2020

Potential Impacts of Artificial Intelligence on Spine Imaging Interpretation and Diagnosis

David Howard Durrant
Walden University

Walden University

College of Health Sciences

This is to certify that the doctoral dissertation by

David H. Durrant

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Raymond Thron, Committee Chairperson, Health Services Faculty
Dr. Cynthia Newell, Committee Member, Health Services Faculty
Dr. Vibha Kumar, University Reviewer, Health Services Faculty

Chief Academic Officer and Provost
Sue Subocz, Ph.D.

Walden University
2020

Abstract

Potential Impacts of Artificial Intelligence on Spine Imaging Interpretation and Diagnosis

by

David H. Durrant

DC, Logan College, 1982

BS, Logan College, 1980

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Health Sciences

Walden University

February 2020

Abstract

Spine and related disorders represent one of the most common causes of pain and disability in the United States. Imaging represents an important diagnostic procedure in spine care. Imaging studies contain actionable data and insights undetectable through routine visual analysis. Convergent advances in imaging, artificial intelligence (AI), and radiomic methods has revealed the potential of multiscale in vivo interrogation to improve the assessment and monitoring of pathology. AI offers various types of decision support through the analysis of structured and unstructured data. The primary purpose of this qualitative exploratory case study was to identify the potential impacts of AI solutions on spine imaging interpretation and diagnosis. Selected constructs from the diffusion of innovations theory and the technology acceptance model provided the conceptual framework. Data were acquired from 4 consensus-based white papers, researcher reflective journaling, and 2 homogenous focus group sessions comprising radiologists and AI experts. Content and thematic analyses of acquired data were performed with ATLAS.ti. Three primary themes emerged from qualitative analysis: patient-based decision support, population-based decision support, and application-based decision support. Subthemes include multiscale in vivo analysis, naturally language processing, change analysis, prioritization, and ground truth. The results suggest how further development of AI could fundamentally alter how spine pathology is detected, characterized, and classified. The study also addresses the potential impact of AI on in vivo tissue analysis, the differential diagnosis, and imaging workflow. This includes introducing the concept of the virtual biopsy and its use in spine imaging.

Potential Impacts of Artificial Intelligence on Spine Imaging Interpretation and Diagnosis

by

David H. Durrant

DC, Logan College, 1982

BS, Logan College, 1980

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Health Sciences

Walden University

February 2020

Dedications

I dedicate this work to my loving wife Christine, for her unconditional love and support, and to my beautiful daughters Amanda, Alyssa, Heather, and Hannah who inspire me each day. My family's enduring love will forever motivate me to fulfill my potential for the benefit of others. My greatest wish is to be there for them and to make them proud.

I dedicate this work to my beloved father Richard Durrant, who is greatly missed, and to my mother Patricia Durrant-Stuffle and stepfather Earl Stuffle for their love and guidance throughout the years. Without their support, this work and academic journey may not have been possible.

I dedicate this work to all those patients who will benefit from the future use of artificial intelligence in health care.

Acknowledgments

The pursuit of truth through education and rigorous research is a noble undertaking. I have been able to participate in this process because of the support of many individuals including family, colleagues, and professors. I am grateful for the opportunity to be able to play a part in the development of artificial intelligence for the converging fields of radiology, pathology, and genomics. I look forward to working with others to help the shape the future of health care.

I would like to thank my professors and the administrators at Walden University for providing the educational process that prepared me for this pioneering work and for future scholarly endeavors. I would like to offer heartfelt acknowledgment to Robert Hoye, PhD, who served as the first chair of my dissertation committee prior to his retirement. Dr Hoye offered practical advice and taught me to work smarter, not harder. He helped set the stage for my academic success in the doctoral program. I would also like to thank Raymond Thron, PhD who assumed the role as my dissertation committee chair. Dr Thron was always available when I needed him. He provided invaluable insights and support which helped guide me through the final stages of the research and dissertation process. I would like to acknowledge the roles of Cynthia Newell, PhD (committee member) and Vibha Kumar, Dr.PH, MD (university research reviewer). Dr. Newell and Dr Kumar provided direction throughout the dissertation process, which helped me cross the finish line with scholarly work. Thank you to Jeff Zuckerman, MA, who reviewed my final dissertation draft in the last hours and ensured my formatting met the criteria.

I am thankful for all of the AI experts and radiologists who gave up time in their busy schedules to participate in the focus group sessions. I am also appreciative of the following individuals whose pioneering work and insights inspired and directed me during this work. This includes Hugo Aerts, PhD, director of the Computational Imaging and Bioinformatics Lab at Harvard; Marilyn Giger, PhD, professor of radiology at the University of Chicago; and Robert Giles, PhD, senior member of the Moffitt Cancer Center. I would also like to thank Anne Osborne, MD, neuroradiologist, past president of the American Society of Neuroradiology, and professor of radiology at the University of Utah for her periodic guidance and support.

I am grateful to my family who made many selfless sacrifices over the past few years that allowed me to purse the topic of this work and to meet the stringent requirements of a PhD program. We all gave up some precious time together. The sacrifices laid the foundation for future contributions to the field of artificial intelligence and to the ongoing development of the virtual biopsy. I look forward to spending more time with my family and joining them on the journey forward into the era of artificial intelligence.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction to the Study

Introduction

The volume of data in health care is exploding exponentially. This trend is supported by digitization of health care data from different sources along with advances in imaging technology and molecular diagnostic methods (Murdoch & Detsky, 2013; Raghupathi & Raghupathi, 2014). The overwhelming volume of data in health care contributes to chaos and uncertainty that can lead to diagnostic errors with devastating consequences such as chronic pain, disability, or death (Gupta et al., 2017; Saber-Tehrani et al., 2013). The radiologist is rapidly becoming one of the most important curators and gatekeepers of big data. Despite this important role, the complexity and volume of available data during the interpretive stage of radiology workflow outpaces the individual radiologist's capacity to make fully informed and timely decisions (Thrall et al., 2016). The rate of diagnostic error associated with the interpretation of abnormal radiology studies has been reported as high as 30% with retrospective review of diagnostic imaging studies yielding even higher error rates (Berlin, 2007; Brady et al., 2012; Donald & Barnard, 2012).

The causes of errors during the interpretive stage of radiology workflow include exposure to complex data from new imaging technology, limited access to relevant non-imaging data, increasing workload, physiological fatigue, and human bias (Brady et al., 2012; Donald & Barnard, 2012; C. S. Lee, Nagy, Weaver, & Newman-Toker, 2012). Without adequate technological assistance, the human interpretive process within radiology workflow will become progressively more inaccurate, inefficient, and untimely

(Croskerry 2013; Manrai et al., 2014; Obermeyer & Emanuel, 2016; Ragupathi & Ragupathi, 2014). The solution requires the adoption and implementation of new data management and decision support solutions such as artificial intelligence (AI).

AI offers a variety of approaches to assist in problem solving and prediction (Pannu, 2015). The process encompasses different computational methods such as machine learning, deep learning, and cognitive computing. Using any of these AI solutions during radiology workflow might improve diagnostic accuracy and timing. AI is capable of revealing biological variability, heterogeneity of pathology, and comorbidities, all of which support more personalized and precise diagnosis (Gillies, Kinahan, & Hricka, 2016; Limkin et al., 2017; Pannu, 2015; Yip & Aerts, 2017). AI will fundamentally alter the field of radiology by facilitating more predictive, preventive, and participatory health care (Ghasemi et al., 2016; Hillman & Goldsmith, 2011; Jha & Topol, 2016). The potential impact of AI in radiology must be better understood to be adopted, implemented, and supported. Research will contribute to this process and contribute to positive social change by introducing decision support solutions capable of improving diagnostic accuracy and reducing health care costs.

My primary goal for this qualitative exploratory research study was to identify the potential impact of AI on spine imaging interpretation and diagnosis. Chapter 1 includes the fundamental elements of the study. These include the background, problem statement, and research questions. The chapter also includes an introduction to relevant terms and definitions to enhance the reader's understanding of the research topic. I present the theoretical perspectives and conceptual framework used for the design of the study. I also discuss relevant assumptions, delimitations, and limitations in the study. Chapter 1 ends with a discussion of the social significance of the research study along with a transitional summary to the next chapter.

Background

The success of diagnostic imaging is highly dependent on technology and protocol. Technological advances in radiology result in a perpetual cycle of discovery, disruption, opportunity, and adaptation. Current innovations are rapidly transforming radiology from a qualitative to a quantitative science with a growing capacity to obtain molecular and physiologic measures (Jha & Topol, 2016; Quer et al., 2017). This evolutionary process has led to unprecedented growth in the volume of multidimensional data which has and will continue to have an impact on radiology workflow and the type of decisions which have to be made. Exposure to an increasing quantity of complex data increases the likelihood of diagnostic errors (Obermeyer & Ezekierl, 2016; Ragupathi & Ragupathi, 2014). The growing demand to read imaging studies faster also influences interpretive accuracy. For example, the average emergency radiologist may read up to 50,000 to 100,000 images per day, resulting in an average of 2 to 3 seconds spent on each image (Syeda-Mahmood, 2015). Sokolovskaya et al. (2015) reported that radiologists pushed to read complex diagnostic imaging studies at a faster speed made over twice as many interpretive errors as those who read the same studies at a normal rate. Emerging health care system productivity goals often drive the radiologist to interpret imaging studies in a shorter period of time. The time constraints combined with the growing complexity of studies adds to the burden of decision making.

The error rate in radiology has been a concern for decades. Pioneering research performed by Garland (1949) over 60 years ago revealed an estimated 30% level of inaccuracy regarding the interpretation of abnormal chest radiographs. This landmark study prompted follow-up research. The error rate of radiologists on the interpretation of abnormal imaging studies between 1949 and 1992 was estimated at 30%, consistent with Garland's work decades prior (Berlin, 1994; Renfrew et al., 1992). Additional studies have revealed error rates for the interpretation of different types of abnormal imaging studies in the range of 15–35% (Elmore, et al., 1994; Lehr et al., 1976; Janjul et al., 1998). Interpretive errors in the current health care environment are potentially more influential than they have been in the past because of widely distributed results within electronic medical record (EMR) systems. Interpretive errors embedded within EMR can lead to exposure at multiple levels along the health care chain, setting the stage for additional errors in subsequent testing, clinical application, and judgment.

Paralleling the burden of big data in radiology there is a growing demand for more precise imaging interpretation and concise reporting to support personalized care. To successfully meet this demand, new decision support solutions have to be embedded within the interpretive stage of radiology workflow. Success will require the integration of human and machine attributes, a form of collective intelligence (CI). Radiomics, a rapidly emerging AI supported process, is capable of performing high-throughput automated analysis of imaging data through a series of sequential steps such as pathology detection, feature extraction, feature characterization, and analysis (Lambin et al., 2017; Lee et al., 2017; Limkin et al., 2017). Pathology features include physiological, molecular, morphological, statistical, and textual attributes (Aerts, 2017). Successful use of radiomics has the potential to help detect, characterize, and classify disease. It is capable of revealing new signatures of pathology not available through visual interpretation and contributing to probability-based decision support.

IBM developed a cognitive computing system in 2007 referred to as Watson, which introduced to the public in 2011 as a potential AI solution for decision support in the presence of complex data and uncertainty (Hoyt, Snider, Thompson, & Mantravadi, 2016). Watson was designed to augment the role of the radiologist during the interpretive stage of radiology workflow and during the differential diagnostic process (Brink et al., 2017; Jha & Topol, 2016; Kharat & Singhal, 2017; Jiang et al., 2017; Ranschaert, 2016). AI solutions such as IBM Watson have the potential to improve diagnostic accuracy and timing through the use of automated pattern detection, evidence analysis, and probability-weighted scoring (Doyle-Lindrud, 2015; Ferrucci, 2012; Kohn et al., 2014). Some AI systems are capable of learning with exposure to structured and unstructured data. This includes exposure to data from radiomic methods, prior imaging studies, electronic health records, disease databases, computational disease models, and peer-reviewed literature.

Interpretive and diagnostic errors in radiology occur due to many conditions including technological limitations, system inefficiencies, and human error (Berlin, 2013; Thammasitboon, Thammasitboon, & Singhal, 2013). Human perceptual errors occur more frequently than cognitive errors (Berlin, 2013; Bruno, Walker, & Abujudeh, 2015; Donald & Barnard, 2012). The causes of perceptual errors include reader fatigue, distractions, and the presence of hidden or subtle disease patterns overlooked with visual analysis methods (Bruno et al., 2015; C. S. Lee et al., 2012). Potential solutions for reducing errors include redesign of the interpretive stage of radiology workflow to include AI solutions, better integration of radiology-pathology data, 3D viewing options, and structured reporting (C. S. Lee et al., 2012). The use of AI has already demonstrated that it can improve the accuracy and quality of some decisions in radiology and in other areas of health care (Brink et al., 2017; Jiang et al., 2017). The potential role of AI in radiology is relatively new, subsequently; research and development will be influenced by needs analysis and meaningful use cases. Research on the use of AI in spine care, and more specifically spine imaging is extremely limited, despite its potential for having a favorable impact on one of the most common causes of pain and disability.

The volume of personalized data acquired with advanced diagnostic imaging is growing as the result of more detailed tissue interrogation, expanded fields of view, thinner slice thickness, and the use of multimodality and multiparametric methods (Aerts, 2017; Hillman & Goldsmith, 2011; Jha & Topol, 2016; Quer et al., 2017). This trend will influence all areas of radiology including spine imaging. The increased volume of accessible personalized data contributes to the complexity of decision making surrounding individual patient care and to inconsistencies and variability in radiology reports.

According to Lundstrom, Gilmore, and Ros (2017), the interpretive stage of radiology workflow may serve as the primary hub for the convergence and analysis of imaging data, pathology data, and genomic data. This would empower the radiologist as a curator of integrated patient data and as a clinical consultant. New decision support solutions accessible during image interpretation will have a significant impact on the timing and precision of the diagnosis process and subsequently decisions at the point-of care. Research has demonstrated that AI use offers “higher than clinician-grade accuracy” in numerous fields such as in dermatology, ophthalmology, and radiology (Syeda-Mahmood, 2018, p. 573). AI has the potential to support more accurate, efficient, and timely decision making in radiology. Its role in spine imaging needs to be addressed. AI applications could help overcome errors associated with human bias while democratizing expert decision support. I designed this study to identify how AI could improve personalized spine care and help direct further investigation on what is required to achieve this goal.

Problem Statement

The unprecedented growth of data in radiology is driven by new imaging technology and methods capable of performing whole body surveys and evaluating tissues at multiple length scales such as anatomic, physiologic, cellular, and molecular levels (Aerts, 2017; Hillman & Goldsmith, 2011, Murdoch & Detsky, 2013; Raghupathi & Raghupathi, 2014). Radiology is rapidly evolving into a dynamic, whole systems diagnostic discipline capable of interrogating multiple dimensions of pathology in vivo (D. Y. Lee & Li, 2009). The growing complexity and volume of imaging and non-imaging data available during the interpretive stage of radiology workflow is exceeding the individual radiologist’s ability to make fully informed decisions (Thrall et al., 2016). Radiologists are finding it increasingly difficult to determine what imaging findings are clinically significant and meaningful. Big data in radiology has become a burden which can be overcome to create new opportunities for patient care. The primary problem addressed by this research was defining how using AI during the interpretive stage of spine imaging could augment the role of the radiologist and improve diagnostic precision.

As previously stated, the rate of error on abnormal radiology studies has been reported as high as 30%; retrospective review of diagnostic imaging studies yielding even higher error rates (Berlin, 2007, 2013; Brady et al., 2012, Donald & Barnard, 2012). An overwhelming amount of data contributes to missed patterns and relationships resulting in diagnostic errors with deleterious consequences (Gupta et al., 2017; Saber-Tehrani et al., 2013). The causes of error in radiology include the burden of complex data, increasing workloads; limited time, physiological fatigue, and human bias (Brady et al., 2012; Donald & Barnard, 2012; C. S. Lee et al., 2012)**.** The high level of inconsistency and variability in radiology workflow is influenced by what data are acquired, how the data are analyzed, and how the results are reported (Aerts et al., 2013). Several authors have suggested that without adequate technological assistance, the human interpretive process within radiology workflow will become progressively more inaccurate, inefficient, and untimely (Croskerry 2013; Manrai et al., 2014; Obermeyer & Emanuel, 2016; Ragupathi & Ragupathi, 2014). Few studies have addressed how to reduce interpretive errors in radiology (Fitzgerald, 2005; Lee, 2017).

I was unable to identify any peer-reviewed articles in which the authors specifically address how to reduce interpretive errors associated with big data in spine imaging. I was also unable to locate peer-reviewed publications in which the authors address the use of AI methods such as radiomics combined with text analysis to reduce interpretive errors in spine imaging. The spine-related complaints represent one of the most common causes of chronic pain and disability in the United States (Hurwitz et al., 2018; Loney & Stratford, 1999; Ricca et al., 2006). The American College of Radiology (2016) recently acknowledged the need for further research to help identify how AI could be used to access meaningful data, enhance the interpretive phase of radiology workflow, improve diagnostic accuracy, and reduce errors.

The use of decision support solutions such as AI during the interpretive stage of spine imaging workflow could lead to more accurate detection, characterization, and diagnosis of pathology. Despite heightened awareness of the need for new decision support solutions and knowledge of the potential benefits of AI use in radiology adequate research on the potential impact on spine imaging has not been performed. Qualitative exploratory research is required to lay the foundation for further research surrounding the use of AI and to identify how benefits can be achieved.

Purpose of the Study

There is a considerable amount of meaningful data and insight embedded within medical images, undetectable through routine visual analysis, and therefore not considered in the diagnostic process (Gillies et al., 2016; V. S. Lee, 2017). Such data embedded within non-imaging records could be made available during the interpretive stage of radiology workflow for correlation with imaging findings. Missed data or misinterpretation of data may lead to an error in diagnosis. Overlooked data and data patterns can also lead to missed opportunities in care. The primary purpose of this research study was to explore the potential of AI to favorably impact diagnostic accuracy and precision during the interpretive stage of spine imaging workflow. The approach includes the potential role of natural language processing and radiomic methods to assist in the detection, characterization, and monitoring of pathology.

Research Questions

The primary research question for this study was: What are the opinions of experts regarding the potential use and impact of AI intelligence during the interpretive stage of spine imaging workflow?

The subquestions for this study follow:

1. How could the use of AI-supported methods (auto detection, segmentation, radiomics, natural language processing) during the interpretive stage of spine imaging influence the differential diagnostic process?
2. How could the use of AI-supported methods during the interpretive stage of spine imaging influence disease classification and staging?
3. What AI solutions could be used to create interpretive priority in spine imaging?
4. What will the future of spine imaging interpretation workflow look like?
5. Could AI-supported solutions such as radiomics be used to interrogate spinal tissue in vivo and eventually lead to a virtual (digital) biopsy?
6. What are some potential advantages of an in vivo virtual “digital” biopsy over a traditional needle biopsy in spine care?
7. How could the use of AI-supported augmented reality (AR) or virtual reality (VR) enhance the evaluation of pathology in spine imaging?

8. What are some potentially “meaningful use” applications of AI in spine imaging?
9. Which construct of the technology acceptance model (TAM) will likely have a greater impact on AI adoption during the interpretive stage of spine imaging: perceived benefits or perceived ease-of-use?
10. Which characteristics of innovations proposed by the diffusion of innovation theory (DOI) will likely have the greatest impact on AI adoption during the interpretive stage of spine imaging: complexity, compatibility and interoperability, observed effects or trialability?

Theoretical Perspectives and Conceptual Framework

Adopting and using new technology is influenced by many factors, including awareness of technology’s potential, ease of use, and interoperability with existing technology and workflow (Khan & Woosley, 2011; Oye, Iahad, & Abrahim, 2012). The potential impact of new technology must be studied in-depth using a flexible and iterative approach. A rigid theoretical framework would not allow for adequate exploration of this topic of study. Exploratory research offers a descriptive and inductive approach which supports discovery, and helps reveal potential advantages and disadvantages surrounding the use of new technology (Creswell, 2009; Patton, 1990; Yin, 2014). Effective qualitative research design requires a conceptual framework to help align theoretical perspectives, the research purpose, and research strategies **(**Creswell, 2007; Marshall & Rossman, 2011; Patton, 1990). Conceptual boundaries can be developed through the integration of ontological, epistemological, methodological, and structural perspectives.

This research study was inductive, beginning with a set of assumptions and research questions leading to data acquisition and analysis. This study would have been restricted by the rigid application of an a priori theory or hypothesis. I did not use a deductive approach. A qualitative exploratory approach supports the investigation of possibilities within the framework of different contexts and realities (Creswell, 2013). Because of the potential depth and breadth of this topic of study, it would have been unfeasible to perform the exploratory study through a single theoretical lens or worldview.

AI represents a complex technology associated with numerous processes; therefore, its potential impact must be evaluated from a pluralistic perspective. No single theory can be used to fully explore the potential role and impact of AI during the interpretive stage of spine imaging workflow. I used an interpretivist-constructivist epistemological view to address the topic of study. The interpretivist perspective helped operationalize how and why questions surrounding the use of AI. A constructivist perspective supported the development of a proposed model for using AI solutions during the interpretive stage of spine imaging workflow. Consistent with the work of Creswell (2013) and Patton (1990), I used theoretical perspectives to help develop the conceptual framework of the study and to develop research questions which aligned with the research methodology and purpose.

The diffusion of innovations theory (DOI) proposed by Rogers (1995) represents one of the most widely accepted theories for exploring how the attributes of innovative technology influence its adoption, use, and distribution (Kahn & Woosley, 2011).

Furthermore, DOI offers insight about the stages of adoption and adopter categories (Dillon & Morris, 1996; Rogers, 2003). The characteristics of innovations that influence adoption proposed by Rogers (2003) include relative advantage, interoperability with existing systems, the potential for technology evolution, reinvention, complexity, and trialability. Additional factors that influence adoption include awareness of benefits (knowledge), knowledge of advantages and disadvantages (decision), method of use (implementation), and clinical utility (confirmation). DOI includes five adopter categories: innovators, early adopters, early majority, late majority, and laggards (Rogers, 2003). The literature review summarized in Chapter 2 includes information regarding the use of theoretical constructs and related conceptual perspectives introduced in this treatise. I designed this study to obtain consensus-based perspectives from published white papers and insights from purposively selected innovators and early adopters in focus group sessions.

The technology acceptance model (TAM) proposed by Venkatesh (2008) is used to address behavior and perceptions that influence willingness to use new technology. The two primary constructs of TAM are perceived usefulness (PU) and perceived ease-of-use (PEOU). The synthesis of DOI and TAM constructs have been previously used to frame exploratory research of computer and information technology (Carter & Belanger, 2005; Y. Lee, Hsieh, & Hsu, 2011; Legris, Ingram, & Colerette, 2003**).** DOI represents an effective theory for addressing variables surrounding technology adoption, whereas TAM offers insight about post-adoption use and support (Hameed & Arachchilage, 2016). Select constructs from DOI were combined with constructs from TAM to guide the development of research questions to be used in focus group sessions and to help develop an initial a priori list of thematic coding categories for data analysis.

A well-developed conceptual framework provides boundaries that help direct the acquisition, management, and analysis of data during research. A conceptual framework can also be used to assess the impact of new technology on professional behavior and workflow (Bogdan & Bilken, 1982; Patton, 1999). The conceptual framework of this study supported a contextual inductive and iterative process for exploring the potential impact of AI on the interpretive stage of spine imaging workflow. I used the conceptual framework for this study to choose data analysis strategies and to help guide the literature search summarized in Chapter 2. I focused on literature that revealed AI technology characteristics and the goals of early adopters of AI in radiology. Research questions were developed to address some of the perspectives offered by DOI and TAM constructs. The chosen method of inquiry in this study required answers to how and why questions surrounding the potential benefits and challenges associated with AI use during the interpretive stage of spine imaging workflow.

Nature of the Study

I performed this qualitative exploratory case study with an in-depth, inductive approach to acquire and analyze data to address the potential impact of AI on spine imaging interpretation and diagnosis. Qualitative exploratory case study research offers an effective method for acquiring contextual knowledge surrounding the use of new technology (Creswell, 2013; Ponelis, 2015; Yin, 2014). The approach is also effective at revealing themes surrounding the potential impact and role of new technology (Bradley, Curry, & Devers, 2007; Patton, 1999). Qualitative exploratory research studies have been successfully used in radiology (Sanberg et al., 2012)**.** The approach in this case offered the level of holistic investigation required to reveal previously unknown variables surrounding the use of AI during the interpretive stage of radiology workflow.

A well-defined unit of study helps direct qualitative research. It also helps frame and inform research strategies. The boundaries of a qualitative case study may consist of a relationship, situation, process or a culture (Miles, Huberman, & Saldana, 2014). The primary unit of study for this research is a process defined as the interpretive stage of spine imaging workflow. This includes the workstation, display technology, embedded AI solutions, available data, and the role of the radiologist. I chose the unit of study to identify how and why AI solutions should be used during the interpretation of advanced spine imaging. I also designed the study to evaluate the potential impact of AI on the current state of the unit of analysis.

Qualitative exploratory research is capable of revealing contextual complexities not adequately addressed by explanatory or quantitative research methods (Ponelis, 2015; Yin, 1984). More specifically, qualitative research can be used to reveal themes and patterns, not revealed with restrictive quantitative methods (Dubin, 1969; Patton, 2002; Ryan & Bernard; 2003). This study required the acquisition of data from multiple sources including expert documents, reflective journaling, and focus group sessions. The focus groups comprised radiologists and AI experts who met study inclusion and exclusion criteria. I used numerous methods to improve the trustworthiness of the study.

Definitions of Relevant Terms

I have listed the operational definitions of selected terms and phrases to help the reader understand the research topic, findings, and conclusions.

Algorithm: An algorithm represents a mathematical formula or program, used as a set of instructions for computational analysis (Kakhani et al., 2017).

Artificial intelligence (AI): Artificial intelligence refers to the use of technology to identify patterns, solve problems, and make predictions using biological-like approaches (Jiang et al., 2017).

Augmented intelligence (AI): Augmented intelligence refers to the use of technology to enhance human capability and to provide decision support (Liew, 2018).

Big data: Big data refers to a large volume of complex data that is difficult to analyze utilizing traditional methods (Farooki, Almeida, & Saltz, 2016).

Collective intelligence (CI): Collective intelligence refers to the combined use of AI solutions and human intelligence to identify a pattern, solve a problem, perform a task, or make a prediction.

*Computer-aided detection (CADe): C*omputer-aided detection refers to the use of a computer system to identify a pattern within structured data, unstructured data, or on a set of diagnostic images (Hillman & Goldsmith, 2011).

Computer-aided diagnosis (CADx): Computer-aided diagnosis refers to the use of a computer system and computational methods of analysis to provide a list of differential diagnostic possibilities or a specific diagnosis (Hillman & Goldsmith, 2011).

Data science: Data science represents an interdisciplinary field that uses computational methods to analyze data, to reveal patterns, study processes or to make predictions (Aerts, 2017).

Deep learning (DL): Deep learning represents a subset of machine learning capable of becoming more accurate and capable with exposure to data through the use of multiple hidden processing layers (Tang et al., 2018).

Differential diagnostic process: A differential diagnostic process is a series of interrelated steps that use probability-based logic or reasoning to differentiate a disease or disorder from others that may have a similar presentation.

Ground truth: Ground truth refers to data assumed or proven to be true (Choy et al., 2018)

Innovation: Innovation is defined as a new or improved method of performing tasks or solving problems (George et al., 2005).

In vivo: In vivo refers to the evaluation of biological elements, processes or systems within a living organism (Lambin et al., 2012; Rizzo et al., 2018).

Machine learning (ML): Machine learning represents a subset of AI that uses a computer system and a set of algorithms to reveal patterns in data with or without handcrafted explicit instructions (Giger, 2018).

Natural language processing (NLP): Natural language processing refers to the use of automated computational methods to analyze and interpret spoken or written language (J. Y. Chen et al., 2017).

Neural networks: The phrase neural networks refer to a sophisticated network of layered interconnected data paths and nodes, which process signals in a manner similar to neurons in a biological system (Tang et al., 2018).

Precision medicine (PM): Precision medicine refers to an objective approach to the diagnosis and delivery of health care which takes into account the heterogeneity of disease along with an individual's unique biology, disease risk, and variable response to treatment (Mesko, 2017).

Radiomics: Radiomics refers to the science of high-throughput data analysis used for identifying, extracting, characterizing, quantifying, and classifying nonvisible elements of pathology within medical imaging data sets (Napel & Giger, 2015).

Segmentation: The definition of the term segmentation in radiology refers to the use of manual, semiautomated or automated methods to identify and outline a region of interest or an abnormality on images (Gillies et al., 2016).

Systems medicine: Systems medicine also referred to as network medicine refers to the science associated with the relationships and interconnections between molecular structures, cells tissues, and organs (McCue & McCue, 2017).

Virtual biopsy: The use of the phrase virtual biopsy in radiology refers to the use of various computational and radiomic methods to reveal, extract, and quantify in vivo characteristics of pathology derived from imaging data guided by a defined region of interest (Echgaray et al., 2016; Lambin et al., 2012; Thrall, 2016).

Assumptions

Assumptions that are true or likely to be true represent beliefs or perspectives. Assumptions influence research designs, methods, and conclusions. It is therefore important to acknowledge assumptions that may influence this research process. I had to consider the potential influence of numerous assumptions in this research study. I assumed that by choosing consensus-based white papers (research documents) published by reputable radiology organizations that I would acquire expert insights regarding the potential role of AI in radiology now and in the future. I assumed that four to six purposively selected participants (experts) placed into one of two homogenous focus group sessions would be enough to reach topic saturation. In addition, I assumed that the research participants would have sufficient knowledge and experience to answer focus group questions and productively contribute to focus group discussions. I also assumed that the research participants would adequately represent the levels of expertise held by others in their relevant fields.

I assumed based on the inclusion and exclusion criteria of this study that the participants would have a comfortable understanding of the research topic and goals. My use of a purposive sampling method increased the likelihood that each research participant had the experience and knowledge required to address the research topic questions. Each of the participants works in the health care field, thus improving the likelihood that they are aware of the importance of an honest and ethical approach to research.

I assumed that participating radiologists had limited knowledge of the potential role of AI in spine care and that AI experts had limited knowledge of the potential impact of AI on spine imaging interpretation. The presumed knowledge and limitations of AI specialists and radiologists opened the door for creative discussions and discovery in each session. The level of dedication required of individuals within each group to reach expert status infers a high level of passion and dedication to the subject; therefore, I assumed that each of the participants could contribute valuable insight and opinions during the focus group sessions. I assumed that, given the preparatory steps taken prior to data acquisition, each of the research participants fully understood their rights and responsibilities in the research project. The research assumptions helped guide me during data collection from focus group sessions and from other sources.

Scope and Delimitations

The topic of delimitations refers to anticipated constraints that may arise during the research design or while conducting research. Several delimitations arose in this case. Research participants voluntarily participated in the focus group sessions; therefore, it was likely that they were interested in the role of AI in radiology and spine care. The defined unit of study and boundaries surrounding this study served as relative constraints to data acquisition and interpretation. Participant inclusion and exclusion criteria also represented delimitations. The moderator guide and questions developed for the focus group sessions served as adaptable delimiters in the study.

The primary purpose of this research study was to explore the potential impacts of AI use during the interpretive stage of spine imaging workflow. I did not investigate the impact of AI on interprofessional relationships within the radiology department or its influence on economic issues surrounding the use of AI in spine care. I also did not address how AI might specifically impact radiologists of different backgrounds, training, and skill levels. I did not design the research project to address the administrative challenges or costs associated with developing, implementing, or supporting AI solutions.

Limitations

Research limitations refer to potential barriers or boundaries that investigators cannot control. Limitations in this study included the available knowledge and expertise from a small study population. Furthermore, the survey sample used in this study was purposive and small, resulting in a data analysis process which is more descriptive than inferential. The focus group participants' levels of expertise influenced research study conclusions. Limitations associated with lack of familiarity with the research topic were relatively low given the highly qualified nature of the experts who participated. The biases associated with the use of a single examiner posed a potential limitation. I reduced this risk with the use of validation methods such as member checking, within and between group analyses, reflective journaling, and triangulation of qualitative data acquired from numerous sources.

This research study consisted of a small purposively chosen population of experts; thereby, limiting the ability to generalize results. Determining whether sample size was adequate was influenced by many factors such as the outcome of the research validation methods. The type of methods used such as triangulation of data, member checking, and peer review must be carefully considered when determining sample size (Patton, 2002).

The use of two research participant categories and two focus group sessions limited the scope of this portion of the study. The potential role of AI in spine imaging is broad; therefore, a single case study cannot address all relevant dimensions of the topic. In addition, there was limited time and resources that restricted the duration of this study, rendering it difficult to study the perceived benefits of AI use during the interpretive stage of spine imaging over time.

Significance of the Study

Medical errors represent the third most common cause of death in the United States (Makary & Daniel, 2016) and one of the most costly and avoidable health care expenses (Saber-Tehrani et al., 2013). Diagnostic errors represent one of the most common reasons for malpractice claims (Andel et al., 2012). Diagnostic imaging represents a widely used method for detecting, characterizing, and monitoring disease. Subsequently, many of the important decisions made in health care arise from diagnostic imaging studies (D. Y. Lee & Li, 2009). The critical role and relevance of diagnostic imaging in health care is increasing. This study addresses a gap in the research literature associated with the potential impacts of AI on spine imaging workflow and the diagnostic process. Using AI during the interpretive stage of spine imaging workflow could help reduce human errors and favorably contribute to a more accurate diagnostic and reporting process, leading to improved patient care. Research in other health care and radiologic specialties has demonstrated that using AI can contribute more consistent, quantitative, and actionable information to the final report, thus influencing point of care decisions (Augimeri et al., 2016; Boone et al., 2015; Mohebian et al., 2017). AI solutions combined with effective data governance and data management in spine imaging could result in fewer interpretive errors and improved diagnostic precision.

Traditionally, radiology has relied upon the visual perception and interpretation of the radiologist (Pinto & Brunese, 2010). Perceptual factors surrounding visual interpretation represent a common cause of error. Factors that contribute to perceptual errors include reader fatigue, distractions, the presence of hidden or subtle disease patterns, and various form of human bias (Bruno et al., 2015; C. S. Lee et al., 2012). A specialized application of AI, referred to as radiomics, has been successfully used by radiologists to reveal the heterogeneity and in vivo features of pathology not detectable by the radiologist during the visual inspection of a study (Aerts, 2017; et al., 2016; H. W. Wu et al., 2012; Yip & Aerts, 2016). Radiomic methods improve the ability to classify disease and stratify treatment approaches, all leading to more precise and personalized patient care. In addition, AI solutions have the potential to improve the diagnostic imaging process through the ability to detect subvisual pathology and to correlate data from other imaging and non-imaging sources (Dreyer & Geis, 2017). This includes data from electronic health records, published literature, genetic databases, as well as from prognostic and predictive disease models. Digitization of in vivo and in vitro pathology supports image sharing and access to remote data analysis and expert consultations. Sharing of digital information also allows for consensus-based decision support.

Advances in medical technology have historically contributed to improved methods of detecting and treating disease (Clinton, 2000). Integrated co-evolution of AI supported methods such as radiomics will contribute to the discovery of new molecular signatures and biomarkers of disease. This will lead to new standards for evidence-based care in all fields including spine care. Successful integration of human and machine intelligence during the interpretive stage of spine imaging workflow has the potential to help control unsustainable health care costs, and support personalized care (Hillman & Goldsmith, 2011; Jha & Topol, 2016; Kressel, 2017; Lee, 2017). In addition, AI is capable of democratizing decision support that will aid underserved facilities, underserved regions, and inexperienced radiologists.

New levels of expectations and knowledge surrounding the use of AI will influence standards of care, which will ultimately benefit individuals, families, and society. Greater use of AI-supported diagnostic methods such as radiomics will expand classifications of disease and improve personalized care. The results of this research study contribute to positive social change by identifying how AI use during the interpretive stage of radiology workflow could favorably shape the future spine care. Successful use of AI during spine imaging interpretation could result in early detection, early intervention, better treatment outcome, and reduced direct, as well as indirect costs associated with chronic pain and disability. Moreover, AI could facilitate collaboration between spine care providers of all disciplines by exposing the fundamental basis for disease, democratizing expert decision support, and by providing evidence-based measures of treatment outcome. Widespread use of AI decision support will help overcome some of the barriers to collaboration associated with human ignorance and biases. I designed this study to reveal potential applications of AI in spine imaging, as well as in other fields and specialties. This study introduces the concept of the digital (virtual) biopsy that could have a profound influence on further investigation of this concept and its application in all fields of health care.

Summary

Given the variability in people and their spine disorders, spine care delivery needs to be more precise and personalized. Better integration of machine and human capabilities during the interpretive stage of radiology workflow will contribute to earlier detection of pathology, better characterization of pathology, and a more precise and timely diagnosis. AI and related solutions used in other fields such as oncology and cardiology can be adapted and used during the interpretive stage of spine imaging. The primary goal of diagnostic imaging in spine care is to provide insight and knowledge that can be used by a provider to deliver care for the right patient, for the right condition, at the right time.

Imaging represents one of the most commonly performed and revealing elements of the diagnostic workup in spine care. The combined burden associated with the growing volume of imaging and non-imaging data available during the interpretive stage of radiology workflow is increasing the complexity of decision making for the radiologist. The missed opportunity and error rate associated with the interpretation of abnormal imaging studies is too high. The growing complexity of data acquired through more advanced imaging technology will only contribute to more complex decisions and higher incidence of error. Applying AI solutions during the interpretive stage of spine imaging workflow has the potential to detect early-stage pathology, offer decision support, and facilitate personalized care.

The primary purpose of this research study was to identify whether the use of AI solutions during the interpretive stage of spine imaging could reduce the risk for diagnostic error and improve the precision of the final diagnosis. The potential AI solutions introduced in this chapter and investigated in the next chapter include natural language processing, radiomics, disease modeling, and computational analysis of structured and unstructured data. The success of the interpretive stage of radiology workflow is dependent on integrated solutions. It is becoming increasingly important to use quantitative measures in diagnostic imaging. The integration of various decision support solutions will likely change the landscape of radiology and spine care.

An accurate and precise diagnosis is dependent on heightened awareness of possibilities and the ability to investigate the possibilities by analyzing available data. Too often knowledge of differential diagnostic possibilities is limited by human experience and limited access to technologies required to identify subtle or hidden patterns within medical records or in the imaging data. This leads to errors in diagnosis and clinical judgment. AI technologies such as natural language processing and radiomics offer potential solutions. In Chapter 2, I address how current applications and contributions of AI in radiology could be adapted or further developed for use during the interpretive stage of spine imaging.

In this chapter, I introduced the research problem, research purpose, and research significance, to help guide my literature search and review addressed in Chapter 2. The subsequent chapter acknowledges current and potential applications of AI in radiology. In Chapter 2, I synthesized the research purpose and goals from this chapter with the results of an extensive literature search. This helped inform the research design and methodology introduced in Chapter 3.

Chapter 2: Literature Review

Introduction

Exposure to large volumes of complex data in radiology increases the level of uncertainty and complex decision making. This challenge combined with human limitations and bias increases the risk for errors and oversights. The field of radiology has always been dependent on the use of technology and has long served as a pillar in health care for the acquisition, analysis, and management of complex data **(**Thrall et al., 2016). Imaging has become one of the most important sources of data and diagnostic information in health care (Aerts et al., 2013; Gillies et al., 2016: Kim et al., 2013; Kinahan, & Hricak, 2016). The radiologist is poised to become a gatekeeper of big data and a disease consultant (di Piro et al., 2017). This role will be augmented with the convergence and integration of imaging, laboratory, genetic, and pathology data at the radiology workstation**.**

The field of radiology is rapidly transforming from a predominantly qualitative to a quantitative science supporting the rising demand for a more personalized diagnoses (Jha & Topol, 2016; Quer et al., 2017). In addition to the demand for a more precise diagnostic process, the radiologist is burdened with a growing quantity of complex data arising from advances in whole body imaging, molecular diagnostic methods, and the integration of multimodality and multiparametric imaging approaches (Murdoch & Detsky, 2013; Raghupathi & Raghupathi, 2014). New forms of decision support are required during radiology workflow to reduce the potential for interpretive and diagnostic errors. The results of an extensive literature search revealed how AI solutions have begun to have a significant impact on disease detection, characterization, and surveillance in health care.

In this chapter, I discuss the literature search strategy that I used to address the topic of study. I also address the fundamental concepts and research findings that I identified in the literature that support qualitative exploration of the potential role of AI use during the interpretative stage of spine imaging. Spine and spine-related research involving the use of computational decision support such as AI were limited in number and scope. The majority of published research and reviews addressed the role of AI solutions in other fields such as oncology and neuroimaging. The literature review identified prior methods of inquiry and research used to perform the studies. It also revealed gaps in the literature regarding the use of AI during the interpretive stage of non-spinal and spinal imaging. Scholarly publications provided the rationale for refining the research problem, as well as developing the research questions and methodology for this study. Given the limited number of publications referencing the use of AI in spine imaging, the literature search was expanded to include research on the use of AI during the interpretive stage of imaging other regions of the body. I chose and organized the topics of this chapter to present relevant findings from the literature. The literature search established a scholarly foundation for the research design and methodology covered in Chapter 3.

Literature Search Strategy

I performed an extensive literature search to explore existing perspectives and facts surrounding the topic of study. My helped to identify relevant theories, conceptual frameworks, and research strategies that were applied in this study. The literature search revealed current expectations and standards associated with the use of AI in radiology.

My literature search concentrated on articles that were published within 5 years of the anticipated completion date of this work. My search was primarily limited to peer-reviewed scholarly publications. I retrieved journal articles from the following research databases; PubMed, EBSCO, ProQuest, and Google Scholar. I used key terms and phrases along with different techniques to perform and refine the search process. Common search terms and phrases included *artificial intelligence, deep learning, machine learning, radiomics, natural language processing, interpretive radiology, spine, spine imaging, virtual biopsy, in vivo, voxel-wise detection, computer-aided detection, computer-aided diagnosis,* and *biometric analysis*. I used independent and combined search terms to improve the investigative process.

The literature search was highly iterative to achieve adequate saturation of the topic. I used single and combined Boolean operators, truncation, and wild card symbols in the search process. I reviewed publication abstracts to ascertain article relevance. All relevant research and review articles were printed, read, and saved. I evaluated the bibliographic reference lists of seminal articles for additional works. My literature search identified research studies, as well as consensus-based documents and position papers published by respected institutions and organizations. During the literature review process I was able to identify key experts on different research topics. I performed author-based searches to look for relevant material they may have authored or contributed to. Research librarians provided access to published articles I was unable to find through more traditional methods. Thus, the literature search on the topic was exhaustive.

Applied Theoretical and Conceptual Framework

Successful adoption and use of new technology are influenced by knowledge of its utility and its role within existing workflow (Khan & Woosley, 2011; Oye, Iahad, & Abrahim, 2012). This knowledge is acquired through reading published works, listening to colleagues or through hands-on experience. A rigid theoretical framework cannot be successfully used to study the role of AI in radiology due to its numerous components, rapid evolution and the complexity of its impact on the spectrum of workflow. As previously stated, AI represents a complex technology associated with numerous processes; therefore, its potential impact must be evaluated from a pluralistic perspective. Exploration using a conceptual framework developed from the synthesis of constructs and perspectives from different theories and models supported an adaptive and iterative approach to addressing the potential role of AI in this study. This approach is consistent with the work of Creswell (2009), Patton (1990), and Yin (2014).

The synthesis of DOI and TAM constructs have been used in numerous research studies to explore the potential applications and benefits of new technology (Carter & Belanger, 2005; Y. Lee et al, 2011; Legris, Ingram, & Colerette, 2003**).** In this case, the synthesis of constructs from DOI and TAM led to the use of practical descriptive categories such as perceived usefulness, perceived ease-of-use, clinical utility, diagnostic accuracy, interoperability, and workflow compatibility. Each of these perspectives is applied to the role of AI in radiology and is addressed through a variety of headings in this chapter. In summary, my use of DOI and TAM perspectives supported the development of a guiding conceptual framework used to help direct the literature review, create focus group questions, and analyze acquired data. These conceptual perspectives were also used to help frame the results of the literature search presented in this chapter.

The DOI proposed by Rogers (1995) acknowledges various categories of technology adopters. These categories include the innovator, early adopter, early majority, late majority, and laggards. Scholarly investigation of publications by innovators and early adopters of AI use in non-spine related specialty fields of radiology provided some of the insights required to address the topic of this research study.

Technology-Induced Transitions

The field of radiology has and will continue to be shaped by technology development and its evolution (Hillman & Goldsmith, 2011). Advances in imaging technology influence the timing and type of decisions, which have to be made during the interpretive stage of radiology workflow. Innovations also restructure the realm of expectations surrounding the analysis and flow of imaging data. Historically, transformative technological advances in radiology have included radiology information systems (RIS), picture archiving communication systems (PACs), natural language processing (NLP), and more recently AI solutions (Hillman & Goldsmith, 2011). Each of these technologies has improved some aspect of radiology workflow including the detection, characterization, and reporting of disease (Sardanelli, 2017). The rapid evolution of technology and data management in radiology is supported by the miniaturization of materials, increased computer processing power, improved data storage, enhanced connectivity, as well as greater access to disease models and registries.

Technological advances that offer decision support such as AI has and will continue to influence every stage of radiology workflow. Computational decision support influences the roles and responsibilities of radiologists. Widespread adoption and use of AI-based decision support will require efficient integration with existing workflow and legacy systems (Bauer, 2017; Dreyer & Geis, 2017). Learning how AI might improve spine imaging interpretation initially requires exploratory research.

Computer systems and related AI solutions are often associated with numerous components and processes, each of which contributes to new insights and technological advances. This process is often referred to as co-evolution. Arthur (2009) introduced the supporting premise that "existing technologies beget further technologies" (p. 21). The impact of new technologies and their spinoffs is often underestimated, whereas the speed of implementation is often overestimated. Knowledge and appreciation for AI technology co-evolution is required to predict its impact on interpretive workflow.

Technological advances in radiology lead to perpetual cycles of discovery, disruption, and adaptation. This unstable process introduces threats to established protocols and standards (Lai, 2017). Each time new technology emerges, its potential impact on the delivery of care has to be evaluated, and its clinical utility determined (Kressel, 2017). In addition, the benefits and risks associated with an innovation have to be disclosed and discussed. The attributes and influences of new technology such as AI must be considered when evaluating its potential impact during the interpretive stage of spine imaging workflow.

Potential AI solutions in radiology cannot be studied in isolation. They must be studied in the context of other technologies and processes embedded into workflow. This includes assessing its potential impact on diagnostic accuracy and efficiency. Zhang et al. (2004) introduced a hierarchical system of human and technological relationships that can contribute to or amplify medical error. The hierarchy includes the "individual, individual-technology interaction, distributed systems, organizational structures, institutional functions and overarching national regulations" (Zhang et al., 2004, p. 194). The impact of new technology and decision support can have both favorable and unfavorable outcomes. For this reason, the potential advantages and disadvantages of AI use during the interpretive stage of radiology workflow must be considered within the context of hierarchical professional and technological relationships.

Big Data Attributes and Related Burdens in Radiology

Diagnostic images are more than pictures. They contain massive quantities of minable and potentially meaningful data, often not considered during the interpretive process. Medical imaging is estimated to represent as much as 90% of all stored medical data, contained within billions of images (Lambin et al., 2017). Imaging data is present in many forms including symbols, words, images, and binary digits. Individual datum and aggregations of data have unique characteristics that influence how it is acquired, analyzed, and applied. The primary attributes of datum and data include value, variability, veracity, velocity, and volume (Gandomi & Haiderm, 2015; Nendaz &

Perrier, 2012). The primary sources of data available during the interpretive stage of radiology workflow are the study requisition, electronic medical records, diagnostic imaging, and disease registries (Brown, 2014). In the near future there will be greater access to relevant pathomic and genomic data at the radiology workstation. Human interpretation of diagnostic images is often limited to visual analysis and qualitative descriptive reporting (Thrall et al., 2016). Additional methods are required to analyze textual and nonvisible data in radiology.

Data are commonly classified as structured or unstructured. Structured data have well-defined form and context, whereas unstructured data tend to have inconsistent form, rendering it more difficult to analyze. The most common type of unstructured data is language, whether printed or verbal (Raghupathi & Ragupathi, 2014). The majority of personalized data in health care are unstructured, in the form of text, represented in medical records and reports.

Emerging computational methods are being used in radiology to help transform data and related patterns to meaningful information or knowledge that has clinical utility. Determining data relevancy in radiology requires knowledge of its validity and clinical utility. The majority of acquired data in radiology are considered noise and are not relevant to clinical care. This perspective is influenced by current limitations in pattern detection and limited human capacity to analyze nonvisual imaging data (Kohn et al., 2014). The large volume of data acquired during current imaging studies only represents a fraction of what will be acquired, mined, and transformed into action in the near future

(Kohn et al., 2014). What are considered irrelevant and meaningless data now may represent actionable data in the future.

A considerable amount of data and insight are embedded within medical images, but remain undetectable through traditional visual analysis (Gillies et al., 2016; Lee et al., 2017). As a result, radiologists face immense challenges during the interpretive stage of radiology workflow. Despite these current challenges advances in diagnostic imaging continues to create progressively larger and more complex data sets. Without adequate technological assistance human interpretation of complex imaging data will become progressively more inefficient, inaccurate, and untimely (Crosskery 2013; Manrai et al., 2014; Murdoch & Detsky, 2013; Obermeyer & Ezekierl, 2016; Ragupathi & Ragupathi, 2014; Weber et al., 2017). As previously stated, new solutions are required. In the future, whoever has access to the best data and best interpretive solutions will likely provide the best care.

Decision-Making Processes

Clinical decision making in health care including radiology is often challenging, and associated with complex nonlinear information (Hussain & Oestreicher, 2017). Available methods to simplify the process include the use of published guidelines, professional collaboration, crowd-sourcing, and computational decision support (Nendaz & Perrier, 2012; Phua & Tan, 2013). Many health care providers, including radiologists, are often confronted with decisions they are unprepared or unqualified to make. Radiologists like other health care providers tend to look for what they know, identify what they are familiar with, and render decisions based on experience. Variables such as the quantity and quality of data, experience, professional knowledge, incentives, and access to technological support influence human decisions (Gandomi & Haider; Weber & El-Kareh, 2017; H. W. Wu et al., 2012). Accurate and timely decisions require awareness of a well-defined problem, knowledge of options and alternatives, and the capacity to evaluate the problem (Croskerry, 2013; Kohn et al., 2014). In addition to addressing diagnostic variables, a radiologist's decisions must meet the standard of care and be consistent with a patient's needs, values, and expectations. The diagnosis offered by a radiologist after interpreting a set of images should be delivered in a manner that provides adequate decision support for the referring health care provider at the point of care.

Decision Support Solutions

Decision support refers to a process or technique used to help determine the right course or courses of action. The primary purpose of clinical decision support (CDS) in radiology is to avoid errors, improve diagnostic accuracy, and improve the quality of care delivered to the patient. A successful decision support system requires various attributes such as availability, ease-of-use, accuracy, and consistency (Kahn, 1994). Decisions made during image interpretation, the reporting process, and at the-point-of care (Raghupathi & Raghupathi, 2014). The rate of discovery and knowledge creation generally outpaces the individual health care provider's ability to keep up to date and make fully informed decisions (Kohn et al., 2014; Thrall et al., 2016). This phenomenon applies to data intensive specialties such as radiology; therefore, the radiologist must remain aware of decision support options.

Numerous solutions have been proposed to address complex decision making in radiology. One of the more recent solutions is the computerized decision support system (CDSS). CDSS solutions are placed into one of two categories; knowledge-based systems or non-knowledge-based systems (Stivaros et al., 2010). Knowledge-based systems contain programmed rules, an inference engine, and a well-defined communication mechanism. Non-knowledge-based systems often use machine-learning techniques or algorithms that learn from the ground up through the exposure, assimilation, and analysis of available data. Effective use of decision support in radiology will help deliver the right information, in the right format, at the right level of workflow to the right person. In summary, access to decision support can augment the role of the radiologist.

Heuristics

Heuristics refers to the application of rules or processes to simplify decision making. Radiologists often rely on heuristic methods such as mental shortcuts to minimize delay, reduce task complexity, and to simplify decisions (Itri & Patel, 2018; (Tversky & Kahneman, 1974). Heuristic methods are used to improve the accuracy, as well as the efficiency of the interpretive process in radiology. Inappropriate use of heuristics or the use of inaccurate heuristic methods can result in errors that adversely influence patient care.

The three principal categories of heuristics are anchoring, availability, and representative (Tversky & Kahneman, 1974). Anchoring heuristics refers to the limitation of further considerations due to perceived truth. In contrast, availability heuristics refers to the assignment of value based upon an individual’s memory or recall. This form of heuristics can result in errors secondary to limited or selective memory of prior events or outcomes. Representative heuristics is characterized by the use of categories to simplify data, information, and knowledge. This approach may be used to simplify a process leading to oversight and conjunction fallacy.

Dual Processes: Intuition and Analytics

The dual process theory of decision making consists of intuitive processing (type I) and analytic (type II) approaches (Croskerry, Petrie, Reily, & Tait, 2014). An intuitive response is characterized by a low cognitive demand and rapid application, whereas an analytic approach is characterized by a high cognitive demand, a slow process, and greater reliance on working memory. There are risks associated with isolated application of one decision-making process over another (Phua & Tan, 2013). Croskerry (2014) acknowledged the importance of discriminant use of both methods during complex decision making. The combined use of problem-solving methods increases the likelihood of a good decision and a good outcome. This perspective applies to human and machine-based approaches.

Sources of Interpretive and Diagnostic Errors in Radiology

Diagnostic errors are often underreported and underappreciated due to a lack of standards in defining, recognizing, and acknowledging their presence (C. S. Lee et al., 2012). It is difficult to estimate the impact of radiology errors due to the reasons mentioned and the limited capacity to measure their short and long-term impact on overall health (C. S. Lee et al., 2012). Errors can occur anywhere along the path of radiology workflow (Huassian & Oestreicher, 2017; Kassier & Kopeman, 1989). Errors that occur during the interpretive stage of radiology workflow are likely to have the greatest impact on the final diagnosis and report.

Health care providers too often make point of care decisions with irrelevant, incomplete or incorrect information (Kelly & Hamm, 2013). In addition, many physicians have limited experience with uncommon diseases or complex presentations associated with coexistent pathology (Manrai et al., 2014). This condition can influence the radiologist during image interpretation and can influence the referring physician who reads the report at the point of care. According to Latts (2016). it is impossible for a single health care provider to stay up-to-date in their field and to remain aware of all of the relevant data and variables associated with any particular disease process or state. This position supports the need for better decision support along the path of care.

Types of Errors

A medical error represents a deviation from a consensus opinion or a standard of care. Errors may occur secondary to missed presentations, oversights, or mistakes of judgment, all of which could lead to failure to implement a process or a plan of action (Andel et al., 2012; Makary & Daniel, 2016). One of the most common forms of error in radiology is diagnostic error, surfacing as a missed diagnosis, wrong diagnosis, or an untimely diagnosis. Diagnostic errors in radiology have generally been classified as perceptual or cognitive (Berlin, 1996, 2013). Several researchers have acknowledged perceptual error as the most common cause of diagnostic errors in radiology (Berlin, 2013; Bruno et al., 2015). In support of this premise, a large radiographic research study revealed that 80% of diagnostic errors were perceptual and 20% were interpretive

(Donald & Barnard, 2012). Causes of nondiagnostic errors in radiology include failure to recommend or perform an indicated test or test protocol and failure to address clinical concerns or reported patient presentations on imaging test requisitions.

Y. W. Kim and Mansfield (2014) proposed one of the most widely accepted classification of errors in health care. The approach represented an expansion of classifications previously proposed by Renfrew (1992), a few decades earlier. According to Kim and Mansfield (2014) common causes of error in radiology include faulty reasoning, lack of knowledge, satisfaction of search bias, miscommunication, and an acquired inaccurate or incomplete history. Cognitive based-errors include limited skills, stress, bias, faulty heuristics, memory loss, and inattention (Zhang et al., 2004). All types of errors are amplified in the presence of large volumes of complex data. In addition to the reasons given, diagnostic errors may occur as the result of technological limitations, restricted access to data, and system failure (Berlin, 2013; Thammasitboon et al., 2013). Errors in radiology may also occur secondary to cognitive bias rather than the result of a perceptual error or lack of knowledge (Hussain & Oestreicher, 2017; Nendaz & Perrier, 2012). The cause of interpretive error in radiology is often the result of more than one factor.

Expanding domain knowledge combined with human variables such as limited time and preoccupation with the care of complex patients, contributes to a high incidence of diagnostic errors within any specialty field (Weber & El-Kareh, 2017). Radiologists are more likely to make wrong decisions in the presence of signs, symptoms or conditions, which they have, limited experience with or knowledge of (Weber & El-

Kareh, 2017). Overwhelming data combined with limited knowledge and competing environmental pressures contribute to chaos and confusion, which can lead to oversights and diagnostic errors (Gupta et al., 2017; Saber-Tehrani et al., 2013). Access to computational decision support such AI increases the potential for problem solving in the presence of human limitations, as well as cultural and environmental pressures.

Human Limitations

Human decisions made in the presence of high degrees of complexity and uncertainty increase the likelihood of error (Stivaros et al., 2010). The ability to process information is limited by physiological mental capacity, a phenomena often referred to as the cognitive threshold. Radiologists interpret images using a variety of cognitive methods such as visual detection, pattern recognition, memory, and reasoning. A radiologist's cognitive performance is influenced by personal attributes, physiology, and professional skills. It is also influenced by environmental factors such as ambient noise, workload intensity, and workflow distractions. Stress influences interpretive accuracy. Stress may be associated with uneven work distribution, poor reimbursement, limited time, increasing liability pressures, along with heightened awareness of the complexity and heterogeneity of pathology.

Humans are flawed in their capacity to process large volumes of multidimensional or deeply nested data (Hatt et al., 2017; Wolf et al., 2015). The human brain is also limited in its capacity to perform highly scalable functions that involve voluminous or unrecognized confounding variables (Wolf et al., 2015; H. W. Wu et al., 2012). Humans are limited in their capacity to perform accurate complex data analysis in a relatively short period of time. Factors that contribute to these limitations include physiological fatigue, cognitive bias, limited knowledge, and distractibility (H. W. Wu et al., 2012). A conscious or unconscious response to self-limitations may result in the use of heuristic methods that introduce bias to a decision-making process.

Increased complexity in a work environment increases the likelihood of human error compounded by deficiencies of the system (Institute of Medicine, 2000). For this reason, health care specialists responsible for analyzing large quantities of complex data such as radiologists are often exposed to high cognitive demands and subsequently high rates of diagnostic error (Crosskery 2013; Nendaz & Perrier, 2012; Obermeyer & Ezekierl, 2016; Ragupathi & Ragupathi, 2014; Weber et al., 2017). Technological solutions can be implemented to reduce and simplify the differential diagnostic process by performing pre-analytic functions prior to human interpretation of images.

Inconsistency and Variability

Conventional imaging interpretation and reporting methods are highly subjective; and subsequently, associated with a high degree of variability (Bosmans, Weyler, DeSchepper, & Parizel, 2011; Bruno et al., 2015). Most radiology reports primarily consist of subjective narrative descriptions of normal and abnormal findings (J. Y. Chen et al., 2017). Moreover, radiologists vary in their use of interpretive descriptors and reporting structures (Napel & Giger, 2015). This personalized approach to reporting contributes to inconsistencies within and between radiology reports. In support of this premise, radiology reports have been described by numerous researchers as incomplete, inconsistent, and inconclusive (Bosmans et al., 2011; J. Y. Chen, Sippel-Schmidt, Carr, &

Kahn, 2017). In one particular example, variability of the interpretation of spine imaging by different radiologists was reflected by a misinterpretation rate of 43.6% plus or -11.7 (Herzog et al., 2017). The study was based on the interpretation of 10 MRI studies performed on the same patient at 10 MRI centers, read by 10 independent radiologists. The degree of inconsistency and variability occurring during the interpretive stage of radiology workflow in all fields including spine care must be improved.

In addition to variability of the reporting process, there is also a high level of data management and data access variability during interpretive workflow (Aerts et. al., 2013). Factors, which adversely influence the flow of data and the pattern of data access during radiology workflow, include incomplete access to medical records and prior imaging reports, inadequate imaging protocols, human error, and technical workstation deficiencies. Image interpretation and the description of pathology often vary between radiologists**.** The degree of interpretive variability is influenced by a radiologist's level of experience, the time allowed for the interpretive process, the complexity of the study, and the presence of human bias (Napel & Giger, 2015). Interpretive variability is also influenced by reporting requirements. AI has the potential to improve patient care by improving the flow of data and access to data across the spectrum of radiology workflow (Augimeri et al., 2016; Boone et al., 2015). AI can also influence how radiology reports are structured.

Cognitive Bias

Cognitive bias represents an error in reasoning. Over 100 different forms of cognitive bias have been identified in health care (Croskerry, 2017). Common forms of bias in radiology include availability bias, alliterative bias, anchoring bias, framing bias, satisfaction of search bias, and pro-innovation bias. Availability bias occurs when a decision is influenced by experiences, whereas alliterative bias occurs when an individual's judgment is influenced by another. Anchoring bias refers to limiting the search for additional possibilities due to the belief that a prior assumption is correct or that a current diagnosis fully explains a patients presentation (Tversky & Kahneman, 1974). Framing bias refers to the use of a limited perspective. Satisfaction of search bias refers to the assumption that the diagnostic process is complete due to lack of knowledge of other differential diagnostic possibilities. Pro-innovation bias refers to assigning value to the role of new technology or the data it provides without consideration for potential inaccuracies or inconsistencies (Bauman & Martigoni, 2012; Rogers, 2003). Radiologists may not be aware of their own cognitive bias under different circumstances. Machine-based decision support is not subject to most forms of human bias, and therefore can be used to help avoid or overcome adverse consequences of human bias.

Research has revealed that radiologists, like experts in other fields, are subject to a phenomenon referred to as inattentional blindness, characterized by missing what should be obvious due to search bias (Drew et al., 2013; Memmert, 2006). For example, Drew et al., (2013**)** revealed that 83% of radiologists asked to review computed tomography lung scans for nodules and other abnormalities failed to identify a gorilla image located in the lung field. The gorilla image was more than 40 times larger than the average nodule. This acclaimed research study confirmed that a prioritized search for specific pathology could lead to blinding of other significant findings.

Cognitive bias can influence data acquisition, data analysis, and data interpretation during the course of radiology workflow. The primary factors that influence cognitive bias include poor training, lack of experience, stress, uncertainty, physiological fatigue, and incomplete information (H. W. Wu et al., 2012). Different forms of bias may overlap or coexist within the same decision-making process. For example, anchoring bias may be amplified by confirmation bias leading to premature diagnostic closure (Hussain & Oestreicher, 2017). The use of detrimental heuristics may complicate cognitive bias and result in higher risk for diagnostic error.

Process and Workflow Error

Any situation that disrupts or interrupts the interpretive process during radiology workflow can lead to human distraction and diagnostic error. Examples include slow data access, difficulty accessing prior imaging studies or records, and lack of familiarity with complicated workstation technology. Additional distractions include phone calls, interventional procedures, and conversations with health care providers (Schemmel et al., 2016). Numerous disruptive factors commonly occur simultaneously or within a short time frame in the radiology setting.

The absence of technological decision support in the presence of complex data can result in an inefficient, inaccurate, and untimely diagnostic process. Potential solutions include the use of embedded AI decision support, access to integrated radiology and pathology data, physical workflow modification, and simplified workstation interfaces (Bruno et al., 2015; C. S. Lee et al., 2012). It is important to embed technological solutions within radiology workflow to simplify the interpretive process and to reduce the risk for diagnostic errors.

Artificial Intelligence: An Introduction

Artificial intelligence (AI) refers to technology that exhibits biological-like properties to assist, augment or replace human processes or actions (Pannu, 2015). The Turing test developed decades ago offered an operational definition of AI, which required that a machine possess certain human-like attributes and capabilities that could be used for problem solving (Turing, 1950). The two principal forms of AI are machine learning and natural language processing (Jiang et al., 2017). Natural language processing (NLP) is used for textual analysis, whereas the use of machine learning (ML) in radiology supports computer-aided disease detection, characterization, and monitoring. ML can be used to assist in the differential diagnostic process.

The three primary forms of AI are assisted intelligence, augmented intelligence, and autonomous intelligence (Bothum & Lancefield, 2017). Rapid advances in computer technology, software programming, and algorithm development have accelerated the evolution of AI in the direction of autonomy for some tasks. Assisted intelligence refers to the use of technology to improve a process a human is capable of performing. In contrast, augmented intelligence refers to the use of technology to enhance human potential. Autonomous intelligence refers to the use of technology to perform a task that exceeds human capability. One of the most important attributes of AI is speed. It can perform most tasks much faster than humans can. Some AI solutions will evolve from offering assistance to becoming autonomous. The primary elements of an intelligent system include infrastructure, algorithms, data, software, and an ecosystem.

The broad topic of AI encompasses different computational methods such as machine learning (ML), deep learning (DL), and cognitive computing (CC). The numerous subcategories of AI, each of have different potential applications and response characteristics (Figure 1). The basic elements of expert machine systems include a knowledge base, an inference engine, and established rules operationalized by algorithms (Salem, 2017). Algorithms represent digital rules used to perform an automated task or operation. DL algorithms have many applications in radiology. For example, they can be used to reveal new features of disease, not previously identified. Conventional ML algorithms are linear, whereas DL and CC algorithms are more abstract, characterized by a hierarchy of increasing complexity. Leading radiology companies have adopted different terminologies for their AI solutions. For example, Phillips refers to its AI solution as adaptive intelligence, General Electric refers to its solution as applied intelligence, and IBM refers to its solution as cognitive computing (Freiherr, 2018). Despite the use of different terms, the goal is to use technology to enhance or replace human performance for improving a process and/or an outcome.

Traditional computer programming requires the use of explicit rules to perform tasks. In contrast, ML uses statistical techniques and algorithms that do not require explicit rules (Cai et al., 2016). The processes associated with ML performance can be classified into three primary categories, which are supervised learning, unsupervised learning, and semi-supervised learning. Supervised learning refers to the use of expert derived handcrafted rules that serve as map for the flow of data between input, output, and ground truth. Unsupervised learning refers to the use of one or more algorithms designed to reveal patterns within data without a priori rules or human intervention. Semi-supervised learning represents a combination of both approaches. Supervised learning is often used to train a model to make a prediction, whereas unsupervised learning methods are often used to explore data without a preconceived determination. Fundamental ML data analysis methods include classification, regression, clustering, pattern matching, density estimation, and dimensionality reduction (Kotsiantis, 2007). ML systems can be used in radiology to detect patterns, aggregate data, and classify disease features.

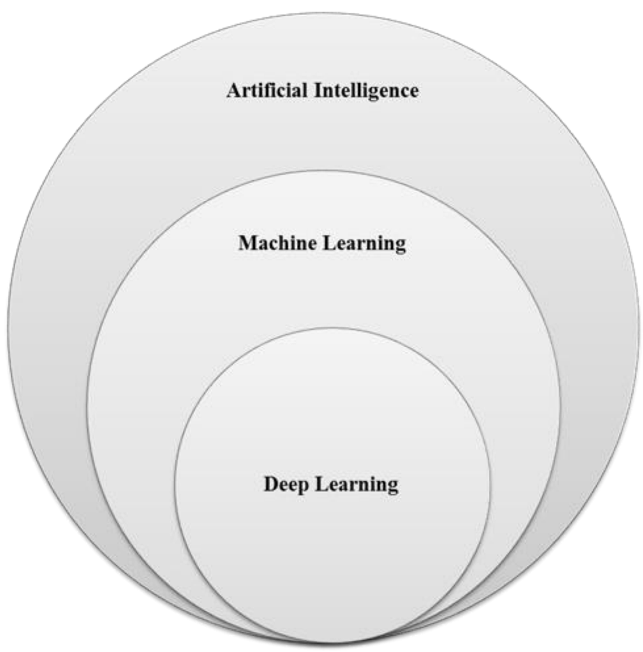


Figure 1. There are different forms and applications of artificial intelligence. Each subtype in the above figure is less dependent on handcrafted rules and more capable of detecting patterns in complex data and learning with exposure to data.

Deep learning (DL) represents a subset of machine learning, modeled after neurological signal transmission in the nervous system (Zaharchuk et al., 2018). One of the most common forms of DL is referred to as convolution neural networks (CNN), characterized by an advanced network of connectivity capable of performing complex parallel and serial processing (Cascianelli et al., 2017; Dreyer & Geis, 2017; Zaharchuk et al., 2018). DL methods are able to detect patterns within high-dimensional data sets using layered processes that impart logic (Lakhani & Sundaram, 2017). Neural networks have generally outperformed individual algorithms in the analysis of complex data (Lerner et al., 2018). DL solutions tend to become more accurate with exposure to data due to their ability to detect patterns and create supportive algorithms (Zaharchuk et al., 2018). DL methods have been successfully used in different health care fields including neuroradiology. For example, Gao et al., (2017) used CNN for the automatic classification of 285 non-contrast brain CT examinations into one of three categories of pathology. The autonomous capacity of DL systems to learn and to construct pattern detection algorithms renders them capable of detecting previously unrecognized patterns within complex imaging datasets.

Algorithms represent a programmed set of rules delivered through a sequence of operations to manipulate and analyze data to achieve a desired outcome (Obermeyer & Emanuel, 2016). They may be handcrafted or created by a computer. Algorithms essentially represent mathematical formulas that are used as instructions for a digital process (Mapoka, Masebu, & Zuva, 2013). They are used to detect patterns, sample variables, register instances, highlight structures, and create reference maps. Algorithms may be deterministic, logical, or recursive. Their success is dependent upon their relevance, accuracy, and speed. Algorithms render computer systems capable of augmenting the role of the radiologist.

The most common types of algorithms used in radiology are k-nearest neighbors, convolutional neural networks, fuzzy logic, support vector machines, decision trees, and Naïve Bayes algorithms (Adduru et al., 2017; Erickson, Korfiatis, Akkus, & Kline, 2017; Fan, Lin, & Tang, 2017). Algorithmic outcomes can be expressed in many forms, which include statistics, natural language, flowcharts, diagrams, and a list of probability-based differential diagnostic possibilities. The handcrafted algorithm sometimes serves as a computational bridge between human and machine processes to transform complex data into practical information (Beam & Kohane, 2017; Obermeyer & Ezekiel, 2016). Algorithms can be used to sort through the millions of variables and patterns within advanced imaging datasets. They can also be used to correlate structured and unstructured data within radiology workflow.

AI uses two primary types of algorithms during computational analysis. These algorithms are referred to as classifiers and controllers. Classifiers are used for pattern detection and pattern matching, whereas controllers are used to assign an action or task to computational outcomes. Special digital filters combined with algorithmic functions have been used to mine and interrogate data (Obermeyer & Emanuel, 2016). Algorithms are not subject to cognitive bias but have the potential to introduce machine bias during their creation and evolution. Machine bias can lead to false negative or false positive outcomes.

Introduction to IBM Watson: A Cognitive Computing System

IBM Watson is a cognitive computing system that was introduced in 2015 to offer solutions for problem solving and decision making in the presence of complexity and uncertainty (Hoyt, Snider, Thompson, & Mantravadi, 2016). The system is capable of generating “ descriptive, predictive, and visual analytics, thus, reducing the risk for error” (Hoyt et al., 2016, p. e165.). Cognitive computing is capable of analyzing structured and unstructured data; thereby, increasing its potential utility in radiology.

IBM Watson represents the first large-scale integrated clinical diagnostic support system (CDSS) capable of analyzing structured and unstructured data, and rendering a differential diagnostic list based on pattern detection and probabilistic calculations. The system uses a series of interdependent computational methods to respond to inquiries. The process generally begins with a question followed by hypothesis development, evidence analysis, and probability assignment (Deloitte, 2015; Ferrucci, 2012). With further development, IBM Watson may become capable of augmenting the role of the radiologist during image interpretation (Brink et al., 2017; Jha & Topol, 2016; Kharat & Singhal, 2017; Jiang et al., 2017; Ranschaert, 2016). In summary, IBM Watson has the potential to auto detect abnormalities, filter irrelevant or normal imaging data, characterize disease, and assist in the differential diagnostic process with probability assignment. Watson is capable of offering data and knowledge-driven decision support that can be used to assist, augment or in some cases replace the role of the radiologist (Kohn et al., 2014). Furthermore, IBM Watson is a pioneering solution capable of integrating ML and NLP applications during the interpretive stage of radiology workflow

(Jiang et al., 2017). IBM technology simply serves as an example of what AI is capable of achieving in health care. Many products will be developed to perform the same or similar tasks. It is too early to tell which technologies will meet the reproducibility and validation demands of future research and regulatory requirements.

Successful adoption and use of AI solutions such as IBM Watson in radiology requires validation studies, heightened awareness of its potential value, and an adequate state of readiness. AI solutions have the potential to reveal biological variability and the features of pathology in vivo in a manner which exceeds human capabilities (Gillies et al., 2016; Limkin et al., 2017; Pannu, 2015; Yip & Aerts, 2016). AI solutions also have the potential to calculate and assign probability to differential diagnostic possibilities based on analysis of structured and unstructured data. For these reasons, AI solutions should be developed to improve the precision and personalization of the diagnostic process.

Collective Intelligence: A Collaborative Approach

Data can be analyzed and health care decisions can be made using collective intelligence (CI), a process referring to the integration of human expertise and computer analytics. The capabilities of CI include reducing the complexity of data, revealing patterns within the data, assigning value to classifications of data, and offering decision support (Hoyt, Snider, Thompson, & Mantravadi, 2016). The future impact of human-machine collaboration is difficult to predict because human performance tends to improve in a linear and relatively predictable fashion, whereas the capabilities of AI grow exponentially. Advanced AI systems can learn from mistakes and successes and are therefore less likely to repeat mistakes than humans are. Humans add an element of creativity and intuition to AI output.

Unlike computers, humans have a limited capacity to analyze complex data. Humans also have difficulty correlating multidimensional nonlinear variables, performing syntactic transformations, and revealing patterns within variable high velocity data (Nendaz & Perrier, 2012). Humans are prone to physiologic limitations and fatigue, whereas computational technology is stable and consistent (El-Kareh et al., 2013; Russel & Norvig, 2010). Human intelligence is highly dependent on experience, analytic skills, intuition, and motivations (El-Kareh et al., 2013). In contrast, AI systems are highly dependent upon access to annotated training data, ground truth, and validation methods (Russel & Norvig, 2010). Human intelligence has many characteristics not offered by AI such as adaptability, intuition, creativity, flexibility, and the ability to plan.

AI systems are not capable of replacing the full breadth and depth of human reasoning and judgment. For example, the human role in radiology offers the benefits of unique experiential insight, empathy, and thoughtful decisions. The combination of human and artificial intelligence offers a collaborative approach with a greater chance of success then either isolated approach in some settings. A collaborative relationship supports what the radiologist does best and combines it with what AI does best.

Radiology Workflow Defined

The radiology workflow environment is complex. Each stage of radiology workflow has unique elements or processes (Figure 2). Radiology workflow has been divided into two primary stages referred to as interpretive workflow and non-interpretive workflow (Lee et al., 2017; Schemmel et al., (2016). I divided the non-interpretive stage into two subsets referred to as the pre-interpretive and post-interpretive stages. AI is poised to play a significant role within all stages of radiology workflow, although its greatest potential is likely within the interpretive stage, the focus of this research study. AI solutions can be used during the interpretive stage of radiology workflow to detect, characterize, classify, and monitor disease. Reasoning during the diagnostic process will involve various types of diagnostic inference and decision support. AI support during the interpretive stage of radiology workflow can also be used to provide a probability-based differential diagnostic list for the attending radiologist.

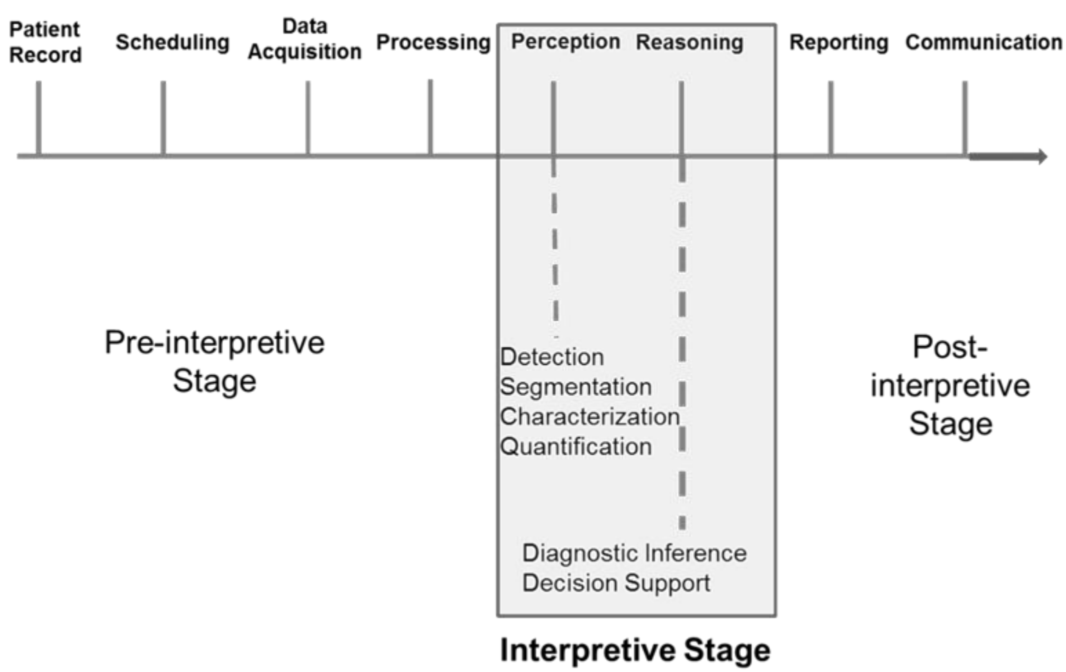


Figure 2. The three stages of radiology workflow that are pre-interpretive stage, interpretive stage and post interpretive stage. The figure also highlights elements of the interpretive stage of workflow, which can benefit from AI support.

Many radiologists spend a greater amount of time performing non-interpretive tasks than interpretive tasks during the course of radiology workflow (Dhanoa et al., 2013). This pattern of labor contributes to inefficient and occasionally inaccurate interpretive outcomes. Limited time combined with exposure to large volumes of complex data exposes the radiologist to a higher degree of uncertainty and subsequently to greater potential for the influence of human bias and error during the interpretive stage of radiology workflow. There are numerous forms of human bias, which can take place during the interpretive stage of radiology workflow (Figure 3). Quite often more than one form of bias will be present. AI-based decision support can reduce the impact of human bias.

Non-interpretive human tasks during radiology workflow include setting image protocols, supervising studies, directly caring for patients, accessing analytic tools, performing image-guided intervention, and consulting with health care providers. Human tasks typically performed during the interpretive and post-interpretive stages of radiology workflow include visual evaluation of images and qualitative report generation. The use of AI methods during the interpretive stage of radiology workflow could help reduce the incidence and prevalence of interpretive and diagnostic errors (Busby, Coutier, & Glastonbury, 2018; Itri & Patel, 2018). Radiomic methods are increasingly becoming a more important quantitative measure in diagnostic imaging, one that may change the landscape of health care. The success of AI use in radiology will be dependent on its ease-of-use, utility, and ability to be integrated into existing workflow.

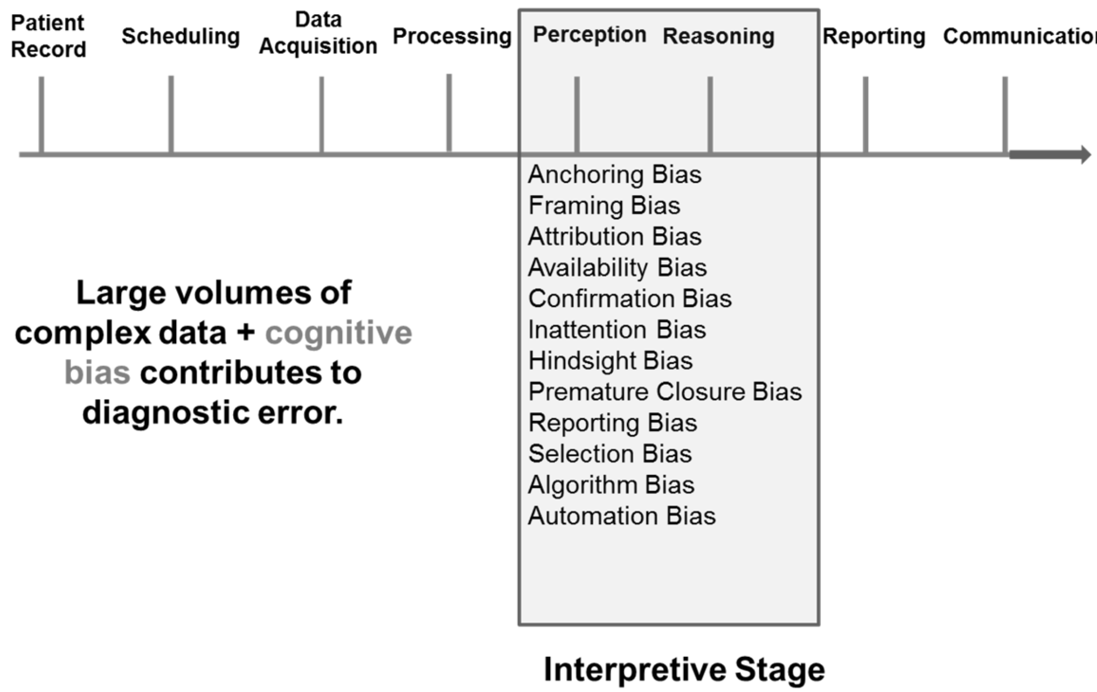


Figure 3. Types of human bias that can occur during the interpretive stage of radiology workflow.

Emerging Options for Interpretive Radiology Workflow

The current tasks performed during the interpretive stage of radiology workflow must be adapted or replaced to support a more accurate and timely diagnostic process in the presence of complex and high velocity data. The potential for reducing the incidence of diagnostic errors in radiology must be addressed at all stages of radiology workflow, including the final report, to have maximum impact on patient care (Bruno et al., 2015). AI has the potential to assist and augment the radiologist during interpretation through its capacity to access, analyze, and correlate data acquired from numerous sources (Dreyer & Geis, 2017). This includes prior imaging studies, electronic health records, laboratory studies, genetic profiles, published literature, and disease registries. AI can be used to pre-analyze data, flag abnormal presentations, and subsequently prioritize the order of what needs to be interpreted by the radiologist (Kahn, 2017). A single health care provider such as a radiologist or pathologist cannot keep track of all relevant information on a particular patient or on a disease process. AI solutions can help overcome human limitations such as lack of experience or unawareness of possibilities.

Radiology workstations of the future will be required to meet new standards based upon performance, reliability, and scalability. Effective performance will require easy access to an AI menu of options, technological interoperability, and seamless networking with external resources (Kahn, 2017). Adaptive workstations will be used by radiologists to leverage the experience of peers from remote locations and to democratize expert decision support with the help of AI. It will also be used by the radiologist to acquire knowledge-based assistance from databases and published resources almost simultaneously. In summary, for radiology workflow to be successful it must support the role of the radiologist and lead to better patient care.

Natural Language Processing

The science of linguistics encompasses the form, meaning, and context of sounds and symbols used to communicate (Zipf, 2012). Natural language processing (NLP) represents a solution for linguistic analysis. It can be used to assign meaning or value to unstructured (textual) data allowing it to be analyzed with computational methods (Pons, Braun, Hunick, & Kors, 2016). The majority of digitized health care data within a patient's medical records and within the realm of population databases is unstructured. The presence of pathology often reported semantically in radiology reports (Acharya et al., 2017). NLP systems can access prior pathology and radiology reports to reveal prior evidence and characteristics of pathology. NLP can therefore be used to leverage information within electronic medical record systems to generate an active problem list. It can also be used to provide access to prior imaging study findings for comparative analysis and to provide decision support during radiology workflow (Massat, 2018). Biomedical text mining with NLP leads to knowledge discovery that can aid imaging interpretation.

NLP uses different methods to analyze unstructured data. The fundamental mechanisms include pattern matching, parsing, and statistical approaches (Cai et al., 2016). Hassanpour and Langlotz (2016) demonstrated that ML combined with NLP could be effectively used to analyze textual data acquired from medical records including prior radiology reports to assist in the differential diagnostic process. NLP represents an important data management tool within the interpretive stage of radiology workflow. AI-supported NLP can only reveal what was previously described, not the current state of pathology evaluated through imaging. Characterization of an individual's current state of pathology requires in vivo imaging and quantitative measures using various techniques such as radiomics. The combined use of NLP and radiomic methods will improve the radiologists' access to relevant information.

Radiomics and Radiogenomics

Radiomics refers to the science associated with high-throughput extraction and quantitative analysis of non-visible data acquired from images to characterize pathology (Lambin et al., 2017; Lee et al., 2017; Limkin et al., 2017; Griethuysen et al., 2017; Yip & Aerts, 2017). Radiomics, first reported in 2012, has evolved quickly during the last couple of years (Verma et al., 2017). Automated radiomic functions include pathology detection, segmentation, feature extraction, and feature analysis **(**Lambin et al., 2017; Limkin et al., 2017; Zaharchuk et al., 2018)**.** Radiomic methods can be used to expand available data within three-dimensional space on imaging studies to characterize pathology and subsequently refine the diagnostic process (Gillies et al., 2015; D. Kumar et al., 2012; Lee et al., 2017; Lambin et al., 2017; Limkin et al., 2017; Peeken et al., 2018). Expanded dimensionality amplifies spatial heterogeneity (Cook et al., 2014). One of the goals in radiomics is to identify disease features and biomarkers that have greater causal rather than correlative relationships (Sanduleanu et al., 2018). Radiomic methods have the potential to expand sub classifications of disease and support more personalized care.

Radiomic measures can be used to quantify characteristics of pathology that are not visible to the radiologist. This form of analysis addresses features of pathology such as shape, volume, edge characteristics, texture, and other statistical measures (Aerts, 2016; Court et al., 2016; Gillies et al., 2016; V. Kumar et al., 2012; Lambin et al., 2012). Radiomics is a discovery process rather than a validation process performed using one of two approaches. The first approach focuses on mining images for predetermined patterns of disease. The second approach uses deep learning methods to discover and learn disease features not currently known. Radiomic methods like other analytic approaches tested for reliability and validity prior to clinical use. Testing should include evaluation of sensitivity and specificity along with the reliability of predicting positive and negative values. The testing must be disease-specific and take into account different cohorts. The primary goals of radiomic development are to help detect, characterize, and monitor pathology. This includes classifying and staging disease.

Radiomic methods have been successfully used to detect and characterize some diseases using different imaging methods such as CT, MRI, and PET (Acharya et al., 2018; Parekh & Jacobs, 2016). Radiomics has been successfully used in different research settings to evaluate breast cancer (Li et al., 2016), brain tumors (Li et al., 2017), lung disease (Bak et al., 2018; Vallieres et al., 2015), liver disease (Naganawa et al., 2018), brain metastasis (Ortiz-Ramon et al., 2017), and prostate cancer (Tanadin-Lang, 2018). Most of the important contributions to radiomics have come from the field of oncology (Gillies et al., 2016). Radiomic methods show promise in the field of breast imaging for the differentiation of benign versus malignant tumors (Hui et al., 2016). Specific radiomic features such as enhancing tumor volume and texture features are emerging as discriminatory factors in the differential diagnostic workup of breast cancer (Drukker et al., 2018; Hui et al., 2016). Ongoing research is required to identify the potential applications and benefits of radiomics in different fields. Future research will reveal whether some of the methods used in other fields may be adapted and used in spine care.

Radiomic methods are not limited to the initial diagnostic process. They can also be used to help monitor disease progression and the response of pathology to treatment (Vargas et al., 2017; Zhou et al., 2017). Radiomic methods can be used for the surveillance of early-stage pathology prior to intervention. In addition, radiomic signatures can ”…be used as precision biomarkers for the prognosis of individual patients" (Castiglioni & Gilardi, 2018, p.412). The clinical role of radiomics is based on the premise that imaging of disease requires a large volume of data that reflects multi-scale pathological mechanisms, not detectable through routine visual assessment of images.

The specialized field of radiogenomics sometimes referred to as imaging genomics refers to the science associated with the correlation of anatomic characteristics (phenotype) of pathology with genetic (genotype) data (Pinker et al., 2017). Radiogenomic methods have been successfully used in the evaluation of lung cancer, glioblastoma multiforme, kidney cancer, prostate cancer, and liver cancer (Incoronato et al., 2018). Radiomic measures have proven useful for revealing phenotypic manifestations of genetic expression (Giger, 2018; Panth et al., 2015). Phenotypic characterization is important for there are many non-genetic determinants in pathology, especially involving age-related diseases (Oakden-Rayner et al., 2017). Radiogenomics has the potential to support precise classification of disease and subsequently the personalized delivery of care (Bai et al., 2016). This conclusion is based on the premise that alterations of genetic expression influence pathology represented by phenotypes revealed through diagnostic imaging methods.

Volumetrics: New Realities and 3D Perspectives

Radiologists have a unique opportunity to combine the use of AI and volumetric datasets to display pathology in three dimensions (3D) to help inform the diagnostic process and therapeutic planning (Farahani et al., 2017). A few years ago, Denzel et al. (2014) acknowledged the importance of in vivo interrogation of pathology within the framework of volumetric imaging and display. A study limited to planar or 2-D imaging can limit the diagnostic process.

Advanced imaging methods such as MRI and CT are capable of acquiring high-resolution volumetric data sets that can be formatted to create 3D perspectives (Douglas et al., 2016). The use of augmented reality and 3D viewing can improve the conspicuity of pathology features (Douglas et al., 2016; Hamacher et al., 2016). Segmented pathology can be manipulated and interrogated using virtual cut plane technology. Multidimensional imaging data sets can be used to create virtual reality (VR) and augmented reality (AR) based representations of pathology. Virtual reality offers an immersive experience that can be used for planning and training. In contrast, augmented reality offers digital images or prompts in the physical world. AR can be used to project nonvisible perspectives of pathology during an exploratory biopsy or over a surgical field, thus, limiting attention shifts between available imaging and the patient. The augmented virtual or interactive display of pathology may soon be used to guide the digital (virtual) biopsy and acquisition of radiomic measures.

Disease Modeling and Computational Diagnosis

Mathematical formulas can be used to help study normal biological and disease processes (Mapoka et al., 2013). In silico modeling sometimes referred to as ex vivo modeling refers to the use of quantitative imaging data to create comprehensive mathematical models which can be used within the context of a computer system to study the characteristics and behavior of pathology (Jeanquartier, Jean-Quartier, Cemernek, & Holzinger, 2016; Louis et al., 2016; Mapoka et al., 2013). Mathematical models support the use of simulated variables to predict the behavior of disease under different circumstances (Louis et al., 2016). Radiomic features acquired in vivo can be used to help develop in silico models (Echegaray et al., 2016). The use of mathematical disease models built from data acquired over time through monitoring of a disease process could help support predictive modeling.

The benefits associated with in silico modeling include the ability to study disease, zoom in on pathology subsystems, and manipulate time scales, all of which reduce the need for laborious and costly biological experimentation (Jeanquartier, et al., 2016). Computational disease models can be used to test assumptions and hypotheses. They can also be used to help predict outcomes and to reduce the frequency and magnitude of uncertainties surrounding disease behaviors (Mapoka et al., 2013). Access to mathematical disease models from the radiology workstation can augment the differential diagnostic process. In the future, mathematical models will be used to help identify ground truth and to train AI systems.

Computational diagnostics is the science associated with the combined use of algorithms, mathematical formulas, and computer systems to detect and study disease. Computational pathology refers to the use of a digital disease model to evaluate assumptions and to study disease behavior under simulated circumstances. The convergence of interoperable technologies offers an unprecedented opportunity to correlate pathomic, radiomic, and genomic features to create computational disease models. In the near future, computational analysis may become as important to radiology as the microscope is to pathology (Louis et al., 2016). Access to AI supported computational diagnostic tools during the interpretive stage of radiology workflow will enable radiologists to address the complexities of pathology.

Workflow Optimization and Data Integration

The differential diagnostic process in radiology typically requires correlation of imaging findings with data from other sources such as prior imaging studies, medical records, peer-reviewed literature, and disease registries. Radiology workstations will eventually consist of embedded AI solutions and networked functions to perform interpretive functions, to obtain second opinions, and to support multidimensional viewing (Morgan, Branstetter, Mates, & Chang, 2006). Radiomic workflow includes data acquisition, pathology segmentation, feature extraction, feature analysis, and correlation with computational disease models. The radiologist will serve as the expert on the in vivo diagnostic process and as a gatekeeper to the flow of data and access to knowledge from diverse data sources.

The radiology workstation of the near future will likely represent the principle hub of converging data from the fields of radiology, pathology, and genetics (Gillies et al., 2016; Jha & Topol, 2016). Computational methods will be required to analyze data from these diverse sources. Unlike the radiologist, AI solutions do not represent a single layer of interpretation. It represents multilayered approaches, consisting of logical analysis and problem solving, not subject to the variability and inconsistencies of a biological system.

Data acquired from nonimaging sources need to be annotated or tagged in a manner that prepares it for sorting, aggregation, and analysis. Complex data often has to be conditioned and prepared for exposure to algorithms and computational analysis. The curating process requires well-defined steps to identify inconsistencies, to standardize symbolic representations, and to transform data to a uniform interoperable and interpretable language. This process can be augmented through the identification of common data elements (CDEs). CDEs help integrate disparate clinical, phenotypic, and genotypic data. The consistent use of CDEs supports the development and use of disease registries (Rubin & Kahn, 2017). Widespread application of CDEs will support the development of integrated knowledge databases accessible by AI for use during the interpretive stage of imaging. Curated data sets are required to train AI.

Access to Disease Databases and Registries

The increasing ability to extract pathology features during imaging procedures will result in new forms of data storage and disease registries (Rastegar-Mojarad et al., 2017). Digital tissue samples can be stored in a manner that will maintain accurate spatial and relational data. Unlike in vitro tissue samples, in vivo digital representation of pathology will not degrade. Digital representation of in vivo characteristics can be acquired and stored in a manner that maintains three-dimensional relationships. The radiologist will have increasing access to digital pathology registries and mathematical disease model libraries from their workstation to augment the differential diagnostic process.

Structured Reporting

The radiology report reflects what took place during the interpretive stage of radiology workflow. The decisions made influence the content and structure of the final report. Structured reporting refers to the use of a standardized format, quantitative measures, and consistent descriptions. Structured radiology reports can be operationalized with quantitative measures derived from AI applications (J. Y. Chen et al., 2017; Dreyer & Geis, 2017). Growing use of widely accepted terminology and standardized disease nomenclature will support more consistent diagnostic reporting and reliable NLP applications. Using keywords and phrases on reports can serve as common data elements used to trigger prospective or retrospective interpretive or comparative processes. For example, describing focal pathology in a prior radiology report may direct auto detection and quantitative radiomic measures within the same volume of interest in a subsequent study. Structured reporting provides more concise decision support at the point of care (Alkasab et al., 2017). According to Schwartz et al. (2011), "Referring clinicians and radiologists found that highly structured reports had better content and greater clarity than conventional reports" (p. 174).

In the near future, AI will offer a full spectrum of solutions that will begin with pathology detection and end with automated reporting within a structured format (Zaharchuk et al., 2018). The combined application of ML techniques and NLP to extract patterns from current imaging studies and prior radiology reports will improve imaging interpretive accuracy (J. Y. Chen et al., 2017). Text clues within a prior radiology report can be used to trigger access to current imaging and prior non-imaging data to support the interpretive process.

Digital Exploration: The Virtual Biopsy

Growing use of AI combined with radiomic methods will expose new features of disease and expand the spectrum of pathology. This process will lead to a longer list of disease subtypes and differential diagnostic possibilities during the interpretative stage of imaging, thereby, increasing the complexity of decision making. Advances in technology will support the exploration and interrogation of in vivo pathology with visual, augmented, and subvisual techniques. Augmented visual approaches include the use of interactive 3-D displays of pathology.

Subvisual Tissue Interrogation

Radiomic methods can be used to detect and characterize features of pathology from imaging studies in a manner undetectable by traditional visual interpretation (Aerts, 2017; Gillies et al., 2016). Advanced algorithms can be used to correlate subvisual pathology features with other sources of data such as histological features from pathology and genetic profiles (Bucking et al., 2017). Echegaray et al. (2016) introduced the concept of the “digital biopsy” which refers to the targeted non-invasive acquisition of pathology features in vivo” (p. 283). In another field, Mancini et al. (2018) introduced the concept of the “digital liver biopsy” referring to the use of multimodality and multiparametric imaging of the liver (p. 3). The use of a digital or virtual biopsy approach has the potential to interrogate and map the entire landscape or volume of a pathological state, whereas the traditional needle biopsy offers limited characterization of pathology (Echegaray et al., 2016; Lambin et al., 2012; Thrall et al., 2016). It is important to evaluate and characterize a whole region of pathology to reduce or eliminate sampling bias. Under sampling may poorly represent the entire state or spectrum of pathology present and subsequently misdirect treatment. Traditional and virtual biopsy results can be combined to better characterize pathology. The virtual biopsy can be also used to augment or direct the traditional needle biopsy.

Radiomics is not limited to the detection and extraction of pathology features. It can be used to create or reveal new quantitative descriptors of pathology. Subsequently, radiomic methods will continue to contribute to the development of new molecular biomarkers and imaging signatures of pathology (J. Wu et al., 2018). For this reason, radiomic methods are being considered for the expansion of the traditional tumor-node-metastasis (TNM) staging process in oncology (Lai-Kwon, Siva, & Lewin, 2018). Digital exploration of pathology can be enhanced through partitioning of an image and by digitally extracting tissues or structures from the field of view, which can enhance regions of interest. Special digital filters can also be used to increase the dimensionality of a volume of interest (VOI), thereby improving tissue feature analysis and the ability to sub-classify disease (Parekh & Jacobs, 2016). Edge detection algorithms is used to help identify the boundaries between healthy and diseased tissues. Contour analysis is used to segment and highlight the borders of pathology. Knowledge of the true border of pathology supports more accurate volumetric measures and feature mapping of pathology. Radiomics is capable of providing an automated solution for the assessment of disease characteristics and evolution.

Radiomic methods do not always serve as the final or conclusive method of pathology assessment. They can be used to help identify the signature for an aggressive region of pathology that can serve as a high-risk target for a traditional needle biopsy within a well-defined region of pathology (Sala et al., 2017). Radiomic feature analysis using multiparametric or hybrid imaging technologies such as PET/CT and PET/MRI can be used to interrogate large volumes of tissue in vivo with subcentimeter resolution, which enhances the accuracy of the virtual biopsy (Kressel, 2017). Integrated data from these approaches can also be used to better guide traditional biopsy methods.

Pathology Feature Extraction and Analysis

Radiomic methods are effective for revealing geometric, statistical, and textual features of pathology from imaging data. Over 440 radiomic features of pathology have been acknowledged in the literature (Wu et al., 2018). Common categorical features of pathology include texture, edge-attributes, volume measures, molecular relationships, contrast kinetics, uptake values, and subvisual pathology boundaries (Bai et al., 2016). Individual voxels and three-dimensional matrices comprised of an array of voxels provide the volumes of interest (VOI) for radiomic assessment. Texture analysis refers to the quantitative evaluation of grayscale intensities and their relationships within and between well-defined three-dimensional space referred to as voxels (Davnall et al., 2012; Lubner et al., 2017). Texture may correspond to various pathological processes. Specialized digital filters can be used to help reveal unique features of pathology such as entropy and uniformity (Nandu, Wen, & Huang, 2018). The greater the number of clinically relevant pathology features, the greater the potential for differentiating and classifying pathology.

Diagnostic images are comprised of first, second, and third order data which when exposed to advanced computational methods can detect patterns, relationships, and trends that support health care decisions (Gillies et al, 2016; Huang et al., 2017). First-order statistics include mean gray-level intensity, standard deviations, entropy, skewness, kurtosis and uniformity. Second-order statistics include local homogeneity, dissimilarity, correlation, and angular second moment energy (Gillies et al., 2016). Higher order statistics include coarseness, contrast, and complexity (Davnall et al., 2012). Data can be acquired and analyzed during the interpretive stage of radiology workflow with handcrafted and/or deep learning radiomic methods.

Multidimensional 3D Exploration

A focal region of pathology such as a tumor often has a high degree of spatial and temporal heterogeneity, which limits the usefulness of conventional structural imaging and traditional biopsy results. A digital (virtual) biopsy can be performed using an in vivo voxel by voxel (voxel-wise) interrogation process across the entire volume of pathology in multiple dimensions. Voxel-wise classifiers can be used to help reveal varying stages and subtypes of pathology within a single lesion (Bucking et al., 2017; Ng et al., 2013). Radiomic methods have been successfully used to help differentiate benign and malignant characteristics of tumors and to provide prognostic insights (Sala et al., 2017). Further research will expand knowledge of the molecular attributes and radiomic signatures of aggressive pathology within different biological systems and tissues.

Comprehensive in vivo assessment of pathology requires analysis across the three-dimensional volume of pathology rather than from a two-dimensional plane

(Echegaray et al., 2016; Nandu, Wen, & Huang, 2018). As data acquisition and computational methods become faster, the trend will be to auto interrogate larger volumes of pathology and auto map pathology features using in vivo data (Zhao et al., 2016). Virtual exploration supports interrogation of pathology within an augmented or immersive 3-D environment.

Pathology does not exist in isolation. It functions and interacts within a biological ecosystem that includes the surrounding microenvironment. Within this ecosystem, there is both phenotypic and genotypic plasticity that contributes to the evolution of pathology. The use of multiparametric imaging methods and radiomic measures can help expand and reveal feature data sets from the whole region of pathology including the surrounding microenvironment. This represents a significant advantage of multidimensional interrogation of pathology. It also offers an advantage over lab (blood) studies that represent circulating biomarkers, which do not localize pathology within an organ or tissue. In the near future, whole pathology slide mounts will be matched with image slice acquisitions and voxel by voxel radiomic measures. This approach will be used to help build computational disease models and disease detection algorithms.

Disease Screening: Discovery Radiomics

Early detection and characterization of pathology influences patient care and treatment outcome. Handcrafted algorithms and rules used with AI are often limited in their capacity to reveal subtle or unknown characteristics of healthy and disease states. D. Kumar et al. (2017) introduced the concept of “discovery radiomics” which refers to the use of high-throughput analysis of imaging data to detect early-stage pathology and to help predict the outcome of an asymptomatic disease process. A specialized approach to data analysis referred to as "evolutionary deep radiomic sequencing" offers a rapidly evolving method for detecting pathology, which is not dependent on a priori knowledge of disease criterion (Shafiee et al. 2017). In summary, **v**arious forms of radiomic applications can be used to provide a low-cost, fast, and potentially reliable method of screening for pathology.

Reimagining the Differential Diagnostic Process

The differential diagnostic process is comprised of a series of interrelated steps applied to a patient's presentation, which uses probability-based logic or reasoning to differentiate a disease or disorder from others that may have a similar presentation. The differential diagnostic process is often dependent upon different sources of information such as the history, physical examination, laboratory evaluation, imaging, and other specialized forms of testing. In general, the more complicated a patients' presentation, the greater the list of possible causes and contributing factors. Advances in diagnostic imaging have contributed to a growing appreciation for the complexity and heterogeneity of disease at anatomic, cellular, molecular, and genetic levels (Aerts, et al., 2013; Davnall et al., 2012; Lubner et al., 2017; Sala et al., 2017; Yip & Aerts, 2017). Radiomic measures can be used to " bridge evidence across different biological scales" in a manner which can inform the differential diagnostic process (Hsu, Markey, & Wang, p. 1010). It is important to discover new dimensions and features of pathology that have meaningful impact on patient care.

Heightened awareness of the complexity of disease expands the list of differential diagnostic considerations. Advances in diagnostic imaging will continue to reveal unique heterogeneic features that can be used to classify and subtype pathology (McCue & McCue, 2017). The heterogeneity of pathology is associated with evolutionary changes at the cellular level associated with genotypic and phenotypic plasticity. Heightened awareness of this process improves diagnostic and prognostic accuracy. For example, high levels of tumor heterogeneity have been associated with greater variability of treatment outcome and generally a poorer prognosis (Lundstrom, Gilmore & Ros, 2017; Sala et al., 2017). Successful delivery of more precise and personalized health care requires knowledge of biological differences and pathological variability along with relevant decision support at the radiology workstation.

The current diagnostic process is very nuanced, and influenced by a provider's familiarity with disease and related testing. In addition to assisting with disease detection and characterization, AI can be used to identify appropriate diagnostic tests and related protocols to help achieve a precise diagnosis (Baldwin, Guo, & Syeda-Mahmoood, 2017). There are differential diagnostic possibilities for every patient presentation (Hussain & Oestreicher, 2017). The individual health care provider typically relies on an intuitive diagnostic approach limited to their familiarity of two to six diseases during the initial differential diagnostic process (Phua & Tan, 2013). This level of awareness is often insufficient for unusual or complex conditions. Limited awareness of differential diagnostic possibilities leads to errors and missed treatment opportunities.

Traditionally, radiology has relied upon the visual perception of the radiologist, which limits their capacity to consider microscopic and molecular level differential diagnostic considerations (Pinto & Brunese, 2010; Yip & Aerts, 2016). Radiomic methods can be used to overcome visual limitations and expose measurable characteristics of subvisual pathologic states, thereby, adding value to the differential diagnostic process. Improvement of the differential diagnostic process is required in all areas of diagnostic imaging including spine care.

The application of new quantitative imaging (QI) measures and standards will refine the diagnostic process and lead to expanded criterion and classifications of disease (Farooki et al., 2016; Gillies et al., 2016; Kharat & Singhal, 2017; V. Kumar et al., 2012; Parekh & Jacobs, 2016). The expansion of objective measures of pathology will support AI solutions. Disease states and their evolution vary between individuals because of molecular, genetic (genotypic), structural (phenotypic), and biological diversity. For this reason, a one-size-fits-all approach to diagnosis or treatment does not work for everyone. Radiomic analysis, a specialized application of QI, can be used to characterize biological variability and pathological heterogeneity at subvisual levels independent of the radiologist's interpretation (Larue et al., 2016; Yip & Aerts, 2017). Radiomic data will assist the radiologist in the differential diagnostic process. The adoption of QI will also support further development of AI by providing objective labels used to annotate training and validation data.

The term *probability* in the diagnostic process refers to measures of the likelihood of a disease or pathologic process being present. AI can expose radiologists to differential diagnostic considerations (diseases) for which they have limited knowledge and experience. Probability is assigned to disease patterns and differential diagnostic possibilities associated with imaging presentations derived from hundreds of thousands of patients within a database or from case-based publications.

In summary, treatment success is dependent upon an accurate and efficient differential diagnostic process. The current diagnostic process remains too imprecise and inconsistent to adequately identify and subtype disease in many situations. This often results in a “one-size-fits-all” treatment approach. AI solutions have the potential to use probability calculations to help identify and prioritize diagnostic possibilities. AI will continue to influence the steps, as well as the sequence of steps used during image interpretation and the differential diagnostic process. The combined use of AI-supported data management solutions such as natural language processing, radiomic methods, machine learning, and advanced computational analysis at the radiology workstation might support improved accuracy and precision of the differential diagnostic process.

The Radiologist: New Roles and Responsibilities

The principal role of the radiologist is to detect, characterize, and report on disease processes in a manner that provides decision support at the point of care. For this reason, there is a growing demand for radiologists to become more involved in consultation and patient care (Ranschaert, 2016). Historically, radiologists are skilled in the evaluation of pathology associated with structural changes and are less skilled in the early detection of pathology based on subvisual criteria revealed by deep learning and radiomic methods (Malone & Newton, 2018). This inadequacy reinforces the need for computational assistance. AI has the potential to augment the role of the radiologist during the interpretive stage of radiology workflow and with final reporting. More specifically, AI may improve the performance of the radiologist by enabling earlier disease detection, offering more precise disease characterization, providing probability based differential diagnostic considerations, and reducing diagnostic error rates (Zherhoni, 2017). Radiologists are rapidly becoming among the most important data managers, knowledge brokers, and primary gatekeepers of big data and curators of knowledge for treatment planning and disease surveillance in health care (Hillman & Goldsmith, 2011; Jha & Topol, 2016). Using AI during the interpretive stage of radiology workflow can help identify meaningful incorporated into structured reports.

Despite these important roles, the interpretation of diagnostic images is still primarily limited to the visual detection and characterization of pathology (Pinto & Brunese, 2010). This approach is no longer adequate. In some areas of radiology AI has proven it has the potential to augment the role of the radiologist in the detection and interpretation of subvisual data (Farooki et al., 2016; Larue et al., 2016). AI has the potential to empower the radiologist to work better, faster, and smarter.

AI can augment the role of the radiologist and empower their role as a clinical consultant in numerous ways. For example, AI can help access, analyze, and correlate non-imaging data with imaging findings during the interpretive stage of radiology workflow. AI can also be used to flag subtle or visually hidden pathology and address mundane and redundant work, thus, freeing radiologists up to interpret pathology and better communicate with referring physicians and other members of the health care team

(Recht & Bryan, 2017). AI may overcome human limitations such as fatigue, lack of experience, unawareness of possibilities, and bias. To maintain relevance the radiologist must be willing to embrace AI, adapt to its use, and contribute to its development. Additional research is required to address how AI applications augment the role of the radiologist and improve the delivery of more precise and personalized care (Sutton et al., 2017).

Artificial Intelligence in Radiology: Current Applications

The individual radiologist is often overwhelmed with the growing burden of high volume complex data. This dilemma has led to the pursuit of different forms of decision support. This includes AI solutions. Most of the research surrounding the use of AI in radiology has been limited to highly specialized fields. There has been little research surrounding its role in spine imaging. Successful non-spinal applications will pave the way for use in spine care.

Non-Spinal Imaging

Because of the complexity of available data and the criticality of decisions, most research on how AI is used in radiology is limited to cardiology, neurology, and oncology. AI solutions are used in other specialties of radiology albeit to a limited degree. Deep learning methods have been successfully used to auto detect pulmonary tuberculosis on chest radiographs (Lakhani & Sundaram, 2017) and to auto detect disease states in neuroimaging such as intracerebral hemorrhage, stroke, and mass effects (Maier et al., 2015; Prevedello et al., 2017; Scherer, 2016). Computer-aided diagnostic systems help detect and characterize some neurodegenerative disorders (Cascianelli et al., 2016).

AI has evolved to a level where it has outperformed the human expert in radiology in some settings (Augimeri et al., 2016; Boone et al., 2015; Mohebian et al., 2017). Research addressing the use of AI has demonstrated its potential to augment the role of the radiologist.

The utility of AI is not limited to radiology. For example, deep learning algorithms have demonstrated greater accuracy than a panel of pathologists in the detection of lymph node metastasis in women with breast cancer (Bejnordi, Veta, & van Diest, 2017). Extensive research is also underway to develop methods for automatically extracting relevant information from unstructured reports in breast imaging (Gupta, Banerjee, & Rubin, 2018). In another study, convolutional neural networks outperformed cardiologist’s interpretation of echocardiographic images with 98% accuracy (Mandani et al., 2017). Successful use of AI solutions in one field of radiology may be adapted to meet the needs in another field such as spine imaging. The adoption of AI solutions requires adequate research to identify meaningful applications, as well as to confirm clinical utility and validity.

Spine Imaging

Diagnostic imaging is often a fundamental and influential component of spine care. Imaging findings influence decision making at all levels of care. Subsequently, interpretive imaging errors and missed opportunities have a profound impact on treatment planning and treatment outcomes. An extensive literature search revealed limited research and real-world applications of AI in spine imaging and spine care. It is common for early applications of AI to be applied to basic steps such as anatomic localization and labeling prior to implementation of more detailed applications. For example, machine-learning algorithms with different imaging modalities can auto-identify vertebral levels (Daenzer et al., 2014; Hetherington et al., 2017). Another research study demonstrated that AI could be used to segment and label vertebral bodies (El-Helo et al., 2013). Auto segmentation and labeling of anatomic regions has to be perfected before tissue features can be auto extracted and characterized.

AI use in spine care has not been entirely limited to anatomic localization. Machine learning methods have also been used to determine bone density, as well as to detect and categorize vertebral compression fractures on computerized tomography (Burns, Yao, & Summers, 2017; Doi, 2007; Hetherington et al., 2017). In another study El-Helo et al. (2013) demonstrated that AI performed with greater than 90% accuracy in the detection of vertebral compression deformities. This condition is relatively common and places a significant financial burden on the health care system. Vertebral compression deformities often missed on non-spinal imaging studies could be detected with AI supported methods. For example, N. Kim et al. (2004) reported that approximately 50% of vertebral compression fractures that presented on routine lateral chest radiographic studies were either missed or underreported.

A few research studies have exposed the potential utility of molecular imaging in spine care. For example, in vivo quantitative voxel-based mapping is used to evaluate the microstructural and molecular attributes of degenerative intervertebral discs and of the spinal cord in cervical spondylotic myelopathy (Grabhar et al. 2015; Grunert et al., 2014).

AI could be used to help identify the presence of subtle spine pathology that may be missed during non-spinal imaging studies, which include spine data in the background.

Most diseases progress through an asymptomatic (subclinical) period. For example, early spinal cord compromise (myelopathy) secondary to degenerative stenosis and compression often results in asymptomatic changes in regional biochemistry, blood flow, and tissue architecture preceding the onset of clinical signs and symptoms (Durrant & True, 2012). These changes are often not evident on routine imaging studies. The visual presence of pathology within the spinal cord on advanced imaging studies is often associated with end-stage pathology and permanent neurological deficits. Successful detection of early stage myelopathy will require non-visual analysis of data acquired form the spinal cord with molecular imaging and AI-supported radiomics. Researchers have begun to address the possibilities. For example, a specialized form of MRI referred to as diffusion tensor imaging (DTI) combined with machine learning classifiers has been successfully used to detect early stage spinal cord compromise secondary to degenerative narrowing of the central spinal canal in the neck, a condition referred to as cervical spondylotic myelopathy (Wang et al., 2015; Wang, Hu, Shen, & Li, 2018). The concept of discovery (screening) radiomics can applied to any bodily region including the spine and spinal cord.

Computational AI approaches have been used to auto classify intervertebral discs as either normal or degenerative based upon the analysis of tissue features such as signal intensity, texture, and shape (Ghosh & Chaudhary, 2014; Oktay, Albayrak, & Akgul, 2014). In another study, AI assisted the automated detection and characterization of lumbar neuroforaminal stenosis on MRI studies with greater than 90% accuracy (Han et al., 2018). Further research is required to determine how AI-supported methods could augment the role of the radiologist during the interpretive stage of spine imaging. Further research on the use of AI in radiology will shape the future spine care.

Spine Pathology and AI: Meaningful Use Considerations

The capabilities of deep learning AI systems are highly dependent on exposure to adequate levels of annotated training data and the establishment of ground truth. The process is often expensive, tedious, and time-consuming. Due to this level of commitment, it is imperative for stakeholders, including clinicians and radiologists, to help identify meaningful use applications prior to investing in the process. Meaningful use in this context refers to the application of an AI solution to address a condition or disease which is prevalent, not adequately assessed with normal methods and which has a significant impact on the individual and society. The institutional definition of meaningful use applications in this context may include economic value assigned to new solutions. An example of successful AI application with a favorable outcome is referred to as a meaningful use case.

Some spine disorders, if left undetected and untreated, can lead to devastating consequences that place an unnecessary burden on the individual, their family, and society. The evaluation of AI use in spine imaging should start with the most devastating, prevalent, and costly spine disorders. Examples include the 550,000 to 700,000 vertebral compression fractures which occur annually in the United States secondary to osteoporosis (Kondo, 2008) and the two-thirds of patients with cancer who will develop bone metastasis, with the spine representing the most common location (Maccauro et al., 2011). Intervertebral disc degeneration represents one of the most common causes of back pain, which afflicts approximately 80% of the adult population throughout their lifetime (Suthar et al., 2015). The annual prevalence of spinal cord compromise (myelopathy) secondary to degenerative changes is estimated to be about 196,000 in North America (Nouri et al., 2015). Each of the above conditions occurs in well-defined anatomic regions of the spine, which can be interrogated in vivo at multidimensional levels using emerging AI-supported methods

Some of the most easily segmented and labeled anatomic regions of the spine house some of the most devastating types of pathology. These regions include the bone marrow microenvironment within the vertebral body and the spinal cord microenvironment within the spinal canal. In each case with the exception of trauma, severe pathology begins with nonvisible, asymptomatic, tissue changes. Early detection and intervention may lead to better patient outcome. AI-supported interrogation of vulnerable anatomic regions could help detect and characterize aggressive pathology and lead to early-personalized intervention.

The bone marrow microenvironment within the vertebral body is involved in many different disease processes. For example, changes within the bone marrow often precede the development of degenerative disc disease, as well as the development of compression deformities and fractures. The bone marrow space is also a common location for metastatic disease. Approximately two thirds of individuals with cancer will develop bone metastasis with the spine representing the most common site (Maccauro et al., 2011). Micro metastasis or early metastatic disease is often subclinical and difficult to detect with traditional anatomic imaging methods. Approximately 36% of vertebral lesions associated with spine metastasis are asymptomatic and discovered incidentally on spine imaging for other disorders (Maccuaro et al., 2011). In some cases evidence of metastasis to vertebral bone marrow may represent the first indicator that cancer is present somewhere else.

The vertebral bodies and bone marrow are well visualized on routine spine MRI and CT studies; therefore, rendering it possible for AI supported screening methods to be applied to detect subtle or early stage pathology. Prior to implementing AI supported solutions such as radiomics ground truth must be established for normal and abnormal states. Discovery radiomics and in vivo interrogation of vertebral body microenvironments may help reveal early stage osteoporosis, micro metastatic disease, and subchondral degenerative changes that often precede the development of intervertebral disc disease. Early detection may lead to early intervention.

Some advanced imaging methods and protocols have demonstrated the ability to reveal micro pathology within the bone marrow of the spine (Long, Yablon, & Eisenberg, 2010; Park et al., 2015). AI-supported radiomic methods have demonstrated the ability to detect nonvisible evidence of pathology in other fields. Its potential role in spine care must be investigated further. In the future, radiomic methods will likely detect and characterize vertebral bone marrow pathology such as myeloproliferative disorders, subchondral degeneration, osteonecrosis, infection, and tumor (Long et al., 2010). It is imperative to diagnose pathology within the vertebral bone marrow environment at the earliest stage possible.

The first step in pathology detection using AI methods is auto assessment and labeling of anatomic structures. The second step is to segment a region of interest. The third step is to apply specialized AI applications such as radiomic methods to detect and characterize pathology. AI has already been successfully used to perform automated vertebral boundary detection and surface texture analysis using advanced context-encoding features (Mirzaalian et al., 2013). This provides the digital framing and segmentation required to isolate bone marrow. With the exception of one study, I was unable to identify any significant research surrounding the use of radiomic methods to evaluate tissues of the spine or spine pathology. In the referenced study, quantitative voxel-based feature detection and analysis was performed using computed tomography studies of the spine to assess true marrow space, fat composition, and mineral-based marrow density (Pena et al., 2016). Pena et al. (2016) found that radiomic methods offer advanced tissue differentiation and characterization that could be applied to the spine.

The same spine disease or disorder within different individuals may vary in presentation, severity, and progression due to molecular, genetic, and structural diversity. For this reason, a one-size-fits-all approach to spine imaging diagnosis or to treatment may not work. AI-supported methods such a radiomics can help classify and stratify disease and therefore help identify the best-personalized treatment plan. This research study may represent one of the first to address the potential impact of AI in spine imaging along with the potential role of radiomics and the virtual biopsy for the in vivo interrogation of spine pathology.

Combinatorial Evolution of Artificial Intelligence

The pace of technological development and the rapid evolution of AI will continue to increase. As technology advances the duration between innovations and new applications often becomes shorter. In radiology, the phenomenon is amplified by co-evolution of integrated technologies such as computer processing, imaging modalities, database networking, and AI workflow solutions. As AI provides better analytic insights and decision support there will be a greater push for more advanced imaging technology, further complicating the decision-making process. Once human limitations are overcome, there will a rising demand for the acquisition and analysis of more complex data to support more precise and personalized care. Limited decision support for complex data was previously a barrier to technology development and technology adoption in radiology.

Perpetual revisions of disease criterion and classifications combined with ongoing disease biomarker discovery will result in recurrent cycles of disruption and adaptation in radiology and related clinical workflow. The three principal influences of an innovation are differentiation, precision, and speed. Each of these factors will be pushed to the limits in radiology by AI. There is a recursive relationship between technology, data, and human contributions, all which collectively influence diagnostic decisions and the delivery of care. The co-evolution of technologies at the radiology workstation surrounding the management of high-throughput data will lead to unprecedented methods of pathology assessment. AI and related technology support will lead to expanded knowledge of disease pathophysiology and new standards of evidence-based care. Exploratory research methods will help reveal the potential use of AI in radiology and heighten awareness of and readiness for what is to come.

Imaging With AI: The Key to Precision and Collaborative Spine Care

Using big data and AI will transform the entire field of health care (Obermeyer & Emanuel, 2016), including spine care. The use of AI solutions in radiology will forever change how decisions are made in all health care specialties (Brink et al., 2017; Jha & Topol, 2016; Jiang et al., 2017; Ranschaert, 2016). The unprecedented paradigm shift associated with molecular level diagnostics and AI decision support will include spine care. Personalized spine care requires identification of biological differences and pathological variability, topics AI can help address.

Various forms of AI have the ability to improve the precision of the diagnostic process in radiology by revealing unique attributes of disease (Syed-Mahmood, 2018; Wang et al., 2015; Wang, et al., 2018; Zherhoni, 2017). For this reason, successful applications of AI in radiology can have a favorable impact on the delivery of predictive, pre-emptive and personalized health care (Augimeri et al., 2016; Brink et al., 2017; Jha & Topol, 2016; Jiang et al., 2017; Ranschaert, 2016). The spine is intricate and complex. The individual spine has many unique structural and biomechanical attributes that influence the impact of disease. Greater knowledge of the unique attributes of an individual's spine pathology will inform all members of the spine care team. It will support more precise and personalized spine care. Fundamental shared knowledge will also facilitate more effective collaborative multidisciplinary care.

Growing use of AI in spine care will result in greater appreciation for the spectrum of pathology, and for the need for a more timely and precise diagnosis. Future AI applications in spine care will also expose common ground, redefine expert boundaries, dampen existing professional turf wars, and lead to a few new ones. Successful application of AI during interpretive stage of radiology workflow in any specialty field of health care has the potential to transform the delivery of care while supporting a more precise and personalized diagnostic process **(**Acharya et al., 2018; Ghasemi et al., 2016; Hillman & Goldsmith, 2011; Jha & Topol, 2016). In addition to improving the accuracy of the diagnostic process, AI may serve as a catalyst to new levels of multidisciplinary collaboration in spine care. Some of the current applications of AI in oncology, neuroradiology, and breast imaging will likely be adapted for use in spine imaging.

A better understanding of the molecular basis of spine pathology using deep learning and radiomic methods will lead to earlier detection and more precise subtyping of pathology. This process will have an impact on the standard of care and will perpetually reset clinical expectations. With heightened awareness of available decision support radiologists, spine care providers, and the public will become less tolerant of complacency and errors made during the interpretation of spine imaging. New expectations and standards in spine imaging interpretation and reporting will have a significant impact on evidence-based spine care.

The potential benefits of multidisciplinary care are well established. The success of intervention is influenced by greater awareness of the heterogeneity of pathology, acceptance of new disease criterion and classifications, and evolving standards in decision support (Aerts, 2016; Collins & Varmus, 2015; Hood & Aufrey, 2013). Collaborative spine care will become progressively more dependent on the central role of the radiologist as a gatekeeper of big data and as a clinical consultant. The radiologist's role and impact is strengthened by improved diagnostic methods and better decision support (Castenda, et al., 2015; Kressel, 2017). A more precise diagnostic process during imaging workflow will help expose the fundamental basis of disease and subsequently help bridge-the-gap between disciplines working at different points along the spectrum of pathology (Brink et al., 2017; Kressel, 2017; Jiang et al., 2017). The collective use of human and machine intelligence during the interpretive stage of spine imaging workflow may help resolve multidisciplinary discordance by democratizing decision support and by exposing standardized terminology and disease criterion.

Ethical Issues Surrounding the Use of AI in Radiology

Numerous ethical considerations surround the use of AI and big data in radiology that could influence patient care (Mittelstadt & Floridi, 2016). Expanded use of AI will subsequently require the creation of new ethical standards surrounding (Mesko, 2017). Lee et al. (2017) raised concerns about potential challenges associated with deep learning in radiology. The challenges included a growing dependency on large volumes of training data and the black box nature of the technology. The later refers to the lack of transparency of computational methods used to provide decision-support. The steps associated with deep learning analysis of complex nested layers of computational processes hidden. The lack of transparency limits validation and may lead to unprecedented categories of liability.

Challenges associated with rapid adoption and evolution of AI in radiology includes the possibility of over fitting results due to innovation bias, unnecessary hype, and unrealistic expectations. Misdirected hype surrounding AI use in radiology could have many direct and indirect adverse consequences on a health care system, such as draining capital, unproductive reassignment of expertise, and the generation of false expectations. An AI supported diagnostic process restricted to data mining can lead to unproductive steps in workflow. Thus, Mayo et al. (2016) introduced the concept of "farming of data," in contrast to mining of data (p. 261). Data farming is characterized by actively harvesting necessary data, locating missing data, picking the best data, and weeding out unnecessary data. Ethical matters must be addressed to improve meaningful use and the clinical utility of AI.

Widespread adoption of AI could result in deskilling of health care providers including radiologists. In addition, the use of AI may result in a growing level of technology codependency, thus, diminishing human influence in decision making. Widespread AI adoption could also introduce automation and/or technology bias. Earlier detection of pathology could result in unnecessary testing and treatment exposure. Access to AI enhances authority and expertise and will afford the user with an advantage that can have a significant impact on leadership roles, collaborative efforts, and the equality of care. For the reasons stated, it is necessary to perpetually explore and address ethical considerations that may arise because of the development and use of AI during the interpretive stage of spine imaging workflow.

The Research Approaches

Researchers in the disciplines of computer science and radiology have approached the potential role of AI in radiology from numerous perspectives. The approaches have identified what may be missed with traditional anatomic imaging and qualitative interpretation. Additional strengths are related to the investigation of narrow (disease feature specific) applications of AI using methods such as radiomics and natural language processing in isolated fields such as oncology. The weakness of this approach is the limited knowledge acquired surrounding proposed AI, its interoperability with current workflow, clinical utility, and ease-of-use.

There have been a limited number of qualitative exploratory studies on the potential role of AI during the interpretive stage of radiology. The absence of qualitative insights has resulted in numerous reductionist approaches to research on the topic with limited capacity for generalization and practical clinical application. Many of the research studies referenced in this work have not led to the development of an adequate concept map or blueprint of the sequence of research required to address the potential role and impact of AI on the interpretation of imaging studies and on the final reporting process. Exploratory research will help identify needs and potential and reveal the sequence of research studies and methodologies required to achieve desired results.

The Literature Gap

The literature clearly establishes numerous variables in radiology, including spine imaging, which complicate the interpretive and diagnostic process. This includes the growing burden of complex data, human bias, individual biological variability, heterogeneity of pathology, and the multifocal nature of spine pathology. Additional challenges associated with the use of AI and radiomic methods in spine care include the intricacy of structures, proximity of anatomic elements, and limited access to relevant databases and computational disease models. Technological variables include the lack of standards and wide range of differences between imaging modalities and protocols. The literature review revealed an absence of gold standards in radiomics. The potential role and impacts of radiomics is underexplored in all imaging specialties (Oakden-Rayner et al., 2017). An exhaustive literature search revealed a growing number of research studies designed to address the role of AI decision support in radiology. The search revealed some of the challenges associated with imaging interpretation. Many of the published research studies address narrow applications of AI and therefore do not address the challenges associated with its adoption, consistent use, and support.

The diagnostic process in spine care is primarily limited to the history, physical examination, electrodiagnostic testing, and diagnostic imaging. The literature search established the absence of reliable serum biomarkers for confirming the presence or progression of spine disorders. The search also confirmed that traditional needle biopsies are rarely performed on spine pathology. Neurologic deficits associated with a spine disorder often represent end-stage pathology and a certain amount of permanency is likely. For the reasons stated spine care providers of all disciplines are highly dependent on diagnostic imaging reports for decision support at the point of care. Improved delivery of spine care will subsequently require the use of advanced imaging and related decision support for the radiologist resulting in more precise and personalized reporting.

The literature search and review performed for this study supported the need to improve the differential diagnostic process during the interpretive stage of spine imaging workflow. Image interpretation and reporting in other fields has also been described as incomplete, inconsistent, and inconclusive (Bosmans et al., 2011; J. Y. Chen et al., 2017). Wu et al. (2018) discussed the “unmet need for methods that allow more comprehensive disease characterization and reliable prediction or early assessment of treatment response and prognosis toward the goal of personalized or precision medicine” (p. 125). Spine disorders represent one of the most common causes of pain and disability; therefore, radiologist’s should do what is necessary during the interpretive stage of radiology workflow to improve the accuracy of the diagnostic process and help ensure successful delivery of personalized spine care.

AI offers potential solutions for spine imaging, although, little attention has been paid to its potential. The use of AI in radiology has generally been limited and slow due to the challenges associated with its development and validation (Aerts, 2016). I was unable to identify any scholarly research articles that addressed the potential applications or impacts of the digital (virtual) biopsy in spine care. Knowledge about how to implement radiomic measures into routine radiology practice is also limited (Vallieres et al., 2017). The potential role of radiomics needs to be addressed within the context of spine imaging. The literature reveals many of the challenges associated with its use in other specialties including lack of standardized imaging protocol, limited access to annotation training datasets, defined meaningful use applications, clinical utility, and contouring regions of interest (Lai-Kwon, Siva, Lewin, 2018). The literature review revealed successful use of AI in many areas of non-spine imaging. Established success, although limited, has involved the use of natural language processing, radiomics, and computational diagnostics. The success of AI solutions in radiology requires scalable applications, clinical utility, and seamless integration into radiology workflow (Court et al., 2016; Syeda-Mahmood, 2018). There is a gap in the literature surrounding the use and potential use of AI and AI-supported methods during spine imaging workflow. This includes the topics of NLP and radiomics.

The role of AI in some fields of radiology such as oncology has advanced more than in spine imaging. Published research surrounding the role of AI use in other subspecialty fields of radiology provide the foundation for the discussion of its potential role in spine care and the design of this research study. The limited research on spine imaging has been associated primarily with automated identification of normal anatomy rather than addressing the detection and characterization of disease states. A published letter in 2016 representing the position of leaders of the American College of Radiology, a cultural authority in the field of radiology addressed the general gaps in knowledge and research surrounding the use of AI in radiology. Written to the U.S. Office of Science and Technology Policy, the authors summarized the gaps in the literature with acknowledgment of the need for further research to help identify how AI can be used to access meaningful data, enhance the interpretive phase of radiology workflow, improve diagnostic accuracy, and reduce errors. This request applies to all areas of radiology including spine imaging. As noted prior, the results of the literature search confirm the paucity of studies addressing gaps in knowledge regarding the role of AI in spine imaging.

Summary

The growing appreciation for the heterogeneity and complexity of pathology has led to the realization that more precise personalized care is possible with the right decision support. To achieve this goal, the diagnostic process must be more comprehensive and classifications of pathology expanded. The standards in radiology must change. The interpretive approach can no longer be limited to one individual's visual assessment of overwhelming volumes of two-dimensional images and the generation of highly variable qualitative reports. The data are too complex and the stakes are too high. Automated methods of disease detection and characterization are required to augment the role of the radiologist. The clinical utility associated with the adoption of AI solutions in radiology, and, more specifically spine care, must be determined. Timely personalized patient care should take priority.

An extensive scholarly literature review revealed that AI systems are capable of integrating and analyzing structured and unstructured data to refine the diagnostic process. The AI methods required to accomplish this goal include quantitative imaging with feature analysis (radiomics), acquisition analysis of qualitative imaging features from records (natural language processing), unsupervised feature learning (text and images), and scaling of AI solutions to accommodate multiplatform data (distributed computational models). Successful integration of AI solutions during the interpretive stage of spine imaging workflow will reduce errors and result in unprecedented diagnostic capabilities. AI can provide new perspectives of disease that will lead to more efficient and effective care. Success requires that gatekeepers of big data, such as radiologists, must accept new responsibilities, assume new roles, and embrace AI decision support.

This study focused on the potential role and impacts of AI applications during the interpretive stage of spine imaging. The results of the literature review served as critical determinants of the potential applications of AI in spine care. The literature review established the benefits of using AI supported method such as radiomics, natural language processing, and computational disease modeling in other fields of radiology to achieve a more precise probability-based diagnosis. It is evident based upon an extensive review of the literature that in the near future, the interpretive stage of spine imaging will likely rely on the use of collective intelligence derived from the integration of human and machine intelligence. This research study was designed to help determine how and when this might occur.

The published research suggests the radiology workstation will become a hub of convergent information from other patient diagnostic procedures and databases, including laboratory, genetic and pathology test results. Databases will include accessible computational disease models and disease registries. Integrating AI and radiomic methods will fundamentally alter how disease is diagnosed, classified, and treated (Langs et al., 2018; Shaikh et al., 2018). Further development of radiomic methods will support applications for the assessment of non-cancer-related spine pathology.

Current research acknowledges that emerging AI solutions are capable of providing radiologists with contextually relevant and probability-based differential diagnostic considerations during the interpretive stage of imaging workflow. This process improves diagnostic precision and therefore can help overcome human limitations and bias. In the future, whoever has access to the best data and the best decision support will likely provide the best care.

Ongoing advances in diagnostic imaging will continue to challenge and expand our current understanding of disease and related diagnostic criterion. The differential diagnostic process will soon no longer be limited to the expertise and skills of an individual; instead, a whole systems process will involve collective intelligence derived from the contributions of humans and machines. For this outcome to be possible, many unknown factors must be addressed. This includes what defines meaningful use and adequate training of an AI system. Heightened awareness of improved accuracy and efficiency associated with AI decision support will drive computational analytics and radiomic methods to the forefront of radiology, supporting their eventual role as routine procedures. Multidisciplinary research will lay the foundation and pave the way for the transformative process.

Heightened awareness of AI potential and clinical utility in spine imaging is required to further the research and development process. The endeavor will require the insights and participation of numerous experts such as AI developers, physicists, radiologists, pathologists, key influencers, and early adopters of AI in radiology. The establishment of expert opinions and predictions can help direct further discussion and research of the topic. A well-designed exploratory case studies can provide this necessary foundation.

Chapter 3 introduces the research design and methodology used to explore the potential impacts of AI on the interpretive stage of spine imaging and on the differential diagnostic process. The results of the extensive literature search reported in this chapter are used to support the chosen research methodology and strategies acknowledged in Chapter 3. The chapter addresses my role as a researcher, as well as the data acquisition and data analysis strategies used in the study. Special attention is placed on the implementation of steps to protect research participants and to improve the credibility and trustworthiness of the study.

Chapter 3: Research Method

Introduction

The primary goal of this study was to establish the potential impact of AI solutions on spine imaging interpretation and diagnosis. I placed special emphasis on the potential role of radiomics. The unit of study was the interpretive stage of spine imaging workflow used to detect, characterize, and monitor pathology. The sources of data included document review, reflective journaling, and focus group sessions. Focus groups are an effective method for exploring attitudes, expectations, and potential applications associated with emerging technologies and related processes (Kitzinger, 1995).

In Chapter 3 I introduce the research design and methods used to address the topic of study. In this discourse I address my role as a researcher and the role of research participants. I also address the methods used for data acquisition and analysis. The chapter indicates the research steps implemented to improve the trustworthiness of the study, as well as the processes used to help ensure the ethical treatment of participants and the ethical management of data.

Research Design and Rationale

I used a qualitative exploratory case study design to investigate the potential impacts of AI on the interpretive stage of spine imaging workflow. This approach offered a flexible and inductive method for acquiring holistic and in-depth insight. The primary purpose of qualitative exploratory research is to reveal the potential contributions and influence of an emerging technology or process (Baxter & Jack, 2008). My chosen research design was used to address how and why questions surrounding the potential use

AI applications such as radiomics, NLP, and diagnostic inference methods using deep learning approaches. Exploratory approaches are often used to lay the foundation for additional methods of inquiry such as quantitative and mixed method research (Creswell, 2013; Patton, 1990). AI represents a bridging technology comprised of an assemblage of evolving elements and processes that are sometimes difficult to identify and assess. This scenario contributes to complexity and uncertainty in research. A qualitative exploratory approach is able to reveal contextual relationships not adequately addressed by more restrictive explanatory or quantitative research methods (Ponelis, 2015; Yin, 1984). It was necessary for me to offer an inductive contextual perspective of potential AI applications which might be of value to radiologists and other stakeholders in the field.

The chosen study design supported the triangulation of qualitative data acquired from numerous sources, which helped to improve study validity and trustworthiness. The qualitative research approach supported purposive sampling, the acquisition of expert insight from different sources, inductive investigation, and the ability to formulate a contextual narrative summary. The primary research question was: What are the opinions of experts regarding the potential use and impact of AI during the interpretive stage of spine imaging workflow? I added supportive research questions to address determinants of AI adoption and various applications of AI such as radiomics and natural language processing.

I acquired qualitative data from expert documents in the form of white papers published by thought leaders and radiology organizations which addressed the evolution of AI in radiology. The documents I used represented consensus opinions on the use and potential use of AI. I framed the method of inquiry with insight acquired from an extensive literature search and from perspectives offered by a consensus-based prospective document prepared by the Spinecare Data Science Committee of the American Academy of Spine Physicians (AASP). The committee provided a list of high-priority topics (needs analysis) related to the potential use of AI during the interpretive stage of spine imaging. I considered the proposed topics during my development of research strategies and in my preparation for the focus group sessions.

Qualitative data acquisition occurred in the following sequential stages: an extensive literature review, review of a prospective document from the AASP Spinecare Data Science Committee, review of expert documents, and the use of two focus group sessions, one consisting of radiologists and the other AI experts. I performed reflective journaling during the entire data acquisition and data analysis process. I used focus groups to acquire expert knowledge, opinions, perceptions, and predictions relevant to the topic of study. The research process was designed to reveal themes, noteworthy quotes, and new perspectives surrounding the use of AI during spine imaging interpretation. Prior research had established that exploratory focus groups can be used to identify attitudes, discover opportunities, generate ideas, and frame new questions for future inquiry (Breen, 2006). I used focus group sessions to facilitate creative discussion and to expose the potential benefits associated with AI use at the radiology workstation in spine care. The process revealed new insights and exposed culturally formed attitudes and opinions surrounding the potential applications of AI. The insights I acquired from the focus group sessions helped me predict the type of synergies, controversies, and debates which may arise surrounding this topic in other research settings. The design of this research study supported the transition from shared experiences and insights to higher levels of abstraction and application.

Role of the Researcher

The role of the researcher is important in any research study but particularly important in qualitative exploratory studies because of the subjective nature of the process. Researcher experience, motives, and bias can all have a significant impact on the research process including the analysis and interpretation of acquired data (Durdella, 2019). A qualitative researcher often assumes a primary role in the acquisition and analysis of data. I assumed the role of the sole researcher in this study. As the sole researcher, I represented the primary instrument for the collection and analysis of data. My clinical experience combined with my desire to help identify new forms of decision support in spine imaging motivated me to pursue the topic of study. It also increased the risk for professional bias during the research. I subsequently implemented numerous steps in the research process to reduce my potential for introducing bias into the study. The steps taken to reduce my personal or professional (researcher) bias and to improve the trustworthiness of the study included the use of independent experts to review the focus group moderator guide, the use of member checking (respondent validation), the application of within group and between group analysis, and triangulation of data. Field testing of focus group protocols and related research questions was performed to help establish credibility and the relevance of the approach. I used reflective journaling to help reveal my biases, thoughts, and opinions throughout the research process and provide the basis for some of the decisions I made during the research process.

Prior to this study I had extensive clinical experience in diagnostic neurology, neuroradiology, and spine care. I also had extensive academic training in AI and radiomics. My professional experience as a clinician combined with my familiarity with AI provided me with the insights required to develop and implement effective exploratory and analytic strategies. Prior to and during the study I prioritized conducting myself and the research process in accordance with acceptable scientific methods and in accordance with the Walden University Institutional Review Board (IRB) guidelines. I implemented numerous methods to help support scientific data analysis the disclosure of research conclusions in an unbiased and objective manner. I implemented the previously disclosed strategies to help ensure that I was reflective and transparent throughout the research process.

Personal and Professional Relationships

Prior to or during the course of this research study I did not have any formal business relationship with IBM, any other AI-related company, or professionals who participated in the focus groups sessions. I did not pursue or accept research participants with whom I had any prior business relationship. During the proposal stage of the dissertation process I participated in numerous conference calls with IBM staff including data scientists to discuss gaps in the research, research strategies, and the management of research related data. During the early stage of the dissertation process I used a key contact from IBM to help identify a few renowned AI experts who met the study inclusion and exclusion criteria. I chose to approach IBM due to their performance record, current market position, and their potential for developing AI solutions for radiology.

Researcher Bias

Among the many potential sources of bias in a qualitative exploratory research study, some involve the researcher. Bias can occur in many forms and can influence different phases of the research process such as the development of research questions, participant recruitment, expert interviews, data acquisition, and data analysis (Creswell, 2013). Potential sources of bias in this study include the effects of the researcher on the study and the effects of the research process on the researcher. Researcher bias is possible whenever research relationships could lead to future business opportunities. For this reason, I did not accept any proposals or entertain discussions about potential future relationships.

I was not offered any position with research participants and/or companies or institutions they were been affiliated with prior to or during the course of the research study. I pursued the research topic and study with bias in favor of the eventual use of AI solutions to improve diagnostic accuracy and the interpretive diagnostic process in spine imaging. I fully disclose that I do not fully understand how AI could or should be used. As a practicing neurologist, I acknowledge that the accuracy of the diagnostic process in spine imaging must be improved. Approximately two years ago I sat at a prototypical IBM Watson AI workstation and experience its potential contribution to the differential diagnostic process in radiology. With the exception of the isolated experience with IBM

Watson, I have no other practical hands-on experience with the use of AI in radiology or spine care. I am therefore not biased toward the adoption of AI in radiology based on personal hands-on experience.

Ethical Issues Surrounding the Researcher

As the researcher in the study, I anticipated and addressed potential ethical issues that may have arisen prior to, during, or after the research process. Prior to designing the study, I became familiar with the Belmont Report (1979) published by the U.S. Department of Health, Education, and Welfare, which acknowledged ethical principles and guidelines which can be used to protect human subjects while conducting research. In addition, during the course of my PhD studies at Walden University, I completed the National Institutes of Health (NIH) web-based training program titled “Protection of Human Research Participants” (Certificate # 2872343). I assumed the duty as the primary researcher in this study to handle myself in a scholarly fashion and to treat all research participants in a professional and ethical manner. This required the implementation of steps to ensure participants well-being while protecting their rights and minimizing their exposure to potential harm.

Research Methods

The dissertation proposal was accepted and Walden University IRB approval was obtained prior to beginning the research process. The IRB reviewed the research plan and the research methodology, as well as all pertinent documents. The approval process was completed prior to research participant recruitment and the acquisition of data. I implemented steps to disclose and address researcher bias, potential conflicts of interest, competing motives, and potentially detrimental power relationships that could have developed during the course of the study.

Study Population and Sampling

The study population consisted of stakeholders involved in or influenced by the interpretation of spine imaging. This includes data scientists, device manufacturers, AI programmers, radiologists, spine care providers, and other health care providers. It was important to identify the subpopulation of stakeholders most capable of addressing the exploratory research topic within focus group settings. The success of this study depended on my ability to recruit research participants experienced and knowledgeable on topics related to the potential impact of AI used during the interpretive stage of spine imaging workflow. The focus group study population was limited to radiologists and AI experts. The members of each category of participants were intricately involved in processes which took place within the parameters of the unit of study, which was the interpretive stage of spine imaging. I implemented steps that required that all research participants met strict research inclusion and exclusion criteria.

A certain degree of homogeneity or similarity within a focus group session combined with purposive sampling of professional participants has proven to enhance the potential for exploring new technology (Kitzinger, 1995). Kitzinger (1995) also demonstrated that a group of professionals with common knowledge along with similar training and experience are more likely to engage in in-depth discussions. To facilitate this approach, I placed AI experts in one focus group session and radiologists in a distinct and separate focus group.

Population sampling for the research study was convenient and purposive so that small groups of confirmed experts could discuss and explore the research topic. Experts recommend purposive sampling strategy to help ensure that participants have the level of experience and expertise required to contribute to the topic of study in group sessions (Creswell, 2012; Patton, 2002). Subsequently, I used purposive sampling in this study to help ensure that all of the research participants had an adequate level of expertise, interest, and experience surrounding the development or application of AI solutions during the interpretive stage of diagnostic imaging. The use of convenience sampling combined with purposeful sampling brought together like-minded professionals who were familiar with the topic of study and categorically with role of AI experts and radiologists in spine care.

The population sampling strategy included estimation of the sample size required to achieve the level of representation and topic saturation required to explore the potential impact of AI on spine imaging interpretation and diagnosis. A renowned key AI contact was used to help identify qualified AI professionals to participate in the study. I used key radiology contacts to help identify radiologists qualified for participation in this study. I provided each of my contacts with the purpose of the research study, as well as participant inclusion and exclusion participation criterion prior to asking for their recommendations. I confirmed that all potential participants met inclusion and exclusion criterion prior to their acceptance into the study.

The ideal size of a focus group is often five to eight participants, unless more are required to address a complex topic (Krueger & Casey, 2015). Qualitative research experts have acknowledged that four to 10 individuals are often adequate for homogenous sampling of professionals (Creswell, 2013; Krueger & Casey, 2015). I limited the number of expert participants to eight to 12, or four to six in each of the two focus group sessions, to facilitate creative and in-depth discussion of a complex and contemporary topic. The small focus group sizes helped me as the moderator better manage research topics and the flow of discussion and ensure that each participant had adequate opportunities to share their expertise and insights. The use of two homogenous focus groups comprised of confirmed experts was large enough for this study to provide a diversity of perspectives and opinions.

The research participants accepted for participation in this study were from different clinical and professional settings. I recruited research participants through personal contacts. I followed up with potential participants by of email and phone calls. I did not offer material or monetary incentives during the recruitment process. Research subjects who agreed to participate in a focus group session were asked to complete a brief survey prior to the focus group session (Appendix B). I used a survey to acquire participant demographic information such as their current position, background, and experience.

Prior to convening for focus group sessions each research participant received an acceptance letter which included an introduction to the research project, a focus group agenda, a list of what was expected of them and a research participation consent form to review, sign and return. Each research participant received email notification of the scheduled focus group session along with an invitation to participate in person or via a prearranged teleconference option. I arranged for the teleconference option through Zoom, a highly respected and secure service. Email notifications included the focus group facility address, focus group directives, and access information for the teleconference option. Each participant received periodic email reminders of their scheduled focus group session. I requested that research participants register online in advance of the focus group sessions.

Sample size in qualitative research is often not as important as the chosen methods of data acquisition, data analysis, and data validation (Njie & Asimiran, 2014). The potential use of AI during the interpretive stage of spine imaging is both a new and complex topic. Subsequently, I chose small sample sizes to help achieve expert in-depth discussions and topic saturation during the focus group sessions. Data saturation is reached in a qualitative study when a coherent and consistent perspective is reached (Guest, Bounce, & Johnson, 2006). I established the criteria for data saturation in this study prior to the focus group sessions and included the definition on in the moderator guide (Appendix D). I developed open-ended and probing research questions to help achieve data saturation during the focus group sessions.

Research Participant Inclusion and Exclusion Criteria

In qualitative research, research participants must be capable of contributing to the study with the chosen methods of inquiry (Creswell, 2013), particularly when addressing a complex and rapidly evolving technology such as AI. I used convenient and purposeful sampling methods to select research participants for this study. The potential research participants were subjected to explicit study exclusion and inclusion criteria. The inclusion criterion addressed the potential participant’s current professional status, background, and experience. The inclusion criterion for AI expert participation was a minimum of five years of experience in health care AI development or applications. In addition, each AI participant was required to have a minimum of a bachelor’s degree in a related field such as informatics, data science, computer science, or AI. AI participants were also required to be actively working in an AI field.

The inclusion criteria for radiologists were a minimum of 10 years of experience in spine imaging interpretation. In addition, each participant was required to hold a doctoral degree and to be actively working in a diagnostic radiology capacity. Each radiologist was required to be board-certified in a field related to the topic of the study. Study exclusion criteria for the AI expert and the radiologist included a history of or current employment with the Chicago Neuroscience Institute (CNI) or the American Academy of Spine Physicians (AASP), both of which I am affiliated with. I prohibited key contacts for research participant recruitment from participating in the study. I also prohibited professionals who played a role in field testing of focus group questions and strategies from participation in the research study.

Data Collection Instruments and Processes

The use of predetermined or validated data collection instruments can help direct the data acquisition and data management processes. I acquired and developed numerous data collection instruments for use in this research study. This included the use of qualitative research software, a brief qualitative survey, and a focus group moderator guide. I developed a brief survey and gave it to each participant to complete prior to participation in a focus group session (Appendix B). This document and the consent form were used to confirm that the various experts met the criterion for participation. I developed a focus group moderator guide to help manage time, topic discussions, and the method of inquiry during the focus group sessions (Appendix F). Published research has demonstrated that the use of a focus group moderator guide helps ensure efficient and systematic in-depth coverage of research topics and related questions (Fraenkel & Wallen, 2003). The moderator guide I used in this study consisted of carefully crafted open-ended questions and probing semistructured questions.

I developed questions for the focus groups with the assistance of insight acquired from an extensive literature search and from the needs analysis document provided by the AASP Spinecare Data Science Committee (Appendix F). Each question was placed in the focus group moderator guide. My development of the moderator guide was overseen by independent expert prior to and after field testing. I made all necessary changes to the guide. This iterative process of assessment helped me reduce the risk for an inappropriate or biased approach to research question development and delivery. It also helped me refine the methods of inquiry I used during each focus group session. The focus group moderator guide consisted of an agenda, an introduction, along with a list of PowerPoint concept slides, and research questions followed by closing remarks. (Appendix D). The moderator guide identified the order of topic presentation and inquiry. I used the guide to help set the tone for each focus group session and for guiding the order of the process.

Consistent with the recommendations of C. L. Lee et al. (2015) I developed data coding guidelines to help ensure analytical and categorical consistency during content and thematic analysis (Appendix C). My coding guidelines consisted of predetermined codes capable of being adapted, modified or replaced during data analysis. I developed a priori codes consistent with the conceptual framework of the study utilizing theoretical perspectives from DOI and TAM. The a priori codes aligned with the research topic, research purpose, and research questions. During the course of the entire research process, I made regular entries in a reflective journal to memorialize my biases, impressions, and insights.

I used numerous expert documents (white papers) to help identify current consensus-based opinions and positions surrounding the use of AI in radiology and oncology. At the time of this study, there were no white papers published on the potential role or impacts of AI in spine care. I used a few published papers, which addressed narrow applications of AI in spine care. The consensus-based “white papers” I used for this study were published by nationally and internationally recognized organizations: the Canadian Association of Radiology, the American College of Radiology, the French Radiology Community, and the European Society of Radiology. I also used seminal publications of leading experts in the field. I analyzed, thematically coded, and triangulated the content of the expert documents with data from other sources to improve the consistency and relevancy of the study’s conclusions. This methodical and transparent analysis process helped improve the trustworthiness and validity of the study. The document provided by the AASP Spinecare Data Science Committee offered a list of potentially meaningful applications of AI during the interpretative stage of spine imaging workflow. The AASP document was developed by a multidisciplinary group of spine care experts independent of the research process. I used this document along with insight acquired from an extensive literature search to guide the development of research questions I used in the focus group sessions.

The research process was not be limited by fixed guidelines or rules, subsequently allowing for inductive assessment of emerging topics and trends. As stated previously, focus group research represents a well-established and disciplined scientific method for acquiring in-depth insight surrounding the use of new technology and related processes (Krueger & Casey, 2015). I applied the concept of data saturation during focus group sessions and during thematic data analysis of expert documents. I used an introductory PowerPoint slide program at the beginning of each focus group session (Appendix E). I conducted topic-specific discussions during the focus group sessions until reasonable topic saturation was achieved. I arranged for a recording of all of the contributions during each focus group session. I also arranged for verbatim transcription of the recorded sessions to avoid misinterpretation or misrepresentation. I performed document analysis until I achieved topic saturation. I triangulated the data from the different research sources to improve the internal, as well as external validity of the study. A concise and comprehensive informed consent form was developed and used to protect the rights of participants and to encourage unfettered contribution to the research process.

Document review and analysis offers a unique and often critical contribution to qualitative exploratory research (Creswell, 2013; Patton, 1999). I used position papers and consensus-based summaries published by reputable organizations and highly regarded experts in this study to help address the potential impact of AI during spine imaging workflow. I compared and contrasted the data acquired through expert documents with data acquired through other expert sources such as the focus group sessions and reflective journaling to improve the internal validity and transferability of the results.

Data Acquisition

I initiated the data collection and analysis process after I received dissertation proposal approval and Institutional Review Board (IRB) approval from Walden University. I created a data acquisition flow diagram to help guide me in the research process. The method of inquiry I used in the focus group session was field tested with two independent experts, one meeting AI expert participation criteria and the other radiologist criteria. I did not accept the experts who assisted me with field testing as research participants. I used field testing to evaluate the focus group protocols and strategies I used in the Focus Group Moderators Guide. I made minimal modifications, as a result of the field testing.

The focus group sessions each lasted approximately 90 minutes. I achieved an acceptable degree of data and topic saturation in each session. I arranged for each focus group session to be digitally recorded, transcribed verbatim, and stored securely. Emotional responses, body language, and nonverbal forms of communication between the participants in a focus group setting can be important (Bunnick et al., 2017). I subsequently recorded any participant behavior during the focus group sessions I felt was relevant to the study purpose. I led each focus group session with the assistance of the moderator guide. I developed my focus group approach guided by Krueger’s categorical strategies that included the use of an opening question, introductory questions, transitional questions, and probing questions (Krueger, 2000). I used probing questions to help actively engage each of the participants in topic discussions.

Scholarly publications have established that reflective journaling offers the researcher with a powerful inductive method for recording observations ideas, and insights using an active voice (Janesick, 2011). I performed reflective journaling during the course of the research process. Journaling served many purposes. It allowed me to identify my initial and evolving perspectives, expectations, and biases associated with the research topic and the research process. Review of journal entries gave me the opportunity to engage a higher level of critical thinking and implement methods to reduce my personal influence on the research process and outcome. Journaling included my impressions of verbal, as well as nonverbal communication during focus group sessions. This included recoding of body language and expressions. I also recorded the level and nature of agreements or disagreements that occurred during each focus group session.

Research Participant Debriefing and Follow-Up

At the completion of each focus group session, I reminded participants of the purpose of the research study and informed them how I would manage and analyze the data acquired. In informed the research participants that they would receive an overview of the focus group data analysis in the form of a thematic summary and a list of supportive quotes for review, a process referred to as member checking or respondent validation. In addition, I informed each research participant that the records of the research study including their consent forms would be stored in a secure location for a minimum of 5 years, after which time they would be properly destroyed. I informed each participant they would receive notice when the dissertation was published. I also assured each research participants that they would be provided with access to the published work when it was available.

Data Analysis

I used a multistep process to analyze acquired data. I implemented an inductive and iterative data analysis process as soon as data was acquired. My analysis process continued throughout the entire research study. I recorded a chain of evidence to memorialize the process and analyzed the focus group transcripts with an exhaustive, inductive, and iterative process of coding for themes. Descriptive codes were clearly established and defined consistent with the work of Glaser and Laudel (2013). I used a hybrid approach to coding, allowing for aggregation, subtraction, combining, and expanding of code categories when necessary. Qualitative data coding offers an effective method for revealing emergent ideas, themes, and relationships (Rubin & Rubin, 1995; Strauss & Corbin, 1998). My evaluation of the focus group transcripts included content analysis and thematic coding. Content analysis is used in the social sciences and in qualitative research to structure information (Krippendorf, 2004). The analysis of focus group data should include identification of noteworthy quotes, as well as identification of outlying factors and unexpected findings consistent with published works (Breen, 2006). I identified and labeled all noteworthy findings, comments, and quotes derived from my assessment of the various sources of research data. I performed the data coding process until topic saturation was achieved.

The first step in the coding process was to become familiar with the data and systematically reduce its complexity. I developed a few provisional (a priori) codes to initiate axial coding. I developed the provisional codes with insights acquired from my extensive literature search along with the influence of theoretical constructs from DOI and TAM. Some of the provisional codes aligned with theoretical constructs of DOI and TAM, such as relative advantage, interoperability, complexity, ease-of-use, and perceived usefulness. I replaced, revised or modified many of the provisional codes during data analysis to better describe and label acquired data. I expanded, contracted, replaced, and modified the coding categories many times throughout my analysis process. Thematic coding arose from the integrated applications of provisional coding, open coding, in vivo coding, axial coding, and selective coding.

My analyses of focus group data included within and between group analysis. I took into account the unique experiences and backgrounds of the participants in each focus group session revealed by their demographic surveys. I displayed the results of my analysis of the acquired research data in many different ways, including a contrast table. Contrast tables offer an effective method for looking at relationships between exemplars, extremes, and outliers (Miles, Huberman, & Saldana, 2014). Combining content analysis and thematic coding supported the development of a concept map depicting the potential relationships between processes and technologies, used during the interpretive stage of radiology workflow. I created a concept map to reveal relationships between the flow of data and processes during the interpretive stage of radiology workflow. I used the concept map to help transform tacit knowledge into a practical resource and a foundation for further discussion and research.

Concept Mapping

Data analysis in qualitative research involves many steps that include data reduction, data organization, data interpretation, and data display. Concept mapping has been successfully used to graphically organize and depict relationships between elements of a system or process (Baugh, McNallen, & Frazelle, 2014). This includes the relationships between data, individuals, and technology (Baugh et al., 2014). Concept mapping has also been used in qualitative research to reveal themes and to depict workflow (Daley, 2004; Novak, 1998). One of my goals in this research study was to identify themes that could be used to create one or more concept maps depicting the potential role of AI solutions during the interpretive stage of spine imaging workflow. A concept map helps depict stages of a complex process and reveal technological relationships to achieve desired goals (Daley, 2004; Novak, 1998; Wheeldon & Faubert, 2009). I subsequently developed a concept map to reveal the flow of data and role of potential AI applications during the differential diagnostic process associated with interpreting spine images.

I used concept mapping in this study to facilitate a shared vision, to help direct subsequent research, and to inform further technology development. I also used it to help determine how to embed AI technology into existing radiology workflow. A concept map can be augmented with the use of numerous elements such as linked tasks, labeled processes, communication pathways, and hierarchies of priority (Wheeldon & Faubert, 2009). In my concept map I used map lines to link mapped elements and I used directional arrows to reflect the flow of data and/or the implementation of a process. I developed a concept map in this study to complement and enhance textual conclusions. This helped me present research findings in an accurate, concise, and effective manner.

Trustworthiness of the Study

The trustworthiness of a research process is influenced by the role of the researcher, the source of the data, the management of the data, and the approaches used to improve study validity and reproducibility (Connelly, 2016; Mays & Pope, 2000; Shenton, 2004). Qualitative exploratory case study research is inductive and subjective and therefore requires high levels of trustworthiness, reliability, and validity to be influential (Creswell, 2013). The attributes of validity and reliability are operationalized in different ways in qualitative versus quantitative research studies (Mays & Pope, 2000). The primary risks associated with qualitative case study research include over generalization of results, researcher bias, inadequate interpretation of data, poor integration of data, and research question mismatch with methodology. I took extra precautions and implemented steps throughout this research process to improve the trustworthiness of the study and its conclusions (Figure 5).

Qualitative studies that include the use of focus group sessions must meet extremely high standards to be reliable and valid. Research study trustworthiness is determined by its credibility, confirmability, dependability, and transferability (Lincoln & Guba, 1985; Shenton, 2004). I addressed each of these elements in this study along with reliability. I implemented numerous steps d to reduce the risk for interjecting personal bias and to support evidence-based conclusions.

Credibility

I implemented numerous steps to improve the credibility of this research study. I used respondent validation also referred to as member checking to help confirm the accuracy of my thematic conclusions. I also provided research participants with a list of the supportive quotes I acquired from the focus group transcripts. Member checking is an important step in qualitative research, because it provides participants with an opportunity to affirm the accuracy of focus group data acquisition, analysis, and interpretation (Creswell, 2013). Member checking in this study served as an effective method for establishing interpretive and descriptive validity. It also helped reduce the impact of my personal bias as the sole researcher.

To help further reduce personal bias during data acquisition and analysis, each focus group session was recorded and transcribed verbatim. I used a few open-ended questions during the focus group discussions to help reduce the risk of framing bias. I read the focus group transcripts numerous times to ensure comprehension of the material prior to initiating descriptive labeling and coding of data. I used an inductive and iterative process of hierarchical coding to avoid rigid misclassification of data.

I used a well-defined unit of analysis to help direct the research process and the flow of data. This approach improved study credibility. Consistent with the work of Mays and Pope (2000) I performed reflective journaling to expose how my role as the sole researcher may have influenced the research process and outcomes. I used reflective journaling to record my thoughts and opinions during the entire research process. The journaling process helped reveal my perspective and beliefs and how they may have had an impact on data analysis, research design, and research conclusions.

Triangulation of data acquired from different sources improved the credibility and internal validity of this research study. Published work has demonstrated that the triangulation of data acquired from a diverse set of expert sources contributes to the authenticity, plausibility, and validity of qualitative research (Greenlaugh & Singlehurst, 2011). Triangulation of acquired data from the focus group sessions, reflective journaling, and from consensus-based white papers in this study improved the credibility of the research conclusions. My use of theoretical constructs from DOI and TAM combined with insights acquired from an exhaustive literature search helped reduce personal bias during my formulation of research questions and the interpretation of the responses. In summary, the methods I used to improve the internal validity of this study included reflective journaling, respondent validation (member checking), inductive coding, and triangulation of data from diverse sources.

Transferability

Transferability refers to the ability of a reader to apply the research process or the research results to another situation or setting. The success of this process is dependent on transparency and adequate description of research boundaries, parameters, and processes (Connelly, 2016; Lincoln & Guba, 1985; Shenton, 2004). In contrast to transferability, generalizability refers to the ability to apply the results from a research sample to a broader population. In order for research to be transferable, the results must be reproducible in different cultural settings with common variables (Shenton, 2004). In this study, I used sequential steps, thick descriptions, a data acquisition flow chart, a focus group moderator guide, and qualitative coding software, all of which contributed to a high degree of transparency, which supports transferability**.**

I used numerous redundant and overlapping methods to improve external study validity and transferability. I disclosed unexpected and conflicting results along with unforeseen challenges in the research. My research conclusions include alternative and rival explanations surrounding the potential impact of AI use during the interpretive stage of spine imaging workflow. The use of two focus group sessions each comprised of homogenous groups of experts from two related fields supported the detection of patterns and themes within and across groups. I compared the findings of the focus group sessions to the themes that emerged from consensus-based white papers, which served as research documents in this study.

Dependability and Confirmability

The attributes of dependability and confirmability are important elements of trustworthiness that influence the ability to replicate the research process and to test related assumptions. The dependability of qualitative research improves with transparent strategies such as thematic coding, content analysis, and the generation of thick descriptions (Shenton, 2004). My use and disclosure of a field-tested focus group moderator guide and data-coding guide offered the level of transparency required to facilitate accurate interpretation and/or replication of this research. I used valid tools and measures available on the Atlas.ti, Version 8 software, to analyze and manage the research data in this study. I made journal entries of influential issues surrounding the integrity or quality of data used for analysis.

I used clear and concise descriptions of the research process and the flow of data to improve the ability for others to critique or replicate the study. I also recorded the chain of evidence and performed content analysis and thematic coding, which I acknowledged in detail with the help of a hierarchical coding table. I developed a data-coding guide that includes a list of code categories and their definitions and the criteria and methods I used to achieve and define data saturation during the analysis process. I performed within and between case analyses with the help of computational methods to reduce my potential bias.

Ethical Procedures

The basic ethical principles acknowledged by the Belmont Report (1979) are respect for individuals, beneficence, and justice. Beneficence refers to the treatment of individuals in an ethical manner by respecting their decisions, securing their safety, prioritizing their well-being, and protecting them from harm. Justice refers to the equal and fair management of research participants. To confirm adherence to these basic principles I treated each research participant equally. I provided each research participant with the same documents, had them sign the same consent forms, and exposed them to the same data collection processes. The research protocol and conduct in this study conformed to the tenants of the Belmont Report and to Walden University IRB requirements.

Documents and Agreements

I performed the research in an ethical and honest manner to ensure the integrity of the study and to minimize any potential harm or risk to research participants. I used predetermined protocols and preapproved documents helped to ensure that an appropriate and ethical approach was used throughout the entire research process. Research in the Walden University doctoral program requires oversight by the IRB to help ensure the integrity of the research process, as well as the safety and privacy of all research participants. The Walden University IRB approval number assigned to this study was 02-13-19-0129405.

Prior to collecting the data, I provided each research participant with preapproved documents, which included an overview of the study, a focus group agenda, participant expectations, and an informed consent form, which included confidentiality terms. The documents safeguarded the consistent and ethical treatment of each research participant and the ethical management of research data. The agreements disclosed any anticipated or potential exposure to risk. The documents also acknowledged the voluntary nature of study participation.

All of the research participants were required to sign an IRB approved consent form prior to participating in the study. The form included confidentiality agreements. Consent included my responsibility as the researcher to keep confidential the personal identities and contributions of all research participants. I informed all participants of the purpose and scope of the study, as well as the expectations for their participation. In addition, I informed the research participants that they would each receive a nominal stipend of $25 for participating in the study.

Treatment of Research Participants

I treated all of the research participants with the utmost respect. Research participants should also be treated as autonomous agents (Kaiser, 2009). The safety and rights of research participants must be prioritized at all times (Belmont, 1979). In qualitative research, the researcher assumes a unique responsibility for protecting each research participant (Orb, Eisenhauer, & Wynaden, 2001). Compliance with well-established ethical principles such as autonomy, beneficence, and justice helps ensure proper care of research participants (Lorell et al., 2015; Orb et al., 2001). Consistent with the recommendations of Kaiser (2009), the methods and forms I used to obtain informed consent in this study were adapted to the type of research and type of research participants required. My use of field testing and an independent review of the focus group moderator guide helped guarantee appropriate treatment of research participants in this study.

I informed all of the research participants of their right to withdraw from the research study at any time and for any reason. In addition, I provided participants with the option to withdraw verbally or in writing from the study without any repercussions. I informed all of the research participants that their names would be kept confidential and their identities would remain anonymous. I removed research participant names from focus group transcripts and replaced them with unique and anonymous identifiers to help ensure confidentiality of their names and contributions. Examples include Participant 1 (P1), Participant 2 (P2), and so forth.

Treatment of Data

To help safeguard trust and to protect the privacy and contributions of each research participant during the focus group sessions, participants were informed that they are not to disclose the names or contrition of participants outside the research setting. I deleted all names from the final focus group transcripts. The research records will be securely stored for 5 years from the time of completion of the research study to protect the rights of all participants. Proper storage of research records will support authorized access for auditing or for review by qualified individuals. I will take proper steps to discard all participant records after 5 years.

The Potential for Research Impact on Social Change

The primary purpose of this research study is to explore the potential impacts of AI on the interpretive stage of spine imaging and to reveal its social implications. The study addressed the potential influence on standards of care, technology development, systems applications, and public expectations. In Chapter 5, I expand the discussion of the research results to include it potential impacts on various levels of society.

Meaningful use of AI during the interpretive stage of spine imaging will augment the role of the radiologist by reducing data complexity, characterizing pathology, and offering decision support. The process will empower the radiologist as a gatekeeper of big data and facilitate their leadership role. Improved availability of AI decision support will increase the demand for remote access teleradiology services. The process has the potential to support improved democratization of decision support surrounding diagnostic image interpretation. It will subsequently offer a potential solution for underserved professionals, facilities, institutions, and geographic locations.

Widespread use of AI use during radiology workflow may alter the health care landscape, especially in the areas of image analysis, disease characterization, disease monitoring, decision support, and final report generation. AI could offer radiologists new solutions, capable of improving their ability to detect early stage pathology. Most diseases are recognized at advanced stages, thus, resulting in high costs and poor treatment outcomes. Early disease detection would contribute to more efficient care at lower costs. These outcomes would all have a favorable social impact at many levels.

Successful use of AI during the interpretive stage of radiology workflow will require redesign and co-evolution of supportive technologies to benefit the broader field of health care and society. Supportive solutions will include new levels of interoperability between databases and management systems (Tang et al., 2018). A successful co-evolutionary process will require the development of unifying platforms, which facilitate sharing of data and support more consistent use of disease criteria, disease classifications, and computational disease models. The summary discussion addresses the potential relationship between AI and relevant emerging technologies. For example, block chain technology has the potential to provide proof-of-work validation while recording the flow of data and computational steps across an AI-based network (Kuo, Kim, & Ohno-Machado, 2017; Mamoshina et al., 2018). The eventual convergence of AI and block chain technology has the potential to decentralize intelligent decision support and offer increased access to computational disease models.

In summary, AI decision support could overcome human bias, reduce interpretive error, and enhance the potential for a more precise and timely diagnosis in spine imaging. Meaningful use of AI during the interpretive stage of spine imaging and at other levels of radiology workflow could have a favorable impact on the role of radiologists, as well as on the co-evolution of decision support technology, standards of care, delivery of care, and public expectations. I address the potential social consequences of AI development and use in spine imaging in the conclusion of this study.

Summary

I designed this research study to identify how the use of various AI solutions could impact data management and the differential diagnostic process during the interpretive stage of spine imaging. AI involves a rapidly evolving set of technologies associated with numerous processes. Its role in radiology is difficult to define because of its wide range of potential applications. It was necessary to address the potential impact of AI with a qualitative exploratory case study approach to identify possibilities worthy of further discussion and investigation. The design of this study supported the acquisition of expert insights and data from different sources. This research design also supports the development of a concept map representing potential AI applications and contributions during the interpretive stage of spine imaging workflow. The research design and methods I detailed in this chapter were supported by expert sources, an extensive literature search, and a review of consensus-based documents surrounding the use of AI in radiology.

An individual radiologist can no longer be required to function with precision accuracy in the face of overwhelming, high velocity, and complex data. Radiologists require new data analysis and decision support systems during the interpretive stage of imaging in all fields including spine care. Successful use of AI will likely result in earlier disease detection, better disease characterization, less diagnostic errors, and shorter lengths of care (Kohn et al., 2014; Lee, 2017).

The research methods introduced in this chapter provide a trustworthy approach and a scholarly foundation for further discussion and research surrounding the development and use of AI during the interpretive stage of spine imaging. I designed this research study to introduce the role of radiomics and the concept of the digital (virtual) biopsy, in a manner that could be applied in spine care, as well as in other fields of health care.

There is a growing demand for health care to become more predictive and preemptive. Success requires a more deliberate approach to the comprehensive and objective analysis of actionable data. My acquisition and analysis of data in this research study revealed concepts and themes that can be used to develop additional research strategies to pursue the role of AI solutions during the interpretive stage of spine imaging workflow. Discoveries associated with this research can be applied to other areas of radiology.

Chapter 4 presents the results of the research study. I further discuss how I acquired and analyzed the data and how I improved the trustworthiness of the study and related data.

Chapter 4: Results

Introduction

The primary purpose of this qualitative, exploratory case study was to explore the potential impacts of artificial intelligence on spine imaging interpretation and diagnosis. I designed the study to acquire and analyze expert opinions from different sources. My goals for the study included identifying how AI solutions might improve the accuracy and efficiency of interpretive workflow and the differential diagnosis process in spine imaging. I implemented qualitative research methods to explore the possibilities associated with computational decision support and to establish a thematic basis for further discussion and research on the topic.

This study is one of the first to address the potential role of AI in spine care and the concept of the digital (virtual) biopsy characterized by multiscale in vivo interrogation of pathology. I initiated this study with the fundamental belief that images are rich in metadata and that diagnostic imaging represents a core diagnostic process in spine care. Select constructs of the TAM and DOI were used to guide the process of data acquisition and analysis. During focus group sessions, I asked open-ended and probing questions to address radiomics, interpretive workflow, the differential diagnostic process, clinical utility, and determinants of AI adoption and use.

The volume and complexity of data acquired with advanced diagnostic imaging methods has created a burden and exposed unprecedented opportunities for radiologists. Big data has exceeded the ability of a radiologist to make fully informed decisions (Aerts, 2017; Gilles et al., 2016). Radiologists require augmentation of their role to reduce errors and to take advantage of new opportunities for rendering a more precise and personalized diagnosis. Without adequate technological assistance, the human interpretive process within radiology workflow will become progressively more inaccurate, inefficient, and untimely (Croskerry 2013; Manrai et al., 2014; Obermeyer & Emanuel, 2016; Ragupathi & Ragupathi, 2014). Diagnostic imaging represents one of the single most important methods for detecting and characterizing pathology in all biological systems including the spine. This study addresses the potential for AI to reveal actionable data from imaging studies while augmenting the role of the radiologist. My primary motivation for performing this exploratory study was to acquire insight and provide direction for the development of decision support solutions to support better spine care.

In this chapter, I address numerous topics such as the research purpose, the research setting, field testing, research participant demographics, data collection, data analysis, trustworthiness of the study, and the research results. I laid out this chapter in a manner consistent with the chronological stages of the research process. In this chapter, I reveal the themes and subthemes which emerged from triangulation of data and data analysis. I also provide supportive evidence for the iterative process. This chapter indicates how the research design and the use of strategic methods improved the trustworthiness of the research results. In addition, I provide an overview of my reflective journaling, which includes disclosure of its impact on the research process and results. The conclusion provides a summary of the research findings along with a transition to Chapter 5.

Field Testing

I field tested the moderator guide to help establish the required level of appropriateness, clarity, and relevance of the approach used during focus group sessions. I achieved these goals through independent review of the strategies and resources outlined in the focus group moderator guide, including the agenda, topic introduction, ground rules for participation, open-ended questions, and a script for the conclusion of the session. I also field tested the appropriateness and clarity of my introductory PowerPoint slides I used during focus group sessions.

I performed the field testing, on separate occasions, with one radiologist and one AI expert. The experts who participated in the testing met study inclusion criteria but did not serve as research participants. The field testing process allowed for peer-review of the focus group protocols and resources. The experts who participated in the field testing were not asked to answer or respond to any research questions. I did not have an employment or consulting relationship with the field testing experts.

I used field-testing to help establish appropriate and relevant focus group protocol. The process included assessment of the methods of data acquisition and the pattern of inquiry. I made no significant changes as a result of field testing, with the exception of the order and clustering of questions to be used during the focus group sessions. Nor did I make contextual revisions to the primary focus group questions. I field tested which PowerPoint concept slides were the most neutral and concise to help guide focus group discussions on complex topics. I removed a few slides from the presentation. I made no content changes to the remaining PowerPoint slides.

My field testing helped identify the order of the focus groups. I determined that the radiology focus group session should take place first, followed by the AI expert focus group session. The independent experts believed this order would support more progressive, in-depth coverage of the research topic. I thought it was necessary to utilize field testing to reduce bias, ensure professionalism, and to efficiently operationalize available multimedia and data acquisition methods.

Research Setting

I held the focus group sessions at an independent and professional location. The setting supported physical participation and teleconference access for all participants. I provided each of the research participants the option to be physically present during the focus group session or to access the session using Zoom, an independent well-established teleconferencing solution. Each research participant chose to access the focus group session with the teleconferencing solution.

During the live focus group sessions, each participant using Zoom had access to an online image gallery for intimate real-time viewing and interaction with all participants. Each participant also had the independent option of engaging a speaker highlight function that prioritized the participant actively contributing. I provided each participant with access to a dynamic online gallery to facilitate efficient communication. I served as the sole moderator for each focus group session. I moderated each session from a conference room designed for focus groups. The room consisted of a boardroom table and chairs, professional audio system, and a large format wall-mounted screen with a camera.

I made sure that the focus group settings and technologies were used as planned and as approved. The method for accessing the focus group sessions was clear and consistent during the study. I did not make any changes in the research setting that would have interfered with the research process or with the contributions of those who took part. Participants and I experienced no technical difficulties during either of the focus group sessions. There were no changes in budget or technical support made during the research process.

A few conditions in the research setting may have influenced focus group participants. For example, some of the participants may have been somewhat unfamiliar with the technical element of the video conference process, which could have led to initial confusion or hesitancy during the early stage of each session. This possibility did not have had any obvious impact on the focus group discussions or the acquisition of data. The low stipend of $25 offered to research participants necessitated the need to offer teleconference options in order to attract renowned experts from various geographic locations throughout the United States and Europe. The participants did not acknowledge the low stipend as a barrier to participation. In fact, some of the experts refused to accept any stipend for their participation.

Demographics

Qualitative researchers need to acquire demographic information that describes the individuals who take part in a study. Examples of demographic information include gender, educational status, employment status, expertise, and duration of experience. In this study, I disclosed basic demographic information about myself and the research participants who served as focus group participants. This offered contextual insights about the source and management of data. The disclosure of demographic information was designed to help others critique, interpret, and duplicate the research study. I placed the demographic information into one of two categories: researcher demographics and participant demographics.

I served as the sole researcher in this study and as the sole moderator during the focus group sessions. At the time of the study, I was a licensed and practicing board certified chiropractic neurologist with more than 25 years of clinical experience. I had extensive experience in the interpretation and clinical correlation of spine imaging results along with academic experience in molecular imaging and the use of AI in neuroimaging. At the time of the study, I was serving as director of the CNI and as president of the AASP. My academic knowledge of AI combined with my training and clinical experience in neurology and neuroimaging, offered me a unique position for acquiring and analyzing data and for moderating the focus group discussions with AI experts and radiologists.

I purposely selected prominent experts from the fields of AI and radiology to participate in one of two homogenous focus group sessions. My purposive recruitment methods helped to assure that participants had the level of experience and expertise required to contribute to research topic discussions. I recruited research participants through personal contact and with the assistance of key contacts in the respective fields. I made some of the initial contacts at professional symposia such as at the annual convention of the Radiological Society of North American (RSNA).

I contacted a total of 23 AI experts and radiologists and invited them to participate in the focus group sessions. Eleven of the individuals I contacted consented to participate. Of the 11 individuals who agreed to participate, nine were able to participate in the focus group studies on the scheduled dates. Originally, six AI experts consented to participate in the focus group session. Unforeseen circumstances prevented two participants from attending . Subsequently, the AI focus group comprised four participants, which met the minimum research requirements for the study. The radiology focus group comprised five participants, which met the research requirements for the study. All of the radiologists who committed to participate attended the focus group session.

The pattern of participation was similar between the AI and radiology groups. Approximately half the professionals invited to participate in this study from each group consented to do so. Among reasons potential participants chose not to participate were limited available time and limited knowledge of the research topic.

All of the radiologists and AI experts who participated in the focus group sessions met research inclusion criteria and were highly qualified to address the topic of study. The research participants collectively represented the spectrum of expertise and experience required to explore the research topic. Five participants were board-certified radiologists and four participants were AI experts. The radiology group consisted of four men and one woman representing four board-certified medical radiologists and one board certified chiropractic radiologist. Two of the medical radiologists had specialized training and certification in neuroradiology. All of the radiologists who participated in the study had a doctoral degree, were board-certified in radiology, and were actively working as a radiologist within or for a clinical setting at the time of the study. The AI group consisted of three men and one woman. All of the AI experts had a minimum of 2 years of experience with AI in radiology, a minimum of bachelor's degree in a related field such as informatics, data science, or computer science, and were actively working in AI-related health care field at the time of the study.

I implemented numerous safeguards to ensure that the demographic information of each research participant remained confidential. I assigned each research participant with a unique identifier for use on audio transcripts and in this published work to prevent identification.

Prior to participating in this study, some of the research participants may have been aware of my experience and role in neurology and spine care. I have spoken at many national venues, have held a number of prominent positions, and have numerous publications in the fields of neurology and imaging. It is also possible that some of the participants were familiar with my neurology textbook titled *Myelopathy, Radiculopathy and Peripheral Entrapment Syndromes.*

Data Collection

I collected data from three primary sources. The first source consisted of four published expert documents, each representing consensus-based white papers, which addressed current and future roles of AI in radiology. Each of the expert documents were published by highly respected national and/or international radiology associations or societies. The white papers used as expert documents this study were published by the Canadian Association of Radiology, the American College of Radiology, the French

Radiology Community, and the European Society of Radiology. The second category of data collection consisted of two homogenous focus group sessions. The first focus group session was comprised of five radiologists. The second focus group session was comprised of four AI experts. The third category of data collection was reflective journaling, which I performed during the research process.

I used a predefined sequence of data acquisition, data analysis, and process validation during this research study (Figure 4). I initially acquired information from the AASP Data Science Committee in the form of a committee summary, representing consensus opinions to help inform data acquisition strategies (Appendix F). The committee comprised a multidisciplinary group of spine care providers not employed by the academy.

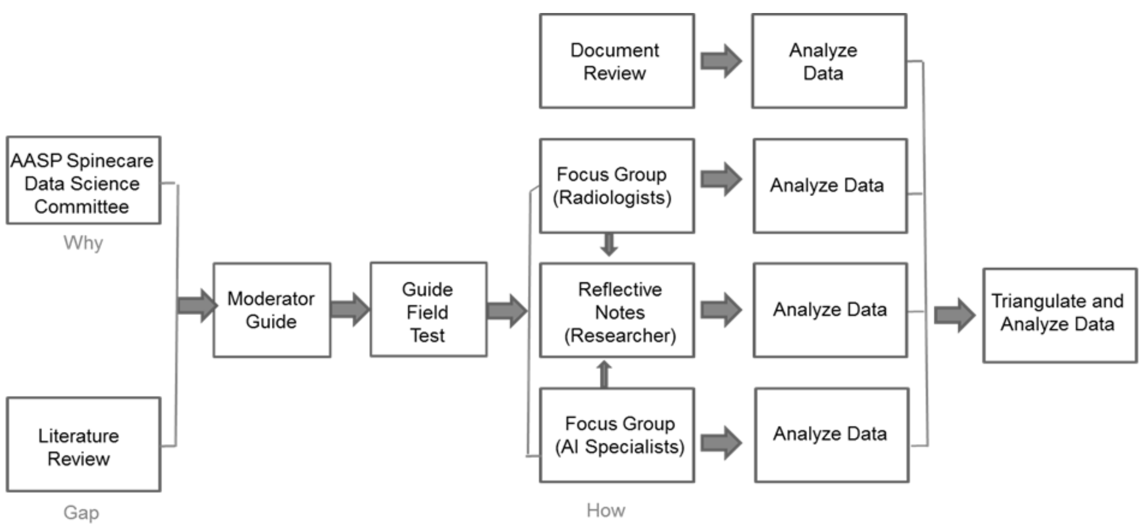


Figure 4. Steps in the research process. The arrows depict the flow of data.

I performed an exhaustive literature review prior to the research study to reveal gaps in knowledge surrounding the potential roles and impacts of AI use during spine imaging workflow. In addition, I developed and field tested a moderator guide prior to acquiring data from the focus group sessions Research data was acquired from consensus-based white papers, reflective researcher notes, and two focus homogenous focus group sessions. I analyzed data from each source and triangulated the data for further analysis.

I acquired the research data in six overlapping phases. The first phase consisted of an exhaustive literature review leading to topic saturation. This phase of data collection spanned about two years and influenced the dissertation research design. The second phase of data collection consisted of obtaining a brief consensus opinion from the AASP Spinecare Data Science Committee regarding the potential applications of AI in spine imaging. The duration of this period was approximately one month. The third phase of data collection consisted of accessing four consensus-based white papers published by reputable radiology organizations. This process took place over a period of approximately three months. The fourth phase of data collection consisted of obtaining brief demographic surveys from each of the research participants. The duration of this phase of data collection was approximately four weeks. The fifth phase of data collection consisted of performing two independent focus group sessions, which took place over a period of two weeks. The sixth and final phase of data collection consisted of completion of a reflective journal, the duration of which overlapped all of the other phases of data collection.

I commenced research participant recruitment and data collection after receiving institutional review board (IRB) approval. My approval included acceptance of the participant recruitment process and the participant consent process. I sent each potential research participant a consent form. I sent each participant who returned a signed consent form a brief survey to complete and return prior to the scheduled focus group session. I designed the survey to acquire demographic information, as well as an overview of the participant's relevant level of experience and expertise on the research topic.

I acquired focus group data from nine participants. The first focus group session was comprised of five radiologists. The second focus group session was comprised of four AI experts. Each focus group session lasted 90 minutes. Pre-focus group activities included scheduling a time and place for each session. I facilitated each focus group session with the help of a field-tested moderator guide. I informed the research participants about focus group ground rules at the beginning of each session. I provided each focus group participant with the same agenda, the same instructions, and essentially the same primary open-ended research questions. I used a few tailored semistructured probing questions during each focus group session to facilitate topic discussion surrounding open-ended questions. I moderated each focus group session from the same physical location. I used the same teleconference recording technology and protocols during each focus group session.

The two focus group sessions consisted of scholarly interaction and a shared exchange of concepts and opinions. The resources I used to facilitate individual contributions and group interaction included advanced videoconferencing tools, a field-tested moderator guide, a concise introductory PowerPoint program, and a select group of neutral PowerPoint concept slides. I used the latter to help focus the discussion of complex subjects. All of the participants openly contributed during each focus group session. I achieved topic saturation for each open-ended question presented during the focus group sessions. I was prepared to hold additional focus groups if necessary. The study had international representation. I conducted each focus group session in English. All participants were fluent in English. I arranged for each focus group session to be recorded and transcribed in their entirety. I analyzed each focus group transcript for relevant concepts, patterns, and themes.

I served as the sole moderator for each focus group session. I set a friendly and informal tone, which motivated open and progressive discussions. I successfully acquired complete answers and in-depth participation by probing with semi-structured questions. I developed some questions in advance to help facilitate discussion during anticipated contributory downtime and made sure each focus group participant had a chance to participate. I did not allow any one professional to dominate a focus group session.

Data Recording Methods

I arranged for a complete recording of each focus group session, each of which was recorded in real-time using secure Zoom cloud-based technology. I also arranged for a backup recording of each session using technology in the conference room. I forwarded Zoom audio files in their entirety to a contracted transcriptionist for final document production. I also used Zoom audio files to create a voice actuated machine-based transcript that served as a backup and compared the final transcribed documents with the automated Zoom file transcripts to check accuracy. I deleted the focus group audio files from the Zoom cloud service in their entirety after cross checking the focus group transcripts and confirming the accuracy of the final documents. I retained and stored a copy of the audio files in a secure location consistent with my IRB requirements.

The Zoom teleconference solution I used for this study complied with SOC 2, the de facto assurance standard for cloud service providers. At the time of this study, Zoom conformed to a high level of privacy practices and technical security measures. The data in transit were protected by TLS 1.2 and at rest using 256-bit advanced encryption standards (AES-256). The only individuals who had access to the Zoom audio files from the focus group sessions were the transcriptionist and me. I granted the transcriptionist with live access to each focus group session to expose her to the context of the discussions. I thought access would result in more accurate transcription. I informed the participants of each focus group of the transcriptionist's presence, although I blocked her visual profile and did not allow her to participate in the discussions. I stored the audio and digital versions of the transcripts on a password-protected computer and the printed versions of the transcripts in a safe fireproof cabinet consistent with University IRB requirements.

The option of accessing the focus group session through a teleconference solution reduced the participant's costs for attending the session, thereby providing access to world-class experts. Some of the disadvantages of this approach included the possibility of greater difficulty moderating and controlling the group, although this did not become an issue. I recognized that it was impossible to overcome all potential disadvantages of a focus group approach. Subsequently, I established focus group ground rules, tailored to the teleconference setting, to improve the data acquisition process. The focus group approach supported r collective knowledge construction and predictions that would have been difficult to achieve through individual interviews.

I implemented numerous steps to help ensure the privacy of each research participant during the data collection process. I removed personal identifiers from focus group transcripts and replaced them with an anonymous identifier consisting of the expert class combined with a participant number. The transcriptionist signed a confidentiality form prior to participating in the study and turned over all audio files and transcripts to me. I managed the focus group sessions in a deliberate, sequential, and purposeful manner consistent with the recommendations of Kruger and Casey (2015).

Variations in Data Collection

I originally anticipated the physical presence of one to three research participants in each focus group session with the rest accessing the session via the remote Zoom teleconference option. Because of the geographic barriers, the low stipend, and the level of experts' professional responsibilities, the participants could not travel and be present physically in the session. Each research participant was willing and able to access the focus group sessions through the available Zoom teleconference solution. The participants represented renowned experts in their fields and were located throughout the United States and abroad. If I had restricted participation to individuals close enough to travel, it may have limited the level of expertise in the focus group sessions.

The original proposed stipend for each expert who participated in a focus group session was $250. I reduced the amount of the stipend to $25 to minimize the potential for or the appearance of an inducement. With the exception of one participant who refused a stipend, each research participant received an equal stipend of $25 on a prepaid Visa card after completing the member-checking phase of the research study. The stipend I provided was consistent with common research practice in the United States.

I modified the consent form prior to completing the participant recruitment process and holding the focus group sessions. I expanded the inclusion criteria for radiologists from a doctoral degree in medicine to a doctoral degree, giving me the option to include board-certified radiologists from different disciplines, such as osteopathy or chiropractic. I submitted this change to the IRB for review. The proposed changes were reviewed and accepted.

Unusual Circumstances

I encountered no disruptive or unusual circumstances during the data collection process, nor did I need to vary from the data collection process I proposed in Chapter 3 of this dissertation. I field-tested the focus group moderator guide and all of its elements. This included field-testing of open-ended focus group questions. The data collected from research participants was limited to their returned demographic survey and their contributions during the focus group sessions. I did not perform any repeat interviews or focus group sessions.

Data Saturation

My goal during each step of the data collection process was to acquire enough representative information to achieve data saturation. Research has demonstrated that data triangulation is an important step toward achieving data saturation (Fusch & Ness, 2015). The definition of saturation varies between research methods and designs (Guest et al., 2006). I triangulated data from multiple sources in this study to support comprehensive exploration of the research topic from numerous perspectives.

The process of data acquisition and analysis should be operationalized in a manner consistent with the research topic and research method (Saunders et al., 2018). I operationalized various strategies to achieve data saturation in this study. This included operationalization of the principal research question and subquestions in the context of the study. I considered data saturation achieved if additional research inquiry was unnecessary to further research questions. In addition, data saturation was achieved if further attempts at acquiring and analyzing data did not lead to new perspectives or thematic conclusions. I introduced my operational definition of data saturation at the beginning of each focus group session. I informed research participants that I would allow discussion of each topic until we achieved data saturation.

Member Checking

This qualitative research study included a research participation review and validation process. I arranged for all discussions and contributions made during each focus group session to be recorded and transcribed verbatim. I gave each research participant the opportunity to review thematic summaries along with supportive quotes derived from the focus group transcripts. Each participant received the information for review approximately six weeks after the focus group session. The short time frame helped ensure a more accurate and contextual review process. This method of follow-up gave each participant the opportunity to assess the accuracy and validity of the focus group data acquisition and analysis process. I did not provide research participants with a complete copy of their focus group transcript, for I felt that it might lead to overcorrection and bias. Research has demonstrated that withholding the full unabridged focus group transcript from participants can help reduce the risk for reflective over revision and subsequent compromise of the research results (Krueger & Casey, 2015). I provided the participants in the AI focus group with their thematic summary and the members of the radiology focus group with their thematic summary.

I provided each focus group participant with the same instructions for reading and responding to the thematic review documents. The documents sent to each participant consisted of a cover letter with instructions, an eight-page preliminary thematic analysis, a table depicting the coding (labeling) hierarchy used for data analysis, and a research participant survey (Appendix G). The preliminary thematic analysis document consisted of primary themes, subthemes, and supportive quotes acquired from each of the focus group session transcripts. I asked each research participant to review the submitted material and to respond to three survey questions by placing an “x” next to each statement they agreed with. I provided each of the participants the opportunity to clarify their survey responses and to submit comments. The first statement on the survey was “The results of thematic analysis reflect opinions offered during the focus group sessions." The second statement was "The focus group session quotes help support the result of thematic analysis." The third and final statement on the survey was "I agree with the results of thematic analysis." I methodically developed each survey statement to acquire a simple and concise measure of confirmation and/or feedback from the thematic summaries.

The research participants all responded with a completed and returned survey within 10 days of receiving the information. Each participant agreed that the results of my thematic analysis accurately reflected opinions offered during the focus group session and that the focus group session quotes helped support the results of thematic analysis. Each research participant also agreed with the results of the thematic analysis. One AI expert offered clarification of the difference between radiomic and deep learning for in vivo data analysis. The research participants did not offer any other comments about the research process or the results of thematic analysis. In summary, I provided all of the research participants with an equal opportunity to affirm whether my analysis of the acquired data accurately reflected their contributions and opinions during the focus group sessions. The respondent validation process provided each participant with the opportunity to assess the adequacy of the data and the reasonableness of my data analysis and interpretive process. I provided each research participant with the opportunity to correct errors, to offer clarification, and to challenge my data analysis process, as well as my interpretation of emergent content and themes. The methodical respondent validation process confirmed the accuracy of data acquisition and analysis thereby, reducing my potential bias and improving the trustworthiness and transferability of my research study results.

Data Analysis

Thematic analysis represents an established and rigorous methodical method for revealing meaningful results in qualitative research (Braun & Clark, 2006). It also represents one of the most common methods of analysis in qualitative research (Quest, 2012). I was able to identify emergent themes in this study using iterative content analysis. I performed content and thematic analysis with the assistance of tools available through Atlas.ti (Version 8), a highly respected qualitative research computer program. I imported all of the research documents into Atlas.ti: the consensus-based white papers, focus group transcripts, and the content of my reflective journal. I used a hybrid approach of content followed by thematic analysis. Content and thematic analysis provides different types of conclusions. Content analysis provides more quantitative and objective perspectives, whereas thematic analysis results in a qualitative set of conclusions. I initially used content analysis to help identify themes and subthemes.

The Coding Process and Derivation of Themes

Consistent with the recommendations of King (2004), I developed a few predefined (a priori) codes and used them to guide content and thematic analysis. Provisional coding consisted of an initial deductive approach with a start list. I developed provisional codes with the insight acquired from my initial literature search, DOI and TAM theoretical perspectives, and the context of my research questions. I modified and revised my provisional codes during the data analysis process.

I implemented six iterative and overlapping phases of data analysis during content and thematic analysis. I applied this approach to each research document including the white papers and focus group transcripts. During the first phase of analysis, I familiarized myself with the data. The second phase consisted of rereading the material and assigning initial codes to label chunks of relevant data. This included the use of some provisional codes. The third phase of data analysis involved a broad search for themes and involved refining (expanding, combining and collapsing) existing code labels to better identify clusters and patterns of meaningful information. During the fourth phase of analysis, I reviewed working themes that included reducing and refining thematic parameters. The fifth phase of data analysis consisted of finalizing and organizing emergent themes and subthemes. The sixth and final phase of my analysis process consisted of triangulation of data from all of my research sources. I concluded the process with further iterative analysis and the creation of composite themes and subthemes.

I synthesized the themes that emerged from the focus group sessions with themes identified from the analysis of consensus-based white papers. I did not consider the themes and subthemes final until I had achieved an exhaustive iterative analysis of triangulated research data. The data acquired from different research sources helped support and validate my thematic conclusions. I kept detailed records of the development, evolution, and application of codes and their relationship to emergent themes throughout my data analysis process.

During my iterative process of reading, annotating, clustering, and rereading the textual data, new codes were applied and existing codes were collapsed or refined. Data sources should be coded using established line-by-line content analysis methods (Saldana, 2016). “Codes are labels that assign symbolic meaning to the descriptive or inferential information compiled during a study” (Miles, Huberman, & Saldana, 2014, p. 71). I developed a coding tree (hierarchy) during qualitative data analysis. My use of descriptive coding allowed for clustering of similar topics to help develop themes.

In summary, I used an iterative process of coding to evaluate the focus group transcripts and consensus-based research documents until thematic saturation was achieved. I analyzed the acquired data with a combination of a priori, open, and in vivo coding. My inductive method of content and thematic analysis offered a highly flexible and iterative approach for evaluating patterns within the data. The freedom to adapt the coding process during data analysis helped me reduce the risk for a priori driven bias. I took the focus group results, compared them and triangulated them with data acquired from published white paper documents and with my reflective journaling. I identified emergent themes, which aligned with research questions.

Data Analysis Software

I uploaded all of the research documents and focus group transcripts into the Atlas.ti software. Atlas.ti offered effective tools for detecting patterns, applying labels, and identifying clusters of similar topics in a transparent manner. The software supported multilevel nesting of similar topics and allowed for efficient application of codes and for the development of hierarchical relationships. I performed coding with all of the available tools, which included open, in vivo, and list coding options. I analyzed each of the research documents and transcripts until thematic saturation was achieved. I used the software to display unique data and data relationships.

Discrepant Cases or Perspectives

The use of iterative content and thematic coding helped exclude perspectives or topics that had limited frequency or inconsistent presentations. My primary goal in this study was not to identify isolated discrepant cases or instances but to present highly supported themes and subthemes. During the course of the study, I was unable to identify any highly discrepant or contradictory opinions, perspectives, or cases deserving mention.

Evidence of Trustworthiness

I used numerous approaches to help establish trustworthiness of this qualitative research study and the results. A high level of trustworthiness is required to improve the value of qualitative research (Yin, 2014). Thick descriptions, well defined research steps, and transparency increase trustworthiness. I used overlapping strategies in this study to improve its credibility, transferability, dependability, and confirmability. In this section of the chapter I highlight the steps taken to improve trustworthiness.

Credibility in this context of this study refers to the level of truth associated with research. Consistent with the published recommendations of Lincoln and Guba (1985) I used various techniques used to help establish credibility in the study. These techniques included; persistent observation, reflective journaling, triangulation of data, participant checking, and the assessment of referential adequacy. I enhanced the credibility of the research though the acquisition and triangulation of data from different expert sources.

The sources of data included two expert focus group sessions, four consensus-based white papers, and my reflective journal.

I developed a predefined sequence of data acquisition, data analysis, and process validation for this research study (Figure 5). I implemented a series of purposeful sequential steps to improve the trustworthiness of the study. These steps included (a) field testing of the elements of the moderator guide, (b) participant checking with response validation, (c) within group and between group analysis, (d) triangulation of data, and (e) transparency offered with the reflective journal. I performed reflective journaling paralleling all of the other steps to improve the trustworthiness of the study.

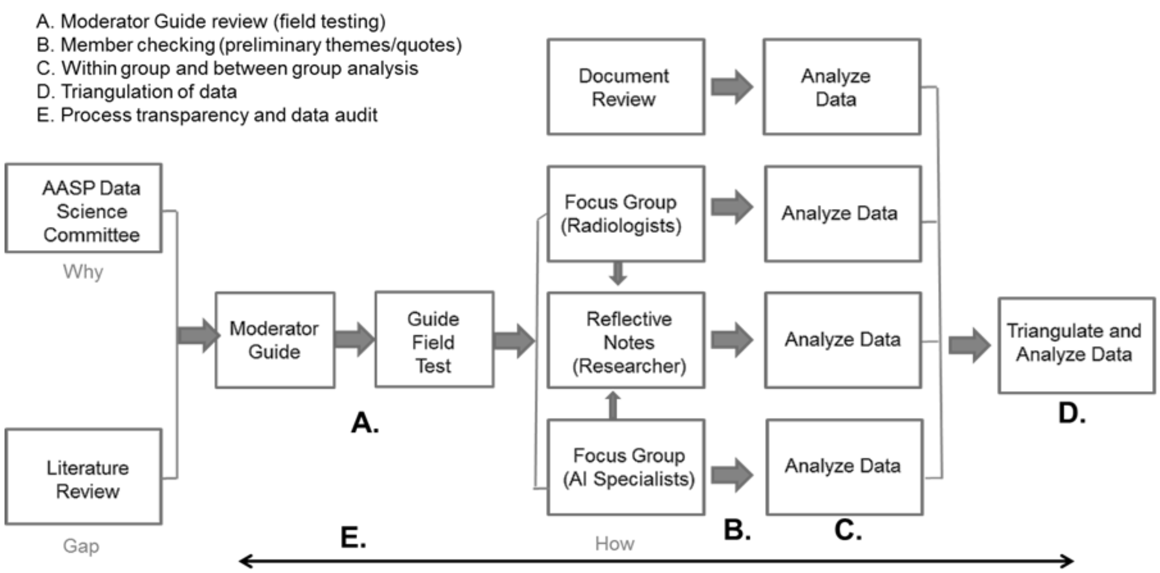


Figure 5. Steps used in the research process to improve trustworthiness of the study.

I used the Atlas.ti program to perform transparent multilevel nested coding. I fully disclosed the process of code development, which included a summary of my coding framework and an audit trail of code generation. I created a chart to depict the final code framework and the relationships between labeled data categories, subthemes, and themes.

The chart reflects topic relationships and thematic coherence. I supported this process with a coding guide comprised of concise definitions. As the sole researcher in this study, I kept a reflective journal that disclosed my perspectives, opinions, and potential biases throughout the entire research process. Disclosure of potential interpretive bias through reflective journaling helps improve qualitative research credibility (Creswell, 1998).

I was able to improve the credibility of the focus group results by facilitating active participation of all of the research participants. I purposefully selected each of the research participants because they represented the level of expertise and the range of demographics necessary to interact and adequately address the research topic. I used multiple sources of data to ensure that I could integrate the opinions of numerous experts and expert sources. In addition, I ensured that the study met the criteria of the COREQ qualitative research checklist that consists of 32 items used to assess the credibility of interview and focus group research studies (Tong, Sainsbury, & Craig, 2007).

I used thick descriptions and flow diagrams to depict the data analysis process and to establish both transparency and transferability, consistent with the recommendations of Lincoln and Guba (1985). The topic of transferability refers to the capacity to generalize the pattern of inquiry and the research results (Nowell et al., 2017). I developed flow diagrams to serve as a blueprint or audit trail of the data acquisition and analysis process. The illustrative approach combined with descriptive disclosures of each step of the research process provides the level of detail required for scholarly critique, as well as recreation of the research process.

The dependability of a qualitative research study influences the ability of other researchers to duplicate the research design (Yin, 2014). The triangulation of data and development of overarching themes from multiple data sources enhanced the dependability of this study. The correlation of data from different sources offered an effective method for evaluating the representativeness of emergent concepts and themes. To help ensure dependability, the accounts of discordant and deviant cases were exposed and addressed if they occurred.

I kept a reflective journal throughout the research process to expose my role and potential influence on research design, data acquisition, data analysis, and data interpretation. My reflective journal offered the level of self-disclosure required to improve the dependability and credibility of the study. I used my entries in the reflective journal to help expose and address any personal bias that might influence the research process. My journal entries provided some of the basis for methodological decisions, as well as the rationale for data coding and thematic analysis. In addition to offering transparency to the research process, reflective journaling helped me identify potential personal in a manner allowing me to take the necessary steps to reduce its influence.

A member checking process often improves the credibility of qualitative research (Lincoln & Guba, 1985). I performed research participant (member) checking as an external check for my data analysis and final thematic conclusions. I provided each of the focus group research participants the opportunity to review themes, subthemes, topics and supportive quotes that I derived from the focus group transcripts. This process represented a form of participant debriefing. This added step also offered “referential adequacy” which helped me confirm the trustworthiness of the research process and its conclusions.

In addition to reflective journaling, I kept a process log, which included the production of notes about the activities, events, and significant decisions that occurred during the study. I used the process log to help create a clear and concise schematic of the research study stages. This served as a pictorial audit trail of the methods used and the order in which they were implemented. In summary, I kept detailed records of the procedures, methods, and decisions made during the research study to improve its trustworthiness.

The Results

I acquired data from internationally recognized and published consensus-based white papers, two expert focus group sessions, and from personal reflective journaling. Qualitative analysis included triangulation of data from all of the research sources. This lead to the emergence of themes and subthemes supported by contributions and quotes from focus group research participants, as well as from the content of consensus-based white papers. I present the research results using thematic summaries, calculated frequency of coded categories, and charted relationships between data categories. I also include tables, figures, and cognitive maps to illustrate the data analysis methods and to present research results. In review, I designed this research study to investigate the potential impacts of AI on the interpretative stage of spine imaging. This includes its impacts on the differential diagnostic process and radiology workflow.

I organized this section of the chapter by the three primary themes that emerged from qualitative analysis of the research data: patient-based decision support, population-based decision support, and application-based decision support. I assigned numerous subthemes to each theme. The primary themes account for interdependent levels of AI decision support that can augment the role of the radiologist during the interpretive stage of spine imaging. I expose how each of the thematic perspectives enhance the differential diagnosis process leading to more precise and personalized spine care.

The subthemes are supported by the synthesis of data acquired from consensus-based “white paper” analysis and focus group contributions. Supportive quotes from focus group participants are included in each of the subtheme summaries. The use of quotes offers an effective method for highlighting important topics and emergent themes (Bloomberg & Volpe, 2008; Miles & Huberman, 1994). I identified quotes by assigning the participant’s expert class followed by a participant number (Class-Px) to maintain the confidentiality of the focus group participants. Evidence-based diagnostic decision support in spine imaging requires the integration of patient and population data through specialized application of technologies. I identified the following themes and subthemes that emerged from my analysis of the research data.

Theme 1: Patient-Based Decision Support

Patient-based decision support in the context of this study refers to the use of data and knowledge about a specific patient acquired with diagnostic imaging and/or personal medical records. This approach helps detect, characterize, and monitor data unique to the patient and their disease process. Patient-based decision support offers the radiologist the ability to formulate a personalized diagnosis, to develop a treatment plan, and to assess treatment outcome. Patient information can be integrated with population-based knowledge during radiology workflow to support a probability-based differential diagnostic process and to help predict treatment outcomes. Subthemes of patient-based decision support include multiscale in vivo analysis, natural language processing, change analysis, prioritization, and immersive data display.

Subtheme: Multiscale in vivo analysis. Radiologists and AI experts acknowledged that diagnostic imaging data in spine care is underutilized and that the assessment of pathology must extend beyond human visual and cognitive limitations. The study confirmed the belief that AI-supported molecular and radiomic diagnostic methods could be used to improve how disease is detected, characterized, and monitored. One of the first steps of in vivo analysis after the acquisition of imaging data is the identification of an ROI followed by targeted segmentation. Research participants concluded that this step determines the location where additional information will be acquired and analyzed. The Canadian Radiology Association acknowledged the importance of accurate detection and segmentation of pathology on imaging studies for disease characterization and monitoring (Tang et al., 2018). AI expert P4 highlighted the importance of getting this step right when they stated "It's very important to define an acquisition protocol and the detection protocol that we are going to apply to each image." AI experts and radiologists in the focus group sessions unanimously agreed that detection criteria should not be limited to human visual interpretation. Radiologist P2 supported this position with "Wouldn't it be nice to know that there is an abnormality, even if we can't visualize it?"

This led to further discussion about how advances in AI support will alter how spine imaging will be performed and how imaging data will be interpreted.

AI experts acknowledged the potential impact of semiautomated and automated tissue segmentation to help speed up and standardize the process of defining a ROI. Obtaining the right perspective using manual, semiautomated, and automated methods is critical to diagnostic precision. Radiologists and AI experts concluded that accurate segmentation is critical for multiscale in vivo tissue interrogation using molecular, radiomic, and deep learning diagnostic methods. Members of the AI focus group discussed numerous challenges associated with the variability of manual segmentation. There was a consensus among AI experts in the focus group that automated segmentation of ROIs would improve consistency. They also concluded that its success would require access to large volumes of curated AI training and testing data. This position was reinforced by AI expert P4, who stated, "Getting the right level, depending on the field of view is certainly a challenge to any fully automated application." The Canadian Radiology Association addressed the importance of using ground truth to guide automated segmentation, detection, and characterization of pathology because of the impact on treatment (Tang et al., 2018). Considering the contributions of AI experts and radiologists it became quite evident that the radiologist' preferred approach is to have access to automated segmentation with the option of manually manipulating ROI parameters.

Many of the radiologists and AI experts who participated in the focus group sessions agreed that non-visible features of pathology can be revealed through the application of engineered hard-coded algorithms supported by domain knowledge or through the application of deep learning methods capable of detecting patterns that represent characteristics of pathology within a defined region of the spine. Experts also agreed that AI solutions could be used to enhance the role of radiomic and deep learning methods. This premise was clarified by AI expert P3, who said,

> Radiomics can give us some kind of insights that cannot be appreciated with the human eye because we cannot interpret or define the statistical appearance. . . . I think radiomics and other things will be able to help use, you know get more information about underlying patterns, statistical patterns that are related to different voxel intensities and how they are distributed

Radiomics is typically distinct from deep-learning approaches. While deep learning encodes image properties in a large number of “deep layers,” radiomics represents a more explicit analysis of specific image properties (shape, intensity, texture, etc.).

Examples of the potential applications of radiomics in spine imaging were discussed during the radiology and AI expert focus group sessions. A prominent neuroradiologist acknowledged that most spinal cord disorders progress and evolve over a long period before clinical signs and symptoms become present. The neuroradiologist also stressed the importance of being able to detect abnormal signals within neurological tissues on advanced imaging studies that are not visible with traditional anatomic imaging and which precede the development of obvious clinical signs and symptoms. Radiologist P2 offered a supportive quote:

> I think that maybe there's more going on in the spinal cord, nerve roots or the cauda equine than we ever imagined, and AI might open up a whole window of opportunity to see [which level] is affected. . . . So just sitting here, I see a lot of potential and what could be, as I used to say, a very boring lumbar spine could suddenly turn into something wonderful and challenging.

Research in other areas of radiology have demonstrated that machine learning (ML) algorithms can be used to identify patterns of disease which occur beyond the threshold of human detection (Tang et al., 2018). AI experts in the focus group addressed the possibility of adapting or improving radiomic methods used in other fields like oncology for use during spine imaging.

Participants in the radiology focus group emphasized the importance of taking advantage of all available imaging data to improve the precision of the diagnostic process in spine imaging. They agreed that identifying all actionable imaging data on a spine imaging study exceeded human potential; thus, creating new challenges and opportunities during the interpretive process. This perspective was supported by radiologist P2, who stated, "Subvisual in vivo identification and characterization of pathology within the spine could turn what appears to be a routine image into a wonderful diagnostic challenge". The spine can be subdivided into clearly identified three-dimensional spaces referred to as voxels that can be used to direct tissue interrogation, as well as map or reference specific findings. AI expert P3 stated, "I think radiomics and other things will be able to help us, you know, get more information about the underlying patterns, statistical patterns that are related to different voxel intensities and how they are distributed." Growing application of voxel-based feature mapping will enhance three-dimensional interpretation of spine pathology.

AI experts discussed the possibility of using radiomic methods to extract and analyze features of spine pathology through the analysis of subvisual voxel-wise statistical and textural features. AI expert P3 supported this premise with the statement that "Radiomics can give us some kind of insights that cannot be appreciated with the human eye because we cannot interpret or define the statistical appearance of pathology." Radiologists and AI experts agreed that successful use of validated radiomic methods would improve the ability to screen for and characterize spine pathology. The potential to improve the detection of early stage pathology through the evaluation of nonvisual patterns of disease is reflected by the following statement about the spinal cord offered by radiologist P2:

> I like the idea of revealing changes detected inside of the cord that are not visible by anatomic imaging, because we know that specifically in the cervical spine, if these changes occur over a long period of time, the spinal cord can accommodate and literally be a ribbon before the patient has any symptoms.

AI experts and radiologists concluded that radiomic and deep learning methods have the potential to be further developed for assessing margins, shape, volume, and heterogeneity of spine pathology at multiple biological scales. Furthermore, the participants acknowledged that multiscale in vivo tissue analysis could offer an effective holistic, systems perspective for characterizing and monitoring spine pathology.

Many of the AI experts, as well as the radiologists in the focus group sessions, agreed that classification and staging of pathology represents one of the final and most important tasks during the interpretive stage of spine imaging workflow. This premise was supported by the following statement offered by radiologist P5: “There's so many things that go through your mind and it would be nice to have help to really sort of categorize it better.” The response reflected the interest in using available AI methods to augment the classification of spine pathology during image interpretation. Participants in the AI expert and radiology focus groups discussed the potential for the combined use of multiparametric and multiscale in vivo analysis to serve as a digital (virtual) biopsy. One of the AI experts believed that the concept of the “digital biopsy,” could be developed and applied to the spine. AI expert P4 further elaborated by stating, “I think that more and more the digital biopsy will be used more than the traditional biopsy.” The expert also acknowledged that, “We know we can characterize voxel by voxel the tumor. So with the digital biopsy, I think that more and more these kinds of biopsies are going to be done first rather than the traditional one.” Other AI experts within the same group acknowledged that the concept of the digital biopsy would not be limited to tumors but could be applied to any pathology within the spine as well as in other tissues.

Advances in spine imaging data acquisition and interpretation will lead to more personalized care. The Canadian Radiology Society reported that personalized health care is going to be dependent on in vivo characterization of various molecular, cellular, textual and structural attributes of pathology (Tang et al., 2018). The European Society of Radiology also highlighted the role of in vivo analysis for the accurate classification and stratification of disease to support the right treatment choices (European Society of Radiology, 2015). Early detection and characterization of spine pathology will help lead to better treatment outcomes. The European Society of Radiology and the French Radiology Association both addressed the importance of using computational support to screen for and auto detect asymptomatic preclinical disease with non-invasive in vivo analysis (European Society of Radiology, 2015; French Radiology Community, 2018). These consensus-based opinions support many of the contributions made during the focus group sessions. Multiscale in vivo tissue interrogation will co-evolve and converge with AI technology to a play critical role in the future of spine care.

Subtheme: Natural language processing. Patient data needs to be better integrated into radiology workflow to enhance the diagnostic process. Natural language processing (NLP), a form of AI, can be used to locate and extract relevant information from various sources of unstructured data such as a patient's electronic medical records (EMR). This includes a patient's problem list, prior radiology reports, pathology results, genetic profiles, and general clinical information.

Some of the AI experts in the focus group reported that using NLP could be used to identify active problems, contextual data, and current diagnoses from EMR to create a "need-to-know" list of information for use at the time of spine image interpretation. Radiologists acknowledged that timely access to relevant non-imaging data at the radiology workstation would improve the accuracy and efficiency of the interpretive process in spine imaging. Radiologist P2 agreed and stated, "If they could just cherry pick the older stuff, it would it would make a big difference for sure." The radiologists in the focus group agreed that access to an accurate problem list would offer the contextual insights required to improve the accuracy and personalization of the diagnostic process.

All of the radiologists agreed that limited time can have an adverse impact on the interpretive process. AI expert P4 demonstrated their knowledge of this challenge when they stated "natural language processing with imaging is always beneficial. . . . A lot of times the radiologist won't even have given enough time to look for or through the EMR, for all, you know, the history, et cetera." AI experts and radiologists addressed which medical record information would be the most helpful during the interpretation of spine imaging studies. Radiologist P5 responded with, "The main thing I think would help us is to be able to get a clinical summary for specific symptoms." A couple of radiologists questioned whether NLP could be used to search the electronic medical records for relevant spine information. All of the AI experts acknowledge that NLP could be used to search electronic medical records for contextually relevant spine care information. As part of this discussion, AI expert P2 said:

> I think that natural language processing can automatically look through the summary of, a relevant summary, for that patient of everything that would be related to the kind of condition, the kind of images, would definitely help as mentioned.

The majority of the radiologists in the focus group session acknowledged they do not have time to review the full depth and breadth of prior radiology reports. Radiologist P3 believed that, "If you can have automated relevant summaries of the prior reports that would be helpful for us to decide which reports to read in detail and what things to focus on. . . . A succinct summary of prior reports would be helpful." All of the radiologists in the focus group session felt that having access to diagnostic impressions and quantitative measures from prior radiology reports would improve the differential diagnostic process during spine imaging.

Participants in the AI and radiology focus group sessions addressed the importance of extracting and correlating relevant information from textual records and from acquired imaging data to improve diagnostic precision. One of the AI experts addressed the potential role of AI for supporting the integration of NLP and radiomic assessment during spine imaging interpretation. AI expert P2 acknowledged that NLP extraction of information from medical records could favorably inform the diagnostic process, but believed that in vivo assessment of pathology offered the most up-to-date and relevant information. AI expert P2 stated:

> I think the image is still probably a better source of information. But I think it could be complemented by NLP. I think the other way around is maybe a little bit less likely because of the incompleteness of what's in the EMR.

AI experts acknowledged that NLP could be used to help overcome a radiologist's limited knowledge by identifying differential diagnostic possibilities and/or specific features of pathology from population-based data and from published literature.

Radiologists addressed how NLP could help overcome interpretive errors in spine imaging. This concern was highlighted by radiologist P5, who commented, "We have tremendous problems with voice recognition errors; for example, little decimal points can make a big difference as we've found out." The same radiologist stated that a "report could be verified through the use of AI and corrections can be made before the report is released." This became a topic of great interest during the radiology focus group session. The group concluded that NLP could be used to check the accuracy of spine radiology reports during the interpretive and postinterpretive stage of spine imaging workflow.

During the radiology focus group session, participants addressed the need to have assistance in prioritizing various elements of the differential diagnostic process. Radiologist P5 contributed, "We have radiologists that just describe everything but never give any differentials and other radiologists that list five, ten or fifteen things so they do need to be prioritized. We could certainly use assistance with that." Radiologist P5 also stated, "If I could snap my fingers and get whatever I wanted, I would want all the clinical information that I could [get]. Then I would certainly love to have some differential diagnostic assistance." AI experts discussed the ability of AI to assign probability to differential diagnostic possibilities supported by the analysis and correlation of data from numerous sources including multiscale in vivo imaging, electronic medical records, and published literature.

The potential role of NLP during the interpretive stage of radiology is becoming widely accepted. The Canadian Radiology Society acknowledged that NLP is capable of converting unstructured text into a structured form, which can be mined and analyzed using AI methods (Tang et al., 2018). The French Radiology and Canadian Radiology Associations highlighted the potential for NLP to provide contextual insight from radiology reports and other medical records for the interpreting radiologist (French Radiology Community, 2018; Tang et al., 2018). In summary, unstructured information in the EMR is an invaluable source of insight for the radiologist during the interpretive stage of workflow.

Subtheme: Change analysis. Spine disorders are often insidious, progressing without obvious signs or symptoms. Timely detection and characterization of the progression of pathology is required to improve the potential for good therapeutic outcomes. All of the radiologists in the focus group session acknowledged the importance of objectively monitoring potentially serious spine disorders such as bone marrow abnormalities, spinal cord compression, vertebral deformities, and fractures. Radiologist P3 stated, “I think having some sort of objective finding that we can assess over time from the previous study might be helpful because a lot of times just eyeballing it, is to subjective.” In addition one of the participating radiologists offered the following statement:

> If we can objectively quantify things like neural foramen stenosis and spinal canal stenosis and compare those quantities over time that might be helpful because a lot of times you know we just make a subjective assessment. . . . It would be nice to have a reproducible number and more reproducibility of findings.

All of the radiologists in the focus group session concluded that consistent surveillance of small seemingly insignificant pathology can sometimes be important. This is particularly relevant for the assessment of small masses, tumors, and spinal cord compression secondary to stenosis. Radiologist P5 addressed the importance of consistent monitoring of suspicious pathology. “We've all seen these patients that fell through the cracks because of reporting of a small mass a year and a half ago and nobody followed it up.” AI experts concluded that spine disease surveillance would require accurate and consistent co-registration of tissue for change algorithms to be successful. AI expert P4 said, “It’s very important to define a consistent acquisition protocol” to monitor spine pathology. AI could be used to auto compare and measure pathology on serial spine imaging studies. The process would require consistent imaging protocol, accurate co-registration of tissue, labeling of spine levels, and the use of validated change analysis methods

AI experts and radiologists alike determined that quantitative measures such as radiomic methods could be developed to help monitor subtle and non-visible changes in spine pathology over time. This included the use of AI-supported change analysis to help differentiate incidental findings from significant early stage pathology. Radiologists discussed the potential for using combined radiomic and deep learning methods to detect and measure change. Radiologist P4 said, “The quicker that we can find injury to the cord, to the nerve root, the quicker we can maybe offset some of the debilitating problems.” Experts in both focus groups acknowledged that structural monitoring may not be enough to detect and characterize changes in pathology on serial spine imaging studies. Participants in both focus groups agreed that statistical and textural features of pathology acquired through voxel-wise measures could be compared over time to assess disease progression and/or treatment outcomes. This concept was supported by the following statement offered by AI expert P3: “I think radiomics and other things will . . . get more information about the underlying patterns, statistical patterns that are related to different voxel intensities and how they are distributed.” Radiomics offers the potential to detect and quantify non-visible changes in spine pathology.

Members of the AI focus group concluded that anatomic, statistical, and textural features of spine pathology could be automatically compared over time to assess spine disease progression, spine disease evolution, and/or treatment outcome. They further proposed that this could be facilitated by temporal subtraction methods applied to successive spine imaging studies. AI participants acknowledged that 3-dimensional voxel-wise analysis could be used to demonstrate change. The concept of using a virtual biopsy to evaluate change was also discussed. The potential for AI methods to help detect and reveal non-structural biomarkers of aggressive spine pathology was supported radiologist P5:

> There are separate microhabitats evolving on their own and just watching them structurally doesn't necessarily change the treatment, whereas if there were different signatures associated with different levels of aggressiveness, that might change the treatment. I don't see how a human can assimilate all that information.

Spinal vertebrae offer rigid bone boundaries that can be used as points of reference to help register and co-register spatial relationships between different spinal tissues for auto segmentation and change analysis. AI expert P2 referred to the spine as one of the "easiest parts of the body to co-register because the vertebrae are very, I would say very rigid, the bone is seen very well on each scanner assessment." The AI experts collectively agreed that the technologies required for developing these solutions are available, and are currently being used in other specialized areas of radiology. Objective change analysis offers predictive insight while providing an effective method for assessing treatment outcome.

The European Society of Radiology and the Canadian Radiology Association both acknowledged the potential benefits of using AI for disease surveillance and monitoring of treatment outcomes (European Society of Radiology, 2015; Tang et al., 2018). Experts have proposed that non-invasive in vivo interrogation of tissues in different dimensions with imaging will improve the characterization of pathology (European Society of Radiology, 2015). The society also conceded that the benefits associated with traditional anatomic imaging are limited. Many spine disorders have a long subclinical history prior to clinical detection. Diagnostic imaging limited to anatomic perspectives can overlook evidence of early-stage pathology. Advances in diagnostic imaging and decision support have improved the ability to detect subclinical pathology (European Society of Radiology, 2015). The combined application of enhanced visual analysis and non-visual data analysis will enhance the potential to detect subclinical spine pathology and to monitor change as the result of disease progression or treatment outcomes.

Subtheme: Automated prioritization. A radiologist has limited time and therefore it must be used wisely. Intelligent solutions can help eliminate redundant tasks and augment the role of the radiologist. AI can be used to achieve these goals, as well as help prioritize the interpretive process and allocate a spine imaging study to a particular level or path of interpretation. This includes prioritizing the worklist and prioritizing the interpretation of specific images or pathology. Radiologists who participated in the focus group session agreed that prioritizing the interpretation of spine imaging studies has been long overlooked and is critically important. Many of the radiologists in the group were enthusiastic about the possibility of AI prioritizing interpretation of spine studies based on prior image abnormalities, clinical presentation, and current pre-interpretive image analysis. Radiologist P2, for example, stated, “I think prioritizing would be an advantage and even identification of a little hint of what was in the past and what you're looking for on a follow up study, sure.” Successful prioritization of imaging would focus the radiologist’s attention on that which is most likely abnormal and clinically relevant.

One of the radiologists suggested that AI solutions could improve the efficiency of trauma assessment. They suggested that AI-supported screenings could locate and label spine injury features such as cortical disruption, dislocations, fractures, and the presence of edema within ligamentous complexes. Radiologist P3 claimed, “highlighting cortical disruption on the images and then prioritizing those to the top of the list would be helpful.” A member of the group acknowledged that AI could identify distinguishing features of aggressive or high-risk spine pathology that might require immediate attention. Radiologist P2 argued, “AI is going to add the icing on the cake, sort of like mammography where you press the button and then the arrow goes, hey did you look at that.” Participants in the AI expert and radiology focus groups both acknowledged the potential benefits of technology driven prioritization of interpretive focus and time.

Numerous AI experts acknowledged that machine learning could prioritize the analysis of specific spine images or a region within a spine image. They also believed that deep learning radiomics could detect, characterize, and prioritize the evaluation of pathology or a particular set of features within a region of pathology. Many of the radiologists and AI experts proposed that AI solutions could be developed to auto compare and measure the state of pathology between imaging studies and subsequently prioritize significant change. Many of the radiologists and AI experts also concluded that molecular measures, radiomic methods, and deep learning solutions could enhance the spine imaging screening and prioritization process. They also acknowledged that NLP could also be combined with pre-interpretive image analysis to enhance the prioritization process.

Spine imaging studies are complex, often including many of the soft tissues of the chest, abdomen or pelvis. Radiologists in the focus group acknowledged that computational support might be helpful for detecting the presence of subtle lesions and for evaluating the presence of multiple lesions within the spine and surrounding spinal tissues. Radiologist P1 stated, “You can't focus your attention on one thing. In radiology, you’ve got to be really out there looking at everything.” The participant also addressed the possibility that AI methods could be used to perform an automated background screening or analysis of e spine and extra spinal tissues prior to or paralleling visual interpretation of the study. Many of the radiologists shared frustration associated with the responsibility of having to evaluate all of the tissues and related data on an imaging study.

It has become progressively more difficult for a single radiologist to focus his or her attention on the detailed interpretation of all aspects of an advanced spine imaging study. The French Radiology Association acknowledged the importance of using AI tools to help prioritize diagnostic imaging studies and regions of pathology before interpretation by the radiologist (French Radiology Community, 2018). The Canadian Radiology Association supported this premise by acknowledging that AI could be used to auto detect and prioritize the interpretation of critical findings on diagnostic imaging studies (Tang et al., 2018). Machine and deep learning applications can be used to screen for preclinical disease, to classify pathology, and to alter the level of reading priority, all in a manner, which cannot be achieved with human visual interpretation (European Society of Radiology, 2015; Tang et al, 2018). In summary, AI solutions could highlight or flag all suspicious areas on a spine imaging study prior to interpretation.

Subtheme: Immersive data display. The spine is intricate and complex. The method of displaying data can influence the accuracy and efficiency of the interpretive process. Two-dimensional (2D) views of spine pathology are often insufficient for a precise diagnosis and for the support of personalized treatment planning. The evaluation of spine pathology in 3D space offers a more comprehensive perspective of pathology than 2D assessment.

Many of the radiologists and AI experts who participated in the focus group sessions agreed that the method used to display imaging data could influence the accuracy of spine image interpretation and the description of findings. Some of the research participants reported that evaluation of complex pathology in 3D space is more revealing than 2D assessment because it offers more perspectives. During the course of discussion AI expert P1 clarified that “What comes to mind here again is the need to look at things in open 3D space. Because when you use 2D views, for example, you can only go through the displays in orthogonal directions.” Another AI expert in the focus group acknowledged that a multidimensional display of data would give the radiologist the opportunity to appreciate the heterogeneity and spatial relationships of pathology including its relationship to surrounding tissues and blood flow. Numerous members of the AI focus group acknowledged that 3D imaging could help reveal atypical or anomalous structural relationships, to better identify, and characterize boundaries or transitional zones between normal and diseased tissue.

AI experts discussed the potential for molecular and radiomic features to be integrated with or mapped onto 3D volumes of pathology, allowing for immersive and volumetric characterization. AI expert P1 concluded that in the near future voxel-wise biomarkers could be embedded into or mapped onto 2D or 3D renderings of spine pathology and color-coded to enhance the interpretive process. The participant further acknowledged that, "It would be terrific, of course, if AI and radiomics, etcetera, are applied to volumes." AI expert P3 responded, "voxel-based biomarker information can be perfectly plotted in the kind of environment you're suggesting." The following statement made by AI expert P1 further supported this possibility:

> So if you're looking at things in open 3D space, then you can kind of swim through the object and find what I call key bookmark views, the key places to really analyze, and then apply the AI and radiomics to the key views, which can really give you significant directions moving forward.

A couple of AI experts in the focus group acknowledged that 3D data displays enhance or support the concept of the digital "virtual" biopsy. The European Society of Radiology acknowledged that 3D representation of pathology along with the use of volumetric measures could be very helpful in the evaluation of pathology progression and response to treatment (European Society of Radiology, 2015). Virtual reality (VR) and augmented reality (AR) were also proposed to have the potential to improve digital multidimensional exploration of pathology, to augment contextual pathology assessment, and to help guide invasive diagnostic and interventional approaches.

Theme 2: Population-Based Decision Support

Population-based decision support in the context of this study refers to the use of data and/or knowledge stored in a database about patients with similar backgrounds, histories, comorbidities, and/or disease states. These data are often used to assist in the differential diagnostic process, making predictions, and rendering a prognosis. Population-based decision support is knowledge and model driven. Population-based information must be acquired from a database or computational disease model built over time with information acquired from numerous individuals and locations. Population-based information is highly dependent on knowledge sharing, knowledge access, and knowledge preservation (Greens, 2014). Patient information can be integrated with population-based knowledge during radiology workflow to formulate a differential diagnosis, to help predict disease progression, and evaluate treatment outcome. Subthemes under this heading include ground truth and knowledge database.

Subtheme: Ground truth. Truth represents a verifiable fact or set of facts derived through scientific methods. Ground truth represents fundamental facts required to make complex observations and decisions. An accurate differential diagnostic process requires various types of decision making resulting in truthful conclusions. Ground truth in radiology can help correlate imaging findings with other sources of data representing pathology or biological states. Potential sources of ground truth include clinical, histologic, laboratory, and genetic workups. Expert P2 stressed that “the key thing is to have a source of truth of your training data.” Ground truth may also be achieved through correlation with other imaging findings (in vivo) or from computational models of disease (in silico). Ground truth is required to assign relevance to molecular signatures, radiomic features, and other imaging biomarkers of pathology. The quality and volume of training data must be adequate to establish ground truth and its relevance. Expert P2 acknowledged, “In theory, given enough images and outcomes you could have some sort of a ground truth . . . but it’s always difficult.” Numerous AI experts and radiologists in the focus group sessions proposed that the accuracy of the spine imaging differential diagnostic process and staging of pathology could expand knowledge of the correlative relationships between biological states and AI-derived biomarkers of health and disease.

AI solutions developed for use during the interpretive stage of spine imaging require proper training and validation. AI experts in the focus group all stressed the importance of having access to ground truth for AI training data. It was determined that one of the challenges associated with AI training for spine imaging is the source of ground truth. AI expert P2 highlighted the importance: “In terms of the applications, you have to think about what is the ground truth that I’m using to train my data? That’s key.” All of the participating AI experts acknowledged that ground truth is established through the correlation of imaging finding with other sources of data representing pathology. They also agreed that this often presents a challenge because the results of “other sources of data” are typically recorded in non-imaging databases and systems. Participating radiologists agreed with the opinions of participating AI experts that the interpretive process in spine imaging would be enhanced with better knowledge of the biological correlates of imaging.

The research participants of both focus groups concluded that better application of ground truth and quantitative imaging measures will improve the clinical utility of the final radiology report. The convergence of pathology and radiology data combined with the use of structured reporting, quantitative measures, and standardized disease classifications will help establish ground truth for AI training and testing. Radiologist P3 proposed that "if everybody used a standardized structure for their reports and adapted to that, that would make things a lot easier for this sort of analysis." Radiology reports provide a source of common data elements for establishing ground truth and for AI training.

The Canadian Radiology Association acknowledged that ground truth often lies on a continuum (Tang et al., 2018). Ongoing discoveries in the fields of genetics, pathology, and radiology will continue to influence the criteria for disease and subsequently the continuum of ground truth for AI training, validation, and applications. The thresholds between the various elements of the continuum will also change with new insights. One of the primary challenges is determining what truth is most important and relevant to the diagnostic process. Radiologist P3 recommended investing the time and money required to obtain ground truth. The radiologist further commented, "We have to address the possibility that getting all this additional data from the imaging is actually something that's useful and will affect the outcome." Many of the AI experts and radiologists agreed that the ability to prioritize the search for ground truth is highly influenced by awareness of its potential clinical utility.

Subtheme: Knowledge database. The individual radiologist brings limited experience and knowledge to the differential diagnostic process. The success of AI-based decision support during the interpretive stage of spine imaging will be highly dependent on access to relevant knowledge and computational analysis at the radiology workstation. The availability of knowledge is dependent upon an interdependent cycle of knowledge generation, validation, management, and application. A knowledge database refers to a virtual or real platform used to transform structured and unstructured data from different sources into actionable intelligence for problem solving. It essentially converts big data into big insights.

AI and radiology experts in the focus group sessions independently came to the consensus that the success of AI-based decision support in spine imaging will be highly dependent on access to disease models and to an adequate volume of curated training data with annotated images. Radiologist P2 said, “The bigger the pool of information, the better we're going to be.” The AI experts acknowledged in their group discussion that the development and integration of knowledge databases in other fields has proven to enhance the role of AI decision support. Radiologist P2 addressed how the use of robust data has transformed the field of genomics and how it could be applied in the field of radiology. The same radiologist illustrated the point with the following statement:

> Just look at a DNA analysis, how, it started kind of slowly and now once they developed these large databases it's just advanced by leaps and bounds, and if you can do this with AI, we would all be very grateful.

Radiologists and AI experts in the focus groups independently acknowledged the importance of having access to data, as well as knowledge to make informed decisions.

Knowledge is derived from the synthesis of information from numerous trusted sources. Participating AI experts and radiologists agreed that information can be acquired from many sources including omics-based disease models. The computational disease model, a form of population-based data offers knowledge about pathology at multiple levels and scales. Population data should include the variability required for disease model training. AI expert P3 acknowledged, “You need to build a very robust dataset. And when I say, robust, I mean a dataset that is representative of the variability of your problem to help ensure the success of AI-based decision support.” The same participant said, “The most trustworthy approach nowadays is having variability represented in the dataset that you will use to train your models.” The success of AI use during the interpretive stage of spine imaging will be dependent on the radiologist having on-demand access to an evolving knowledge database that offers adequate decision support.

Radiologists represent one of the most successful knowledge brokers in health care. They are rapidly becoming the gatekeeper of big data and complex decisions. In the near future, the majority of multidimensional perspectives of disease will be evaluated at the radiology workstation. To achieve these goals knowledge must be readily accessible through the integration of AI technologies, patient data, population data, and research data at the workstation. The latter includes computational disease models. Many AI applications will require large training datasets validated at the population level, thus requiring interoperable data flow between repositories (Tang et al., 2018). The French Radiology Association reported that annotated images are required to train AI solutions and to achieve higher levels of knowledge about a disease process (French Radiology Community, 2018). The Canadian Radiology Association concluded that limited access to representative and properly curated training data constitutes one of the most common obstacles to the development of image-related knowledge and valid AI applications in all fields of health care (Tang et al., 2018). The participating radiologists admitted that they must let go of old concepts and access new knowledge in order to embrace emerging opportunities and to remain a valuable member of the health care team.

Theme 3 Application-Based Decision Support

Application-based decision support in the context of this study refers to the use of structured processes and technological solutions to solve problems during the interpretive stage of spine imaging workflow. This approach can augment the role of the radiologist by providing access to data, knowledge, and clinically relevant AI applications. Specialized AI applications could integrate patient and population-based data to provide a more precise and personalized spine diagnosis.

Application-based decision support enhances radiology workflow by offering access to a pipeline of data and accesso a menu of narrow AI solutions. Meaningful applications of AI are based on clinical utility and demand. They include the technology attributes and processes required to augment the role of the radiologist. Important determinants of technology adoption include interoperability, ease-of-use, and benefits of use. Specialized applications of AI and related technologies are required to provide decision support during spine imaging workflow. Subthemes under this heading include clinical utility and technology attributes.

Subtheme: Clinical utility. Clinical utility refers to the usefulness and potential benefits of a technology, process or intervention in patient care. Spine disorders considered a high priority for AI development and use during the interpretive stage of spine imaging workflow include spinal cord pathology, spine tumors, bone marrow disorders, intervertebral disc pathology, fractures, and spine pain syndromes. Each of these disorders is prevalent and requires early detection, accurate characterization, and timely intervention to avoid poor clinical outcomes.

During a focus group session, one of the radiologists acknowledged the potential role of AI for detecting and prioritizing multilevel spine and spinal cord pathology. The expert also acknowledged the importance of early detection of pathologic changes in the neurological elements of the spine before structural changes are evident on routine diagnostic images. Radiologist P4 offered the following statement in supported of the premise, "The quicker that we can find injury to the cord, to the nerve root, the quicker we can maybe offset some of the debilitating problems." Undetected progression of spinal cord and/or nerve root pathology can lead to permanent functional deficits and subsequent disability.

Radiologists discussed the potential role of AI-supported solutions such as radiomics and deep learning to acquire more insight about abnormal bone marrow signals and other suspected regions of pathology on advanced imaging studies such as MRI and CT. This included sharing concerns about the increased prevalence of bone marrow disease including cancer metastasis. During the focus group discussion, radiologist P2 addressed concerning statistics claiming “with the aging patient population, we're seeing more metastatic disease and there’s incidents of multiple myeloma, which is sometimes really tough to identify.” The radiologist’s contribution prompted further discussion about the use of radiomics to help differentiate various types of pathology in bone marrow, in the spinal cord, and elsewhere in the spine. AI experts acknowledged that the use of consistent imaging protocols combined with co-registration of vertebral bodies would allow for targeted serial comparison of bone marrow and spinal cord signal changes between diagnostic imaging studies.

AI experts in the focus group session addressed the potential for AI to be used for the prediction of bone pathology such as fracture. AI expert P4 addressed the possibility of performing “radiomic analysis in order to predict if a new patient is going to suffer this kind of fracture.” During the discussion another AI expert acknowledged the current availability of technology that can be used to auto detect vertebral fractures. All of the focus group AI experts recognized the importance of developing solutions that auto label vertebrae, identify vertebral deformities, detect fractures, and highlight suspicious regions of cortical disruption. They acknowledged that success would require accurate registration and co-registration of imaging data. AI expert P4 referred to the spine was as “one of the [easiest] parts to co-register because they are very, I would say . . . very rigid, the bone is very well seen in each scanner.” One of the radiologists acknowledged that

AI-supported quantitative assessment of vertebral body morphology (dimensions) would offer early evidence of compression deformity and subsequently expose the risk for fracture and bony displacement.

During an extended discussion of clinical utility, one of the AI experts acknowledged that deep learning networks could establish relationships between imaging characteristics and the presence of spine pain. An AI expert offered clarification stating, “In theory, given enough images and outcomes you could have some sort of a ground truth on pain, but it’s always difficult. But in theory, a deep learning network could help establish relationships between imaging and pain.” Success would require ground truth and adequate AI training. AI experts agreed that this approach would be a challenging project, but if successful would have huge impact to society.

The clinical utility of AI applications in spine imaging will be dependent on their capacity to reliably improve the detection and characterization of spine pathology, as well as help predict outcomes. Radiologist P3 acknowledged, “We have to address the possibility that getting all this additional data from the imaging is actually something that's useful and will affect the outcome.” Clinical utility includes the ability to reveal new molecular and radiomic signatures of disease that can be used to better classify and stage spine pathology.

Consistent use of disease criteria and related terminology is required for AI to achieve widespread clinical utility. AI solutions could help achieve these goals. AI could augment the role of the radiologist by auto labeling anatomic structures, detecting and characterizing regions of abnormality, and by providing probability-based differential diagnostic support. Members of the radiology focus group were encouraged to hear fellow radiologist P4 state, “I’m happy to hear people say that the radiologist will still be involved, but I do think AI is going to add the icing on the cake.”

The Radiological Association of North America and collaborating societies acknowledged that identifying fundamental common data elements (CDEs) and prioritizing meaningful use cases for AI applications not only addresses whether an algorithm can be built, but whether it should be built (Allen et al., 2019). Success requires that radiologists and other health care providers be involved in determining which disorders require better decision support. The Canadian Radiology Association concedes that the best way to approach the research and development of clinical applications of AI is to identify and classify meaningful use cases (Tang et al., 2018). Clinical utility is influenced by need, workflow, technology applications, and classes of use (Tang et al., 2018). The decision whether to develop an AI solution is also influenced by cost-benefit analysis, ethical considerations, and establishes needs of health care providers (Allen et al., 2019).

Subtheme: Technology attributes. Many variables influence whether new AI and related technologies are adopted, used, and supported. Participants of the AI focus group session concluded that perceived usefulness will represent one of the most important drivers of AI adoption during the interpretive stage of spine imaging workflow. Radiologist P3 supported this premise. “We have to address the possibility that getting all this additional data from the imaging is actually something that's useful and will affect the outcome.” Participants in the AI and radiology focus group sessions stressed the importance of identifying what AI applications would have the greatest impact on patient care and what would be required for their adoption and use.

Participants in the radiology and AI focus groups addressed the important influence of ease-of-use on the adoption of AI solutions. AI expert P3 said, “The perceived ease of use is key.” on the relationship between perceived benefits and usefulness. “When I see usefulness, I think of clinical efficacy” and “When I see ease of use, I think of workflow.” Another AI expert responded by acknowledging that for an AI solution to be adopted it must be easy to use and have a seamless application in radiology workflow. This is reflected by the perspective offered by AI expert P3, “application must be seamless. So thankfully, AI is very good at this. It’s very good at automatic procedures.” Participating AI experts and radiologists agreed that AI solutions must be clinically relevant, immediately accessible, and easy to use.

Participants of the AI focus group concluded that perceived ease-of-use and perceived benefits are both powerful determinants of the AI technology adoption process. The majority of the AI experts acknowledged that the use of AI during the interpretive stage of spine imaging workflow will ultimately depend on awareness of its clinical benefits. One of the study participants projected that clinical outcome measures will be used for training AI solutions. AI expert P3 addressed the relationship between ease-of-use and perceived benefits with, “You’re adding value to the clinical workflow. But if you offer usefulness and you don’t address perceived ease of use, you are dead.” The same expert further elaborated by acknowledging, “all of our applications need to be seamless and automatic. AI needs to be perfectly integrated in the workflow of the radiologist." Many of the participating AI experts and radiologists acknowledged that unresolved disruption of radiology workflow would adversely influence the adoption and use of AI solutions.

Numerous participants in the AI and radiology focus group sessions believed that heightened awareness of the relative advantages of AI solutions will have a significant impact on whether AI technology is adopted for use during the interpretive stage of spine imaging workflow. Radiologist P5 said, "We can use all the help we can get in my opinion" highlights the importance of addressing new solutions." It is therefore important to identify what help is required. Some of the radiologists felt that successful AI adoption requires that the applications be interoperable and seamlessly woven into the fabric of existing spine imaging workflow. Heightened awareness of the role of proposed AI solutions requires education. Radiologist P1 stated, "At every level there has to be education of what to do." AI experts and radiologists independently concluded that to be successful AI applications must have a proven positive impact on patient care, which could not be achieved without its use.

The positions of many of the AI experts and radiologists who participated in this study are supported by the published consensus-based opinions of numerous respected radiology organizations. The Canadian Radiology Association acknowledged that one of the most important factors in the adoption of AI solutions is the ability to integrate emerging applications into existing technology at the workstation (Tang et al., 2018). The association also recognized that successful adoption requires that AI solutions meet clinical needs, enhance the efficiency of interpretive workflow, and improve the accuracy of the diagnostic process. Successful applications of AI and related technologies will require seamless integration of AI tools into existing workflow (Allen et al., 2019). Successful AI adoption will also require that the solutions are able to address unmet clinical needs with user-friendly interfaces (Allen et al., 2019; Tang et al., 2018). In summary, various AI applications can be developed and used to augment the role of the radiologist during spine image interpretation. They can also be used to help establish more efficient workflow and best practices in spine imaging.

Data Results

I present the research results using two primary methods, which are thematic summaries and calculations of coded topic frequencies. In included the later approach to reveal how topics were prioritized in the literature, which topics are more developed, and which topics may be generating more interest. The combined approach helps provide a foundation for further discussion and research. I used tables to provide a practical overview of the research results. Table 1 identifies the themes, subthemes, and subtheme topics that emerged during the course of qualitative data analysis. The primary themes that emerged from the triangulation and analysis of research data in this study were patient-based decision support, population-based decision support, and application-based decision support. Each category of decision support has the potential to augment the role of the radiologist during spine image interpretation and diagnosis. The table was used to represent the relationships between coding subtheme topics, subthemes, and themes that emerged from qualitative data analysis in this research study.

Table 1

Themes That Emerged From Qualitative Data Analysis

Topics	Subthemes	Themes	Goal
Detection Segmentation Characterization Monitoring	Multiscale In Vivo Analysis	Patient-Based Decision Support	Augmented Role of the Radiologist *(Improved Diagnostic Precision & Personalization)*
Problem List Prior Radiology Reports Electronic Health Records	Natural Language Processing		
Structural Features Radiomic Features Molecular Features	Change Analysis		
Worklist Triage Image Triage Pathology Triage	Prioritization		
3D Images AR/VR Images Feature Mapping	Immersive Data Display		
In Vitro/ Ex Vivo In Vivo In Silico Clinical	Ground Truth	Population-Based Decision Support	
Disease Model (Omics) Annotated Data Training Data Validation Data	Knowledge Database		
Spinal Cord Disorders Bone Marrow Pathology Fractures Spine Pain	Clinical Utility	Application-Based Decision Support	
Perceived Benefits Ease-of-Use Interoperability Relative Advantage	Technology Attributes		

Note. The hierarchical relationships that emerged with thematic coding.

The primary themes are supported by subthemes that represent technologies, methods, or resources, which can be used individually, or in an integrated fashion to augment the role of the radiologist. This approach has the potential to improve the accuracy and efficiency of the differential diagnostic process in spine imaging. Table 2 provides an overview of the definitions of the codes used to label topics, subthemes, and themes, in this study.

Table 2

Data Analysis Code Definitions

Data Analysis Codes	Code Definitions
Patient-Based Decision Support	Data acquired from a patient used to make health care decisions
Population-Based Decision Support	Data acquired from population data used to benefit a patients care
Application-Based Decision Support	Technology and/or processes used to help personalize patient care
Multiscale in Vivo Analysis	Multidimensional characterization of pathology in a living system
Natural Language Processing	Use of computational methods to analyze textual data
Change Analysis	Multiscale objective measures of interval change in pathology
Automated Prioritization	Use of automated AI methods to rank image interpretation priority
Immersive Data display	Interactive multidimensional display of diagnostic imaging results
Ground Truth	Confirmed factual relationship between the real world and AI data
Knowledge Database	Collection of interdependent facts and info to support decisions
Clinical Utility	Relevance and use of AI applications in the care of specific disorders
Technology Attributes	Various determinants of technology adoption and use

Note. The table above identifies the contextual definitions of theme and subtheme headings used in this study.

The subthemes associated with patient-based decision support were multi-scale in vivo analysis, natural language processing, change analysis, prioritization, and immersive data display. The subthemes associated with population-based decision support were ground truth and knowledge database. The subthemes associated with application-based decision support were clinical utility and technology attributes. The research study including contributions during the focus group sessions identified the interdependent relationships between the thematic topics and their subthemes. The co-evolution and convergence of the processes and technologies referred to in the subthemes of the study can be integrated and used to construct an AI ecosystem in radiology. I performed within and between group thematic topic frequency analysis to illustrate the interest in and/or knowledge of the topics that emerged during the study.

I subjected consensus-based white papers and focus group session transcripts to content and thematic coding with the use of Atlas.ti software. The iterative process led to the development of code categories such as topics, subthemes, and themes. The first theme referred to as patient-based decision support was comprised of five subthemes. Each of the subthemes comprised three to four categories of relevant topics. I referenced the frequency of the subthemes and the relevant coded topics in Table 3 to illustrate patterns of interest and the priority placed on topics related to patient-based decision support during radiology workflow. The subtheme coded with the greatest frequency was radiomics followed by natural language processing. A close third was change analysis, which refers to quantitative serial monitoring of pathology. Under the heading of radiomics, the topic with the highest coding frequency was the detection of pathology followed by its characterization. Under the heading of NLP the area of greatest coding frequency was the ability to extract contextually relevant information about the patient from the electronic medical records and prior radiology reports. Table 3 reflects the importance of knowing what to look for and finding it.

Table 3

Patient-Based Decision Support Subtheme Topic Frequency

Topics	Radiomics	NLP	Change Analysis	Prioritization	Immersive Display
Detection	32				
Segmentation	18				
Characterization	27				
Monitoring	25				
Electronic Medical Records		14			
Prior Radiology Reports		14			
Problem List		6			
Publication List		3			
Structural Features			11		
Molecular Features			15		
Statistical Features			7		
Worklist Triage				1	
Image Triage				6	
Pathology Triage				7	
3D Imaging					11
AR/VR					1
Pathology Feature Mapping					2
Total	102	37	33	14	13

Note. Frequency of code labeling during the analysis of data acquired from consensus-based white papers and focus group transcripts performed with Atlas.ti Version 8 software.

The second theme referred to as population-based decision support was comprised of two subthemes. Each of the subthemes was comprised of three categories of relevant subtopics. I referenced the frequency of the subthemes and coded subtopics in a table to illustrate patterns of interest and the priority placed on them during radiology workflow (Table 4). The subtheme that received the highest frequency of coding was knowledge database followed by ground truth, although both topics were nearly equal in coding frequency. Under the subtheme titled knowledge database, the subtopic that received the highest frequency of coding was annotated data for training AI. The next subtopic in order of frequency was validating the AI process. Table 4 reflects the importance of training and validating AI applications prior to clinical use.

Table 4

Population-Based Decision Support Subtheme Code Frequency

Topics	Ground Truth	Knowledge Database
In Vitro/Ex Vivo	5	
In Vivo	8	
In Silico	5	
Disease Models		5
Annotated Training Data		14
Validation Data		8
Total	18	27

Note. The frequency of code labeling during the analysis of data acquired from consensus-based white papers and focus group transcripts performed with Atlas.ti Version 8 software.

Under the subtheme heading titled ground truth there was relatively equal coding frequency between the subtopics titled in vitro, in vivo, and in silico sources of data for developing ground truth. The subtopics with the least coding frequency under the subtheme ground truth was in silico and under the subtheme knowledge database was disease models. These topics are interrelated and represent important elements of diagnostic workflow that are just beginning to emerge in radiology. The underdeveloped status of these areas may account for their low coding frequency and some of the current related challenges in the evolving AI field.

The third theme titled application-based decision support is comprised of two subthemes. The two subthemes are technology attributes and clinical utility. Each of the subthemes was comprised of four subtopics. I referenced the frequency of the coded subtopics in Table 5 to illustrate the patterns of interest and the priority placed on the topics related to the use of AI during radiology workflow. The subtheme titled technology attributes received the highest subtheme coding frequency followed by the subtheme titled clinical utility. Under the heading of technology attributes the subtopic that received the highest coding frequency was interoperability, followed by ease-of-use, and relative advantage. The subtopics of relative advantage and perceived benefits had a similar pattern of coding frequency. Under the subtheme clinical utility, spinal cord disorders and spine pain had the highest frequency of subtopic coding.

Table 5

Application-Based Decision Support Subtheme Code Frequency

Topics	Technology Attributes	Clinical Utility
Perceived Benefits	9	
Ease-of-Use	18	
Interoperability	25	
Relative Advantage	13	
Spinal Cord Disorders		10
Bone Marrow Pathology		6
Fractures		7
Spine Pain		10
Total	65	33

Note. The table above represents the frequency of code labeling during the analysis of data acquired from consensus-based white papers and focus group transcripts performed with Atlas.ti Version 8 software.

The code frequency in this study likely parallels the level of interest in AI-related themes, subthemes and subtheme topics. Code frequency analysis reflects the level of awareness and/or importance of topics related to the potential use of AI during the interpretive stage of spine imaging workflow. The research results indicate the level of interest in the interoperability of technology and the capacity to integrate and embed numerous AI solutions into existing radiology workflow. The results also reflect interest in using AI to better assess spinal cord disorders, bone marrow pathology, fractures, and spine pain. The insights from code frequency tables and thematic analysis could be helpful in conceptually developing an AI ecosystem and designing future research.

The thematic category of patient-based decision support was of greatest interest in both the AI expert and radiology focus group sessions. A calculation of the frequency and percentage of coded response from each homogenous focus group was performed (Table 6). The assessment revealed that AI experts had a higher level of interest on the topics of radiomics and immersive display categories associated with advanced in vivo pathology assessment. These categories represent emerging technologies that radiologists may be less familiar with. In contrast, the radiologists had a higher frequency response for natural language processing, change analysis, and, prioritization solutions that would immediately augment the role of the radiologist during the interpretive stage of spine imaging. This table represents the importance of the two specialists working together to fully develop compatible and interoperable solutions for interpretive spine imaging workflow.

Table 6

Between Group Analysis of Patient-Based Decision Support

Focus Group	Radiomics	NLP	Change Analysis	Prioritization	Immersive Display
AI Experts	69	29	20	6	73
Radiologists	31	71	80	94	27

Note. The table above represents the percentage of coded responses from each homogenous focus group for each category assigned to the subtheme Patient-Based Decision support. The table may reflect interest in and/or knowledge of the subject matter. NLP = Natural Language Processing.

The acquired research data clearly established that radiologists require, as well as desire technological assistance to help overcome the growing burdens associated with interpreting complex data in spine imaging. Radiologists are required to make important decisions, often in the presence of overwhelming data and incomplete clinical information about the patient. This combination of challenges increases the risks for error and missed diagnostic opportunities. AI-supported solutions such as multiscale in vivo tissue interrogation, natural language processing, prioritization, and immersive data displays can identify and present actionable data to the radiologist at the time of image interpretation.

The results of this study support the premise that AI-supported deep learning and radiomic methods will become major determinants in the spine imaging differential diagnostic process. The approach will be further developed to better detect and characterize non-visible features of spine pathology resulting in a more precise and timely diagnosis. This will lead to the development of new disease criteria, better stratification of disease, and a movement away from traditionally accepted spine pathology models.

This study indicated that there is much work to do. Additional AI technologies and protocols must be developed, validated, and integrated to fully enhance the differential diagnostic process in spine imaging. The study also revealed that automated mining and analysis of imaging data must be embedded into radiology workflow. The radiologist of the future should have access to a menu of narrow AI applications that can be used on an as needed basis during image interpretation. These tools could be used to perform targeted in vivo tissue interrogation (virtual biopsy), to retrieve relevant contextual information from the medical records, to prioritize the interpretive process, perform change analysis, and to display data in a unique manner which improves the diagnostic process. Successful application of these solutions will require access to ground truth and large curated spine imaging datasets for AI training and validation.

My analysis of the research data revealed the importance of identifying and prioritizing the clinical needs in spine care that could benefit from AI applications. This included the use of AI-supported methods to augment the role of the radiologist in the evaluation of spinal cord pathology, bone marrow pathology, fractures, and spine pain. The study findings indicated that spinal cord pathology and spine pain were prioritized, possibly due to the potential for disability associated with undiagnosed progression. The study further revealed that the primary clinical goal of AI applications in spine imaging are the early detection, characterization, and objective monitoring of pathology. The research indicated the desire to use AI-supported methods to identify and detect multiscale imaging biomarkers that can make a precise diagnosis, measure treatment outcome, train AI, and help build computational models of disease.

My research confirmed that the methods currently used to interpret spine imaging studies are inadequate and must be improved to favorably impact the delivery of personalized spine care. The results support the concept of categorizing and allocating spine imaging data to a specified hierarchical level or path of interpretation to overcome the limitations of single human interpretation of some studies. This potential action step represents intelligent workflow allocation based on study complexity or severity. My analysis of the research data revealed that AI could be further developed and used to help deliver the right spine imaging study, with the right level of interpretive decision support, at the right time.

I was unable to identify any significant controversies, discrepant perspectives or conflicting opinions from the research sources in this study. There was general agreement regarding the co-evolving processes and technologies surrounding AI and its potential for use in radiology, more specifically spine imaging. The research revealed different opinions regarding the magnitude and methods of AI training required and the steps required for the validation of AI applications prior to their use in clinical and radiology settings. The scope of this research study did not allow for the exploration of the moral and ethical impact of the use of AI during spine imaging workflow.

The opinions of spine care providers of various disciplines and other experts are required to help direct the development of AI and related technologies. Professionals such as radiologists and AI experts are required to work together and with their organizations to reveal the fundamental basis for investigating the potential role and impacts of AI during the interpretive stage of radiology. “AI developments are currently being driven largely by computer scientists, informaticians, engineers and business people, with much less direct participation by radiologists” (Rubin 2014, p. 1309). This research study laid the foundation for members of the spine care, radiology, and AI communities to contribute valuable insight about the potential utility of AI solutions during the interpretive stage of spine imaging workflow.

The results indicate that for AI tools to be successful there must be access to relevant data, a positive impact on care, and the technologies must capable of being seamlessly woven into existing spine imaging workflow. Successful AI applications require clinical relevance and must be easy to use. The co-evolution of AI and related technologies will eventually augment the role of radiologists and will contribute to shaping the future of spine care. The path to innovation will lead to unprecedented challenges and discoveries at many levels. The outcome of this research study can be summarized by the work of Rudie et al. (2019), who stated “AI methods, given their ability to discern patterns and combine information in a way that humans cannot, show substantial promise for the future of radiology in precision medicine” (p. 616). Radiologists must embrace the potential of AI and help direct its development and evolution.

Reflective Journal Summary

Qualitative research is subject to researcher bias. I kept a reflective journal throughout the research process. Journaling helped me identify how my perspectives, presumptions, and bias might influence the study. It also allowed me to reflect on how my role as the sole researcher might influence the acquisition, analysis, and interpretation of data. I made regular journal entries throughout the research process. I reflected on and reviewed the entries on a regular basis. This process offered the level of self-refection and transparency required for me to implement steps to improve the trustworthiness of the study.

My background and experience influenced some of the decisions I made during the research design and research process. I came into this research study with over 25 years of experience that included the interpretation and clinical correlation of advanced spine and neuroimaging studies. During that time, I had the opportunity to collaborate with numerous radiologists of various backgrounds and specialties. I became well familiar with their role, as well as their challenges and strengths. During the same timeframe, I received academic training in molecular imaging, quantitative imaging, and computational decision support in radiology. As an experienced clinician and neurologist, I was astutely aware of the complex decisions associated with advanced imaging and the differential diagnostic process. I was also aware of the potential adverse impacts of human bias and limited knowledge on diagnostic decisions.

I designed this study with the hope that the results and insights acquired would have a favorable impact on further AI development for use in spine imaging, as well as in other specialties of radiology. I understood that this perspective increased my risk for introducing personal bias into the study. Heightened awareness of this possibility led me to develop and implement various solutions to reduce my influence. The solutions included field testing of the focus group moderator guide, member checking, and a high level of commitment for identifying the most credible experts and expert resources on the research topic. To help achieve the later goal I purposefully recruited renowned AI experts and radiologists for the focus group sessions. I recognized that the combination of member checking and triangulation of data obtained from highly respected consensus-based white papers would help reduce the potential influence of personal bias. To further reduce my influence and potential bias as the moderator of the focus group sessions, I developed and used a few neutral concept slides to focus the discussion on complex subjects. These concept slides were exposed to expert field testing for neutrality and relevancy.

Throughout my reflective journaling process, I recognized that I occasionally succumbed to the hope and hype surrounding the emergence of new technology. My deep-seated desire for having access to more accurate and detailed interpretation of spine imaging studies had the potential to influence my analysis of data in this study. To help reduce this risk I implemented painstaking efforts to improve the iterative process of triangulating and analyzing data from a diverse set of expert sources. Implementation of participant response validation also reduced the risk for exposing the research process to personal bias during data analysis, data interpretation, and in the presentation of the research results. With the help of reflective journaling, I was able to continuously question and crosscheck my beliefs with the methodological and operational choices I made during the research process. I found that one of the most exciting elements of this research study was introducing and discussing the concept of the virtual (digital) the biopsy. This subject represents my single greatest topic of interest arising from the study. The concept of the digital or virtual biopsy has the potential to transform the field of imaging.

Summary

The primary purpose of this research study was to explore the potential impact of AI on spine imaging interpretation and diagnosis. The study revealed themes, subthemes, and supportive topics for each subtheme. The three primary themes which emerged from the analysis of data were patient-based decision support, population-based decision support, and application-based decision support. The subthemes that emerged identified potentially interdependent technologies and processes, which are co-evolving and converging. The result will lead to methods of data analysis and decision support that will augment the role of the radiologist in all specialties including spine care.

Further development of AI for use with patient-based decision support solutions such as multiscale in vivo tissue interrogation, natural language processing, change analysis, prioritization, and immersive data displays will improve the accuracy and efficiency of the differential diagnostic process in spine imaging. These solutions will support improved detection and characterization of visible and nonvisible features of pathology, which will result in greater knowledge of the fundamental basis for pathology and will likely expand disease classification and staging.

Further development of AI use with population-based decision support solutions such as operationalizing ground truth and expanding knowledge databases will provide the radiologist with the opportunity to integrate this information with patient-based data. This process will support a probability driven differential diagnostic process. It will also help support the development of predictive and prognostic perspectives.

Expanded capabilities for analyzing data acquired from multimodal and multiparametric imaging studies will lay the foundation for the development of the digital “virtual” biopsy for use in spine and other areas of imaging. The virtual biopsy, unlike the traditional needle biopsy is capable of in vivo characterization of the full volume of spine pathology, including surrounding tissues. In the near future, AI supported natural language processing will provide on-demand context from medical records to assist in the interpretation of acquired in vivo spine data.

Radiologists require access to computational decision support and unique displays of data to enhance professional productivity and to improve diagnostic accuracy. The study revealed that the use of interactive immersive displays with virtual reality, augmented reality or combination of the two could help the radiologist assess pathology. The study also revealed that the use of an interactive 3D display of data in some situations would likely enhance appreciation for the various features of pathology at multiple biological scales.

The participants of the focus groups sessions in this research study identified and prioritized clinical applications for AI. The spine disorders addressed included spinal cord pathology, bone marrow pathology, fracture detection, and evaluation of spine pain. The area of greatest interest was assessment of the spinal cord due to the potential for devastating consequences of undiagnosed pathology. The study also revealed determinants of AI adoption for clinical use. Important technology attributes included interoperability, ease-of-use, and potential benefits. The latter topic encompassed clinical utility.

In summary, the study revealed that AI can augment the role the radiologist. Success requires AI supported methods which can integrate patient and population-based decision support with bridging technologies. Numerous AI applications and related technologies under each thematic heading are co-evolving and converging in a manner that will result in the development of an AI ecosystem. Successful adoption and use of AI applications will impact many of the core responsibilities of the radiologist, such as the detection, characterization, and monitoring of pathology. AI will replace redundant and mundane tasks currently performed by radiologists. This will free them up to perform more intuitive tasks and for consulting with referring health care providers. My research revealed that AI will improve the radiologist's productivity along with the precision and personalization of the diagnostic process in spine care. Success will require the use of validated solutions embedded into existing workflow, are easy to use, and that offer significant clinical utility.

Chapter 5 addresses the interpretation of research findings and provides recommendations for further research. The chapter also highlights the limitations of the study and identifies the potential impact the research results may have on spine care and social change. To help achieve these goals the chapter synthesizes the research results with the conceptual framework used for the study and the findings of the literature review. I used numerous figures and flow diagrams to illustrate concepts.

Chapter 5: Discussion, Conclusions, and Recommendations

Introduction

The primary purpose of this qualitative research study was to explore the potential of AI to augment the role of the radiologist during the interpretive stage of spine imaging workflow. I designed the research process to acquire data from multiple expert sources including two focus group sessions, four consensus-based white papers, and a reflective journal to reveal concepts and establish themes that can be used to expose the developmental requirements and potential benefits of AI use. I performed this study to establish a foundation for further research on the use and potential applications of AI in spine imaging.

My research identified numerous potential applications of AI during spine imaging. This included the use of AI-supported radiomics and deep learning methods in spine imaging to improve the detection, characterization, and monitoring of pathology at scales, which exceed human perception. My research also introduced numerous AI applications that could be integrated with multiscale in vivo interrogation to improve the accuracy of the differential diagnostic process. These methods included the use of NLP to provide relevant context, the use of change analysis algorithms to improve disease surveillance, prioritization of imaging studies to focus the attention of the radiologist, and novel methods for displaying data in multiple dimensions.

The three principle themes that emerged from qualitative data analysis were patient-based decision support, population-based decision support, and application-based decision support. In Chapter 4, I addressed the theme, subthemes and subtopics that emerged in the research in Chapter 4. The clinical conditions prioritized for AI use include spinal cord pathology, bone marrow pathology, vertebral fractures, and spine pain. The research study revealed a desire on the part of professionals, facilities, and organizations to improve the interpretive process during radiology workflow. It also exposed widespread recognition that the role the radiologist must be augmented with computational support to meet the growing need for a more precise and personalized diagnosis in spine care, as well as in other specialties. Access to AI support would allow radiologists to become more informed spine care consultants. Radiologists already represent one of the most successful knowledge brokers in health care. The majority of multidimensional perspectives of disease will likely arise from the radiologist workstation. This position will become empowered for they will become one of the primary gatekeepers of big data and decision support for other health care providers.

In this chapter, I address the relationships between the research results, the conceptual framework used for the study, and the findings of the literature review addressed in Chapter 2. The contextual interpretation of the research findings is followed by a discussion of the study’s limitations, as well as the methodical, theoretical, and social implications of the research results. This chapter concludes with recommendations for further research and considerations for various applications of the research findings. In this chapter, I also offer predictions about the future use of AI in spine imaging along with projected applications in the broader fields of radiology and in pathology.

Interpretation of Findings

Numerous researchers have identified high error rates in the interpretation of abnormal imaging studies (Elmore, et al., 1994; Lehr et al., 1976; Janjul et al., 1998). their work also identified the potential for missed diagnostic opportunities when using a limited although traditional anatomic interpretive approach. The complexity, intricacy, and redundancy of the spine render it a greater interpretive challenge than many other bodily regions. Subsequently, the risk for misdiagnosis or missed opportunity may be higher than expected.

This research study confirmed the results of the literature search that revealed heightened level of awareness regarding the underutilized data and insights embedded within imaging studies that could improve the diagnostic process. The study also confirmed knowledge of the diagnostic limitations associated with current imaging protocols, data processing methods, and interpretive measures. These conclusions are consistent with findings in other fields of radiology acknowledged by Gillies et al. (2016). My exhaustive literature search that preceded this research study confirmed that the unprecedented increase in imaging data volume and complexity during the last decade has placed an extraordinary burden on the individual radiologist (Obermeyer & Ezekierl, 2016; Ragupathi & Ragupathi, 2014). The literature search established that computational decision support is required to improve the differential diagnostic process and to reduce the risk for interpretive errors in in all areas of radiology including spine imaging (Hillman & Goldsmith, 2011; Jha & Topol, 2016; Kressel, 2017; Lee, 2017). This research study offers an updated status on these topics and applies the insight to the field of spine imaging.

The literature search revealed numerous variables that limit the acquisition and analysis of data in radiology. This includes biological variability of the individual patient, variability of data acquisition, heterogeneity of pathology, human bias, and the individual radiologist's limited capacity to address overwhelming data. As previously acknowledged in this disertation, Dreyer and Geis (2017) reported that AI augments the role of the radiologist during the interpretative stage of workflow. This includes informing the radiologist during the differential diagnostic process. These goals can be achieved through the development of computational solutions that are able to access and analyze vast quantities of actionable data from imaging, as well as non-imaging sources. The focus group sessions in this research study revealed some of the desires and unmet needs of radiologists who interpret spine imaging studies. I asked AI experts during the focus group session to offer predictions and solutions. This research study established the presence of numerous variables unique to the interpretation of spine imaging studies, as well as some of the unique diagnostic opportunities.

Radiomic methods can detect and extract features of pathology from imaging data, undetectable by traditional visual interpretation (Aerts, 2017; Gillies et al., 2016). Numerous researchers have proposed that the digital or virtual biopsy has the potential to interrogate and map the entire landscape or volume of a pathological state (Echegaray et al., 2016; Lambin et al., 2012; Thrall et al., 2016). In this study, I was able to conclude that AI-supported deep learning and radiomic methods could be further developed for use in the detection and characterization of spinal cord disorders, bone marrow pathology, and spine trauma including fractures. These methods could be developed for early detection and characterization of asymptomatic stages of spine disorders that could result in early intervention, biological intervention, and reduced risk for chronic pain and disability.

This study revealed that the unmet needs and interests of radiologists who interpret spine imaging are similar to those who interpret other types of studies. The technical challenges associated with developing valid AI solutions in spine care are also similar to those in other health care fields with a few exceptions. The unique challenges associated with establishing tested and validated AI solutions and ground truth in spine care include the intricacy of spine structures, inadequate volumes of annotated (curated) AI training data, and limited access to computational disease models and population data. Expert sources in this study predict that each of these challenges could eventually be overcome. Success would lead to practical applications of NLP, deep learning, radiomics, prioritization algorithms, immersive data displays, and computational disease models. Research participants also addressed the importance of adopting and adapting AI applications used in other fields of radiology, such as oncology, for use during the interpretive stage of spine imaging workflow.

The research confirmed that constructs of the TAM such as perceived ease-of-use and perceived usefulness will play a significant role in the decision to adopt AI and related technologies during spine imaging workflow. Constructs of the DOI such as interoperability and relative advantage will also represent key determinants of technology adoption and use. Interoperability is an important concept for the study. It applies to the integration of AI solutions into existing workflow. My research addresses the importance of being able to integrate various AI supported methods and technologies in a manner that is both practical and clinically useful. The success of AI use of spine during spine imaging workflow will be highly dependent on computational and technological interoperability and the impact on clinical outcomes. My analysis of research data revealed that the clinical utility of AI was prioritized over "ease-of-use," although both attributes were considered critical determinates of AI adoption and use. I listed below important concepts that arose from synthesizing information from the literature search with the research results.

In Vivo Interrogation of the Spine: The Virtual Biopsy

The goal of any diagnostic procedure is to obtain information about a disease process, which can inform care. Historically, looking directly at tissue has offered the highest yield of specific information about pathology. The procedure that supports this approach is the biopsy. In most cases, a biopsy is performed with a specialized needle to extract a small amount of tissue from a targeted region. The tissue is prepared, placed on a slide, and stained. A detailed microscopic examination is performed or supervised by a pathologist to assess disease features and to render a descriptive personalized diagnosis. This process is rarely used to evaluate the spine. A needle biopsy offers a limited perspective of disease. The diagnostic process begins after the tissue has been extracted from its normal biological microenvironment. Once a sample is removed (ex vivo) from the body, the ability to detect and characterize dynamic (living) biological attributes and the relationship between the sample and neighboring tissues is removed.

In contrast to the traditional biopsy, noninvasive imaging can assess a whole region of pathology and perform longitudinal disease surveillance of living tissue in its normal biological environment (Grossman et al, 2017; Patriarche & Erickson, 2007). It can also evaluate regional pathology and its relationship to surrounding tissues. AI-supported radiomics can be used for high-throughput evaluation of statistical, textual, and morphological values acquired from 2D regions of pathology referred to as pixels or from via 3D regions of pathology referred to voxels (Figure 6).

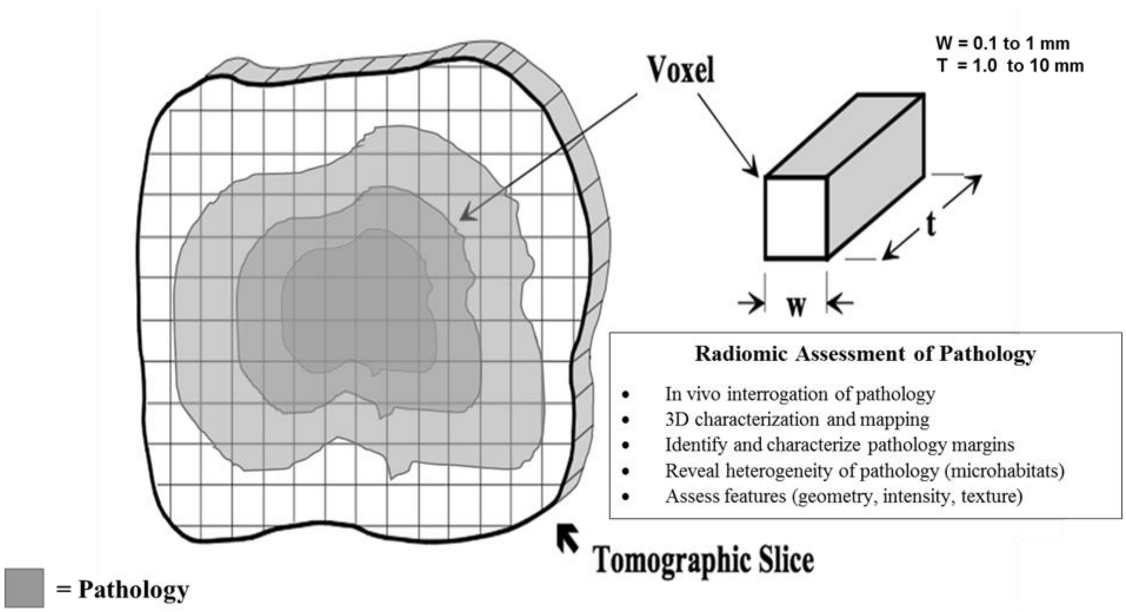


Figure 6. A voxel represents a three-dimensional volume of tissue in radiology. The figure above represents an artist's rendition of a slice of tissue using computed tomography (CT). The image depicts a focal region of pathology with transition to normal tissue. Radiomic methods can be used to extract non-visible characteristics of healthy or diseased tissue from one or more voxels. Permissions for use and adaptation of graphic from Wikipedia Commons Domain.

The approach shown in Figure 6 can identify and characterize non-visible features of pathology, as well as evaluate disease margins and transitional zones associated with pathology. The approach could help classify and stratify spine pathology.

The needle biopsy represents a limited sampling of pathology and therefore offers a limited approach for disease characterization, whereas radiomic methods can be used to provide more comprehensive characterization of the full volume of pathology (Figure 7).

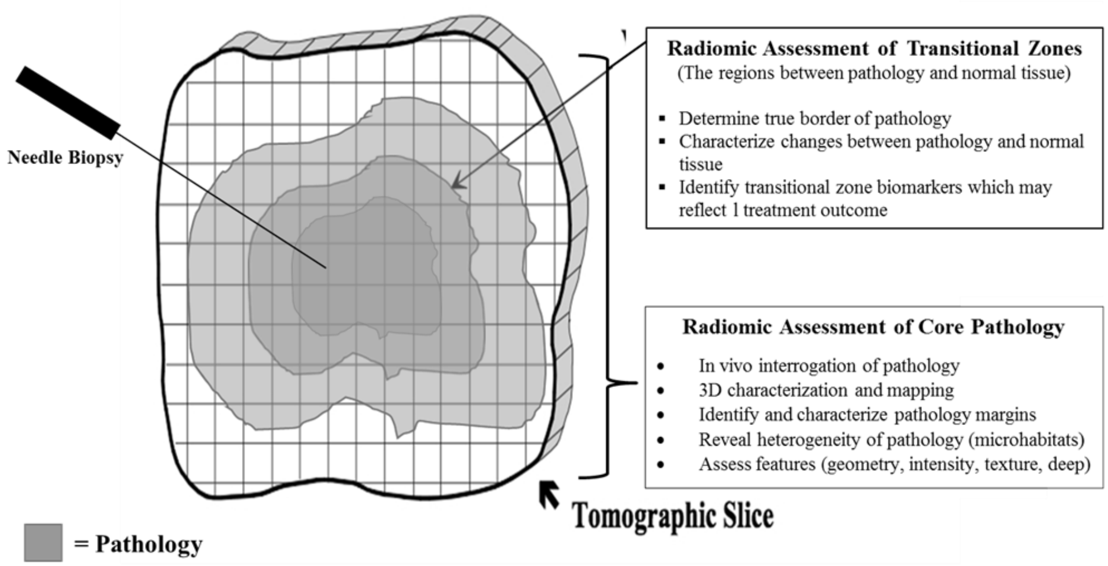


Figure 7. A needle biopsy can be used to extract a small sample of tissue from a targeted region of pathology. Radiomic methods can be used to characterize an entire region of pathology including neighboring tissues, a concept referred to as a virtual biopsy. Permissions for use and adaptation of graphic from Wikipedia Commons Domain.

Radiomic methods can help direct the traditional needle biopsy. It is important to evaluate and characterize a whole region of pathology to reduce or eliminate sampling bias that could adversely influence treatment planning and outcome measures. In support of this premise, Parekh and Jacobs (2019) acknowledged that most pathology in health care is under sampled. In addition, the traditional needle biopsy also fails to evaluate the biological features within a transitional zone surrounding a region of pathology. In addition, the needle biopsy has limited predictive value when compared to in vivo whole pathology analysis. It is not routinely performed on the spine due to potential complications. AI supported radiomic methods offer the ability to interrogate the spine and related tissue in vivo without disrupting or injuring tissue. One of the goals of diagnostic imaging in spine care is to derive as much actionable insight as possible. Radiomic methods can assess the non-visible features of spine pathology. The approach requires determination of a region of interest (ROI) followed by a series of well-defined steps to detect, characterize, and classify pathology (Figure 8).

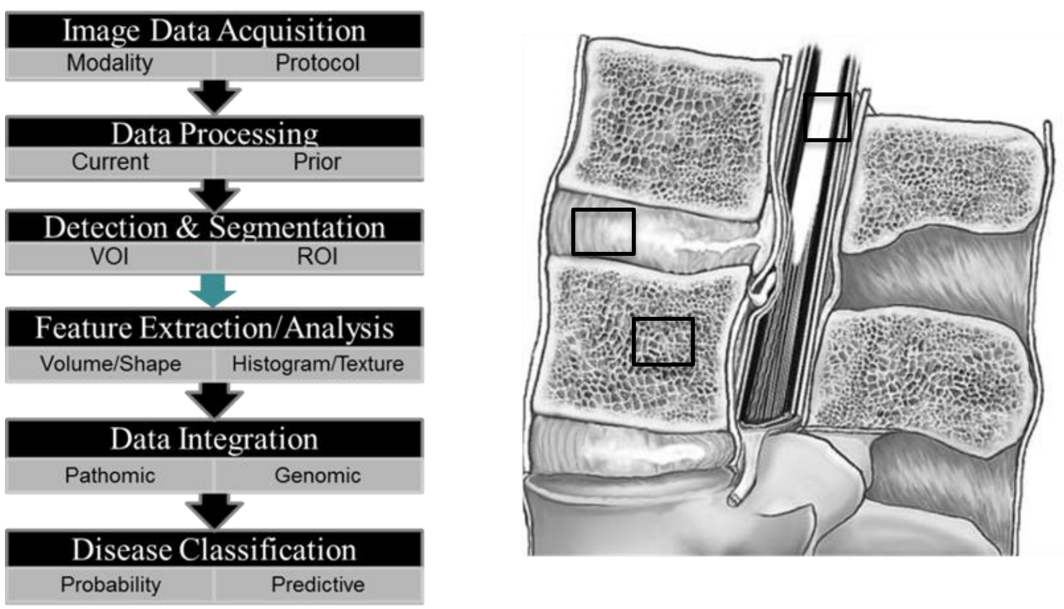


Figure 8. The sequence of steps that could be used during radiomic assessment of the spine. The approach offers new solutions for in vivo detection, characterization, and monitoring of pathology. Spine image used and adapted with permissions of Scholars Consortium LLC. Durrant, D. H., & True, J. M. (2012). Myelopathy, radiculopathy and peripheral entrapment syndromes. Palm City, Florida: Scholars Consortium, LLC.

Each modality used to image the spine such as MRI, CT, PET, and hybrid combinations provide different types of data on different biological scales. The use of combined imaging and related data analysis methods is capable of improving disease detection and characterization. Traditional radiomic methods are characterized by a series of sequential steps influenced by handcrafted algorithms, whereas deep learning methods are capable of pattern detection without the need for preprogrammed rules (Hosney et al., 2018). In the near future, radiomic methods could be combined with deep learning methods to enhance the assessment of pathology (Figure 9). The combined approach will offer an effective framework for diagnostic decision support in radiology (Parekh & Jacobs, 2019). The fusion or synthesis of multimodality and multiparametric spine imaging data subjected to analysis with AI methods could improve diagnostic imaging accuracy.

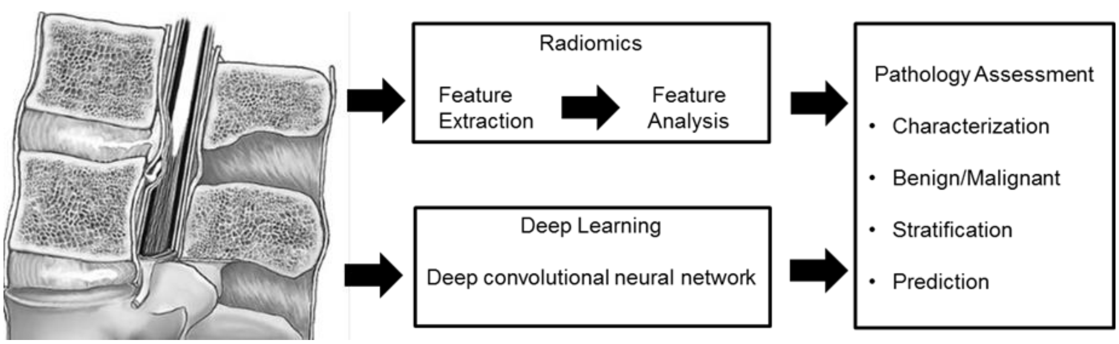


Figure 9. The primary difference between radiomic and deep learning methods for evaluating spine pathology in vivo. Radiomic methods are guided by programmed steps, whereas deep learning uses evolving multilayered neural networks methods to detect patterns without instructions.

My research revealed that data acquired through in vivo tissue interrogation with radiomic methods could be processed through a series of analytic steps developed to provide decision support for the radiologist (Figure 10). Aggregation and clustering of data integrated from a patient's electronic medical records, population databases, and disease models can be used to create a correlation matrix and heat map using numerous variables. A heat map refers to the visual representation of complex statistics and data relationships through plotting of color assigned to statistical criteria. Heat maps can be used as part of the differential diagnostic process. They have the potential to reveal new subvisual characteristics of disease.

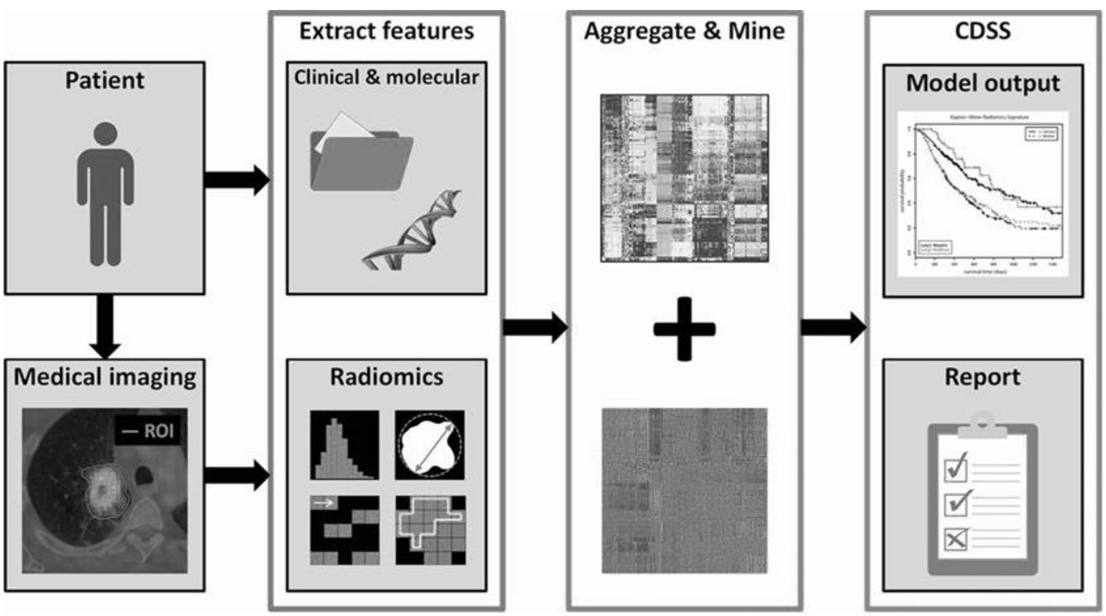


Figure 10. Sequential steps in the acquisition and analysis of non-visible data from a region of interest on an imaging study. The figure depicts the use of heats maps for aggregation and mining of data followed by the use of clinical decision support systems (CDSS). Permissions for graphics use granted by Nancy International Ltd, subsidiary of AME Publishing Company.

Multiscale in vivo tissue interrogation of pathology supports the conceptual development of a digital "virtual" biopsy. Echegaray et al. (2016) introduced the concept of the "digital biopsy," which refers to the targeted, non-invasive acquisition of pathology features in vivo (p. 283). Deep learning and radiomic methods are capable of performing simultaneous analysis and measures of many parameters across the entire realm of an imaging study data in a manner that exceeds human potential (Aerts, 2017; Scheckenbach, 2018). The combined capabilities support the concept of a virtual biopsy. AI-experts and radiologists who participated in this study concluded that the elements of this process could be further developed and applied to the spine. Successful use would likely expose new features of spine disease, expand the spectrum of pathology, and influence disease classification. The approach could map features of pathology across the full volume of a region of interest. Further development and application of a digital or virtual biopsy would help detect and characterize asymptomatic early-stage pathology, not detectable through normal anatomic imaging approaches. Success would require extensive training and validation of emerging AI methods. It would also require referential ground truth based upon greater knowledge of the biological correlates of imaging results.

Researchers have proposed that it will be essential in some cases that the spine be assessed and visualized in three dimensions (Cramer, Quickley, Hutchens, & Shah, 2017). With the use of advanced computational methods, the spine can be partitioned or segmented into a 3D matrix of well-defined spaces referred to as voxels or volumes of interest (VOI). This pattern of tissue registration and division supports 3D spatial mapping of pathology features. This process could be developed for use during a virtual biopsy or to help direct and guide a traditional needle biopsy. Multidimensional pathology feature mapping will support change analysis and inform treatment methods.

Spine pathology does not exist in isolation. It is interdependent with other tissues and systems. The state of spine health is subsequently influenced by surrounding microenvironments that comprise a biological ecosystem, which could be evaluated through multiscale in vivo tissue interrogation. In the near future, the virtual biopsy combined with interactive and immersive 3D displays of data could better define the margin, volume, and heterogeneity of pathology. In the future, greater emphasis will be placed on the assessment and treatment of transitional regions between the core of pathology and healthy tissues. This level of tissue insight is not available through methods other than the traditional biopsy that is limited in its scope. The use of voxel-based encoding will add spatial dimension to data and support further development of pathology mapping and modeling (Naselaris, Kay, Nishimoto & Gallant, 2011). Quantitative measures acquired through radiomic and/or or deep learning methods may be used in the future to calculate a pathology margin score, heterogeneity score, and an overall score of aggressiveness.

The desires of radiologists combined with the growing demand for more precise and personalized care will likely lead us in the direction of automated and manually prompted in vivo tissue interrogation and feature mapping. This approach combined with immersive data displays will allow the radiologist to obtain new actionable perspectives of spine pathology. Success will require the integration of patient, population, and application-based decision support during the interpretive stage of spine imaging workflow.

Differential Diagnosis: A Probabilistic Approach

The differential diagnosis process represents one of the most complex decision-making tasks in health care. The process refers to the development of a list of possible diseases or disorders felt to be the underlying cause of presenting signs or symptoms. The differential diagnostic list is created using a series of cognitive steps by a radiologist. The process is limited by the radiologist's knowledge of diseases, the finite nature of human competence, and the limited ability to address vast amounts of data and possibilities. The differential diagnostic process is complicated by the scale of acquired data and the level of available decision support.

The heuristic problem-solving approach used by humans exposed to a high degree of uncertainty is often weak and susceptible to variation and inconsistency. Disease-specific knowledge rather than general intelligence is required to be more accurate during the differential diagnostic process. To perform the task accurately the radiologist requires knowledge of all relevant diseases and their characteristic features. In the real world this situation does not exist, thus the reason for occasional misdiagnoses. Access to AI decision support during the interpretive stage of spine imaging can expand available knowledge and be used to assign weighted values to signs, symptoms, imaging features, and other evidence of pathology to develop a probability-based differential diagnostic list.

The spine is both intricate and complex. Many disorders that afflict the spine do not have a consistent pattern of signs and symptoms. To complicate matters further, many patients have more than one condition, therefore signs and symptoms can overlap, rendering the differential diagnosis more challenging. In confusing or complex cases, the differential diagnosis process often requires a combined probabilistic and deterministic approach. Advances in diagnostic imaging has and will continue to heightened awareness of the complexity and heterogeneity of disease at different scales, thus, influencing the differential diagnostic process (Aerts, et. al., 2013; Davnall et al., 2012; Lubner et al., 2017; McCue & McCue, 2017; Sala et al., 2017; Yip & Aerts, 2017). This is compounded by growing demands for a more personalized and precise diagnosis within the back drop of an expanding spectrum of pathology. The differential diagnostic process in spine imaging could integrate patient and population data, as well related decision support.

The complexity of decision making during the differential diagnostic process of spine imaging will continue to increase with growing awareness of the features of disease at multiple biological scales revealed with advanced imaging. The discovery of new imaging biomarkers and molecular signatures of spine pathology will complicate the differential diagnostic process and render a radiologist’s experience less relevant. Technological assistance is subsequently required during spine image interpretation to provide decision support and improve the accuracy of the differential diagnostic process in the presence of growing complexity and uncertainty.

The current differential diagnostic process in spine imaging is influenced by a radiologist’s familiarity with diagnostic possibilities and accessibility to imaging and non-imaging data at the time of image interpretation. Interpretive bias and/or limited awareness of differential diagnostic options leads to errors and/or oversights, which can result in missed spine care opportunities. This research study identified various AI applications that could augment the role of the radiologist in the differential diagnostic process. These applications include radiomic analysis of nonvisible data, application of change analysis algorithms, use of immersive data displays, and NLP access to contextual patient data from EMR. The integration and correlation of patient and population-based data supported by AI can score or assign probability to each element of the differential diagnosis. The combined use of human and machine intelligence, a process referred to as collective intelligence (CI) has the potential to provide valuable support during the diagnostic process.

Figure 11 represents how the flow of data and sequential applications of AI can be used to assist in the differential diagnostic process. This includes using natural language processing, radiomics, deep learning, and computational disease models. Structured and unstructured data can flow from the medical records and diagnostic tests to AI applications or processes used to identify and analyze actionable data at the radiology workstation. AI-supported methods can be used to detect and characterize disease, to assign values to disease features, to compare results to computational disease models, and to create a probability-based differential diagnosis. In the near future radiologists will use AI to augment the differential diagnosis of unusual, atypical, complex, ill-defined, and polymorphic disease.

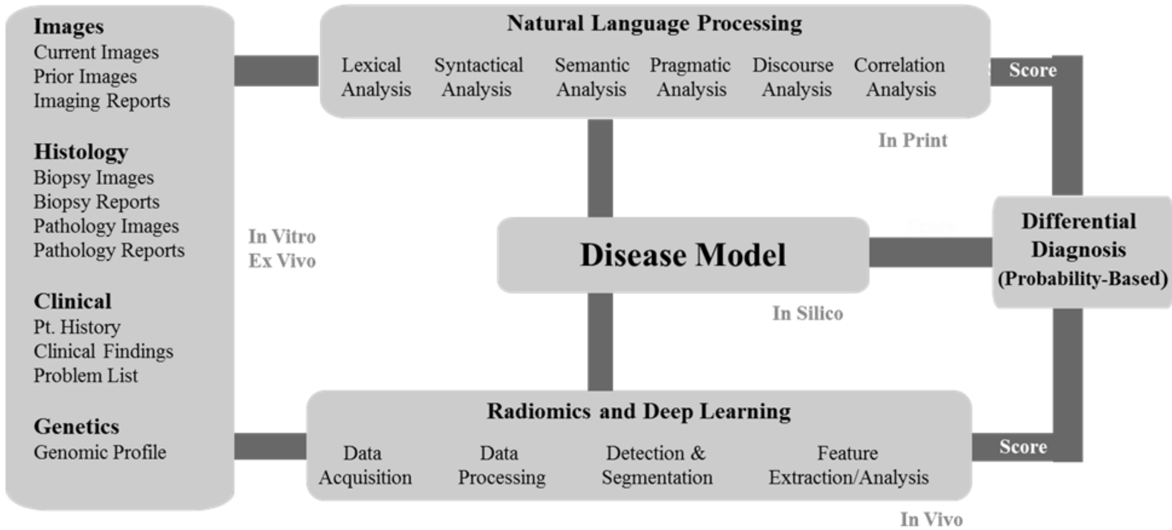


Figure 11. The flow of patient data from left to right exposing the data to decision support to achieve a probability-based differential diagnosis in spine imaging. The list of data to the left represents sources of structured and unstructured data.

There will be a transition from rule-based systems to autonomous deep learning systems capable of becoming more accurate with exposure to big data. During the differential diagnostic process, potentially relevant diseases can be scored and ranked in order of probability. Expanded disease criteria and new diagnostic thresholds will be developed to improve this process.

AI solutions have the capacity to narrow the number of diagnostic possibilities for the radiologist's consideration. AI methods also have the potential to identify diseases the radiologist is unaware of. Deep learning applications evaluate the complexities associated with atypical presentations, as well as coexistent and overlapping pathologies. The combined use of the attributes of human and machine intelligence will improve the differential diagnostic process during the interpretive stage of spine imaging. The collective approach will also provide the radiologist with data and insights that they are not currently receiving with traditional structural imaging approaches used in spine care.

Combinatorial AI Evolution and Spine Imaging

This research study revealed a high level of professional interest in various potential applications of AI during the interpretive stage of spine imaging. The majority of the research participants in this study were aware of the use of AI applications in other fields such as cardiology, oncology, and neurology. The greatest use has been in the specialties of brain and breast imaging.

AI and related technologies are evolving rapidly through a process often referred to as combinatorial evolution, also referred to as technology co-evolution. This refers to the paralleling development, convergence, and integration of different technologies and solutions. Focus group participants in this study addressed numerous potential applications of AI during the interpretive stage of spine imaging. Many AI and related technologies are being developed on a parallel tract and subsequently will converge to become part of an integrated solution. This concept is supported by the work of Tarassoli (2019), who noted that technologies do not remain singular but inevitably merge with other technologies leading to new applications and outcomes.

The phenomenon of combinatorial evolution not only applies to technology. It also applies to the use of data that has become evident in the converging fields of AI, genetics, radiology, and pathology. As technology advances, and the demands for its use and benefits increase, the duration between innovations becomes shorter. AI development and its use in spine imaging will be amplified by heightened awareness of the complexity of the spine and by the growing demand for more precise and personalized care.

AI is well suited to perform tasks that are too complex or time-consuming for a radiologist to perform during a single interpretive session. When research begins to reveal that AI provides clinically relevant decision support during the interpretive stage of radiology workflow there will be a greater push for more advanced imaging technology. This co-evolutionary process will lead to the acquisition of new actionable data, thus, adding to the degree of uncertainty and further complicating the decision making process. This pattern of recursive discovery, disruption, and adaptation will lead to new expectations and standards. Every time the need to interpret complex spine imaging data exceeds human limitations, new forms or levels of decision support will follow. The co-evolution of AI and related technologies combined with advances in diagnostic imaging will transform the field of spine care. The process will also support convergence of the fields of pathology, radiology, and genetics. It will also lead to new forms of professional collaboration that will influence how spine care will be delivered.

Multilevel Pathology Interpretation in Spine Imaging

Diagnostic radiology has long been recognized as an important element of spine care. Most spine imaging studies are interpreted by one radiologist who renders a qualitative report. Exposure to larger quantities of complex spine and related pathology data will place new burdens on the radiologist. Human judgment during conditions of complexity and uncertainty is often suboptimal. New interpretive solutions will be required. To resolve this matter, radiologists will require access to new levels of data analysis and decision support, similar to approaches that have been used in the field of histopathology.

A tiered hierarchy of human and machine intelligence will be necessary for the analysis of complex multidimensional spine imaging data. Complex or voluminous imaging data may have to be separated and allocated for machine or remote interpretation. The definition of multiple levels of decision support in this context refers to analysis that occurs in addition to that performed by the primary radiologist. Multiple levels of pathology assessment and interpretation can help overcome limited expertise, limited experience, and limited capabilities often attributed to single level interpretation (Williams, 2017). The primary goal of decision support is to offer assistance with problem solving and to help direct human actions (Greens, 2014). Widespread adoption of multilevel solutions will help democratize expert decision support.

Humans and computers each have unique attributes that can collectively be used during radiology workflow to analyze data, solve problems, render a diagnosis, and contribute to the final reporting process. Human intelligence is characterized by unique qualities such as intuition, abstraction, and adaptability, whereas AI offers an endless capacity for consistently detecting and characterizing patterns within vast amounts of data. The combined use of the attributes of human intelligence with those of AI will result in a form of collective intelligence resulting in more accurate disease detection, characterization, and monitoring (Figure 12).

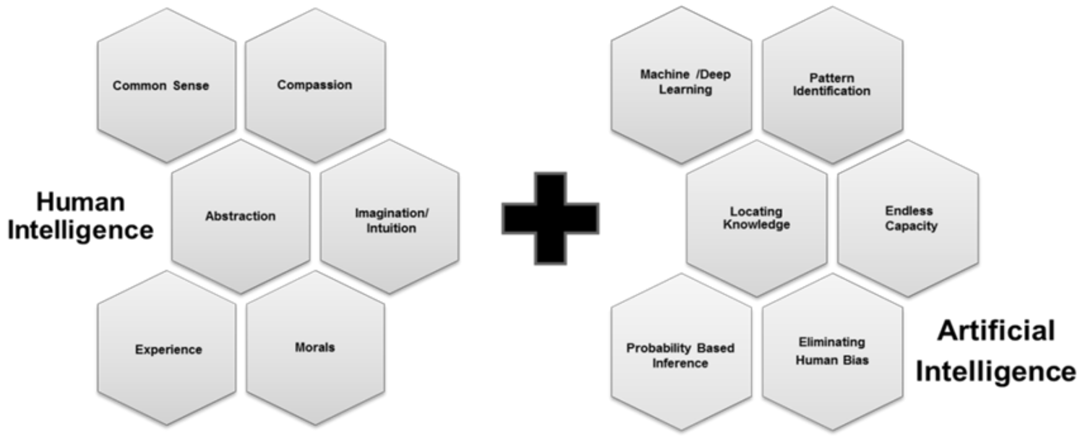


Figure 12 . Humans and computers each have unique attributes, that when combined, can improve decision making during the interpretive stage of imaging.

Diagnostic imaging represents the most commonly performed procedure used to detect and characterize spine pathology. With the exception of the biopsy, it represents the only procedure used to view spine and related tissues. Data acquired with multimodal imaging reveals characteristics of spine pathology at different biological scales that has to be interpreted. This may include a combination of mathematical (statistical), molecular, metabolic, and anatomic perspectives. The ability to acquire multiscale in vivo data from normal and abnormal biological states represents a form of digital pathology. Acquired digital tissue information can be mobilized and analyzed. Future interpretive workflow in spine imaging will subsequently begin to look more like the workflow in traditional histopathology. The workflow in the field of pathology is designed to provide access to multilevel decision support. In select cases, all or a portion of acquired data may be forwarded to a computational system and/or to one or more human specialists or subspecialists for advanced interpretation. The path of interpretation in some cases may involve a group of human experts in the form of a pathology board similar to the tumor boards used for the review of challenging oncology cases.

A single radiologist at a well-defined point of workflow currently interprets spine imaging studies in most settings. Spine imaging data can be processed, analyzed, and/or interpreted at many levels to render a more precise diagnosis or to measure change in challenging or complex studies (Figure 13). Access to tiered interpretive options could be automated or manually prompted by the attending radiologist. In the near future, the various paths of decision support depicted in Figure 13 may be integrated or may represent distinctly separate options at a radiology workstation. Continuous evolution and refinement of radiology workflow will result in the development of new interpretive options and paths not reflected in Figure 13. The relationship between human and machine interpretation of imaging data will change.

Selected spine imaging data could be allocated across institutional, regional, and national networks for remote analysis and/or interpretation. The potential benefits of this approach include broadened access to specialized levels of expertise and to state-of-the-art AI solutions. Additional benefits include access to population-based data and computational disease models. Multilevel decision support offers various benefits such as improved patient safety, increased diagnostic efficiency, rapid case archiving, rapid case retrieval, and timey diagnosis of urgent cases (Williams, 2017). Multilevel data interpretation could also provide access to emerging methods of disease scoring and staging, thus offering predictive and prescriptive value. This may include access to linear and nonlinear predictive modeling (Lakhani et al., 2017). Well-defined multilevel interpretive workflow would offer a built-in second opinion, diagnostic audit trails, and unique research opportunities.

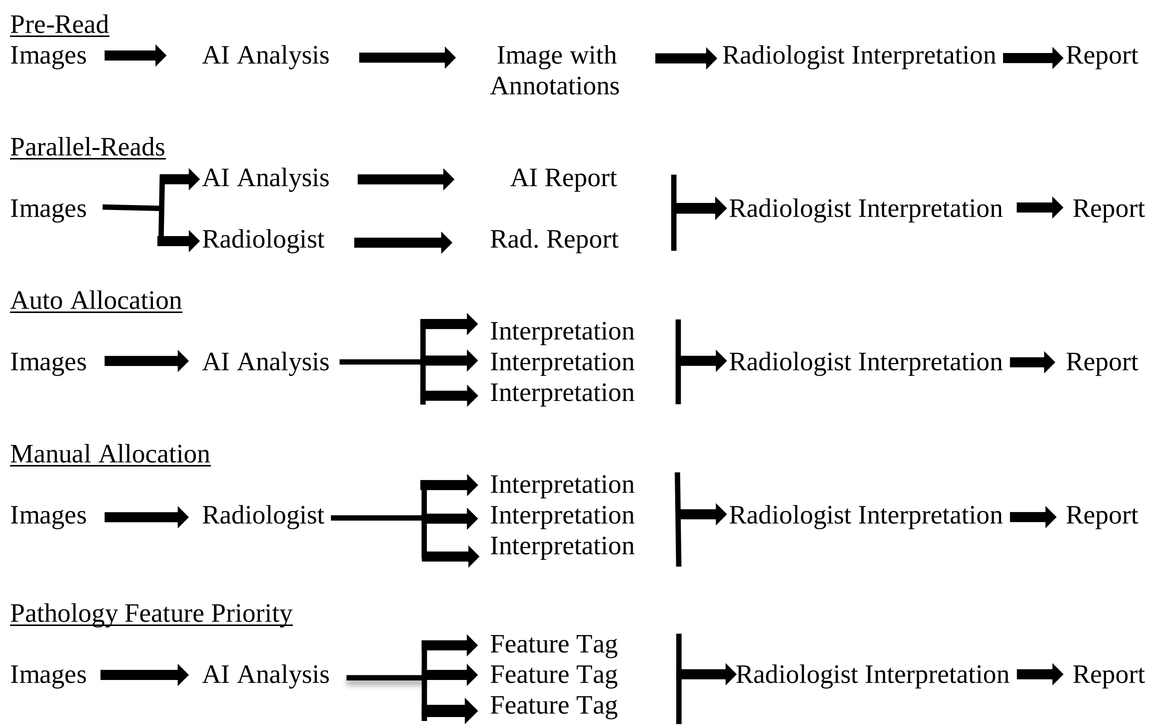


Figure 13. Many different paths can be followed for interpreting spine images. The process diagrams above represent conceptual examples of different methods of data flow and interpretation during spine imaging workflow. Future interpretive workflow will likely use numerous data paths.

The clinical utility of imaging data is highly dependent on the level and detail of the interpretive process. Clinical utility refers to the relevance, usefulness, and benefits of a technology, process or intervention in patient care (Lesko, Zineh, & Huang, 2010). Access to multiple levels of interpretive workflow in spine imaging could have a significant impact on the final radiology report. Future reports will be quantitative and structured. Some reports will offer a combination of anatomic, molecular, metabolic and statistical (mathematical) characterization of pathology. In the near future, there may be numerous expert signatures at the end of a spine imaging report.

An accurate pathology diagnosis is fundamental to effective spine care. At the current time, acquiring a second opinion for spine surgery or other forms of spine care is common. In the future, similar priorities may be assigned to the interpretation of complex spine imaging data. There will no longer be a need to limit the interpretation of complex spine data to human efforts or local talent. As previously stated, digital imaging data can be mobilized and allocated to numerous locations and resources for interpretation. The use of multiple levels of interpretation will help overcome the deficiencies and limitations of a single interpretive approach to spine imaging. The option of multilevel interpretation of a spine imaging study will augment the role of the radiologist for the benefit of the patient.

Future Predictions: Spine Imaging Workstation and Workflow

There have been unprecedented advances in diagnostic imaging capabilities during the last decade. The growing volume and complexity of data has created unique challenges for the individual radiologist in all fields including spine care. This evolutionary process has led to the development of new forms of data analysis and decision support.

The rapid combinatorial evolution of AI and related technologies will continue to disrupt the status quo while offering unprecedented opportunities for the radiologist during the interpretive stage of imaging. Those prepared for the paradigm shift will be in a position to offer better care to their patients. Radiologists who interpret spine imaging must embrace the potential role of AI and contribute to its development. Expanded knowledge of spine diseases and their characteristic imaging features will complicate the interpretive and reporting process. Spine care professionals who receive more detailed and structured reports from radiologists must also be prepared to use more detailed and objective diagnostic insights at the point of care.

The ongoing development of diagnostic imaging technology, imaging protocols, and AI methods will lead to new and integrated applications of AI during the various stages of spine imaging workflow. This will include the adoption of overlapping or coexistent AI applications during the pre-interpretive, interpretive, and post-interpretive stages of radiology workflow. Some of the applications will be automated providing the radiology was a manual override option. Other methods may require manual activation by the attending radiologist (Table 7).

It is important to predict the influence AI may have on the future role of the radiologist and on various stages of spine imaging workflow. Projected models of workflow can enhance readiness for change and facilitate early adoption of emerging solutions. I included this section of the chapter to provide a glimpse into the potential future of spine imaging. My basis for the conceptual description of the future radiology workstation and related workflow was derived from an extensive literature search and from the results of this research study. My proposed concepts are supported by current applications of AI in other fields of radiology. I developed a table to demonstrate the relationships between potential AI applications during the pre-interpretive, interpretive, and post-interpretive stages of radiology workflow (Table 7). The primary focus of this study was on the potential impacts of AI during the interpretive stage of spine imaging workflow.

Table 7

Application of AI During the Stages of Radiology Workflow

Application	Pre-interpretive Stage	Interpretive Stage	Post-Interpretive Stage
Patient scheduling	X		
Image protocoling	X		
Image acquisition	X		
Image quality analytics	X		
Post-processing registration of images	X		
NLP access to EMR	X	X	
Auto detection of abnormal	X	X	
Auto segmentation of pathology	X	X	
Characterization of pathology	X	X	
Change analysis	X	X	
Quantitative measures	X	X	
Targeted digital (virtual) biopsy		X	
Image/data display		X	
Integration of multi-omics data		X	
Computational decision support		X	
Structured reporting			X
Report accuracy analytics			X
Image and report archiving			X
AI training data			X

Note. The various processes or tasks of each stage of radiology workflow. Some of the AI applications may be found in more than one stage.

Future requisitions for spine imaging will contain required data that can be l be used to trigger automated NLP access to a patient's relevant medical records including prior radiology reports, pathology reports, and an active problem list. Elements of the diagnostic imaging requisition will be used to automate some of the imaging protocols. This process will likely precede or parallel automated radiomic and/or deep learning assessment of acquired imaging data. Both approaches could be performed during the pre-interpretive stage of workflow prior to human engagement. The radiologist will be able to manually activate either of the methods during the interpretive stage of imaging workflow. A menu of change analysis algorithms will be available at the workstation for automated disease surveillance. Some of AI-supported change analysis methods will be automated. Change analysis solutions may be applied during the pre-interpretive and/or interpretive phase of spine imaging workflow.

Prior to interpreting spine images, the radiologist of the future may have access to an initial list of differential diagnostic possibilities based on automated pre-interpretive analysis. Each of the differential diagnostic possibilities may be assigned a level of statistical probability. The list will be modifiable by the radiologist after interpretation of the images and imaging data. Prior to interpreting the spine imaging study the radiologist could have access to flagged regions of abnormality, an initial draft of a structured radiology report, and contextual information from the medical records. In the near future, the radiologist will have access to a menu of narrow AI applications at the workstation that could purposefully interrogate regions of interest and perform quantitative measures. The menu of options will likely include the ability to perform multilevel in vivo tissue interrogation (virtual biopsy), as well as having access to change analysis tools to evaluate pathology. Current research demonstrates that change detection algorithms can be used to evaluate pathology over time (Patriarche & Erickson, 2007). The menu of decision support tools may include special AI applications developed for use on specific tissues, anatomic regions, or for the assessment of specific disorders involving the spinal cord, bone marrow, cerebrospinal fluid, spinal muscles, and intervertebral disc. The latter may involve unique combinations of algorithms or deep learning paths designed for specific pathology.

The radiologist of the future will have access to radiomic and deep learning solutions designed to complement the visual assessment of pathology during the interpretive stage of spine imaging. Research has already demonstrated the success of using a hybrid approach with a combination of hand-crafted radiomic (HCR) and deep-learning radiomic (DLR) methods to characterize nonvisible features of pathology in other tissues (Bizego et al, 2019; Bodalal, Trebeschi, & Beets-Tan, 2018; van Griethuysen et al., 2017). Spine imaging studies represent a mineable database. Subsequently, a continuous loop of automated mining of spine imaging data may be embedded into radiologic workflow. Automated voxel-wise analysis may be used to free up the radiologist to focus on the interpretation of manually selected regions of interest and to consult with referring providers.

Future radiology workstations will provide the radiologist with access to AI-supported data analysis methods that can be used to reveal information about various dimensions of pathology spanning from anatomic to molecular levels. AI methods will reveal "data blind spots" and help identify or retrieve missing information from medical records or from imaging data. The future radiology workstation will likely provide access to computational disease models and population-based information, which can help classify spine pathology and offer predictive insight.

In the near future, the radiology workstation will offer numerous options to the radiologist for applying AI solutions (Figure 14). A menu of narrow AI applications could screen for and characterize specific spine disease features such as edema, fibrosis, stenosis, vertebral deformities, bleeds, ischemia, fracture, and demyelination. The menu of options may also include specialized deep learning applications and/or algorithms developed to target a specific tissue or disease process. The workstation may also have the option of broad or general AI applications capable of detecting and characterizing a variety of different diseases and disorders.

The radiologist will likely have the option to choose from a menu of AI applications that can expand or confirm an automated process. One of the options will be a digital “virtual” biopsy tool that could further interrogate a known region of pathology or another defined region of interest. The radiologist will also have access to 3D pathology feature mapping tools to enhance the diagnostic process. These tools will be capable of mapping molecular relationships, radiomic features, and blood flow.

Continued use of traditional anatomic and qualitative interpretive methods in spine imaging will result in overlooked data and missed features of pathology. Some features of pathology and their relationships to surrounding tissues are unseen or misinterpreted with traditional 2D anatomic visualization. A limited display of data can compromise a radiologist’s ability to make fully informed diagnostic decisions and treatment recommendations. Radiologists require new solutions. In the near future, the radiologist will likely have the option of choosing an immersive display of data to explore a region of interest from different visual perspectives and at different biological scales. Future solutions will include access to virtual reality (VR), augmented reality (AR), and interactive mixed reality (IMR) applications at the radiology workstation. Interaction with a 3D display of data will enhance human insight and intuition (Figure

15). Insight refers to a deep understanding of a topic, whereas intuition refers to ability to respond to the environment without using conscious reasoning.

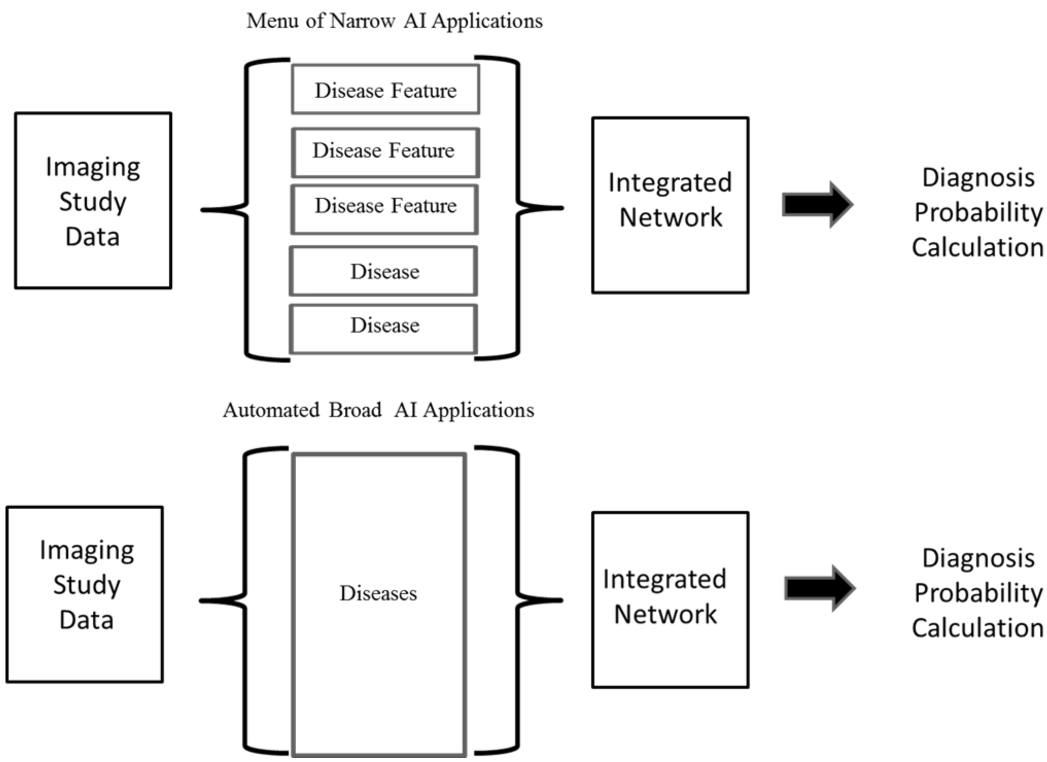


Figure 14. Two general categories of AI applications at the radiology workstation. This includes a menu of narrow AI options and automated broad application of AI.

Real-time interaction with a 3D display of data will enhance interpretive capabilities and enhance the ability for a radiologist to provide the patient with more precise and personalized solution for care. This can be helpful when evaluating the spine that is both intricate and complex. Features of pathology derived from AI-supported molecular, radiomic, or deep learning methods, may be embedded within or color mapped onto 3D depictions of pathology, thus supporting the in vivo interrogation (virtual biopsy) of the full volume of pathology.

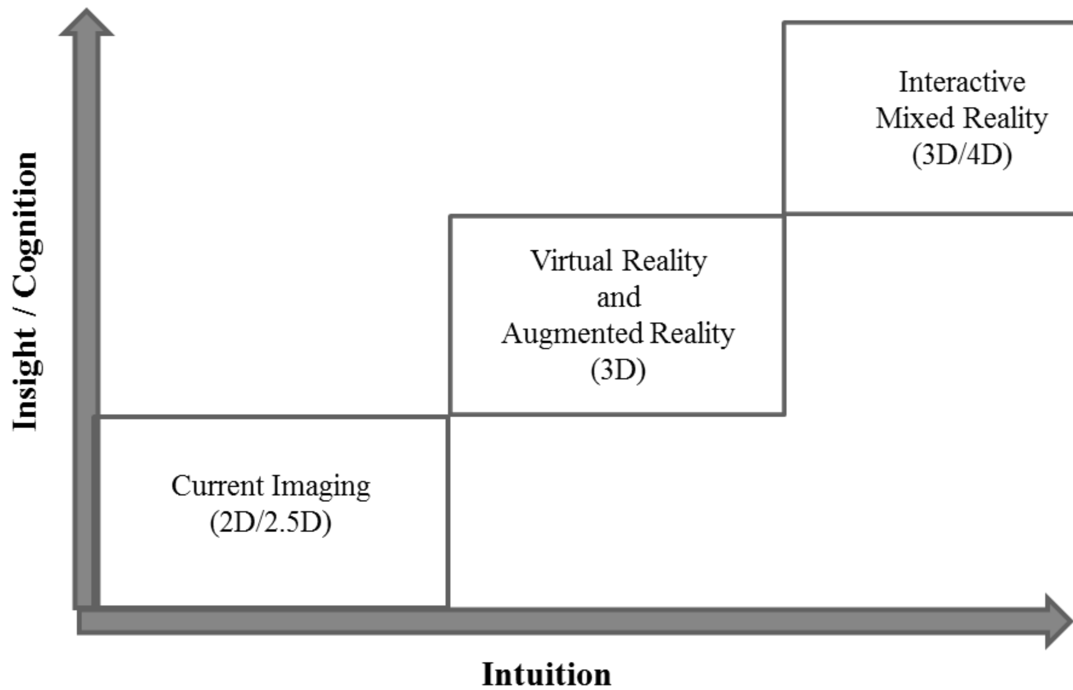


Figure 15. A conceptual depiction of how an immersive display of complex imaging data could impact a health care provider's insight and intuition about a patient's health status. Interactive mixed reality (IMR) refers to the blended use of virtual and augmented reality to display the features of normal tissue and/or pathology derived from advanced imaging data.

Data derived from computational disease models may be fused with imaging data in an immersive display to project the evolution of pathology or to predict a post-interventional outcome. In spine care, the use of immersive 3D displays of data may be helpful to providers such as interventional radiologists and surgeons to help plan and/or guide their approach. Special software may become available for performing pathology margin (edge) assessment and measures. Research has already demonstrated that radiomic profile maps can be used to evaluate specific zones of pathology in some tissues (McGarry et al., 2019).

The field of spine imaging is evolving rapidly, supported by innovative methods of data acquisition, data analysis, and decision support. I propose that AI will remove redundant and time-consuming tasks from radiologists' workloads allowing them to engage more with the patient and health care providers. The three principle types of AI are assisted intelligence, augmented intelligence, and autonomous intelligence (Figure 16). There will be a perpetual transition from assisted intelligence, to augmented intelligence, and autonomous intelligence. The future workstation will likely have

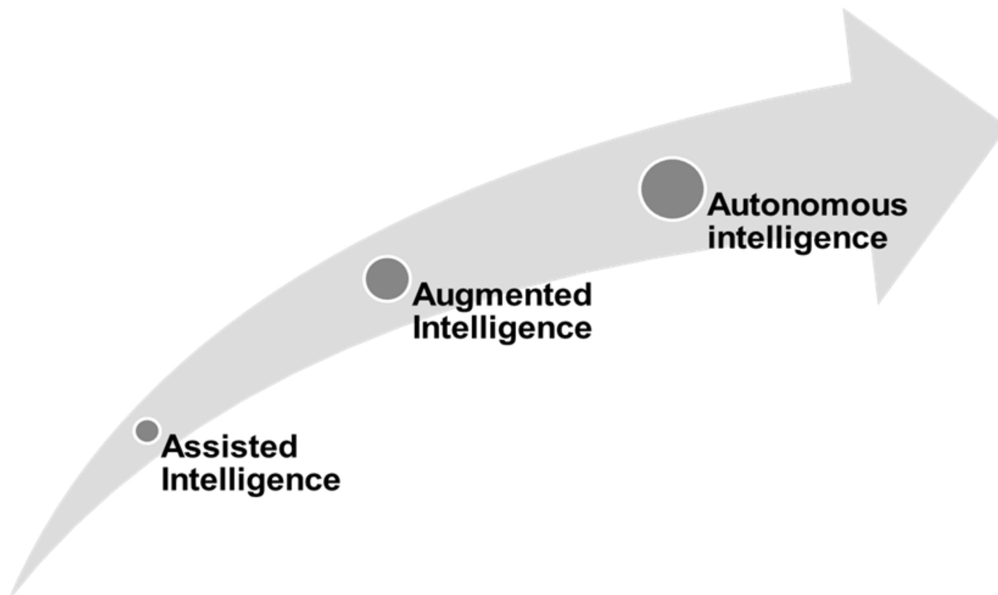


Figure 16. Perpetual advances in computer technology and algorithm development supports the rapid evolution of artificial intelligence. The trend is moving toward a elements of each form of AI. It is quite likely that the concept of the digital "virtual" biopsy will become widely accepted and eventually embedded into the imaging workflow of all specialties. greater dependency on machines to perform complex tasks and to augment decision support.

The radiology workstation of the future will be designed to augment the role of the radiologist. It will incorporate patient-based decision support with population-based decision support with bridging technologies to provide a more precise and personalized diagnosis. The implementation of networking and teleradiology solutions at the radiology workstation of the future will provide access to multilevel pathology interpretation. In the future it will not be necessary for an individual radiologist to assume the entire burden of making highly complex decisions by themselves. The role of the radiologist will be augmented by access to multiple levels of integrated decision support. The collective intelligence of human experts and machines will lead to better care than could be provided by either one alone (Lakhani et al., 2017). Despite the important role of technologies, the human element will always remain the most critical step in the final diagnostic process of spine imaging.

In summary, the radiology reading room of the future will be transformed into a diagnostic hub embedded within an AI ecosystem with access to radiomic, pathomic, and genomic data. It will serve as the epicenter of complex decision support in spine care and in other specialty fields. The future radiology workstation will empower information technologies, as well as the collective roles of human and artificial intelligence, thus offering a more precise and personalized diagnosis.

Limitations of the Study

This research study had numerous limitations. The relatively small sample size in each focus group session limited the ability to generalize the results, although the smaller size offered more in-depth contributions from participants. The relatively short duration of this study during a period of unprecedented development and evolution of AI in radiology may overshadow the relevance of some of the study findings at the time of publication. For this reason, the conclusions of the study include the predicted impact of

AI applications on various aspects of radiology workflow and on the radiology workstation. This includes the potential impact of the conceptual digital “virtual” biopsy.

I purposefully selected focus group participants based on their level of experience and expertise. However, the sample was not representative of all of the experts required to develop, implement, and support the development and implementation of AI solutions in the radiology setting. This process often involves other disciplines such as physicists, bioinformaticists, imaging technologists, computer engineers, and health care administrators. Furthermore, the opinions of experts may differ across health care and research systems.

The absence of direct communication between AI experts and radiologists in this study did not allow for collaborative problem solving or predictions, but it did allow for more in-depth discussion within each specialized domain. Working with high-profile experts can influence research conclusions due to their ability to over intellectualize contributions during focus group sessions. I was able to reduce the impact of this effect through triangulation of data from numerous experts, expert resources, and expert publications.

Recommendations for Further Research

There has been limited research on the potential impacts of AI use during the interpretive stage of spine imaging workflow. Recent work in other fields has demonstrated its potential for use in spine imaging. In limited settings, AI applications have been successfully used to localize and label structures, segment tissues, and provide diagnostic decision support (Galbusera, Casaroli, & Bassani, 2018). This qualitative exploratory research study was designed to offer insight about the potential impacts of AI during the interpretive stage of spine imaging workflow. The study revealed expert derived themes and subthemes that provide a fundamental basis for future discussion and research. Meaningful clinical applications of AI were disclosed during the course of the study along with various AI solutions that could be used to better detect, characterize, and monitor spine disorders.

This research study confirmed that one of the biggest challenges for developing AI solutions in spine care is gaining access to large volumes of curated imaging data from different sources for AI training and validation. Further research is required to identify efficient methods for acquiring AI training data. Research is also required to determine how spine images could be efficiently annotated including the possibility of using computational reference models along with automated or semiautomated labeling methods.

The converging fields of radiology, pathology, genetics, and computer science will lead to unprecedented challenges and opportunities in radiology, including the field of spine imaging. The results of this convergence will be disruptive to the status quo. This opens the door for research in many areas. Studies are required identify the biological correlates of pathology features acquired through radiomic and deep learning methods. Research is required to establish biologic correlates to statistical, structural, and textural radiomic features of spine pathology. This represents a necessary step in the development of ground truth. It is also essential to explore how AI could address the relationships between probability and causality as it relates to the differential diagnostic process and spine disorders.

Radiomics has demonstrated some clinical utility in oncology. In isolated cases, it has been successfully used for cancer diagnosis, tumor detection, tumor classification, attribute scoring, survival prediction, recurrence prediction, and disease staging. Future research is required to identify whether similar approaches could be adapted or developed for use in spine care. The consistency and relevance of AI-supported deep learning and radiomic methods must be validated for use in spine care. This is also required of methods used to perform a digital "virtual" biopsy. Research is also required to identify clinically relevant radiomic and deep learning biomarkers and r signatures of of pathology. Further research is required to help identify whether the use of AI will lead to better patient care outcomes. This approach may include identifying the best method to integrate human and machine intelligence, taking into account the strengths and weaknesses of each.

Radiomics is a logical process "explainable" to humans, whereas deep learning is not as transparent, rendering it difficult to determine how it works in most cases. Greater transparency of the "black box" methods of deep learning need to be achieved. There are ways to help determine how "deep learning" processes achieve a particular outcome. An example is the use of probability heat maps that may reveal how algorithms assigned higher probability of a disease or disease features. This work will have ethical and legal ramifications. Research is required to identify methods that may be used to provide improved transparency of the "black box" approach offered by deep learning methods.

Successful development and use of AI solutions in future spine imaging will require additional research using various methods. Applied research is necessary to help determine what is required to meet the criteria for “seamless integration” of valid AI solutions in workflow and to define what is required to simplify its use. Systems research will be essential to help identify the necessary components and relationships within an effective AI ecosystem. Additional research is also required to investigate various sociotechnical ramifications associated with AI use in spine imaging. This should include refining the criteria and goals of clinical utility and meaningful use applications.

Scientific investigation is required to determine how the use of AI solutions will impact the structured reporting process and point of care decision making. This includes addressing how common data elements used in structured reporting could help identify ground truth and train AI systems. Greater knowledge of the economic impact, ethical dimensions, and liabilities associated with AI is required. Additional qualitative exploratory research studies should help conceptualize the radiology workstation (diagnostic cockpit) of the future.

Reflexive Statement

The results of analyzing data from consensus-based white papers, focus group sessions, and reflective journaling confirmed that radiologists desire to render a more accurate imaging diagnosis with the assistance of various forms of decision support. There was good correlation between white paper predictions and the discussions that took place in the AI expert and radiology focus group sessions. There was also good content and thematic alignment between the literature search and all other sources of data in this research study.

The consensus-based white papers placed emphasis on population-based decision support involving system architecture and curated data pipelines. The radiology focus group participants placed a high priority on patient-based decision support at the level of the radiology workstation and with structured reporting. The discussions that took place in the AI focus group sessions tended to be more consistent with the challenges raised in the consensus-based white papers and less focused on specific clinical applications.

Radiologists have expressed their desire for pre-interpretive data analysis and immediate access to relevant contextual information from a patient's electronic medical records prior to beginning the interpretive process. During the beginning of the focus group session the radiologists initially focused on missing contextual information from the medical records and the potential application of AI for structured reporting. By the end of the radiology focus group session there was greater interest in the potential benefits of using AI-supported methods to perform multiscale in vivo tissue interrogation and to obtain probability-based differential diagnostic support. There was also a significant level of interest in the concept of the digital "virtual" biopsy.

Implications for Social Change

Spine disorders represent one of the leading causes of chronic pain and disability, both that have a devastating impact on the individual, family, and on society. Any diagnostic solutions that lead to more preventive, preemptive or personalized care will help reduce the burdens associated with spine and related disorders. One of the most prevalent and important diagnostic steps in spine care is imaging. Spine care providers of all disciplines are highly dependent on insights obtained through radiology reports. Better decision support with the help of AI during the interpretive stage of spine imaging workflow has the potential to render a timelier and more precise diagnosis, thus, reducing the incidence of diagnostic errors and missed opportunities. More accurate and descriptive radiology reports that support more precise and personalized intervention will contribute to better spine care outcomes

Training with annotated data on a massive scale will enable robust AI solutions and help democratize decision support (Hosney et al., 2018). The adoption of multilevel pathology interpretation in spine imaging will help democratize diagnostic decision support. Improving teleradiology platforms will help mobilize spine imaging data to off-site interpretive solutions that may include artificial and/or human intelligence driven decision support. Access to this level of interpretive assistance can help overcome professional deficiencies and resource inequalities. It could to provide interpretive services to underserved facilities or regions.

Successful use of AI during the interpretive stage of spine imaging workflow will have a profound and favorable impact on point of care decisions. It will influence how patients are evaluated, treated, and managed. Successful use of AI decision support will help reduce the complications and costs associated with diagnostic errors and missed opportunities. AI decision support will also facilitate the development of new imaging technologies and protocols capable of acquiring more data, thus leading to more personalized care. New methods will be capable of detecting and interrogating nonvisible features of spine pathology in vivo. This will support the discovery of new biomarkers and molecular signatures of spine pathology. New pathology insights and measures will facilitate new methods of intervention and treatment expectations.

AI-based prioritization of diagnostic imaging interpretation in emergency settings will improve the potential for timely intervention and improved therapeutic outcomes. Automated surveillance of disease with quantitative change analysis algorithms will help reveal early evidence of disease progression; thus, alerting the radiologist to make timely recommendations to referring health care providers. It will provide new dimensions of treatment outcome measures. The development of successful AI applications in spine imaging can be adapted and have a positive impact in other specialties of radiology. The potential societal impact of early disease detection, timely intervention, and evidence-based spine care is unmeasurable.

Successful use of AI supported molecular, radiomic, and deep learning methods during the interpretive stage of spine imaging will offer precise insight for treatment planning and surgical approaches. Three-dimensional in vivo assessment of spine pathology will also have a profound impact on how spine care is delivered. Furthermore, the integration of deep learning and radiomics will establish new imaging biomarkers and other measures of pathology that will be recognized and respected by all spine care disciplines. The process will help reveal the fundamental basis of spine disease, thus, supporting improved multidisciplinary collaboration. Expanded disease classifications and improved stratification of spine pathology will set the stage for more personalized care.

The opinions and views of experts in this study represent an important contribution to the concept of the digital (virtual) biopsy that can be further developed and used in all fields of radiology. Future application of the digital "virtual" biopsy could revolutionize how pathology is evaluated and characterized. New standards would evolve. Successful use of a digital "virtual" biopsy approach would help overcome the challenges associated with limited sampling of pathology performed with the traditional needle biopsy. Non-invasive, in vivo characterization of pathology will provide an effective method for evaluating the full volume of pathology. The concept of the virtual biopsy is not limited to the spine but could be applied to other areas of the body and regions of pathology.

Successful use of AI solutions will support the development of "best practice models." This will facilitate further adoption and use of the technologies. Successful use of AI will also support additional innovations resulting in a compounding effect on related AI solutions and technologies. This can result in efficient use of imaging data during radiology workflow. The projected co-evolution and compounding effect of AI development is consistent with many of the theoretical constructs of the DOI theory and the technology acceptance model. This includes perceived usefulness, ease-of-use, normalization, and interoperability.

Many challenges and opportunities are associated with the emergence of AI solutions in radiology. Radiologists will need to sort through the hype to identify whether proposed solutions have been validated. This will in part require reliance on peer-reviewed publications of best practice outcomes and consensus-based opinions. Narrow applications of AI solutions will be developed and accepted for use in radiology prior to the adoption of broad or general AI applications. Radiologists and all spine care providers should embrace current and future applications of AI and participate in related educational and training opportunities.

Conclusions

Spine disorders represent one of the most common causes of pain and disability. With the exception of the history and clinical examination, spine imaging often represents the single most important diagnostic procedure in spine care. Successful treatment outcome is dependent upon an accurate and timely diagnosis. The current state of spine imaging interpretation is relatively imprecise, inconsistent, and often limited to the qualitative description of late stage disease. Spine imaging data are complex, and the stakes are high, thus supporting the development of new methods of data acquisition, data analysis, and decision support.

AI solutions help overcome human factors that contribute to interpretive error in diagnostic imaging such as cognitive limitations and bias, compounded by increasing study complexity and volume. The incidence of missed abnormalities and misdiagnosis is too high in radiology. The ability to characterize and monitor pathology using traditional anatomic imaging is limited. AI can be used during the interpretive stage of radiology workflow to detect patterns and combine information in a way that exceeds human potential. AI and related computational decision support is required to augment the role of the radiologist during the interpretive stage of spine imaging. Successful adoption and use of validated AI applications has the potential to reduce interpretive errors, overcome missed opportunities, and help provide a more accurate and personalized spine diagnosis.

Further development of AI will support the ability to perform multiscale in vivo detection and characterization of spine pathology at multiple biological scales. Visible and nonvisible patterns in the data will need to be analyzed and interpreted. The process will contribute to the development of automated screening methods and the digital "virtual" biopsy. AI will also be used to evaluate temporal changes in pathology and proactively prioritize spine imaging studies that require immediate or focused attention. Co-development and integrated use of AI-supported methods such as radiomics, deep learning, natural language processing, change analysis, and immersive data displays will be developed for use during the interpretive stage of spine imaging workflow. AI solutions will improve the ability to detect, characterize, and monitor spine pathology. AI supported methods will also improve the ability to classify and stage spine pathology. AI will support the development of computational disease models that will help assign value to disease features, which could factor into diagnostic and prognostic conclusions.

Bridging technologies and platforms will to improve access to images and facilitate the flow of imaging and related data. The radiologist will subsequently have access to numerous paths of image interpretation and decision support for unusual or complex pathology, including integrated forms of human and artificial intelligence, referred to as collective intelligence. In addition, a hierarchy of imaging interpretation will likely become available to offer secondary and tertiary opinions on challenging cases. Teleradiology and shared databases will democratize expert decision support.

Widespread adoption of successful interpretive solutions will support further development of imaging technology and protocols for acquiring new data.

In the near future, the radiology workstation will transform into a digital diagnostic hub, a virtual platform used to aggregate and analyze multi-omics data to provide more personalized spine care. The differential diagnostic process in spine imaging will expand beyond the expertise and skills of the individual radiologist. AI methods will be used to integrate spine imaging data with non-imaging data to support a probability-based differential diagnostic process. The combined use of patient and population-based insights will refine the approach. A better understanding of the fundamental basis of spine disorders combined with shared solutions for decision support and common diagnostic criteria will facilitate multidisciplinary collaboration.

Widespread adoption of AI for use during the interpretive stage of spine imaging workflow will require heightened awareness of its reliability, as well as knowledge of validated applications and its clinical utility. Prior to use, the clinical utility, accuracy and robustness of AI applications must be validated (Galbusera, Casaroli, & Bassani, 2018). Success requires access to an adequate volume of curated training data, ground truth, and collaboration between stakeholders to help ensure ethical and trustworthy use.

Diagnostic imaging will play a progressively more fundamental role in the evaluation and care of the spine. Structured imaging reports will evolve as mineable patient data spaces, adding to decision support at the point of care. Advances of multiscale in vivo tissue interrogation will reveal new features of pathology, which will lead to a better understanding of disease processes and more biological solutions. AI supported approaches will be used to help detect, characterize, and monitor the mathematical, molecular, microscopic, and macroscopic features of spine disorders. New imaging biomarkers and signatures of pathology will emerge.

Convergence of the fields of pathology and radiology combined with advances in computer science will render AI progressively more influential in decision support. In the near future, the use of radiology in spine care will no longer be limited to visual interpretation of images and the provision of a subjective and qualitative report. The process of spine imaging interpretation will transform into a computational and quantitative science. Imaging biomarkers and radiomic signatures of pathology will assist spine care providers in choosing the best evidence-based treatment options. Better detection and characterization of early-stage spine pathology will support pre-emptive and conservative care (Hussain et al., 2019). An earlier spine diagnosis will also support minimally invasive intervention.

This research study revealed a desire on the part of radiologists to improve interpretive spine imaging workflow and accuracy. AI experts acknowledged that the needs of radiologists could be met. Prioritized clinical applications included spinal cord disorders, bone marrow pathology, vertebral compression deformities, fractures, and spine pain. The spinal cord was given the greatest attention. The primary clinical goals were early detection and characterization of pathology. The study revealed an interest in developing more sensitive and specific imaging biomarkers with good biologic correlation. There is a well-defined need to detect changes in tissue structure, composition and function at non-visible scales. This includes the ability to detect evidence of chemical shifts, ischemia, inflammation, neoplasia, and microstructural changes in tissues. There is also an interest in better characterization of pathology volume, heterogeneity, and margin contours.

The co-evolution of AI and related technology will facilitate the transition from hype to hope for practical application. Radiologists and spine care providers must both embrace the potential of AI and participate in its development. AI will transform the field of spine imaging into a more quantitative and objective specialty. The radiologist will become one of the primary gatekeepers of big data and decision support in spine care. The radiologist will also become highly informed, sought-after clinical consultants. Spine care providers of all disciplines will benefit from the augmented role of the radiologist. They will be empowered at the point of care with more accurate and descriptive imaging conclusions. An AI ecosystem will evolve providing new perspectives of spine pathology. The process will support the transition to preventive, preemptive, and personalized spine care. Prior to this, there is much work to be done developing, training, validating, and supporting AI applications.

References

Acharya, U. R., Hagiwara, Y., Sudarshan, V. K., Chan, W. Y., & Ng, K.H. (2018). Towards precision medicine: From quantitative imaging to radiomics. *Journal of Zhejiang University, 19*(1), 6-24. http://doi.org/10.1631/jzus.B1700260

Adduru, V., Michael, A. M., Helguera, M., Baum, S. A., & Moore, G. J. (2017). Leveraging clinical imaging archives for radiomics: Reliability of automated methods for brain volume measurements. *Radiology, 284*(3), 862-869.

Aerts, J. W. L. (2016). The potential of radiomic-based phenotyping in precision medicine. *Journal of the American Medical Association Oncology, 2*(12), 1636-1643. http://doi.org/10.1001/jamaoncolo.2016.2631

Aerts, J. W. L. (2017). Data science in radiology: A path forward. *Clinical Cancer Research,* epub ahead of print. http://doi.org/10.1158/1078-0432.CCR-17-2804

Aerts, J. W. L., Velazques, E. R., Leijenaar, R. T. H., Parmar, C., Grossmann, P., Cavalho, S., . . . Lambin, P. (2014). Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nature Communications*, 5, 4644. http://doi.org/10.1038/ncomms5006

Alkasab, T. K., Bizzo, B. C., Berland, L. L., Nair, S., Pandharipande, P. V., & Harvey, H. B. H. (2017). Creation of an open framework for point-of-care computer-assisted reporting and decision support tools for radiologists. *American College of Radiology, 14*, 1184-1189. http://doi.org/10.1016/j.jacr.2017.04.031

Allen, B., Seltzer, S. E., Langholz, C. P., Dreyer, K. P., Summers, R. M., Petrick, N . . . Kandarpa, K. (2019). A road map for translation research on artificial intelligence in medical imaging: From the 2018 National Institute for Health/RSNA/ACR/The Academy workshop. *Journal of the American College of Radiology*, 1-11.

American College of Radiology (2016). *Letter to the United States Office of Science and Technology Policy.* Retrieved from https://www.acr.org/-/media/ACR/Files/Advocacy/Meaningful-Use/acr_comments_ostp_rfiai_7222016.pdf

Andel, *C.*, Davidow, S. L., Hollander, M., & Moreno, D.A. (2012). The economics of health care quality and medical errors. *Journal of Health Care Finance, 39*(1), 39-50.

Augimeri, A., Cherubini, A., Cascini, G. L., Galea, D., Caligiuri, M. E., Barbgallo, G., . . . Quattrone, A. (2016). CADA: Computer-aided DaTSCAN analysis. European *Journal of Nuclear Medicine and Molecular Imaging Physics, 3*(4), 1-13. http://doi.org/10.1186/s40658-016-0140-9

Avanzo, M., Stancanello, J., & Naga, I. (2017). Beyond imaging: The promise of radiomics. *European Journal of Medical Physics, 38*, 122-139. http://doi.org/10.1016/j.ejmp.2017.05.071

Bak, S. H., Park, H., Lee, H. Y., Kim, Y., Kim, H. L., Jung S. H., . . . Park, K. (2018). Imaging genotyping of functional signaling pathways in lung squamous cell carcinoma using a radiomics approach. *Scientific Reports, 8*(1), 3284. http://doi.org/10.1038/s41598-018-21706-1

Balagurunathan, Y., Gu, Y. H., Wang, H., Kumar, V., Grove, O., Hawkins, S . . . Gillies, R. J. (2014). Reproducibility and prognosis of quantitative features extracted form CT images. *Translational Oncology, 7*(1), 72-87. doi.org/10.1593/tlo.13844

Baldwin, T., Guo, Y., & Syeda-Mahmoood, T. (2017). Automatic generation of conditional diagnostic guidelines. *AMIA Annual Symposia Proceedings*, 295-304.

Barbour, R. S. (2001). Checklists for improvising rigour in qualitative research: A case of the tail wagging the dog. *British Journal of Medicine, 322,* 1115-1117.

Bauer, J. C. (2017). 2017 annual meeting of the American College of Radiology-Moreton lecture: Forecasting the futures of radiology. *Journal of the American College of Radiology, 14,* 1524-1529. http://doi.org/10.1016/j.jacr.2017.08.020

Baugh, N., McNallen, A., & Frazelle, M. (2014). Concept mapping as a data collection and analysis tool in historical research. *The Qualitative Report, 19*(13), 1-10. Retrieved from http://nsuworks.nova.edu/tqr/vol19/iss13/3

Baxter, J., & Jack, S. (2008). Qualitative case study methodology: Study design and Implementation for novice researchers. *The Qualitative Report, 13*(4), 544-559. Retrieved from http://nsuworks.edu/tqr/vol13/iss4/2

Beam, A. L., & Kohane, I. S. (2018). Big data and machine learning in health care. *Journal of the American Medical Association, 319*(13) 1317-1318. http://doi.org/10.1001/jama.2017.18391

Bejnordi, B. E., Veta, M., & van Diest, P. J. (2017). Diagnostic assessment of deep learning algorithms for detection of lymph node metastasis in women with breast cancer. *Journal of the American Medical Association, 318*(22), 2199-2210. http://doi.org/10.1001/jama.2017.14585

Belmont Report (1979). *The Belmont Report: Ethical principles and guidelines for the protection of human subjects of research.* Retrieved from hhs.gov/ohrp/humansubjects/guidance/belmont.html

Berlin, L. (1996). Perceptual errors. *American Journal of Radiology, 167,* 587-590.

Berlin, L. (1994). Reporting the "missed" radiological diagnosis: Medicolegal and ethical considerations. *Radiology, 192,* 183-187.

Berlin, L. (2007). Radiological errors and malpractice: A blurry distinction. *American Journal of Roentgenology, 189*(3), 517-522.

Berlin, L. (2013). Radiologic errors, past, present and future. *Diagnosis, 1*(1), 79-84. http://doi.org/10.1515/dx-2013-0012.

Bizzego, A., Bussola, N., Salvari, D., Chierici, M., Maggio, V., Jurman, G. & Furlanello, C. (2019). Integrating deep and radiomics features in cancer bioimaging. *BioRxiv,* 108. https:doi.org/10.1101/568170

Bloomberg, L. D., & Volpe, M. F. (2008). *Completing your qualitative dissertation: A roadmap from beginning to end.* Los Angeles, CA: Sage.

Bodalal, Z., Trebeschi, S., & Beets-Tan, R. (2018). Radiomics: A critical step towards integrated health care. *Insights into Imaging, 9*(6), 911.

Boone, D., Mallet, S., McQuillan, J., Taylor, S. A., Altman, D. G., & Halligan, S. (2015). Assessment of the incremental benefit of computer-aided detection (CAD) for interpretation of CT colonography by experienced and inexperienced readers. *PLOS ONE, 10*(9), e013662. http://doi.org/10.1371/journal.pone.0136624

Bosmans, J. M. L., Weyler, J. J., De Schepper, A. M., & Parizel, P. M. (2011). The radiology report as seen by radiologists and referring clinicians: Results of the COVER and ROVER surveys. *Radiology, 259*(1), 184-195.

Bradley, E. H., Curry, L. A., & Devers, K. J. (2007). Qualitative data analysis for health services research: Developing taxonomy, themes, and theory. *Health Services Research, 42*(4), 1758-1772. http://doi.org/10.1111/j.1475-6773.2006.00684.x

Brady, A., Laoide, R. O., McCarthy, P., & McDermott, R. (2012). Discrepancy and error in radiology: Concepts, causes and consequences. *Ulster Medical Journal, 81*(1), 3-9.

Braun, V., & Clark, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3,* 77-101.

Breen, R. L. (2006). A practical guide to focus-group research. *Journal of Geography in Higher Education, 30*(3), 463-475. http://doi.org/10.1080/03098260600927575

Brink, J. A., Arenson, R. L., Grist, T. M., Lewin, J. S., & Ezmann, D. (2017). Bits and bytes: The future of radiology lies in informatics and information technology. *European Radiology, 27*(9), 3647-3651. doi:10.1007/s00330-016-4688-5

Brown, A. D. & Marotta, T. R. (2017). Using machine learning for sequence-level automated MRI protocol selection in neuroradiology. *Journal of the American Medical Informatics Association,* [Epub ahead of print]. doi:10.1093/jamia/ocx125

Brown, M. S. (2014). *Data mining for dummies.* Hoboken, NJ: John Wiley & Sons.

Bruno, M. A., Walker, E. A., & Abujudeh, H. H. (2015). Understanding and confronting our mistakes: The epidemiology of error in radiology and strategies for error reduction. *Radiographics, 35*, 1668-1676. http://doi.org/10.1148/rg.2015150023

Bucking, T. M., Hill, E. R., Robertson, J. L., Maneas, E., Plumb, A. A., & Nikitichev, D. I. (2017). From medical imaging data to 3D printed anatomical models. *PLoS One, 12*(5), e10178540. http://doi.org/10.1371.journal.pone.0178540

Bunnick, E. M., van Bodegom, L., Pinxten, W., de Beaufort, I. D., & Vernooji, M. W. (2017). Ethical framework for the detection, management and communication of incidental findings and imaging studies, building on an interview study of researchers' practices in perspectives. *Biomed Central Medical Ethics, 18*(1). http://doi.org/10.1186/s12910-017-0168-y

Burns, J. E., Yao, J., & Summers, R. M., (2017). Vertebral body compression fractures and bone density: Automated detection and classification on CT images. *Radiology*, *284*(3), 788-797. http://doi.org/10.1148.radiolo.2017162100

Busby, L. P., Coutier, J. L., & Glastonbury, C. M. (2018). Bias in radiology: The how and why of misses and misinterpretations. *Radiographics, 38*(1), 23-247. http://doi.org/10.1148/rg.2018170107

Cascianelli, S., Scialpi, M., Amici, S., Forini, N., Minestrini, M., Fravolini, M. L., . . . Palumbo, B. (2016). Role of artificial intelligence techniques (automatic classifiers) in molecular imaging modalities in neurodegenerative diseases. *Current Alzheimer's Research,* Epub ahead of print

Castaneda, C., Nalley, K., Mannion, C., Bhattacharyya, P., Blake, P., Pecora, A., . . . Suh, K. S. (2015). Clinical decision support systems for improving diagnostic accuracy and achieving precision medicine. *Journal of Clinical Bioinformatics, 5* (4), 1-16. http://doi.org/10.1186/s13336-015-0019-3

Castiglioni, I., & Gilardi, M. C. (2018). Radiomics: Is it time to compose the puzzle. *Clinical and Translational Imaging, 6,* 411-413. http://doi.org/10.1007/s40336-018-0302-4

Chartrand, G., Cheng, P. M., Vorontsov, E., Drozdzal, M., Turcotte, S., Pal, C. J., Kadoury, S., & Tang, (2017). Deep learning: A primer for radiologists. *Radiographics*, *37,* 2113-2131. http://doi.org/10.1148/rg.2017170077

Chen, J. Y., Sippel-Schmidt, T. M., Carr, C. D., & Kahn, C K. (2017). Enabling the next-generation radiology report: Description of two new system standards. *Radiographics*, *37*, 2106-2112.

Chen, P. H., Zafar, H., Galperin-Aizenberg, M., & Cook, T. (2017). Integrating natural language processing and machine learning algorithms to categorize oncologic responses in radiology reports. *Journal of Digital Imaging,* [Epub ahead of print]. http://doi.org/10.1007/s10278-017-0027-x

Chen, S. Liu, Z., Guan, M., Song, J., Lian, Y., Wang, S., . . . Tian, J. (2017). 2D and 3D CT radiomics features prognostic performance comparison in non-small cell lung cancer. *Translational Oncology*, *10*(6), 886-894. http://doi.org/10.1016/j.tranon.2017.08.007

Clinton, W. J. (2000). *President Clinton's address to Caltech on science and technology.* Retrieved from http://resolver.caltech.edu/CaltechCampusPubs:20140227-191407954

Collins, F. S., & Varmus, H. (2015). A new initiative on precision medicine. *New England Journal of Medicine, 372,* 793-795. http://doi.org/10.1056/NEJMp1500523

Connelly, L. M. (2016). Trustworthiness in qualitative research. *MedSurg Nursing, 25*(6), 435-436. Retrieved from https://search.proquest.com/openview/44ffecf38cc6b67451f32f6f96a40c78/1?pq-origsite=gscholar&cbl=30764

Cook, J. R., Siddique, M., Taylor, B. P., Yip, C., Chicklore, S., & Goh, V. (2014). Radiomics in PET; Principles and applications. *Clinical and Translational Imaging, 2,* 269-276. http://doi.org/10.1007/s40336-014-0064-0

Court, L. E., Fave, X., Mackin, D., Lee, J., Yang, J., & Zhang, L. (2016). Computational resources for radiomics. *Translational Cancer Research, 5*(4), 1-18. Retrieved from http://tcr.amergroups.com/issue/view/189

Cramer, J., Quickley, E., Hutchins, T. & Shah, L. (2017). Educational material for 3D visualization of spine procedures: Methods for creation and dissemination. *Journal of Digital Imaging, 30*, 296-300.

Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Croskerry, P. (2013). From mindless to mindful practice-cognitive bias and clinical decision-making. *New England Journal of Medicine, 368,* 2445-2448. http://doi.org/10.1056/NEJMp1303712

Croskerry, P. (2014). Bias: A normal operating characteristics of the diagnosing brain. *Diagnosis (Berlin, Germany), 1* (2), 23-27. doi:1515/dx-2103-0028

Croskerry, P., Petrie, D. A., Reilly, J. B., & Tait, G. (2014). Deciding about fast and slow decisions. *Academy of Medicine, Journal of the association of American Medical colleges, 39*(2), 197-200. doi:1097/ACM.0000000000000121

Croskerry, P., Singhal, G., & Mamede, S. (2013). Cognitive debiasing 1: Origins of bias and theory of debiasing. BMJ *Quality and Safety, 22,* ii58-ii64. doi:1136/bmjqs-20120- 001712

Croskerry, P., Singhal, G., & Mamede, S. (2013). Cognitive debiasing 2: Impediments to and strategies for change. *BMJ Quality and Safety, 22,* ii65-ii72. http://doi.org/10.1136/bmjqs-2012-001713

Daenzer, S., Freitag, S., von Sachsen, S., Steinke, H., Groll, M., Meixensberger, J., & Leimert, M. (2014). VolHOG: A volumetric object recognition approach based on bivariate histograms of oriented gradients for vertebra detection in cervical spine MRI. *Medical Physics, 41*(8), 082305. http://doi.org/10.1118/1.4890587

Daley, B. J. (2004). Using concept maps in qualitative research. In: Concept maps: Theory, methodology, technology. *Proceedings of the First International Conference on Concept Mapping.* Universidad Publica de Navarra, Pamplona, Spain, 191-199.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13(3)*, 319-340.

Davnall, F., Yip, C. S., Ljungqvist, G., Selmi, M., Ng, F., Sanghera, B., . . . Goh, V. (2012). Assessment of tumor heterogeneity: An emerging imaging tool for clinical practice? *Insights Imaging, 3*(6), 573-589. doi:10.1007/s13244-012-0196-6

Dhanoa, D., Dhesi, T. S., Burton, K. R., Nicolaou, S., & Liang, T. (2013). The evolving role of the radiologist: The Vancouver workload utilization evaluation study. *Journal of the American College of Radiology, 10*(10), 764-769. http://doi.org/10.1016/j.jacr.2013.04.001

DiPiro, P. J., Krajewski, K. M., Giardino, A., A., Braschi-Amirfarzan, M., & Ramaiya, N. H. (2017). Radiology consultation in the era of precision oncology: A review of consultation models and services in the tertiary setting. *Korean Journal of Radiology, 18*(1), 18-27. http://doi.org/10.3348/kjr.2017.18.1.18

Doi, K. (2007). Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. *Computerized Medical Imaging and Graphics, 31*(4-5), 198-211. http://doi.org/10.1016/j.compmedimag.2007.02.002

Donald, J. J., & Barnard, S. A. (2012). Common patterns in 558 diagnostic radiology errors. *Journal of Medical Imaging and Radiation Oncology, 56*(2), 173-178. http://doi.org/10.1111/j.1754-9485.2012.02348.x

Douglas, D. B., Boone, J. M., Petricoin, E., Liotta, L., & Wilson, E. (2016). Augmented realty imaging system: 3D viewing of a breast cancer. *Journal of Nature and Science, 2*(9), 1-6.

Douglas, D. B., Petricoin, E. F., Liotta, L., & Wilson, E. (2016). D3D augmented reality imaging system: Proof of concept in mammography. *Medical Devices: Evidence and Research, 9,* 277-283.

Doyle-Lindrud, S. (2015). Watson will see you now: A supercomputer to help clinicians make informed treatment decisions. *Clinical Journal of Oncology Nursing, 19*(1), 31-32. http://doi.org/10.1188/15.CJON.31-32

Drew, T., Vo., M. L. H., & Wolfe, J. M. (2013). The invisible gorilla strikes again: Sustained inattentional blindness in expert observers. *Psychology Science, 24*(9), 1848-1853. http://doi.org/10.1177/0956797613479386

Dreyer, K. J. & Gies, J. R. (2017). When machines think: Radiology's next frontier. *Radiology, 285*(3), 713-718.

Drukker, K., Antropova, N., Edwards, A., Papaionnou, J. & Giger, M. J. (2018). Most-enhancing tumor volume by MRI radiomics predicts recurrence-free survival "early on" in neoadjuvant treatment of breast cancer. *Cancer Imaging, 18*(12). http://doi.org/10.1186/s40644-018-0145-9

Dubin, R. (1969). *Theory building.* New York, NY: Free Press.

Durdella, N. (2019). *Qualitative dissertation methodology: A guide for research design and methods.* Los Angeles, CA: Sage.

Durrant, D. H., & True, J. M. (2012). *Myelopathy, radiculopathy, and peripheral entrapment syndrome.* Palm City, FL: Scholars Consortium

Echegaray, S., Nair, V., Kadoch, M., Leung, A., Rubin, D., Gevaert, O., & Napel, S. (2016). A rapid segmentation-insensitive "digital biopsy" method for radiomics feature extraction: Method and pilot study using CT images of non-small cell lung cancer. *Tomography*, *2*(4), 283-294. http://doi.org/10.18383/j.tom.2016.00163

El-Helo, S., Alomari, R. S., Ghosh, S., Chaudhary, V., Dhillon , G., Al-Zoubi, M. B., . . . Hamtini, T. M. (2012). Compression fracture diagnosis in lumbar: A clinical CAD system. *International Journal of Computer Assisted Radiology and Surgery, 8*(3), 461-469. http://doi.org/10.1007/s11548-012-0796-0

Elmore, J.G., Wells, C. K., Lee, C. H., Howard, D. H., & Feinstein, A. R. (1994). Variability in radiologists interpretation of mammograms. *New England Journal of Medicine, 331*, 1493-1499.

Erickson, B. J., Korfiatis, P., Akkus, Z., & Kline, T. L. (2017). Machine learning for medical imaging. *Radiographics*, *37*, 505-515.

European Society of Radiology (2015). Medical imaging in personalized medicine, A white paper of the research committee of the European Society of Radiology. *Insights Imaging, 6,* 141-155. http://doi.org/10.1007/s13244-015-0394-0

Farahani, N., Braun, A., Jutt, D., Huffman, T., Reder, N., Liu, Z., Yagi, Y., & Pantanowitz, L. (2017). Three-dimensional imaging and scanning: Current and future applications for pathology. *Journal of Pathology Informatics, 8,* 36. http://doi.org/10.4103/jpi_32_17

Farooki, S., Almeida, J., & Saltz, M. (2016). The impending “3D transformation” of radiology informatics. *Applied Radiology*, 6-8. Retrieved from www.applied radiology.com

Ferrucci, D. A. (2012). Introduction to “This is Watson.” *IBM Journal of Research and Development*, *56*(3.4), 1-15.

Fitzgerald R. (2005). Radiological error: Analysis, standard setting, targeted instruction and teamworking. *European Radiology, 15*(8), 1760-1767.

Fraenkel, J. R., & Wallen, N. E. (2003). *How to design and evaluate research in education.* (5th ed.). New York, NY: McGraw Hill.

French Radiology Community. (2018). Artificial intelligence and medical imaging 2018: French Radiology Community white paper. *Diagnostic and International Imaging, 99*, 727-742. http://doi.org/10.1016/j.diii.2018.10.003

Fusch, P. I & Ness, L. R. (2015). Are we there yet? Data saturation in qualitative research. The *Qualitative Report, 20*(9), 1408-1416. Retrieved from http://www.nova.edu/ssss/QR20/9/fusch1.pdf

Galbusera, F., Casaroli, G., & Bassani, T. (2018). Artificial intelligence and machine learning in spine research. *JOR Spine*, *2,* e1044, 1-20. http://doi.org/10.1002/jsp2.1044

Gao, X. W., Hui, R., & Tian, Z. (2017). Classification of CT brain images based on deep learning networks. *Computer Methods and Programs in Biomedicine*, *138,* 49-56. doi:10.1016/j.cmpb.2016.10.007.

Garland, L. H. (1949). On the scientific evaluation of diagnostic procedures. *Radiology*, *52*, 309-328.

Garland, L. H., (1959). Studies on the accuracy of diagnostic procedures. *American Journal of Radiology*, *82*, 25-38.

Ghasemi, M., Nabipour, I., Omrani, A., Alipour, Z., & Assadi, M. (2016). Precision medicine and molecular imaging: New targeted approaches toward cancer therapeutic and diagnosis. *American Journal of Medical and Molecular Imaging*, *6*(60), 310-327. Retrieved from www.ajnmmi.us/ISSN:2160-8407?ajnmmi0039837

Ghosh, S., & Chaudhary, X. (2014). Supervised methods for detection and segmentations of tissues in clinical lumbar MRI. *Computerized Medical Imaging and Graphics*, *38*(7), 639-649. http://doi.org/10.1016/j.compmedimag.2014.03.005

Giger, M. L. (2018). Machine learning in medical imaging. *Journal of the American College of Radiology*, *15*(3 PT B), 512-520. http://doi.org/10.1016/j.jacr.2017.12.028

Giger, M., Chan, H., & Boone, J. (2008). Anniversary paper: History and status of CAD and quantitative image analysis: The role of medical physicians and AAPM. *Medical Physics*, 35(12), 5799-5820. http://doi.org/10.1118/1.3013555

Gillies, R. J., Kinahan, P. E., & Hricak, H. (2016). Radiomics: Images are more than pictures, They are data. *Radiology*, *2789*(2), 563-577.

Glaser, J. & Laudel, G. (2013). Life with and without coding: Two methods for early-stage data analysis in qualitative research aiming at causal explanations. *Forum: Qualitative Social Research, 14*(2), 1-27.

Grabher, P., Mohammadi, S., Trachsler, A., Friedi, S., David, G., Sutter, R., …Freund, P.(2016). Voxel-based analysis of grey and white matter degeneration in cervical spondylotic myelopathy. *Scientific Reports, 6:*24636, 1- 10. http://doi.org/10.1038/srep24636

Greenhalgh, T., & Singlehurst, D. (2011). Studying technology use as social practice: The untapped potential of ethnography. *BMC Medicine, 9,* 45. http://doi.org/10.1186/1741-7015-9-45

Greenes, R. A. (2014). *Clinical decision support.* (2nd ed.). Boston, MA: Elsevier.

Grossmann, P., Stringfield, O., El-Hachem, N., Bui, M. M., Velazquez, E. R., Parmar, C., …Aerts, H., JWL. (2017). Defining the biological basis of radiomic phenotypes and lung cancer. *eLife, 6,* e23421. http://doi.org/10.7554/eLife.23421

Grunert, P., Hudson, K. D., Macielak, M. R., Aronowitz, E., Borde, E., Alimi, M., . . . Hartl, R. (2014). Assessment of intervertebral disc degeneration on quantitative MRI analysis: An in vivo study. *Spine, 39*(6), E369-e378. http://doi.org/10.1097/BRS.0000000000000194

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods, 18,* 59-82. http://doi.org/10.1177/1525822X05279903

Gupta, A., Banerjee, I., & Rubin, D. L. (2018). Automatic information extraction from unstructured mammography reports using distributed semantics. *Journal of Biomedical Information, 78,* 78-86. http://doi.org/10.1016/j.jbi.2017.12.016

Gupta, A., Snyder, A., Kachalia, A., Flanders, S., Saint, S. & Chopra, V. (2017). Malpractice claims related to diagnostic errors in the hospital. *BMJ Quality & Safety,* Epub ahead of print. http://doi.org/10.1136/bmjqs-2017-006774

Hainc, N., Federau, C., Stieltjes, B., Blatow, M., Bink, A., & Stippich, C. (2017). The bright, artificial intelligence-augmented future of neuroimaging reading. *Frontiers in Neurology, 8*(489), 1- 3. http://doi.org/10.3389/fneuro.2017.00489

Hamacher, A., Kim, S. J., Cho, S. T., Pardeshi, S., Hyun, S. H., Eun, S., & Whangbo, T. K. (2016). Application of virtual, augmented, and mixed reality to urology. *International Neurourology Journal, 20,* 172-181. http://doi.org/10.5213/inj.1632714.357

Hameed, M. A., & Arachchilage, N. A. G. (2016). A model for the adoption process of information system security innovations in organizations: A theoretical perspective. *Australian Conference on Information Systems,* 1-12.

Han, Z., Wei, B., Leung, S., Nachum, I. B., Laidley, D., & Li, S. (2018). Automated pathogenesis-based diagnosis of lumbar neural foraminal stenosis via deep multiscale multitask learning. *Neuroinformatics,* (Epub ahead of print]. http://doi.org/10.1007/s12021-018-9365-1

Hassanpour, S., & Langholtz, C. (2016). Information extraction from multi-institutional radiology reports. *Artificial Intelligence in Medicine, 66,* 29-39. http://doi.org/10.1016/j.artmed.2015.09.007

Hatt, M., Tixier, F., Visvikis, D., & Le Rest, C. C. (2017). Radiomics in PET/CT: More than meets the eye? *Journal of Nuclear Safety, Society of Nuclear Medicine, 58*(3), 365-366. http://doi.org/10.2967/jnumed.116.184655.hal-01525677

Herzog, R., Elgort, D. R., Flanders, A. E., & Moley, P. (2017). Variability in diagnostic error rates of 10 MRI centers performed lumbar spine MRI examinations on the same patient within a 3-week period. *The Spine Journal, 17,* 554-561. doi10.1017/j.spinee.2016.11.009

Hetherington, J., Lessoway, V., Gunka, V., Abolmaesumi, P., & Rohling, R. (2017). *International Journal of Computer Assisted Radiology and Surgery, 12*(7), 189-1198. http://doi.org/10.1007/s11548-017-1575-8

Hillman, B. J., & Goldsmith, J. C. (2011). *The sorcerer's apprentice.* New York, NY: Oxford University Press.

Holden, R. J., & Karsh, B. (2010). The technology acceptance model: Its past and its future in health care. *Journal of Biomedical Informatics, 43,* 159-172.

Hood, L., & Auffray, C. (2013). Participatory medicine: A driving force for revolutionizing health care. *Genome Medicine, 5,* 110. doi.org/10.1186/gm514

Hosney, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, J. W. L. (2018). Artificial intelligence in radiology. *Nature reviews. Cancer. 18*(8), 500-510. http://doi.org/10.1038/s41568-018-0016-5

Hoyt., R. H., Snider, D., Thompson, C., & Mantravadi, S., (2016). IBM Watson Analytics: Automating visualization, descriptive, and predictive statistics. *JMIR Public Health Surveillance, 2(*2), e157.

Hsu, W., Markey, M. K., & Wang, M. D. (2013). Biomedical imaging informatics in the era of precision medicine: Progress, challenges, and opportunities. *Journal of the American Informatics Association, 20,* 1010-1013. http://doi.org/10.1136/amiajnl-2013-002315

Hurwitz, E. L., Randhawa, K., Torres, P., Yu, H., Verville, L., Hartvigsen, J., . . . Haldeman, S. (2017). The global spine care initiative: A systematic review of individual and community-based burden of spinal disriders in rural populations in low and middle income communites. *European Spine Journal.* Epub ahead of print. http://doi.org/10.1007/s00586-017-5393-z

Hussain, A., & Oestreicher, J. (2017). Clinical decisoin-making: Heuritics and cognitive biases for the opthamologst. *Survey of Opthamology, 63,* 119-124. http://dx.doi.org//10.1016/j.survophthal.2017.08.007

Hussain, I., Schmidt, F. A., Kirnaz, S., Wipplinger, C., Schwartz, T. H., & Hartl, R. (2019). MIS approaches in the cervical spine. *Journal of Spine Surgery, 5* (suppl 1), S74-S83. http://doi.org/10.21037/jss.2019.04.21

Incoronato, M., Aiello, M., Infante, T., Cavaliere, C., Grimaldi, A. M., Mirabelli, P., . . . Monti, S. (2017). Radiogenomic analysis of oncological data: A technical Survey. *International Journal of Molecular Sciences, 18,* 805. http://doi.org/10.3390/ijms18040805

Institute of Medicine. (2000). *To err is human: Building a safer health system.* Washington, DC: National Academy Press.

Itri, J. N., & Patel, S. H. (2018). Heuristics and cognitive error in medical imaging. *American Journal of Roentgenology, 12,* 1-9. http://doi.org/10.2214?AJR.17.18907

Janesick, V. J. (2011). *Stretching: Exercises for qualitative researchers* (3rd ed.). Thousand Oaks, CA: Sage.

Janjul K. J., Sugrue, M., & Deane, S.A. (1998). Prospective evaluation of early missed injuries and the role of tertiary trauma survey. *Journal of Trauma, 44,* 1000-1007.

Jeanquarter, F., Jan-Quartier, C., Cemernek, D., & Holzinger, A. (2016). In silico modeling for tumor growth visualization. *Biomed Central Systems Biology, 10*(59), 1-15. http://doi.org/10.1186/s12918-016-0318-8

Jha, S., & Topol, E. J. (2016). Adapting to artificial intelligence: Radiologists and pathologists as information specialists. *Journal of the American Medical Association, 316*(22), 2353-2354. http://doi.org/10.1001/jama.2016.17438

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., . . . Wang, Y. (2017). Artificial intelligence in health care: Past, present and future. *Stoke and Vascular Neurology, 0,* e000101. http://doi.org/10.1136/svn-2017-000101

Kahn, C. E. (2017). From images to actions: Opportunities for artificial intelligence in radiology. *Radiology, 285*(3), 719-720.

Kaiser. (2009). Protecting respondent confidentiality in qualitative research. *Qualitative Health Research, 19*(1), 1632-1641. http://doi.org/10.1177/1049732309350897

Kakhani, P., Prater, A. B., Hutson, R. K., Andriole, K. P., Dreyer, K. J., Morey, J….Hawkins, C. M. (2017). Machine learning in radiology: Applications beyond image interpretation. *Journal of the American College of Radiology, 17,* 31287-5. http://doi.org/10.1016/j.jacr.2017.09.044

Kansagra, A. P., Yu, J. P., Chatterjee, A. R., Lenchik, L., Chow, D.S., Prater, A. B….Ali, S. (2016). Big data and the future of radiology informatics. Academic Radiology, 23(1), 30-42. http://doi.org/10.1016/j.acra.2015.10.004

Kelly, J. E., & Hamm, S. (2013). *Smart machines: IBM's Watson and the era of cognitive computing.* New York, NY: Columbus Business School.

Khan, A., & Woosley, J. M. (2011). Comparison of contemporary technology acceptances models and evaluation of the best fit for health industry organizations. *International Journal of Computer Science & Engineering Technology,1*(11), 709-717.

Kharat, A. T., & Singhal, S. (2017). A peek into the future of radiology using big data. *Indian Journal of Radiological Imaging, 27*(2), 241-248. http://doi.org/10.4103/ijri.IJRI_493_16

Kim, N., Choi, J., Yi, J., Choi, S., Park, S., Yi, J., . . . Seo, J. B. (2013). An engineering view on megatrends in radiology: Digitization to quantitative tools of medicine. *Korean Journal of Radiology, 14*(2), 139-153. http://doi.org/10.3348/kjr.2013.14.2.139

Kim, N., Rowe, B. H., Raymond, G., Jen, H., Colman, I., Jackson, S. A., . . . Majumdar, S. R. (2004). Underreporting of vertebral compression fractures on routine chest radiography. *American Journal of Radiology, 182*, 297-300. doi:10.2214/AJR.182.2.1820297

Kim Y. W., & Mansfield, L. T. (2014). Fool me twice: Delayed diagnoses in radiology with emphasis on perpetuated errors. *American Journal of Roentgenology, 202*(3), 203-210.

Kitzinger, J. (1995). Introducing focus groups. *British Journal of Medicine, 311*, 299-302.

Kohn, M. S., Sun, J., Knoop, S., Shabo, A., Carmeli, B., Sow, B., . . . Rapp, W. (2014). IBM's health analytics and clinical decision support. *IMIA Yearbook of Medical Informatics*, 154-162. http://doi.org/10.15265/14-2014-0002

Kondo, K. J. (2008). Osteoporotic vertebral compression fractures and vertebral augmentation. *Seminars in Interventional Radiology, 25,* 413-424. http://doi.org/10.1005/s-0028-1103000

Kressel, H. Y. (2017). Setting sail. *Radiology, 282,* 4-6. http://doi.org/10.1148/radiol.2016162471

Krippendorf, K. (2004). *Content analysis: An Introduction to its methodology* (2nd ed.). Thousand Oaks, CA: Sage.

Krueger, R. A., & Casey, M. A. (2015). *Focus groups: A practical guide for applied research.* (5th ed.). Los Angeles, CA: Sage.

Kumar, D., Chung, A. G., Shaifee, M. J., Khalvati, F., Haider, M. A., & Wong, A. (2017). Discovery radiomics for pathologically proven computed tomography lung cancer prediction. *Computer Vision and Pattern Recognition,* 1-8. Retrieved from https://arxiv.org/abs/1509.00117

Kumar, V., Gu, Y., Basu, S., Berglund, A., Eschrich, S. A., Schabath, M. B….Gillies, R. J. (2012). QIN "Radiomics: The process and the challenges." *Magnetic Resonance Imaging, 30*(9), 1234-1248. http://doi.org/10.1016/j.mri.2012.06.010

Kuo, T. T., Kim, H. E., & Ohno-Machado, L. (2017). Blockchain distributed ledger technologies for biomedical and health care applications. *Journal of the American Medical Informatics Association, 24*(60), 1211-1220.

Lai-Kwon, J., Siva, S. S., & Lewin, J. (2018). Assessing the clinical utility of computed tomography-based radiomics. *The Oncologist, 23,* 103. http://doi.org/10.1634/theoncologist.2018-0193

Lakhani, P., Prater, A. B., Hutson, R. K., Andriole, K. P., Dreyer, K. J., Morey, J.…Hawkins, C. W. (2017). Machine learning in radiology: Application beyond image interpretation. *Journal of the American College of Radiology, 15*, 350-359. http://doi.org/10.1016/j.jacr.2017.09.044

Lakhani, P., & Sundaram, B. (2017). Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology, 284*(2), 574-582.

Lambin, P., Leijenaar, R. T. H., Deist, T. M., Peerlings, J., de Jong, E. E. C., van Timmeren, J., . . . Walsh, S. (2017). Radiomics: The bridge between medical imaging and personalized medicine. *Nature Reviews. Clinical Oncology*. Epub ahead of print. doi:101038/nrclinonc.2017.141

Lambin, P., Rios-Valezquez, E., Leijenaar, R., Carvalho, S., van Stiphout, R. G., Granton, P., . . . Aerts, H. (2012). Radiomics: Extracting more information from medical images using advanced feature analysis. *European Journal of Cancer, 48*(4), 441-446.

Langs, G., Rohrich, S., Hofmanninger, J., Prayer, F., Pan, J. Herold, C. & Prosch, H. (2018). Machine learning: From radiomics to discovery and routine. *Radiologe*, 1-6. http://doi.org/10.1007/s00117-018-0407-3

Larue, R, T. H. M., Defraene, G., De Ruysscher, D., Lambin, P., & Van Elmpt, W. (2016). Quantitative radiomics studies for tissue characterization: A review of technology and methodological practices. *British Journal of Radiology, 90*, 20160665.

Latts, L. (2016). The age of big data and the power of Watson. IBM. Retrieved from http://www.ema.europa.eu/docs/en_GB/document_library/Presentation/2017/01/WC500219353.pdf

Lee, C. I., Khodyakov, D., Weidmer, B. A., Wenger, N. S., Timbie, J. W., Brantley, B. . . . Kahn, K. L. (2015). Radiologist's perception of computerized decision support: A focus group study from the Medicare imaging demonstration project. *American Journal of Roentgenology, 205*, 947-955. http://doi.org/10.2214/AJR.15.14801

Lee, C. S., Nagy, P. G., Weaver, S. J., Newman-Toker, D. E. (2012). Cognitive and system factors contributing to diagnostic errors in radiology. *American Journal of Roentgenology, 201*, 611-617. http://doi.org/10.2214/AJR.12.10375

Lee, D. Y., & Li, K. C. P. (2009). Systems diagnostics: the systems approach to molecular imaging. *American Journal of Roentgenology, 193,* 287-294. http://doi.org/10.2214/AJR.09.2866

Lee, G., Lee, H. Y., Ko, E. S., & Jeong, W. K. (2017). Radiomics and imaging genomics in precision medicine. *Precision and Future Medicine, 1*(1), 10-31. https://doi.org/10.23838/pfm.2017.00101

Lee, J., Jun, S., Cho, Y. Lee, H., Kim , G., Seo, J. B., & Kim, N. (2017). Deep learning in medical imaging: General overview. *Korean Journal of Radiology, 18*(4), 570-584. http://doi.org/10.3348/kjr.2017.18.4.570

Lee, M. H., Schemmel, A. J., Pooler, B. D., Hanley, T., Kennedy, T. A., Field, A. S. …Yu, J. J. (2017). Workflow dynamics and the imaging value chain: Quantifying the effect of designating a nonimage-interpretive task workflow. *Current Problems in Diagnostic Radiology*, *46*(4), 275-281.

Lee, V. S. (2017). Annual oration: Driving value through imaging. *Radiology, 285*(1), 3-15.

Lee, Y., Hsieh, Y., & Hsu, C. (2011). Adding innovation diffusion theory to the technology acceptance model: Supporting employees' intentions to use e-learning systems. *Education Technology & Society, 14*(4), 124-137.

Lehr, J. L., Lodwick, G. S., Farrell, C., Braaten, O., Vitama, P., Kolvisto, E. L. (1976). Direct measurement of the effect of film miniaturization on diagnostic accuracy. *Radiology, 118*, 257-263.

Legris, Ingham, & Colertte, (2003). Why do people use information technology? A critical review of the technology acceptance model. *Information and Management, 41*, 795-804.

Lesko, L. J., Zineh, I., & Huang, S. M. (2010). What is clinical utility and why should we care? Clinical Pharmacology & Therapeutics, 88(6), 728-733.

Li, H., Zhu, Y., Burnside, E. S., Huang, E., Drukker, K., Hoadley, K. A., . . . Giger, M. L. (2016). Quantitative MRI radiomics in the prediction of molecular classifications of breast cancer subtypes in the TCGA/TCIA data set. *Nature Partner Journals; Breast Cancer, 2,* 16012. http://doi.org/10.1038/npjbcancer.2016.12

Liew, C. (2018). The future of radiology augmented with artificial intelligence: A strategy of success. *European Journal of Radiology, 102,* 152-156. http://doi.org/10.1016/j.ejrad.2018.03.019

Limkin, E. J., Sun, R., Dercle, L., Zacharaki, E. I., Borbert, C., Reuze, S., . . . Ferte, C. (2017). Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. *Annuals of Oncology, 28*(6), 1191-1206.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry.* Newbury Park, CA: Sage.

Loney, P. L., & Stratford, P. W. (1999). The prevalence of low back pain in adults: A methodical review of the literature. *Journal of the American Physician Therapy Association*, *79*(4), 384-396

Long, S. S., Yablon, C. M., & Eisenberg, R. L. (2010). Bone marrow signal alteration in the spine and sacrum. *American Journal of Roentgenology, 195,* W178-W200. http://doi.org/10.2214/AJR.o9.4134

Lorell, B. H., Mikita, J. S., Anderson, A., Hallinan, Z. P., & Forrest, A. (2015). Informed consent in clinical research: Consensus recommendations for reform identified by an expert interview panel. *Clinical Trials 12*(6), 692-695. http://doi.org/10.1177/174077451594362

Louis, D. N., Feldman, M., Carter, A. B., Dighe, A. S., Pfiefer, J. D., . . . Becich, M. J. (2016). Computational pathology: A path ahead. *Archives of Pathology & Laboratory Medicine,* 140(1), 41-50. http://doi.org/10.5858/arpa.2015-0093-SA

Louis, D. N., Gerber, G. K., Baron, J. M., Bry, L., Dighe, A. S., Getz, G., . . . Golden, J. A. (2014). Computational pathology: An emerging definition. *Archives of Pathology & Laboratory Medicine*, 138, 1133-1138. http://doi.org/10.5858.arpa.2014-0034-ED

Lubner, M. G., Smith, A. D., Sandrasegaran, K., Sahani, D. V., & Pickhardt, P. J. (2017). CT texture analysis: Definitions, applications, biologic correlates, and challenges. *Radiographics, 37,* 1483-1503. http://doi.org/10.1148/rg.2017170056

Lundstrom, C., Gilmore, H. L., & Ros, P. R. (2017). Integrated diagnostics: The computational revolution catalyzing cross-disciplinary practices in radiology, pathology and genomics. *Radiology, 285,* 12-15. http://doi.org/10.1148/radiolo.2017170062

Maccauro, G., Spinelli, M. S., Mauro, S., Perisano, C., Graci, C., & Rosa, M. A. (2011). Physiopathology of spine metastasis. *International Journal of Surgical Oncology*, 1-8, 107969. http://doi.org/10.1155/2011/107969

Madani, A., Arnaout, R., Mofrad, M., & Arnout, R. (2017). Fast and accurate view classifications of echocardiograms using deep learning. *Digital Medicine, 1*, 6. http://doi.org/10.1038/s41746-017-0013-1

Maier, O., Schroder, C., Forkert, N. D., Martinez, T., & Handles, H. (2015). Classifiers for ischemic stroke lesion segmentation: A comparison study. *Plos One, 10*(12), e0145118. http://doi.org/10.1371/jurnal.pone,0145118

Makary, M. A. & Daniel, M. (2016). Medical error-the third leading cause of death in the US. *British Medical Journal, 353,* i2139. http://doi.org/10.1136/bmj.i2139

Malone, D., & Newton, I. G. (2018). Molecular imaging: The convergence of form and function. *Applied Radiology,* 47(5), 14-24.

Mamoshina, P., Ojomoko, L., Yanovich, Y., Ostrovski, A., Botezatu, A., Prikhodko, P….Zhavoronkov. (2018). Converging blockchain and next-generation artificial intelligence technologies to decentralize and accelerate biomedical research and health care. *Oncotarget, 9*(5), 5665-5690.

Mancini, M., Summers, P., Faita, F., Brunetto, M. R., Callea, F., De Nicola, A….Bonino, F. (2018). Digital liver biopsy: Bioimaging of fatty liver for translational and clinical research. *World Journal of Hepatology, 10*(2), 231-24. http://doi.org/10.42534/wjh.v10.i2.231

Manrai, A. K., Bhatia, G., Strymish, J., Kohane, I. S., & Jain, S. H. (2014). Medicine's uncomfortable relationship with math: Calculating positive predictive value. *JAMA Internal Medicine, 174*(6), 991-993. http://doi.org/10.10001/jamainternmed.2014.1059

Mapoka, K., Maseby, H., & Zuva, T. (2013). Mathematical models and algorithm challenges. *International Journal of Control theory and Computer Modeling, 3(*6), 21-28. http://doi.org/10.5121.ijctcm.2013.3602

Marshall, C., & Rossman, G. B. (2011). *Designing qualitative research* (5th ed.). Thousand Oaks, CA: Sage.

Massat, M. B. (2018). Artificial intelligence in radiology: Hype or hope? *Applied Radiology, 3,* 22-26.

Mayo, C. S., Kessler M., L., Eisbruch, A., Weyburne, G., Feng, M. Hayman, J. A., . . . Ten Hanken, R. K. (2016). The big data effort in radiation oncology: Data mining or data farming. *Advances in Radiation Oncology, 1,* 260-271.

Mays, N., & Pope, C. (1995). Rigour and qualitative Research. *British Medical Journal, 311*, 109-112.

Mays, N., & Pope, C. (2000). Assessing quality in qualitative research. *British Journal of Medicine, 320*, 50-52.

Mays, N., & Pope, C. (2000). Qualitative research in health care: Assessing quality in qualitative research. *British Medical Journal, 320*(72260), 50-52.

McCue, M. E., & McCoy, A. M. (2017). The scope of big data in one medicine: Unprecedented opportunities and challenges. *Frontiers in Veterinary Science, 4*(194), 1-23. http://doi.org/10.3389/fvets.2017.00194

McGarry, S. D., Bukowy, J. D., Iczkowski, K. A., Unteriner, J. G., Duvnjak, P., Lowman, A. K., . . . Laviolette, P. S. (2019). Gleason probability maps: A radiomics tool for mapping prostate cancer likelihood in MRI space. *Tomography*, *5*(1), 127-134. http://doi.org/10.18383/j.tom.2018.00033

Memmert, D. (2006). The effects of eye movement, age, expertise on inattentional blindness. *Conscious Cognition*, *15*(3), 620-627.

Mesko, B. (2017). The role of artificial intelligence in precision medicine. Expert Review of *Precision Medicine and Drug Development*, *2*(5), 239-241. http://doi.org/10.1080/23808993.2017.1380516.

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data Analysis: A methods sourcebook.* (3rd ed.). Los Angeles, CA; Sage.

Miller, J. G. (1978). *Living systems.* New York: McGraw-Hill.

Mirzaalian, H., Wels, M., Heimann, T., Kelm, B. M., & Suehling, M. (2013). Fast and robust 3D vertebra segmentation using statistical shape models. *Conference Proceedings of IEEE Engineering in Medicine and Biology Society*, *2013*, 3379-3382. http://doi.org/10.1109/EMBC.2013.6610266

Mittelstadt, B. D., & Floridi, L. (2016). The ethics of big data: Current and foreseeable issues in biomedical contexts. *Science and Engineering Ethics*, *22*(2), 303-341. http://doi.org/10.1007/s11948-015-9652-2

Mohebian, M. R., Marateb, H. R., Mansourian, M., Mananas, M. A., & Mokarian, F. (2017). A hybrid computer-aided-diagnosis system for prediction of breast cancer recurrence (HPBCR) using optimized ensemble learning. *Computational and Structural Biotechnology Journal*, *15*, 75-85.

Morgan, M. B., Branstetter, B. F., Mates, J., & Chang, P. J. (2006). Flying blind: Using a digital dashboard to navigate a complex PACS environment. *Journal of Digital Imaging, 19*(1), 69-75.

Murdoch, T. B., & Detsky, A. S. (2013). The inevitable application of big data to health care. *Journal of the American Medical Association, 309*(13), 1351-1352. http://doi.org/10.1001/jama.2013.393

Naganawa, S., Enooku, K., Tateisha, R., Akai, H., Yasaka, K., Shibahara, J., . . . Kiryu, S. (2017). Imaging prediction of nonalcoholic steatohepatitis using computed tomography texture analysis. *European Radiology, [*Epub ahead of print]. http://doi.org/10.1007/s00330-017-5270-5

Nandu, H., Wen, P. Y., & Huang, R. Y. (2018). Imaging in neuro-oncology. *Therapeutic advances in neurological disorders. 11*, 1-19. http://doi.org/10.1177/1756286418759865

Napal, S., & Giger, M. (2018). Radiomics and imaging genomics: Quantitative imaging for precision medicine. *Journal of Medical Imaging, 2*(4), 1-3.

Naselaris, T., Kay, K. N., Nishimoto, S. & Gallant. (2011). Encoding and decoding in fMRI. *NeuroImage, 56,* 400-411. http://doi.org/10.1016/j.neuroimage.2010.07.073

National Science and Technology Council (2016). The National Artificial Intelligence Research and Development strategic plan. 1-48. https://www.nitrd.gov/PUBS/national_ai_rd_strategic_plan.pdf

Nendaz, M., & Perrier, A. (2012). Diagnostic errors and flaws in clinical reasoning: Mechanisms and prevention in practice. *Swiss Medical Weekly, 142,* w13706. http://doi.org/10.4414/smw.2012.13706

Ng, F., Kozarski, R., Ganeshan, B., & Goh, V. (2013). Assessment of tumor heterogeneity by CT texture analysis: Can the largest cross-sectional area be used as an alternative to whole tumor analysis? *European Journal of Radiology, 82,* 342-348. http://doi.org/10.1016/j.ejrad.2012.10.023

Njie, B., & Asimiran, S. (2014). Case study as a choice in qualitative methodology. *ISOR Journal of Research & Method in Education, 4*(3), 35-40.

Nouri, A., Tetreault, L., Singh, A., Karadimas, S., & Fehlings, M. (2015). Degenerative cervical myelopathy: Epidemiology, genetics, and pathogenesis. *Spine, 40*(12), E675-E693.

Novak, J. (1998). *Learning, creating and using knowledge: Concept Maps as facilitative tools in schools and corporations.* Mahwah, NJ: Erlbaum Associates.

Nowell, L., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods, 16,* 1-13. http://doi.org/10.1177/1609406917733847

Oakden-Rayner, L., Carneiro, G., Bessen, T., Nascimento, J. C., Bradley, A. P., & Palmer, L. J. (2017). Precision radiology; Predicting longevity using feature engineering and deep learning method in a radiomics framework. *Scientific Reports,* 1648. http://doi.org/10.1038/s41598-017-01931

Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the future: Big data, machine learning, and clinical medicine. *New England Journal of Medicine, 375*(13), 1216-1219. http://doi.org/10.1056/MEJMp1606181

Oktay, A. B., Albayrak, N. B., & Akgul, Y. S. (2014). Computer aided diagnosis of degenerative intervertebral disc diseases from lumbar images. *Computerized Medical Imaging and Graphics, 38*(7), 613-619. http://doi.org/10.1016/j.compmedimag.2014.04.006

Orb, A., Eisenhauers, L., Wynaden, D. (2001). Ethics in qualitative research. *Journal of Nursing Scholarship, 33*(1), 93-96.

Ortiz-Ramon, R., Larroza, A., Arana, E., & Moratal, D. (2017). A radiomics evaluation of 2D and 3D MRI texture features to classify brain metastasis form lung cancer and melanoma. *Conference Proceedings of the IEEE Engineering and Biology Society*, 493-496. http://doi.org/10.1109/EMBC.2017.8036869

Pannu, A. (2015). Artificial intelligence and its application in different areas. *International Journal of Engineering and Innovation Technology, 4*(10), 1-6.

Panth, K. M., Leijenar, R. T. H., Carvalho, S., Lieuwes, N. G., Yaromina, A., Bubois, L., & Lambin, P. (2015). Is there a causal relationship between genetics changes and radiomics-based image features? An in vivo preclinical experiment with doxycycline inducible GADD34 tumor cells. *Radiotherapy and Oncology, 116*, 4620466.

Papanicolas, I., Woskie, L. R., & Jha, A. K. (2018). Health care spending in the United States and other high-income countries. *Journal of the American Medical Association, 319*(10), 1024-1039. http://doi.org/10.1001/jama.2018.1150

Parasuraman, A., & Colby, L. C. (2001). *Techno-ready marketing.* New York, NY: The Free Press.

Parekh, V., & Jacobs, M. A. (2016). Radiomics: A new application from established techniques. *Expert Review of Precision Medicine and Drug Delivery, 1*(2), 207-226. http://doi.org/10.1080/23808993.2016.1164013

Parekh, V., & Jacobs, M. A. (2019). Deep learning and radiomics in precision medicine. *Expert Review of Precision Medicine and Drug Development, 4*(2), 59-72. http://doi.org/10.1080/23808993.2019.1585805

Park, S. H., Eber, M.R., Widner, D. B., & Shiozawa, Y. (2018). Role of bone marrow microenvironment in the development of painful complications of skeletal metastasis. *Cancers, 10*(5), pii E141. doi:103390/cancers/0050141.review

Parmar, C., Grossmann, P., Bussink, J., Lambin, P., & Aerts, H. J. (2015). Machine learning methods for quantitative radiomic biomarkers. *Scientific Reports, 5,* 13078. http://doi.org/10.1038/srep13087

Patriarche, J. W., & Erickson, J. (2007). Part 1. Automated change detection and characterization and serial MRI studies of brain-tumor patients. *Journal of Digital Imaging, 20*(3), 203-222.

Patriarche, J. W., & Erickson, J. (2007). Part 2. Automated change detection and characterization applied to serial MRI of brain tumors may detect progression earlier than human experts. *Journal of Digital Imaging 20*(3), 203-222.

Patton, M. Q. (1990). *Qualitative research & evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Patton, M. Q. (1999). Enhancing the quality and credibility of qualitative analysis. *Health Services Research, 34,* 1189-2018

Peeken, J. C., Bernhofer, M., Wiestler, B., Goldberg, T. Cremers, D., Rost, B., . . . Nusslin, F. (2018). Physical Medicine, 48, 27-36. Radiomics in radiooncology: Challenging the medical physicist. *Physics in medicine, 48*, 27-36. http://doi.org/10.1016/j.ejmp.2018.03.012

Pena, J. A., Thomsen, F., Damm, T., Campbell, G. M., Bastgen, J., Barkman, R., & Gluer, C. C. (2016). Bone-marrow densitometry: Assessment of marrow space of human vertebrae by single energy high resolution-quantitative computed tomography. *Medical Physics, 43*(7), 4174. http://doi.org/10.1118/1.4950874

Phua, D. H., & Tan, N. C. K. (2013). Cognitive aspect of diagnostic errors. Annuals of the *Academy of Medicine, Singapore, 42*, 33-41.

Pinker, K., Shitano, F., Sala, E., Do, R. K., Young, R. J., Wibner, A. G., . . . Morris, E. A. (2017). Background, current role, and potential applications of radiogenomics. *Journal of Magnetic Resonance Imaging.* Epub ahead of print. http://doi.org/10.1002/jmri.25870

Pinto, A., & Brunese, L. (2010). Spectrum of diagnostic errors in radiology. *World Journal of Radiology, 2*(1), 377-383. http://doi.org/10.4329/wjr.v2.i10.377

Ponelis, S. R. (2015). Using interpretive qualitative case studies for exploratory research in doctoral studies: A case of information systems research in small and medium enterprises. *International Journal of Doctoral Studies, 10,* 535-550. http://ijds.org/Volume10/IJDSv10p535-550Ponelis0624.pdf

Pons, E., Braun, L. M., Hunink, G., & Kors, J. A. (2016). Natural language processing in radiology: A systematic review. *Radiology, 279*(2), 329-343. http://doi.org/10.1148/radiol.16142770

Quest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough. An experiment with data saturation and variability. *Field Methods, 18*(1), 59-82.

Quer, G., Muse, E. D., Nikzad, N., Topol, E. J., & Steinhubl, S. (2017). Augmenting diagnostic vision with AI. *The Lancet, 390*(10091), 221. http://doi.org/10.1016/S0140-6736(17)31764-6

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in health care: Promise and potential. *Health Information Science and Systems, 2*(3), 1-10. http://doi.org/10.1186.2047-2501-2-3

Rancher, C. E., Shoemaker, J. M., Petree, L. E., Holdsworth, M., Phillips, J. P., & Helitzer, D. L. (2016). Disclosing neuroimaging incidental findings: A qualitative thematic analysis of health literacy challenges. *BMC Medical Ethics, 17*(58). doi.101186/s12910-016-0141-1

Ranschaert, E. R. (2016). The impact of information technology on radiology services: An overview. *Journal of the Belgian Society of Radiology, 100*(1), 1-6. http://doi.org/10.5334/jbr-btr.1184

Rastegar-Mojarad, M., Sohn, S., wang, L., Shen, F., Bleeker, T. C., Cliby, W. A., & Liu, H. (2017). Need of informatics in designing interoperable clinical registries. *International Journal of Medical Informatics, 108,* 7884. http://doi.org/10.1016/j.ijmedinf.2017.10.004

Recht, M., & Bryan, N. (2017). Artificial intelligence: Threat or boon to radiologists. *Journal of the American College of Radiology, 14,* 1476-1480.

Renfrew, D. L., Franken, E. A., Berbaum, K. S., Weigelt, F. H., & Abu-Yousef, M. M. (1992). Error in radiology: Classification and lessons in 182 cases presented at problem case conference. *Radiology, 183,* 145-150.

Rizzo, S., Botta, F., Raimondi, S., Origgi, D., Fanciullo, C., Morganti, A. G., & Bellomi, M. (2018). Radiomics: The facts and challenges of imaging analysis. *European Radiology Experimental, 2*(36), 1-8. doi.org/:10.1186/s41747-018-0068-z

Rogers, E. M. (1995). *Diffusion of innovations.* (4th ed.) New York, NY: Free Press.

Rogers, E. M. (2003). *Diffusion of innovations.* (5th ed.) New York, NY: Free Press.

Rosenkrantz, A. B., Mendiratta-Lala, M., Bartholmai, B. J., Ganeshan, D., Abrahamson, R. G., Burton, K. R., Lenchik, L. (2015). Clinical utility of quantitative imaging. *Academic Radiology, 22*(1), 33-49. http://doi.org/10.1016/j.acra.2014.08.011

Rubin, D. L. (2019). Artificial intelligence in imaging: The radiologists role. *Journal of the American College of Radiology*. *16*(9 Pt B), 1309-1317. http://doi.org/10.1016/j.jacr.2019.05.036

Rubin, D. L., & Kahn, C. E. (2017). Common data elements in radiology. *Radiology, 283*(3), 837-844.

Rubin, H. J. & Rubin, I. S. (1995). *Qualitative interviewing: The art of hearing data.* Thousand Oaks, CA: Sage

Rudie, J. D., Rauschecker, A. M., Bryan, R. N., Davatzikos, C., & Mohan, S. (2019). Emerging applications of artificial intelligence in neuro-oncology. *Radiology, 290*(3), 607-618. http://doi.org/10.1148/radiol.2018181928

Ryan, G. W., & Bernard, H. R. (2003). Techniques to identify themes. *Field Methods, 15*(1), 85-109.

Saber-Tehrani, A. S., Lee, H., Mathews, S. C., Shore, A., Makary, M. A., Pronovost, P. J., & Newman-Toker, D.E. (2013). 25-Year summary of US malpractice claims for diagnostic errors 1986-2010: An analysis from the national practitioner database. *BMJ Quality & Safety*, *22*(8), 672-680. http://doi.org/10.1136/bmjqs-2012-001550

Sala, E., Mema, E., Himoto, Y., Veeraraghavan, H., Brenton, J. D., Snyder, A., . . . Vargas, H. A. (2017). Unraveling tumour heterogeneity using next generation imaging: Radiomics, radiogenomics, and habitat imaging. *Clinical Radiology, 72*(1), 3-10. https://doi.org/10.1016/j.rad/2016.09.013

Saldana, J. (2016). *The coding manual for qualitative researchers.* Los Angeles, CA: Sage.

Sanberg, J. C., Nguyen, H. T., Arcury, T. A. Johnson, A. J., Hwang, W., Gage, H.D., . . . Carr, J. J. (2012). Insight into the sharing of medical images. *Applied Clinical Informatics, 3,* 475-487. http://doi.org/10.4338/ACI-2012-06-RA-0022

Sanduleanu, S., Woodruff, H. C., de Jong, E. E. C., van Timmeren, J. E., Jochems, A., Bubois, L., & Lambin, P. (2018). Tracking tumor biology with radiomics: A systematic review utilizing a radiomics quality score. *Radiotherapy and Oncology,* 127(3), 349-360. http://doi.org/10.1016/j.radonc.2018.03.033

Sardanelli, F. (2017). Trends in radiology and experimental research. *European Radiology Experimental, 1*(1), http://doi.org/10.1186/s41747-017-0006-5

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J. Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: exploring its conceptualization and operationalization. *Quality & Quantity, 52*, 1893-1907. http://doi.org/10.1007/s11135-017-0574-8

Scheckenbach, K. (2018). Radiomics: Big data Instead of biopsies in the future? *Laryingorhinootologie, 97,* S129-S141. http://doi.org/10.1055/s-0043-1211964

Schemmel, A., Lee, M., Hanley, T., Pooler, D., Kennedy, T. Field, A., . . . Yu, J. J. (2016). Radiology workflow disruptors: A detailed analysis. *Journal of the American College of Radiology, 13*(10), 1210-1214.

Scherer, M., Cordes, J., Younsi, A., Sahin, Y., Gotz, M., Mohlenbruch, M….Orakcioglu, B. (2016). Development and validation of an automated segmentation algorithm for quantification of brain hemorrhages. *Stroke*, *47*, 2776-2782. http://doi.org/10.1161/strokeaha.116.013779.)

Schwartz, L. H., Panicek, D. M., Berk, A. R., Yuelin, L., & Hricak, H. (2011). Improving communication of diagnostic radiology findings through structured reporting. *Radiology, 260*(1), 174-181.

Shafiee, M. J., Chung, A. G., Khalvati, F., Haider, M. A., & Wong, A. (2017). Discovery radiomics via evolutionary deep radiomic sequencer discovery for pathologically proven lung cancer detection. *Journal of Medical Imaging 4*(4), 041305. http://doi.org/10.1117/I.JMI.4.4.041305

Shaikh, F. A., Kolowitz, B. J., Awan, O., Aerts, H. J., von Reden, A., Halabi, S., . . . Deible, C. (2017). Technical challenges in the clinical application of radiomics. *JCO Clinical Cancer Informatics,* 1-8.Sharma, A., Wang, K., & Siegal, E. (2017). Radiologist digital workspace use and preference: A survey-based study. *Journal of Digital Imaging, 30,* 687-694.

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research project. *Education for Information,* 63-75.

Shortlife, E. H., & Cimino, J. J. (2006). Biomedical Informatics: Computer Applications in Health care and Biomedicine (3rd ed,), New York, NY: Springer.

Sokolovskaya, E., Shinde, T., Ruchman, R. B., Kwak, A. J., Lu, S., Shariff, Y. K., . . . Talangbayan, L. (2015). The effect of faster reporting speed for imaging studies on the number of misses and interpretation error: A pilot study. *Journal of the*

American College of Radiology, *12*(7), 683-688. http://doi.org/10.1016/j.jacr.2015.03.040

Strauss, A., & Corwin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques.* Newbury Park, CA: Sage.

Stivaros, S. M., Gledson, A., Nenadic, G., Zeng, X-J., Keane, J. & Jackson, A. (2010). Decision support systems for clinical radiological practice: Towards the next generation. *The British Journal of Radiology, 83,* 904-914. http://doi.org/10.1259/bjr/33620087

Stroman, P. W., Wheeler-Kingshott, C., Bacon, M., Schwab, J. M., Bosma, R., Brooks, J. …Tracey, I. (2014). The current state-of-the art of spinal cord imaging: Methods . *Neuroimage, 84*, 1070-1081. http://doi.org/10.106/j.neuroimge.2013.04.124

Suthar, P., Patel, R., Mehta, C. & Patel, N. (2015). MRI evaluation of lumbar disc herniation. *Journal of Clinical Diagnostic Research, 9*(4), TCO4-9. http://doi.org/10.7860/JCDR/2015/11927.5761

Sutton, E. J., Huang, E. P., Drukker, K., Burnside, E. S., Hui, L., Net, J. M. . . . Morris, E. A. (2017). Breast MRI radiomics: Comparison of computer- and human-extracted imaging phenotypes. *European Radiology Experimental, 1*(22), 1- 9. http://doi.org/10.1186/s41747-017-0025-2

Syeda-Mahmood, T. (2015). Role of machine learning in clinical decision support. Plenary presentation from SPIE medical imaging 2015 [Video file]. Retrieved from https://www.youtube.com/watch?v=A4Uk88-DOBA

Syeda-Mahmood, T. (2018). Role of big data and machine learning in diagnostic decision support in radiology. *Journal of the American College of Radiology, 15*(3 Pt B), 569-576. http://doi.org/10.1016/j.jacr.2018.01.028

Tang, A., Tam, R., Cadrin-Chenevert, A., Guest, W., Chong, J., Barfett, J., Chepeev, L.…Gies, R. (2018). Canadian association of radiologist's white paper on artificial intelligence in radiology. *Canadian Association of Radiologists Journal, 69,* 1-16. https://doi.10.1016/j.carj.2018.02.002

Tang, Q., Lui, Y., & Liu, H. (2017). Medical image classification via multiscale representation learning. *Artificial Intelligence Medicine, 79*, 71-78. http://doi.org/10.1016/j.art,ed.2017.06.009

Tarassoli, S. P. (2019). Artificial intelligence, regenerative surgery, robots? What is realistic for the future of surgery. *Annals of Medicine and Surgery, 41,* 453-55. https//doi.org/10.1016/j.amsu.2019.04.001

Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care, 19*(6), 349-357. http://doi.org/10.1093/intqhc/mzm042

Thammasitboon, S., Thammasitboon, S., & Singhal, G. (2013). System-related factors contributing to diagnostic errors. *Current Problems in Pediatric and Adolescent Health Care, 43*(9), 242-247.

Thrall. J. H. (2016). Trends and developments shaping the future of diagnostic medical imaging: 2015 Annual oration in diagnostic radiology. *Radiology, 279*(3), 660-666.

Thrall, J. H., Li, X., Li, Q., Cruz, C., Do, S., Dreyer, K., & Brink, J. (2018). Artificial intelligence and machine learning in radiology: Opportunities, challenges, pitfalls, and criteria for success. *Journal of the American College of Radiology*, S1546-1440(17)30671-X. [Epub ahead of print] . http://doi.org/10.1016/j.jacr.2017.12.026

Turing, A. (1950). Computer machinery and intelligence. *Mind.* Lex (236), 433-460. http://doi.org/10.1093/mind/lix236.433

Vallieres, M., Freeman, C. R., Skamene, S. R., & Naga, E. (2015). A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastasis in soft-tissue sarcomas of the extremities. *Physical Medicine and Biology*, *60*, 5471-5496. http://doi.org/10.1088/0031-9155/60/14/5471

Vallieres, M., Zwanenburg, A., Badic, B., Cheze-Le Rest, C., Viskikis, D. & Hatt, M., (2017). Responsible radiomics research for faster clinical translation. *Journal of Nuclear Medicine*, Epub ahead of print. http://doi.org/10.2967/jnumed.117.200501

van Griethuysen, J. J. M., Fedorov, A., Parmar, C., Hosyn, A., Aucoin, N., Narayan, V., . . . Aerts, H. J. W. L. (2017). Computational radiomics system to decode the radiographic phenotype. *Cancer Research, 77*(21), e104-e107. http://doi.org/10.1158//0008-5472.CAN-17-0339

Vargas, H. A., Huang, E. P., Lakhman, Y., Ippolito, J. E., Bhosale, P., Mellnick, V., . . . Sala, E. (2017). Radiogenomics of high-grade serous ovarian cancer: Multireader multi-institutional study from the cancer gnome atlas ovarian cancers imaging research group. *Radiology, 285*(3), 482-492.

Venkatesh, V. (2008). Technology acceptance model 3 and a research agenda on intervention. *Decision Sciences, 39*(2), 273-315. http://doi.org/10.1111/j.1540-5915.2008.00192.x

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science, 46*(2), 186-2014.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425-478.

Verma, V., Simone, C. B., Krishan, S., Lin, S. H., Yang, J. & Hahn, S. M. (2017). The rise of radiomics and implications for oncological management. *Journal of the National Cancer Institute, 109*(7). http://doi.org/10.1093/jnci/djx055

Viswanath, S. E., Tiwari, P., Lee, G., & Madabhushi, A. (2017). Dimensionality reduction-based fusion approaches for imaging and non-imaging biomedical data: Concepts, workflow, and use-cases. *Biomed Central Medical Imaging, 17*(2), 1-17. http://doi.org/10.1186/s12880-016-0172-6

Waite, S., Scott, J. M., Legasto, A., Kola, S., Gale, B., & Krupinski, E. A. (2017). Systemic error in radiology. *American Journal of Roentgenology, 20*(3), 629-639. http://doi.org/10.2214/AJR.16.17719

Wang, S., Hu, Y., Shen, Y., & Li, H. (2018). Classification of diffusion tensor metrics for the diagnosis of a myelopathic cord using machine learning. *International Journal of Neural Systems, 28*(2), 1750036. http://doi.org/10.1142/SO129065717500368

Wang, S., Li, X., Cui, J. L., Li, H.X., Luk, K. D., & Hu, Y. (2015). Prediction of myelopathic level in cervical spondylotic myelopathy using diffusion tensor imaging. *Journal of Magnetic Resonance Imaging,41*(6), 168201688. http://doi.org/10.1002/jmri.24709.

Wang, S., & Summers, R. M. (2012). Machine learning and radiology. *Medical Image Analysis, 16*(5), 933-951. http://doi.org/10.1016/j.media.2012.02.005

Wheeldon, J., & Faubert, J. (2009). Framing experience: Concept maps, mind maps, and data collection in qualitative research. *International Journal of Qualitative Methods, 8*(3), 68-83.

Williams, B. J. (2017). Future-proofing pathology: The care for clinical adoption of digital pathology. *Journal of Clinical Pathology, 70,* 1010-1018. http://doi.org/10.1136/jclinpath-2017-204644

Wu, H. W., Davis, P. K., & Bell, D. S. (2012). Advancing clinical decision support using lessons from outside of health care: An interdisciplinary systematic review. *BMC Medical informatics and Decision Making, 12*(20), 1-10. http://biomedcentral.com/1472-6947/12/90

Wu, J., Khin, Tha, K. K., Xing, L., & Li, R. (2018). Radiomics and radiogenomics for precision radiotherapy. *Journal of Radiation Research, 59*(S1), i25-i31. http://doi.org/10.1093/jrr/rxx2102

Yin, R. (2014). *Case study research: Design and methods* (5th ed.). Los Angeles, CA: Sage.

Yip, S. F., & Aerts, J. W. L. (2016). Applications and limitations of radiomics. *Physics in Medicine and Biology, 61,* R150-166. http://doi.org/10.1088/0031-9155/61/13/R15

Zaharchuk, G., Gong, E., Wintermark, M., Rubin, D., & Langlotz, C. P. (2018). Deep learning in neuroradiology. *American Journal of Neuroradiology,* 1-9. http://doi.org/10.3174/ajnr.A5543

Zhang, J., Patal, V. L., Johnson, T. R., & Shortliffe, E. H. (2004). A cognitive taxonomy of medical errors. *Journal of Biomedical Informatics, 37,* 193-204. http://doi.org/10.1016/j.jbi.2004.04.004

Zherhouni, E. A. (2017). Imaging innovation in the 21st century. Biomedicine-challenges and opportunities, RSNA Annual Meeting. Symposia conducted at the meeting of the Radiological Society of North America, Chicago Illinois.

Zhao, B. S., Tan, Y. Q., Tsai, W. Y., Qi, J., Xie, C., Lu, L. & Schwartz, L. H. (2016). Reproducibility of radiomics for deciphering tumor phenotype with imaging. *Scientific Reports, 6,* 23428. http://doi.org/10.1038/srep23428

Zhou, M., Scott, J., Chaudhury, B., Hall, L., Goldgof, D., Yeom, K. W., . . . Gatenby, R. (2017). Radiomics in brain tumor: Image, assessment, quantitative feature descriptors, and machine -learning approaches. *American Journal Neuroradiology.* http://doi.org/10.3174/ajnr.A.5391

Appendix A: NIH Certificate

Certificate of Completion

The National Institutes of Health (NIH) Office of Extramural Research certifies that **David Durrant** successfully completed the NIH Web-based training course "Protecting Human Research Participants."

Date of Completion: 07/30/2018

Certification Number: 2872343

NIH National Institutes of Health
Office of Extramural Research

Appendix B: Research Participant Survey

Research Participant Survey

Name: ______________________________ Street Address: _______________________________
City: _________________________ State:___________________ Country __________________
Personal Phone: _________________ Personal Email Address: ____________________________

Primary Professional Role: ___ Administrator ___Researcher ___ Clinician ___ Consultant
Degrees (check all that apply): ___ BS ___MS ___PhD ___ MD ___ Other _____________
Specialty: __Artificial Intelligence __ Radiology __ Bioinformatics __Computer Science
Experience in Specialty: __________ years

Are you familiar with the concept of the "Virtual Biopsy" as used in the field of radiology?
___ Yes
___ No

Rate your level of familiarity with radiomics.
___ Unfamiliar
___ Familiar
___ Expert

Rate your level of familiarity with natural language processing.
___ Unfamiliar
___ Familiar
___ Expert

Do you believe that the use of AI solutions during the interpretive stage of radiology workflow would augment the role of the radiologist and reduce the risk for error?
___ Yes
___ No

Do you believe that the use of AI solutions during the interpretive stage of radiology workflow could improve diagnostic precision?
___ Yes
___ No

Please identify why decided to participate in this research study?

Appendix C: Qualitative Coding Guide

Qualitative Coding Guide

Data Coding and Analysis:
A multi-step process will be used to analyze acquired data form different expert sources in this exploratory qualitative case study. Data analysis will begin as soon as data is acquired and will continue throughout the research process. Focus group transcripts and consensus-based white papers will be imported into Atlas.TI (Version 8) and subjected to an inductive and iterative process of content analysis and thematic coding. A hybrid approach will be used allowing for aggregation, subtraction, and expanding of code categories. The coding process will be used to identify patterns, relationships, themes, and subthemes. The analysis of focus group data will include identification of noteworthy quotes, outlying factors, and unexpected findings. The coding process will be performed until topic saturation is achieved.

The application of data saturation in this study was operationalized by the principal research question and the context of the research study. In this study data saturation was achieved if additional analysis of the research sources was unnecessary to address the research questions. This is confirmed if further attempts at acquiring and analyzing data do not lead to new perspectives or thematic conclusions

The first step in the coding process will be to systematically reduce the complexity of acquired data. Provisional codes were developed a priori to assist in this process. The provisional codes were developed with insights acquired from an extensive literature search, the research questions, and theoretical concepts derived from the diffusion of innovations theory (DOI) and the technology acceptance model (TAM). The provisional codes will likely be replaced, revised or modified during data analysis to better address the acquired data.

The combination of content analysis and thematic coding may support the development of a concept map that could be used to depict the relationships between steps in radiology workflow. A concept map could be adapted to create a logic map representing more specific relationships between the flow of data and technological processes during the interpretive stage of spine imaging workflow. The use of concept or logic maps can help transform tacit knowledge into a practical resource.

Provisional Codes/Coding Categories:

Impact
Impact on radiologist
Impact on diagnostic precision
Impact on workflow
Impact on reporting
Impact on patient care

Behavioral Intention
Perceived Usefulness
Perceived Ease of use
Fear of use

Diagnostic Applications
Pathology detection
Pathology segmentation
Pathology feature extraction
Pathology feature analyses
Diagnostic inference

Decision Support
Differential diagnosis
Personalized care plan
Pathology surveillance

Adoption and Use	Pathology Diagnosis
Subjective Norm	Early stage detection
Job Relevance	Virtual biopsy
Results Demonstrability	In vivo evaluation
Social Influence	Whole pathology assessment
Supportive Conditions	Computational modeling

Appendix D: Focus Group Moderator Guide

Focus Group Moderator Guide

Dissertation Title: Artificial Intelligence: Potential Impact on Spine Imaging Interpretation and Diagnosis.
Researcher: Dr. David H. Durrant, board certified chiropractic neurologist and PhD candidate.
Institution: Walden University
Focus Group Location(s): One Collective
Focus Group Access: In person or through teleconference.

Focus Group Guidelines and Scripts

Introduction: Explanation of Study

I would like to thank all of you for participating in this focus group session. The goal of this session is to obtain your feedback, opinions, and perspectives regarding the potential impact of artificial intelligence (AI) on the interpretive stage of spine imaging workflow. The session will be recorded in its entirety after the introductory slide program. All of your answers will be used for research. Your contributions to the session will remain confidential and your anonymity will be protected. Any quotes used to support emergent themes in the final dissertation document will be placed into quotations and identified with expert class followed by an assigned participant number. Your participation in focus group discussion today will contribute to the future development of AI solutions in radiology.

Focus Group: Explanation of Purpose

This focus group session is designed to support a free flowing, creative, and scholarly discussion. There are no desirable or undesirable answers. You can disagree with each other, and you can change your mind during the course of discussion. I would like each of you to feel comfortable contributing what you think and what you know about the topic. We have a lot of expertise, and experience in this session today; therefore, we should have fun discussing possibilities.

General Instructions and Directives

During the next ninety minutes I will serve as the moderato of this focus group session.

All of your contributions and responses during the focus group session will be recorded verbatim. I will receive two sets of transcripts. The transcript data will be analyzed. A few weeks after this focus group session you will receive summary of themes and subthemes derived from the focus group transcripts along with the list of supportive quotes for your review .You will be given two weeks to the review the documents and to provide your feedback.

During the course of this focus group session, I ask that you give each other a chance to express your opinions. Be courteous and do interrupt or talk over one another. If I feel that anyone is taking too much time on a topic I will intervene. I also request that each of you respect each other's confidential participation and contributions to this study.

I encourage everyone to contribute. You are each welcome to respond to my questions and to respond to each other's contributions.

Participant Introductions

I would like each of you to provide an introduction consisting of your degree, your expertise, general job description, and interest surrounding use of artificial intelligence in radiology. Please identify whether you have a specific interest regarding the use of AI in spine imaging or in spine care.

Introductory Slide Program

I have prepared a brief PowerPoint program that will be used to introduce topics to be addressed during the focus group sessions. The program introduces general concepts surrounding the use of AI in radiology. The approach is designed to help focus the discussion of complex topics. The introductory slide programs designed to take about 10 minutes to present.

Steps Taken to Prepare for Focus Group Sessions

A moderator guide was developed to help me lead this session. I will use it to help prepare you for the session and to present questions. The primary questions will be open-ended. I may occasionally is a semi-structured probing question to facilitate more comprehensive discussion of the topic. The introductory PowerPoint program and questions were both subjected to independent review and field testing to help ensure their appropriateness and relevance to the topic of study.

Focus Group Ground Rules

General Considerations: The following ground rules have been established for use in focus group sessions. The ground rules were designed to facilitate scholarly discussion and to help protect the rights of research participants. The ground rules will be presented at the beginning of each focus group sessions.

Ground Rules

- Participation in this focus group is entirely voluntary and is based on consent.
- You each have the right to leave the session at any time and for any reason.
- The focus group session will take approximately 90 minutes. I may allow the session to go a few minutes little longer if necessary.
- It is okay to abstain from participating in the discussion of select topics if you are not comfortable contributing.
- Please turn off or silence all mobile phones.
- You are welcome to access refreshments or use the restroom at any time during the session.
- There is no right or wrong answers. Every participant's contribution is valuable.

- Please respect the opinions of other research participants even if you do not agree with them.
- It participant names are used during the focus group session they will be removed from audio transcripts to maintain participant privacy outside the study.
- Please do not interrupt a research participant when they are speaking.
- It is important that only one person speak at a time.
- You do not have to speak in any particular order.
- Please speak in a clear voice, loud enough for everyone to hear, and precise enough to be audio recorded accurately.
- This focus group session will be audio recorded in its entirety. A complete transcript will be created and used for content and thematic analysis
- A professional stenographer and transcriber will be present during the entire focus group session to ensure accurate transcripts.
- Is important that each of you respect and protect each other's anonymity and the confidentiality of contributions outside the study.
- As the moderator of this focus group session I may occasionally intervene to facilitate topic discussion or to ensure ground rules are followed. My primary role as the moderator is to ensure professional and topic specific discussion. Because of the limitations in time I may occasionally have to facilitate or re-direct the discussion with probing questions.

Focus Group Research Questions

The Primary Research Question:

What are the opinions of experts regarding the potential use and impact of artificial intelligence (on spine image interpretation diagnosis) during the interpretive stage of spine imaging workflow?

Subtopics: Open-Ended and Probing Questions:

Differential Diagnostic Process

How could the use of AI-supported methods (auto detection, segmentation, radiomics, natural language processing) during the interpretive stage of spine imaging influence the differential diagnostic process?

How could the use of AI-supported methods during the interpretive stage of spine imaging influence disease classification and staging?

Interpretive Workflow

What AI solutions could be used to create interpretive priority in spine imaging?

What will the future of spine imaging interpretation workflow look like?

Radiomics/Virtual Biopsy

Could AI-supported solutions such as radiomics be used to interrogate spinal tissue in vivo and eventually lead to a virtual (digital) biopsy?

What are some potential advantages of an in vivo virtual "digital" biopsy over a traditional needle biopsy in spine care?

Artificial Intelligence and Augmented Reality/Virtual Reality

How could the use of AI-supported augmented reality (AR) or virtual reality (VR) enhance the evaluation of pathology in spine imaging?

Meaningful Use Applications (Specific Spine Disorders)

What are some potentially "meaningful use" applications of AI in spine imaging?

AI Adoption and Use

Which construct of the technology acceptance model (TAM) will likely have a greater impact on AI adoption during the interpretive stage of spine imaging: perceived benefits or perceived ease-of-use?

Which characteristics of innovations proposed by the diffusion of innovation theory (DOI) will likely have the greatest impact on AI adoption during the interpretive stage of spine imaging: complexity, compatibility and interoperability, observed effects or trialability?

Determination of Data Saturation

The application of data saturation in this study was operationalized by the principal research question and the context of the research study. In this study data saturation was achieved if additional analysis of the research sources was unnecessary to address the research questions. This is confirmed if further attempts at acquiring and analyzing data do not lead to new perspectives or thematic conclusions

Closing Discussion (Script)

Thank you very much for choosing to participate in this focus group session. Is there any additional information we may have missed that may be important for this study? What is the most important topic we discussed today? Your time is very much appreciated your contributions helpful.

Within the next few weeks you will receive a letter and a copy of your transcript from today's focus group session. Please review the document with your signature of approval within the time allotted.

Appendix E: Focus Group Introductory Slide Program

Focus Group Session Introduction

Dr. David H. Durrant
(PhD Candidate)
Researcher and Moderator

PhD Dissertation Research

Artificial Intelligence: Potential Impact on Spine Imaging Interpretation and Diagnosis

WALDEN UNIVERSITY
A higher degree. A higher purpose.

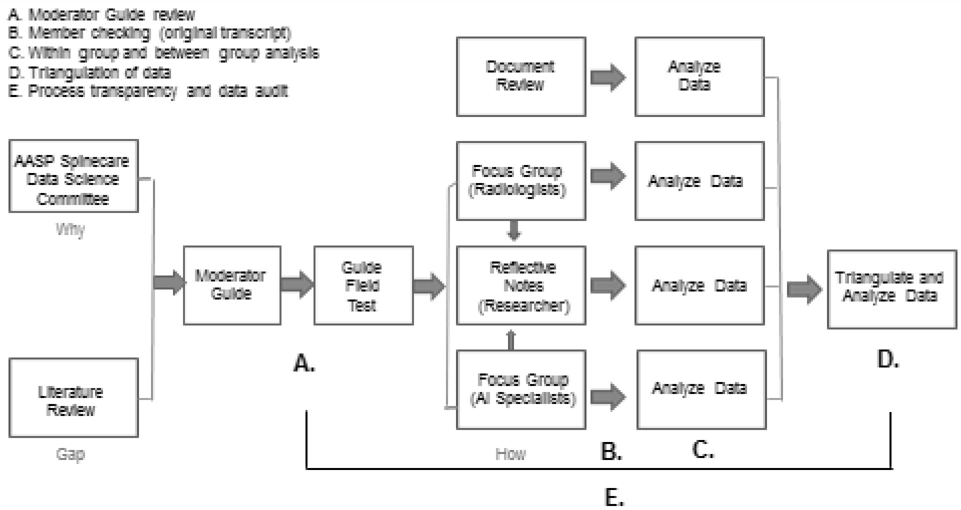
Focus Group Guidelines

- It is important that only one person speak at a time.
- Please speak in a clear voice, loud enough to hear and for audio recording.
- This focus group session will be audio recorded in its entirety. A complete transcript will be created and used for content and thematic analysis.
- Participant names are used during the focus group session they will be removed from audio transcripts to maintain participant privacy.
- There is no right or wrong answers. Every participant's contribution is valuable.
- My role as a moderator is to direct an open discussion

2

WALDEN UNIVERSITY
A higher degree. A higher purpose.

Research Process and Data Flow



  WALDEN UNIVERSITY *A higher degree. A higher purpose.*

Research Study Purpose

To identify the potential impact of artificial intelligence on spine imaging interpretation and diagnosis

This includes addressing its potential to
improve workflow efficiency,
reduce interpretive errors,
and improve diagnostic accuracy.

  WALDEN UNIVERSITY *A higher degree. A higher purpose.*

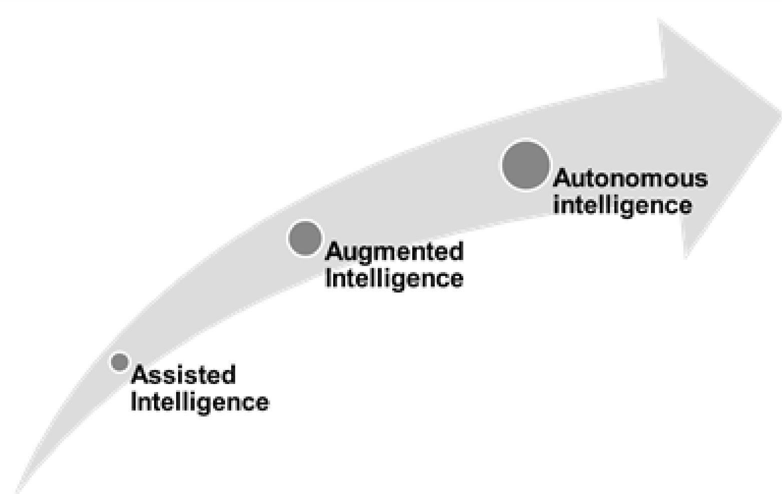
Background: Big Data and Complex Decisions

- Meaningful data and insight is embedded within medical images, undetectable through routine visual analysis, and therefore not considered in the diagnostic process (Gillies, Kinahan, Hricak, 2016; Lee, 2017).

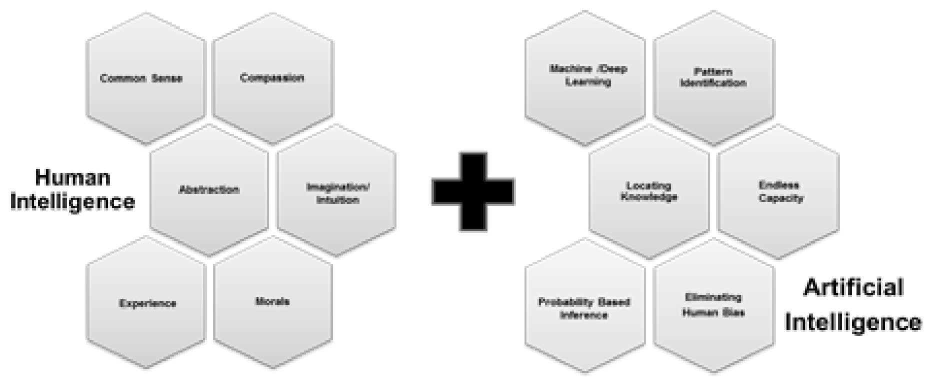
- Without adequate technological assistance the human interpretive process within radiology workflow will become progressively more complex, inefficient, inaccurate, and untimely (Croskerry 2013; Obermeyer & Emanuel, 2016; Ragupathi & Ragupathi, 2014).

- To improve diagnostic accuracy the radiologist requires new solutions such as AI-supported radiomics for the detection, characterization, classification and quantification of disease in vivo (Aerts et al., 2014).

Concepts: Projected AI Evolution

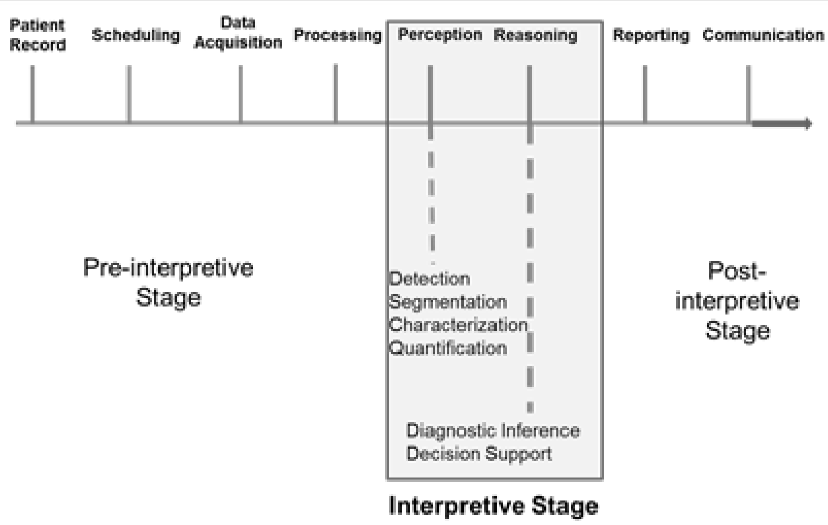


Collective Intelligence (CI)



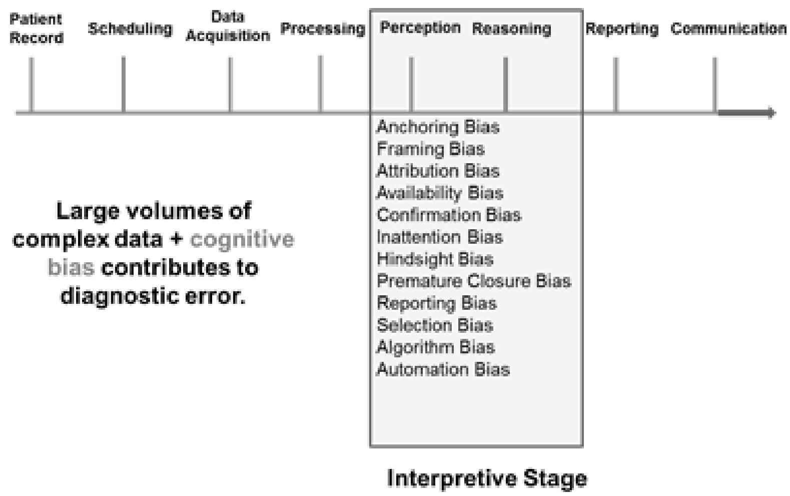
7 © 2019 Dr. David H. Durrant WALDEN UNIVERSITY *A higher degree. A higher purpose.*

Radiology Workflow: Interpretive Stage



8 © 2019 Dr. David H. Durrant WALDEN UNIVERSITY *A higher degree. A higher purpose.*

Radiology Workflow: Interpretive Stage Bias



9

WALDEN UNIVERSITY
A higher degree. A higher purpose.

Introduction: Biopsy

Biopsy:

an examination of tissue removed from a living body to discover the presence, cause or extent of a disease.

Digital (Virtual) Biopsy

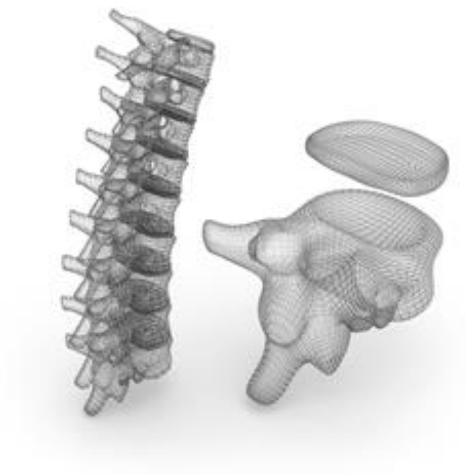
an examination of tissue (data) removed from a living body to discover the presence, cause or extent of a disease.

10

WALDEN UNIVERSITY
A higher degree. A higher purpose.

Introduction: Localizing Pathology

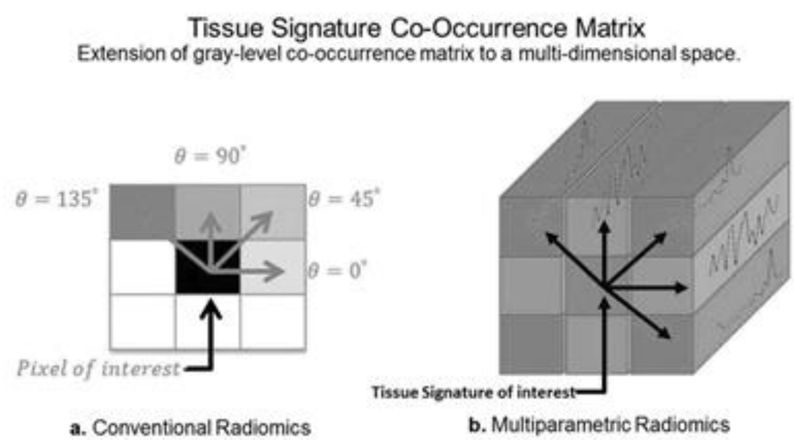
- A voxel represents a 3D volume of space on a regular grid which contains data and values.
- Non-visible data/patterns can be extracted from voxels and analyzed using AI (deep learning) pattern detection methods



11 Permission for use of graphic provided by Alamy Reference OY27054567

WALDEN UNIVERSITY
A higher degree. A higher purpose.

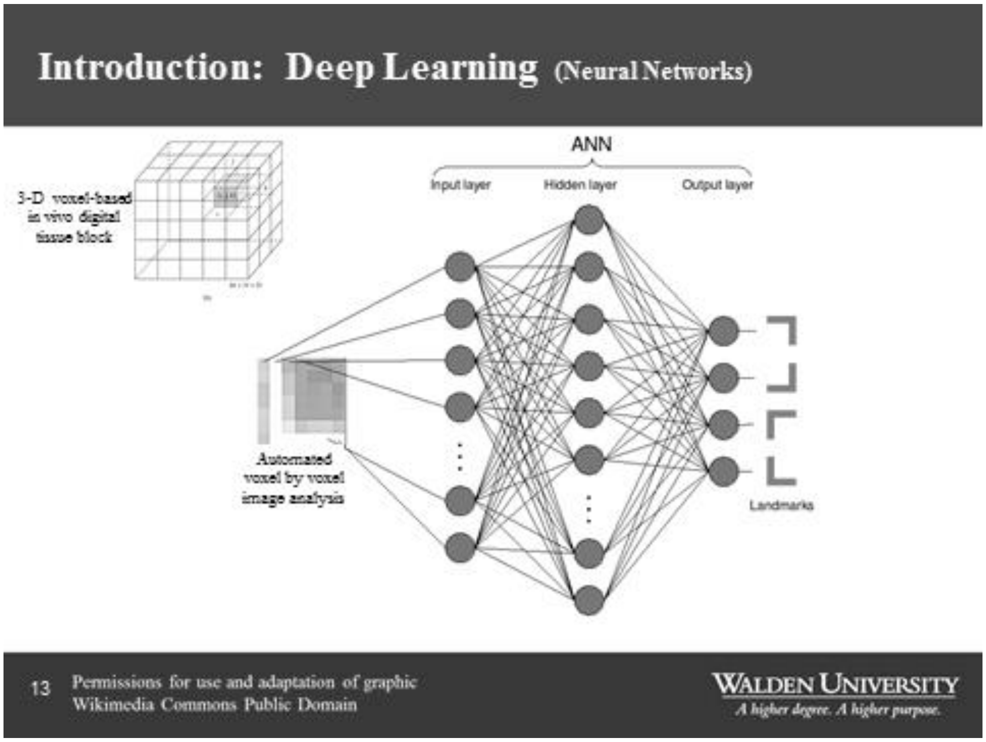
Introduction: Voxel-Based Interrogation

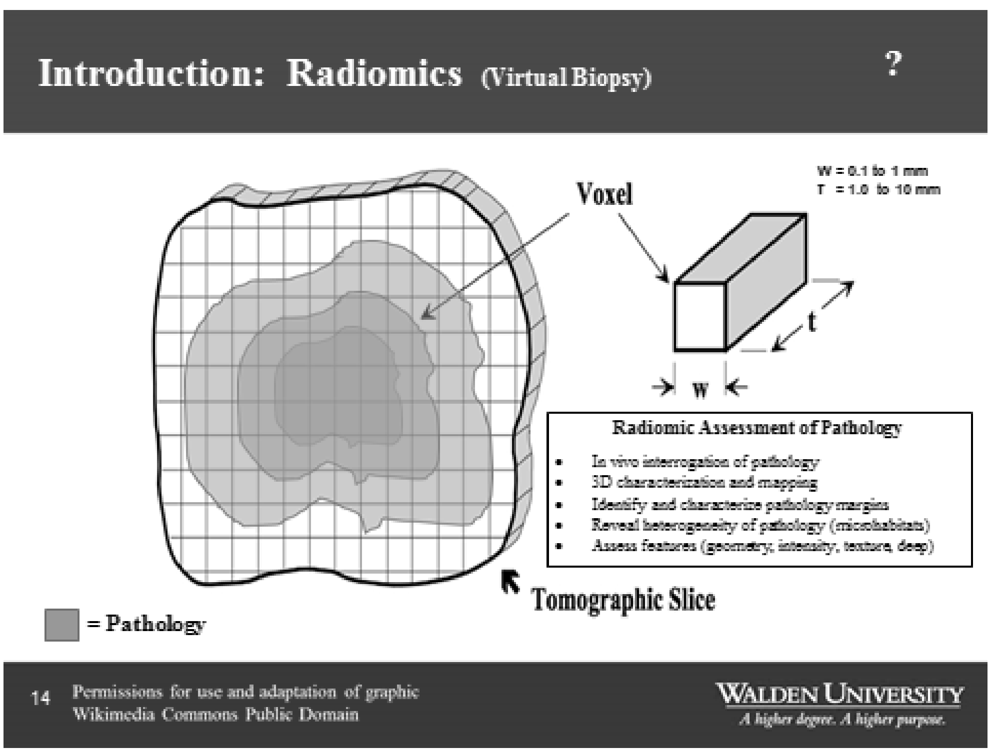


Single image radiomics features extract the inter-pixel relationships in-plane of a radiological image whereas, multiparametric radiomics extract the inter-tissue-signature relationships across multiple radiological images.

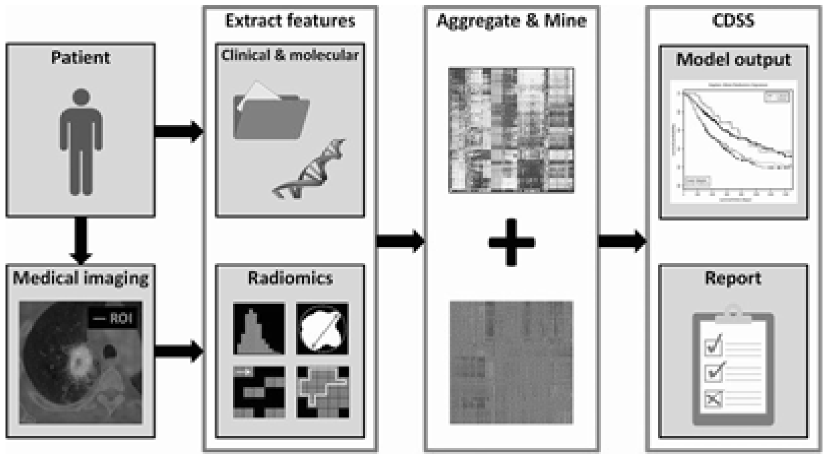
12 Used with permissions via Creative Commons Attribution License. Parekh, V., & Jacobs, M. A. (2019). Deep learning and radiomics in precision medicine. *Expert Review of Precision Medicine and Drug Development*, *4*(2), 59-72. doi:10.1080/23808993.2019.1585805

WALDEN UNIVERSITY
A higher degree. A higher purpose.


Introduction: Deep Learning (Neural Networks)
ANN
Input layer
Hidden layer
Output layer
3-D voxel-based in vivo digital tissue block
Automated voxel by voxel image analysis
Landmarks
13
Permissions for use and adaptation of graphic
Wikimedia Commons Public Domain
WALDEN UNIVERSITY
A higher degree. A higher purpose.


Introduction: Radiomics (Virtual Biopsy)
?
Voxel
W = 0.1 to 1 mm
T = 1.0 to 10 mm
t
w
Radiomic Assessment of Pathology
• In vivo interrogation of pathology
• 3D characterization and mapping
• Identify and characterize pathology margins
• Reveal heterogeneity of pathology (microhabitats)
• Assess features (geometry, intensity, texture, deep)
Tomographic Slice
= Pathology
14
Permissions for use and adaptation of graphic
Wikimedia Commons Public Domain
WALDEN UNIVERSITY
A higher degree. A higher purpose.

Radiomics: The Sequential Process



15 Permissions for graphics use granted by Nancy International Ltd, subsidiary of AME Publishing Company

WALDEN UNIVERSITY
A higher degree. A higher purpose.

Relevant Spine Facts

- Low back pain is the #1 cause of disability worldwide. (1)
- The spine is the most common site of bone metastasis. (2)
- About 25% of all post-menopausal women will develop a vertebral compression fracture. (3)
- MRI signs of cervical spinal cord compression increase with age from about 30% in the fifth decade to about 70% in the eighth decade. Most presentations are asymptomatic at the time of diagnosis. (4)
- The prevalence of degenerative disc disease (over the entire spine) is 70-80% in adults <50 years of age, and >90% of adults >50 years of age. (5)
- The lifetime prevalence of significant low back pain is > 80%. (6)

The human spine is intricate and complex. The majority of pathology occurring in the spine is to small to be detected by traditional anatomic imaging.

16 Refer to reference slide for sources of facts.

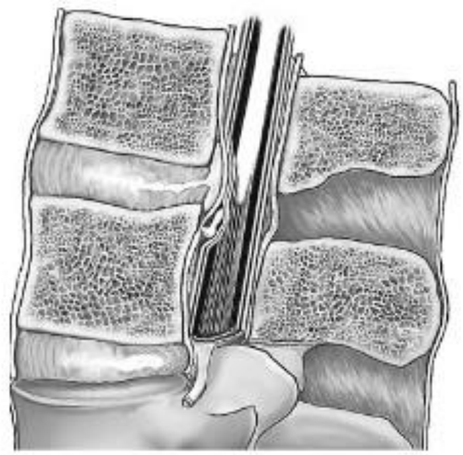
WALDEN UNIVERSITY
A higher degree. A higher purpose.

AI and Radiomics: Meaningful use in Spine Care

?

Meaningful Use Applications?

- Vertebral Localization & Labeling
- Spinal Cord Pathology
- NLP → Fracture Detection
- Intervertebral Disc Pathology
- Bone Marrow Pathology
- Spinal Tumors
- Post-Operative Spine

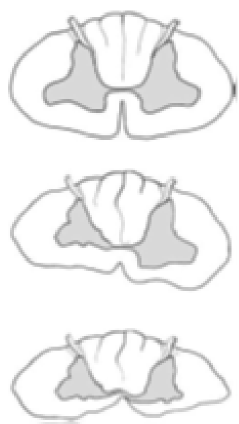


17 Image used with permissions of Scholars Consortium LLC.
Durrant, D. H., & True, J. M. (2012). Myelopathy, radiculopathy and peripheral entrapment syndromes. Palm City, Florida: Scholars Consortium, LLC.

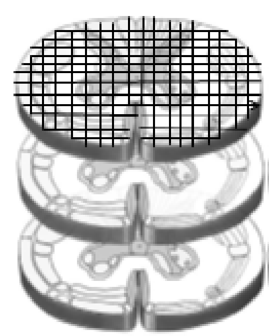
WALDEN UNIVERSITY
A higher degree. A higher purpose.

Spinal Cord Pathology: Radiomics

Spinal Stenosis



Spinal Cord Compression

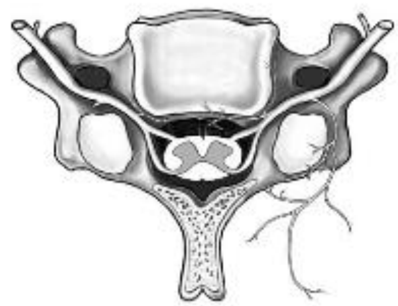


Spinal Cord Radiomics

18 Images used and adapted with permissions of Scholars Consortium LLC.
Durrant, D. H., & True, J. M. (2012). Myelopathy, radiculopathy and peripheral entrapment syndromes. Palm City, Florida: Scholars Consortium, LLC.

WALDEN UNIVERSITY
A higher degree. A higher purpose.

Spinal Cord Pathology: Quantitative Measures

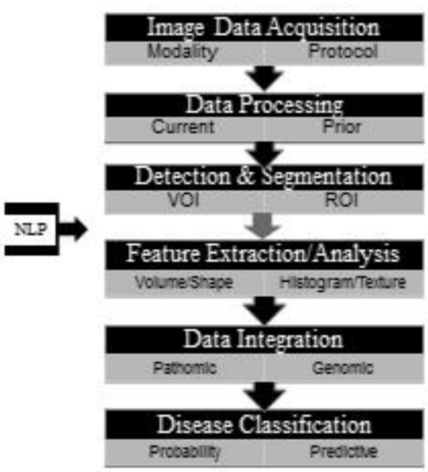


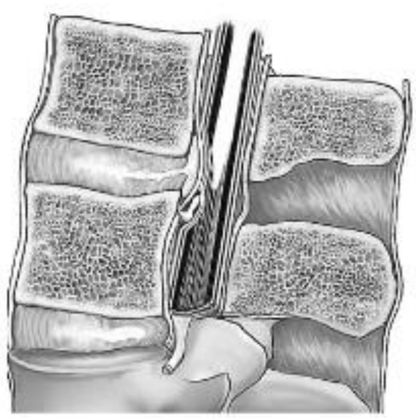
19 Images used and adapted with permissions of Scholars Consortium LLC. Durrant, D. H., & True, J. M. (2012). Myelopathy, radiculopathy and peripheral entrapment syndromes. Palm City, Florida: Scholars Consortium, LLC.

WALDEN UNIVERSITY
A higher degree. A higher purpose.

?

AI and Radiomics: Meaningful use in Spine Care





20 Image used with permissions of Scholars Consortium LLC. Durrant, D. H., & True, J. M. (2012). Myelopathy, radiculopathy and peripheral entrapment syndromes. Palm City, Florida: Scholars Consortium, LLC.

WALDEN UNIVERSITY
A higher degree. A higher purpose.

Immersive Technologies: AR, VR and Mixed Reality

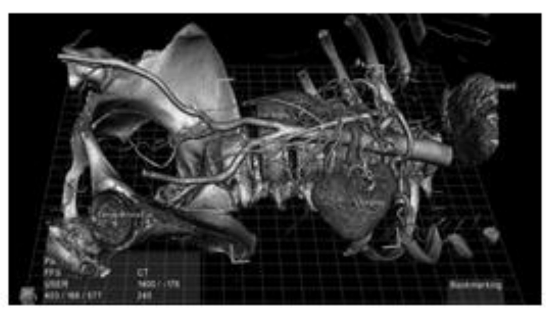
<u>Potential benefits of immersive technologies include:</u>

- enabling physicians to interact with imaging data
- visualization of pathology/3D feature mapping
- enhanced pre- and intra-operative planning
- 3-D integration and display of anatomic, molecular and radiomic data
- high-yield biopsy guidance
- surgical simulation training
- support for head tracking and motion control
- Integration with AI for bioprinting, prosthetic development and robotic procedures
- instantaneous and precise visualization of 3D spatial relationships
- multidimensional disease surveillance
- improved patient selection for care

21 Images used with the permissions of Echopixel

WALDEN UNIVERSITY
A higher degree. A higher purpose.

Immersive Technologies: AR, VR and Mixed Reality ?



22 Images used with the permissions of Echopixel

WALDEN UNIVERSITY
A higher degree. A higher purpose.

Concepts: Auto-Enabled Radiology Support

Detect and Characterize → Alert and Prioritize → Compare

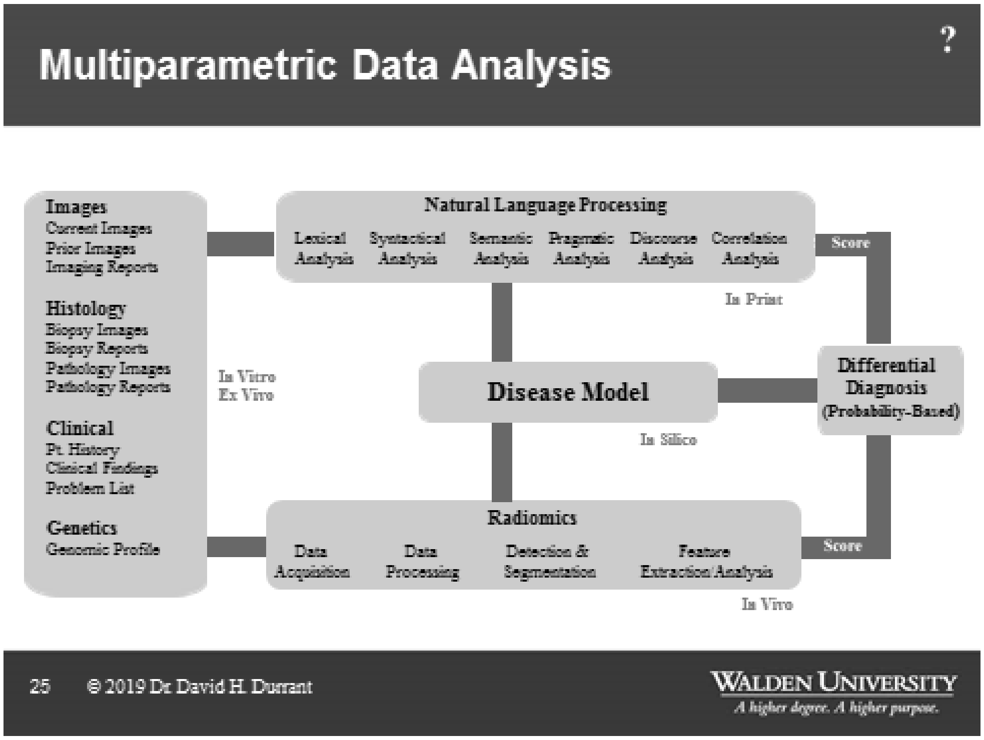
23

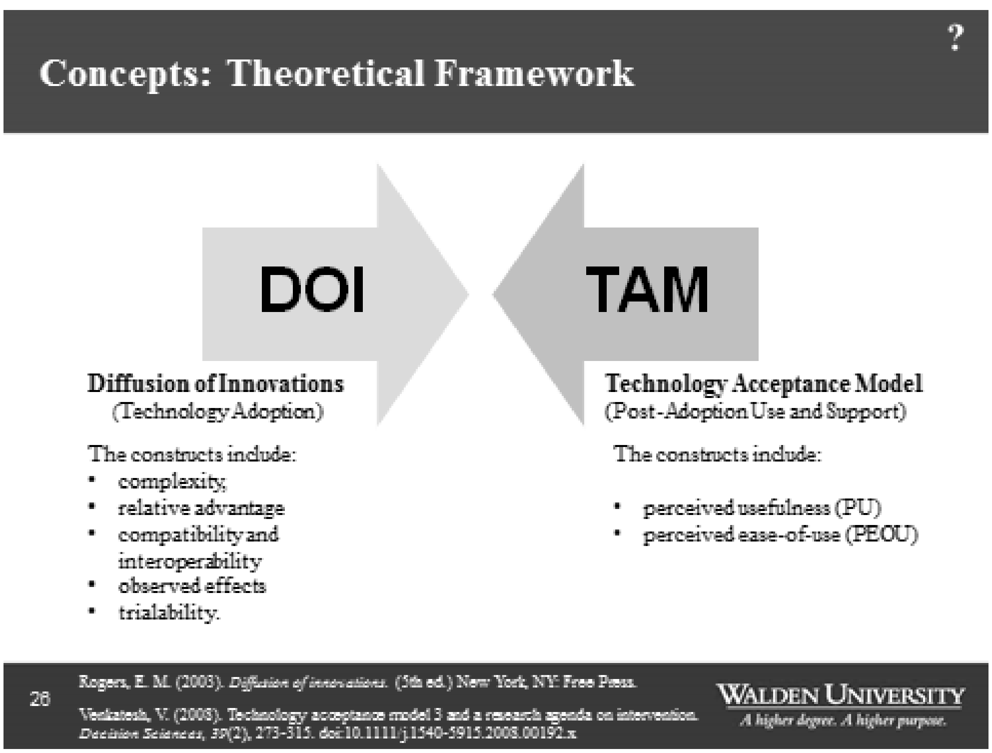
WALDEN UNIVERSITY
A higher degree. A higher purpose.

Radiology Workstation of the Future ?

24

WALDEN UNIVERSITY
A higher degree. A higher purpose.


?
Multiparametric Data Analysis
Images
Current Images
Prior Images
Imaging Reports
Histology
Biopsy Images
Biopsy Reports
Pathology Images
Pathology Reports
Clinical
Pt. History
Clinical Findings
Problem List
Genetics
Genomic Profile
In Vitro
Ex Vivo
Natural Language Processing
Lexical Analysis
Syntactical Analysis
Semantic Analysis
Pragmatic Analysis
Discourse Analysis
Correlation Analysis
Score
In Print
Disease Model
In Silico
Differential Diagnosis
(Probability-Based)
Radiomics
Data Acquisition
Data Processing
Detection & Segmentation
Feature Extraction/Analysis
Score
In Vivo
25
© 2019 Dr. David H. Durrant
WALDEN UNIVERSITY
A higher degree. A higher purpose.


?
Concepts: Theoretical Framework
DOI
TAM
Diffusion of Innovations
(Technology Adoption)
The constructs include:
• complexity,
• relative advantage
• compatibility and interoperability
• observed effects
• trialability.
Technology Acceptance Model
(Post-Adoption Use and Support)
The constructs include:
• perceived usefulness (PU)
• perceived ease-of-use (PEOU)
26
Rogers, E. M. (2003). Diffusion of innovations. (5th ed.) New York, NY: Free Press.
Venkatesh, V. (2008). Technology acceptance model 3 and a research agenda on intervention. Decision Sciences, 39(2), 273-315. doi:10.1111/j.1540-5915.2008.00192.x
WALDEN UNIVERSITY
A higher degree. A higher purpose.

Thank you for your participation!

27 Image used with the permissions of iStock ID 1022892498

WALDEN UNIVERSITY
A higher degree. A higher purpose.

Member Checking

You will have an opportunity to review the summary of thematic analysis and supportive quotes from the focus group sessions prior to publication.

During the document review process you will be asked to identify misinformation, proprietary information, or potentially incriminating information you do not want included in the final analysis of data.

You will also have an opportunity to offer comments on the validity of the thematic conclusions.

28

WALDEN UNIVERSITY
A higher degree. A higher purpose.

Spine Fact References

(1) Vos T, Barber RM, Bell B et al. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet 2015; 386*:743–800. 10.1016/S0140-6736(15)60692-4

(2) Maccauro, G., Spinelli, M. S., Mauro, S., Perisano, C., Graci, C., & Rosa, M. A. (2011). Pathophysiology of spine metastasis *International Journal of Surgical Oncology, 2011*(ID 107969), 1-8. doi: 10.1155/2011/10969

(3) Alexandru, D., & Su, W. (2012). Evaluation and management of vertebral compression fractures. *The Permanente Journal, 16*(4), 46-51.

(4) Kovalova, I., Kerkovsky, M., Kadanka, Z., Kadanka, Z., Nemec, M., Jurova, B., Dusek, L.,Jarkovsky, J., & Bednarik, J. (2016). Prevelance and imaging characteristics of nonmyelopathic and myelopathic spondylotic cervical cord compression. *Spine, 41*(24), 1908-1916

(5) Teraguchi, M., Yoshimura, N, Hashizume, H., Yamada, H., Minamide, A. Oka, H.…Yoshida, M. (2014). Prevalence and distribution of intervertebral disc degeneration over the entire spine in a population-based cohort: The Wakayama spine study. *Osteoarthritis and Cartilage, 22*(1), 104-110.

(6) Rubin DI. (2007). Epidemiology and risk factors for spine pain. *Neurologic Clinics, 25*(2):353-71.

Appendix F: AASP Spinecare Data Science Committee Summary

AMERICAN ACADEMY OF SPINE PHYSICIANS

Spinecare Data Science Committee Summary

The volume and complexity of data in spinecare is growing exponentially. The unprecedented growth of data is arising from many sources including diagnostic methods, such as genetics, laboratory studies, and advanced imaging. The increasing burden of overwhelming data adds to the complexity of decision-making during the diagnostic process and at the point-of-care. The unprecedented challenge has led to a search for new decision-support solutions, such as artificial intelligence (AI), which can be implemented at different stages along the spectrum of spinecare.

The American Academy of Spine Physicians (AASP) plans to lead the effort in the exploration of potential applications of computational support and AI in spinecare with an emphasis on spine imaging and pathology characterization. To help achieve this goal, the AASP developed the Spinecare Data Science Committee to serve as a multidisciplinary forum to support collaboration between spinecare providers, AI developers, physicists, data scientists, and other industry leaders. The goals of the committee include:

- Identification of areas of spinecare that could benefit from the use of AI solutions and computational methods.
- Exploration of the potential role and impact of auto detection and characterization of pathology (radiomics/deep learning/spectroscopy) in spine imaging
- Discuss the required elements of meaningful use cases for the development of AI algorithms
- Contribute to the development of new spine disease taxonomies and classifications (subtyping) that could be used to annotate spine images and refine the differential diagnostic process.
- Create opportunities for acknowledging the effectiveness of AI in clinical settings.
- Identify and prioritize the clinical application of AI use during the interpretation of spine imaging studies.
- Organize and participate in multidisciplinary focus group sessions to explore the potential use and impact of AI in spine imaging and spinecare.
- Help identify common data elements (CDE) within the pipeline of spinecare, which could be used as input and output variables for algorithm development and implementation.
- Help develop implementation strategies for the application of AI solutions in spinecare.
- Heighten the state of awareness and readiness for AI use at all levels of spinecare

- Investigate the potential applications of computational 3D imaging and interactive virtual reality for tissue interrogation and surgical planning.
- Discuss the importance of protocol standards and image registration for the serial assessment of pathology
- Address the importance of AI supported augmented reality/virtual reality solutions for tissue (whole pathology) interrogation and surgical planning

The AASP Data Science Committee is committed to empowering the advancement, validation and implementation of algorithms and artificial intelligence solutions in spinecare for the benefit of patients, spinecare professionals and society.

The AASP Data Science Committee is comprised of experienced professionals from different disciplines within the spinecare field including physicians, data scientists, surgeons, radiologists, physicists and other nonsurgical spinecare professionals.

Appendix G: Participant Response Validation (Member Checking)

Dear Research Participant,

Thank you for your participation and valuable contributions to this research study. I have attached two documents for your review. The first document titled "Preliminary Thematic Summary" represents an overview of the themes that emerged during analysis of the audio transcripts from the AI expert and radiology focus group sessions. This document includes a list of supportive quotes from the focus group sessions. The second document titled "Table X" represents a table that reveals the code hierarchy used to label data during analysis. Please review the documents and complete the form titled "Research Participant Response Survey."

This research participant review process represents a form of member checking (participant validation), a technique commonly used in qualitative research to help improve the trustworthiness of the study. The themes from the focus group sessions will be synthesized with those from other data sources in the study.

Please return the completed Research Participant Response Survey to me within five business days so that I can finalize the data analysis process. Send the document to me via fax (847-888-1836) or email (neurodoc92@comcast.net).

Thank you again for your time and contributions to the research study. You will be receiving a stipend for participation within the next few weeks. I will provide you with access to the dissertation when it is completed and published.

Please contact me if you have any concerns or questions. You can each me at 847-888-1811 or on my cell phone at 847-385-8799.

Sincerely,

Dr. David H. Durrant
PhD Candidate

Table X

Themes That Emerged from Qualitative Data Analysis

Categories	Subthemes	Themes	Goal
Detection Segmentation Characterization Monitoring	Multiscale In Vivo Analysis	Patient-Based Decision Support	Augmented Role of the Radiologist *Improved Diagnostic Precision & Personalization*
Problem List Prior Radiology Reports Electronic Health Records	Natural Language Processing		
Structural Features Radiomic Features Molecular Features	Change Analysis		
Worklist Triage Image Triage Pathology Triage	Prioritization		
3D Images AR/VR Images Feature Mapping	Immersive Data Display		
In Vitro/ Ex Vivo In Vivo In Silico Clinical	Ground Truth	Population-Based Decision Support	
Disease Model (Omics) Annotated Data Training Data Validation Data	Knowledge Database		
Spinal Cord Disorders Bone Marrow Pathology Fractures Spine Pain	Clinical Utility	Application-Based Decision Support	
Perceived Benefits Ease-of-Use Interoperability Relative Advantage	Technology Attributes		

RESEARCH STUDY: PRELIMINARY THEMATIC SUMMARY

Study Title: Artificial Intelligence: Potential Impact on Spine Imaging Interpretation and Diagnosis
Researcher: Dr. David H. Durrant

Results

Three primary themes emerged from the qualitative analysis of data derived from two expert focus group sessions. One focus group consisted of radiologists and the other consisted of AI experts. The three primary themes were patient-based decision support, population-based decision support, and application-based decision support. Numerous subthemes were assigned to each theme. The primary themes account for interdependent levels of decision support required for AI solutions to augment the role of the radiologist during the interpretive stage of spine imaging.

The elements of each subtheme are supported by the synthesis of information acquired from consensus-based white paper analysis and focus group contributions. Supportive statements (quotes) from focus group participants are listed below the subtheme summaries. To maintain confidentiality of focus group participants, quotes are identified by the participant's expert class followed by a participant number (Px). This is followed by an overview of supportive information from consensus-based white papers. Diagnostic decision support in spine imaging requires the integration of patient and population data through specialized application of technologies. The following themes and subthemes address how this might be achieved.

Theme 1: Patient-Based Decision Support

Patient-based decision support in the context of this study refers to the use of data and knowledge about a specific patient acquired through the use of diagnostic imaging and/or personal medical records. This approach is used to help detect, characterize, and monitor information unique to the patient and their disease status. Patient-based decision support is often used to formulate a personalized diagnosis and to assess an individual's treatment outcome. Patient information can be integrated with population-based knowledge during radiology workflow to support a probability-based differential diagnostic process. Subthemes of patient-based decision support include multiscale in vivo analysis, natural language processing, change analysis, prioritization, and immersive data display.

Subtheme: Multiscale in Vivo Analysis

Diagnostic imaging data in spine care is underutilized. The assessment of pathology must extend beyond human visual and cognitive limitations. The spine can be divided or mapped into 2D and 3D partitions referred to as pixels or voxels. AI-supported molecular and radiomic assessment could be used to improve how spine disease is detected, characterized, and monitored. One of the first steps in this process is identification of a region of interest (ROI) followed by targeted segmentation. Manual segmentation is a time-consuming task. AI has the capability of facilitating fast and accurate semiautomated and automated segmentation in spine imaging. Success requires access to an adequate volume of annotated training data. Engineered hard-coded algorithms, deep learning, and radiomic methods can be used to assess nonvisible features of spine disease within a defined ROI. It could also be used to help define the margins, shape, volume, and heterogeneity of spine pathology. The combined use of multiparametric and multiscale in vivo analysis with radiomic methods has the potential to serve as a digital (virtual) biopsy. This approach is used in other specialties of health care and could be further developed or adapted for use in spine imaging. The virtual biopsy may eventually take priority and be used to guide the traditional biopsy when necessary. Multiscale in vivo tissue analysis would offer a holistic, systems perspective for characterizing and monitoring spine pathology.

> *"Wouldn't it be nice to know that there is an abnormality, even if we can't visualize it…?" (Radiologist – P2).*
>
> *" getting the right level, depending on the field of view is certainly a challenge to any fully automated application" (AI Expert-P2).*
>
> *"Segmentation is a very time-consuming task. I think that with AI and algorithms it is going to be done faster" (AI Expert - P4).*
>
> *"I like the idea of clinical changes detected inside of the cord that are not visible by traditional anatomic imaging, because we know that specifically in the cervical spine, if these changes occur over a long period of time, the spinal cord can accommodate and literally be a ribbon before the patient has any symptom." (Radiologist - P2)*
>
> *"Radiomics can give us some kind of insights that cannot be appreciated with the human eye because we cannot interpret or define the statistical appearance." (AI Expert - P3).*
>
> *"subvisual in vivo identification and characterization of pathology within the spine could turn what appears to be a routine image into a wonderful diagnostic challenge. (Radiologist - P2).*
>
> "*I think that maybe there's more going on in the nerve roots or the cauda equina than we ever imagined, and AI might open up a whole window of opportunity to reveal what was previously invisible… So just sitting here, I see a lot of potential and what could be, as I used to say, a very boring lumbar spine could suddenly turn into something wonderful and challenging" (Radiologist - P2).*
>
> *"We know we can characterize pathology voxel by voxel. So with the digital biopsy, I think that more and more these kinds of biopsies are going to be done first rather than the traditional one" (AI Expert – P4).*
>
> *"I think that more and more the digital biopsy will be used more than the traditional biopsy" (AI Expert-P4).*

Subtheme: Natural Language Processing

Patient data needs to be better integrated into radiology workflow. Natural language processing (NLP) can be used to locate and extract relevant information from various sources of unstructured data such as a patient's electronic medical records (EMR). This includes access to a patient's problem list, prior radiology reports, pathology results, genetic profiles, and general clinical information. AI-supported NLP can be used to create a prioritized list of "need-to-know" information for the radiologist at the time of spine image interpretation. This would provide access to imaging-specific contextual perspectives that would enhance interpretive efficiency and diagnostic precision. NLP could also be used to help overcome a radiologist's limited knowledge by identifying differential diagnostic possibilities from published literature. NLP could be used to check the accuracy of the spine radiology report during the post-interpretive stage of spine imaging workflow. This includes verification of statistical measures.

> *"natural language processing with imaging is always beneficial. …a lot of times the radiologist won't even have enough time to look for or through the EMR, for all, you know, the history, etcetera" (AI Expert - P2).*

> *"I think that natural language processing can automatically look through a relevant summary, for that patient of everything that would be related to the kind of condition, the kind of images, would definitely help as mentioned" (AI Expert - P2).*

> *"If they could just cherry pick the relevant stuff, it would it would make a big difference for sure" (Radiologist - P1).*

> *"Yeah, if I could snap my fingers and get whatever I wanted, I would want all the clinical information that I could regarding particular study in a particular patient at the time. Then I would certainly love to have some differential diagnostic assistance" (Radiologist - P5).*

> *"If you can have automated relevant summaries of the prior reports that would be helpful for us to decide which reports to read in detail and what things to focus on… a succinct summary of prior reports would be helpful" (Radiologist - P3).*

> *"if everybody used a standardized structure for their reports and adapted to that, that would make things a lot easier for this sort of analysis" (Radiologist – P3).*

> *"We have radiologists that just describe everything but never give any differentials and other radiologists that list five or ten things so they do need to be prioritized. We could certainly use assistance with that" (Radiologist - P5).*

> *"A report could be verified through the use of AI and corrections can be made before the report is released" (Radiologist - P5).*

> "*We have tremendous problems with voice recognition errors; for example, little decimal points can make a big difference as we've found out" (Radiologist - P5).*

Subtheme: Change Analysis

Spine disorders are often insidious, progressing without obvious signs or symptoms. Early detection and timely intervention is required to improve the potential for good therapeutic outcomes. Successful spine care requires adequate detection, characterization, and monitoring of spine disorders. Pathology given priority included bone marrow abnormalities, spinal cord compression, vertebral deformities, intervertebral disc pathology, degenerative disorders, and fractures. The surveillance of small seemingly insignificant pathology could prove to be important. Quantitative approaches such as radiomic measures could be developed to help monitor non-visible and subtle structural changes in spine pathology. AI-supported change analysis could also be used to differentiate an incidental finding from significant early stage pathology. Anatomic, statistical, and textural features of pathology could be automatically compared over time to assess disease progression and/or treatment outcome. This can be facilitated by temporal subtraction methods applied to successive spine imaging studies. Voxel-wise analysis could be implemented to demonstrate change. Successful spine disease surveillance requires consistent imaging protocols, accurate co-registration of tissues, and appropriate implementation of validated change analysis algorithms. Spinal vertebrae offer rigid bone boundaries that can be used as points of reference to help register and co-register spatial relationships for auto segmentation and change analysis. The technologies required for developing these solutions are available. Objective change analysis offers predictive insight and also provides an effective method for assessing treatment outcome.

> *"Yes, I think having some sort of objective finding that we can over time from the previous study might be helpful because a lot of times just eyeballing it, is to subjective" (Radiologist-P3).*

> *"We've all seen these patients that fell through the cracks because of reporting of a small mass a year and a half ago and nobody followed it up" (Radiologist - P5).*

> *"I think radiomics and other things will be able to help us, you know, get more information about the underlying patterns, statistical patterns that are related to different voxel intensities and how they are distributed. (AI Expert - P3).*

> *"I think if we can objectively quantify things like neural foramen stenosis and spinal canal stenosis and compare those quantities over time that might be helpful because a lot of times you know we just make a subjective assessment of how bad the stenosis is. It would be nice to have a reproducible number and more reproducibility of findings" (Radiologist - P3).*

> *"The quicker that we can find injury to the cord, to the nerve root, the quicker we can maybe offset some of the debilitating problems" (Radiologist- P4).*

> *"In cancer there are separate microhabitats evolving on their own and just watching them structurally doesn't necessarily change the treatment whereas if there were different signatures associated with different levels of aggressiveness, that might change the treatment. I don't see how a human can assimilate all that information" (Radiologist - P5).*

> *"It's very important to define an acquisition protocol"* to monitor spine pathology *(AI-Expert - P4).*

> *". . .the spine is one of the easiest parts of the body to co-register because the vertebrae are very, I would say very rigid, the bone is seen very well on each scanner assessment" (AI Expert - P2)."*

Subtheme: Prioritization

The radiologist has limited time and it must be used wisely. AI can be used to help prioritize the interpretive process and to allocate a spine imaging study to a particular level or path of interpretation. This includes prioritizing the worklist, as well as prioritizing the interpretation of specific images or pathology. NLP could be combined with pre-interpretive image analysis to enhance the prioritization process. Diagnostic imaging provides real time information about pathology, whereas medical records refer to prior findings. Subsequently, a high level of relevance should be assigned to current imaging results for prioritization. Molecular measures and radiomic methods could be implemented to enhance the screening and prioritization process. AI methods could be further developed to perform an automated screening of spinal and extra spinal tissues on studies prior to or paralleling visual interpretation. Screening could be performed to locate and label spine injury features such as cortical disruption, dislocations, fractures, and the presence of edema within ligamentous complexes. AI could also be used to identify distinguishing features of aggressive or high-risk spine pathology that require immediate attention.

> *You can't focus your attention on one thing. In radiology, you've got to be really out there looking at everything" (Radiologist - P1).*
>
> *"Well, I think prioritizing would be an advantage and even identify a little hint of what was in the past and what you're looking for a follow up study, sure" (Radiologist - P2).*
>
> *"Radiomics can give us some kind of insights that cannot be appreciated with the human eye because we cannot interpret or define the statistical appearance." (AI Expert - P3).*
>
> *"I think the image is still probably a better source of information. But I think it could be complemented by NLP. I think the other way around is maybe a little bit less likely because of the incompleteness of what's in the EMR" (AI Expert-P2).*
>
> *"I do think AI is going to add the icing on the cake, sort of like mammography where you press the button and then the arrow goes, hey did you look at that, that type of thing" (Radiologist - P2.)*
>
> *"AI could highlight wherever there's a cortical disruption and then bring that up to the top of the list so that study gets looked at first" (Radiologist - P3).*

Subtheme: Immersive Data Display

The spine is intricate and complex. How data is displayed can influence the accuracy and efficiency of the interpretive process. Two-dimensional (2D) views of spine pathology are often insufficient for a precise diagnosis and personalized treatment planning. The evaluation of spine pathology in three-dimensional (3D) space would offer a more comprehensive perspective of pathology than 2D assessment. A multidimensional display of data would give the radiologist the opportunity to better appreciate the spatial relationships of pathology. It could help reveal atypical or anomalous structural relationships and be used to identify boundaries or transitional zones between normal and diseased tissue. Molecular and radiomic features of pathology could be integrated with or mapped onto 3D images allowing for volumetric characterization of pathology. Voxel-wise biomarkers could be mapped on 2D or 3D renderings of pathology and color-coded to enhance the interpretive process. All of these approaches could be used to enhance or support the digital “virtual” biopsy. Virtual reality (VR) and augmented reality (AR) have the potential to improve digital multidimensional exploration of pathology and to help guide invasive diagnostic and interventional approaches.

> *“What comes to mind here again is the need to look at things in open 3D space. Because when you use 2D views, for example, you can only go through the displays in orthogonal directions” (AI Expert - P1).*

> *“You can’t focus your attention on one thing. In radiology, you’ve got to be really out there looking at everything” (Radiologist - P1).*

> *“I think radiomics and other things will be able to help us, you know, get more information about the underlying patterns, statistical patterns that are related to different voxel intensities and how they are distributed. (AI Expert - P3).*

> *“So if you’re looking at things in open 3D space, then you can kind of swim through the object and find what I call key bookmark views, the key places to really analyze, and then apply the AI and radiomics to the key views, which can really give you significant directions moving forward.” (AI Expert - P1)*

> *“It would be terrific, of course, if AI and radiomics, etcetera, are applied to volumes” (AI Expert - P1).*

> *“voxel-based biomarker information can be perfectly plotted in the kind of environment you’re suggesting” (AI Expert - P3).*

Theme 2: Population-Based Decision Support

Population-based decision support in the context of this study refers to the use of data and/or knowledge stored in a database about patients with similar backgrounds, histories, comorbidities, and/or disease states. This approach is used to help assign probability to differential diagnostic considerations and to formulate a prognosis. Population-based decision support is knowledge and model driven. In addition to assisting the diagnostic process, it can be used to help identify treatment options for a precise and personalized diagnosis. Patient information can be integrated with population-based knowledge during radiology workflow to help classify disease, predict disease progression, and evaluate treatment outcome. Subthemes of population-based decision support include ground truth and knowledge database.

Subtheme: Ground Truth

Truth represents a verifiable fact or set of facts derived through correlative methods. Ground truth represents fundamental facts required make complex observations and decisions. An accurate differential diagnostic process requires truthful decision-support. Ground truth in radiology can be established through the correlation of imaging findings with other sources of data representing pathology. Potential sources include clinical, histologic, laboratory, and genetic workups. Ground truth may also be achieved through correlation with other imaging findings (in vivo) or from computational models of disease (in silico). Ground truth is required to assign relevance to molecular signatures, radiomic features, and other imaging biomarkers of pathology. The differential diagnostic process and accurate staging of spine pathology could be improved through expanded knowledge of the correlative relationships between biological states and AI-derived imaging biomarkers. This a laborious process which requires validation. Better application of ground truth and quantitative imaging measures will improve the clinical utility of the radiology report. The convergence of pathology and radiology combined with the use of structured reporting and standardized disease classifications will help establish ground truth for AI training and testing in spine imaging.

> *"The key thing is to have a source of truth of your training data" (AI Expert - P2).*

> *"in terms of the applications, you have to think about what is the ground truth that I'm using to train my data? That's key" (AI Expert - P2).*

> *"We have to address the possibility that getting all this additional data from the imaging is actually something that's useful and will affect the outcome" (Radiologist - P3).*

> *"In theory, given enough images and outcomes you could have some sort of a ground truth… but it's always difficult (AI Expert - P2).*

> *"That's one of the reasons a lot of people are pushing structured reporting so that if everybody used a standardized structure for their reports and adapted to that, that would make things a lot easier for this sort of analysis" (Radiologist - P3).*

Subtheme: Knowledge Database

The individual radiologist brings limited experience and knowledge to the differential diagnostic process. The success of AI-based decision support in spine imaging is highly dependent on access to knowledge and data analysis methods. A knowledge database refers to a virtual or real platform used to transform integrated structured and unstructured data from different sources into actionable intelligence for problem-solving. It helps convert big data into big insights. Knowledge derived from population data should include the variability required for disease model training. Knowledge can be acquired from many sources including omics-based disease models. The success of AI use in spine imaging will be dependent on having access to an adequate and evolving knowledge database. The development and integration of knowledge databases in other fields has proven to enhance the role of AI decision support. The radiologist would benefit from on-demand access to knowledge and decision support during the interpretive stage of spine imaging.

> *"The bigger the pool of information, the better we're going to be" (Radiologist - P2).*

> *"You need to build a very robust dataset. And when I say, robust, I mean a dataset that is representative of the variability of your problem" to help ensure the success of AI-based decision support (AI Expert - P3).*

> *". . .the most trustworthy approach nowadays is having variability represented in the dataset that you will use to train your models" (AI-Expert - P3).*

"Just look at a DNA analysis, how, it started kind of slowly and now once they developed these large databases it's just advanced by leaps and bounds, and if you can do this with AI, we would all be very grateful" (Radiologist - P2).

Theme 3 Application-Based Decision Support

Application-based decision support in the context of this study refers to the use of structured processes and technological solutions to overcome challenges and solve problems during radiology workflow. This approach enhances the role of the radiologist by providing access to data, knowledge, and AI applications to improve the clinical utility of spine imaging data. This form of decision support encompasses technology and workflow attributes such as interoperability, ease-of-use, and benefits of use. The co-evolution and integration of AI and related technologies supports, as well as facilitate the interdependence of patient and population-based decision-support. Subthemes of application-based decision support include clinical utility and technology attributes.

Subtheme: Clinical Utility

Improved decision-support during the interpretive stage of spine imaging would enhance the potential for better spine care outcomes. Clinical utility refers to the usefulness and potential benefits of a technology, process or intervention in patient care. Spine disorders considered a high priority for AI development and use during spine imaging workflow include spinal cord pathology, spine tumors, bone marrow disorders, intervertebral disc pathology, fractures, and spine pain syndromes. Each of these spine disorders is prevalent and requires early detection and intervention to avoid poor clinical outcomes. Meaningful utility of AI applications in spine imaging is dependent on their ability to reliably improve the detection and characterization of spine pathology, as well as help predict outcomes. This includes the ability to reveal new radiomics biomarkers or molecular signatures which can be used to better classify and stage pathology. Consistent use of disease criteria and related terminology is required for AI to achieve widespread clinical utility. AI methods can be used to auto label anatomic structures of the spine and to identify regions of abnormality during automated pre-interpretive analysis of imaging data.

"Wouldn't it be nice to know that there is an abnormality, even if we can't visualize it…?" (Radiologist – P2).

"We have to address the possibility that getting all this additional data from the imaging is actually something that's useful and will affect the outcome" (Radiologist - P3).

"The quicker that we can find injury to the cord, to the nerve root, the quicker we can maybe offset some of the debilitating problems that are ongoing after the surgery or after conservative care" (Radiologist - P4).

"One thing is very common as you see extruded discs that are sitting in the canal and you always wonder, well, is it a disk or is it a tumor. So I think that AI would be very helpful" (Radiologist-P1).

"with the aging patient population, we're seeing more metastatic disease and … there's incidents of multiple myeloma, which is sometimes really tough to identify" (Radiologist - P2).

"radiomic analysis in order to predict if a new patient is going to suffer this kind of fracture" (AI Expert - P4).

"But in theory, a deep learning network could help establish relationships between imaging and pain. But I think it's a much bigger project, certainly in terms of the impact to society" (AI Expert - P2).

"You obviously have to call things like you see them but I think we could really use some consistency and I think AI may provide that" (Radiologist - P5).

"I'm happy to hear people say that the radiologist will still be involved, but I do think AI is going to add the icing on the cake" (Radiologist - P2).

Subtheme: Technology Attributes

There are many factors which influence whether new technologies such as AI are adopted, used, and supported. Perceived usefulness will be an important driver of AI adoption during the interpretive stage of spine imaging workflow. Knowledge of clinical utility is critical to AI adoption and use. The theoretical construct ease-of-use will also play an important role in the decision to implement AI solutions. Successful adoption requires that AI applications be interoperable with existing spine imaging workflow. It must be seamlessly woven into the fabric of spine imaging workflow. Heightened awareness of the relative advantages of AI solutions, including clinical utility, will also have a significant impact on whether the technology is adopted. To be successful AI applications must have a positive impact on patient care which could not be achieved without its use. Publication of best practice models representing augmentation of the role of the radiologist during the differential diagnostic process in spine imaging will support further research and development.

"When I see usefulness, I think of clinical efficacy" (AI Expert - P1).

"When I see ease of use, I think of workflow" (AI Expert - P1).

"*I will take perceived usefulness as the final goal, but perceived ease of use is the train that is going to bring you to this goal. If you don't have both, you're dead" (AI Expert – P3).*

"You're adding value to the clinical workflow. But if you offer usefulness and you don't address perceived ease of use, you are dead" (AI Expert - P3).

"Your application must be seamless. So thankfully, AI is very good at this. It's very good at automatic procedures" (AI Expert - P3).

"AI needs to be perfectly integrated in the workflow of the radiologist" (AI Expert - P3).

"We can use all the help we can get my opinion" highlights the importance of addressing new solutions. (Radiologist - P5).

" At every level there has to be education of what to do" (Radiologist - P1)

RESEARCH PARTICIPANT RESPONSE SURVEY

Research Study Title: Artificial Intelligence: Potential Impact on Spine Imaging Interpretation and
Diagnosis
Researcher: Dr. David H. Durrant
Research Participant:
Focus Group Session: Radiology / AI Expert

SURVEY

Instructions: Please review the document titled "Preliminary Thematic Summary" and complete the survey below. Place an "x" next to all responses which apply. Provide clarification and comments where appropriate.

___ The results of thematic analysis reflect opinions offered during the focus group session.

If you did not place an "x" next to the statement above, please clarify.

___ The focus group session quotes help support the results of thematic analysis.

If you did not place an "x" next to the statement above, please clarify.

___ I agree with the results of thematic analysis.

If you did not place an "x" next to the statement above, please clarify.

General Comments (Optional):