UNIVERSITA' DEGLI STUDI DI SIENA
DIPARTIMENTO BIOTECNOLOGIE MEDICHE
/DEPARTMENT OF MEDICAL BIOTECHNOLOGY

DOCTORAL THESIS IN

MEDICAL BIOTECHNOLOGY

CYCLE XXXIII

COORDINATOR: PROF. LORENZO LEONCINI

TITLE OF THESIS:

INVESTIGATING CLINICALLY RELEVANT METHODS OF ASSESSING THE QUALITY OF THREE-DIMENSIONAL SURFACE SCAN DATA IN DENTISTRY

SCIENTIFIC-DISCIPLINARY SECTOR: MALATTIE ODONTOSTOMATOLOGICHE MED/28

PhD STUDENT:

CECILIE ANNETH OSNES

SUPERVISORS:

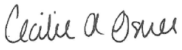
PROF. MARCO FERRARI

DR ANDREW KEELING

Academic Year in which the PhD was awarded:
2021

Declaration of own work

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to these works have been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.



X Signed 30/04/2021

Cecilie A Osnes
PhD student

Contributions

Chapter 2: Full arch precision of six intraoral scanners in vitro

C. Osnes: Design of the research project, data acquisition, data processing, analysis and interpretation, drafting and manuscript completion

J. Wu: Statistical analysis

P. Venezia: Data acquisition and manuscript completion

M. Ferrari: Design of the research project, data acquisition and manuscript completion

A. Keeling: Design of the research project, analysis and interpretation, manuscript completion

Chapter 3: Investigating three methods of assessing the clinically relevant trueness of two intraoral scanners

C. Osnes: Design of the research project, data acquisition, data processing, analysis and interpretation, drafting and manuscript completion

J. Caplan: Data acquisition and manuscript completion

M. Ferrari: Design of the research project and manuscript completion

A. Keeling: Design of the research project, analysis and interpretation, manuscript completion

Chapter 4: Sources of Error in Maximum Intercuspation from Complete Dentate Full Arch Intraoral Scans in vitro

C. Osnes: Design of the research project, data acquisition, data processing, analysis and interpretation, drafting and manuscript completion

J.Wu: Statistical analysis

T. Joda: Design of research project and manuscript completion

M. Ferrari: Design of the research project and manuscript completion

A. Keeling: Design of the research project, analysis and interpretation, and manuscript completion

Chapter 5: Optimal Use of Physical Centric Relation Records for Digital Workflows

C. Osnes: Design of the research project, data acquisition, data processing, analysis and interpretation, drafting and manuscript completion

J. Patel: Design of research project, data acquisition, and manuscript completion

A. Wagstaff: Data acquisition and manuscript completion

A. Keeling: Design of the research project, analysis and interpretation, and manuscript completion

All custom software credited to Leeds Digital Dentistry was written solely by Andrew Keeling, but collaboratively developed by CO and AK.

Abstract

Title: **Investigating clinically relevant methods of assessing the quality of three-dimensional surface scan data in dentistry**

Motivation:
Having an awareness of the quality of three-dimensional (3D) scan data produced by a dental scanner can affect the clinician's treatment plan, and potentially whether they choose to invest in a 3D scanner for clinical use. Assessing the quality of 3D scan data can therefore be of great value both to practitioners and scanner manufacturers.

Statement of problem:
Assessing the quality of 3D data is a challenge; assessing the quality of 3D data for clinical use is even more so. As a result, there is no standardised method of assessing, or reporting, the quality of 3D data within the field of dentistry. This research aimed to investigate methods of assessing the trueness and precision of scanners with clinical application in mind.

Method and summary of publications:
This research resulted in four publications. All data collection was undertaken *in vitro*. Digital processing, measurements and analyses were all performed digitally, in most cases using automated methods.

Chapter 2 compares full arch edentulous scans produced by six intraoral scanners [IOS], with focus on identifying full arch error which may not be identified if analysed using commonly used methods such as measuring the mean distance deviation between a scan-pair. The proposed method suggests reporting the unsigned distance of the median value of the upper 1% most deviating aspect of a repeated scan.

Chapter 3 investigates three methods of analysing scan data produced by two IOSs, applying the *upper-bound* method presented in the previous chapter in a more approachable manner, by reporting the percentage of a scan deviating beyond 0.1mm. This paper also presents the use of a *virtual key point* method, by which topologically similar key points can be robustly identified across differing meshes.

The virtual key point method is further investigated in Chapter 4, in which the accuracy of the virtual occlusion of an IOS is investigated. By using virtual key points, the proportion of error

produced by the arch scan and the proportion of error which is introduced during the occlusion stage are identified.

Lastly, in Chapter 5 the key point method is used to investigate the precision of physical interocclusal records and the difference in precision between virtual articulation of dental models using un-clamped, scanned bite records, and traditional, manual articulation done by an experienced technician. Having determined the accuracy with which an IOS can record occlusion in the previous chapter, we use the same scanner to digitally record the technician's manual articulations and, from there, identify the portion of error which was introduced by the manual articulation.

Results:

Chapter 2 finds that the *upper-bound* method may provide a clinically useful metric with which to gain an insight into the precision of full arch scans produced using different IOSs. The results indicate that three of the six scanners investigated would likely produce scans appropriate for clinical use where the full arch is required, while the latter three produced errors deviating beyond 0.3mm, hence proving to be unlikely to be appropriate for clinical use.

Chapter 3 concludes that the Primescan produces significantly truer scans than the Omnicam, regardless of the method used to analyse the scan data. Furthermore the 'standard' analysis method might incorrectly infer that Omnicam produced clinically acceptable full arch scans. The proposed novel methods, measuring the intermolar-width and proportion beyond 0.1mm, may give a clinically relevant insight into the quality of scan data. These novel analyses reveal clinically unacceptable limitations for Omnicam.

Our findings in Chapter 4 conclude that while the virtual bite records were relatively precise ((never deviating beyond 0.022mm) the error produced during the creation of the full arch scan negatively impacts the virtual occlusion.

Lastly, Chapter 5 indicates that the digital articulation method using un-clamped, scanned bite records is significantly more precise than the traditional articulation method when considering precision (or lack of dislocation) along the anteroposterior axis.

Discussion:

All papers, with the exception of Chapter 5, investigate aspects of IOSs ability to accurately record full arch scans. All investigations presented are *in vitro*; as a result one may assume the quality of *in vivo* scan data will be worse than the findings reported herein. Chapter 3, which investigates the quality of full arch data produced by the Omnicam and Primescan scanners,

highlights the importance of using the appropriate measurement method during investigation, as seen in Figure 3:2. Chapters 4 and 5 highlight the variable nature of digital methods, with the IOS virtual occlusion feature being negatively affected by error produced in the full arch scan, making the method less likely to be clinically reliable. Whereas the virtual method in Chapter 5 indicates that virtual articulation, using a high quality, clamp-less laboratory scanner is likely to reproduce articulation more precisely than traditional, physical methods.

Conclusion:

This thesis concludes that efforts to gain a clinically relevant insight into the quality of scan data are challenging. It finds finds that there is no *one-size-fits-all* when assessing 3D data in a clinically relevant manner but suggests some newer methods that go some way to addressing this.

Standard surface comparison methods, borrowed from Engineering and used extensively in dental research, almost invariably produce overly optimistic results. Given that dental audiences are generally less well versed in mathematical 3D analysis, there is a real risk that clinical applicability of some digital techniques may be advocated in error.

The findings also show that the quality of 3D data within digital dentistry varies widely. A paradigm shift from *digital dentistry* being considered as multiple methods all producing data and clinical work of similar quality, to digital dentistry being considered an umbrella term covering a spectrum of workflows, all of *highly varying* quality, is needed.

Acknowledgements

Thanks to Marco Ferrari for enabling me to do this research. Thanks, also, to all co-authors for your collaboration and endless patience all those times I caused yet another deadline to whoosh by.

Thank you, Andy, for guiding me home.

List of Contents

List of Figures

1. Introduction

1.1. Intention

Digital dental solutions have in recent years been adapted by clinicians and technicians across the globe. Intraoral (IOS) and indirect scanners, dental computer-aided design (CAD) packages, and millers and 3D printers (computer-aided manufacture, or, CAM) can all be used within the clinical workflow (Van Noort, 2012; Rekow, 2020; van der Zel, 2020).

The digital dental workflow always requires a data *acquisition* stage, where an impression is taken and this, or the consequent model, is 3D scanned. Alternatively, an intraoral scanner can be used to create a direct digital impression. (Or in some cases, such as for implantology, computed tomography scan data can also be used, but this is often in combination with higher resolution surface scan data). Once digitised, the scan data can be brought into a dental CAD package. Crowns, bridges, prosthetic implants, partial dentures, bite splints and retainers are just some of the devices currently commonly designed using CAD. The final design can then be manufactured through various methods, commonly milling for crowns and prosthetics, and 3D printing or vacuum forming for orthodontic appliances.

So, the first stage of any digital dental workflow is scanning the patient's dental arch, impression, or model. That the scan *data* is a true representation of the scan *object* is critical for the success of any clinical appliance, crown or prosthesis designed on this data. Most commercially available dental laboratory scanners cite a "certified" accuracy of <0.01mm or less (see Figure 1:1 Example of a scanner manufacturer specifying the accuracy of a dental laboratory scanner. The claimed accuracy of 0.004 mm might hold very little meaning to a clinician attempting to decide on the appropriate scanner for their use yet the literature commonly report scan errors ten times this distance (Keul and Güth, 2020, being but one example). This suggests that the methods and standards used by manufacturers to validate their devices do not align with the methods clinicians and researchers use to measure the data produced by the same device.

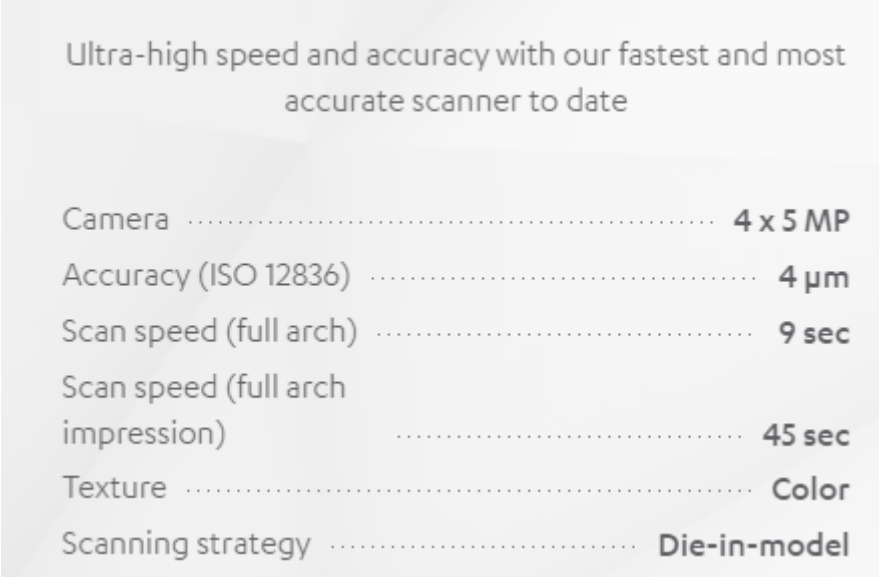


Figure 1:1 Example of a scanner manufacturer specifying the accuracy of a dental laboratory scanner. The claimed accuracy of 0.004 mm might hold very little meaning to a clinician attempting to decide on the appropriate scanner for their use, and rarely reflects the findings produced by academic studies investigating the accuracy of scanners.

Assessing and robustly measuring the trueness and precision[1] of the data produced by a 3D scanner in a clinically relevant way is an actively investigated, but unsolved, problem. The most common method for assessing scanner accuracy, by aligning multiple scans and measuring the average distance deviation between scan-pairs, carries numerous risks and pitfalls in that it relies on correct global alignment; a computational problem with no single correct solution if the two scans are even remotely different (which will always be the case with scan data) (Güth *et al.*, 2017; O'Toole *et al.*, 2019a).

Taking the problem of relying on global alignments (also known as *best fit* alignment) and mean distance measurements as a starting point, this thesis aims to investigate methods of robustly assessing the ability of dental scanners to correctly record scan data, particularly with clinical application in mind.

The dimensional accuracy of full arch scans (Ender and Mehl, 2011; Ender et al., 2019), the accuracy of virtual bite registration (DeLong et al., 2002; Gintaute et al., 2020) and the quality of a scanned margin of a tooth preparation (Keeling et al., 2017) have been identified as key issues which could benefit from further investigation. In particular, this thesis will focus on the issues of full arch scan accuracy and occlusal accuracy, since these are fundamental to the

[1] The combination of *trueness* and *precision* is from here-on referred to as *accuracy,* as according to BS ISO 5725-1:1994.

production of good quality digital study models, from which any future prosthodontic treatment can be planned and designed.

Gaining an awareness of the achievable quality of data within the clinical dental workflow would be of great advantage to clinicians when treating patients using digital solutions as well as to clinicians and dental technicians when investing in digital systems.

Being a laboratory-based, non-clinical research project, this work will focus on determining the quality of data which is *technically achievable*, and leave the task of determining the quality of data which is *clinically necessary* to future researchers.

1.2. An introduction to 3D scanning in dentistry

There are a handful of tangible differences between indirect scanners (often referred to as *laboratory* scanners) and intraoral scanners; solving the problem of identifying the location of 3D points in the real world is greatly complicated by not knowing the location of the sensor (normally a camera) which is collecting the information. As such, an intraoral scanner is faced with a greater computational challenge than a laboratory scanner, which can be calibrated to know the exact location of its sensors during manufacture. Additionally, intraoral scanners are used in the inherently challenging environment of the patient's mouth, with moisture, fogging and optically complex surfaces such as enamel further adding to the challenge of collecting valid scan data. As a result, laboratory scanners are generally assumed to produce more accurate and higher quality (in terms of point density) data than intraoral scanners.

All publications within this thesis (with some exception in the final publication) focus on intraoral scanners. It is worth noting that all the methods developed and explored herein could also be applied to data produced by indirect scanners, as many of the challenges in investigating scan data apply to all optical surface scanners, regardless of acquisition method.

1.2.1. Digital impressions

It is assumed that an intraoral scanner needs to produce digital impressions of similar quality to that produced by traditional impressions, or better. There is limited agreement as to how best to assess the quality of a digital impression. Further, there is limited data when it comes to determining the clinically required quality of an impression, be it digital or traditional, partly due to the lack of a standardised metric for assessing such qualities.

Investigations into the precision and trueness of IOSs have been undertaken regularly for the last decade. A 2011 study reported some IOSs to record full arch impressions more precisely than alginate impressions, but less precisely than silicone impressions (Ender and Mehl, 2011). A 2012 publication observed an increase in distortion across the arch, both in real world distance, and angulation distortion of the scan artefact, indicating that a single quadrant scan was likely to be more accurate than a full arch scan (van der Meer *et al.*, 2012).

Many studies have concluded with similar findings since van der Meer *et al.*'s publication: that short distance scans, such as those required for single preparations or partial prostheses, may be acceptable (Syrek et al., 2010), but that IOSs fail to produce clinically acceptable full arch scans (Ender et al., 2015; Chochlidakis et al., 2016; Tsirogiannis et al., 2016; Ender et al., 2016; Ahlholm et al., 2018). Guth *et al.* found in their 2016 study that intraoral scanning systems showed similar or higher accuracy than conventional polyether impression with subsequent indirect digitalisation (Güth *et al.*, 2016). Kuhr *et al.*, on the other hand, reported the very same year that conventional polyether impressions, analysed by pouring of the model and measuring these with a coordinate measuring machine, were still truer than digital impressions when performed *in vitro* (Kuhr *et al.*, 2016). Mangano *et al*. concluded, regardless of analysis method, that the 12 IOSs under investigation produced widely varying results when scanning implant impressions (Mangano et al., 2020).

It becomes apparent at this stage that the lack of agreement in methods used when investigating the accuracy of scanners obfuscates the field, and that measuring 3D data can be done in many ways, all producing different results and, therefore, a lack of clarity.

Further, IOSs have not yet truly penetrated the market, and the user-group may be self-selecting to enthusiasts willing to invest in IOS. This is likely to result in studies investigating IOSs being instigated by investigators who may be biased towards digital solutions. To illustrate the tunnel-vision of the general digital audience: Dentsply Sirona's flagship scanner at the time of writing, the Primescan, is heavily promoted based on the results presented in Ender *et al.*'s study from 2019; this study did in fact show that conventional silicone impressions are still truer and more precise than all of the scanners investigated in the *in vitro* experiment, echoing their findings eight years prior (Ender et al., 2019).[2]

[2] Further, with the study being *in vitro,* one may argue that the study favours IOSs by design: IOSs perform better *in vitro* than under the challenges posed *in vivo.* Silicone, on the other hand, is designed to work in a slightly moist environment, and thus might, perversely, perform worse *in vitro.* It might be

One fact which does become clear through much of the literature, and as specifically concluded by Mangano *et al.* in their 2020 publication, is that the quality of data produced by IOSs varies widely between devices. This is shown throughout the field and is reflected in this thesis.

While findings such as the conflicting views on the precision of conventional impressions in comparison to IOS could be the result of inconsistent methods of investigation and measurement, these divergences could also be the result of using different laboratory scanners when scanning the impression/model against which to compare the IOS scans. The Ender *et al. (2019)* study used a 5-axis laboratory scanner to scan impressions. In contrast, the study which concluded that IOSs produced impressions of a similar or higher quality than conventional methods relied on an older 3-axis scanner (Güth et al., 2016). It may be the case that the inability of a scanner to produce full-coverage of the scan object will negatively affect any study using surface deviation as a metric: dental models are inherently complex shapes containing undercuts and crevices[3]; poor scan coverage, as result of limited range of freedom by the scanner, may result in holes. As well as negatively affecting global alignment, holes or artificial "hole patching" performed by scanner software on an incompletely scanned object (unbeknownst to the user) would result in standard methods of measuring precision, such as mean deviation between repeated scans, reporting greatly varying values; and as such, a lack of precision. This could erroneously result in the conclusion that an indirectly digitised conventional impression is less precise than a virtual impression produced by an IOS, while the true cause of error is the inability of the laboratory scanner to correctly digitise the dental model.

1.2.2. Digital methods of occlusion and bite registration

Having produced a digital version of a patients' dental arches, by whichever method deemed appropriate by the clinician, a record of the patient's occlusion is usually required. Having an awareness of the patient's habitual intercuspation (occlusion) and dynamic paths of function (articulation) could potentially prevent future tooth fractures (Keeling, 2016). If part of the

argued that the study would likely have had greater promotional value had it been used by the poly-vinylsiloxane material manufacturer whose materials proved the most precise in Ender *et al.'s* study.

[3] Impressions are even more of a challenge to scan. The complex topology, and tendency of dentate impressions to have undercuts and narrow crevices, means moving a scanner head or scan object into a position where a direct scan view is possible is a challenge, if not impossible. The result is often either a scan with holes, or a scan with artificially repaired areas to close up any holes.

treatment is digital (such as CAD/CAM crowns or larger appliances such as digitally manufactured bite splints) the occlusion may need to be digitised. While dynamic articulation falls beyond the scope of this thesis, (and IOSs currently have only limited capabilities of recording this) static occlusal registration is offered by most IOS (and most indirect scanners) and could benefit from further investigation.

There are several methods which can be used to create a digital record, ranging from scanning analogue bite records, scanning articulated models in a laboratory scanner or using an IOS. The challenges that come to light when developing an experimental protocol with which to assess the quality of articulation, be it analogue or digital, are evident in the literature (Rhee *et al.*, 2015; Solaberrieta *et al.*, 2015; Solaberrieta, Garmendia, *et al.*, 2016; Alghazzawi, 2016). Evaluation of contact points is a common method, as seen in (Bohner et al., 2016), but quantitatively assessing the quality of contact points is a challenge, as this may fail to indicate perforation of teeth, if overclosure has occurred. Overclosure is, in most cases, an artifact of the nature of virtual data (which can intersect freely, unlike real world objects), and is unlikely to be encountered in analogue studies (Gintaute et al., 2020). Furthermore, contact point analysis requires some form of arbitrary parameterization. Is a contact point valid if there is a 0.001 mm separation between the teeth? Or 0.01 mm? Or 0.35 mm? The latter value has been used previously (DeLong et al., 2002) but would a clinician really accept articulating paper which is over one third of a millimeter thick? Typical clinical articulating paper is less than 0.1mm thick, and usually less than half of that value. Therefore, when assessing the body of literature relating to digital dental occlusion, one must be mindful that the measurement metrics are not ideal.[4]

[4] While DeLong settled for 0.35mm distance between cusp tips, Straga compared physical intersections (<0.000 mm) between opposing teeth in the digital arm of their study (Straga, 2009). Shimstocks and articulating foil were used as the control in this study. This is a problematic approach, as seen on page 41 in Straga's study: the digital version underestimates the number of contact points. This may be the result of the missing contact point being just less than the thickness of the shimstock/articulating foil away from the opposing tooth. This would result in the contact being identified in the conventional arm, but not in the digital arm which required physical intersection with the opposing tooth.

Alternative reasons for reduced contact points in the digital study include misalignment of the model during scan creation, misalignment of one or both models against the buccal scans, or a lack of resolution in the digital mesh produced by the scanner. A simplified, or low resolution mesh (in comparison to Type IV stone, for example) would likely negatively affect the trueness and precision of virtual articulation. Cusp tips, due to their relatively smooth surface (as opposed to margins of preparations), may suffer from aggressive simplification during scanning and mesh generation. This will be discussed in more detail, but in relation to virtual articulation, such simplification may result in digital cusp tips consisting of fewer points and larger face triangles — resulting in a faceted appearance. In

With these caveats in mind, digital recording of maximum intercuspation [ICP] using IOS has been reported to produce favourable results compared to analogue methods, with Jaschouz and Mehl reporting a reproducibility in ICP registration using an IOS of 0.042(±0.034) mm as opposed 0.135(±0.077) mm from an analogue method (Jaschouz and Mehl, 2014). This study investigated the precision of reproducibility in buccal scan alignment, as opposed to the alignment of a full arch scan using the buccal scans as reference. We show, in Chapter 4, that the results produced when introducing the full arch scan differ greatly from the optimal buccal bite alignments. The results presented by Jaschouz and Mehl must therefore be considered optimal results under ideal conditions, and clinical ICP registration using IOSs is likely to produce inferior results.

Assessing the ability with which a scanner is able to correctly record occlusion is a challenge, and is complicated, in IOSs, by any arch distortion produced during the original scan. Work investigating the challenge of identifying, and separating, arch scan error from bite scan error is presented in Chapter 4. This paper could be read as a continuation of the work presented in (Gintaute *et al.*, 2020).

1.2.3. Current methods for measuring scanner accuracy

Determining the accuracy of dental scanners is not a new topic, and although technologies are constantly being improved, the quality of dental scan data, and methods with which to investigate this, has been "under investigation" since 1987, if not longer (Rekow *et al.*, 1987).

In 2003, Luthard *et al.* claimed to present a novel method for investigating dimensional changes in dental materials, by using an optical scanner. Though presented as an investigation into material shrinkage, rather than the scanner, this study is an earlier example of the method used subsequently by many to measure the accuracy of 3D data. This method proceeds by performing best-fit registration on multiple data-sets before measuring the mean distance between scan pairs (by searching for the closest point of the opposing mesh from each vertex of the target mesh). While there was little noticeable change in material shrinkage, the authors report, following best-fit registration of the scanned stone dies and the master die, a mean deviation between test groups of:

"about 10 µm. However, the maximum deviations reached up to 100 µm"

---

fact, example scans provided from a Carestream IOS contained cusp tip triangle edges of 0.4mm. This would obviously have a negative effect on any attempts of recreating correct articulation.

(Luthardt *et al.*, 2003).

In 2012, Van der Meer *et al.* raised the issue of the trend in the literature of reporting a single figure to give an indication of the quality of a scan, as seen in the quoted example from Luthard *et al.*(van der Meer *et al.*, 2012). This is a salient point, as the most common method of reporting scanner accuracy in the literature is a single figure representing the mean deviation of the complete scan from either a repeated scan (to assess precision) or a scan produced by a validated, traceable scanner (to assess trueness). Similarly, most commercial scanners are released with a statement claiming a "validated *x* micron accuracy", which one might argue holds little informative value to an end-user (see also *1.2.4: The issue with standards*).

Most current methods for measuring scanner precision rely on *black box solutions*[5], often using, or repurposing, commercial engineering and metrological software packages, which may not have been designed for this purpose. This may, as a result, lead to erroneous conclusions. For instance, collaborative work undertaken with investigators at Kings College London found that measuring erosion using commercial software, not designed with this task in mind, may under-report dental erosion (O'Toole *et al.*, 2019b; Marro *et al.*, 2020). This problem is not solely restricted to the virtual quantification of tooth wear. The key issue is exclusively relying on global alignment of scan data prior to analysis. This is discussed in more detail in section 1.3.1 and in Chapter 2.

An alternative to relying on alignments to investigate the accuracy of 3D data is by measuring real-world distances, such as the width of a tooth, or an arch, or any linear distance between multiple key points, as seen in (van der Meer et al., 2012; Braian and Wennerberg, 2019; Mangano et al., 2020). The precision with which a scanner can reproduce data which measures the identical distance across repeated scans might be a more conceptually accessible metric than other more statistical and process-heavy methods, or indeed, a single

[5] *Black box* is a term commonly used within the digital field to describe tools and software in which parameters and settings have been 'hard coded' into the software, meaning that the end-user has less control, and often little knowledge, of the underlying methods and processes which they are using. While often convenient for the user, this is problematic as it may affect results and impede reproducibility.

In addition, black box packages (be it the scanner software or analysis software used by investigators) may modify the data, unbeknownst to the user. This might take the shape of automatically closing scan holes, sharpening edges (in the case of dental scanner software) and remeshing and/or simplifying the data in question. All are actions which result in the data produced being slightly different to the data the user fed into the software — and which the user assumes they are working on…

mean value measurement. Kuhr *et al.* presented a method of investigating the trueness of impressions and IOSs *in vivo,* by attaching metal spheres to the cusp tips of the test subjects, and measuring the distance between spheres, and deviation in angle of virtual planes fitted to these spheres, across various scanner systems. This is one of few examples of methods with which to assess the *in vivo trueness* of IOSs (Kuhr *et al.*, 2016). The inherent risk of such methods is that the results are operator dependent, unlike the mean distance measurement between scans, which is calculated by the computer, with little human input.

Due to the financially driven nature of technological developments within dentistry[6], the risk remains of quantitative data being interpreted incorrectly to present enhanced, biased, or even falsified, results. This risk becomes all the greater with much research being undertaken by clinicians and other researchers whom may not necessarily have a deep understanding of the low-level workings of 3D scanners and 3D data.[7] There is a risk of well-meaning research foundering in the mists of data analysis, and producing flawed or erroneous data due to lack of awareness of underlying processes within scanners, and analysis software. More conceptually accessible methods of evaluating 3D data than those currently in use would therefore be of great value to the field.

1.2.4. The issue with standards

There is no single standardised method for measuring 3D data within the dental field, nor is there a standardised method for measuring the trueness and precision of intraoral dental scanners. The International Organization for Standardization's (ISO) effort to provide a standard by which to assess the accuracy of digital scanning devices may offer an introduction into the real challenges of quantifying and assessing the accuracy of 3D surface data. BS EN ISO 12836:*2015 (Dentistry — Digitizing devices for CAD/CAM systems for indirect dental restorations — Test methods for assessing accuracy*)(British Standards Institution, 2015) can be used to verify the accuracy of laboratory scanners. This standard has occasionally, somewhat misleadingly, been used by researchers investigating intraoral scanners, despite the standard explicitly stating that it does not cover handheld scanning devices. The obvious reason for this misconduct is that there (at the time of writing) is no standard available for the validation of handheld scanning devices, as previously discussed by (Braian, 2018). (It is also

[6] Brontes Technologies, the start-up which developed what eventually became known as the Lava Chairside Oral Scanner, was acquired by 3M for $95 million in 2006 (Writer, 2006). In 2009, 3M announced that the Lava platform (in USA and Canada) had been sold (3M, n.d.). While the Lava scanner is still commercially available, it is no longer actively promoted, and unlikely to have been considered a successful investment.

[7] Myself included, at the beginning of this project.

worth noting that most manufacturers of indirect scanners fail to document their ISO 12836 verification in a way that it is available to the customer, despite the standard stating this as a requirement[8]). An alternative standard is the VDI 2634, though this is generally used by more general, not dental-specific, 3D scanners (Luhmann and Wendt, 2000).

It has become evident that passing the requirements for ISO 12836 verification is considered crucial by most dental scanner developers, even for IOS. One lead engineer described this, in conversation, as "jumping through hoops", and a tick-box exercise for which the team would "tune" their scanners. While not necessarily a damaging approach, whether the ISO 12836 metric provides a valid measure of a scanner's accuracy when scanning clinical topology, as opposed the proposed standard scan objects, —and whether such "tuning" improves the performance of the scanner, or not— is a matter very much worthy of further investigation, but falls beyond the scope of this thesis[9]. The fact that the entire IOS industry relies on, and is shaped by, a standard which specifically states it *does not* cover hand-held devices, gives an insight into how financial motivations may — blindly — steer technological developments within the field.

Standardisation is never simple. It became clear early on during this research that developing and proposing a standardised method for measuring scanner accuracy would provide very little, except yet another proposed standard (Figure 1:2) [10]. As such, though this work does introduce a handful of new methods for assessing the quality of data, it aims to investigate the clinical relevance of various methods of measuring scan data, traditional and novel, with the hope of raising an awareness of the problematic nature of an industry or field placing its trust blindly in standardisation procedures and gold-standard analysis methods.

[8] I have, as fits a tunnel-visioned PhD researcher, tried to obtain the ISO 12836:2015 test documentation from several of the large manufacturers within the field without any luck (but with a fair share of ambivalence).

[9] An anecdotal footnote: In computer science, specifically the field of 3D face prediction, benchmark data has traditionally been made available to the community. As a result, many developers "tune" their algorithms to work optimally with the benchmark data, as this results in optimal results in publication. This often means that the algorithms have been optimised to such an extent they fail to perform "in the wild" — despite this being their originally intended purpose. There are moves in the field to use "hidden" benchmark data, against which developers can submit their algorithm to be tested by independent researchers. Anecdotally, this has resulted in a dramatic reduction in optimally performing algorithms, and in the number of new publications.

[10] See also Whitehouse, D.J. 1982. THE PARAMETER RASH - IS THERE A CURE? *Wear: an international journal on the science and technology of friction lubrication and wear*. **83**, pp.75–78. (Whitehouse, 1982)

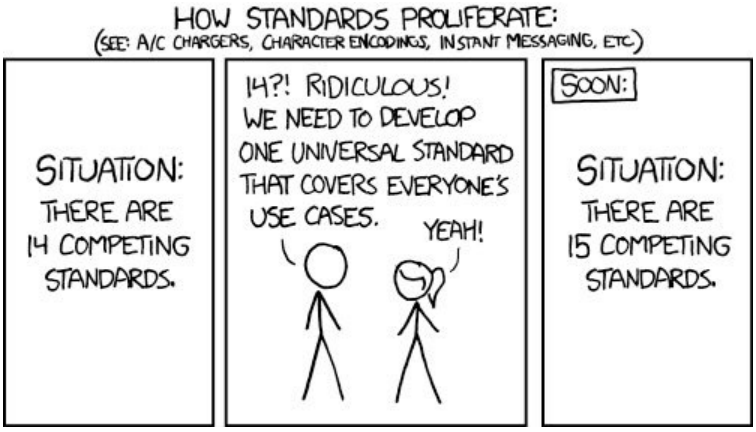


Figure 1:2 XKCD by Randall Munroe

Much of the unpublished preliminary work undertaken as part of this research focused on the current standards and methods for assessing trueness. It quickly became clear that investigating quantitative trueness is a highly problematic topic, as summed up in Hasong Chang's deliberations on scientific measurement:

"Among physicists, those who are involved in the testing of complicated and advanced theories by means of elementary observations would be in a relatively straightforward epistemic position [...]. But for those who try to justify the reasoning that justifies the elementary observations themselves, it is very difficult to escape circularity. The basic problem is clear: empirical science requires observations based on theories, but empiricist philosophy demands that those theories should be justified by observations. ***And it is in the context of quantitative measurement, where the justification needs to be made most precisely, that the problem of circularity emerges with utmost and unequivocal clarity****."* (Chang, 2005)[11]

Or in simpler terms: traceability. One can only ever assume that a measurement is true to within the margin of error of the device against which it was calibrated and/or manufactured. For a PhD student on the quest to frame and narrow down their research, this is a tempting academic booby-trap and time-drain, which I am grateful to have escaped mostly unscathed. I thus leave the challenge of truly determining the trueness of 3D scanners to a braver, and more patient, researcher.

[11] page 221, Chapter 5: Measurement, Justification and Scientific Progress. Emphasis in quotation is my own.

1.3. A brief introduction to 3D data and 3D scanners

Up until this point we have considered the current state of 3D scanning for clinical use, with a brief introduction into methods commonly used to measure 3D scan data, and some of the issues related to this. At this stage, it may be appropriate to briefly cover some of the fundamentals of 3D data and 3D scanners.

The first stage of any digital workflow is data acquisition, or, getting the relevant information digitised. Within dentistry this is achieved through 3D scanning.[12] The exact method by which scanners obtain the relevant 3D information and converts this to a presentable 3D file (a clean, manifold mesh, in most cases saved as an STL or PLY file before reaching the end-user) depends on the scanner type and the manufacturers' design decisions. No matter the scan method, the goal is to accurately, virtually reproduce the real-world object. (Whether the data acquisition method of the various scanners on the market affect the quality of the scan data is *mostly* beyond the scope of this work, and did not make an appearance in any of the publications enclosed, but will be briefly revisited in Chapter 6.)

Current scanning technologies have been built on work developed in applied geometry (Besl and McKay, 1992; Gruen and Akca, 2005), optics (Datta *et al.*, 2009; Zhou *et al.*, 2013; Schmalz *et al.*, 2015), metrology, and computer graphics research (Ahuja and Coons, 1968; Zorin *et al.*, 1996). Factors which are considered priority within these fields, such as speed, and financial and processing costs, may not be of equal importance to a clinician, whose priorities may differ greatly from those of, say, a theoretical mathematician. Unfortunately, most clinical end-users are unlikely to be aware of these differences in priorities and are therefore unable to critically assess how such design decisions may affect the digital data presented to the user. This will be explored further, following an introduction to the 3D scanning process. An introduction to the 'black box' of scanning will allow the reader to get a more rounded understanding of how 3D scanning works, and from this, what some of the most common pitfalls and shortcomings of 3D scanners are.[13]

[12] Other fields relating to less complex subject matters with less variation can occasionally (re)produce the relevant information using CAD.

[13] It may be worth noting that any references to algorithms and publications in the following sections are *examples* of processes that may be taking place within dental scanners. A well-performing alignment or surfacing algorithm may well result in a scanner performing better than the competitor. It is therefore nearby impossible for an independent researcher to uncover the exact processes in use in scanners currently available on the market.

Any reader familiar with the literature on the quality of dental scan data would be excused for believing that the "STL" or mesh available following scanning is the raw scan data. This is not the case. To gain an understanding of what can, and does, go wrong prior to producing the STL file, one needs a deeper understanding of the underlying processes which take place while the scanner is calculating the final mesh. This involves numerous stages, all out of sight of the user. It is noteworthy that all these stages may result in a degree of data degradation, as points are averaged, resampled, recalculated or simply removed. And to understand how, and why, these processes take place, a summary of the very basics of 3D data will now be described.

A 3D dataset, at its very lowest level, consists of vertices, or, 3D coordinates. A vertex is colloquially referred to as a *point*, and a dataset consisting of multiple vertices is a *point cloud*[14].

A 3D scanner — *scanning* — collects samples of coordinates which produce a point cloud. Due to the many challenges that present themselves when trying to determine the location of any given point in space, a large number of samples may be collected, before an average point is generated, to represent the sample. Hence, even at this very early stage, filtering is occurring, and an average estimate point cloud is generated (Le *et al.*, 2017).

Point clouds are of little use for clinical — or any physical — use, where a continuous surface is required. Therefore, the point cloud, is *surfaced* to produce a 3D mesh. This is a collection of triangles[15], connecting vertices, to create a continuous surface. While one may conceptually think of this as *joining the dots*, it may be worth pointing out that most surfacing algorithms re-sample the point clouds, resulting in a mesh whose vertices are in different locations to those of the original point cloud (Kazhdan *et al.*, 2006)[16].

Surfacing, while a necessity, brings with it several issues which can negatively affect the clinical usability of the data. Namely, producing sharp corners and edges is a challenge to a surfacing algorithm, whose very nature is to average any wildly disagreeing values (such as two points on a corner, facing in very different directions) to produce a concordant surface (Demarsin *et*

[14] Much like their 2D equivalent, pixels, vertices are not little cubes or spheres, *just* coordinates (Smith, 1995). However, for one's sanity, and teaching purposes, vertices are often represented as little squares — like postage stamps — as this allows us to consider the concept of a 3D coordinate possessing a back and front face (which is indicated by the vertex and/or triangle face *normal)*.

[15] Or, in inefficient cases, more complex planar shapes. But most, if not all, modern dental software use triangle-based meshes.

[16] …which was already an average estimate of the raw point data.

al., 2007; Cao *et al.*, 2012; Le *et al.*, 2017)[17]. The result is often a rounded edge or corner, much to the disadvantage of the clinician (Keeling et al., 2017). Nedelcu *et al.* report that the finish line distinctness of various intraoral scanners vary greatly, and that scanners which produced meshes with non-uniform tessellation tended to produce scans with a more distinct finish line; highlighting the fact that the method with which scanners surface their data is not always the same, and does affect the clinical usability of the data produced (Nedelcu *et al.*, 2018). Nedelcu *et al.*'s findings suggests that algorithms which rely on heavily sampling scan data in areas which contain interesting features, such as margins, and preserving fewer points in areas with little of interest may produce a more clinically useful scan, without creating an unnecessarily large data file. While a seemingly sensible approach, this does yet run the risk of affecting the clinical suitability of the data. As previously mentioned, accurate virtual occlusion could be negatively affected by an (even modestly) simplified mesh, when a patient is able to feel 0.01mm of difference between their teeth (Tryoe et al., 1962). Much work has been invested in solving the problems that come with surfacing and the topics of edge sharpening and the *acuity* of scan data are revisited in Chapters 3 and 6.

1.3.1. Alignment in data generation and measuring

Regardless of whether 3D data has been generated by a laboratory scanner or an IOS, alignment (or *registration)* is required to assemble the many scan views to produce the final scan (See Figure 1:3). Alignment also plays a key part in many methods currently used to investigate data quality, as highlighted in section 1.1. Alignment is a very challenging problem within computer science, in that there is no single true alignment solution. Identical scan data aligned multiple times may produce different outputs.

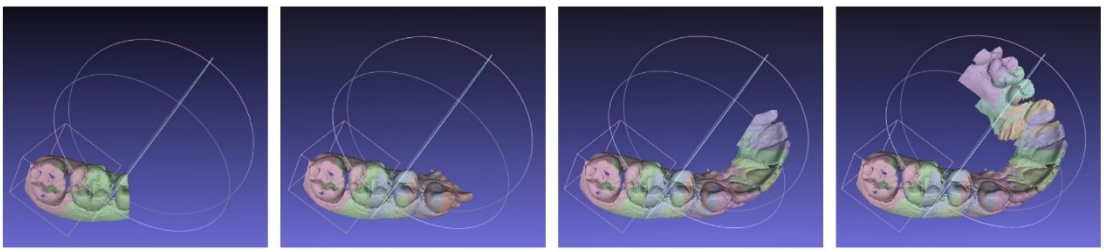


Figure 1:3: Illustration simulating data acquisition in an intraoral scanner. With thanks to Andy Keeling.

[17] A technically alternative, (but conceptually similar, to the non-programming end-user) approach with which to reconstruct scan data is presented in (Curless and Levoy, 1996).

In addition, ideal digital alignment is not the same as ideal clinical alignment. Much of the work presented within this thesis explores the consequences of relying on the problematic assumption that digital alignment produces clinically relevant results. Iterative closest point (ICP) algorithms work by taking a large number of distance measurements between two (rigid) meshes and minimizing the *average* distance between the meshes. As a result of this, the algorithm allows the meshes to perforate through one another if this reduces the average distance between the two. (See Figure 1:4). Clinically, such a scenario (where two physical objects perforate, non-destructively) is not physically possible, meaning that an ideal digital alignment can be less than ideal in a clinical setting. Thus, a clinical ideal alignment must take protrusions and undercuts into account and requires a determined path of insertion. The first positive error encountered along the path of insertion is the clinical fit, be it for a crown or a denture. (See Figure 1:5).

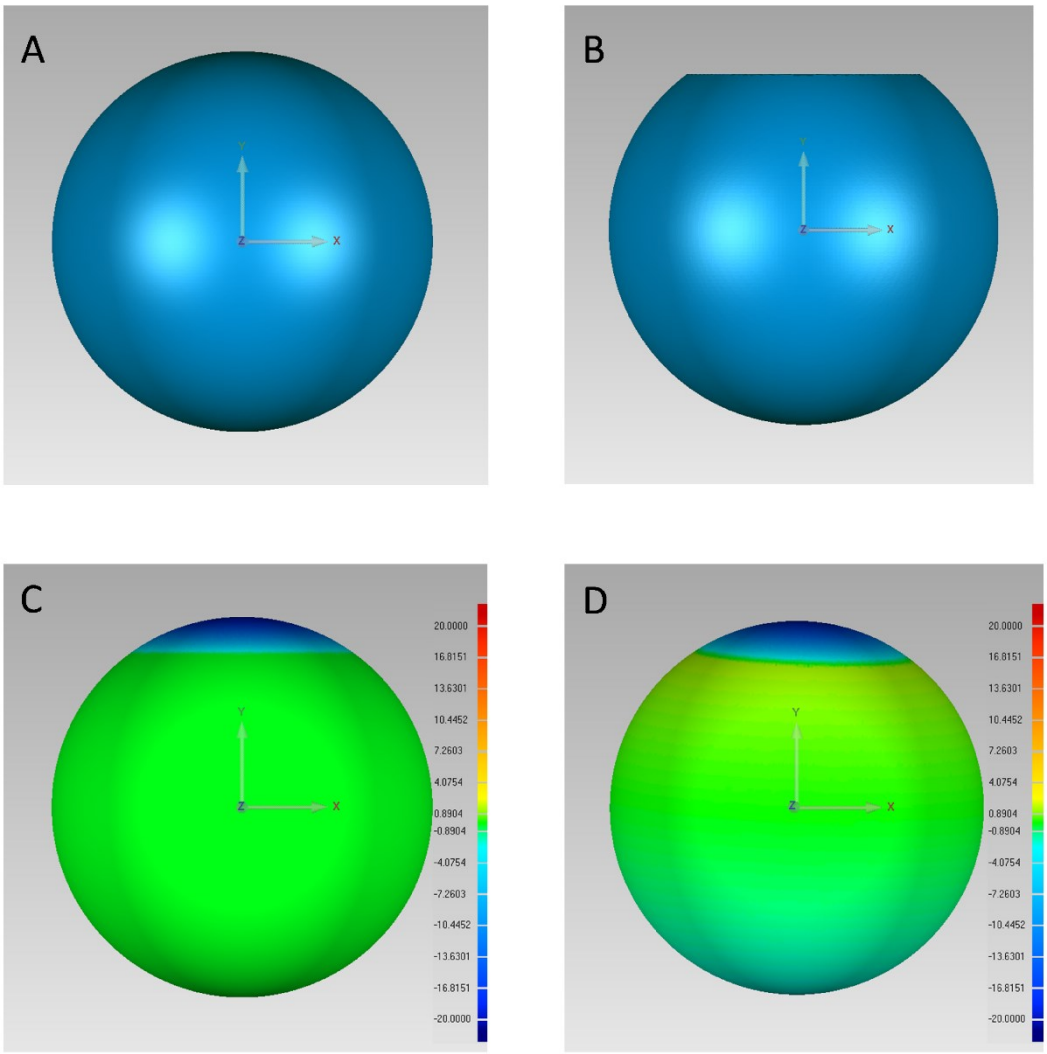


Figure 1:4: Illustrating the issue with global alignment A and B are perfectly identical spheres, except the top of sphere B has been removed. C shows perfect alignment of sphere A and sphere B, with the apices of the spheres displaying a difference between the two meshes, but all else being identical. D shows global alignment working to reduce the maximum deviation at any point between the two meshes, and as a result pulling sphere B upwards, introducing error at the base of the spheres, as illustrated by the change in colour. The yellow band towards the top indicates the sphere B perforating through sphere A producing a 'positive' error. This experiment was done using Geomagic, which is, undoubtedly, the most common software package used within the dental literature for 3D analysis. Thanks to Saoirse O'Toole for the screenshots.

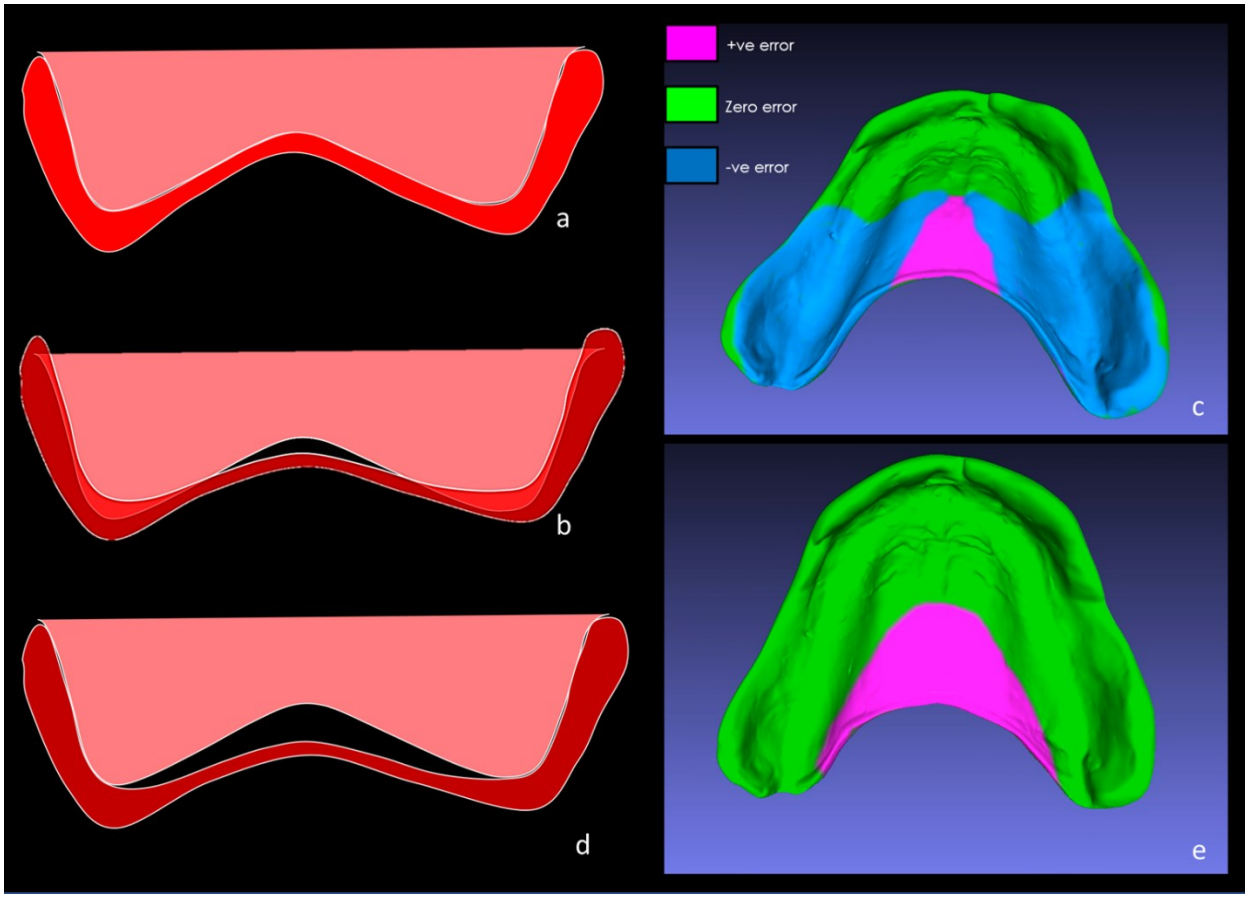


Figure 1:5 Illustrating the difference between optimal digital and optimal clinical alignment of a denture in the patient's mouth. a) illustrates optimal fit, requiring a perfectly fitting denture. b) illustrates an imperfectly fitting denture having been "optimally" aligned. The computer assumes this is as perfectly aligned as it is possible to fit two differently shaped objects. This alignment is not physically possible and is show in c) as a colourmap of where the denture and patient's gums interact. Blue indicates areas where the denture has passed beyond the patient's gums, pink shows areas where there is a gap between the denture and gum and green shows ideal fit. d) and e) illustrate optimal clinical alignment, where no area of the denture passes through the patient. The ICP alignment algorithm would report this as incomplete alignment. Thanks to Andy Keeling for illustrations and screenshots.

The selection criteria by which the ICP (Iterative Closet Point) algorithm selects its sample points between which to take each measurement depends on the implementation of the algorithm. Besl and McKay propose measuring between a vertex on mesh A to the closest vertex on mesh B, while Chen and Medioni suggest measuring from a vertex on mesh A to the closest intersecting plane on mesh B (Chen and Medioni, 1991; Besl and McKay, 1992). Chen and Medioni's point to plane method is generally considered likely to produce improved convergence and accuracy (Pomerleau *et al.*, 2013). While a seemingly abstract, and minor, difference, the two methods would produce somewhat different results if used to measure the same two meshes, and of some importance when investigating data which may present seemingly minute, but potentially relevant, errors. This also highlights the issue of using black box systems to align and investigate scan data, as these rarely allow the investigator the

opportunity to make active decisions regarding the most appropriate method with which to align data, and rather presents a “one button does it all” with limited user input, as previously discussed in (O’Toole *et al.*, 2019b).

1.3.2. Summary of working with 3D data

It is not my intention to cover all aspects of the 3D data pipeline which takes place in 3D scanners. However, there are a couple of key messages that lay the foundations for the following work:

- All 3D scanners work by calculating spatial coordinates which are stored as a raw point cloud.
- Scan data undergoes multiple processing steps (including, but not limited to: cleaning, alignment, simplification and surfacing) before the final mesh (STL or PLY) is produced.
- Every processing stage may result in data degradation.
- Alignment is necessary both during scan generation and during most methods of analysis.
- There is no single, true solution to surface alignment. A computer aligning the same pair of scans multiple times may produce divergent solutions. (The only guaranteed alignment is in the mathematically trivial case of aligning two identical, but displaced, meshes (each with identical vertices) – a situation which will never occur when using different scans.)

1.4. Scope

The focus of this work is to explore novel methods for quantifying the quality of 3D surface scan data in dentistry. Much of the work uses data produced by IOSs, though the methods can also be applied to data produced by extraoral (laboratory) scanners.

In particular, this thesis seeks to determine how to measure 3D data in a clinically relevant way. To that end, the basics tenets of good prosthodontic treatment are investigated – namely high quality full arch models, correctly occluded in the desired relationship (inter-cuspal position or centric relation). Further prosthodontic requirements (for example, accurate crown preparation models with sharp margins and good emergence profiles) are beyond the scope of this thesis but are touched upon when relevant.

Illuminated by the findings of these methods, and the intuition that *not all digital data is equal,* an optimal digital method is proposed in Chapter 5 for the prosthodontic situation of aligning dental models in the centric relation position (useful, for example, in rehabilitative treatment). This method overcomes the largest single barrier to dentists 'going digital' – that they do not own an IOS – by keeping the clinical workflow entirely traditional and thus achievable. The novel measurement method developed throughout this thesis (key point analysis) is applied to the final experiment. It is shown that a simple modification to current laboratory practice (by digitally scanning the inter-occlusal records using an inexpensive, but high quality, clamp-less scanner) significantly improves the precision of the articulated digital models, compared to the standard practice of physically articulating the models, *then* scanning them. The experiment also hints that the precision exceeds that of IOS bite scans from a modern IOS and might therefore represent the state-of-the-art in high precision digital articulation for centric relation models, whilst remaining accessible to all dentists rather than the minority equipped with IOSs.

1.5. Statement of problems

- There is no *gold standard* method used to measure or report the quality of 3D data in dentistry
- Current standard methods of reporting scanner accuracy are often too abstract to be clinically relevant

- Novel, clinically relevant measurement methods are required. These will pave the way for developing digital (or hybrid analogue/digital) workflows which truly provide the best possible clinical standards of treatment

1.6. Aims, goals and objectives

The aim of this research is twofold. Firstly, attempts will be made to establish methods, both current and novel, for testing the accuracy of 3D dental arch scans in *a clinically relevant way,* with particular focus on full arch scanning for large appliance and virtual articulation. Secondly, using these methods, investigations will be made to establish the accuracy of various intraoral scanners, and a suggestion for a more precise method to digitally occlude models in centric relation will be explored.

Objectives include:

- Methods commonly used within the field to investigate scanner accuracy will be identified and tested
- Alternative methods with which to assess the quality of scan data will be proposed and tested
- All results will be considered with clinical applicability in mind

1.7. Indication of research approach

All work presented takes the form of *in vitro* laboratory studies. One notable difference from much *in vitro* work which has been undertaken within the field, is the tendency towards robustly performed investigations with a large number of repetitions wherever possible, with the aid of automation tools and batch processing.

1.8. Statement of contribution

There are several novel concepts presented in the following pages:

i. The "meaningless mean" and the inherent risks of relying on mean and median values of scan data are mainly, but not solely, explored in Chapter 2. While the problematic nature of relying on mean values in data analysis has been raised in other fields (Matejka and Fitzmaurice, 2017), this issue is rarely mentioned in the dental field. In this paper, we propose the use of an “upper-bound value” to reach a more clinically relevant value when investigation scan errors.

ii. A robust method for selecting features on similar, but inherently different, meshes is presented in Chapter 3. The upper-bound value method, presented in the previous chapter, is also simplified and made more conceptually approachable, by reporting the percentage of a scan deviating beyond 0.1mm. See Appendix 8.2 for a detailed description of the key point method.

iii. A method for measuring full arch distances using the aforementioned key point method is presented in Chapter 3 and again in Chapter 4. This method can give an indication of the location of distortion within an arch scan. Unlike previous physical feature-reliant methods (by adhering spheres or bars to the teeth for reference, for instance), this method can be applied to data collected *in vivo*, as well as *in vitro.*

iv. A method for apportioning sources of error in the inter-cuspal occlusion created from intraoral scans, using the key point method, is presented in Chapter 4. This allows us to identify the proportion of scan error which is caused by the arch scan, and the proportion of error introduced during the bite registration stage.

v. Chapter 5 uses the key point method, and the findings on arch accuracy of the Omnicam presented in Chapter 4, to investigate the accuracy of using bite records to occlude dental models in centric relation, using a traditional and a virtual (hybrid) method of articulation. This method, using a clamp-less laboratory scanner, could be seamlessly integrated into a clinician’s workflow without the practitioner having to change their clinical routine.

vi. Lastly, in section 6.3: *The current state of full arch scanning and the effect of different measures of 'accuracy'* the results found in Chapter 3 are revisited, highlighting the real value of using appropriate means of measurement, as illustrated by the three different methods producing three different results, and therefore conclusions, as to whether the device under investigation produces clinically appropriate scan data, or not.

1.9. List of references

1. 3M n.d. 3M™ True Definition Intraoral Scanner acquired by Midmark in U.S. and Canadian Markets. *3m.com*. [Online]. [Accessed 11 December 2020]. Available from: https://www.3m.com/3M/en_US/dental-us/expertise/digital-dental-impressions/.

2. Ahlholm, P., Sipila, K., Vallittu, P., Jakonen, M., Kotiranta, U., Sipilä, K., Vallittu, P., Jakonen, M., Kotiranta, U., Sipila, K., Vallittu, P., Jakonen, M. and Kotiranta, U. 2018. Digital Versus Conventional Impressions in Fixed Prosthodontics: A Review. *Journal of Prosthodontics*. **27**(1), pp.1–7.

3. Ahuja, D. V. and Coons, A. 1968. Interactive Graphics in Data Processing: Geometry for construction and display. *IBM Systems Journal*. **3.4**, pp.188–205.

4. Alghazzawi, T.F. 2016. Advancements in CAD/CAM technology: Options for practical implementation. *Journal of Prosthodontic Research*. **60**(2), pp.72–84.

5. Besl, P.J. and McKay, N.D. 1992. A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. **14**(2), pp.239–256.

6. Bohner, O.L.L., Neto, P.T., Ahmed, A.S., Mori, M., Lagana, D.C. and Sesma, N. 2016. CERECChairside Systemto Registerand Designthe Occlusion in Restorative Dentistry : A Systematic Literature Review. *Journal of Esthetic and Restorative Dentistry*. **28**(4).

7. Braian, M. 2018. *Studies on the trueness and precision of additive manufacturing and intraoral scanning*.[Online] Malmö University. Available from: https://www.diva-portal.org/smash/get/diva2:1404654/FULLTEXT01.pdf.

8. Braian, M. and Wennerberg, A. 2019. Trueness and precision of 5 intraoral scanners for scanning edentulous and dentate complete-arch mandibular casts: A comparative in vitro study. *Journal of Prosthetic Dentistry*. **122**(2), pp.129–136.

9. British Standards Institution 2015. BS EN ISO 12836:2015: Dentistry - Digitizing devices for CAD / CAM systems for indirect dental restorations — Test methods for assessing accuracy

10. Cao, J., Wushour, S., Yao, X., Li, N., Liang, J. and Liang, X. 2012. Sharp feature extraction in point clouds. *IET Image Processing*. **6**(7), pp.863–869.

11. Chang, H. 2005. *Inventing Temperature: Measurement and Scientific Progress*.

12. Chen, Y. and Medioni, G. 1991. Object Modeling by Registration of Multiple Range Images. *Proceedings - IEEE International Conference on Robotics and Automation*., pp.2724–2729.

13. Chochlidakis, K.M., Papaspyridakos, P., Geminiani, A., Chen, C.J., Feng, I.J. and Ercoli, C. 2016. Digital versus conventional impressions for fixed prosthodontics: A systematic review and meta-analysis. *Journal of Prosthetic Dentistry*. **116**(2), pp.184–190.
14. Curless, B. and Levoy, M. 1996. Volumetric method for building complex models from range images. *Proceedings of the ACM SIGGRAPH Conference on Computer Graphics*., pp.303–312.

15. Datta, A., Kim, J. and Kanade, T. 2009. Accurate camera calibration using iterative refinement of control points. *Computer Vision Workshops (…*., pp.1201–1208.

16. DeLong, R., Ko, C.C., Anderson, G.C., Hodges, J.S. and Douglas, W.H. 2002. Comparing maximum intercuspal contacts of virtual dental patients and mounted dental casts. *Journal of Prosthetic Dentistry*. **88**(6), pp.622–630.

17. Demarsin, K., Vanderstraeten, D., Volodine, T. and Roose, D. 2007. Detection of closed sharp edges in point clouds using normal estimation and graph theory. *CAD Computer Aided Design*. **39**(4), pp.276–283.

18. Ender, A., Attin, T. and Mehl, A. 2016. In vivo precision of conventional and digital methods of obtaining complete-arch dental impressions. *Journal of Prosthetic Dentistry*. **115**(3), pp.313–320.

19. Ender, A. and Mehl, A. 2011. Full arch scans : Conventional versus digital impressions - An in-vitro study. *International Journal of Computerized Dentistry*. **14**(21).

20. Ender, A., Zimmermann, M., Attin, T. and Mehl, A. 2015. In vivo precision of conventional and digital methods for obtaining quadrant dental impressions. *Clinical Oral Investigations*. **20**, pp.1495–1504.

21. Ender, A., Zimmermann, M. and Mehl, A. 2019. Accuracy of complete- and partial-arch impressions of actual intraoral scanning systems in vitro. *International journal of computerized dentistry*. **22**(1), pp.11–19.

22. Gintaute, A., Keeling, A.J., Osnes, C.A., Zitzmann, N.U., Ferrari, M. and Joda, T. 2020. Precision of maxillo-mandibular registration with intraoral scanners in vitro. *Journal of Prosthodontic Research*. **64**(2), pp.114–119.

23. Gruen, A. and Akca, D. 2005. Least squares 3D surface and curve matching. *ISPRS Journal of Photogrammetry and Remote Sensing*. **59**(3), pp.151–174.

24. Güth, J.F., Edelhoff, D., Schweiger, J. and Keul, C. 2016. A new method for the evaluation of the accuracy of full-arch digital impressions in vitro. *Clinical Oral Investigations*. **20**(7), pp.1487–1494.

25. Güth, J.F., Runkel, C., Beuer, F., Stimmelmayr, M., Edelhoff, D. and Keul, C. 2017. Accuracy of five intraoral scanners compared to indirect digitalization. *Clinical Oral Investigations*. **21**(5), pp.1445–1455.

26. Jaschouz, S. and Mehl, A. 2014. Reproducibility of habitual intercuspation in vivo. *Journal of Dentistry*. **42**(2), pp.210–218.

27. Kazhdan, M., Bolitho, M. and Hoppe, H. 2006. Poisson surface reconstruction *In*: *Proceedings of the fourth Eurographics symposium on Geometry processing*.

28. Keeling, A. 2016. *Determination of the transverse horizontal axis and interocclusal registration using a novel optical technique*. University of Leeds.

29. Keeling, A., Wu, J. and Ferrari, M. 2017. Confounding factors affecting the marginal quality of an intra-oral scan. *Journal of Dentistry*. **59**, pp.33–40.

30. Keul, C. and Güth, J.F. 2020. Accuracy of full-arch digital impressions: an in vitro and in vivo comparison. *Clinical Oral Investigations*. **24**(2), pp.735–745.

31. Kuhr, F., Schmidt, A., Rehmann, P. and Wöstmann, B. 2016. A new method for assessing the accuracy of full arch impressions in patients. *Journal of Dentistry*. **55**, pp.68–74.

32. Le, M.T., Chen, L.C. and Lin, C.J. 2017. Reconstruction of accurate 3-D surfaces with sharp edges using digital structured light projection and multi-dimensional image fusion. *Optics and Lasers in Engineering*. **96**(April), pp.17–34.

33. Luhmann, T. and Wendt, K. 2000. Recommendations for an acceptance and verification test of optical 3-D measurement systems. *International Archives of Photogrammetry and …*. **XXXIII**, pp.493–500.

34. Luthardt, R.G., Kühmstedt, P. and Walter, M.H. 2003. A new method for the computer-aided evaluation of three-dimensional changes in gypsum materials. *Dental Materials*. **19**(1), pp.19–24.

35. Mangano, F.G., Admakin, O., Bonacina, M., Lerner, H., Rutkunas, V. and Mangano, C. 2020. Trueness of 12 intraoral scanners in the full-arch implant impression: a comparative in vitro study. *BMC oral health*. **20**(1), p.263.

36. Marro, F., Jacquet, W., Martens, L., Keeling, A., Bartlett, D. and O'Toole, S. 2020. Quantifying increased rates of erosive tooth wear progression in the early permanent

dentition. *Journal of Dentistry*. **93**(October 2019), pp.103-282.

37. Matejka, J. and Fitzmaurice, G. 2017. Same Stats , Different Graphs : Generating Datasets with Varied Appearance and Identical Statistics through Simulated Annealing *In*: *CHI 2017*. Denver, CO, USA.

38. van der Meer, W.J., Andriessen, F.S., Wismeijer, D., Ren, Y., Meer, W.J. Van Der, Andriessen, F.S., Wismeijer, D. and Ren, Y. 2012. Application of intra-oral dental scanners in the digital workflow of implantology. *PLoS ONE*. **7**(8), pp.1–8.

39. Nedelcu, R., Olsson, P., Nyström, I. and Thor, A. 2018. Finish line distinctness and accuracy in 7 intraoral scanners versus conventional impression: An in vitro descriptive comparison. *BMC Oral Health*. **18**(1), pp.1–11.

40. Van Noort, R. 2012. The future of dental devices is digital. *Dental Materials*. **28**(1), pp.3–12.

41. O'Toole, S., Osnes, C., Bartlett, D. and Keeling, A. 2019a. Investigation into the accuracy and measurement methods of sequential 3D dental scan alignment. *Dental Materials*. **35**(2019), pp.8–13.

42. O'Toole, S., Osnes, C., Bartlett, D. and Keeling, A. 2019b. Investigation into the validity of WearCompare, a purpose-built software to quantify erosive tooth wear progression. *Dental Materials*. **35**(10), pp.1408–1414.

43. Pomerleau, F., Colas, F., Siegwart, R. and Magnenat, S. 2013. Comparing ICP variants on real-world data sets. *Autonomous Robots*. **34**(3), pp.133–148.

44. Rekow, E.D. 2020. Digital dentistry: The new state of the art — Is it disruptive or destructive? *Dental Materials*. **36**(1), pp.9–24.

45. Rekow, E.D., Erdman, A.G. and Speidel, T.M. 1987. Review of Developments in Computer-Based Systems To Image Teeth and Produce Dental Restorations. *Proceedings - Annual Symposium on Computer Applications in Medical Care*., pp.566–570.

46. Rhee, Y.K., Huh, Y.H., Cho, L.R. and Park, C.J. 2015. Comparison of intraoral scanning and conventional impression techniques using 3-Dimensional superimposition. *Journal of Advanced Prosthodontics*. **7**(6), pp.460–467.

47. Schmalz, C., Angelopoulou, E., Schmalz, C. and Hw, C.T.T. 2015. Camera calibration : active versus passive targets Camera calibration : active versus passive targets.

48. Smith, A.R. 1995. A pixel is not a little square. *Pixar Technical Memo*. (6), pp.1–11.

49. Solaberrieta, E., Garmendia, A., Brizuela, A., Otegi, J.R., Pradies, G. and Szentpétery, A. 2016. Intraoral Digital Impressions for Virtual Occlusal Records: Section Quantity and Dimensions. *BioMed Research International*. **2016**.

50. Solaberrieta, E., Otegi, J.R., Goicoechea, N., Brizuela, A. and Pradies, G. 2015. Comparison of a conventional and virtual occlusal record. *Journal of Prosthetic Dentistry*. **114**(1), pp.92–97.

51. Straga, R.W. 2009. *Comparison of occlusal contacts on mounted models to contacts identified on digital 3D models using a new virtual alignment method*. University of Alberta.

52. Syrek, A., Reich, G., Ranftl, D., Klein, C., Cerny, B. and Brodesser, J. 2010. Clinical evaluation of all-ceramic crowns fabricated from intraoral digital impressions based on the principle of active wavefront sampling. *Journal of Dentistry*. **38**(7), pp.553–559.

53. Tryoe, G., Frydenberg, O. and Brill, N. 1962. An assessment of the tactile sensibility in human teeth: An evaluation of a quantitative method. *Acta Odontologica Scandinavica*.

54. Tsirogiannis, P., Reissmann, D.R. and Heydecke, G. 2016. Evaluation of the marginal fit of single-unit , complete-coverage ceramic restorations fabricated after digital and conventional impressions : A systematic review and meta-analysis. *The Journal of Prosthetic Dentistry*. **116**(3), pp.328-335.e2.

55. Whitehouse, D.J. 1982. The parameter rash — is there a cure? *Wear*. **83**(1), pp.75–78.

56. Writer 2006. 3M Snaps Up Brontes Technologies. *The New York Times*. [Online]. Available from: https://dealbook.nytimes.com/2006/10/17/3m-snaps-up-brontes-technologies/.

57. van der Zel, J.M. 2020. Review & Reports Bites for the Computer : Thirty Years of Digitization in Dentistry. . **2**(3), pp.1–7.

58. Zhou, F., Cui, Y., Wang, Y., Liu, L. and Gao, H. 2013. Accurate and robust estimation of camera parameters using RANSAC. *Optics and Lasers in Engineering*. **51**(3), pp.197–212.

59. Zorin, D., Schröder, P. and Sweldens, W. 1996. Interpolating Subdivision for meshes with arbitrary topology *In*: *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques* [Online]. ACM, pp.189–192. Available from: http://portal.acm.org/citation.cfm?doid=237170.237254.

60.

2. Full arch precision of six intraoral scanners *in vitro*

C.A.Osnes, J.H.Wu, P.Venezia, M. Ferrari, A.J.Keeling

Journal of Prosthodontic Research

Volume 64, Issue 1, January 2020, Pages 6-11

DOI: doi.org/10.1016/j.jpor.2019.05.005

2.1. Abstract

Introduction: Intraoral scanners may be used as an alternative to traditional impressions. That intraoral scanners produce precise scans is essential. Popular methods used to evaluate precision tend to be based on mean distance deviation between repeated scans. Mean value measurements may underestimate errors resulting in misleading conclusions which may influence clinical decisions. This study investigated the precision of six intraoral scanners using the traditional method of measuring signed mean error, and a proposed method considering only the most extreme and clinically relevant aspects of a scan.

Method: A metal edentulous model was scanned five times using six different intraoral scanners. The repeated scans were aligned, uniformly trimmed and mean surface deviation measured across all 20 possible scan combinations within each scanner group. All scan combinations were then measured using the proposed alternative method: by arranging all trimmed scan vertices from greatest to smallest unsigned distance from its compared scan and measuring the median value within the 1% of most greatly deviating points. 1% of a scan was estimated to represent a cumulative surface area of 37.5mm^2. Traditional mean deviation results, and upper-bound deviations, were compared.

Results: Measuring the upper-bound deviation within a scan reported scan errors up to two times greater than those found when measuring global mean distances. Results revealed clinically relevant errors of more than 0.3mm in scans produced by the Planmeca and Dentalwings scanners. These findings were not seen when using the traditional method of measuring mean distance error of the complete scan.

Conclusion: The upper-bound deviation of a cropped scan may provide a clinically useful metric for scanner precision. The Aadva, 3Shape, CEREC and TDS produced scans potentially appropriate for clinical use while Planmeca and Dentalwings produced deviations greater than 0.3mm, when measured using the upper-bound deviation method of analysis.

2.2. Introduction

Replacing traditional impressions and models with digital scans offer potential benefits for dental practitioners and patients. Previous studies have shown that digital scanners reduce the cost of storing and transporting impressions, and may also be preferred by patients (Wismeijer *et al.*, 2014). That a scanner produces precise and true virtual models is key.

Much work has been undertaken to validate the trueness and precision of intraoral scanners (IOSs). Precision, defined as the variability of repeated measurements, is commonly used when the real measurement value (trueness) is difficult to ascertain. Recent reviews show a tendency for digital impressions to be precise over short distances, but to fall short of their physical counterparts over larger arch spans (Chochlidakis *et al.*, 2016; Tsirogiannis *et al.*, 2016). Ahlholm *et al.* (2018) report that IOSs may result in acceptable scans when recording a single or partial fixed dental prosthesis preparation. Ender *et al.* (2016) found IOSs to be capable of recording clinically satisfactory quadrant impressions. Both Ender *et al.* and Ahlholm *et al.* found a significant difference in accuracy between traditional and digital impressions of full arches: digital impressions being less precise than high precision traditional materials. Ahlholm *et al.* conclude with the recommendation of the continued use of traditional impression techniques for full-arch impressions.

Nevertheless, recent attention has been directed towards the challenges surrounding edentulous scanning using IOSs. These include difficulties in recording muco-compressive impressions and functional depth and width of sulci, and the lack of interesting topology in edentulous regions causing cumulative errors in the stitching process of data during scanning (Gan *et al.*, 2016). Suggestions to add features clinically (such as painting pressure indicating paste or adding composite spheres to the mucosa (Kuhr *et al.*, 2016; Lee, 2017; Fang *et al.*, 2018), have been shown to help in recording a digital impression, but this is arguably no longer a true impression of the patient; and could be construed as flawed attempts to try and fit digital technology to a purpose for which other techniques might be more appropriate.

With the high rate of development within digital solutions, and as manufacturers respond to clinical requirements, software updates may improve data acquisition methods and stitching processes, calling for ever newer investigations.

However, the handling and analysis of the data extracted from, and for, scan comparisons vary within the literature. A key challenge is to distinguish between methodology-induced noise and valid, but erroneous, data. Another is to determine the threshold at which an error is clinically significant. There is a lack of agreement of clinically relevant metrics by which to compare 3D data within the dental field which could benefit from further investigation.

A common metric for precision is the mean distance deviation between scan-pairs. This method requires caution as data typically contains many thousands of points, and scan alignment algorithms serve solely to minimise the mean distance between two sets of such points (regardless of clinical fit). This phenomenon has been reported to result in the underestimation of tooth wear when standard alignment methods are used (O'Toole *et al.*, 2019a). Even if the alignment was perfect, the resulting regression to the mean of such large datasets can underestimate clinically relevant errors when using mean distance deviations. Small areas of significant inaccuracy (such as at a crown margin or an overextended sulcus) can be 'drowned out' by large regions of accurate smooth surface alignment (Keeling et al., 2017).

Several studies specify a confidence interval and perform all analysis on the data within this envelope, on the assumption that this will remove erroneous data (scanner noise) and retain valid scan data. Some investigators remove as much as the maximum and minimum 10% before analysis. Since 3D data from a scanner has already undergone statistical outlier removal before surfacing it may be counterproductive to remove such large quantities of potentially valid data and further compound the regression to the mean problem (Ender and Mehl, 2011; Ender et al., 2016). In summary, while maximum value removal excludes noise and outliers from the analysis, it may also exclude valid data; disregarding that the greatest true error in a scan can be of great clinical relevance.

The issue of relying on confidence intervals has been discussed previously (Güth *et al.*, 2013). In some cases, methods may allow for scan data to be consistently pruned or cropped to eliminate scan noise prior to analysis. Based on this premise, some investigators have chosen to analyse all data points, as opposed to applying a confidence interval(Güth *et al.*, 2013; Güth *et al.*, 2016; Güth *et al.*, 2017). In the case of the 2013 study evaluating 'un-cropped' data, findings were reported to show similar results to a cited study which relied on a "80-20 percentile method." This is to be expected when comparing mean distance deviation, as discussed above, with mean values potentially drowning out clinically relevant differences.

In this paper we compare the precision of six IOSs in scanning a replica edentulous maxillary arch. We use this model as it is likely to provide a challenge for IOSs in terms of smooth

anatomy. We assess the precision first by measuring conventional signed mean deviations over the full surface. We then assess the unsigned median error over the poorest 1% of the surface, and we examine whether this data is likely to be scanner noise, or valid data. We calculate the surface area that this 1% represents on our model to assess its clinical relevance.

Null Hypotheses

(A) That there are no significant differences between the precision of edentulous scans using six different intraoral scanners when assessed using standard signed mean deviation.

(B) That there are no significant differences between the precision of edentulous scans using six different intraoral scanners when assessed using the poorest fitting 1% unsigned median deviation.

(C) That there are no significant differences in the clinical acceptability of each scanner under both analyses, when considering a threshold of < 0.3mm as a clinically acceptable error.

2.3. Method

A metal edentulous model was scanned by the same operator using six different intra oral scanners, namely: True Definition Scanner [*TDS*] SW 5.1.1 (3M, Minnesota, United States), *Planmec*a Emerald SW 4.6 (Planmeca, Helsinki, Finland), *Omnicam* SW 4.5.2 (Sirona, now Dentsply Sirona, Pennsylvania, United States), Straumann Cares Intraoral Scanner, previously Dental Wings [*DWIO],* SW 2.1 (Straumann, Basel, Switzerland), TRIOS Model s1P, SW 1.4.7.5 (*3Shape,* Copenhagen, Denmark), and *Aadva* iOS100 (GC Corporation, Tokyo, Japan). Each scan was repeated five times and exported as STL files. All scans were aligned into a common coordinate frame with custom alignment software using the Generalized Iterative Closest Point (GICP) algorithm (Segal *et al.*, 2009) which has been shown to be more robust than standard ICP; our implementation was adapted from the PointCloudLibrary (Rusu and Cousins, 2011).

The first scan from each of the six scanners were all aligned. The four subsequent repeat scans were then aligned to the first scan within each group and mesh distances measured as detailed below. This process was repeated using the second, third, fourth and fifth scan as the 'base scan' for all scan sets, resulting in 20 possible combinations per scanner, and a total of 120 alignments over all six scanners.

To measure the surface deviations, a custom set of trimming planes were created and applied to all scans; this ensured that all scans were cropped identically, retaining only the clinically relevant surface extending to the functional depth and width of the sulcus and to the post dam region. This plane crop method trimmed through triangles of the mesh to prevent artificially induced errors caused by removing entire edge triangles. The mean surface area of the individual trimmed scan, and the mean surface area of each scan group, was calculated using Meshlab (Cignoni *et al.*, 2008), so that the magnitude of 1% of the surface area could be determined.

The scanning precision of each test group was assessed by two methods. Firstly, by measuring the signed mean surface deviation between each pair of scans, and the signed standard deviations between each scan-pair within an IOS group. The mean values for these metrics, over the 20 alignment pairs, were reported for each group. Secondly, the 1% of vertices which deviated by the largest unsigned amount were identified. For each scan-pair, these unsigned deviations were ranked and plotted to identify signs of noise (extreme or erratic errors) versus genuine erroneous data (smoothly decreasing errors). The median error value of these poorest 1% was noted to be clearly within 'valid' data for each scanner and was recorded for each alignment [referred to *as the upper-bound deviation* from here on]. The results from each comparison method were compared across scanners using ANOVA and Multiple Comparisons with Bonferroni correction.

The agreement between scanners was then tested by aligning and measuring all possible scan combinations across the six groups. This led to 25 separate alignments per group-pair (e.g. Planmeca 1-5 to Trios 1-5), and 30 possible group pairings; 750 alignments in total. The upper-bound deviation was recorded for each group-pair and assessed using ANOVA and Multiple Comparisons with Bonferroni correction.

Statistical significance was defined as $p < 0.05$ in all cases.

2.4. Results

The mean surface area of the trimmed model was 3753 ± 16.9mm^2. Therefore, 1% of the vertices represented approximately 37.5 ± 0.17mm^2.

Figure 1 shows colormaps of the trimmed surface deviation comparisons for representative samples from each scanner group. Aadva consistently displayed the lowest deviations while DWIO and Planmeca showed the largest. Mean positive and negative surface deviations

remained below 100 µm for all scanners as did standard deviations for individual scans. Table 1 shows global positive and negative mean deviations, and positive and negative standard deviations for all scanners.

Global positive and negative mean distances with standard deviations have been plotted in Figure 2a. See Appendix Table A1 for statistical significance of global mean measurements.

Figure 3 shows the ranked unsigned deviations for the poorest fitting 1% of vertices for each scanner. In all cases, disproportionately larger errors were noted for only the first small fraction of vertices – less than one tenth of the total 1% sample. This is indicated by the "L"-shaped descent of the poorest fitting vertices. The number of vertices comprising 1% of the scan varied between scanners, ranging from approximately 900 for DWIO up to 2000 for TDS. The mean distance of the medians of the poorest fitting 1% of vertices were: 0.103mm (TDS), 0.531mm (Planmeca), 0.153mm (Omnicam), 0.452mm (DWIO), 0.092mm (3Shape), 0.040mm (Aadva); these are plotted with standard deviations in 2b. These errors exceeded 0.3mm in two groups; DWIO (0.45 ± 0.19mm) and Planmeca (0.53 ± 0.21mm).

The comparisons across IOSs, as given by median of poorest 1% of vertices, are reported in Table 2.

2.5. Discussion

This study investigated the precision of different IOSs using two measurement methods *in vitro*.

There was a significant difference between some scanners investigated when assessing repeated scan precision using both the signed mean deviation method and the upper-bound deviation method. While the upper-bound deviation reported that the Planmeca and DWIO scanners exceeded the threshold of < 0.3mm, these findings were not seen using the standard signed mean deviation. Thus, all null hypotheses were rejected.

Several studies demonstrate that intraoral scanning *in vivo* reduces scan accuracy due to movement restrictions and the optically challenging environment within the oral cavity. Results obtained *in vitro* can therefore be assumed to be an optimal scenario and real-world clinical precision may be lower(Flügge et al., 2013; Kim et al., 2015; Keeling et al., 2017).

Despite the varying number of points produced per mesh from each scanner, the automated crop method used for this study resulted in a similar surface area for each scan (mean surface

area 3753 ± 16.9mm^2) as expected. This indicates that the alignment and cropping tools used are unlikely to have had much, if any, unfavourable effect on the total scan volume data prior to measurement, as this would likely have been reflected by greater variability in mesh surface area. However, a small scan/alignment error could conceivably have a disproportionate effect on the largest surface deviations following cropping. For example, a vertical surface that just missed a cropping plane in one model, but was just included in another, might produce a few very large error distances. For this reason, we visually assessed all the errors in the poorest 1% (Fig 3) to exclude any apparent outliers. An alternative would be to always compare the cropped source surface against a full version of the target mesh.

The current study specified that an error greater than 0.3mm at the 99.5% most deviating aspect of the scan would be considered likely to have a clinical impact. The exact percentage value would depend on the surface area of the scan and the intended clinical procedure. Errors below a maximum of 0.2mm have previously been reported as clinically acceptable for complete dentures (Mowery *et al.*, 1958). More recently, deviations in the posterior region of maxillary and mandibular dentures when flasked traditionally approached 0.25mm (Zampieri *et al.*, 2014). We therefore consider 0.3mm to be clinically relevant and inferior to current standard practice.

Signed mean distance measurements revealed that Aadva produced the most precise data, with statistically significant signed mean differences compared to all other scanners tested. The signed standard deviation of the Aadva was only significant compared to Planmeca and DWIO. This low mean deviation may be an indication towards the underlying algorithm used during data collection: a centre-of-voxel-based point-set simplification algorithm might create datasets with minimal variation across scan repetitions. While the precision of resulting scans is high, the trueness of a scanner using such an algorithm, impetuously, is questionable. However, trueness validation is beyond the scope of this study. Interestingly, the plotted max 1% points illustrated greater variation between Aadva scans than the 3Shape and TDS (Fig 3), suggesting more greatly deviating outliers, and hence a reduced precision by the Aadva than that suggested by the global mean values. Empirically, the Aadva scans appeared ‘simplified’, i.e. lacking surface detail, which may also imply an algorithm which collects the regional mean of multiple vertices. This would increase the precision but reduce the resolution (the ability to discern fine detail), in the same way that a medium bodied silicone impression might be precise, but not reveal the fine detail of a light-bodied wash.

Much less than the greatest 0.5% of the data appeared to be spurious (Fig 3). It was decided that the median of the greatest 1% of points, over the 20 comparisons for each scanner would be a safe indicator of the clinically relevant deviation produced by each scanner, albeit a likely underestimate of the real value. Plotting the first percent of the greatest error of each scan gave a clear indication of both the proportion of noise in the scan and the greatest error within the scan. In this experiment, 1% of the scan object was estimated to represent a cumulative surface area of 37.5 ± 0.17mm^2. This is equivalent to more than a 6mm x 6mm patch, larger than the occlusal surface of a premolar if all erroneous points were to be located within close proximity, and potentially clinically relevant.

While scanner trueness is beyond the scope of this study, comparing the variation between scanners may give an indication of trueness if several scanners reach similar conclusions. Whilst there were statistically significant differences in the upper bound error between all scanners, Omnicam, 3Shape and Aadva showed clinically acceptable variations (consistently < 0.3mm) when compared to each other. Whilst TDS precision was good (Fig 2b), there was a systematic discrepancy between its scans and those produced by Omnicam, 3Shape and Aadva. We may speculatively assume that these latter three scanners produced the truest scans. Planmeca and DWIO produced clinically intolerable disagreements with all other scanners (consistently > 0.3mm) and up to 1.16mm in the case of DWIO compared to TDS; casting doubt on the trueness of these two scanners (Table 2).

These findings may indicate that there is merit in evaluating the maximum deviation within scan data as an adjunct to global mean distance measurements, to provide a more clinically applicable assessment. Further, as seen in the results from the Aadva, immoderate use of processing algorithms may 'game the system' and produce erroneous conclusions. This highlights the need for further investigation into edge sharpness, acuity, and clinical applicability of scanners in relation to their reported precision and resolution. Notably, in our experience we have yet to see an IOS capable of producing the detail of a light-bodied silicone wash, such as the bur marks typically seen in a dental crown model. Accurate mesh vertices would be required every 10-20 µm for this, whilst typical scans currently show triangles with edge lengths often exceeding 100 µm. Further work is required to investigate the *resolution* of IOSs, and indeed, whether there is any clinical detriment in this mesh simplification. It is quite possible that there is no clinical disadvantage in working with simplified data, in which case analogue dentists could consider abandoning the use of light-bodied washes, saving cost and time.

The apparent high precision of the Aadva scanner introduces the factor of the underlying 'black-box' algorithms used to collect, align, and surface scan data. These proprietary algorithms are rarely discussed, validated or developed within the dental field. As such, processing speed may be prioritized over edge definition, or reduced file size over global trueness. Such factors would not be identifiable in a precision experiment such as this but could potentially have great impact on scanner trueness and clinical application. Further investigations are required to directly assess the factors of mesh simplification and edge definition.

Full arch scanning has been shown to be less accurate than conventional impressions(Ahlholm *et al.*, 2018).This is likely to be related to the problem of error accumulation and propagation whilst stitching multiple smaller scans together. Modern algorithms use loop-closure (a process where start and end points of a circular scanning path are stitched together to minimize this accumulated error). This process is simpler in the upper arch, where the palate is also scanned. It would be interesting to repeat this experiment using a lower arch form, which does not lend itself as readily to loop closure and might be expected to show poorer precision.

2.6. Conclusion

Both the traditional standard trimmed signed mean deviation method and the upper-bound method revealed a significant difference between the precision of some of the six intraoral scanners investigated.

The greatest global mean errors were produced by Planmeca and DWIO, but these fall just below our clinically relevant threshold of 0.3mm in optimal measurement conditions. The upper-bound deviations of both scanners produced clinically relevant errors greater than 0.3mm. Trios, Aadva, Omnicam and TDS all produced clinically acceptable scans according to both metrics.

We suggest future studies should report both mean distance measurements and upper-bound deviation to ensure inter-study comparability and promote clinically relevant investigations.

2.7. Figures

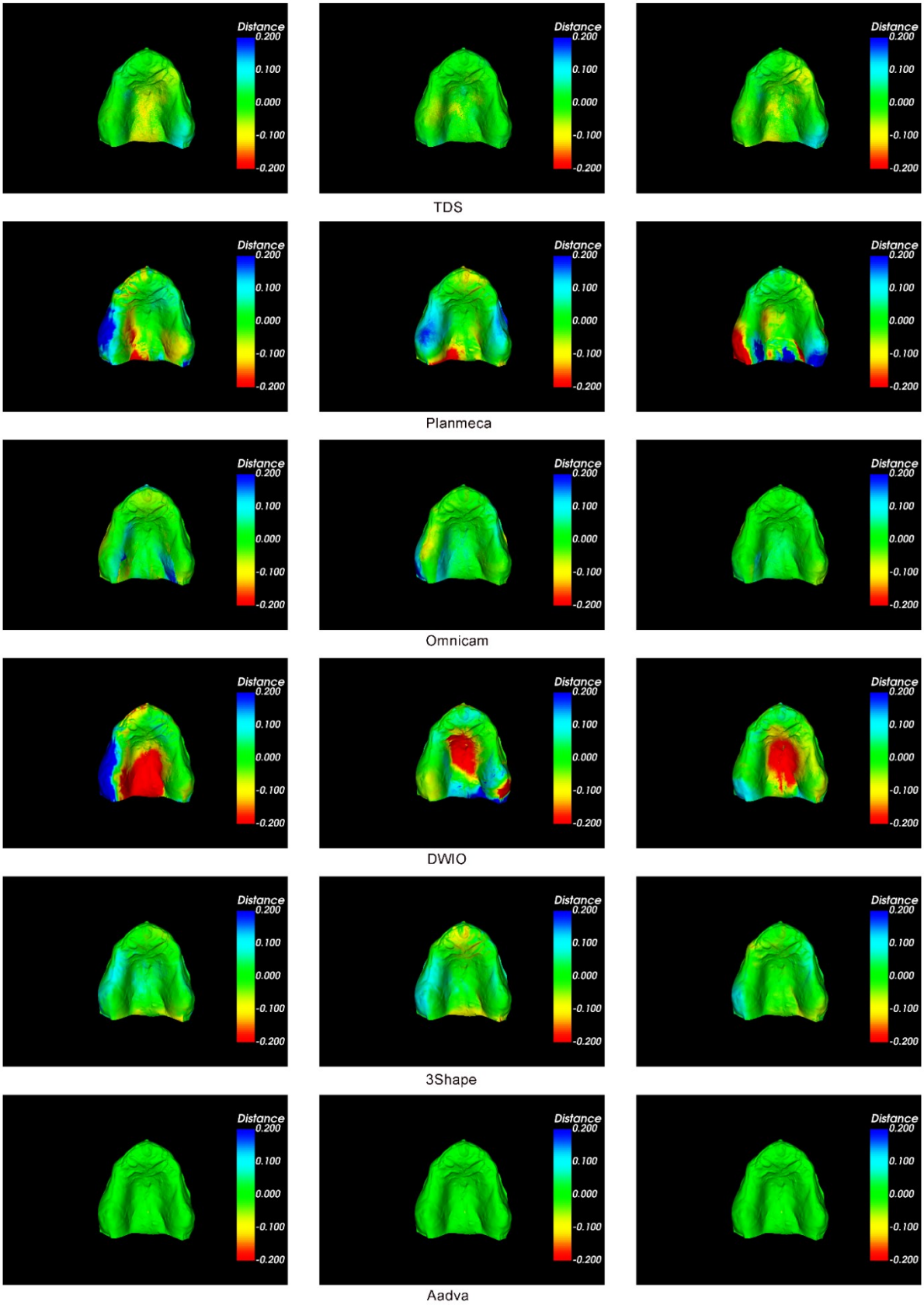


Figure 2:1 Colour map showing (from left to right) 2nd, 3rd and 4th scan compared to the 1st scan for each scanner, with distance measured in millimeters.

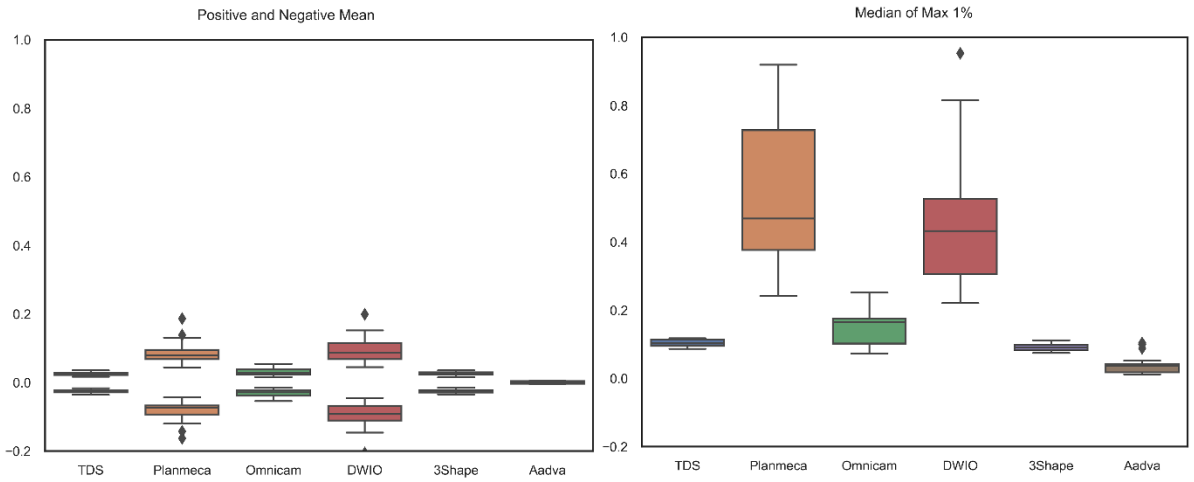


Figure 2:2: 2a Global positive and negative mean for each scanner.The line indicates mean value, the box upper and lower quartile, while the whiskers show overall distribution. Outliers are indicated with a diamond. Figure 2b Median of greatest 1%. Distance in mm. The line indicates mean value, the box upper and lower quartile, while the whiskers show overall distribution. Outliers are indicated with a diamond. [Unpublished correction: the line indicates median value, not mean, in both figures.]

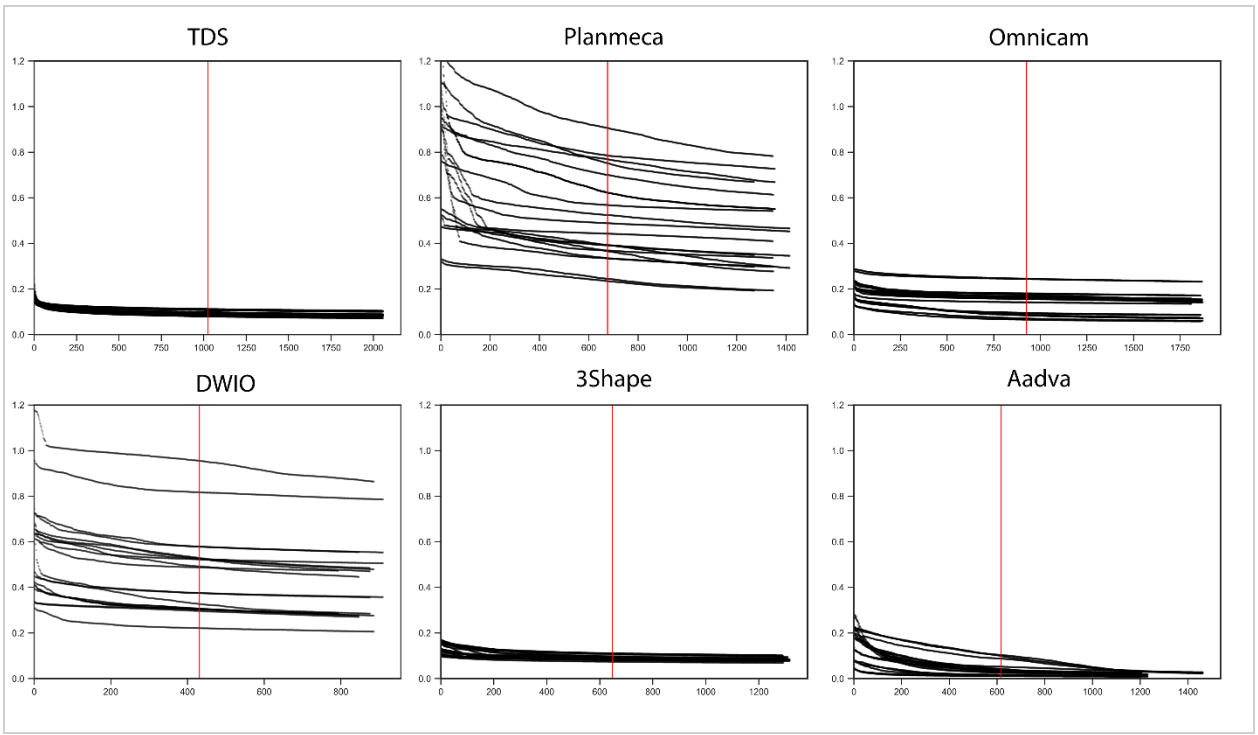


Figure 2:3 Maximum 1% of each scan-pair measurement plotted. Distance in mm.

Table 1 Global positive and negative mean for each scanner, see results plotted in Fig 2. (See supplementary material A1 for statistical analysis)

	Mean pos (SD)	Mean neg (SD)	Pos mean SD	Neg mean SD
TDS	0.025 (0.005)	-0.025 (0.005)	0.02	0.02
Planmeca	0.087 (0.034)	-0.084 (0.031)	0.098	0.092
Omnicam	0.032 (0.011)	-0.031 (0.011)	0.028	0.028
DWIO	0.097 (0.038)	-0.096 (0.039)	0.093	0.091
3Shape	0.026 (0.005)	-0.026 (0.005)	0.019	0.019
Aadva	0.003 (0.001)	-0.003 (0.001)	0.007	0.005

Table 2 *Comparison between the upper-bound deviation (mm) (median value of most greatly deviating 1% of each scan) across all scanners based on all 750 scan combinations.*

Upper-bound deviations (mm) compared across all scanners							
	TDS	Planmeca	Omnicam	DWIO	3Shape	Aadva	p-value
TDS		0.44 (0.20)	0.31 (0.05)*	1.15 (0.27)*	0.5 (0.1)*	0.26 (0.03)*	<0.001
Planmeca	0.39 (0.17)		0.40 (0.13)*	0.80 (0.25)*	0.48 (0.12)*	0.31 (0.13)*	<0.001
Omnicam	0.31 (0.04)*	0.41 (0.15)*		0.69 (0.20)*	0.15 (0.04)*	0.11 (0.04)*	<0.001
DWIO	1.16 (0.18)*	0.80 (0.25)*	0.70 (0.19)*		0.69 (0.18)*	0.77 (0.22)*	<0.001
3Shape	0.48 (0.09)*	0.50 (0.14)*	0.15 (0.05)*	0.65 (0.16)*		0.15 (0.01)*	<0.001
Aadva	0.26 (0.03)*	0.32 (0.14)*	0.12 (0.04)*	0.76 (0.22)*	0.16 (0.01)*		<0.001

Note: p-value represents the overall significance between scanners for a given model.

*denotes statistical significance at the 0.05 level with post hoc Bonferroni correction

2.8. Appendix [Unpublished]

Appendix/Table A1 *Comparison of positive and negative mean distance (mm) and standard deviation amongst scanners with ANOVA*

Multiple comparison of mean difference with Bonferroni correction						
Positive Mean	TDS	Planmeca	Omnicam	DWIO	3Shape	Aadva
TDS		.06160*	.00667	.07158*	.00119	-.02175*
Planmeca	-.06160*		-.05493*	.00999	-.06040*	-.08334*
Omnicam	-.00667	.05493*		.06492*	-.00547	-.02841*
DWIO	-.07158*	-.00999	-.06492*		-.07039*	-.09333*
3Shape	-.00119	.06040*	.00547	.07039*		-.02294*
Aadva	.02175*	.08334*	.02841*	.09333*	.02294*	
Negative Mean	TDS	Planmeca	Omnicam	DWIO	3Shape	Aadva
TDS		-.05900*	-.00624	-.07068*	-.00095	.02216*
Planmeca	.05900*		.05275*	-.01168	.05804*	.08116*
Omnicam	.00624	-.05275*		-.06444*	.00529	.02840*
DWIO	.07068*	.01168	.06444*		.06972*	.09284*
3Shape	.00095	-.05804*	-.00529	-.06972*		.02312*
Aadva	-.02216*	-.08116*	-.02840*	-.09284*	-.02312*	
Positive SD	TDS	Planmeca	Omnicam	DWIO	3Shape	Aadva
TDS		.07813*	.00824	.07325*	-.00043	-.01290
Planmeca	-.07813*		-.06989*	-.00488	-.07856*	-.09104*
Omnicam	-.00824	.06989*		.06500*	-.00867	-.02115
DWIO	-.07325*	.00488	-.06500*		-.07367*	-.08615*
3Shape	.00043	.07856*	.00867	.07367*		-.01248
Aadva	.01290	.09104*	.02115	.08615*	.01248	
Negative SD	TDS	Planmeca	Omnicam	DWIO	3Shape	Aadva
TDS		.07242*	.00782	.07075*	-.00043	-.01461
Planmeca	-.07242*		-.06460*	-.00167	-.07285*	-.08703*
Omnicam	-.00782	.06460*		.06293*	-.00825	-.02243
DWIO	-.07075*	.00167	-.06293*		-.07118*	-.08536*

3Shape	.00043	.07285*	.00825	.07118*		-.01418
Aadva	.01461	.08703*	.02243	.08536*	.01418	
*. The difference is significant at the 0.05 level, p-value was adjusted with Bonferroni correction.						

2.9. List of references

1. Ahlholm, P., Sipila, K., Vallittu, P., Jakonen, M., Kotiranta, U., Sipilä, K., Vallittu, P., Jakonen, M., Kotiranta, U., Sipila, K., Vallittu, P., Jakonen, M. and Kotiranta, U. 2018. Digital Versus Conventional Impressions in Fixed Prosthodontics: A Review. *Journal of Prosthodontics*. **27**(1), pp.1–7.

2. Chochlidakis, K.M., Papaspyridakos, P., Geminiani, A., Chen, C.J., Feng, I.J. and Ercoli, C. 2016. Digital versus conventional impressions for fixed prosthodontics: A systematic review and meta-analysis. *Journal of Prosthetic Dentistry*. **116**(2), pp.184–190.

3. Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F. and Ranzuglia, G. 2008. Meshlab: an open-source mesh processing tool. *Eurographics Italian*.

4. Ender, A., Attin, T. and Mehl, A. 2016. In vivo precision of conventional and digital methods of obtaining complete-arch dental impressions. *Journal of Prosthetic Dentistry*. **115**(3), pp.313–320.

5. Ender, A. and Mehl, A. 2011. Full arch scans : Conventional versus digital impressions - An in-vitro study. *International Journal of Computerized Dentistry*. **14**(21).

6. Ender, A., Zimmermann, M., Attin, T. and Mehl, A. 2016. In vivo precision of conventional and digital methods for obtaining quadrant dental impressions. . **20**, pp.1495–1504.

7. Fang, J.H., An, X., Jeong, S.M. and Choi, B.H. 2018. Digital intraoral scanning technique for edentulous jaws. *Journal of Prosthetic Dentistry*. **119**(5), pp.733–735.

8. Flügge, T. V., Schlager, S., Nelson, K., Nahles, S. and Metzger, M.C. 2013. Precision of intraoral digital dental impressions with iTero and extraoral digitization with the iTero and a model scanner. *American Journal of Orthodontics and Dentofacial Orthopedics*. **144**(3), pp.471–478.

9. Gan, N., Xiong, Y. and Jiao, T. 2016. Accuracy of intraoral digital impressions for whole upper jaws, including full dentitions and palatal soft tissues. *PLoS ONE*. **11**(7), pp.1–15.

10. Güth, J.F., Edelhoff, D., Schweiger, J. and Keul, C. 2016. A new method for the evaluation of the accuracy of full-arch digital impressions in vitro. *Clinical Oral Investigations*. **20**(7), pp.1487–1494.

11. Güth, J.F., Keul, C., Stimmelmayr, M., Beuer, F. and Edelhoff, D. 2013. Accuracy of digital models obtained by direct and indirect data capturing. *Clinical Oral Investigations*. **17**(4), pp.1201–1208.

12. Güth, J.F., Runkel, C., Beuer, F., Stimmelmayr, M., Edelhoff, D. and Keul, C. 2017. Accuracy of five intraoral scanners compared to indirect digitalization. *Clinical Oral Investigations*. **21**(5), pp.1445–1455.

13. Keeling, A., Wu, J. and Ferrari, M. 2017. Confounding factors affecting the marginal quality of an intra-oral scan. *Journal of Dentistry*. **59**, pp.33–40.

14. Kim, J., Kim, K., Kim, S., Kim, W., Kim, H. and Kim, J. 2015. Quantitative evaluation of common errors in digital impression obtained by using an LED blue light in-office CAD/CAM system. *Quientessenz International*. **46**(5).

15. Kuhr, F., Schmidt, A., Rehmann, P. and Wöstmann, B. 2016. A new method for assessing the accuracy of full arch impressions in patients. *Journal of Dentistry*. **55**, pp.68–74.
16. Lee, J.H. 2017. Improved digital impressions of edentulous areas. *Journal of Prosthetic Dentistry*. **117**(3), pp.448–449.

17. Mowery, W.E., Burns, C.L., Dickson, G. and Sweeney, W.T. 1958. Dimensional stability of denture base resins. *Journal of the American Dental Association (1939)*. **57**(3), pp.345–353.

18. O'Toole, S., Osnes, C., Bartlett, D. and Keeling, A. 2019. Investigation into the accuracy and measurement methods of sequential 3D dental scan alignment. *Dental Materials*. **35**(2019), pp.8–13.
19. Rusu, R.B. and Cousins, S. 2011. 3D is here: Point Cloud Library (PCL) *In*: *IEEE International Conference on Robotics and Automation (ICRA)*.

20. Segal, A. V, Haehnel, D. and Thrun, S. 2009. Generalized-ICP *In*: *Robotics: Science and Systems*. Seattle, WA, USA.

21. Tsirogiannis, P., Reissmann, D.R. and Heydecke, G. 2016. Evaluation of the marginal fit of single-unit , complete-coverage ceramic restorations fabricated after digital and conventional impressions : A systematic review and meta-analysis. *The Journal of Prosthetic Dentistry*. **116**(3), pp.328-335.e2.

22. Wismeijer, D., Mans, R., van Genuchten, M. and Reijers, H.A. 2014. Patients' preferences when comparing analogue implant impressions using a polyether impression material versus digital impressions (Intraoral Scan) of dental implants. *Clinical Oral Implants Research*. **25**(10), pp.1113–1118.

23. Zampieri, M.H., Consani, R.L.X., Ferraz, M.M., Marchini, L. and dos Santos, F.M.B. 2014. Effect of Flask Types, Post-pressing Times and Base Regions on the Adaptation of Complete Dentures. *Journal of Research and Practice in Dentistry*. **2014**(2014).

3. Investigating three methods of assessing the clinically relevant trueness of two intraoral scanners

C.A Osnes, J. Caplan, M. Ferrari, A.J Keeling

Journal of Osseointegration

Accepted January 2021

3.1. Abstract

Intraoral scanners (IOS) are increasingly used for a wide range of treatments. Most IOSs produce data appropriate for local work, such as crowns, but evidence suggests that full-arch scans result in more erroneous scans, which may affect the fit of clinical appliances. There are no standardised methods for assessing the quality of IOSs. Though many studies have investigated the accuracy of scanners, one may find the reported values are difficult to interpret in a clinical context.

This study investigated the trueness of two IOSs, using three metrics. The clinical value of each metric is discussed. A dentate model was scanned 10 times using two intraoral scanners. Three methods were used to assess the trueness of the scans against a scan produced in a laboratory scanner.

The mean unsigned distance deviation between a laboratory scan and the Primescan scans was 0.016(±0.006)mm. The mean unsigned distance deviation between the laboratory scan and the Omnicam scans was 0.116(±0.01)mm. The arch width between molars was 55.44mm for the Solutionix scan. The arch width of the Primescan was 55.439(±0.075)mm, while the Omnicam reported 54.672(±0.065)mm. The mean proportion of the Primescan scans deviating beyond 0.1mm when compared against the Solutionix was 0.7(±2.0)%. The equivalent for the Omnicam was 42.1(±2.5)%.

All methods indicated significantly different results between the scanners. The Primescan produced truer scans than the Omnicam, regardless of measurement method. The intermolar-width and proportion beyond 0.1mm methods may give more clinically relevant insight into the trueness of scan data than current gold-standard methods.

List of abbreviations

IOS: Intraoral Scanner

STL: Stereolithographic (file format)

3D: Three-dimensional

Keywords:

Intraoral Scanners, Trueness, Dentistry

3.2. Introduction

Intraoral scanners (IOS) can be of great convenience to the dental practitioner and their patients. Interest in capturing full arch scans has been growing over the past decade in an effort to increase the range of treatment modalities offered by the digital workflow. Investigations report a tendency for intraoral scanners to produce clinically acceptable digital impressions over short distances, while complete arch scans may suffer from distortions at a scale which may have potential clinical implications(Chochlidakis et al., 2016; Tsirogiannis et al., 2016; Ender et al., 2016; Ahlholm et al., 2018). Despite this, dental design software offers the ability to provide full arch prostheses based on intraoral scans. The onus is on the end-user to decide on appropriate use, based on the published evidence. Unfortunately, there is a lack of consensus on both the degree of trueness required for a particular procedure, and the metric by which to measure said trueness.

For example, the level of trueness required for full-arch implant work is unresolved, with studies citing a range of 10 to 150 microns as minimal tolerance for the passive fit of an implant framework (Kan *et al.*, 1999). Wismejer *et al.* report that "CAD/CAM technology has not eliminated the risks for hardware-related complications" in implant-supported reconstructions, implying that though single crowns and abutments can be reliably produced using CAD/CAM solutions the current performance of the complete CAD/CAM workflow does not fall within the tolerance required for optimal full-arch implant work (Wismeijer, Daniel; Buser, Daniel; Chen, 2019). As such, an awareness of the full-arch accuracy of a scanner could play a deciding part in whether or a not a clinician chooses to rely on a virtual impression as means of data acquisition in the digital treatment workflow.

An important question arises in how best to measure the quality of full arch dental scans. Any measurements should give a good clinical indication of the potential quality of fit of a prosthesis. The precision (repeatability) of a scanner is a common metric, whilst trueness is also useful when correct values are known *a priori* (though this is rarely the case *in vivo*). However, how best to measure and report accuracy (meaning both trueness and precision) is much disputed; and robustly assessing the quality of 3D (three-dimensional) scan data in a clinically relevant manner is an unsolved problem.

The authors have previously presented evidence to highlight the inherent flaws in using the commonly used metric of global mean deviation between repeated scans as a measure of clinical appropriateness (Osnes *et al.*, 2019). Mean deviation is likely to report smaller scan errors than might be present in a scan, leading to an overly optimistic appraisal. Likewise, as all

scan data produced by dental scanners have already undergone filtering and noise removal prior to output mesh generation (generally as an STL file), removing a portion of the most extremely deviating values of a data-set, as is often reported in the field, may result in artificially precise data prior to analysis.

Further, any measurement relying on aligning multiple scans will be subject to a margin of error; as there is rarely a single, true solution to 3D alignment problems. Investigating errors such as those accumulated over a full-arch scan can be a challenge greatly affected by the artificial minimisation of error, as investigated by O'Toole et al (O'Toole *et al.*, 2019a), and may result in analyses underreporting global errors. This may further bring the relevance of reporting the mean deviation of meshes into question, as this value may reflect more on the success of the alignment algorithm to minimise deviation than the quality of the scan data.

Ender et al (2019) investigated full arch versus segment scan (Ender et al., 2019). The authors report a generally higher precision in posterior segment scans. However, it might be suggested that the mean distance deviation reported for scan segments is likely to provide favourable results for the posterior segments; as straighter sections, as opposed to the curved anterior sections, are more likely to align in such a manner as to minimize any distance deviation between repeated scans. Because of this, aligning and measuring isolated segments discards most, if not all, cross arch error through optimal (though potentially incorrect) mesh alignment.

One possible solution to overcome the minimization of error caused by relying on scan alignments may be to forego measurements relying on global alignment, and instead measure distances between robustly identified key points within a single scan, as seen in (Kühle, 2003; Güth *et al.*, 2016; Gintaute *et al.*, 2020). This could provide an insight into any arch distortion introduced by the scanner during the scanning process, including location specific errors, without suffering from alignment minimisation artifacts. When considering the clinical fit of a full arch prosthesis, a metric such as cross-arch distance error might be considered more clinically relevant. The 2019 study (Gintaute *et al.*, 2020) used specific mesh vertices as virtual key points, as opposed to introducing physical features of interest onto the scan object to investigate virtual occlusion (Güth *et al.*, 2016; Kuhr *et al.*, 2016).

This study compared the trueness values reported using three different analysis methods on the same two sets of intraoral scan data captured from two contemporary IOSs, with the aim of gaining an insight in the clinical applicability of the various methods. The IOSs used were Primescan and Omnicam (Dentsply Sirona) and the methods compared were A) the unsigned

mean distance deviation, B) the linear distance between (virtual) key features on the dental arch, as described above [and referred to as *Inter-molar width* from here-on**]**, and C) the percentage surface area of a scan deviating beyond 0.1mm (a simplified version of the method reported in [6]).

Null hypotheses

i) That there is no significant difference between trueness, as measured using mean surface deviation compared to a reference scan, between the two test IOSs
ii) That there is no significant difference between trueness, as measured using linear cross-arch distance deviation compared to a reference scan, between the two test IOSs
iii) That there is no significant difference between trueness, as measured using the percentage of the surface deviating beyond 0.1mm compared to a reference scan, between the two test IOSs

3.3. Materials and methods

A dentate type IV stone maxillary model was scanned ten times with a Primescan, CEREC 5.0.0 (Dentsply Sirona) [P1 -P10], and ten times with an Omnicam, CEREC 4.6 (Dentsply Sirona) intraoral scanner [O1 -O10] , using the manufacturer's recommended scanning strategies. The scanned model had been poured more than 30 days prior to scanning.

All scans were recorded in one session by an experienced operator. Both scanners had been calibrated prior to scanning. All scans were exported as high-resolution STL (stereolithography) files.

To produce an indication of a trueness metric, the model was scanned once using a verified [VDI 2634/2] lab scanner (Rexcan DS2, Solutionix) which has a quoted resolution of <10µm when measured against the industry standard. All Primescan (P1–P10) and Omnicam (O1-O10) scans were aligned to the Solutionix scan. The alignment algorithm used was iterative closest point implemented using the freely available Open3D software(Zhou *et al.*, 2018), following a subsampling of all scans to produce pointclouds with a point distance no greater than 25 microns. Once aligned, the scans were reverted to their original point spacing. All

meshes were cropped identically using a cropping lasso [using Contour Select, LDD] defined on P1, and applied to all 20 meshes. This ensured that all future measurements would be taken from identical regions across all scans (Fig. 1).

Topologically identical key points were identified on the upper right second molar [UR7] on the Solutionix scan and all 20 IOS scans using the method outlined in (Gintaute *et al.*, 2020). Three more key points were similarly identified on UR3, UL3 and UL7.

Statistical analysis was done using SPSS Statistics 26 (IBM). Independent two-sample t-tests were used to assess the difference in trueness between the two intraoral scanners, across the three metrics investigated; mean deviation, Inter-molar width and proportion beyond 0.1mm.

3.4. Results

Surface comparison against Solutionix

The mean unsigned distance deviation between the verified Solutionix scan and the ten Primescan scans was 0.016(±0.006)mm. The mean signed standard deviation for the Primescan was 0.021(±0.009)mm .

The mean unsigned distance deviation between the verified Solutionix scan and the ten Omnicam scans was 0.116(±0.01)mm. The mean signed standard deviation for the Omnicam was 0.158(±0.025)mm.

There was a significant difference between the unsigned distance deviations produced by the two intraoral scanners $p<0.001$).

Inter-molar width

The arch width between the left and right molars was 55.44mm for the Solutionix scan. The arch width between the left and right molars was 55.439(±0.075)mm for Primescan. The mean perimeter distance of the Primescan was 152.38(±0.076)mm for the Primescan.

This same distance was 54.672(±0.065)mm for Omnicam. The perimeter distance of the single Solutionix scan was 152.40mm. The mean perimeter distance for the Omnicam was 151.29(±0.06)mm. See table i.

There was a significant difference between the arch widths produced by the two intraoral scanners ($p<0.001$).

Upper bound deviation against Solutionix

The mean proportion of the Primescan scans deviating beyond 0.1mm when compared against the Solutionix was 0.7(±2.0)%.

The mean proportion of the Omnicam scans deviating beyond 0.1mm when compared against the Solutionix was 42.1(±2.5)%.

The difference between the two intraoral scanners in proportion of scan deviating beyond 0.1mm from the Solutionix was significant ($p<0.001$).

Table ii Mean difference in key point distance from Solutionix scan (mm)

	UR7 to UR3	**UR3 to UL3**	**UL3 to UL7**	**UL7 to UR7 (Inter-molar width)**
Primescan	-0.012 (±0.012)	-0.006 (±0.010)	-0.006 (±0.014)	-0.001 (±0.075)
Omnicam	-0.065 (± 0.018)	-0.199 (±0.017)	-0.085 (±0.012)	-0.768 (±0.065)

3.5. Discussion

This study investigated the full arch trueness of two intraoral scanners using three different methods of assessment. There was a significant difference in global mean unsigned deviation between the two scanners ($p<0.001$). There was a significant difference in inter-molar width recorded by the two scanners ($p<0.001$). There was a significant difference in proportion of scan deviating beyond 0.1mm from the Solutionix scan $p<0.001$). Thus, the null hypotheses must all be rejected.

The Primescan produced data significantly closer to the verified lab scanner (an indication of trueness) compared to Omnicam, both when assessing using mean unsigned deviation and linear cross arch distance. Omnicam consistently under-reported the linear intermolar width, in effect, narrowing the arch form.

The Primescan reported only fractional amounts of scan data deviating beyond 0.1mm (0.7(±2.0)%) from the Solutionix scan, whereas an average of 42.1(±2.5)% of each Omnicam scan deviated beyond this distance.

All three metrics indicated that the Primescan produced truer data than the Omnicam. However, unlike the mean distance metric, the key point method gave a better intuition as to the potential quality of fit of a cross-arch prosthesis. For example, the casual reader might consider the mean unsigned deviation error of the Omnicam (0.116 +/- 0.01mm) to be clinically tolerable, envisaging that the fit of a cross-arch framework would require only a small adjustment. Conversely, the cross-arch linear error metric revealed that Omnicam consistently under-estimated the intermolar width by a much larger 0.768 (± 0.065)mm. This degree of framework inaccuracy would require significant chairside adjustment, or more likely, remaking. By contrast, the Primescan linear cross-arch error averaged -0.001 (±0.075) mim, which could more confidently be assumed to produce a well-fitting full arch prosthesis. Hence the key point method appears to discriminate better between IOSs and their likely clinical potential.

Our key point method requires no physical placement of landmarks, making it simpler to implement than previous studies, which required fixed reference objects such as metal bars or spheres (Güth *et al.*, 2016; Kuhr *et al.*, 2016). Interestingly in the latter study, using spheres *in vivo* on 50 test subjects, an intermolar error of 0.828(±0.265)mm for the Omnicam (measurement D1_4 in their paper) was reported. This agrees well with our value of 0.768 (±0.065)mm, with the slightly poorer trueness in the Kuhr *et al.* study perhaps being due to a

combination of older Omnicam software and the fact their study was performed *in vivo.* This might also explain the lower precision in their study (as given by the standard deviation).

Our key point method could easily be employed to measure precision *in vivo*, which may be a more clinically informative metric than the commonly used mean surface deviation. However, the problem remains in employing our metric to measure trueness *in vivo,* in that we have no reference values for the key point separation distances. Given the evidence of numerous papers regarding good conventional impressions outperforming IOSs over full arches, it would seem appropriate to use physical silicone or polyether impressions as a 'gold standard' in future work, when attempting to assess IOS trueness over a full arch (Ender and Mehl, 2011; Ender et al., 2016; Kuhr et al., 2016; Ender et al., 2019).

Our third test metric, percentage of surface lying beyond 0.1mm, may also hold value as a broad comparison of IOS accuracy and the user may select a threshold value appropriate to their needs. Here, we report that 42.1 ± 2.5%, of the Omnicam scan surface lay beyond 0.1mm of the true value. That almost half the entire scan is poorer than 0.1mm might allow a clinician to make an informed choice on appropriate use. Conversely, Primescan (0.7 ± 2.0%) revealed a strong improvement in trueness as judged by this metric.

It is interesting to note that the noise in both scanners, as measured via the standard deviations across scans, did not differ significantly.

An inherent challenge in accuracy validation of intraoral scanners *in vivo* is the lack of a measurable reference. Hence, while *in vivo* scans can be used to measure precision and repeatability, trueness validation of *in vivo* scans can be challenging. As a result, a large number of intraoral scanner accuracy studies rely on *in vitro* studies. A number of previous studies demonstrate that intraoral scanning reduces scan accuracy, due to movement restrictions and the optically challenging environment within the oral cavity (Flügge et al., 2013; Kim et al., 2015; Keeling et al., 2017). Results obtained *in vitro* can therefore be assumed to be an optimal scenario and likely to produce artificially favourable conclusions.

Visual comparison between the Omnicam and Primescan scans made it evident that the Primescan data had undergone significant edge sharpening (Fan and Jin, 2014), resulting in artificially sharp margins and severe mesh artifacts. One such artifact, a tunnel burrowing half-way across the distal aspect of an anterior tooth would potentially have interfered with any CAD design, had the artifact occurred on a prepared tooth. There seems to be a commercial drive to make IOS scans appear better using digital enhancements(Kim *et al.*, 2020).These

algorithms are potentially risky, because clinicians will see a sharp looking scan, but it may no longer actually represent the patient. Our metrics will not inform with regards to these, and so further work is required to assess the local effects (for example on crown margin trueness) of edge enhancement and interproximal sharpening.

3.6. Conclusion

We present a comparison of three methods for assessing the quality of 3D data produced by two IOSs. The *virtual keypoint*, and *percentage of scan deviating beyond 0.1mm* methods may both give a clearer insight into clinical scanner trueness than the commonly reported unsigned mean surface deviation. Due to the virtual method of keypoint creation, the method can be used on scan data obtained both *in vitro* and *in vivo.*

Primescan produced significantly truer results than Omnicam, under all three metrics. Its clinical use over full arches would appear to be more appropriate than Omnicam.

However, the Primescan was found to perform notable edge-sharpening, to the point of data deterioration; the clinical effect of this aspect of data manipulation should be investigated further.

3.7. Figures

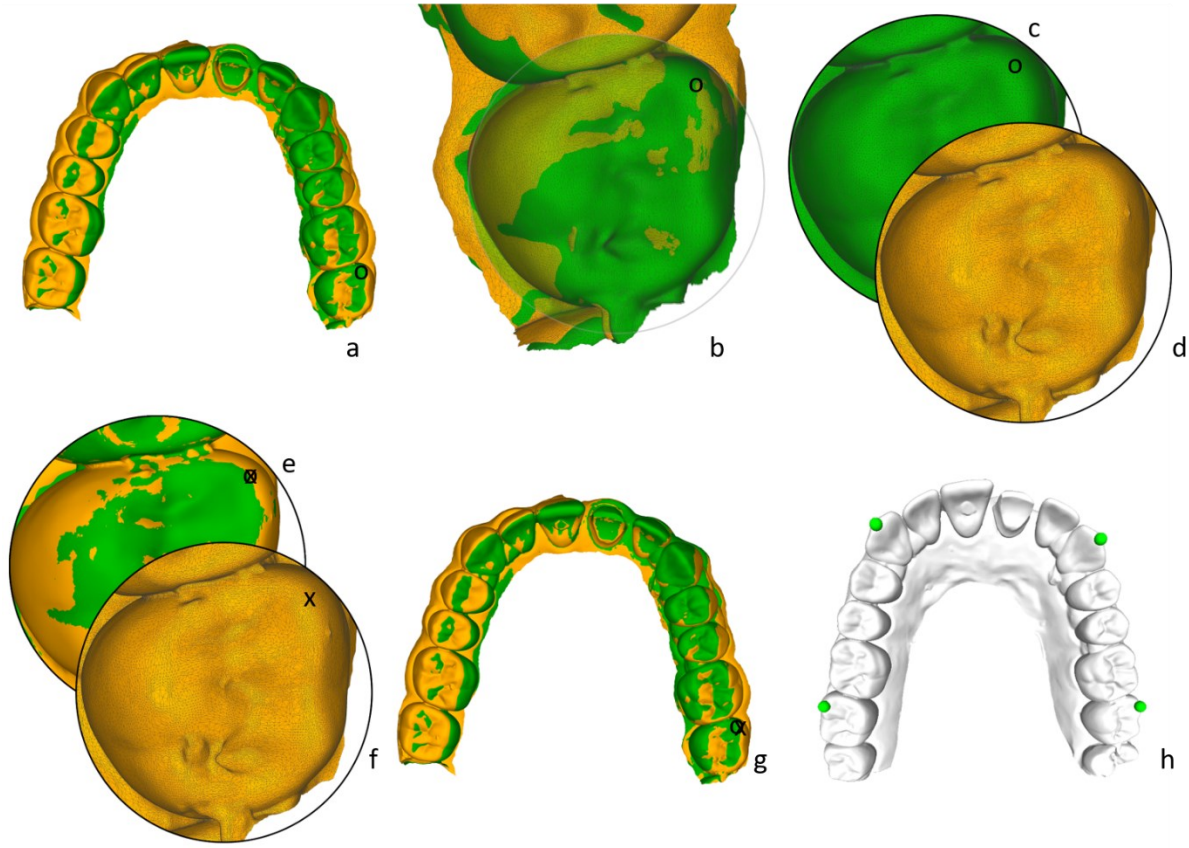


Figure 3:1 Identifying anatomically identical keypoints on different scans to enable direct inter-arch distance comparisons. a) Following global alignment of a pair of scans, a keypoint, UL7, has been identified on one scan (labelled "o") and we wish to identify this virtual keypoint on all subsequent upper arch scans from both IOSs. The arch width discrepancy in full arch IOS scans leads to poor alignments if performed over the full arch. In step b) we take a small section (10mm radius) of the source arch, and a similar selection from the test arch (c, d) and finely align these two (e). The precise location of the UL7 keypoint is then mapped onto the test arch (e, f), (labelled "x"), and an inverse transformation is then applied to carry the keypoint back to the untransformed test arch (g). This process was repeated for all four keypoints across the arch (h) on all scans."

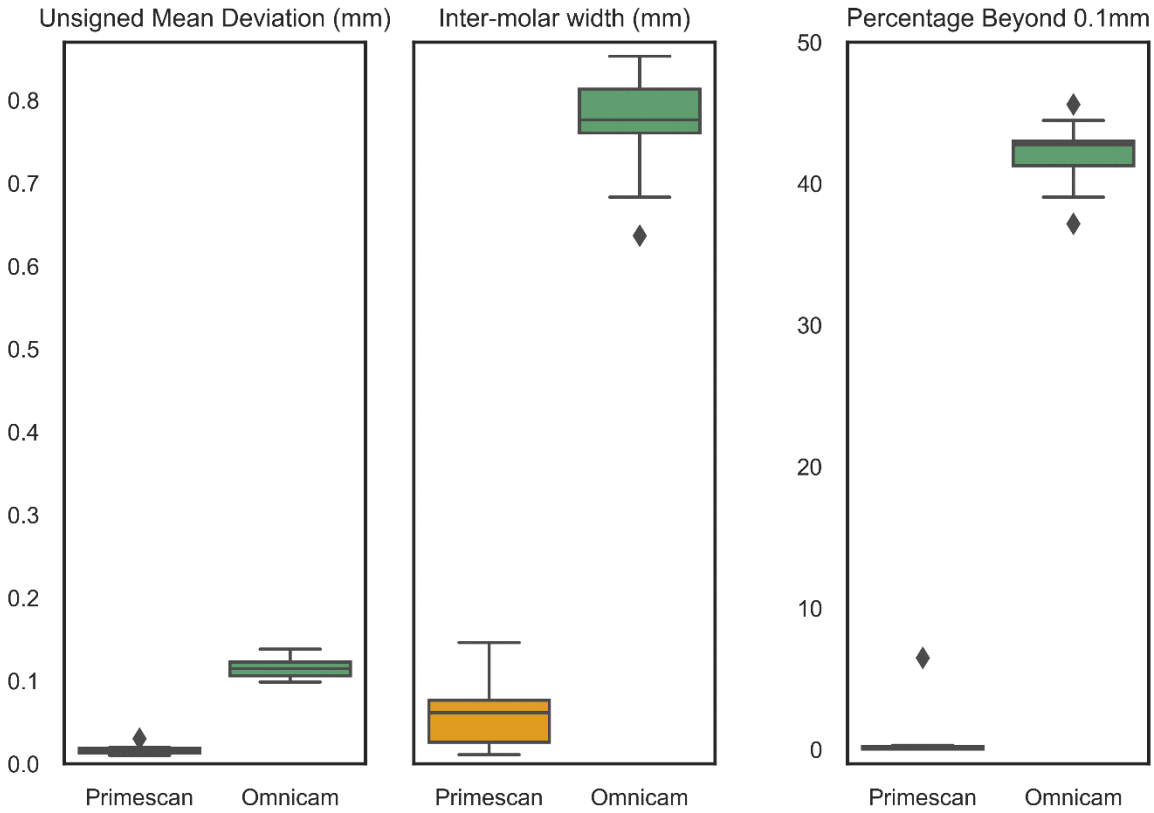


Figure 3:2 Three methods with which to measure scan accuracy. The line indicates median value, the box upper and lower quartile, while the whiskers show overall distribution. Outliers are indicated with a diamond.

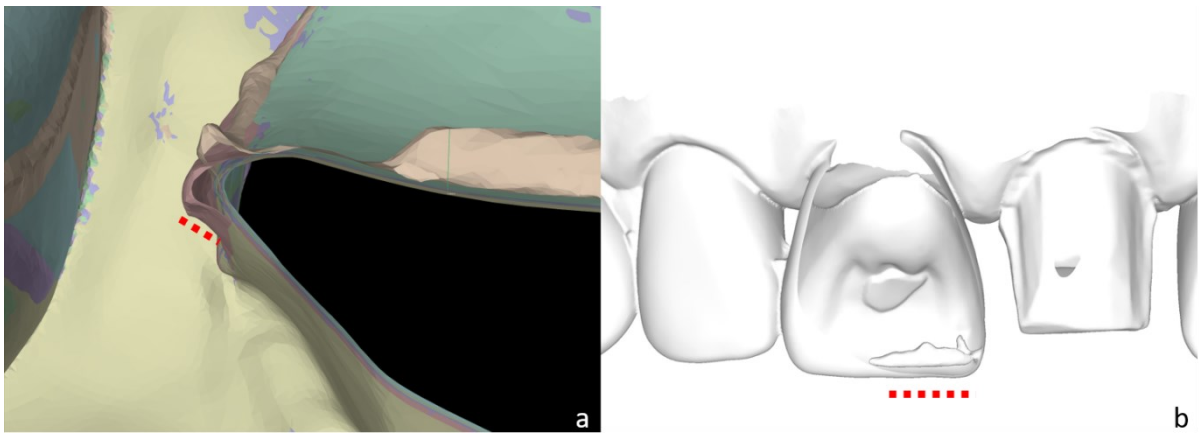


Figure 3:3. Screencaptures displaying edge sharpening artefacts occurring on the Primescan scans. a) shows an occlusal view looking down on the prep and edge sharpening occurring on the distal aspect of the tooth preparation on two of the five scans displayed. The red dotted line is ≈ 0.2mm. This error would lead to a much larger cement thickness than desired in any crown produced, but the error would be difficult to detect. b) shows a cross-section of a Primescan scan, with a tunnel burrowing ≈ 4.4mm (red dotted line) across the mesial aspect of an anterior tooth (UR1). This would potentially have interfered with any CAD design, had the artefact occurred on a prepared tooth.

3.8. List of references

1. Ahlholm, P., Sipila, K., Vallittu, P., Jakonen, M., Kotiranta, U., Sipilä, K., Vallittu, P., Jakonen, M., Kotiranta, U., Sipila, K., Vallittu, P., Jakonen, M. and Kotiranta, U. 2018. Digital Versus Conventional Impressions in Fixed Prosthodontics: A Review. Journal of Prosthodontics. 27(1), pp.1–7.

2. Chochlidakis, K.M., Papaspyridakos, P., Geminiani, A., Chen, C.J., Feng, I.J. and Ercoli, C. 2016. Digital versus conventional impressions for fixed prosthodontics: A systematic review and meta-analysis. Journal of Prosthetic Dentistry. 116(2), pp.184–190.

3. Ender, A., Attin, T. and Mehl, A. 2016. In vivo precision of conventional and digital methods of obtaining complete-arch dental impressions. Journal of Prosthetic Dentistry. 115(3), pp.313–320.

4. Ender, A. and Mehl, A. 2011. Full arch scans : Conventional versus digital impressions - An in-vitro study. International Journal of Computerized Dentistry. 14(21).
5. Ender, A., Zimmermann, M. and Mehl, A. 2019. Accuracy of complete- and partial-arch impressions of actual intraoral scanning systems in vitro. International journal of computerized dentistry. 22(1), pp.11–19.

6. Fan, R. and Jin, X. 2014. Controllable Edge Feature Sharpening for Dental applications. Computational and Mathematical Methods in Medicine. (2014).

7. Flügge, T. V., Schlager, S., Nelson, K., Nahles, S. and Metzger, M.C. 2013. Precision of intraoral digital dental impressions with iTero and extraoral digitization with the iTero and a model scanner. American Journal of Orthodontics and Dentofacial Orthopedics. 144(3), pp.471–478.

8. Gintaute, A., Keeling, A.J., Osnes, C.A., Zitzmann, N.U., Ferrari, M. and Joda, T. 2020. Precision of maxillo-mandibular registration with intraoral scanners in vitro. Journal of Prosthodontic Research. 64(2), pp.114–119.

9. Güth, J.F., Edelhoff, D., Schweiger, J. and Keul, C. 2016. A new method for the evaluation of the accuracy of full-arch digital impressions in vitro. Clinical Oral Investigations. 20(7), pp.1487–1494.

10. Kan, J.Y., Rungcharassaeng, K., Bohsali, K., Goodacre, C.J. and Lang, B.R. 1999. Clinical methods for evaluating implant framework fit. The Journal of prosthetic dentistry. 81(1), pp.7–13.

11. Keeling, A., Wu, J. and Ferrari, M. 2017. Confounding factors affecting the marginal quality of an intra-oral scan. Journal of Dentistry. 59, pp.33–40.

12. Kim, J., Kim, K., Kim, S., Kim, W., Kim, H. and Kim, J. 2015. Quantitative evaluation of common errors in digital impression obtained by using an LED blue light in-office CAD/CAM system. Quientessenz International. 46(5).

13. Kim, T., Cho, Y., Kim, D., Chang, M. and Kim, Y. 2020. Tooth Segmentation of 3D Scan Data Using Generative Adversarial Networks. *Applied Sciences*. **10**(490), pp.1–17.

14. Kühle, A. 2003. Calibration Procedures for Atomic Force Microscopes. Advanced Techniques for Assessment Surface Topography: Development of a Basis for 3D Surface Texture Standards 'Surfstand'., pp.175–194.

15. Kuhr, F., Schmidt, A., Rehmann, P. and Wöstmann, B. 2016. A new method for assessing the accuracy of full arch impressions in patients. Journal of Dentistry. 55, pp.68–74.

16. O'Toole, S., Osnes, C., Bartlett, D. and Keeling, A. 2019. Investigation into the accuracy and measurement methods of sequential 3D dental scan alignment. Dental Materials. 35(2019), pp.8–13.

17. Osnes, C., Wu, J., Venezia, P., Ferrari, M. and Keeling, A. 2019. Full arch precision of six intraoral scanners in vitro. Journal of Prosthodontic Research. 64(64), pp.6–11.

18. Tsirogiannis, P., Reissmann, D.R. and Heydecke, G. 2016. Evaluation of the marginal fit of single-unit , complete-coverage ceramic restorations fabricated after digital and conventional impressions : A systematic review and meta-analysis. The Journal of Prosthetic Dentistry. 116(3), pp.328-335.e2.

19. Wismeijer, Daniel; Buser, Daniel; Chen, S. 2019. Biological and Hardware Complications in Implant Dentistry [Online]. Quintessence Publishing. Available from: https://play.google.com/store/books/details?id=5iu6DwAAQBAJ].

20. Zhou, Q.-Y., Park, J. and Koltun, V. 2018. Open3D: A Modern Library for 3D Data Processing.

Investigating three methods of assessing the clinically relevant trueness of two intraoral scanners

1.

4. Sources of error in maximum intercuspation from complete dentate full arch intraoral scans *in vitro*

C.A. Osnes, J. Wu, M Ferrari, T. Joda, A.J. Keeling

International Journal of Computerized Dentistry

Accepted November 2020

4.1. Abstract

Aim: Recording maximum intercuspation (ICP) is critical for many dental procedures. Digital ICP from intraoral scanners (IOS) produces variable results. This study investigated the sources of error in recording ICP using an IOS and a recently reported method.

Materials and methods: A set of dentate models was scanned three times in a Rexcan DS2 scanner. The models were then scanned six times with CEREC Omnicam. For each scan ten bilateral 'bite' scans were performed (n=6x10 bite registrations). Three key-points were identified on the first IOS scan, and automatically transplanted onto all subsequent scans. The key-point method was validated by using a "secondary" key-point transplantation from each scan back to the three laboratory scans, where the location of each point was compared by one-way ANOVA. IOS full arch error was identified by comparing the intermolar key-point distance on all IOS scans against the gold standard model scans. Precision of the virtual occlusion was identified by comparing the distance between all upper-lower key-point pairs for all IOS scans, using intra-class correlation.

Results: Automatic key-points were transplanted to model scans with standard deviations in location of ≤0.003mm [upper] and ≤0.004mm [lower] arch. IOS intermolar width had a mean error of 0.183(±0.061)mm [upper arch] and 0.017(±0.092)mm [lower arch]. Inter-occlusal key-point separation showed poor reliability across groups, but good precision (s.d.<0.022mm) within groups.

Conclusion: Automatic key-points allowed valid linear distance comparisons across repeated scans. Poor trueness and precision in full arch intraoral scans adversely affected inter-occlusal registrations. Bite scan precision had a less detrimental effect on inter-occlusal registration.

Keywords

Occlusion, Articulation, Intraoral Scanner, Accuracy, precision, measurement

4.2. Introduction

The ability to accurately record maximum intercuspation [ICP] is critical for many aspects of treatment planning and prosthodontic rehabilitation(Squier, 2004). The preferred analogue method does not use any bite registration material, but this relies on the presence of a clear and stable ICP(Walls et al., 1991). If patients do not have a stable ICP, perhaps due to loss of teeth or tooth wear, a bite recording material is required, but the error in ICP increases(Walls et al., 1991; Eriksson et al., 2002).

Digital ICP registrations from intraoral scanners [IOS] have shown promise in quadrant dentistry, with one study reporting a reproducibility of 0.042(±0.034) mm from the IOS compared to 0.135(±0.077) mm from analogue(Jaschouz and Mehl, 2014). However, early work on full arch IOS occlusions revealed large cross-arch errors in the region furthest from the unilateral buccal bite scan(Iwaki et al., 2013). This led to the recommendation of using bilateral (or more) buccal bite scans to optimise the full arch digital ICP(Solaberrieta, Arias, et al., 2016; Solaberrieta, Garmendia, et al., 2016), a feature which is now common in modern IOSs. Despite this enhanced protocol, current evidence still suggests there is a large variation between the full arch ICPs produced by a range of contemporary IOSs(Park et al., 2018; Gintaute et al., 2020; Mangano et al., 2020).

The sources of error in ICP from an IOS may originate from two aspects: the quality of the buccal bite scan or the quality of the arch/quadrant scans. The reproducibility of unilateral buccal bite scans has been investigated for use with quadrant scans(Ueda et al., 2014). This work showed the buccal bite scan to have a larger influence in quadrant ICP precision than the actual quadrant arch scans. Conversely, it is known that the precision of full arch IOS scans is variable and poorer than conventional impressions(Rhee et al., 2015; Chochlidakis et al., 2016; Ender et al., 2016; Kuhr et al., 2016; Tsirogiannis et al., 2016; Vögtlin et al., 2016; Treesh et al., 2018; Ahlholm et al., 2018; Ender et al., 2019).

Thus, in the case of full arch IOS scans, it is unclear whether the reported errors in ICP are caused primarily by variability in the buccal bite scans, or by variation in the individual full arch scans. Knowledge of the source of error would be a first step in improving the accuracy of full arch digital ICP records, and in suggesting optimal clinical technique. For example, if the bite scans were found to be more variable, it could be recommended to take multiple bite scans and have the computer calculate the mean occlusion.

Further complication is introduced by the challenge of robustly assessing scan data, and correctly measuring the same point across several scans, regardless of whether we are

measuring arch distortion or error in ICP. Investigations such as (Kuhr et al., 2016; Güth et al., 2016) used metal spheres, or a metal bar spanning the width of the arch, in the case of the second study, to introduce reference points from which to take measurements when measuring the arch width. An alternative to this may be to use virtually identified topological key points, as previously presented in (Gintaute et al., 2020).

The aim of this study was to investigate the sources of error in recording maximum intercuspation using an IOS. Using a recently reported novel method to robustly identify topological key points in scans, the relative contributions of buccal bite scan precision, and full arch trueness and precision were investigated.

This study was instigated based on the following three null hypotheses. Firstly, that there will be no difference in the cross arch inter-molar width created from matching key points within multiple IOS arch scans. Secondly, that there will be no difference in the interocclusal separation of matching key points across a pair of full arch IOS scans when the occlusion is created by multiple different bilateral buccal bite scans. Thirdly, that there will be no difference in the interocclusal separation of matching key points recorded from multiple full arch IOS scans.

4.3. Materials and method

A pair of dental stone models (type IV gypsum, John Winters, Halifax, UK) were poured from agar impressions of a fully dentate typodont set (Frasaco, Tettnang, Germany). Three steel spheres (diameter = 2mm) were cemented to the occlusal surfaces of 18 and 28, and the cingulum of 11 using epoxy resin cement. This had the effect of creating a stable three-point occlusion and the models were mounted on an average value articulator in this relationship.

The upper and lower models were scanned three times in a model scanner (Rexcan DS2, Solutionix, Seoul, Korea).

The upper and lower models were each scanned six times (labelled A-F) using a CEREC Omnicam (Dentsply Sirona, Charlotte, USA) with the full arch Ortho 1.1.2 software. Next, ten bilateral buccal bite scans were performed for each of the six pairs of arch scans (labelled 1-10), with the resulting occluded scans being exported after each bite calculation. All bite scans were recorded without touching or moving the articulated models. This produced 60 sets of occluded digital models, with each group of ten comprising identical STL files of the upper and lower arches, but with slightly different occlusal alignments.

The CEREC software maintains the position of the lower arch and moves only the upper arch, based on the buccal bite scans. Therefore, all 60 lower scans were aligned and the transformation matrix for each alignment was applied to the corresponding upper arch. This brought all 60 model-pairs into the same alignment, whilst preserving idiosyncratic differences in the individual occlusions.

Experiment 1: Validating the automatic key point method

Three key points were visually identified on the first upper scan (mesh A) in the regions of the metal spheres near the molars on both sides and one point anteriorly. Three closely related opposing key points were likewise identified in the lower scan. The molar key points allowed intermolar width measurements for both the upper and lower arches (Fig 1).

Next, corresponding key points on the remaining five upper and lower meshes (B-F) were automatically identified using the method presented in [9]. The ten identical meshes from group A all used the original, manually identified vertex key point throughout the experiment. All other groups used the manually identified key point from the first scan in group A (A1) as the donor to calculate the equivalent recipient key point.

The level of trueness of the automatic key point method was confirmed using the three scans from the model scanner (Rexcan DS2, Solutionix, Seoul, Korea) which has a stated accuracy of < 0.01 mm under VDI 2634/2. Using the key point method described in [9], the variation in inter-molar width on the three scans, as calculated using donor key points derived from each of the first IOS scans in the groups A-F, were compared (n=18 upper and n=18 lower) (Fig 2). Thus, if the key point generation method was accurate, we would expect little variation in inter-molar width using key points derived from different IOS groups (A-F) within each of the three model scans. In addition, if the Solutionix scans were accurate, we would expect little variation in inter-molar width across the three scans. Hence, in all cases, the outcome variable for assessing key point validation was the linear intermolar width. To test these, two separate one-way ANOVA tests with Bonferroni correction were performed. One fixed the Solutionix scanner groups (S1-3) while the other fixed the donor groups (A-F). Both had significance set at $p < 0.001$.

Experiment 2: Identifying full arch IOS error

Next, the trueness and precision of the IOS intermolar width across the six scans was assessed. The mean (sd) errors in IOS inter-molar width were reported, using the mean Solutionix inter-molar width as a gold standard.

Experiment 3: Identifying bite scan error

The magnitude of the 3-dimensional vector between each pair of upper-lower key points on each IOS sample was recorded. This yielded 60 results per key point, or 180 interocclusal measurements in total (one anterior, and two posterior sets of measurements). The three outcome measures were the linear distances between the three pairs of key points. Variations in inter-occlusal distances were assessed within the six groups (A-F) and across the six groups using intraclass correlation (ICC) to assess the relative influence of different buccal bite scans and different full arch scans on the precision of the occlusion.

4.4. Results

One buccal bite scan (A3) failed to export correctly so group A contained only 9 samples, while groups B to F contained 10 samples, each representing a unique buccal bite scan aligning the respective IOS models.

i) Experiment 1 – Key point validation

The inter-molar widths recorded on the three Solutionix scans, using donor key points from the first of each of the IOS groups are shown in Table 1. The largest standard deviation was 0.003 mm for the lower arch and 0.004 mm for the upper arch.

The one-way ANOVA, using the three separate Solutionix scans as factors, revealed no significant differences between the upper arch intermolar widths on the Solutionix scans ($p=0.611$). The lower arch showed a statistically significant variation between scans ($p<0.001$), but the overall range was extremely low at 0.011 mm (45.105 – 45.126 mm).

The one-way ANOVA, using the six key point donors (A-F) as factors, revealed significant differences between the upper arch intermolar widths on the Solutionix scans ($p<0.001$). However, the range within each group (A-F) never exceeded 4 microns across the three scans. The lower arch showed no statistically significant variation between scans ($p=0.926$).

ii) Experiment 2: Full arch accuracy

The variance of intermolar width within each IOS group was extremely low, never reaching 1 micron for any of the upper or lower scan groups indicating that the key points were valid within each the group.

The absolute error in IOS intermolar width (as compared to the Solutionix mean) is shown in Figure 3 below. The upper arch had a mean (sd) intermolar error of 0.183(±0.061) mm. The lower arch intermolar error was 0.017(±0.092) mm.

iii) Inter-occlusal analysis

The inter-occlusal distances for the right molar, left molar and anterior regions across the 6 IOS groups are shown in Figs 4, 5, 6 and in Table 2.

The intra-class correlation (ICC) average measures value for the right molar inter-occlusal distances was low at 0.099, implying poor correlation across the six groups. The ICC average measures for the left molar and anterior inter-occlusal distances were negative, implying poor reliability across groups (-2.738 and -0.912 respectively).

The standard deviation within any group never rose above 0.022mm, indicating relatively good precision of multiple bite scans applied to the same model.

4.5. Discussion

This experiment investigated the sources of error in full arch ICP occlusions generated by an IOS. It also confirmed a previously reported technique to ensure precise measurements from anatomically identical key points in meshes when the typical triangle edge length of the mesh is larger than the expected error under investigation.

Firstly, considering key point creation and validation: IOSs typically produce meshes with mean triangle edge lengths of >0.1mm. This does not imply that this is the limit of their trueness and precision, because the (flat) face of each triangle may pass close to the slight curvature of the real surface in that region. However, when identifying corresponding key points across repeated scans, the manual selection of the ‘closest’ vertex on the target mesh to a key point vertex on the source mesh is very unlikely to be the closest anatomical match. Our automatic

method attempted to identify the best anatomical match, regardless of whether it lay at a vertex, or on a face (or edge) of a triangle.

Our key point method was able to 'donate' key points to the model scan (Solutionix) with standard deviations of only 0.004 mm within a single scan, and a standard deviation of only 0.007 mm in measured inter-molar width across multiple scans. The key points on the IOS scans were even more consistent, even when donated onto slightly displaced models (the 10 separate buccal bites in each group). This is probably because the IOS scans received 'first-hand' key point donations (all from A1), whereas the Solutionix received 'second-hand' donations (eg A1->B5->Solutionix).

It would seem this method precisely transplants anatomically equivalent points between meshes and provides a 'markerless' way of comparing similar scans. This may be important for future *in vivo* work, negating the need to place metal spheres or similar reference objects in the mouth, and thus simplifying study design.

The first experiment investigated the trueness and precision of the intermolar width: A clinically significant intermolar error was noted in the upper arch, with the Omnicam overestimating the intermolar width by a mean of 0.183(±0.061) mm. The mean lower arch intermolar error was much smaller at 0.017 mm but the precision was still clinically relevant at 0.093 mm standard deviation. Thus, the null hypothesis stating that there would be no difference in the cross arch inter-molar width created from matching key points within multiple IOS arch scans is rejected empirically, although we note that further statistical analysis is not possible.

It is not clear why the upper arch width was consistently overestimated while the lower arch width was not. This may be due to different habitual wand paths used when scanning the articulated upper and lower models, although with the CEREC Ortho software, a guided path is employed. Regardless, even the truer lower arch showed large standard deviations, implying any single scan could differ from the true value by a clinically noticeable amount.

The second experiment explored interocclusal precision and the influence of different factors: The true interocclusal distances between the three pairs of key points was not known in this experimental set up. Therefore, it is only possible to assess the precision (variation) in key point separation, and the possible sources of this variation. It should also be remembered that the linear distance of the key point separation is a crude measure, since we are not isolating

vertical and lateral errors and indeed the two may cancel each other out, causing a potential underestimate of the true clinical error.

Considering the Influence of the buccal bite scans on ICP precision*:* The precision, as assessed by standard deviation of separation distance, within each group and for each key point, was very good, never rising above 0.022 mm. It would therefore seem that multiple buccal bite scans of the same model tend to produce consistently similar occlusions around the arch. Previous studies showed similar small variations caused by multiple optical bite registrations[10] in dentate arches. Our second null hypothesis stated that there would be no difference in the interocclusal separation of matching key points across a pair of full arch IOS scans when the occlusion is created by multiple different bilateral buccal bite scans. Our findings confirm this hypothesis.

The contrast with analogue dentistry is noteworthy here. Often, a physical bite registration (or inter-occlusal record) is subject to variability, leading to clinical recommendations of using a 'multiple check bite' procedure. Here the clinician records several bite registrations, and the technician confirms them all for agreement during articulation of the models. This is time-consuming and labour-intensive. However, with digital bite registrations, the idea of performing multiple registrations becomes more viable because of the speed and ease of data capture, and the subsequent ease of digital model articulation. Unfortunately, our study shows that, while easy to achieve, the use of multiple digital check bites might yield little or no clinical benefit because the digital buccal bite scans are very precise already.

Considering the Influence of the full arch scans on ICP precision: The low intra-class correlation for the right molar groups (0.099), and the negative values for the other two groups point to poor reliability and poor correlation across the groups. This may mean there is a larger influence on occlusion caused by the quality of the full arch scan, rather than the buccal bite scan. For example, the overall values of inter-occlusal separations for the left molar ranged from 0.02 mm (B1) to 0.188mm (F9). In the presence of full arch inaccuracies, the ability of the computer to align the bilateral buccal bites consistently will be impaired (the arch distortions will mean that there is no true mathematical solution to the bite alignment, and that the two buccal bite scans will be 'competing' during the alignment). Our third null hypothesis, that there would be no difference in the interocclusal separation of matching key points recorded from multiple full arch IOS scans, must therefore be rejected.

These results suggest that the largest factor influencing full arch occlusal precision from the tested IOS was a low trueness and precision for capturing full arch data, while the buccal bite scans produced consistent occlusions. An improvement in full arch IOS scanning would likely bring the concomitant benefits of better occlusal registration, and better fitting large-span prostheses. For maximum clinical benefit, manufacturers efforts should be focussed towards improvements in this aspect of their IOS.

It should be noted that our results are limited to one IOS – others may show different patterns of trueness and precision. Despite this shortcoming, the Omnicam used here is a popular IOS, and generally performs comparably well amongst its peers in the literature. We can therefore consider the results presented here as a reasonable bellwether for IOSs in general.

It should also be noted that these results were captured *in vitro*, and that, for any IOS*, in vivo* results will likely be poorer. The clinical observation that patients may tend to bite down more firmly on the side where the scanner wand is recording the bite (and have a slight separation on the contralateral side) may add further difficulties in the *in vivo* situation. Similarly, the fact that the investigation was undertaken using typodont dentition, with an ideal occlusion, may mean that the conclusions reached are not directly comparable to in vivo studies. It would be interesting to investigate whether anatomical variation in occlusion would affect the IOS's virtual articulation.

Further, investigating the quality of articulation, be it analogue or digital, is a challenge; as made evident by the literature (Solaberrieta, Arias, et al., 2016; Alghazzawi, 2016). While the Solutionx scans could be used as a "gold standard" for the arch width, and thus give an indication of the level of trueness produced by the Omnicam, there was no gold standard for the occlusion. One issue with the virtual key point method is that it uses virtual landmarks based on vertex location. Thus, it would be impossible to identify a feature on the physical model and use this exact location for the key point measurements. Therefore, any attempt at measuring the true interocclusal distance would be likely to produce different results from our virtual measurements; not because one method is superior at measuring distances, but because it is nearby impossible to measure a distance between the exact same two points, first on the physical model, and then on its virtual equivalent.

4.6. Conclusions

An automatic method for identifying anatomically identical key points on meshes enables accurate measurements when analysing IOS scans. A lack of trueness and precision in full arch scans produced by the CEREC Omnicam adversely affected inter-occlusal registrations. The bilateral buccal bite scans do not adversely influence the occlusal errors to the same degree.

Clinical relevance: The fundamental cause of the full arch ICP variation reported in the literature appears to come from errors in the full arch scans themselves. Thus, a multiple check bite scan strategy (analogous to the gold standard analogue method) cannot be recommended, since the bite scan does not seem to be at fault. Improvements in full arch scan accuracy should yield improvements in ICP full arch registrations.

4.7. Figures

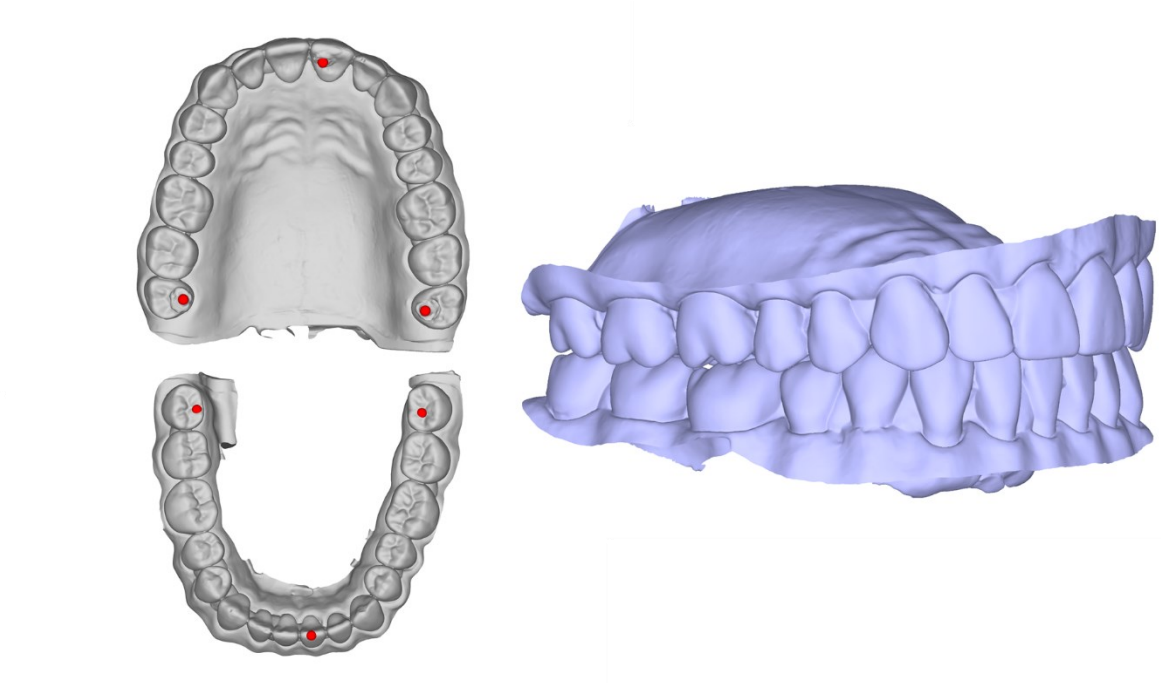


Figure 4:1: The occluded model: Key point locations on UR8, UL8, UL1, LR8, LL8 and LL1 shown in red (left), and an orthographic view of the first pair of scans, A1, in occlusion (right).

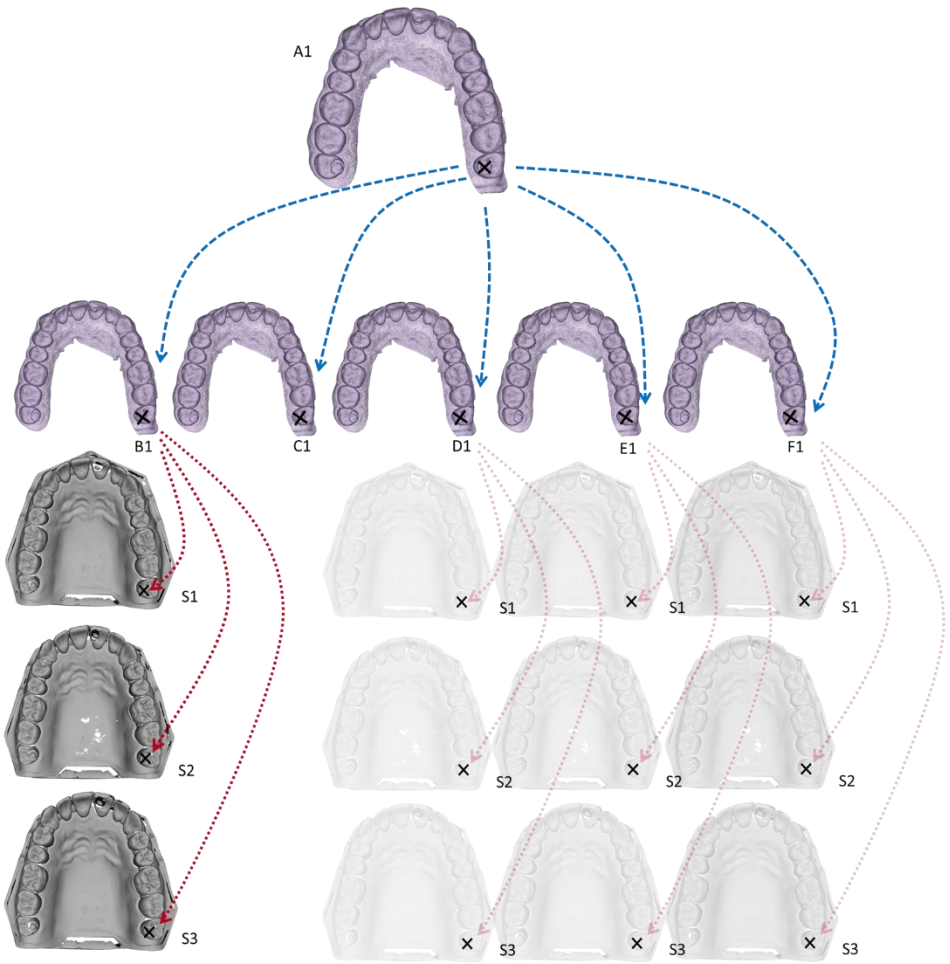


Figure 4:2: Key point validation method used in experiment 1: A key point was manually identified on scan A1. This was automatically transplanted onto all scans in groups B-F, (with all 10 scans within each group being identical meshes). To confirm the accuracy of the key point location, the new key points on all IOS scans (A-F) were then transplanted onto each of the three Solutionix scans (S1-S3) and their location reported. This was repeated for each of the three points in the upper arch and the three points in the lower arch.

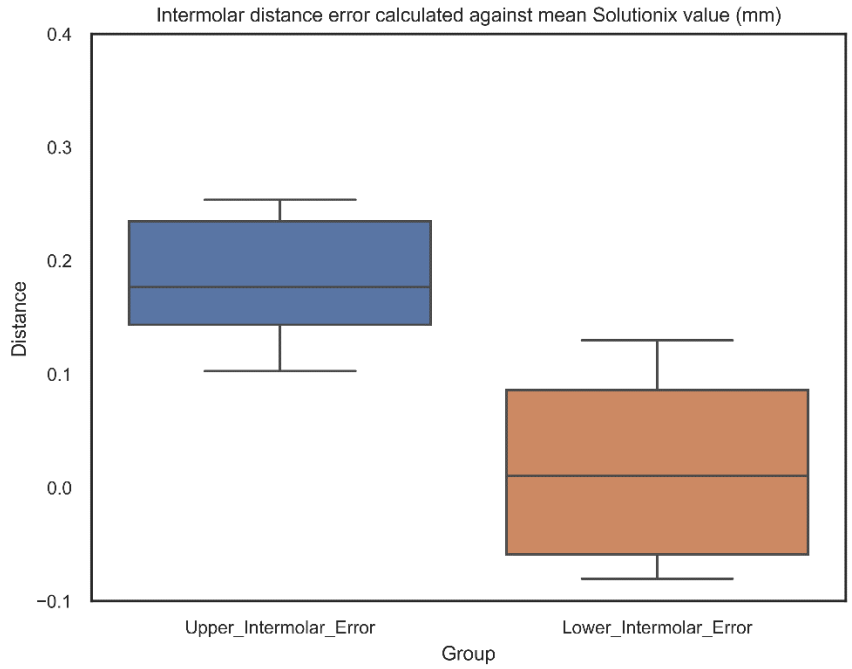


Figure 4:3: Intermolar distance error over the six IOS scans, as calculated against the mean gold standard Solutionix value. The line indicates median value, the box upper and lower quartile, while the whiskers show overall distribution.

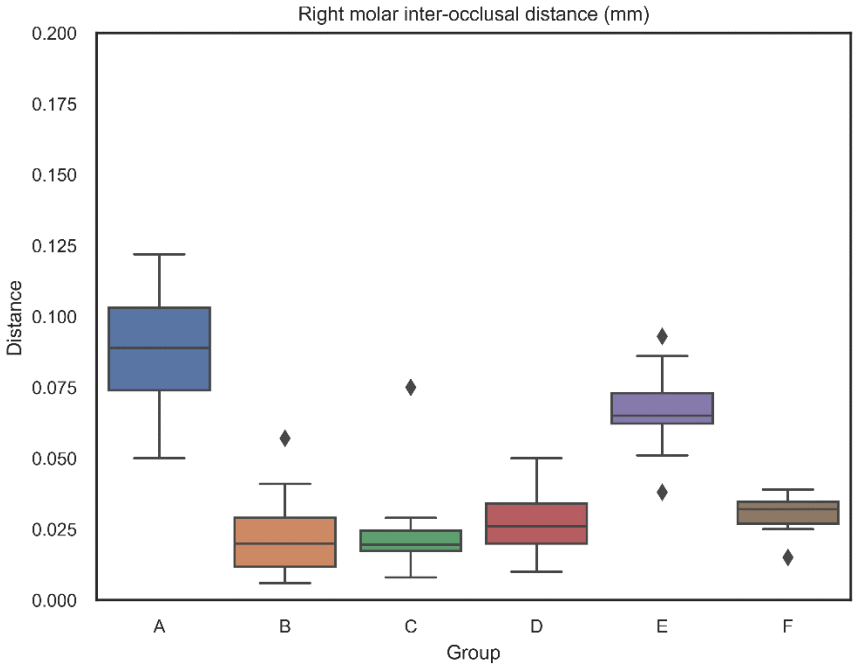


Figure 4:4: Right molar inter-occlusal distances across the six scan groups. The line indicates median value, the box upper and lower quartile, while the whiskers show overall distribution. Outliers are indicated with a diamond.

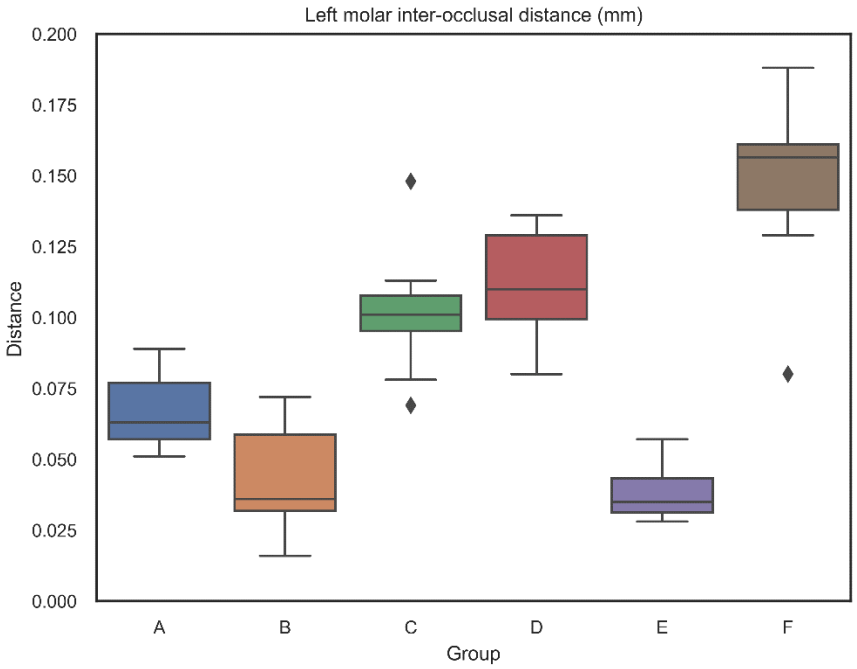


Figure 4:5: Left molar inter-occlusal distances across the six scan groups. The line indicates median value, the box upper and lower quartile, while the whiskers show overall distribution. Outliers are indicated with a diamond.

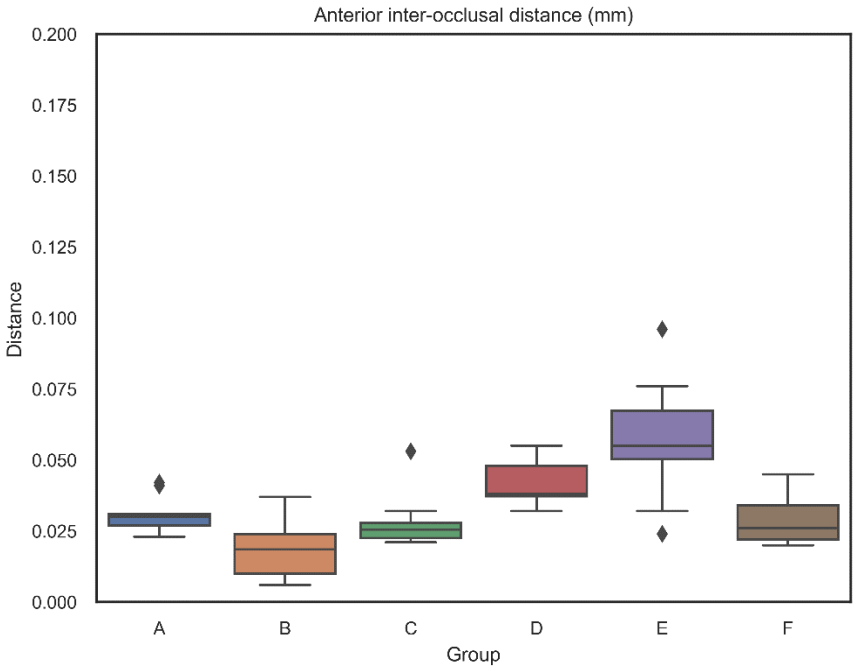


Figure 4:6: Anterior inter-occlusal distances across the six scan groups. The line indicates median value, the box upper and lower quartile, while the whiskers show overall distribution. Outliers are indicated with a diamond.

Table 3. Descriptive statistics for inter-molar widths recorded on three upper and three lower Solutionix scans using donor key points from the first of each IOS group, giving an indication of the trueness of the key point method.

Descriptive Statistics						
	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Solutionix Lower 1	6	45.105	45.113	45.10700	.002967	.000
Solutionix Lower 2	6	45.117	45.123	45.11829	.002186	.000
Solutionix Lower 3	6	45.121	45.128	45.12324	.002231	.000
Solutionix Upper 1	6	46.369	46.379	46.37420	.004307	.000
Solutionix Upper 2	6	46.370	46.379	46.37528	.003679	.000
Solutionix Upper 3	6	46.369	46.377	46.37312	.002977	.000

Table 4 Key point separation between UR8 and LR8; UL8 and LL8; and UL1 and LL1 (mm) (standard deviation)

Key point separation (Mean(±standard deviation) (mm)			
	UR8 - LR8	UL8 - LL8	UL1 - LL1
A	0.088(±0.022)	0.067(±0.013)	0.031(±0.007)
B	0.024(±0.016)	0.45(±0.018)	0.020(±0.010)
C	0.026(±0.019)	0.100(±0.022)	0.029(±0.010)
D	0.025(±0.010)	0.109(±0.018)	0.040(±0.008)
E	0.066(±0.017)	0.039(±0.009)	0.058(±0.022)
F	0.032(±0.005)	0.151(±0.030)	0.028(±0.009)

4.8. List of references

1. Ahlholm, P., Sipila, K., Vallittu, P., Jakonen, M., Kotiranta, U., Sipilä, K., Vallittu, P., Jakonen, M., Kotiranta, U., Sipila, K., Vallittu, P., Jakonen, M. and Kotiranta, U. 2018. Digital Versus Conventional Impressions in Fixed Prosthodontics: A Review. *Journal of Prosthodontics*. **27**(1), pp.1–7.

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

3. Chochlidakis, K.M., Papaspyridakos, P., Geminiani, A., Chen, C.J., Feng, I.J. and Ercoli, C. 2016. Digital versus conventional impressions for fixed prosthodontics: A systematic review and meta-analysis. *Journal of Prosthetic Dentistry*. **116**(2), pp.184–190.

4. Ender, A., Attin, T. and Mehl, A. 2016. In vivo precision of conventional and digital methods of obtaining complete-arch dental impressions. *Journal of Prosthetic Dentistry*. **115**(3), pp.313–320.

5. Ender, A., Zimmermann, M. and Mehl, A. 2019. Accuracy of complete and partial arch impressions of actual intraoral scanning systems in vitro. *International journal of computerized dentistry*. **22**(1), pp.11–19.

6. Eriksson, A., Ockert-Eriksson, G., Lockowandt, P. and Eriksson, O. 2002. Clinical factors and clinical variation influencing the reproducibility of interocclusal recording methods. *British Dental Journal*. **192**(7), pp.395–400.

7. Gintaute, A., Keeling, A.J., Osnes, C.A., Zitzmann, N.U., Ferrari, M. and Joda, T. 2020. Precision of maxillo-mandibular registration with intraoral scanners in vitro. *Journal of Prosthodontic Research*. **64**(2), pp.114–119.

8. Güth, J.F., Edelhoff, D., Schweiger, J. and Keul, C. 2016. A new method for the evaluation of the accuracy of full-arch digital impressions in vitro. *Clinical Oral Investigations*. **20**(7), pp.1487–1494.

9. Iwaki, Y., Wakabayashi, N. and Igarashi, Y. 2013. Dimensional accuracy of optical bite registration in single and multiple unit restorations. *Operative Dentistry*. **38**(3), pp.309–315.

10. Jaschouz, S. and Mehl, A. 2014. Reproducibility of habitual intercuspation in vivo. *Journal of Dentistry*. **42**(2), pp.210–218.

11. Kuhr, F., Schmidt, A., Rehmann, P. and Wöstmann, B. 2016. A new method for assessing the accuracy of full arch impressions in patients. *Journal of Dentistry*. **55**, pp.68–74.

12. Mangano, F.G., Admakin, O., Bonacina, M., Lerner, H., Rutkunas, V. and Mangano, C. 2020. Trueness of 12 intraoral scanners in the full-arch implant impression: a comparative in vitro study. *BMC oral health*. **20**(1), p.263.

13. Park, J.-M., Jeon, J. and Heo, S.-J. 2018. Accuracy comparison of buccal bite scans by five intra-oral scanners. *Journal of Dental Rehabilitation and Applied Science*. **34**(1), pp.17–31.

14. Rhee, Y.K., Huh, Y.H., Cho, L.R. and Park, C.J. 2015. Comparison of intraoral scanning and conventional impression techniques using 3-Dimensional superimposition. *Journal of Advanced Prosthodontics*. **7**(6), pp.460–467.

15. Solaberrieta, E., Arias, A., Brizuela, A. and Garikano, 2016. Determining the requirements , section quantity , and dimension of the virtual occlusal record. *The Journal of Prosthetic Dentistry*. **115**(1), pp.52–56.

16. Solaberrieta, E., Garmendia, A., Brizuela, A., Otegi, J.R., Pradies, G. and Szentpétery, A. 2016. Intraoral Digital Impressions for Virtual Occlusal Records: Section Quantity and Dimensions. *BioMed Research International*. **2016**.

17. Squier, R.S. 2004. Jaw relation records for fixed prosthodontics. *Dental Clinics of North America*. **48**(2), pp.471–486.

18. Treesh, J.C., Liacouras, P.C., Taft, R.M., Brooks, D.I., Raiciulescu, S., Ellert, D.O., Grant, G.T. and Ye, L. 2018. Complete-arch accuracy of intraoral scanners. *The Journal of Prosthetic Dentistry*. **120**(3), pp.382–388.

19. Tsirogiannis, P., Reissmann, D.R. and Heydecke, G. 2016. Evaluation of the marginal fit of single-unit , complete-coverage ceramic restorations fabricated after digital and conventional impressions : A systematic review and meta-analysis. *The Journal of Prosthetic Dentistry*. **116**(3), pp.328-335.e2.

20. Ueda, Y., Okada, D., Ogura, R., Komada, W., Otake, S., Ikeda, M. and Miura, H. 2014. Evaluation of occlusal relationship reproducibility with CAD / CAM techniques. *Asian Pacific Journal of Dentistry*. **14**, pp.35–40.

21. Vögtlin, C., Schulz, G., Jäger, K. and Müller, B. 2016. Comparing the accuracy of master models based on digital intra-oral scanners with conventional plaster casts. *Physics in Medicine*. **1**, pp.20–26.

22. Walls, A.W., Wassell, R.W. and Steele, J.G. 1991. A comparison of two methods for locating the intercuspal position (ICP) whilst mounting casts on an articulator. *Journal of Oral Rehabilitation*. **18**(1), pp.43–48.

4.9.

5. Optimal Use of Physical Centric Relation Records for Digital Workflows

C.A Osnes, J Patel, A Wagstaff, M Ferrari, A.J Keeling

European Journal of Prosthodontics & Restorative Dentistry

Submitted December 2020, resubmitted following review January 2021

5.1. Abstract

Correctly articulated dental casts are essential for aspects of dental treatment. Articulation can be “traditional”, using a physical articulator, or “virtual”, using 3D scanning and only virtual articulation; many “digital” methods currently rely on traditional articulation prior to scanning. This study compared the precision in articulation performed traditionally and virtually - without the use of prior physical articulation.

Articulated dental casts and centric relation records were obtained. 12 pairs of records were recorded from the articulated casts. Virtual method: all records were scanned, unclamped, in a custom laboratory scanner. The casts were aligned to each scanned record in turn to create virtual articulations. Traditional method: each record was used to physically articulate the casts. Each articulation was recorded using an intraoral scanner. The mean inter-arch separation between three key-points on each cast were used to determine differences in occlusal separation in three anatomical directions, and the precision of the two methods.

For the traditional articulations, standard deviations in key-point distance never exceeded 0.102mm. The virtual equivalent was 0.059mm. Statistically significant differences ($p<0.05$) between all anteroposterior separation distances were found between the two methods, and in three of six lateral/vertical separations.

Virtual articulation was significantly more precise than traditional articulation.

5.2. Introduction

Centric relation records are an important facet of diagnosis and rehabilitative dental treatment, aiding the correct orientation of the upper and lower study casts in centric relation for planning occlusal reorganisation therapy. The centric relation records can be used to articulate dental casts traditionally: by using a physical articulator, or virtually: by 3D scanning the records and aligning the scanned casts to these. However, encouraged by scanner manufacturer protocols, the most common way for a dental laboratory to utilize traditional centric relation registrations in a digital workflow is to physically articulate the dental casts first, and then scan the physically articulated set-up.

Both traditional and virtual articulation methods see the dentist record the patient's centric relation, before the records, along with impressions or dental casts, are passed on to the technician to be articulated. The technician may then choose to articulate the case using a physical articulator, or use a range of digital methods by which to digitally reproduce the articulation from either the centric relation records directly [virtually articulation] or from the traditional articulator. Since clinically recording centric relation is time consuming and challenging, the clinician is unlikely to take multiple centric relation records of a patient, to ensure precision, despite this being the biggest factor in correctly reproducing the patients' occlusal relationship (Eriksson et al., 2002). Further, the process of physically articulating dental casts can be technique sensitive and at risk of poor reproducibility(Breeding and Dixon, 1992; Chai et al., 1994; Campos and Nathanson, 1999; Ockert-Eriksson et al., 2000; Baumann, 2009; Patel and Alani, 2015). Traditional articulation methods have been reported to have low precision, resulting in occlusal adjustments being necessary in most cases(Utz et al., 2002). Traditional methods are also time consuming for the dental technician, further reducing the likelihood of multiple articulations being done for the sake of best precision. As a result - that the method used to articulate the patient's dental casts is precise is essential.

In this paper we present and investigate the precision of a "hybrid" workflow, which uses the centric relation records, but omits the physical articulation, unlike most current digital workflows. This investigation therefore uses the term "traditional" when referring to models being articulated using a physical articulator, even if this is prior to scanning, and therefore part of a digital workflow. The term "virtual articulation" is reserved for scenarios where a physical articulator has not been used at any stage.

Few dentists are equipped to virtually record the patient's centric relation, due the high cost of intraoral 3D scanners, despite much of contemporary treatment being digital during the laboratory stage: CAD/CAM crowns being just one example. As a result, technicians commonly need to produce digitised articulations using centric relation records to be able to produce the requested dental work. This may involve scanning casts individually, and then mounted on the articulator, before aligning all scans(Alghazzawi, 2016). Scanning articulated casts is problematic, requiring rigid clamping, and either a large scan-space to fit the articulator, or more commonly, custom locking plates and complex mechanisms to fit the articulated casts into the dental scanner. This process introduces all errors produced in traditional articulation, as well as the risk of movement of the dental casts during scanning into the resulting digital articulation. (See (Lepidi et al., 2020) for an overview over the most common current methods of digital articulation.) An alternative technique involves scanning casts individually, and independently scanning the interocclusal record (Alghazzawi, 2016); the casts are then digitally aligned to the scan of the occlusal record. Conventionally, this technique requires that all records are stabilised and clamped into position, which introduces distortion (Rhee et al., 2015). By using a clamp-less, custom made scanner (Keeling and Osnes, 2019), the current study aimed to investigate the precision of virtual versus traditional articulation, assuming little, or no distortion to the occlusal records during scanning. If scanning the centric relation records in isolation, before virtually articulating the casts were shown to be true and precise, and since the majority of dentists do not have access to intraoral scanners, yet many laboratories have model scanners, it would be useful to harness a hybrid, 'semi-digital' workflow; where the clinician follows the conventional method of taking centric relation records, before the technician scans the records and virtually articulates the dental cast scans against these, without needing to use a physical articulator.

Traditionally, high levels of surface detail captured using centric relation records may preclude the records from seating fully on the dental casts (Walls et al., 1991). This has led to the clinical advice to trim the bite registrations significantly – a technique sensitive procedure. Virtual articulation, where the centric relation records are scanned in isolation, may offer a method of articulating dental casts more accurately. This is because the digital alignment of scanned centric relation records (without fixing the records onto the cast before scanning) differs from traditional articulation in that digital alignment may benefit from maximising the available information (i.e. having a larger interocclusal record that records the embrasures and tooth bulbosities). Traditional articulation tolerates these features poorly as the record will often fail to fully seat on casts if gross undercuts or narrow embrasures are incorporated into

the interocclusal record; these can be difficult to fully remove while retaining sufficient positive 'fit' of the record on the casts. Virtual articulation methods alleviate this issue, as alignment disregards any penetrating meshes. This also results in plaster pearls or other minor, but potentially clinically relevant, imperfections on the physical cast not negatively affecting the virtual articulation in the same way it might affect the traditional method.

While the virtual methods may forgo some of the error-inducing aspects of the traditional articulation methods as listed above, virtual methods rely on repeatable and accurate digital alignments. Previous work shows that obtaining and assessing clinically correct digital alignment is a challenge (Gkantidis et al., 2020). As such, both methods under investigation contain their own challenges which impact the precision of articulation.

This study investigated the precision of two methods of articulating casts using centric relation records: a) traditional articulation using vinyl polysiloxane bite registration material and b) virtual articulation using digital alignment of scanned, unclamped, centric relation records and the scanned dental casts. Unlike other digital methods which rely on physical articulation, and may include inherent errors introduced at this stage when digitised, we ensure that none of the error potentially introduced during the laboursome process of using a traditional articulator is accidentally included in the virtual workflow, by not using a physical articulator at any point. This method would not require any change to the clinicians' workflow. An intraoral scanner was used to record the traditional articulation to enable a comparison between the two methods.

Null hypothesis

i) There is no statistically significant difference in the *precision* of traditional articulation and virtual articulation methods using vinyl polysiloxane centric relation records, measured as the variation in arch separation at three locations
ii) There is no statistically significant difference in the mean arch separation, measured at three locations, between traditional articulation and virtual articulation methods using vinyl polysiloxane centric relation records

5.3. Method

Impression and centric relation records: A set of traditional upper and lower fully dentate impressions were taken using a dual phase polyvinylsiloxane (Affinis heavy and medium body, Coltene Group Ltd) in rigid stock trays. The impressions were cast in type III stone. The respective casts were verified clinically to assess for errors in critical areas involved in the occlusion. Centric relation records were taken using an anterior jig (Duralay, Reliance, IL, USA) according to standard clinical technique with an incisal separation of approximately 3mm, and using two posterior vinyl polysiloxane (Blumousse, Parkell Inc.) registrations; these extended from the first molar to the canine teeth bilaterally. The casts were articulated on a semi-adjustable articulator (Denar® Mark II, WhipMix, KY, USA) using average values for the position of the upper cast and the settings of the condylar housings. The centric relation record was used to mount the lower cast, and the articulator pin remained set at the height of these records to maintain the incisal separation. Next, 12 centric relation records were taken on the articulated casts, using vinyl polysiloxane. All records were assessed for drag and over-extension into undercuts. The records were all trimmed with a scalpel to remove fine detail, in accordance with good clinical practice.

Virtual articulation: The unarticulated upper and lower dental casts were scanned using a custom lab scanner (Leeds Digital Dentistry). Each pair of centric relation records were scanned, unclamped, using the same scanner. See Fig 2.

To virtually articulate the scans, the upper and lower jaw and each of the 12 pairs of the centric relation records were aligned. The two records were individually aligned to the upper cast, using global alignment in WearCompare (Keeling and Osnes, 2019), before the alignment was refined using Meshlab(Cignoni et al., 2008). This alignment process was then repeated, but this time aligning the lower cast to the two repositioned centric relation records, bringing all four meshes into “articulation”. The alignment transformation was saved. This was repeated for all 12 pairs of scanned centric relation records. See Fig 3.

Measurement of these alignments proceeded as described later.

Traditional articulation: Following scanning, all twelve centric relation record pairs were used to physically re-articulate the lower cast on the articulator. This process involved use of low expansion mounting plaster. Each cast was assessed visually for correct seating into the centric relation record. A facebow record was not utilised and the upper cast remained mounted throughout, with the articulator pin unmoved and set at the original incisal separation. The

protrusive, immediate side shift and progressive side shift adjustments remained unchanged at factory settings throughout. See Fig 4.

In order to measure the traditional articulations, it was necessary to digitise each dataset. To achieve this, each set of articulations were scanned with two optical buccal bite records whilst mounted on the articulator using an intraoral scanner; Omnicam (Ortho 1.1.2). The full arch casts were also scanned with the Omnicam once, and this same set of scans was used during each articulation, to minimise confounding variables. See Fig 5.

The precision of this Omnicam method for measuring traditional articulations was investigated by scanning two buccal bite records on the same articulated casts five times without opening the casts between scans.

The CEREC Ortho software maintains the position of the lower arch and moves only the upper arch, based on the buccal bite scans. Therefore, all lower scans were aligned to the cast scanned in the custom Leeds scanner and the transformation matrix for each alignment was applied to the corresponding upper arch. This brought scan-pairs into the same alignment, whilst preserving idiosyncratic differences in the individual occlusions.

Measurement of all articulation methods: Three pairs of vertex points were identified across the upper and lower arch on the scans produced by the laboratory scanner in Meshlab, located on UR1,LR1; UR7,LR7 and UL7, LL7. The coordinate of each of the three points per arch, were recorded using custom software (Leeds Digital Dentistry), as outlined in (Gintaute et al., 2020). The identical key points were then identified on all twelve scan pairs produced by the Omnicam (the “Traditional” articulation), as outlined in (Gintaute et al., 2020). The “Virtual” arm of the study (using the clamp-less scanner) consisted of one set of casts with twelve different transformation matrices. Thus, the new location of the originally identified key points were identified by vertex-ID and recorded.

The X, Y and Z values for the displacement between upper and lower key point pairs were assessed separately, for each pair of key points. All casts were oriented such that Z represented the anteroposterior direction, X was lateral, and Y was superior-inferior.

The standard deviation was used as a proxy for precision, whilst a comparison of mean values indicated whether the methods produced a different occlusal result in the three anatomical directions.

Statistical analysis:

The precision of using the Omnicam bite scan to record the physical articulation group, was reported using the standard deviations for the directions in X, Y and Z between upper and lower key points, for the three key point pairs (Anterior, Left, Right). This initial assessment used the five repeated buccal bite scans from a single instance of the articulated casts (where the articulator had not been moved in between each scan).

Next, differences in the physical and virtual methods were assessed, using the 12 different polyvinylsiloxane centric relation records as follows.

The mean separation between upper and lower key points in the directions X, Y and Z were compared for the Traditional and Virtual groups using Students paired t-test. Beforehand, homogeneity of variance for the differences between each group was confirmed using Shapiro-Wilk. Significance was set at $p<0.05$.

The precision of the articulation was assessed using Levene's Test to compare the variances in the separation between upper and lower key points in the directions X, Y and Z. Significance was set at $p<0.05$.

5.4. Results

The Omnicam precision experiment produced standard deviations in key point distance which never exceeded 0.032mm laterally (X), 0.033mm vertically (Y) and 0.03mm anteroposteriorly (Z). See Table i for all findings.

For the traditionally articulated casts, standard deviations in key point distance never exceeded 0.076mm laterally (X), 0.102mm vertically (Y) and 0.073mm anteroposteriorly (Z). The virtual equivalent was 0.027mm laterally (x), 0.059mm vertically (y) and 0.024mm anteroposteriorly. Statistically significant differences ($p<0.05$) between mean values for Traditional and Virtual articulations were seen in all anteroposterior separations, and three of the six lateral and vertical separations. Statistically significant differences in variance (precision) were seen in all anteroposterior measurements, and all lateral measurements. See Figure 3 and Table ii for all results.

5.5. Discussion

This study investigated the precision of traditional and virtual articulation using centric relation records.

The first null hypothesis posed no statistically significant difference in precision between the articulation methods. In this experiment, the precision was defined as the standard deviation of the variation in upper and lower key point separation in the anteroposterior (Z), lateral (X) and vertical (Y) directions. Significant differences in anteroposterior and lateral directions, for all key points pairs across the two methods, lead us to reject this hypothesis.

The second null hypothesis stated that there would be no statistically significant difference in articulations produced by both methods. The articulations were defined in terms of upper/lower key point separations in the three anatomical directions noted above. In six of the nine test cases, a different mean value was recorded. In particular, all anteroposterior means differed significantly between Traditional and Virtual groups, with digital consistently producing smaller values. Therefore, the second null hypothesis is also rejected.

In considering the Traditional measurements using the Omnicam, it is worth noting that the precision measurements (standard deviations) for the main experiment were consistently double the values found in the preliminary experiment. Our preliminary experiment (which investigated the precision of the Omnicam in measuring the same articulation multiple times), resulted in standard deviations of around 0.03mm in each of X, Y and Z. These values agree well with previously published results (Osnes et al., 2020), although in the current experiment the precision is slightly poorer. This may be due to the increased separation between the teeth when recording centric relation registrations vs recording intercuspal position. The intra-oral scanner collects less data per-frame in the former case, because of the increased gap between the upper and lower teeth, and therefore has slightly poorer data with which to ‘glue’ the upper and lower arches together. Overall, while some of the variance in the Traditional articulations can be explained by variance in our measurement method, an equal amount of the variation cannot, and must be attributed to genuine differences in the physical articulation of the casts.

Note also, that these centric relation records were all recorded on the casts, not in the mouth. This was to reduce the variation in mandibular position, which might confound the results. However, in general, *in vivo* records will not necessarily seat as perfectly on *in vitro* casts, because of variations in the level of detail recorded. Thus, in our experiment the Traditional arm is likely to perform better than would be the case in clinical practice.

The centric relation records were trimmed of extraneous detail before use, as is often recommended clinically. This aids the full seating of such records on stone casts. However, this process would not be needed for the virtual method (because 3D meshes can pass through each other), and it could be argued that such trimming should be avoided, since useful detail is lost. Therefore, our results may be an under-estimate of the precision of the Virtual method.

The clinical implications of this experiment point towards an improved method of articulating dental casts in centric relation, which uses entirely familiar and available chairside techniques (in contrast to the low availability of intraoral scanners). It requires that the dental laboratory has sufficient cast scanning capabilities, but this tends to be more common in laboratories than in dental practices.

Further, this 'hybrid' of traditional chairside techniques coupled with modern high-resolution digital techniques may offer the best of both approaches. It is cost effective, simple, and may offer a better quality full-arch articulation than that produced by an intra-oral scanner alone (as seen here with Omnicam).

We cannot assess the trueness in this experiment, but high precision is desirable in any clinical technique. In particular, the anteroposterior precision of repeated centric relation records using our Virtual method never exceeded 0.025mm, while the Traditional articulations were consistently more than 0.060mm. It may also be worth noting that the precision of the five *identical* Omnicam repeats produced articulations less precise than the precision of the 12 different articulations produced using the digital method in the main experiment. The precision with which the Traditional articulations were recorded may have been improved by using a more recent intraoral scanner. The use of an extra-oral scanner (NextEngine, CA, USA) to digitise the articulations was investigated, but the results proved less reliable than the Omnicam scanner and were therefore not included in this manuscript. Trueness could be investigated, perhaps, by 3D printing a Michigan Splint and fitting this onto the articulated casts to test the number of occluding units on the splint. This would give a clinically relevant indication of the occlusal fit of both methods under investigation. This falls beyond the scope of the current experiment but is suggested as worthy of further investigation in the future.

It is likely that digitizing the centric relation records and mounting the casts virtually will confer a clinically detectable benefit over scanning a physical version of the articulated casts (as is currently the most common practice). Further work should be undertaken to compare this method with intraoral scanners, particularly in the use case of the centric relation record as opposed to an intercuspal record.

5.6. Conclusion

The precision of a virtual method for articulating dental casts using a traditional centric relation registration was compared to the traditional plaster mounting on an articulator. A significantly improved precision in articulation was noted for the Virtual group. This method does not require the clinician to modify their chairside technique or invest in expensive chairside scanning equipment.

5.7. Figures

Table v Precision of the Omnicam method for assessing the Traditional arm of the trial, across three axes and across the three key points.

Precision of Omnicam for Measuring Traditional Articulation (mean mm)(± standard deviation)			
	X	Y	Z
Anterior	0.051 (±0.032)	1.802 (±0.033)	0.289 (±0.012)
Right	-2.651 (±0.022)	2.192 (±0.030)	2.117 (±0.030)
Left	-0.672 (±0.022)	2.782 (±0.015)	7.135 (±0.013)

Table vi Key point distance for both experiments in mm(± standard deviation)

Key point distances for the two methods (mean mm)(±standard deviation)						
	Trad X	Virtual X	Trad Y	Virtual Y	Trad Z	Virtual Z
Anterior	-0.056 (±0.067)	-0.058 (±0.028)	1.972 (±0.102)	1.907 (±0.059)	0.462 (±0.063)	0.353 (±0.017)
Right	-2.617 (±0.076)	-2.720 (±0.019)	2.171 (±0.053)	2.270 (±0.034)	2.179 (±0.073)	2.038 (±0.023)
Left	-0.670 (±0.071)	-0.617 (±0.018)	2.841 (±0.057)	2.872 (±0.034)	7.396 (±0.068)	7.228 (±0.024)

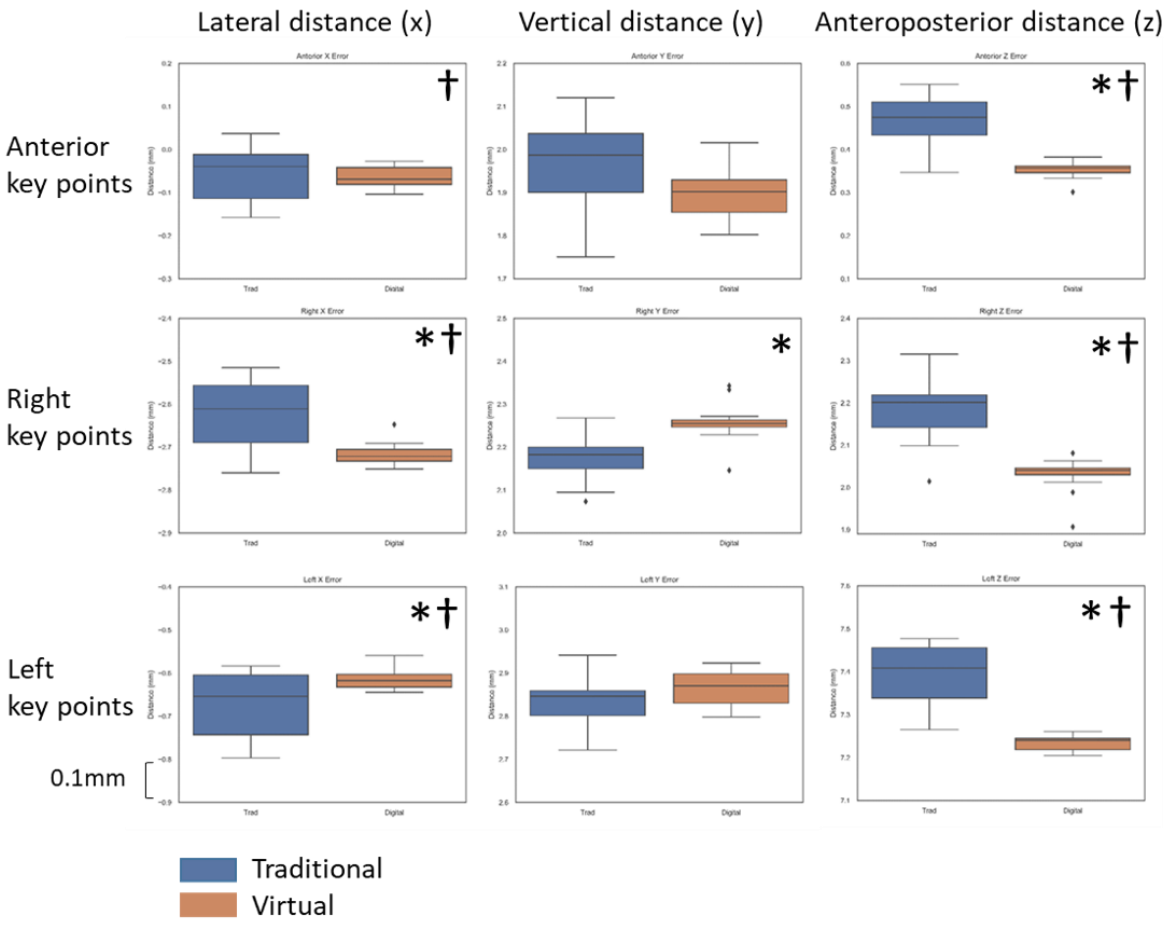


*Figure 5:1: Variation in key point distance for all three key points across the traditional group (blue) and the virtual group (orange). Y axis scale = 0.1mm per unit. * denotes a statistically significant difference (p < 0.05). † denotes a significant difference in variance (precision).*

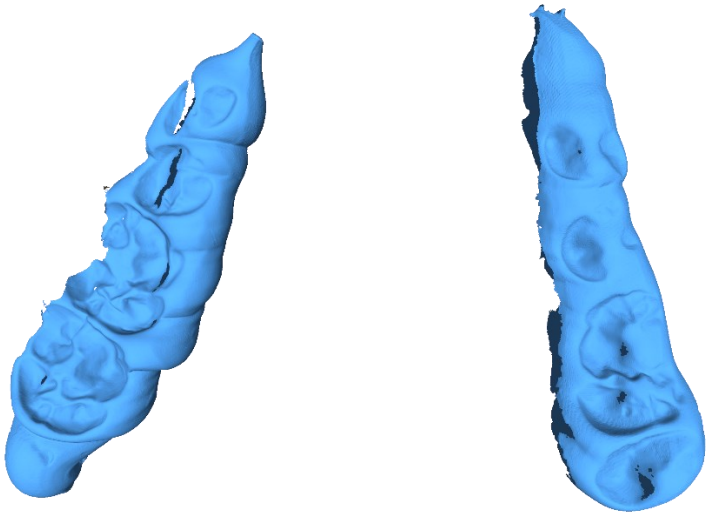


Figure 5:2: Centric relation records

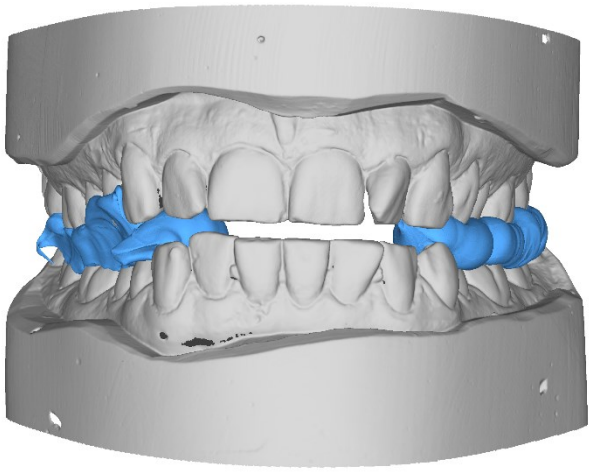


Figure 5:3: The cast scans having been articulated by aligning to the scans of the left and the right centric relation records (Virtual group).

Figure 5:4: Physical articulation (Traditional group only).

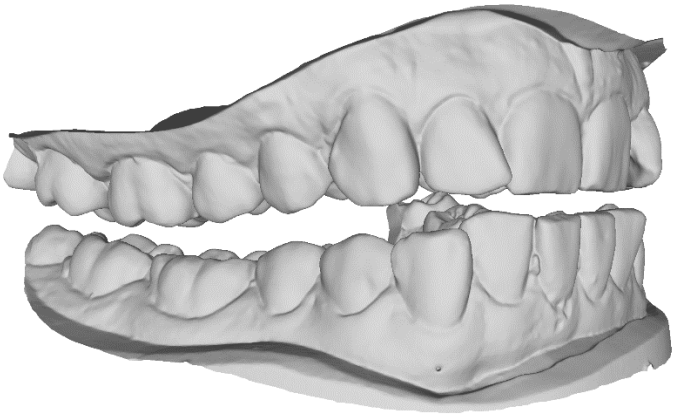


Figure 5:5: Articulated scans produced by the Omnicam, based upon physical articulation (Traditional group only).

5.8. List of References

1. Eriksson A, Ockert-Eriksson G, Lockowandt P, Eriksson O. Clinical factors and clinical variation influencing the reproducibility of interocclusal recording methods. Br Dent J. 2002 Apr;192(7):395–400.
2. Patel M, Alani A. Clinical issues in occlusion – Part II. Singapore Dent J [Internet]. 2015;36:2–11. Available from: http://dx.doi.org/10.1016/j.sdj.2015.09.004
3. Chai J, Tan E, Pang J-C. A study of the surface hardness and dimensional stability of several intermaxillary registration materials. Int J Prosthodont. 1994;7(6):538–42.
4. Breeding LC, Dixon DL. Compression resistance of four interocclusal recording materials. J Prosthet Dent. 1992;68(6):876–8.
5. Campos AA, Nathanson D. Compressibility of two polyvinyl siloxane interocclusal record materials and its effect on mounted cast relationships. J Prosthet Dent. 1999;82(4):456–61.
6. Ockert-Eriksson G, Eriksson a, Lockowandt P, Eriksson O. Materials for interocclusal records and their ability to reproduce a 3-dimensional jaw relationship. Int J Prosthodont. 2000;13(2):152–8.
7. Baumann TM. Effect of interocclusal recording materials on mounted casts. ProQuest Diss Theses. 2009;64.
8. Utz KH, Müller F, Lückerath W, Fuß E, Koeck B. Accuracy of check-bite registration and centric condylar position. J Oral Rehabil. 2002;29(5):458–66.
9. Alghazzawi TF. Advancements in CAD/CAM technology: Options for practical implementation. J Prosthodont Res [Internet]. 2016;60(2):72–84. Available from: http://dx.doi.org/10.1016/j.jpor.2016.01.003
10. Lepidi L, Galli M, Mastrangelo F, Venezia P, Joda T, Wang HL, et al. Virtual Articulators and Virtual Mounting Procedures: Where Do We Stand? J Prosthodont. 2020;30:24–35.

11. Rhee YK, Huh YH, Cho LR, Park CJ. Comparison of intraoral scanning and conventional impression techniques using 3-Dimensional superimposition. J Adv Prosthodont. 2015;7(6):460–7.

12. Keeling A, Osnes C. Leeds Digital Dentistry [Internet]. Website. 2019. Available from: https://leedsdigitaldentistry.com/

13. Walls AW, Wassell RW, Steele JG. A comparison of two methods for locating the intercuspal position (ICP) whilst mounting casts on an articulator. J Oral Rehabil. 1991;18(1):43–8.

14. Gkantidis N, Dritsas K, Ren Y, Halazonetis D, Katsaros C. An accurate and efficient method for occlusal tooth wear assessment using 3D digital dental models. Sci Rep. 2020;10(1):1–9.

15. Cignoni P, Callieri M, Corsini M, Dellepiane M, Ganovelli F, Ranzuglia G. Meshlab: an open-source mesh processing tool. Eurographics Ital. 2008;

16. Gintaute A, Keeling AJ, Osnes CA, Zitzmann NU, Ferrari M, Joda T. Precision of maxillo-mandibular registration with intraoral scanners in vitro. J Prosthodont Res [Internet]. 2020;64(2):114–9. Available from: https://doi.org/10.1016/j.jpor.2019.05.006

17. Osnes C, Wu J, Ferrari M, Joda T, Keeling A. Sources of Error in Maximum Intercuspation from Complete Dentate Full Arch Intraoral Scans in vitro. Int J Comput Dent. 2020;(Accepted, in press).

6. Discussion

6.1. Summary of work presented

This thesis investigated various ways with which to assess the quality of 3D data produced by dental scanners with focus on the precision of full arch scanning and articulation. In Chapter 2, the core issue of investigators' tendency to rely on presenting a single number to represent their findings, as previously raised by van der Meer *et al. (2012*) is discussed. A novel 'upper bound deviation' metric was proposed and shown to potentially give a more clinically relevant insight into the quality of the data under investigation. However, it still relies on a correct alignment and appropriate cropping of the scan data prior to analysis and is quite complex to implement.

Chapter 3 simplifies the implementation of the upper-bound metric by reporting the percentage value of the mesh deviating beyond a set threshold. Furthermore, it presents a novel method using topologically identical key points with which to measure cross-arch error and which foregoes relying on global alignment and the need to clean or crop the scan data. Results may be more clinically (and conceptually) accessible than the upper-bound deviation values presented in the previous chapter.

Chapter 4 uses the key point distance measurement method from the previous chapter to assess the accuracy with which the Omnicam is able to correctly record occlusion. Findings conclude that the buccal scans, and subsequent occlusion, were relatively precise, with key points never deviating beyond 0.022 mm; but that significant error was introduced by error in the full arch scan.

Chapter 5 uses the key point method presented in Chapter 3, and the findings from Chapter 4, to compare the precision between traditional articulation of dental models and a novel virtual method relying on a custom, clamp-less laboratory scanner. Findings indicate that the 'hybrid digital' method is significantly more precise than the traditional equivalent. In addition, it hints that it may also be optimal, compared to IOS inter-occlusal records, and thus represents the state-of-the-art in quality digital inter-occlusal registration, despite being perceivably less glamorous than a fully digital workflow.

6.2. Critical Reflection

All experiments within this thesis relied on *in vitro* experiments. It would be interesting to apply some of the methods introduced to IOS scans gathered *in vivo,* to gain an insight into the level of deterioration of scan data which would likely occur when using the devices for their intended purpose.

At the time of writing, Chapter 2 has been cited 28 times, indicating a certain relevance of its contents and the fast moving and highly productive nature of the field, currently. While the focus of this, and all other publications herein, focus on clinical relevance, the method presented in Chapter 2 was both conceptually and computationally challenging. As a result of this, the methods presented in subsequent manuscripts were intentionally more accessible to the reader. However, to make the use of these methods more approachable all custom software should be made openly available. *Using* custom software *and making it available to an unfamiliar end-user* are two very different things, however; and as a result most of the custom programs discussed require some degree of refinement before eventual public release (Keeling and Osnes, 2019).

With exception of the final publication, all papers presented investigated various aspects of full arch scanning using intraoral scanners. Chapter 2 concluded that only three of the six scanners tested showed clinically acceptable variations of less than 0.3mm when scanning an edentulous model. It must be noted at this stage, that “clinically acceptable” in this instance was defined in very loose terms, and with removable complete prosthodontics in mind. 0.3 mm errors would of course be deemed entirely unacceptable in fixed prosthodontics, and certainly problematic even in removable prosthodontics. (Hyde et al., 2014) found that patients (statistically significantly) prefer dentures made from silicone impressions over alginate impressions. With (Ender and Mehl, 2011) identifying mean arch errors of 0.15 mm from alginate impressions, our 0.3 mm cut-off value is clearly too great. Clinical work produced on scanners producing errors of 0.3 mm is likely to be negatively received by the patient, even in somewhat more lenient fields such as removable prosthodontics.

This brings us to the issue of a lack of a clinically acceptable baseline. The first three papers in this thesis, much in line with many digital research papers in the field, fail to include an analogue baseline against which to compare our findings. Working on the assumption that digital acquisition techniques are currently striving to produce quality of a similar standard to conventional silicone impressions, all papers presented could have been improved by introducing a “conventional arm”, a silicone impression scanned in a high quality 3D scanner.

This would have given an insight into whether silicone outperforms all scanners under investigation — in which case, digital is currently not quite up to standard for any clinical work requiring "silicone quality"— or whether some of the devices under investigation are capable of producing data which would be considered clinically acceptable, in comparison to silicone. Such a finding could have identified if any devices currently on the market *out-perform* the highest quality conventional standard – this would have been a very valuable finding, indeed.

6.3. The current state of full arch scanning and the effect of different measures of 'accuracy'

In Chapter 3, the reader would be forgiven for drawing the 'headline' conclusion that the Primescan produced statistically significantly more precise scans than the Omnicam over a full arch when scanning a dentate model. However, the real insight in this paper relates to the effect of different outcome measures, and the risk of the casual reader drawing incorrect (and clinically dangerous) conclusions, as illustrated by *Figure 3:2 Three methods with which to measure scan accuracy.* Let us now consider the results for the Omnicam full arch accuracy, measured in the three reported ways:

Firstly, the standard mean surface distance measure was used. The reader is reminded that this is the outcome measure used in *almost all* digital dental research over the past two decades. The Omnicam produced a mean full arch error of 0.116 mm. A casual reader might see this number and consider, "if on average, my cross-arch errors are only about 0.1 mm, I can fix that with chairside adjustments." ***They conclude that Omnicam is clinically acceptable over full arches.***

Secondly, the novel upper-bound method was used. Here, it was found that 42.1% of the Omnicam scan deviated by > 0.1 mm from the true value. The casual reader might now take notice, considering that nearly half of their impression will be inaccurate by > 0.1 mm. ***Some may conclude that the Omnicam is not acceptable over full arches.***

Thirdly, the novel 'topologically identical key point' method was used to measure inter-molar arch width error. Here, the Omnicam was found to shrink the intermolar width by, on average, **0.768 mm**. Our reader would now surely sit up and take stock. Few clinicians would tolerate errors of over three-quarters of a millimetre. ***They conclude that Omnicam is certainly unacceptable over full arches.***

It is important to note that the above results are not meant to vilify the Omnicam. Rather they show the importance of the metric by which data is measured. Identical scan data was used in all three analyses above, and yet the 'concluding value' has moved from around 0.1 mm to > 0.75 mm. This must surely reduce the risk of the dental audience innocently misinterpreting results and thus using digital technology in inappropriate clinical situations. Parallels can also be drawn to the issue of ISO standards for model scanners outlined in the introduction, where we see scanners advertised with '4 microns' accuracy. The wise reader will know that this means much larger errors will be incurred in the scan (let alone regions where the scanner fails to see at all – and produce holes – which may get artificially filled without the user knowing).

Returning to the results of Chaper 3, the Primescan did indeed perform well under all three measurement metrics. One might argue it has therefore 'passed the test', under the most stringent of analyses, and produces clinically acceptable full arch scans (in vitro). However, the issue of applying a 'single number' to the scan measurements is still present. The Primescan showed evidence of artificial enhancements to the detriment of the scan data on some occasions. While these errors were visually identified, it is worth noting that the novel method of measuring the upper-bound deviation of meshes, as presented in Chapter 2 would likely indicate the presence (but not location) of such over-sharpening artefacts, unlike more traditional methods. The key point method, on the other hand, would naturally give no indication of the presences of such problematic artefacts.

6.4. Underlying wand technology

Section 1.3 *A brief introduction to 3D data and 3D scanners* briefly touched on the various acquisition methods used by IOS. While this was mostly considered beyond the scope of this thesis, it is interesting to note that while the original Dentsply Sirona scanner, the Omnicam, used triangulated structured light, their more recent flagship product, the Primescan, relies on confocal scanning. This must be appreciated as a drastic and strategic shift — reflected by the increase in cost, doubling the purchase price of the device. However, our findings do indicate that the shift in technology may have improved scanner accuracy, as seen in Chapter 3, and as there are clear similarities between the Omnicam and Primescan software packages, one may

assume the improvement is the result of hardware, not software, changes.[18]

However, this is not the only change in the Primescan. The wand head is much larger than its predecessor, leading to an increased field of view. This may well improve full arch accuracy by increasing the size of each individual scan, making global alignment less challenging and more robust.[19] While a larger scanner head is likely to improve scanning *in vitro*, it is also likely to impede aspects of *in vivo* scanning, as regions which are difficult to access with an optical device, such as last standing molars, are less likely to be captured by a larger scanner head. Future work investigating scan coverage, *in vivo* accuracy and the accuracy of virtual bite registration using the Primescan would be very interesting.

6.5. The lack of acuity measurements of scan data

The trueness and precision of dental scanners have been widely explored, and the dimensional accuracy of scan data produced by various scanners, has been considered in depth. Despite numerous studies investigating the dimensional accuracy of scanners, little work has been undertaken to explore scanners' ability to reproduce surface detail; for the lack of a uniformly accepted term, we might call this 'sharpness' or 'acuity'. In fact, the ISO standard for optical scanners, ISO 12836:*2015,* fails to acknowledge that scanners may produce scans of various levels of acuity at all. Hence, there is currently no universal standard, or metric for measuring and validating, scan acuity within dentistry.

The acuity (or resolution) of scan data may have a direct impact on the quality of clinical work designed on it; for instance, a blunt crown margin, would negatively affect the fit and longevity of a crown designed on the scan data. However, this issue is rarely discussed within the digital field, where dimensional accuracy tends to be in focus. One might compare this to the decision the clinician makes in choosing the appropriate impression material during treatment. While certain cases may call for an alginate impression, other scenarios require light-bodied silicone to capture the necessary features for successful treatment. (See Figure

[18] Although, a 2020 publication did find somewhat improved scanner accuracy, of a single tooth preparation, following a software update in the Omnicam (Patzelt and Hack, 2020) and a 2019 study found significant differences in arch width precision between scanner software packages using the same intraoral scanner (Jablonski et al., 2019), highlighting that the software generating the scan data also has an effect on the quality of data produced.

[19] Much like a jigsaw puzzle where fewer, larger pieces make the puzzle easier to solve.

6:1.) This approach may translate well to digital, where certain scanners may well be appropriate for some use cases which require less detail but fail to perform when a high level of detail, and acuity, is required. Unfortunately, the tendency to assume that an IOS may be appropriate for work based on short span scans such as single preparation and bridges, purely because it has been proven dimensionally accurate to this extent, is problematic, as global accuracy and acuity are separate issues. A scanner which has only been deemed appropriate for short span scanning, does not necessarily produce data of appropriate quality for clinical use, even when only used for shorter scans.

The level of detail required for clinical use is uncertain, though a century's worth of technological development has determined the necessity for fine grit stone in dental models; with a 2020 study reporting 10% of particles to be smaller than 0.002mm (Nagasawa et al., 2020). The triangle edge length in IOS scans is generally 0.1mm, though this may vary depending on the scan region (as stated by (Nedelcu et al., 2018)).

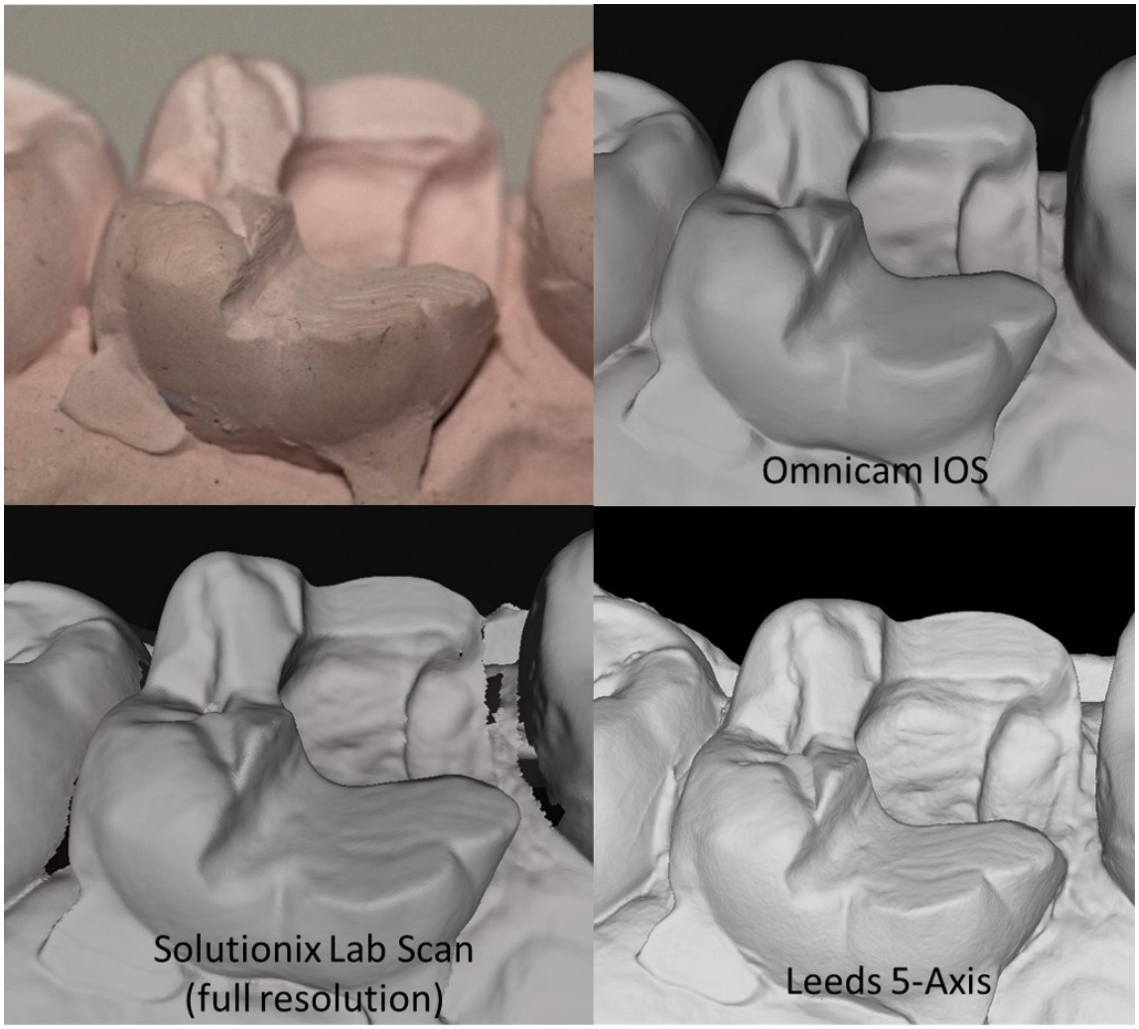


Figure 6:1 Visual comparison of acuity across a stone dental model and three scanners : Top left: photo of a stone model, top right: the scan produced by an Omnicam IOS, bottom left: the scan produced by a high-resolution lab scanner, bottom right: the scan produced by the custom lab scanner used in Chapter 5. Note the difference in the level of detail produced by the various scanners. Most of the quantitative methods presented in this thesis would fail to highlight the wide range in surface quality presented across these scans. Whether being able to capture fine detail such as those seen in the bottom right have clinical implications, may be assumed, but lacks evidence.

The sharpness, or lack thereof, produced by a scanner may depend on a range of factors, many of which have been discussed previously in this thesis. Surfacing a point cloud is prone to smoothing corners, as discussed in Chapter 1 and, again, in relation to the evidence of the Primescan artificially sharpening edges to overcome this issue, in Chapter 3. Conceptually, it is very unlikely that a scanner will acquire a data point at the very furthest point on a sharp edge or corner, or that this collection of points will make it through the many processing steps within the scanner (Le *et al.*, 2017). As a direct consequence of this, we see evidence of attempts to artificially enhance scan areas that may have undergone “blunting”, as seen in Chapter 3.

Artificial enhancement of data, as is seen in Chapter 3, with the occasional over-sharpening of cusps and margins, is now suspected to be common across many scanners on the market, as indicated by a recent publication outlining how to improve scanner data using a generative adversarial network to enhance scan data (Kim *et al.*, 2020). These black-box artificial enhancements are commercially driven (to produce nicer looking scans) but are not talked about in the sales brochures. As such, is likely that most end-users are unaware of these features being implemented (across the industry). In analogue terms, this is equivalent to the dentist electing to record their crown impression in only a heavy-bodied silicone and asking the technician to sharpen the margins on the model with a scalpel… This is clearly problematic and the aspiration is to record the margin correctly in the first place (using a light-bodied wash and a Type IV die stone).

Therefore, such artificial digital enhancements should be implemented with caution: a crown designed on an artificially sharpened preparation could have negative effects on the success of the crown and long-term treatment (typically, sharpening algorithms increase the height of the margin, which will either result in a larger marginal gap on the prosthesis, or a larger cement spacing inside the prosthesis). Future work seeking a method with which to measure sharpness of 3D scan data could become highly valuable in the near future, as a means to determine whether data may have been *over-sharpened*, and thus inappropriate for certain clinical use-cases.

As a non-clinical researcher, no attempts have been made to identify the quality of scan data which would be needed to be considered clinically acceptable. This would of course depend on the use case and clinical scenario in question. The comparison between scanner acuity and various impression materials does raise some interesting questions: It is a known fact that IOSs produce relatively low resolution scan data (normally with triangle edges around 0.1mm long) due to the computational (and financial/market) challenges faced by the device manufacturers. As such, clinical work produced on IOS scan data may be compared to conventional clinical work produced using, perhaps, heavy-bodied silicone impressions, when considering the level of detail captured. It would be interesting to know the longevity and success-rate of dental work produced using IOSs (as compared directly to conventional methods; as touched on by (Roggendorf et al., 2012; Aziz et al., 2019)). If the success-rate of work produced is no poorer than traditional methods, one may ask whether light-bodied silicone, and other high-fidelity materials, and scanners, are necessary. This would be an interesting opportunity for digital dentistry to inform (and benefit) conventional dentistry. Analogue dentists would no longer need to purchase light or medium bodied silicones for

prosthodontic work, since digital dentistry would have shown there is no advantage. This would represent a step-change in prosthodontic training and treatment.

If, on the other hand, the clinical outcomes of work produced by IOSs is poorer than traditional methods, perhaps “hybrid methods” such as scanning a silicone impression or model in a lab scanner may be concluded as an optimal clinical workflow, in terms of quality, in some scenarios. This would certainly be the case for clinical treatments which require high resolution data, such as articulation (where the patient can feel 0.01mm discrepancies), as shown in Chapter 5, and likely crown work, where the longevity of the crown may depend on the quality of the marginal fit.

6.6. Contributions

In Chapter 2, we raise the issue of the "meaningless mean" and the risk of relying on mean and median values when measuring 3D data. The standard methods with which investigators evaluate the quality of scan data does not reflect the any clinically useful metric. We propose the use of an "upper-bound value" to gain a more clinically relevant insight into the data under analysis. While we use 0.3mm deviation as our threshold for clinical acceptance, this value would differ for different disciplines within dentistry.

In Chapter 3, the upper-bound value method, presented in the previous chapter, is made more conceptually accessible, by reporting the percentage of a scan deviating beyond 0.1mm. The virtual key point method is also introduced, by which identical features on inherently different meshes can be identified. This method does not rely on global alignment (which pitfalls are outlined *in 1.3.1 Alignment in data generation and measuring)* and can be applied to data generated both *in vitro* and *in vivo.* The precision of the key point method is identified in the preliminary study in Chapter 4 as never deviating beyond 0.004 mm. See Appendix 8.2 for a detailed description of the key point method.

Chapter 3 also explores the key issue of different measurement methods producing different results, as illustrated in Figure *3:2 Three methods with which to measure scan accuracy*., where the same scan data produce concluding figures ranging from 0.1mm to >0.75mm scan errors depending on the analysis method used. This highlights the importance of choosing the appropriate method for measuring data to produce clinically valid conclusions.

Having identified the level of precision with which the Omnicam is able to virtually occlude dental models in the previous study, we present, in the final experiment, a method of virtually occluding dental models which is significantly more precise than traditional methods. Moreover, the suggested method requires no modification to the clinician's workflow.

To conclude, this thesis presented four publications investigating various methods of measuring the quality of data produced by 3D dental scanners, with an eye on clinical relevance and applicability. One novel measure (anatomically identical key points) is subsequently used to investigate a proposed enhancement to recording traditional inter-occlusal registrations, which would be available to all dentists (including the great majority who do not own IOSs) but which is shown to benefit from a post-hoc digitization stage in the dental laboratory. These findings become obtainable through the results identified in Chapter 4 – the level of precision possible when using the Omnicam to record virtual occlusion. Armed with this metric, we are able to critically evaluate the methods investigated in Chapter 5 and suggest an optimal method beneficial to the general clinical audience. This illustrates how outcome measures must not be considered purely as a passive measure of the current status quo but should rather inform circular feedback with which to inform developments of the next generation of digital technologies and techniques to improve clinical treatment.

6.7. Future Work

The works presented could benefit from future investigation into comparisons against conventional methods. With the working assumption that digital acquisition techniques need to be of the same or higher standard as conventional silicone impressions, the ability to rank the quality of data produced by a method or device in comparison to the current "gold standard" of a good silicon impression (scanned in a high quality 3D scanner) would arguably be the most accessible method by which to assess the quality of data produced by different methods and devices.

In relation to the work presented in Chapter 3, indicating that potential merits of the Primescan scanner, future work *in vivo* would be particularly interesting. In fact, regardless of the scanner in question, investigations into the difference in data produced by IOS *in vitro* vs *in vivo* would be highly beneficial to the field as a whole. The key point method would be an appropriate tool for such a study.

The acuity of dental scan data has been discussed briefly, but the much-needed thesis on the topic remains unwritten. A standard by which to assess the level of acuity and surface detail present in a scan would be highly beneficial to the field. Related to this, methods by which to identify artificial sharpening present in scan data — and an awareness of at which point scan enhancements become detrimental to clinical work— would also be advantageous.

While comparative investigations into the longevity of dental work produced using digital and conventional methods have been done (Roggendorf et al., 2012; Aziz et al., 2019), further work is ever required as the technology changes, and could shed light into the current quality of work produced using digital, and hybrid, methods.

6.8. List of references

1. Aziz, A., El-Mowafy, O., Tenenbaum, H.C., Lawrence, H.P. and Shokati, B. 2019. Clinical performance of chairside monolithic lithium disilicate glass-ceramic CAD-CAM crowns. Journal of Esthetic and Restorative Dentistry. 31(6), pp.613–619.

2. Ender, A. and Mehl, A. 2011. Full arch scans : Conventional versus digital impressions - An in-vitro study. International Journal of Computerized Dentistry. 14(21).

3. Hyde, T.P., Craddock, H.L., Gray, J.C., Pavitt, S.H., Hulme, C., Godfrey, M., Fernandez, C., Navarro-Coy, N., Dillon, S., Wright, J., Brown, S., Dukanovic, G. and Brunton, P.A. 2014. A Randomised Controlled Trial of complete denture impression materials. Journal of Dentistry. 42(8), pp.895–901.

4. Jablonski, R., Elneklawy, M., Osnes, C., Ferrari, M., Wu, J. and Keeling, A. 2019. Guided Scanning Improves Full Arch Precision of an Intra-Oral Scanner In: IADR. Vancouver.

5. Keeling, A. and Osnes, C. 2019. Leeds Digital Dentistry. Website. [Online]. Available from: https://leedsdigitaldentistry.com/.

6. Kim, T., Cho, Y., Kim, D., Chang, M. and Kim, Y. 2020. Tooth Segmentation of 3D Scan Data Using Generative Adversarial Networks. *Applied Sciences*. **10**(490), pp.1–17.

7. Le, M.T., Chen, L.C. and Lin, C.J. 2017. Reconstruction of accurate 3-D surfaces with sharp edges using digital structured light projection and multi-dimensional image fusion. Optics and Lasers in Engineering. 96(April), pp.17–34.

8. Nagasawa, Y., Hibino, Y., Shigeta, H., Eda, Y., Matsumoto, S. and Nakajima, H. 2020. Characteristics of a new dental stone mixed by shaking. Dental Materials Journal. 39(3), pp.355–366.

9. Nedelcu, R., Olsson, P., Nyström, I. and Thor, A. 2018. Finish line distinctness and accuracy in 7 intraoral scanners versus conventional impression: An in vitro descriptive comparison.

BMC Oral Health. 18(1), pp.1–11.

10. Patzelt, S.B.M. and Hack, G. 2020. Evaluation of the Accuracy of an Intraoral Scanner : An Update on the Omnicam Global Dentistry Evaluation of the Accuracy of an Intraoral Scanner : An Update on the Omnicam. . (November).

11. Roggendorf, M.J., Kunzi, B., Ebert, J., Roggendorf, H.C., Frankenberger, R. and Reich, S.M. 2012. Seven-year clinical performance of CEREC-2 all-ceramic CAD/CAM restorations placed within deeply destroyed teeth. Clinical Oral Investigations. 16(5), pp.1413–1424.

7. Conclusion

The studies presented in this thesis show that to produce valuable conclusions, investigators must decide on the appropriate metric for their use-case. This research shows that conclusions reached are dictated by the means of investigation, as illustrated in Figure 3:2 *Three methods with which to measure scan accuracy.* Likewise, Chapter 3 indicates the Primescan to be among the most appropriate of the devices under investigation, within this thesis as a whole, in recording the full arch, but the study also shows that significant sharpening artifacts may occur, which, at the time of writing, may potentially make it less suitable for crown work and other procedures which require accurate edges and margins. Any study investigating average global deviation, as discussed in Chapter 2 — and as is currently the standard method of analysis within the field — is unlikely to give any indication of local errors such as these sharpening artifacts, despite these potentially being of a magnitude considered likely to affect the clinical fit of a crown or other small prosthesis. As such, whether or not an investigator decides that a device is appropriate for clinical use is entirely dependent on the metric of measure used.

Our findings indicate there is no *one-size-fits-all* when assessing 3D data with the aim to provide a clinically relevant insight into the data under investigation. It follows that there is likely to be no one scanner that performs better than all other scanners in all use-cases.

This thesis has presented several novel concepts. The overarching conclusion which it is hoped a reader might gain from the studies presented is that any investigator within any clinical field should critically consider the motivations for undertaking a study and appropriately design the experiment, and methods of data analysis, to ensure that the findings are clinically relevant and accessible. The facilitation of findings to inform new methods and research should be encouraged: the scientific method applies even in the financially — and emotionally— driven field of digital dentistry.

Secondly, the studies within this thesis illustrate that the quality of data produced within the digital dentistry field, as a whole, is a gamut ranging from poor to clinically appropriate, and that *not all digital is equal.* This is worth laboring, as there is currently a tendency to approach “digital”, as a single, uniform method and clinical approach. We find, however, that the quality of digital data produced by scanners can range from poor to excellent, much like one might expect from traditional methods. Much like clinicians appreciate the difference between

alginate and silicone impressions, one must appreciate that the quality of digital methods, and devices, may range similarly. That the quality of input data is adequate for the success of clinical treatment using digital solutions is therefore crucial, just as is the case in conventional dentistry.

For technology to be worth adopting, it should be better than its conventional equivalent. So while complex, and various, methods of establishing a numerical value with which to identify the quality of 3D data can hold some merit, the true measure as to whether a digital device is appropriate for clinical use, may simply be to rank it against a (well scanned) silicone impression. If silicone continues to outperform digital methods — but if one assumes that scanning a digital impression is less skill-dependent than taking a conventional impression is — perhaps a good quality digital impression ought to, for the time being, be considered a safe and consistent *baseline* for clinical work. Clinical work requiring the *very best*, on the other hand, may still be more appropriately performed using a good silicone impression and either conventional, or selected, precise digital methods, resulting in a hybrid conventional/digital workflow, tailored to the specific clinical scenario.

Digital dentistry is an exciting, innovative field with great potential to help clinicians work faster and better, and to reduce the amount of time the patient must spend in the chair overall. However, this is not always the case today, nor will it be in the future, unless any digital device or method introduced into the practitioner's workflow is approached with rationality and a healthy portion of cynicism, and with robust and appropriate evidence to support its use.

8. Appendix

8.1. Corrections

- Figure 2 in section 2.7 incorrectly stated in publication that the horizontal black line indicated *mean* value. The line indicates *median* value, not mean, in all figures throughout this thesis.

8.2. A summary of the automatic key point method

While the automatic key point method was developed for the current body of research and used in Chapters 3, 4 and 5, it was first published in the collaborative work titled *Precision of maxillo-mandibular registration with intraoral scanners in vitro* (Gintaute *et al.*, 2020). This means that a clear description of the exact method is not included in any of the publications herein. I will therefore briefly summarise the method here.

A key issue when comparing 3D meshes is that each mesh is unique, even if the scan object recorded was the same. This is partly due to variations introduced during data acquisition, and partly due to the fact that surfacing algorithms may produce slightly differing meshes each time the algorithm is run. (See Figure 8:1 *Variation in position of vertices in scans*.) In addition to this, it was discovered that while the Cerec Omnicam software did not recalculate its meshes each time an occlusal bite scan was repeated in the experiments for Chapter 4, the software did occasionally re-number its vertices, meaning that relying on the vertex ID of the scan for analysis was no longer a reliable method. As a result, it was clear that a method with which to reliably identifying virtual key features across varying meshes would be beneficial.

The automatic key point method works as follows. (See also Figure 8:2 Key point method.)

1. A key point is identified on Mesh A. This is the feature we would like to identify on all meshes.
2. Mesh B has already been roughly aligned to mesh A. The user specifies a search radius, and an x mm sphere (normally 10mm when working with dental models) is cropped from mesh A, centered around the key point.
3. The spherical subsection from Mesh A is finely aligned to Mesh B.

4. A new vertex point is created on Mesh B in the exact location of where keypoint A was located after alignment.

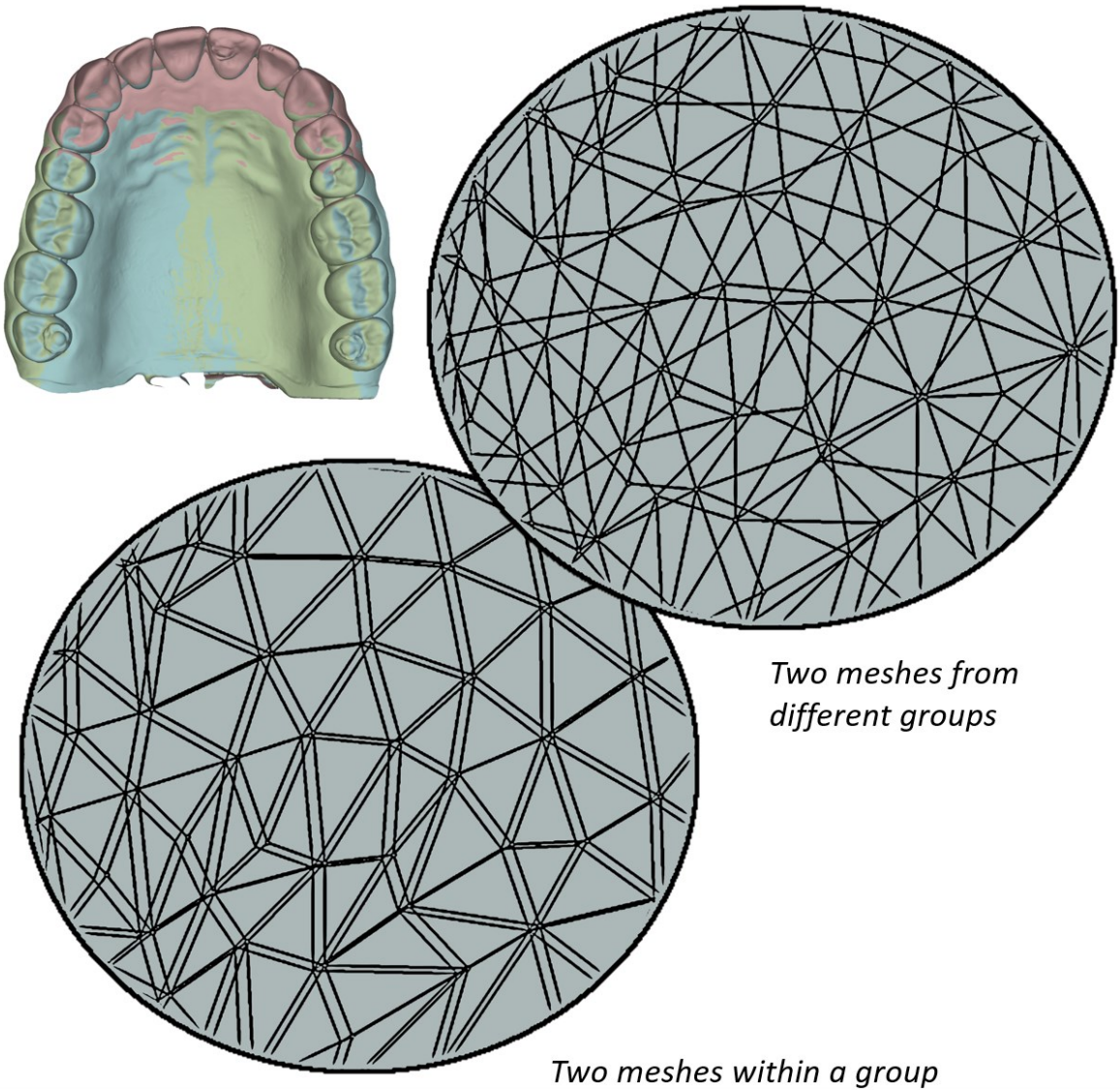


Figure 8:1 Variation in position of vertices in scans. Initially illustrated for Chapter ***Error! Reference source not found.****. The meshes from "within a group" consist of scan data which has only undergone surface reconstruction once, and thus consist of identically "constructed" meshes, although the global location of each individual mesh may differ. The meshes from different groups are the result of the surface reconstruction having calculated two different solutions (not necessarily from different scans). Here, the vertices are entirely random, meaning that making a robust measurement between two scans is complicated. With exception of the experiments investigating the accuracy of the Omnicam's virtual bite record feature in Chapter* ***Error! Reference source not found.****, all investigations in this thesis, and most studies in the field, work with the latter kind of meshes.*

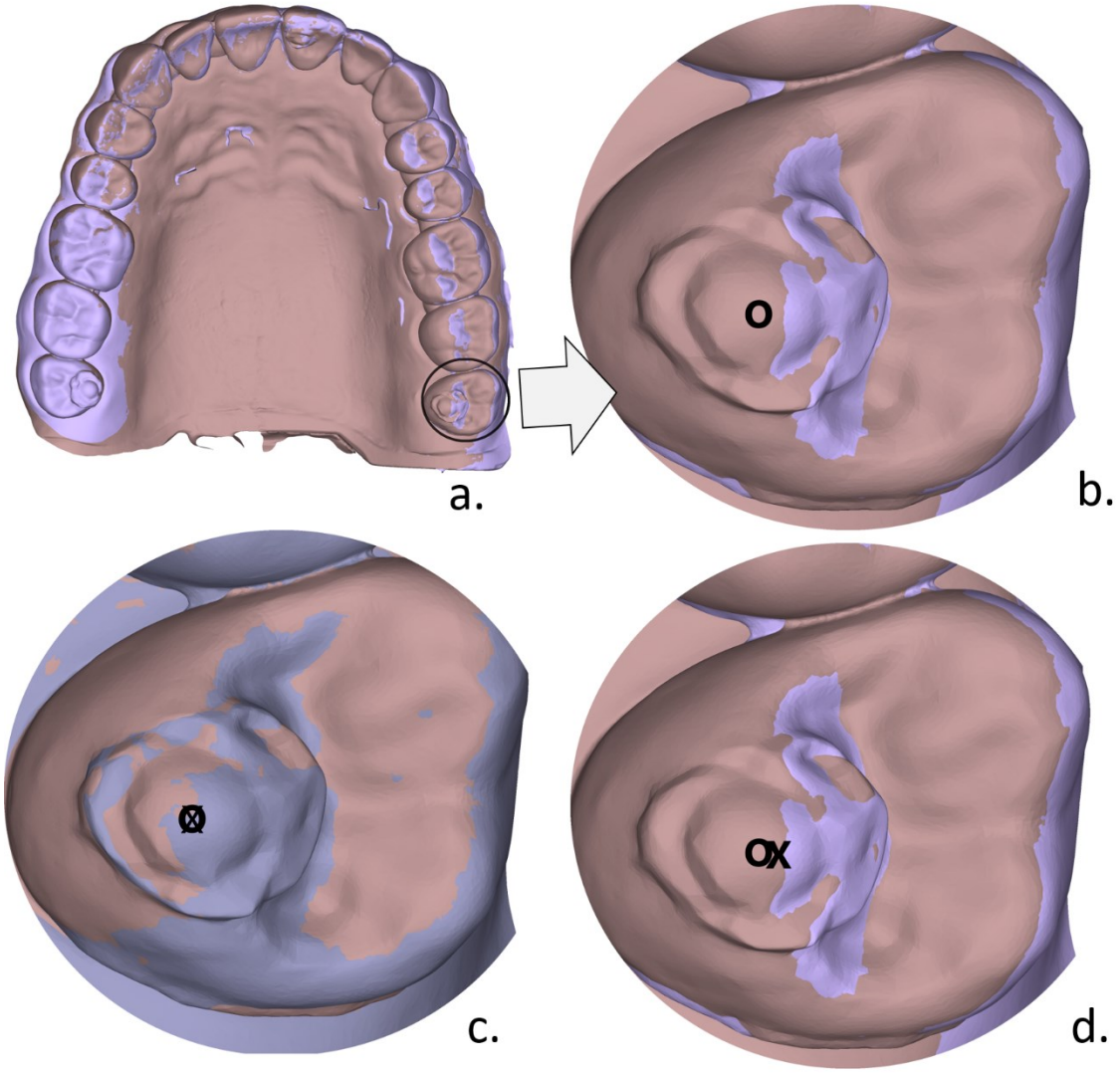


Figure 8:2 Key point method: To identify the exact location of a key point on a separate mesh, a section (10 mm radius) around the manually selected key point is created on the source scan (a). This "patch" is finely aligned to the target scan , before the precise location of the key point from the source mesh is mapped onto the target mesh (c). (d) shows the source scan back in its original location and the target scan with its topologically identical key point labelled as "x."

8.3. Additional publications and presentations of relevance

3D Scanning:

1. Gintaute A, Keeling AJ, **Osnes C**, Zitzmann NU, Ferrari M, Joda T. Precision of maxillo-mandibular registration with intraoral scanners in vitro. J Prosthodont Res 2019. doi:10.1016/j.jpor.2019.05.006.
2. Jablonski RY, **Osnes C**, Khambay BS, Nattress BR, Keeling AJ. Accuracy of capturing oncology facial defects with multimodal image fusion versus laser scanning. J Prosthet Dent 2019:1–6. doi:10.1016/j.prosdent.2018.10.017.
3. Jablonski RY, **Osnes C**, Khambay BS, Nattress BR, Keeling AJ. ScienceDirect The Surgeon, Journal of the Royal Colleges of Surgeons of Edinburgh and Ireland An in-vitro study to assess the feasibility , validity and precision of capturing oncology facial defects with multimodal image fusion. Surg 2017:1–6. doi:10.1016/j.surge.2017.11.002.

Haptic simulation development and education research:

4. **Osnes C**, Duke A, Wu J, Franklin P, Mushtaq F, Keeling A. Investigating the construct validity of a haptic virtual caries simulation for dental education. BMJ Simul Technol Enhanc Learn 2020:1–5.
5. Al-Saud L M., Mushtaq F, Mann, Mirghani I, Balkhoyor A, Harris R, **Osnes C**, Keeling A, Mon-Williams M A, Manogue M. Early assessment with a virtual reality haptic simulator predicts performance in clinical practice. BMJ Simul Technol Enhanc Learn 2019. doi:10.1136/bmjstel-2018-000420.
6. Mirghani I, Mushtaq F, Balkhoyor A, Al-Saud L, **Osnes C**, Keeling A, *et al.* The factors that count in selecting future dentists: sensorimotor and soft skills. Br Dent J 2019;226:417–21. doi:10.1038/s41415-019-0030-3.
7. **Osnes C**, Keeling AJ. Development of haptic caries simulation for dental education. J Surg Simul 2017;4:29–34.

Measuring erosive tooth wear:

8. Kumar S, Keeling A, **Osnes C**, Bartlett D, O'Toole S. The sensitivity of digital intraoral scanners at measuring early erosive wear. J Dent 2019;81. doi:10.1016/j.jdent.2018.12.005.

9. O'Toole S, **Osnes C**, Bartlett D, Keeling A. Investigation into the accuracy and measurement methods of sequential 3D dental scan alignment. Dent Mater 2019:8–13. doi:10.1016/j.dental.2019.01.012.
10. O'Toole S, **Osnes C,** Bartlett D, Keeling A. Investigation into the validity of WearCompare , a purpose-built software to quantify erosive tooth wear progression. Dent Mater 2019;35:1408–14. doi:10.1016/j.dental.2019.07.023.

Digital dentures:

11. Marie A, Keeling A, Hyde TP, Nattress BR, Pavitt S, Murphy RJ, Shary, TJ, Dillon, S, **Osnes, C**, Wood, DJ. Deformation and retentive force following in vitro cyclic fatigue of cobalt-chrome and aryl ketone polymer (AKP) clasps. Dent Mater 2019:1–9. doi:10.1016/j.dental.2019.02.028.
12. Davda K, **Osnes C**, Dillon S, Wu J, Hyde TP, Keeling A. An Investigation into the Trueness and Precision of Copy Denture Templates Produced by Rapid Prototyping and Conventional Means. Eur J Prosthodont Restor Dent 2017;19:67–71. doi:10.1922/EJPRD_01716Davda07.

Anatomical education:

13. Rea P, Livingstone D, Jocks IT, **Osnes C**. Digitisation of Anatomical Specimens and Historical Pathology Specimens for Educational Benefit. Serious Games Edutainment Appl 2017:101–19.

Presentations of note:

- The Meaningless Mean? Cecilie Osnes. British Society of Prosthodontic Research Annual Conference 2019. London, 2019. *This presentation was based on the initial results and findings of the project which became Chapter 2:* Full arch precision of six intraoral scanners in vitro.
- Efficacy of Haptic Caries. Cecilie Osnes. COPDEND Dental Educators Conference 2018. Birmingham, November 2018. Invited speaker.

Conference Posters:

- 3D-printed Dentures Deviate from the CAD File. C. Osnes*, S. Khalid, A. Keeling 2020 IADR/AADR/CADR General Session (Washington, D.C., USA) [Cancelled due to Covid-19]
- Edentulous Hybrid 3D Scanning Yields Maximal Surface Area Capture. K. Davda*, S. Khalid, C. Osnes, A. Keeling. 2020 IADR/AADR/CADR General Session (Washington, D.C., USA) [Cancelled due to Covid-19]
- Investigating Trueness and Precision of 3D Printed Denture Fitting Surfaces. S. Dillon*, C. Osnes, T.P. Hyde, A. Keeling. 2020 IADR/AADR/CADR General Session (Washington, D.C., USA) [Cancelled due to Covid-19]
- Optimising Contemporary Workflows to Improve Digital Dentures. T.P, Hyde*, K. Davda, S. Khalid, S. Dillon, C. Osnes, A. Keeling. 2020 IADR/AADR/CADR General Session (Washington, D.C., USA) [Cancelled due to Covid-19]
- Guided Scanning Improves Full Arch Precision of an Intra-Oral Scanner. R. Jablonski*, M. Elneklawy, C. Osnes, M. Ferrari, J. Wu, A. Keeling, Presented at IADR, 19th -22nd June 2019. Vancouver.
- Clasp Deformation: CoCr and Ultaire™ AKP Clasps Following Cyclical Loading. C. Osnes, A.K. Marie, S. Dillon, D.J. Wood, T.P. Hyde*, B. Nattress, A. Keeling. Presented at IADR. 25th-28th July 2018. London
- Measurement of Early Tooth Wear Using Profilometry and Intraoral Scanning. S. Kumar*, A. Keeling, C. Osnes, D. Bartlett, S. O'Toole. Presented at IADR. 25th-28th July 2018. London
- Digital comparison software significantly underestimates erosive tooth wear. S. O'Toole*, C. Osnes, A. Keeling. Presented at IADR. 25th-28th July, 2018. London
- A low-cost augmented reality system to assist student evaluation of typodont crown preparations. C. Osnes*, A. Keeling. Presented at ADEE. 24th of August, 2017. Vilnius.
- Feasibility and accuracy of capturing oncology facial defects with surface scanning: a proof of concept study. R. Jablonski*, A. Keeling, B. Khambay, C. Osnes, B. Nattress. BSSPD. April 2017. London.
- Developing Patient-Specific Haptic Simulation for Undergraduate Prosthodontic Training. A. Keeling*, C. Osnes, C. Hird, P. Franklin. ADEE. August 2016. Barcelona.

- An investigation into the accuracy of copy denture templates produced by rapid prototyping and conventional means. K. Davda*, A. Keeling, S. Dillon, C. Osnes, P. Hyde. BSSPD. May 2016. Liverpool.

8.4. Comprehensive list of references

1. 3M n.d. 3M™ True Definition Intraoral Scanner acquired by Midmark in U.S. and Canadian Markets. *3m.com*. [Online]. [Accessed 11 December 2020]. Available from: https://www.3m.com/3M/en_US/dental-us/expertise/digital-dental-impressions/.

2. Ahlholm, P., Sipila, K., Vallittu, P., Jakonen, M., Kotiranta, U., Sipilä, K., Vallittu, P., Jakonen, M., Kotiranta, U., Sipila, K., Vallittu, P., Jakonen, M. and Kotiranta, U. 2018. Digital Versus Conventional Impressions in Fixed Prosthodontics: A Review. *Journal of Prosthodontics*. **27**(1), pp.1–7.

3. Ahuja, D. V. and Coons, A. 1968. Interactive Graphics in Data Processing: Geometry for construction and display. *IBM Systems Journal*. **3.4**, pp.188–205.

4. Alghazzawi, T.F. 2016. Advancements in CAD/CAM technology: Options for practical implementation. *Journal of Prosthodontic Research*. **60**(2), pp.72–84.

5. Aziz, A., El-Mowafy, O., Tenenbaum, H.C., Lawrence, H.P. and Shokati, B. 2019. Clinical performance of chairside monolithic lithium disilicate glass-ceramic CAD-CAM crowns. *Journal of Esthetic and Restorative Dentistry*. **31**(6), pp.613–619.

6. Baumann, T.M. 2009. Effect of interocclusal recording materials on mounted casts. *ProQuest Dissertations and Theses*., p.64.

7. Besl, P.J. and McKay, N.D. 1992. A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. **14**(2), pp.239–256.

8. Bohner, O.L.L., Neto, P.T., Ahmed, A.S., Mori, M., Lagana, D.C. and Sesma, N. 2016. CERECChairside Systemto Registerand Designthe Occlusion in Restorative Dentistry : A Systematic Literature Review. *Journal of Esthetic and Restorative Dentistry*. **28**(4).

9. Braian, M. 2018. *Studies on the trueness and precision of additive manufacturing and intraoral scanning*.[Online] Malmö University. Available from: https://www.diva-portal.org/smash/get/diva2:1404654/FULLTEXT01.pdf.

10. Braian, M. and Wennerberg, A. 2019. Trueness and precision of 5 intraoral scanners for scanning edentulous and dentate complete-arch mandibular casts: A comparative in vitro study. *Journal of Prosthetic Dentistry*. **122**(2), pp.129–136.

11. Breeding, L.C. and Dixon, D.L. 1992. Compression resistance of four interocclusal recording

materials. *The Journal of Prosthetic Dentistry*. **68**(6), pp.876–878.

12. British Standards Institution 2015. BS EN ISO 12836:2015: Dentistry - Digitizing devices for CAD / CAM systems for indirect dental restorations — Test methods for assessing accuracy

13. Campos, A.A. and Nathanson, D. 1999. Compressibility of two polyvinyl siloxane interocclusal record materials and its effect on mounted cast relationships. *The Journal of prosthetic dentistry*. **82**(4), pp.456–461.

14. Cao, J., Wushour, S., Yao, X., Li, N., Liang, J. and Liang, X. 2012. Sharp feature extraction in point clouds. *IET Image Processing*. **6**(7), pp.863–869.

15. Chai, J., Tan, E. and Pang, J.-C. 1994. A study of the surface hardness and dimensional stability of several intermaxillary registration materials. *International Journal of Prosthodontics*. **7**(6), pp.538–542.

16. Chang, H. 2005. *Inventing Temperature: Measurement and Scientific Progress*.

17. Chen, Y. and Medioni, G. 1991. Object Modeling by Registration of Multiple Range Images. *Proceedings - IEEE International Conference on Robotics and Automation*., pp.2724–2729.

18. Chochlidakis, K.M., Papaspyridakos, P., Geminiani, A., Chen, C.J., Feng, I.J. and Ercoli, C. 2016. Digital versus conventional impressions for fixed prosthodontics: A systematic review and meta-analysis. *Journal of Prosthetic Dentistry*. **116**(2), pp.184–190.

19. Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F. and Ranzuglia, G. 2008. Meshlab: an open-source mesh processing tool. *Eurographics Italian*.

20. Curless, B. and Levoy, M. 1996. Volumetric method for building complex models from range images. *Proceedings of the ACM SIGGRAPH Conference on Computer Graphics*., pp.303–312.

21. Datta, A., Kim, J. and Kanade, T. 2009. Accurate camera calibration using iterative refinement of control points. *Computer Vision Workshops (…*., pp.1201–1208.

22. DeLong, R., Ko, C.C., Anderson, G.C., Hodges, J.S. and Douglas, W.H. 2002. Comparing maximum intercuspal contacts of virtual dental patients and mounted dental casts. *Journal of Prosthetic Dentistry*. **88**(6), pp.622–630.

23. Demarsin, K., Vanderstraeten, D., Volodine, T. and Roose, D. 2007. Detection of closed sharp edges in point clouds using normal estimation and graph theory. *CAD Computer Aided Design*. **39**(4), pp.276–283.

24. Ender, A., Attin, T. and Mehl, A. 2016. In vivo precision of conventional and digital methods of obtaining complete-arch dental impressions. *Journal of Prosthetic Dentistry*. **115**(3), pp.313–320.

25. Ender, A. and Mehl, A. 2011. Full arch scans : Conventional versus digital impressions - An in-vitro study. *International Journal of Computerized Dentistry*. **14**(21).

26. Ender, A., Zimmermann, M., Attin, T. and Mehl, A. 2015. In vivo precision of conventional and digital methods for obtaining quadrant dental impressions. *Clinical Oral Investigations*. **20**(7), pp.1495–1504.

27. Ender, A., Zimmermann, M. and Mehl, A. 2019. Accuracy of complete- and partial-arch impressions of actual intraoral scanning systems in vitro. *International journal of computerized dentistry*. **22**(1), pp.11–19.

28. Eriksson, A., Ockert-Eriksson, G., Lockowandt, P. and Eriksson, O. 2002. Clinical factors and clinical variation influencing the reproducibility of interocclusal recording methods. *British Dental Journal*. **192**(7), pp.395–400.

29. Fan, R. and Jin, X. 2014. Controllable Edge Feature Sharpening for Dental applications. *Computational and Mathematical Methods in Medicine*. (2014).

30. Fang, J.H., An, X., Jeong, S.M. and Choi, B.H. 2018. Digital intraoral scanning technique for edentulous jaws. *Journal of Prosthetic Dentistry*. **119**(5), pp.733–735.

31. Flügge, T. V., Schlager, S., Nelson, K., Nahles, S. and Metzger, M.C. 2013. Precision of intraoral digital dental impressions with iTero and extraoral digitization with the iTero and a model scanner. *American Journal of Orthodontics and Dentofacial Orthopedics*. **144**(3), pp.471–478.

32. Gan, N., Xiong, Y. and Jiao, T. 2016. Accuracy of intraoral digital impressions for whole upper jaws, including full dentitions and palatal soft tissues. *PLoS ONE*. **11**(7), pp.1–15.

33. Gintaute, A., Keeling, A.J., Osnes, C.A., Zitzmann, N.U., Ferrari, M. and Joda, T. 2020. Precision of maxillo-mandibular registration with intraoral scanners in vitro. *Journal of Prosthodontic Research*. **64**(2), pp.114–119.

34. Gkantidis, N., Dritsas, K., Ren, Y., Halazonetis, D. and Katsaros, C. 2020. An accurate and efficient method for occlusal tooth wear assessment using 3D digital dental models. *Scientific Reports*. **10**(1), pp.1–9.

35. Gruen, A. and Akca, D. 2005. Least squares 3D surface and curve matching. *ISPRS Journal of Photogrammetry and Remote Sensing*. **59**(3), pp.151–174.

36. Güth, J.F., Edelhoff, D., Schweiger, J. and Keul, C. 2016. A new method for the evaluation of the accuracy of full-arch digital impressions in vitro. *Clinical Oral Investigations*. **20**(7), pp.1487–1494.

37. Güth, J.F., Keul, C., Stimmelmayr, M., Beuer, F. and Edelhoff, D. 2013. Accuracy of digital models obtained by direct and indirect data capturing. *Clinical Oral Investigations*. **17**(4), pp.1201–1208.

38. Güth, J.F., Runkel, C., Beuer, F., Stimmelmayr, M., Edelhoff, D. and Keul, C. 2017. Accuracy of five intraoral scanners compared to indirect digitalization. *Clinical Oral Investigations*. **21**(5), pp.1445–1455.

39. Hyde, T.P., Craddock, H.L., Gray, J.C., Pavitt, S.H., Hulme, C., Godfrey, M., Fernandez, C., Navarro-Coy, N., Dillon, S., Wright, J., Brown, S., Dukanovic, G. and Brunton, P.A. 2014. A Randomised Controlled Trial of complete denture impression materials. *Journal of Dentistry*. **42**(8), pp.895–901.

40. Iwaki, Y., Wakabayashi, N. and Igarashi, Y. 2013. Dimensional accuracy of optical bite registration in single and multiple unit restorations. *Operative Dentistry*. **38**(3), pp.309–315.

41. Jablonski, R., Elneklawy, M., Osnes, C., Ferrari, M., Wu, J. and Keeling, A. 2019. Guided Scanning Improves Full Arch Precision of an Intra-Oral Scanner *In*: *IADR*. Vancouver.

42. Jaschouz, S. and Mehl, A. 2014. Reproducibility of habitual intercuspation in vivo. *Journal of Dentistry*. **42**(2), pp.210–218.

43. Kan, J.Y., Rungcharassaeng, K., Bohsali, K., Goodacre, C.J. and Lang, B.R. 1999. Clinical methods for evaluating implant framework fit. *The Journal of prosthetic dentistry*. **81**(1), pp.7–13.

44. Kazhdan, M., Bolitho, M. and Hoppe, H. 2006. Poisson surface reconstruction *In*: *Proceedings of the fourth Eurographics symposium on Geometry processing*.

45. Keeling, A. 2016. *Determination of the transverse horizontal axis and interocclusal registration using a novel optical technique*. University of Leeds.

46. Keeling, A. and Osnes, C. 2019. Leeds Digital Dentistry. *Website*. [Online]. Available from:

https://leedsdigitaldentistry.com/.

47. Keeling, A., Wu, J. and Ferrari, M. 2017. Confounding factors affecting the marginal quality of an intra-oral scan. *Journal of Dentistry*. **59**, pp.33–40.

48. Keul, C. and Güth, J.F. 2020. Accuracy of full-arch digital impressions: an in vitro and in vivo comparison. *Clinical Oral Investigations*. **24**(2), pp.735–745.

49. Kim, J., Kim, K., Kim, S., Kim, W., Kim, H. and Kim, J. 2015. Quantitative evaluation of common errors in digital impression obtained by using an LED blue light in-office CAD/CAM system. *Quientessenz International*. **46**(5).

50. Kim, T., Cho, Y., Kim, D., Chang, M. and Kim, Y. 2020. Tooth Segmentation of 3D Scan Data Using Generative Adversarial Networks. *Applied Sciences*. **10**(490), pp.1–17.

51. Kühle, A. 2003. Calibration Procedures for Atomic Force Microscopes. *Advanced Techniques for Assessment Surface Topography: Development of a Basis for 3D Surface Texture Standards 'Surfstand'*., pp.175–194.

52. Kuhr, F., Schmidt, A., Rehmann, P. and Wöstmann, B. 2016. A new method for assessing the accuracy of full arch impressions in patients. *Journal of Dentistry*. **55**, pp.68–74.

53. Le, M.T., Chen, L.C. and Lin, C.J. 2017. Reconstruction of accurate 3-D surfaces with sharp edges using digital structured light projection and multi-dimensional image fusion. *Optics and Lasers in Engineering*. **96**(April), pp.17–34.

54. Lee, J.H. 2017. Improved digital impressions of edentulous areas. *Journal of Prosthetic Dentistry*. **117**(3), pp.448–449.

55. Lepidi, L., Galli, M., Mastrangelo, F., Venezia, P., Joda, T., Wang, H.L. and Li, J. 2020. Virtual Articulators and Virtual Mounting Procedures: Where Do We Stand? *Journal of Prosthodontics*. **30**, pp.24–35.

56. Luhmann, T. and Wendt, K. 2000. Recommendations for an acceptance and verification test of optical 3-D measurement systems. *International Archives of Photogrammetry and …*. **XXXIII**, pp.493–500.

57. Luthardt, R.G., Kühmstedt, P. and Walter, M.H. 2003. A new method for the computer-aided evaluation of three-dimensional changes in gypsum materials. *Dental Materials*. **19**(1), pp.19–24.

58. Mangano, F.G., Admakin, O., Bonacina, M., Lerner, H., Rutkunas, V. and Mangano, C. 2020.

Trueness of 12 intraoral scanners in the full-arch implant impression: a comparative in vitro study. *BMC oral health*. **20**(1), p.263.

59. Marro, F., Jacquet, W., Martens, L., Keeling, A., Bartlett, D. and O'Toole, S. 2020. Quantifying increased rates of erosive tooth wear progression in the early permanent dentition. *Journal of Dentistry*. **93**(October 2019), p.103282.

60. Matejka, J. and Fitzmaurice, G. 2017. Same Stats , Different Graphs : Generating Datasets with Varied Appearance and Identical Statistics through Simulated Annealing *In*: *CHI 2017*. Denver, CO, USA.

61. van der Meer, W.J., Andriessen, F.S., Wismeijer, D., Ren, Y., Meer, W.J. Van Der, Andriessen, F.S., Wismeijer, D. and Ren, Y. 2012. Application of intra-oral dental scanners in the digital workflow of implantology. *PLoS ONE*. **7**(8), pp.1–8.

62. Mowery, W.E., Burns, C.L., Dickson, G. and Sweeney, W.T. 1958. Dimensional stability of denture base resins. *Journal of the American Dental Association (1939)*. **57**(3), pp.345–353.

63. Nagasawa, Y., Hibino, Y., Shigeta, H., Eda, Y., Matsumoto, S. and Nakajima, H. 2020. Characteristics of a new dental stone mixed by shaking. *Dental Materials Journal*. **39**(3), pp.355–366.

64. Nedelcu, R., Olsson, P., Nyström, I. and Thor, A. 2018. Finish line distinctness and accuracy in 7 intraoral scanners versus conventional impression: An in vitro descriptive comparison. *BMC Oral Health*. **18**(1), pp.1–11.

65. Van Noort, R. 2012. The future of dental devices is digital. *Dental Materials*. **28**(1), pp.3–12.

66. O'Toole, S., Osnes, C., Bartlett, D. and Keeling, A. 2019a. Investigation into the accuracy and measurement methods of sequential 3D dental scan alignment. *Dental Materials*. **35**(2019), pp.8–13.

67. O'Toole, S., Osnes, C., Bartlett, D. and Keeling, A. 2019b. Investigation into the validity of WearCompare, a purpose-built software to quantify erosive tooth wear progression. *Dental Materials*. **35**(10), pp.1408–1414.

68. Ockert-Eriksson, G., Eriksson, a, Lockowandt, P. and Eriksson, O. 2000. Materials for interocclusal records and their ability to reproduce a 3-dimensional jaw relationship. *The International journal of prosthodontics*. **13**(2), pp.152–8.

69. Osnes, C., Wu, J., Ferrari, M., Joda, T. and Keeling, A. 2020. Sources of Error in Maximum Intercuspation from Complete Dentate Full Arch Intraoral Scans in vitro. *International Journal of Computerized Dentistry*. (Accepted, in press).

70. Osnes, C., Wu, J., Venezia, P., Ferrari, M. and Keeling, A. 2019. Full arch precision of six intraoral scanners in vitro. *Journal of Prosthodontic Research*. **64**(64), pp.6–11.

71. Park, J.-M., Jeon, J. and Heo, S.-J. 2018. Accuracy comparison of buccal bite scans by five intra-oral scanners. *Journal of Dental Rehabilitation and Applied Science*. **34**(1), pp.17–31.

72. Patel, M. and Alani, A. 2015. Clinical issues in occlusion – Part II. *Singapore Dental Journal*. **36**, pp.2–11.

73. Patzelt, S.B.M. and Hack, G. 2020. Evaluation of the Accuracy of an Intraoral Scanner : An Update on the Omnicam Global Dentistry Evaluation of the Accuracy of an Intraoral Scanner : An Update on the Omnicam. . (November).

74. Pomerleau, F., Colas, F., Siegwart, R. and Magnenat, S. 2013. Comparing ICP variants on real-world data sets. *Autonomous Robots*. **34**(3), pp.133–148.

75. Rekow, E.D. 2020. Digital dentistry: The new state of the art — Is it disruptive or destructive? *Dental Materials*. **36**(1), pp.9–24.

76. Rekow, E.D., Erdman, A.G. and Speidel, T.M. 1987. Review of Developments in Computer-Based Systems To Image Teeth and Produce Dental Restorations. *Proceedings - Annual Symposium on Computer Applications in Medical Care*., pp.566–570.

77. Rhee, Y.K., Huh, Y.H., Cho, L.R. and Park, C.J. 2015. Comparison of intraoral scanning and conventional impression techniques using 3-Dimensional superimposition. *Journal of Advanced Prosthodontics*. **7**(6), pp.460–467.

78. Roggendorf, M.J., Kunzi, B., Ebert, J., Roggendorf, H.C., Frankenberger, R. and Reich, S.M. 2012. Seven-year clinical performance of CEREC-2 all-ceramic CAD/CAM restorations placed within deeply destroyed teeth. *Clinical Oral Investigations*. **16**(5), pp.1413–1424.

79. Rusu, R.B. and Cousins, S. 2011. 3D is here: Point Cloud Library (PCL) *In*: *IEEE International Conference on Robotics and Automation (ICRA)*.

80. Schmalz, C., Angelopoulou, E., Schmalz, C. and Hw, C.T.T. 2015. Camera calibration : active versus passive targets Camera calibration : active versus passive targets.

81. Segal, A. V, Haehnel, D. and Thrun, S. 2009. Generalized-ICP *In*: *Robotics: Science and*

Systems. Seattle, WA, USA.

82. Smith, A.R. 1995. A pixel is not a little square. *Pixar Technical Memo*. (6), pp.1–11.

83. Solaberrieta, E., Arias, A., Brizuela, A., Garikano, X. and Pradies, G. 2016. Determining the requirements, section quantity, and dimension of the virtual occlusal record. *Journal of Prosthetic Dentistry*. **115**(1), pp.52–56.

84. Solaberrieta, E., Garmendia, A., Brizuela, A., Otegi, J.R., Pradies, G. and Szentpétery, A. 2016. Intraoral Digital Impressions for Virtual Occlusal Records: Section Quantity and Dimensions. *BioMed Research International*. **2016**.

85. Solaberrieta, E., Otegi, J.R., Goicoechea, N., Brizuela, A. and Pradies, G. 2015. Comparison of a conventional and virtual occlusal record. *Journal of Prosthetic Dentistry*. **114**(1), pp.92–97.

86. Squier, R.S. 2004. Jaw relation records for fixed prosthodontics. *Dental Clinics of North America*. **48**(2), pp.471–486.

87. Straga, R.W. 2009. *Comparison of occlusal contacts on mounted models to contacts identified on digital 3D models using a new virtual alignment method*. University of Alberta.

88. Syrek, A., Reich, G., Ranftl, D., Klein, C., Cerny, B. and Brodesser, J. 2010. Clinical evaluation of all-ceramic crowns fabricated from intraoral digital impressions based on the principle of active wavefront sampling. *Journal of Dentistry*. **38**(7), pp.553–559.

89. Treesh, J.C., Liacouras, P.C., Taft, R.M., Brooks, D.I., Raiciulescu, S., Ellert, D.O., Grant, G.T. and Ye, L. 2018. Complete-arch accuracy of intraoral scanners. *The Journal of Prosthetic Dentistry*. **120**(3), pp.382–388.

90. Tryoe, G., Frydenberg, O. and Brill, N. 1962. An assessment of the tactile sensibility in human teeth: An evaluation of a quantitative method. *Acta Odontologica Scandinavica*.

91. Tsirogiannis, P., Reissmann, D.R. and Heydecke, G. 2016. Evaluation of the marginal fit of single-unit , complete-coverage ceramic restorations fabricated after digital and conventional impressions : A systematic review and meta-analysis. *The Journal of Prosthetic Dentistry*. **116**(3), pp.328-335.e2.

92. Ueda, Y., Okada, D., Ogura, R., Komada, W., Otake, S., Ikeda, M. and Miura, H. 2014. Evaluation of occlusal relationship reproducibility with CAD / CAM techniques. *Asian*

Pacific Journal of Dentistry. **14**, pp.35–40.

93. Utz, K.H., Müller, F., Lückerath, W., Fuß, E. and Koeck, B. 2002. Accuracy of check-bite registration and centric condylar position. *Journal of Oral Rehabilitation*. **29**(5), pp.458–466.

94. Vögtlin, C., Schulz, G., Jäger, K. and Müller, B. 2016. Comparing the accuracy of master models based on digital intra-oral scanners with conventional plaster casts. *Physics in Medicine*. **1**, pp.20–26.

95. Walls, A.W., Wassell, R.W. and Steele, J.G. 1991. A comparison of two methods for locating the intercuspal position (ICP) whilst mounting casts on an articulator. *Journal of Oral Rehabilitation*. **18**(1), pp.43–48.

96. Whitehouse, D.J. 1982. The parameter rash — is there a cure? *Wear*. **83**(1), pp.75–78.

97. Wismeijer, Daniel; Buser, Daniel; Chen, S. 2019. *Biological and Hardware Complications in Implant Dentistry* [Online]. Quintessence Publishing. Available from: https://play.google.com/store/books/details?id=5iu6DwAAQBAJ].

98. Wismeijer, D., Mans, R., van Genuchten, M. and Reijers, H.A. 2014. Patients' preferences when comparing analogue implant impressions using a polyether impression material versus digital impressions (Intraoral Scan) of dental implants. *Clinical Oral Implants Research*. **25**(10), pp.1113–1118.

99. Writer 2006. 3M Snaps Up Brontes Technologies. *The New York Times*. [Online]. Available from: https://dealbook.nytimes.com/2006/10/17/3m-snaps-up-brontes-technologies/.

100. Zampieri, M.H., Consani, R.L.X., Ferraz, M.M., Marchini, L. and dos Santos, F.M.B. 2014. Effect of Flask Types, Post-pressing Times and Base Regions on the Adaptation of Complete Dentures. *Journal of Research and Practice in Dentistry*. **2014**(2014).

101. van der Zel, J.M. 2020. Review & Reports Bites for the Computer : Thirty Years of Digitization in Dentistry. . **2**(3), pp.1–7.

102. Zhou, F., Cui, Y., Wang, Y., Liu, L. and Gao, H. 2013. Accurate and robust estimation of camera parameters using RANSAC. *Optics and Lasers in Engineering*. **51**(3), pp.197–212.

103. Zhou, Q.-Y., Park, J. and Koltun, V. 2018. Open3D: A Modern Library for 3D Data Processing.

104. Zorin, D., Schröder, P. and Sweldens, W. 1996. Interpolating Subdivision for meshes

with arbitrary topology *In*: *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques* [Online]. ACM, pp.189–192. Available from: http://portal.acm.org/citation.cfm?doid=237170.237254.

1)

8.5. Glossary of terms

Alignment	The process of finding a spatial transformation (e.g., scaling, rotation and translation) that aligns two three-dimensional data sets
IOS	Intraoral Scanner
Mesh	A collection of vertices, edges and faces that define the shape of a three-dimensional object in computer graphics
Occlusion	The contact relationships of maxillary and mandibular teeth as they bite together
PLY	Digital file format storing 3D data.
Precision	The closeness of agreement between test results
Registration	The process of finding a spatial transformation (e.g., scaling, rotation and translation) that aligns two three-dimensional data sets
Stereographic	Relating to a pair of lenses/sensors, which enables 3D vision.
STL	Digital file format storing 3D data.
Transform/Transformation	The act of rotating and/or translating an object in 3D space
Trueness	The closeness of agreement between the arithmetic mean of a large number of test results and the true or accepted reference value