ASDJ
AINSHAMS DENTAL JOURNAL

Print ISSN 1110-7642
Online ISSN 2735-5039

AIN SHAMS DENTAL JOURNAL
Official Publication of Ain Shams Dental School
March 2025 • Vol. 37

Optimizing Data Management: A Cornerstone for AI

Mohammad Nabawy Tawfik[1], Amr Essam Ramah[2], Esraa Ahmed Hassan[2], Sahar Mohamed Sameer[3], Walaa Mohamed Hamed[4,5]

Aim: This study aimed to explore the challenges faced by postgraduate researchers at Ain Shams University in accessing and managing patient data from the Oral and Maxillofacial Radiology (OMFR) department, and to assess the potential impact of implementing a specialized data management software. This can be generalized to all OMFR workers in different locations.
Materials and Methods: A descriptive survey was conducted among 235 postgraduate researchers using Google Forms from August 5 to August 28, 2024. The survey comprised four sections addressing participants' specialties, experiences with OMFR data, and challenges in data collection. Descriptive statistics, including frequencies and percentages, were used to analyze the data.
Results: The majority of participants (83.3%) used CBCT imaging in their research, yet over 65% reported moderate-to-high difficulty in collecting data, with key challenges including tracking patient history (44.4%) and incomplete patient information (33.3%). Concerns over data availability and accessibility deterred 41.3% of respondents from conducting retrospective studies.
Conclusions: The findings highlighted the need for a dedicated data management system to streamline research processes. Respondents valued easy access to patient data and comprehensive datasets, including clinical and radiographic records, which are essential for AI applications in dental diagnostics. Implementing such a system would improve data standardization, enhance research efficiency, and support the effective use of AI models in OMFR.

Keywords: Data Management, Information Storage and Retrieval, Artificial Intelligence, Surveys and Questionnaires, Medical Records.

1. Teaching Assistant of Oral and Maxillofacial Radiology, Faculty of Dentistry, Egyptian Russian University, Egypt.
2. Assistant Lecturer of Oral & Maxillofacial Radiology, Faculty of Dentistry, Ain Shams University, Egypt.
3. Associate Professor of Oral & Maxillofacial Radiology, Faculty of Dentistry, Ain Shams University, Egypt.
4. Professor of Oral & Maxillofacial Radiology, Faculty of Dentistry, Ain Shams University, Egypt.
5. Vice Dean for Community Service and Environmental Development Affairs, Faculty of Dentistry, Ain Shams University
Corresponding author: Mohammad Nabawy Tawfik, email: dr.mohammad.nabawy@gmail.com

Introduction

Artificial Intelligence (AI) is rapidly becoming a cornerstone in modern healthcare, particularly in dental diagnostics. AI's potential to enhance diagnostic accuracy and streamline patient care is immense. A simple example of the crucial role played by the artificial intelligence in the dental field is displayed in the automated cephalometric analysis required for orthodontic treatment. Artificial intelligence-based programs were obviously time saving computing the cephalometric analysis in an extremely short time with highly accepted accuracy. Moreover, the recorded data can be stored, retrieved and used at any time, saving time and space needed for image storage. [1,2] However, the effectiveness of AI systems is directly tied to the quality and quantity of data used to train and validate these models. In Oral and Maxillofacial Radiology, where detailed and accurate patient data are critical, the need for an efficient data management system is paramount.

Before embarking on the development of a data management system, the Oral and Maxillofacial Radiology department at Ain Shams University conducted a comprehensive survey among dental professionals across various specialties, including Oral Radiology, Oral Surgery, Implantology, and Endodontics. The survey aimed to identify the specific challenges and needs related to data management in research and clinical practice.

The survey results highlighted significant obstacles in accessing and managing patient data. Many professionals expressed hesitation in pursuing retrospective studies due to concerns over the large amount of data required, data unavailability, and inaccessibility. Additionally, a considerable number of respondents who have used data from the Oral & Maxillofacial Radiology (OMFR) department at Ain Shams University (ASU) underscored the inefficiencies in current data retrieval methods, which often rely on paper records, spreadsheets, and basic electronic health record systems.

These findings underscored the urgent need for a dedicated data management system that could streamline the collection and retrieval of patient data. In response, the Oral and Maxillofacial Radiology department developed a groundbreaking software application aimed at optimizing patient data management. This software is designed to facilitate the systematic collection of personal data, medical history, and radiological findings from various imaging modalities, including periapical, panoramic, and CBCT images. By providing advanced search and retrieval functions, secure data storage, and easy access to comprehensive patient data, this software addresses the specific needs identified in the survey.

Literature Review

The integration of AI into dental diagnostics is contingent on the availability of high-quality datasets. Previous studies have consistently emphasized the importance of standardized and comprehensive patient databases for the effective training of AI models. [3] However, the lack of efficient data management systems in dental institutions has been a significant barrier. The survey conducted by Ain Shams University revealed that dental professionals are often hesitant to undertake research involving retrospective data due to concerns over the complexity and inefficiency of current data retrieval methods.

Moreover, the literature underscores the critical role of data management in the successful implementation of AI in healthcare. He et al [4] highlights that well-organized, accessible, and secure patient data are essential for developing reliable AI models.[4] The software developed by ASU's Oral and Maxillofacial Radiology department directly addresses these needs by offering

features that align with the priorities identified in the survey, such as easy access to patient data, advanced search capabilities, and secure data management.

The survey further revealed that features like full patient data, including clinical and radiographic findings, and the availability of cases at an affordable cost, are particularly valued by professionals. These insights have informed the development of the software, ensuring that it meets the practical needs of its users while facilitating the efficient collection and analysis of data crucial for AI research.

Research Question

How does the implementation of a specialized software application for patient data management in Oral and Maxillofacial Radiology, informed by user needs through a preliminary survey, impact the effectiveness of AI models in dental diagnostics, particularly regarding data quality, standardization, and research efficiency?

Materials and Methods

This is a descriptive study in the form of a survey proposed by the department of Oral and Maxillofacial Radiology, Faculty of Dentistry, Ain Shams University. This survey aims to assess the difficulties encountered by postgraduate researchers in the process of data collection and retrieval of cases required for their research work. The target population in this study was all postgraduate researchers of Faculty of Dentistry, Ain Shams University. A convenient sampling design was used in this study meaning that no random sampling or assignment was performed.

The survey was designed and published using Google Forms on August 5, 2024, and data collection remained open until August 28. During this period, the survey was announced to the entire population of over 300 postgraduate researchers through the official online pages of postgraduate affairs of the faculty and that of the Oral and Maxillofacial Radiology Department. Participants were informed that their participation was voluntary, and their responses would be kept confidential. The total number of participants was 235.

The survey was divided into four sections: The first section consisted of 5 questions including, single response or multi-select multiple-choice questions and dichotomous questions. All participants were asked to answer the questions of this section and according to the answer to the last question they pass directly to one of the following two sections.

The second section was supposed to be answered only by those who previously used or are currently using our department's data in their research work. It consisted of 5 questions including single response or multi-select multiple-choice questions and a Likert scale question (Very difficult- difficult- moderate- easy- very easy).

Researchers who never used our department's data before automatically bypassed the second section to answer the questions of the third one. It consisted of a single multi-select multiple-choice question that clarifies the reasons for not using our data. The fourth section was to be answered by all participants. It consisted of 2 multi-select multiple-choice questions.

Data Analysis

Data was analyzed in terms of frequencies and percentages using Excel 360 Office ADD-INS real statistics. Descriptive statistics included measurements of frequencies, percentages, means, and standard deviations. A Likert-type scale falls within the ordinal level of measurement. That is, response categories have a rank order, but the intervals between the values cannot be equal.

Results

Questions of the first section:

Question 1: What is your specialty?

This question revealed the percentages of the respondents` specialty in the following descending order: Operative (24.04%), Prosthodontics (22.12%), Periodontics and Orthodontics (16.35%), Pedodontics (13.46%), Oral radiology (10.58%), Oral medicine (5.77%), Oral pathology (3.85%), Dental materials (1.92%), Implantology, Endodontics, Oral surgery, and Oral biology all are equally (0.96%). Obviously, this is not an indication of the need of each of these specialties for the data provided by the oral radiology department, it is just a report for those who participated in this survey randomly.

Question 2: What is your academic affiliation?

With a slight difference in percentages, most of the respondents were not staff members at all (29.81%), followed by those who are staff members in a private university (28.85%), then staff member in Ain Shams University (ASU) (27.88%), and the least percentage was for staff members in another governmental university (13.46%).

Question 3: Are you a post graduate candidate in ASU?

The majority of the respondents were candidates in ASU (85.6%) while (14.4%) were not.

Question 4: Have you ever hesitated about choosing a retrospective study?

41.3% of the respondents chose Yes, for fear of data unavailability, 32.7% chose Yes, for fear of data inaccessibility, 29.8% chose Yes, for fear of large amount of data required, and only 26.9% chose NO.

Question 5: Have you ever used the data of Oral & Maxillofacial Radiology (OMFR) department in your research work?

82.7% of the respondents chose Yes, while 17.3 % chose NO

Questions of the second section:

Question 1: What is/are the type of imaging modality/modalities involved in your study?

Most of the answers were CBCT (83.3%), followed by Periapical (27.8%), then Bitewing and Panorama equally (22.2%), Lateral Cephalometric (16.7%), and finally Extraoral Bitewing chosen by no one (zero%).

Question 2: How many cases did you need to retrieve?

38.89% of the respondents chose (Less than 50), 27.78% chose (100 – 300), 22.22% chose (50 – 100), and 11.11% chose (More than 300)

Question 3: How long did you spend to collect the full set of data required?

1 month and 6 - 12 months were the most commonly chosen answers with 27.78% for each, followed by 3 - 6 months (22.22%), 1 - 3 months (16.67%), and More than 12 months (5.56%).

Question 4: How difficult was your required data collection?

The responses to this question are presented in figure 1. Assuming the following scale: Very Easy= 1, Easy= 2, Moderate= 3, Difficult= 4, & Very Difficult= 5, the descriptive analysis of the data indicates the following:

- Mode=3, indicates that most of the respondents estimated the process of data collection to be of moderate difficulty.
- Mean = 3.3, indicates that the average of the responses is between moderate and difficult.

- Standard deviation= 0.76, indicates very limited variation between the respondent's opinion.
- Maximum = 5 and Minimum = 2, indicates that the choice (Very Easy) was not chosen by any of the respondents.

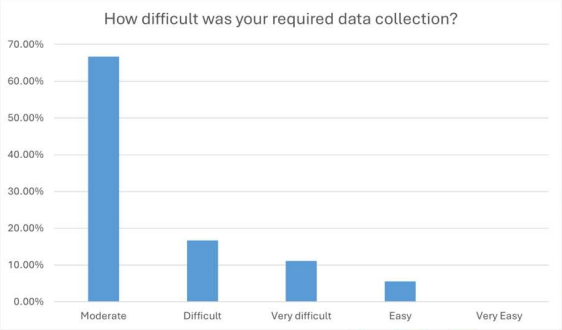

Figure 1: A bar graph showing the percentage of answers to question 4 in section 2

Question 5: What are the most common difficulties you encounter when collecting and managing patient data for your research?

Difficulty in tracking patient history was the most chosen response (44.44%), followed by -Incomplete patient information- (33.33%), then -Lack of integration between different data sources- (11.11%), and finally equal distribution between -MRI criteria along with CBCT finding- and -Time-consuming data entry processes- (5.56% for each). The responses to this question are presented in figure 2.

Questions of the third section:

Why didn`t you use OMFR department`s data before?

58.82% of the respondents chose (I did not need this type of data in my research), 29.41% of them chose (I had no idea about your database and its` availability), 6.86% chose (I tried to use but I faced difficulties in data collection), 2.94% chose (I asked for it but it was not available at that time), and finally 1.96% added another answer saying (I asked for some data but the department refused to help).

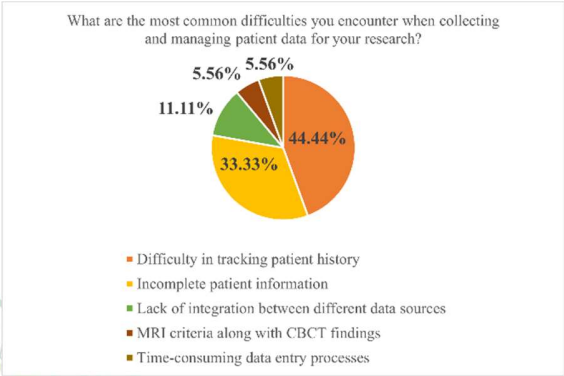

Figure 2: A pie chart showing the percentage of answers to question 5 in section 2

Questions of the fourth section:

Question 1: How do you retrieve the required cases for your research without a dedicated system?

The answers were: Paper records (41.00%), Manual entry into multiple databases (24.00%), Basic electronic health record systems (20.00%), and Spreadsheets (15.00%).

Question 2: If we offer a new service in our department, what features are most valuable for you?

Easy access to patient`s data was the most chosen answer (42.3%), followed by -Full patient`s data including clinical and radiographic findings- (38.5%), then -Availability of cases in an affordable cost- (8.7%), the least chosen responses were -Secure data storage and access controls- and -Advanced search and retrieval functions- (6.87% and 3.8% respectively). The responses to this question are presented in figure 3.

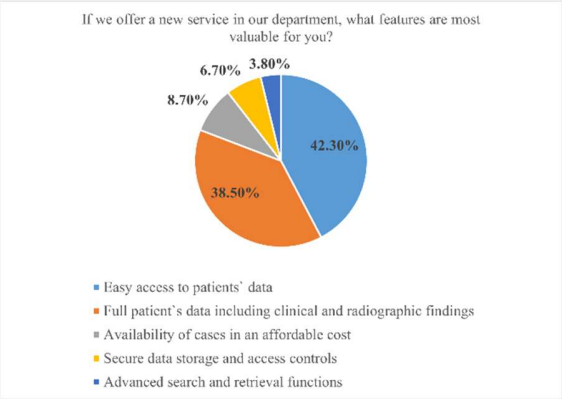

Figure 3: A pie chart showing the percentage of answers to question 2 in section 4

Discussion

Data management is one of the fundamental ethical issues associated with acquiring, assessing and storing data. Dentists should use the collected data only to improve the patients' health, as well as to improve the dental practice. That's why we suggested that our proposed software will have a great role in data collection in dentistry. Moreover, it is unethical to use data in ways that may harm or in any way adversely affect patients.

A convenient sampling design was used in this study meaning that no random sampling or assignment was performed. This type of sampling was used because of the time constraints imposed on this study, the researchers had only three weeks for data collection.

The survey was divided into four sections with a total of 13 close-ended questions in accordance with Roganović J, et al.[5] The first section was mainly related to the participants' specialty, affiliation, and if they used our department's data before. While the second section was mainly measuring the degree of difficulty encountered by the participants during their research work in our department and specifies these difficulties aiming to solve it.

The third part was designed to be answered by those researchers who never used our department's data before and finally the fourth section we were looking for a solution to the problems encountered by the researchers by giving them a chance to select from the answers what could fulfil their needs.

The survey was designed and published using Google Forms on August 5, 2024, and data collection remained open until August 28, when no additional responses were reported for two days. We were in hurry to find out the results so that we could recommend prolonged timing of the survey. The main limitation of this survey is the design of the questions that focused only on the difficulties encountered by the researchers on data collection, without focusing on the requirements they need to face these difficulties. Moreover, the questions design did not allow proper statistical data analysis due to lack of correlation. We suggest that the use of open-ended questions could have provided more information about the requirements needed by the researcher to facilitate data collection and retrieval, which is the main aim of our study.

Regarding our results, more than 65% of the participants agreed that the regular patient data collection was moderately difficult. These difficulties were also confirmed and documented in the book by Carmel D, et al.[6] In our results 44.44% of the responses reported difficulties in tracking the patient's history. As well as 33.33% reported incomplete patient information. These results were matching with the systematic review results on difficulties of diagnosis by Kostopoulou O, et al.[7]

At the end of our survey, we asked about the most convenient feature that will facilitate database collection, the answers were easy access to the data followed by the full data set. These features are the first goal of our proposed database collection software and we thus working on enhancing these features and all other organizing features.

Limitations and Recommendations

The constricted survey time of this survey limited the candidates' responses. It is recommended to provide enough timing for data collection for next research. Moreover, it is better to use more Likert-type scale questions and reduce the use of multiple-answer multiple choice questions to allow for better correlation and appropriate statistical analysis

Conclusion

This study identified significant challenges faced by postgraduate researchers at Ain Shams University in accessing and managing patient data from the Oral and Maxillofacial Radiology (OMFR) department. Key difficulties included tracking patient history, incomplete data, and the lack of integration between data sources, which resulted in moderate to high levels of difficulty in data collection. These barriers have adversely impacted research efficiency, data quality, and the adoption of AI in dental diagnostics.

The survey findings underscore the need for a dedicated data management system with features such as easy access to patient data and comprehensive datasets, including clinical and radiographic records. Such a system would improve data standardization, facilitate retrospective studies, and enable more efficient research processes, thereby enhancing the practical implementation of AI in the field.

The study also highlighted areas for improvement in survey design, suggesting that future coming research should include longer survey periods, more Likert-scale questions, and open-ended responses to gain deeper insights into researcher needs. Addressing these challenges through a tailored data management system is essential to improving data quality, supporting advanced research, and ultimately enabling more accurate AI-driven diagnostic models in OMFR. The software can be generalized for all OMFR workers in any location.

Ethical approval

Ethical approval was granted by the local Ethics Committee of the faculty of dentistry, Ain Shams University (FDASU-Rec PC 122485).

Competing interests

The authors declare they have no conflict of interest.

Funding

No funding is subjected to the research reported in this manuscript.

Availability of data and materials

All data included in this study are available from the corresponding author upon request.

References

1. Saifeldin H. Comparison between manual lateral cephalometric analysis and artificial intelligence driven platforms. Ain Shams Dental Journal (Egypt). 2023;29(1).
2. Hesham N, Ashmawy M, Samir S. Reliability of Artificial Intelligence in Lateral Cephalometric Analysis. Ain Shams Dental Journal (Egypt). 2024;33(1).
3. Choi E, Schuetz A, Stewart WF, Sun J. Using recurrent neural network models for early detection of heart failure onset. Journal of the American Medical Informatics Association. 2017 Mar 1;24(2):361–70.
4. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. Vol. 25, Nature Medicine. 2019.
5. Roganović J, Radenković M, Miličić B. Responsible Use of Artificial Intelligence in Dentistry: Survey on Dentists' and Final-Year Undergraduates' Perspectives. Healthcare (Switzerland). 2023 May 1;11(10).
6. Carmel D, Yom-Tov E. Estimating the Query Difficulty for Information Retrieval. Synthesis Lectures on Information Concepts, Retrieval, and Services. 2010 Jan;2(1):1–89.
7. Kostopoulou O, Delaney BC, Munro CW. Diagnostic difficulty and error in primary care - A systematic review. Vol. 25, Family Practice. 2008.