Clinical Decision Support Systems in Orthodontics: A Narrative Review of Data Science Approaches

Najla N. Al Turkestani, Jonas Bianchi, Romain Deleat-Besson, Celia Le, Li Tengfei, Juan Carlos Prieto, Marcela Gurgel, Antonio C.O. Ruellas, Camila Massaro, Aron Aliaga Del Castillo, Karine Evangelista, Marilia Yatabe, Erika Benavides, Fabiana Soki, Winston Zhang, Kayvan Najarian, Jonathan Gryak, Martin Styner, Jean-Christophe Fillion-Robin, Beatriz Paniagua, Reza Soroushmehr, Lucia H.S. Cevidanes

Najla N. Al Turkestani

DDS, MS; PhD Student

Department of Orthodontics and Pediatric Dentistry, School of Dentistry, University of Michigan, Ann Arbor, MI, United States

Department of Restorative and Aesthetic Dentistry, Faculty of Dentistry, King Abdulaziz University, Jeddah, Saudi Arabia

Jonas Bianchi

DDS, MS, Ph.D.; Assistant Professor

Department of Orthodontics and Pediatric Dentistry, School of Dentistry, University of Michigan, Ann Arbor, MI, United States

Department of Orthodontics, University of the Pacific, Arthur A. Dugoni School of Dentistry, San Francisco, CA, USA.

Romain Deleat-Besson

BS; Research Fellow

Department of Orthodontics and Pediatric Dentistry, School of Dentistry, University of Michigan, Ann Arbor, MI, United States

Celia Le

BS; Research Fellow

Department of Orthodontics and Pediatric Dentistry, School of Dentistry, University of Michigan, Ann Arbor, MI, United States

Li Tengfei

Ph.D., Assistant Professor

Department of Biostatistics, University of North Carolina, Chapel Hill, NC, United States

Juan Carlos Prieto

Ph.D., Assistant Professor

Department of Psychiatry, University of North Carolina, Chapel Hill, NC, United States

Marcela Gurgel

DDS, MS; PhD Student

Department of Orthodontics and Pediatric Dentistry, School of Dentistry, University of Michigan, Ann Arbor, MI, USA.

Antonio C. O. Ruellas

DDS, MS, PhD; Professor

Department of Orthodontics, School of Dentistry, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

Department of Orthodontics and Pediatric Dentistry, School of Dentistry, University of Michigan, Ann Arbor, MI, USA.

Camila Massaro

DDS, MS, PhD

Department of Orthodontics, Bauru Dental School, University of São Paulo, São Paulo, Brazil

Department of Orthodontics and Pediatric Dentistry, School of Dentistry, University of Michigan, Ann Arbor, MI, USA.

Aron Aliaga Del Castillo

DDS, MS, PhD

Department of Orthodontics, Bauru Dental School, University of São Paulo, São Paulo, Brazil

Department of Orthodontics and Pediatric Dentistry, School of Dentistry, University of Michigan, Ann Arbor, MI, USA.

Karine Evangelista

DDS, MS, PhD; Assistant Professor

Department of Orthodontics, School of Dentistry, University of Goias, Goiania, Brazil

Department of Orthodontics and Pediatric Dentistry, School of Dentistry, University of Michigan, Ann Arbor, MI, USA.

Marilia S. Yatabe

DDS, MS, PhD; Clinical Assistant Professor

Department of Orthodontics and Pediatric Dentistry, School of Dentistry, University of Michigan, Ann Arbor, MI, United States

Erika Benavides

DDS, MS, PhD; Clinical Professor

Department of Periodontics and Oral Medicine, School of Dentistry, University of Michigan, Ann Arbor, MI, United States

Fabiana Soki

DDS, MS, PhD; Clinical Assistant Professor

Department of Periodontics and Oral Medicine, School of Dentistry, University of Michigan, Ann Arbor, MI, United States

Winston Zhang

MS; Ph.D. Student

Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, United States

Kayvan Najarian

Ph.D.; Professor

Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, United States

Jonathan Gryak

Ph.D.; Assistant Research Scientist

Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, United States

Martin Styner

Ph.D.; Associate Professor

Departments Psychiatry and Computer Science, University of North Carolina, Chapel Hill, NC, United States

Jean-Christophe Fillion-Robin

BS, MS; Principal Engineer

Kitware Inc., Carrboro, NC, United States

Beatriz Paniagua

BS, MS, Ph.D.; Assistant Director of the Medical Computing

Kitware Inc., Carrboro, NC, United States

Reza Soroushmehr

Ph.D., Assistant Professor

Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, United States

Lucia H. S. Cevidanes

DDS, MS, PhD; Associate Professor

Thomas and Doris Graber Endowed Professor of Dentistry

Department of Orthodontics and Pediatric Dentistry, School of Dentistry, University of

Michigan, Ann Arbor, MI, United States

Corresponding Author

Najla Al Turkestani DDS, MS

University of Michigan School of Dentistry

1011 North University Avenue

Ann Arbor, MI 48109 -1078

United States

Telephone: (734) 936 6282

E-mail: alnajla@umich.edu

Authors' Contribution

Najla N. Al Turkestani: Formal analysis, Investigation, Methodology, Writing original draft. **Jonas Bianchi, Marcela Gurgel, Camila Massaro, Aron Aliaga Del Castillo, Karine Evangelista, Marilia Yatabe:** Writing – review & editing. **Romain Deleat-Besson, Celia Le, Juan Carlos Prieto, Winston Zhang, Kayvan Najarian, Jonathan Gryak, Martin Styner, Jean-Christophe Fillion-Robin, Beatriz Paniagua, Reza Soroushmehr:** software engineering, Writing – review & editing. **Li Tengfei**: statistician, Writing – review & editing. **Antonio Ruellas:** Methodology, Writing – review & editing. **Erika Benavides, Fabiana;** radiological image interpretation , Writing – review & editing. **Lucia Cevidanes:** Conceptualization, Formal analysis, Investigation, Methodology, Writing – review & editing.

Conflict of Interest Statement

The authors declare no potential conflicts of interest with respect to the authorship and/or publication of this article.

Data Availability Statement

The data that support the findings of this study are available upon reasonable request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

DR NAJLA AL TURKESTANI (Orcid ID : 0000-0002-7650-3638)

DR JONAS BIANCHI (Orcid ID : 0000-0002-3749-0918)

DR CAMILA MASSARO (Orcid ID : 0000-0001-6011-7946)

DR ARON ALIAGA-DEL CASTILLO (Orcid ID : 0000-0003-3963-1742)

DR LUCIA CEVIDANES (Orcid ID : 0000-0001-9786-2253)

Article type   : Review Article

Clinical Decision Support Systems in Orthodontics: A Narrative Review of Data Science Approaches

Abstract

Advancements in technology and data collection generated immense amounts of information from various sources such as health records, clinical examination, imaging, medical devices, as well as experimental and biological data. Proper management and analysis of these data via high-end computing solutions, artificial intelligence and machine learning approaches can assist in extracting meaningful information that enhances population health and well-being. Furthermore, the extracted knowledge can provide new avenues for modern healthcare delivery via clinical decision support systems. This manuscript presents a narrative review of data science approaches for clinical decision support systems in orthodontics. We describe the fundamental components of data science approaches including (1) Data collection, storage and management; (2) Data processing; (3) In-depth data analysis; and (4) Data communication. Then, we introduce a web-based data management platform, the Data Storage for Computation and Integration, for temporomandibular joint and dental clinical decision support systems.

Keywords: Artificial Intelligence; Decision Support Systems; Orthodontics; Machine Learning.

1.Introduction

Over the past decades, the digital revolution has transformed every facet of our world, including dental practice. Adopting modern technology in orthodontics has not only provided us with advanced diagnostic and treatment tools, but also given us the chance to shift clinical practice from a "malocclusion-centered" to a "patient-centered" model.[1] Establishing optimum and personalized orthodontic care requires (1) analysis of large and complex data sets derived from different sources, such as clinical examination, diagnostic images, biological and genetic data[2-4], and (2) identification of patterns/associations that turn the individual's big data into knowledge for precise decisions and outcomes prediction.[5] Big data analysis is associated with numerous challenges that necessitate the use of high-end computing solutions, advanced analytical methods[6] (e.g. artificial intelligence and machine learning algorithms) and data science approaches before communicating the acquired knowledge to healthcare providers via decision support systems.[3]

Clinical Decision Support Systems (CDSSs) are computer programs developed to provide expert support for healthcare providers in making decisions regarding prevention, diagnosis and treatment of health diseases.[7] Lusted and Ledley were pioneers in explaining how the reasoning behind the foundation of diagnosis and treatment in medicine can be investigated and solved accurately using mathematical tools.[8] Since then, researchers have been using different methods to supply clinical applications with knowledge. In orthodontics, CDSSs are being developed to reduce subjectivity, decrease errors, save time and increase the efficiency of diagnosis and treatment planning among clinicians.[9] Examples of CDSSs include systems that aid in detecting cephalometric landmarks, determining the need for extractions, identifying the maturation stage of cervical vertebrae and predicting facial soft tissue changes following treatment.[10] Over the last 5 years, our clinical research team, the Dental and Craniofacial Bionetwork for Image Analysis (DCBIA), has addressed knowledge gaps in dentistry that require CDSSs (Figure 1). In this manuscript, we present a narrative review of the data science approaches that are required to develop clinical decision support systems in orthodontics. We also describe a web-based data management platform for the temporomandibular joint and the dental clinical decision support systems.

2. Development of clinical decision support systems (CDSSs)

2.1 Types of CDSSs

CDSSs have been categorized according to: (1) function of the support: provision of alerts (active) or reaction to patients' information/providers' inputs (passive); (2) time of support delivery: before, during or following decision making; and (3) method of development: knowledge or non-knowledge based.

Knowledge-based CDSSs contain compiled data entered directly by users or extracted from patients' electronic records, phone apps[12] and data about medications[13] or clinical protocols and guidelines based on the intent of using a clinical decision system.[4,13] Such systems will deliver support to users, usually following the IF-THEN rule. For instance, in a system for detection of medication interactions, the rule will be IF drug X and IF drug Y are used, THEN alert the clinician.[13] Non-knowledge based CDSSs also require a data source; however, decisions are made with statistical pattern recognition or machine learning (ML) approaches. Consequently, computers learn from previous experiences, discover patterns within the data and eliminate the need for expert inputs or rules. Artificial neural networks and genetic algorithms compose the main types of non-knowledge based CDSSs.[14-15]

2.2 Data science approaches for the development of CDSSs

CDSSs development is a highly challenging and multidisciplinary task that integrates clinical knowledge with decision science to adapt clinical practice and workflow with the decision system.[16] Here, we provide a simplified overview of steps involved in creating a clinical decision support system (Figure 2A).

2.2.1 Data collection, storage and management

The advancement of information technology promoted exponential growth of health data and generation of big data[16]; i.e. large volume data that is produced at high speed and integrates different types of data, such as: (a) clinical data (e.g. orthodontic diagnosis and treatment progress notes, imaging, health records), (b) omics data (e.g. genomics, proteomics, metabolomics), (c) patient-generated data (e.g. wearable devices and scanners, social media) and (d) normative data sets (e.g. data collected in clinical trials or nationwide surveys).[2,17-18] Current evolution of data capturing, storage and analytical methods will allow us to transform the wealth of knowledge in big data into actionable plans

to overcome challenges in clinical decision, deliver personalized care and improve the population's health.[19-20]

Collection of patient diagnosis and treatment progress data is considered valuable when it is done in a systematic way following inter-linkable and coherent data standards that produce high-quality information.[18-21] Clinical data constitutes an essential resource for medical and health research - electronic health records (EHR) are one of the major types of clinical data.[22] The use of EHR is highly encouraged to improve clinicians' compliance with documentation and enable data sharing among different members of the healthcare team.[17] In 2009, the Health Information Technology for Economic and Clinical Health (HITECH) Act was created to support the adoption and meaningful use of health information technology, including a provision of financial incentives for using EHR systems.[23] By 2015, the adoption of certified EHR programs by hospitals and office-based physicians reached 96% and 78%, respectively.[24-25] Nevertheless, the quality of data within EHR was affected by duplication, missing information, fragmentation and inconsistent organization. The accuracy and reliability of data can influence the development and use of CDSSs, even in the presence of sophisticated advanced technologies. Thus, there is a need to standardize the terminology within the dental and orthodontic fields using well-structured forms and templates to assist in ensuring consistency of the collected data.[17] There have been several governmental efforts to standardize the data within different EHR systems; however, at this time no federal or professional association program has been able to produce universally accepted high-quality data. Indeed, data governance requires policies for care providers and auxiliary staff, hands-on training, a culture of responsibility and the right tools to improve and monitor data quality.[26]

Multicenter data collection is commonly performed to construct clinical prediction models. Although such data structures create additional challenges for data analysis, they cover a broader population and can improve the generalizability of the artificial intelligence models.[27] Digital data repositories provide web-based platforms that enable researchers from multiple institutions to access and manage their data securely. Data repositories are designed to store large amounts of data, ranging from thousands of data set repositories, supported by funding and government agencies, to small data sets, supported by a research team for a certain study.[28]

"BigMouth" is an example of an oral health data repository that contains EHR from 11 dental schools across the United States. It provides access to the demographic, dental, and medical data of

over 3 million patients to facilitate the advancement of research and patient care outcomes.[29-30] Defining and establishing a code of conduct is important for big data collaborations to guide the ethical and meaningful use of shared data. When preparing data for sharing, it is essential to strike a balance between privacy protection (e.g., de-identification & anonymization of patients' data, security measures, controlled data sharing) and maintenance of data utility.[17,31]

2.2.2 Data processing

After defining the proper source for data collection, it is essential to verify the quality of data and prepare a final data set for analysis or machine training.[32]

In 1955, a mathematician named John McCarthy coined the term artificial intelligence (AI) to describe the ability of machines to conduct tasks that lie within the range of intelligent activities. Afterward, Richard Bellman defined AI as the ability to automate activities with human thinking capabilities; for example, problem-solving, learning, and decision making.[33] Machine learning (ML) is a subfield of AI, whereby algorithms are utilized to find structures and patterns within data. Consequently, machines will learn to predict similar patterns on unseen data, and their actions will improve each time new data is introduced without human inputs.[9,34-35]

High-quality data sets are essential for developing effective machine learning models. To build a ML model, three non-overlapping data sets are utilized for training, validation, and testing. The training data set is used to develop the ML model. Then, model parameters are adjusted with the validation data set. The last step involves testing the model performance on unseen data i.e., test data set.[36] In ML, raw data is not usually suitable for learning; features/variables should be identified and extracted from the raw data via data mining. However, raw data should be cleaned and prepared for data mining; noise and errors within raw data might confuse the data mining process and thus result in faulty detection of patterns.[37] The aims of cleaning and enriching data are primarily data normalization, elimination of redundant features, data standardization, removal of duplicates, resolution of inconsistent data, management of missing data, and data matching across multiple sources.[36,38] Images normalization is an example of a data preparation procedure required before applying statistical or ML algorithms. Putting several images in a common statistical distribution based on the size and pixel values (spatial normalization) will enable reliable detection of changes between several individuals or within the same individual at different time points.[39-41] Additionally, normalization of images' intensity should be

performed when planning to develop a model that classifies a disease, to avoid biasing the results, or building a model for image synthesis e.g., segmentation and transformation. Evidence demonstrated the accuracy of the image synthesis model is greatly affected by standardizing the intensity features across the training's input images.[42-43]

Data mining involves analyzing large data sets to extract unknown patterns and comprehensible information from large data sets.[44] Several functionalities or tasks can specify the knowledge found in data mining, such as regression, clustering, and classification.[37] These findings can summarize the input data or be utilized in additional analyses such as machine learning and predictive analytics. Although algorithms are used in data mining and ML, outputs from data mining help optimize decisions e.g. detecting a valuable clinical information will help the practitioner make better decisions and increase the quality of care. However, training a machine with the extracted knowledge will enable predicting the diagnosis or prognosis of a new patient.[9]

Dimensionality reduction is performed before machine training to eliminate irrelevant and redundant data, improve learning accuracy, and enhance output comprehensibility. The main types of dimensionality reduction are feature selection and feature extraction. Feature selection involves selecting data that contains the most relevant information for solving a particular problem; a subset of the original data is maintained and used for machine training.[45] For instance, Bianchi et al.[46], conducted a study aiming to detect temporomandibular joint osteoarthritis (TMJ OA). Only the most robust features were selected for machine training out of the collected 52 clinical, biological, and radiomic markers and 1326 interactions. Similarly, in AI systems for two-dimensional cephalometric analyses, the rates of success of landmark detection and classification of skeletal and dental problems depend on the proper selection of the diagnostic variables.[47] On the other hand, feature extraction transforms the original features into a new, smaller set of more significant features.[45] In the study of Bianchi et al.[46], twenty three-dimensional imaging features of bone texture and bone morphometry, that quantify the initial morphological changes in the condylar trabecular bone, were extracted and used for machine training instead of analyzing the whole Cone-Beam Computed Tomography (CBCT) scan grey level voxels.

After implementing and tailoring the previous steps according to the aim of the ML model, the data is considered ready for advanced analysis and machine learning.

2.2.3 In-depth data analysis

Following data preparation, each clinical application of artificial intelligence in orthodontics requires the selection of the proper machine learning methodology, training the ML model and evaluation of the developed model's performance.

The success of the ML algorithms depends on the thorough comprehension of what algorithms can provide, the limitations of algorithms, and how that will support and fit into clinical care. Hence, communication between data analysts, data scientists, and clinicians is important during all phases of CDSSs development.[48] Within dentistry, different ML algorithms have been utilized based on the size of the data, variables/features to analyze, and the objective of the model. ML can involve unsupervised or supervised learning.[49]

In unsupervised learning, algorithms detect hidden patterns within an unlabeled data set. That means all variables within the training dataset are utilized as inputs, and the machine will automatically discover structures/patterns within that dataset without receiving instructions about the desired outcomes.[49-50] Based on the problem at hand, unsupervised learning algorithms will split the data set into groups (clustering) or find rules representing the relationship between variables within a data set (association).[50] For example, in a study conducted by Auconi et al.[51], combinations of variables (inputs) were provided to the fuzzy cluster, which detected the best phenotypic factors to group a sample of Class- III patients into subjects with increased mandibular dimensions, subjects with increased maxillomandibular divergence, and subjects with intermediate characteristics between the two groups.

Supervised learning algorithms, on the other hand, analyze training datasets with predetermined inputs and outputs. Consequently, the inferred ML model can predict the outputs of new data. Common tasks of supervised learning algorithms include classification and regression.[50] The classification task aims to detect a function (discrete value) that aids in splitting the data set into classes based on various parameters.[52] For example, using a classification algorithm and a training data set consisting of patients' intra-oral and cephalometric findings, the supervised ML model can determine (i.e. classify) the cases that need or do not need a tooth extraction for orthodontic reasons.[53] In the regression task, the correlation between dependent and independent variables is detected during the

machine training, and the developed model can predict continuous variables.[52] For instance, a ML model trained with a regression algorithm can predict dental age (continuous variable) from the pulp-to-tooth ratio of the canines.[54] When training with supervised ML algorithms, it is possible to generate an over-fit model that performs well only on the training data set. Thus, it is essential to evaluate the developed model's generalizability and tune the model parameters through validation methods, e.g., cross-validation.[55] Last, the performance of the model should be evaluated using the test data set. For that, various methods exist and should be selected based on the task of the ML algorithms (classification or regression) and the type of outcomes.[56]

Interestingly, a review conducted by Asiri et al.[9] revealed that most of the ML applications in orthodontics have utilized supervised machine learning algorithms to automate clinical procedures that execute or assist in diagnosis and treatment planning. Radiographs were commonly targeted in developing those ML models as they are considered essential tools for orthodontic diagnosis, treatment planning, and evaluation of treatment outcomes. Discrepancies in landmarks identification have been recognized as a critical source of error in cephalometric analyses. Since the analysis' diagnostic value relies on the reproducibility and precision of landmarks identification, interests to develop an automated approach have increased to reduce the laboriousness of the task and subjectivity of the analysts. Park et al. [57] reported that training the machine with the YOLOv3 algorithm for automatic labeling of 80 landmarks resulted in small error plots and 5% improvements in the accuracy compared to top benchmarks reported in the literature. Additionally, the mean computational time consumed per image was only 0.05 seconds. Similarly, Kunz et al.[58] reported promising results obtained with an AI algorithm that can analyze new cephalometric X-rays with precision comparable to the gold standard, i.e., experienced human examiners. Furthermore, Lee et al.[59] presented an automated framework for cephalometric landmarks detection with the implementation of confidence regions (95%) around the estimated positions of the landmarks. This will allow clinicians to gauge the accuracy of the size and location of the calculated landmarks.

Besides using the two-dimensional radiographs for the cephalometric analysis, accurate measurements can be attained with CBCT imaging modality; CBCT provides an accurate three-dimensional spatial representation of the oral and craniofacial structures. However, the accuracy of manual landmarks plotting on the CBCT requires substantial effort, experience, and time. Gupta et al. reported an automatic knowledge-based landmark detection algorithm able to produce accurate

cephalometric measurements comparable to those computed from manual identification.[60] Automation of the radiographic analysis has also involved attempts to estimate the skeletal maturation that provides the best estimate of the individual's biological age and aid treatment planning.[61] Kashif et al.[62] used a classifier algorithm and developed a tool that can help predict the bone age from hand radiographs with a mean error of .6 years compared to the average reading of two experienced radiologists. Kok et al.[63], on the other hand, reported different algorithms that predict the skeletal maturation from the cervical vertebrae on lateral cephalograms. ML algorithms have also been utilized to develop an automated imaging system that provides objective morphological facial assessment during the orthodontic diagnosis process.[64] Furthermore, various methods for automatic volumetric segmentation of CBCT images were developed using ML algorithms. This will allow objective generation of three-dimensional models for advanced diagnosis and treatment planning while saving the time and efforts required for manual segmentation.[65]

Growth and development of orofacial complex is influenced by interaction of the genetic and environmental factors. Orthodontists are mainly using phenotype-driven diagnostic tools like cephalometric analyses to predict the growth in individuals with Class II or class III skeletal malocclusions. However, future studies of genetics, epigenetics and metabolic pathways that utilize advanced machine learning tools will transform the process of orthodontic diagnosis and treatment planning.[48,66]

2.2.4 Data communication

The next step after developing the ML-based model is to communicate the ML findings with the clinicians. Production of predictive models in a real-world healthcare setting is more challenging than developing models in an experimental environment. Therefore, it is important for clinical experts who were not involved in tools development to test and validate the system's performance. Following the deployment of the ML model, it should be monitored for reliability and correction of errors since clinical protocols and populations are changing over time.[67]

The CDSSs, reviewed in this manuscript, will improve Orthodontic care only if clinicians utilize ML and AI tools to analyze the inter-relationships among the dentition, craniofacial skeleton and soft tissues. Then, translate the acquired knowledge towards advancement of orthodontic diagnosis,

treatment planning, evaluation of growth and development, assessment of treatment progression, outcomes and stability.

3. Data Storage for Computation and Integration

Our clinical research team, the DBCIA[11], has developed a web-based system called Data Storage Computation and Integration (DSCI) for management of data science approaches in TMJ and dental CDSSs. The DSCI allows clinicians and researchers to store and share de-identified data between multiple clinical centers. In addition, it allows data processing, in-depth data analysis with several machine learning algorithms and outcomes communication with the users (Figure 2B).[68] The security and privacy of the access to the DSCI are handled using Jason Web Tokens, with encryption of each user who requests to log in. The DSCI uses Amazon Web Services which enable distributed computing across multi-site clinical centers. Furthermore, the web data management server architecture facilitates scalability and inclusion of plugins or processing pipelines to exploit data sets stored in the web system resources.[69]

3.1 <u>TMJ Clinical Decision Support System</u>

Rationale: Osteoarthritis of the TMJ (TMJ OA) is a chronic debilitating disease that affects millions of people and poses a burden on public health globally.[45] It is a multifactorial disease that results from biological and mechanical events that destabilize normal coupling of synthesis and degradation of the subchondral bone and the articular cartilage.[70] Diagnosis of the TMJ OA is currently based on pre-existent clinical signs and symptoms/imaging markers following the recommendations of the Diagnostic Criteria for Temporomandibular Disorders (DC/TMD).[71] However, several studies showed that clinical diagnosis is poorly correlated with the bony changes in CBCT images. Therefore, new assessment tools are needed for the precision of the diagnoses.[70]

TMJ decision optimization: The CDSS for the TMJ, deployed in the DSCI, enables reliable detection of the TMJ OA and visualization of surface changes of the affected condyles (Figure 3). The TMJ CDSS has components for data storage of biological, clinical and imaging data (e.g., magnetic resonance images, panoramic images, CT and CBCT scans). Furthermore, it possesses a tool for data processing (TMJseg) that allows automatic segmentation of the condyles from the CBCT scans.[72]

Regarding the in-depth analytics component of the CDSS, the DSCI's statistical analysis, cross-validation and machine learning (Light GBM and XGBoost) tools permit users to integrate patient-specific multi-source data and to obtain a diagnosis of the TMJ condition, i.e. healthy or diseased. A recent study by Bianchi et al.[45] showed that the use of the DSCI tools facilitated the diagnosis of TMJ OA in its initial stages with an accuracy of 0.823. Moreover, they demonstrated that the interaction of the biomolecular features has a large contribution to the prediction of the TMJ OA status. Furthermore, the DSCI has a machine learning model that detects changes of the condyles' surfaces (Shape Variation Analyzer) which aids in classifying the TMJ OA disease into different categories based on the degree of the condylar degeneration.[73]

3.2 <u>Dental Clinical Decision Support System</u>

Rationale: Currently, the commercial companies that fabricate clear aligners are utilizing data from digital dental models and applying AI algorithms to predict and plan teeth movement, as well as to undertake teeth segmentation. However, such AI algorithms have not been validated and require caution by clinicians in terms of utilizing the provided predictions as well as monitoring treatments' results.[57] Moreover, these technological advancements also require the integration of multi-source data capture, including clinical information and three-dimensional imaging exams such as CBCT, digital dental models (DDMs), photographs, lateral cephalogram and panoramic x-rays.[67]

Dental decision optimization: The CDSS for dental applications deployed in the DSCI integrates dental crowns' and root canals' relevant clinical information from the DDMs and CBCT scans, respectively (Figure 4). In addition, it provides tools for automatic segmentation of root canals (RootCanalSeg), teeth and gums (DentalModelSeg).[74] Overcoming the registration challenges created by merging the information from different imaging modalities and DDM, a work in progress, will enable reliable quantitative assessments of teeth movement (Figure 4).

The DSCI AI tools incorporate a full spectrum of data science approaches that provide optimized clinical information and promote personalized care delivery across various fields of dentistry.

4. Conclusion

CDSSs incorporate knowledge with patient-specific data to serve clinicians with tools that enhance their clinical decision-making process. Thorough understanding of the steps involved in developing DSS and communication between clinicians, data scientists and analysts are keys to creating successful tools that fit into the clinical workflow.

References

1. Mascitti M, Campisi G. Dental Public Health Landscape: Challenges, Technological Innovation and Opportunities in the 21st Century and COVID-19 Pandemic. *Int J Environ Res Public Health*. 2020;17(10):3636. doi:10.3390/ijerph17103636

2. Finkelstein J, Zhang F, Levitin S, Cappelli D. Using Big Data to Promote Precision Oral Health in the Context of a Learning Healthcare System. *J Public Health Dent*. 2020;80(S1). doi:10.1111/jphd.12354

3. Panahiazar M, Taslimitehrani V, Jadhav A, Pathak J. Empowering Personalized Medicine with Big Data and Semantic Web Technology: Promises, challenges, and Use Cases. *IEEE International Conference on Big Data (Big Data)*, Washington, DC, USA (2014);pp. 790-795, doi: 10.1109/BigData.2014.7004307.

4. Dash S, Shakyawar S, Sharma M, Kaushik S. Big Data in Healthcare: Management, Analysis and Future Prospects. *J Big Data*. 2019;6(1). doi:10.1186/s40537-019-0217-0

5. Pastorino R, De Vito C, Migliara G et al. Benefits and Challenges of Big Data in Healthcare: an Overview of the European Initiatives. *Eur J Public Health*. 2019;29(Supplement_3):23-27. doi:10.1093/eurpub/ckz168

6. Frank M, Drikakis D, Charissis V. Machine-Learning Methods for Computational Science and Engineering. *Computation*. 2020;8(1):15. doi:10.3390/computation8010015

7. Vikram K, Karjodkar F. Decision Support Systems in Dental Decision Making: An Introduction. *Journal of Evidence Based Dental Practice*. 2009;9(2):73-76. doi:10.1016/j.jebdp.2009.03.003

8. Fieschi M, Gouvernet J. Reasoning Foundations of Medical Diagnosis Revisited. *Yearb Med Inform*. 1999;08(01):78-82. doi:10.1055/s-0038-1637913

9. Asiri S, Tadlock L, Schneiderman E, Buschang P. Applications of Artificial Intelligence and Machine Learning in Orthodontics. *APOS Trends in Orthodontics*. 2020;10:17-24. doi:10.25259/apos_117_2019

10. Khanagar S, Al-Ehaideb A, Vishwanathaiah S et al. Scope and Performance of Artificial Intelligence Technology in Orthodontic Diagnosis, Treatment Planning, and Clinical Decision-Making - A Systematic Review. *J Dent Sci*. 2021;16(1):482-492. doi:10.1016/j.jds.2020.05.022

11. Dental and Craniofacial Bionetwork for Image Analysis, DCBIA, https://sites.google.com/a/umich.edu/dentistry-image-computing/, accessed on Mar 11th, 2021

12. Wiharto W. Clinical Decision Support Systems Theory and Practice. *Journal Teknosains*. 2018;7(2):148. doi:10.22146/teknosains.38641

13. Helmons P, Suijkerbuijk B, Nannan Panday P, Kosterink J. Drug-Drug Interaction Checking Assisted by Clinical Decision Support: a Return on Investment Analysis. *Journal of the American Medical Informatics Association*. 2015;22(4):764-772. doi:10.1093/jamia/ocu010

14. Sutton R, Pincock D, Baumgart D, Sadowski D, Fedorak R, Kroeker K. An Overview of Clinical Decision Support Systems: Benefits, Risks, and Strategies for Success. *NPJ Digit Med*. 2020;3(1). doi:10.1038/s41746-020-0221-y

15. Zikos D, DeLellis N. CDSS-RM: A Clinical Decision Support System Reference Model. *BMC Med Res Methodol*. 2018;18(1). doi:10.1186/s12874-018-0587-6

16. Joda T, Waltimo T, Probst-Hensch N, Pauli-Magnus C, Zitzmann N. Health Data in Dentistry: An Attempt to Master the Digital Challenge. *Public Health Genomics*. 2019;22(1-2):1-7. doi:10.1159/000501643

17. Raghupathi W, Raghupathi V. Big Data Analytics in Healthcare: Promise and Potential. *Health Inf Sci Syst*. 2014;2(1). doi:10.1186/2047-2501-2-3

18. Nanayakkara S, Zhou X, Spallek H. Impact of Big Data on Oral Health Outcomes. *Oral Dis*. 2018;25(5):1245-1252. doi:10.1111/odi.13007

19. Lu J, Liu A, Song Y, Zhang G. Data-Driven Decision Support Under Concept Drift in Streamed Big Data. *Complex & Intelligent Systems*. 2019;6(1):157-163. doi:10.1007/s40747-019-00124-4

20. Joda T, Waltimo T, Pauli-Magnus C, Probst-Hensch N, Zitzmann N. Population-Based Linkage of Big Data in Dental Research. *Int J Environ Res Public Health*. 2018;15(11):2357. doi:10.3390/ijerph15112357

21. Library Guides: Data Resources in the Health Sciences: Clinical Data. Guides.lib.uw.edu. https://guides.lib.uw.edu/hsl/data/findclin. Published 2021. Accessed February 26, 2021.

22. Cohen M. Impact of the HITECH Financial Incentives on EHR Adoption in Small, Physician-Owned Practices. *Int J Med Inform*. 2016;94:143-154. doi:10.1016/j.ijmedinf.2016.06.017

23. Henry J, Pylypchuk Y, Searcy T, Patel V. *Adoption of Electronic Health Record Systems Among U.S. Non-Federal Acute Care Hospitals: 2008–2015.* Washington, DC: Office of the National Coordinator for Health Information Technology; May, 2016

24. Jamoom E, Yang N. *Table of Electronic Health Record Adoption and Use Among Office-Based Physicians in the U.S. by State: 2015 National Electronic Health Records Survey.* 2016

25. Electronic Health Record Data Governance and Data Quality in the Real World. HIMSS. https://www.himss.org/resources/electronic-health-record-data-governance-and-data-quality-real-world. Published 2021. Accessed March 3, 2021.

26. Wynants L, Kent D, Timmerman D, Lundquist C, Van Calster B. Untapped Potential of Multicenter Studies: A Review of Cardiovascular Risk Prediction Models Revealed Inappropriate Analyses and Wide Variation in Reporting. *Diagn Progn Res*. 2019;3(1). doi:10.1186/s41512-019-0046-9

27. Perazzo J, Rodriguez M, Currie J, Salata R, Webel A. Creation of Data Repositories to Advance Nursing Science. *West J Nurs Res*. 2017;41(1):78-95. doi:10.1177/0193945917749481

28. Favaretto M, Shaw D, De Clercq E, Joda T, Elger B. Big Data and Digitalization in Dentistry: A Systematic Review of the Ethical Issues. *Int J Environ Res Public Health*. 2020;17(7):2495. doi:10.3390/ijerph17072495

29. Johnny Phan J. BigMouth Dental Data Repository. Bigmouth.uth.edu.https://bigmouth.uth.edu/. Published 2021. Accessed February 26, 2021.

30. Tucker K, Branson J, Dilleen M et al. Protecting Patient Privacy When Sharing Patient-Level Data from Clinical Trials. *BMC Med Res Methodol*. 2016;16(S1). doi:10.1186/s12874-016-0169-4

31. de Dumast P, Mirabel C, Cevidanes L et al. A Web-based System for Neural Network Based Classification in Temporomandibular Joint Osteoarthritis. *Computerized Medical Imaging and Graphics*. 2018;67:45-54. doi:10.1016/j.compmedimag.2018.04.009

32. Brosset S, Dumont M, Cevidanes L, Soroushmehr R, Bianchi J, Gurgel M, Deleat-Besson R, Le C, Ruellas A, Yatabe M, Rosas Liliane, Goncalves J, Najarian K, Gryak J, Paniagua B, Styner M, Prieto JC. Web infrastructure for data management, storage and computation, Proc. SPIE 11600, Medical Imaging 2021: Biomedical Applications in Molecular, Structural, and Functional Imaging, 116001N (14 February 2021); https://doi.org/10.1117/12.2582283

33. Schwendicke F, Samek W, Krois J. Artificial Intelligence in Dentistry: Chances and Challenges. *J Dent Res*. 2020;99(7):769-774. doi:10.1177/0022034520915714

34. Artificial Intelligence: Implications for the Future of Work | | Blogs | CDC. Blogs.cdc.gov. https://blogs.cdc.gov/niosh-science-blog/2019/08/26/ai/. Published 2021. Accessed March 2, 2021.

35. V. Gudivada, A. Apon, and J. Ding. Data Quality Considerations for Big Data and Machine Learning: Going Beyond Data Cleaning and Transformations. *International Journal on Advances in Software* 10.1 (2017), pp. 1 - 20

36. Tallon-Ballesteros A, Riquelme J. Deleting or keeping Outliers for Classifier Training? 2014 Sixth World Congress on Nature and Biologically Inspired Computing (NaBIC 2014). 2014. doi:10.1109/nabic.2014.6921892

37. Al-Jabery K, Obafemi-Ajayi T, Olbricht G, Wunsch D. Computational Learning Approaches to Data Analytics in Biomedical Applications. *Academic Press*, an Imprint of Elsevier, 2020.

38. Moore C. Image Normalization | Radiology Reference Article | Radiopaedia.org. Radiopaedia.org. https://radiopaedia.org/articles/image-normalization?lang=us. Published 2021. Accessed March 6, 2021.

39. Chapter 33 - Spatial Normalisation Using Basis Functions, Editor(s): Richard S.J. Frackowiak, Karl J. Friston, Christopher D. Frith, Raymond J. Dolan, Cathy J. Price, Semir Zeki, John T. Ashburner, William D. Penny,Human Brain Function (Second Edition),*Academic Press*,2004,Pages 655-672,ISBN 9780122648410

40. Ioshida M, Muñoz B, Rios H et al. Accuracy and Reliability of Mandibular Digital Model Registration with Use of the Mucogingival Junction as the Reference. *Oral Surg Oral Med Oral Pathol Oral Radiol*. 2019;127(4):351-360. doi:10.1016/j.oooo.2018.10.003

41. Reinhold J, Dewey B, Carass A, Prince J. Evaluating the Impact of Intensity Normalization on MR Image Synthesis. Medical Imaging 2019: Image Processing. 2019. doi:10.1117/12.2513089

42. Bianchi, J. et al. 3D Slicer Craniomaxillofacial Modules Support Patient-Specific Decision-Making for Personalized Healthcare in Dental Research. Multimodal Learn. Clin. Decis. Support Clin. image-based Proced. 10th Int. Work. ML-CDS 2020, 9th Int. Work. CLIP 2020, held conjunction with MICCAI 2020, Lima, Peru, Oct. 4-8, ... 12445, 44–53 (2020).

43. Data mining. En.wikipedia.org. https://en.wikipedia.org/wiki/Data_mining. Published 2021. Accessed March 6, 2021.

44. Liu H. (2011) Feature Selection. In: Sammut C., Webb G.I. (eds) Encyclopedia of Machine Learning. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-30164-8_306

45. Bianchi J, de Oliveira Ruellas A, Gonçalves J et al. Osteoarthritis of the Temporomandibular Joint can be Diagnosed Earlier Using Biomarkers and Machine Learning. *Sci Rep*. 2020;10(1). doi:10.1038/s41598-020-64942-0

46. Hwang HW, Moon JH, Kim MG, Donatelli RE, Lee SJ. Evaluation of automated cephalometric analysis based on the latest deep learning method. Angle Orthod. 2021 Jan 12. doi: 10.2319/021220-100.1

47. Thomas Blanchard D. Machine Learning Algorithms: A Clinician and Data Analyst Partnership. Acumenmd.com. https://acumenmd.com/blog/machine-learning-algorithms-a-clinician-and-data-analyst-partnership/. Published 2021. Accessed March 8, 2021.

48. Allareddy V, Rengasamy Venugopalan S, Nalliah R, Caplin J, Lee M, Allareddy V. Orthodontics in the Era of Big Data Analytics. *Orthod Craniofac Res*. 2019;22(S1):8-13. doi:10.1111/ocr.12279

49. Alloghani M, Al-Jumeily D, Mustafina J, Hussain A, Aljaaf A. A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science. *Unsupervised and Semi-Supervised Learning*. 2019:3-21. doi:10.1007/978-3-030-22475-2_1

50. Auconi P, Scazzocchio M, Cozza P, McNamara J, Franchi L. Prediction of Class III Treatment Outcomes Through Orthodontic Data Mining. *Eur J Orthod*. 2014;37(3):257-267. doi:10.1093/ejo/cju038

51. Regression vs Classification in Machine Learning - Javatpoint. www.javatpoint.com. https://www.javatpoint.com/regression-vs-classification-in-machine-learning. Published 2021. Accessed March 6, 2021.

52. Jung S, Kim T. New Approach for the Diagnosis of Extractions with Neural Network Machine Learning. *American Journal of Orthodontics and Dentofacial Orthopedics*. 2016;149(1):127-133. doi:10.1016/j.ajodo.2015.07.030

53. Farhadian M, Salemi F, Saati S, Nafisi N. Dental Age Estimation Using the Pulp-to-Tooth Ratio in Canines by Neural Networks. *Imaging Sci Dent*. 2019;49(1):19. doi:10.5624/isd.2019.49.1.19

54. Sidey-Gibbons J, Sidey-Gibbons C. Machine Learning in Medicine: A Practical Introduction. *BMC Med Res Methodol*. 2019;19(1). doi:10.1186/s12874-019-0681-4

55. Jiang T, Gradus J, Rosellini A. Supervised Machine Learning: A Brief Primer. *Behav Ther*. 2020;51(5):675-687. doi:10.1016/j.beth.2020.05.002

56. Park J, Hwang H, Moon J et al. Automated Identification of Cephalometric Landmarks: Part 1—Comparisons Between the Latest Deep-Learning Methods YOLOV3 and SSD. *Angle Orthod*. 2019;89(6):903-909. doi:10.2319/022019-127.1

57. Faber J, Faber C, Faber P. Artificial Intelligence in Orthodontics. *APOS Trends in Orthodontics*. 2019;9:201-205. doi:10.25259/apos_123_2019

58. Lee J, Yu H, Kim M, Kim J, Choi J. Automated Cephalometric Landmark Detection with Confidence Regions Using Bayesian Convolutional Neural Networks. *BMC Oral Health*. 2020;20(1). doi:10.1186/s12903-020-01256-7

59. Gupta A, Kharbanda O, Sardana V, Balachandran R, Sardana H. Accuracy of 3D Cephalometric Measurements Based on an Automatic knowledge-based Landmark Detection Algorithm. *Int J Comput Assist Radiol Surg*. 2015;11(7):1297-1309. doi:10.1007/s11548-015-1334-7

60. Buschang P, Roldan S, Tadlock L. Guidelines for assessing the growth and development of orthodontic patients. Semin Orthod. 2017;23(4):321-335. doi:10.1053/j.sodo.2017.07.001

61. Kashif M, Deserno T, Haak D, Jonas S. Feature Description with SIFT, SURF, BRIEF, BRISK, or FREAK? A General Question Answered for Bone Age Assessment. *Comput Biol Med*. 2016;68:67-75. doi:10.1016/j.compbiomed.2015.11.006

62. Kök H, Acilar A, İzgi M. Usage and Comparison of Artificial Intelligence Algorithms for Determination of Growth and Development by Cervical Vertebrae Stages in Orthodontics. *Prog Orthod*. 2019;20(1). doi:10.1186/s40510-019-0295-8

63. Murata S, Lee C, Tanikawa C, Date S. Towards a Fully Automated Diagnostic System for Orthodontic Treatment in Dentistry. 2017 IEEE 13th International Conference on e-Science (e-Science). 2017. doi:10.1109/escience.2017.12

64. Bayirli B, Kim-Berman H, Puntillo A. Embracing Novel Technologies in Dentistry and Orthodontics. Hdl.handle.net. http://hdl.handle.net/2027.42/153991. Published 2021. Accessed March 9, 2021.

65. Varma G, Harsha B, Palla S, Sravan S, Raju J, Rajavardhan K. Genetics in an Orthodontic Perspective. *Journal of Advanced Clinical and Research Insights*. 2019;6(3):86-90. doi:10.15713/ins.jcri.267

66. Wiens J, Saria S, Sendak M et al. Do no harm: a Roadmap for Responsible Machine Learning for Health care. *Nat Med*. 2019;25(9):1337-1340. doi:10.1038/s41591-019-0548-6

67. Li P, Kong D, Tang T et al. Orthodontic Treatment Planning based on Artificial Neural Networks. *Sci Rep*. 2019;9(1). doi:10.1038/s41598-018-38439-w

68. Heinrichs J, Lim J. Integrating Web-based Data Mining Tools with Business Models for Knowledge Management. *Decis Support Syst*. 2003;35(1):103-112. doi:10.1016/s0167-9236(02)00098-2

69. Khanagar S, Al-ehaideb A, Maganur P et al. Developments, Application, and Performance of Artificial Intelligence in Dentistry – A systematic Review. *J Dent Sci*. 2021;16(1):508-522. doi:10.1016/j.jds.2020.06.019

70. Su N, Liu Y, Yang X, Shen J, Wang H. Correlation between oral health-related quality of life and clinical dysfunction index in patients with temporomandibular joint osteoarthritis. *Journal of Oral Science*. 2016;58(4):483-490. doi:10.2334/josnusd.16-0224

71. Schiffman E, Ohrbach R, Truelove E et al. Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) for Clinical and Research Applications: Recommendations of the International RDC/TMD Consortium Network* and Orofacial Pain Special Interest Group†. *J Oral Facial Pain Headache*. 2014;28(1):6-27. doi:10.11607/jop.1151

72. Brosset S, Dumont M, Bianchi J et al. 3D Auto-Segmentation of Mandibular Condyles. *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. 2020. doi:10.1109/embc44109.2020.9175692

73. Tubau N, de Dumast P, Yatabe M et al. Shape variation analyzer: a classifier for temporomandibular joint damaged by osteoarthritis. *Medical Imaging 2019: Computer-Aided Diagnosis*. 2019. doi:10.1117/12.2506018

74. Dumont M, Prieto J, Brosset S et al. Patient Specific Classification of Dental Root Canal and Crown Shape. *Shape in Medical Imaging*. 2020:145-153. doi:10.1007/978-3-030-61056-2_12

Figure Legends

Figure 1. Gaps of knowledge in **(A)** TMJ and **(B)** Dental diagnosis and prognosis assessment tools.
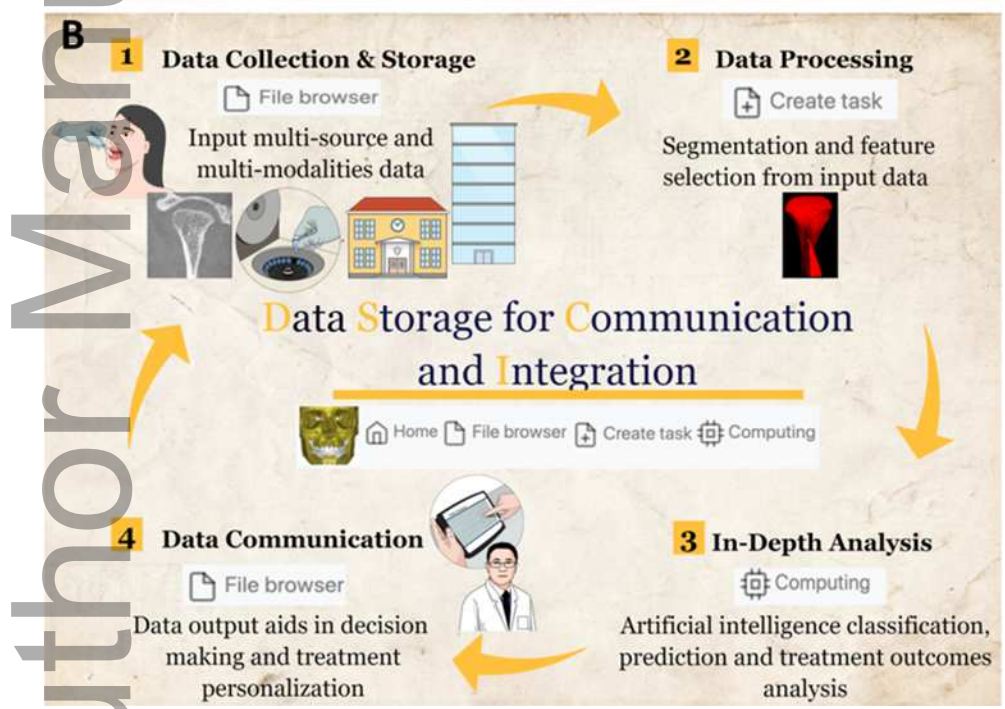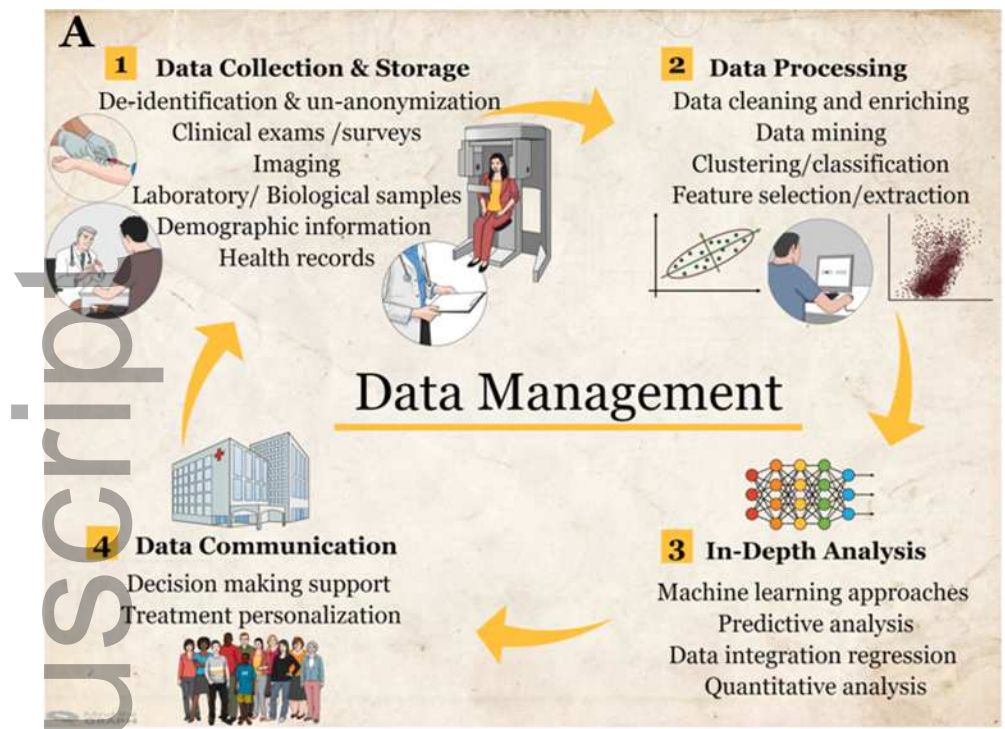
Figure 2. Overview of steps involved in developing clinical decision support systems (CDSSs). **(A)** General spectrum of data science approaches in the CDSSs. **(B)** Implementation of robust data management, Dental Storage for Computation and Integration, in the CDSSs.

Figure 3. Data science approaches in the TMJ clinical decision support system.

Figure 4. Sequence of data science approaches in the dental decision support system.
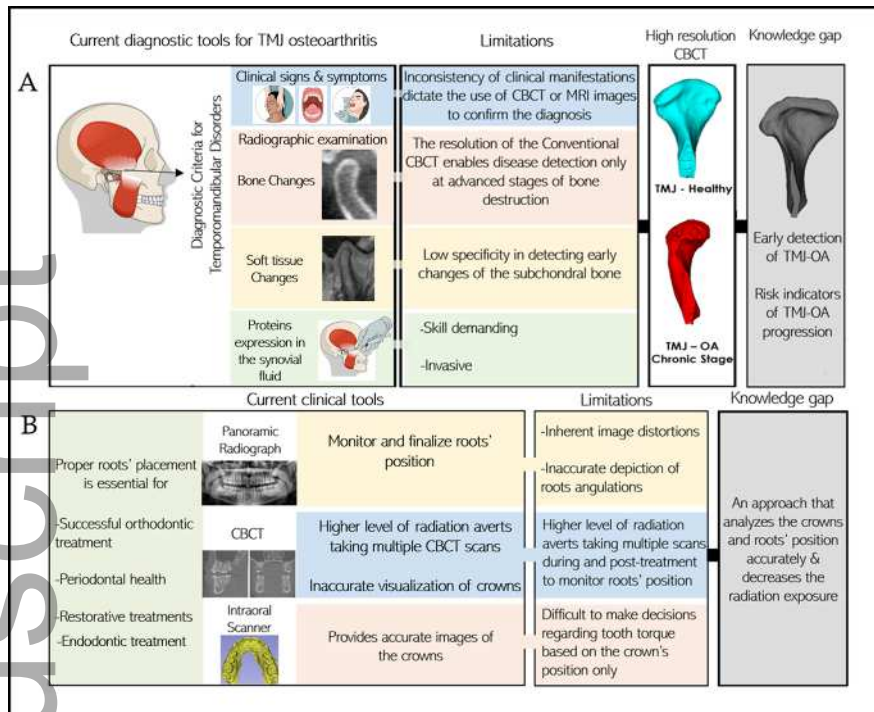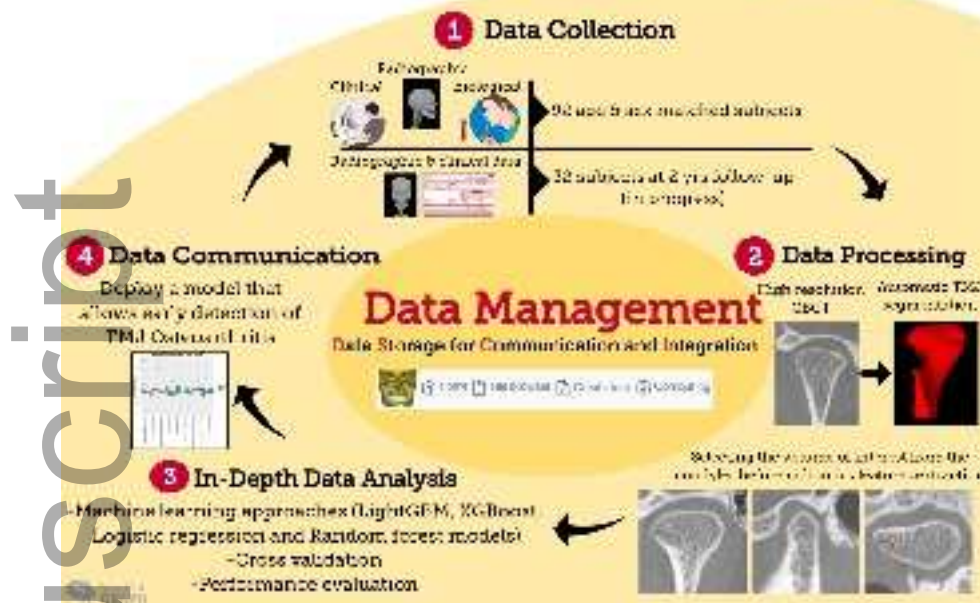
ocr_12492_f1.tif

ocr_12492_f2.tif

ocr_12492_f3.tiff

ocr_12492_f4.tiff