

Aus der Poliklinik für Zahnerhaltung und Parodontologie
Klinikum der Ludwig-Maximilians-Universität München

Direktor: Prof. Dr. Reinhard Hickel



***Die Erkennung von Zahnpathologien auf apikalen
Röntgenaufnahmen - Ergebnisse einer Reliabilitätsstudie***

Dissertation

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

vorgelegt von

Theresa Meusburger

aus

Vorarlberg/Österreich

Jahr

2023

Mit Genehmigung der Medizinischen Fakultät der
Ludwig-Maximilians-Universität zu München

Erster Gutachter: Prof. Dr. Jan Kühnisch
Zweiter Gutachter: Prof. Dr. Roswitha Heinrich - Weltzien
Dritter Gutachter: PD Dr. Dr. Wenko Smolka

Dekan: Prof. Dr. med. Thomas Gudermann

Tag der mündlichen Prüfung: 12. Dezember 2023

Affidavit



Eidesstattliche Versicherung

Meusburger, Theresa

Name, Vorname

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel:

„Die Erkennung von Zahnpathologien auf apikalen Röntgenaufnahmen - Ergebnisse einer Reliabilitätsstudie“

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.

München, 28.02.2024

Theresa Meusburger

Ort, Datum
Doktorand

Unterschrift Doktorandin bzw.

Inhaltsverzeichnis

Affidavit	3
Inhaltsverzeichnis	4
Abkürzungsverzeichnis	5
Publikationsliste	6
Ihr Beitrag zu den Veröffentlichungen	7
1. Einleitung	8
2. Zielsetzung	10
3. Projektrealisierung	11
4. Ergebnisse	13
5. Diskussion	15
6. Zusammenfassung	18
Abstract	19
Paper I	20
Paper II	21
Literaturverzeichnis	22
Danksagung	27

Abkürzungsverzeichnis

aOR	adjusted (angepasste) Odds Ratio
PBL	Periodontal bone loss
LMU	Ludwig-Maximilians-Universität München
OR	Odds Ratio
CK	Cohens Kappa
95 %-CI	95%-Konfidenzintervall
InterUR	Inter-Untersucher-Reliabilität
IntraUR	Intra-Untersucher-Reliabilität
OPG	Orthopantomogramm

Publikationsliste

Meusburger, T., Wülk, A., Kessler, A., Heck, K., Hickel, R., Dujic, H., Kühnisch, J. (2023). The Detection of Dental Pathologies on Periapical Radiographs-Results from a Reliability Study. Journal of Clinical Medicine. 12(6), 2224. <https://doi.org/10.3390/jcm12062224>

Wülk, A., Meusburger, T., Dujic, H., Hickel, R., Kessler, A., Heck, K., Kühnisch, J. (2023). The Dentist's Ability to Detect Different Restorative Procedures on Periapical Radiographs-Results from a Reliability Study. International Journal of Environmental Research and Public Health. 20(3), 2619. <https://doi.org/10.3390/ijerph20032619>

Ihr Beitrag zu den Veröffentlichungen

Anlässlich der Promotion waren folgende Abschnitte erforderlich, die in der Studiengruppe gemäß der nachstehenden Tabelle bearbeitet wurden.

Publikation: “The dentist’s reliability to detect dental pathologies on periapical radiographs”

	Theresa Meusburger	Prof. Dr. Kühnisch	Co- Autoren
Projektidee	25%	50%	25%
Literaturrecherche	100%	-	-
Datenmanagement	80%	-	20%
Statistische Analyse	90%	-	10%
Training und Kalibrierung der Arbeitsgruppe	25%	50%	25%
Tabellen und Abbildungen	100%	-	-
Manuskript-Management	80%	10%	10%
Bildselektion	100%	-	-

Publikation: “The dentist’s ability to detect different restorative procedures on periapical radiographs - Results from a reliability study”

	Annika Wülk	Prof. Dr. Kühnisch	Co- Autoren
Projektidee	25%	50%	25%
Literaturrecherche	100%	-	-
Datenmanagement	80%	-	20%
Statistische Analyse	90%	-	10%
Training und Kalibrierung der Arbeitsgruppe	25%	50%	25%
Tabellen und Abbildungen	100%	-	-
Manuskript-Management	80%	10%	10%
Bildselektion	100%	-	-

1. Einleitung

Karies und parodontale Erkrankungen gehören zu den weltweit am weitesten verbreiteten dentalen Erkrankungen (GBD et al., 2020; Kassebaum et al., 2017; Foros et al., 2021; Pihlstrom et al., 2005). Entsprechend ist es von großer Bedeutung die klinische Untersuchung durch zuverlässige röntgenologische Diagnostik indikationsgerecht zu optimieren. Alle diagnostischen Verfahren müssen sich dabei durch eine hohe Präzision und gute Befundreliabilität auszeichnen, was Gegenstand zahlreicher Studien in der Vergangenheit war (Kühnisch et al., 2016; Heinrich-Weltzien et al., 2005; Litzemberger et al., 2021). Dies erfordert eine strukturierte Erfassung relevanter pathologischer Befunde, welche im klinischen Alltag mehr oder weniger gut gelingt. Die digitale Auswertung radiologischer Befunde mit Hilfe von künstlicher Intelligenz (KI) verspricht hierbei Vorteile, da Bildinformationen automatisiert und in kurzer Zeit ausgewertet werden können (Hwang et al., 2019; Ossowska et al., 2022; Tian et al., 2021). Neue Entwicklungen verbessern insbesondere die (zahn)ärztlichen Diagnosemöglichkeiten (Alghazzawi et al., 2016; Kühnisch et al., 2022; Schönewolf et al., 2022; Thurzo et al., 2022). Um das maschinelle Lernen voranzutreiben, benötigt es umfangreiches Daten- bzw. Bildmaterial, in dem wichtige und gewünschte pathologische Befunde von trainierten Untersuchern bzw. Annotatoren ausgewertet wurden (Alves et al., 2021; Vela et al., 2022).

Das vorliegende Publikationsprojekt untersuchte die Reliabilität von zahnärztlichen Befundungen in den Parametern Karies, apikale Parodontitiden, parodontaler Knochenabbau und Paro-Endo-Läsionen. Damit die Erfahrung und individuelle Beurteilung der jeweiligen Untersucher die diagnostische Entscheidung nicht beeinflussten, wurden diese zuvor geschult, um eine Konsistenz aller diagnostischen Entscheidungen sicherzustellen. Alle Ergebnisse wurden aus der vorangegangenen Trainings- und Kalibrierungsphase zusammengefasst und als InterUR und IntraUR dargestellt.

Die Literatur zeigte zahlreiche Studien, die sich mit der Karieserkennung auf Bissflügelaufnahmen beschäftigt haben (Janjic et al., 2021; Kapor et al., 2021; Litzemberger et al., 2018; Newman et al., 2009; Senneby et al., 2016; Blumer et al., 2021).

Zu den apikalen Parodontitiden gab es allerdings eine deutlich geringere Anzahl an Veröffentlichungen, die Vergleiche ermöglichen, obwohl apikale Röntgenaufnahmen als bevorzugte Methode zur Erkennung von apikaler Parodontitis dienen. Eine Studie von Patel et al. (2000) beispielsweise, zielte darauf ab, apikale Parodontitiden unter bestimmten Bedingungen auf apikalen Zahnfilmen zu erkennen und dokumentierte eine gute InterUR.

Die vorhandene Literatur zu parodontalem Knochenabbau und Paro-Endo-Läsionen war begrenzt mit nur wenigen Studien, welche die InterUR und IntraUR dediziert bewertet haben.

2. Zielsetzung

Die vorliegende Dissertation hatte zum Ziel, die InterUR und IntraUR für wichtige röntgenologische Befunde auf apikalen Zahnfilmen zu untersuchen. Dabei sollten Kappa-Werte nach Cohen´s (CK) für Karies, apikale Parodontitis, parodontalen Knochenabbau und Paro-Endo-Läsionen ermittelt werden. Im Fokus standen folgende Fragestellungen:

1. Mit welcher Zuverlässigkeit gelingt die diagnostische Beurteilung von Pathologien auf apikalen Röntgenbildern, speziell die InterUR und IntraUR von Karies, apikale Parodontitiden, parodontaler Knochenabbau und Paro-Endo-Läsionen?
2. Wie gut bewerten erfahrene Zahnärzte im Vergleich zu Zahnärzten mit weniger als zwei Jahren Berufserfahrung die ausgewählten pathologischen Befunde?
3. Welchen signifikanten Einfluss hat die Variable Evaluierungsrunde auf die Ergebnisse?

3. Projektrealisierung

Diese Studie war Teil des Studienprojektes „Diagnostik mit künstlicher Intelligenz“ an der Poliklinik für Zahnerhaltung und Parodontologie der LMU München (Projektnummer 20-0798). Das Projekt verfolgte das Ziel, die Stärken und Schwächen in der röntgenologischen Diagnostik von verschiedenen Zahnärzten zu erkennen, um erforderliche Kalibrierungen zielorientiert umsetzen zu können.

Insgesamt waren 14 Zahnärzte an der Diagnostikstudie beteiligt, darunter elf mit weniger als zwei Jahren Berufserfahrung. Voraussetzung für die Projektteilnahme war es, dass alle Untersucher mit den festgelegten Diagnostikkriterien und den Röntgenbildern vertraut waren, damit jeder Teilnehmer die ausgewählten apikalen Zahnfilme anschließend befunden konnte. Alle gewählten diagnostischen Standards entsprachen den international anerkannten Empfehlungen. Wenn proximale oder okklusale Kariesläsionen jenseits der Schmelz-Dentin-Grenze oder sogar bis in den inneren Dentin- und Pulpabereich reichten, wurden sie als Dentinkaries registriert. Schmelzkaries wurde nicht berücksichtigt (Marthaler et al., 1966; Ekstrand et al., 1997; Pitts et al., 2009). Gleichzeitig wurde keine Unterscheidung zwischen verschiedenen Graden der Progression gemacht. Eine apikale Parodontitis wurde diagnostiziert, wenn die Radioluzenz im Bereich des Apex an mindestens einem Zahn mit der doppelten Breite des parodontalen Ligaments sichtbar war (Low et al., 2008; Bornstein et al., 2011). Der periapikale Index von Orstavik et al. (1986) klassifiziert apikale Parodontitiden in fünf verschiedene Läsionsstadien. Eine Differenzierung zwischen Läsionsstadien wurde in der vorliegenden Studie allerdings nicht vorgenommen. Zur Beurteilung des parodontalen Knochenabbaus wurde die Schmelz-Zement-Grenze, der Limbus alveolaris und der Apex herangezogen. Erreichte der röntgenologische Knochenabbau mindestens das zervikale Wurzel Drittel (15-33%), so wurde von einem parodontalen Knochenabbau ausgegangen. Hierbei wurde nicht zwischen koronalem, mittlerem oder apikalem Wurzel Drittel unterschieden (Papapanou et al., 2018; Shaker et al., 2021; Fiorellini et al., 2021). Paro-Endo-Läsionen wurden festgestellt, wenn die apikale Parodontitis an einem Zahn pulpace und parodontale Strukturen einschloss. Die Klassifizierung von Simon et al. (1972)

diente als Referenz für die Erfassung von Paro-Endo-Läsionen. Wesentliches Kennzeichen war das Vorhandensein einer Radioluzenz, die das parodontale Ligamentum von der apikalen Region bis zum Sulcus gingivae umfasste.

Um sicherzustellen, dass von Beginn an kalibrierte Rahmenbedingungen vorlagen, besuchten alle teilnehmenden Zahnärzte zuvor einen zweitägigen Workshop. Während des Workshops wurden die genannten Diagnostikkategorien vorgestellt und besprochen. 150 apikalen Zahnfilme wurden zur Beurteilung aus einem erweiterten Bildersatz (N >500) ausgewählt. Von Bedeutung war der Einschluss korrekt belichteter und angulierter apikaler Zahnfilme.

Um die Erinnerung der einzelnen Zahnärzte möglichst umfassend zu reduzieren, wurden die ausgewählten Röntgenbilder nach vier Wochen erneut diagnostisch bewertet. Die Beurteilung erfolgte mit einer dichotomen Entscheidung (ja/nein). Nach Abschluss der beiden Auswertungsrunden wurden die Bilder ein weiteres Mal gemeinsam in der Gruppe analysiert und ein Konsensbefund als Referenzstandard festgelegt. Die statistische Auswertung wurde mittels Excel und SPSS durchgeführt (SPSS Statistics 27, 2020, IBM Corporation, Armonk, NY, USA). Zur Veranschaulichung der Untersucherübereinstimmung wurde der Kappa-Werte nach Cohen's herangezogen. Die Kalkulation erfolgte zwischen den Untersuchern sowie in Relation zum Referenzstandard (Cohen et al., 1960; Fleiss et al., 1973; Landis et al., 1977). In der vorliegenden Publikation bewegten sich die Reliabilitätswerte von moderater Übereinstimmung der CK-Werte (0,41 – 0,60), über gute Übereinstimmung der CK-Werte (0,61 – 0,80), bis zu perfekter Übereinstimmung der CK-Werte (0,81 – 1,00).

4. Ergebnisse

Die Ergebnisse zeigten insgesamt ein Spektrum zwischen moderater (CK 0,435) und perfekter Übereinstimmung (CK 0,840) für die gewählten diagnostischen Parameter (Tabelle 1). Vergleicht man die Intra-Untersucher-Reliabilitäts-Werte miteinander, erreichten die Paro-Endo-Läsionen eine perfekte Übereinstimmung von 0,840. Insgesamt schnitten die Kategorien apikale Parodontitis und parodontaler Knochenabbau am ungünstigsten ab – bei der apikalen Parodontitis mit einer InterUR von nur 0,435 (Runde 1 im Vergleich zum Referenzstandard) und der parodontale Knochenabbau erzielte eine ähnliche InterUR von 0,454 (Runde 1). Bei der IntraUR lagen die Werte hingegen zwischen einer guten und perfekten Übereinstimmung (Tabelle 1).

Tabelle 1. Kappa Werte nach Cohen´s (CK) steht für die InterUR und IntraUR der vier zahnärztlichen Pathologien

Variablen	InterUR (Runde 1)	InterUR (Runde 2)	IntraUR	InterUR (Runde 1/Referenzstan dart)	InterUR (Runde 2/Referenzstan dart)
Karies	0,704 (0,602-0,740)	0,659 (0,592-0,713)	0,778 (0,565-0,960)	0,748 (0,587-0,865)	0,724 (0,563-0,879)
Apikale Parodontitis	0,643 (0,374-0,718)	0,611 (0,380-0,688)	0,768 (0,388-1,000)	0,435 (0,323-0,901)	0,442 (0,265-0,901)
Parodontale r Knochenab bau	0,454 (0,311-0,563)	0,482 (0,361-0,580)	0,739 (0,455-1,000)	0,575 (0,375-0,762)	0,598 (0,361-0,580)
Paro-Endo- Läsion	0,702 (0,309-0,781)	0,689 (0,319-0,760)	0,840 (0,627-0,930)	0,780 (0,397-0,968)	0,755 (0,319-0,767)

Zur weiteren Untersuchung des Datensatzes wurde ein binäres logistisches Regressionsmodell angewandt. Tabelle 2 zeigt, dass die Evaluierungsrunden keinen signifikanten Einfluss auf die Zuverlässigkeit hatten (aOR 0,99, p-Wert < 0,711). Außerdem bewerteten Zahnärzte mit längerer klinischer Erfahrung (>2 Jahre) die periapikalen Röntgenbilder präziser (aOR 1,98, p-Wert < 0,001). Bei der Nutzung von Karies als Referenzwert, erzielten der parodontale Knochenabbau (aOR 0,57, p-Wert < 0,001) und die apikale Parodontitis (aOR

0,34, p-Wert < 0,001), auch im statistischen Modell, die geringeren Zuverlässigkeitswerte. Paro-Endo-Läsionen zeigten hingegen auch hier die besten Werte auf (aOR 1,54, p-Wert < 0,001).

Tabelle 2. Berechnung der Odds Ratios (OR) mit entsprechenden 95%-Konfidenzintervallen (CI) und p-Werten auf Grundlage des logistischen Regressionsmodells

Variablen	Gruppen	Odds Ratio	95% CI	p-Werte
Diagnostikkate gorien	Karies	1	-	-
	Apikale Parodontitis	0,34	0,30-0,37	<0,001
	Parodontaler Knochenabbau	0,57	0,51-0,64	<0,001
	Paro-Endo- Läsionen	1,54	1,34-1,77	<0,001
Evaluationsrun de	Runde 1	1	-	-
	Runde 2	0,99	0,91-1,07	0,711
Berufserfahrun g	< 2 Jahre	1	-	-
	> 2 Jahre	1,98	1,77-2,22	<0,001

5. Diskussion

Die in Tabelle 2 anhand des binären logistischen Regressionsmodells gerechneten Werte zeigten signifikante Abweichungen bei der Detektion der gewählten zahnärztlichen Pathologien. Die Reliabilitätswerte belegten einen klaren Unterschied zwischen den verschiedenen Kategorien und den Teilnehmern mit Kappa-Werten nach Cohen's im moderaten bis perfekten Bereich.

Des Weiteren zeigte die Datenauswertung wesentliche Unterschiede zwischen erfahrenen und weniger erfahrenen Untersuchern bei der Beurteilung pathologischer Befunde auf apikalen Zahnfilmen. Höhere Reliabilitätswerte wurden bei Untersuchern mit mehr praktischer Erfahrungen festgestellt, was sich in Übereinstimmung zu weiteren Veröffentlichungen fand (Diniz et al., 2010; Hellén-Halme et al., 2010; McCaul et al., 2001). Im Gegensatz zu den untersuchten Pathologien der vorliegenden Studie wurden für restaurative Befunde, direkte und indirekte Restaurationen, wurzelkanalbehandelte Zähnen und Implantaten, günstigere Ergebnisse ermittelt (Wülk et al. 2023). In diesem vergleichbaren Studienprojekt gab es keine Unterschiede zwischen den Untersuchern, deren Berufserfahrung und der Erst- bzw. Zweitbefundung. Grund hierfür könnte sein, dass Implantate (Kuusisto et al., 2015), direkte- und indirekte Restaurationen (Salzedas et al., 2006) und wurzelkanalbehandelte Zähne (Carvalho-Junior et al., 2007) eine gute Röntgensichtbarkeit aufwiesen.

Eine weitere wesentliche Erkenntnis war, dass in der vorliegenden Studie und in der Studie von Wülk et al. (2023) im Vergleich zur Veröffentlichung von Sebring et al. (2021) keine signifikante Verbesserung von der ersten zur zweiten Evaluierungsrunde festgestellt wurde. Dies lässt sich möglicherweise darauf zurückführen, dass alle Untersucher (einschließlich solcher mit weniger als zwei Jahre Berufserfahrung) bereits vor Beginn der Studie einen zweiwöchigen Workshop als Vorbereitung besuchten.

Bei der detaillierten Gegenüberstellung der Kariesdaten mit anderen Studien derselben Methodik, wie Freire et al. (2022), ergaben sich ähnliche Ergebnisse für apikale Röntgenbilder. Da aus diesem Artikel jedoch keine InterUR und IntraUR entnommen werden konnte, sollten zusätzliche Reliabilitätsstudien zur Karieserkennung auf Bissflügelaufnahmen berücksichtigt werden, wie zum

Beispiel die Veröffentlichung von Litzemberger et al. (2018), die eine IntraUR von 0,77 erzielte. Die Ergebnisse wurden auch von den kürzlich veröffentlichten Daten aus zwei systematischen Übersichten und Metaanalysen zur Okklusal- und Approximalkarieserkennung unterstützt (Kapor et al., 2021; Janjic et al., 2021).

Obwohl der apikale Zahnfilm, neben dem Orthopantomogramm (OPG), die bevorzugte Projektionstechnik zur Erkennung apikaler Parodontitis ist, standen für die betrachtete Fragestellung nur wenige vergleichbare Veröffentlichungen zur Verfügung. Die Studie von Patel et al. (2000) beurteilte apikale Röntgenbilder mit Verdeckung der Zahnkronen, um mögliche diagnostische Verzerrungen auszuschließen, was in unserer Untersuchung nicht berücksichtigt wurde. Dabei dokumentierte die Autorengruppe eine moderate InterUR (CK 0,26) unter den Teilnehmern. Ähnliche Beobachtungen stellten Sebring et al. (2021) und Saunders et al. (2000) fest, die beide ebenso wie in der vorliegenden Studie moderate Übereinstimmungen erzielten. Im Gegensatz dazu erbrachten Untersucher in der Veröffentlichung von Patel et al. (2012) nahezu perfekte Übereinstimmungen bei der Erkennung von apikalen Parodontitiden (CK 0,878), wobei erwähnt werden sollte, dass die Untersucher alle geschulte Endodontologen waren, die nur 30 apikale Röntgenbilder bewertet haben.

Für parodontalen Knochenabbau und Paro-Endo-Läsionen wurden keine vergleichbaren Studien für die zuverlässige Erkennung von Pathologien auf apikalen Zahnfilmen in den einschlägigen Literaturdatenbanken gefunden. Allerdings lagen wissenschaftliche Daten zur Beurteilung des parodontalen Knochenabbau in Verbindung mit OPGs vor, die den generalisierten Knochenabbau übersichtlich darstellten (Persson et al., 2003; Semenov et al., 2011; Molander et al., 1991).

Auch die vorliegende Studie wies Stärken und Schwächen auf. Das umfangreiche Bilderspektrum mit 150 periapikalen Röntgenbildern und die vier diagnostischen Kriterien konnten als Stärke festgehalten werden und sind relevant in der täglichen Routinearbeit des Zahnarztes. Zudem gehört das gleichzeitige Screening und die Erkennung verschiedener Pathologien auf Röntgenbildern in der zahnärztlichen Praxis zur alltäglichen Realität.

Als Limitierung gilt, dass das Studienmaterial aus einem Bildarchiv stammte, in dem klinische Informationen der Patienten fehlten, was die diagnostische Interpretation der Zahnfilme negativ beeinflusst haben könnte. Zudem war die Anzahl der Untersucher begrenzt. Eine größere Anzahl an Untersuchern mit verschiedenen Graden an Berufserfahrung wäre vorteilhaft gewesen, um einen relevanten Vergleich aufzustellen. Die berufliche Erfahrung erwies sich als signifikanter Einflussfaktor und könnte in einer größeren Teilnehmergruppe zu potenziellen Abweichungen führen. Es sollte auch berücksichtigt werden, dass die radiologischen Auswertungen unter verschiedenen Bedingungen durchgeführt wurden, was die Ergebnisse beeinflussen könnte. In anderen Publikationen wurde ebenfalls belegt, dass unterschiedliche äußere Umstände zu Diskrepanzen in der Bewertung führen könnten (Lima et al., 2020; Patel et al., 2000).

6. Zusammenfassung

Die vorliegende Studie dokumentierte, dass es Diskrepanzen in der Befundung von apikalen Röntgenbildern zwischen erfahrenen und unerfahrenen Zahnärzten gab. Ebenso zeigten die Ergebnisse Unterschiede zwischen den diagnostischen Kategorien. Insbesondere für die Untersuchung von Röntgenbildern auf apikalen Parodontitiden und parodontalen Knochenabbau, bedarf es an verbesserten Schulungen, um die Reliabilität der Befundungen von Pathologien auf ein einheitlich hohes Niveau zu fördern. Demgegenüber stehen die Ergebnisse bei der Befundung von restaurativen Behandlungsmaßnahmen aus der Studie von Wülk et al. (2023). Diese zeigten, dass die Erkennung der untersuchten restaurativen Behandlungsmaßnahmen für erfahrene und unerfahrene Zahnärzte als gleich souverän und zuverlässig eingestuft werden kann. In beiden Studien wurden keine Unterschiede zwischen den Evaluierungsrunden festgestellt, was die Replizierbarkeit der Arbeit verstärkt. Die Annahme, dass es keine Unterschiede in der Reliabilität der Befundung von Pathologien zwischen den Kategorien oder den Erfahrungsgraden der Untersucher gab, wurde hiermit entsprechend widerlegt.

Ein Ansatz für weitere Forschungsprojekte ist die zusätzliche Bewertung von verschiedenen Kategorisierungen, wie sie bei apikalen Parodontitiden unterteilt werden (Orstafik et al. (1986)). Darüber hinaus könnte der röntgenologische Vergleich von Zahnfilmen mit OPG's und Bissflügeln helfen, potenzielle Abweichungen oder Unstimmigkeiten in der Diagnose zu identifizieren und eine höhere diagnostische Genauigkeit zu gewährleisten.

Abstract

The goal of the presented study was to illustrate the inter- and intra-examiner reliability of pathologies on apical radiographs and compare them with similar studies. Specifically, the differences between diagnostic categories, evaluation rounds and the professional experience of the examiner were studied to check for statistical relevance. For the detection of pathologies, varying Cohens Kappa (CK) values were recorded in the moderate to perfect range, indicating a significant difference in the assessment of the diagnostic categories and the experience of the examiner. In contrast, CK values were consistently in the perfect range regardless of the professional experience of the examiner and categories for the detection of restorative measures on periapical radiographs, as evidenced in the study by Wülk et al. (2023). Endo-perio-lesions had the best CK values of 0,840 CK and were thus rated most reliable, whereas periapical lesions and periodontal bone loss scored the least with CK values of 0,435 and 0,454 respectively. Additionally, the odds ratios (OR) indicate a large disparity between the experienced and less experienced examiners, while showing almost identical values across the evaluation rounds. The hypothesis that there is no difference in the reliability of assessment between diagnostic categories and the experience of examiners was therefore rejected in this study, while it was confirmed in the research by Wülk et al. (2023). A main driver for this discrepancy could be that pathologies are generally more difficult to assess in radiographs requiring a higher level of experience in the examiner than restorative measures. Differences between evaluation rounds were not detected in both studies, ensuring the replicability of the studies. The results suggest that enhanced university teaching is recommended to improve education and training programs for the detection of pathologies on apical radiographs, especially for periapical lesions and periodontal bone loss. Additional research on the quality of detection for additional diagnostic categories on apical radiographs, OPGs and bitewings is recommended to aid the identification and elimination of further uncertainties. The additional research should also address all potential weaknesses in radiological diagnostics to support automated and standardized results. With the rapid advancements in artificial intelligence (AI) use-cases in dentistry, the generation of high-quality data for AI models must be ensured.

Paper I

Meusburger, T., Wülk, A., Kessler, A., Heck, K., Hickel, R., Dujic, H., Kühnisch, J. (2023). The Detection of Dental Pathologies on Periapical Radiographs - Results from a Reliability Study. *Journal of Clinical Medicine*, 12(6), 2224. <https://doi.org/10.3390/jcm12062224>

Paper II

Wülk, A., Meusburger, T., Dujic, H., Hickel, R., Kessler, A., Heck, K., Kühnisch, J. (2023). The Dentist's Ability to Detect Different Restorative Procedures on Periapical Radiographs-Results from a Reliability Study. *International Journal of Environmental Research and Public Health*, 20(3), 2619. <https://doi.org/10.3390/ijerph20032619>

Literaturverzeichnis

Alghazzawi, T. F. (2016). Advancements in CAD/CAM technology. Options for practical implementation *Journal of prosthodontic research*. 60(2), 72–84.

Alves, V. M., Auerbach, S. S., Kleinstreuer, N., Rooney, J. P., Muratov, E. N., Rusyn, I., Tropsha, A., & Schmitt, C. (2021). Curated Data In - Trustworthy In Silico Models Out: The Impact of Data Quality on the Reliability of Artificial Intelligence Models as Alternatives to Animal Testing. *Alternatives to laboratory animals: ATLA*. 49(3), 73–82.

Blumer, S., Kharouba, J., Kats, L., Schachter, D., & Azem, H. (2021). Visual Examination, Fluorescence-Aided Caries Excavation (FACE) Technology, Bitewing X-Ray Radiography in the Detection of Occlusal Caries in First Permanent Molars in Children. *The Journal of clinical pediatric dentistry*. 45(3), 152–157.

Bornstein, M. M., Lauber, R., Sendi, P., & von Arx, T. (2011). Comparison of periapical radiography and limited cone-beam computed tomography in mandibular molars for analysis of anatomical landmarks before apical surgery. *Journal of endodontics*. 37(2), 151–157.

Carvalho-Junior, J. R., Correr-Sobrinho, L., Correr, A. B., Sinhoreti, M. A., Consani, S., & Sousa-Neto, M. D. (2007). Radiopacity of root filling materials using digital radiography. *International endodontic journal*. 40(7), 514–520.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1), 37–46. Diniz, M. B., Rodrigues, J. A., Neuhaus, K. W., Cordeiro, R. C., & Lussi, A. (2010). Influence of examiner's clinical experience on the reproducibility and accuracy of radiographic examination in detecting occlusal caries. *Clinical oral investigations*. 14(5), 515–523.

Ekstrand, K. R., Ricketts, D. N., & Kidd, E. A. (1997). Reproducibility and accuracy of three methods for assessment of demineralization depth of the occlusal surface: an in vitro examination. *Caries research*. 31(3), 224–231.

Fiorellini, J. P., Sourvanos, D., Sarimento, H., Karimbux, N., & Luan, K. W. (2021). Periodontal and Implant Radiology. *Dental clinics of North America*, 65(3). 447–473.

Fleiss, J. L., & Cohen, J. (1973). The Equivalence of Weighted Kappa and the Intraclass Correlation Coefficient as Measures of Reliability. *Educational and Psychological Measurement*, 33(3) 613–619.

Foros, P., Oikonomou, E., Koletsi, D., & Rahiotis, C. (2021). Detection Methods for Early Caries Diagnosis: A Systematic Review and Meta-Analysis. *Caries research*. 55(4), 247–259.

Freire, R. T., Prata-Júnior, A. R., Pinho, J. N. A., & Takeshita, W. M. (2022). Diagnostic Accuracy of Caries and Periapical Lesions on a Monitor with and without DICOM-GSDF Calibration Under Different Ambient Light Conditions. *Journal of digital imaging*. 35(3), 654–659.

GBD 2017 Oral Disorders Collaborators, Bernabe, E., Marcenes, W., Hernandez, C. R., Bailey, J., Abreu, L. G., Alipour, V., Amini, S., Arabloo, J., Arefi, Z., Arora, A., Ayanore, M. A., Bärnighausen, T. W., Bijani, A., Cho, D. Y., Chu, D. T., Crowe, C. S., Demoz, G. T., Demsie, D. G., Dibaji Forooshani, Z. S., ... Kassebaum, N. J. (2020). Global, Regional, and National Levels and Trends in Burden of Oral Conditions from 1990 to 2017: A Systematic Analysis for the Global Burden of Disease 2017 Study. *Journal of dental research*. 99(4), 362–373.

Heinrich-Weltzien, R., Kühnisch, J., Iffland, S., Tranaeus, S., Angmar-Månsson, B., & Stösser, L. (2005). Detection of initial caries lesions on smooth surfaces by quantitative light-induced fluorescence and visual examination: an in vivo comparison. *European journal of oral sciences*. 113(6), 494–498.

Hellén-Halme, K., & Petersson, G. H. (2010). Influence of education level and experience on detection of approximal caries in digital dental radiographs. An in vitro study. *Swedish dental journal*. 34(2), 63–69.

Hwang, J. J., Jung, Y. H., Cho, B. H., & Heo, M. S. (2019). An overview of deep learning in the field of dentistry. *Imaging science in dentistry*. 49(1), 1–7.

Janjic Rankovic, M., Kapor, S., Khazaei, Y., Crispin, A., Schüller, I., Krause, F., Ekstrand, K., Michou, S., Eggmann, F., Lussi, A., Huysmans, M. C., Neuhaus, K., Kühnisch, J. (2021). Systematic review and meta-analysis of diagnostic studies of proximal surface caries. *Clinical oral investigations*. 25(11), 6069–6079.

Kapor, S., Rankovic, M. J., Khazaei, Y., Crispin, A., Schüller, I., Krause, F., Lussi, A., Neuhaus, K., Eggmann, F., Michou, S., Ekstrand, K., Huysmans, M. C., Kühnisch, J. (2021). Systematic review and meta-analysis of diagnostic methods for occlusal surface caries. *Clinical oral investigations*. 25(8), 4801–4815.

Kassebaum, N. J., Smith, A. G. C., Bernabé, E., Fleming, T. D., Reynolds, A. E., Vos, T., Murray, C. J. L., Marcenes, W., & GBD 2015 Oral Health Collaborators (2017). Global, Regional, and National Prevalence, Incidence, and Disability-Adjusted Life Years for Oral Conditions for 195 Countries, 1990-2015: A Systematic Analysis for the Global Burden of Diseases, Injuries, and Risk Factors. *Journal of dental research*. 96(4), 380–387.

Kuusisto, N., Vallittu, P. K., Lassila, L. V., & Huuonen, S. (2015). Evaluation of intensity of artefacts in CBCT by radio-opacity of composite simulation models of implants in vitro. *Dento maxillo facial radiology*. 44(2), 20140157.

Kühnisch, J., Meyer, O., Hesenius, M., Hickel, R., & Gruhn, V. (2022). Caries Detection on Intraoral Images Using Artificial Intelligence. *Journal of dental research*. 101(2), 158–165.

Kühnisch, J., Söchtig, F., Pitchika, V., Laubender, R., Neuhaus, K. W., Lussi, A., & Hickel, R. (2016). In vivo validation of near-infrared light transillumination for interproximal dentin caries detection. *Clinical oral investigations*. 20(4), 821–829.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*. 33(1), 159–174

Lima, C. A. S., Nascimento, E. H. L., Gaêta-Araujo, H., Oliveira-Santos, C., Freitas, D. Q., Haiter-Neto, F., & Oliveira, M. L. (2020). Is the digital radiographic detection of approximal caries lesions influenced by viewing conditions?. *Oral surgery, oral medicine, oral pathology and oral radiology*. 129(2), 165–170.

Litzenburger, F., Heck, K., Pitchika, V., Neuhaus, K. W., Jost, F. N., Hickel, R., Jablonski-Momeni, A., Welk, A., Lederer, A., & Kühnisch, J. (2018). Inter- and intraexaminer reliability of bitewing radiography and near-infrared light transillumination for proximal caries detection and assessment. *Dento maxillo facial radiology*. 47(3), 20170292.

Litzenburger, F., Schäfer, G., Hickel, R., Kühnisch, J., & Heck, K. (2021). Comparison of novel and established caries diagnostic methods: a clinical study on occlusal surfaces. *BMC oral health*. 21(1), 97.

Low, K. M., Dula, K., Bürgin, W., & von Arx, T. (2008). Comparison of periapical radiography and limited cone-beam tomography in posterior maxillary teeth referred for apical surgery. *Journal of endodontics*. 34(5), 557–562.

Marthaler T. M. (1966). A standardized system of recording dental conditions. *Helvetica odontologica acta*. 10(1), 1–18.

McCaul, L. K., McHugh, S., & Saunders, W. P. (2001). The influence of specialty training and experience on decision making in endodontic diagnosis and treatment planning. *International endodontic journal*. 34(8), 594–606.

Molander, B., Ahlqwist, M., Gröndahl, H. G., & Hollender, L. (1991). Agreement between panoramic and intra-oral radiography in the assessment of marginal bone height. *Dento maxillo facial radiology*. 20(3), 155–160.

Newman, B., Seow, W. K., Kazoullis, S., Ford, D., & Holcombe, T. (2009). Clinical detection of caries in the primary dentition with and without bitewing radiography. *Australian dental journal*. 54(1), 23–30.

Orstavik, D., Kerekes, K., & Eriksen, H. M. (1986). The periapical index: a scoring system for radiographic assessment of apical periodontitis. *Endodontics & dental traumatology*. 2(1), 20–34.

Ossowska, A., Kusiak, A., & Świetlik, D. (2022). Artificial Intelligence in Dentistry-Narrative Review. *International journal of environmental research and public health*. 19(6), 3449.

Papapanou, P. N., Sanz, M., Buduneli, N., Dietrich, T., Feres, M., Fine, D. H., Flemmig, T. F., Garcia, R., Giannobile, W. V., Graziani, F., Greenwell, H., Herrera, D., Kao, R. T., Kerschull, M., Kinane, D. F., Kirkwood, K. L., Kocher, T., Kornman, K. S., Kumar, P. S., Loos, B. G., ... Tonetti, M. S. (2018). Periodontitis: Consensus report of workgroup 2 of the 2017 World Workshop on the Classification of Periodontal and Peri-Implant Diseases and Conditions. *Journal of clinical periodontology*. 45 Suppl 20, S162–S170.

Patel, N., Rushton, V. E., Macfarlane, T. V., & Horner, K. (2000). The influence of viewing conditions on radiological diagnosis of periapical inflammation. *British dental journal*. 189(1), 40–42.

Patel, S., Wilson, R., Dawood, A., & Mannocci, F. (2012). The detection of periapical pathosis using periapical radiography and cone beam computed tomography - part 1: pre-operative status. *International endodontic journal*. 45(8), 702–710.

Patel, S., Wilson, R., Dawood, A., Foschi, F., & Mannocci, F. (2012). The detection of periapical pathosis using digital periapical radiography and cone beam computed tomography - part 2: a 1-year post-treatment follow-up. *International endodontic journal*. 45(8), 711–723.

Persson, R. E., Tzannetou, S., Feloutzis, A. G., Brägger, U., Persson, G. R., & Lang, N. P. (2003). Comparison between panoramic and intra-oral radiographs for the assessment of alveolar bone levels in a periodontal maintenance population. *Journal of clinical periodontology*. 30(9), 833–839.

Pihlstrom, B. L., Michalowicz, B. S., & Johnson, N. W. (2005). Periodontal diseases. *Lancet* (London, England). 366(9499), 1809–1820.

Pitts N. B. (2009). How the detection, assessment, diagnosis and monitoring of caries integrate with personalized caries management. *Monographs in oral science*. 21, 1–14.

Salzedas, L. M., Louzada, M. J., & de Oliveira Filho, A. B. (2006). Radiopacity of restorative materials using digital images. *Journal of applied oral science : revista FOB*. 14(2), 147–152.

Saunders, M. B., Gulabivala, K., Holt, R., & Kahan, R. S. (2000). Reliability of radiographic observations recorded on a proforma measured using inter- and intra-observer variation: a preliminary study. *International endodontic journal*. 33(3), 272–278.

Schönewolf, J., Meyer, O., Engels, P., Schlickerrieder, A., Hickel, R., Gruhn, V., Hesenius, M., & Kühnisch, J. (2022). Artificial intelligence-based diagnostics of molar-incisor-hypomineralization (MIH) on intraoral photographs. *Clinical oral investigations*. 26(9), 5923–5930.

Sebring, D., Kvist, T., Buhlin, K., Jonasson, P., EndoReCo, & Lund, H. (2021). Calibration improves observer reliability in detecting periapical pathology on panoramic radiographs. *Acta odontologica Scandinavica*. 79(7), 554–561.

Semenoff, L., Semenoff, T. A., Pedro, F. L., Volpato, E. R., Machado, M. A., Borges, A. H., & Semenoff-Segundo, A. (2011). Are panoramic radiographs reliable to diagnose mild alveolar bone resorption?. *ISRN dentistry*. 2011, 363578.

Senneby, A., Elfvin, M., Stebring-Franzon, C., & Rohlin, M. (2016). A novel classification system for assessment of approximal caries lesion progression in bitewing radiographs. *Dento maxillo facial radiology*. 45(5), 20160039.

Shaker, Z. M. H., Parsa, A., & Moharamzadeh, K. (2021). Development of a Radiographic Index for Periodontitis. *Dentistry journal*. 9(2), 19.

Simon, J. H., Glick, D. H., & Frank, A. L. (1972). The relationship of endodontic-periodontic lesions. *Journal of periodontology*. 43(4), 202–208.

Thurzo, A., Urbanová, W., Novák, B., Czako, L., Siebert, T., Stano, P., Mareková, S., Fountoulaki, G., Kosnáčová, H., & Varga, I. (2022). Where Is the Artificial Intelligence Applied in Dentistry? Systematic Review and Literature Analysis. *Healthcare (Basel, Switzerland)*. 10(7), 1269.

Tian, Y., Chen, C., Xu, X., Wang, J., Hou, X., Li, K., Lu, X., Shi, H., Lee, E. S., & Jiang, H. B. (2021). A Review of 3D Printing in Dentistry: Technologies, Affecting Factors, and Applications. *Scanning*. 2021, 9950131.

Vela, D., Sharp, A., Zhang, R., Nguyen, T., Hoang, A., & Pianykh, O. S. (2022). Temporal quality degradation in AI models. *Scientific reports*. 12(1), 11654.

Wülk, A., Meusburger, T., Dujic, H., Hickel, R., Kessler, A., Heck, K., & Kühnisch, J. (2023). The Dentist's Ability to Detect Different Restorative Procedures on Periapical Radiographs-Results from a Reliability Study. *International journal of environmental research and public health*. 20(3), 2619.

Danksagung

Es ist Zeit ein großes Dankeschön an folgende Menschen auszusprechen, ohne die das alles niemals möglich gewesen wäre.

An erster Stelle möchte ich Herrn Prof. Dr. Jan Kühnisch erwähnen, ein Doktorvater wie es im Bilderbuche steht. Mit seiner Geduld, seinem Wohlwollen, seiner sehr angenehmen Art und seinem immerzu offenen Ohr, hat er uns/mir die Unterstützung gegeben, um mit dieser Arbeit erfolgreich voranzukommen und das Ziel nie aus den Augen zu verlieren.

Ich bedanke mich auch herzlichst bei meinen Co-Autoren und Promotionskolleginnen, Annika Wülk und Helena Dujic für die ausgezeichnete Teamarbeit, den kompetenten Austausch und sehr kollegialen Umgang miteinander. Neben der gemeinsamen Zielverfolgung habe ich mit Ihnen Freunde fürs Leben gefunden.

Vielen Dank auch an alle anderen Doktoranten, Professoren und Oberärzte für die gemeinsame Zeit und deren Engagement das bestmögliche aus diesem gemeinsamen Projekt hervorzubringen.

Schließlich möchte ich ein außerordentliches Dankeschön an Mama und Däta aussprechen, ohne Euch wäre ich beruflich aber vor allem persönlich keineswegs dort, wo ich jetzt bin. Ein großer Dank gilt auch meinen geliebten Geschwistern Elisabeth und Lukas, sowie meiner besten Freundin Veronika und meinem Verlobten Tag. Sie haben mich bestärkt und immer wieder motiviert, mir stets nur das Beste gewünscht.

Danke...