

**Beyond the Archive: Using Excavation Photography to Assess Skeletal Preservation at
Khuzhir-Nuge XIV, an Early Bronze Age Cemetery on Lake Baikal**

by

Ruth Charlotte Urlacher

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Arts

Department of Anthropology
University of Alberta

© Ruth Charlotte Urlacher, 2017

ABSTRACT

The purpose of this research is to examine how accurately skeletal preservation can be estimated using in-situ photographs of human burials taken for documentary purposes during archaeological excavation. Skeletal completeness and fragmentation were estimated on an element-by-element basis using excavation photographs of 66 burials from the Bronze-Age cemetery Khuzhir-Nuge XIV, excavated by the Baikal-Hokkaido Archaeology Project between 1997 and 2001. Estimated preservation scores were then compared to preservation scores observed in the field during the original excavations, and the level of agreement between the two datasets was assessed in order to determine if estimated scores were accurate. The results indicate that when elements are clearly visible in excavation photographs, completeness can be estimated relatively accurately; however, estimations of fragmentation were found to be somewhat less reliable. The main impediment to accurate estimations of element completeness was caused by factors like body position and commingling, resulting in skeletal elements obscuring one another. It is suggested that taking photographs of each layer of a burial as skeletal elements are removed would remedy this problem. The under-utilisation of archaeological photography as a research resource, and advances in visual documentation of archaeological sites are also discussed.

ACKNOWLEDGEMENTS

I would first like to thank my thesis advisor, Dr. Andrzej Weber of the Department of Anthropology, both for his guidance and patience as I slowly worked out the kinks of my thesis, and for providing me with the opportunity to spend two fantastic summers excavating at the Hamanaka-II site on Rebun Island as a member of the Baikal-Hokkaido Archaeology Project team. I would also like to thank Dr. Lesley Harrington of the Department of Anthropology for being a supportive and friendly face whenever I showed up in her office looking for help, something which I wish I had done more often. Many thanks as well to the other members of my thesis committee for their time, and for their helpful and insightful comments and questions: Dr. Mark Gierl of the Department of Educational Psychology; and my committee chair, Dr. Ruth Gruhn of the Department of Anthropology.

This thesis was partially supported by a Joseph-Armand Bombardier Canada Graduate Scholarship awarded by the Social Science and Humanities Research Council, and a Queen Elizabeth II Graduate Scholarship awarded by the Faculty of Graduate Studies and Research. I gratefully acknowledge this support, and the support of the Department of Anthropology as well.

I would like to give a massive thank you to Andrea Hiob, BHAP secretary/centre of the universe, for always having time for me, and for the many illicit beverages that we most definitely did not drink in the BHAP office. To my office mates, Ben Osipov and Jacob Conner, thanks for always making and/or buying the coffee. RIP to all the coffee makers that lived and died in our office. Massive thanks as well to Victoria van der Haas and Zoë Eddy for their invaluable friendship and emotional/moral/spiritual support in the field in Japan – you ladies are the best!

Finally, thank you to my Mum and Step-dad, for their unwavering, if at times somewhat bemused support; and my siblings Derek and Jude for all the memes. All those dog pictures really kept me going. Last, but far from least, I want to thank my two best friends in the universe: my spiritual twin, and constant cheerleader, Anne Rudzinski; and the cornerstone of my existence as a functional human being, my husband, Jordan Urlacher. You are both so special and precious to me, I don't know how (or if) I would have gotten through this process without your support – I love you so much!

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGEMENTS	iii
LIST OF TABLES	v
LIST OF FIGURES	viii
INTRODUCTION	1
CHAPTER 1: THEORETICAL BACKGROUND	3
1.1 PHOTOGRAPHY IN ARCHAEOLOGY: A BRIEF HISTORY	3
1.2 THE UTILITY OF EXCAVATION PHOTOGRAPHS BEYOND THE ARCHIVE.....	9
1.3 RESEARCH OBJECTIVES	22
CHAPTER 2: REVIEW OF RESEARCH IN CIS-BAIKAL	27
2.1 GEOGRAPHIC OVERVIEW OF CIS-BAIKAL	27
2.2 MAJOR MIDDLE HOLOCENE MORTUARY COMPLEXES IN CIS-BAIKAL.....	28
2.3 PAST AND CURRENT RESEARCH.....	34
2.2.2 <i>Culture History of Cis-Baikal: 1950-1995</i>	34
2.2.3 <i>Recent Research in Cis-Baikal: 1995-2015</i>	37
2.4 KHUZHIR-NUGE XIV	47
CHAPTER 3: MATERIALS AND METHODS	55
3.1 MATERIALS: PHOTOGRAPHY AND DOCUMENTATION AT KNXIV	55
3.2 METHODS: DATA COLLECTION.....	57
3.3 STATISTICAL ANALYSES	65
3.3.1 <i>Data Considerations</i>	66
3.3.2 <i>Element Level Analysis</i>	69
3.3.3 <i>Burial Level Analysis</i>	74
CHAPTER 4: RESULTS AND DISCUSSION	79
4.1 PRESENTATION OF RESULTS	79
4.2 DISCUSSION.....	97
4.2.1 <i>Visibility</i>	97
4.2.2 <i>Agreement - Completeness</i>	102
4.2.3 <i>Agreement - Fragmentation</i>	106
4.2.4 <i>ANOVA & Multiple Regression</i>	110
4.3 SUMMARY OF RESULTS	121
4.4 RESEARCH QUESTIONS	127
CHAPTER 5: CONCLUSIONS AND FUTURE RESEARCH	133
5.1 CONCLUSIONS	133
5.2 FUTURE CONSIDERATIONS IN ARCHAEOLOGICAL PHOTOGRAPHY	135
BIBLIOGRAPHY	143
APPENDICES	156

LIST OF TABLES

2.1 Cultural Chronology as described by Weber in 1995	37
2.2 Cultural Chronology as described by Weber, Link, and Katzenberg in 2002	44
2.3 Cultural Chronology as described by Weber, McKenzie, and Buekens in 2010.....	45
2.4 Cultural Chronology as described by Weber et. al. in 2016	47
3.1 Agreement categories based on Bias scores	73
3.2 Minimum distribution thresholds for Limits of Agreement categorisation	73
4.1 Percentage Disagreement and Krippendorff's Alpha for all elements.....	79
4.2 Percentage Disagreement and Krippendorff's Alpha by Element Side	79
4.3 Percentage Disagreement and Krippendorff's Alpha by Element Type	79
4.4 Percentage Disagreement and Krippendorff's Alpha by Element Group	79
4.5 Percentage Disagreement and Krippendorff's Alpha values by Burial	80
4.6 Difference between Observed and Estimated Completeness scores by Element Type	81
4.7 Difference between Observed and Estimated Fragmentation Ratio scores by Element Type	82
4.8 Difference between Observed and Estimated Completeness scores by Element Group	83
4.9 Difference between Observed and Estimated Fragmentation Ratio scores by Element Group	84
4.10 Completeness Bias and percentages of elements within Limits of Agreement per burial, with bias agreement categories noted.....	85
4.11 Fragmentation Ratio Bias and percentages of elements within Limits of Agreement per burial, with bias agreement categories noted.....	86
4.12 Burial Agreement Categories for Completeness, based on bias calculated using all scored elements; bias calculated using only elements scored as 'Visible' by both observers; and bias (all elements) adjusted for Limits of Agreement percentages.....	87

4.13 Burial Agreement Categories for Fragmentation Ratio, based on bias calculated using all scored elements; bias calculated using only elements scored as 'Visible' by both observers; and bias (all elements) adjusted for Limits of Agreement percentages.....	88
4.14 Mean Observed and Estimated Completeness and Fragmentation Ratio scores by Element Type	89
4.15 Mean Observed and Estimated Completeness and Fragmentation Ratio scores by Element Group	89
4.16 Mean Observed and Estimated Completeness and Fragmentation Ratio scores for all elements scored, and using only elements scored as 'Visible' by both observers.....	89
4.17 Completeness and Fragmentation Bias by Element Type, calculated using all elements scored, and only elements scored as 'Visible' by both observers.....	90
4.18 Completeness and Fragmentation Bias by Element Group, calculated using all elements scored, and only elements scored as 'Visible' by both observers.....	90
4.19 Completeness and Fragmentation Bias for all elements scored, and only elements scored as 'Visible' by both observers	90
4.20 Percentages of Variance accounted for by categorical Independent Variables in one-way ANOVA tests	91
4.21 Percentage of Variance accounted for by continuous Independent Variables in single-regression tests	91
4.22 Multiple Regression Model 1 - Completeness Bias with all Independent Variables	92
4.23 Multiple Regression Model 2 - Fragmentation Bias with all Independent Variables	92
4.24 Multiple Regression Model 3a - Difference in RI scores with continuous Independent Variables only	93
4.25 Multiple Regression Model 3b - Difference in RI scores with all Independent Variables	93
4.26 Multiple Regression Model 4a - Difference in Completeness scores with continuous Independent Variables only.....	94
4.27 Multiple Regression Model 4b - Difference in Completeness scores with all Independent Variables	94

4.28 Multiple Regression Model 5a - Difference in Fragmentation Ratio scores with continuous Independent Variables only.....	95
4.29 Multiple Regression Model 5b - Difference in Fragmentation Ratio scores with all Independent Variables	95
4.30 Single Regression Model for predicting Fragmentation Ratios of right-sided elements from the Fragmentation Ratios of left-sided elements	96
4.31 Single Regression Model for predicting Completeness of right-sided elements from the Completeness of left-sided elements.....	96
4.32 Single Regression Model for predicting the Fragmentation Ratios of right-sided cranial elements from the Fragmentation Ratios of left-sided cranial elements.....	96
4.33 Single Regression Model for predicting the Completeness of right-sided cranial elements from the Completeness of left-sided cranial elements.....	96

LIST OF FIGURES

3.1 Cropped image scanned with Nikon Super Coolscan 4000-ED at 1100DPI.....	60
3.2 Cropped image scanned with Nikon Super Coolscan 4000-ED at 1100 DPI, enlarged after scanning	60
3.3 Cropped image scanned with Epson Perfection V700 at 700 DPI, enlarged during scanning	61
3.4 Bland-Altman plot for Burial 23 showing the distribution of Completeness and Fragmentation Ratio scores within the Limits of Agreement	72
4.1 Scatterplot with regression line showing the relationship between Krippendorf's alpha values and the difference in Observed vs. Estimated RI scores by Burial.....	99
4.2 Scatterplot with regression line showing the relationship between Observed RI scores and the difference in Observed vs. Estimated RI scores by Burial.....	100
4.3 Scatterplot with multiple regression lines showing interactions between Krippendorf's Alpha values, difference in RI scores, and categorical levels of the variable 'Number Interred', by Burial	101
4.4 Means Plot showing Average Completeness by Number of Angles	112
4.5 Means Plot showing Average Fragment Count by Number of Angles	112
4.6 Means Plot of Average Fragment Count by Number of Angles, with outliers removed.....	113

Introduction

This thesis examines the utility of archaeological photography in assessing the completeness and preservation of in-situ human remains photographed during the excavation of burials. Its objective is to assess if estimations of preservation made based on photographic documentation of burials are accurate enough to be useful for research purposes, given that situations often arise wherein access to the skeletal remains themselves may be limited or impossible. The materials and comparative dataset used in this study are drawn from the BHAP's archive of research on the Bronze Age cemetery Khuzhir-Nuge XIV, located on the western coast of Lake Baikal, Siberia; particularly, Angela Lieveise's work on human taphonomy and skeletal preservation.

Chapter One examines the background and context of this research, beginning with an examination of the historical interconnections between photography as an artistic discipline and archaeology in its modern, scientific form. Following this overview, the remainder of this chapter is comprised of a literature review discussing the ways in which archaeological photography can be utilised beyond a simple documentary archive, and a detailed description of the research questions this study seeks to answer through examining the photographic archive from Khuzhir-Nuge XIV.

Chapter Two presents an overview of the body of archaeological research concerning the cultural chronology of the Cis-Baikal region of Siberia, where Khuzhir-Nuge XIV is located. This chapter is intended to place Khuzhir-Nuge XIV in geographic, cultural, and scientific contexts; and as such features a brief description of the geographic particulars of the region, followed by an explanation of the major Middle-Holocene mortuary complexes, and a historical

overview of major research in the region since 1950. Finally, a detailed overview of research concerning Khuzhir-Nuge XIV itself is also presented.

Chapter Three describes the materials and methods used in this study. First, the technical details regarding the scanning and preparation of photographic materials are presented, followed by an explanation of how completeness and fragmentation were assessed; and a discussion of the statistical analyses conducted on the resulting data.

Chapter Four begins with the presentation of the results of the statistical analyses discussed in Chapter Three. These are followed by a detailed discussion of the results, which are then summarized and applied to the research questions that were presented in Chapter One.

Finally, Chapter Five presents a summary of the conclusions drawn from this research and its results, followed by a brief discussion of advances in the visual documentation of archaeological sites for consideration on future excavations.

Chapter One: Theoretical Background and Research Objectives

1.1 Photography in Archaeology: A Brief History

The value of photography as a means of visual documentation of archaeological sites and monuments was apparent to archaeologists from the moment the first cameras became publicly available. The first photographic process, invented in France by Louis Daguerre in 1839, used a *camera obscura* to capture light and fix an image on a silvered copper plate, producing a left-right reversed image. The use of these images, known as daguerreotypes, was enthusiastically recommended by French archaeologist Dominique François Arago that very same year in a report to the French Chamber of Deputies. Seen as a way of recording ancient inscriptions on Egyptian antiquities, Arago felt that the daguerreotype was superior in both the detail captured and the time taken to capture the image than traditional drawings or sketches (Allen, 1996; Bohrer 2005; Falconer 2010).

Two years later, in 1841, William Henry Fox Talbot invented the calotype process, a forerunner of later negative-positive processes, in which negative images captured on paper treated with silver-chloride were then contact-printed on new sheets of paper to obtain positive images. Despite the cumbersome equipment and complex developing processes required by early versions of the camera, the ability of photographs to capture images quickly and accurately led to its immediate and enthusiastic adoption by many archaeologists. Both the calotype and the daguerreotype were used in the field from as early as 1841 in the documentation of antiquities and monumental architecture in Egypt, India, South America, and the Near East (Allen, 1996; Dorrell, 1989; Falconer, 2010).

The timing of the advent of photography was fortuitous, as it provided a technological means by which archaeologists could achieve one of the major goals of archaeology as it sought to establish itself as a legitimate scientific discipline. In fact, it was only two years prior to the invention of the daguerreotype that the term ‘archaeology’ began to be used in its modern sense, referring to the "scientific study of the remains and monuments of the prehistoric past" (Bazin, 1989 in Bohrer, 2011: 27). However, although these technologies purported to provide an objective, scientific means of documenting archaeological finds, one of their major limitations lay in the fact that the images they produced could not be easily reproduced and shared. Daguerreotypes could only be printed on glass; and while calotypes could be printed on paper, there was no affordable way to mass-produce these prints until 1852, when Louis Desire Blanquart-Evrard started the first large-scale photographic printing house in France. Thus, many publications from this time period contain engravings or illustrations copied from photographs as opposed to the photographs themselves (Hallotte, 2007; Bohrer, 2011; Falconer, 2010).

By the late 1800s, with the introduction of dry-gelatin emulsion during the 1870s and celluloid film in 1889, images no longer needed to be developed immediately after being taken, a feature which made taking photographs in the field significantly easier, as it removed the need for a mobile darkroom (Hallotte, 2007). However, although photographic technology continued to improve throughout the 19th century and into the 20th, the technical difficulty of actually capturing images persisted. While the perception of the camera’s objectivity – its ability to capture a “true reproduction” of reality (Arago, 1839 in Bohrer, 2005: 183) – was a fundamental draw of photography for many archaeologists, others maintained that the medium was inadequate as a documentary tool. For example, British artist Frederick Catherwood, tasked with producing drawings of stone monuments as part of an archaeological expedition in the Yucatan

in 1843, complained that the daguerreotypes produced by his colleagues at the same time were fundamentally flawed, as the interplay of light and shadow on any given surface highlighted the details of certain areas whilst simultaneously obscuring that of others. Similarly, many photographers documenting architectural monuments in Egypt and the near East complained that the camera simply could not do justice to the grand scale of the structures they were attempting to record (Bohrer, 2011).

However, the idea that photographic technology allowed archaeologists to effectively and efficiently capture images that were objective reproductions of reality lent photography a certain authority as a scientific method of documenting exactly what was seen in the field. This authority stemmed from a sense that photography, as a mechanical and chemical means of capturing and producing images, removed any human elements from the process; instead "presenting images of unfiltered reality free of the embellishment and invention prone to the work of even the most practiced artist" (Bohrer, 2011: 28; Shanks, 1997; Parno, 2012).

Despite its apparent superiority over other methods of visual documentation, the early technological frustrations of archaeological photography were often compounded by anxieties surrounding the legitimacy of archaeology as a science, a situation which led to some archaeologists displaying a certain ambivalence towards photography even as it became a permanent fixture in the archaeological toolkit (Bohrer, 2005). For example, in the chapter on photography in his 1904 guide to archaeological excavation, *Methods and Aims in Archaeology*, W.M. Flinders Petrie, despite providing detailed instructions on the basics of archaeological photography, nevertheless reflects Catherwood's earlier opinion on the superiority of drawings, noting that while photographs are useful for confirming the accuracy of such artistic

documentation, drawings are “the more useful edition for most purposes.” (Flinders Petrie, 1904: 73).

Although published over 100 years ago, today’s principles of archaeological photography still hew closely to those laid out in Petrie’s guide. Indeed, there are now entire books – such as Peter Dorrell’s *Photography in Archaeology and Conservation* (1989) – devoted to the minutiae of camera settings, focal lengths, lighting, etc.; and how to capture the best survey, architectural, site, *in-situ*, and various other types of archaeological photograph. While anybody with a camera can take a picture, capturing ‘good’ images of archaeological finds that adequately reflect the scientific documentary goals of archaeology was, and remains, a technically complex procedure. In terms of technical complexity, archaeological photographer Aaron Levin (1986) noted that it is often easier to teach excavation methodology to a photographer than it is to teach photographic methodology to an archaeologist. In fact, even contemporary archaeological field manuals suggest that archaeological excavation teams should ideally be joined by a dedicated photography professional, or, at the very least, a “photography enthusiast,” rather than risking “a photographically untrained archaeologist [taking] unpublishable site photographs” (Tassie and Owens, 2010: 428).

However, while there is a large amount of literature concerning the best way to take different kinds of archaeological photographs, the question of what, exactly, to take photographs of is generally left up to the discretion of the archaeologists conducting the excavation. As Dorrell notes, “the definition of significance...varies widely from one type of dig to another,” (1989: 148). Similarly, while Dorell (1989) does note that *in situ* photographs of artifacts are important for establishing contexts of and relationships between objects, he states that these images are “usually a matter of common sense” (1989: 149).

With Petrie's (1904) guidelines, the documentary nature of archaeological photography led to certain conventions in the way these images were – and generally still are – set up, captured, and understood. Today, archaeological photographs are clinically composed with an almost forensic eye, maintaining an aesthetic of "pure, documentary evidence" (Bohrer, 2011: 141) divorced from any sense of artistry that might lead to accusations of subjectivity or misrepresentation on the part of the photographer, or the archaeologist reporting on the excavation. Human figures -- or even evidence of human activity -- feature very rarely in archaeological photographs. Artifacts and the units that contain them are brushed clean; and while certain details of an object or a trench wall may be highlighted with water, all traces of the excavation process that led up to the moment the image is captured are removed from the frame so as not to distract from the object itself (Parno, 2012; Dorrell, 1989).

However, as well as playing a fundamental role in how archaeologists communicate their findings to one another through academic channels, photography has also been the major way in which archaeology as a discipline engages wider audiences. As Bohrer (2011) points out, archaeology has always been a very public science, both in terms of popular interest in archaeological sites and finds, and in how archaeological expeditions are often funded by government bodies or similar sponsors. Thus, although the goals of documenting archaeological finds and communicating them to wider audiences overlap, they also result in a clear dichotomy between the types of photographs that are shared amongst archaeologists in academic publications and the images shared with the public in popular literature (Bohrer, 2011; Parno, 2012).

For example, the photographs taken by Howard Burton of the treasures discovered in the tomb of King Tutankhamun in Upper Egypt's Valley of the Kings in 1923 and subsequently

published in the *The Illustrated London News* created intense public interest, partly because of the way in which Burton photographed the artifacts. These images were, for the most part, the only way in which the public could experience these objects; and they illustrate what Hallote (2007: 129) refers to as the “artistic conundrum” of archaeological photography – namely, the fact that archaeologically valuable objects may not necessarily always be visually interesting or artistically valuable *subjects* to photograph. Although Burton’s photographs are beautiful, and no doubt played a large role in making King Tutankhamun’s tomb a globally famous site, they display a clear sense of artistry over documentation. Objects from the tomb are isolated and photographed from dramatic angles, while the scales, north arrows, and mug boards ever-present in modern excavation photography are absent.

British archaeologist Mortimer Wheeler considered photography to be a fundamentally important tool in engaging larger/public audiences with archaeology. While this sentiment may not be incorrect, it creates a mindset wherein archaeological photography is often thought of only in terms of its illustrative value as it relates to publications, exhibits, and other media directed towards public relations (Bohrer, 2011; Parno, 2012). Even in academic publications, images of excavations and artifacts tend to operate simply as visual footnotes, acting as proof that an excavation actually occurred, or that an object actually exists. While *in-situ* excavation images or photographs of artifacts may be used deliberately in academic publications to support a particular statement by the author, in a survey of the use of photographs in historical archaeology journals, Allen (1996) noted that visual documentary evidence (i.e., photographs and other images) was generally used very infrequently, and often as filler that was not referenced in the article text (Allen, 1996; Shanks, 1997; Parno, 2012).

In their documentary, illustrative function, archaeological photographs are a means to an end, acting as both witnesses to the archaeological process and as stand-ins for the material objects uncovered by that process. They allow archaeologists to, in a sense, quote reality by importing sites and objects into a text-based medium (Allen, 1996; Shanks, 1997). However, to understand archaeological photographs, particularly *in-situ* images and images of in-progress excavations, in purely visual terms may be short-sighted, as a number of studies show that these types of images can also act as valuable sources of archaeological data once an excavation is complete.

1.2 The Utility of Excavation Photographs Beyond the Archive

It is perhaps interesting to note that while Shanks (1997), Bohrer (2005, 2011), and others (Dorrell, 1989, Hodder, 1999, Roskams, 2001) take pains to disabuse readers of the assumption that archaeological photography is any more objective than written documentation, others note that written documentation can be just as subject to bias as photographs. Banks and Snortland (1995) observe that while written documents may be limited by the author's conscious and unconscious biases, photographs may contain unintentionally captured background details that can prove more revealing than their accompanying written documents. A similar point is also noted by Allen (1996), in her discussion of photographic intent as it relates to interpreting the resulting image, in which she notes that photographs "contain information that is not consciously intended and it is this information which renders the photograph useful for other purposes" (Allen, 1996: 21).

Although Shanks (1997) and Bohrer (2005, 2011) argue that archaeological photographs are primarily understood by archaeologists as a form of visual note-taking, with the resulting images functioning as *aide-memoires* more than anything else (Shanks, 1997), there are some

(although not many) examples of researchers using archaeological images as analytical tools. In the few studies of this type that have been published, photographs, both archaeological and historic, have been used in innovative ways to gather information that it would not have been possible to obtain from written documents.

For example, in their 1995 study of Northern Plains Indian tipi sites, Banks and Snortland used historic photographs of occupied tipi sites as an analytical tool to help them determine patterns in the placement and structural morphology of these camps. By looking for these patterns in photographs of extant sites, Banks and Snortland (1995) hoped to gain practical insight that could then be applied to archaeological tipi ring sites, as these features, though common throughout the Northern Plains, are notoriously difficult to interpret due to their general lack of artifacts, dateable material, or identifiable features or activity areas.

In this study, the authors studied images including historic photographs taken between the 1850s and 1950s, drawings and sketches, paintings, ledger art, and rock art, looking at certain features of the tipi camps which they recorded as specific variables in a database. These features, or variables, included: camp type (group, cluster, circular, or linear); associated features (hearths, shade structures, travois, drying racks, sweatlodges, etc.); seasonal variability (presence or absence of snow or vegetation; types of clothing worn in photographs); topographic setting (upland, valley, river, or woodland); orientation of doorways; method of securing tipi cover (pegs, logs, stones, sod, or a combination); and associated communal camp features (council, sun dance, women's, and medicine lodges; communal shades). After analyzing the content of the photographs, the authors then used statistical analysis to answer questions, particularly whether certain camp types are limited to specific topographic settings; and if specific features have typical, set positions in which they always appear within a camp.

Based on their analysis, Banks and Snortland (1995) were able to make predictions about where in an archaeological tipi site certain features are most likely to be found. For example, the photographs show that hearths are most likely to be found in front of a tipi, while evidence of travois and drying racks are most likely to occur along one of the sides. The authors also determined that smaller rings within a camp may be evidence of children's play tipis, or the homes of single elderly people, rather than evidence of older occupations, as had previously been thought (Banks and Snortland, 1995).

In a study by Zejdlik (2014), in-situ photographs of human burials were used to identify and re-associate skeletal elements from a collection that had become both commingled and disparate, with parts of the collection curated at least five different institutions. Within the known curated collection -- a skeletal assemblage excavated from the Aztalan archaeological site in southeastern Wisconsin -- there are 12 individual burials, 11 cremated individuals, and 3,000 pieces of human bone that were isolated and commingled at excavation (Zejdlik, 2014: 178). Despite the site being excavated a number of times since 1919, almost no written documentation regarding any excavation at the site has survived, leaving excavation photographs, a single site monograph, *Ancient Aztalan*, published by Samuel Barrett in 1933, and a site map drawn in 1999 by Lynne Goldstein (Zejdlik 2014: 179) as the only available resources from which the original contexts of skeletal remains could be determined.

In order to identify commingled elements, Zejdlik used a simple photo-matching system wherein individual elements were matched to in-situ photographs of bones or skeletons taken during the original excavation. Matches between photographs and bones were made macroscopically, using a 10x hand lens; and were based on a minimum of two matching characteristics, including morphology, pathological lesions, and fracture patterns. Despite loss

and damage to the written site records, the site photographs remained in excellent condition; and the quality of photographs printed from original glass-slide negatives was extremely high (Zejdlik, 2014).

Using this method, Zejdlik (2014) was able to identify which remains in the collection had become commingled during storage; and which were commingled, fragmentary, isolated, and disarticulated when excavated. By matching isolated skeletal elements to photographs of articulated in-situ remains, Zejdlik was able to re-articulate the bones of an individual within the collection known as the 'Aztalan Princess', a young adult female who had been given an unusually elaborate burial, and whose remains had become mixed up with the portion of the collection that was commingled when excavated. Using catalogue numbers as a guide, the Aztalan Princesses' skull and scapulae were located in a single drawer along with three sets of arm bones – photo-matching confirmed that two of the three sets of upper limbs also belonged to the Aztalan Princess. In addition, a number of elements stored in a bag with no catalogue number or recorded provenience were also identified through this technique as belonging to the Aztalan Princess, leading to the full re-articulation of her skeleton for the first time since 1933 (Zejdlik, 2014).

As well as demonstrating the importance of visual documentation of sites in terms of simple curation of human remains and other artefacts once an excavation is complete, Zejdlik's study also shows the important contributions that contextual and background information contained within these types of photographs can make to interpretations (or re-interpretations) of a site. In addition to re-articulating entire individuals, Zejdlik (2014) was also able to re-articulate an isolated hand that had been found in what was thought to be a baking pit, a placement which had led Barrett to speculate in *Ancient Aztalan* that the hand's location was

evidence of processing or cooking for cannibalistic purposes. However, once it was re-articulated, the hand showed no evidence of being subject to the type of processing or heat exposure that would indicate this type of activity, meaning that this interpretation is likely incorrect.

A second example of this issue in interpretation involves a photograph of commingled remains that were found beneath a potsherd, and originally identified as belonging to a sub-adult human. However, Zejdlik determined that the bones in the photograph displayed full epiphyseal fusion; therefore it was not possible for them to belong to a juvenile, which indicates that they were likely faunal remains. Although the bones from the photograph could not be located within the collection, the potsherd was present; and by using this as a scale, Zejdlik was able to confirm that the bones were far too small to be fully-fused, adult human remains, an observation which supports the conclusion that they are probably faunal, rather than the bones of a human child.

In addition, the background of excavation photographs can contain important taphonomic information, which can also prove valuable in terms of interpreting a site or a particular feature. In the case of Aztalan burial 6, the individual was re-articulated through photo-matching; and discovered to be missing the right radius and ulna, despite the skeleton being otherwise almost complete. Upon examining the in-situ photographs of burial 6, Zejdlik (2014) determined that a post had, at some point, been driven into the ground above the burial, destroying the forearm but leaving the rest of the body intact. According to Zejdlik, there are two possible interpretations of this action: either this individual, a female, was buried early enough in the site's occupation history that later occupants forgot or did not know that her body was buried in that location; or the post was deliberately placed there with the intent to destroy part of her body. Regardless of which interpretation is correct, the additional information provides a richer insight into

behaviours and their potential social contexts at the site, one which would not have been available without studying the site photographs.

Zejdlik notes two similar examples of studies in which the use of excavation photographs allowed researchers to make important discoveries about particular skeletal collections that had significant impacts on their interpretations of these assemblages. In one, a study of the human remains from the Angel Mounds site in southeastern Indiana by Marshall in 2011, researchers observed that juvenile remains tended to display better preservation than the remains of adults from the same site. By examining the excavation photographs from the site, the authors determined that this difference was due to the fact that adult remains seemed to have been left exposed in-situ for some time before being removed – probably in order for photographs to be taken of the spatial relationships between burials – while juvenile remains were immediately removed from the ground upon excavation. The result of being exposed to the sun was significantly poorer preservation of the adult remains, a factor which negatively impacted Marshall's ability to extract DNA and interpret the demographic profile of the site (Zejdlik, 2014).

In the second study, referenced by Zejdlik, Marden and Ortner in 2011 utilized a similar photo-matching technique to rectify a cataloguing error wherein elements belonging to a single individual from the Pueblo Bonito site in Chaco Canyon, New Mexico, had been given multiple catalogue numbers after excavation. Because of this error, and general commingling amongst the rest of the collection, four individuals had previously been observed as displaying pathological lesions indicating that they may have suffered from treponematosi s – a frequency which was considered to be relatively high. Using a combination of photo-matching and matching of bones by size, robusticity, taphonomic condition, and articular fit, Marden and Ortner were able to re-

associate all of the bones displaying these lesions to a single individual. This analysis reduced the number of individuals from Pueblo Bonito thought to have suffered from trepanemotosis from 4 to 1, resulting in a significant re-interpretation of health in the prehistoric population of this site (Marden and Ortner, 2011; Zejdlik, 2014).

Despite being a simple method, Zejdlik's photo-matching technique proved incredibly valuable both in the practical outcome of allowing a disorganised and commingled skeletal collection to be re-catalogued correctly, and in terms of the volume of interpretive information that the photographs themselves provided. The studies by Marshall (2011) and Marden and Ortner (2011) show similar results in terms of the practical and interpretive value of site photographs. Likewise, Banks and Snortland (1995) demonstrated the analytical value of historical photographs of occupied sites in the interpretation of the archaeological remains of sites belonging to similar cultures. Though this latter study was quite different in terms of the researcher's goals, and how it was carried out, it nevertheless shows similar efficacy in terms of the large amounts of useful, usable information that the authors were able to gather from studying photographs, whether of archaeological or historical sites. The apparent scarcity of studies like these shows that despite the fact that archaeological sites are often extensively documented in visual mediums, the resulting images seem to be a singularly underappreciated, underused resource.

While studies utilizing archaeological photographs in the ways discussed above seem to be relatively rare, the use of excavation photographs to investigate taphonomic factors affecting burials is not wholly unknown in archaeology. In recent years, Liv Nilsson-Stutz (1998, 2003, 2006, 2008) and others (Willis and Tayles, 2009; Harris and Tayles, 2012; Roksandic, 2002) have pioneered the use of in-situ photographs of skeletal remains within burial features in

interpreting mortuary ritual and behaviour at prehistoric burial sites. This methodology, a post-hoc application of the analytical principles behind the French archaeological method known as *anthropologie de terrain*, or archaeoethanatology, (Duday, 2009) relies on analysis of the spatial and anatomical relationships of skeletal elements present within a burial to reconstruct on theoretical grounds the original context of a grave at the moment of inhumation. By doing so, researchers can then infer information about the burial's original state, in particular, the original position of the body; whether or not any kind of burial container was used; and the presence of organic artifacts or objects that may have since decayed. This type of information is important, as it can provide insight about specific funerary gestures and the social role that these rites served within a society (Duday and Guillon, 2006; Nilsson-Stutz, 2003, Nilsson, 1998).

Archaeoethanatology, originally developed as part of a French archaeological technique pioneered by Henri Duday, has received very little attention in English publications (Nilsson-Stutz, 2003). The method can be broken down into two essential parts, excavation and interpretation, with the interpretive aspect relying on careful excavation and meticulous documentation of mortuary features in order to preserve the context as fully as possible. The approach to excavation is based on another French technique, *décapages*, wherein large areas known as *sols*, or living surfaces, are fully exposed, leaving all associated artifacts and materials left in place to be mapped and photographed. This technique supposedly allows spatial patterns that may be important to the interpretation of a site to be identified later, during the analysis stage (Nilsson-Stutz, 2003: 150).

In archaeoethanatology, mortuary features are exposed in a similar manner, keeping in mind what Duday and Guillon (2006) refer to as the "double preoccupation" (p. 120) of exposing bones as completely as possible, while also leaving them exactly in position. Once exposed,

bones are recorded, mapped, photographed, and recovered in stages, depending on how many individuals a grave contains, their bodily position, and, if applicable, the level of commingling. For example, in a grave containing a single individual in the extended supine position, there would be very few instances of bones lying on top of one another. Therefore, assuming that the only instance of overlap between bones would occur at the neck, with the skull obscuring the cervical vertebrae, the skull would be recorded, mapped, photographed as one layer; then, following the skull's recovery, the rest of the skeleton, including the now-visible cervical vertebrae, would be recorded, mapped, and photographed as a second, separate layer (Duday and Guillon, 2006).

As part of the recording process, the position of each bone and any artifacts associated with the burial are recorded in three dimensions prior to removal, with an additional depth measurements recorded for the soil underneath each bone after it is removed. Each skeletal element is identified, sided, and catalogued during removal; and as well as recording its positional data, information about the state of preservation, slant, orientation, and which side of the bone faces up is also documented. Duday and Guillon caution that all of this information must be documented in a logical and consistent manner, because "the dataset must be clear for any person who will use it later, even a long time after the digging if there is no possibility of consulting anyone who was in the field at the time" (Duday and Guillon, 2006: 124).

Although the level of detail that goes into recording burials with this method is extremely time-consuming, it serves the purpose of preserving the context of the burial as much as possible so that the archaeologist, during the interpretive stage, can reconstruct the burial at the time of excavation; and from there work towards reconstructing the burial at the time of inhumation in order to identify and interpret specific funerary gestures or rituals (Nilsson-Stutz, 2003). The

visual documentation of burials is very important in the interpretive stage because, as Nilsson-Stutz (1998, 2003) and others (Willis and Tayles 2009, Harris and Tayles 2012) have shown, it is possible to analyze photographs of burials using archaeoethanatology principles without necessarily having access to the type of written documentation that Duday and Guillon (2006) suggest recording.

All of the data recorded during excavation are put to use during the analytical phase of this method, when researchers combine excavation photographs, maps, and positional data to reconstruct burials exactly as they were during excavation. From that reconstruction, by applying certain principles relating to the process of decomposition and the actions of gravity on sediments and skeletal elements, researchers can work backwards to reconstruct the original burial context. The main assumption on which archaeoethanatology analysis of burials is based is the idea that temporary spaces open up around a buried body as its soft tissues decompose, and the presence of these spaces causes skeletal elements to shift out of their original positions to reach gravitational equilibrium (Willis and Tayles 2009; Duday and Guillon 2006). The amount of movement that occurs is dependent on the characteristics of the sediment in which the burial occurs, and how the body itself is buried.

For example, when a body is buried in dry, fine-grained sediment, this soil will continuously replace soft tissues as they decompose, a process which 'fixes' the skeletal elements in position. Evidence of this phenomenon can be seen in burials in which joint articulations that are known to break down relatively early in the decomposition process -- termed 'labile' joints by Duday (Duday and Guillon, 2006) -- such as the phalanges of the hands and feet, remain completely articulated (Willis and Tayles, 2009; Nilsson-Stutz, 2003, 2006). In contrast, if a body is buried inside a sturdy structure, like a coffin, which takes a long time to decay (if it

decays at all), the decomposing body is surrounded by empty space, a situation which causes the unsupported skeletal elements to disarticulate and fall away from the median axis of the body. According to Duday (Duday and Guillon, 2006), the continuous presence of empty space around the body during decomposition is indicated by the disarticulation of what he terms 'persistent' joint articulations. Persistent articulations take a long time to decay after death, and they are usually found between joints that are subject to high biomechanical loads during life that undergo little to no movement due to very tight joint capsules; for example, the sacro-iliac joints of the pelvis. (Willis and Tayles, 2009).

A physical anthropologist or osteoarchaeologist with sufficient understanding of the details of these processes should be able to identify bones that appear to have shifted out of their expected position, whether viewing the bones in-situ in the field, or in a photograph. This situation means that even if no useful data are recorded about a burial during its excavation, the interpretive aspects of the archaeoethanatology can still be applied after the fact if the visual documentation is sufficient, although this procedure is not the ideal application of the method (Nilsson-Stutz, 2003).

In her study of the burials at Zvejnieki, a mesolithic-neolithic cemetery in northern Latvia, Nilsson-Stutz (2006) applied the interpretive principles of archaeoethanatology to in-situ photographs of the burials in order to determine if they showed evidence of having been wrapped in burial shrouds at the time of inhumation. One of the major indicators of body wrapping is referred to as the *effet de parois*, or wall effect, in which, due to pressure on all sides from a tightly-wrapped covering, the body, once exposed, appears to be compressed by invisible walls. The shoulders and clavicles are often forced upwards, with the humeri pressed closely against the thoracic cage; and the anterior portions of the ribs will tend to fall towards the medial axis of

the body, in front of the vertebral column, making the ribcage appear very narrow. The legs will often be positioned close together, with the knees and feet touching and the patella still articulated. This bi-lateral compression is caused when the body decomposes inside of a tight covering which is gradually replaced by the surrounding sediment as it decays, a phenomenon which keeps the body in its original position. In contrast, when buried bodies are surrounded by open space, for example in a coffin burial, the skeletal elements will rotate outwards and fall away from the medial axis as the soft tissues of the body decay because there is nothing left to hold them in position (Nilsson-Stutz, 2003, 2006).

Although the Zvejnieki cemetery was not excavated using Duday's recommended field method, Nilsson-Stutz (2006) was still able to apply the interpretive principles in analyzing the burials, showing that the absence of the meticulous data recording that takes place when this method is employed in the field does not necessarily mean that a site or burial cannot then be analyzed using the interpretive aspects of the archaeoethanatology approach (Willis and Tayles, 2009). However, as Nilsson-Stutz (2006) points out, not having certain data available, particularly depth measurements for both skeletal elements and the grave itself, and information about the slope and walls of a grave, make it more difficult to reconstruct the original positions of bones, as this information is not something that can be easily discerned (if it can be at all) from a two-dimensional image.

Nevertheless, despite these limitations, and the fact that a number of burials had to be removed from analysis due to some photographs not being clear enough, Nilsson-Stutz (2006) was still able to develop new and potentially valuable insights into the burial practices at Zvejnieki. With this study, Nilsson-Stutz (2006) was able to provide taphonomic evidence to support the general hypothesis that bodies were often tightly wrapped before burial at many sites

in Northern Europe during the Mesolithic. She suggests that at Zvejnieki, bodies may have been wrapped in birch bark, as evidence for this practice has been documented at other Mesolithic cemeteries in the region, including Skateholm II, in Scania in southern Sweden, where individuals in graves in which the presence of wood was confirmed through chemical analysis were also identified as displaying the characteristics of having been wrapped before burial (Nilsson-Stutz, 2006; Nilsson, 1998).

In addition, through identifying instances in which bodies were wrapped before burial, Nilsson-Stutz (2006) was also able to draw a connection between the burials at Zvejnieki and Stone Age clay figurines from the nearby Åland Islands that depict limbless human figures with well-defined eyes, and cross-hatching across their torsos. According to Nilsson-Stutz, these figurines closely match certain burials at Zvejnieki in which, as well as showing evidence for wrappings, metal and amber rings were placed over the eyes of the dead. This custom could reflect the persistence of certain mortuary practices and beliefs from earlier periods into the Mesolithic, and could even reflect the existence of a “general circumpolar ideology spread throughout the Scandinavia and the Eastern Baltic during the Mesolithic” (Nilsson-Stutz, 2006: 231).

Harris and Tayles (2012), in their analysis of burial containers at Ban Non Wat, Thailand, used the same technique as Nilsson-Stutz (1998, 2003, 2006), performing archaeothanalogical analysis of in-situ burials using excavation photographs in order to identify changes in burial patterns over time. Their findings show a wide variety of burial contexts during the Bronze Age, with coffin and jar burials as well as burials in both loose and tight wrappings. Later, during the Iron Age, burial contexts become more uniform, with loose wrappings or coffins making up the majority of burial contexts; however, there is a greater diversity in grave goods. Harris and

Tayles (2012) suggest that this phenomenon may be related to the emergence of more complex socio-political systems, with the homogeneity of styles reflecting a sense of community identity while the diversity in grave goods possibly reflected emerging social hierarchies.

As with Nilsson-Stutz's (2006) study of the burials at Zvejnieki, Harris and Tayles (2012) note that in a number of cases their conclusions were tentative because the burial context was ambiguous. For example, they note that the use of open burial containers, such as hollowed-out tree trunks, is not distinguishable from photographs alone. Because the soil-infilling pattern would be the same in a burial with an open container as it would for a burial with no container – the body is in direct contact with the sediment in both cases – it is impossible to tell the difference without the additional context that might be provided by written documentation of associated findings within the burial (i.e., the presence of wood, indicating an open coffin).

Both Nilsson-Stutz (2003, 2006) and Harris and Tayles (2012) demonstrate that archaeothanatology can be applied post-hoc to sites that have already been excavated, as long as these sites are reasonably well-documented visually. Even without the additional context that in-depth documentation of burials and graves provides, systematic analysis of images of burials can still provide a wealth of useful information that may significantly add to the interpretation of a burial in terms of its wider social context. Indeed, all of the studies discussed above reveal the wealth of information that excavation photographs can contain, and the various different ways in which this information can be collected and applied, a fact which gives weight to the argument that archaeological photographs may be a resource that is significantly underused.

1.3 Research Objectives

Given the apparent wealth of contributions that archaeological photography can potentially make to the understanding and interpretation of archaeological sites, it is perhaps

pertinent to ask exactly how much information a photograph can provide. Is it possible to quantify how accurately a photograph of an in-situ burial or artifact reflects reality? It is this question which informs the study described in this thesis.

The main goal of this study is to answer questions about what type of information regarding skeletal preservation can be obtained from photographs of in-situ burials of human remains, and how reliable this information is. Determining the accuracy of data gathered using photographs as a resource necessarily requires that similar data, gathered in the field at the time the photographs were originally taken, already exist, so that there is something to compare the photographic data to. The materials in this study were provided by the Baikal Archaeology Project, which excavated a number of prehistoric hunter-gatherer cemeteries throughout the Cis-Baikal region of Siberia between 1997 and 2008, all of which were extensively documented both visually and in terms of written records.

The cemetery Khuzir-Nuge XIV (KNXIV) was chosen for this study, as the burials there were the subject of a dedicated taphonomic study by Angela Liverse (1999, 2007a, 2007b), who collected extensive data on the skeletal completeness and preservation of each burial, as it was excavated, during five field seasons between 1997 and 2001. The extensive photographic documentation of 74 burials at KNXIV, combined with the large amounts of quantitative data collected on-site by Liverse, resulted in a large, extremely well-documented sample with which a study of this kind could be performed.

The primary objective of Angela Liverse's original study at KNXIV (1999) was to examine the taphonomic factors responsible for skeletal condition (i.e. fragmentation, completeness, and articulation) and to identify which, if any, were responsible for the unusual level of variability in preservation that was noted at this cemetery. Liverse focused her research

on factors that influence human remains once they are buried, and which she classified into three major groups: intrinsic (i.e. skeletal), extrinsic natural, and extrinsic cultural.

Factors categorized as intrinsic are those that relate to individual skeletal characteristics, i.e. bone shape, size, and density; age at death; disease; trauma; nutritional status; and chemical composition. These factors were grouped into three intrinsic independent variables for use in the study: element type (long bones, short bones, flat bones, irregular bones, and mandibles), element side (left and right, for paired elements), and age at death (infant, child, adolescent, young adult, middle adult, old adult, and 'adult').

Extrinsic natural factors are the effects of both biotic and abiotic agents, while extrinsic cultural factors are effects related to human activity. Biotic agents includes both macroscopic flora and fauna (e.g. burrowing animals, scavenging carnivores), and microscopic organisms like bacteria and fungi. Abiotic factors are environmental elements such as water, temperature, sediment composition, oxygen, sunlight, and gravity. Factors that can be categorized as cultural include cause and manner of death (which may or may not be reflected as indications of trauma or disease), post-mortem treatments of the body, burial practices, and excavation and curation techniques. From these extrinsic factors, a combination of eight natural and cultural factors were identified as independent variables: element charring (present/absent), body position (supine or flexed), burial type (single, double, triple), burial depth, distance between burial and bedrock, presence of birch bark in grave, burial integrity (primary or secondary), and burial disturbance (disturbed or undisturbed).

Using these independent variables, Lieverse investigated four specific hypotheses regarding variable skeletal condition at KNXIV:

- 1) Archaeological age is responsible for variable skeletal condition

- 2) Intrinsic skeletal factors affect the variability of skeletal condition
- 3) Extrinsic natural factors affect the variability of skeletal condition
- 4) Variability can be accounted for by previously identified cultural factors, in particular the Glazkovo mortuary protocol.

Lieverse's fourth hypothesis included the examination of five variable dimensions of the Glazkovo mortuary protocol (see Chapter 2 for a discussion of mortuary complexes in Cis-Baikal): Body preparation (specifically with regard to cremation), body position, interment type (single vs multiple interments), paving stone arrangement, and type/amount of grave goods. Of these, only the first 3 factors were considered relevant to the study and included as extrinsic independent variables. In order to investigate the independent variables that she identified, Lieverse (1999) collected extensive data on skeletal preservation during the process of excavation throughout five seasons of fieldwork at KNXIV. It is this dataset that informs the current study, as it provided a sample to which data collected from the photographic archive could be compared for accuracy. A large portion of Lieverse's data collected in the field concerned the completeness and fragmentation of skeletal elements within each burial at KNXIV. For example, Lieverse recorded the presence or absence of individual or small groups of bones, and their levels of completeness and fragmentation (Lieverse, 1999). Using the photographs of these burials to collect the same kind of data allowed for a number of questions to be considered, the most basic of which concerned how similar is what a viewer sees in a photograph of an in-situ burial to what was seen and recorded in the field at the time the photograph was taken.

- Can the same elements that were recorded as being present during excavation be seen in the photographs?

- Are estimates of completeness made based on photographs similar to those made when viewing the bones during excavation?
- Are estimates of fragmentation made based on photographs similar to those made when viewing the bones during excavation?
- What factors affect whether or not elements can be seen or identified in photographs?
- What factors affect whether or not completeness and fragmentation estimates made from photographs accurately reflect those made directly in the field?

Assuming that excavation photographs provide a relatively accurate representation of reality, they nevertheless present a two-dimensional, static reflection, a feature which is a considerable disadvantage over viewing remains in person in the field. Acknowledging that there would always be certain skeletal elements that were impossible to view in a still image raised a second set of questions regarding whether or not it is reasonable to make assumptions about the preservation of skeletal elements that are partially obscured, based on the preservation of similar, nearby elements.

- Can the completeness of visible skeletal elements on one side of the body be used to predict or infer the completeness of those known to be present, but which are not visible, on the other side?
- Can the fragmentation level of visible skeletal elements on one side of the body be used to predict or infer the completeness of those known to be present, but which are not visible, on the other side?

Chapter Two: Review of Research in Cis-Baikal 1950-2015

2.1.1 Geographic Overview of Cis-Baikal

In terms of geographic area, the Cis-Baikal, as defined by Michael in 1958, encompasses the regions north and west of Lake Baikal including the Angara River Valley, reaching from the river's source at the western end of Lake Baikal to Ust'-Ilimsk, where the Angara meets the Ilim river some 990 km north; the Upper Lena River Valley, which follows the drainage of the Upper Lena river to where it meets the mouth of the Kirenga River; and the western coast of Lake Baikal itself, including Ol'khon Island. For analytical purposes, the research area covered by the Baikal-Hokkaido Archaeology Project is sub-divided into four geographic micro-regions, due to the variability of topography, geology, flora, and both terrestrial and aquatic fauna across the region. The four micro-regions are: Southwest (SW) Baikal, which extends from the western tip of the Lake to the mouth of the Selenga River; the Little Sea, which encompasses the middle of the northwest coast of Lake Baikal and includes Ol'khon Island; and the Angara and Upper Lena river valleys. Of the four micro-regions, the Little Sea, and the Angara and Upper Lena river valleys have been the subject of extensive archaeological exploration, while, with the exception of the cemetery Shamanka II, the SW Baikal micro-region remains relatively unknown (Weber, 2003; Weber and Bettinger, 2010).

The northwest and northern coasts of Lake Baikal are bounded by the Primorskii and Baikalskii mountain ranges; and the southwest edge of the region is marked by the Eastern Sayan Mountains, to the west of the southwest tip of the lake. The mountainous areas of the region display alpine characteristics, while the remainder of the Cis-Baikal is largely grassland steppe and southern boreal forest (taiga), with the river valleys acting as transition zones in which the taiga opens into the steppe landscapes with rolling hills that are characteristic of the Central Siberian Plateau. The climate of Cis-Baikal is continental, with long, cold winters averaging five

months duration; and short, dry summers that last for around two months. Climate specifics vary somewhat across the four micro-regions – the area immediately surrounding Lake Baikal is typically milder than elsewhere, due to the lake’s store of thermal energy. This factor results in cooler summers and warmer winters along the coast of the lake compared to along the Angara River, while the Upper Lena river valley tends to be colder in winter than the other micro-regions (Weber et. al., 2002; Weber and Bettinger, 2010).

2.2 Major Middle Holocene Mortuary Complexes in Cis-Baikal

Archaeological research on middle Holocene hunter-gatherers in Cis-Baikal is dominated by the examination of mortuary sites, as although the region has many cemeteries, habitation sites are few and far between. This situation has meant the cultures that existed within Cis-Baikal at various times throughout prehistory are generally identified primarily by their mortuary practices via the excavation of cemetery sites. While this has allowed for detailed analysis of groups that existed during periods of high cemetery use, periods during which formal cemeteries appear to have fallen out of common usage remain relatively unknown (Weber and Bettinger, 2010; Weber et. al., 2016).

Although the Baikal-Hokkaido Archaeology Project (hereafter referred to as BHAP) makes radiocarbon dating a cornerstone of its approach to the analysis of archaeological materials from Cis-Baikal, in typological terms BHAP’s culture-history model follows the Russian tradition of defining eras of prehistory by technology. The Russian tradition defines the Neolithic as being characterised by the appearance in the archaeological record of pottery, ground stone tools, and the bow and arrow, all of which are absent in the Mesolithic, which is defined by certain changes in the stone industry, including the increased use of microlithic technology. The Bronze Age is marked by the appearance of copper and bronze objects. Western

models, on the other hand, tend to mark the boundary between the Mesolithic and the Neolithic by the appearance of evidence suggesting agriculture and sedentism, with bronze objects and evidence of increasing social complexity and sedentism marking the Bronze Age (Weber et. al, 2016; Weber, 1995).

There are four major mortuary complexes that have been identified in Cis-Baikal: Kitoi, Isakovo, Serovo, and Glazkovo. Two other complexes, Khin' and Shivera, are also identified in Russian literature; however, they are much less well known, as burials that can be definitively identified as belonging to these cultures are much less common than those associated with the other four. Prior to the advent of radiocarbon dating techniques, the chronology of the culture sequence in Cis-Baikal was determined by typological criteria, a practice which led to a certain level of disagreement amongst Russian scholars as to where in the sequence certain mortuary complexes, primarily the Kitoi, should be placed (see section 2.3.1 for detail). Today, radiocarbon dating has definitely proved that the Kitoi complex dates to the Early Neolithic (EN), with the Serovo and Isakovo occurring contemporaneously in the Late Neolithic (LN), and the Glazkovo dating to the Early Bronze Age (EBA). Shivera materials are placed at the end of the sequence, following the Glazkovo complex, while Khin materials are thought to pre-date the Kitoi complex. Due to the relative scarcity of materials associated with the Khin and Shivera complexes, only the four major mortuary traditions are discussed here.

The majority of graves assigned to the Kitoi mortuary complex were found in the Angara micro-region at Lokomotiv, Kitoi, Galashikha, and Ust'-Belaia cemeteries; and in SW Baikal at Shamanka II. There are also a number of isolated Kitoi graves scattered along the Angara River. Kitoi burials typically feature heavy use of red ochre, with bodies sometimes covered from head-

to-toe in the pigment; and in fact this feature may be considered the primary defining characteristic of this mortuary tradition (Weber et. al., 2016).

Generally, body position in Kitoi burials is supine with the head oriented north, although there are examples of prone, flexed, and bundle burials that are considered Kitoi as well. Although single burials are typical, multiple interments are common; and when they occur, they tend to be stacked, indicating successive rather than simultaneous interment. In terms of grave goods, assemblages from Kitoi burials are the most variable in terms of both numbers and types of objects, ranging from no objects at all to hundreds. Bifacially shaped arrowheads and composite fishhooks are the most common objects found in Kitoi burial assemblages; and bone, stone, and antler tools, ornamental beads and pendants, zoomorphic *objets d'art*, and objects made from green nephrite also frequently occur. Pottery is rare in Kitoi graves; when it does occur, it is usually in the form of complete, net-impressed, mitre-shaped pots (Weber et. al., 2016).

Outside of the Angara and SW Baikal micro-regions, EN burials are more variable, with some 'classic' Kitoi traits occurring, but with others absent. For example, at Fofanovo cemetery in the Lower Selenga, there is heavy use of red ochre, but burials are flexed with the heads oriented to the southeast; while at Kurma XI in the Little Sea all the burials are oriented with the heads to the north, but there is much less use of red ochre and the burial assemblages appear typologically Mesolithic. Similarly, cemeteries in the Upper Lena micro-region also display high variability, with each EN cemetery appearing to differ from others in the region in some way, something which can make them difficult to definitively classify using strictly typological criteria (Weber et. al., 2016).

The LN mortuary traditions of the Serovo and Isakovo frequently co-occur, with cemeteries in the Angara containing both types of graves. In the Upper Lena there appears to be a distinct local variation of the Serovo tradition, and this phenomenon may also occur in the Little Sea as well. While Serovo graves appear throughout Cis-Baikal, Isakovo graves appear to be restricted to the Angara valley. The major feature that distinguishes Isakovo and Serovo burials is their orientation with regard to the Angara river: Graves containing Isakovo burials are oriented parallel to the river with the heads pointing downstream (usually north), while graves with Serovo burials are typically perpendicular to the river with the heads pointing away from it (Weber et. al., 2016).

Isakovo burials appear to be fairly uniform in terms of body position, grave architecture, and grave goods, exhibiting much less variation than both Kitoi and Serovo burials. Both Isakovo and Serovo graves feature stone structures on the surface, which are not seen in EN graves; however, the stone structures are confined to the surface in Isakovo graves, while they extend into the grave pit in Serovo graves. Single burials are the most common type in both the Isakovo and Serovo traditions, although multiple burials also occasionally occur in both complexes. Red ochre does not occur in Isakovo burials; and while it is common in Serovo burials, it seems to have been used much more judiciously than in Kitoi burials, occurring in small, isolated patches rather than covering the entire body (Weber et al., 2016).

In terms of grave goods, assemblages in Isakovo graves are much less diverse than those associated with Kitoi, with 20 – 25 categories of objects recorded for Isakovo graves vs. 60 – 65 for Kitoi graves. Approximately 70 percent of Isakovo burials contain mitre-shaped, net-impressed pottery, making this one of the most common grave goods. Other items that regularly occur in Isakovo assemblages include bone or antler points and shafted, double-sided tools, and

harpoons, while lithics, *objets d'art*, and fishing tackle are extremely rare, as they are in Serovo graves as well. Assemblages from Serovo graves are somewhat more variable, and tend to be larger and richer than Isakovo assemblages; however, they are still less diverse than Kitoi assemblages. As with the Isakovo, pottery is the most frequently found item in Serovo graves; however, Serovo pots are egg-shaped rather than mitre-shaped. In addition to pots, which are found in all Serovo graves in the Angara valley, large lithic biface spearheads are also extremely common. Stone, bone, and antler tools are also prevalent; and composite points found in Serovo graves differ morphologically from those found in Isakovo assemblages (Weber et. al., 2016).

In terms of distinct local variations of the Serovo tradition, many Serovo graves in the Upper Lena and Little Sea micro-regions feature burials that are wrapped in birch bark; and often also show evidence of fires being set in the grave pit after the body was placed inside it. Upper Lena Serovo burials also differ slightly in terms of their assemblages, with fishing tackle more common in burials here than in other micro-regions; and antler picks, which do not occur at all in assemblages from outside the Upper Lena. In addition, some of the pottery, ground stone tools, and composite knives found in Serovo graves on the Upper Lena are morphologically similar to those found in Isakovo graves rather than being distinct. However, these differences aside, general patterns of grave architecture, body position, red ochre use, and grave orientation with regard to bodies of water hold true throughout the entire Cis-Baikal region where Serovo graves are found.

The EBA Glazkovo complex is the only burial tradition that has so far been documented in all four of Cis-Baikal's micro-regions. Glazkovo burials display a wide range of variability across micro-regions; and while they can be distinguished from other traditions by the presence of copper or bronze objects in burial assemblages, the occurrence of these objects is quite rare. In

the absence of metal artifacts, Glazkovo burials are still distinguishable by their orientation relative to burials associated with other traditions. While Glazkovo burials are typically supine or flexed like the other mortuary traditions, seated burials have also been documented; and so far seem to appear exclusively in the Glazkovo tradition (Weber et. al., 2016).

In the Angara river valley, graves containing Glazkovo burials are typically oriented parallel to the river with the heads pointed upstream, a feature which distinguishes them from Isakovo burials in the same area, which have the heads oriented downstream. This positioning relative to the flow of the river also occurs in the Upper Lena river valley, where the heads again point upstream. In the Little Sea, the rule of orienting burials relative to bodies of water is slightly different, as Lake Baikal appears to be the point of reference rather than a river. Here, burials appear to be oriented along the long, SW/NE axis of the lake, with the heads typically pointing SW. In the South Baikal micro-region, the absence of a river and the complexity of Lake Baikal's coastline seem to nullify the water-referencing tradition. In this micro-region, Glazkovo graves have been documented in only two cemeteries: Shamanka II and Fofanovo, in which they are clearly distinguished from the more numerous Kitoi burials by their lack of red ochre, grave and burial orientation and grave good assemblage (Weber et. al., 2016).

While red ochre is generally absent, some Glazkovo burials in the Little Sea exhibit very heavy use of the pigment. A large number of EBA burials in the Little Sea are also extensively disturbed, and exhibit charring due to fires being set inside the grave pits. However, despite these regional variations, generally, in terms of body position, grave architecture, number interred, use of red ochre, and types of grave goods, Glazkovo burials bear much more similarity to Isakovo and Serovo burials than to those associated with the Kitoi. Interestingly, Kitoi graves are often found, undisturbed, side-by-side with Glazkovo graves. This placement suggests that the later

inhabitants of Cis-Baikal were aware of the locations of these older graves, indicating that they may have been marked in some way that was no longer apparent by the time they were excavated in the modern era.

2.3 Past And Current Research in Cis-Baikal

2.3.1 Culture History of Cis-Baikal, 1950 - 1995

The cultural chronology of Middle-Holocene Cis-Baikal has been a subject of debate since the 19th century, the intensity of which was only increased after A.P. Okladnikov published his synthesis of the region's culture history in the 1950s. Okladnikov's work was based on archaeological materials collected during the second half of the 19th century and during the 20th century prior to the beginning of the Second World War, much of which was collected by Okladnikov himself during his fieldwork in the 1930s. Okladnikov's model was largely based on burial data collected from approximately 300 Neolithic and Early Bronze Age graves; however, this material is no longer available for analysis, as its whereabouts are unknown (Weber, 1995; Weber, Link, and Katzenberg, 2002).

Fieldwork continued to be carried out around Lake Baikal throughout the 1950s, becoming more extensive during the 1960s and 70s as the construction of three hydroelectric dams on nearby rivers spurred archaeologists to salvage as much material as possible. Much of the material collected in the 1950s remains unpublished, with the exception of Fofanovo cemetery (Gerasimov and Chernykh, 1959; Konev, 1996). Weber (1995) suggests that the volume of material collected during this period may have actually hindered scholars in terms of theoretical reflection, stagnating progress in the analysis and interpretation of the region's prehistory. This stagnation is perhaps best reflected in the persistence of the notion of unilinear

cultural progression from simple to complex that characterises much of the Russian scholarship on Cis-Baikal's cultural chronology, despite evidence to the contrary.

As is common in Russian archaeology (as well as many other archaeological traditions) Okladnikov's materials were classified typologically; and his proposed chronological model suggested a unilinear progression of cultures in the region, with four developmental stages of increasing complexity. Okladnikov suggested that the Isakovo culture appeared first, followed by the Serovo, then the Kitoi, and finally, the Glazkovo. The model was comprehensive, with each culture defined by specific artifact forms, mortuary practices, social structures, technology, and subsistence strategies; and it dominated Baikal scholarship for decades (Weber, 1995).

Okladnikov's placement of the Kitoi culture towards the end of the sequence was a source of criticism, with M.M. Gerasimov arguing that this complex should have been placed at the beginning. Okladnikov's rationale for his placement of the Kitoi was based on the presence in burial assemblages of fishing gear, green nephrite, and fine bone tools, all of which suggested a level of cultural complexity similar to that of the Glazkovo culture, which was placed at the end of the sequence, based on the presence of copper and bronze objects in burial assemblages. However, Gerasimov argued that Okladnikov's sequence selectively ignored data that did not fit his model. According to Gerasimov, not only did the Isakovo and Serovo cultures bear more similarity to the Glazkovo in terms of material culture and burial rituals, but they were also much more alike in a biological sense (based on craniometric analyses) than were the Glazkovo and the Kitoi (Weber, 1995).

The placement of the Kitoi within Cis-Baikal's cultural chronology remained a major topic of debate amongst Baikal scholars through the 1960s and into the 1980s with a number of revisions to the chronology proposed based on pottery typology, as well as attempts to correlate

burial materials with materials from habitation sites. Early radiocarbon dating attempts were inconclusive, with dates originating from burials with questionable contexts and unclear cultural associations. It was not until 1989 that a radiocarbon dating study by Mamonova and Sulerzhitskii definitively confirmed that the Kitoi culture pre-dates the Isakovo and thus belongs at the beginning of the chronological sequence (Weber, 1995).

As well as confirming that the Kitoi culture was the oldest in the sequence, Mamonova and Sulerzhitskii's study also confirmed that all the burial complexes were in fact older than had previously been assumed, and that the Glazkovo period lasted for approximately two to three times longer than previous estimates. Mamonova and Sulerzhitskii dated the Kitoi complex to the sixth millennium BC, followed by the Isakovo from the end of the fifth to the beginning of the fourth millennium BC. The Serovo were thought to have overlapped with the Isakovo, appearing during the fourth millennium BC, while the Glazkovo overlapped with the Serovo, appearing at the end of the fourth millennium BC and lasting until the end of the third millennium BC. However, despite evidence to the contrary given the new placement of the Kitoi culture, Mamonova and Sulerzhitskii still presented the cultural chronology as representing a unilinear evolution of culture, in keeping with the dominant academic tradition (Weber, 1995).

It was not until 1995 that the suggestion was made, by A.W. Weber, that the unilinear model was not necessarily the most appropriate framework for interpreting Cis-Baikal's cultural chronology. In his review of previous research on the cultural chronology of Cis-Baikal, Weber (1995) demonstrated the utility of a multi-disciplinary approach, synthesizing a number of different lines of evidence – ceramic, funerary, radiocarbon, and osteological – to produce a revised chronology of the area that was based on what the evidence itself suggested rather than a pre-conceived idea of what such evidence *should* demonstrate.

According to Weber’s (1995) revised chronology, the region’s cultures were not continuous, as had previously been suggested. In fact, there was a distinct gap of approximately 700-years between the disappearance of the Kitoi culture and the appearance of the Serovo, wherein the use of formal cemeteries appears to have ceased, suggesting that the region may have become severely depopulated during that time. In addition, the groups on either side of this hiatus, which spans the Middle Neolithic, appear to be both culturally and biologically distinct. Weber suggested a number of possible scenarios to explain the discontinuity, calling for further investigation of existing archaeological materials as well as more fieldwork aimed specifically at answering questions surrounding this period. Indeed, answering questions relating to how and why this ‘biocultural’ change and discontinuity occurred is one of the primary research objectives of BHAP, whose work over the following twenty years is discussed in more detail later in the chapter.

Age	Culture	Date (cal. BP)
Late Mesolithic	Khin (?), Early Kitoi	8800 – 7800
Early Neolithic	Kitoi	7800 – 6900
<i>Unclear</i>	<i>Area depopulated?</i>	6900 – 6200
Middle Neolithic	Serovo	6200 – 5000
Chalcolithic/Early Bronze Age	Glazkovo	5400 – 3700

Table 2.1 Cultural Chronology as described by Weber in 1995.

2.3.2 Review of Recent Research in Cis-Baikal, 1995-2015

Breaking with traditional Russian scholarship on the regional middle Holocene prehistory, BHAP’s approach to interpreting the data from Cis-Baikal emphasizes the cyclical nature of culture change, examining both the similarities between the EN, LN, and EBA groups that existed in the region at different times as well as their differences. In their 2002 overview of BHAP’s research since the beginning of its existence as an official project, Weber et al. (2002) identified the three main questions that inform BHAP’s research agenda: (1) the origin and extent of middle Neolithic (MN, Table 2.2) biocultural discontinuity between the Kitoi and later

groups, (2) The origins of the Serovo and Glazkovo peoples, and (3) the validity of the separation of Serovo and Glazkovo into separate groups/cultures.

In attempting to answer these questions, BHAP has conducted studies in various areas of research, integrating multiple lines of data to inform their interpretation of the archaeological materials excavated within the Cis-Baikal region. In addition to analysing archaeological materials, BHAP has also conducted genetic (Mooder et. al. 2010), stable isotope (Katzenberg et. al., 2010; Weber et. al., 2011), climate (White and Bush, 2010), and osteological (Lieverse 2007a, 2007b, 2010) research as well as producing and analysing hundreds of radiocarbon dates. The bulk of BHAP's archaeological material upon which these studies are carried out comes from five major cemeteries that make up the core of BHAP's research focus: Khuzir-Nuge XIV, Kurma XI, Lokomotiv, Shamanka II, and Ust'Ida I. However, as well as these five 'main' cemeteries, there are a further 179 known cemeteries in the region. In total, the 184 cemeteries of the Cis-Baikal have yielded 1026 graves, and the remains of 1182 individuals. Taken together, this collection of mortuary assemblages represents a massive body of data, the analysis of which has, and continues to be, an incredibly large and complex task (Weber and Bettinger, 2010, Weber et. al., 2016).

Rather than following traditional analytical frameworks that assume that cultural change occurs in a relatively predictable, linear fashion, BHAP's approach to studying culture change in Cis-Baikal instead begins with accepting the premise that cultural change is dynamic, and often occurs rather suddenly. In traditional analytical frameworks, a gradual, linear progression towards greater cultural and technological complexity is often assumed; and such progressions tend to be defined in a purely typological manner. In reality, complexity fluctuates – a culture can become more *or less* complex over time, with social structures and technologies arising or

being lost depending on the circumstances under which a given culture exists. More contemporary approaches, like New Archaeology and interpretive archaeological approaches touch on this notion of culture changing in response to external circumstance; but they tend to take a very systematic view, understanding culture change as a kind of homeostatic response to fluctuations in the system into which a culture has been placed (Weber, Jordan, and Kato, 2013).

In the New Archaeology approach, cultural change is positioned as something that occurs in response to changes in the local environment, like shifting climate patterns or other ecological disruptions, while the roles of the internal social dynamics of a culture and human agency are downplayed. Meanwhile, in interpretive archaeological approaches, the effect of human agency and social practices in driving cultural change are prioritized at the expense of acknowledging the potential effects of a cultural group's interactions with the local climate and environment. BHAP argues that in order to fully understand the mechanism of cultural change, interpretations of cultural change and the systems and contexts within which cultures operate need to be less rigid. BHAP has thus drawn on approaches that utilise Neo-Darwinian theory to develop an analytical framework that emphasizes both the life-histories and social dynamics of individuals and their communities, as well as the reconstruction of an individual's biographical information and their interactions with their environment. This 'life history approach', or bioarchaeology of individual life histories, allows researchers, by aggregating the data collected from each individual, to identify long-term cultural transitions at the community level, and to explore population dynamics, interaction and migration patterns, subsistence, health, and demography (Weber, Jordan, and Kato, 2013).

Using this approach, BHAP has been able to identify some general correlations between cultural and environmental variables during different time periods in Cis-Baikal. The EN is

characterised by cultural heterogeneity amongst groups of Kitoi people, and poorer overall health than in later periods. These factors correlate with uneven distributions of both people and aquatic resources, but with heavy reliance on fishing for subsistence, which in turn seems to be related to heavier workloads, unequal distributions of labour, and more male travel. The LN and EBA, in contrast, exhibit even distributions of game resources and people; and the people appear to be much more culturally homogeneous and to rely on game hunting rather than fishing for subsistence. This lifeway correlates with better health, lower workloads, less male travel, and more equal distribution of labour amongst the sexes (Weber and Bettinger, 2010).

Although Cis-Baikal suffers from a relatively low number of dwelling sites in its archaeological record, the examination of mortuary sites in terms of the distribution, number, and size of cemeteries across Cis-Baikal allows archaeologists to estimate approximate population sizes, while osteobiographical studies provide clues to the demographic makeup of these populations in terms of age and sex distributions. In addition, within individual cemeteries, the assessment of mortuary variation in terms of the number and spatial arrangement of graves, and numbers of individuals interred in each one, provides clues to how individual communities may have been organised socially. Using these data, it is possible to track how population sizes and social organization may have changed over time. In Cis-Baikal, it appears that in the Angara micro-region population sizes remained relatively stable from the EN to the LN/EBA; however the social organization changed from a small number of large groups in the EN to smaller, but more numerous groups in the LN/EBA (Weber and Bettinger, 2010; Weber and Goriunova, 2013).

In terms of overall health, osteological studies have shown that the health of both EN and LN/EBA groups appears to have been relatively good overall, based on dental evidence and

examination of skeletal elements for evidence of injury and disease. Although neither EN nor LN/EBA populations appear to have been subject to excessive physical stress, EN groups tend to display higher incidence of enamel hypoplasia (EH), indicating that they were more likely to suffer from serious insults to their health during childhood than were groups in the LN/EBA. In terms of disease, both groups exhibit similar rates of osteoarthritis (OA); however, there are distinct differences in which joints are affected in which groups. In the EN, males display more OA in their vertebrae and knees than females, while there is no real difference between EBA males and females in the incidence of OA in these joints. EBA females show more vertebral degeneration than EN females, and EBA males show less degeneration of the knees than do EN males. This finding indicates that EBA females may have participated more in physical labour than females of the EN, indicating a more equitable distribution of labour (Weber and Bettinger, 2010; Lieverse 2007b).

As well as examining demographics and health in Middle Holocene Cis-Baikal, BHAP researchers have also utilised bio- and geochemical data to investigate dietary and migration patterns across the region over time (Haverkort et. al., 2010; Katzenberg et. al., 2010). Stable isotope evidence shows that EN groups relied heavily on aquatic resources, particularly freshwater fish; and that this reliance increased over time, possibly as a response to a decrease in the numbers of local terrestrial mammals. EN groups also exhibit more dietary variability across micro-regions, likely due to differences in the abundance and species diversity of riverine fishes in different micro-regions. In the LN/EBA, evidence indicates that there was much less dietary variability throughout Cis-Baikal; and higher reliance on the consumption of terrestrial mammals, although populations of this period also appear to have moved towards greater reliance on aquatic resources over time. However, the LN/EBA is also characterized by varying

complexity of dietary signatures across micro-regions, a feature which seems to be related to the consumption of Baikal seals. Some LN/EBA groups included Baikal seals in their diet while others did not, and this distinction may have been used as a social identifier when these individuals migrated between regions. The details of this topic are beyond the scope of this review, however (c.f., Weber et. al., 2011; Weber, Jordan, and Kato, 2013; Weber et. al., 2016).

As well as demographic, social, and dietary differences, genetic studies (Mooder et al, 2010; Schurr et al, 2010) have shown that the EN population of Cis-Baikal is genetically distinct from the later LN/EBA populations, a conclusion which is consistent with earlier craniometric studies. Populations from the LN also exhibit some genetic differences from later EBA groups; however, these differences are much smaller than the difference between either of these two groups and the EN population. Chromosomal analyses indicate that the EN population may have descended from Western and Northern Eurasian groups, while the LN/EBA population bears more genetic similarity to later, modern indigenous Siberian groups (Weber and Bettinger 2010; Mooder et al, 2010; Schurr et al, 2010).

Although preliminary genetic studies indicate that the EN population and LN/EBA populations are probably two genetically as well as culturally distinct groups, the issue of separation or continuity between the LN Isakovo and Serovo, and the EBA Glazkovo is less clear. At various times, Isakovo and Serovo, and Isakovo, Serovo, and Glazkovo have been treated as single units in Cis-Baikal's culture chronology, due to the number of similarities these mortuary complexes have been seen to exhibit in terms of grave architecture, body position and grave orientation, and burial assemblages (Tables 2.2 and 2.3). For example, Glazkovo and Isakovo burials are both oriented parallel to the Angara River, but with the heads pointing in opposite directions, a difference which could indicate some form of cultural affiliation or

continuity between these groups. Likewise, aside from grave orientation, which is markedly different, Isakovo and Serovo burials in the Angara valley are otherwise quite similar in terms of burial positions and grave goods, a similarity which could again indicate some kind of cultural affiliation (Weber, Link, and Katzenberg, 2002; Weber and Bettinger, 2010).

Changes in the understanding of Cis-Baikal's cultural chronology are reflected in Tables 2.1 – 2.4, which show the major revisions that have occurred in the established sequence as more radiocarbon dates have been collected and analysed over time. As it forms a major pillar of its analytical approach, BHAP's strategy for radiocarbon dating has been subject to many updates and refinements over the years, resulting in changes to the chronological boundaries. Most recently, a process has been developed by BHAP researchers to identify and correct potential inaccuracies in their radiocarbon dates that are the result of a Freshwater Reservoir Effect (FRE), which operates similarly to the much better documented Marine Reservoir Effect (Nomokonova et. al., 2013; Weber et. al., 2016).

Table 2.1 (after Weber, 1995) is based on a series of radiocarbon dates that, although quite large, was of questionable quality, and Weber's (1995) in-depth analysis of available typological evidence from mortuary and habitation sites described in Russian literature. This model combines the Isakovo and Serovo into a single analytical unit and places it in the MN, while the hiatus is not associated with a specific cultural period, and the EBA overlaps with the Chalcolithic. Table 2.2 (after Weber, Link, and Katzenberg, 2002) retains very similar date ranges to Table 2.1, but is based on the analysis of a large series of radiocarbon dates collected and analysed by BHAP in the intervening years. In this model, the age and culture associations have shifted somewhat, so that the hiatus is now associated with the MN, and the term Chalcolithic is no longer used – instead the LN and EBA are separated into two distinct ages. In

addition, the Isakovo, Serovo, and Glazkovo are all identified as distinct entities, but are grouped together into a single analytical unit that spans the LN and the BA.

Age	Culture	Date (cal. BP)
Late Mesolithic	Early Kitoi	8800 – 7800
Early Neolithic	Late Kitoi	7800 – 6900
Middle Neolithic	<i>Hiatus</i>	6900 – 6200
Late Neolithic	Early Glazkovo/Serovo/Isakovo	6200 – 5400/5000
Bronze Age	Late Glazkovo/Serovo/Isakovo	5400/5000 – 3000

Table 2.2 Cultural Chronology as described by Weber, Link, and Katzenberg in 2002.

Table 2.3 reflects the increased complexity of BHAP’s radiocarbon studies between 2002 and 2010, including the addition of greater numbers of radiocarbon dates to the BHAP database, a focus on dates obtained from samples with high collagen yields (higher collagen yields are considered to result in more reliable dates), and attempts to make finer-grained distinctions between micro-regions in terms of their associated temporal boundaries. The study that produced the time ranges in Table 2.3 sorted radiocarbon dates from three of the four micro-regions into separate analytical units in order to examine differences in how long certain cultures persisted in different areas of the Cis-Baikal region. Compared to the model illustrated by Table 2.2., the LM is shown to be approximately 200 years shorter, while the EN both begins and ends slightly earlier. The MN hiatus in this model is considerably longer than in earlier models, starting up to 300 years earlier and ending 400 years later, although the authors note that a flattening of the calibration curve for this time period complicates the interpretation of radiocarbon data dating to this era (Weber et. al., 2010).

In addition, the Glazkovo complex has been separated from the Isakovo and Serovo, and placed by itself in the EBA, which starts at approximately the same time as in previous models but now ends 600 – 1000 years earlier, while the LN starts slightly later than in previous models but ends at around the same time. While there is some overlap in the radiocarbon date sequence

between LN Serovo and Isakovo burials and EBA Glazkovo burials, the calibrated dates from the 2010 study showed that the majority of Glazkovo burials were in fact considerably younger than those identified as Isakovo or Serovo. Given this revised timeline, and the typological differences noted in the mortuary complexes, the group that earlier had formally been referred to as the Isakovo/Serovo/Glazkovo complex was therefore separated into two distinct groups that existed in different periods of time (Weber et. al., 2010).

Age	Culture	Dates (Cal. BP)		
		<i>Upper Lena</i>	<i>Angara</i>	<i>Little Sea</i>
Late Meso.	Khin (?)	8800 – 8000	8800 – 8000	8800 – 8000
Early Neo.	Kitoi, Others (?)	8000 – 7200	8000 – 7000/6800	8000 – 7200
Middle Neo.	<i>Hiatus</i>	7200 – 6000/5800	7000/6800 – 5800	7000/6800 – 5800
Late Neo.	Serovo, Isakovo	6000/5800 – 5200/5000	5800 – 5200/5000	6000/5800 – 5200/5000
Early Bronze	Glazkovo	5200/5000 – 3400	5200/5000 – 4000	5200/5000 – 4000

Table 2.3 Cultural Chronology as described by Weber, McKenzie and Buekens in 2010.

The most recent update to the model, presented in Table 2.4 (after Weber et. al., 2016) is the result of two new methods that have been integrated into BHAP’s analysis of radiocarbon data. First, a number of dates ($n = 42$) from the five major BHAP cemeteries were adjusted using a linear regression equation in order to correct for the FRE identified by Nomokonova et. al. in their 2013 study of the Little Sea habitation site, Sagan-Zaba II. The second new approach is the integration of Bayesian statistical methods into the modelling of temporal boundaries between groups of radiocarbon dates. In this case, only dates to which the FRE correction had been applied were used in the models to generate temporal ranges known as highest posterior distribution (HPD) intervals using OxCal (see Weber et. al. 2016 for technical details). The date ranges produced using this approach suggest that entire chronological sequence in Cis-Baikal is a few hundred years younger than previously thought, and result in some significant changes to the lengths of some of the time periods in the model.

Previously, there was a small number of burials ($n = 10$) that were considered to be somewhat ambiguous in terms of whether or not they were Kitoi and if they truly belonged in the LM or in the EN. According to the new model, it is highly likely that these burials date to the LM, well before the Kitoi period, however the cemeteries where these burials were located were likely used into the EN period. The re-assignment of these burials from the EN to the LM resulted in the revised date range for the EN being significantly different to that of previous models. Rather than enduring for 1000 – 1200 years, it appears that the EN lasted for approximately 500 years only, starting much later than previously thought and ending slightly earlier. The LN also starts later than previously assumed, and ends later as well, although its length remains similar to previous models at approximately 1000 years. However, the authors of this study (Weber et. al., 2016) note that these boundaries, particularly of the beginning of LN, may yet be subject to further revision, as dating of materials continues and the FRE correction process is refined further. As with the EN, the EBA in the new model starts much later and is considerably shorter, spanning approximately 900 years, making this period up to 800 years shorter than previously believed, depending on the micro-region (Weber et. al., 2016).

The effect of the shifting temporal boundaries of the EN and the LN also means that the MN hiatus now appears to be much longer than in previous models. Due to the lack of data for this period, the boundaries of the MN must necessarily be defined by those of the periods on either side of it, which in this model means that the hiatus is approximately 1400 years long, compared to 1000 – 1200 years in previous models. Despite the huge body of work generated by BHAP about the Cis-Baikal region over the past 20 years, the MN hiatus remains something of a mystery, especially in terms of how the transition from the EN to the MN occurred, and what type of socio-economic system existed during this period. Although it is generally assumed that

the population of Cis-Baikal was probably much smaller during the MN than at other times, it is unlikely that the region was completely depopulated. However, the lack of formal cemeteries during this period suggests that the groups that remained in the region were likely less socially complex than the EN and LN/EBA populations. If the formal cemeteries of other time periods represent eras of increased social complexity as a result of large, sedentary populations settling in specific areas for extended periods of time, it is possible that the MN was a period during which these trends reversed and the population dispersed across the region into smaller, nomadic groups. When, or indeed if, an archaeological identity for the MN is discovered, it can be expected that it will be quite distinct from those of the EN and LN/EBA, and that it will be marked by evidence of major cultural change (Weber et. al., 2016; Weber et. al., 2005; Weber, Link, and Katzenberg, 2002).

Age	Culture	Dates (mean HPD* cal. BP)
Late Mesolithic	Khin (?)	8277±176 – 7503±14
Early Neolithic	Kitoi	7503±14 – 7027±33
Middle Neolithic	<i>Hiatus</i>	7027±33 – 5571±88
Late Neolithic	Serovo, Isakovo	5571±88 – 4597±76
Early Bronze Age	Glazkovo	4597±76 – 3726±34

Table 2.4 Cultural Chronology as described by Weber et. al. in 2016, reflecting the use of Bayesian methods to calibrate radiocarbon date ranges.

2.4 Khuzhir-Nuge XIV

KNXIV is located approximately 3 km southwest of the mouth of the Sarma River in the Little Sea micro-region. The cemetery is located on the slope of a hill in an open, semi-arid grassy plain on the west coast of the Little Sea, near the southern end of Ol'Khon Island. It lies between two large outcroppings of bedrock, and spans an area of about 200 m northeast-southwest and 30 m northwest-southeast. There are 79 graves at KNXIV, containing the remains of 89 individuals, making it the largest documented LN/EBA cemetery both in the Little Sea micro-region, and in the Cis-Baikal region as a whole. In comparison, the next

largest cemeteries of the same age in the area, the nearby Uliarba and Sarmanskii Mys, consist of around 30 graves each (Weber and Goriunova, 2008).

The site was discovered during an archaeological survey conducted by Irkutsk State University in 1991. This survey revealed the presence of at least 50 graves at the site, which were identified at the surface level by the presence of slightly elevated clusters of flat rocks. A test excavation was conducted in 1993, during which five of these graves were excavated; with the materials contained within them suggesting that the cemetery was associated with the Glazkovo tradition, dating the site to the Early Bronze Age. This dating was later confirmed via radiocarbon analysis. A second survey conducted prior to the beginning of BHAP excavations at KNXIV in 1997 revealed a further 10-20 potential graves. By the completion of fieldwork in 2001 a total of 79 graves – including the five excavated in 1993 – had been excavated.

The graves in KNXIV are oblong pits filled with loamy sand and rocks, and they are generally no more than approximately 30 – 60 cm deep due to the proximity of bedrock to the surface (the bottoms of graves often rest directly on the bedrock). Like most EBA graves, the graves at KNXIV were marked with stone cairns made up of slabs of rock of varying sizes covering the entire surface of the grave. Of the 79 graves excavated at KNXIV, 70 were single burials, with 7 double interments, and 2 triple interments, giving a total of 89 identifiable individuals. In most of the graves containing multiple burials, the individuals were placed side-by-side, with the exception of two cases, grave no. 58 and grave no. 59, in which the individuals were stacked (Weber and Goriunova, 2008). Many of the graves were disturbed, and preservation of the burials was generally poor. Only 33 individuals were well-preserved enough for sex to be determined with any degree of certainty, and only 20 individuals had all 3 molars

present, plus an intact femur, for use in a strontium sampling study (Weber and Goriunova, 2013).

As the largest cemetery in the Cis-Baikal region, KNXIV is one of BHAP's core analytical units and thus has a large body of research associated with it, including osteoarchaeological analyses of bone preservation, health, and demography (Lieverse, 2007a; 2007b); spatial analyses of burial patterns (Weber and Goriunova, 2013); and mortuary variability (McKenzie, Weber, and Goriunova, 2008; Weitzel and Weber, 2008); analysis of grave disturbance patterns (Robertson, Weber, and Drouin, 2008); biogeochemical studies of dietary and migration patterns (Katzenberg et. al., 2010; Weber et. al., 2011,); and sedimentary analyses (Dlussky et. al, 2008), amongst others.

In one of these studies, Weber et al. (2005) hoped to determine with relative precision the likely duration of cemetery use, by utilising Bayesian modeling to analyse a set of 87 radiocarbon dates representing 79 individuals from 70 graves at KNXIV. However, the later discovery of the FRE (Nomokonova et. al., 2013) and the subsequent revision of the dates associated with the cultural chronology for the region as a whole (Weber et. al., 2016) means that the dates obtained from this study are now assumed to be incorrect. BHAP has submitted new radiocarbon samples from KNXIV to Oxcal for reanalysis, and these dates will subsequently be analysed using the revised Bayesian approach described in Weber et al's 2016 study. Therefore, although the dates for which KNXIV was assumed to be in use following the 2005 study are available, they are not presented here as they are now assumed to be incorrect; and will almost certainly be revised in future publications (Weber, 2016, personal communication).

A number of studies were conducted on the skeletal remains from KNXIV, including analyses of skeletal preservation as well as osteobiographical studies of the age, sex, and health

of each individual (Lieverse 1999, 2007a, 2007b). Skeletal preservation at KNXIV was variable, though in general it was rather poor, especially when compared to skeletal remains from other Cis-Baikal cemeteries like Ust'-Ida I and Shamanka II. Although remains of 89 individuals were recovered from KNXIV in total, osteobiographical studies were conducted only on the 84 individuals excavated by BHAP between 1997 and 2001; and of these 84, two could not be assessed for sex or age due to their extremely poor preservation. Of the remaining 81 individuals, 17 were subadults aged 15 or younger and thus could not be assessed for sex, as the secondary sex characteristics used for sex determination do not develop until late adolescence. For the remaining 64 individuals aged 15 and older, sex was indeterminate in 31 cases, while of the 55 individuals determined to aged 20 or older, skeletal preservation was too poor in 17 cases for a more specific age category to be assigned (Lieverse, 2007b).

Angela Lieverse's (1999, 2007a) studies of the taphonomic effects of various aspects of mortuary protocol at KNXIV found that completeness and fragmentation of skeletal elements were highly correlated with burial depth, burial disturbance and integrity, and the number and size of paving stones placed on top of the grave. Deeper graves with more protection (i.e., stone cairns on the surface of the grave), primary burials, and burials that were undisturbed tended to display higher levels of completeness and less fragmentation. Disturbed burials, secondary burials, and burials that displayed evidence of the use of fire as part of the burial ritual displayed high levels of fragmentation and disarticulation, and had higher proportions of missing elements.

As well as these extrinsic factors, the intrinsic qualities of skeletal elements themselves were also shown to affect preservation. Large, dense, regularly-shaped bones were generally more complete and less fragmented than smaller, irregularly shaped bones. This observation applies both in terms of absolute size of an element, and in terms of overall body size, with

smaller-bodied individuals (children, older adults) displaying lower preservation than those with larger bodies (adolescents, young and middle adults). Preservation was generally higher for individuals identified as male than it was for females, a finding which may also be a function of body size and bone density, both of which are typically higher in males than in females.

Although the majority of individuals for which sex could be determined were male, binomial statistical tests indicated a high likelihood that the male-female sex ratio of the population was probably fairly even; therefore it is likely that many of the individuals for which sex could not be determined due to poor preservation were female (Lieverse, 1999, 2007a, 2007b).

Amongst the children at KNXIV (individuals aged less than 12 years old), the sharpest increase in mortality was observed for children aged between two and four; however, with one exception there were no children under the age of two buried at KNXIV, and the single neonate that was recovered was interred with two adolescents. Given the otherwise total lack of very young children at KNXIV, it is possible that the neonate that was found was the near-term fetus of one of the adolescents with whom it was buried, and that its interment was unintentional.

The bones of infants and very young children are much more susceptible to taphonomic factors due to their size and density, making them less likely to be recovered than the bones of older children; it is also possible that children who died before reaching a certain age were disposed of in a different manner to other members of the group. Therefore, while Lieverse (2007b) suggests that the increase in mortality for children aged two to four may reflect weaning stress, it is difficult to be certain without knowing the relative mortality of children aged younger than two. Amongst KNXIV's adults, the highest mortality was observed amongst individuals aged between 35 and 50, an observation which suggests that it was uncommon for individuals in this group to reach old age. However, data from Ust'-Ida I, another EBA cemetery, indicate that

this pattern was not necessarily common to all EBA groups, as amongst Ust'-Ida I individuals the largest age group is formed by adults aged 50 and above (Lieverse, 2007b).

In terms of health, the most common conditions identified at KNXIV were osteoarthritis (OA), dental pathology, and enamel hypoplasia (EH). OA was the most common ailment observed, occurring in approximately 42% of the 60 individuals considered well-preserved enough to be assessed for health markers. The most common joints affected were vertebral and pedal, indicating habitual activities that stressed the back and feet in particular; for example, frequent carrying of heavy loads over long distances. OA is typically associated with age, and this pattern was reflected in the fact that none of the sub-adults assessed showed any signs of the condition, while all of the observable adults aged 50+ did (Lieverse, 2007b).

Dental pathology was the second most common condition identified, with periodontitis and ante-mortem tooth loss (AMTL) the most common manifestations. Again, these conditions primarily affected older individuals, a pattern which likely reflects the chronic nature of these conditions – the instances of AMTL are thought to reflect very advanced cases of periodontitis. None of the observable sub-adults showed signs of periodontitis or AMTL, while approximately 47% of adults exhibited one or the other, with adults over the age of 35 displaying this condition more often than younger adults. High levels of AMTL in adults typically reflect a diet heavy in tough, unprocessed foods. This diet results in dental attrition (the wearing away of the chewing surfaces of the teeth), a development which allows bacterial plaque to form and inflame the gums, resulting in periodontitis. Indeed, dental attrition is severe amongst the individuals from KNXIV. Interestingly, caries and abscesses are totally absent amongst this population, indicating a diet that contained little to no carbohydrates, a feature which is reflected in stable isotope data

that show that the diets of the individuals at KNXIV were primarily made up of fish and ungulate and seal meat (Lieverse, 2007b).

The third most common health issue was enamel hypoplasia (EH); and unlike the other two most common health conditions, this issue tended to affect more sub-adults than adults. Of the 20 observable sub-adults, 7 exhibited EH, compared to only 4 of the 24 observable adults. The defects in the teeth caused by EH form when the body undergoes physiological stress during childhood, so its prevalence in the younger individuals from KNXIV suggests that individuals who were subjected to such stressors may have been less likely to survive into adulthood (Lieverse, 2007b).

Other skeletal pathologies at KNXIV were uncommon, with only 6 of the 60 observable individuals displaying any other identifiable pathological markers. Two individuals exhibited osseous tarsal coalition (fusion of tarsal bones), and there were two cases of localized periostitis. One individual exhibited a number of lytic lesions on the cranial vault, tibiae, and femora, as well as an irregularly shaped femoral shaft, which was thought to be caused by multiple myeloma, and another individual showed bilateral defects of the tibial tuberosities, which were likely the result of Osgood-Schlatter disease. Finally, there was also one instance of a cortical depression in the shaft of a long bone, which may have been the result of a benign tumor. There was also a single incidence of skeletal trauma observed in an individual with healed fractures of the second, third, and fourth metatarsals of the left foot, likely the result of an accident rather than interpersonal violence (Lieverse, 2007b).

Although the incidence of certain chronic conditions like OA and AMTL was somewhat high, the relative absence of acute illness and infectious diseases – both specific (i.e. leprosy, tuberculosis) and non-specific (periostitis, osteomyelitis) – indicates that the KNXIV population

was generally quite healthy and carried a low pathogen load. This observation suggests a low population density and high group residential mobility rather than large, sedentary groups, amongst whom such infectious diseases are more common due to living in close proximity with others. In addition, despite some instances of EH, there is no evidence that the groups who utilised KNXIV suffered from chronic malnutrition, which indicates that they also had reliable access to high quality food resources and did not suffer from prolonged periods of hunger or other physiological stressors.

Chapter Three: Materials and Methods

3.1 Materials: Photography and Documentation of human burials at Khuzhir-Nuge XIV

Data for this project were collected using photographs taken during excavations at Khuzhir-Nuge XIV (KNXIV) between 1997 and 2001. All human burials at KNXIV were photographed with as many skeletal elements in situ as possible, and multiple images were taken of each burial on both black-and-white (BW) film and colour slides. As a rule, photography protocol in the field dictated that 35mm BW film was used to record horizontal, oblique, and close-up images of burials, and 35mm colour slides were used for oblique and close-up images. Images from other angles and views were also taken as necessary (Weber and Goriunova, 2008).

Photographs were always taken using a cord release trigger, and either a tripod, for oblique and close-up images, or a 3 m long steel monopod, for horizontal images. The tripod and monopod allowed photographs of each burial to be taken from the same height and angle, ensuring that image perspective and coverage were consistent from burial to burial. Most burials were recorded using the following sequence of images: horizontal and oblique shots showing the perimeters of the burial feature, a head-to-toe shot of the entire skeleton in situ, a $\frac{3}{4}$ -length head-to-mid-femur shot, and a 'passport' shot showing the skull and upper torso. Any additional images were usually of the skull, or, occasionally, the lower legs and feet, and artefact clusters depending on the specific archaeological context.

Khuzhir-Nuge XIV was excavated during five field seasons between 1997 and 2001. Seventy-four graves were excavated, with the remains of 89 individuals recovered in total (Lieverse, 2007). During excavation, as well as being photographed, each individual within a given burial was extensively documented while still in situ, with the aim of determining how much certain taphonomic variables affect skeletal preservation (Lieverse, 1999). Detailed

discussions of the results of these efforts are presented in a number of publications by Angela Lieverse (1999, 2006, 2007).

Information recorded prior to removal of skeletal elements from a grave included: body position, orientation of the body and face, articulation and position of individual skeletal elements, burial type, burial disturbance, maximum length, width, and depth of the grave, metric and non-metric osteobiographical indicators (such as mid-shaft diameters of long bones and subpubic concavities when elements were highly fragmented); and completeness of each skeletal element (expressed as a percentage). The number of fragments for each element was also recorded as skeletal materials were removed from graves (Lieverse, 1999).

The two variables of interest for this project were skeletal completeness and fragmentation. During excavation at KNXIV, skeletal completeness was visually estimated, taking into account both the relative completeness of the entire skeleton, and the completeness and fragmentation of individual elements (Lieverse, 1999: 68). Completeness was recorded as a percentage, with zero indicating that an element was completely absent, and 100% indicating a complete element. While Lieverse (1999) did not use a specific protocol for estimations of completeness (such as, for example, the protocol suggested by Buikstra and Ubelaker [1994]) her approach was systematic; and was consistently applied across all burials and field seasons (Lieverse, 1999; 2007). As all the burials at KNXIV were photographed in situ, completeness could be visually estimated from these photographs in much the same way as was done in the field, producing directly comparable data.

During excavation, fragmentation of skeletal elements was recorded by counting the number of fragments comprising an element as that element was removed from the grave and prepared for transport to the laboratory on site (Lieverse, 1999). For the purposes of data

analysis, fragmentation was expressed as a ratio between the completeness of an element and the number of fragments comprising it. While this protocol was obviously impossible to replicate completely in this study, the high quality of photographic documentation of the burials generally allowed for reasonable estimations of the number of fragments per element, and therefore calculation of fragmentation ratios, again producing directly comparable data.

3.2 Methods: Data Collection

All photographs taken during fieldwork at KNXIV are stored at the University of Alberta, Edmonton, as hard copies in the form of BW photographic negatives and colour slides, and as digital images on the BHAP's server. Though originally scanned at high resolutions (typically between 1000 and 4000 DPI) for printing purposes, the digital images stored on the BHAP server were not large enough when viewed at full size to allow skeletal fragments to be easily counted, so a number of images were selected to be re-scanned at a larger size in order to improve the amount of detail visible when viewed on a computer screen at full size (i.e., zoomed in to 100%).

Although the digital images from KNXIV typically had very large DPI (dots per inch) values, it is incorrect to interpret this characteristic as a measure of the resolution of digital images. DPI, or "dots per inch" refers to the printed resolution of a given image, based on the number of dots of ink per inch to which a printer converts the image during the printing process. This resolution is based both on the desired size of the printed image, and the DPI capabilities of the printer itself. As digital images are measured in pixels, it would be more correct to use the term "pixels per inch" or PP; however, even this term is not quite correct, as a digitally stored image is not a tangible object, and thus cannot be measured in inches. The use of the term DPI is simply a holdover from the days of print photography that can be thought of as a conversion factor that determines the size at which a scanned image can be printed while maintaining its

details and clarity. As a digital image cannot be measured in inches until it is printed, its DPI is essentially meaningless in the context of images viewed solely on a computer screen, because it is the computer's screen resolution, not the printer's DPI, which determines the quality of the image viewed on the screen.

Thus, when viewing images on a screen, the size and resolution of the display determine the PPI, while the pixel dimensions of the image determine how large it will appear on the screen. All images assessed during this study were viewed on a Dell 22" LCD monitor with a display resolution of 1920 x 1080 pixels. The screen dimensions of 18.74" x 10.54" mean that at the maximum screen resolution of 1920 x 1080 pixels, this monitor has a PPI of approximately 100¹. Thus, an image with pixel dimensions of 800 x 500 pixels would be displayed at the same size (taking up roughly 50% of the screen) and resolution (100 PPI) regardless of the DPI at which it was scanned, because the screen can display no more than 100 pixels per inch. Thus, the size of an image in terms of its pixel dimensions is more important than the scanned resolution of the image, because this feature is what allows the image to appear larger (or smaller) on-screen without causing distortion. The loss of detail that occurs when images are enlarged in editing software occurs because many programs tend to automatically resample resized images, adding extra pixels to maintain PPI at the new size via a process called interpolation. This procedure adds noise, and causes fine details to appear blurry. Enlarging images during scanning ensures that the resulting image has larger pixel dimensions, so that it takes up more of the screen, while maintaining the clarity of the image.

For this study, slides were scanned with an Epson Perfection V700 flatbed scanner. Individual slides were inserted into a slide carrier, which was then placed face-down on the

¹ PPI is calculated by dividing the diagonal pixel resolution of a display by its diagonal size in inches. The diagonal pixel resolution is calculated as $\sqrt{w^2 + h^2}$, where w^2 is the width of the display in pixels and h^2 is the height in pixels.

scanning bed, allowing twelve slides to be scanned at a time. Adjustments to the target size and resolution of the scanned images were made using the manufacturer-provided Epson Scan software. The manufacturer-recommended DPI of 300 is generally considered to be the optimal print resolution for most common print sizes, and enlarging images as they are scanned does not typically require a change in DPI. This change is only required if the size of the scanned images are going to be increased in photo-editing software; in this scenario Epson recommends increasing DPI by the amount by which the image is expected to be scaled up (e.g. for a 200% increase in size, the DPI should be increased from 300 to 600) (cite Epson manual/guidelines). With this in mind, images were enlarged during scanning by scaling up the target size to 250%, or 275% when a grave contained highly fragmented remains (as indicated by the scan preview); and the DPI was also increased to 700 as a precaution in case the image size needed to be increased further after scanning.

The quality of the hardware used to scan the image also makes a difference in the quality of both the digital image produced and the final printed image (Epson Canada, n.d.). When BHAP's slides were originally digitised, they were scanned with a Nikon Super Coolscan 4000-ED scanner, which was produced between 2001 and 2003, and has a maximum optical resolution of 4000 dpi. The Epson Perfection V700 scanner, which was released in 2006 and is still in production as of 2016, has a maximum optical resolution of 6400 dpi. Although the Nikon scanner was top-of-the-line, professional quality equipment when it was originally produced, improvements in technology and software and subsequent reductions in the cost of manufacturing mean that many newer, high-end flatbed scanners designed for home use match or even exceed the scanning capabilities of older professional models, despite being much cheaper

machines (Morris, 2007). These advances are evident in comparisons between original digital images stored on the BHAP server, and the newly scanned images used in this study:



Figure 3.1 Cropped image scanned with Nikon Super Coolscan 4000-ED at 1100DPI. Original pixel dimensions: 545 x 441.



Figure 3.2 Cropped image scanned with Nikon Super Coolscan 4000-ED at 1100 DPI, enlarged after scanning. Original pixel dimensions: 713 x 547.



Figure 3.3 Cropped image scanned with Epson Perfection V700 at 700 DPI, enlarged during scanning.

When considered in isolation, Figure 3.1, part of an original image taken from the BHAP server, appears reasonably sharp and clear. However, this image could not be enlarged any further without losing detail and sharpness, as can be seen in figure 3.2, which is the same image enlarged to similar dimensions as figure 3.3. In figure 3.3, the re-scanned image was enlarged during scanning so that the starting pixel dimensions were larger. Even though it was scanned at a lower DPI than the image in figures 3.1 and 3.2, the image in 3.3 is sharper and clearer, indicating that the newer scanner is able to pick up and resolve fine details just as well, if not better than the older model.

Images to be scanned were selected with the aim of ensuring that there were as many clear images of each burial as possible. Full-length images of each skeleton were re-scanned, as well as any $\frac{3}{4}$ -length, passport, and other close-up images associated with a particular interment. The process of scanning BW images from negatives proved to be very slow, as it required entire strips of film containing five or six frames to be scanned in order to obtain images from one or

two frames of interest. Because of this problem, scanning of BW film was abandoned and only colour slides were scanned.

In total, 473 images were re-scanned and enlarged. The process of scanning slides and organizing the resulting image files took approximately 30 hours. The number of images per burial was recorded in order to assess if image quantity had any effect on the accuracy of data collected. The number of image types associated with each burial was also recorded for the same reason, using the following categories: single angle – full body only (“1”); single angle – isolated element (“2”); two angles – full body and 3/4 length *or* passport (“3”); two angles – full body and isolated element/s (including skull) (“4”); three angles – full body, 3/4 length, *and* passport (“5”); three angles – full body, 3/4 *or* passport, isolated element/s above waist (“6”); three angles; full body, 3/4 *or* passport, isolated element/s below waist (“7”); four angles -- full body, 3/4, passport, *and* lower body (“8”); four angles – full body, 3/4, passport, and isolated region of upper body (e.g., pelvis) (“9”); four angles; full body, 3/4, passport, and other isolated elements or region (e.g., forearm, single hand) (“10”); and five angles – full body, 3/4, passport, lower body, and other isolated element or region (“11”). Angle categories were based on both number and type of angles, so that categories could be ordered to the extent that it was possible for statistical purposes; and so that any differences between categories with the same number of angles but different combinations of angle types would be more apparent.

After scanning, images were viewed and processed using the open-source software Image J (Rasband, 2014), an image-processing program designed for processing and analyzing scientific and medical images. Processing of images was minor, consisting of small adjustments to sharpness, contrast, and brightness as necessary, with the aim of making skeletal elements stand out from the image background. As well as allowing users to make basic adjustments to

images in terms of contrast, brightness, colour balance, and sharpness, Image J allows the user to install plugins that perform complex image analyses; for example, the identification of all objects of a specific size and shape in an image, or the estimation of the volume of blood vessels in 3D images of vascular networks. One of these plugins, Cell Counter (DeVos, 2010), allows the user to mark different categories (e.g., skeletal elements) of objects within an image with coloured markers; and to record the number of objects counted within each category. This software provided a simple way for each individual fragment of all visible skeletal elements in a single image to be visually identified and counted.

Skeletal completeness was always estimated and recorded prior to the counting of skeletal fragments. As well as estimates of completeness of each element, which was recorded as a percentage, the apparent presence or absence of elements was recorded under the category 'Visibility'. Visibility was recorded as either "0", for elements that appeared to be completely absent; or "1" for elements that were observable or partially observable. These data were then used to calculate an average measure of skeletal completeness, called the Representation Index (RI), which was calculated by dividing the number of visible elements by the total number of elements scored for each burial (Littleton et al. 2012).

A similar method was also used to create an average measure of articulation for each burial, called the Articulation Index (AI), calculated by dividing the number of elements scored as "1" for the variable 'Articulation' by the total number of elements scored (elements for this variable could be scored as "0" for no data, "1" for articulated elements, or "2" for elements that were present but not articulated). Due to the way Lieverse (1999) originally defined articulated vs. non-articulated elements, it was felt that results for this variable obtained from photographic documentation would generally not prove to be significantly different from what was

documented in the field. Therefore, AI was calculated only from Lieverse's (2006) data, as this variable was not included in the collection of data from photographic documentation.

Completeness of elements was estimated in a way that followed Lieverse's (1999, 2006) protocol as closely as possible. Estimates of element completeness were expressed as percentages to the nearest multiple of 5, and were based on the expected size of the element relative to both the observed portion and the size of the opposing and surrounding elements (Lieverse, 2006). Similarly, the completeness of elements that were only partially visible was estimated based on the completeness of the portion that was observable. As well, small groups of elements of a similar size and shape were combined into single units, with their completeness expressed as an average calculated by dividing the sum of the completeness estimate for each element by the total number of elements in the group. Elements were grouped together as follows: cervical vertebrae, thoracic vertebrae, lumbar vertebrae, left and right ribs, left and right carpals, left and right metacarpals, left and right proximal manual phalanges, left and right intermediate manual phalanges, left and right distal manual phalanges, left and right tarsals (excluding the calcaneus and talus, which were scored separately), left and right metatarsals, left and right proximal pedal phalanges, left and right intermediate pedal phalanges, and left and right distal pedal phalanges.

Fragmentation was assessed using Image J's Cell Counter plugin (DeVos, 2010). After initializing the Cell Counter for a particular image, all identifiable fragments of a skeletal element were marked by clicking on the fragment to overlay it with a coloured dot. The Cell Counter recorded these fragments under the label 'Type 1'. To mark the number of fragments of a new element in the same image, a new category was added to the existing overlay ('Type 2', 'Type 3' etc.); and the process of marking fragments by clicking was repeated. New categories

were added until all fragments in an image were counted, with each element's fragments represented by a different coloured marker. Generally, full-length burial images were used for counting fragments in the feet and legs, $\frac{3}{4}$ length images were used for counting fragments in the pelvis and upper body, and passport images were used to count fragments in the cranium. Following Lieverse (1999) again, fragment counts were recorded and used to calculate a fragmentation ratio for each element by dividing the element's completeness score by the total number of fragments identified as belonging to that element. Fragments that could not be confidently assigned to a particular element were not marked.

3.3 Statistical Analyses

Statistical analysis had two primary goals: to measure how closely data collected from photographic documentation agreed with data collected in the field, and to assess what factors influence the accuracy of data collected from photographic documentation. Agreement was assessed at two levels. At the individual element level, the variables Visibility, Completeness, and Fragmentation Ratio (FR) were compared. At the burial level, average measures for these three variables were calculated, with Completeness representing the average completeness score per visible element within a burial; FR, the average fragmentation ratio per visible element; and Representation Index (RI) – calculated from the Visibility scores – used as a measure of overall skeletal completeness.

The contributions of each independent variable (IV) to the variance of a series of different dependent variables (DVs) were assessed using one-way ANOVA for categorical IVs, and simple linear regression (also known as ordinary least squares, or OLS, regression) for continuous IVs. Multiple regression (MR) was then used to assess the effects of a combination of multiple IVs on a series of different DVs in order to try and isolate which factors most affected

the accuracy of Completeness and FR estimates. All regressions were conducted with burial-level data, using various DVs calculated during the analysis of agreement for both Completeness and FR.

A secondary question that arose during the process of data collection was also addressed using simple linear regression. When scoring cranial elements for completeness, it was noted that it was not possible for the entire cranium to ever be completely visible in a two-dimensional image, since it is a globular, three-dimensional structure composed of many separate bones. This observation raised the question of whether or not, when using photographic documentation in this way, it is reasonable to assume that elements on the side of the cranium that are not visible are in a similar state of preservation to their counterparts on the visible side. This question was addressed using linear regression to assess how well the completeness of left-sided cranial elements (the side with the highest total of observable elements) could be used to predict the completeness of right-sided cranial elements.

3.3.1 Data Considerations

There were two major considerations that influenced the choice of statistical methods used to analyse the data in terms of agreement. First, at the individual element level in both datasets, there were large amounts of missing data in which elements were either not present or otherwise unobservable for either completeness, fragmentation, or both. In many scenarios, variables or cases with missing data are simply excluded from statistical analysis, as it was assumed that nothing can be learned from data that do not exist (Field, 2005). However, in this case, a “0” in the data did not always necessarily indicate a lack of information. For example, with the Visibility variable, it was equally important to know for comparative purposes when an element was not visible, i.e., scored as “0”, as it was to know which elements were given a score

of “1”. Similarly, when comparing the variables Completeness and FR, it would not make sense to remove a case from the dataset in which one observer had given an element a score of “0” and the other observer had scored the same element differently, as such differences formed the main objects of interest.

The large number of zeros in the dataset meant that transforming by the commonly-recommended methods of using the natural log or square root to achieve normality had very little effect on the distribution, as the natural log and the square root of zero is still zero; thus, even when the overall range of measurements was reduced, the distribution remained resolutely non-parametric. For the Completeness and FR variables, removing cases in which both observers had scored an element as “0” resulted in an overall distribution that was closer to normal. However, this effect was not consistent across all burials, as it was dependent on the number of elements that remained once the paired zero observations were removed, and therefore a non-parametric method of measuring agreement was determined to be the most appropriate approach.

The second consideration stemmed from the interpretation of the term ‘agreement’ as a statistical concept. According to Bland and Altman (1986), researchers tend to conflate the concepts of agreement and correlation, using measures of correlation as indicators of agreement when in fact these two things are not quite the same. A correlation between two variables indicates that they share a linear relationship, but this correlation does not necessarily mean that the data agree at the level of the individual pairs of observations. If the two variables in question are measurements of the same thing by two different observers, then the expectation of a linear relationship between them should be obvious; however, two variables can be highly correlated without any single pair of observations being in absolute agreement. Thus, correlations merely show the presence, direction, and strength of a relationship between two variables; they do not

indicate to what extent the data agree with one another in terms of absolute differences between pairs of observations (Bland and Altman 1986).

With these considerations in mind, Bland and Altman's (1986) Limits of Agreement (LOA) procedure was selected in order to assess agreement at the level of individual elements on a burial-by-burial basis. The LOA procedure is a simple method of assessing the agreement between two sets of observations; using this method, agreement is quickly and easily interpreted visually via a scatterplot. It was developed for use in medical research in order to quantify how much a new method or instrument differs from one currently in use, in order to determine if the new one is accurate enough to satisfactorily replace the old.

Generally, the LOA are calculated from the mean difference (\bar{d}) between two variables, using the standard deviation (sd) of these differences to estimate the 95% confidence limit, which, if the differences are normally distributed, will be approximately $\bar{d} \pm 2$ s.d. The LOA are then easily visualized with a scatterplot of the differences between the two variables plotted against their means, with reference lines added to indicate the mean difference – here referred to as the bias – and the confidence interval (Figure 3.4). Although Bland and Altman (1986) assume that the differences between two variables will be normally distributed, they note that it is not necessary for the variables themselves to follow a normal distribution. By looking at the measurement error (the differences) rather than the measurements themselves, the variation is reduced and thus the data are more likely to be normally distributed (though this observation is not always the case).

For this study, due to the large amount of variation between burials in terms of both the number of elements and the range of differences, non-parametric LOA were determined to be more appropriate than the more typical parametric version, as they could be consistently applied

in the same way to each burial. When using LOA and following the assumption of a normal distribution, non-uniform differences can result in sometimes complicated relationships between the differences between two variables and their means, differences which must be accounted for when calculating the bias and the LOA (Bland and Altman, 1999). Given the idiosyncrasies of the data being analysed, this situation meant that if the assumption of normality was followed, the bias and LOA would potentially have to be calculated differently for some burials than they were for others, which could make them difficult to compare directly. Therefore, in order to avoid such a situation, all burials were treated as if they had non-parametric distributions.

3.3.2 Element-level Analysis

Visibility was analysed using Krippendorff's alpha (α) as a simple measure of inter-observer reliability. Krippendorff's alpha is a measure of inter-observer reliability developed by Klaus Krippendorff in response to the perceived flaws of other, more common measures (Hayes and Krippendorff, 2007). For binary data with two observers, like the visibility data from this study, Cohen's Kappa (κ) is a commonly-used measure of inter-observer reliability. However, κ suffers from some widely acknowledged flaws, including the lack of a standardized interpretive scale; and a tendency for the value of κ to become inflated when systematic disagreements occur between observers. This inflation occurs because the value of κ is heavily dependent on the distribution of the marginal totals, and this dependence results in a further weakness in that values of κ cannot be reliably compared between samples unless the marginal distributions are similar (Laerd Statistics, 2015). In addition, the zero-point of κ is more akin to that of correlation statistics like Spearman's rho as, although a rare outcome, κ can have a negative value which, combined with the fact that there is no agreed-upon interpretive scale, can make the interpretation of κ somewhat ambiguous, especially at lower values. In contrast, α is designed to

be statistically independent of the data it is being used to describe, and it can be simply interpreted on a scale from 0 to 1, in which 1 indicates perfect, or 100% agreement, and 0 indicates no agreement whatsoever (Hayes and Krippendorf, 2007).

Because of the way it is calculated, by assessing disagreements directly rather than correcting agreement percentages, α is able to cope with missing data in the form of both missing observations and missing categories (i.e., if an observer fails to assign any observations to a particular category); and it can be used with any number of observers, and with nominal, ordinal, interval, and ratio data. Hayes and Krippendorf (2007) argue that its properties make α a good candidate for use as a standard measure of reliability; however, it remains relatively unknown due to its lack of inclusion in common statistical software packages. To that end, Hayes developed a freely available macro for SPSS that is easily implemented and straightforward to use. As well as calculating the value of α , the macro allows the user to utilise bootstrapping of samples in order to calculate 95% confidence intervals for fine-grained interpretation of α values. However, this option was not utilized in this study as that level of detail was not considered necessary for general estimates of the reliability of a single variable.

An overall value of α was calculated for the entire dataset (approximately 2100 individual elements from 66 individuals from 59 graves; Table 4.1), as well as for smaller sub-groups of data, including by element side (Table 4.2), element type (Table 4.3), element type (Table 4.4), and for each of the 66 individuals included in the study (Table 4.5). Element types were the same as those used by Lieverse (1999): long bones, short bones, flat bones, irregular bones, and mandibles. Element groups were based on the region of the body to which an element belonged: skull/cranium, axial skeleton (ribs, vertebrae), arm/shoulder (clavicles, scapulae, humeri, radii, ulnae), hand/wrist (carpals, metacarpals, manual phalanges), leg/pelvis (ilia, ischia, pubis,

femora, tibiae, fibulae), and foot/ankle (tarsals, metatarsals, pedal phalanges). Element side was again based on Lieverse's (1999) classification, with left, right, and unsided elements. In addition, frequency tables were also constructed to assess, using percentages of the total element count, how often and by how much the differences between real and estimated Completeness and FR scores differed from zero by element type (Table 4.6, 4.7) and element group (Table 4.8, 4.9).

Non-parametric LOA were calculated for all 66 individuals for both the Completeness and FR scores. Rather than using a 95% confidence interval as the upper and lower LOA, non-parametric limits are based on the proportion of differences that fall within a certain reference range. Following the example shown in Bland and Altman (1999), the reference range for the LOA was set with three levels: ± 5 , ± 10 , and ± 15 , representing excellent, good, and moderate agreement. These levels are visually represented as reference lines on the scatterplots for each individual that was assessed, as shown in the LOA plot for Burial 23 (Figure 3.4; also see Appendix B).

Difference Between Element Scores vs. Mean of Element Scores
Feature 23

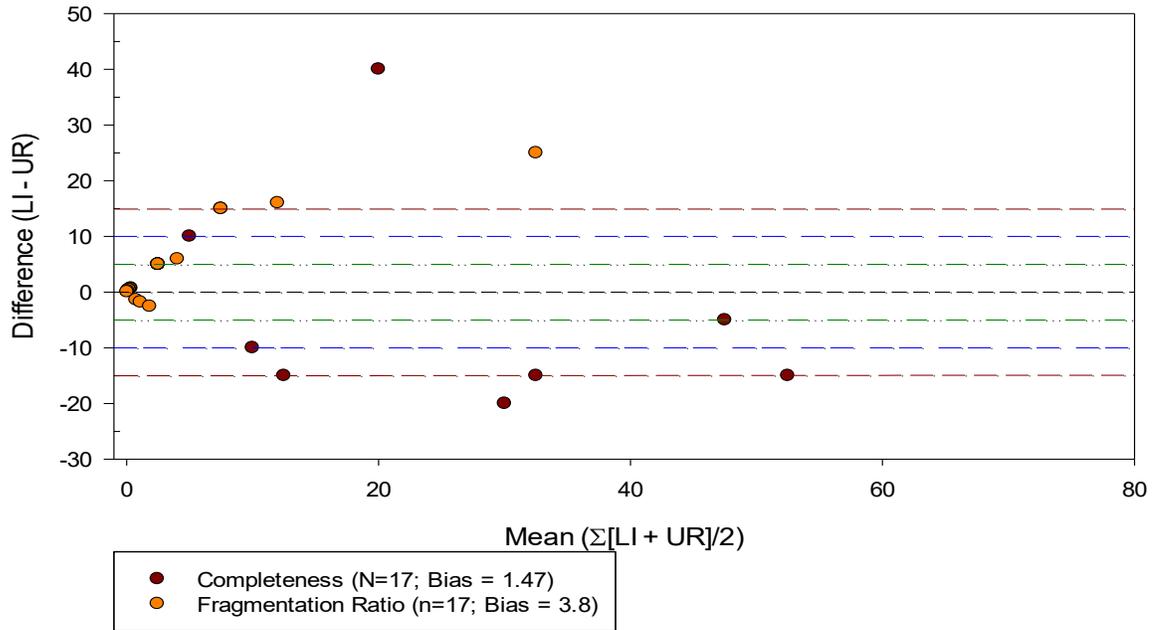


Figure 3.4 Bland-Altman plot for Burial 23 showing the distribution of Completeness and Fragmentation Ratio scores within the Limits of Agreement

To aid in interpretation of the scatterplots, the cumulative percentages of elements showing a difference from zero of ± 5 , ± 10 , and ± 15 were calculated for each individual for both Completeness and FR, and these were entered into tables along with the average difference per element (Table 4.10, 4.11). The average difference per element – referred to as the bias – was calculated by taking the mean of the sum of differences for each individual (Bland and Altman, 1989). A bias of 0 ± 5 indicated that the average difference per element observed between Lieveise’s and Urlacher’s scores was no more than 5 points, a result which was considered to indicate excellent agreement. A bias of $\pm 5.01 - 10$ was considered to indicate good agreement, $\pm 10.01 - 15$ moderate agreement, $\pm 15.01 - 20$ poor agreement, and ± 20.01 or more very poor agreement (Table 3.1).

Bias	Agreement
0±5	Excellent
±5.01-10	Good
±10.01-15	Moderate
±15.01-20	Poor
±20.01	Very poor

Table 3.1: Agreement categories based on Bias scores

% 0±5	% 0±10	% 0±15	Agreement
55	65	75	Excellent
45	55	65	Good
35	45	50	Moderate
15	20	25	Poor
<15	<20	<25	Very Poor

Table 3.2: Minimum distribution thresholds for Limits of Agreement categorisation

Using the data in Tables 4.10 and 4.11, the agreement category indicated by the bias was then revised based on the cumulative percentages of elements that exhibited excellent, good, and moderate agreement, using the distribution thresholds noted in Table 3.2 (i.e., the total percentage of elements exhibiting a difference from zero of ± 5 , ± 10 , and ± 15). Following Table 3.2, a burial with 55% or more cases exhibiting a difference from zero of ± 5 , 65% or more cases a difference ± 10 , and 75% or more cases a difference ± 15 would be classified as displaying excellent agreement. A burial with more than 45% of cases but fewer than 55% exhibiting a difference of ± 5 , more than 55% of cases, but fewer than 65% a difference of ± 10 , and more than 65% but fewer than 75% a difference of ± 15 would be classified as showing good agreement.

In cases in which the value of the overall bias matched the agreement classification defined by Table 3.2, no adjustment to the agreement category was made; however, where the two classifications did not match, the distribution of difference was used to revise the agreement category indicated by the bias alone. For example, if a burial's overall bias was in the range of 0 ± 5 , but the distribution of cases was in the poor agreement range (i.e., fewer than 50% of cases

with a difference from zero of ± 15 , fewer than $45\% \pm 10$, and fewer than $35\% \pm 5$), then its agreement classification would be revised from excellent to moderate or poor, depending on which threshold values the distributions of cases were closer to (Table 4.12, 4.13).

Finally, the real and estimated mean completeness and FR were calculated for all elements, and for each element group and element type using all elements scored. The means were then re-calculated using only elements for which both Lieverse and Urlacher assigned a Visibility score of “1”; and these were compared with the values obtained using all elements scored by each observer, in order to assess how they differed (Table 4.14, 4.15, 4.16).

Completeness and FR biases and average scores were also recalculated for all elements, by element group, and by element type using only elements that both observers recorded as being visible; and these were compared with the original bias scores generated using every single element scored by each observer (Table 4.17, 4.18, 4.19).

3.3.3 Burial-level Analysis

Examination of relationships and correlations between variables was conducted at the burial level, incorporating data produced by examination of differences at the level of individual skeletal elements. Average completeness per element in a burial was calculated by taking the mean completeness for all scored visible elements, and average FR per element was calculated the same way. New variables for the difference between averages for completeness and fragmentation were then calculated by subtracting the average Completeness and FR scores calculated from photograph-collected data (hereafter referred to as ‘estimated’ scores) from the average scores calculated from field-collected data (hereafter referred to as ‘observed’ scores).

Although these difference variables may seem similar to the bias per element variables calculated during analysis of agreement, the bias takes into account the difference in the number

of Observed vs. Estimated elements within a burial as well as differences in their Completeness scores, while the average difference is simply a measure of how much the average completeness and FR scores differ between observers in absolute terms. It is thus unaffected by the difference in observed numbers of elements. RI was calculated as the percentage of elements scored as Visible out of a potential total number of elements, and thus it can also be interpreted as a basic measure of overall skeletal completeness (as opposed to the average completeness of individual elements represented by the variable Completeness).

One-way ANOVA was used to assess the individual contributions of categorical variables to a series of different DVs including Completeness Bias, FR Bias, Difference RI Number of Images, and Angles, as well as Estimated Completeness, Estimated FR, and Estimated RI (Table 4.20; See Appendix C for list and explanation of all variables). Simple linear regression was used to assess the individual contributions of continuous variables to the same DVs (Table 4.21), with the exceptions of Number of Images and Angles. Number of Images and Angles were excluded from this step of analysis because as categorical variables, they cannot be used as DVs in regression models. However, because of its similarity to a continuous variable, it was possible to use Number of Images as a continuous IV in MR models, because each category of the variable has a numerically meaningful value in which the number of a given category directly represents the number of images associated with a burial. This meaningful representation was not the case for Angles, where each numeric category represented a different combination of angle types. Due to the large number of categories this variable represented, it was not possible to use dummy coding and include each level of the variable as a separate IV, as was done with other categorical variables (Angles consists of 11 categories, whereas the other categorical variables had a maximum of 3 categories each). Because of this

factor, Angles was excluded from MR models; and its effect on the DVs was visually assessed using means plots instead (see Figure 4.4, 4.5, 4.6 and Appendix B).

During MR analyses, estimated scores were excluded from the list of IVs because they were assumed to share a strong linear relationship with IVs representing observed scores. Including variables that share close linear relationships in the same MR model results in collinearity, which makes the extent of the effect of either variable difficult to distinguish. As the DVs used in the MR models – Completeness Bias, FR Bias, and Difference Completeness, FR, and RI – are all derived from the differences between observed scores and their corresponding estimates, they can therefore act as a measure of the accuracy of the estimated scores. Because the observed scores existed first and represent the ‘true’ values of a burial’s Completeness, FR, or RI, they can be assumed to affect estimated scores, a situation which will be reflected in the size of the bias or the difference. As the opposite (i.e., estimated scores affecting observed scores) is not possible, it therefore makes sense to exclude the estimated variables from MR models and retain the observed variables.

In a number of cases, the presence of collinearity amongst variables resulted in heteroscedasticity, the presence of correlated errors. This type of violation of the assumption of normality usually renders a model invalid because the presence of heteroscedasticity makes it difficult to correctly calculate standard errors and significance values for the variables in which it is present. However, because this factor is a fairly common issue in OLS regression, Hayes and Cai (2007) have developed a family of heteroscedasticity-consistent standard error estimating equations, which are freely available and easily implemented as an SPSS macro called HCREG. When the macro is used to calculate regression parameters, if the presence of heteroscedasticity is detected in a variable, the standard error terms and significance value are automatically

corrected, while variables that do not violate this assumption are unaffected. Typically, the only variable that seemed to be strongly affected by heteroscedasticity was Integrity (see Discussion), which displayed very large standard errors when MR models run with the normal SPSS procedure were corrected using the HCREG macro (a procedure which was carried out for every regression model).

Variables were grouped together for block entry in hierarchical MR models based on whether they were categorical or continuous, and on which variables were of most interest in terms of their effects on the DV. A series of MR models were produced. In the main models, Model 1 and Model 2, the DVs were Completeness Bias (CMP BIAS) and FR Bias (FR BIAS) respectively. Rather than using stepwise, forward, or backwards entry or removal of variables based on their effect on the value of F, all variables of interest were forced into the model regardless of their significance. This procedure was carried out because the presence and direction of relationships between independent variables and the dependent variables were considered to be of greater interest than the significance value alone; and with all of the independent variables left in the model, each variable's contribution could be assessed in terms of its potential real-world effects, regardless of its statistical significance.

Based on the results of Models 1 and 2, three further MR models were constructed to assess the effect of the IVs on the DVs Difference Completeness (DIFF CMP), Difference Fragmentation (DIFF FRAG), and Difference RI (DIFF RI). Each of these three models was constructed hierarchically, in order to assess the effects of the continuous IVs with and without the influence of the categorical IVs. During analysis, the reliability of estimated FR scores was found to be unsatisfactory due to issues with the calculation used to transform raw fragment counts into FR values (see Discussion). Because of this problem, Difference FR was replaced in

the final MR model (Model 5) with Difference Fragmentation, which represents the difference between Observed and Estimated average raw fragment counts per element.

Chapter Four: Results and Discussion

The results of this research are presented below in Section 4.1 as a series of tables summarising the outcomes of statistical analyses. A detailed discussion follows immediately in Section 4.2.

4.1 Presentation of Results

	Ntotal	Ndisagree	%	K α
Total	5290	759	14.35	0.674

Table 4.1 Percentage Disagreement and Krippendorf's Alpha for all elements

Element Side	Ntotal	Ndisagree	%	K α
Unsided	928	154	16.59	0.568
Left	2181	314	14.40	0.684
Right	2181	291	13.34	0.7

Table 4.2 Percentage Disagreement and Krippendorf's Alpha by Element Side

Element Type	Ntotal	Ndisagree	%	K α
Long bone	937	126	13.45	0.699
Short Bone	1072	113	10.54	0.501
Flat Bone	1205	254	21.08	0.565
Irregular Bone	2009	257	12.79	0.625
Mandible	67	9	13.43	0.701

Table 4.3 Percentage Disagreement and Krippendorf's Alpha by Element Type

Element Group	Ntotal	Ndisagree	%	K α
Cranial	1540	252	16.36	0.56
Axial	592	92	15.54	0.554
Arm/Shoulder	615	101	16.42	0.664
Hand/Wrist	679	101	14.87	0.406
Leg/Pelvis	938	128	13.65	0.722
Foot/Ankle	926	85	9.18	0.7554

Table 4.4 Percentage Disagreement and Krippendorf's Alpha by Element Group

Burial	Ntotal	Ndisagree	%	K α
7	79	15	18.99	0.509
9	79	20	25.32	0.49
10	79	8	10.13	0.502
11	77	20	25.97	0.25
12	79	7	8.86	0.542
14	79	23	29.11	0.27
15	79	7	8.86	0.824
16	80	11	13.75	0.652
19	79	13	16.46	0.659
21	79	5	6.33	0.515
22	79	9	11.39	0.509
23	79	11	13.92	0.444
24	79	6	7.59	0.786
25	79	6	7.59	0.741
27.01	79	12	15.19	0.698
27.02	79	15	18.99	0.559
27.03	79	9	11.39	0.509
28	78	6	7.69	0.754
29	76	15	19.74	0.541
32	79	15	18.99	0.622
33	79	6	7.59	0.364
34	79	14	17.72	0.643
35.01	79	16	20.25	0.582
35.02	79	13	16.46	0.505
36.02	79	9	11.39	0.601
37.01	79	12	15.19	0.694
37.02	79	14	17.72	0.647
38	79	13	16.46	0.672
39	79	6	7.59	0.707
44	79	11	13.92	0.704
45	79	19	24.05	0.442
46	79	12	15.19	0.451
47	79	9	11.39	0.729
48	79	27	34.18	0.116

Burial	Ntotal	Ndisagree	%	K α
49	79	12	15.19	0.658
50	79	13	16.46	0.599
51	79	8	10.13	0.799
53	80	15	18.75	0.617
55	79	12	15.19	0.694
57.01	79	16	20.25	0.524
57.02	80	23	28.75	0.164
58.01	78	13	16.67	0.656
58.02	79	4	5.06	0.897
59.01	79	2	2.53	0.844
59.02	79	20	25.32	0.497
60	79	13	16.46	0.672
61	79	7	8.86	0.771
63	81	18	22.22	0.558
64	79	22	27.85	0.446
66	79	23	29.11	0.363
68	79	22	27.85	0.268
71	79	4	5.06	0.888
72	79	7	8.86	0.689
73	79	14	17.72	0.601
75	79	1	1.27	0.883
76	79	6	7.59	0.8
77	79	1	1.27	0.927
78	80	9	11.25	0.731
79	79	0	0.00	1
80.02	79	17	21.52	0.444
81	80	7	8.75	0.79
83	77	1	1.30	0.795
84	79	2	2.53	0.862
85	79	1	1.27	0.883
86	79	15	18.99	0.584
87	79	17	21.52	0.554

Table 4.5 Percentage Disagreement and Krippendorf's Alpha values by Burial

Diff. Complete	All Elements		Long Bone		Short Bone		Flat Bone		Irregular Bone		Mandible	
	N	%	N	%	N	%	N	%	N	%	N	%
+80.01-100	108	5.23	9	1.32	9	5.14	39	6.41	51	9.26	0	0.00
+60.01-80	130	6.29	27	3.95	14	8.00	41	6.74	43	7.80	4	8.16
+40.01-60	150	7.26	36	5.27	22	12.57	42	6.91	46	8.35	4	8.16
+25.01-40	203	9.83	40	5.86	31	17.71	55	9.05	73	13.25	4	8.16
+15.01-25	212	10.26	55	8.05	44	25.14	54	8.88	50	9.07	9	18.37
+10.01-15	126	6.10	35	5.12	16	9.14	32	5.26	42	7.62	1	2.04
+5.01-10	146	7.07	52	7.61	7	4.00	40	6.58	41	7.44	6	12.24
+0.01-5	244	11.81	120	17.57	8	4.57	64	10.53	46	8.35	6	12.24
0	208	10.07	86	12.59	7	4.00	68	11.18	40	7.26	7	14.29
-0.01-5	197	9.54	103	15.08	3	1.71	54	8.88	34	6.17	3	6.12
-5.01-10	119	5.76	56	8.20	2	1.14	38	6.25	23	4.17	0	0.00
-10.01-15	66	3.19	19	2.78	8	4.57	29	4.77	17	3.09	1	2.04
-15.01-25	75	3.63	24	3.51	3	1.71	22	3.62	14	2.54	3	6.12
-25.01-40	53	2.57	12	1.76	1	0.57	21	3.45	17	3.09	0	0.00
-40.01-60	15	0.73	6	0.88	0	0.00	5	0.82	4	0.73	0	0.00
-60.01-80	9	0.44	2	0.29	0	0.00	4	0.66	3	0.54	0	0.00
-80.01-100	5	0.24	1	0.15	0	0.00	1	0.16	3	0.54	0	0.00
Total	2066	100.00	683	100	175	100	608	100	551	100	49	100

Table 4.6 Difference between Observed and Estimated Completeness scores by Element Type

Diff. Fragmentation	All Elements		Long Bone		Short Bone		Flat Bone		Irregular Bone		Mandible	
	N	%	N	%	N	%	N	%	N	%	N	%
+80.01-100	133	6.44	1	0.15	54	30.86	16	2.64	62	11.23	0	0.00
+60.01-80	77	3.73	8	1.17	26	14.86	12	1.98	31	5.62	0	0.00
+40.01-60	78	3.78	11	1.61	11	6.29	15	2.48	41	7.43	0	0.00
+25.01-40	102	4.94	17	2.49	12	6.86	26	4.30	45	8.15	2	4.08
+15.01-25	91	4.41	15	2.20	4	2.29	28	4.63	42	7.61	2	4.08
+10.01-15	66	3.20	16	2.34	5	2.86	20	3.31	25	4.53	0	0.00
+5.01-10	131	6.35	34	4.98	5	2.86	47	7.77	42	7.61	4	8.16
+0.01-5	394	19.09	137	20.06	8	4.57	166	27.44	69	12.50	14	28.57
0	80	3.88	9	1.32	14	8.00	21	3.47	33	5.98	3	6.12
-0.01-5	378	18.31	213	31.19	7	4.00	107	17.69	42	7.61	9	18.37
-5.01-10	130	6.30	61	8.93	7	4.00	42	6.94	16	2.90	4	8.16
-10.01-15	87	4.22	43	6.30	0	0.00	25	4.13	17	3.08	2	4.08
-15.01-25	102	4.94	40	5.86	3	1.71	34	5.62	23	4.17	2	4.08
-25.01-40	89	4.31	28	4.10	6	3.43	18	2.98	32	5.80	4	8.16
-40.01-60	60	2.91	22	3.22	7	4.00	13	2.15	18	3.26	0	0.00
-60.01-80	53	2.57	21	3.07	6	3.43	13	2.15	10	1.81	1	2.04
-80.01-100	13	0.63	7	1.02	0	0.00	2	0.33	4	0.72	2	4.08
Total	2064	100	683	100	175	100	605	100	552	100	49	100

Table 4.7 Difference between Observed and Estimated Fragmentation Ratio scores by Element Type

Diff. Complete	Cranium		Axial		Arm/Shoulder		Hand/Wrist		Leg/Pelvis		Foot/Ankle	
	N	%	N	%	N	%	N	%	N	%	N	%
+80.01-100	72	14.49	5	3.14	5	1.24	8	5.44	7	1.18	11	4.12
+60.01-80	54	10.87	6	3.77	13	3.23	12	8.16	26	4.38	19	7.12
+40.01-60	41	8.25	10	6.29	23	5.72	17	11.56	38	6.40	21	7.87
+25.01-40	47	9.46	15	9.43	32	7.96	30	20.41	42	7.07	37	13.86
+15.01-25	33	6.64	13	8.18	40	9.95	36	24.49	56	9.43	34	12.73
+10.01-15	21	4.23	8	5.03	28	6.97	15	10.20	32	5.39	22	8.24
+5.01-10	29	5.84	19	11.95	30	7.46	5	3.40	44	7.41	19	7.12
+0.01-5	36	7.24	26	16.35	70	17.41	8	5.44	81	13.64	23	8.61
0	58	11.67	6	3.77	48	11.94	1	0.68	71	11.95	24	8.99
-0.01-5	36	7.24	14	8.81	40	9.95	1	0.68	88	14.81	18	6.74
-5.01-10	15	3.02	9	5.66	35	8.71	1	0.68	47	7.91	12	4.49
-10.01-15	16	3.22	7	4.40	14	3.48	0	0.00	21	3.54	8	3.00
-15.01-25	13	2.62	11	6.92	9	2.24	7	4.76	21	3.54	14	5.24
-25.01-40	13	2.62	8	5.03	6	1.49	4	2.72	17	2.86	5	1.87
-40.01-60	4	0.80	2	1.26	6	1.49	2	1.36	1	0.17	0	0.00
-60.01-80	5	1.01	0	0.00	2	0.50	0	0.00	2	0.34	0	0.00
-80.01-100	4	0.80	0	0.00	1	0.25	0	0.00	0	0.00	0	0.00
Total	497	100	159	100	402	100	147	100	594	100	267	100

Table 4.8 Difference between Observed and Estimated Completeness scores by Element Group

Diff. Fragmentation	Cranium		Axial		Arm/Shoulder		Hand/Wrist		Leg/Pelvis		Foot/Ankle	
	N	%	N	%	N	%	N	%	N	%	N	%
+80.01-100	40	8.06	7	4.40	0	0.00	54	36.73	5	0.84	27	10.07
+60.01-80	12	2.42	5	3.14	7	1.74	22	14.97	10	1.69	21	7.84
+40.01-60	24	4.84	10	6.29	5	1.24	12	8.16	8	1.35	19	7.09
+25.01-40	34	6.85	11	6.92	12	2.99	10	6.80	9	1.52	26	9.70
+15.01-25	34	6.85	10	6.29	13	3.23	2	1.36	12	2.03	20	7.46
+10.01-15	21	4.23	7	4.40	11	2.74	4	2.72	13	2.20	10	3.73
+5.01-10	52	10.48	11	6.92	24	5.97	1	0.68	24	4.05	19	7.09
+0.01-5	102	20.56	35	22.01	110	27.36	8	5.44	120	20.27	19	7.09
0	36		1	0.63	3	0.75	9	6.12	12	2.03	19	7.09
-0.01-5	59	11.90	22	13.84	95	23.63	7	4.76	177	29.90	18	6.72
-5.01-10	16	3.23	9	5.66	36	8.96	4	2.72	53	8.95	12	4.48
-10.01-15	15	3.02	5	3.14	21	5.22	0	0.00	38	6.42	7	2.61
-15.01-25	12	2.42	12	7.55	25	6.22	5	3.40	39	6.59	10	3.73
-25.01-40	17	3.43	9	5.66	14	3.48	3	2.04	29	4.90	17	6.34
-40.01-60	7	1.41	1	0.63	14	3.48	4	2.72	18	3.04	16	5.97
-60.01-80	11	2.22	3	1.89	12	2.99	2	1.36	18	3.04	7	2.61
-80.01-100	4	0.81	1	0.63	0	0.00	0	0.00	7	1.18	1	0.37
Total	496	100	159	100	402	100	147	100	592	100	268	100

Table 4.9 Difference between Observed and Estimated Fragmentation Ratio scores by Element Group

Burial	Bias	% +/-5	% +/-10	% +/-15	N
7	10.96	21.43	42.86	53.57	28
9	25.52	24.44	40	44.44	45
10	20.54	38.46	46.15	46.15	13
11	33.54	57.14	61.42	64.29	70
12	20.17	25	33.33	50	12
14	4.26	29.41	35.29	41.17	34
15	21.19	26.83	36.59	51.22	41
16	19.35	11.11	18.52	29.63	27
19	11.42	23.68	39.47	57.89	38
21	55	0	0	0	8
22	22	20	20	26.67	15
23	1.47	52.94	64.71	88.24	17
24	-0.53	42.86	61.9	71.43	21
25	7.2	11.76	23.53	35.29	17
27.01	9.92	45.23	59.52	66.67	42
27.02	10.94	52	72	80	25
27.03	-23.67	40	40	53.33	15
28	7.89	55.56	66.67	66.67	19
29	24.66	22.22	29.63	37.04	54
32	17.13	25.54	31.91	48.94	47
33	26.43	28.57	28.57	42.86	7
34	23.29	21.42	42.86	50	42
35.01	8.12	32.5	52.5	62.5	40
35.02	12.63	31.58	47.37	52.63	19
36.02	-12.33	40	60	66.67	15
37.01	14.41	33.33	48.72	56.41	39
37.02	9.62	25	52.08	64.58	48
38	22.83	25.53	36.17	40.42	47
39	7	40	66.67	66.67	15
44	14.29	43.64	50.91	65.55	55
45	28.68	20.59	29.41	29.41	34
46	19.03	45.71	52.86	57.14	70
47	17.81	25	46.43	53.57	28
48	20.82	11.76	14.71	17.65	34

Burial	Bias	% +/-5	% +/-10	% +/-15	N
49	13.71	40.63	56.25	59.36	32
50	17.49	44.83	51.72	68.97	29
51	12.86	32.56	53.49	58.14	43
53	9.42	16.98	32.08	43.4	53
55	9.12	36.59	53.66	60.98	41
57.01	11.68	37.5	53.13	75	32
57.02	34.31	24.14	24.14	31.03	29
58.01	18.28	26.31	39.47	63.16	39
58.02	10.28	42.86	62.86	68.57	35
59.01	1.25	0	12.5	12.5	8
59.02	21.8	30	42	48	50
60	20.28	35.42	45.83	56.25	48
61	8.97	33.33	54.17	75	24
63	22.3	31.38	43.14	56.86	51
64	19.42	17.31	35.29	43.14	52
66	36.83	20	21.67	25	60
68	27.71	37.14	41.43	45.71	70
71	1.37	55.17	72.41	72.41	29
72	12.41	17.65	58.82	70.59	17
73	12.8	39.39	57.58	63.64	33
75	-5	40	60	60	5
76	7.61	30.43	47.83	52.17	23
77	8.5	25	37.5	37.5	8
78	16.22	32.14	50	64.29	28
79	0.63	37.5	50	62.5	8
80.02	32.24	32.14	39.26	46.43	28
81	17.21	29.63	40.74	51.85	27
83	-5	33.33	66.67	66.67	3
84	8.22	44.44	77.78	88.89	9
85	10	40	60	100	5
86	18.82	22.86	37.14	60	35
87	9.81	28.13	43.75	59.36	32

Agreement Categorization (Bias)	
0±5	Excellent
±5.01-10	Good
±10.01-15	Moderate
±15.01-20	Poor
±20.01	Very poor

Table 4.10 Completeness Bias and percentages of elements within Limits of Agreement per burial, with bias agreement categories noted

Burial	Bias	% +/-5	% +/-10	% +/-15	N
7	-3.9	85.71	85.71	92.86	26
9	7.01	26.67	42.22	42.22	45
10	7.06	61.53	69.23	84.62	13
11	22.92	34.29	45.71	45.71	70
12	3.09	66.67	83.33	83.33	12
14	2.73	47.06	58.82	64.71	34
15	10.22	26.83	41.46	51.22	41
16	13.27	37.04	55.56	62.96	27
19	-0.32	50	63.16	76.32	38
21	29.35	37.5	37.5	50	8
22	13.57	46.67	60	66.67	15
23	3.8	76.47	82.35	88.24	17
24	3.3	71.43	85.71	90.48	21
25	1.8	82.35	100	100	17
27.01	3	33.33	50	59.52	42
27.02	2.29	44	72	96	25
27.03	-5.49	46.67	80	86.67	15
28	-5.4	68.42	84.21	89.47	19
29	15.32	40.74	50	57.41	54
32	2.84	55.32	65.96	76.6	47
33	6.11	71.43	85.71	85.71	7
34	4.88	61.9	73.81	73.81	42
35.01	0.58	70	75	77.5	40
35.02	0.53	100	100	100	19
36.02	-0.19	73.33	86.67	93.33	15
37.01	5.18	76.92	82.05	84.62	39
37.02	6.22	45.83	64.58	68.75	48
38	17.31	51.06	55.32	59.57	47
39	-11.72	46.67	60	60	15
44	-2.31	32.73	41.82	47.23	55
45	5.59	38.24	50	64.71	33
46	11.89	34.26	44.29	51.43	70
47	9.87	17.86	42.86	53.57	28
48	10.18	52.94	67.65	73.53	34

Burial	Bias	% +/-5	% +/-10	% +/-15	N
49	7.17	59.38	71.86	78.13	32
50	-0.77	55.17	65.52	79.31	29
51	1.18	39.53	51.17	67.44	43
53	4.69	28.3	37.74	47.17	53
55	-1.53	46.34	65.85	80.49	41
57.01	2.66	21.86	37.5	43.75	32
57.02	12.41	41.38	51.72	62.07	29
58.01	3.49	41.03	52.63	63.16	39
58.02	-1.51	51.43	62.86	68.57	35
59.01	-4.25	25	37.5	37.5	8
59.02	6.83	20	36	40	50
60	4.98	35.42	45.83	63.83	48
61	1.28	20.83	41.67	62.5	24
63	23	35.29	45.1	47.06	51
64	0.12	32.69	44.23	46.15	52
66	24.07	10	38.33	43.33	60
68	17.57	21.42	31.43	37.14	70
71	-6.43	41.38	58.62	62.07	29
72	3.61	64.71	76.47	76.47	17
73	7.27	45.45	54.55	72.73	33
75	-1.77	60	60	60	5
76	-15.13	8.7	8.7	17.39	23
77	-14.04	12.5	25	37.5	8
78	36.99	25	35.71	35.71	28
79	1.75	62.5	75	75	8
80.02	16.19	32.14	42.86	50	27
81	3.57	25.93	37.04	51.85	27
83	-6.55	33.33	66.67	100	3
84	-9.52	11.11	22.22	44.44	9
85	-2.59	80	100	100	5
86	0.46	25.71	37.14	48.57	35
87	3.99	46.86	68.75	78.13	32

Agreement Categorization (Bias)	
0±5	Excellent
±5.01-10	Good
±10.01-15	Moderate
±15.01-20	Poor
±20.01	Very poor

Table 4.11 Fragmentation Bias and percentages of elements within Limits of Agreement per burial, with bias agreement categories noted

Bias per Burial	All Elements	Visibility = 1	Bias + LOA
0±5	14, 23, 24, 59.01, 71 , 75, 79, 83	7, 14, 16, 19, 27.01, 27.02, 28, 33, 35.01, 35.02, 37.02, 39, 44, 48, 49, 50, 51, 55, 58.01, 59.02, 61, 63, 60, 71 , 72, 73, 76, 79, 80.02, 84, 87	23, 71
±5.01-10	25, 27.01, 28, 35.01, 37.02, 39, 55, 53, 61, 76, 77, 84, 85 , 87	9, 10, 22, 24, 25, 27.03, 36.02, 45, 46, 53, 58.02, 59.01, 64, 77, 85 , 86	24, 27.02, 27.01, 28, 36.02, 39, 58.02, 61, 75, 83, 84, 85
±10.01-15	7, 19, 27.02, 35.02, 36.02, 37.01, 44, 49, 51, 57.01, 58.02, 72, 73	11, 12, 15, 23, 32, 34, 37.01, 47, 57.02, 68, 75, 78, 81, 83	7, 11, 19, 35.01, 37.01, 37.02, 44, 46, 47, 49, 50, 51, 55, 57.01, 72, 73, 76, 78, 79, 87
±15.01-20	16, 32, 46, 47, 50, 58.01, 64, 78, 81, 86	--	9, 10, 12, 14, 15, 22, 25, 27.03, 29, 32, 33, 34, 38, 45, 53, 58.01, 59.02, 60, 63, 64, 68, 77, 80.02, 81, 86
±20.01	9, 10, 11, 12, 15, 21 , 22, 27.03, 29, 33, 34, 38, 45, 48, 57.02, 59.02, 60, 63, 66 , 68, 80.02	21 , 29, 38, 66	16, 21 , 48, 59.01, 57.02, 66

Table 4.12 Burial Agreement Categories for Completeness, based on bias calculated using all scored elements; bias calculated using only elements scored as 'Visible' by both observers; and bias (all elements) adjusted for Limits of Agreement percentages

Bias per Burial	All Elements	Visibility = 1	Bias +LOA
0±5	7, 12 , 14, 19, 23, 24, 25 , 27.01, 27.02, 32 , 34, 35.01, 35.02, 36.02 , 44, 50, 51, 53, 55, 57.01, 58.01, 58.02, 59.01, 60, 61, 64, 72, 75, 79 , 81, 85 , 86, 87	10, 11, 12 , 15, 22, 24, 25, 27.02 , 27.03, 29, 32, 33, 35.01, 35.02, 36.02 , 37.02, 46, 49, 53, 58.02, 60, 63, 72, 75, 79, 85, 87	7, 10, 12 , 19, 23, 24, 25, 27.02, 28, 32, 33, 34, 35.01, 35.02, 36.02 , 37.01, 49, 50, 55, 72, 79, 85, 87
±5.01-10	9, 10, 27.03, 28, 33, 37.01, 37.02, 45, 47, 49, 59.02, 71, 73, 83, 84	14, 16, 19, 23, 27.01, 37.01, 38, 45, 47, 48, 51, 55, 58.01, 59.02, 61, 66, 68	14, 27.03, 37.02, 38, 48, 58.01, 58.02, 60, 73, 83
±10.01-15	15, 16, 22, 39, 46, 48, 57.02, 77	7, 28, 34, 44, 50, 59.01, 64, 71, 73, 80.02, 81, 83	16, 22, 27.01, 29, 39, 44, 45, 51, 53, 57.02, 61, 64, 71, 75, 81, 86
±15.01-20	29, 38, 68, 76, 80.02	9, 21, 57.01, 57.02	9, 11, 15, 21, 46, 47, 57.01, 59.01, 59.02, 63, 66, 68, 77, 80.02, 84
±20.01	11, 21, 63, 66, 78	39, 76, 77, 78, 79, 84, 86	76, 78

Table 4.13 Burial Agreement Categories for Fragmentation Ratio, based on bias calculated using all scored elements; bias calculated using only elements scored as 'Visible' by both observers; and bias (all elements) adjusted for Limits of Agreement percentages

Element Type	Completeness		Fragmentation	
	Observed	Estimated	Observed	Estimated
Long Bone	47.39	41.00	10.02	13.69
Short Bone	7.83	3.25	11.16	4.55
Flat Bone	25.72	17.03	8.52	7.16
Irregular Bone	15.41	8.99	13.07	8.37
Mandible	52.76	39.14	11.75	16.60

Table 4.14 Mean Observed and Estimated Completeness and Fragmentation Ratio scores by Element Type

Element Group	Completeness		Fragmentation	
	Observed	Estimated	Observed	Estimated
Cranium	21.15	12.10	10.51	6.69
Axial	8.69	5.75	5.74	3.56
Arm/Shoulder	36.77	30.52	9.10	11.05
Hand/Wrist	9.69	3.62	15.83	5.29
Leg/Pelvis	38.69	32.15	8.99	12.18
Foot/Ankle	16.31	10.47	15.47	10.84

Table 4.15 Mean Observed and Estimated Completeness and Fragmentation Ratio scores by Element Group

	Completeness		Fragmentation	
	Observed	Estimated	Observed	Estimated
Total (all elements)	22.36	15.71	11.09	8.37
Total (Visibility = 1)	65.84	59.48	28.03	31.73

Table 4.16 Mean Observed and Estimated Completeness and Fragmentation Ratio scores for all elements scored, and using only elements scored as 'Visible' by both observers

Element Type	Completeness				Fragmentation Ratio			
	All Elements		Visibility=1		All Elements		Visibility=1	
	N	Bias	N	Bias	N	Bias	N	Bias
Long	683	8.77	559	3.56	683	-5.04	571	-7.77
Short	175	28.18	72	19.85	175	40.49	68	-10.92
Flat	609	17.24	369	4.55	604	2.71	380	-5.44
Irregular	552	23.39	310	9.63	552	17.11	307	-3.98
Mandible	49	18.62	40	12.56	49	-6.64	36	-11.32

Table 4.17 Completeness and Fragmentation Bias by Element Type, calculated using all elements scored, and only elements scored as 'Visible' by both observers

Element Group	Completeness Bias				Fragmentation Ratio Bias			
	All Elements		Visibility = 1		All Elements		Visibility = 1	
	N	Bias	N	Bias	N	Bias	N	Bias
Cranial	498	28.04	254	5.08	496	11.88	249	-5.32
Axial	159	10.96	87	7.08	159	8.11	93	-2.66
Arm/Shoulder	402	9.56	303	5.01	402	-3.04	312	-6.69
Hand/Wrist	147	28.07	49	16.13	147	48.65	46	-6.70
Leg/Pelvis	594	10.34	468	4.01	592	-5.07	479	-9.25
Foot/Ankle	268	20.20	189	13.05	268	16.03	183	-2.62

Table 4.18 Completeness and Fragmentation Bias by Element Group, calculated using all elements scored, and only elements scored as 'Visible' by both observers

	Completeness		Fragmentation Ratio	
	N	Bias	N	Bias
Total (All Elements)	2068	17.04	2064	6.98
Total (Visibility = 1)	1350	6.36	1362	-6.52

Table 4.19 Overall Completeness and Fragmentation Bias, calculated using all elements scored, and only elements scored as 'Visible' by both observers

Independent Variable	Dependent Variable							
	Completeness Bias	Fragmentation Bias	Difference RI	Estimated RI	Estimated Completeness	Estimated Fragmentation	Number of Images	Angles
Position	5.64	2.29	8.63*	6.39	5.99	1.95	8.14*	5.52
Disturbance	8.66	12.02	3.31	10.85	5.12	2.98	0.5	0.06
Integrity	2.1	0.15	3.86	5.1	3.76	0.22	1.04	2.1
Number Interred	3.27	7.38	1.51	1.92	0.13	0.17	17.41*	0.34
Number of Images	21.96	15.54	29.92	42.28	22.41	23.8	--	51.75
Number of Angles	14.23	20.304	45.56	49.98	25.66	20.01	47.61	--

Table 4.20 Percentages of Variance accounted for by categorical Independent Variables in one-way ANOVA tests

Independent Variable	Dependent Variable					
	Completeness Bias	Fragmentation Bias	Diff RI	Estimated RI	Estimated Completeness	Estimated Fragmentation
Observed RI	17.6	18.1	69.9	86.2	16.5	29.4
AI	2.3	0.9	9.3	0.13	0.4	3
Observed Completeness	27	21.5	33.6	43.6	49.2	53.5
Observed Fragmentation	20.9	40.8	45	36.3	45.9	77.1
Diff RI	19.9	18.3	--	19.9	7.4	6.6

Table 4.21 Percentage of Variance accounted for by continuous Independent Variables in single-regression tests

	R-sq	F	df1	df2	p
CMP BIAS	0.728	12.749	12.000	53.000	0.000

	Coeff	SE (HC)	t	p
<i>Constant</i>	-16.106	8.383	-1.923	0.060
Obs RI	-0.158	0.134	-1.180	0.243
Obs FR	-0.099	0.116	-0.8532	0.397
Obs Cmp	0.452	0.165	2.745	0.008
Diff RI	1.018	0.174	5.853	0.000
AI	0.026	0.038	0.678	0.501
Semi-flexed	4.468	4.016	0.565	0.575
Pos Unknown	2.358	4.142	0.569	0.572
Double	0.526	2.600	0.202	0.841
Triple	-6.052	6.514	-0.929	0.357
Secondary/Unknown	34.522	256.033	0.135	0.893
Disturbed	-2.429	3.091	-0.786	0.436
No Img	-0.032	0.320	-0.099	0.913

Table 4.22 Multiple Regression Model 1 - Completeness Bias with all Independent Variables

	R-sq	F	df1	df2	p
FR BIAS	0.723	7.263	13.000	52.000	0.000

	Coeff	SE (HC)	t	p
<i>Constant</i>	-4.372	6.867	-0.637	0.527
Obs RI	-0.2528	0.119	-2.117	0.039
Obs FR	0.4608	0.157	2.779	0.008
Obs Cmp	-0.052	0.130	-0.400	0.6905
Diff RI	0.7022	0.111	6.319	0.000
AI	-0.0304	0.049	-0.625	0.535
Semi-flexed	4.168	2.837	1.499	0.148
Pos Unknown	-7.587	6.241	-1.216	0.230
Double	-0.800	1.789	-0.447	0.657
Triple	-3.542	4.625	0.766	0.447
Secondary/Unknown	21.490	320.053	0.067	0.947
Disturbed	-5.324	2.622	-2.030	0.047
No Img	0.9021	0.522	1.729	0.090
Obs Frag Avg	0.0736	0.055	1.343	0.185

Table 4.23 Multiple Regression Model 2 - Fragmentation Bias with all Independent Variables

	R-sq	F	df1	df2	p
DIFF RI	0.565	12.412	5.000	60.000	0.000

	Coeff	SE (HC)	t	p
<i>Constant</i>	14.159	5.645	2.509	0.015
Obs RI	0.344	0.053	6.442	0.000
Obs FR	0.117	0.065	1.799	0.077
Obs Cmp	-0.249	0.083	-2.991	0.004
AI	-0.009	0.031	-0.276	0.783
No Img	-0.630	0.309	-2.037	0.046

Table 4.24 Multiple Regression Model 3a - Difference in RI scores with continuous Independent Variables only

	R-sq	F	df1	df2	p
DIFF RI	0.605	6.171	11.000	54.000	0.000

	Coeff	SE (HC)	t	p
<i>Constant</i>	13.772	6.190	2.225	0.030
Obs RI	0.373	0.073	5.128	0.000
Obs FR	0.107	0.070	1.520	0.134
Obs Cmp	-0.281	0.089	-3.157	0.003
AI	0.007	0.044	0.163	0.871
Semi-flexed	-3.538	2.094	-1.690	0.097
Pos Unknown	-1.400	4.200	-0.333	0.740
Double	1.134	3.038	0.373	0.710
Triple	-2.118	2.386	-0.888	0.379
Secondary/Unknown	9.136	34.874	0.262	0.794
Disturbed	1.558	2.167	0.719	0.475
No Img	-0.672	0.430	-1.564	0.124

Table 4.25 Multiple Regression Model 3b - Difference in RI scores with all Independent Variables

	R-sq	F	df1	df2	p
DIFF CMP	0.115	1.342	5.000	60.000	0.259

	Coeff	SE (HC)	t	p
<i>Constant</i>	7.568	4.602	1.645	0.105
Obs FR	0.100	0.082	1.223	0.226
Obs Cmp	0.017	0.075	0.2309	0.818
AI	-0.022	0.043	-0.505	0.615
N vis=1	0.029	0.117	0.251	0.803
No Img	-0.417	0.341	-1.221	0.227

Table 4.26 Multiple Regression Model 4a - Difference in Completeness scores with continuous Independent Variables only

	R-sq	F	df1	df2	p
DIFF CMP	0.270	0.718	11.000	54.000	0.717

	Coeff	SE (HC)	t	p
<i>Constant</i>	7.495	5.194	1.443	0.155
Obs Cmp	0.006	0.083	0.074	0.941
Obs FR	0.060	0.081	0.736	0.465
AI	-0.009	0.043	-0.201	0.841
N Vis=1	0.076	0.101	0.744	0.460
Semi-flexed	-1.699	4.243	-0.400	0.690
Pos Unknown	-1.205	2.853	-0.423	0.674
Double	-2.830	2.025	-1.398	0.168
Triple	-1.803	3.380	-0.533	0.596
Secondary/Unknown	21.706	224.026	0.097	0.923
Disturbed	0.319	2.040	0.157	0.876
No Img	-0.285	0.352	-0.809	0.422

Table 4.27 Multiple Regression Model 4b - Difference in Completeness Scores with all Independent Variables

	R-sq	F	df1	df2	p
DIFF FRAG	0.972	75.203	6.000	59.000	0.000

	Coeff	SE (HC)	t	p
<i>Constant</i>	-9.182	2.187	-4.199	0.000
Obs Cmp	0.163	0.036	4.5886	0.000
AI	-0.020	0.020	-1.022	0.311
N vis=1	-0.029	0.048	-0.610	0.545
Obs Frag Avg	0.921	0.060	15.265	0.000
No Img	-0.712	0.215	-3.381	0.002

Table 4.28 Multiple Regression Model 5a - Difference in Fragmentation Ratio scores with continuous Independent Variables only

	R-sq	F	df1	df2	p
DIFF FRAG	0.979	47.193	12.000	53.000	0.000

	Coeff	SE (HC)	t	p
<i>Constant</i>	-6.050	2.815	-2.149	0.036
Obs Cmp	0.116	0.045	2.578	0.022
AI	-0.028	0.023	-1.200	0.235
N Vis=1	0.037	0.075	0.494	0.623
Pos Unknown	1.441	1.836	0.785	0.436
Semi-flexed	-2.625	2.235	-1.174	0.245
Double	-1.773	1.656	-1.070	0.289
Triple	4.854	2.024	2.398	0.020
Secondary/Unknown	-3.562	60.039	-0.059	0.953
Disturbed	3.841	1.600	2.401	0.020
Obs Frag Avg	0.910	0.047	19.200	0.000
No Img	-1.039	0.275	-3.777	0.000

Table 4.29 Multiple Regression Model 5b - Difference in Fragmentation Ratio scores with all Independent Variables

	R-sq	F	df1	df2	p
FR Left	0.409	1505.605	1	2177	0.000

	Coeff	SE	t	p
Constant	4.823	0.389	12.406	0.000
FR Right	0.644	0.029	21.863	0.000

Table 4.30 Single Regression Model for predicting Fragmentation Ratios of right-sided elements from the Fragmentation Ratios of left-sided elements

	R-sq	F	df1	df2	p
Cmp Left	0.601	3275.141	1	2178	0.000

	Coeff	SE	t	p
Constant	6.049	0.473	12.782	0.000
Cmp Right	0.777	0.016	49.829	0.000

Table 4.31 Single Regression Model for predicting Completeness of right-sided elements from the Completeness of left-sided elements

	R-sq	F	df1	df2	p
FR Left	0.692	1198.299	1	533	0.000

	Coeff	SE	t	p
Constant	1.832	0.498	3.680	0.000
FR Right	0.831	0.045	18.678	0.000

Table 4.32 Single Regression Model for predicting the Fragmentation Ratios of right-sided cranial elements from the Fragmentation Ratios of left-sided cranial elements

	R-sq	F	df1	df2	p
Cmp Left	0.806	2220.032	1	533	0.000

	Coeff	SE	t	p
Constant	1.810	0.469	3.857	0.000
Cmp Right	0.884	0.028	31.358	0.000

Table 4.33 Single Regression Model for predicting the Completeness of right-sided cranial elements from the Completeness of left-sided cranial elements

4.2 Discussion

At first glance, the results appear to indicate that estimated completeness scores do not agree particularly well with the observed data, while estimates of fragmentation seem more reliable; however, a deeper analysis of the data indicates that the opposite is actually true. In fact, when Visibility does not act as a confounding factor, completeness estimates were found to generally be fairly close to observed scores. On the other hand, the way that FR is calculated introduces multiple opportunities for inaccuracy to occur – either through inaccurate completeness estimates, inaccurate fragment counts, or both – and these are then masked by the nature of the equation used to transform raw fragment counts into the FR variable.

4.2.1 Visibility

Overall disagreement between observers, calculated from discordant pairs using the crosstabs procedure, was 14.35% (Table 4.1). When elements were grouped by side, disagreement for un-sided elements was 16.59%, 14.4% for left-sided elements, and 13.34% for right-sided elements (Table 4.2). Grouped by element type, disagreement was calculated at 13.45% for long bones, 10.54% for short bones, 21.08% for flat bones, 12.79% for irregular bones, and 13.43% for mandibles (Table 4.3). Results by element group were similar, with 16.36% disagreement between observers for cranial elements, 15.54% for axial elements, 16.42% for arm and shoulder elements, 14.87% for hand and wrist elements, 13.65% for leg and pelvic elements, and 9.18% for foot and ankle elements (Table 4.4).

Using the interpretive scale developed for assessing agreement between observers regarding completeness and fragmentation biases (Table 3.1), the overall agreement exhibited between observers can be described as moderate. The majority of categorical groups showed moderate to poor agreement, with the exception of foot and ankle elements, which at 9.18% disagreement indicated generally good agreement between observers. At the burial level, seven

burials displayed disagreements of 0–5%, indicating excellent agreement, while a further 14 burials showed good agreement, with disagreements of 5.01-10%. Of the remaining 45 burials, 10 showed moderate agreement, with disagreements of 10.01-15%, while the remaining 35 showed poor (15.01-20%) to very poor (>20.01%) agreement (Table 4.5).

Reliability tests using Krippendorff's Alpha (α) generally reflected the percentage disagreements, with higher α values correlating with lower percentage disagreements (figure 4.1). However, there was some variability within this trend, as occasionally burials with a low percentage difference (and therefore high agreement) exhibited low reliability scores. Where this difference occurred, it appeared to be related to the total number of visible elements recorded; for example, α for Burial 33 was 0.364, despite a percentage difference of only 7.59% (see Table 4.5). However, there were only seven visible elements recorded within this burial; and of those seven elements Lieverse and Urlacher actually agreed on just two, so while this burial exhibited high agreement on elements scored 0 for visibility, the agreement on elements scored 1 was actually relatively low, resulting in a low reliability score.

Overall, reliability was fairly low, even with a generous cut-off of 0.75 as an acceptable α reliability score. Only 16 out of 66 burials (Table 4.5), and one element group (foot and ankle bones, Table 4.4) exhibited an α of above 0.75, while α reliability for all elements was 0.674 (Table 4.1). As with the individual burials, individual categories for element type, element group, and element side generally followed the pattern of higher α scores for lower percentage disagreements, with some deviations. Again, this result is likely related to the distribution of marginal totals, as described above.

Difference in Representation Index scores (Difference RI) was highly correlated with both Observed RI and α . Regression scatterplots showed a strong negative correlation between

Difference RI and α ($R^2 = .491$) (Figure 4.1), and a strong positive correlation between Observed RI and Difference RI ($R^2 = .489$) (Figure 4.2). The negative relationship between Difference RI and α indicates that high α -values are generally associated with lower levels of disagreement between Observed and Estimated RI scores, while the positive correlation between Difference RI and Observed RI suggests that the potential for disagreement between observers regarding element visibility rises as the number of elements present in a burial increases:

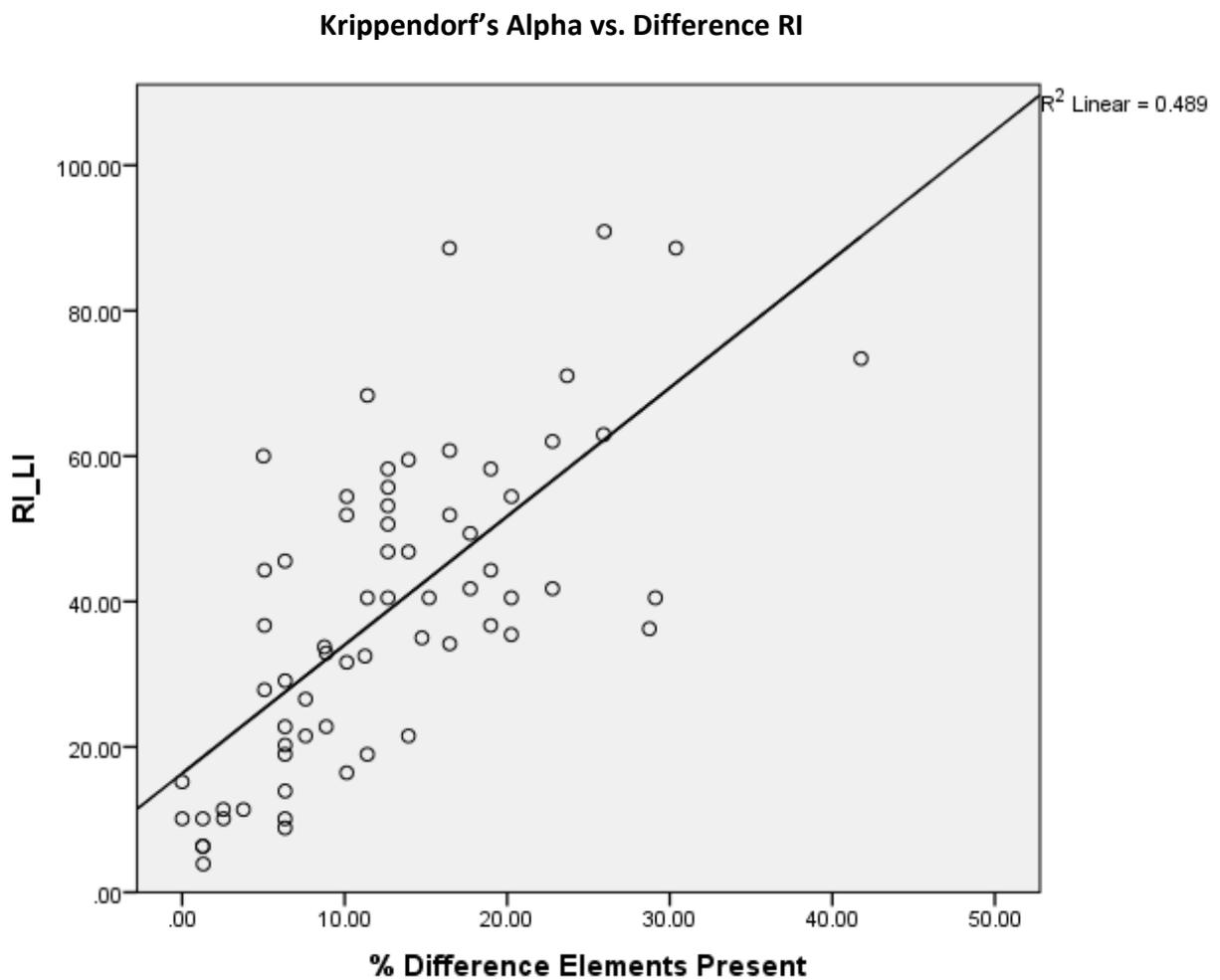


Figure 4.1 Scatterplot with regression line showing the relationship between Krippendorff's alpha values and the difference in Observed vs. Estimated RI scores by burial

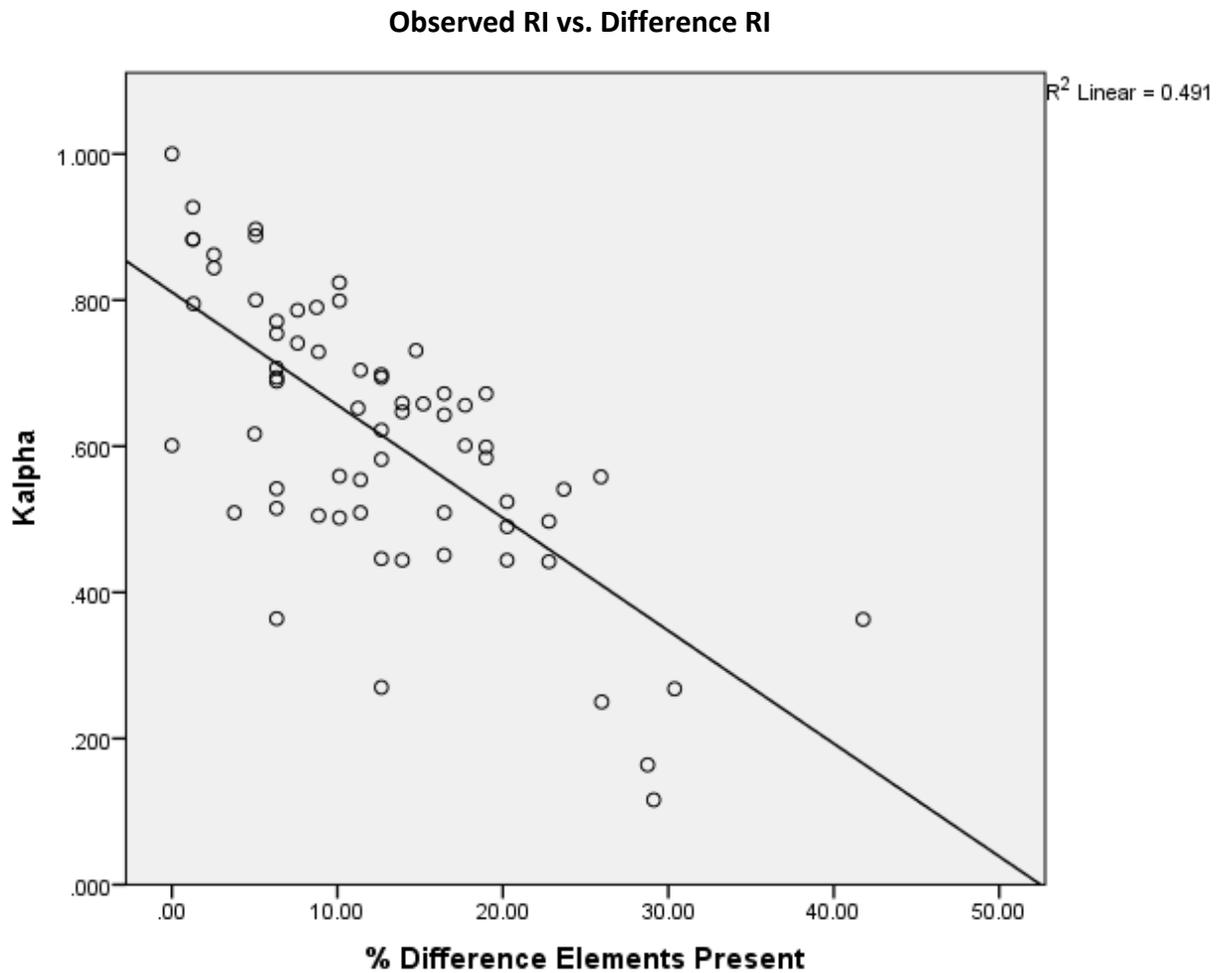


Figure 4.2 Scatterplot with regression line showing the relationship between Observed RI scores and the difference in Observed vs. Estimated RI scores by burial

However, these relationships are not necessarily straightforward, and may be moderated to some extent by certain factors. This moderation is reflected by differences between slopes when separate regression lines are plotted for each level of a categorical variable. For example, when Difference RI is regressed against α , the slopes for each level of the categorical variables Position, Disturbance, and Integrity, while all oriented in the same direction differ slightly in terms of their angles (see Appendix B). This observation indicates that the although the directionality of the relationships between the Difference RI α is the same across all categories of

these variables, the strength of the relationship varies somewhat depending on the exact body position, disturbance, or integrity of a burial. However, in the regression scatterplot for Number Interred, while the slopes for Single and Double burials are roughly parallel, suggesting little difference in the effect of the relationship between α and Difference RI on these two categories, the slope for triple burials runs in the opposite direction, indicating that at this level of the variable the typical relationship between α and Difference RI is reversed (Figure 4.3).

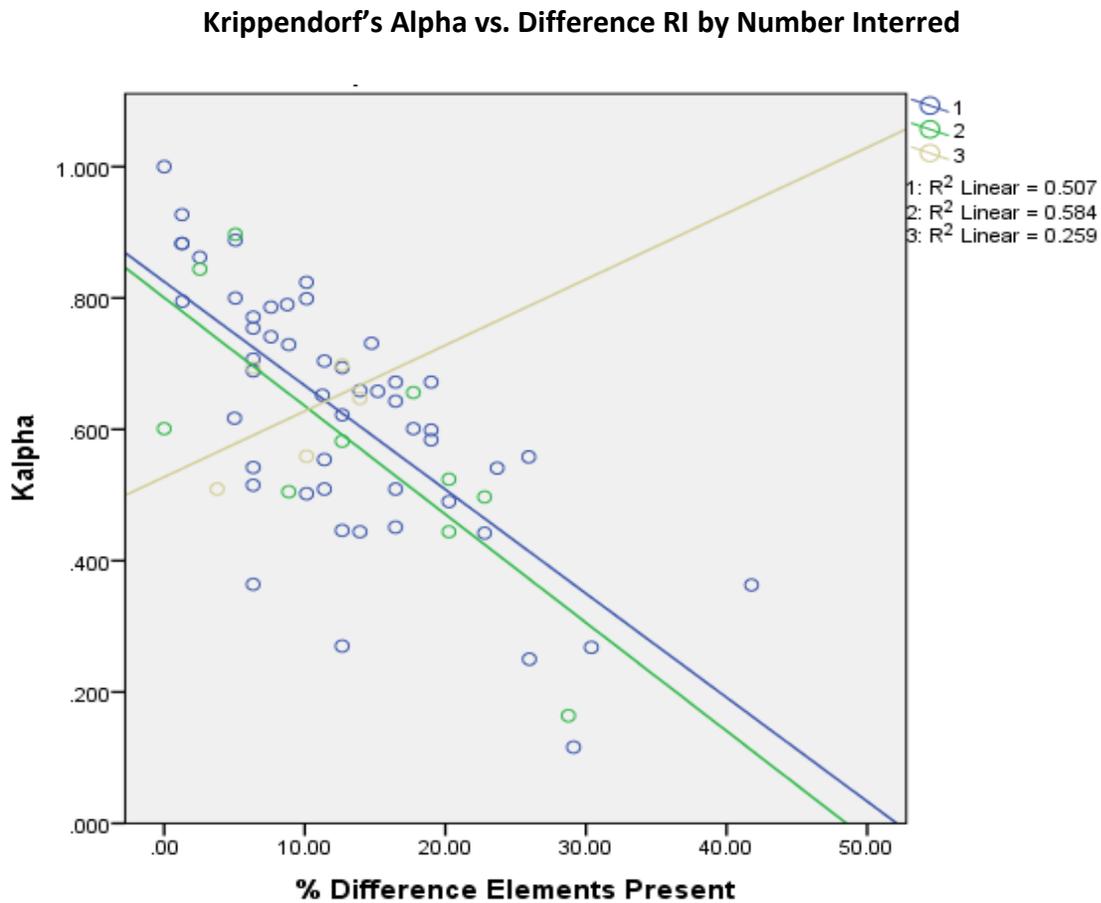


Figure 4.3 Scatterplot with multiple regression lines showing interactions between Krippendorff's Alpha values, difference in RI scores, and categorical levels of the variable 'Number Interred', by Burial

4.2.2 Agreement - Completeness

The average difference (bias) in Estimated Completeness and Observed Completeness was 17.04 across all elements. The positive score indicates that Observed Completeness scores were consistently higher than Estimated Completeness scores (negative scores would indicate that estimated scores were higher than observed scores). An average bias of 17.04 indicates poor agreement overall, an observation which is reflected in the fact that only 53.35% of scores were observed to exhibit a difference within the range of 0 ± 15 , with 44.24% exhibiting a difference of 0 ± 10 , and 34.41% a difference of 0 ± 5 . Just 10.06% of all pairs (208 out of a total of 2066) exhibited zero difference in Observed vs. Estimated Completeness scores (Table 4.6).

In terms of element types, mean Estimated Completeness was highest for long bones, followed consecutively by mandibles, flat bones, irregular bones, and short bones; while the first two categories were reversed for Observed completeness, with mandibles exhibiting higher average completeness than long bones (Table 4.14). While 61.05% of long bones were estimated to within 0 ± 10 of their original completeness score, this outcome was true for only 44.9% of mandibles. In addition, the completeness scores of the same proportion of mandibles (44.9%) were significantly underestimated, exhibiting a difference between Estimated Completeness and Observed Completeness scores of +10.01 to +100 (Table 4.6).

These differences are likely due to the particular shape of the mandible and how it appears in photographs, features which can make it difficult to assess completeness accurately relative to long bones. The mandible, unless very fragmented, is generally easy to identify as it is quite distinct in both size and shape; and it is therefore difficult to confuse with any other cranial elements, especially when it is in the correct anatomical position. However, the shape and position of the mandible when it is articulated often meant that it was only partially visible in photographs; thus completeness could only be estimated from the visible portion, leading to

inaccuracy. In contrast, long bones were rarely obscured; and because of their size (with the exception of the fibula, which would often be entirely obscured by the tibia), when they were partially obscured by other elements, the majority of the element would still be visible, allowing for an estimation of completeness to be made that was still relatively accurate.

Although short bones exhibited both the lowest Estimated and the lowest Observed Completeness scores (Table 4.14), they were also the element type with the largest bias (Table 4.19), indicating that the level of agreement at the individual element level was actually very poor. The high level of disagreement at the individual element level indicated by the bias does not appear to be due to disagreement over whether or not these elements were visible, as short bones also had the lowest percentage disagreement of all element types in terms of Visibility (Table 4.3). Rather, frequency tables indicate that the completeness of a large proportion of short bones was significantly underestimated, with 77.71% of all short bones displaying a difference of +10.01 or higher, and only 15.43% estimated to within 0 ± 10 of the corresponding Observed Completeness score (Table 4.6).

This result is likely due to the fact that although it was generally possible to identify short bones when they were present in photographs, they were often difficult to see in detail unless they were the subject of a close-up image, or they appeared in close-up images of nearby elements. In addition, phalanges, particularly of the hand, were often disarticulated and somewhat obscured, typically by other manual elements, a situation which added to the difficulty in accurately estimating their completeness. This observation also explains the relatively larger bias for irregular bones compared to other element types, as many irregular bones, such as the carpals, are also small and thus difficult to see; or positioned in such a way that they are easily obscured by nearby elements.

Completeness biases for each element group were 28.04 for cranial bones, 10.96 for axial bones, 9.56 for bones of the arm and shoulder, 28.07 for bones of the hand and wrist, 10.34 for bones of the leg and pelvis, and 20.20 for bones of the foot and ankle (Table 4.20). Perhaps unsurprisingly, element groups composed primarily of long bones (i.e., the arms, legs, and ribs) exhibited much lower mean differences than element groups with higher proportions of short and irregular bones. Long bones, due to their size and density, tended to exhibit a higher overall level of preservation (Lieverse, 1997). As discussed above, their size also generally meant that long bones were very easily identified in photographic images, even when not articulated, and this feature made it much easier to estimate accurately their completeness relative to certain other elements.

Although these results indicate generally poor agreement overall, the bias may be somewhat misleading in terms of how accurately the completeness of individual elements was estimated, due to the influence of disagreements between observers regarding the visibility or presence of elements. Considering only elements visible to both observers (i.e., only those elements given a visibility score of 1 by both observers), completeness bias dropped significantly across all element types and groups (Table 4.17, 4.18) The overall completeness bias fell from 17.04 to just 6.36 (Table 4.19), moving from poor to good agreement. For long bones and flat bones, the bias fell to less than 5, indicating excellent agreement; and for irregular bones, the bias fell to 9.63, indicating good agreement. The bias also decreased for mandibles and short bones, though less significantly, with the agreement for short bones rising from very poor to poor, and increasing for mandibles from poor to moderate. Similarly, by element group, Completeness Bias decreased to less than 10 for cranial, axial, arm and shoulder, and leg and

pelvic elements, with a somewhat less significant decrease for elements of the hand and wrist and of the foot and ankle (Table 4.17, 4.18).

The fact that the biases indicate very little disagreement in the Estimated and Observed Completeness scores for most groups of elements when observers agree that the elements in question are present suggests that a large proportion of the overall bias is a result of differences in the perception of Visibility, causing extreme disagreements in Estimated vs. Observed Completeness scores, an outcome which skews the bias upwards. The position of the cranium within a burial often rendered all of the bones on one side of the skull completely invisible, leading to scores of 0 for both Visibility and Estimated Completeness, even when the elements in question were in fact present and 100% intact. In contrast, short bones and mandibles were much less likely to be obscured in this way; and erroneously given a Visibility score of 0 despite being present in a burial. For these elements, it appears that extreme differences in Estimated vs. Observed Completeness scores were instead caused by an inability to see the elements in enough detail to estimate completeness accurately, even though they were easily identified as being present within an image. Similarly, at the burial level, when all elements were considered, only 8 burials exhibited a bias of 0 ± 5 , indicating excellent agreement between observers; but this number rose to 31 burials when only elements which observers agreed were present were used to calculate the bias (Table 4.12). This result again suggests that discrepancies in the number of visible elements play a large role in determining the bias, rather than differences in estimates of completeness at the individual element level.

As Table 4.12 shows, the level of agreement by burial changed depending on how the bias was calculated and interpreted. When the bias was calculated only from elements visible to both observers, the vast majority of burials exhibited excellent to moderate agreement (i.e., a bias

within 0 ± 15). Using all scored elements to calculate the bias, approximately half of the total number of burials exhibited excellent to moderate agreement, with the remainder falling into either the poor or very poor agreement categories. When agreement was assessed based on a combination of the bias and the total proportion of elements with differences of 0 ± 15 (Table 3.2), roughly the same number of burials exhibited excellent to moderate agreement; but a large number of burials were placed into a different category relative to that which they were assigned using the bias alone. In trying to account for biases affected by large numbers of extreme differences by looking at the distributions of these differences in addition to the bias, some burials with excellent agreement in terms of their bias alone were re-classified as exhibiting good, moderate, or even poor agreement and vice versa.

4.2.3 Agreement - Fragmentation

Overall agreement in terms of Fragmentation Ratio (FR) was good, with a bias of 6.98 across all elements ($n=2064$; Table 4.21). While only 3.88% of cases showed zero difference between observations, 37.4% exhibited excellent agreement, displaying differences within the range of 0.01 ± 5 . In total, 61.3% of cases displayed differences between Estimated FR and Observed FR that fell within the range 0 ± 15 , indicating excellent to moderate agreement (Table 4.7).

By element type, short bones contained the highest proportion of elements with zero difference between FR scores (8%), but long bones exhibited the highest proportion of elements with a difference within 0.01 ± 5 (51.2%), even though only 1.32% of long bones exhibited a difference of 0. In fact, long bones showed the best agreement overall, with 75.1% of these elements showing a difference within 0 ± 15 . In contrast, only 26.3% of short bones fell into this

range. Flat bones exhibited the second highest level of agreement, followed by mandibles, and then irregular bones (Table 4.7).

By element group, the leg and pelvis, and the arm and shoulder showed very similar agreement, with almost equal proportions of cases exhibiting a difference in FR scores within the range of 0 ± 5 , and falling within the range of $\pm 5.01-15$. The hand and wrist group of elements contained the lowest proportion of elements within the range of 0 ± 5 , as well as the smallest proportion of cases exhibiting a difference within the range of 0 ± 15 . Following elements of the leg and pelvis and the arm and shoulder, cranial elements exhibited the next highest proportion of cases within the ranges of 0 ± 5 and 0 ± 15 , followed by axial elements, and elements of the foot and ankle (Table 4.9).

Similar to with Completeness Bias, mandibles and long bones tended to show the smallest difference between Estimated and Observed FR, exhibiting smaller FR Biases than short and irregular bones, which had the highest. In general, with the exception of short bones and elements of the hand and wrist, all element groups exhibited much smaller biases for FR than they did for Completeness, a result which was unexpected, as estimating FR was considerably more difficult than estimating Completeness. In addition, the biases were generally positive, a result which was again unexpected as it was assumed that fragments would be less visible in photographs, which would lead to over-estimation of FR, and that this would result in a negative bias (Table 4.17, 4.18). As well, when comparing mean Estimated and Observed FR scores by Element Type and Element Group, the differences tended to be smaller than those seen when comparing mean Completeness scores (Table 4.14, 4.15).

However, when biases were recalculated using only elements which both observers had scored as being visible, the bias for each category switched to the expected negative bias,

suggesting that fragment counts were actually being under-estimated (Table 4.17, 4.18, 4.19). As was seen with the data for Completeness, the presence of certain elements being scored as not visible in photographs when they were scored as present in the original dataset led to a number of extreme disagreements. It is these disagreements that make it appear, when all elements are used to calculate the bias, as if the Observed FR scores are larger than the Estimated FR scores, when generally the opposite is true. Similar to the effect seen with Completeness scores, it appears that when elements were clearly visible in photographic images, Estimated FR scores were in fact fairly close to the Observed FR scores.

An additional factor affecting FR Bias that must be taken into account is the fact that FR is a function of an element's Completeness score, and so the accuracy of an estimated FR is affected by the accuracy of the corresponding estimate of completeness for that element. In general, one would assume that individual fragments would be more difficult to see in a photograph than individual elements, leading to under-estimation of the number of fragments; and thus over-estimation of the FR. Indeed, when only elements scored as visible by both observers are taken into account, this result is generally the case; however, under- or over-estimation of the total number of fragments *or* the completeness of an element both affect the accuracy of the estimated FR. Thus, the apparently higher overall accuracy of these estimates may actually be the result of a quirk in the FR calculation method that allows two very different estimates of completeness and/or numbers of fragments to produce very similar fragmentation ratios. Essentially, the nature of the calculation means that it is much easier to arrive by chance at an accurate figure for estimations of FR than it is to do so with estimates of completeness.

For example, an element estimated as being 50% complete, and consisting of 20 fragments, would produce an FR of 2.5. If the element's true completeness score was 90%, and

the actual number of fragments was 40, this calculation would produce an FR of 2.25. This figure results in a difference between Observed and Estimated FR scores of just 0.25, a number which indicates excellent agreement, even though the Estimated Completeness and the number of fragments counted are actually very different from the Observed scores. While this is not to say that this phenomenon is occurring in every instance in which there is very little difference between the estimated and the real fragmentation ratios, a comparison of a random sample of 100 elements indicates that the observed number of fragments counted was frequently substantially higher than the number estimated from photographic images, even when completeness estimates were similar (see Appendix A1). For example, the difference in FR for the left femur in Burial 12 is only 1.13, even though the real number of fragments counted is almost four times higher than the estimated fragment count; and the difference in completeness scores is only 10.

This outcome occurs because, as the number of fragments gets closer to, and exceeds, the number by which it is being divided (the corresponding Completeness score), the amount by which the FR can shrink gets smaller and smaller. The result of this is that seemingly very small differences in the Estimated vs. Observed FR can obscure what are actually often very large differences in the actual number of fragments counted by each observer. This effect is then compounded when the individual FRs for each element are summed together and then divided again to create an average FR for the entire skeleton. Therefore, although excluding cases with disagreements regarding visibility of elements appears to make the FR Bias more accurate in the sense that it then conforms more closely to expectations, these results nevertheless cannot be assumed to be as reliable as those that were seen for Estimated Completeness. It is recommended that Estimated FR be viewed more as a “ballpark” figure of preservation rather than an accurate estimate of true fragmentation, as these results indicate that as fragmentation increases beyond a

certain point, it may be close to impossible to count accurately the number of fragments present in a photographic image. This disconnect is illustrated by the table in Appendix A2, which indicates that the size of the difference between observed and estimated fragment counts appears to have very little influence on the size and direction of the Bias for any given burial.

Keeping this in mind, at the burial level, 47 burials exhibited excellent (bias of 0 ± 5) or good (bias of $\pm 5.01-10$) levels of agreement when FR Bias was calculated using all elements (Table 4.13). Using only elements which observers agreed were visible to calculate the bias, 44 burials fell into these two categories; however, only 29 of these burials were the same as those that fell into these categories when all elements were used to calculate the bias. Finally, using both the bias (calculated using all elements) and the total proportion of elements with differences within the range of 0 ± 15 (see Table 3.1) to categorize agreement, 33 burials exhibited excellent or good agreement, with the majority of remaining burials divided evenly between the moderate and poor agreement categories. Again, a number of burials were assigned to different agreement categories using this method than when other methods of assessing agreement were used. However, with this method, the distribution of burials between categories was more similar to that of the first method (elements visible to both observers only), while the individual burials that appeared in each category were more similar to the second method (bias calculated from all elements).

4.2.4 ANOVA & Multiple Regression

One-way ANOVA of categorical variables (Table 4.20) indicated that position, disturbance, integrity, and number interred were responsible for approximately 20% of the variance altogether in both Completeness Bias and FR Bias. In both cases, Number of Images and Angles were responsible for a much larger proportion of the total variance than other

variables². For Completeness Bias, Number of Images accounted for almost 22% of the variance alone, while Angles accounted for approximately 20% of the variance in fragmentation bias. OLS regression of each continuous variable (Table 4.21) indicated that AI was not a particularly large contributor to the variance of either Completeness Bias or FR Bias based on the value of R^2 , but all of the other continuous variables each accounted for between 17 and 40 percent of the variance, depending on the dependent variable.

As a documentation-dependent variable, the effect of Angles in terms of the accuracy of estimated completeness and fragmentation ratios was also of interest, and this factor was assessed via means plots. A plot of mean Estimated and mean Observed Completeness per angle category indicated that while completeness was generally somewhat over-estimated regardless of the number of angles, the differences between mean Observed and mean Estimated Completeness were reduced when images were taken of a burial from three or more different angles (Figure 4.4).

² The nature of the variable number of images means that it can act as either a categorical or a continuous variable. While each number of images acts as a distinct, unambiguous category, the fact that the numbers are ordered and meaningful numeric values means that it can also act as a continuous variable in multiple regression models. This is not the case with number of angles. With this variable, categories are ordered in the sense that categories with more angles are given higher numbers than categories with fewer angles; however, the numbers assigned to each category are not meaningful in and of themselves. For example, categories 8, 9, and 11 for the variable number of images contain burials with 8, 9, or 11 images associated with them, while the same-numbered categories in number of angles all contain burials with different combinations of 4 types of angle, and so it cannot act as a continuous variable.

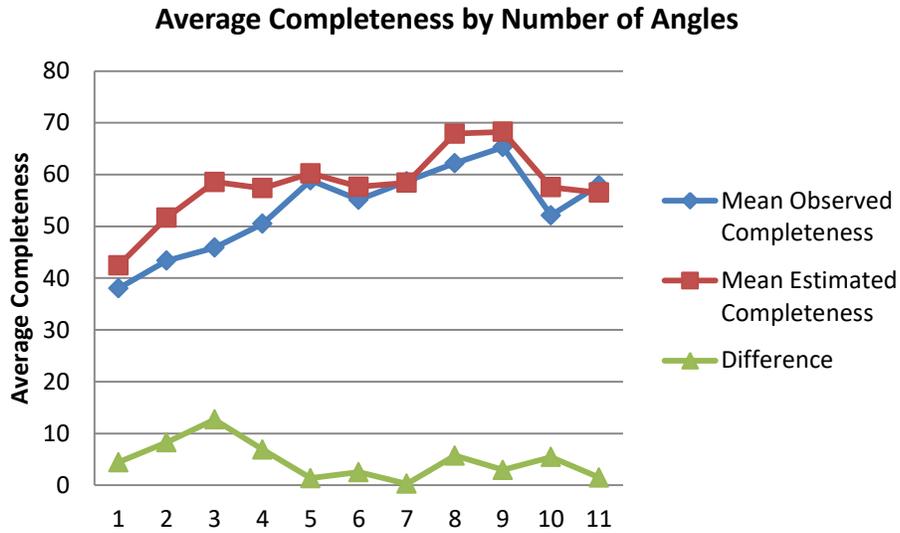


Figure 4.4 Means Plot showing Average Completeness by Number of Angles

A similar plot for Observed and Estimated mean fragment counts exhibited somewhat less of an identifiable pattern in terms of the number of angles and the size of the difference between the Observed and Estimated means, exhibiting a number of peaks and troughs within a slight downward trend (Figure 4.5). However, large peaks in the plot make it difficult to say with confidence if any real pattern exists.

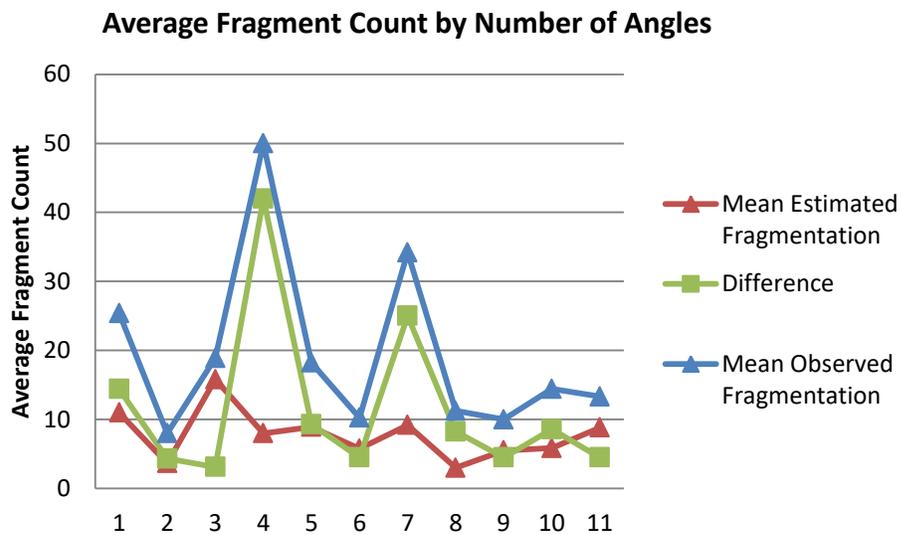


Figure 4.5 Means Plot showing Average Fragment Count by Number of Angles

These peaks, caused by burials with very high Observed fragment counts, are naturally associated with large differences between Estimated and Observed fragment counts; and thus potentially serve to obscure any other trends that may be present. Removing the three largest peaks from the means plot for Number of Angles (Figure 4.6) does not provide any further enlightenment; rather, it appears that in the absence of the outliers, which seem to create the appearance of a slight downward trend, the difference between average observed and average estimated fragment counts fluctuates between approximately four and ten regardless of the number or combination of angles.

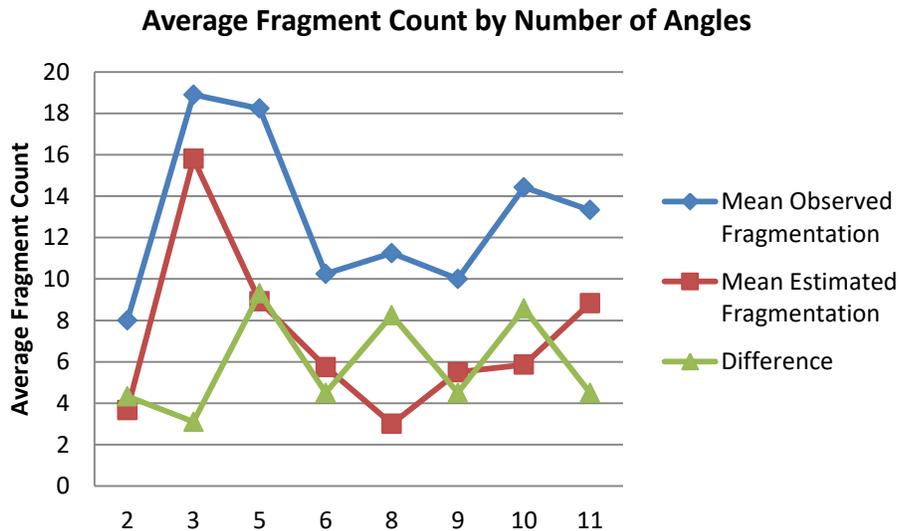


Figure 4.6 Means Plot of Average Fragment Count by Number of Angles, with outliers removed

Despite the low proportion of variance that categorical variables accounted for in ANOVA, they tended to exhibit the largest regression weights in multiple regression (MR) models. However, their effects in terms of statistical significance appears to reflect the results of the ANOVA models, although low statistical significance should not necessarily be taken as an indicator of a lack of real-world effect. For example, in Models 1 (DV = Cmp Bias; Table 4.22) and 2 (DV = FR Bias; Table 4.23), triple burials were associated with a decrease in the bias

relative to the reference category (single burials) of 6.052 and 3.542 respectively (reference categories are not included in the model and their regression weights are therefore assumed to be zero). Although the model indicates that the *statistical* effect is negligible (sig. = .357 in Model 1, and .447 in Model 2), in real-world terms, these regression weights are clearly very different from zero; and represent what would be a noticeable difference between categories. They should not necessarily be disregarded as an unimportant difference on the basis of statistical significance alone. While the categorical variables may not have a statistically significant impact on either Completeness Bias or FR Bias when taken as a whole, the differences in regression weights between different levels of categorical variables nevertheless indicate that differences in body position, the number interred, whether or not a burial is a primary interment, and whether or not it is disturbed can all have noticeable effects on the predicted bias.

While most of the continuous variables accounted for a large proportion of the variance singly, R^2 for Models 1 and 2 was somewhat lower than might be expected, given how much of the variance each of the variables appeared to explain on their own (Table 4.21). This result indicates the presence of collinearity amongst the variables (Field, 2005), suggesting that there may be a number of interactions occurring between variables when they are all used to predict Completeness Bias or FR Bias. In addition, integrity appeared to have a disproportionately large effect on the bias in both models. In Model 1, secondary burials or burials of unknown integrity caused completeness bias to increase by 34.522 relative to primary burials; and in Model 2 the same variable caused fragmentation bias to increase by 21.490. However, the standard errors for this variable were much higher than they were for any other variable (256.033 and 320.053 respectively), indicating that while we may be able to assume that the nature of the relationship

between this variable and the dependent variables indicated by these regression weights is correct, the regression weights themselves may not be particularly reliable.

The large standard errors are likely due to the uneven distribution of cases amongst the levels of this variable, as 63 of 66 burials were classified as primary burials, leaving just two secondary burials and one of unknown integrity. Because these two categories contained only 3 cases between them, they were collapsed into a single category for the purposes of analysis in order to avoid any levels of a variable containing only a single case, as this can make some standard statistical tests difficult to run and to interpret. While it is perhaps reasonable to accept from these results that primary burials may be associated with a lower bias than secondary or burials of unknown integrity, the size of the standard errors indicates that the actual value of the regression weights for this particular variable should be regarded with caution.

In both Model 1 and Model 2, difference in RI was the most significant predictor of bias. Interestingly, Number of Images appeared to have almost no effect on Completeness Bias, and to result in a larger FR Bias, both of which were unexpected. In Model 1, while increases in Observed RI and Observed FR were associated with a decrease in bias, Observed Completeness had a significant positive effect on Completeness Bias, indicating that higher levels of average skeletal completeness were associated with a higher bias. A similar effect occurred in Model 2, with higher Observed RI and Observed Completeness values both resulting in a decrease in the bias, while higher Observed FR was associated with an increase.

The results for these models seem somewhat counter-intuitive, as one might expect a more complete and/or well-preserved burial (in terms of individual element completeness and fragmentation) to be easier to visually assess, thus leading to more accurate estimates (and thus a smaller bias). At least part of this relationship is likely due simply to the fact that biases are

calculated using both Observed and Estimated Completeness and FR values, and so these variables will naturally share a somewhat linear relationship with the variables calculated from them. Interactions between Difference RI and Observed Completeness in Model 1, and Observed FR in Model 2 may also affect the relationship between observed variables and their corresponding biases in the sense that a greater difference in RI may naturally result in greater differences between observed and estimated averages for Completeness and FR.

A similar interaction may also explain why, in Model 2, a greater number of images associated with a burial appear to result in an increased bias. Generally, more complete, well-preserved burials have more images associated with them than burials containing few, highly fragmentary elements. Models 3*a* (Table 4.24) and 3*b* (Table 4.25), with Difference RI as the DV, suggest that the difference between observed and estimated RI increases as Observed RI rises – an effect that is obscured in both Models 1 and 2. Based on this observation, it is possible that the apparent effect of Number of Images on FR bias in Model 2 is actually a result of the fact that better preserved burials tend to have both more images associated with them *and* a higher observed RI – a factor which results in higher biases due to the resulting increased differences between observed and estimated RI.

As Models 1 and 2 establish, differences between observed and estimated RI appear to be the primary contributors to the bias for both Completeness and FR. Model 3 shows the effects of the IVs from the first two models when Difference RI is the DV, both with (Model 3*a*) and without (Model 3*b*) categorical variables. In Model 3*a*, considering only continuous IVs, Observed RI is a significant predictor of Difference RI, with higher Observed RI resulting in larger differences between observed and estimated RI; however, both Observed Completeness and Number of Images are associated with significant decreases in this difference. In Model 3*b*,

the addition of the categorical variables results in a slight increase in the regression weight for Number of Images; however in this version of the model it is no longer statistically significant.

In term of the effects of categorical variables on Difference RI, both semi-flexed burials and burials of unknown position are associated with smaller differences in RI than their reference category, supine burials. In the case of burials of unknown body position, the smaller difference is likely a result of the fact that the body position was generally recorded as 'unknown' due to poor preservation, resulting in lower observed RI; and thus smaller differences between Observed and Estimated RI. As a well-preserved semi-flexed burial would presumably result in a number of elements being obscured due to the position of the body on its side, resulting in a greater difference in RI, we can assume that a similar effect is occurring with this category as occurs for unknown burials; otherwise the associated regression weight would likely be positive rather than negative.

With regard to the number of individuals interred, double burials are associated with an increased Difference RI over the reference category (single burials), while triple burials were associated with a decrease. The positive regression weight for double burials is likely a reflection of the fact that elements contained within a grave featuring a multiple interment may be more likely to be obscured due to commingling. As well as making it difficult to see individual elements, this situation can also make it challenging to assign a given element to the correct individual. While these factors obviously also apply to triple burials, and in fact we might expect the chances of error to increase the more individuals are present in a grave, the triple burials at KNXIV were both associated with large numbers of images, a factor which may have been enough to counteract this effect, resulting in a negative rather than a positive regression weight.

Finally, both disturbed burials, and burials of secondary or unknown integrity, were associated with increases in Difference RI over their reference categories, undisturbed and primary burials. Although the standard error for the integrity category is much smaller in this model than in Models 1 and 2, it is nevertheless still significantly larger than the standard errors for the other variables, so the previous caveat regarding this particular regression weight still applies here. It is perhaps also worth noting that, with the exception of the semi-flexed category, none of the effects of the categorical variables approach statistical significance; and due to their relatively small regression weights their real-world effect is likely also quite small.

In order to remove the effect of differences in RI on other IVs, for Models 4 and 5, the DVs, Difference Completeness and Difference FR, were calculated using only elements to which both observers had assigned a score of 1 for Visibility. Observed Completeness and FR for each burial were also recalculated using this criterion; and a new IV, representing the total number of valid elements per burial, was added to the model in the place of RI. As with Model 3, Models 4 and 5 were run first with continuous variables only, then again with categorical variables added, resulting in *a* and *b* versions of each model. In addition, while the values of the bias can be positive or negative, depending on whether observed or estimated values are higher, for models 3, 4, and 5 all the values of the DV were converted to positive integers in order to make the effect of the IVs easier to interpret in terms of the magnitude of their effects.

With the effects of RI removed, neither Model 4*a* (Table 4.26) nor Model 4*b* (Table 4.27) was significant, with the value of R^2 greatly reduced in both cases compared to that of Model 2. Although none of the continuous variables in Model 4*a* were significant, the value of the coefficient for Number of Images is worth noting. Recall that in Model 2, the effect of number of images on Completeness Bias was negligible, with a coefficient of -0.032 . However, with the

effects of RI removed, the coefficient for Number of Images is noticeably larger, indicating a decrease in Difference Completeness of .417 for each additional image. Although still not statistically significant in this model, the coefficient is large enough that it would likely have a noticeable, albeit small, real-world effect.

When the categorical variables are added into the model, the effects of the continuous variables are reduced further, while the effects of the categorical variables appear to be similar to those seen in Model 3*b* (Table 4.25). The coefficients for both body position categories are negative, indicating that, relative to supine burials, semi-flexed and unknown position burials are associated with smaller differences in completeness. As with Model 3*b*, this result may be due to the fact that these types of burials typically contain fewer elements, a feature which makes them easier to see, thus increasing the potential accuracy of estimates. For double and triple burials, the negative coefficients likely reflect a similar phenomenon to that observed for triple burials in Model 3*b*, in which the increased number of images associated with these burials may counteract the effects of increased numbers of elements and the potential for commingling to cause elements to become obscured. Again, as none of the IVs in this model are statistically significant, any real-world effects on the DV would likely be small and not particularly noticeable.

In Models 5*a* (Table 4.28) and 5*b* (Table 4.29), with Difference FR as the DV, the value of R^2 was significant, indicating that the effects of RI are not as large for estimates of FR as they are for estimates of Completeness. Indeed, this outcome is also evident when comparing the coefficients for Difference RI in Model 2 and Model 1, so we can be confident that this assumption is likely correct. In Model 5*a*, the only significant IV is Observed Completeness, with the coefficient indicating that Difference FR increases slightly as Observed Completeness rises. It is difficult to come to an explanation for why this might be the case, given the

problematic nature of FR comparisons discussed above, especially when we consider that elements with high Completeness scores would generally be expected to be less fragmented, which should make FR easier to accurately estimate. It should also be noted that in Model 5b, the addition of the categorical variables causes the coefficient for Observed Completeness to become non-significant, indicating that some of the variance in this variable is accounted for by the categorical variables.

In both Model 5a and 5b, as with the two variants of Model 4, there is a noticeable difference in the effect of Number of Images once the effect of Difference RI is removed. In Model 2, Number of Images appears to cause an increase in FR Bias of 0.902 for every additional image associated with a burial, a result which is the opposite of what would be expected; i.e., that more images should lead to greater accuracy and smaller biases. However, when we examine the absolute difference between average Observed and Estimated FR, with the effect of differences in RI removed, this relationship reverses, with each additional image resulting in a 0.594 *decrease* in Difference FR. Again, although the coefficients are not statistically significant, they are likely large enough to have a small but noticeable real-world effect on the size of the difference between observed and estimated scores. In addition, the effect of the average Observed fragment count is also much larger and more significant than it is in Model 2; and the regression coefficient indicates an effect much more in line with what might be expected; i.e., that the difference between Observed and Estimated fragment counts rises as the number of Observed fragments increases.

With regards to the categorical variables, in Model 5b, while the model itself is statistically significant in terms of how much of the variance in the DV it explains, the only significant IV is that of disturbed burials. The coefficient for Disturbed indicates that disturbed

burials exhibit an additional difference of 7.031 in Observed vs. Estimated FR compared to undisturbed burials. Burials of unknown body position exhibit a similarly large difference relative to supine burials, but it is not statistically significant. Most of the remaining categorical coefficients are so small that their effects would be negligible, with the exception of that for Secondary/Unknown, although, as in the other models, the large standard error associated with this variable calls the accuracy of the coefficient into question.

Finally, two additional OLS regression analyses were carried out in order to investigate the viability of predicting the completeness and fragmentation ratio of non-visible paired elements from the scores of their visible partners. The results for FR indicate generally poor accuracy, with an R^2 of .409 for all elements (Table 4.30) and .692 for cranial elements only (Table 4.32). Given the previously discussed problems with estimating fragmentation ratio from photographs, assuming the fragmentation of non-visible elements would not be recommended based on these results. For completeness, OLS regression with all elements resulted in an R^2 of .601 (Table 4.31), and an R^2 of .806 for cranial elements only (Table 4.33). This outcome indicates that it may be reasonable to assume the completeness scores of non-visible paired cranial elements from the completeness scores of their visible partners, as this assumption is likely to be correct roughly 80% of the time.

4.3 Summary of Results

Visibility

Total disagreement across all elements with regard to visibility was 14.35%, indicating moderate agreement. In terms of element groups, the foot and ankle exhibited the lowest level of disagreement, while the arm and shoulder exhibited the most. Short bones exhibited the least

disagreement in terms of element type, and the highest disagreement was exhibited by flat bones. Disagreement between observers regarding the presence or visibility of elements was highly correlated with observed RI, with higher RI's associated with more disagreement between observers. However, scatterplots with separate regression lines for each level of the four main categorical variables (position, number interred, disturbance, and integrity) indicate that the relationship between RI and visibility disagreement may differ in strength and direction, depending on the values of other variables.

Slightly fewer than half of all burials (31 out of 66) exhibited excellent to moderate agreement, and burials with high agreement generally also exhibited high α -reliability scores. When agreement was high but reliability was low, or vice-versa, this result appeared to be related to the marginal distributions of visible and non-visible elements in cross-tabs tables. Mean α across all burials was relatively low at .674, with only 16 burials exhibiting an α of .75 or higher.

Agreement

Average agreement in terms of completeness was poor when calculated using all elements; however, it improved significantly when calculated using only elements which were scored as visible or present by both observers (i.e., visible both in the field and in photographs). This effect occurred with the overall average completeness scores as well as across every element group and type category. This observation indicates that completeness estimates are generally accurate; however, extreme differences between observed and estimated completeness arise when observers disagree on the visibility of elements, and this disagreement negatively impacts the overall average.

Estimates of completeness by category generally reflected observed completeness scores in terms of the element groups and types with the highest and lowest completeness scores.

Generally, completeness was estimated more accurately for long bones and groups of elements that contained more long bones than other element types. Some elements, like mandibles, were found to be more likely to have their completeness estimated inaccurately due to their shape, while others, like short bones, were more likely to be scored incorrectly as a result of their size.

For FR, overall agreement was good, but the measure itself was found to be problematic in that the calculation used to transform raw fragment counts into the FR often obscured large differences between observers in the actual number of fragments counted. Using all elements to calculate fragmentation ratio agreement resulted in an average bias that indicated that observed FRs were higher than estimated FRs, while using only elements that were scored as visible by both observers resulted in the opposite. Since it was assumed that estimated fragment counts would be lower than observed fragment counts, it was expected that estimated FRs would be higher; therefore average agreement from the second calculation is assumed to be a somewhat more accurate reflection of reality.

Similar to the results for completeness, long bones exhibited the highest agreement by element type and short bones the lowest, while the leg and pelvis and arm and shoulder element groups exhibited the highest agreement and the hand and wrist the lowest. Approximately half to three-quarters of burials exhibited good to excellent agreement, depending on which method of classifying the bias was used. A greater number of burials exhibited good to excellent agreement on fragmentation ratio than they did for completeness; however, the fragmentation ratio scores were also assumed to be much less reliable.

ANOVA & Multiple Regression

Results of one-way ANOVA indicated that the four main categorical independent variables (IVs) – Position, Number Interred, Disturbance, and Integrity– accounted for approximately 20% of the total variance in both Completeness Bias and FR Bias. For other

dependent variables (DVs) – Observed Completeness, Observed FR, Observed RI, and Difference RI – the categorical IVs were responsible for a similarly small amount of the total variance, although Position was a significant factor by itself when Difference RI was the DV. Number of Images and Angles both had a large influence on the variance of all DVs; and when Number of Images was used as the DV, Number Interred exhibited a significant influence on the variance.

In terms of continuous variables, AI was not an important factor in the variance, no matter which variable was used as the DV. When entered singly into OLS regression models, Observed Completeness, Observed RI, Observed FR, and Difference RI exhibited varying degrees of influence; and were typically responsible for 15– 20% or more of the variance in any given DV. Despite this difference, MR models tended to exhibit lower R^2 values than might be expected when all variables were entered into the models together. This observation indicates the presence of collinearity amongst IVs, which affects the strength and direction of their relationships with the DV, depending on the combination of variables entered into a given model.

Difference RI exhibited the largest influence on both Completeness Bias and FR Bias in terms of statistical significance, while the values of regression coefficients for burials of Secondary or Unknown Integrity were found generally to be unreliable due to large standard errors. Models 1 and 2 exhibited seemingly illogical relationships between Completeness Bias and Observed Completeness in Model 1, and the FR Bias and Observed FR in Model 2. However, this problem is likely due at least in part to the pre-existing linear relationship between these variables that stems from the way the biases were calculated.

MR models with Difference RI as the DV indicated that Number of Images, Observed RI, and Observed CMP had the greatest influence on Difference RI. Higher Observed RI was associated with greater Difference RI; but this affect may be offset by the effects of Observed Completeness and Number of Images, both of which were associated with a decrease in Difference RI. Effects of categorical variables on Difference RI were uniformly small and statistically insignificant, indicating that a greater number of images, and high completeness scores are likely the most important factors in reducing the difference between observed and estimated RI scores.

Removing the effects of variables related to RI seen in Models 1 and 2 resulted in MR models for Difference Completeness that were not statistically significant. This result suggests that differences in RI are the primary contributing factor to large positive or negative biases when it comes to estimating average skeletal completeness, as many of the remaining variables from Model 1 appear to have little to no effect on Difference Completeness. With the effects of RI removed, Number of Images appeared to have a much larger, although still statistically insignificant, effect on differences between observed and estimated completeness scores. This observation indicates that the number of images associated with a burial likely has more of an effect on the accuracy of completeness estimates than it appears to in Model 1. As Angles had too many categories to be included in MR models, a means plot of average Observed Completeness and average Estimated Completeness per Angle category was constructed. The means plot showed a noticeable reduction in the difference between observed and estimated completeness scores when three or more angle types were associated with a burial. These effects appear to support the conclusions drawn from the agreement analysis that, differences in

visibility aside, estimates of element completeness are generally accurate as long as image quality is sufficient.

Given the problematic nature of FR scores discussed above, the difference in average fragments counted per element (Difference Fragmentation) was used as the DV for Models *5a* and *5b*. This substitution of variables was done in order to try and remove the potentially confounding influence of completeness scores used in the calculation of FR. Both models were highly significant, with R^2 values of .972 in Model *5a* and .979 in Model *5b*. These results suggest that, while differences in RI are a significant contributor to FR biases, they are not the only influential factor. In Model *5a*, the regression coefficient for average Observed Fragments per Element (Observed Fragmentation) is .921, indicating that the difference in average observed and estimated fragment counts increases at almost a 1:1 ratio with the average observed fragment count. Number of Images has a significant negative effect on Difference Fragmentation, but it does not fully compensate for the effect of Observed Fragmentation in reducing the difference. Additionally, higher scores for Observed Completeness appear to be associated with an increase in Difference Fragmentation, a result which was unexpected because elements with high completeness scores would generally be expected to exhibit low fragmentation; however, it is possible for elements with high completeness scores to also be highly fragmented. In such cases the greater overall number of fragments present – as a result of the bone itself being more complete – would likely result in a larger Difference Fragmentation score, and this effect may be what the regression coefficient for Observed Completeness reflects.

In Model *5b*, the addition of categorical IVs does not increase the R^2 by a significant amount, which suggests that their effects on the DV in a general sense are rather small. However, the coefficients for both Triple burials and Disturbed burials exhibit significant p values, and are

associated with increases in Difference Fragmentation. This result may reflect lower preservation (and thus higher observed fragment counts) in these types of burials, commingling obscuring elements and making accurate estimated fragment counts difficult, or a combination of both factors. With the addition of the categorical IVs, the effect of Number of Images on the DV appears to increase, suggesting that a greater number of images significantly increases the accuracy of estimated fragmentation counts despite the numerous other factors that act to reduce it. In terms of the effect of Angles on the size of the difference between Observed and Estimated fragment counts, means plots indicated that there is no identifiable effect, even when large, outlying values are removed.

Finally, OLS regression to determine how well visible paired elements predict the completeness and fragmentation on their non-visible counterparts indicate that assuming the completeness of paired cranial elements may be reasonable. Results indicate that the completeness of non-visible paired cranial bones can be predicted with approximately 80% accuracy, but the accuracy for other areas of the body is likely to be lower. Fragmentation ratio cannot be accurately predicted for cranial bones or for any other region of the body.

4.4 Research Questions

1). Can the same elements that were recorded as being present during excavation be seen in the photographic documentation?

Across all elements, disagreement between observers in terms of visibility was 14.35%. Most of the disagreements occurred when the position of the body made certain elements impossible to see in photographs. This was a particular problem with cranial elements, especially the flat bones of the skull; and is reflected by higher disagreement scores for the cranial element group and flat bone element type. In general terms, the answer to this question is yes, as a

disagreement rate of 14.35% means that observers did agree on visibility of elements 85.65% of the time. The fact that body position causes certain elements to become obscured in photographic images should be kept in mind during the photographing and removal of skeletal elements from a grave. A simple solution for this issue would be for the persons responsible for this aspect of photography during excavation to consider taking multiple sets of images of a skeleton *in situ* as each layer of bones is removed and previously obscured elements become visible.

2.) Are estimates of completeness based on photographs similar to those based on direct viewing elements during excavation?

The overall Completeness Bias (calculated using all elements) of 17.04 would indicate that the answer to this question is ‘no;’ however, when only elements which both observers agreed were present were considered, the overall bias was much smaller, as were the biases for all sub-groups of elements. This comparison indicates that, as long as elements are clearly visible in photographic documentation, estimates of their completeness will be relatively accurate. This conclusion is supported by the fact that the most accurate estimates of completeness tended to be made for larger elements, which are generally more visible in photographs. Smaller elements, while they may have still been visible in an image, were not always clearly depicted enough to make an accurate assessment of their completeness.

3.) Are estimates of fragmentation based on photographs similar to those based on direct viewing elements during excavation?

On the surface, it appears that the answer to this question is ‘yes’, as observed and estimated fragmentation ratios appeared to exhibit less difference than observed and estimated

completeness scores; however, as with Completeness Bias, FR Bias was heavily influenced by differences in the number of visible elements. In addition, the way fragmentation ratio is calculated means that estimated fragmentation ratios are also strongly influenced by estimated completeness scores. With elements that are highly fragmented, this condition can lead to large differences in the raw fragment counts being obscured when these numbers are transformed into fragmentation ratios. While completeness estimates of clearly visible elements can generally be assumed to be reliable, fragmentation ratios are perhaps best thought of as educated guesses at the true level of fragmentation, as it seems that they are just as likely to be accurate by chance as by design.

4.) What factors affect whether or not elements can be seen or identified in photographs?

The primary factors that affect visibility of elements, represented in the difference between Observed and Estimated RI scores, appear to be Observed RI, which causes the difference to increase; and Observed Completeness and Number of Images, both of which cause the difference to decrease. Higher Observed RI may result in less accurate estimates of RI due to a higher likelihood of elements obscuring one another when more are present in a grave; however, this effect appears to be offset to some degree in burials where Observed Completeness scores are also high. As well, differences in observed and estimated RI appear to be significantly reduced when there are a larger number of images associated with a burial.

5). What factors affect the accuracy of estimations of completeness based on photographic documentation of burials?

The removal of the effects of RI resulted in models that were not statistically significant when Difference Completeness was the DV. This result suggests that differences in RI may be the most influential factor when it comes to Completeness Bias, reflecting the above conclusion that completeness estimates are generally accurate as long as an element is visible. When working at the burial level rather than the individual element level, since Difference RI appears to be the primary driver of high positive or negative Completeness Bias scores, factors that reduce the difference between average observed and average estimated RI may also result in a reduction of the difference between average observed and average estimated completeness scores. These factors include low RI, high Completeness scores, and a greater Number of Images associated with a burial. The number of angles used to document a burial also appears to be fairly important, as a means plot indicated a reduction in the difference between Observed and Estimated completeness scores when three or more angles were used in documentation.

6). What factors affect the accuracy of estimations of fragmentation ratio based on photographic documentation of burials?

The use of the calculation to transform raw fragment counts into fragmentation ratios was found to be a significant source of inaccuracy, as the influence of completeness scores seems to obfuscate significant disagreements between Observed and Estimated fragment counts. In terms of the accuracy of raw fragment counts, a greater number of Observed fragments per element was associated with a significant increase in the DV Difference Fragmentation. Triple burials and disturbed burials were also significantly associated with increases in Difference Fragmentation, while Number of Images appeared to be the only factor that significantly reduced differences between Observed and Estimated fragment counts. A greater number of angles in the

documentation of burials did not appear to be associated with a decrease in Difference Fragmentation, with a means plot indicating that even with the removal of high outlying fragmentation count values, the number or combination of angles has no identifiable effect on Difference Fragmentation.

7). Can estimations of the completeness of skeletal elements on one side of the body be used to predict or infer the completeness of elements on the other side of the body that are known to be present, but which are not visible in photographs?

OLS regression of right-sided elements on left-sided elements using observed completeness data indicates that, overall for paired elements, the completeness of an element on one side of the body predicts the completeness of the other with 60% accuracy; however, regression conducted using only paired cranial elements, for which this question is the most relevant, indicates that the completeness of elements on one side of the skull predicts the completeness of those on the other side 80% of the time. While this result can be taken as an indication that it may be reasonable to assume the completeness of non-visible, paired cranial elements based on the completeness of visible ones, such assumptions should nevertheless be made with caution, and will not be as likely to be reliable in areas of the body beyond the skull and cranium.

8). Can estimations of the fragmentation ratio of skeletal elements on one side of the body be used to predict or infer the completeness of elements on the other side of the body that are known to be present, but which are not visible in photographs?

OLS regression using observed fragmentation ratio data indicated that the fragmentation of one paired element could be predicted from its mate only 41% of the time overall, and 69% of the time for cranial elements. While predictive accuracy of almost 70% might be considered reasonable by some researchers, assuming the fragmentation ratio of non-visible elements is not recommended, due to the above mentioned difficulties with accurately estimating fragmentation ratio from photographic images.

Chapter Five: Conclusions and Future Research Considerations

5.1 Conclusions

In the above study, photographs of in-situ human burials from the Siberian site KNXIV were assessed for completeness and fragmentation on an element-by-element basis, and the scores compared with similar data recorded in the field at the time of excavation between 1997 and 2001. Analysis of the two datasets showed that estimates of completeness are generally accurate to within 0 ± 10 , unless elements are very small, or are obscured in some way. Additionally, it was found that when estimating completeness for bones of the skull, the completeness of one paired bone predicts the completeness of its mate approximately 80% of the time. In other words, when one side of the skull is not visible in a photograph, the researcher can assume that the completeness of paired elements that cannot be seen is approximately the same as those that are visible with an 80% chance of being correct.

In terms of fragmentation ratio, it was found that not only is it very difficult to accurately count the number of fragments present in elements that are very fragmented, but that the transformation of raw fragment counts into the fragmentation ratio often obscures this inaccuracy. Because of this obfuscation, it may be preferable to use fragmentation ratios as a “ballpark” measure of preservation when working from photographs; or instead to simply report the raw number of fragments identified. As was seen with completeness estimates, the accuracy of fragmentation estimates was reduced when elements were small, or if they were obscured. Because completeness scores were used in the calculation of fragmentation ratios, the accuracy of completeness estimates also had an effect on the accuracy of estimated fragmentation ratios. When the effects of completeness estimates were removed, the number of fragments present was

also found to strongly influence the accuracy of estimates, with the accuracy of estimated fragment counts decreasing as the level of fragmentation increased.

With regard to how visual documentation of burials affected estimates of preservation, for both completeness and fragmentation, the number of images associated with a burial was found to significantly reduce the differences between observed and estimated scores, indicating that having a greater number of images allows for preservation of skeletal elements to be estimated with greater accuracy. The number of angles from which a burial was documented was also taken into account. For estimates of completeness, the relationship between photographic angles and accuracy of estimated scores was straightforward; where three or more angles were captured in the photographic documentation, a general downward trend was observed in the difference between observed and estimated scores. The relationship between angles of documentation and fragmentation was more difficult to ascertain, as the presence of some extremely fragmented burials inflated the average difference between observed and estimated scores for some angle categories, resulting in large outliers. Both with and without the outliers, there appeared to be no identifiable pattern in how the average difference changed between categories, although when the outlying categories were removed, it appeared that documentation that included images of isolated elements or body regions may be associated with smaller differences between observed and estimated fragment counts.

Given that one of the most important factors affecting the accuracy of both estimated completeness and estimated fragmentation was related to differences in the visibility of skeletal elements, the most obvious suggestion to improve such estimates is to ensure that photographs are taken of in-situ burials not only prior to their removal from a grave, but throughout this process as well. Duda and Guillion (2006) recommend removing skeletal elements from graves

in layers, re-documenting the burial with drawings and photographs every time a new set of previously-obscured elements are fully exposed by the removal of those above. Following this procedure then allows researchers to “superimpose the consecutive vertical drawings and to rebuild a synthetic view of the site,” (Duday and Guillon, 2006: 123); however, the drawing of complex features like burials is time-consuming, and risks damage to fragile skeletal remains through exposure to the elements while a burial is re-drawn multiple times. Taking photographs as each layer of elements is removed, however, should generally be a relatively quick and simple procedure that does not create significant delays in removing skeletal elements from the ground.

As well as ensuring greater accuracy in studies of preservation, visually documenting features like human burials more thoroughly increases the utility of such documentation beyond its primary function as a simple archival record of the excavation process. Depending on a researcher’s goals and interests, archaeological photography can be used in a number of different ways, examples of which were presented in Chapter 1. Indeed, recent technological advances allow researchers to enhance digital images using various computer software, giving users the ability to geo-reference images and embed them in site maps; to rectify images so that they are true-to-scale and can be used for metric studies; and to cheaply and easily create detailed 3D representations of sites and features that can be used in public displays. A small number of these post-processing techniques are presented below, with a view to demonstrating how slight changes to photographic documentation workflows in the field and the use of fairly simple computing processes can allow archaeologists to use digital photography to its full potential.

5.2 Future Considerations in Archaeological Photography

As noted in Chapter 1, the utility of archaeological photography extends far beyond its simply acting as a visual accessory to written descriptions of excavations. Indeed, as

photographic and computer technologies continue to improve and to become more accessible in terms of cost, the possibilities for and utility of visual documentation of archaeological sites will only increase. Many computer programs now offer researchers the ability to manipulate digital images across a wide spectrum of potential adjustments aimed at improving image quality, from removing lens distortion and increasing the dynamic range of single images to combining multiple 2D images and building 3D models from them. Many of these methods use software that is freely or cheaply available and which automate most of the image processing, thus requiring very little adjustment to photographic workflows typically used in the field during visual recording of excavation processes.

Photogrammetry, the practice of determining the geometric properties of objects within a photographic image, is commonly used in forensic contexts. Using photogrammetric techniques, researchers can determine the lengths of objects in photographs, and the distances and angles between them, features which have obvious applications to mortuary archaeology in terms of measuring skeletal elements for metric studies. González-Jorge et. al. (2012; 2013) demonstrated that a simple photogrammetric technique, known as single image rectification (SIR), could be effectively utilised in the field with measurement errors of less than 6% when compared to “gold standard” measurements obtained using a terrestrial laser scanner. The process described by González-Jorge et. al. (2012; 2013) utilised a carbon-fibre cross with target points at the end of each arm, which was placed on a flat surface within the area to be photographed. The coordinates of the target points were measured using a Total Station; and these control points were then entered along with the photograph into MATLAB, and a specially developed algorithm was then applied to correct perspective distortions introduced by the camera lens (González-Jorge et. al., 2012; 2013).

While the most accurate results for this method were obtained using a telescopic apparatus constructed by the researchers themselves that allowed them to take aerial images from a height of up to 5 m, the use of known measurements in the form of the target points means that this method does not require the use of specially-calibrated cameras. This method allows SIR to be used with widely available DSLR cameras, and with zoom lenses and autofocus if need be. In addition, as the name implies, SIR requires only a single image, meaning that although the cross needs to be placed in the frame and the position of the target points recorded, this task only needs to be done once for each SIR image; thus the amount of time needed for photography is not extended by an unreasonable amount.

In situations where metric measurements are not required, but researchers want to improve lighting, contrast, and detail in digital images for archival, publication, or research purposes, Watson and Weiland (2013) suggest the use of High Dynamic Range (HDR) imaging. This technique involves ‘stacking’ multiple exposures of the same image in Adobe Photoshop or similar programs in order to increase the dynamic range of an image to more closely resemble how a scene would be perceived by the human eye in terms of detail and resolution. As it is already common practice when taking photographs in the field to ‘bracket’ images by taking three to five exposures of the same frame, this technique should not require any modifications to the photographic workflow.

To create HDR images, three to five exposures of the same image are imported into photo-editing software and ‘stacked’, essentially combining them into a single image that incorporates the entire tonal range of bracketed photographs used to create it. Technically, these images have a higher tonal range than can be captured by the camera’s sensor, and than a computer monitor or printer can reproduce. Fortunately this issue can be resolved by a procedure

called tone-mapping, wherein the photo-editing software applies an algorithm which selectively edits the dynamic range of the stacked image on a region-by-region basis in order to utilize the tonal information in such a way that it provides the most detail possible (Watson and Weiland, 2013).

Watson and Weiland (2013) recommend HDR imaging for burials in particular, as well as for other complex features, as the process creates higher contrast and highlights more details because of how the HDR process captures a wider spectrum of illumination. In complex features like human burials, the number and arrangement of skeletal elements and other objects can make individual items difficult to identify; and this method acts to mitigate the resulting potential data loss by increasing the available detail and resolution of digital images. As well, the combination of multiple exposures of a single image can also have a compensatory effect in cases for which the available lighting was poor.

Watson and Weiland (2013) also note that in addition to stacking digital photographs to create HDR images, multiple HDR images can also be stitched together horizontally to create large, detailed composite images of complex features like burials. Stitching images together in photo-editing software is generally a fairly simple procedure, although the exact details will vary depending on the program used. In order to ensure that composite images are of the best possible quality, the authors suggest that when photographs are taken of an object or feature with the intention to stitch them together later, such photographs are taken so that each frame overlaps the preceding frame by approximately one third. This procedure should require only two to five images, as images composed of more than five frames may end up being too large for publication. Additionally, when stitching together multiple HDR images, the authors note that it

is essential that the same tone mapping settings are used on every HDR image that is intended for use in the final composite image (Watson and Weiland, 2013).

As well as using such methods for enhancing single images for greater clarity and resolution, in recent years a number of archaeological researchers have investigated the utility of 3D models built from photographic images as a faster, cheaper and more portable alternative to terrestrial laser scanners (TLS) for the purposes of creating 3D models of features and sites. While TLS and other ranged data collection systems produce very precise, detailed 3D models, these systems are often expensive; and can be cumbersome to transport into the field. In addition, TLS models require a high level of computing power, meaning that they generally cannot be processed in the field, leaving researchers uncertain about the quality of any scans taken until they are processed on a computer. As well, use of TLS in the field requires an intimate understanding of the geography of a given site, including any exposed architecture or features as well as other contextual information like the location of burials, etc. (Gudjeran and Warden, 2012).

In addition to processing techniques that enhance single images, recent advances have made 3D imaging much more accessible to researchers who may be laypeople in terms of the skills required to capture and build 3D models of archaeological sites or features. In their 2011 Paper, Ducke, Score, and Reeves offer a detailed description of the workflow they have developed for constructing 3D models from multiple overlapping digital photographs using largely open-source, freely available software. The authors demonstrate the efficacy of the technique by presenting an example of a 3D reconstruction of a mass burial excavated in Weymouth, England, which was then used in a public display about the site. Although this process uses multiple computer programs to build 3D models, the majority of the work can be

fully automated; however, the authors do note that when capturing images intended for use in 3D modelling, purposeful photographic practice is important. To this end, they recommend that, when taking photographs for use in a 3D model the photographer walk around the object or feature in question and overlap each frame by approximately 25° - 30°. This technique will ensure that a wide variety of angles and viewpoints are captured, which will provide as much visual data as possible, ensuring a more accurate 3D representation (Ducke, Score, and Reeves, 2011).

Ducke, Score, and Reeves' (2011) workflow first requires images to be imported into a program called SIFT, available for educational use from the University of British Columbia. SIFT extracts features from the input images by identifying thousands of points in each image. A second program, Bundler, is then used to match the points identified in each image by SIFT; and to reconstruct the camera settings and identify the depths of objects and shadows. A third program, Patch-Based Multiview Stereo (PMUS), is then used to generate a dense point cloud from this information. Finally, the point cloud data is imported into a fourth program, Mesh Lab, which allows the user to perform surface reconstruction, colour transfer, and manual data cleanup as needed to generate the final 3D model. An optional additional program, ParaView, can also be used to manipulate and animate models once they are built.

A similar workflow is used at Çatalhöyük, a Neolithic site in Central Anatolia, Turkey, which since 2009 has been the subject of an experimental project aimed at recording every phase of an excavation in 3D, with the goal of essentially making the excavation process 'reversible' in order to mitigate (to the extent possible) the inherently destructive process that is archaeological investigation (Forte, 2012; 2014). By using a combination of laser scanning, photogrammetry, and 3D stereo visualization, the Çatalhöyük project has been able to produce maps, sections,

profiles, and volumetric analyses of the site and its various features with a high degree of accuracy. In addition, the use of digital images to create 3D models has allowed researchers on the project to develop a digital workflow that allows 3D models to be created within hours of data being captured, resulting in 3D visualizations that can be built and viewed within a single day's work (Forte 2012; 2014).

While images built using computer vision (as the process for building models from digital images is known) sacrifice some of the geometric precision offered by laser scanners, the difference is small, with 3D models built using computer vision offering accuracies of 3-5 mm, compared with accuracies within 1 mm for models built from laser scan data. However, the sacrifice of a small amount of precision may be outweighed by other benefits that this approach confers. Not only does it allow for models to be constructed very quickly, the process is simple enough that at Çatalhöyük all of the digital image recording and processing is completed by students attending archaeological field schools at the site. In addition, the speed at which these models can be produced allows researchers to use them to analyse and interpret burials almost in real-time as they are excavated. For example, by using computer vision to create 3D models, the osteological team at Çatalhöyük was able to reconstruct a complex sequence of multiple burials layer by layer, a procedure which allowed them to visualize connections amongst skeletons that were not visible in 2D images (Forte, 2014). Although this benefit is not mentioned in the report, it is easy to see how having access to this detailed a level of information on a daily basis might be useful with regard to informing the progress of excavation on subsequent work days.

As the studies discussed above show, high-quality, detailed visual documentation of archaeological sites can be achieved in a multitude of ways, many of which are increasingly simple and affordable in terms of the skill and technology required. While 2D documentation

with photographs and maps is standard for archival and publication purposes, there is no reason why techniques like HDR imaging and computer vision 3D modeling should not be adopted on a wider scale, as Ducke, Score, and Reeves (2011) propose. The adoption of documentation techniques that maximize detail and minimize data loss are highly desirable, given the destructive nature of archaeological excavation, as is the ability to visually reconstruct and review an entire excavation from start to finish, as is being done at Çatalhöyük (Forte, 2012; 2014). As the ability of researchers to visually capture the excavation process in ever-finer detail increases, so too does the utility of such documentation, as it will allow deeper and more detailed contextual analysis of sites long after fieldwork is complete.

Bibliography

Allen, Caitlin

1996. "A Picture is Worth a Thousand Artefacts: Investigating the Archaeology of Workers' Camps in New South Wales 1872- 1962 using Historic Photographs."

Unpublished thesis. Department of Prehistory and Historical Archaeology, University of Sydney.

Banks, Kimball M., and J. Signe Snortland

1995. Every Picture Tells a Story: Historic Images, Tipi Camps, and Archaeology. *Plains Anthropologist* 40 (152): 125-144.

Bohrer, Frederick N.

2005. Photography and Archaeology: The Image as Object. In: *Envisioning the Past: Archaeology and the Image*, edited by Sam Smiles and Stephanie Moser, pp. 180 – 191.

Blackwell Publishing: London.

Bohrer, Frederick N.

2011. *Photography and Archaeology*. Reaktion Books: London.

Bland, J M, and D G Altman

1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* i: 307–310.

Bland, J Martin, and Douglas G Altman

1999. Statistical Methods in Medical Research. *Statistical Methods in Medical Research* 8(2): 161–179.

Buikstra, Jane, and Douglas Ubelaker, eds.

1994. *Standards for Data Collection from Human Skeletal Remains: Proceedings of a Seminar at the Field Museum of Natural History*. Arkansas Archaeology Research Series

44.

DeVos, Kurt.

2010. *Image J Cell Counter*. University of Sheffield, Academic Neurology

<http://rsb.info.nih.gov/ij/plugins/cell-counter.html>.

Dorrell, Peter G.

1989. *Photography in Archaeology and Conservation*. 2nd Ed. Cambridge Archaeology
Manuals. Cambridge University Press: Cambridge, UK.

Ducke, Benjamin, David Score, and Joseph Reeves

2011. Multiview 3D reconstruction of the archaeological site at Weymouth from image
series. *Computers and Graphics (Pergamon)* 35(2): 375–382.

Dlussky, Konstantin G., Andrzej W. Weber, Nat Rutter, and Dustin White

2008. Sediment from Kuzhir-Nuge XIV Site: Application to Mortuary Rituals. In
*Khuzhir-Nuge XIV, a Middle Holocene Hunter-Gatherer Cemetery on Lake Baikal,
Siberia: Archaeological Materials*, edited by A.W. Weber, O.I. Goriunova, and H.G.
McKenzie, pp. 341-362. Northern Hunter-Gatherers Research Series, Vol. 4, Canadian
Circumpolar Institute Press, University of Alberta, Edmonton.

Duday, Henri

2009. *The Archaeology of the Dead: Lectures in Archaeoethanatology*. Studies in
Funerary Archaeology Vol. 3. Translated by Anna Maria Cipriani and John Pearce.
Oxbow Books: Oxford.

Duday, Henri, and Mark Guillon

2006. Understanding the Circumstances of Decomposition when the Body is
Skeletonized. In: *Forensic Anthropology and Medicine: Complementary Sciences from*

Recovery to Cause of Death, edited by A. Schmitt, E. Cunha, and J. Pinheiro, pp: 117-157. Humana Press: Totowa, NJ.

Epson America, Inc.

2013. Epson Perfection V700 Photo Scanner User's Guide. Electronic Document, https://files.support.epson.com/htmldocs/prv7ph/prv7phug/html/scan1_9.htm

Epson Canada, Ltd.

n.d. Scanner Resolution: Image Quality is More Than a Number. Electronic Document, http://www.epson.ca/cmc_upload/pdf/tech_scanner-resolution.pdf

Falconer, John

2010. Perspectives on Photography's Contribution to Archaeology in Central Asia. In: *Conservation of Ancient Sites on the Silk Road: Proceedings of the Second International Conference of Grotto Sites, Mogao Grottoes, Dunhuang, Peoples Republic of China, June 28-July 3, 2004*, edited by Neville Agnew, pp. 107-116. Getty Conservation Institute: Los Angeles.

Field, Andy

2005. *Discovering Statistics Using SPSS*. 2nd Ed. Sage: London.

Forte, Maurizio

2014. 3D Archaeology: New Perspectives and Challenges – The Example of Çatalhöyük. *Journal of Eastern Mediterranean Archaeology and Heritage Studies* 2(1): 1-29.

Forte, Maurizio, N. Dell'Untob, J. Issavia, L. Onsureza, and N. Lercaria

2012. 3D Archaeology at Çatalhöyük. *International Journal of Heritage in the Digital Era*, 1(3): 351-378.

Flinders Petrie, W. M.

1972. *Methods and Aims in Archaeology*. Reprinted. Benjamin Blom, Inc, New York.
Originally Published 1904.

Gerasimov, M. M., and E. N. Chernykh

1975. Raskopki fofanovskogo mogil'nika v 1959 g [Excavation of the Fofanovo cemetery in 1959]. In: *Pervobytnaia arkheologiya Sibiri*, edited by A. M. Mandel'shtam, pp. 23–48. Nauka: Leningrad.

González-Jorge, H., M. Solla, J. Martínez-Sánchez, and P. Arias

2012. Comparison between laser scanning, single-image rectification and ground-penetrating radar technologies in forensic science. *Measurement: Journal of the International Measurement Confederation* 45(5): 836–843.

González-Jorge, Higinio, Iván Puente, Pablo Eguía, and Pedro Arias

2013. Single-image rectification technique in forensic science. *Journal of Forensic sciences* 58(2): 459–64.

Hallotte, Rachel

2007. Photography and the American Contribution to Early "Biblical" Archaeology, 1870-1920. *Near Eastern Archaeology* 70 (1): 26-41.

Harris, N.J., and N. Tayles

2012. Burial Containers – A Hidden Aspect of Mortuary Practices: Archaeoethanatology at Ban Non Wat, Thailand. *Journal of Anthropological Archaeology* 31: 227–239.

Haverkort, Caroline M., Vladimir I. Bazaliiskii, and Nikolai A. Savel'ev

2010. Identifying Hunter-Gatherer Mobility Patterns Using Strontium Isotopes. In *Prehistoric Hunter-Gatherers of the Baikal Region, Siberia: Bioarchaeological Studies of*

Past Life Ways, edited by Andrzej W. Weber, M. Anne Katzenberg, and Theodore G. Schurr, pp. 217-242. University of Pennsylvania Press, Philadelphia, PA

Hayes, Andrew, and Li Cai

2007. Using Heteroskedasticity-Consistent Standard Error Estimators in OLS Regression: An Introduction and Software Implementation. *Behaviour Research Methods* 39(4): 709-722.

Hayes, Andrew F., and Klaus Krippendorf

2007. Answering the Call for a Standard Reliability Measure for Coding DData. *Communication Methods and Measures*: 77–89.

Hodder, Ian

1999. *The Archaeological Process: An Introduction*. Blackwell: Oxford.

Katzenberg, M. Anne, Vladimiar I. Bazaliiskii, Olga I. Goriunova, Nikolai A. Savel'ev, and Andrzej W. Weber

2010. Diet reconstruction and stable isotope ecology in the Lake Baikal region. In *Prehistoric Hunter-Gatherers of the Baikal Region, Siberia: Bioarchaeological Studies of Past Life Ways*, edited by Andrzej W. Weber, M. Anne Katzenberg, and Theodore G. Schurr, pp. 175–191. University of Pennsylvania Press, Philadelphia, PA

Konev, V.P.

1996. Fofanovskii mogil'nik: Novyi etap isseldovaniia. In: *Arkheologiia, paleokologiia i etnologiia Sibiri i Dal'nego Vostoka*, edited by N. E. Berdnikova, pp. 114 – 116. Irkutskii gosudarstvennyi universitet: Irkutsk.

Laerd Statistics

2016. Cohen's kappa using SPSS Statistics. Statistical tutorials and software guides.

Retrieved from <https://statistics.laerd.com/>

Levin, Aaron M.

1986. Excavation Photography: A Day on a Dig. *Archaeology* 39(1), pp.34-39.

Lieverse, Angela R.

1999. "Human Taphonomy at Khuzhir-Nuge XIV, Siberia." MA Thesis. Department of Anthropology, University of Alberta.

2007a. Human Taphonomy: Variable Skeletal Condition and Its Cultural Implications. In: *Khuzhir-Nuge XIV, A Middle Holocene Hunter-Gatherer Cemetery on Lake Baikal, Siberia: Osteological Materials*, edited by A.W. Weber, M. A. Katzenberg, and O.I. Goriunova, pp 201-216. Northern Hunter-Gatherers Research Series 3. Canadian Circumpolar Institute Press, University of Alberta, Edmonton.

2007b. Human Osteological Evidence: Demography and Health. In: *Khuzhir-Nuge XIV, A Middle Holocene Hunter-Gatherer Cemetery on Lake Baikal, Siberia: Osteological Materials*, edited by A.W. Weber, M. A. Katzenberg, and O.I. Goriunova, pp 217-234. Northern Hunter-Gatherers Research Series 3. Edmonton, AB: Canadian Circumpolar Institute Press.

2010. Health and Behaviour in Mid-Holocene Cis-Baikal: Biological Indicators of Adaptation and Culture Change. In: *Prehistoric Hunter-Gatherers of the Baikal Region, Siberia: Bioarchaeological Studies of Past Life Ways*, edited by Andrzej W. Weber, M. Anne Katzenberg, and Theodore G. Schurr, pp. 135 - 173. University of Pennsylvania

Press, Philadelphia, PA

Lieverse, Angela R., Andrzej W. Weber, and Olga I. Goriunova

2006. Human taphonomy at Khuzhir-Nuge XIV, Siberia: a new method for documenting skeletal condition. *Journal of Archaeological Science* 33(8): 1141–1151.

Littleton, Judith, Bruce Floyd, Bruno Frohlich, Michael Dickson, Tsend Amgalantögs, Sarah Karstens, and Kristen Pearlstein

2012. Taphonomic analysis of Bronze Age burials in Mongolian khirigsuurs. *Journal of Archaeological Science* 39(11): 3361–3370.

Marden, K. and D. J. Ortner

2011. A case of treponematosi from pre-Columbian Chaco Canyon, New Mexico. *International Journal of Osteoarchaeology* 21(1), pp.19-31.

McKenzie, Hugh G.

2010. Variability in Bronze Age Mortuary Practices in the Little Sea Microregion of Cis-Baikal. In *Prehistoric Hunter-Gatherers of the Baikal Region, Siberia: Bioarchaeological Studies of Past Life Ways*, edited by Andrzej W. Weber, M. Anne Katzenberg, and Theodore G. Schurr, pp. 87–106. University of Pennsylvania Press, Philadelphia, PA.

McKenzie, Hugh G., Andrzej W. Weber, and Olga I. Goriunova

2008. Mortuary Variability. In *Khuzhir-Nuge XIV, a Middle Holocene Hunter-Gatherer Cemetery on Lake Baikal, Siberia: Archaeological Materials*, edited by A.W. Weber, O.I. Goriunova, and H.G. McKenzie, pp. 219-266. Northern Hunter-Gatherers Research Series, Vol. 4, Canadian Circumpolar Institute Press, University of Alberta, Edmonton.

Mooder, Karen P., Tia A. Thompson, Andrzej W. Weber, Vladimir I. Bazaliiskii, and Fiona J.

Bamforth

2010. Uncovering the Genetic Landscape of Prehistoric Cis-Baikal. In *Prehistoric Hunter-Gatherers of the Baikal Region, Siberia: Bioarchaeological Studies of Past Life Ways*, edited by Andrzej W. Weber, M. Anne Katzenberg, and Theodore G. Schurr, pp. 107-119. University of Pennsylvania Press, Philadelphia, PA.

Morris, Keith J.

2007. The Scanning of Colour and B&W Film and Photographs for Image Processing, Analysis, and Archiving – On a Tight Budget. In *Focus 5*: 19-33.

Nilsson, Liv

1998. Dynamic Cadavers: A “Field Anthropological” Analysis of the Skateholm II Burials. *Lund Archaeological Review 4*: 5-17.

Nilsson-Stutz, Liv

2003. *Embodied Ritual and Ritualized Bodies: Tracing Ritual Practices in Late Neolithic Burials*. Acta Archaeologica Lundensia Series, No. 46. Almqvist and Wiskell International: Stockholm, Sweden.

2006. Unwrapping the Dead: Searching for Evidence of Wrappings in the Mortuary Practices at Zvejnieki. In: *Back to the Origin: New Research in the Mesolithic-Neolithic Zvejnieki Cemetery and Environment, Northern Latvia*, edited by Lars Larsson and Ilga Zargoska, pp: 217-233. Acta Archaeologica Lundensia Series in 8°, No. 52. Almqvist and Wiskell International: Stockholm, Sweden.

2008. More than Metaphor: Approaching the Human Cadaver in Archaeology. In: *The Materiality of Death*, edited by F. Fahlander and T. Østigaard, pp. 19-28. B. A. R. S1768. Archaeopress: Oxford.

Nomokonova, Tatiana, Robert J. Losey, Olga I. Goriunova, and Andrzej W. Weber

2013. A Freshwater Old Carbon Offset in Lake Baikal, Siberia, and Problems with the Radiocarbon Dating of Archaeological Sediments: Evidence from the Sagan-Zaba II Site. *Quaternary International* 290: 110-125.

Parno, Travis G.

2010. Snapshots of History and the Nature of the Archaeological Image. *Archaeologies: Journal of the World Archaeological Congress* 6 (1): 115-137.

Raband, W. S.

1997-2014. *ImageJ*. U. S. National Institutes of Health, Bethesda, Maryland, USA, <http://imagej.nih.gov/ij/>.

Robertson, Cameron, Andrzej W. Weber, and Bradley Drouin

2008. Grave Disturbance Patterns. In *Khuzhir-Nuge XIV, a Middle Holocene Hunter-Gatherer Cemetery on Lake Baikal, Siberia: Archaeological Materials*, edited by A.W. Weber, O.I. Goriunova, and H.G. McKenzie, pp. 267-318. Northern Hunter-Gatherers Research Series, Vol. 4, Canadian Circumpolar Institute Press, University of Alberta, Edmonton.

Roskams, Steve

2001. *Excavation*. Cambridge Archaeology Manuals. Cambridge University Press: Cambridge, UK.

Roksandic, Mirjana

2002. Position of Skeletal Remains as Key to Understanding Mortuary Behaviour. In: *Advances in Forensic Taphonomy: Method, Theory and Archaeological Perspectives*, edited by William D. Haglund and Marcella H. Sorg, pp. 95-113. Boca Raton, FL: CRC Press.
- Schurr, Theodore G., Ludmilla P. Osipova, Sergey I. Zhadnov, and Matthew C. Dulik
2010. Genetic diversity in native Siberians: Implications for the Prehistoric Settlement of the Cis-Baikal region. In *Prehistoric Hunter-Gatherers of the Baikal Region, Siberia: Bioarchaeological Studies of Past Life Ways*, edited by Andrzej W. Weber, M. Anne Katzenberg, and Theodore G. Schurr, pp. 121-134. University of Pennsylvania Press, Philadelphia, PA.
- Shanks, Michael
1997. Photography and Archaeology. In: *The Cultural Life of Images: Visual Representation in Archaeology*, edited by Brian Leigh Molyneaux, pp. 73-107. Routledge: London and New York.
- Tassie, G. J., and L.S. Owens (editors)
2010. *Standards of Archaeological Excavation: A Field Guide*. Egyptian Cultural Heritage Monograph Series, No. 1. Egyptian Cultural Heritage Organization. Golden House Publications: Peterborough, UK.
- Warden, Robert, and Thomas H Guderjan
2012. Practical Recording Issues At Small and Large Scales in Maya Archaeology. *SAA Archaeological Record* 12(4): 47–51.
- Watson, J. T., and J. Weiland
2013. Documenting Archaeological Mortuary Features using High Dynamic Range

(HDR) Imaging. *International Journal of Osteoarchaeology* 25(3): 366-373.

White, Dustin, and Andrew Bush

2010. Holocene Climate, Environmental Change, and Biocultural Discontinuity in the Baikal Region. *Prehistoric Hunter-Gatherers of the Baikal Region, Siberia: Bioarchaeological Studies of Past Life Ways*, edited by Andrzej W. Weber, M. Anne Katzenberg, and Theodore G. Schurr, pp. 1 - 26. University of Pennsylvania Press, Philadelphia, PA

Weber, Andrzej W.

1995. The Neolithic and Early Bronze Age of the Lake Baikal region, Siberia: a review of recent research. *Journal of World Prehistory* 9; 99–165.

Weber, Andrzej W.

2003. Bio-geographic profile of the Lake Baikal Region. In: *Prehistoric Foragers of the Cis-Baikal, Siberia*, edited by Andrzej W. Weber and Hugh G. McKenzie, pp. 53-76. Canadian Circumpolar Institute Press, University of Alberta, Edmonton.

Weber, Andrzej W. and Robert L. Bettinger

2010. Middle Holocene Hunter-Gatherers of Cis-Baikal, Siberia: An Overview for the New Century. *Journal of Anthropological Archaeology* 29: 491-506.

Weber, Andrzej W, and Olga I. Goriunova

2008. Field work methods at Khuzhir Nuge XIV. In *Khuzhir-Nuge XIV, a Middle Holocene Hunter-Gatherer Cemetery on Lake Baikal, Siberia: Archaeological Materials*, edited by A.W. Weber, O.I. Goriunova, and H.G. McKenzie, pp. 23–32. Northern Hunter-Gatherers Research Series, Vol. 4, Canadian Circumpolar Institute Press, University of Alberta, Edmonton.

Weber, Andrzej W., and Olga I. Goriunova

2013. Hunter-gatherer migrations, mobility and social relations: A case study from the Early Bronze Age Baikal region, Siberia. *Journal of Anthropological Archaeology* 32(3): 330–346.

Weber, Andrzej W., Peter Jordan, and Hirofumi Kato

2013. Environmental change and cultural dynamics of Holocene hunter-gatherers in Northeast Asia: Comparative analyses and research potentials in Cis-Baikal (Siberia, Russia) and Hokkaido (Japan). *Quaternary International* 290–291: 3–20.

Weber, Andrzej W., David W. Link, and M. Anne Katzenberg

2002. Hunter-Gatherer Culture Change and Continuity in the Middle Holocene of the Cis-Baikal, Siberia. *Journal of Anthropological Archaeology* 21(2): 230–299.

Weber, Andrzej W., Hugh G. McKenzie, and Rolf Buekens

2010. Radiocarbon Dating of Middle Holocene Culture History in Cis-Baikal. In *Prehistoric Hunter-Gatherers of the Baikal Region, Siberia: Bioarchaeological Studies of Past Life Ways*, edited by Andrzej W. Weber, M. Anne Katzenberg, and Theodore G. Schurr, pp. 27-49. University of Pennsylvania Press, Philadelphia, PA.

Weber, Andrzej W., Hugh G. McKenzie, Rolf Beukens, and Olga I. Goriunova

2005. Evaluation of Radiocarbon dates from the Middle Holocene hunter-gatherer cemetery Khuzhir-Nuge XIV, Lake Baikal, Siberia. *Journal of Archaeological Science* 32: 1481–1500.

Weber, Andrzej W., Rick Schulting, Christopher Bronk Ramsey, Vladimir I. Bazaliiskii, Olga I. Goriunova, and Natal'ia E. Berdnikova

2016. Chronology of Middle Holocene Hunter-Gatherers in the Cis-Baikal Region of

Siberia: Corrections Based on Examination of the Freshwater Reservoir Effect.

Quaternary International 419: 74 – 98.

Weber, Andrzej W., Dustin White, Vladimir I. Bazaliiskii, Olga I. Goriunova, Nikolai A. Savel'ev, and M. Anne Katzenberg

2011. Hunter–gatherer foraging ranges, migrations, and travel in the middle Holocene Baikal region of Siberia: Insights from carbon and nitrogen stable isotope signatures.

Journal of Anthropological Archaeology 30(4): 523–548.

Weitzel, Misty, and Andrzej W. Weber

2008. Mortuary Use of Red Ochre: An Experimental Approach. In *Khuzhir-Nuge XIV, a Middle Holocene Hunter-Gatherer Cemetery on Lake Baikal, Siberia: Archaeological Materials*, edited by A.W. Weber, O.I. Goriunova and H.G. McKenzie, pp. 319-40.

Northern Hunter-Gatherers Research Series, Vol. 4, Canadian Circumpolar Institute Press, University of Alberta, Edmonton.

Willis, A., and N. Tayles

2009. Field anthropology: application to burial contexts in prehistoric Southeast Asia.

Journal of Archaeological Science 36: 547–554.

Zejdlik, Katie J.

2014. Unmingling Commingled Museum Collections: A Photographic Method. In: *Commingled and Disarticulated Human Remains: Working Toward Improved Theory, Method, and Data*, edited by Anna J. Osterholtz, Kathryn M. Baustian, and Debra L. Martin, pp. 173-192. Springer-Verlag: New York.

**APPENDIX A:
ADDITIONAL TABLES**

A1: Comparison of the differences between Observed and Estimated Completeness, Observed and Estimated, FR, and Observed and Estimated fragment counts (sample of 100 randomly drawn elements)

Burial	Observed Completeness	Estimated Completeness	Diff.	Observed FR	Estimated FR	Diff.	Observed Fragments	Estimated Fragments	Diff.
9	40	40	0	1.33	6.67	-5.34	30	6	24
9	95	95	0	4.75	31.67	-26.92	20	3	17
10	40	65	-25	1	5.91	-4.91	40	11	29
11	95	100	-5	95	100	-5	1	1	0
11	100	95	5	50	95	-45	2	1	1
12	30	40	-10	0.3	1.43	-1.13	100	28	72
14	90	85	5	11.25	7.73	3.52	8	11	-3
15	90	65	25	90	65	25	1	1	0
16	75	85	-10	37.5	42.5	-5	2	2	0
19	65	80	-15	7.22	20	-12.78	9	4	5
19	35	30	5	1.75	1.88	-0.13	20	16	4
19	30	10	20	3.46	2.5	0.96	9	4	5
19	90	40	50	1.5	8	-6.5	60	5	55
23	5	20	-15	0.06	1.43	-1.37	83	14	69
25	15	40	-25	15	10	5	1	4	-3
27.01	75	70	5	4.29	23.33	-19.04	17	3	14
27.01	87.5	80	7.5	5.83	40	-34.17	15	2	13
27.01	46	55	-9	20.91	11	9.91	2	5	-3
27.03	70	70	0	70	70	0	1	1	0
29	90	50	40	22.5	25	-2.5	4	2	2
29	100	80	20	100	100	0	1	1	0
32	95	85	10	47.5	42.5	5	2	2	0
32	80	100	-20	20	100	-80	4	1	3
32	60	65	-5	1.3	1.33	-0.03	46	49	-3
32	35	20	15	5.83	10	-4.17	6	2	4
32	90	90	0	90	90	0	1	1	0
34	90	100	-10	90	33.33	56.67	1	3	-2
35.01	60	67.5	-7.5	7.5	5.19	2.31	8	13	-5
35.01	75	75	0	0.77	2.88	-2.11	97	26	71
35.01	60	30	30	2.14	4.29	-2.15	28	7	21
36.02	10	10	0	0.77	1.67	-0.9	13	6	7
37.01	40	45	-5	0.7	1.13	-0.43	57	40	17
37.01	95	85	10	1.64	3.86	-2.22	58	22	36
37.01	25	30	-5	25	30	-5	1	1	0
37.02	25.5	34	-8.5	4.55	18.89	-14.34	6	2	4
37.02	15	12.5	2.5	1.25	1.14	0.11	12	11	1
38	55	60	-5	5.5	7.5	-2	10	8	2
38	40	40	0	1.43	0.66	0.77	28	61	-33
39	60	35	25	60	2.69	57.31	1	13	-12
39	60	65	-5	3.75	60	-56.25	16	1	15
44	42.86	13.57	29.29	75	94.99	-19.99	1	0	0
44	15	15	0	0.88	0.79	0.09	17	19	-2
44	80	60	20	5.33	3.75	1.58	15	16	-1
44	64	60	4	40	75	-35	2	1	1
46	100	100	0	100	100	0	1	1	0
46	100	100	0	100	100	0	1	1	0
46	95	75	20	95	37.5	57.5	1	2	-1
46	95	95	0	95	47.5	47.5	1	2	-1
46	95	95	0	95	95	0	1	1	0

46	99	54	45	99	67.5	31.5	1	1	0
47	75	5	70	37.5	5	32.5	2	1	1
47	80	85	-5	16	28.33	-12.33	5	3	2
47	70	90	-20	11.67	45	-33.33	6	2	4
48	60	65	-5	6.67	7.22	-0.55	9	9	0
51	70	80	-10	1.84	3.48	-1.64	38	23	15
51	65	30	35	5	10	-5	13	3	10
51	65	20	45	5.42	5	0.42	12	4	8
51	30	30	0	3.75	30	-26.25	8	1	7
53	45	5	40	3.21	1	2.21	14	5	9
53	90	80	10	12.86	80	-67.14	7	1	6
53	85	95	-10	14.17	11.88	2.3	6	8	-2
53	75	85	-10	3.26	7.73	-4.47	23	11	12
55	85	100	-15	85	100	-15	1	1	0
57.01	80	95	-15	13.33	31.67	-18.34	6	3	3
57.02	75	55	20	0.94	27.5	-26.56	80	2	78
58.01	90	70	20	4.74	10	-5.26	19	7	12
58.01	65	35	30	1.44	5	-3.56	45	7	38
58.01	35	20	15	35	20	15	1	1	0
58.02	80	40	40	80	40	40	1	1	0
58.02	80	35	45	80	11.67	68.33	1	3	-2
59.02	95	80	15	6.79	80	-73.21	14	1	13
59.02	95	100	-5	95	100	-5	1	1	0
59.02	70	70	0	14	23.33	-9.33	5	3	2
60	90	80	10	4.74	16	-11.26	19	5	14
61	67.5	56.25	11.25	90	90	0	1	1	0
61	40	30	10	3.08	15	-11.92	13	2	11
63	90	80	10	90	80	10	1	1	0
64	40	60	-20	2	5	-3	20	12	8
66	99	100	-1	24.75	100	-75.25	4	1	3
66	75	40	35	5.77	10	-4.23	13	4	9
68	100	100	0	100	100	0	1	1	0
68	80	65	15	80	65	15	1	1	0
68	80	100	-20	5	50	-45	16	2	14
68	75	30	45	2.88	2.31	0.57	26	13	13
71	85	90	-5	6.07	45	-38.93	14	2	12
71	30.63	71.84	-41.21	81.76	82.11	-0.35	0	1	-1
71	45	40	5	2.5	5.71	-3.21	18	7	11
73	95	90	5	11.88	15	-3.12	8	6	2
77	65	65	0	2.03	4.64	-2.61	32	14	18
77	40	50	-10	40	50	-10	1	1	0
78	85	90	-5	85	45	40	1	2	-1
78	99	95	4	99	95	4	1	1	0
80.02	8	16	-8	40	80	-40	0	0	0
81	60	35	25	1.54	35	-33.46	39	1	38
81	90	50	40	18	50	-32	5	1	4
84	25	35	-10	8.33	35	-26.67	3	1	2
85	80	70	10	1.36	7	-5.64	59	10	49
85	60	45	15	2.73	5.63	-2.9	22	8	14
86	95	60	35	8.64	8.57	0.07	11	7	4
87	45	70	-25	3.21	6.36	-3.15	14	11	3

and Estimated total fragment counts per Burial, with FR

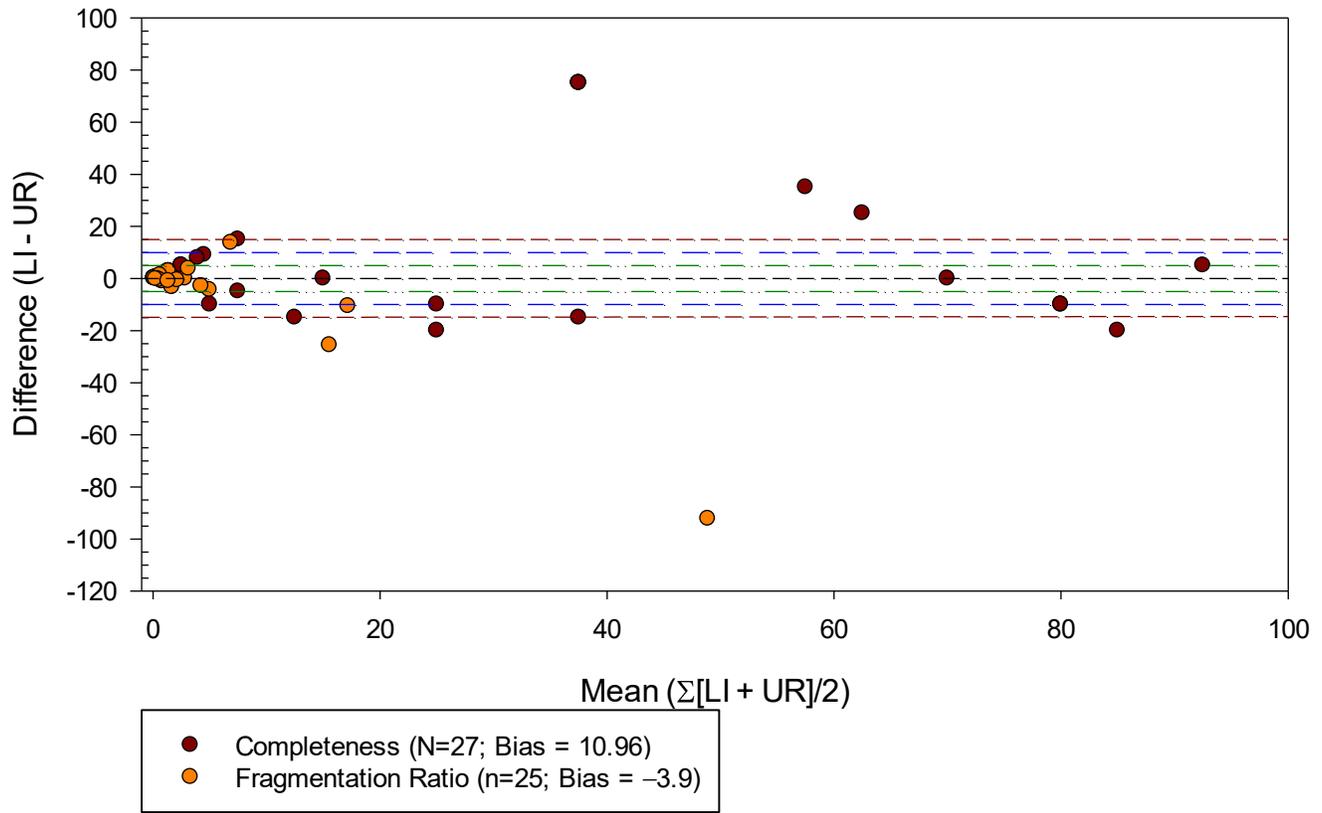
Biases

Burial	Bias FR	Observed Frags	Estimated Frags	Difference	Burial	Bias FR	Total Observed Frags	Total Estimated Frags	Difference
7	-4.20	245	666	421	48	10.18	144	507	363
9	7.01	127	484	357	49	7.17	163	507	344
10	7.06	86	155	69	50	-0.77	126	658	532
11	22.92	71	154	83	51	1.18	232	480	248
12	3.09	133	372	239	53	4.68	195	588	393
14	2.73	256	679	423	55	-1.53	183	562	379
15	10.22	110	258	148	57.01	2.66	31	222	191
16	13.27	129	227	98	57.02	12.41	19	482	463
19	-0.32	235	620	385	58.01	3.49	122	499	377
21	29.35	12	57	45	58.02	-1.51	261	722	461
22	13.57	57	322	265	59.01	-4.25	24	60	36
23	3.80	77	500	423	59.02	6.83	136	718	582
24	3.30	211	584	373	60	4.98	175	538	363
25	1.80	186	335	149	61	1.28	47	172	125
27.01	3.00	152	582	430	63	23.00	209	609	400
27.02	2.29	53	279	226	64	0.12	165	654	489
27.03	-5.49	76	75	-1	66	24.07	81	968	887
28	-5.41	128	293	165	68	17.57	101	567	466
29	15.32	167	412	245	71	-6.43	94	359	265
32	2.84	611	826	215	72	3.61	196	1755	1559
33	6.11	104	114	10	73	7.27	111	300	189
34	16.12	375	660	285	75	-1.77	16	29	13
35.01	0.58	416	1189	773	76	-15.13	29	188	159
35.02	0.53	269	780	511	77	-14.04	21	87	66
36.02	-0.19	153	237	84	78	36.99	57	55	-2
37.01	5.19	413	772	359	79	1.75	75	176	101
37.02	6.22	293	586	293	80.02	16.79	66	291	225
38	17.31	340	586	246	81	3.57	88	866	778
39	-11.72	114	402	288	83	-6.55	14	370	356
44	-2.31	248	619	371	84	-9.52	9	30	21
45	5.77	110	536	426	85	-2.59	36	201	165
46	11.89	95	210	115	86	-0.46	68	561	493
47	9.87	95	135	40	87	3.99	202	650	448

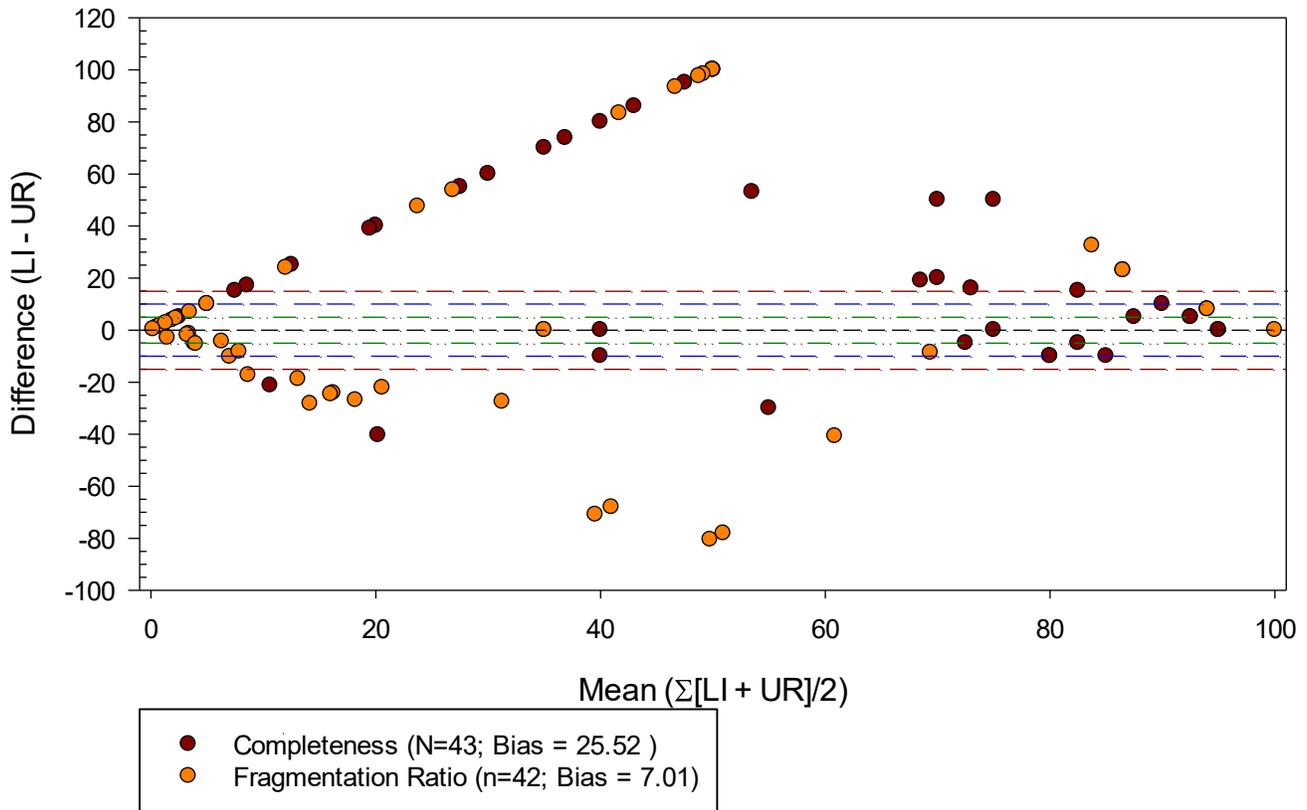
**APPENDIX B:
ADDITIONAL FIGURES**

APPENDIX A1: BLAND-ALTMAN PLOTS

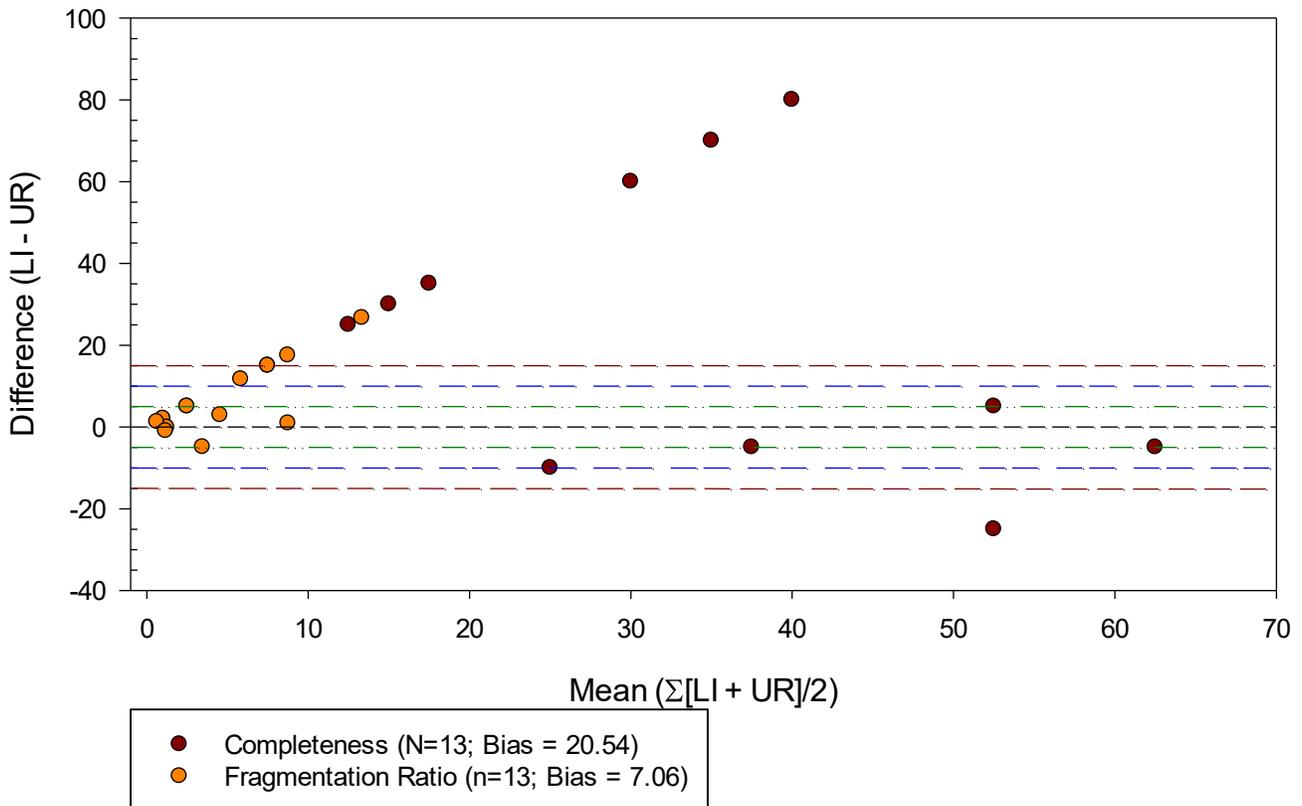
Difference Between Element Scores vs. Mean of Element Scores
Burial 7



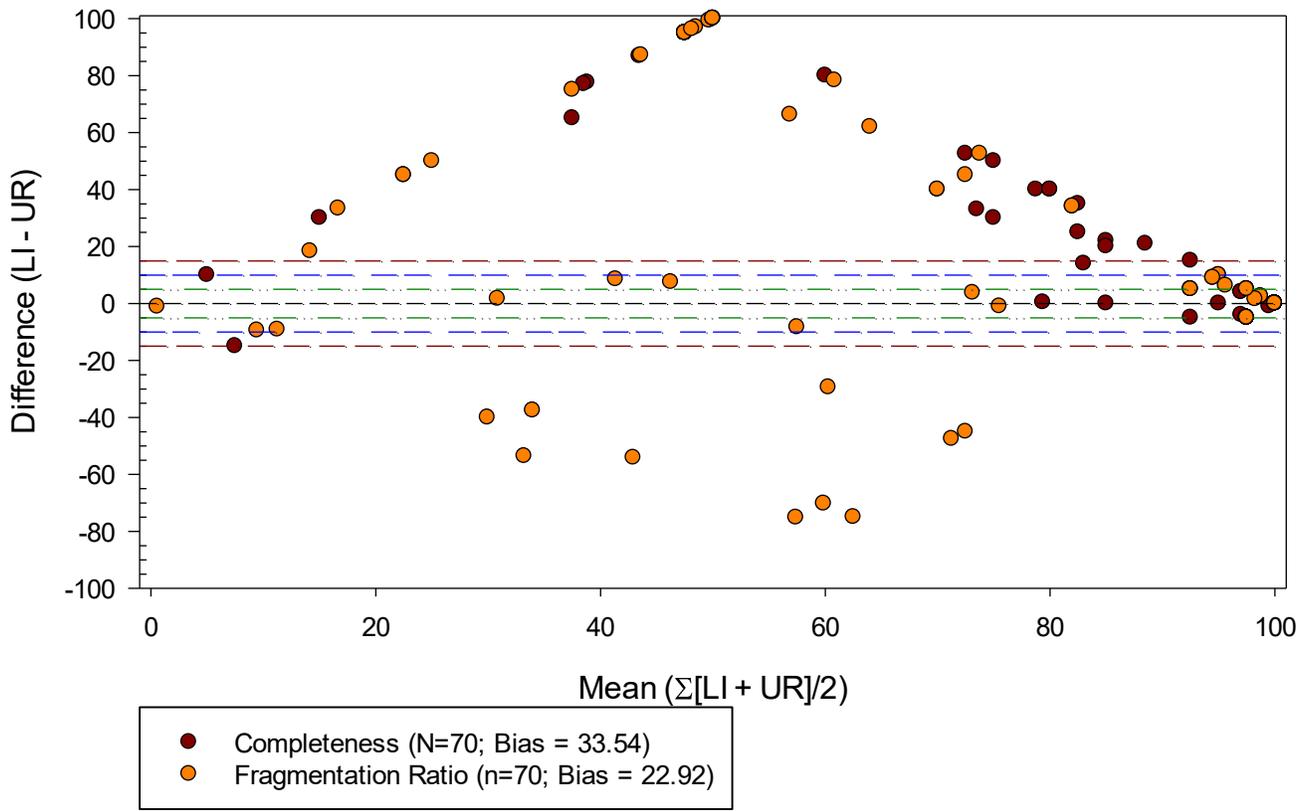
Difference Between Element Scores vs. Mean of Element Scores Burial 9



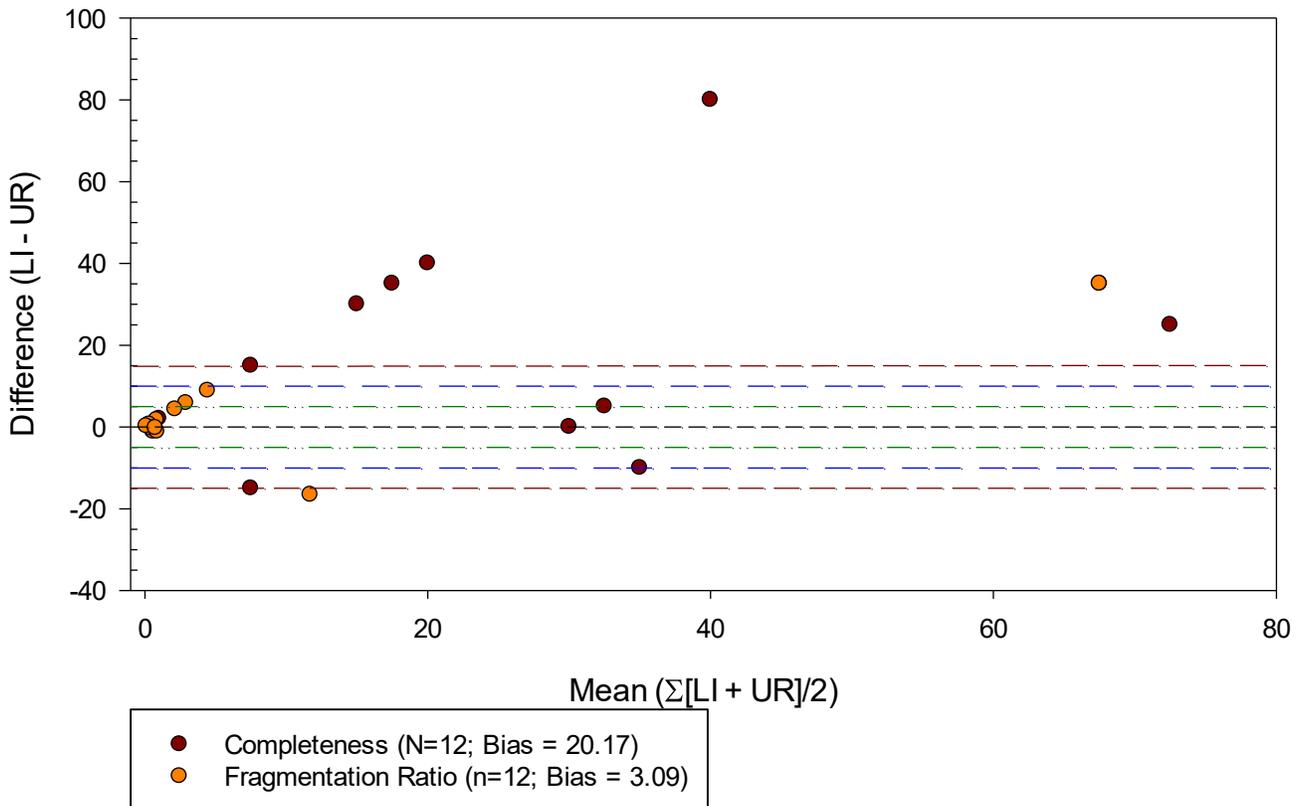
Difference Between Element Scores vs. Mean of Element Scores Burial 10



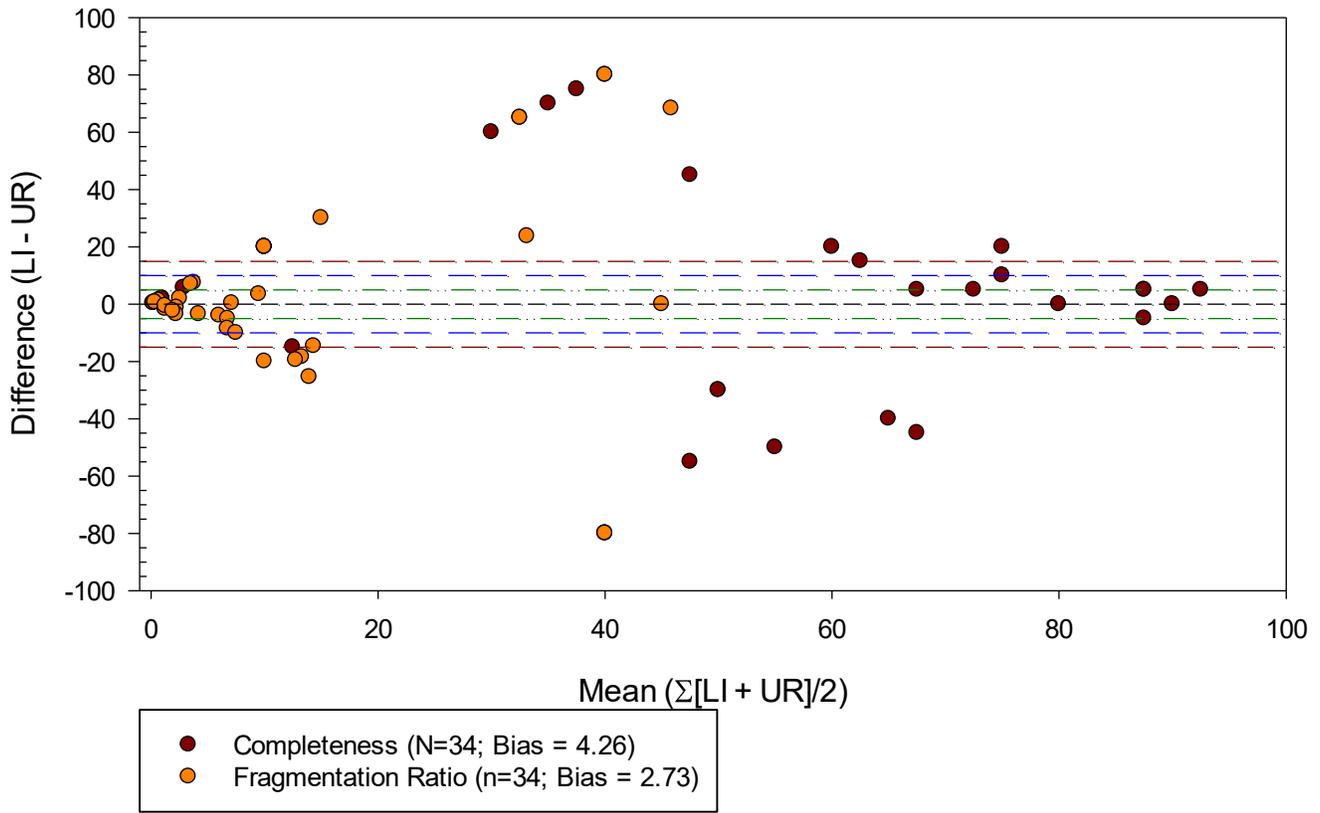
Difference Between Element Scores vs. Mean of Element Scores Burial 11



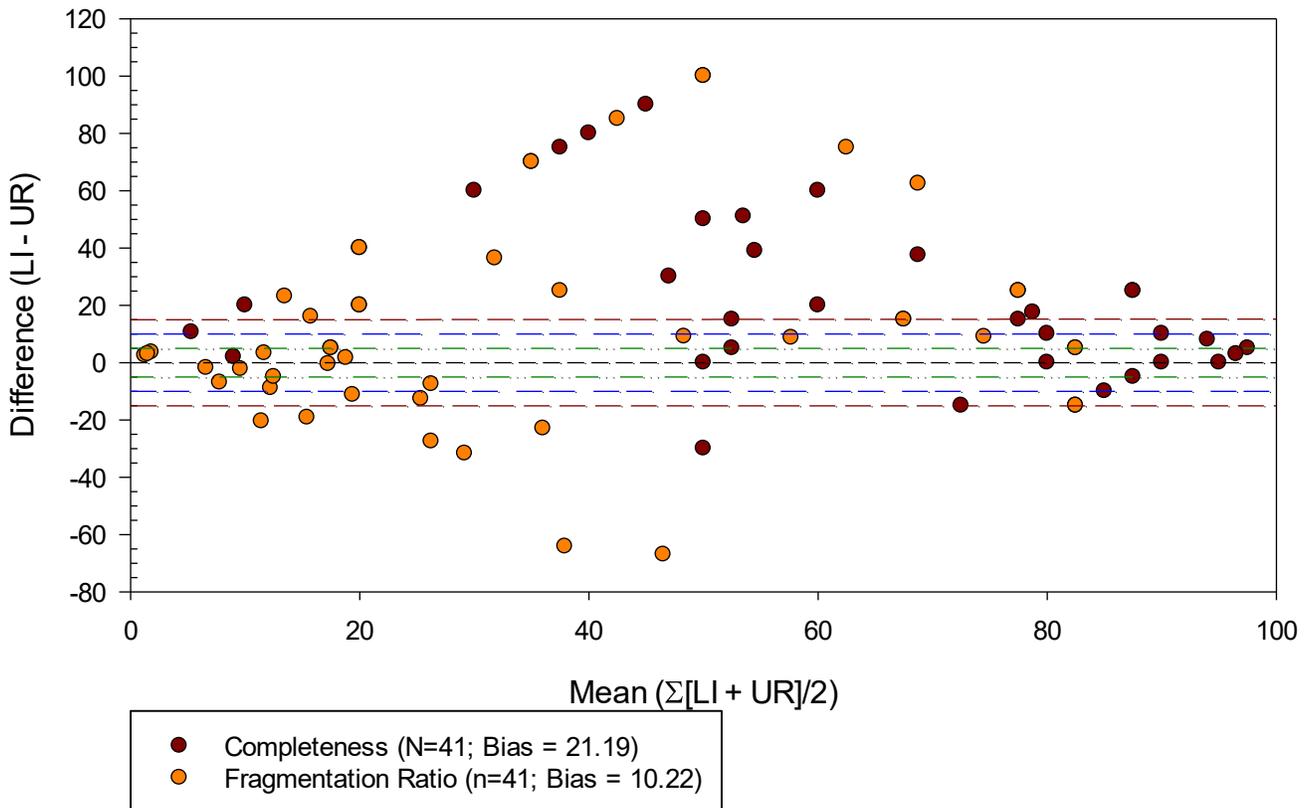
Difference Between Element Scores vs. Mean of Element Scores Burial 12



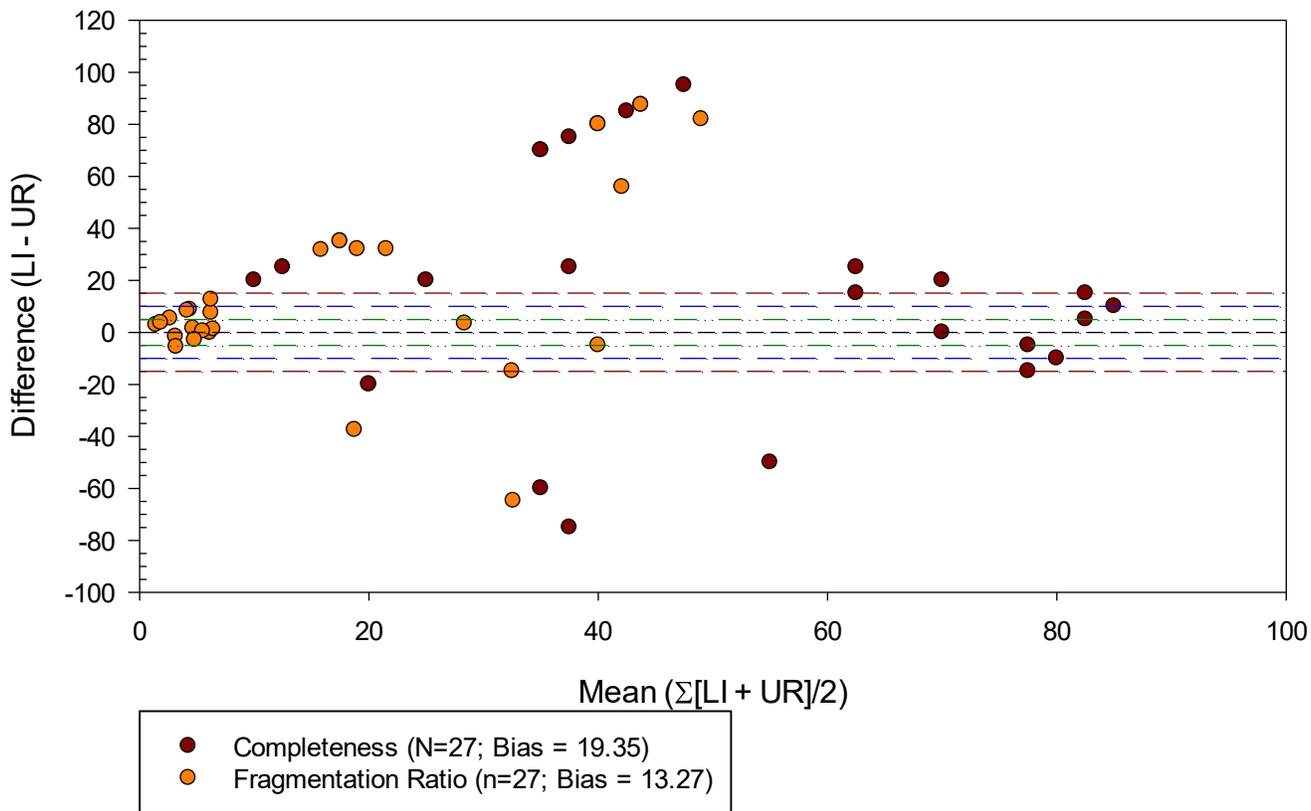
Difference Between Element Scores vs. Mean of Element Scores Burial 14



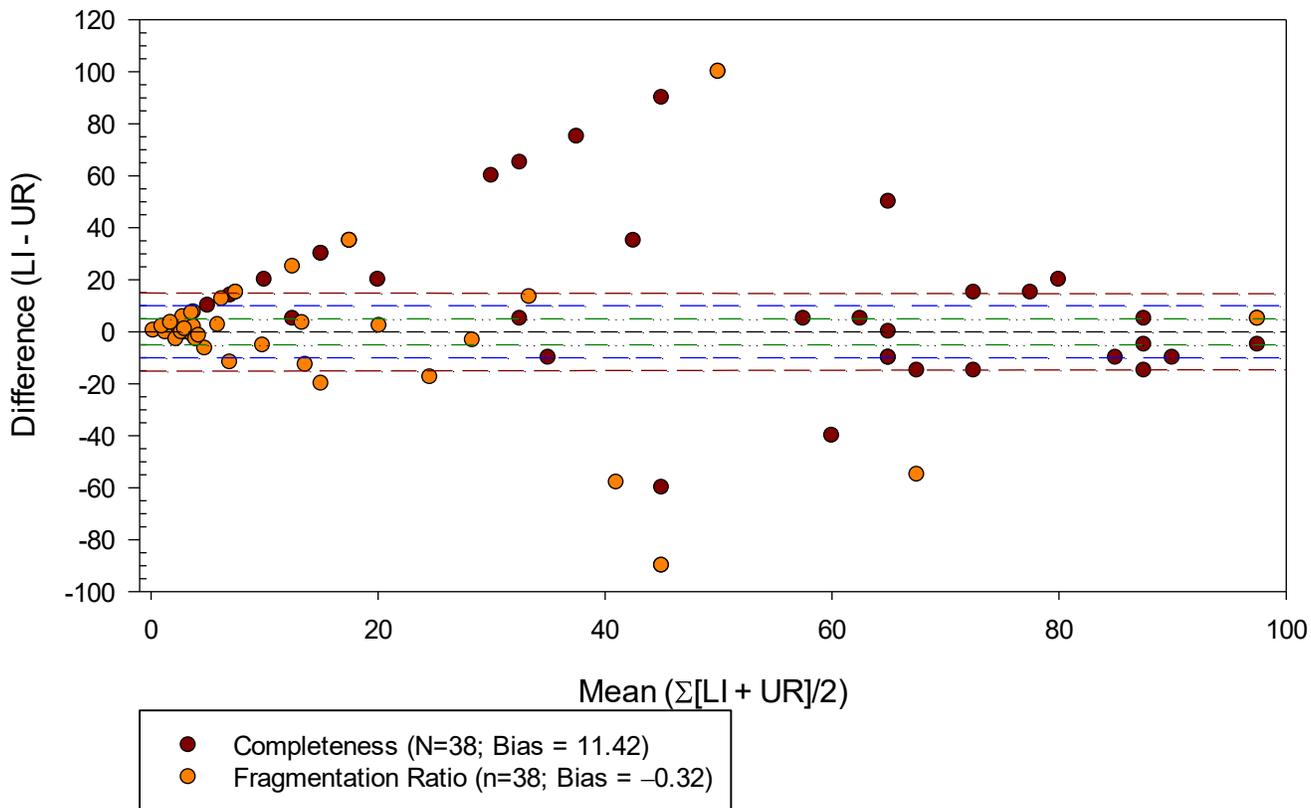
Difference Between Element Scores vs. Mean of Element Scores Burial 15



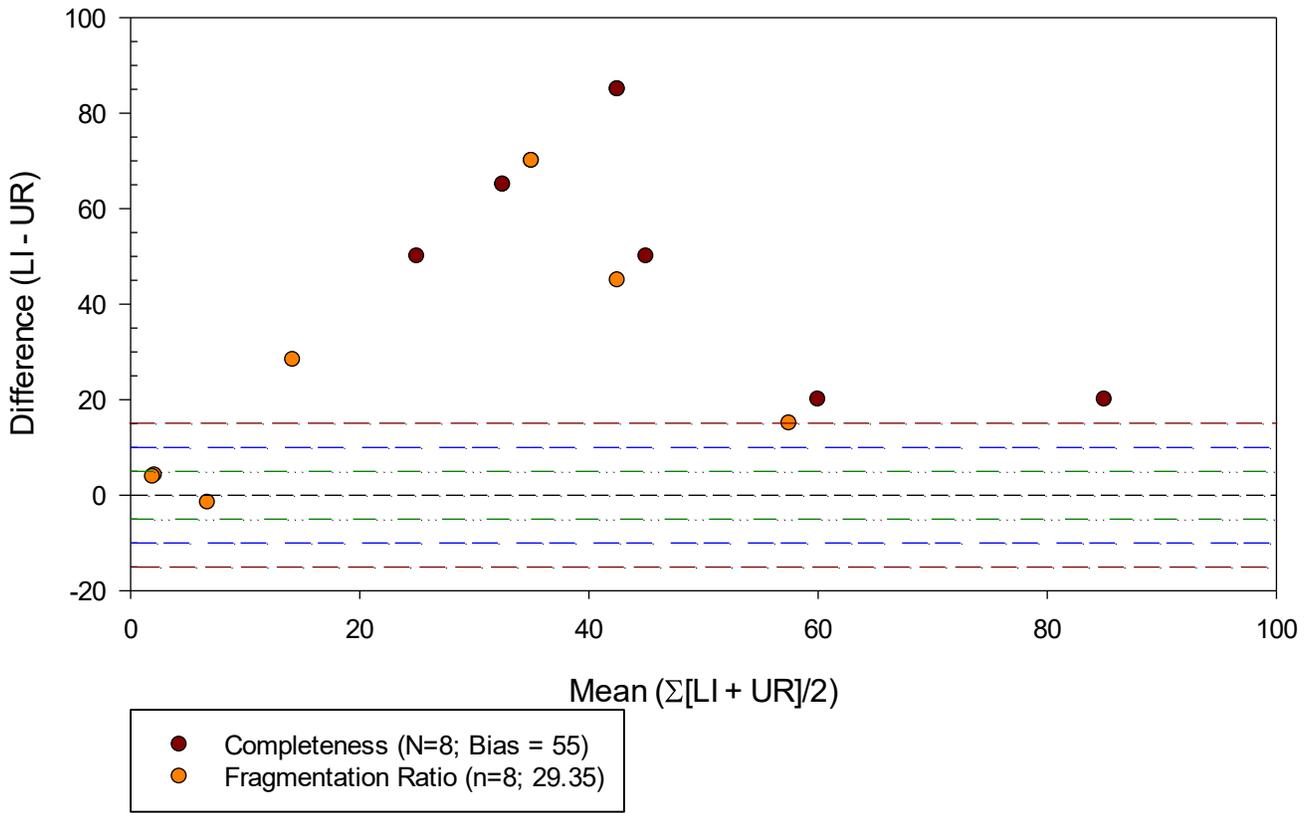
Difference Between Element Scores vs. Mean of Element Scores Burial 16



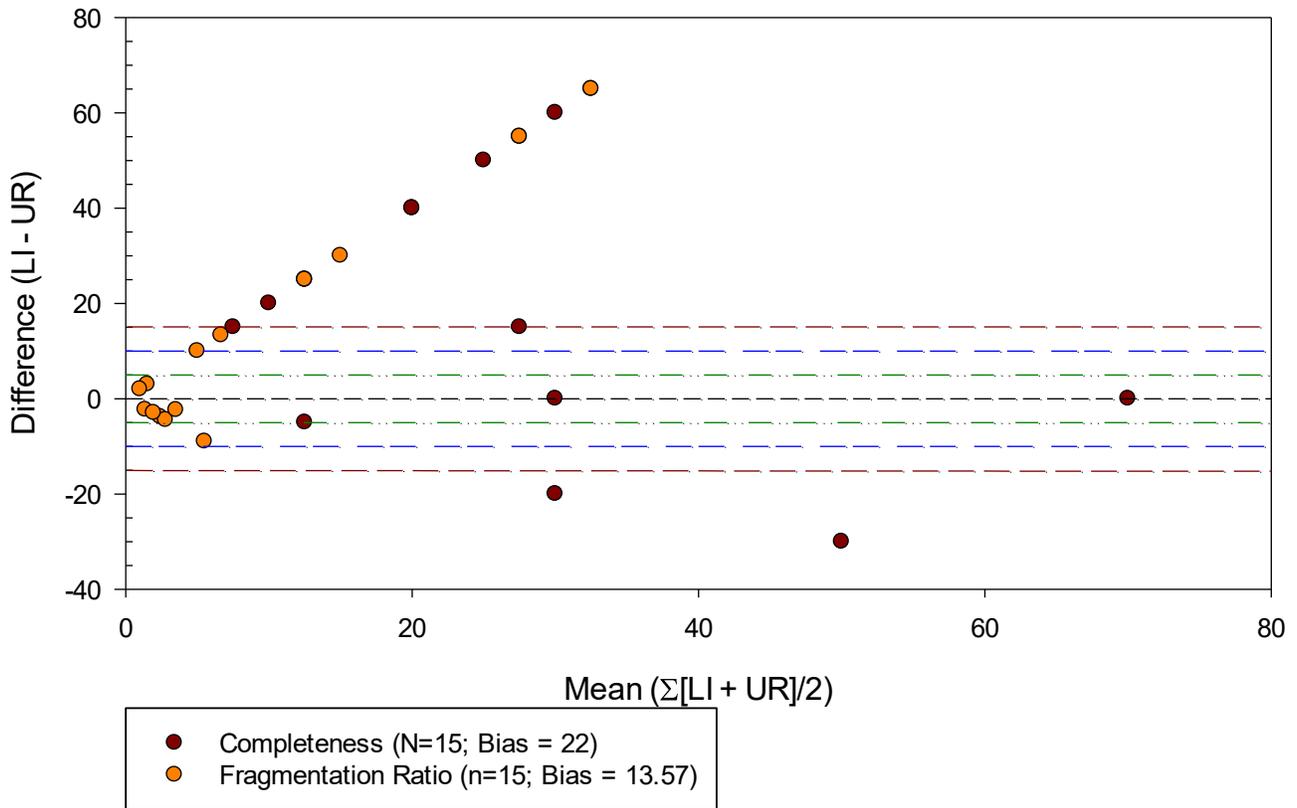
Difference Between Element Scores vs. Mean of Element Scores Burial 19



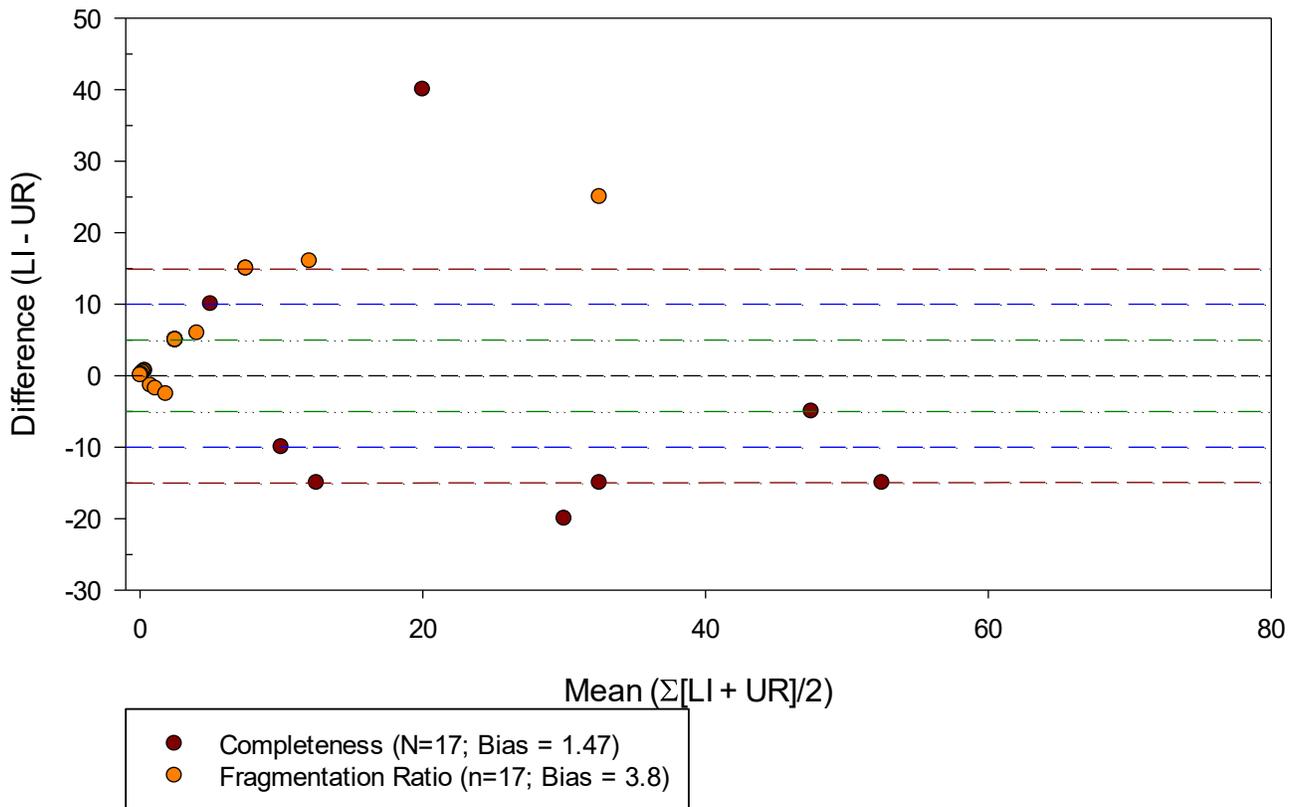
Difference Between Element Scores vs. Mean of Element Scores Burial 21



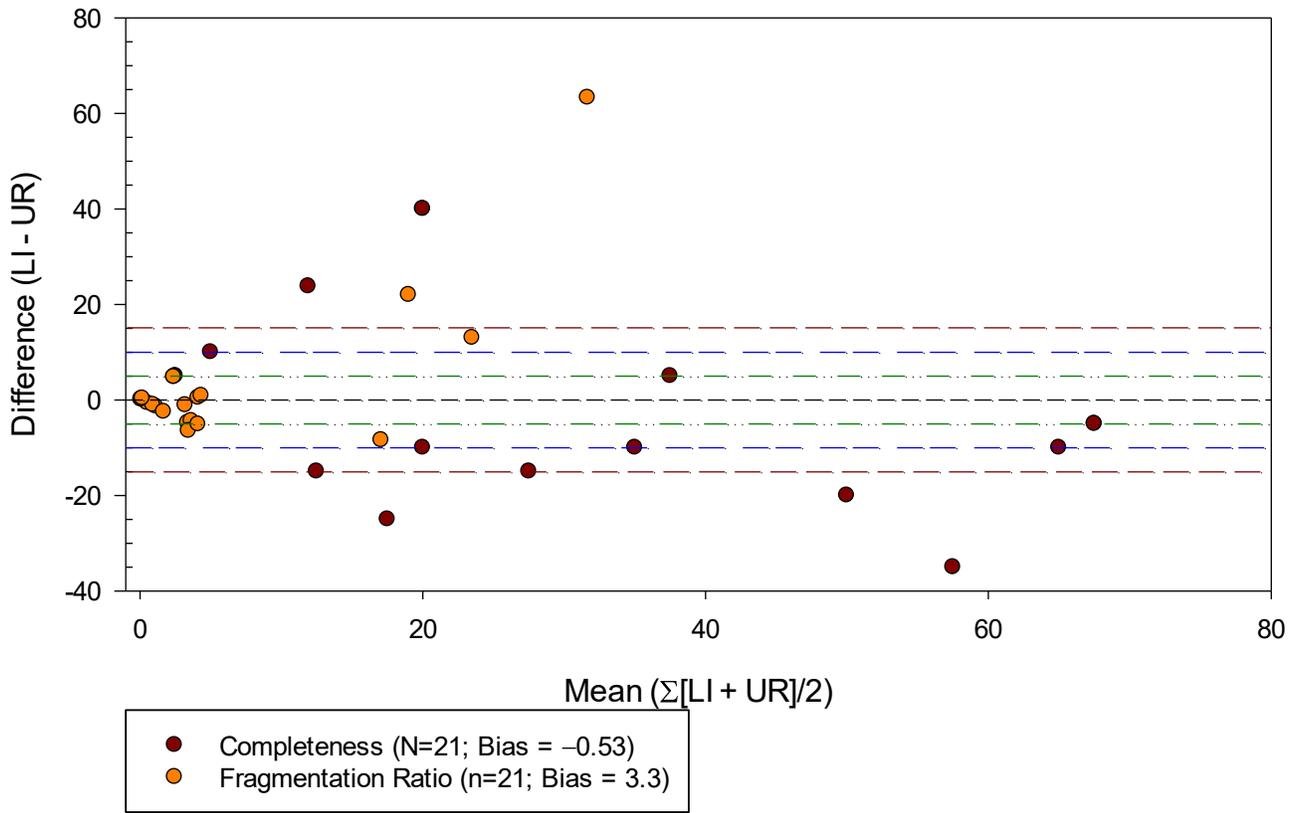
Difference Between Element Scores vs. Mean of Element Scores
Burial 22



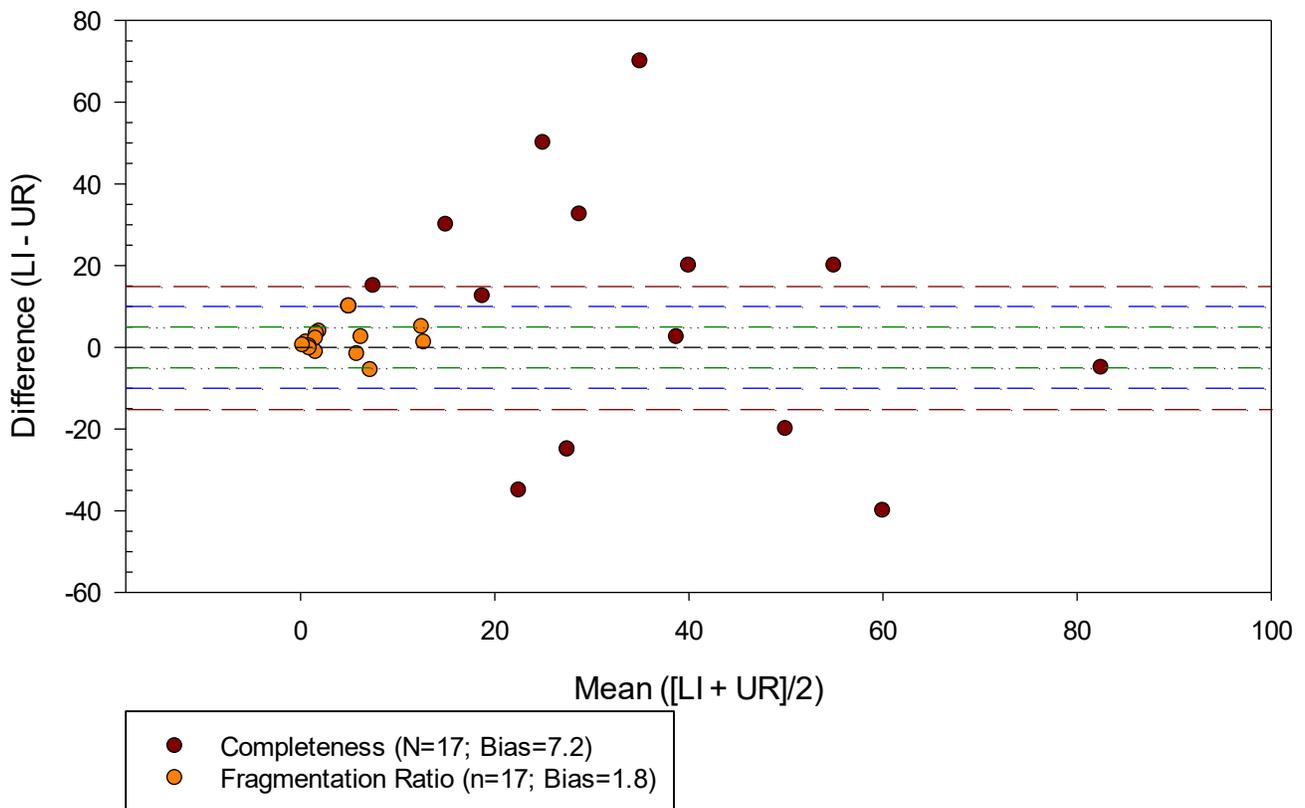
Difference Between Element Scores vs. Mean of Element Scores
Burial 23



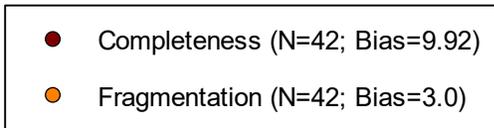
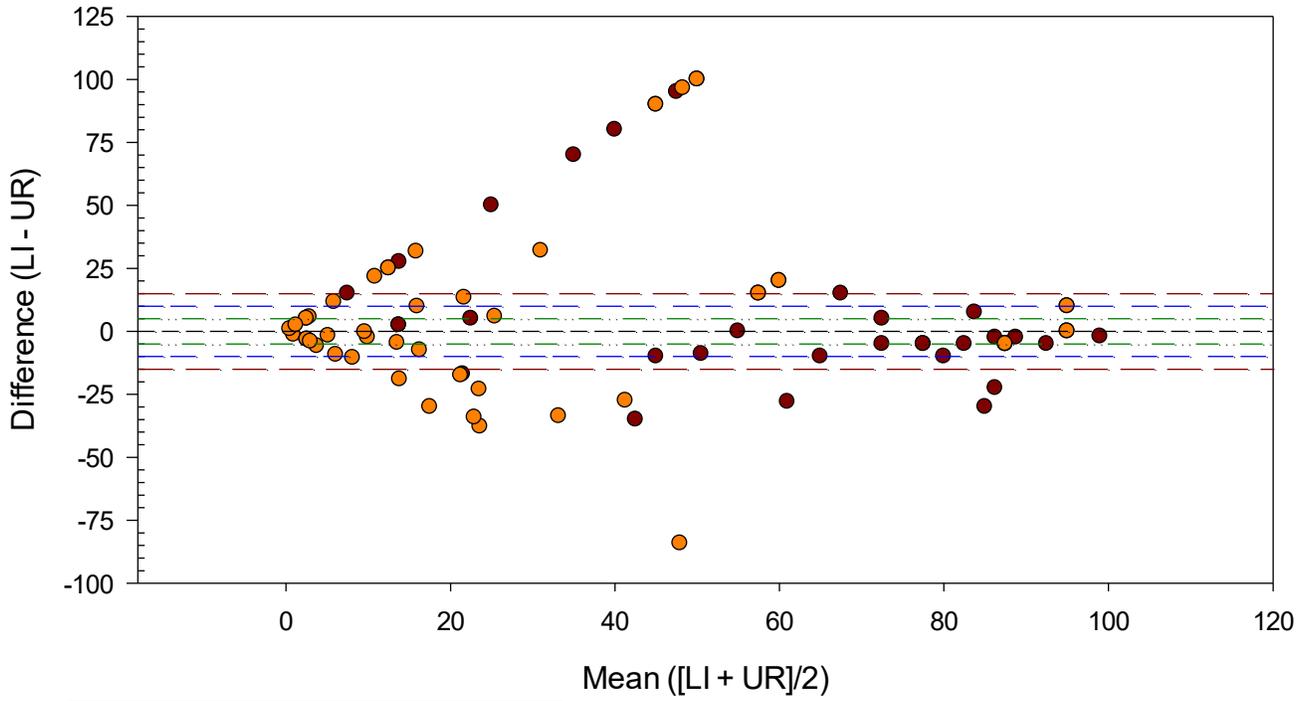
Difference Between Element Scores vs. Mean of Element Scores Burial 24



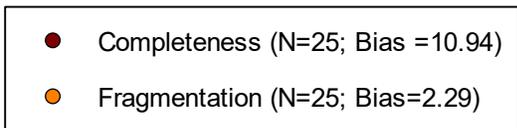
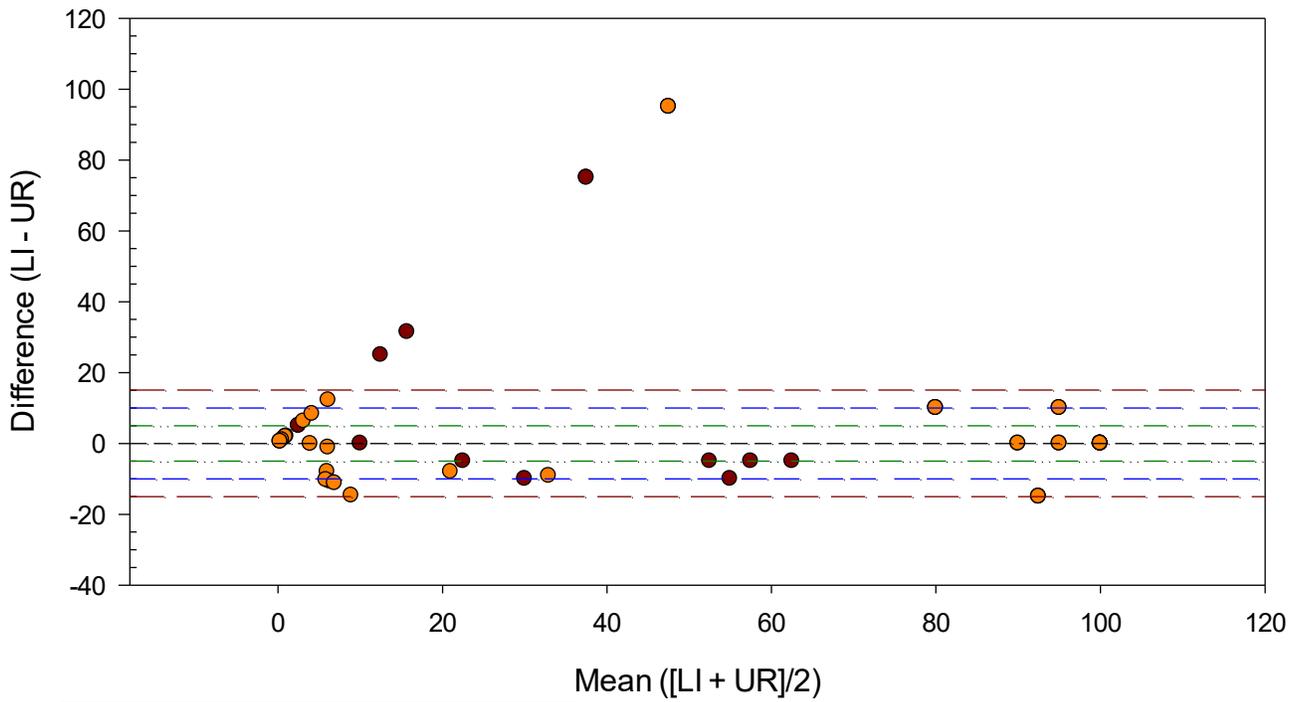
Difference Between Element Scores vs. Mean of Element Scores Burial 25



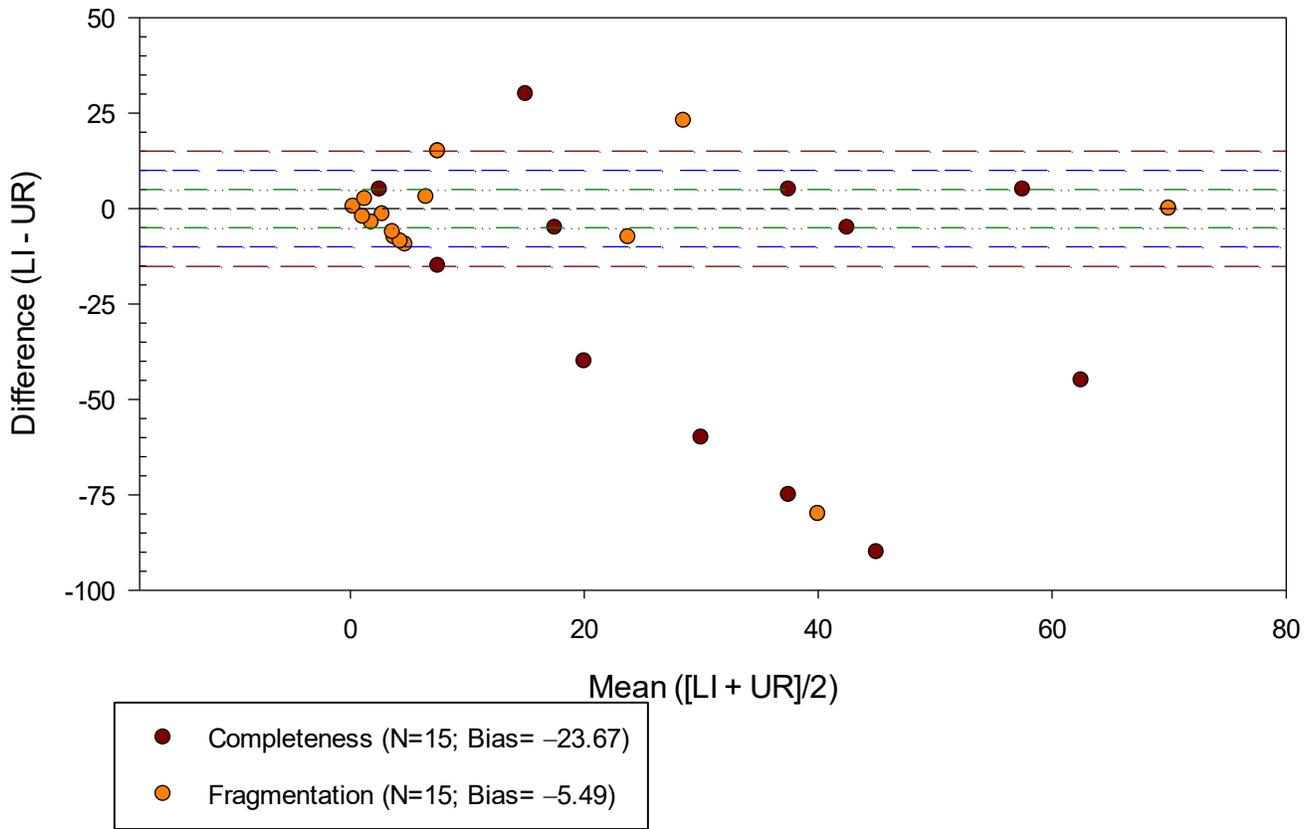
Difference Between Element Scores vs. Mean of Element Scores
Burial 27.01



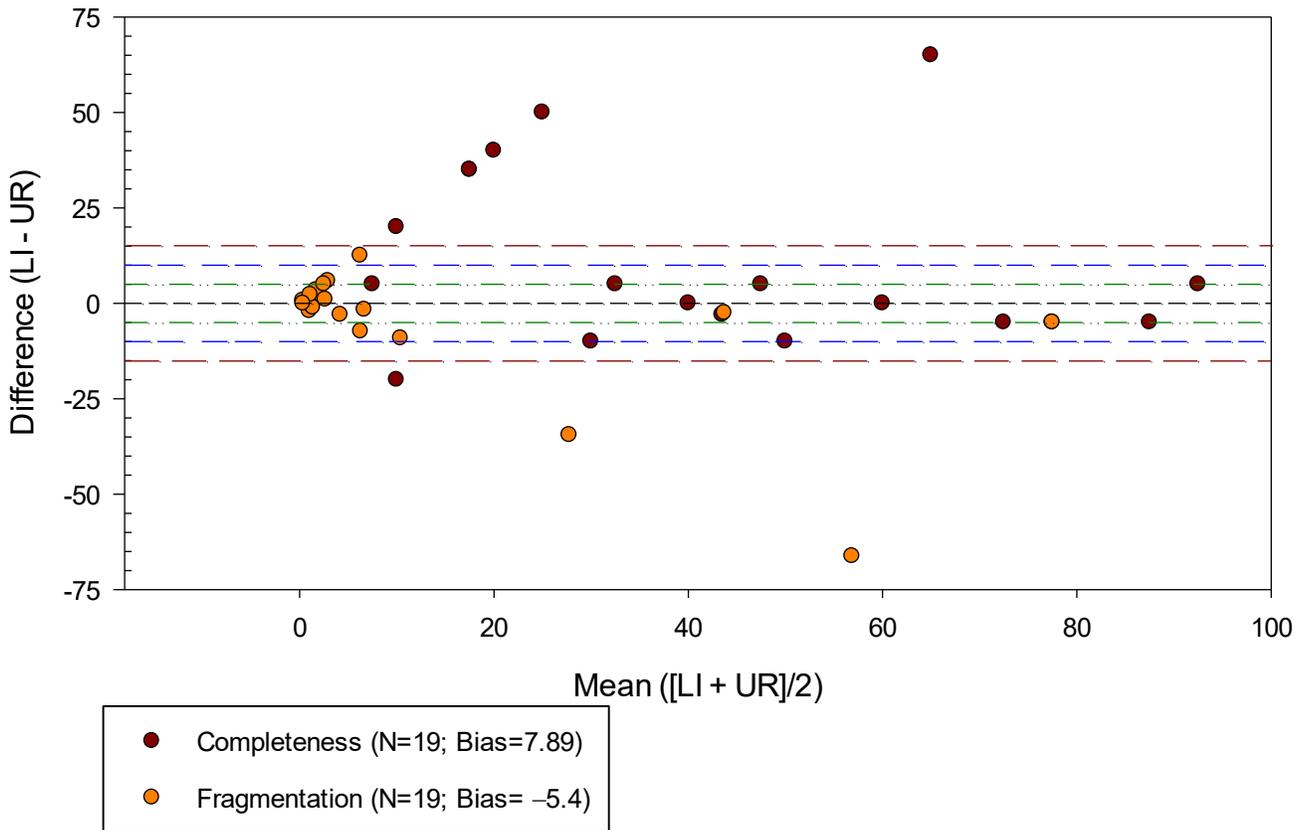
Difference Between Element Scores vs. Mean of Element Scores
Burial 27.02



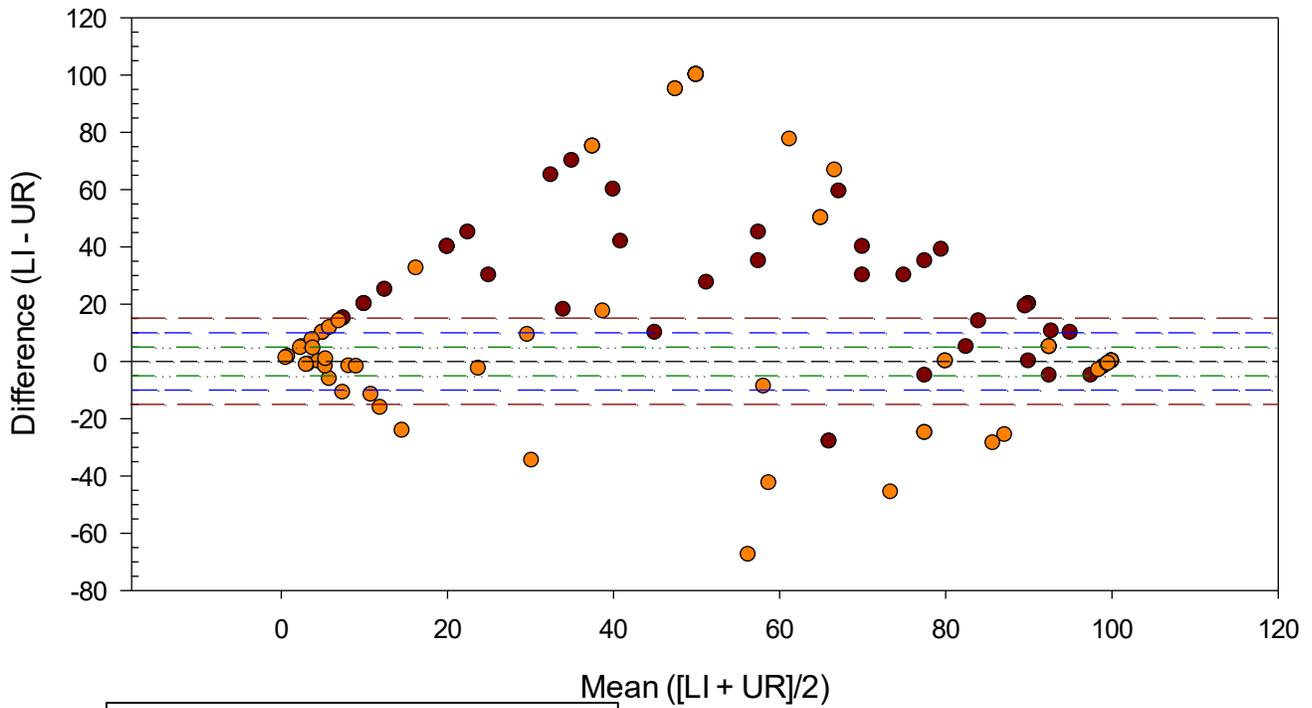
Difference Between Element Scores vs. Mean of Element Scores Burial 27.03



Difference Between Element Scores vs. Mean of Element Scores Burial 28

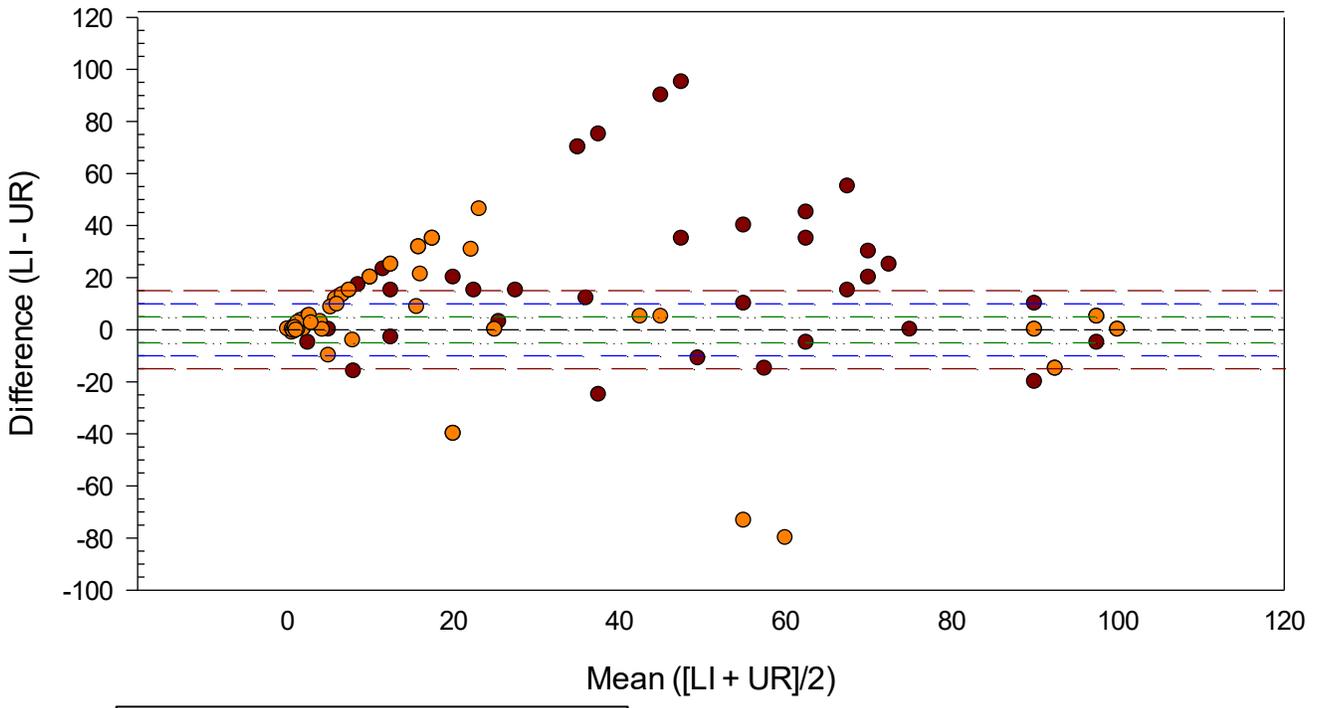


Difference Between Element Scores vs. Mean of Element Scores
Burial 29



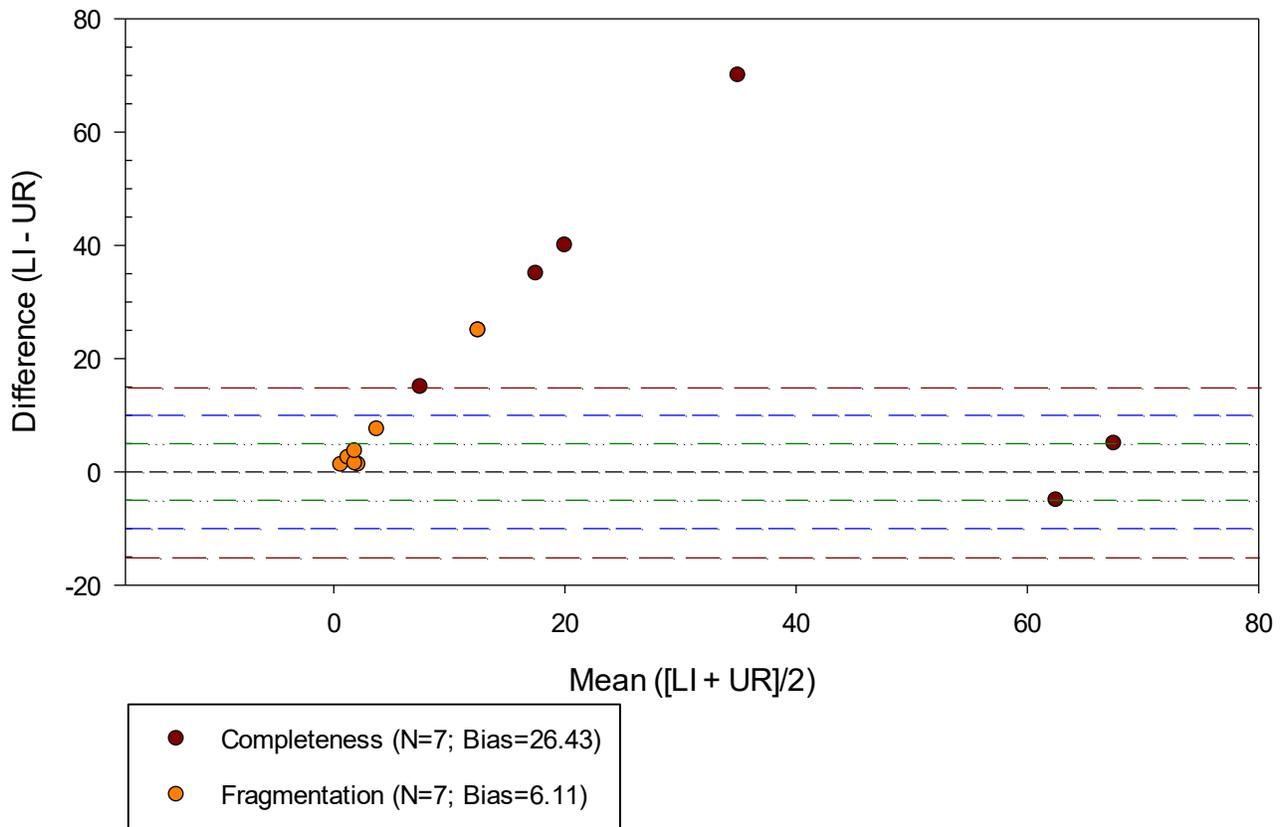
- Completeness (N=54; Bias=24.66)
- Fragmentation (N=54; Bias=15.32)

Difference Between Element Scores vs. Mean of Element Scores
Burial 32

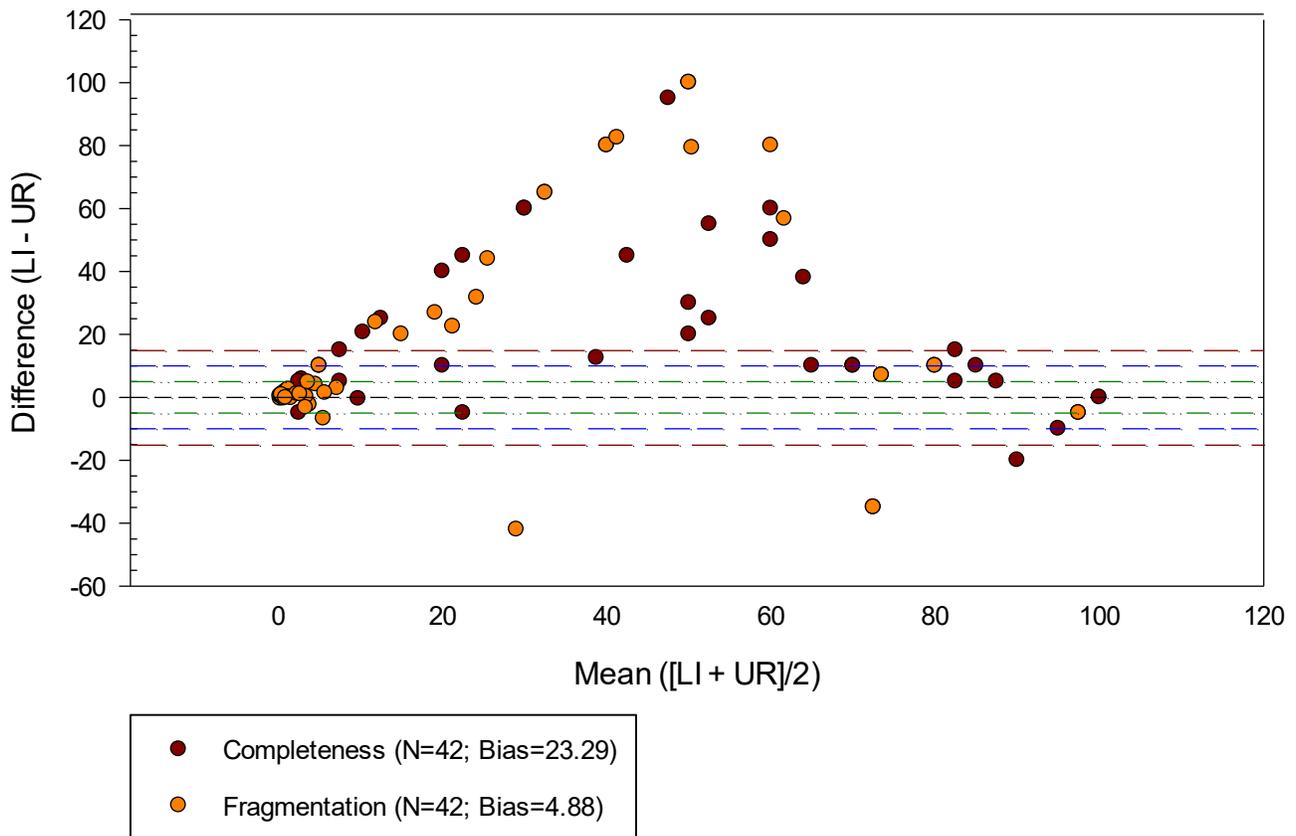


- Completeness (N=47; Bias=17.13)
- Fragmentation (N=47; Bias=2.84)

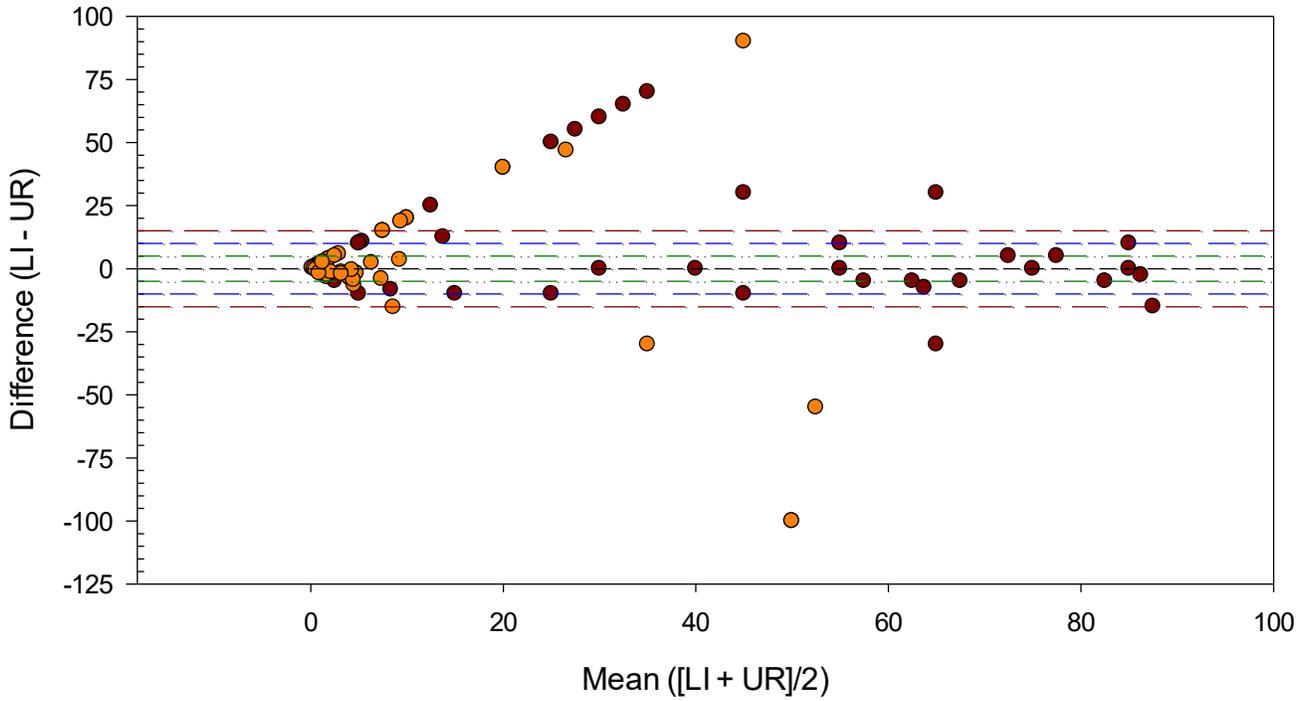
Difference Between Element Scores vs. Mean of Element Scores
Burial 33



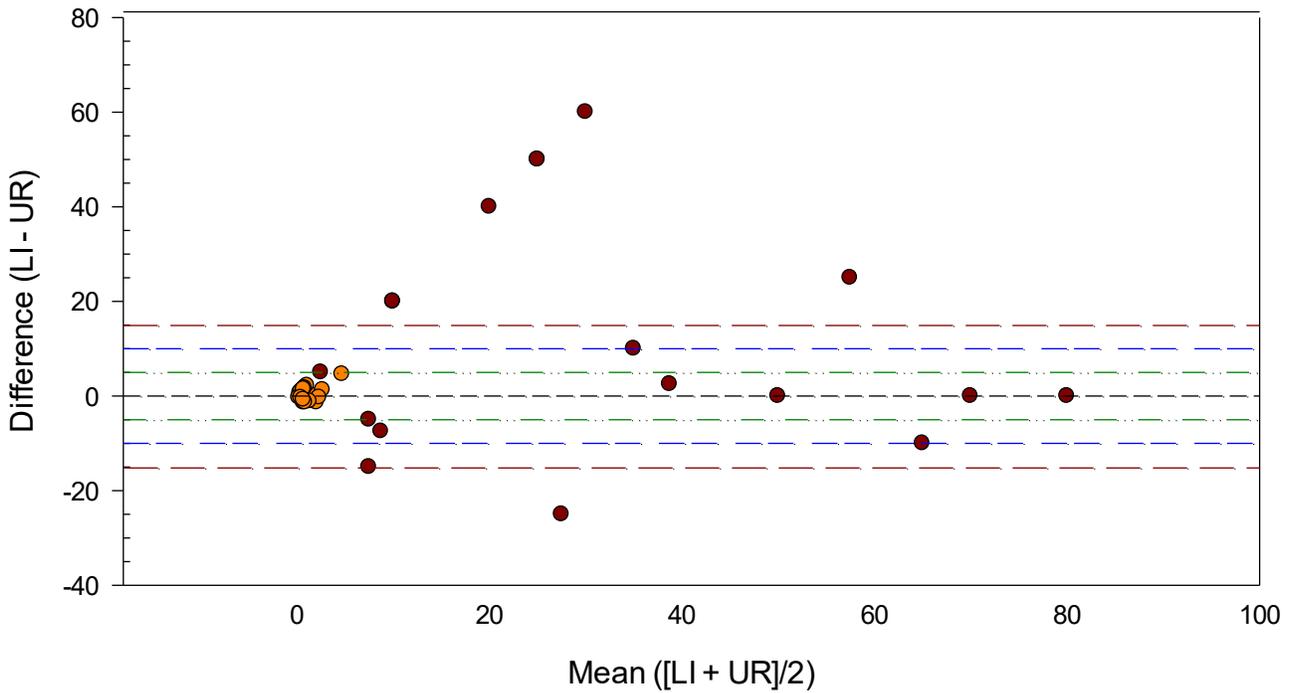
Difference Between Element Scores vs. Mean of Element Scores
Burial 34



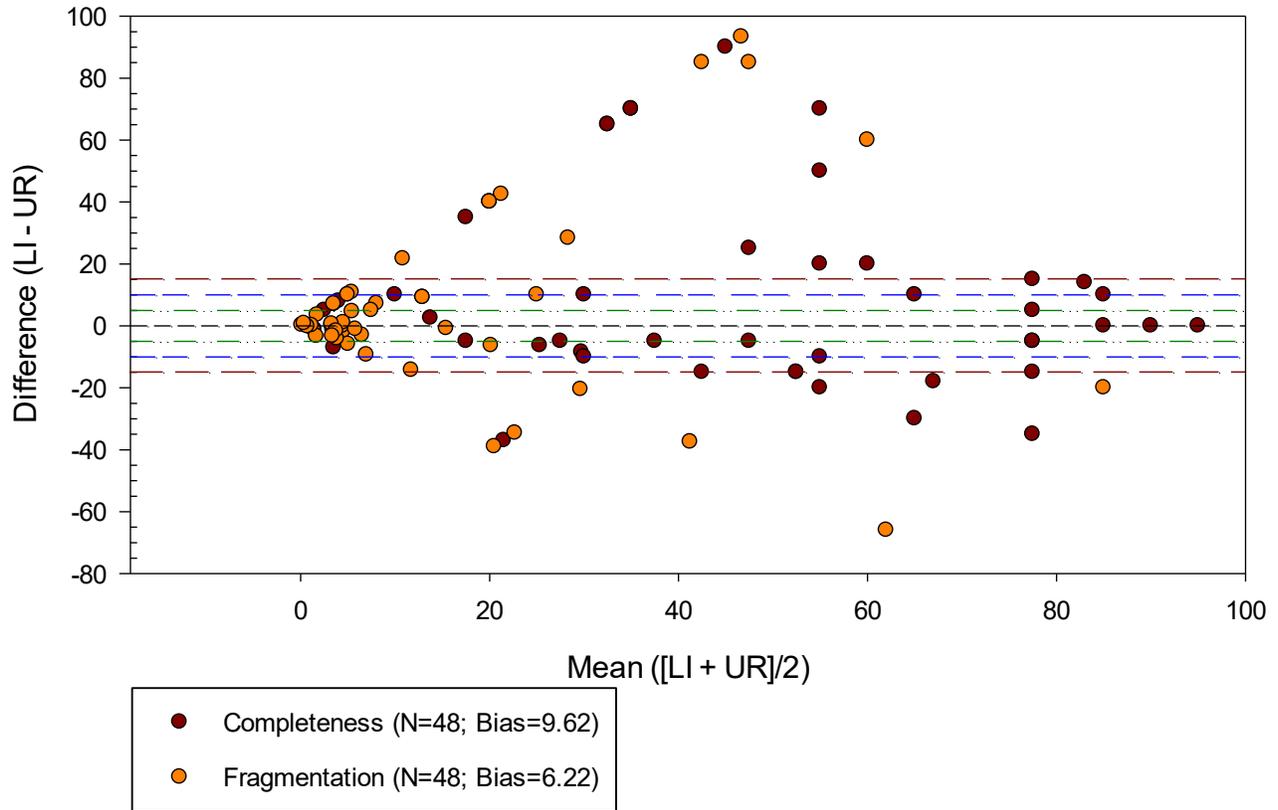
Difference Between Element Scores vs. Mean of Element Scores
Burial 35.01



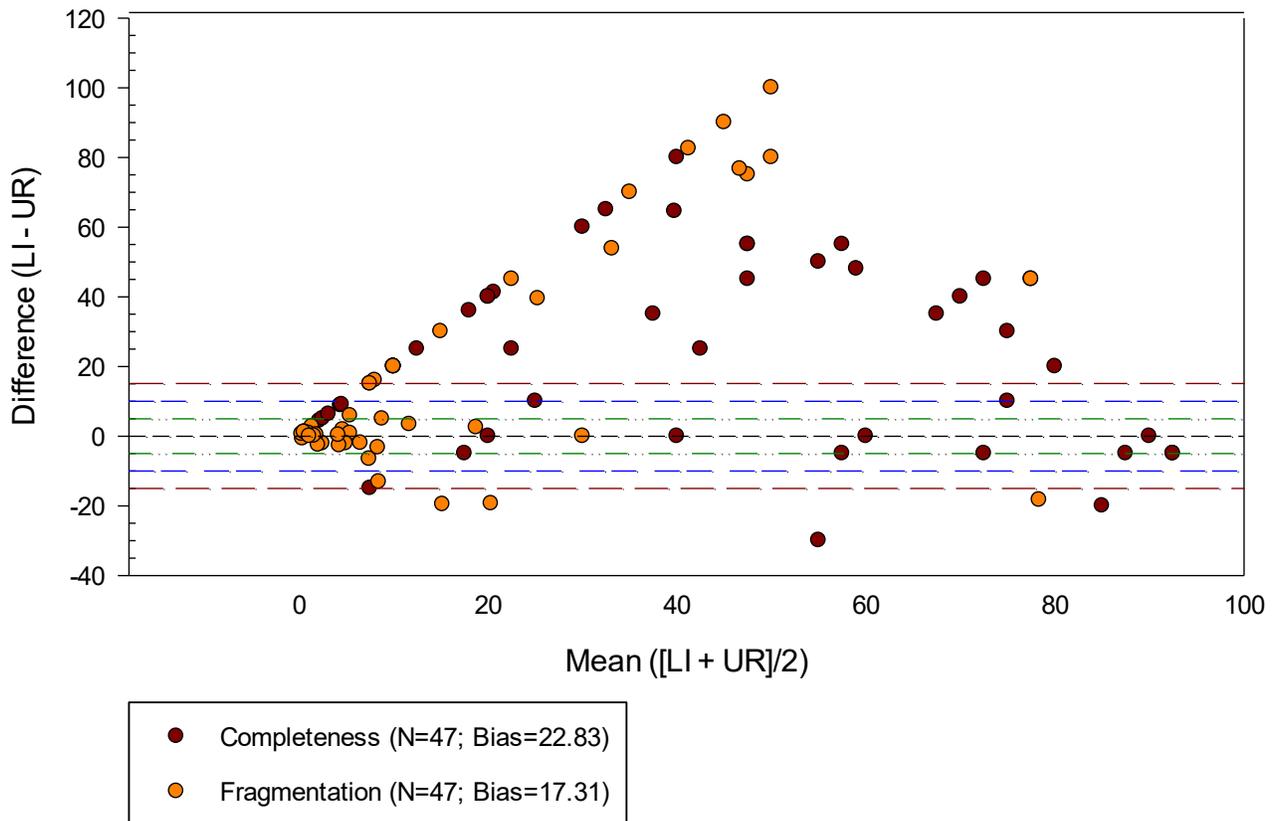
Difference Between Element Scores vs. Mean of Element Scores
Burial 35.02



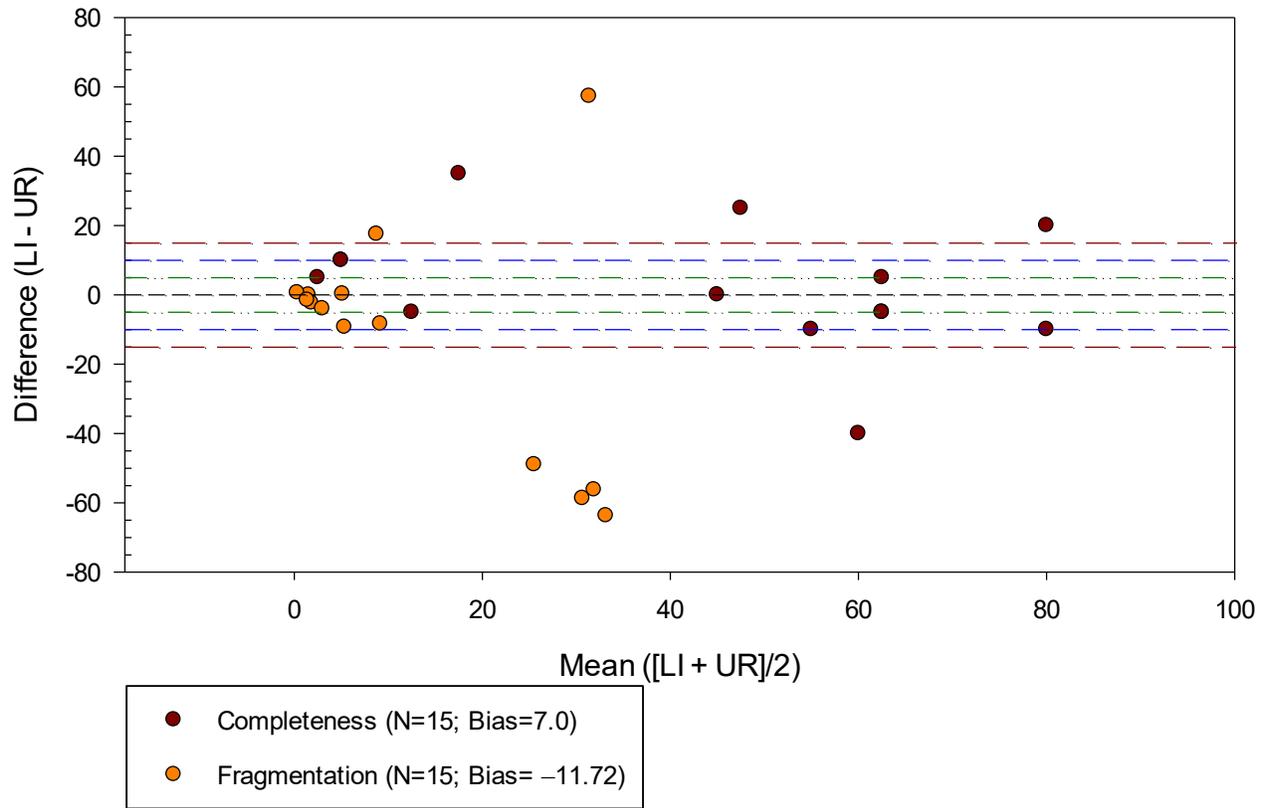
Difference Between Element Scores vs. Mean of Element Scores
Burial 37.02



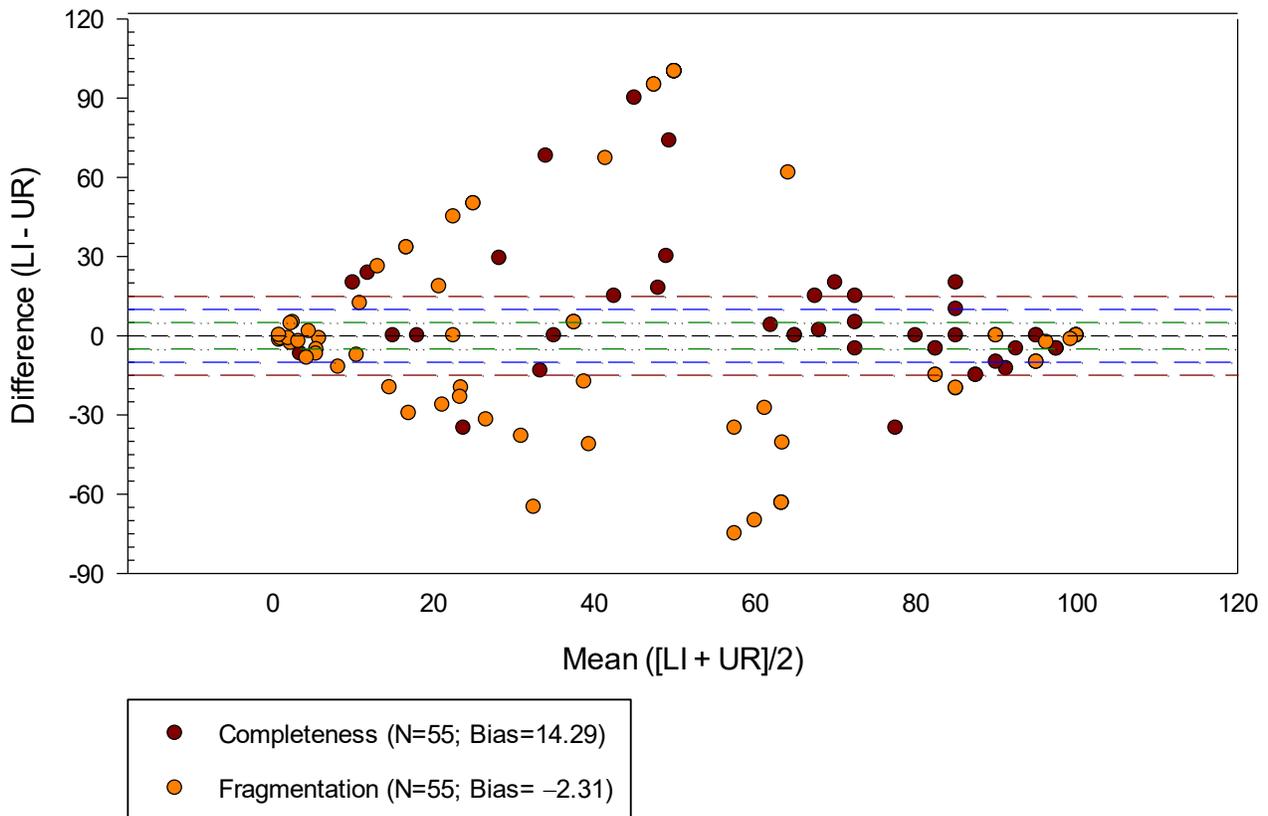
Difference Between Element Scores vs. Mean of Element Scores
Burial 38



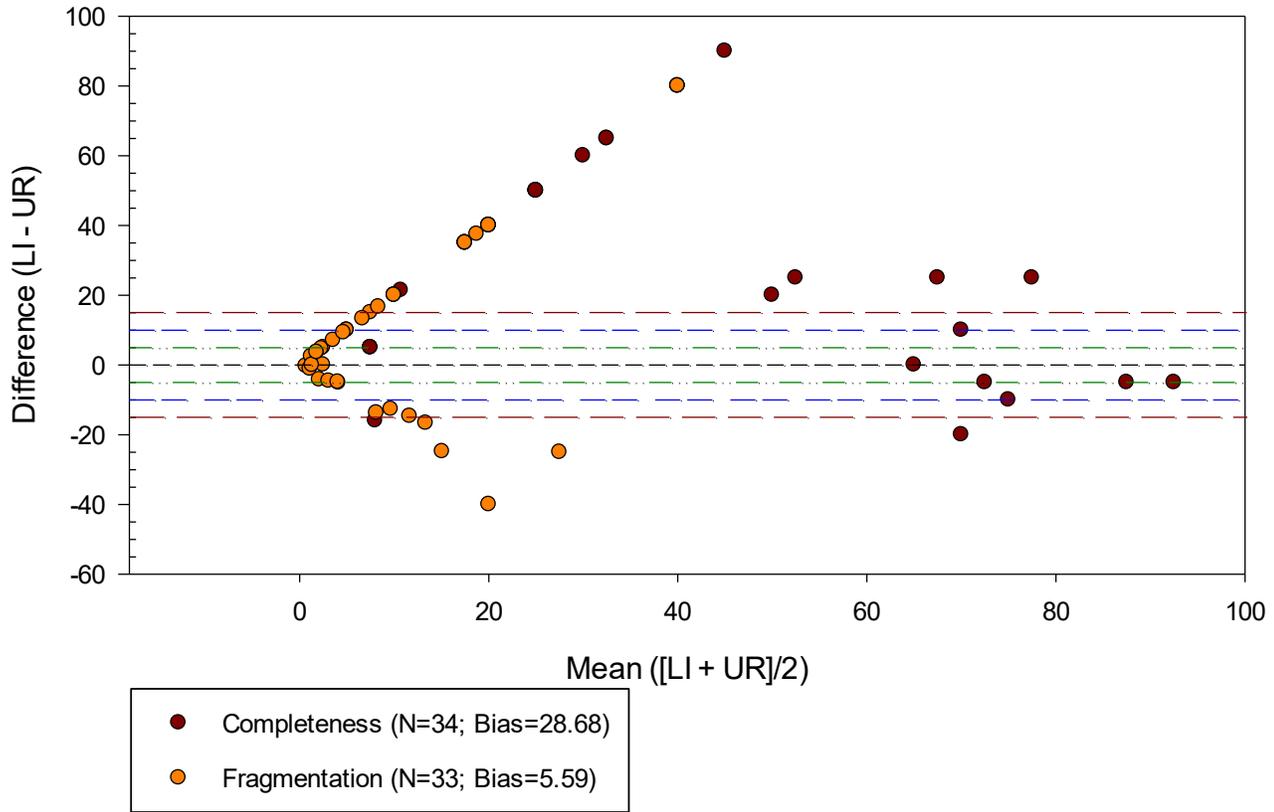
Difference Between Element Scores vs. Mean of Element Scores Burial 39



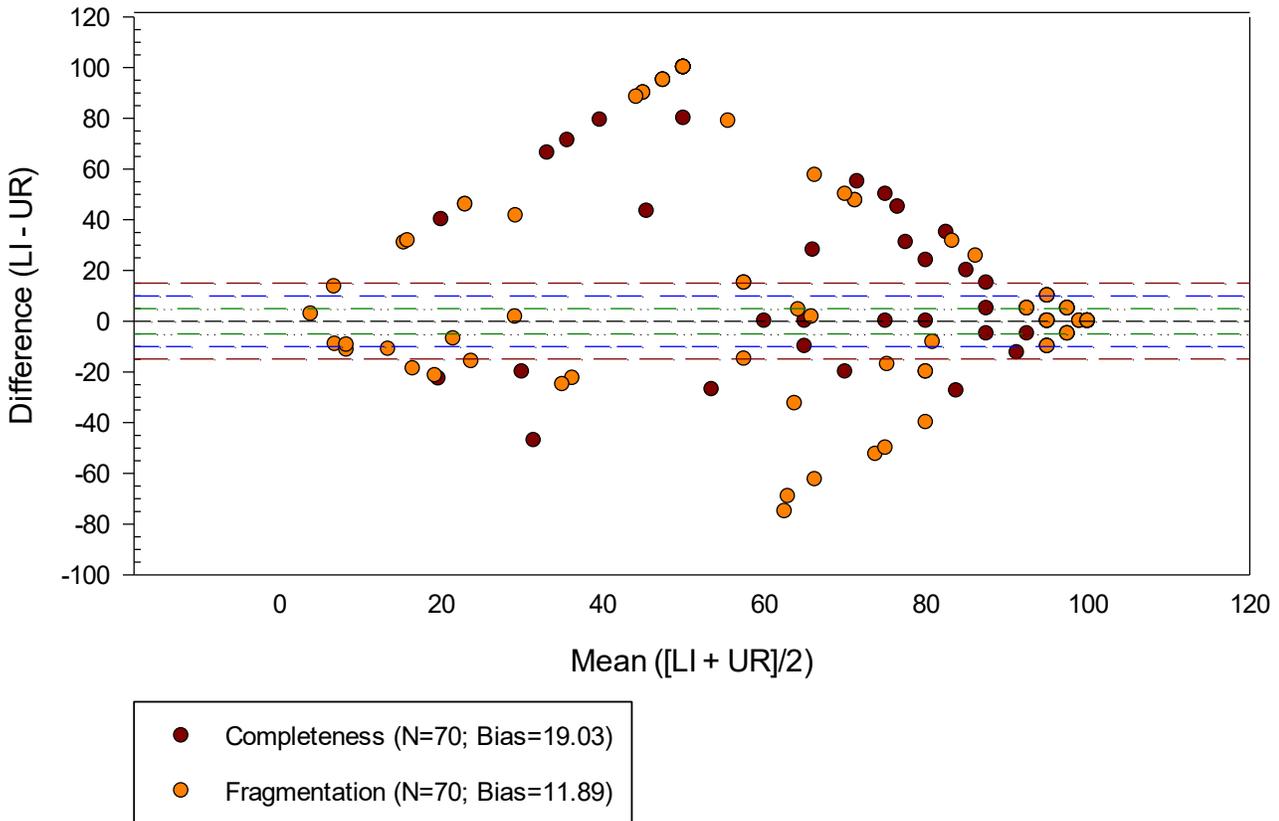
Difference Between Element Scores vs. Mean of Element Scores Burial 44



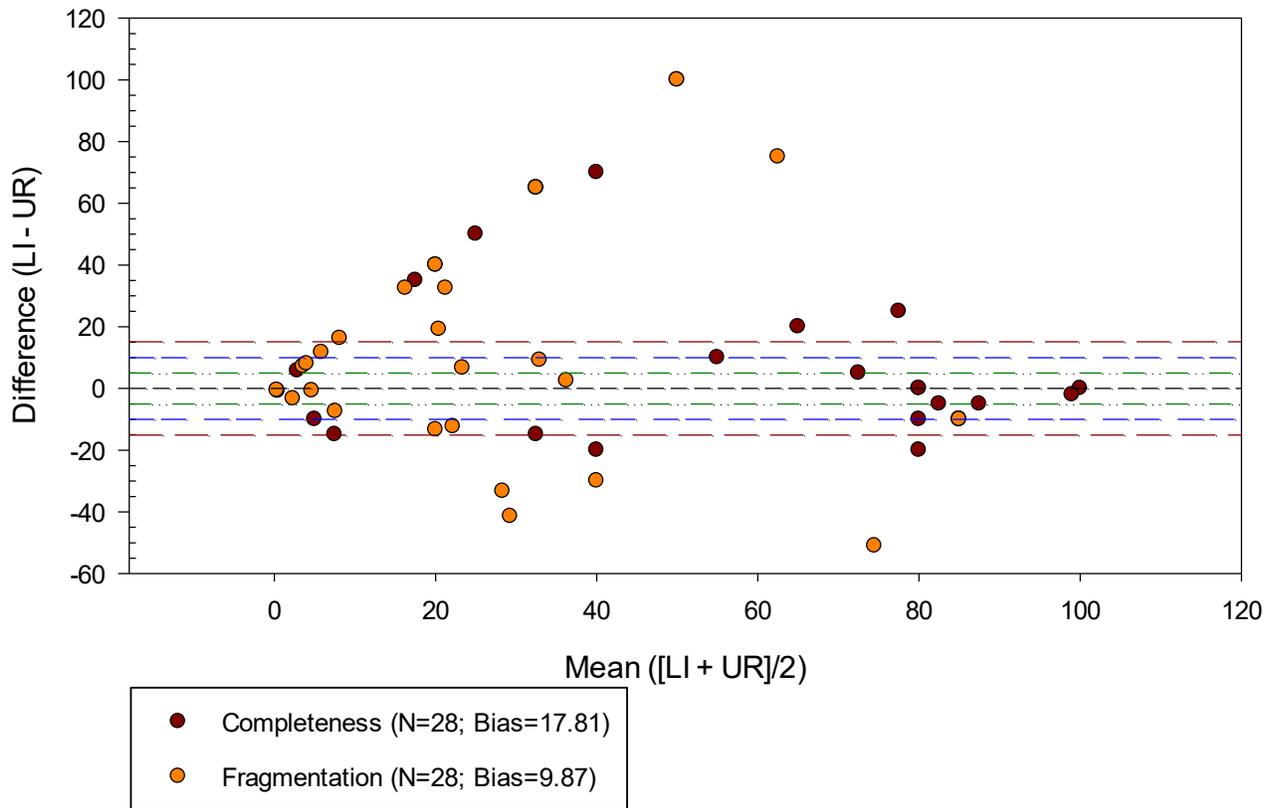
Difference Between Element Scores vs. Mean of Element Scores
Burial 45



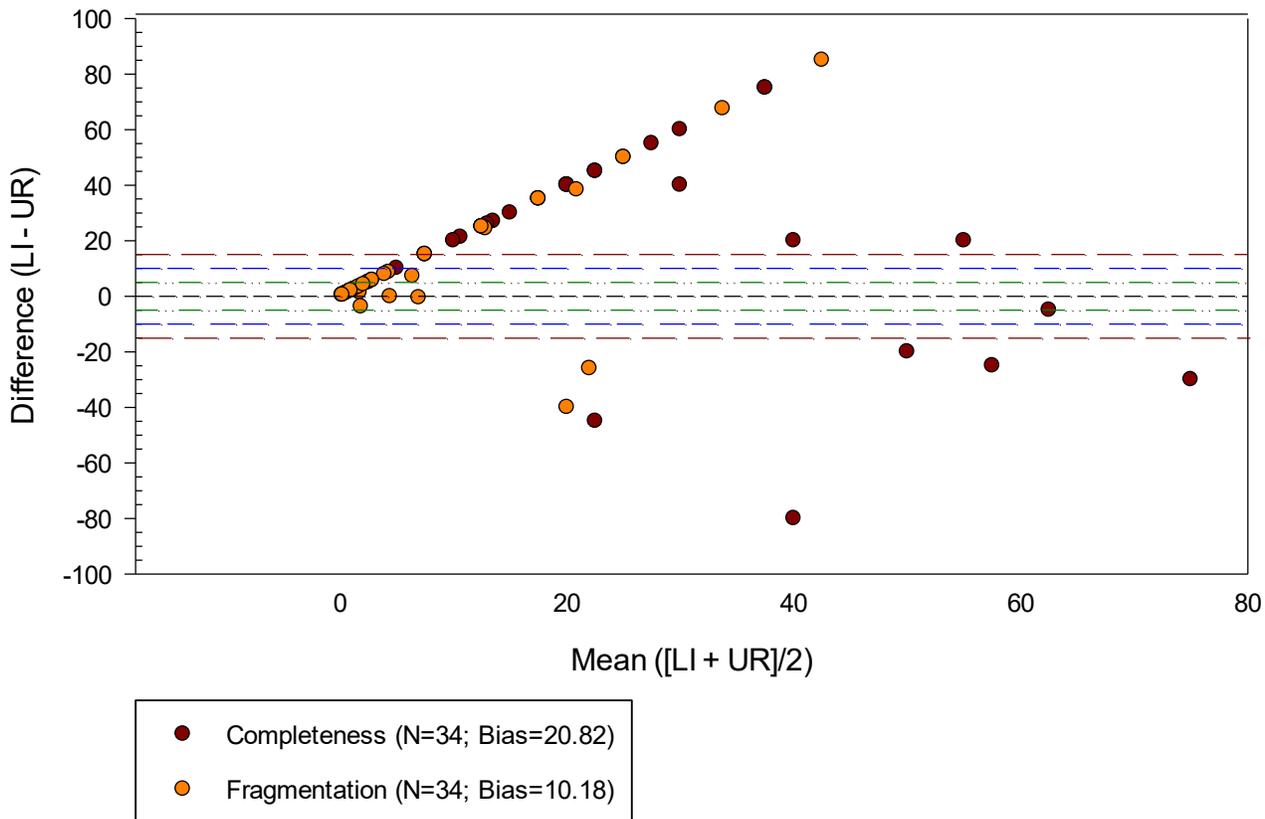
Difference Between Element Scores vs. Mean of Element Scores
Burial 46



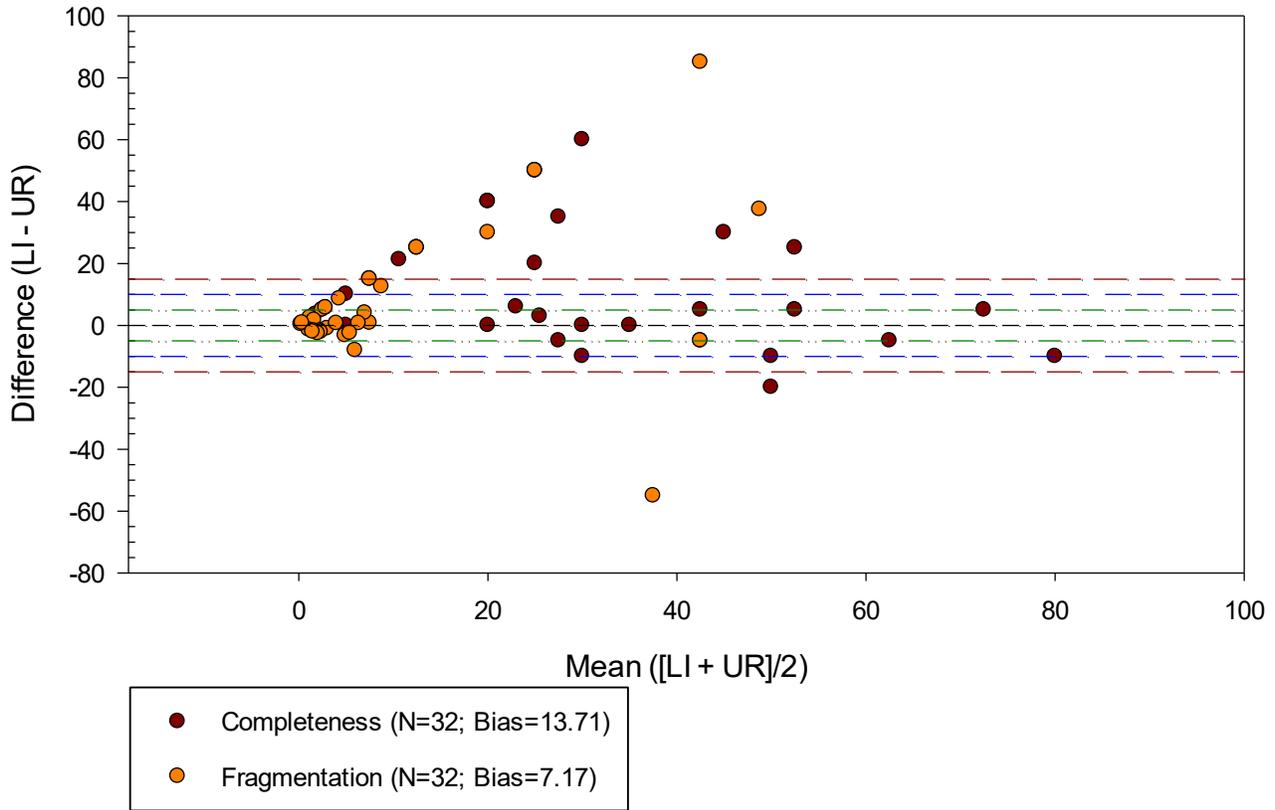
Difference Between Element Scores vs. Mean of Element Scores
Burial 47



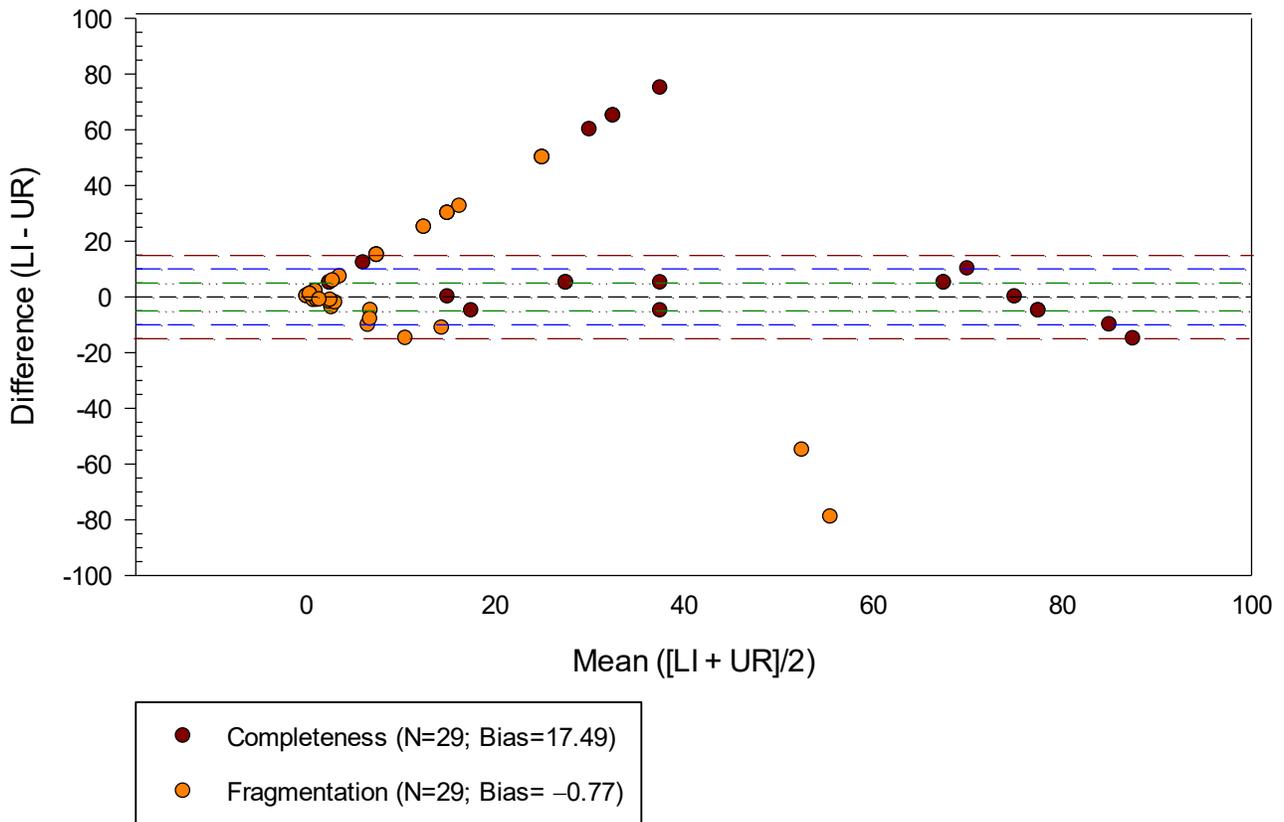
Difference Between Element Scores vs. Mean of Element Scores
Burial 48



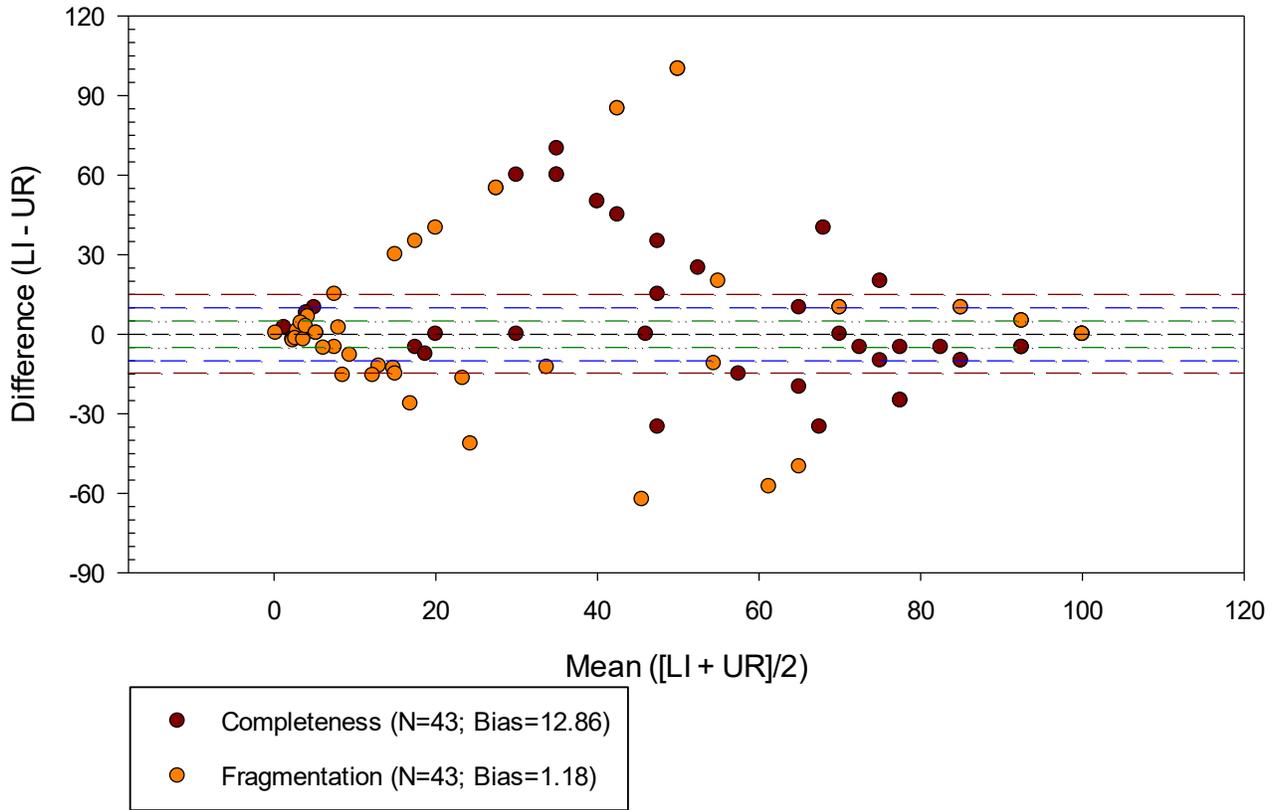
Difference Between Element Scores vs. Mean of Element Scores
Burial 49



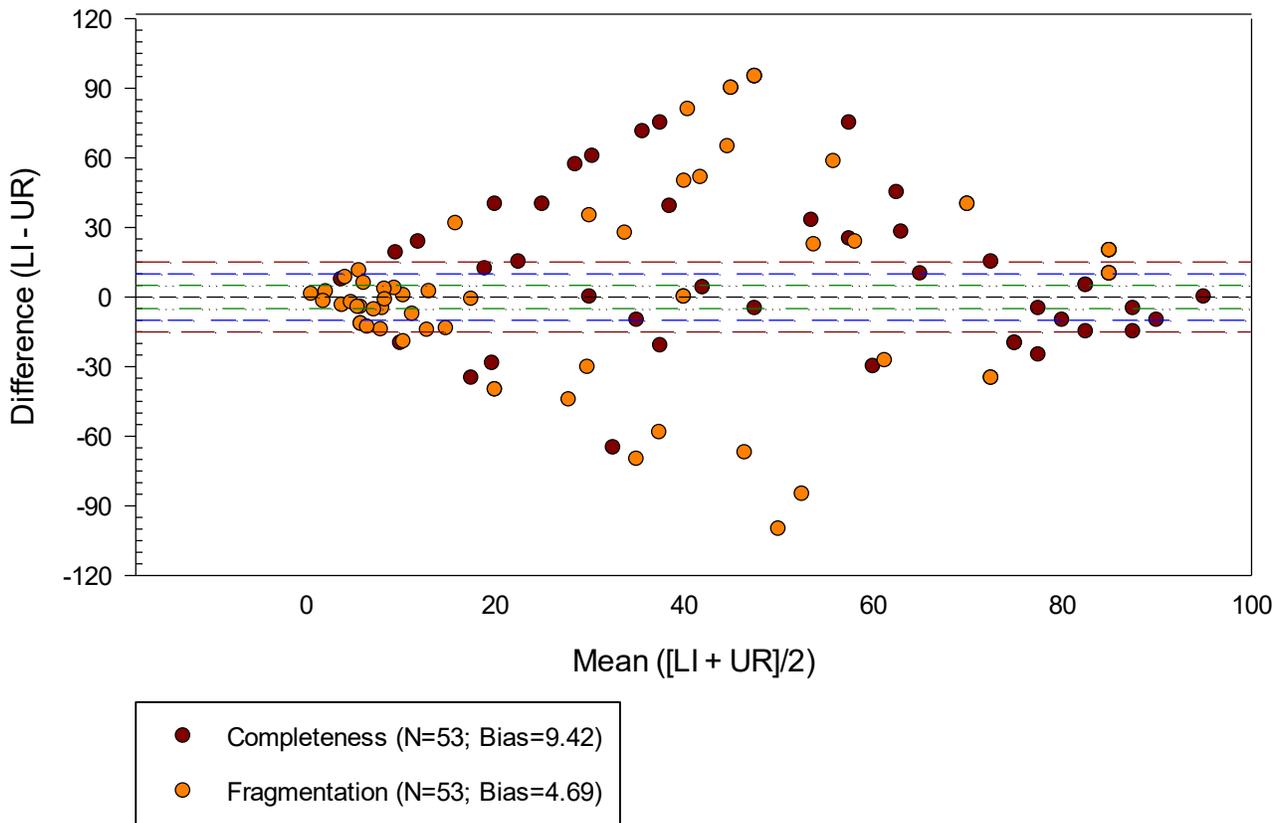
Difference Between Element Scores vs. Mean of Element Scores
Burial 50



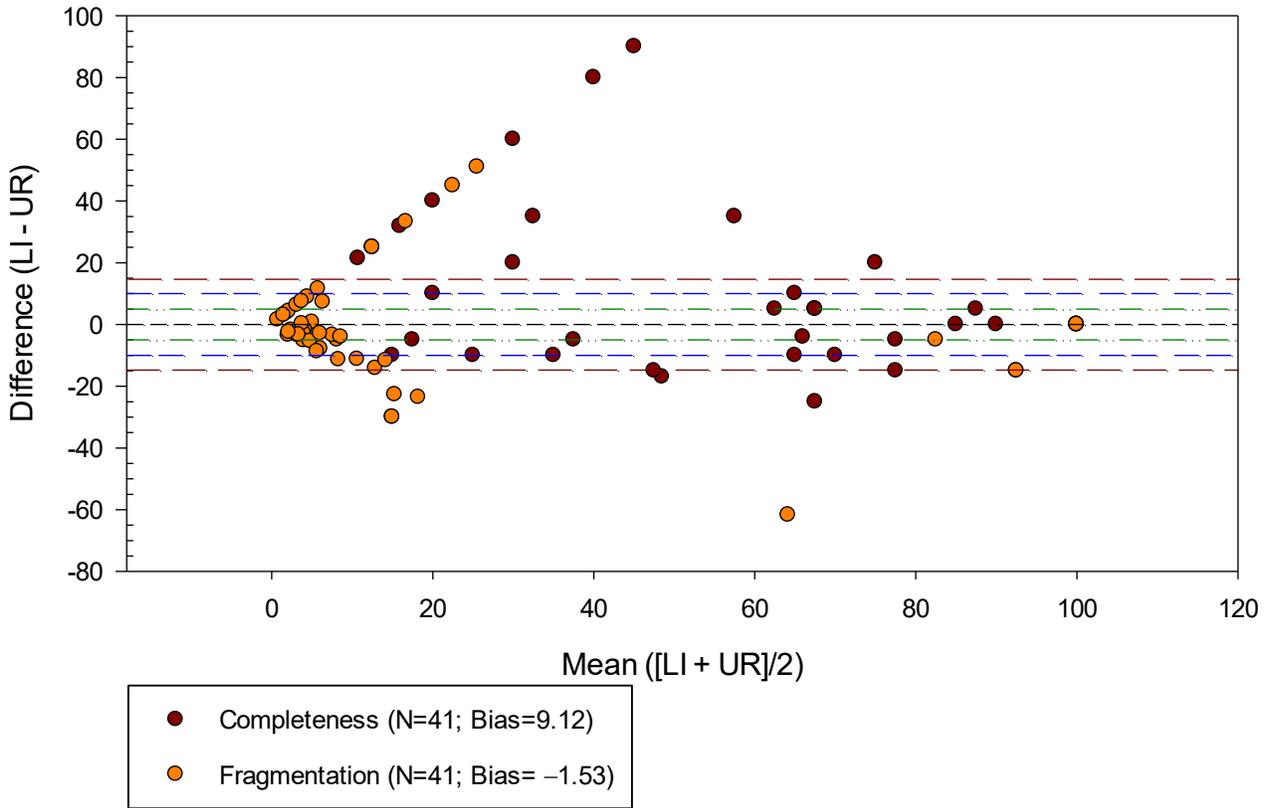
Difference Between Element Scores vs. Mean of Element Scores
Burial 51



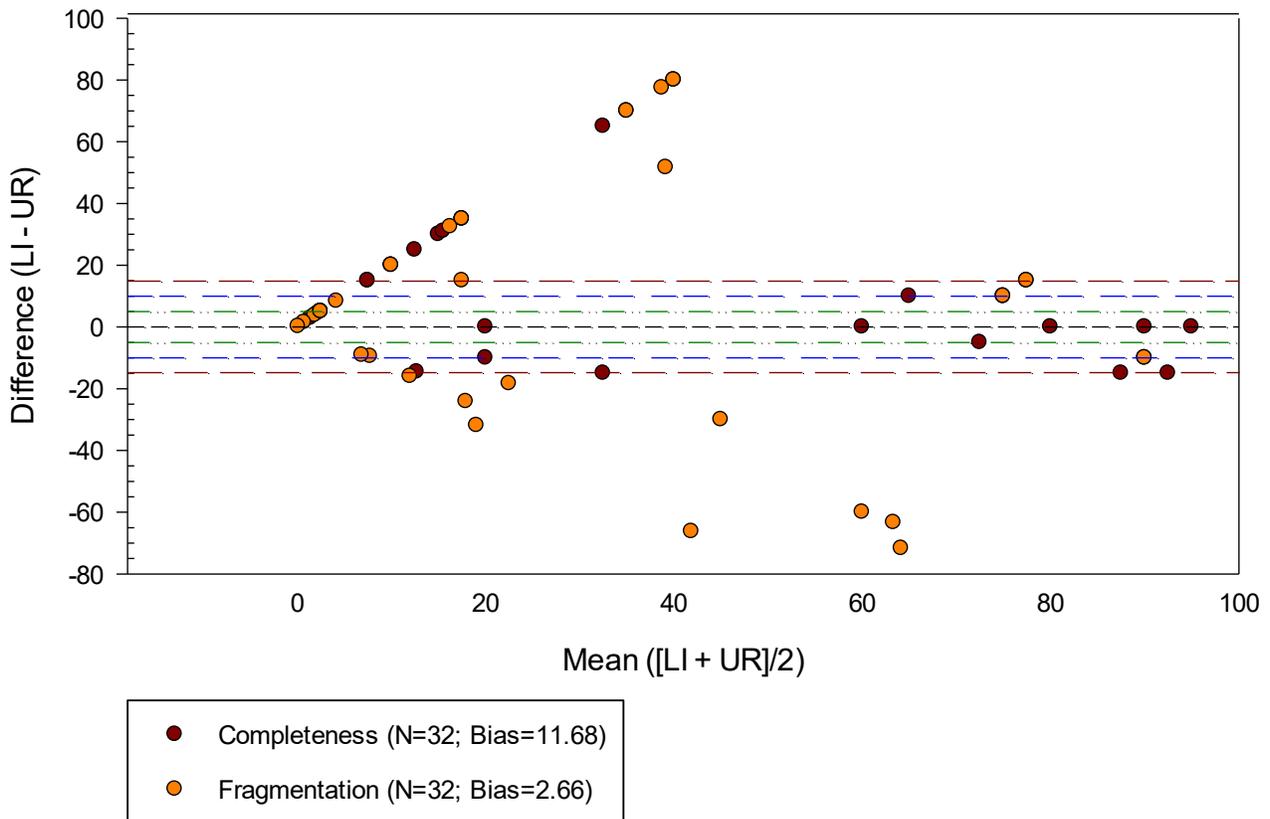
Difference Between Element Scores vs. Mean of Element Scores
Burial 53



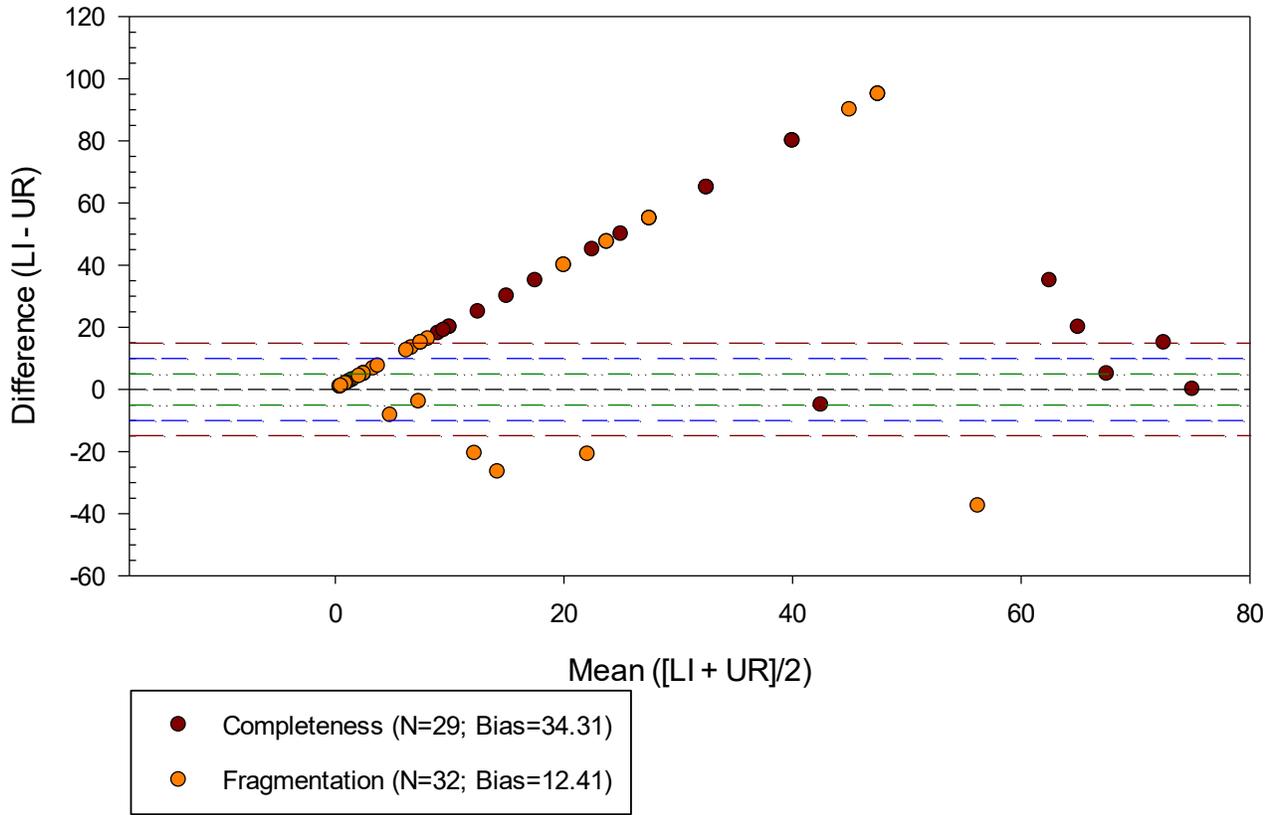
Difference Between Element Scores vs. Mean of Element Scores
Burial 55



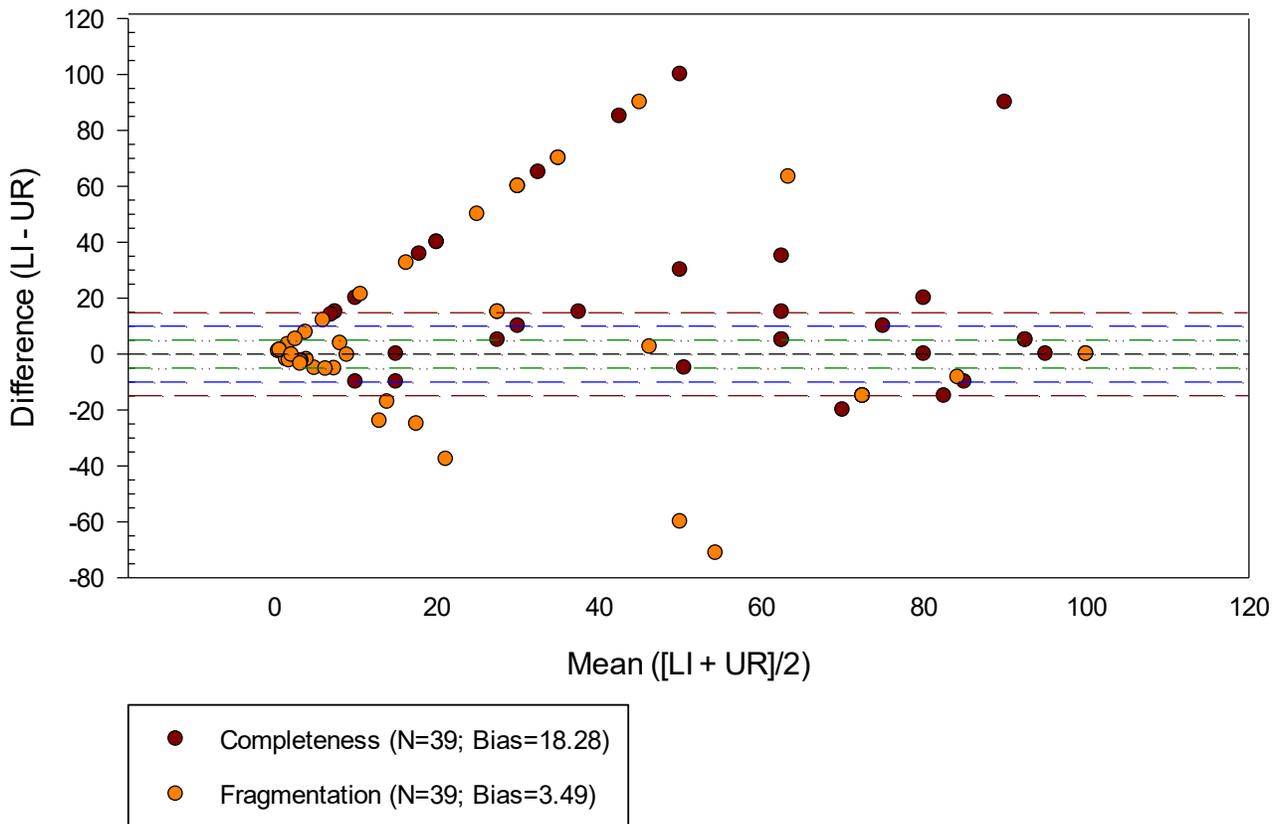
Difference Between Element Scores vs. Mean of Element Scores
Burial 57.01



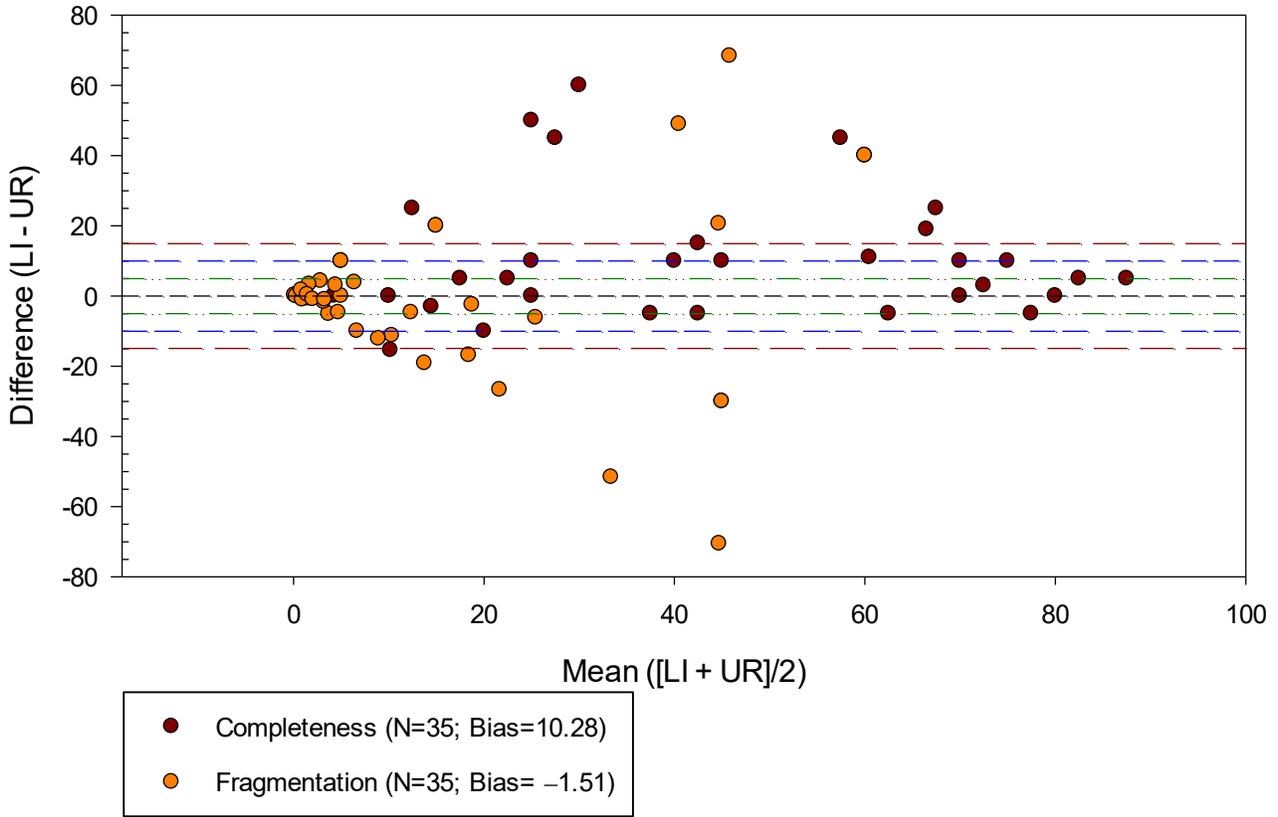
Difference Between Element Scores vs. Mean of Element Scores
Burial 57.02



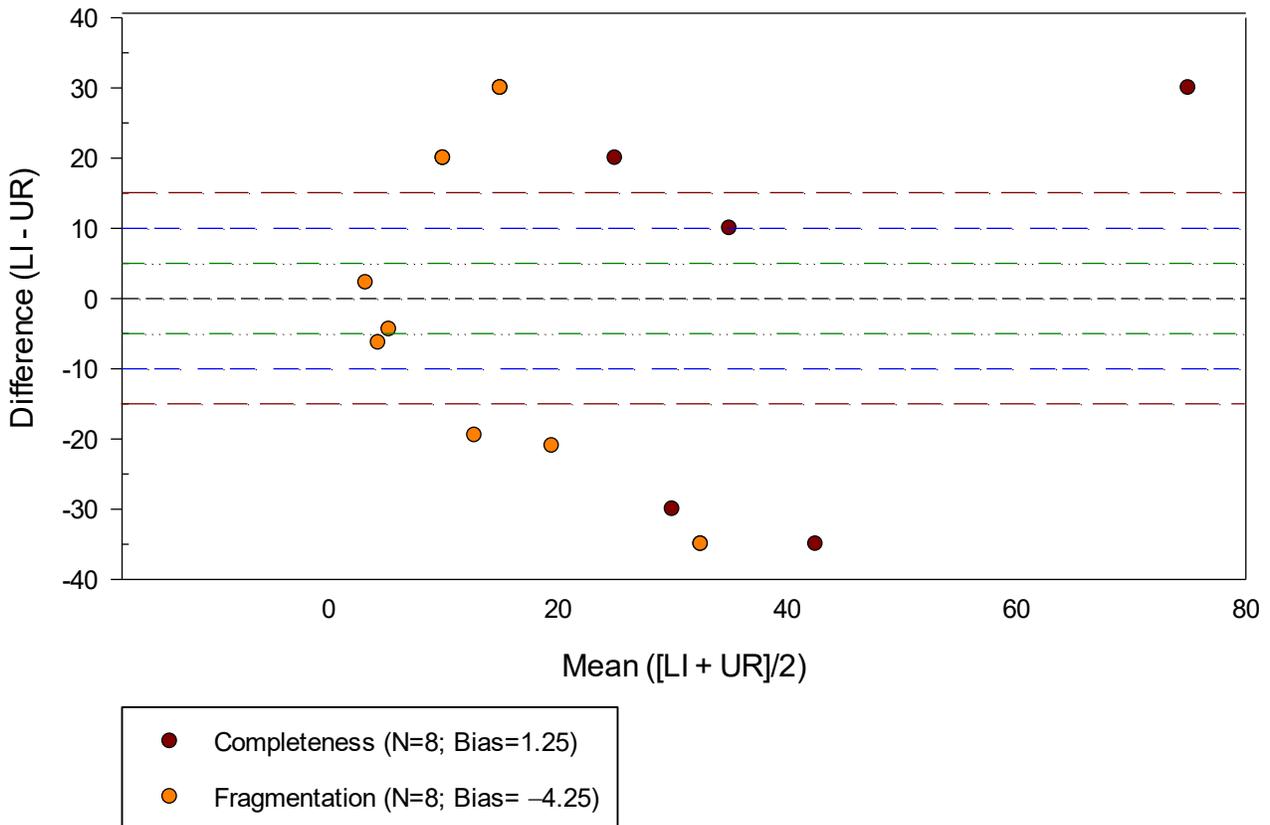
Difference Between Element Scores vs. Mean of Element Scores
Burial 58.01



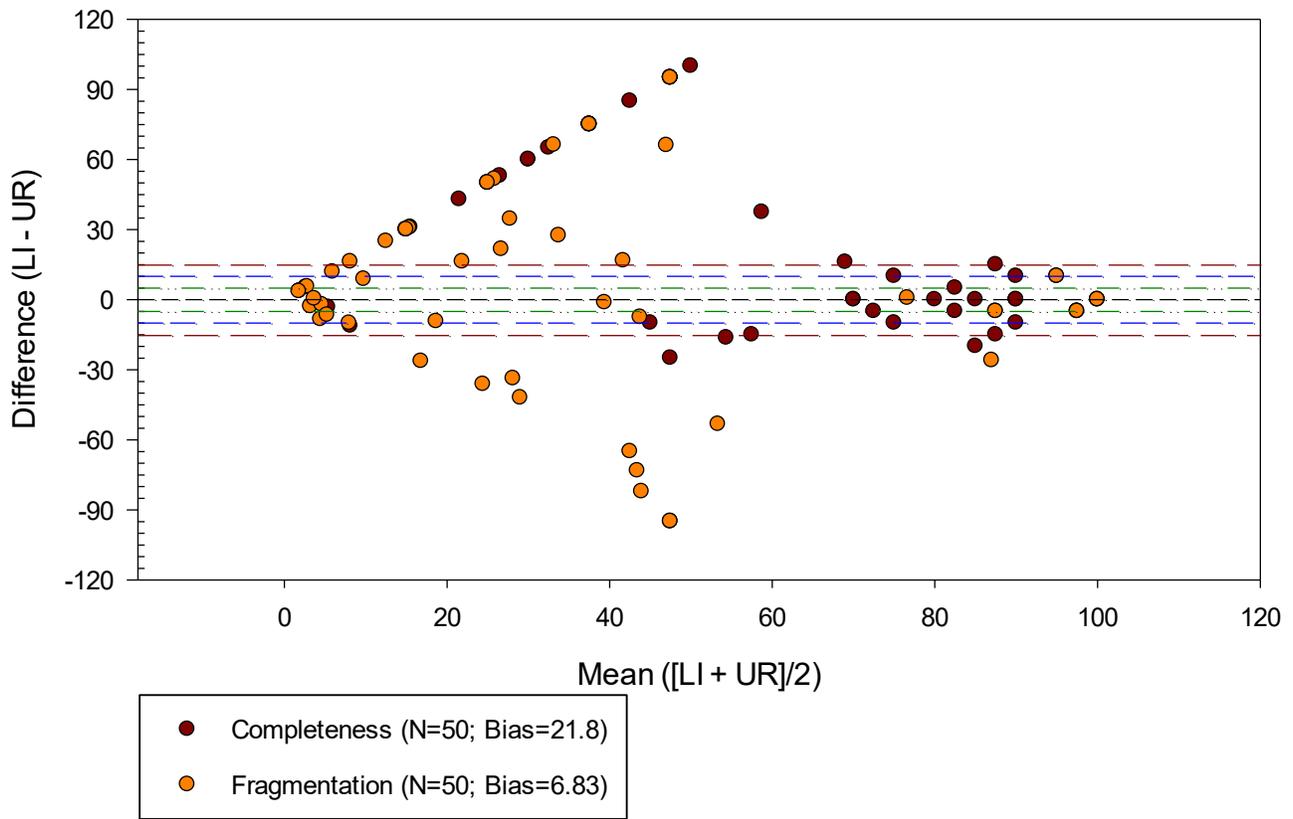
Difference Between Element Scores vs. Mean of Element Scores
Burial 58.02



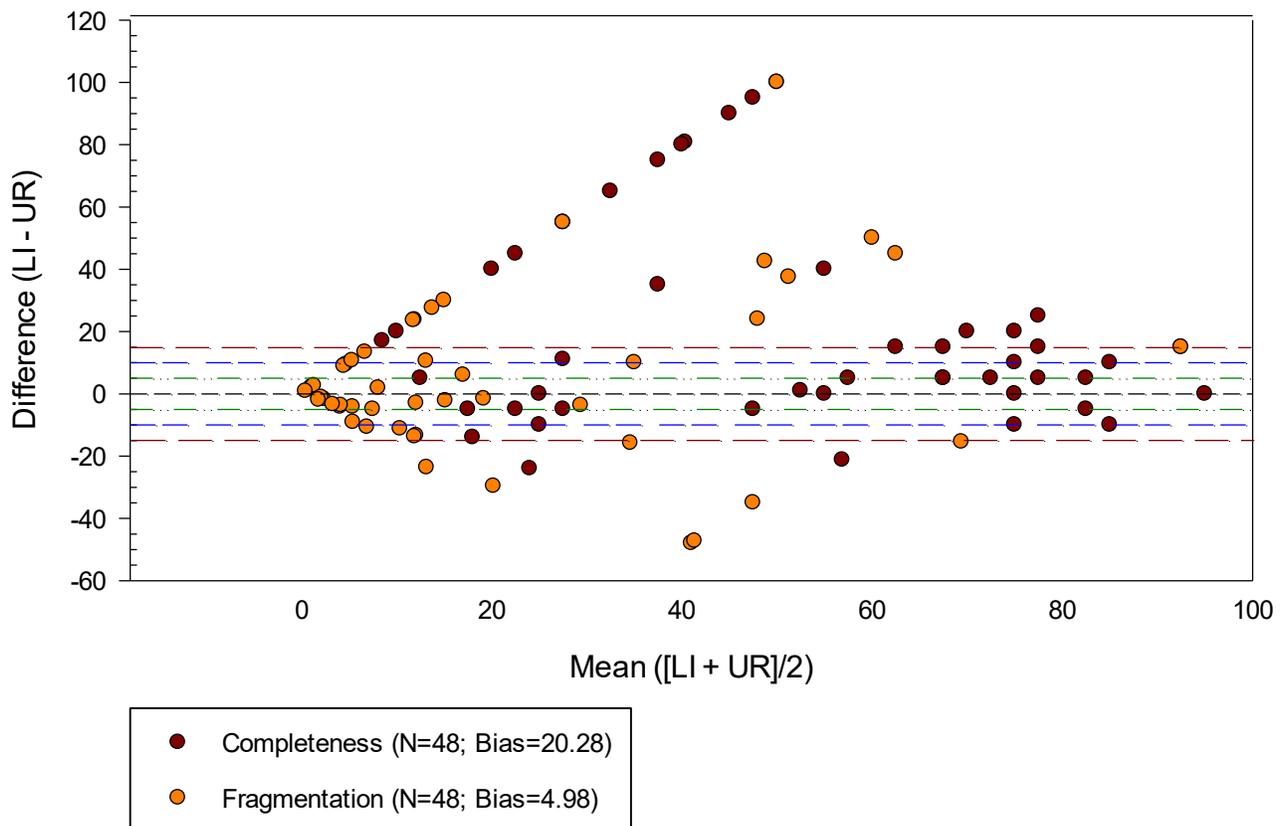
Difference Between Element Scores vs. Mean of Element Scores
Burial 59.01



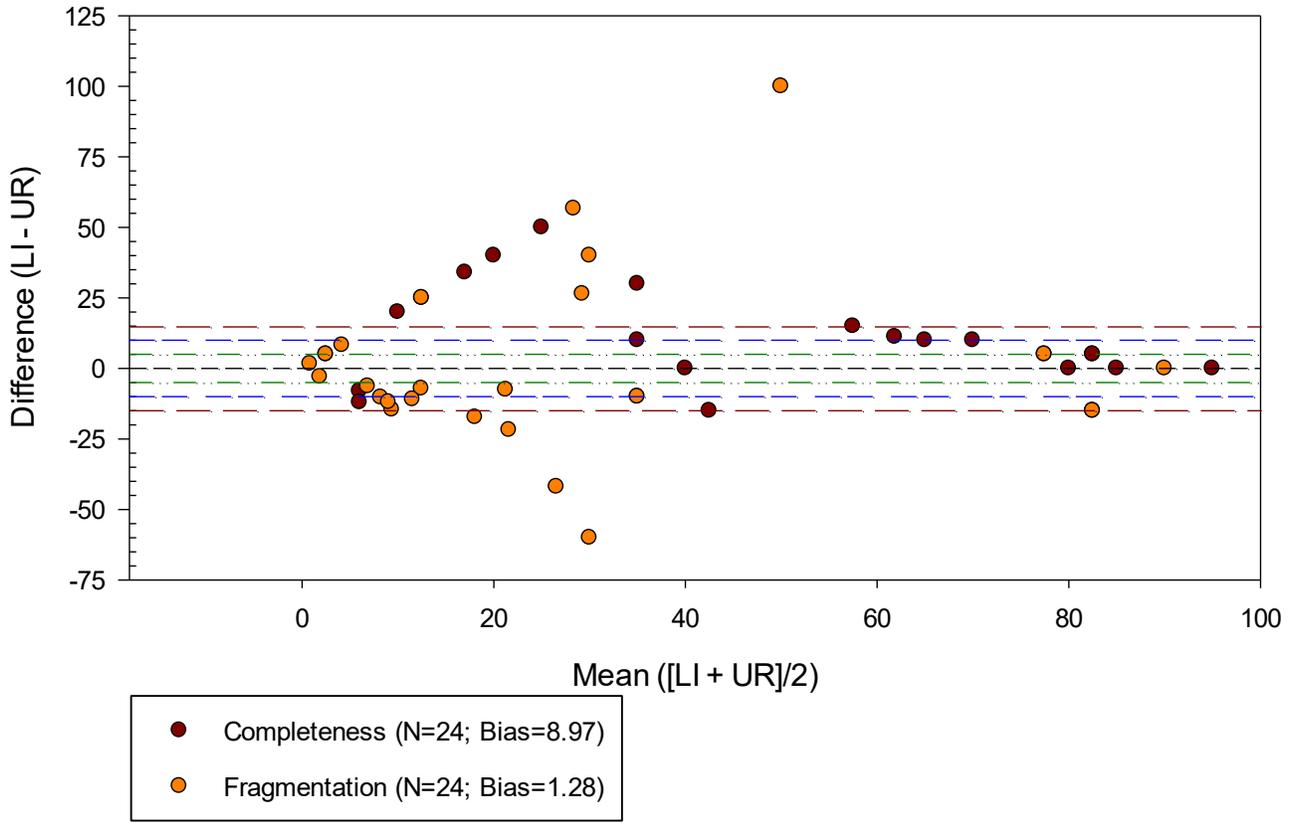
Difference Between Element Scores vs. Mean of Element Scores Burial 59.02



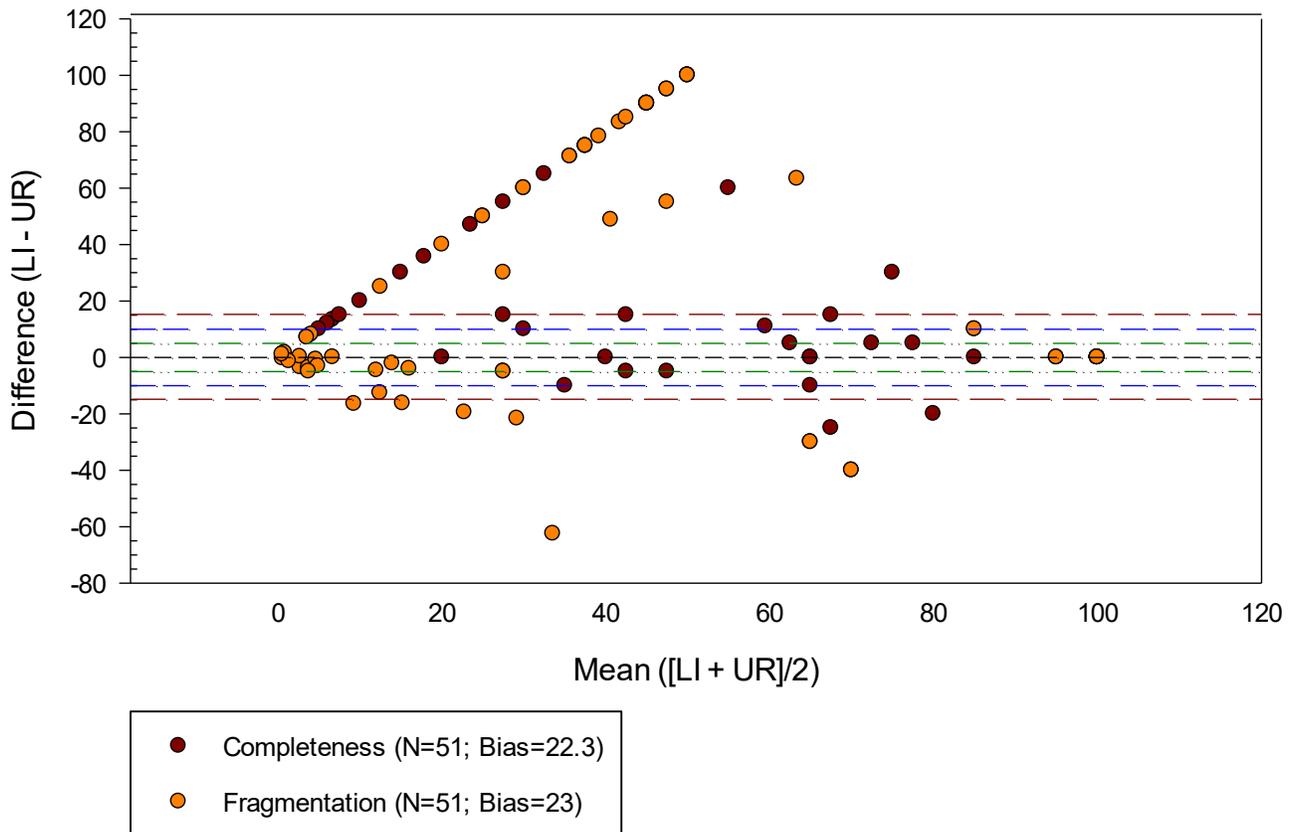
Difference Between Element Scores vs. Mean of Element Scores Burial 60



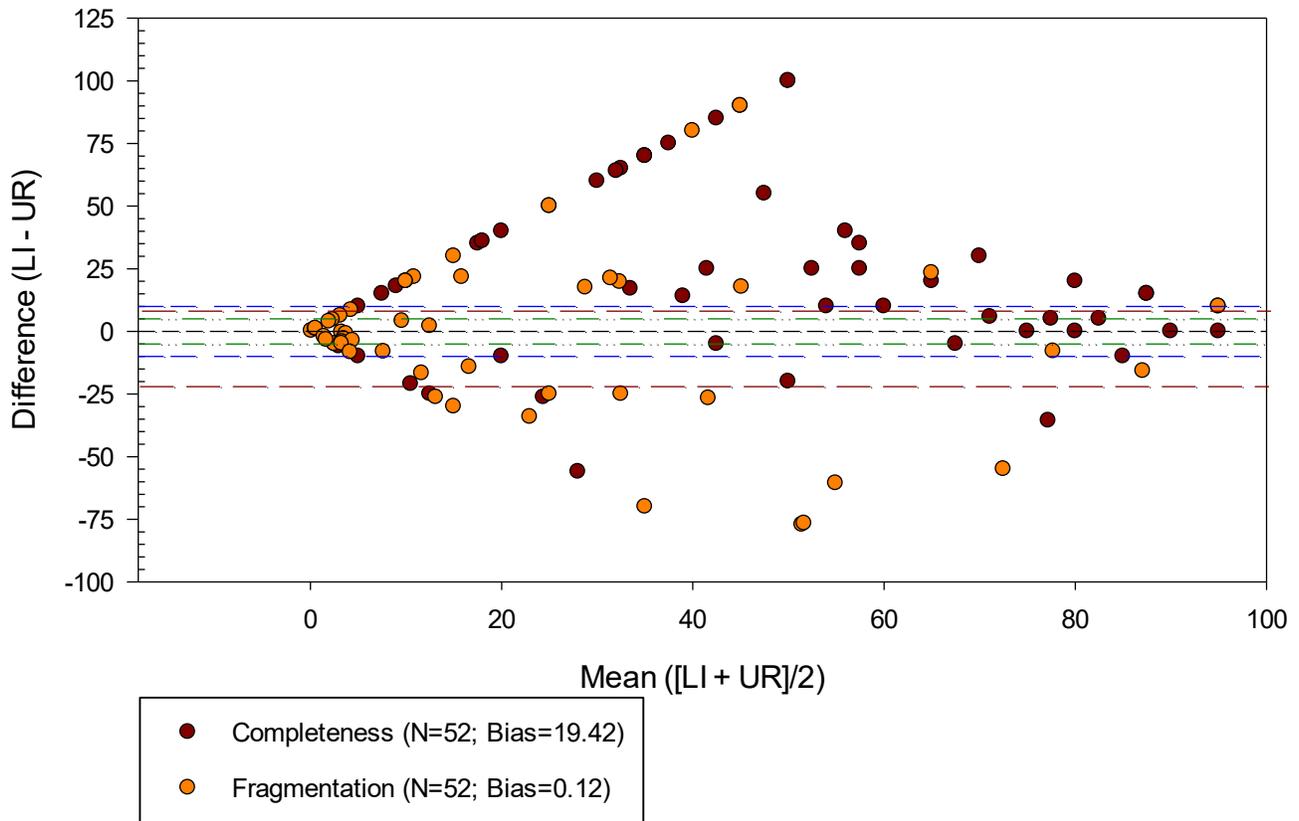
Difference Between Element Scores vs. Mean of Element Scores
Burial 61



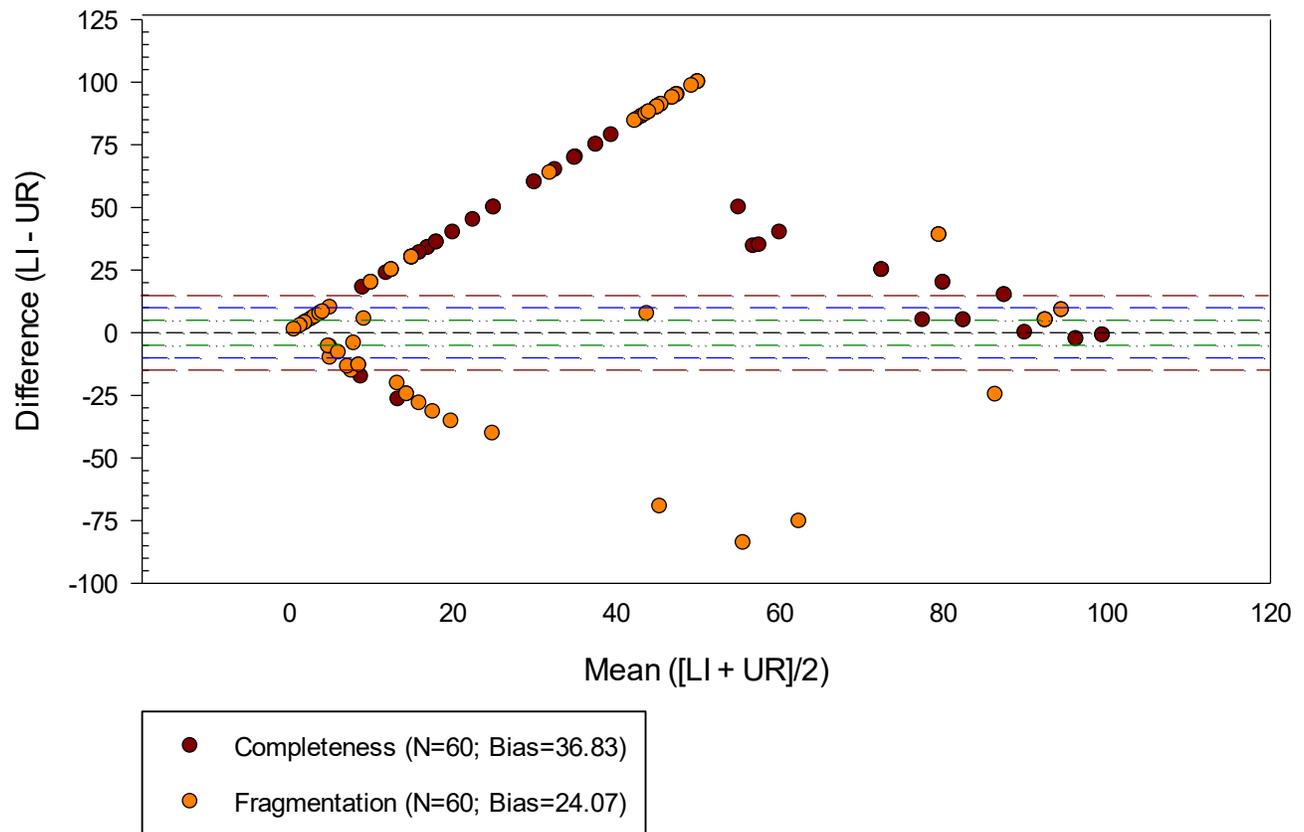
Difference Between Element Scores vs. Mean of Element Scores
Burial 63



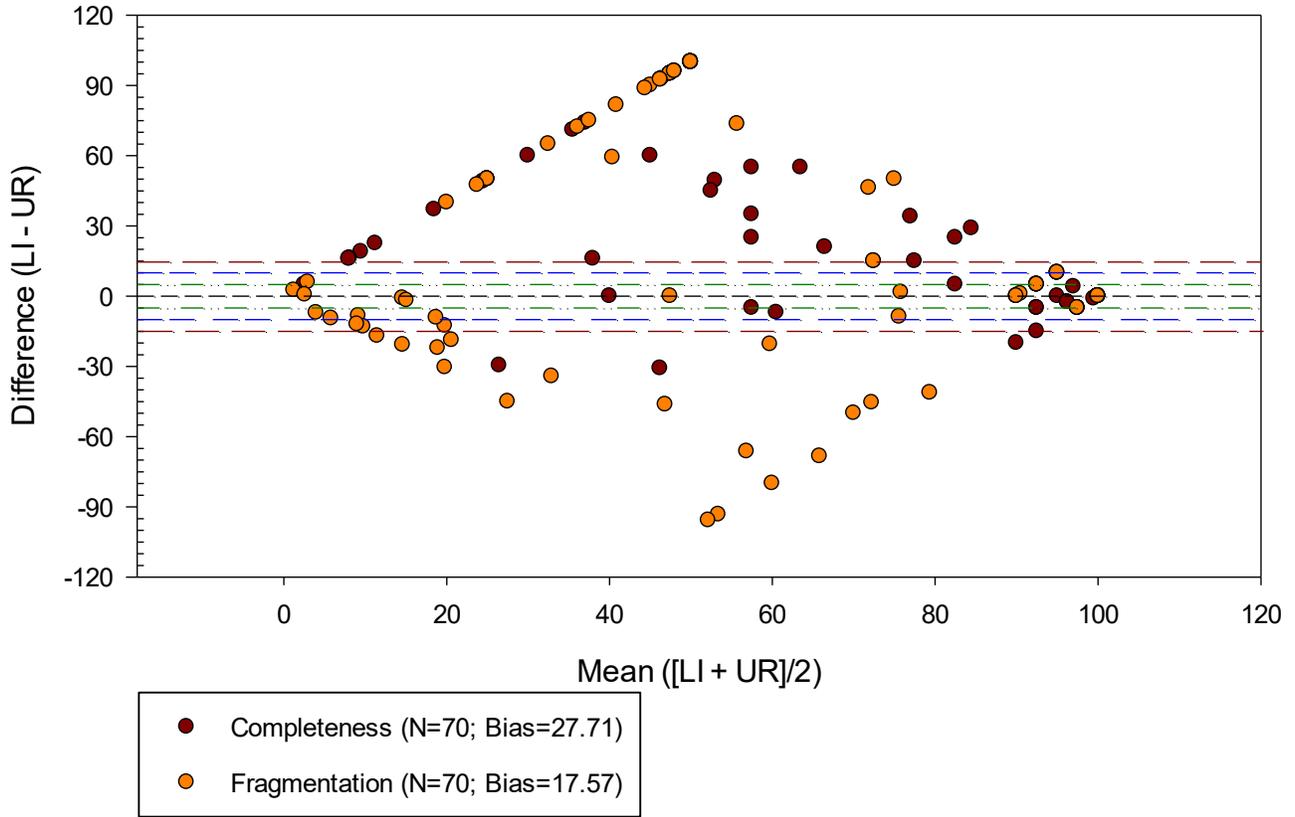
Difference Between Element Scores vs. Mean of Element Scores
Burial 64



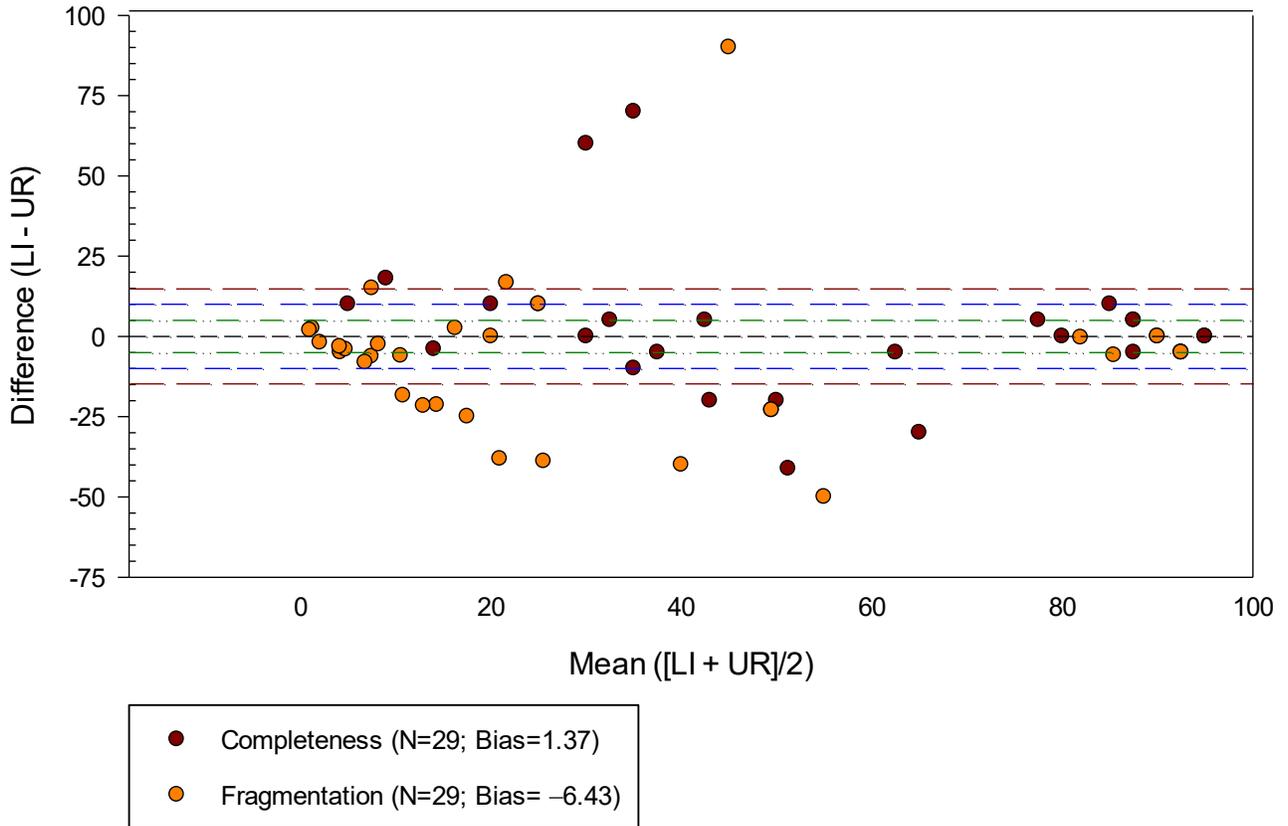
Difference Between Element Scores vs. Mean of Element Scores
Burial 66



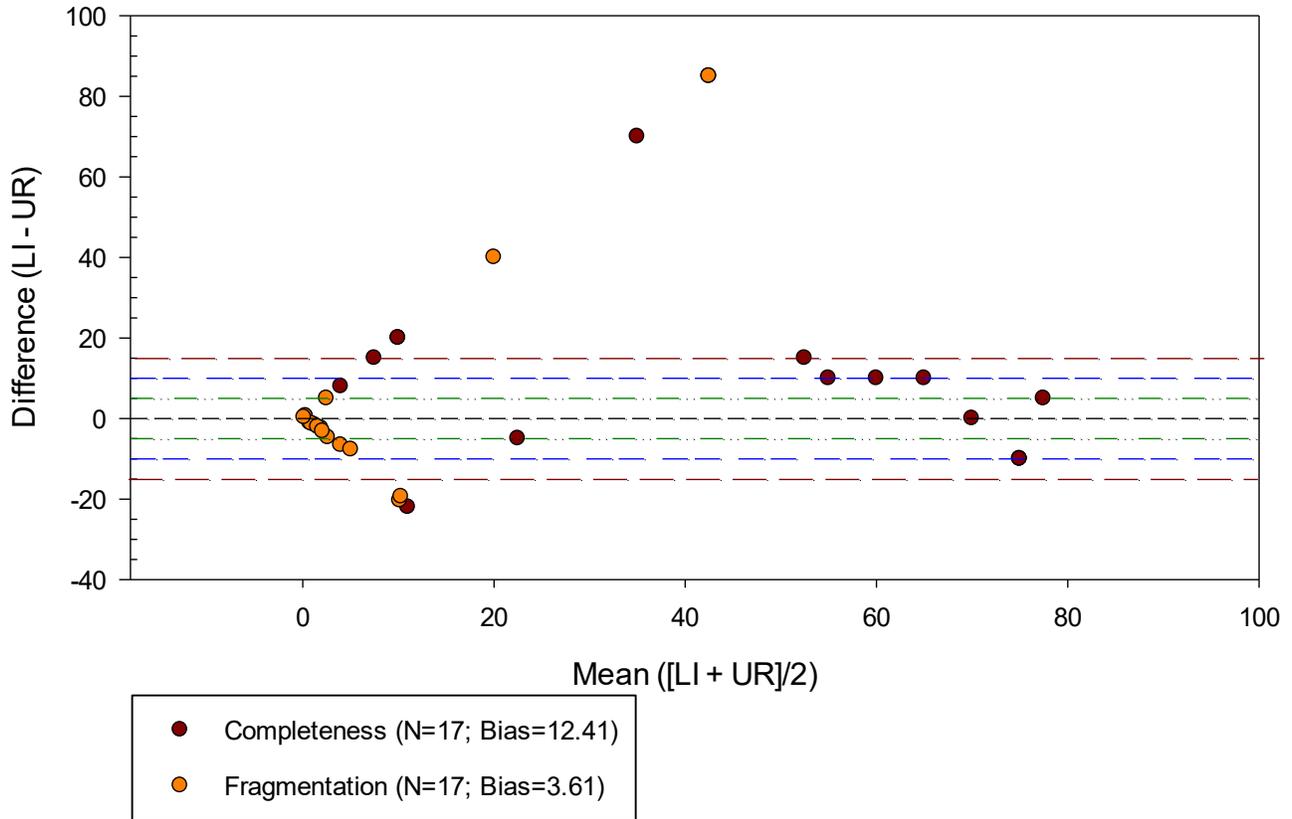
Difference Between Element Scores vs. Mean of Element Scores
Burial 68



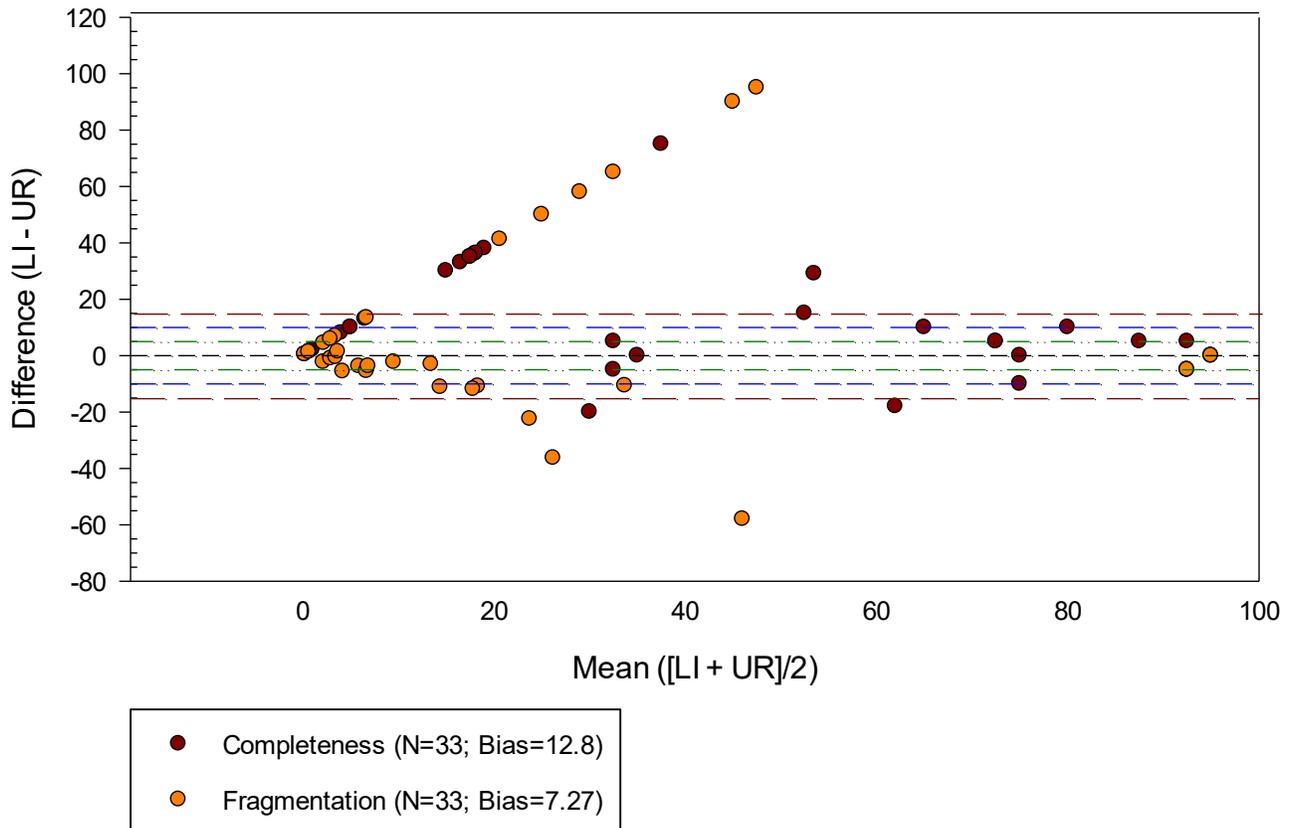
Difference Between Element Scores vs. Mean of Element Scores
Burial 71



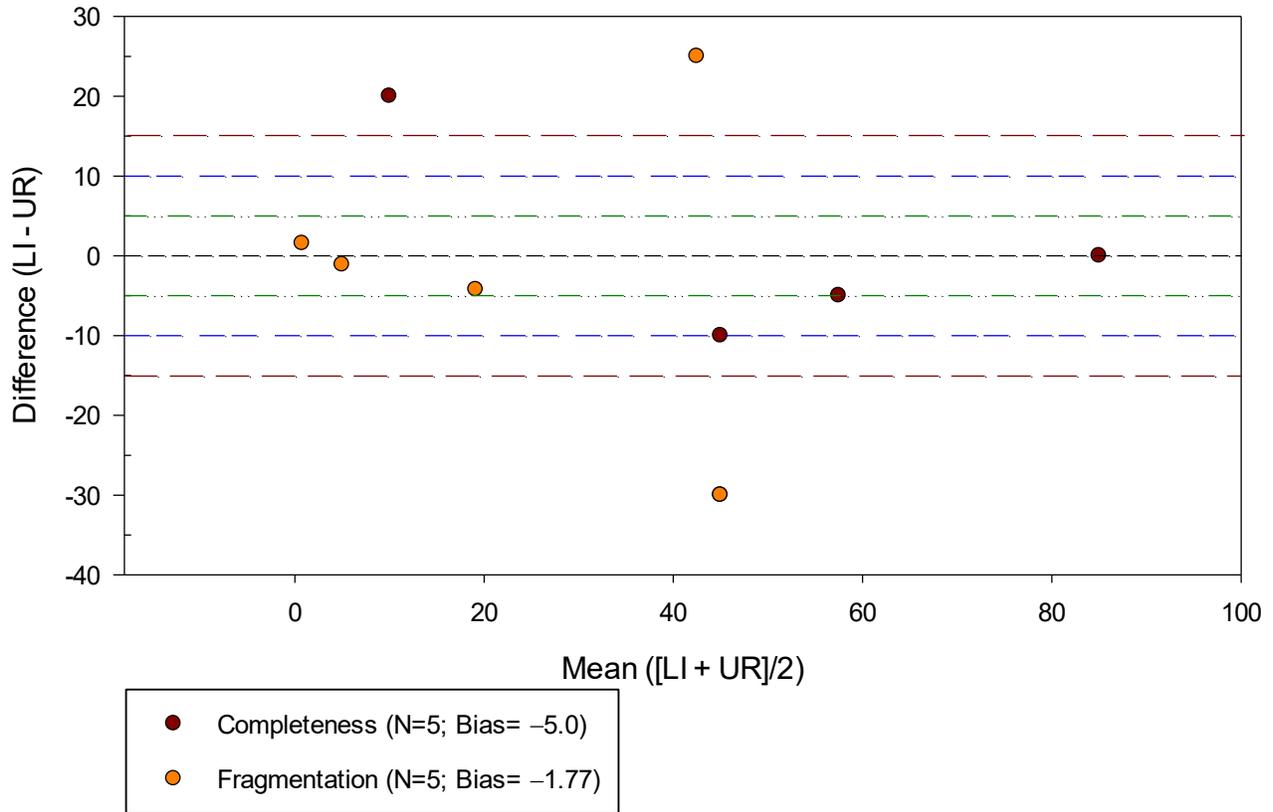
Difference Between Element Scores vs. Mean of Element Scores
Burial 72



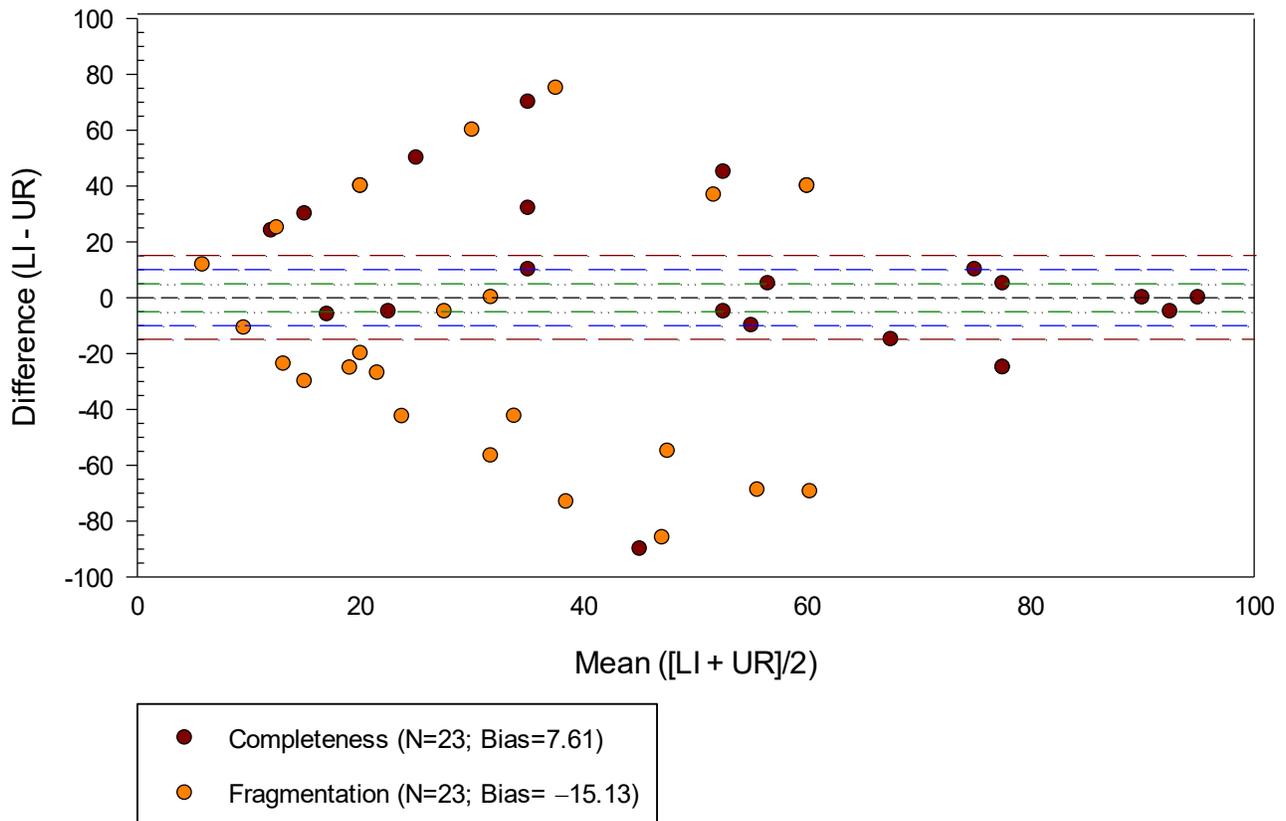
Difference Between Element Scores vs. Mean of Element Scores
Burial 73



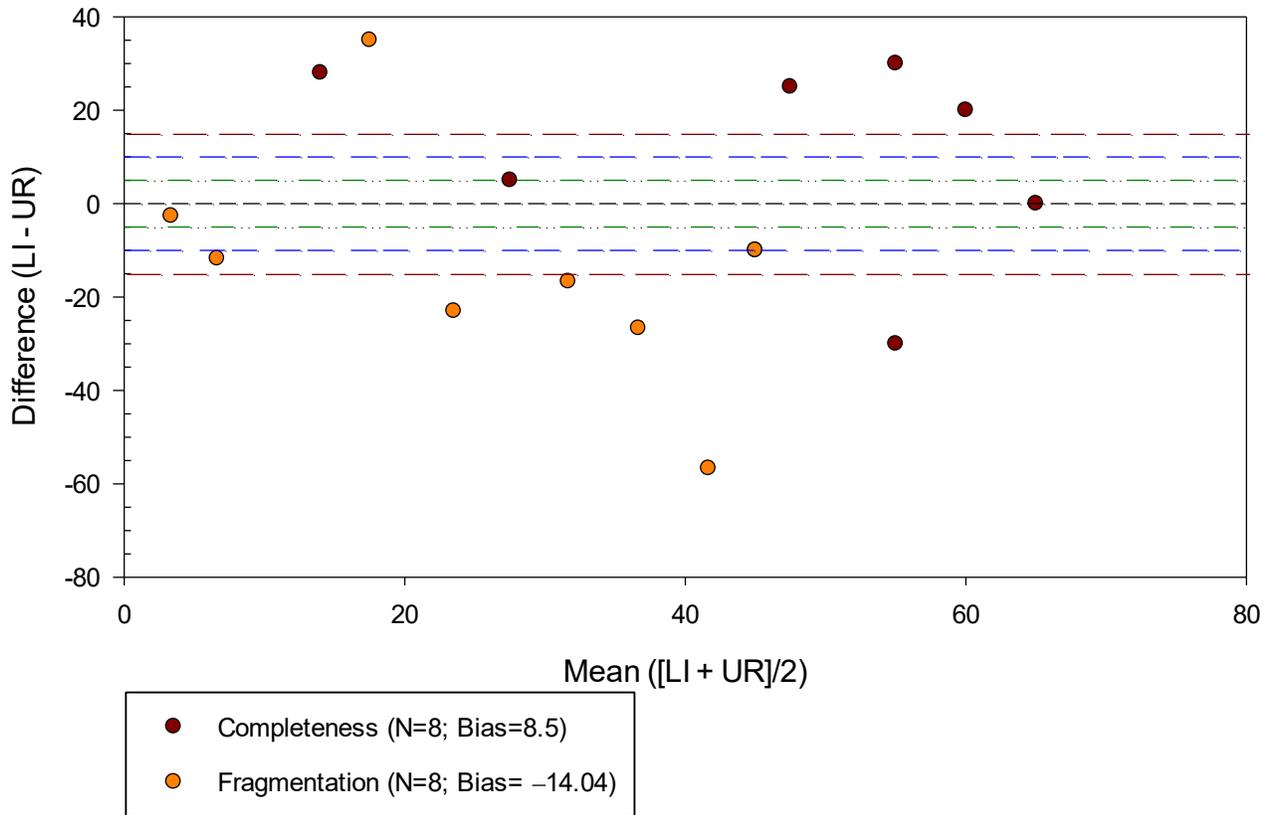
Difference Between Element Scores vs. Mean of Element Scores
Burial 75



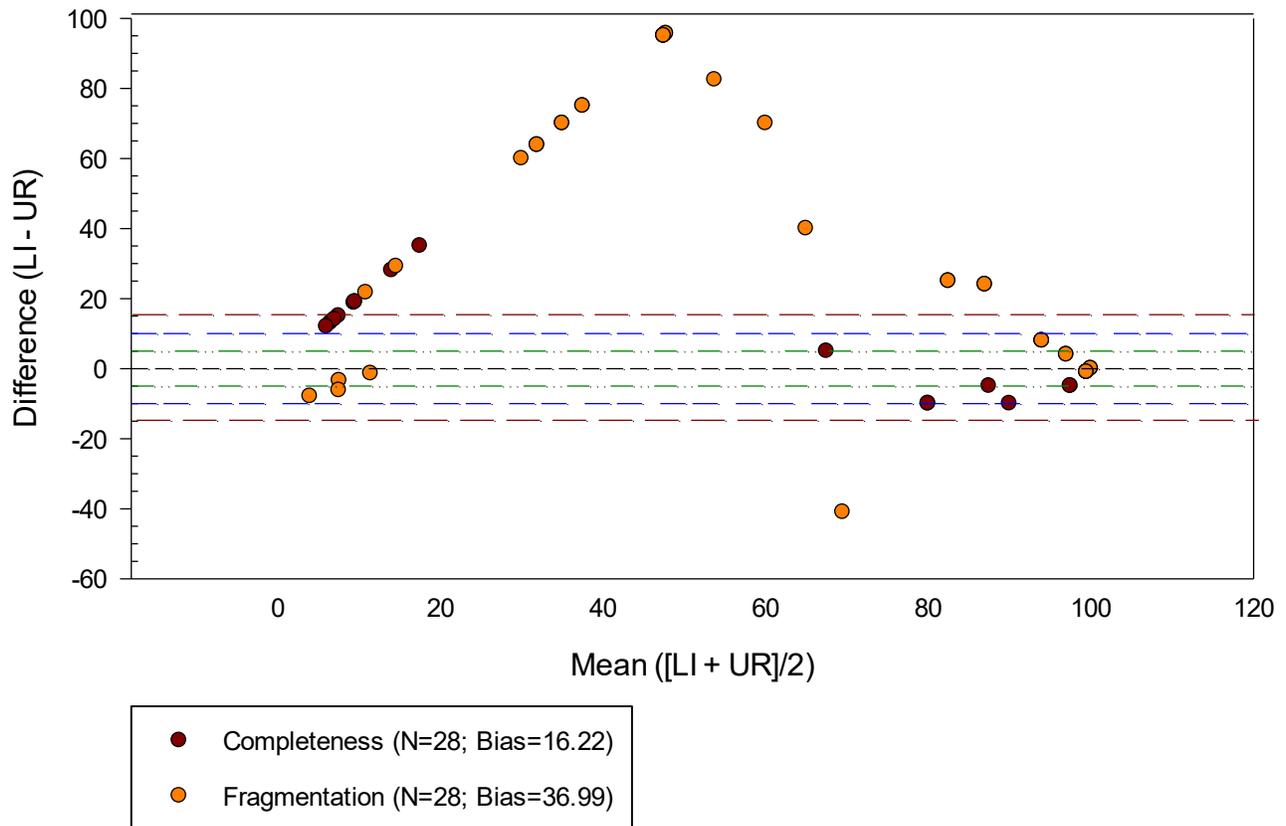
Difference Between Element Scores vs. Mean of Element Scores
Burial 76



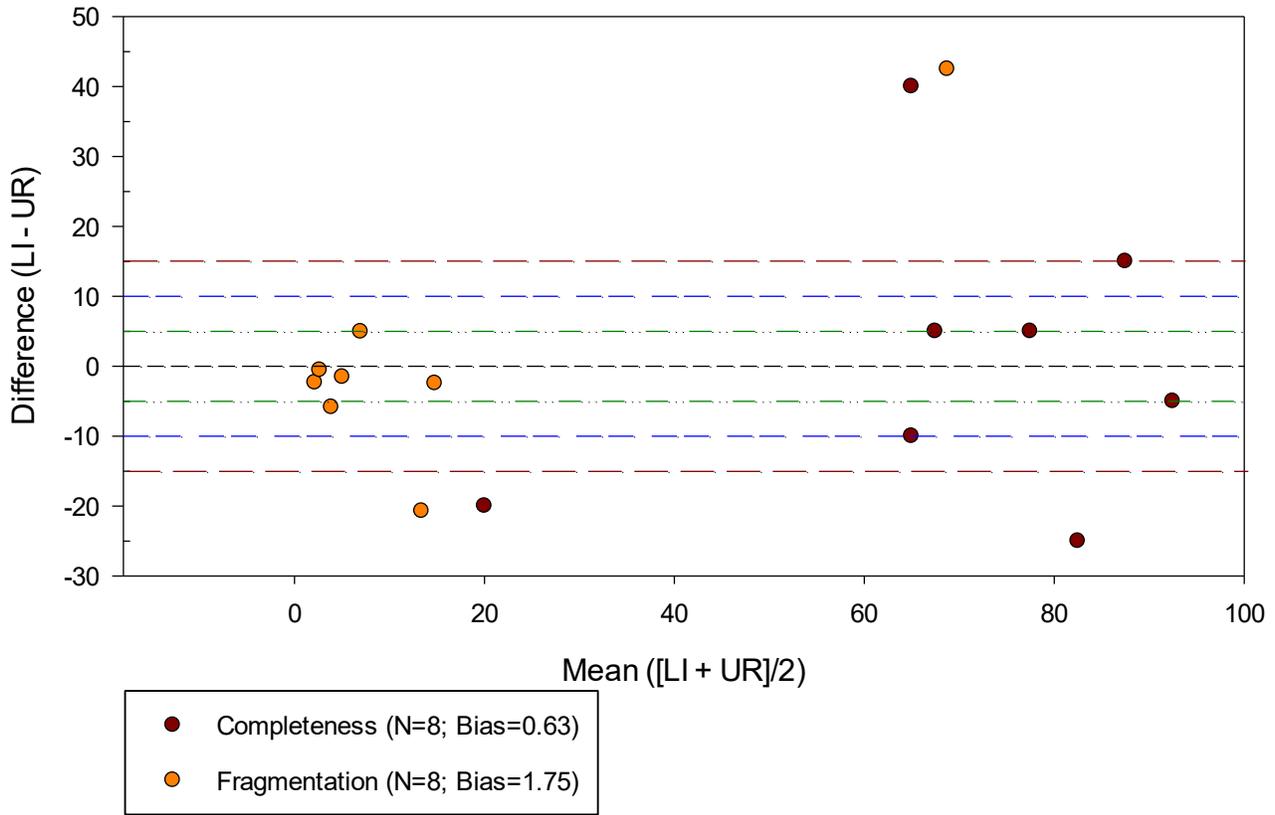
Difference Between Element Scores vs. Mean of Element Scores
Burial 77



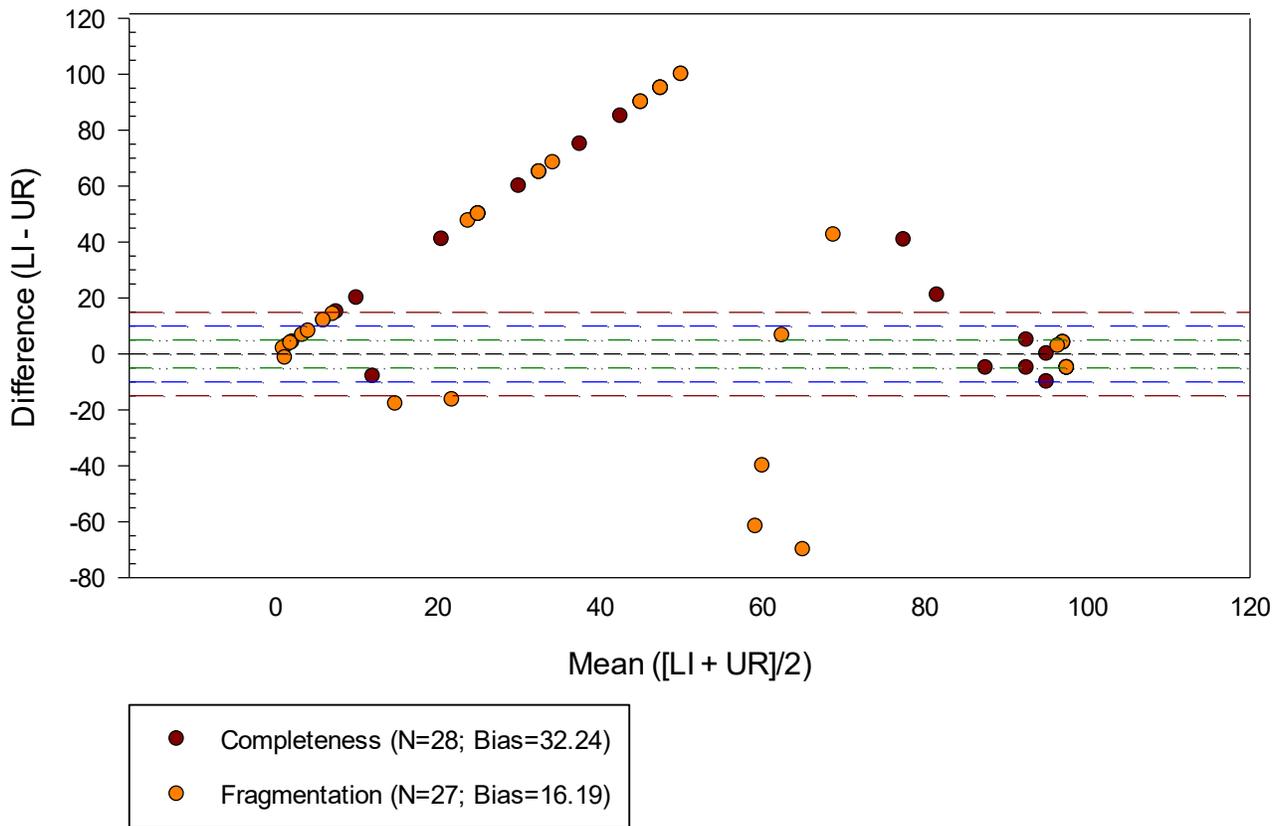
Difference Between Element Scores vs. Mean of Element Scores
Burial 78



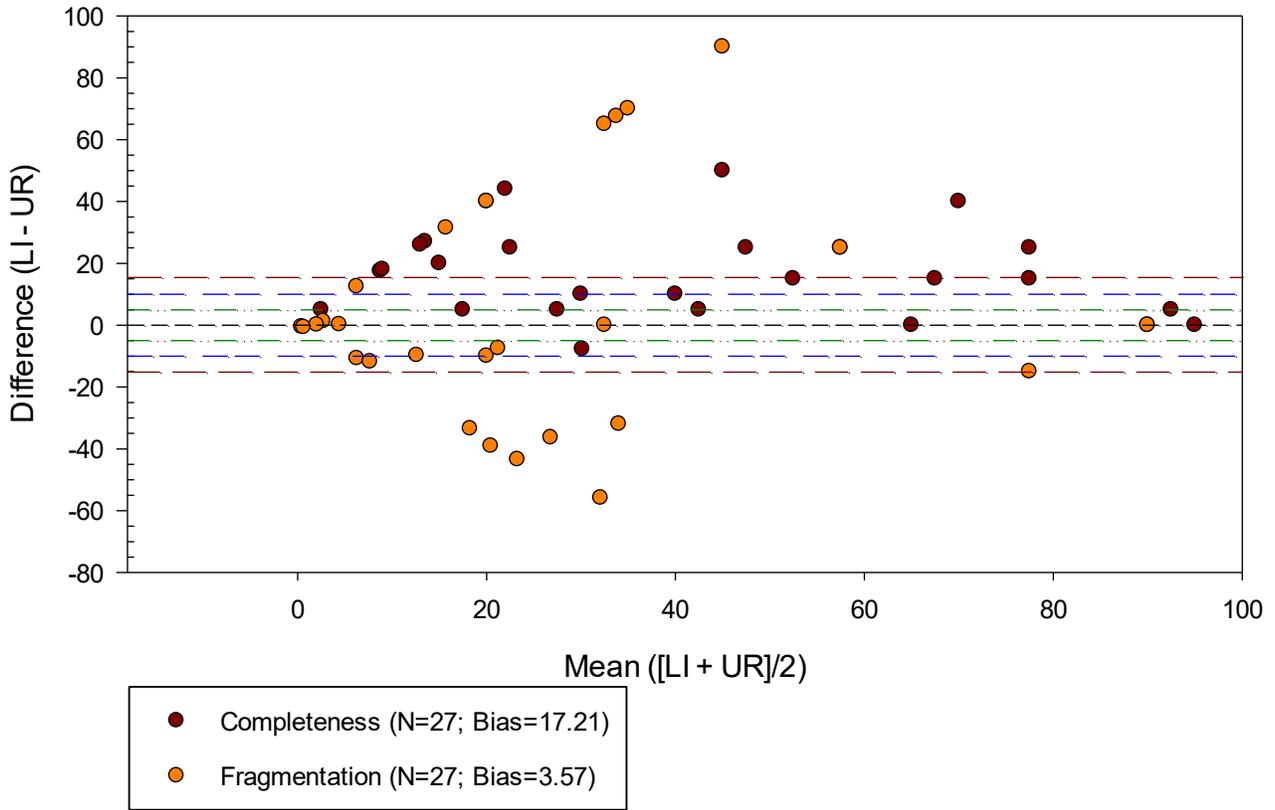
Difference Between Element Scores vs. Mean of Element Scores
Burial 79



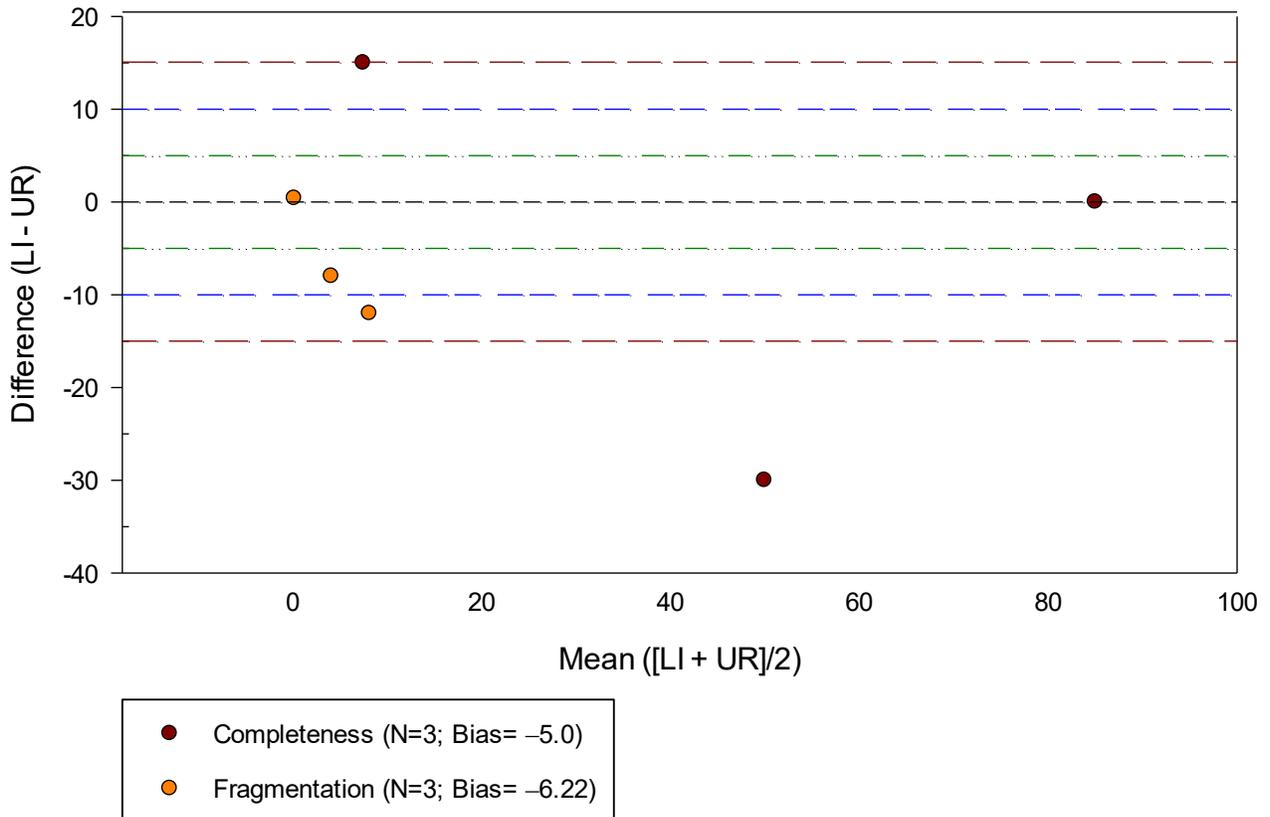
Difference Between Element Scores vs. Mean of Element Scores
Burial 80.02



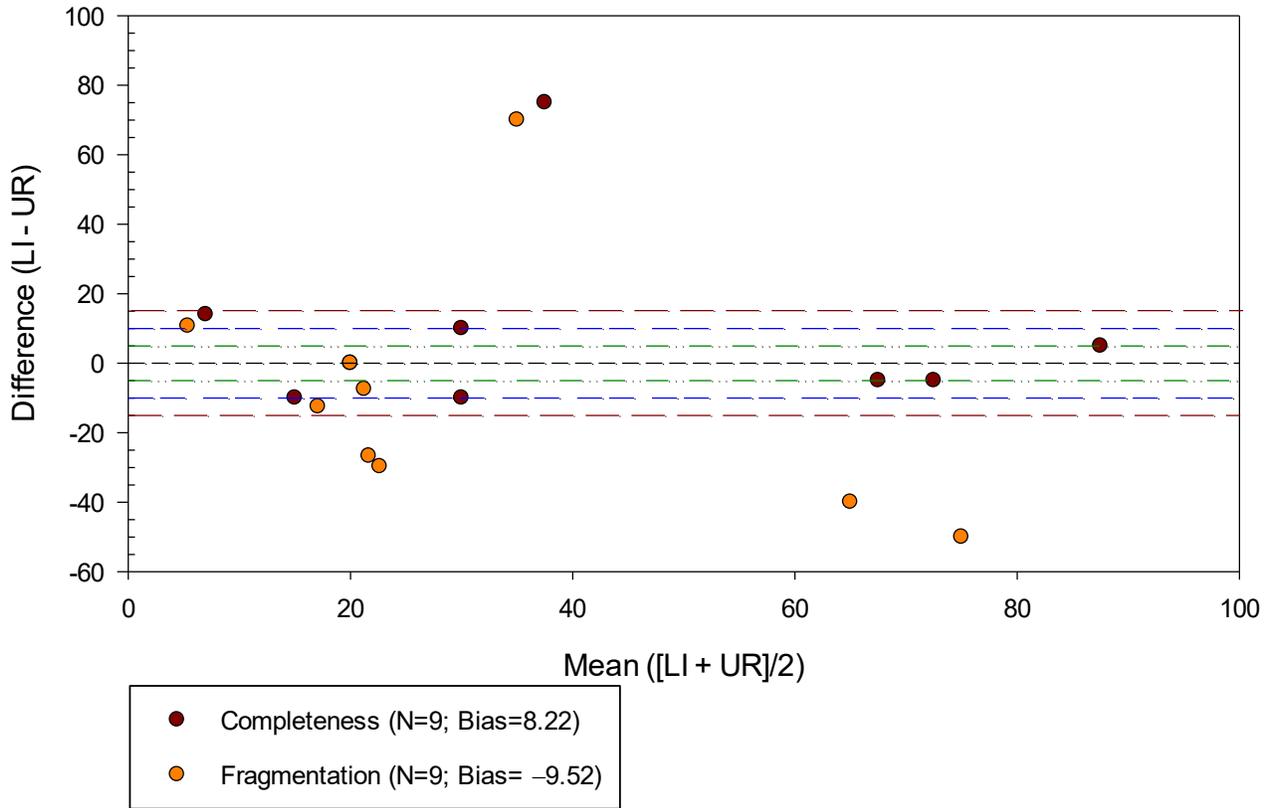
Difference Between Element Scores vs. Mean of Element Scores
Burial 81



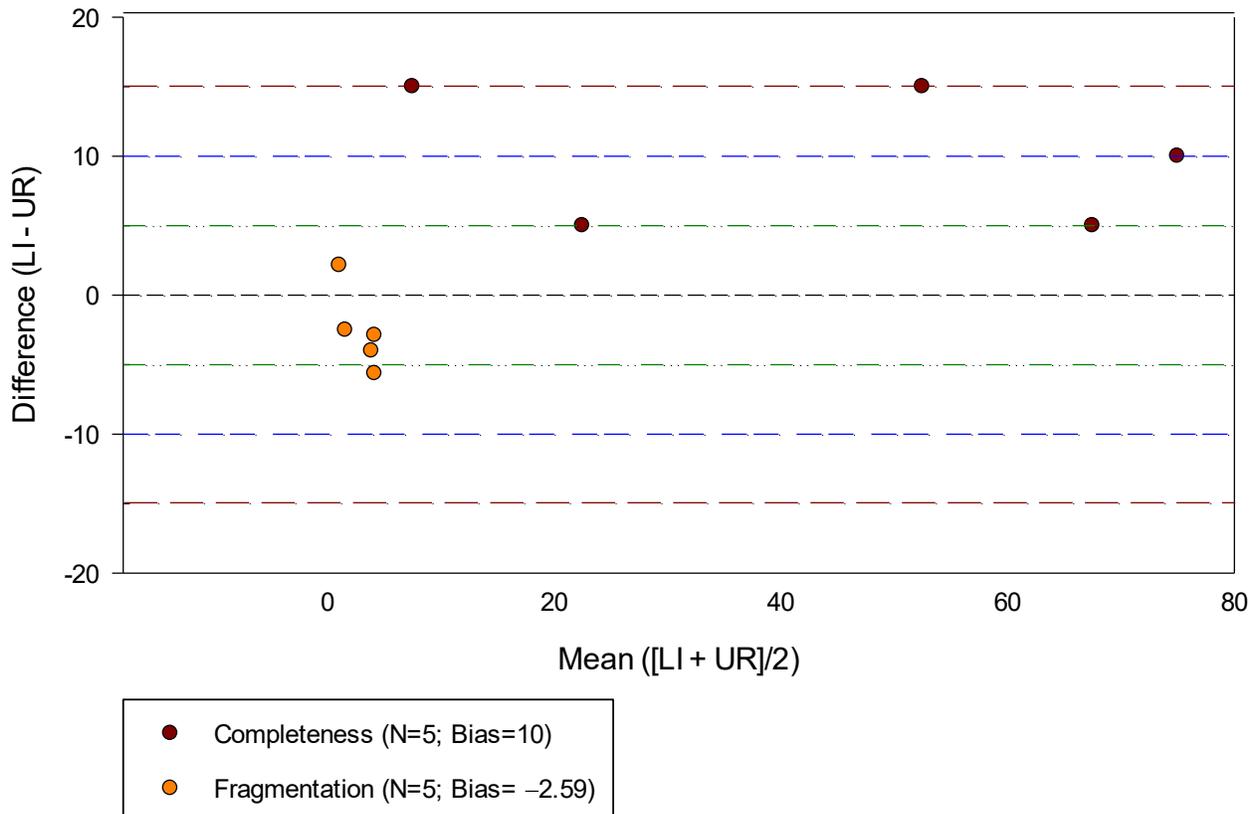
Difference Between Element Scores vs. Mean of Element Scores
Burial 82



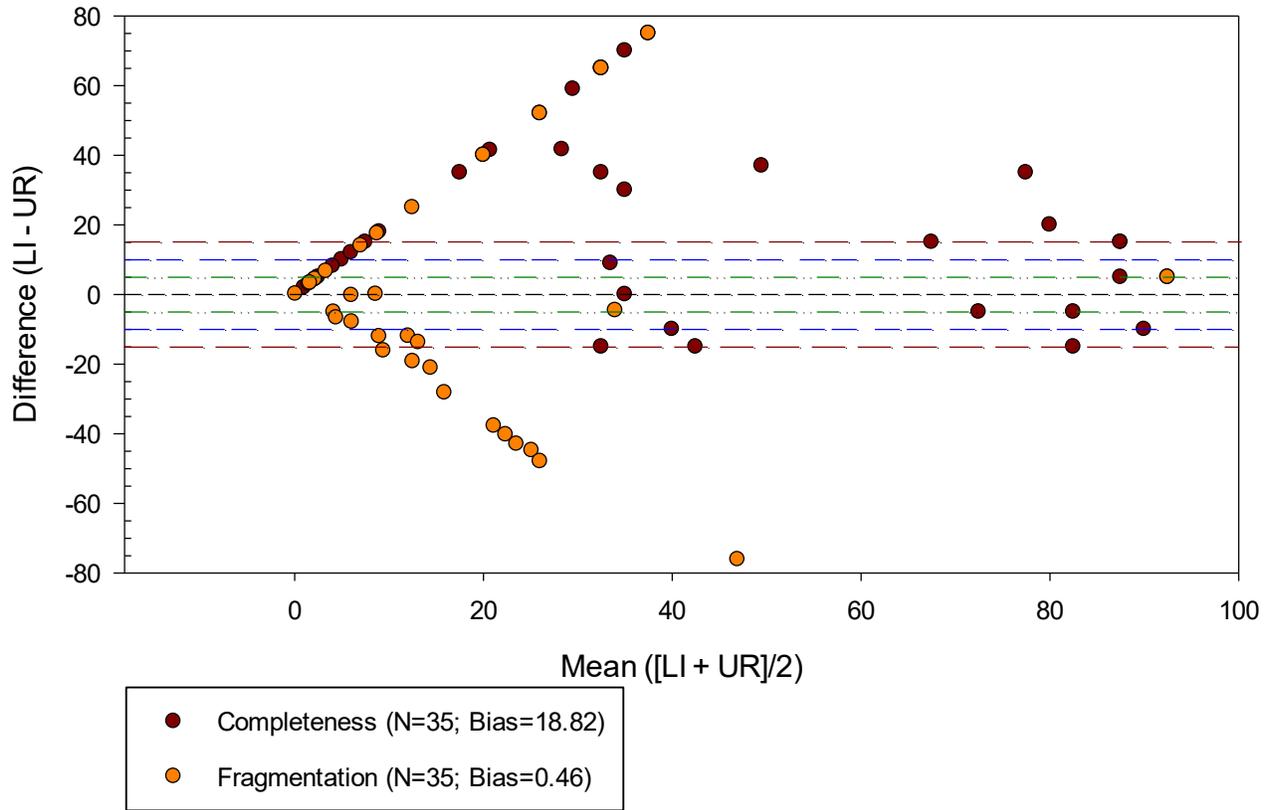
Difference Between Element Scores vs. Mean of Element Scores Burial 84



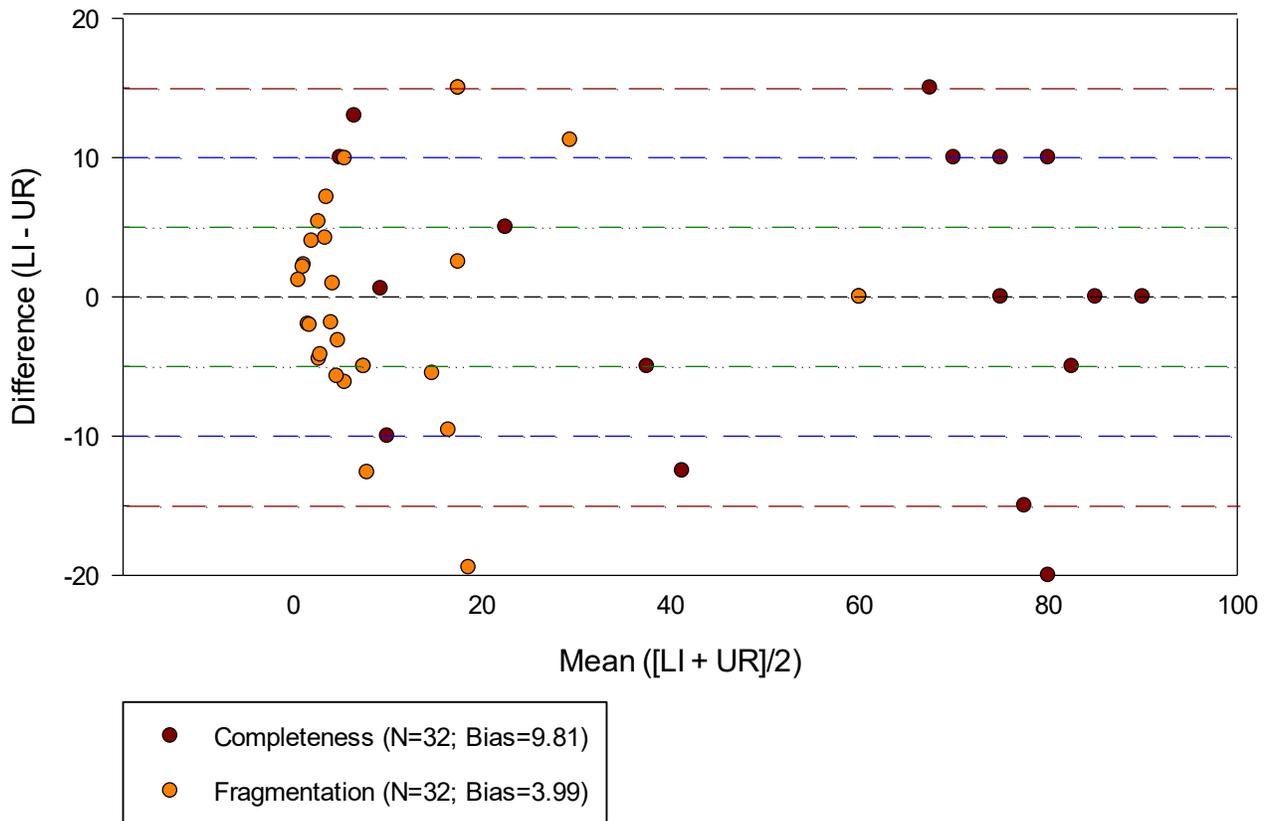
Difference Between Element Scores vs. Mean of Element Scores Burial 85



Difference Between Element Scores vs. Mean of Element Scores
Burial 86

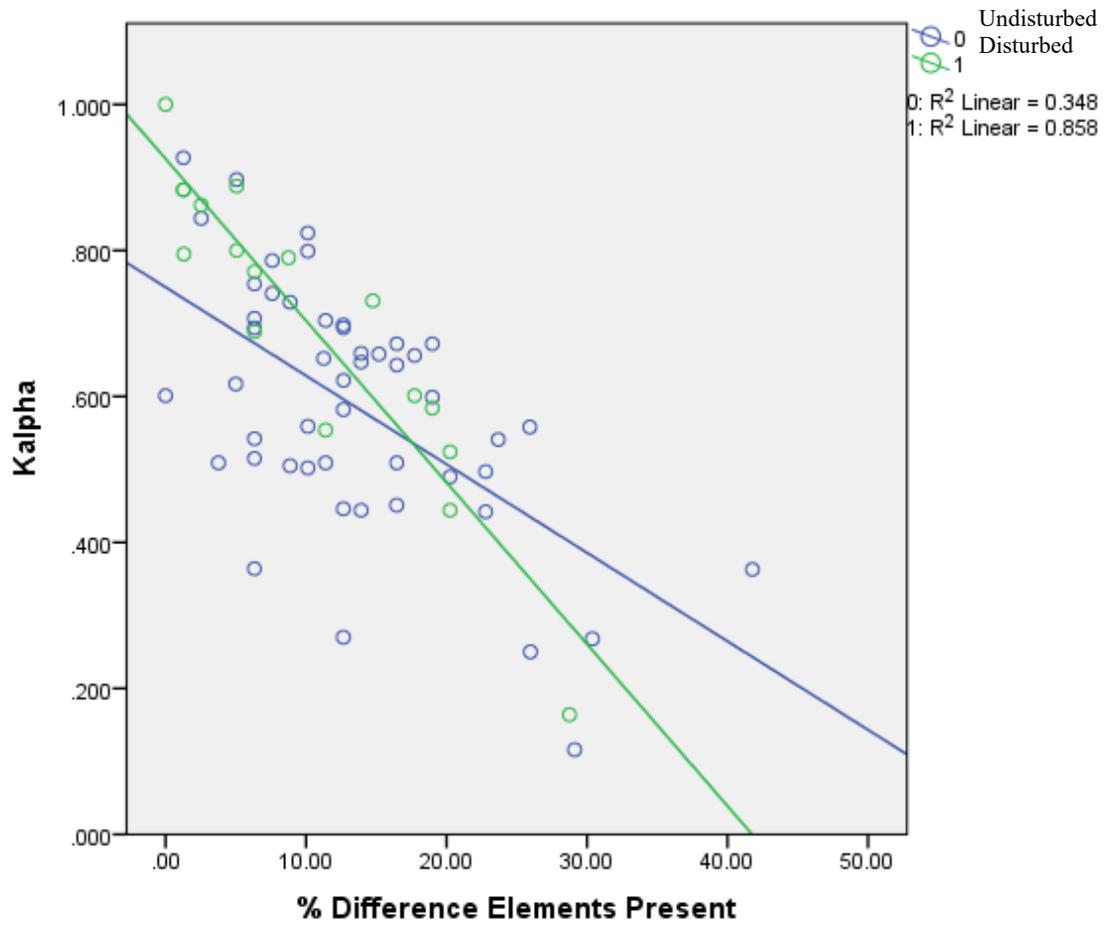


Difference Between Element Scores vs. Mean of Element Scores
Burial 87

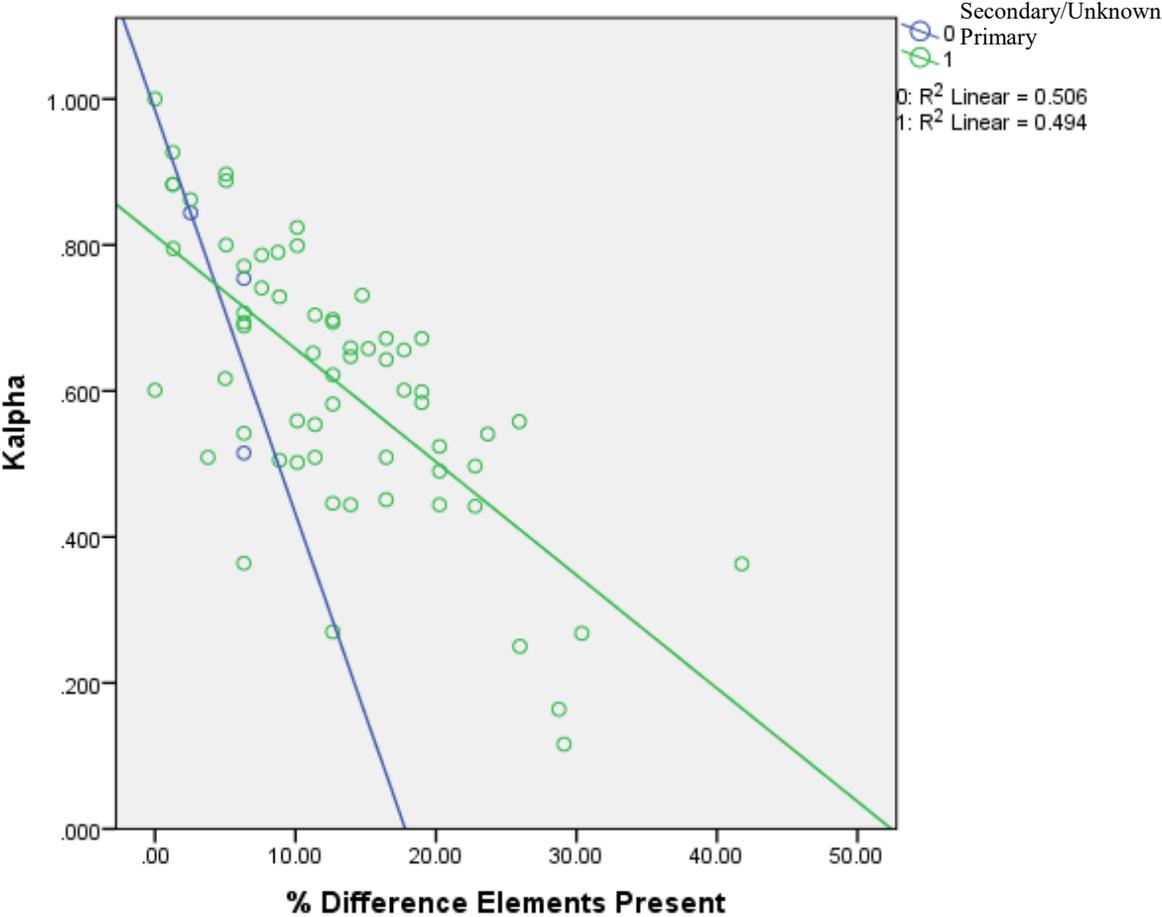


APPENDIX A2: ADDITIONAL SCATTERPLOTS

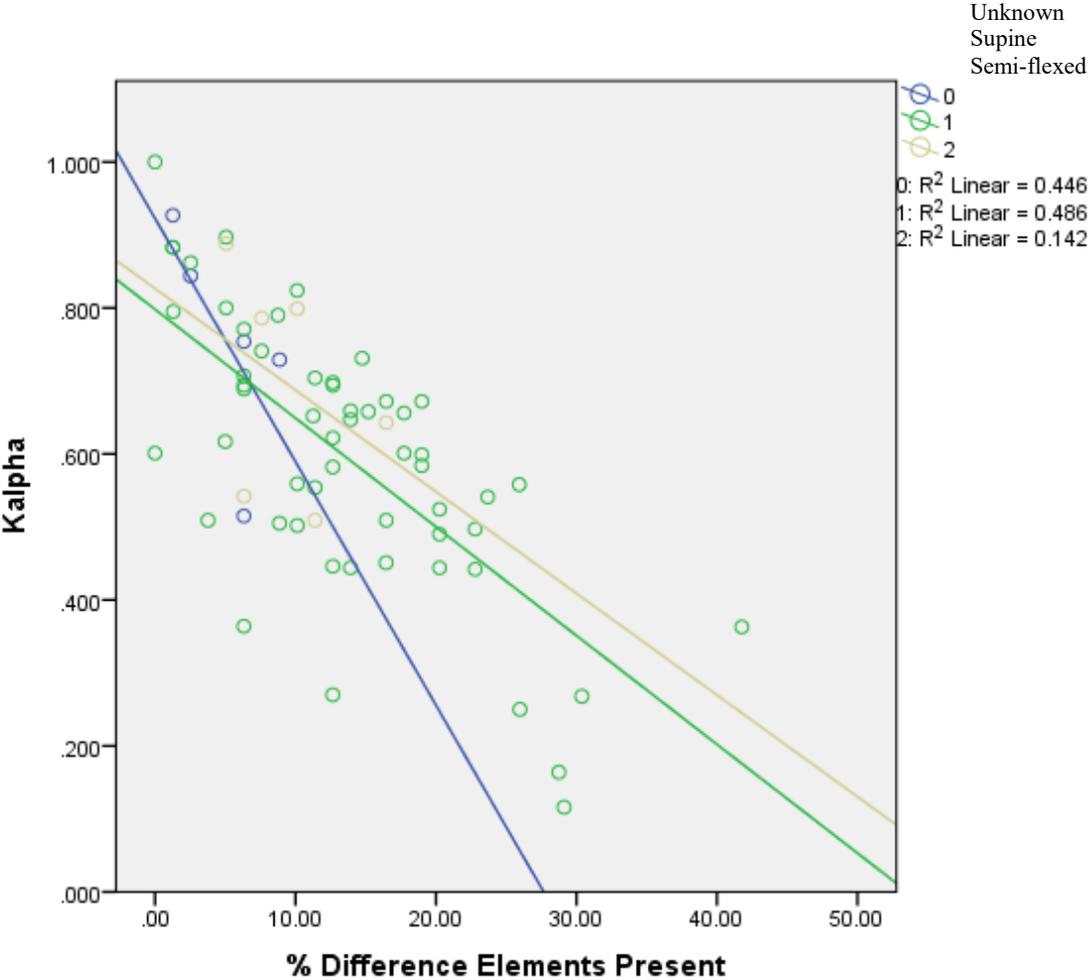
Krippendorf's Alpha vs. Difference RI by Disturbance



Krippendorff's Alpha vs. Difference RI by Integrity



Krippendorff's Alpha vs. Difference RI by Position



APPENDIX C: DATABASES

C1-A: Burial Level Data (part 1 of 2)

Year	Burial	Cmp_UR	Cmp_LI	Avg_Cmp	Diff_Cmp	Cmp_Bias	CmpB_V1	FR_LI	FR_UR	Avg_FR	Diff_FR	FR_Bias	FRB_V1
1997	7.00	48.57	36.55	42.56	-12.02	10.96	-3.08	2.57	12.39	7.48	-9.82	-4.20	-10.45
1997	9.00	67.88	69.32	68.60	1.44	25.52	7.96	34.72	42.32	38.52	-7.60	7.01	-18.69
1997	10.00	50.00	39.77	44.88	-10.23	20.54	-8.00	8.62	4.07	6.35	4.56	7.06	-0.47
1997	11.00	82.71	92.62	87.67	9.90	33.54	13.33	76.99	75.69	76.34	1.30	22.92	-2.67
1997	12.00	37.50	42.45	39.98	4.95	20.17	14.00	10.10	12.33	11.22	-2.23	3.09	3.54
1997	14.00	70.91	53.27	62.09	-17.64	4.26	5.00	13.77	15.80	14.78	-2.03	2.73	6.04
1997	15.00	61.91	71.02	66.46	9.11	21.19	14.65	37.54	33.95	35.75	3.60	10.22	0.51
1997	16.00	61.18	60.10	60.64	-1.08	19.35	-2.81	24.65	16.62	20.63	8.03	13.27	7.52
1997	19.00	65.54	58.49	62.01	-7.05	11.42	-0.80	15.60	22.68	19.14	-7.08	-0.32	-6.84
1997	21.00	48.33	73.13	60.73	24.79	55.00	30.00	39.04	25.83	32.44	13.21	29.35	19.48
1997	22.00	40.00	38.00	39.00	-2.00	22.00	-6.67	15.56	4.98	10.27	10.58	13.57	-4.13
1997	23.00	37.50	14.71	26.10	-22.80	1.47	-13.33	5.66	5.27	5.47	0.39	3.80	6.87
1997	24.00	57.67	40.66	49.16	-17.01	-0.53	-9.00	11.61	11.64	11.63	-0.03	3.30	-0.29
1998	25.00	43.86	35.59	39.73	-8.28	7.21	-5.68	4.82	4.66	4.74	0.16	1.80	0.16
1998	27.01	70.94	63.97	67.45	-6.97	9.92	-2.76	27.65	33.39	30.52	-5.74	3.00	-7.21
1998	27.02	72.65	60.34	66.49	-12.31	10.94	-2.24	37.67	52.03	44.85	-14.36	2.29	-3.50
1998	27.03	55.83	35.00	45.42	-20.83	-23.67	-7.50	17.62	20.07	18.84	-2.45	-5.49	1.82
1998	28.00	51.54	48.72	50.13	-2.82	10.89	-1.50	11.31	23.55	17.43	-12.25	-5.41	-10.86
1998	29.00	67.21	69.46	68.33	2.25	24.66	20.48	49.41	51.14	50.27	-1.73	15.32	3.21
1998	32.00	48.94	56.12	52.53	7.18	17.13	11.21	21.02	23.99	22.50	-2.97	2.84	1.68
1998	33.00	65.00	45.00	55.00	-20.00	26.43	0.00	6.48	1.28	3.88	5.20	6.11	1.42
1998	34.00	53.84	60.63	57.23	6.79	23.29	13.11	29.37	18.84	24.10	10.54	16.12	10.66
1998	35.01	55.00	48.92	51.96	-6.09	8.12	-0.45	9.59	12.76	11.18	-3.17	0.58	0.32
1998	35.02	41.82	38.89	40.35	-2.93	12.63	-1.00	1.59	1.68	1.64	-0.09	0.53	-0.01
1998	36.02	48.75	33.33	41.04	-15.42	-12.33	-6.67	6.47	6.72	6.59	-0.25	-0.19	-2.79

1998	37.01	35.48	46.17	40.82	10.68	14.41	12.07	10.25	5.38	7.82	4.87	5.19	5.54
1998	37.02	53.58	50.87	52.22	-2.71	9.62	-0.48	19.32	16.93	18.12	2.39	6.22	-0.59
1998	38.00	47.66	55.45	51.55	7.79	22.83	21.68	26.04	12.40	19.22	13.64	17.31	9.63
1998	39.00	58.00	45.67	51.83	-12.33	7.00	-1.67	6.79	23.14	14.97	-16.35	-11.72	-20.25
1999	44.00	69.77	72.70	71.23	2.93	14.29	1.05	38.00	46.74	42.37	-8.74	-2.31	-14.04
1999	45.00	52.07	53.23	52.65	1.16	28.69	9.33	12.86	13.88	13.37	-1.02	5.77	-8.22
1999	46.00	80.98	84.97	82.97	3.99	19.03	8.09	69.80	71.11	70.46	-1.31	11.89	0.77
1999	47.00	65.79	67.26	66.52	1.47	17.81	10.42	34.89	33.20	34.04	1.69	9.87	5.11
1999	48.00	55.00	37.59	46.29	-17.41	20.82	0.00	13.84	10.75	12.30	3.09	10.18	6.33
1999	49.00	38.20	37.59	37.89	-0.61	13.71	3.45	13.84	10.68	12.26	3.17	7.17	0.05
1999	50.00	55.36	44.21	49.78	-11.15	17.49	4.38	8.39	18.96	13.68	-10.57	-0.77	-12.20
1999	51.00	60.90	62.43	61.67	1.53	12.86	4.64	27.20	31.97	29.58	-4.77	1.18	-8.85
1999	53.00	57.99	63.55	60.77	5.56	9.42	5.81	33.97	31.41	32.69	2.56	4.68	0.65
1999	55.00	59.33	53.85	56.59	-5.48	9.12	-0.35	17.20	25.03	21.11	-7.83	-1.53	-7.93
1999	57.01	66.56	44.96	55.76	-21.60	11.68	-2.47	28.25	51.18	39.72	-22.93	2.66	-19.92
1999	57.02	58.33	46.37	52.35	-11.96	34.31	11.67	18.47	29.30	23.88	-10.83	12.41	-19.63
1999	58.01	58.92	57.89	58.41	-1.03	20.12	3.00	24.04	32.06	28.05	-8.02	3.49	-8.83
1999	58.02	42.78	48.17	45.47	5.39	10.28	6.93	13.94	17.43	15.69	-3.49	-1.51	-2.19
1999	59.01	43.33	33.75	38.54	-9.58	1.25	-6.67	10.71	19.94	15.32	-9.23	-4.25	-14.00
1999	59.02	76.10	70.39	73.24	-5.71	21.80	-1.91	35.67	45.37	40.52	-9.70	6.83	-9.79
1999	60.00	55.64	60.79	58.22	5.15	20.22	4.36	23.57	25.49	24.53	-1.92	4.98	-1.95
2000	61.00	56.57	53.63	55.10	-2.94	8.97	3.13	27.60	33.56	30.58	-5.96	1.28	-6.23
2000	63.00	63.47	59.63	61.55	-3.83	22.30	3.53	43.38	34.64	39.01	8.74	23.00	2.67
2000	64.00	53.05	63.47	58.26	10.42	19.42	9.00	26.72	33.97	30.34	-7.25	0.12	-10.94
2000	66.00	70.14	68.33	69.24	-1.81	36.83	38.26	41.96	39.58	40.77	2.38	24.07	9.80
2000	68.00	76.03	77.67	76.85	1.64	27.71	11.10	56.73	59.60	58.16	-2.87	17.57	-9.12
2000	71.00	63.47	56.09	59.78	-7.38	1.37	-4.73	23.75	35.01	29.38	-11.26	-6.43	-11.83
2000	72.00	58.36	53.31	55.84	-5.05	12.41	1.50	8.56	6.88	7.72	1.68	3.61	-4.99
2000	73.00	67.37	51.58	59.47	-15.79	12.80	1.37	23.12	27.53	25.33	-4.41	7.27	-10.18
2000	75.00	63.75	46.00	54.88	-17.75	-5.00	-11.25	21.60	29.20	25.40	-7.60	-1.77	-2.59

2000	76.00	61.28	58.09	59.68	-3.19	7.61	3.00	24.88	49.26	37.07	-24.38	-15.13	-31.15
2000	77.00	47.86	50.38	49.12	2.52	8.50	5.71	18.72	37.45	28.08	-18.73	-14.04	-21.05
2000	78.00	90.00	67.65	78.82	-22.35	16.22	14.76	73.27	58.90	66.08	14.37	36.99	23.41
2000	79.00	69.38	70.00	69.69	0.63	0.63	0.63	15.59	13.84	14.72	1.75	1.75	1.75
2000	80.02	83.67	70.10	76.88	-13.57	34.24	3.07	47.24	68.53	57.88	-21.29	16.79	-13.47
2001	81.00	45.95	51.25	48.60	5.30	17.21	14.36	26.92	31.52	29.22	-4.60	3.57	-14.00
2001	83.00	75.00	45.00	60.00	-30.00	-5.00	-15.00	0.88	11.15	6.01	-10.27	-6.55	-10.02
2001	84.00	47.14	44.89	46.02	-2.25	8.22	-2.14	26.68	46.55	36.61	-19.87	-9.52	-23.77
2001	85.00	50.00	50.00	50.00	0.00	10.00	8.75	1.69	5.35	3.52	-3.66	-2.59	-3.77
2001	86.00	57.13	51.46	54.29	-5.67	18.82	9.38	17.35	31.15	24.25	-13.80	-0.46	-21.57
2001	87.00	56.80	50.63	53.72	-6.17	9.81	-1.39	15.86	16.51	16.18	-0.65	3.99	-2.89

C1-B: Burial Level Data (part 2 of 2)

Year	Burial	RI_UR	RI_LI	Avg_RI	Diff_RI	AI	Pos	No_inter	Disturb	Integ	Angle	No_image
1997	7.00	17.72	34.18	25.95	16.46	100.00	1	1	0	1	3.00	5
1997	9.00	34.17	54.43	44.30	20.26	100.00	1	1	0	1	3.00	4
1997	10.00	6.33	16.46	11.39	10.13	100.00	1	1	0	1	3.00	3
1997	11.00	64.94	90.91	77.92	25.97	70.27	1	1	0	1	9.00	6
1997	12.00	7.59	13.92	10.76	6.33	100.00	2	1	0	1	1.00	2
1997	14.00	27.85	40.51	34.18	26.58	100.00	1	1	0	1	9.00	7
1997	15.00	41.78	51.90	46.84	8.86	87.80	1	1	0	1	9.00	6
1997	16.00	21.25	32.50	26.88	11.25	72.00	1	1	0	1	3.00	4
1997	19.00	32.91	46.84	39.87	13.92	100.00	1	1	0	1	7.00	6
1997	21.00	3.80	10.13	6.96	6.33	12.50	0	1	0	0	4.00	4
1997	22.00	7.59	18.99	13.29	11.39	100.00	2	1	0	1	1.00	1
1997	23.00	7.59	21.52	14.56	13.92	100.00	1	1	0	1	1.00	3
1997	24.00	18.99	26.58	22.78	7.59	100.00	2	1	0	1	3.00	3
1998	25.00	13.92	21.52	17.72	7.59	100.00	1	1	0	1	10.00	14
1998	27.01	40.51	53.16	46.84	7.59	100.00	1	3	0	1	5.00	7
1998	27.02	21.52	31.65	26.58	1.27	100.00	1	3	0	1	3.00	7
1998	27.03	15.16	11.39	13.28	3.85	100.00	1	3	0	1	3.00	8
1998	28.00	16.45	22.78	19.62	5.13	0.00	0	1	0	0	6.00	8
1998	29.00	47.37	71.05	59.21	3.95	83.02	1	1	0	1	9.00	7
1998	32.00	43.03	55.70	49.36	6.33	100.00	1	1	0	1	11.00	7
1998	33.00	2.53	8.86	5.70	5.06	100.00	1	1	0	1	3.00	3
1998	34.00	35.44	51.90	43.67	15.19	100.00	2	1	0	1	5.00	5
1998	35.01	34.18	46.84	40.51	12.66	100.00	1	2	0	1	3.00	11
1998	35.02	13.92	22.78	18.35	3.80	100.00	1	2	0	1	5.00	7
1998	36.02	15.19	15.19	15.19	3.80	53.85	1	2	0	1	4.00	6

1998	37.01	39.24	45.57	42.40	2.53	100.00	1	3	0	1	11.00	11
1998	37.02	45.57	59.49	52.53	15.19	100.00	1	3	0	1	11.00	12
1998	38.00	39.24	58.23	48.73	11.39	78.26	1	1	0	1	5.00	5
1998	39.00	12.66	18.99	15.82	7.59	100.00	1	1	0	1	3.00	3
1999	44.00	56.96	68.35	62.66	12.66	98.15	1	1	0	1	7.00	6
1999	45.00	18.99	41.77	30.38	21.52	100.00	1	1	0	1	5.00	4
1999	46.00	72.15	88.60	80.38	12.66	100.00	1	1	0	1	5.00	4
1999	47.00	24.05	32.91	28.48	6.33	50.00	0	1	0	1	10.00	7
1999	48.00	11.39	40.51	25.95	29.87	100.00	1	1	0	1	3.00	4
1999	49.00	25.32	40.51	32.91	15.19	100.00	1	1	0	1	10.00	7
1999	50.00	17.72	36.71	27.21	16.46	93.10	1	1	0	1	5.00	5
1999	51.00	44.30	54.43	49.37	10.13	97.62	2	1	0	1	5.00	4
1999	53.00	55.00	60.00	57.50	6.25	58.33	1	1	0	1	11.00	9
1999	55.00	37.97	50.63	44.30	12.66	100.00	1	1	0	1	9.00	4
1999	57.01	20.25	40.51	30.38	20.25	76.67	1	2	1	1	8.00	5
1999	57.02	7.50	36.25	21.88	28.75	82.14	1	2	1	1	10.00	4
1999	58.01	31.65	49.37	40.51	16.67	100.00	1	2	0	1	8.00	4
1999	58.02	39.24	44.30	41.77	5.06	100.00	1	2	0	1	7.00	5
1999	59.01	7.59	10.13	8.86	2.53	0.00	0	2	0	0	2.00	1
1999	59.02	39.24	62.03	50.63	22.78	0.00	1	2	0	1	10.00	5
1999	60.00	44.30	60.76	52.53	16.46	26.53	1	1	0	1	7.00	5
2000	61.00	22.78	29.11	25.95	6.33	86.96	1	1	1	1	6.00	5
2000	63.00	37.04	62.96	50.00	22.78	66.67	1	1	0	1	5.00	5
2000	64.00	45.57	58.23	51.90	12.66	97.82	1	1	0	1	11.00	6
2000	66.00	31.65	73.42	52.53	16.46	79.31	1	1	0	1	8.00	5
2000	68.00	58.23	88.61	73.42	27.85	94.29	1	1	0	1	8.00	5
2000	71.00	31.65	36.71	34.18	5.06	93.10	2	1	1	1	10.00	6
2000	72.00	13.92	20.25	17.09	6.33	100.00	1	1	1	1	7.00	5
2000	73.00	24.05	41.77	32.91	17.72	96.97	1	1	1	1	9.00	6
2000	75.00	5.06	6.33	5.69	1.27	100.00	1	1	1	1	2.00	3

2000	76.00	22.78	27.85	25.32	5.06	77.27	1	1	1	1	6.00	6
2000	77.00	8.86	10.13	9.49	1.27	75.00	0	1	0	1	2.00	3
2000	78.00	20.25	35.00	27.63	11.25	82.14	1	1	1	1	11.00	8
2000	79.00	10.13	10.13	10.13	0.00	75.00	1	1	1	1	5.00	5
2000	80.02	15.19	35.44	25.31	21.52	55.56	1	2	1	1	5.00	6
2001	81.00	25.00	33.75	29.38	8.75	81.48	1	1	1	1	5.00	7
2001	83.00	2.60	3.90	3.25	1.30	66.67	1	1	1	1	4.00	2
2001	84.00	8.86	11.39	10.13	2.53	66.67	1	1	1	1	1.00	2
2001	85.00	5.06	6.33	5.70	1.27	60.00	1	1	1	1	1.00	3
2001	86.00	25.32	44.30	34.81	18.99	91.43	1	1	1	1	10.00	8
2001	87.00	29.11	40.51	34.81	18.99	100.00	1	1	1	1	5.00	7

C2: Codebook for Table Heading Abbreviations

Abbreviation	Meaning	Range	Additional Notes
RI	Representation Index	0-100	
FR	Fragmentation Ratio	0-100	
Cmp	Completeness	0-100	
RI/FR/Cmp LI	Observed Representation Index/Fragmentation Ratio/Completeness scores recorded during fieldwork by LIEVERSE	0-100	
RI/FR/Cmp UR	Estimated Representation Index/Fragmentation Ratio/Completeness scores recorded from photographic documentation by URLACHER	0-100	
Completeness/FR Bias	Average difference per element in observed vs. estimated Cmp/FR scores	-100-100	Negative values indicate estimated scores greater than observed (over-estimation), and vice versa
AI	Articulation Index	0-100	
Difference RI/FR/Cmp	Absolute difference between average observed and average estimated scores	0-100	
No Img	Number of photographic images associated with burial	1-14	
Angles	Number of angle types used in photographic documentation of burials	1	Single angle – full body only
		2	Single angle – isolated element
		3	Two angles – full body + $\frac{3}{4}$ length <i>or</i> passport
		4	Two angles – full body + isolated element (Incl. skull)
		5	Three angles – full body, $\frac{3}{4}$ length <i>and</i> passport
		6	Three angles – full body, $\frac{3}{4}$ length <i>or</i> passport + isolated elements above waist
		7	Three angles – full body, $\frac{3}{4}$ length <i>or</i> passport + isolated elements below waist
		8	Four angles – full body, $\frac{3}{4}$ length <i>and</i> passport + lower body
		9	Four angles – full body, $\frac{3}{4}$ length <i>and</i> passport + isolated region of upper body (e.g. pelvis)
		10	Four angles – full body, $\frac{3}{4}$ length <i>and</i> passport + other isolated elements or region (e.g. single hand, single foot)
		11	Five angles – full body, $\frac{3}{4}$ length <i>and</i> passport , lower body + other isolated

			element or region
Pos	Body position	0	Unknown
		1	Supine
		2	Semi-flexed
No_inter	Number of individuals interred	1	Single burial
		2	Double Burial
		3	Triple burial
Dist	Burial disturbance	0	Undisturbed
		1	Disturbed
Integ	Burial Integrity	0	Unknown or Secondary
		1	Primary

