CT, Interbeurteilerdiskrepanz.

MS	Beurteiler 1		Beurteiler 2		Beurteiler 3	
	Mittelw.	S.D.	Mittelw.	S.D.	Mittelw.	S.D.
	( $\mu$ m)	( $\mu$ m)	( $\mu$ m)	( $\mu$ m)	( $\mu$ m)	( $\mu$ m)
1	195	10,8	192	6,3	192	10,6
2	262	17,8	253	7,4	246	6,4
3	275	30,1	259	14,5	260	6,1
4	180	10,7	183	6,6	193	6,1
5	172	12,0	177	7,1	187	3,2
6	253	11,2	244	10,8	244	10,8
7	146	10,5	148	8,4	152	9,3
8	181	10,4	185	13,2	185	7,8
9	204	14,1	199	6,6	197	5,8
10	186	10,4	185	12,1	180	4,4



Balkendiagramm 6: Interbeurteilervergleich der Vermessung des PDLs der Probe 4 am  $\mu$ CT.

#### 4.2.2 Ergebnisse des Technical Error of Measurement

Der technische Messfehler betrug für die Methode Foto minimal 12  $\mu\text{m}$  mit einem relativen Fehler von 7,6 % (Probe 3). Der größte Fehler war bei der Probe 2 mit einem TEM von 25  $\mu\text{m}$  und einem relativen TEM von 14 % zu finden.

Die Berechnung des TEM anhand der am DVT erfassten Messwerte ergab, dass ein maximaler Fehler bei der Probe 5 von 42  $\mu\text{m}$  (rTEM: 21,8 %) vorlag. Der kleinste Fehler war bei der Probe 1 mit 18  $\mu\text{m}$  (rTEM: 11,5 %) erfasst worden.

Der technische Fehler der Messungen am  $\mu\text{CT}$  lag bei minimal 10  $\mu\text{m}$  und einem relativen TEM von 6,8 % (Probe 3) und maximal bei 18  $\mu\text{m}$  (rTEM: 12,6 %) bei der Probe 5.

Eine Übersicht der Ergebnisse des Technical Error of Measurement bietet die Tabelle 15.

Tabelle 15: TEM und rTEM der Interbeurteilerdiskrepanz.

Probe	Foto		DVT		$\mu\text{CT}$	
	TEM ( $\mu\text{m}$ )	rTEM (%)	TEM ( $\mu\text{m}$ )	rTEM (%)	TEM ( $\mu\text{m}$ )	rTEM (%)
1	12	8,2	18	11,5	14	9,1
2	25	14,0	33	17,2	16	8,6
3	12	7,6	18	13,6	10	6,8
4	14	7,4	23	10,5	12	5,8
5	15	16,3	42	21,8	18	12,6
<b>Gesamt</b>	16,3	10,8	28,5	15,9	14,0	8,5

Gelb hervorgehoben in der Tabelle 15 sind die Werte der Proben, die einen TEM-Wert von 14  $\mu\text{m}$  überschritten.

### 4.3 Methodenvergleich

Dieses Kapitel beschreibt die Ergebnisse des Methodenvergleichs. Zunächst werden die Ergebnisse der induktiven und deskriptiven Statistik, gefolgt von dem technischen Messfehler, dargestellt. Anschließend werden Punktwolken aus Mittelwerten gezeigt und beschrieben.

#### 4.3.1 Ergebnisse der deskriptiven und induktiven Statistik

Die Ergebnisse der induktiven Statistik in Form des Kruskal-Wallis-Tests ergaben, dass bei nur zwei der insgesamt 50 zu beurteilenden Messstellen kein signifikanter Unterschied zwischen den Messwerten der Methoden vorlag. Im Einzelnen sind dies zum einen die Messstelle 6 der Probe 1 im Vergleich zwischen den Methoden Foto und  $\mu$ CT und zum anderen die Messstelle 3 der Probe 2 im Vergleich zwischen den Methoden Foto und DVT. Bei den restlichen 48 Messstellen lag die Signifikanz bei unter  $p = 0,05$  oder es war aufgrund zu weniger Daten kein Vergleich möglich.

Bei der Errechnung der deskriptiven Statistik wurden folgende Ergebnisse ermittelt:

Der Vergleich zwischen den Messungen am Foto und DVT zeigte, dass im Mittel am DVT der Parodontalspalt um  $+31 \mu\text{m}$  größer vermessen wurde. Besonders groß war die Messdifferenz bei der Probe 4, Messstelle 2 mit  $188 \mu\text{m}$ . Bei nur sechs der insgesamt 25 erfassten Messstellen wurde am Foto ein größerer Wert gemessen als an den äquivalenten Messstellen am DVT. Hierbei lag das Maximum bei  $37 \mu\text{m}$  (Probe 3, Messstelle 3).

Der Vergleich zwischen den Methoden Foto und  $\mu$ CT ergab, dass gemittelt kein Unterschied bzgl. der Messgrößen zu verzeichnen war. Von den insgesamt beurteilbaren Messstellen wurden 14 am Foto größer gemessen, und 13 am  $\mu$ CT. Der Maximalwert des Vergleichs Foto zu  $\mu$ CT wurde bei Probe 3, Messstelle 6 mit  $+72 \mu\text{m}$  ermittelt und der Maximalwert des Vergleichs  $\mu$ CT zu Foto lag bei Probe 5, Messstelle 4, mit  $+77 \mu\text{m}$  vor.

Der Parodontalspalt wurde im Mittel am DVT um  $15 \mu\text{m}$  größer vermessen als am  $\mu$ CT. Eine maximale Differenz lag bei der Probe 5, Messstelle 7 vor. Diese Messstelle wurde am DVT um  $+124 \mu\text{m}$  größer vermessen als am  $\mu$ CT. Maximal größer hingegen wurde beim  $\mu$ CT die

Messstelle 1 der Probe 1 mit +43  $\mu\text{m}$  vermessen. Insgesamt wurden neun der 21 zu vergleichenden Messstellen am DVT kleiner, zwölf hingegen größer vermessen als am  $\mu\text{CT}$ .

Einen detaillierten Vergleich zwischen den Methoden bieten die Tabelle 16 sowie die folgenden Boxplots.

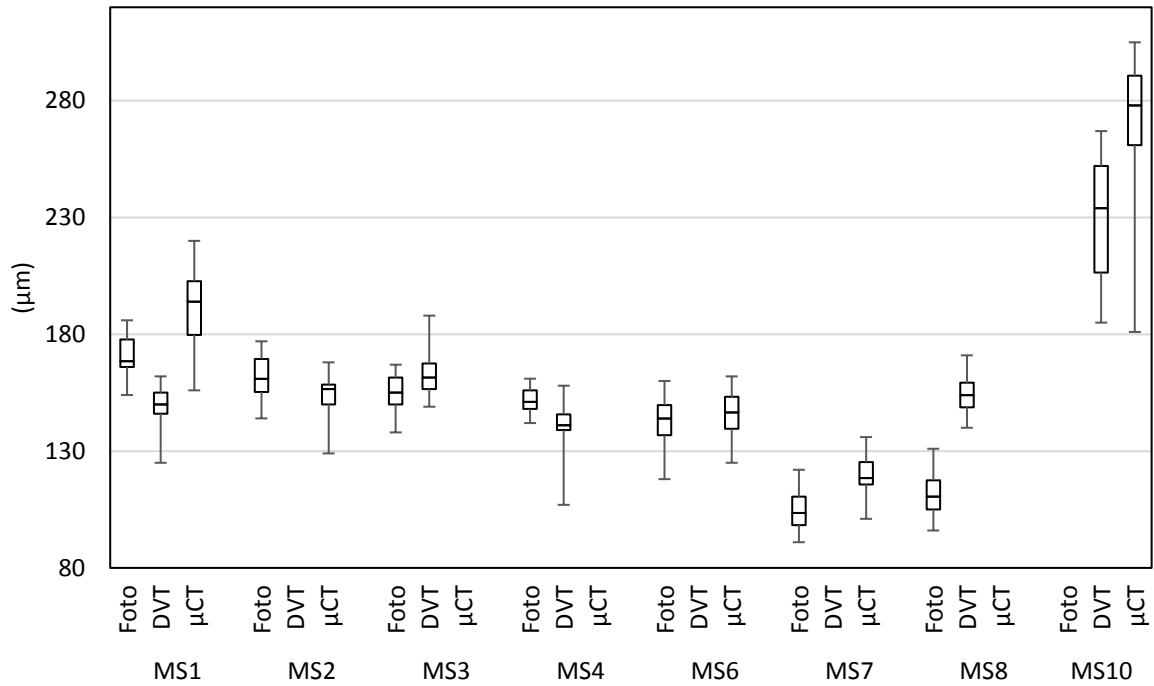
Tabelle 16: Deskriptive Statistik des Methodenvergleichs.

Probe	MS	Foto						DVT						$\mu\text{CT}$					
		N	Mittelw. ( $\mu\text{m}$ )	S.D. ( $\mu\text{m}$ )	Min. ( $\mu\text{m}$ )	Max. ( $\mu\text{m}$ )	Median ( $\mu\text{m}$ )	N	Mittelw. ( $\mu\text{m}$ )	S.D. ( $\mu\text{m}$ )	Min. ( $\mu\text{m}$ )	Max. ( $\mu\text{m}$ )	Median ( $\mu\text{m}$ )	N	Mittelw. ( $\mu\text{m}$ )	S.D. ( $\mu\text{m}$ )	Min. ( $\mu\text{m}$ )	Max. ( $\mu\text{m}$ )	Median ( $\mu\text{m}$ )
1	1	30	171	8,8	154	186	169	30	149	8,3	125	162	150	20	192	16,8	156	220	194
	2	30	162	8,6	144	177	161	0	-	-	-	-	-	20	154	9,9	129	168	157
	3	30	155	8,2	138	167	155	20	163	9,8	149	188	162	0	-	-	-	-	-
	4	20	152	5,6	142	161	151	30	141	9,6	107	158	141	0	-	-	-	-	-
	5	20	131	7,9	111	144	131	0	-	-	-	-	-	0	-	-	-	-	-
	6	30	143	9,9	118	160	144	0	-	-	-	-	-	20	146	9,4	125	162	147
	7	30	104	8,1	91	122	104	0	154	8,1	140	171	154	20	119	8,9	101	136	119
	8	20	112	9,7	96	131	111	0	-	-	-	-	-	0	-	-	-	-	-
	9	20	158	11,8	140	181	157	20	229	24,6	185	267	234	0	-	-	-	-	-
	10	0	-	-	-	-	-	0	-	-	-	-	-	20	273	27,9	181	305	278
2	1	20	192	7,7	178	210	191	20	181	9,2	165	201	179	30	182	8,6	153	194	184
	2	30	194	13,0	174	218	193	20	192	10,8	172	219	194	20	177	10,0	155	194	177
	3	30	184	19,0	154	221	185	20	192	11,7	175	215	190	0	-	-	-	-	-
	4	20	165	9,6	144	185	166	20	171	14,2	139	195	175	20	176	8,9	158	190	176
	5	0	-	-	-	-	-	10	198	28,6	167	256	190	0	-	-	-	-	-
	6	20	174	12,0	142	190	178	0	-	-	-	-	-	0	-	-	-	-	-
	7	20	108	17,0	89	155	100	0	-	-	-	-	-	0	-	-	-	-	-
	8	20	197	9,6	179	220	196	20	191	24,9	132	220	202	20	209	10,1	182	231	209
	9	20	177	12,4	149	195	178	20	193	9,7	178	213	191	20	189	10,1	172	208	189
	10	20	174	7,0	159	191	175	0	-	-	-	-	-	0	-	-	-	-	-
3	1	20	212	10,1	188	225	216	0	-	-	-	-	-	30	154	9,5	125	173	153
	2	30	298	10,7	259	312	300	0	-	-	-	-	-	20	233	15,0	200	264	233
	3	30	212	10,4	192	232	213	20	175	20,5	151	250	170	30	197	12,2	169	222	199
	4	20	146	7,2	131	155	148	20	170	10,2	147	188	170	30	183	9,3	159	195	185
	5	20	127	13,5	108	166	124	0	-	-	-	-	-	30	142	6,2	130	157	142
	6	30	194	10,0	168	219	195	0	-	-	-	-	-	30	122	8,2	105	135	121
	7	20	119	10,1	105	139	117	20	108	7,7	97	127	110	20	112	8,5	98	130	112
	8	20	102	4,7	92	110	101	0	-	-	-	-	-	0	-	-	-	-	-
	9	20	50	13,2	40	95	47	0	-	-	-	-	-	0	-	-	-	-	-
	10	20	46	3,2	39	50	47	0	-	-	-	-	-	20	70	6,3	60	82	70
4	1	20	194	9,7	179	213	195	20	213	13,2	194	241	212	30	193	9,2	174	210	195
	2	30	151	6,9	139	163	150	20	339	26,8	308	406	334	0	-	-	-	-	-
	3	0	-	-	-	-	-	0	-	-	-	-	-	20	259	10,8	235	284	259
	4	30	191	9,2	175	211	190	20	210	14,7	178	237	212	20	182	8,8	167	196	182
	5	30	188	10,4	162	210	187	30	184	11,0	164	210	183	20	174	9,9	145	190	176
	6	20	239	12,0	220	260	238	0	-	-	-	-	-	30	247	11,4	220	265	247
	7	20	161	8,6	143	172	162	30	191	11,3	174	210	189	30	148	9,4	131	167	146
	8	30	166	9,3	148	193	166	30	203	9,1	189	225	201	30	184	10,5	162	204	185
	9	0	-	-	-	-	-	30	180	12,7	155	208	182	30	200	9,7	184	229	199
	10	20	201	21,3	172	269	201	30	206	11,7	184	229	205	30	183	9,6	165	201	182
5	1	20	95	8,0	76	110	922	0	-	-	-	-	-	0	-	-	-	-	-
	2	0	-	-	-	-	-	0	-	-	-	-	-	20	202	21,0	140	235	200
	3	20	98	8,0	85	114	99	20	171	18,0	140	215	167	30	145	8,0	130	165	146
	4	20	96	8,0	78	109	98	20	157	41,0	54	234	155	20	173	9,0	160	195	174
	5	0	-	-	-	-	-	20	168	12,0	140	187	169	0	-	-	-	-	-
	6	20	71	3,0	67	77	70	30	188	17,0	160	232	185	0	-	-	-	-	-
	7	0	-	-	-	-	-	20	213	22,0	184	272	207	20	89	9,0	65	104	91
	8	20	114	6,0	104	124	115	30	197	34,0	159	295	189	30	121	10,0	105	158	120
	9	0	-	-	-	-	-	0	-	-	-	-	-	0	-	-	-	-	-
	10	0	-	-	-	-	-	0	-	-	-	-	-	0	-	-	-	-	-

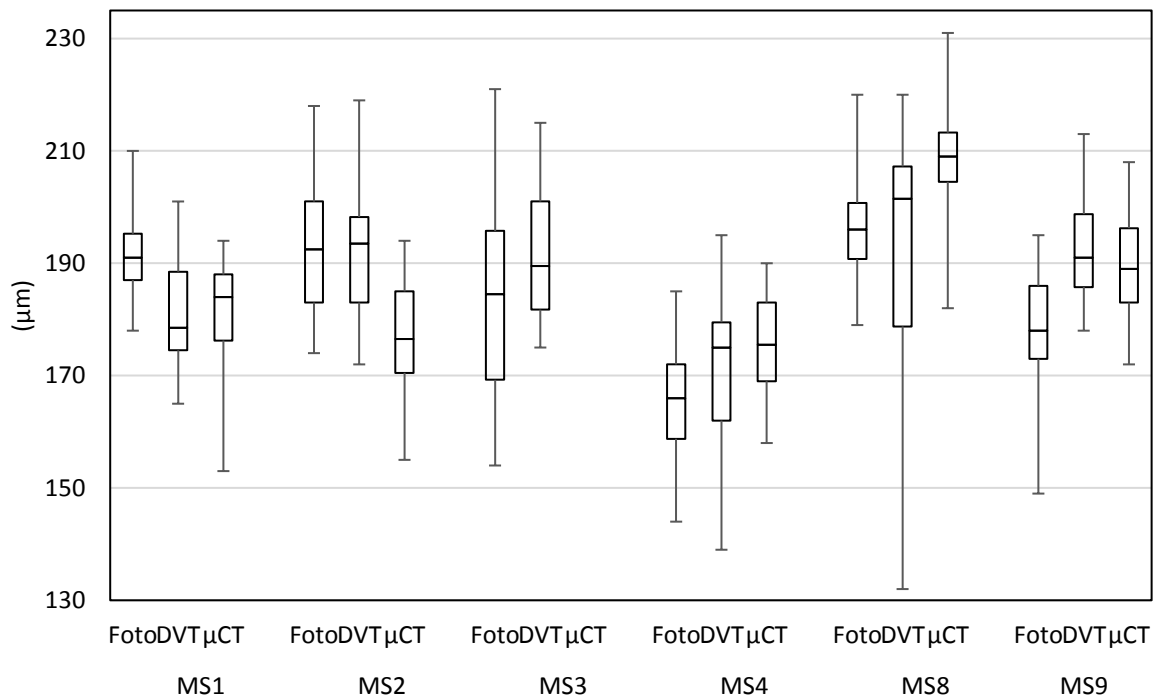


### Boxplots des Methodenvergleichs:

#### Probe 1

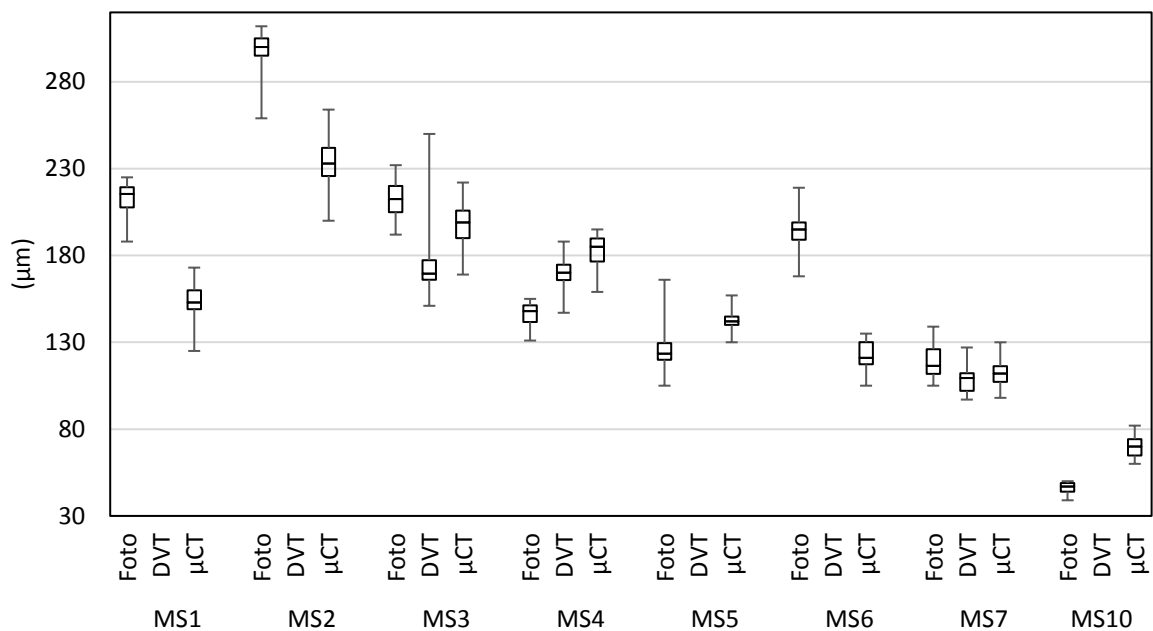


#### Probe 2

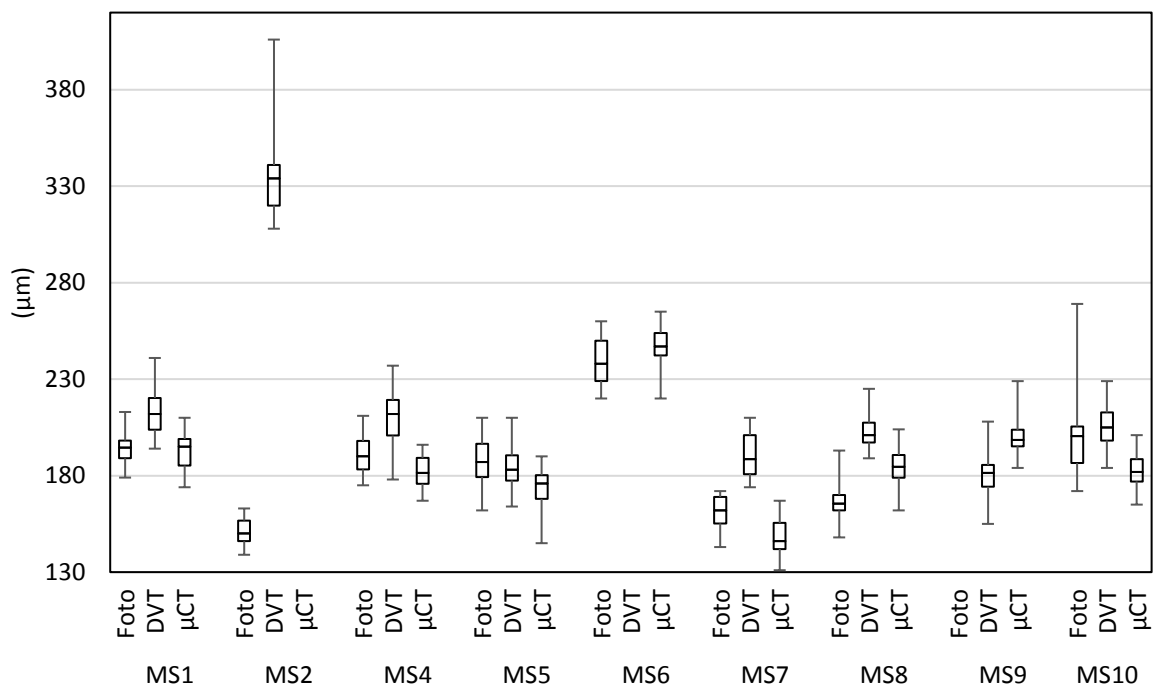


Boxplots 10: Boxplots der Proben 1 und 2 des Methodenvergleichs.

Probe 3

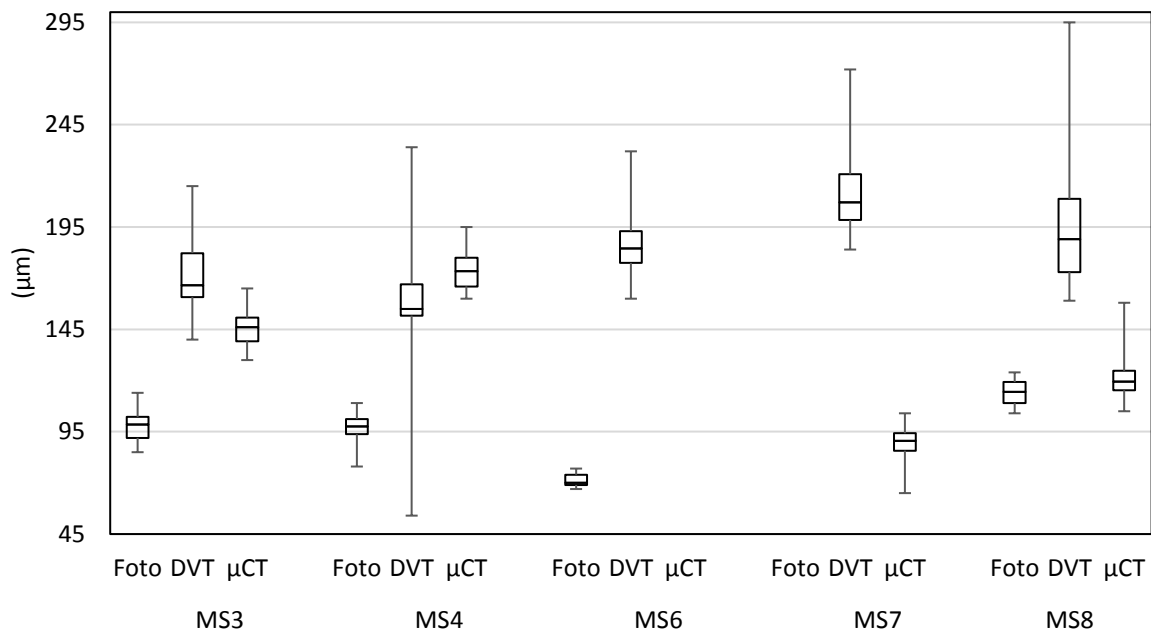


Probe 4



Boxplots 11: Boxplots der Proben 3 und 4 des Methodenvergleichs.

## Probe 5

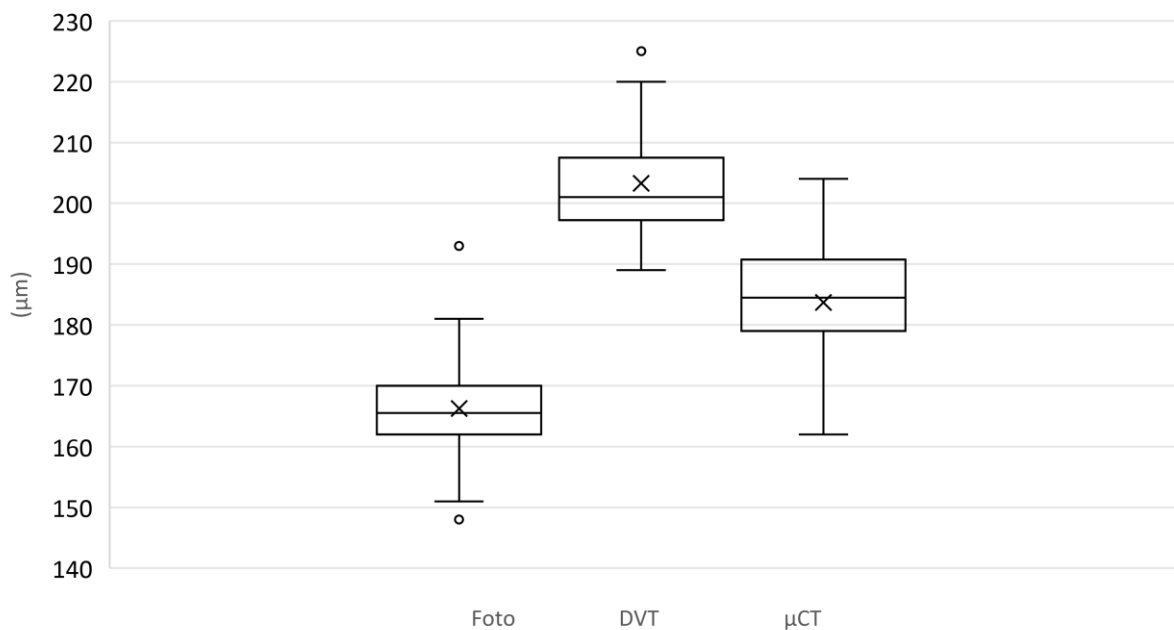


Boxplots 12: Boxplots der Probe 5 des Methodenvergleichs.

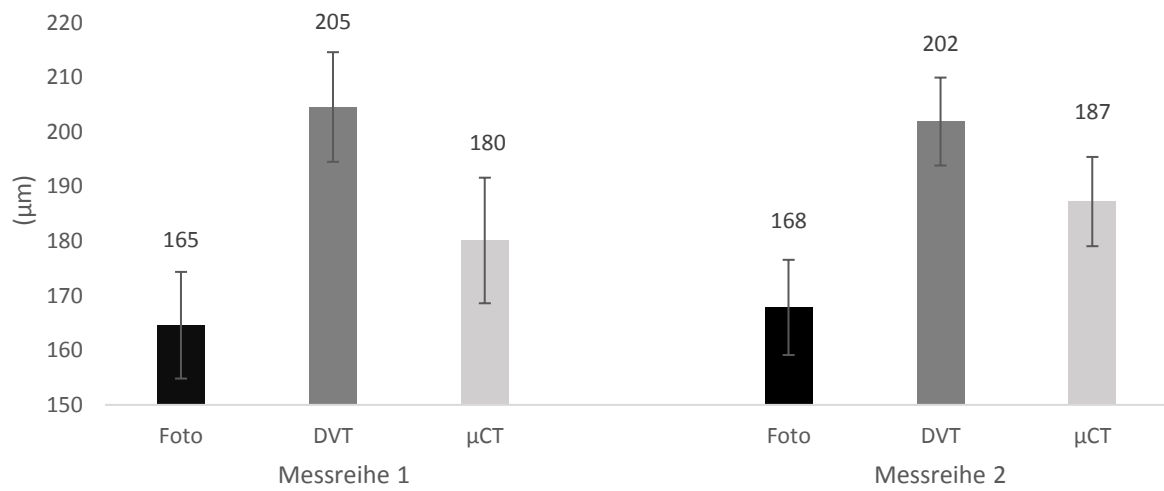
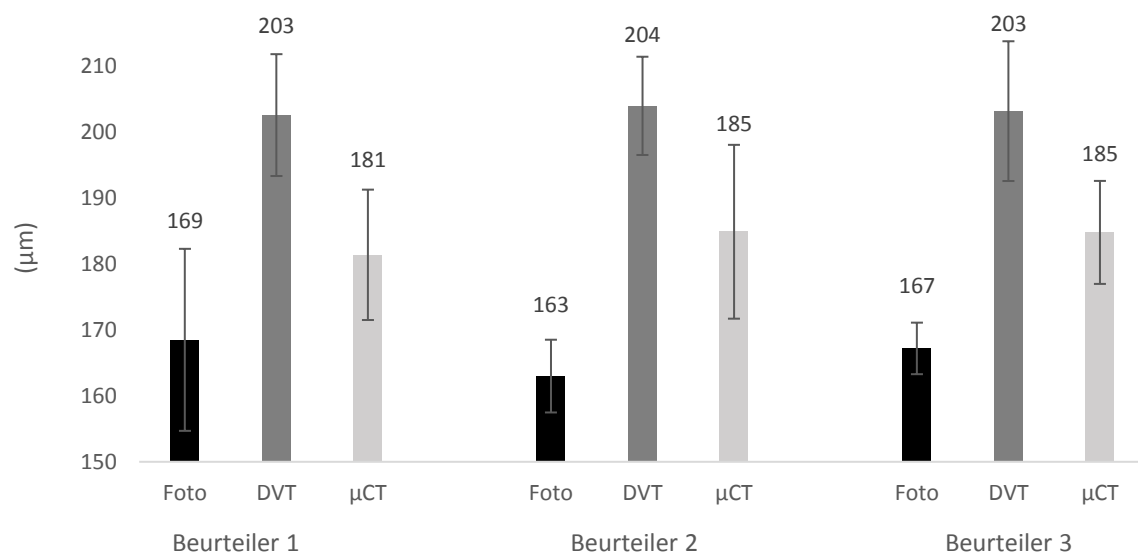
Bei Betrachtung der Messtabelle und der Boxplots des Methodenvergleichs sind die Unterschiede zwischen den drei Methoden deutlich sichtbar. Neben unterschiedlichen Mittelwerten und Medianen sind auch die Standardabweichungen bzw. Minima und Maxima der Boxplots methodenabhängig unterschiedlich groß.

Im Durchschnitt lagen die Standardabweichungen für die Methoden Foto bei  $9,7 \mu\text{m}$  und für die Methode  $\mu\text{CT}$  bei  $10,7 \mu\text{m}$ . Für die Methode DVT ergab sich insgesamt eine Standardabweichung von  $15,9 \mu\text{m}$ .

Im Speziellen wird noch einmal die Messstelle 8 der Probe 4 betrachtet. Dort lag eine Gesamtmeßanzahl von 90 vor (30 pro Methode). Zu erkennen ist, dass diese Stelle am DVT am größten gemessen wurde (203  $\mu\text{m}$ ), gefolgt vom  $\mu\text{CT}$  (184  $\mu\text{m}$ ) und dem Foto (166  $\mu\text{m}$ ) (s. Tabelle 16). Die gezeigten Boxplots (Boxplots 13) veranschaulichen den Methodenvergleich für diese Messstelle unter Einbeziehung beider Messreihen und aller Beurteiler. Die nachfolgenden Balkendiagramme stellen zum einen gesondert den Vergleich der beiden Messreihen an dieser Messstelle, unter Einbeziehung aller Beurteiler (Balkendiagramm 7) und zum anderen den Vergleich zwischen den Messungen der Beurteiler (Balkendiagramm 8), dar. Gezeigt werden die jeweiligen Mittelwerte mit kenntlich gemachten Standardabweichungen.



*Boxplots 13: Boxplots der Messungen der Messstelle 8 der Probe 4 im Methodenvergleich (Ausreißer exkludiert).*

*Balkendiagramm 7: Balkendiagramm der Messstelle 8, Probe 4 im Vergleich der Messreihen.**Balkendiagramm 8: Balkendiagramm der Messstelle 8, Probe 4 im Beurteilervergleich.*

### 4.3.2 Methodenvergleich anhand des Technical Error of Measurement

Die Berechnung des technischen Messfehlers für den Methodenvergleich schloss alle Messungen mit ein, unabhängig davon, ob ein signifikanter Unterschied zwischen den Messreihen oder den Messungen der Beurteiler vorlag. Es wurden somit bei jeder Messstelle 90 Messwerte miteinbezogen (30 je Methode). Der TEM gibt nur einen quantitativen Unterschied zwischen den Methoden an, nicht aber, bei welcher Methode größere, bzw. kleinere Werte gemessen wurden.

Der technische Fehler zwischen den Methoden Foto und DVT betrug  $36 \mu\text{m}$  bei einem relativen Fehler von 22,2 %.

Der technische Fehler zwischen den Methoden Foto und  $\mu\text{CT}$  betrug  $25 \mu\text{m}$  bei einem relativen Fehler von 16,4 %.

Der technische Fehler zwischen den Methoden DVT und  $\mu\text{CT}$  betrug  $28 \mu\text{m}$  bei einem relativen Fehler von 16,4 %.

Tabelle 17: TEM und rTEM des Methodenvergleichs.

	TEM ( $\mu\text{m}$ )	rTEM (%)
FOTO/DVT	36	22,2
FOTO/ $\mu\text{CT}$	25	16,4
DVT/ $\mu\text{CT}$	28	16,4

### 4.3.3 Methodenvergleich anhand von Punktwolken und Regression

In den drei folgenden Koordinatensystemen sind die Mittelwerte aus beiden Messreihen aller Beurteiler der einzelnen Messstellen aufgezeigt. X- und Y-Achse wurden den verschiedenen Methoden zugeordnet und sind dementsprechend gekennzeichnet. Hinzugefügt wurden lineare Trendlinien, die durch den Nullpunkt verlaufen. Die Steigungen der Trendlinien werden in Form von Geradengleichungen angegeben. Zusätzlich werden die Bestimmtheitsmaße ( $R^2$ ) gezeigt.

#### Vergleich zwischen Foto und DVT (Abbildung 16):

Die Trendlinie des Vergleichs zwischen den Mittelwerten der Messungen am Foto und am DVT hat die Steigung  $m = 1,1102$ . Das Bestimmtheitsmaß beträgt  $R^2 = 0,283$ . Drei Punkte haben einen maximalen Abstand von der Trendlinie (rot umkreist). Überhalb der Linie sind es die beiden Punkte mit den Koordinaten (151/308) und (121/277). Diese Koordinaten entsprechen der Messstelle 2 der Probe 4, sowie der Messstelle 2 der Probe 5.

Unterhalb der Trendlinie liegt der Punkt mit den Koordinaten (298/226) maximal von der Linie entfernt. Dieser Punkt entspricht der Messstelle 2 der Probe 3.

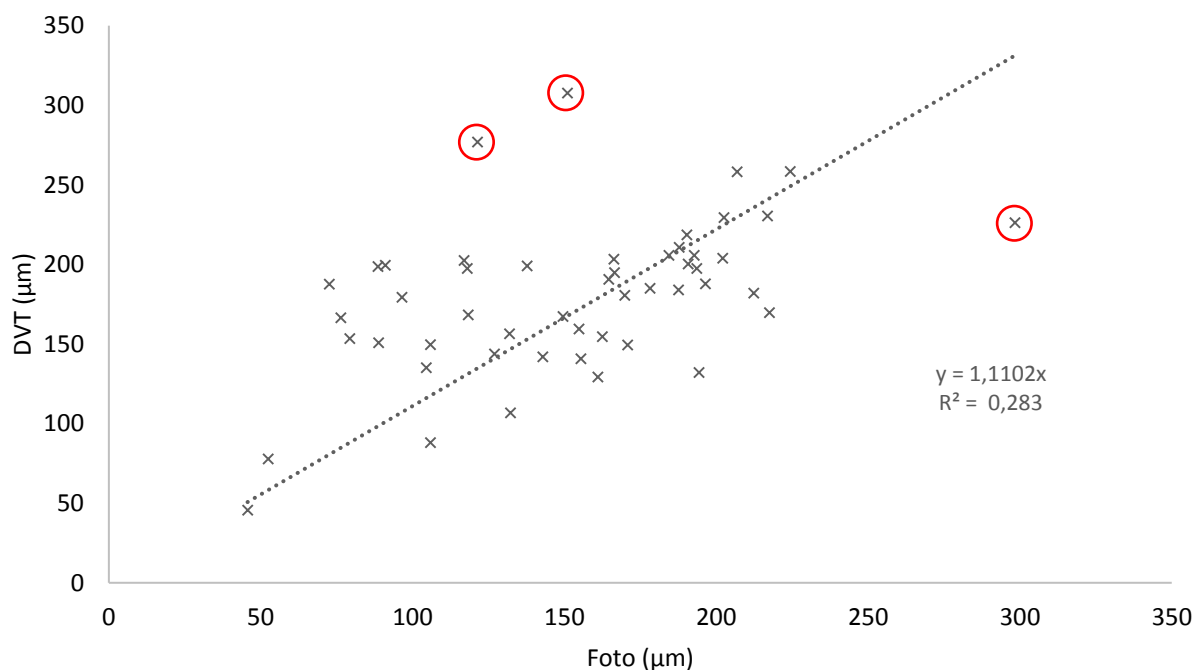


Abbildung 16: Punktwolke aus Mittelwerten der Messungen der Methoden Foto und DVT.

Vergleich zwischen Foto und  $\mu$ CT (Abbildung 17):

Der Vergleich zwischen den Mittelwerten der Messungen am Foto und  $\mu$ CT ergab eine Trendlinie mit der Steigung  $m = 1,035$  und einem Bestimmtheitsmaß von  $R^2 = 0,2831$ . Oberhalb der Trendlinie hat der Punkt mit der Koordinate (151/254) einen maximalen Abstand von der Trendlinie. Dieser Punkt entspricht der Messstelle 2 der Probe 4 (rot umkreist).

Unterhalb der Trendlinie hat der Punkt mit den Koordinaten (194/122) einen maximalen Abstand zu der Linie. Dieser Punkt entspricht der Messstelle 6 der Probe 3 (rot umkreist).

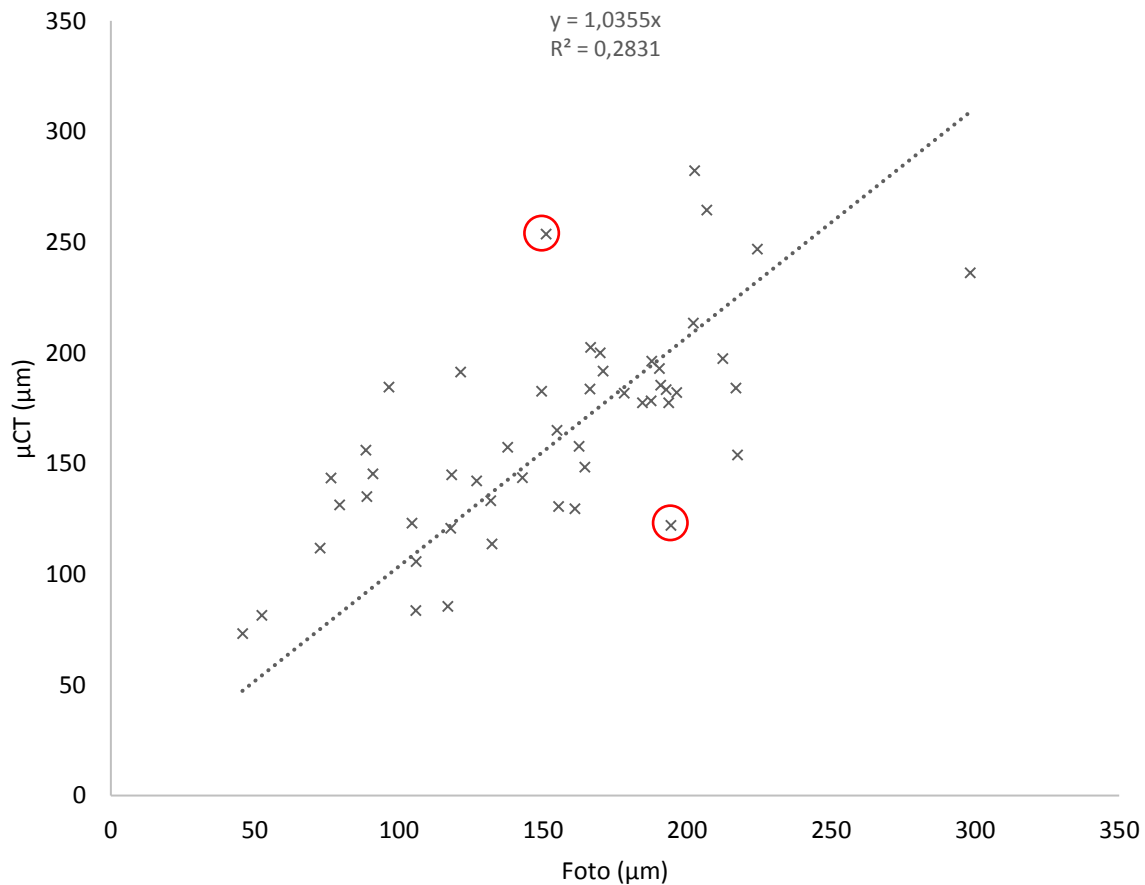


Abbildung 17: Punktwolke aus Mittelwerten der Messungen der Methoden  $\mu$ CT und Foto.



Vergleich zwischen DVT und  $\mu$ CT (Abbildung 18):

Der Vergleich zwischen den Mittelwerten der Messungen am DVT und  $\mu$ CT ergab eine Trendlinie mit der Steigung  $m = 0,9025$  und einem Bestimmtheitsmaß von  $R^2 = 0,5587$ . Oberhalb der Trendlinie hat der Punkt mit der Koordinate (229/282) einen maximalen Abstand von der Trendlinie. Dieser Punkt entspricht der Messstelle 10 der Probe 1 (rot umkreist). Unterhalb der Trendlinie hat der Punkt mit den Koordinaten (202/85) einen maximalen Abstand zu der Linie. Dieser Punkt entspricht der Messstelle 7 der Probe 5 (rot umkreist).

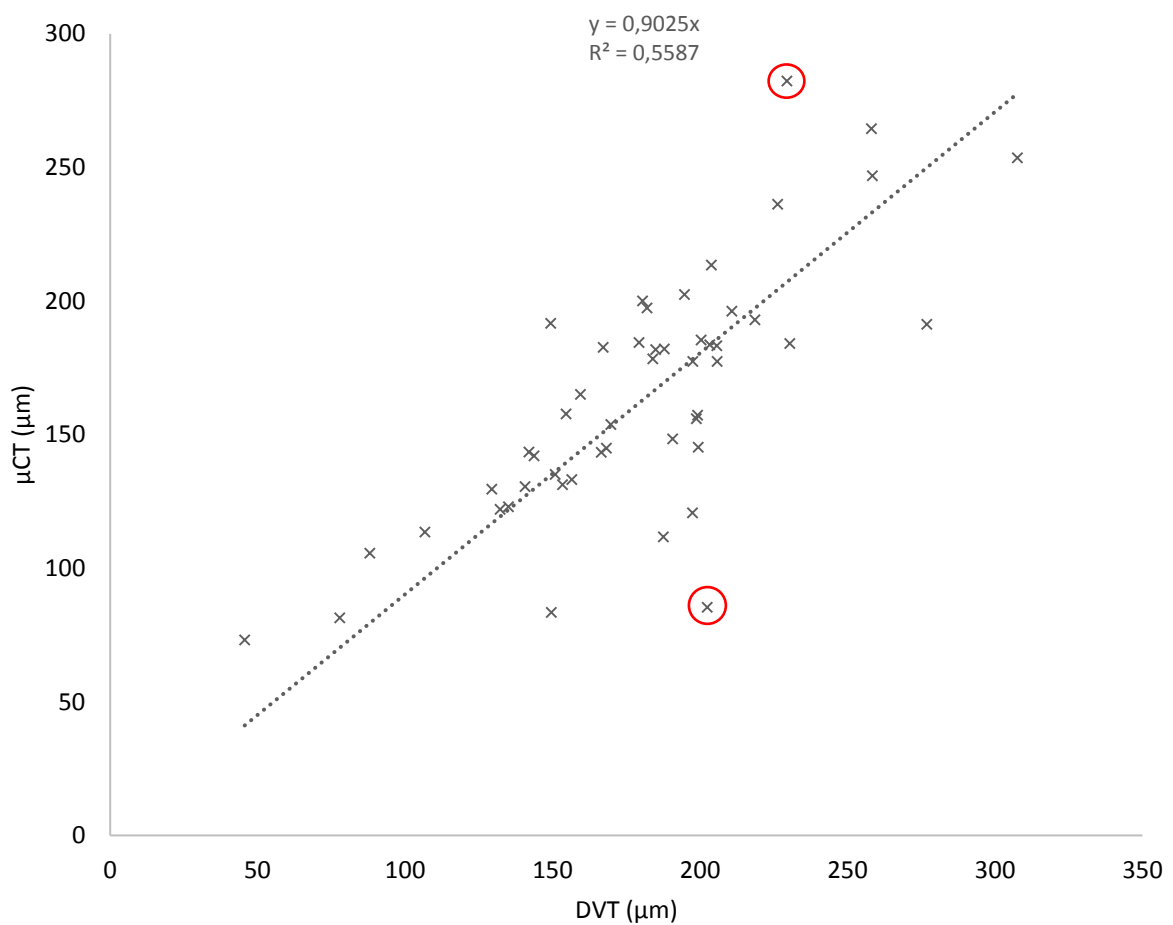


Abbildung 18: Punktwolke aus Mittelwerten der Messungen der Methoden  $\mu$ CT und DVT.

## 5 Diskussion

### 5.1 Diskussion der Methodik

In dieser Studie sollte nicht die Breite und Anatomie des Parodontalspaltes erforscht werden, sondern die Messakkuranz und Qualität des radiologischen Aufnahmeverfahrens der digitalen Volumentomografie validiert werden. Die Technologie der digitalen Volumentomografie stellt einen Paradigmenwechsel in der oralen und maxillofazialen Bildgebung dar (Alsufyani et al. 2012). Wie bereits erwähnt, lieferten viele Studien gute Ergebnisse der anatomischen Vermessung im Millimeterbereich am DVT (Lascalea et al. 2004). Durch das heutige Auflösungsvermögen moderner DVT-Geräte von 90 µm sollte es demnach möglich sein, sehr viel kleinere anatomische Strukturen, wie das parodontale Ligament, exakt zu vermessen.

Um diese These zu überprüfen, musste die Vermessung des PDLs zunächst am Humanpräparat erfolgen. Dies wurde in dieser Studie anhand hochauflösender Fotografien realisiert. Doch auch die Vermessung des PDLs an Fotografien war fehlerbehaftet, wie die Ergebnisse der Intra- und Interbeurteilerdiskrepanz zeigten. Viele Faktoren wie Belichtung, Kontrast und Auflösung konnten die Darstellung des PDLs verfälschen. Außerdem war nicht auszuschließen, dass aufgrund der Transluzens von Dentin, Zement und/ oder Knochen eine Überlagerung der dargestellten Strukturen stattfand.

Daher konnte die Qualität der Darstellung des PDLs am DVT nur in Relation zur Darstellung in der Fotografie gesetzt werden, nicht aber in Bezug zur tatsächlichen Größe des PDLs.

Die Methodik der DVT-Aufnahmen muss ebenfalls kritisch betrachtet werden. Die Aufnahmen wurden ohne mögliche Störfaktoren, wie Bewegungsartefakte oder gewebespezifische Strahlungsschwächung durch Muskelgewebe der Wange oder Zunge, durchgeführt. Dies bot eine maximale Darstellungsqualität der Proben. Ob sich diese Qualität auch am lebenden Menschen erreichen lässt, ist nicht garantiert.

Einen zusätzlichen Vergleich der Messqualität wurde mittels Messungen am µCT erreicht. µCT-Aufnahmen bieten mehr Schärfe und Kontrast als digitale Volumentomografien, was vor allem durch den Einsatz von hohen Strahlendosen und langen Aufnahmezeiten erreicht wird. Dies und vor allem die Tatsache, dass nur Objekte von ca. 2 cm mal 3 cm Größe gescannt werden können, macht diese Methode bis dato am lebenden Menschen nicht anwendbar und ist somit als vergleichende Methode rein akademisch zu betrachten.

---

Schon vor Beginn der Studie wurde vermutet, dass Vermessungen im Mikrometerbereich beurteilerabhängig sein würden. Durch die Miteinbeziehung drei Beurteiler und Durchführung der Messungen in zwei Messreihen konnte diese Vermutung überprüft werden.

## 5.2 Beurteilung der Proben

Die verwendeten Humanpräparate lagen ohne bekannte Anamnese vor. Dies bedeutete, dass eventuelle parodontale Erkrankungen die Beurteilung und Vermessung der Proben erschweren konnten. So könnte beispielsweise eine vorhandene Ankylose, also die Verwachsung von Knochen und Zahn, zu Verwirrungen bei der Messung geführt haben, sollte dies übersehen worden sein. Auch die Beschaffenheit der einzelnen anatomischen Strukturen hatten Einfluss auf Kontrast und Schärfe. Deshalb wurde in dieser Studie nicht nur eine Probe verwendet, sondern fünf. Mithilfe der Intra- und Interbeurteilerdiskrepanz ließ sich bereits erkennen, welche Proben präzise, bzw. reproduzierbar vermessen werden konnten. Geringe Abweichungen zwischen den Messreihen und den Messungen der Beurteiler ließen auf eine gut zu messende Probe schließen.

Tabelle 7 auf Seite 46 zeigt, dass die Probe 5, vermessen am DVT, für die Messungen aller Beurteiler, technische Messfehler zwischen den beiden Messreihen von 16, 18 und 40  $\mu\text{m}$  ergaben. Dieselbe Probe am Foto und  $\mu\text{CT}$  vermessen, ergab geringere technische Messfehler von maximal 17  $\mu\text{m}$ .

Eine Abhängigkeit der Messakkuranz von der Probe lässt auch die Tabelle 11 auf Seite 51 erkennen. Diese Tabelle beschreibt, wie viele Messstellen pro Probe von den Beurteilern different vermessen wurden. Während die Probe 1 am Foto an fünf Messstellen von allen drei Beurteilern ohne signifikante Differenz vermessen wurde, war dies bei der Methode  $\mu\text{CT}$  keinmal der Fall. Daraus lässt sich schließen, dass die Vermessung der Probe 1 am Foto weniger beurteilerabhängig war, als die äquivalenten Vermessungen mit der Methode  $\mu\text{CT}$ .

Im Gegensatz dazu wurde die Probe 3 mit der Methode  $\mu\text{CT}$  an mehr Messstellen (5) von allen Beurteilern indifferent vermessen, als am Foto (3).

Dieser Widerspruch zeigt, dass die Präzision und Reproduzierbarkeit der Messungen auch von der beurteilten Probe abhängig war. Durch die Nichtbetrachtung von Messstellen, die schon signifikante Abweichungen bei der Intra- und Interbeurteilerdiskrepanz ergaben, konnte die Abhängigkeit des Methodenvergleichs von der Probe auf ein Minimum reduziert werden.

Bei der Herstellung der Proben konnte keine „spiegelglatte“ Oberfläche geschaffen werden. Biologisches Material besitzt eine gewisse Eigendynamik. Beispielsweise ist der Zahn in der Alveole nicht starr verankert, sondern in alle Dimensionen beweglich (s. Kapitel 2.1). Somit war nicht auszuschließen, dass Zahn und Knochen nicht mehr exakt in einer Ebene lagen, sondern der Zahn wenige Mikrometer aus dem Knochen rausragte. Dies zeigt die Abbildung 19. Gezeigt wird der sich im Knochen befindliche Zahn mit umgebendem PDL mit skizzierten Sharpey-Fasern. Die Schnittebene des Zahnes liegt minimal höher als die des Knochens. Beim Festlegen der Messebene in den radiologischen Bildgebungen (DVT und  $\mu$ CT) wurde die Schicht verwendet, in der Zahn, Knochen und somit auch das PDL, deutlich erkennbar waren. Diese Ebene konnte wenige Mikrometer innerhalb des Zahnes liegen, um parallel des Knochens zu sein. Die Ebene, in der gemessen wurde, wird in Abbildung 19 mit der roten Linie skizziert. Die grüne Messlinie spiegelt die Messung an der Fotografie wider und ist deutlich kürzer als die rote Linie, was zu kleineren Messwerten führen würde. Die Abbildung 19 verdeutlicht, dass die in den radiologischen Bildgebungen verwendete Messebene unter bestimmten Voraussetzungen in den Fotografien gar nicht sichtbar war. Jedoch ist zu bedenken, dass die hier gezeigte Skizze, im Sinne der Verdeutlichung, übertrieben dargestellt wird. Der „Überhang“ des Zahnes über den PDL (der Bereich des Zahnes zwischen der grünen und roten Linie) wäre hier verantwortlich für die Messdifferenz. Sollte dieser nicht vorhanden sein, bzw. sehr klein sein, wären die Messebenen am Foto und DVT/ $\mu$ CT identisch.

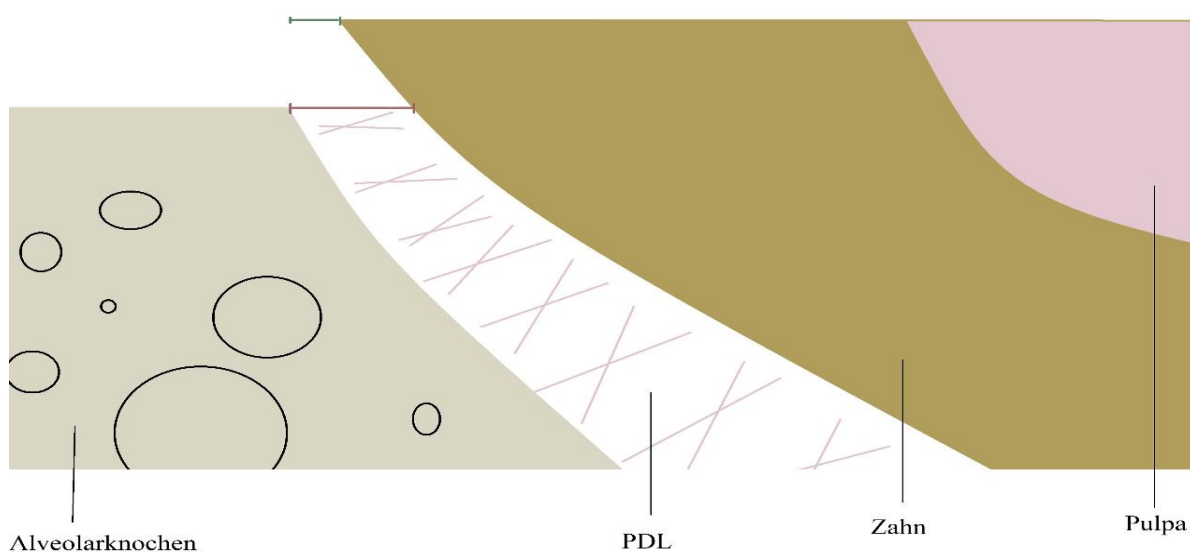


Abbildung 19: Schematische Illustration der Messproblematik am Foto und Röntgenbild.

Ebenfalls problematisch konnten Hohlräume im Trabekelwerk gewesen sein, die bis zum PDL reichten, und/oder genau in der Sägebene lokalisiert waren. Dies konnte zur Folge haben, dass der Parodontalspalt in der Sägebene gar nicht beurteilbar war. Abbildung 20 verdeutlicht diese Problematik. Dargestellt wird der Horizontalschnitt eines Zahnes. Der Parodontalspalt ist deutlich zu erkennen. Die blaue Linie skizziert die Sägebene. Genau in dieser ist, gelb umkreist, ein Hohlraum im Trabekelwerk zu erkennen. Im radiologischen Vertikalschnitt würde sich dieser als extrem großes PDL darstellen oder könnte zur Verwirrung bei der Schnittebenenfindung führen, während das PDL im Foto gar nicht bis schwer auszumachen wäre und somit die Messungen verfälschen würde. Zwar wurde bei der Festlegung der Messpunkte darauf geachtet, solche Bereiche des PDLs zu meiden, es konnte allerdings nicht ganz ausgeschlossen werden.

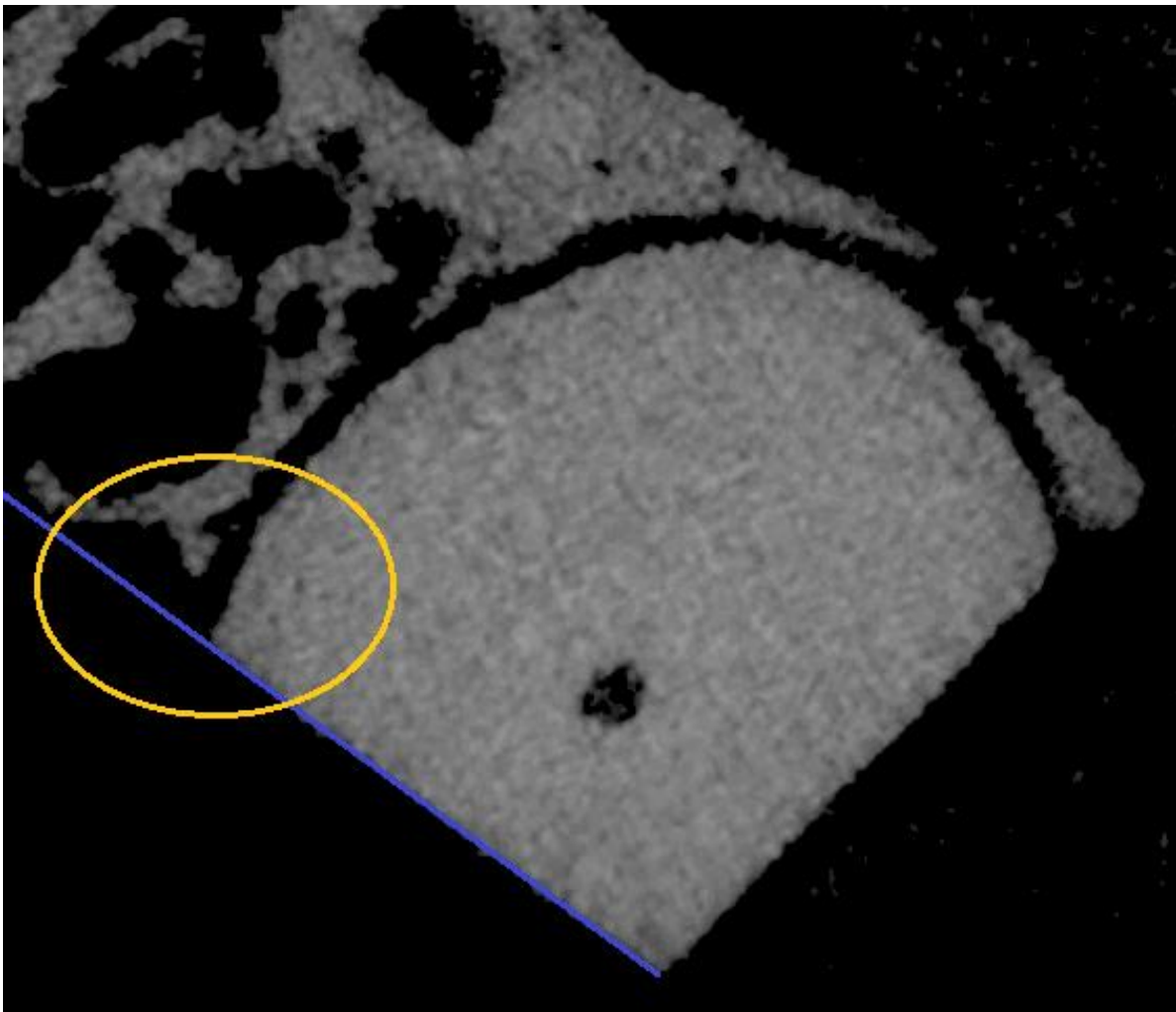


Abbildung 20: Horizontalschnitt eines Zahnes am  $\mu$ CT.

---

Ebenfalls gut zu erkennen in Abbildung 20 sind die kleinen Variationen der Alveoleninnenkortikalis, die von dem  $\mu$ CT sehr gut dargestellt wurden. In entsprechenden DVT-Aufnahmen waren diese nicht auszumachen.

Die erwähnte Tatsache, dass die Humanpräparate keine spiegelglatte Oberfläche besaßen, führte auch dazu, dass eine minimale Winkelabweichung der Röntgenmessebene zur fotografischen Messebene nicht auszuschließen war (s. Kapitel 3.3, S. 23).

Beim Foto lag die Schwierigkeit darin, dass man ein zweidimensionales Bild eines dreidimensionalen Objektes erhielt. Transluzenz von Knochen, Dentin und Schmelz konnten die Differenzierung des PDLs erschweren und verfälschen.

### **5.3 Diskussion der Ergebnisse der Intrabeurteilerdiskrepanz**

Eine mangelhafte Validierung von Messwerten durch die Intrabeurteilerdiskrepanz kann zu erheblichen statistischen Fehlern und Aussagen führen (Lewis 1999). Auch in dieser Studie zeigten sich signifikante Abweichungen zwischen den Messungen der beiden Messreihen der Beurteiler. Hierbei ließ sich klar erkennen, dass ein Messunterschied zwischen den beiden Messreihen bei der Methode DVT am häufigsten vorkam. 28,67 % der Messstellen wurden am DVT in Messreihe 1 signifikant different zur Messreihe 2 vermessen. Am Foto lag dieser Wert bei 12,67 % und für die Methode  $\mu$ CT bei 17 % (s. Tabelle 9, S.48).

Absolut bedeutete dies, dass 43 von 150 mit dem DVT betrachteten Messstellen, für einen Interbeurteiler- und Methodenvergleich nicht mehr in Betracht kamen. Bei der Methode Foto waren dies lediglich 19 Messstellen und bei der Methode  $\mu$ CT 26.

Die maximalen Unterschiede zwischen den beiden Messreihen betrugen an verschiedenen Messstellen für die Methode Foto 61  $\mu$ m, für die Methode DVT 181  $\mu$ m und für die Methode  $\mu$ CT 70  $\mu$ m.

Dieses Ergebnis wurde auch durch die Ergebnisse des TEM und rTEM bestätigt. Die Werte des TEM lagen für die Methode Foto bei jeder Probe unter 15  $\mu$ m. Die Messungen der Methode DVT ergaben höhere technische Messfehler als bei der Methode Foto oder  $\mu$ CT (s. Tabelle 7 S.46).

Die größten technischen Fehler, unabhängig von der Bildgebung, hat der Beurteiler 1 verursacht, wie in Tabelle 8 auf Seite 47 zu erkennen ist.

Im Vergleich zu den Fotografien lassen all diese Ergebnisse auf eine geringere Reproduzierbarkeit der Messungen bei radiologischer Bildgebung schließen.

Besonders das DVT bot den schwächsten Kontrast zwischen Zahn, PDL und Knochen, was die Grenzdefinierung erschwerte und möglicherweise die hohe Fehlerrate erklärt. Die Methode  $\mu$ CT weist zwar höhere Messunterschiede zwischen den beiden Messreihen auf als bei der Fotografie, diese sind dennoch geringer als bei der Methode DVT, was wohl durch höhere Schärfe und Kontrast begründet sein dürfte.

Einen Lerneffekt, dass bei den zuerst vermessenen Proben der Fehler größer war als bei den zuletzt vermessenen, ließen die Daten nicht erkennen.

#### **5.4 Diskussion der Ergebnisse der Interbeurteilerdiskrepanz**

Die Interbeurteilerdiskrepanz gibt Auskunft darüber, wie zuverlässig verschiedene Beurteiler die vorgegebenen Strecken vermessen haben. Die Abweichungen zwischen den Messungen der Beurteiler sollten möglichst gering sein. Eine geringe Abweichung lässt sich als gute Beurteilbarkeit der vermessenen Strecke interpretieren.

Unter diesen Gesichtspunkten und bei Interpretation der Ergebnisse der Interbeurteilerdiskrepanz ist festzustellen, dass die Beurteilbarkeit der Messstrecke, also der Breite des PDLs, beim Foto am besten war. Bei lediglich vier der 43 Messstellen haben alle Beurteiler signifikant unterschiedlich gemessen (9,3 % aller Messstellen). Bei den anderen 39 Messstellen haben mindestens zwei Beurteiler sehr ähnlich große Werte gemessen ( $p > 0,05$ ; Kruskal-Wallis-Test).

Die Messungen an den radiologischen Bildgebungen DVT und  $\mu$ CT ergaben, nach den erwähnten Kriterien, eine deutlich schlechtere Beurteilbarkeit. Jeweils zwölf Messstellen wurden von allen Beurteilern signifikant different vermessen. Diese Zahl entspricht dem Dreifachen des Ergebnisses der Methode Foto. Relativ gesehen sind die zwölf unterschiedlich gemessenen Messstellen jedoch bei der Methode DVT kritischer zu beurteilen, da sie 30 % aller beurteilten Messstellen ausmachten. Fast ein Drittel aller Messstellen wurden mit dem DVT von jedem Beurteiler andersgroß vermessen. Die zwölf Messstellen machten bei der Methode  $\mu$ CT einen Prozentsatz von 27,3 % aus (s. Tabelle 11, S. 51).

Die Prozentzahlen legen nahe, dass die Messungen an radiologischen Bildgebungen eher von der messenden Person abhängen, als Messungen am Foto.

Eine mögliche Erklärung für die Abweichungen zwischen den Beurteilern könnte sein, dass verschiedene Techniken zur Vermessung verwendet wurden, wenngleich diese vor Beginn der Studie erklärt und geübt wurden. Der Parodontalspalt wurde durch einen Übergang von der Farbe weiß, über grau zu schwarz von Zahn und Knochen abgegrenzt. Innerhalb dieses Übergangs hatten die Beurteiler einen gewissen Spielraum, wo sie die Grenze des PDLs setzten. Dadurch kann die gemessene Strecke verkürzt oder verlängert worden sein.

## **5.5 Diskussion der Ergebnisse des Methodenvergleichs**

Evaluiert wurden die Methoden in Bezug auf Akkuranz und Reproduzierbarkeit. Die Anforderungen waren wie folgt zu stellen:

1. Die Messungen an einer Bildgebung müssen von einem Beurteiler zuverlässig reproduzierbar sein (geringe Abweichung der Messwerte zwischen den Messreihen)
2. Mehrere Beurteiler messen identische Strecken ohne signifikante Differenz (geringe Abweichungen zwischen den Messreihen der drei Beurteiler)

Diese beiden Punkte boten einen Ausdruck über die Qualität der Methoden.

Ein wichtiger Indikator für die Präzision der Messungen bot die Standardabweichung. Je größer diese war, desto größer war die Streuung der Messungen. Bei Betrachtung der deskriptiven Statistik der Intraurteilerdiskrepanz (Tabelle 3, S.40 ff.) war festzustellen, dass die Standardabweichungen der Messungen der Methode DVT deutlich höher waren als die der Methoden Foto und  $\mu$ CT. Im Durchschnitt betrugen sie für die Methode Foto 8,25  $\mu$ m, für die Methode  $\mu$ CT 8,1  $\mu$ m und für die Methode DVT 11,2  $\mu$ m. Ähnliche Ergebnisse liefert auch die Tabelle 16 auf den Seiten 64 ff., in der die Messwerte und Standardabweichungen im Methodenvergleich dargestellt werden. Auch hier war die Standardabweichung der Messungen am DVT mit 15,9  $\mu$ m am höchsten (Vergleichend: Foto: 9,7  $\mu$ m und  $\mu$ CT: 10,7  $\mu$ m). Dies bedeutet, dass die Messungen am DVT einer höheren Streuung unterlagen als die Messungen am Foto und  $\mu$ CT und damit weniger präzise waren.

Einen wichtigen Hinweis auf die Reproduzierbarkeit der Messungen boten der technische



---

Messfehler (TEM) und die induktive Statistik. Ein möglichst kleiner technischer Messfehler deutete auf gute Reproduzierbarkeit der Messungen hin. Tabelle 7 auf Seite 46 (*TEM- und rTEM- Übersicht der Intrabeurteilerdiskrepanz*) lässt sich so interpretieren, dass die Reproduzierbarkeit der Messwerte bei der Methode Foto und  $\mu$ CT besser war als bei der Methode DVT. Die TEM-Werte lagen für die Methoden Foto und  $\mu$ CT zwischen 5 und 17  $\mu$ m (Foto: Gesamt TEM 11,1  $\mu$ m, rTEM 7,2 %/  $\mu$ CT: Gesamt TEM 11,1  $\mu$ m, rTEM 6,8 %), während die technischen Messfehler bei der Methode DVT bei bis zu 40  $\mu$ m (Gesamt TEM 17,3  $\mu$ m, rTEM 10,0 %) lagen. Auch die Tabelle 8 auf Seite 47 zeigt, dass der technische Messfehler zwischen den Messreihen von allen drei Beurteilern beim DVT am höchsten war. Ebenfalls am größten war der technische Messfehler für den Vergleich der Messreihen der Beurteiler untereinander beim DVT (s. Tabelle 15, S. 62).

Auch die Ergebnisse der induktiven Statistik ließen den Schluss zu, dass die Messungen am DVT weniger zuverlässig reproduzierbar waren, als die Messungen am Foto und  $\mu$ CT. Dies verdeutlicht die Tabelle 9 auf Seite 48, die zeigt, dass 28,7 % aller Messstellen am DVT in der Messreihe 1 signifikant different zu Messreihe 2 vermessen wurden. Bei der Methode Foto waren dies nur 12,7 % der Messstellen und bei der Methode  $\mu$ CT 17,3 %.

Auch die Tabelle 11 auf Seite 51, die sich darauf bezieht, ob die einzelnen Beurteiler die Messstellen signifikant different voneinander vermessen haben, zeigt, dass am DVT mehr Messstellen unterschiedlich vermessen wurden als bei den Methoden Foto und  $\mu$ CT (12 der 39 zu beurteilenden Messstellen wurden von allen drei Beurteilern am DVT abweichend voneinander vermessen).

Die Ergebnisse, sowohl der deskriptiven, als auch der induktiven Statistik und die Ergebnisse des TEM ließen sich so interpretieren, dass die Messgenauigkeit und Reproduzierbarkeit bei der Methode DVT am schlechtesten war und stärker von dem Beurteiler abhing, als bei den Methoden Foto und  $\mu$ CT.

Relevant in dieser Untersuchung war die Frage, ob der Parodontalspalt an denselben Stellen mit unterschiedlichen Bildgebungen gleich groß vermessen wurde. Die Antwort auf diese Frage lässt sich abhängig von dem Kapitel 4.3.1, Seite 63 ff. diskutieren. Diese Ergebnisse lassen sich in zwei Abschnitte unterteilen:

1. Zunächst wurden die Messwerte der einzelnen Methoden, unter Berücksichtigung der Ergebnisse der Intra- und Interbeurteilerdiskrepanz, nebeneinandergestellt.
2. Anschließend erfolgte ein Vergleich anhand von Punktwolken und dem TEM, bei denen alle Messwerte, unabhängig von den Ergebnissen der Intra- und Interbeurteilerdiskrepanz, miteinbezogen wurden.

Zu 1.:

Wie beschrieben, lag gemittelt die Differenz zwischen den Methoden Foto und DVT bei 31  $\mu\text{m}$  (s. Kapitel 4.3.1, S. 63 ff.). Im Schnitt wurde also am DVT der Parodontalspalt um „nur“ 31  $\mu\text{m}$  größer vermessen als am Foto. Dies mag auf den ersten Blick akzeptabel erscheinen, da dieser Wert unterhalb der Voxelgröße von 90  $\mu\text{m}$  des DVTs lag.

Bei detaillierter Betrachtung der deskriptiven Statistik (Tabelle 16, S. 64) war allerdings festzustellen, dass bei einzelnen Messstellen die Differenz bei bis zu 188  $\mu\text{m}$  lag, was mehr als das doppelte der Voxelgröße ausmachte. Eine mögliche Erklärung könnte sein, dass kleinste Variationen der Struktur der Alveoleninnenkortikalis mit dem DVT, im Gegensatz zum  $\mu\text{CT}$  und der Fotografie nicht, bzw. insuffizient dargestellt wurden.

Eine Differenz der gemittelten Messwerte von Foto zu  $\mu\text{CT}$  war anhand der deskriptiven Statistik nicht auszumachen. Im Mittel wurde also der Parodontalspalt am Foto genauso groß gemessen wie am  $\mu\text{CT}$ . Dass trotzdem am DVT im Schnitt um 15  $\mu\text{m}$  größer vermessen wurde als am  $\mu\text{CT}$  erscheint zunächst paradox, hängt aber damit zusammen, dass aufgrund der „Ausortierung“ fehlerbehafteter Messstellen durch die Intra- und Interbeurteilerdiskrepanz für die Vergleiche der Methoden unterschiedlich viele, bzw. andere Messstellen zur Verfügung standen.

Auf den Seiten 68 f. wurde im speziellen die Messstelle 8 der Probe 4 dargestellt. Diese Messstelle wurde von allen Beurteilern in allen drei Bildgebungen ohne signifikante Differenz vermessen. Daher konnten für jede Methode 30 Messwerte in den Vergleich miteinbezogen werden. Dies war bei keiner anderen Messstelle der Fall. Somit war dieser Messstelle eine besondere Relevanz zuzumessen. Die Messergebnisse dieser Messstelle bestätigten den Gesamteindruck des Methodenvergleichs. Auch bei dieser Messstelle wurden mit der Methode DVT die größten Strecken gemessen. Die Messungen am DVT waren durchschnittlich um

---

36  $\mu\text{m}$  größer als die entsprechenden Messungen beim Foto. Die Differenz zwischen den Messungen am DVT und  $\mu\text{CT}$  lag bei 19  $\mu\text{m}$  und die Differenz zwischen den Methoden Foto und  $\mu\text{CT}$  betrug 18  $\mu\text{m}$ .

Insgesamt lässt sich daraus schließen, dass über viele Messungen der Unterschied zwischen Messungen am DVT, Foto und  $\mu\text{CT}$  gering war, der Unterschied an einzelnen Messstellen jedoch erheblich sein konnte.

Zu 2.:

Um festzustellen, ob auch unter Einbeziehung aller Messstellen, also auch die, die von einem Beurteiler in Messreihe 2 signifikant different zur Messreihe 1 vermessen wurden oder die von den Beurteilern unterschiedlich voneinander vermessen wurden, ein Unterschied zwischen den Messungen an den drei Methoden vorlag, wurden Punktwolken mit allen Mittelwerten der Messreihen erstellt. Anhand der Trendlinien, die eine bestimmte Steigung besaßen, ließ sich ablesen, ob und mit welcher Methode, der Parodontalspalt größer vermessen wurde.

Die Steigung „ $m$ “ betrüge im Idealfall, also wenn exakt identisch vermessen worden wäre,  $m = 1$ . Je größer die Differenz zu 1 ist, desto größer die Abweichung der Messungen der verglichenen Methoden.

Zusätzlich zur Steigung wurde das Bestimmtheitsmaß ( $R^2$ ) angegeben, welches idealerweise ebenfalls 1 betragen sollte (Alle Punkte lägen auf der Trendlinie; keine Streuung der Punkte). Die Trendlinie des Vergleichs Foto zu DVT (Abbildung 16, S.71) hat die Steigung  $m = 1,102$  und ist somit größer als 1. Die Messwerte des DVTs sind auf der Y-Achse aufgetragen. Daraus lässt sich schließen, dass am DVT größere Werte gemessen wurden als am Foto.

Die Punktwolke des Vergleichs Foto zu  $\mu\text{CT}$  (Abbildung 17, S. 72) ergab eine Trendlinie mit der Steigung  $m = 1,0355$ . Dieser Wert weicht nur minimal von 1 ab und zeigt, dass die Abweichungen der Messungen am Foto und  $\mu\text{CT}$  nur minimal voneinander waren.

Die Steigung der Trendlinie des Vergleichs von  $\mu\text{CT}$  zu DVT betrug 0,9025 und war somit kleiner als 1 (Abbildung 18, S. 73). Die Werte des DVTs waren auf der X-Achse aufgetragen. Dies ließ den Schluss zu, dass am DVT größere Werte gemessen wurden als am  $\mu\text{CT}$ .

Bei Betrachtung der Bestimmtheitsmaße fiel auf, dass sowohl der Vergleich von Foto zu DVT, als auch der Vergleich von Foto zu  $\mu\text{CT}$  ein Bestimmtheitsmaß von 0,283 ergab, was auf eine hohe Streuung hindeutet. Der Vergleich von DVT zu  $\mu\text{CT}$  hingegen ergab ein

Bestimmtheitsmaß von 0,559 und deutet auf eine geringere Streuung der Messwerte hin.

Auch der technische Messfehler des Methodenvergleichs (Tabelle 17, S. 70) zeigte, dass die Messdifferenz zwischen den Methoden Foto und DVT mit 36  $\mu\text{m}$  den Ergebnissen der deskriptiven Statistik entsprachen. Der geringste Fehler ergab sich zwischen den Messungen der Methoden Foto und  $\mu\text{CT}$  (25  $\mu\text{m}$ ), wenngleich der Fehler zwischen den Methoden DVT und  $\mu\text{CT}$  nur minimal größer war (28  $\mu\text{m}$ ).

Zusammenfassend stellte sich der Methodenvergleich so dar, dass die Messungen am DVT im Vergleich zu denen am Foto und  $\mu\text{CT}$  zu groß ausfielen und stellenweise erheblich von den Messungen der anderen beiden Methoden abwichen. Zusätzlich waren sie abhängiger von dem Beurteiler. Die Reproduzierbarkeit war bei den Methoden Foto und  $\mu\text{CT}$  größer als bei der Methode DVT. Dennoch lagen die Differenzen zwischen den Methoden unterhalb der Voxelgröße des DVTs von 90  $\mu\text{m}$ .

## **6 Praktischer Bezug und Schlussfolgerung**

Die Validierung der Akkuranz digitaler Volumentomografien spielt in vielen Teilbereichen der Zahn-, Mund- und Kieferkrankheiten sowie in der Forschung eine wichtige Rolle. Es gibt Unmengen an Publikationen, die sich der Technik der dreidimensionalen Bildgebung bedienen. Besonders das Erstellen von Finite-Elemente-Modellen (FEM-Modelle) zur Quantifizierung von durch kieferorthopädischen Geräten induzierten Stress auf das PDL, basiert auf aus dreidimensionalen Bildgebungen gewonnenen Daten (Meyer et al. 2010; Qian et al. 2001). Bisher wurden die Aufnahmen meist mit strahlungsintensiven Dental-CTs in vitro durchgeführt (Holberg et al. 2014). Aufgrund der geringeren Strahlendosis bei einer DVT-Aufnahme und der immer besser werdenden Auflösung, könnten in Zukunft FEM-Modelle auf Basis von bei lebenden Patienten durchgeführten DVT-Aufnahmen erstellt werden. Die geforderte Akkuranz bestätigt diese Studie teilweise. Nicht nur in der Kieferorthopädie, sondern vor allem in der Chirurgie und Endodontie findet die digitale Volumentomografie immer mehr Anwendung. Während die Chirurgie besonders den Vorteil der räumlichen Darstellung von Weisheitszähnen oder für die Positionierung von Implantaten nutzt (Neugebauer et al. 2008), haben Endodontologen ein großes Interesse an der dreidimensionalen Darstellung von

Wurzelkanälen, Wurzelfrakturen und apikalen Parodontitiden (Patel et al. 2015). Uraba et al. fanden 2016 heraus, dass Entzündungen an der Wurzelspitze an DVT-Aufnahmen früher erkennbar waren als in Zahnfilmen (Uraba et al. 2016). Weitere Untersuchungen im Bereich der Endodontie bezüglich Wurzelkanalmorphologie oder Bestimmung des Dentinvolumens nach Entfernung von abgebrochenen Instrumentenspitzen fanden erfolgreich mithilfe von DVT-Aufnahmen statt (Caputo et al. 2016; Garg and Grewal 2016). Diese jüngeren Studien zeigen, dass mit dem DVT immer kleinere Strukturen gemessen werden.

Die Ergebnisse der vorliegenden Studie zeigen jedoch, dass Messungen, bzw. Distanzbestimmungen an DVT-Aufnahmen im Mikrometerbereich kritisch zu betrachten sind. Eine Validierung der Messungen durch Bestimmung der Intra- und Interbeurteilerdiskrepanz ist dringend zu empfehlen.

Auch deuten die Ergebnisse der vorliegenden Studie darauf hin, dass die Benutzer dazu neigten, am DVT zu große Strecken zu messen.

## **7 Zusammenfassung**

Ziel dieser statistisch geprägten Studie war es, die Akkuranz und Qualität von Messungen an digitalen Volumentomografien zu validieren. Außerdem sollte dieses Verfahren mit zwei vermeintlich präziseren bildgebenden Verfahren, der Fotografie und der Mikrocomputertomografie, verglichen werden. Als zu messende anatomische Struktur wurde das parodontale Ligament (PDL) menschlicher Zähne gewählt.

Es wurden fünf, im Knochen gefasste Zähne aus der Mandibula einer menschlichen Leiche gesägt. Die Zahnkronen wurden entfernt und die Zähne entlang der Längsachse zersägt, um den Parodontalspalt freizulegen. Von den zehn Zahnhälften wurden fünf subjektiv als „gut“ beurteilte Zahnhälften ausgewählt. Die resultierenden fünf „Zahn-Knochen-Blöcke“ wurden mit beigefügten Keramiksplittern als Referenzkörper in dem Silikon „Optosil“ fixiert.

Nach Herstellung hochauflösender Fotografien der fünf Proben wurden DVT- und  $\mu$ CT-Aufnahmen der Proben angefertigt. Pro Probe wurden zehn Messstellen entlang des PDLs definiert und mithilfe aufwendiger Techniken und unter Zuhilfenahme der Keramiksplitter sichergestellt, dass die Messstellen in jeder Bildgebung identisch waren. Drei Beurteiler führten unabhängig voneinander zwei Messreihen in einem Abstand von zwei Monaten durch.

Jede Messreihe umfasste die fünfmalige Messung jeder Messstelle pro Probe und Methode.

---

Hieraus ergaben sich insgesamt 4500 Messungen. Mithilfe des statistischen Programms SPSS wurden die Daten ausgewertet. Methodenabhängige Fehler zwischen den Messungen der beiden Messreihen wurden beurteilt (Intrabeurteilerdiskrepanz), sowie Fehler zwischen den Messungen der drei Beurteiler (Interbeurteilerdiskrepanz). Zudem wurden die Messungen der drei Methoden miteinander verglichen (Methodenvergleich). Zum Einsatz kam die deskriptive Statistik mit dem Programm Microsoft Excel 2016, 64 bit, die induktive Statistik und die Berechnung des technischen Messfehlers (TEM).

Die Auswertung der statistischen Verfahren zeigte, dass die Messungen an den DVT-Aufnahmen im Vergleich zu den beiden anderen Methoden, die höchste Abweichung zwischen den beiden Messreihen aufwiesen. Dies ließ auf eine geringere Reproduzierbarkeit der Messungen deuten, was durch eine „unschärfere“ Darstellung des Parodontalspalts erklärbar sein dürfte. Auch die Abweichung zwischen den Messungen der drei Beurteiler war bei der Methode DVT am größten. Daraus ließ sich schließen, dass Messungen am DVT abhängiger von dem jeweiligen Beurteiler sind, als Messungen am Foto oder  $\mu$ CT.

Der Vergleich der äquivalenten Messungen am DVT, Foto und  $\mu$ CT zeigte, dass im Schnitt der Parodontalspalt am DVT um 31  $\mu$ m größer vermessen wurde. Dieser Wert beträgt ca. ein Drittel der Voxelgröße der DVT-Aufnahmen und ist somit akzeptabel.

Insgesamt wurde das PDL mithilfe des DVTs zufriedenstellend dargestellt. Die Messung dessen ließ allerdings aufgrund der im Vergleich zum  $\mu$ CT, mit einer Voxelgröße von 18  $\mu$ m, unscharfen Darstellungen, viel Spielraum. Daher sollten in zukünftigen Studien Messungen an DVT-Aufnahmen stets kritisch hinterfragt und überprüft werden.

---

## 8 Literaturverzeichnis

- Alsufyani NA, Flores-Mir C, Major PW (2012). Three-dimensional segmentation of the upper airway using cone beam CT: a systematic review. *Dentomaxillofac Radiol*; 41(4):276-284.
- Arai Y, Tammsalo E, Iwai K, Hashimoto K, Shinoda K (1999). Development of a compact computed tomographic apparatus for dental use. *Dentomaxillofacial Radiology* (28):245 ff.
- Beertsen W, McCulloch CA, Sodek J (1997). The periodontal ligament: a unique, multifunctional connective tissue. *Periodontol* 2000; 13:20-40.
- Bernhard Böhm, Hirschfelder U (2000). Localization of Lower Right Molars in a Panoramic Radiograph, Lateral Cephalogram and Dental CT. *Journal of Orofacial Orthopedics*; Volume 61(4):237–245.
- Burstone (1962). The Biomechanics of tooth movement. B.S. Kraus, R.A. Riedel (Eds.) *Vistas in orthodontics*:197-213.
- Buzug TM (2004). Einführung in die Computertomographie. Springer-Verlag Berlin Heidelberg GmbH.
- Caputo BV, Noro Filho GA, de Andrade Salgado DM, Moura-Netto C, Giovani EM, Costa C (2016). Evaluation of the Root Canal Morphology of Molars by Using Cone-beam Computed Tomography in a Brazilian Population: Part I. *J Endod*; 10.1016/j.joen.2016.07.026.
- Cohnen M, Kemper J, Mobes O, Pawelzik J, Modder U (2002). Radiation dose in dental radiology. *Eur Radiol*; 12(3):634-637.
- Coppenrath E, Draenert F, Lechel U, Veit R, Meindl T, Reiser M, Mueller-Lisse U (2008). [Cross-sectional imaging in dentomaxillofacial diagnostics: dose comparison of dental MSCT and NewTom 9000 DVT]. *Rofo*; 180(5):396-401.
- De Clerck NM, Meurrens K, Weiler H, Van Dyck D, Van Houtte G, Terpstra P, Postnov AA (2004). High-resolution X-ray microtomography for the detection of lung tumors in living mice. *Neoplasia*; 6(4):374-379.
- Garg H, Grewal MS (2016). Cone-beam Computed Tomography Volumetric Analysis and Comparison of Dentin Structure Loss after Retrieval of Separated Instrument by Using Ultrasonic EMS and ProUltra Tips. *J Endod*; 10.1016/j.joen.2016.06.016.
- Gonzales C, Hotokezaka H, Arai Y, Ninomiya T, Tominaga J, Jang I, Hotokezaka Y, Tanaka M, Yoshida N (2009). An in vivo 3D micro-CT evaluation of tooth movement after the application of different force magnitudes in rat molar. *Angle Orthod*; 79(4):703-714.
- Goz GR, Rahn BA, Schulte-Monting J (1992). The effects of horizontal tooth loading on the circulation and width of the periodontal ligament--an experimental study on beagle dogs. *Eur J Orthod*; 14(1):21-25.

- Heidemann D (1997). Parodontologie, Praxis der Zahnheilkunde 4. Parodontologie, Praxis der Zahnheilkunde 4.]: Urban & Schwarzenberg München Wien Baltimore; p. 3
- Hirschfelder U (2008). Radiologische 3D-Diagnostik in der Kieferorthopädie (CT/DVT). Stellungnahme DGKFO.
- Holberg C, Steinhauser S, Geis P, Rudzki-Janson I (2005). Cone-beam computed tomography in orthodontics: benefits and limitations. *J Orofac Orthop*; 66(6):434-444.
- Holberg C, Rudzki-Janson I, Wichelhaus A, Winterhalder P (2014). Periodontal ligament strain induced by different orthodontic bracket removal techniques: nonlinear finite-element comparison study. *J Orofac Orthop*; 75(4):287-298.
- Hounsfield GN (1973). Computerized Transverse Axial Scanning (Tomography): Part I. Description of System. . *Brit. J. Radiol*; (46 ):1016-1022.
- Idiyatullin D, Garwood M, Gaalaas L, Nixdorf DR (2016). Role of MRI for detecting micro cracks in teeth. *Dentomaxillofac Radiology*; 45(7):20160150.
- Kamburoglu K, Murat S, Kolsuz E, Kurt H, Yuksel S, Paksoy C (2011). Comparative assessment of subjective image quality of cross-sectional cone-beam computed tomography scans. *J Oral Sci*; 53(4):501-508.
- Kapila S, Conley RS, Harrell WE, Jr. (2011). The current status of cone beam computed tomography imaging in orthodontics. *Dentomaxillofac Radiol*; 40(1):24-34.
- Kumar V, Ludlow J, Soares Cevidanes LH, Mol A (2008). In vivo comparison of conventional and cone beam CT synthesized cephalograms. *Angle Orthod*; 78(5):873-879.
- Lascalea CA, Panella J, Marques MM (2004). Analysis of the accuracy of linear measurements obtained by cone beam computed tomography (CBCT-NewTom). *Dentomaxillofac Radiol*; 33(5):291-294.
- Lewis SJ (1999). Quantifying measurement error. In: Anderson S (Ed.). *Current and recent research in osteoarchaeology 2: Proceedings of the 4th, 5th and 6th meetings of the Osteoarchaeological Research Group.*]: Oxbow Books; pp. 54-55.
- Loubele M, Van Assche N, Carpentier K, Maes F, Jacobs R, van Steenberghe D, Suetens P (2008). Comparative localized linear accuracy of small-field cone-beam CT and multislice CT for alveolar bone measurements. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*; 105(4):512-518.
- Meyer BN, Chen J, Katona TR (2010). Does the center of resistance depend on the direction of tooth movement? *Am J Orthod Dentofacial Orthop*; 137(3):354-361.
- Mohan R, Singh A, Gundappa M (2011). Three-dimensional imaging in periodontal diagnosis – Utilization of cone beam computed tomography. *Journal of Indian Society of Periodontology*; 15(1):11-17.
- Mueller HP (2004). *Periodontology: The Essentials*. Thieme.
- Mueller., Martorell. (1988). *Reliability and accuracy of measurement*. Human Kinetics Books. :83-86.



- 
- Naser AZ, Mehr BB (2013). A comparative study of accuracy of linear measurements using cone beam and multi-slice computed tomographies for evaluation of mandibular canal location in dry mandibles. *Dent Res J (Isfahan)*; 10(1):15-19.
- Neugebauer J, Shirani R, Mischkowski RA, Ritter L, Scheer M, Keeve E, Zoller JE (2008). Comparison of cone-beam volumetric imaging and combined plain radiographs for localization of the mandibular canal before removal of impacted lower third molars. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*; 105(5):633-643.
- Patel S, Durack C, Abella F, Shemesh H, Roig M, Lemberg K (2015). Cone beam computed tomography in Endodontics - a review. *Int Endod J*; 48(1):3-15.
- Perini. TA, Oliveira. GLd, Ornellas. JdS, Oliveira. FPd (2005). Technical error of measurement in anthropometry. *Rev Bras Med Esporte*; tem(Vol. 11, Nº 1):86-90.
- Pilon JJ, Kuijpers-Jagtman AM, Maltha JC (1996). Magnitude of orthodontic forces and rate of bodily tooth movement. An experimental study. *Am J Orthod Dentofacial Orthop*; 110(1):16-23.
- Qian H, Chen J, Katona TR (2001). The influence of PDL principal fibers in a 3-dimensional analysis of orthodontic tooth movement. *Am J Orthod Dentofacial Orthop*; 120(3):272-279.
- Renggli HH, Mühlemann HR, Rateitschak KH (1984). *Parodontologie*. Thieme.
- Romero-Delmastro A, Kadioglu O, Currier GF, Cook T (2014). Digital tooth-based superimposition method for assessment of alveolar bone levels on cone-beam computed tomography images. *Am J Orthod Dentofacial Orthop*; 146(2):255-263.
- Scarfe WC, Farman AG, Sukovic P (2006). Clinical applications of cone-beam computed tomography in dental practice. *J Can Dent Assoc*; 72(1):75-80.
- Schnelle C (2001). Vergleich der Strahlenexposition bei der Digitalen-Volumen-Tomographie, der Panoramaschichtaufnahme und der Computertomographie.
- Schulze D, Heiland M, Thurmann H, Adam G (2004). Radiation exposure during midfacial imaging using 4- and 16-slice computed tomography, cone beam computed tomography systems and conventional radiography. *Dentomaxillofac Radiol*; 33(2):83-86.
- Shapiro. SS, Wilk. MB (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*; Vol. 52,(No. 3/4):591-611.
- Silva MA, Wolf U, Heinicke F, Bumann A, Visser H, Hirsch E (2008). Cone-beam computed tomography for routine orthodontic treatment planning: a radiation dose evaluation. *Am J Orthod Dentofacial Orthop*; 133(5):640 e1-5.
- Suomalainen A, Vehmas T, Kortensniemi M, Robinson S, Peltola J (2008). Accuracy of linear measurements using dental cone beam and conventional multislice computed tomography. *Dentomaxillofac Radiol*; 37(1):10-17.
- Uraba S, Ebihara A, Komatsu K, Ohbayashi N, Okiji T (2016). Ability of Cone-beam Computed Tomography to Detect Periapical Lesions That Were Not Detected by Periapical Radiography: A Retrospective Assessment According to Tooth Group. *J Endod*; 42(8):1186-1190.
-

- Weinberg SM, Scott NM, Neiswanger K, Marazita ML (2005). Intraobserver error associated with measurements of the hand. *Am J Hum Biol*; 17(3):368-371.
- White SC, Pharoah MJ (2014). *Cone-Beam Computed Tomography. Oral Radiology: Principles and Interpretation.*] Elsevier Health Sciences; p. 185 ff.
- William., Kruskal., Wallis A (1952). Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association*; 47(260):558-621.
- [www.scanco.ch](http://www.scanco.ch) Scanco Medical  $\mu$ CT40 Datenblatt.
- Zimmermann. K-P, Gehrke1. P, Neugebauer. J (2014). Experimentelle Untersuchung zum Einfluss materialabhängiger Bildartefakte auf die Beurteilbarkeit im digitalen Volumentomogramm. *Deutscher Ärzte-Verlag*; 10.3238/ZZI.2014.0038-0052:38–52.

## 9 Danksagung

Frau Prof. Dr. Andrea Wichelhaus gilt mein besonderer Dank für die Überlassung des Dissertationsthemas und die konstruktiven Gespräche und Ratschläge.

Herrn Prof. Dr. Karl-Heinz Kunzelmann möchte ich meine tiefste Dankbarkeit für die Hilfe bei den  $\mu$ CT-Aufnahmen und die Bereitstellung der ImageJ-Plugins ausdrücken.

Mein weiterer Dank gilt Herrn Dr. Sebastian Stapfner und Herrn Dr. Uwe Baumert für die kompetente Betreuung und die Hilfe bzgl. der statistischen Auswertung.

Im Besonderen möchte ich meine tiefste Dankbarkeit meiner Schwester ausdrücken, die mir stets mit kritischen Ratschlägen in puncto Grammatik zur Seite stand.

Für die tolle Mithilfe möchte ich mich bei meinen Kollegen André Hausner und Marie Groh bedanken.

Abschließend bedanke ich mich bei meinen lieben Eltern, die mir meinen bisherigen Werdegang ermöglichten und mich immer unterstützen.

# Eidesstattliche Versicherung

Berkel, Maximilian Florentin

Name, Vorname

Ich erkläre hiermit an Eides statt,

dass ich die vorliegende Dissertation mit dem Thema

Validierung der Messgenauigkeit von Aufnahmen des Parodontalspaltes an digitalen  
Volumentomografien

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und  
alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als  
solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle  
einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in  
ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades  
eingereicht wurde.

Datteln, 27.06.2018

Ort, Datum

Maximilian Florentin Berkel

Unterschrift Doktorandin/Doktorand