The era of big data, mobile health, and artificial intelligence in dentistry and craniofacial research

The rapid advancement in information technologies, including cloud computing, social media, smartphones, various sensors, and the Internet of Things (IoT), has resulted in an exponential increase in data generation.[1] This surge contributed to the evolution of big data (BD), presenting significant opportunities as well as challenges across numerous research domains, including health care.

The term "big data" refers to any collection of datasets, both structured and unstructured, that is too large, complex, and dynamic for traditional systems of database management, application, and processing. The health care industry constantly generates large, important databases on patient demographics, treatment plans, results of medical examinations, insurance coverage, and more. BD research in medicine focuses on identifying meaningful patterns within extensive, often longitudinal clinical datasets using advanced knowledge-based analytical tools. Indeed, the rapid evolution of technology, artificial intelligence (AI) and its subset, machine learning (ML), have taken a central role in BD and real-world data (RWD) analytics, to generate real-world evidence (RWE).

While endless specific BD medical research topics developed, the medical BD research has primarily focused on three key areas: disease-related topics (eg, obesity, diabetes), technical aspects (eg, data mining, AI, ML), and health care services (eg, personalized medicine and elderly care).[1] The analysis of BD from multiple patients has advanced the identification of individual- and population-level risk factors, enabling the development of personalized strategies for prevention, diagnosis, and disease management. This research fosters large-scale behavior-focused studies (behavioromics), which improve patient outcomes, enhance patient-provider relationships, and reduce health care costs. For these reasons, BD is now recognized as one of the most critical and emerging technological domains, particularly in the health care sector.

However, challenges also exist for using BD analysis including regulation issues related to privacy, security and data sharing, the need for technical expertise, and operational costs. To advance research and face the challenges, many countries are vigorously promoting the fields of medical informatics. A growing number of medical institutions are receiving strategic support and funding to work, research, and analyze BD. Projections indicate a compound annual growth rate (CAGR) of 19.2% for the global BD market in digital health from 2022 to 2032, with its value anticipated to grow from $39.7 billion USD in 2022 to $194.7 billion by 2032.[2]

In dental, oral, and craniofacial research, BD research is a relatively new and developing field. Poor oral health leads to significant morbidity that affects overall wellbeing with a severe burden to society. Over recent decades, numerous systemic conditions have been studied in relation to oral diseases, especially dental caries and periodontitis. However, the findings are conflicting, which may stem from inconsistent definitions of dental and systemic conditions, and unaddressed confounding factors. To address these limitations, large-scale studies adhering to standardized definitions and accounting for confounders are necessary. Data linkage, which integrates separate datasets such as medical, dental, and sociodemographic data, facilitates tracking outcomes, monitoring less studied or rare conditions, and assessing health care costs. It helps explore complex population-level associations, particularly the link between oral and systemic health.

To address this gap in the literature, the corresponding author (GA) established and serves as a principal investigator of the Dental, Oral, Medical Epidemiological (DOME) study.[3] The DOME is an electronic records-based BD study that combines comprehensive structured sociodemographic, medical, and dental databases of a nationally representative sample of military personnel in Israel. The DOME is aimed to study oral-systemic interactions, in order to identify predictors of dental-oral-systemic morbidity.[3-13] A strict protocol used to establish the DOME study was developed and published, including the health informatics standards used for structuring and integrating the database, standardized terminology, data collection, and handling.[3] The first DOME repository, established in 2015, contains de-identi-

fied sociodemographic, dental, and medical records of 200,000 individuals. A subsequent longitudinal dataset, covering 2010 to 2020, was developed with data from 900,000 subjects across 700 variables, including sociodemographics, dental parameters (planned and delivered dental procedures, diagnoses, and attendance patterns), medical diagnoses, laboratory results, and a comprehensive medical survey.[3] To date, the DOME is the first and largest longitudinal nationwide comprehensive big data database combining sociodemographic medical and dental data in Israel, and one of the few in the world.

A series of studies utilizing the DOME repository explored links between dental health and factors such as cognitive function,[4] sociodemographic indicators, health behaviors, and dental attendance.[5] ML and statistical approaches were combined to investigate dental-oral-systemic morbidity, with a focus on components of metabolic syndrome and related conditions including hypertension,[8] obesity,[9] non-alcoholic fatty liver disease,[11] and obstructive sleep apnea.[13]

The "metabolic cluster" was further investigated in the context of periodontal disease[12] and temporomandibular disorders (TMD),[10] and its associations with the accumulation of dental pathologies, referred to as the "dental cluster," was also analyzed.[6] In a series of DOME papers, the term "SOS teeth" was coined, as the first priority teeth for treatment, and its associations with sociodemographic indicators[5] and metabolic outcomes[7] were analyzed. We have also analyzed the associations between C-reactive protein (CRP) and dental parameters and combined the statistically significant variables to create a dental inflammation score (DIS).[8] It is the view of the present authors that additional extensive studies are needed, which should focus on exploiting BD to explore the knowledge hidden in longitudinal, multivariate dental, oral, and craniofacial time-oriented data. Future studies will study longitudinal patients' trajectories to cluster and predict systemic-dental multi-morbidity.

Combining electronic health record (EHR) with other platforms such mobile health (mHealth) tools including smartphones and wearables, can enrich the datasets with valuable information collected outside the clinic boundaries and thus advance the BD research. As the principal investigators of a mHealth project entitled "Rheumatic Monitor," the present authors (GA and YS) developed a mobile application for iPhone and Android to monitor patients with painful inflammatory conditions, focused on autoimmune rheumatologic conditions with craniofacial manifestations (https://www.rheumaticmonitor.org/). The Rheumatic Monitor study focuses on advancing personalized medicine by identifying patterns from big clinical data that are predictive for exacerbations, remissions, and the responses to treatment of rheumatic disease patients. Currently, such predictive patterns are incompletely understood. The data are combined with a longitudinal EHR collected for 20 years and analyzed by the teams. A series of advanced temporal data mining and ML methods are used, including advanced methods developed in one of the present authors' (YS) laboratory.[14] Using the Rheumatic Monitor data, a novel multi-dimensional methodology has recently been developed to discover and validate the optimal number of patient clusters, and has been deployed for the task of clustering fibromyalgia patients.[15]

In summary, emerging technologies such as BD, AI, and mHealth hold significant potential to revolutionize dental, oral, and craniofacial research in the digital age, especially when exploited for analysis of large-scale longitudinal, multivariate time-oriented clinical data. Advancing this field will require international collaboration among multidisciplinary teams to drive global progress.

References

1. Gu D, Li J, Li X, Liang C. Visualizing the knowledge structure and evolution of big data research in healthcare informatics. Int J Med Inform 2017;98:22–32.
2. Fact.MR. Big data analytics in healthcare market. https://www.factmr.com/report/369/big-data-analytics-healthcare-market. Accessed 22 September 2024.
3. Almoznino G, Kedem R, Turgeman R, et al. The Dental, Oral, Medical Epidemiological (DOME) study: protocol and study methods. Methods Inf Med 2020;59:119–130.
4. Abramovitz I, Zini A, Atzmoni M, et al. Cognitive performance and its associations with dental caries: results from the Dental, Oral, Medical Epidemiological (DOME) records-based nationwide study. Biology (Basel) 2021;10:178.
5. Abramovitz I, Zini A, Kessler Baruch O, et al. SOS teeth with advanced caries and sociodemographic indicators, health-related habits and dental attendance patterns: data from the Dental, Oral, Medical Epidemiological (DOME) nationwide records-based study. BMC Oral Health 2021;21:389.
6. Abramovitz I, Zini A, Pribluda P, et al. "Dental cluster" versus "metabolic cluster": Analyzing the associations of planned and delivered dental procedures with metabolic syndrome, utilizing data from the Dental, Oral, Medical Epidemiological (DOME) cross-sectional record-based nationwide study. Biology (Basel) 2021;10:608.
7. Almoznino G, Kessler Baruch O, Kedem R, et al. SOS teeth: first priority teeth with advanced caries and its associations with metabolic syndrome among a national representative sample of young and middle-aged adults. J Clin Med 2020;9:3170.
8. Almoznino G, Zini A, Kedem R, Protter NE, Zur D, Abramovitz I. Hypertension and its associations with dental status: data from the Dental, Oral, Medical Epidemiological (DOME) nationwide records-based study. J Clin Med 2021;10:176.
9. Ben-Assuli O, Bar O, Geva G, Siri S, Tzur D, Almoznino G. Body mass index and caries: machine learning and statistical analytics of the Dental, Oral, Medical Epidemiological (DOME) nationwide big data study. Metabolites 2022;13:37.

10. Chweidan H, Rudyuk N, Tzur D, Goldstein C, Almoznino G. Statistical methods and machine learning algorithms for investigating metabolic syndrome in temporomandibular disorders: a nationwide study. Bioengineering (Basel) 2024;11:134.

11. Ram D, Wilensky A, Zur D, Almoznino G. The triangle of non-alcoholic fatty liver disease, metabolic dysfunction, and periodontitis: analysis of the Dental, Oral, Medical and Epidemiological (DOME) records-based nationwide research. Metabolites 2022;12:1212.

12. Wilensky A, Frank N, Mizraji G, Tzur D, Goldstein C, Almoznino G. Periodontitis and metabolic syndrome: statistical and machine learning analytics of a nationwide study. Bioengineering (Basel) 2023;10:1384.

13. Ytzhaik N, Zur D, Goldstein C, Almoznino G. Obstructive sleep apnea, metabolic dysfunction, and periodontitis-machine learning and statistical analyses of the Dental, Oral, Medical Epidemiological (DOME) big data study. Metabolites 2023;13:595.

14. Lion M, Shahar Y. Implementation and evaluation of a multivariate abstraction-based, interval-based dynamic time-warping method as a similarity measure for longitudinal medical records. J Biomed Inform 2021;123:103919.

15. Goldstein A, Shahar Y, Weisman Raymond M, et al. Multi-dimensional validation of the integration of syntactic and semantic distance measures for clustering fibromyalgia patients in the Rheumatic Monitor big data study. Bioengineering (Basel) 2024;11:97.

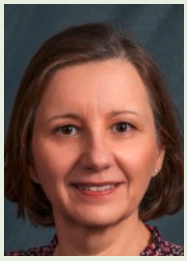
Galit Almoznino, DMD, MSc, MHA, MPH
Associate Professor, Head, Big Biomedical Data Research Laboratory, Hebrew University-Hadassah, School of Dental Medicine, Jerusalem, Israel; and Biostatistician and Epidemiologist, Specialist in Oral Medicine, Head, Orofacial Chemo-Sensory Clinic, Department of Oral Medicine, Sedation & Maxillofacial Imaging, Hebrew University-Hadassah School of Dental Medicine, Jerusalem, Israel.
Email: galit@almoznino.com

Yuval Shahar, MD, PhD
The Josef Erteschik Chair in Information Systems Engineering; Full Professor and Head, Medical Informatics Research Center, The Department of Software and Information Systems Engineering, Ben Gurion University of the Negev, Beer Sheva, Israel

Dorota T. Kopycka-Kedzierawski, DDS, MPH
Professor of Dentistry, Eastman Institute for Oral Health; Node Director, NorthEast Region National Dental PBRN; Director, EIOH Clinical and Translational Research Core, Rochester, USA

Galit Almoznino

Yuval Shahar



Dorota T. Kopycka-Kedzierawski