Check for updates

Received: 27 February 2021 | Revised: 23 June 2021

DOI: 10.1111/ocr.12520

REVIEW ARTICLE

Orthodontics & Craniofacial Research

WILEY

Complexity and data mining in dental research: A network medicine perspective on interceptive orthodontics

Tommaso Gili[1,2] | Gabriele Di Carlo[3] | Silvia Capuani[2] | Pietro Auconi[4] | Guido Caldarelli[2,5] | Antonella Polimeni[3]

[1]Networks Unit, IMT School for Advanced Studies Lucca, Lucca, Italy

[2]CNR-ISC Unità Sapienza, Rome, Italy

[3]Department of Oral and Maxillo-Facial Sciences, Sapienza University of Rome, Rome, Italy

[4]Private Practice Rome, Rome, Italy

[5]Department of Molecular Sciences and Nanosystems, Ca'Foscari University of Venice, Venezia Mestre, Italy

Correspondence
Gabriele Di Carlo, Department of Oral and Maxillo-Facial Sciences, Sapienza University of Rome, Viale Regina Elena 287a, 00161 Rome, Italy.
Email: gabriele.dicarlo@uniroma1.it

Abstract

Procedures and models of computerized data analysis are becoming researchers' and practitioners' thinking partners by transforming the reasoning underlying biomedicine. Complexity theory, Network analysis and Artificial Intelligence are already approaching this discipline, intending to provide support for patient's diagnosis, prognosis and treatments. At the same time, due to the sparsity, noisiness and time-dependency of medical data, such procedures are raising many unprecedented problems related to the mismatch between the human mind's reasoning and the outputs of computational models. Thanks to these computational, non-anthropocentric models, a patient's clinical situation can be elucidated in the orthodontic discipline, and the growth outcome can be approximated. However, to have confidence in these procedures, orthodontists should be warned of the related benefits and risks. Here we want to present how these innovative approaches can derive better patients' characterization, also offering a different point of view about patient's classification, prognosis and treatment.

KEYWORDS

artificial intelligence, big data, complexity, machine learning, musculoskeletal magnetic resonance imaging, network medicine, orthodontics

1 | INTRODUCTION

The technological revolution we have been experiencing in the last decades impacted medical sciences in various ways. Many results from high-throughput experiments and the increased connection between data scientists have made available an unprecedented amount of insights to researchers and practitioners. Consequently, it becomes evident that many biological processes, as craniofacial growth, are prominent examples of complex systems in which the outcome depends upon the interaction of various components.[1,2] Each element depends on its past properties, the properties of the other elements, and their past. This new scientific paradigm triggered the search for mathematical instruments to describe and possibly drive such complex behaviour.[3] A complex system of interacting agents takes a natural mathematical form of a graph where the vertices are the elements of the system, and their complex interaction is put in the form of an edge (directed, weighted, etc).[4] By considering the individual variability within patients, network theory makes it possible to optimize and individualize the diagnosis in the modelling of precision medicine.[5,6] In the last decades, the evolution of disparate, complex biological

Tommaso Gili and Gabriele Di Carlo contributed equally to the work and should be considered joint first author.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2021 The Authors. *Orthodontics & Craniofacial Research* published by John Wiley & Sons Ltd.

systems has been predicted, almost automatically, by using different tools and procedures from Artificial Intelligence (AI).[7-10]

From an initial data set of patient's characteristics ('learning set'), AI algorithms learn how the features relate to and predict the outcome. Machine Learning (ML), a sub-discipline in AI, is instead focussed on the ability to handle noisy or irrelevant data and on the capability to predict the outcome of a disease based on data derived from similar conditions.[11,12] The first attempts to use biomedical data to extract prediction resulted in probabilistic models trained on a series of case studies, tried to match the individual patient's condition with predefined classes of stratified increasing health risks.[13]

Using Artificial Neural Network procedures, the correlation between early craniofacial features and the risk of craniofacial worsening during growth was established in 43 orthodontic patients.[14] As the number of cases available for predicting health outcomes increased, new problems related to data incompleteness and inhomogeneity and ethical aspects of storage arose.[15,16] Using Graph analysis, feature covariance across orthodontic patients allowed to disentangle Class III subjects from the topological patterns of other malocclusions.[17] ML algorithms have been adopted to assist the orthodontist in treatment plan decision, including premolar extraction.[18,19] Indeed, such essential decisions during orthodontic treatments can be subjected to different points of view, as they tend to be based on the practitioner's experience and intuition. Resorting to ML procedures, one can decrease, at least in principle, personal biases in the treatment analysis and choices.[19]

The precision of ML answers crucially depends on the quality of data input, both in terms of number and appropriateness. ML algorithms cannot work if provided with raw data, which must be transformed into domain-specific representative and salient information. Therefore, choosing the best characteristics (the most expressive of the problem) is crucial for predicting. It is also imperative to collect and give the machine data expressed as geometric, physiological, clinical and anamnestic parameters to make the most of what the current technology has to offer.[17,20-22]

The integration of biomedicine and computer science is based on two concepts: Systems of the body have complex and dynamic biological properties that rely on the interaction of molecular agents sustaining the physiological functioning, as well as the pathogenesis of diseases at different scales; if complex interactions cannot be understood or processed by conventional methods, they can be investigated and explained using ML. The quantification of imaging biomarkers able to witness the status of a system of the body is the core business of medical research. Among the multiple choices available in the field, Magnetic Resonance Imaging (MRI)[23] is the one that offers high accuracy and reliability at multiple spatial resolutions, with the invaluable plus to be non-invasive.

This work aimed to present an essential overview of ML techniques and possible applications in the orthodontic discipline, highlighting merits, potentials and potential improvements offered by the inclusion of networked multiscale musculoskeletal data congruently segregated for advanced AI procedures. The integration of biomedical databases and mathematical models in multiscale and multi-physics systems is a well-established research field[24,25] that discusses dynamic and static mechanical dynamics as essential elements in constructing models of living organisms and diseases. Although it is well beyond the ultimate purpose of this paper, we feel confident that the inclusion of such modelling in the fusion of biomedicine and computer science presented here will be mandatory in the next future.

1.1 | The orofacial system as a hierarchical multiscale complex system

Orthodontic researchers and practitioners are interested in extracting the most significant possible amount of information from all potential sources to provide the best diagnostic framing and treatment. It means mining significant patterns from past studies and tracing down the localization of the pathobiology underlying growing trajectories with and without unfavourable dentoskeletal growth. Such difference may point towards the prediction of modifiable skeletal disharmonies.[26] However, it is problematic to understand a priori what would be the best data representation, and transform inputs into information that the machine could understand.

It is well recognized that bones and teeth structural and mechanical properties,[27] and the elastic properties of muscles strongly influence the post-treatment stability.[28]

Once conceptualized as a whole, bones, teeth and muscles represent a complex system whose physiology is related to their multiscale hierarchical structures and the precise organization of inorganic and organic phases at the nanoscale, microscale, up to the macroscopic scale (Figures 1,2).

It is known how the effects of orthodontic treatments can lead to significant or poor results, depending on the period of growth during which it is implemented and on the quality of the bone and bone-tooth interface.[29] These latter factors involve both the mineral part and the collagen of bone marrow (Figure 1).[30-32] Therefore, it is crucial to intervene with orthodontic treatments in the most appropriate period of growth but only to know the characteristics of the cancellous bones. Currently, there are attempts to evaluate both the optimal period of treatment and the health of the cancellous bone related to vertebral maturation. The maturation of the different skeletal segments does not happen at the same time. For example, the upper and lower jaw complete their development at other times.[33] Due to the quality of the bone, the teeth can move or rotate unexpectedly, even with a technically perfect orthodontic procedure. Parameters that quantify the upper and lower jaw spongy bone quality should constitute the dataset to be collected for an appropriate and inclusive treatment plan. It is worth quantifying MRI parameters such as T2* and Apparent Diffusion Coefficient (ADC),[34-38] which are related to spongy bone quality. Efforts to translate new dental and bone microstructural investigations to clinical practice in orthodontics should be devoted to making reliable predictions about the evolution of dentoskeletal tissue. It is crucial to understand how the micro and nanoscopic rearrangement contribute to the macroscopic bone tissue modification.[32]

16016343, 2021, S2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ocr.12520 by CochraneItalia, Wiley Online Library on [21/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

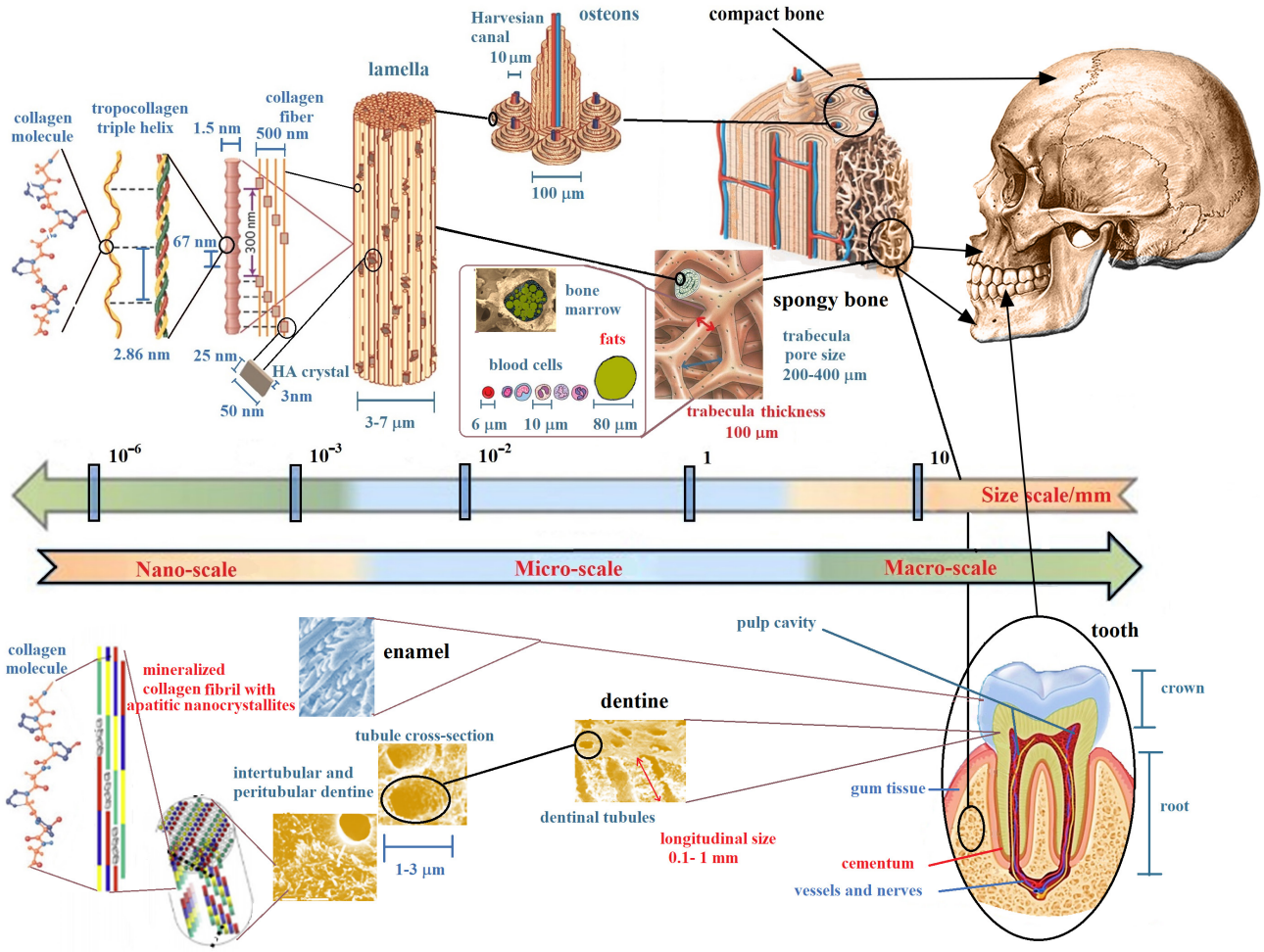

FIGURE 1 Schematic depiction of bone and tooth multiscale hierarchical structure from macroscopic bone (up raw) and tooth (bottom raw) to their nanoscopic elements. This hierarchical organization is the genesis of bone and dentine properties, including stress tolerance, adaptability and development during the growth process. In dentine, tubules are the prominent structural feature at a micro-level, whereas collagen fibres decorated with apatite crystallite platelets dominate the nanoscale. In bone, hydroxyapatite (HA) crystals at nanometer-level periodically are deposited within the gap zones of collagen fibrils during the bone biomineralization process. This hierarchical arrangement produces nanomechanical heterogeneities, which enable a mechanism for high energy dissipation and resistance to fracture. At a micro-macro-level, bone marrow quality in spongy bone and trabecular network rearrangement affects the resistance of bone to mechanical stress. Adapted from reference[24,26,65-67] [Colour figure can be viewed at wileyonlinelibrary.com]

Cells of the musculoskeletal system reside within complex and often interconnected 3D environments, which regulate musculoskeletal physiology and homeostasis (Figure 2). Maxillofacial development is also regulated by the surrounding soft tissues. Muscles affect skeletal components' size and shape with a possible different relationship between the muscular pressure and structural configuration in Class I, II and III malocclusions.[39] Several orthodontic studies suggest the usefulness of monitoring electromyographic (EMG) activity of the jaw muscles. Moreover, a close correlation between skeletal and muscle tissues states has been recently highlighted,[40] and fat increase and decreased muscle function. Parameters such as MRI T2 and ADC have proven useful to evaluate the muscles' state because of the relationship between the microscopic structure and function.[41-44]

As highlighted above, despite the multiscale hierarchical nature of the musculoskeletal tissue, only macroscopic cephalometric variables composed of linear and angular geometric measurements are considered in conventional clinical practice. However, due to the development of in vivo imaging technology, additional methods can be used to obtain a set of multimodal and multi-parametric measurements at different length scales, involving the microstructure, the topology of the craniofacial tissue, and also its physiology.[45] In recent years, in addition to the conventional x-ray computer tomography (CT) and the cone-beam computer tomography (CBCT), MRI has shown great potential in dentistry.[46-50] Furthermore, as most patients receive orthodontic treatment during childhood, further development of radiation-free techniques, such as MRI are highly desirable.[51-53]

16016343, 2021, S2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ocr.12520 by CochraneItalia, Wiley Online Library on [21/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

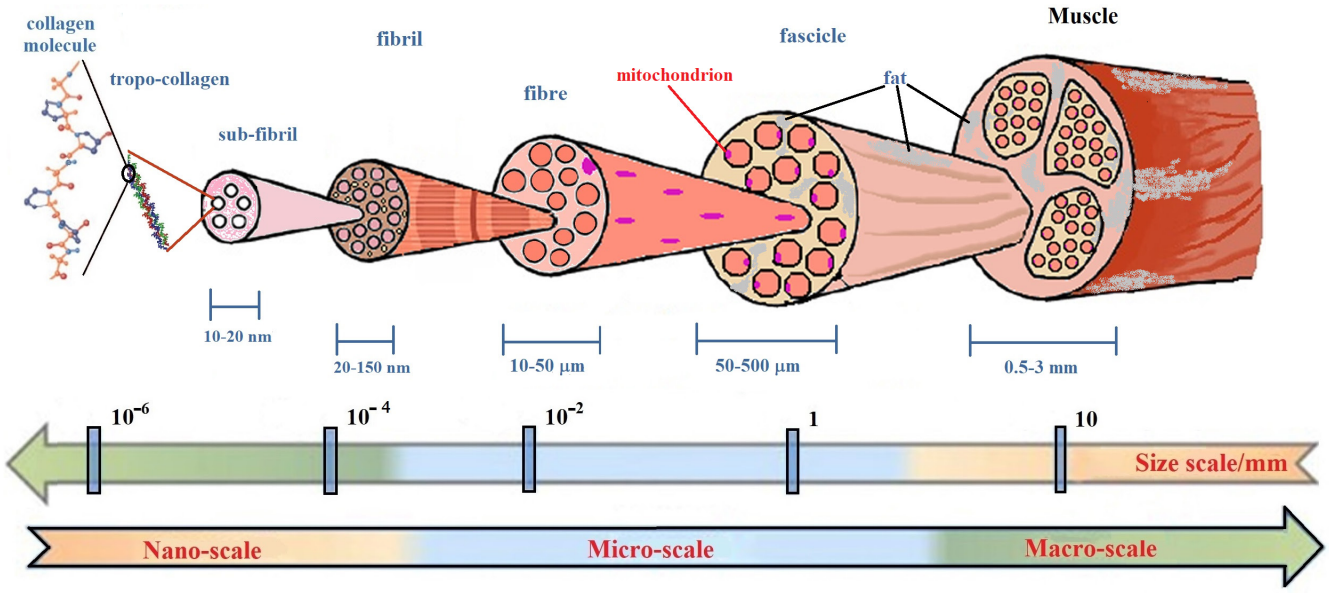

FIGURE 2 Schematic representation of muscle multiscale hierarchical structure. Most of the physiological muscle functions are related to its hierarchical organization and components. Physical inactivity causes a decrease in muscle mass and an increase in fat mass, but a chronic high fat diet also increases muscle fat, limiting full muscle function. Muscles quality is often related to skeletal tissue quality. Parameters that quantify the craniofacial muscles quality of an orthodontic patient should constitute the data set to be collected for planning treatment. Adapted from references[40,41] [Colour figure can be viewed at wileyonlinelibrary.com]

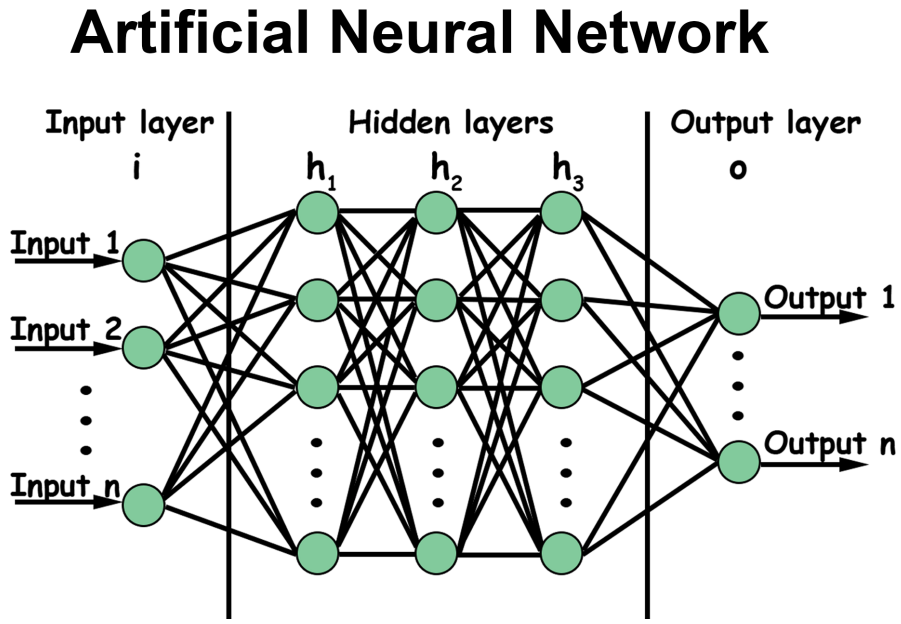

FIGURE 3 Artificial neural network schematization. An artificial neural network is based on a set of connected nodes, where connections, like synapses in a biological brain, can transmit a signal from a node to another. The transmitted signal is a real number, and the output of each node is computed by some non-linear function of the sum of its inputs. Connections (also called edges or links) typically have a weight that adjusts as learning proceeds. Nodes are aggregated into layers, and different layers may perform various transformations on their inputs. Signals travel from the first layer to the last one (the output layer), often after going through the mid-layers (hidden layers) multiple times. According to the number of layers included in the neural network, different machines can be realized, from simple machine-learning engines to deep learning ones [Colour figure can be viewed at wileyonlinelibrary.com]

Together with the measurements provided by the imaging techniques, it is crucial to collect clinical and anamnestic data from patients. Some authors use genetic data to obtain a set of priors that could better supplement the inputs for correct patient classification.[54] Modelling malocclusion progression means taking into account the complexity of the system. The complex interplay of causes

16016343, 2021, S2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ocr.12520 by CochraneItalia, Wiley Online Library on [21/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

behind atypical growth requires a different perspective about the disorders affecting the orofacial biological balance. By enlarging the basin of information about the system's chemical, physical and mechanical properties, it is possible to enable its description and evolution using more in-depth ML and AI procedures. To achieve data more representative of the patient's situation, and therefore, the most significant possible amount of information, different in vitro investigation techniques about orofacial tissue at various scale lengths could be of clinical and research interest (Figure 2).[55,56] Once the most extensive possible spectrum of information has been obtained, deciding the best approach to combine the collected data are required. Here, we present different ML methods that stem from the complex network theory to offer the appropriate patients' classification, without drawing on the choice of a specific and predictive biophysical model.

1.2 | Learning features from data

Biomedical data, such as growth data, are constantly evolving. The related information is disseminated in a network of interconnected pieces of local information (nodes), and memory is encoded in the topology and the strength of such multiple connections (edges), rather than stored in the single information as in statistical databases. The power of such networks resides in their capacity to learn.[57] Both in living beings and computer algorithms, the strengthening of connections is the microscopic mechanism for elaborating the information.[58,59] Artificial neural networks (ANN) use multiple layers of calculations to imitate the human brain's reasoning and draw conclusions from initial information. ANN can deal with complex intertwined problems (Figure 3). Different types of signs, symptoms, X-rays, risk factors, imaging results enter into ANN algorithms to find the most predictive combination of variables. ANN assign weights to some combination of nodes (the features chosen to 'train' the machine) to repeatedly optimize the model's predictive performance.[60] According to the number of layers included in the neural network, different machines can be realized, from simple machine-learning (ML) engines to 'deep' learning ones.[61,62]

The canonical ML workflow involves four steps: (a) data gathering, cleaning and pre-processing, (b) feature extraction, (c) model training, (d) evaluation of results. Refining these steps can be complicated by intuition alone. The ability to separate patient's outcomes requires more targeted diagnostics, as to subgrouping patients with significantly different growth trajectories and clinical courses despite the similar early diagnostic frameworks.[60,63] Once ML approaches are matched to orthodontic data, challenges arise from data incompleteness, high-dimensionality, heterogeneity, dynamicity, sparseness and statistical noise that can be partially mitigated by Network representation.[15,64] Due to such data's complex and interconnected nature, any single model can deepen only a tiny part of the entire orthodontic domain.[65,66] Whilst the conventional medical approach is based on the careful recruitment of clinical and laboratory data, testing of a diagnostic hypothesis, causes and effects of phenomena, significance, checking the initial hypothesis and so on, ML workflow focuses instead on the fast predictive performance of models and iterative improvement of the algorithms, coping with high-dimensional spaces, variability of features and formats.[60,67]

1.2.1 | Data gathering and Feature transformation

Feature engineering (FE) is the process that transforms raw data into features to feed into the prediction models. Therefore, features sit between data and models in the ML workflow (Figure 4). Moreover, being interested in defining the quality of growth, in the beginning, it has to decide how to assign each patient to the proper category: as an example, favourable/unfavourable growth often represents the label of interest. Then, patients' clinical and cephalometric data can be imported into the training data, together with the specific labels. It is the beginning of the supervised learning approach. The algorithm will find the function that links the patient's input data with the outcome, minimizing the number of errors. However, raw data can rarely be entered without a preliminary selection to discard redundancies and simplify the system.[68]

1.2.2 | Learning

Regularization is the crucial attribute for high-quality learning data from a large number of features. The process does not aim to predict labels within data of the learning dataset optimally but rather generalize the prevision about new, previously unseen patients. Accordingly, data regularization is equivalent to imposing a penalty for the system's complexity to improve performance.[69] When the algorithm has few patient samples available from which it can learn the connection between the characteristics and the outcome, it ends up memorizing the training data set instead of learning general features of data (over-adaptation to data, a.k.a. 'overfitting'). For this reason, a model will perform well on the training set and poor with new patient's data. On the other hand, a small number of patient's characteristics in the learning set provides an inadequate description of the problem at hand and may result in learning difficulties ('underfitting'). Thus, there is a trade-off between the model performance and the size of the training set, that is, the patient coverage, the per cent of potentially eligible patients for that risk assessment can be completed.[69]

Supervised ML learning is used when the output is known. The learning procedure primarily deals with regression and classification problems. It focuses on classification, which involves choosing between subgroups to describe a new patient best, and predictions, which consists of estimating an outcome of interest, such as the quality of future craniofacial growth.[63,69]

16016343, 2021, S2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ocr.12520 by CochraneItalia, Wiley Online Library on [21/1/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



FIGURE 4 Machine learning. Different approaches to the learning process produce different machine-learning schemes. According to the number of features used to train the machine, one can obtain supervised learning (labelled data sets are used to train the machine), unsupervised learning (unlabelled data sets are used to train the machine), semi-supervised learning (a mix of labelled and unlabelled data sets are used to train the machine) and reinforcement learning (there are not data sets to train the machine). According to the quality of the target data (continuous or categorical), different tasks can be performed by the learning processes [Colour figure can be viewed at wileyonlinelibrary.com]

In contrast, unsupervised learning is not intended to find outputs to predict, and it tries to find naturally occurring patterns within data.[10] The patterns identified have to be evaluated for utility. Unsupervised models aim to discover groupings from data samples 'x' without knowing the label output 'y'. The algorithm is provided with unclassified data records to recognize whether any existing latent patterns are present. Clustering, dimensionality reduction, like principal component analysis are leading examples of unsupervised learning approaches. The preliminary subdivision into subgroups of patients may lead to the subsequent estimation of specific risk factors. In particular, clustering refers to the extraction of group latent similarities within data that allow subjects to be grouped into subsections. The clustering assumption states that such subsections of subjects often exhibit the same outcome.[70] In orthodontics, because of the complexity of medical data and heterogeneity of patients, identifying subsections of patients by intuition can be difficult.[20]

Semi-supervised learning (SSL) is a mixture of supervised and unsupervised models. It analyses several unlabelled cases (patients) whilst augmenting its pattern recognition capacities with a small quantity of labelled data.[71]

This approach is promising in orthodontic longitudinal studies since it is laborious to find data that refer to many patients followed closely over time, at comparable regular intervals; even more, untreated patients followed longitudinally.

Finally, reinforcement learning (RL) is an approach where software agents take actions in an environment to maximize a cumulative reward. RL differs from SL in not needing labelled input/output pairs to be presented without requiring sub-optimal actions to be corrected. The goal is finding a balance between exploration (of uncharted territory) and exploitation (of current knowledge).[72,73]

1.3 | Learning features from networks

Learning techniques rely on networked environments to conduct the learning process on data of interest. What if data are not generated from grid-like Euclidean structures (like images and videos) and represented as graphs with complex interdependencies between objects? For network-based methods, the learning procedure is performed by navigating in networks built from the input data according to some similarity criterion.[74] As networks naturally contain topological information of data relationship, network-based methods take advantage of typical algorithms that use raw data. It must be stressed that network-based methods can be considered a general solution for learning tasks, even for data sets not represented by networks. One can apply network construction techniques to that data set to generate a network from the input data. Once the network is constructed, the learning process can be run. Patient stratification is a general target of network-based methods of machine learning. For example, cancer somatic mutation profiles are highly coupled with the biomolecular network,[5,75] in fact, somatic mutations of a cancer driver gene may lead cancer genome evolution to mutations in other genes.[76] Therefore, each patient with its somatic mutation profiles can be identified, and the similarity between patients can be used to mine for tumour stratification. Analogously, information about malocclusion within the orofacial system can be extrapolated from the correlation matrix of the orthodontic

16016343, 2021, S2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ocr.12520 by CochraneItalia, Wiley Online Library on [21/1/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

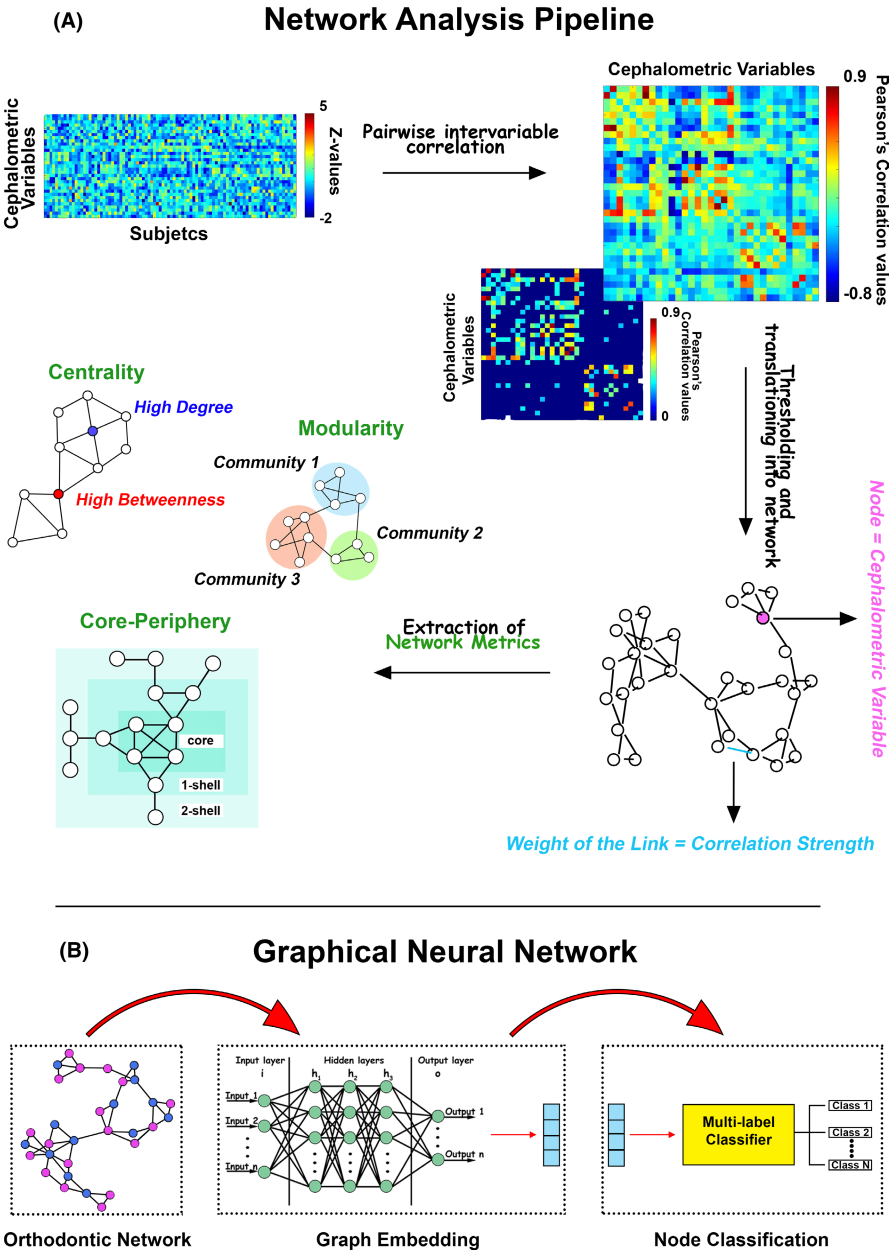

FIGURE 5 Complex Networks and data analytics. (A) Network analysis pipeline for orthodontics data. Once cephalometric variables are standardized to Z-values, they are entered in a cross-correlation process that returns a symmetric matrix, whose entries are the intervariable Pearson's correlation coefficients across subjects. A threshold is set to the matrix according to the *P*-values associated with the coefficients. The final matrix (a weighted adjacency matrix) is translated into a network whose nodes are the cephalometric variables and the weights of the links the Pearson's correlation coefficients that survived the thresholding process. Finally, different metrics have been calculated from the network topology: centrality measures, modules or communities and the core-periphery structure. (B) GNN. Low-dimensional node representations are first learned from networks by graph embedding and then used as features to build specific classifiers for different tasks [Colour figure can be viewed at wileyonlinelibrary.com]

features. The correlation matrix relates to a network whose vertices are associated with the features and the edges to their covariance across patients,[17] allowing the visualization of malocclusion information (Figure 5A). The networks are obtained by fixing a threshold value T: if the Pearson correlation between two vertices (features) is larger than T, they are considered connected ("linked").[17] Several topological network metrics are typically considered to convey biological meaning to cephalometric correlations (Figure 5A), as 'betweenness'[20] (a node centrality index) or modularity (the capacity of nodes to form communities characterized by similar properties).[77] Once the graph is computed, the topological structure of the input is encoded in simpler structures (eg vectors of reals), and then the ML algorithms can be run.

The application of ML to graphs can be generally divided into two categories[78]:

- Node-focussed applications, where the final task is associated with specific properties of each node (examples include node clustering, link prediction, semi-supervised node classification).
- Graph-focussed application, where the model to be realized is dependent on the whole graph structure (examples include estimation of properties of graphs and graph classification).

The machine-learning methods that operate on graph domain are known as Graph Neural Networks (GNNs). Due to its convincing performance, GNN has become widely applied to infer data described by graphs.[78-80] In GNN, the concept of Node Embedding is implemented. It means to map nodes to a d-dimensional embedding space (low dimensional space rather than the actual dimension of the graph) to embed close to each other similar nodes in the graph. The procedure aims to map nodes, so the similarity in the embedding space approximates similarity in the network.[81,82] Training can be unsupervised,

16016343, 2021, S2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ocr.12520 by CochraneItalia, Wiley Online Library on [21/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

semi-supervised or supervised. The supervised setting provides labelled data for training and is used for node classification. The semi-supervised setting gives a small number of labelled nodes and many unlabelled nodes for training. In the test phase, the transductive setting requires the model to predict the labels of the given unlabelled nodes, whilst the inductive setting provides new unlabelled nodes from the same distribution to infer. Most node and edge classification tasks are semi-supervised. The unsupervised setting only offers unlabelled data for the model to find patterns. Node clustering (features/patients grouping) is a typical unsupervised learning task.

1.4 | Perspectives

The parameters necessary to understand and predict the evolution of an orthodontic condition (such as a malocclusion) cannot be captured by a single measurement modality or using a few data type (as an example, geometrical data from cephalometrics, from CBCT or anamnestic data). The dental-musculoskeletal tissue complexity lies in the interconnection of morphological progression, functional and genetic data. Nowadays, new and more sophisticated AI algorithms can deepen these aspects. Unfortunately, they do not give a clue about the reasoning that led to deciding the saliency, the relevance or the causal importance of the variables.[83,84] Interestingly, theoretical advancements with methods from statistical physics will also allow determining partial information as estimating longitudinal data from cross-sectional to create individual prognosis.[70] Humans are brilliant at clinical pattern recognition in dimensions equal to or less than three. However, most medical data's dimension is much higher than three, making cognitive analysis problematic or even impossible.[6]

Multiple perspectives are required to study complex phenomena formed of parts in a non-random organization, in which our knowledge can only be partial and idealized. Forcing biological explanatory pluralism into the narrow computational framework appears to be a misleading strategy. The complexity of the musculoskeletal craniofacial system requires the use of novel sources of information. In vivo, MRI matches this framework optimally, providing the elucidation of geometrical and physiological tissue features, even at length scales smaller than voxel resolution of X-ray CT and CBCT in a radiation-free modality. The computational models applied to orthodontics must include the skeletal characteristics change over time (even at the end of growth) and the relationship between the mechanical properties of the complex system composed of bones and teeth and their multiscale hierarchical structures, whose precise organization orchestrates the whole physiology. The question is to what extent this new 'deep' reasoning leads to reliable and responsible decisions in the global growth of the skeletal bases and the maturation of the dentition. The course of events almost always deviates from what was predicted and planned. There may be areas of local densification of skeletal disharmonies that can act as growth attractors. These areas are difficult to detect clinically and neither be easily detected by computational systems.

The future of orthodontics inexorably will pass through a Copernican revolution that will lead algorithms to optimize therapies ('personalized medicine'). Nonetheless, the process will not be complete until the complexity of the problem is fully addressed. The need for new imaging techniques (in vivo and in vitro) to characterize teeth, bones and muscles structures and physiology, the availability of genetic data, and the filtration of helpful information from them will be the route to a new vision of orthodontics.

ACKNOWLEDGMENT

Open Access funding provided by Universita degli Studi di Roma La Sapienza within the CRUI-CARE Agreement. [Correction added on 27 May 2022, after first online publication: CRUI funding statement has been added.]

CONFLICT OF INTEREST

The authors have no conflict of interests to declare.

AUTHOR CONTRIBUTION

Conceptualization, TG, GDC; methodology, PA, SC.; writing—original draft preparation, TG, GDC; writing—review and editing, GC, AP; supervision, GC.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no data sets were generated or analysed during the current study.

REFERENCES

1. Andersen PW. More is different. *Science*. 1972;177(4047):393-396. https://doi.org/10.1126/science.177.4047.393
2. Polimeni A. *Pediatric Dentistry* (2nd Ed). Edra (Ed); 2020.
3. Yan G, Tsekenis G, Barzel B, Slotine JJ, Liu YY, Barabási AL. Spectrum of controlling and observing complex networks. *Nat Phys*. 2015;11:779-786. https://doi.org/10.1038/nphys3422
4. Scale-Free CG. *Networks: Complex Webs in Nature and Technology*. Oxford University Press; 2007. https://doi.org/10.1093/acprof:oso/9780199211517.001.0001
5. Barabási AL, Gulbahce N, Loscalzo J. Network medicine: A network-based approach to human disease. *Nat Rev Genet*. 2011;12(1):56-68. https://doi.org/10.1038/nrg2918
6. Holzinger A. Extravaganza tutorial on hot ideas for interactive knowledge discovery and data mining in biomedical informatics. In: *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Berlin: Springer; 2014. https://doi.org/10.1007/978-3-319-09891-3_46
7. Angermueller C, Pärnamaa T, Parts L, Stegle O. Deep learning for computational biology. *Mol Syst Biol*. 2016;12(7):878. https://doi.org/10.15252/msb.20156651
8. Bellazzi R, Zupan B. Predictive data mining in clinical medicine: Current issues and guidelines. *Int J Med Inform*. 2008;77(2):81-97. https://doi.org/10.1016/j.ijmedinf.2006.11.006
9. Miotto R, Li L, Kidd BA, Dudley JT. Deep patient: an unsupervised representation to predict the future of patients from the electronic health records. *Sci Rep*. 2016;6(1):26094. https://doi.org/10.1038/srep26094
10. Deo RC. Machine learning in medicine. *Circulation*. 2015;132(20):1920. https://doi.org/10.1161/CIRCULATIONAHA.115.001593

16016343, 2021, S2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ocr.12520 by CochraneItalia, Wiley Online Library on [21/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

11. Il GK, Cusick ME, Valle D, Childs B, Vidal M, Barabási AL. The human disease network. *Proc Natl Acad Sci U S A*. 2007;104(21):8685-8690. https://doi.org/10.1073/pnas.0701361104
12. Lovell DP. Biological importance and statistical significance. *J Agric Food Chem*. 2013;61(35):8340-8348. https://doi.org/10.1021/jf401124y
13. Braitman LE, Davidoff F. Predicting clinical states in individual patients. *Ann Intern Med*. 1996;125(5):406. https://doi.org/10.7326/0003-4819-125-5-199609010-00008
14. Lux CJ, Stellzig A, Volz D, Jäger W, Richardson A, Komposch G. A neural network approach to the analysis and classification of human craniofacial growth. *Growth Dev Aging*. 1998;62(3):95-106.
15. Cios KJ, William MG. Uniqueness of medical data mining. *Artif Intell Med*. 2002;26(1-2):1-24. https://doi.org/10.1016/S0933-3657(02)00049-0
16. Faber J, Faber C, Faber P. Artificial intelligence in orthodontics. *APOS Trends Orthod*. 2019;9:201-205. https://doi.org/10.25259/apos_123_2019
17. Auconi P, Caldarelli G, Scala A, Ierardo G, Polimeni A. A network approach to orthodontic diagnosis. *Orthod Craniofacial Res*. 2011;14(4):189-197. https://doi.org/10.1111/j.1601-6343.2011.01523.x
18. Al Turkestani N, Bianchi J, Deleat-Besson R, et al. Clinical decision support systems in orthodontics: a narrative review of data science approaches. *Orthod Craniofac Res*. 2021. https://doi.org/10.1111/ocr.12492. Online ahead of print.
19. Jung SK, Kim TW. New approach for the diagnosis of extractions with neural network machine learning. *Am J Orthod Dentofac Orthop*. 2016;149(1):127-133. https://doi.org/10.1016/j.ajodo.2015.07.030
20. Auconi P, Scazzocchio M, Cozza P, McNamara JA, Franchi L. Prediction of class III treatment outcomes through orthodontic data mining. *Eur J Orthod*. 2015;37(3):257-267. https://doi.org/10.1093/ejo/cju038
21. Gui C, Chan V. Machine learning in medicine. *Univ West Ont Med J*. 2017;86(2):76-78. https://doi.org/10.5206/uwomj.v86i2.2060
22. Sajda P. Machine learning for detection and diagnosis of disease. *Annu Rev Biomed Eng*. 2006;8(1):537-565. https://doi.org/10.1146/annurev.bioeng.8.061505.095802
23. Seiberlich N, Gulani V, Campbell A, Sourbron S, Doneva MI, Calamante F, Hu HH (eds). *Quantitative Magnetic Resonance Imaging* (Volume 1, 1st Edn). Academic Press; 2020. https://www.elsevier.com/books/quantitative-magnetic-resonance-imaging/seiberlich/978-0-12-817057-1
24. Bassingthwaighte JB. Strategies for the physiome project. *Ann Biomed Eng*. 2000;28(8):1043-1058. https://doi.org/10.1114/1.1313771
25. Hunter P, Coveney PV, De Bono B, et al. A vision and strategy for the virtual physiological human in 2010 and beyond. *Philos Trans R Soc A Math Phys Eng Sci*. 2010;368(1920):2595-2614. https://doi.org/10.1098/rsta.2010.0048
26. Zheng A, Casari A. *Feature engineering for machine learning: Principles and Techniques for Data Scientists*. O'really media,Inc; 2018.
27. Spencer P, Misra A, Ye Q. Fundamentals of the material-tissue interface in dental reconstructions: Structure/property relationships and characterization. In: *Material-Tissue Interfacial Phenomena: Contributions from Dental and Craniofacial Reconstructions*. Woodhead publishing; 2017. https://doi.org/10.1016/B978-0-08-100330-5.00004-2
28. Pepicelli A, Woods M, Briggs C. The mandibular muscles and their importance in orthodontics: a contemporary review. *Am J Orthod Dentofac Orthop*. 2005;128(6):774-780. https://doi.org/10.1016/j.ajodo.2004.09.023
29. King G. Biomedicine in orthodontics: from tooth movement to facial growth. *Orthod Craniofacial Res*. 2009;12(2):53-58. https://doi.org/10.1111/j.1601-6343.2009.01437.x
30. Gao C, Peng S, Feng P, Shuai C. Bone biomaterials and interactions with stem cells. *Bone Res*. 2017;5(1):17059. https://doi.org/10.1038/boneres.2017.59
31. Liu Y, Luo D, Wang T. Hierarchical structures of bone and bioinspired bone tissue engineering. *Small*. 2016;12(34):4611-632. https://doi.org/10.1002/smll.201600626
32. Podshivalov L, Fischer A, Bar-Yoseph PZ. On the road to personalized medicine: multiscale computational modeling of bone tissue. *Arch Comput Methods Eng*. 2014;21(4):399-479. https://doi.org/10.1007/s11831-014-9120-1
33. Manenti G, Capuani S, Fanucci E, et al. Diffusion tensor imaging and magnetic resonance spectroscopy assessment of cancellous bone quality in femoral neck of healthy, osteopenic and osteoporotic subjects at 3T: preliminary experience. *Bone*. 2013;55(1):7-15. https://doi.org/10.1016/j.bone.2013.03.004
34. Chung H, Wehrli FW, Williams JL, Kugelmass SD. Relationship between NMR transverse relaxation, trabecular bone architecture, and strength. *Proc Natl Acad Sci U S A*. 1993;90(21):10250-10254. https://doi.org/10.1073/pnas.90.21.10250
35. Wehrli FW, Hilaire L, Fernández-seara M, et al. Quantitative magnetic resonance imaging in the calcaneus and femur of women with varying degrees of osteopenia and vertebral deformity status. *J Bone Miner Res*. 2002;17(12):2265-2273. https://doi.org/10.1359/jbmr.2002.17.12.2265
36. Wehrli FW. Magnetic resonance of calcified tissues. *J Magn Reson*. 2013;229:35-48. https://doi.org/10.1016/j.jmr.2012.12.011
37. Rebuzzi M, Vinicola V, Taggi F, Sabatini U, Wehrli FW, Capuani S. Potential diagnostic role of the MRI-derived internal magnetic field gradient in calcaneus cancellous bone for evaluating postmenopausal osteoporosis at 3T. *Bone*. 2013;57(1):155-163. https://doi.org/10.1016/j.bone.2013.07.027
38. Capuani S, Manenti G, Iundusi R, Tarantino U. Focus on diffusion MR investigations of musculoskeletal tissue to improve osteoporosis diagnosis: a brief practical review. *Biomed Res Int*. 2015;2015:1-10. https://doi.org/10.1155/2015/948610
39. Gupta E, Sidhu MS, Grover S, Dabas A, Malik V, Dogra N. Measurement of perioral pressures at rest and its correlation with dental parameters in orthodontic patients with different occlusions. *J Clin Diagnostic Res*. 2019;13(6):zc13-zc18. https://doi.org/10.7860/jcdr/2019/41542.12917
40. Ogawa S, Yakabe M, Akishita M. Age-related sarcopenia and its pathophysiological bases. *Inflamm Regen*. 2016;36(1):1-6. https://doi.org/10.1186/s41232-016-0022-5
41. Berquist TH. *MRI of the Musculoskeletal System* (6th Edn). 2012.
42. Ariji Y, Nakayama M, Taguchi A, et al. Intramuscular changes of soft and hard areas after low-level static contraction of the masseter muscle and the correlations with muscle hardness and increase in water content: evaluations with sonographic elastography and magnetic resonance imaging. *Oral Surg Oral Med Oral Pathol Oral Radiol*. 2013;116(3):354-361. https://doi.org/10.1016/j.oooo.2013.05.017
43. Ringleb SI, Bensamoun SF, Chen Q, Manduca A, An KN, Ehman RL. Applications of magnetic resonance elastography to healthy and pathologic skeletal muscle. *J Magn Reson Imaging*. 2007;25(2):301-309. https://doi.org/10.1002/jmri.20817
44. Shiraishi T, Chikui T, Yoshiura K, Yuasa K. Evaluation of T2 values and apparent diffusion coefficient of the masseter muscle by clenching. *Dentomaxillofacial Radiol*. 2011;40(1):35-41. https://doi.org/10.1259/dmfr/15607259
45. Plooij JM, Maal TJJ, Haers P, Borstlap WA, Kuijpers-Jagtman AM, Bergé SJ. Digital three-dimensional image fusion processes for planning and evaluating orthodontics and orthognathic surgery. A systematic review. *Int J Oral Maxillofac Surg*. 2011;40(4):341-352. https://doi.org/10.1016/j.ijom.2010.10.013
46. Di Nardo D, Gambarini G, Capuani S, Testarelli L. Nuclear magnetic resonance imaging in endodontics: a review. *J Endod*. 2018;44(4):536-542. https://doi.org/10.1016/j.joen.2018.01.001
47. Van Vlijmen OJC, Kuijpers MAR, Bergé SJ, et al. Evidence supporting the use of cone-beam computed tomography in orthodontics.

16016343, 2021, S2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ocr.12520 by CochraneItalia, Wiley Online Library on [21/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

J Am Dent Assoc. 2012;143(3):241-252. https://doi.org/10.14219/jada.archive.2012.0148
48. Lee H, Zhao X, Song HK, Zhang R, Bartlett SP, Wehrli FW. Rapid dual-RF, dual-echo, 3D ultrashort echo time craniofacial imaging: a feasibility study. *Magn Reson Med*. 2019;81(5):3007-3016. https://doi.org/10.1002/mrm.27625
49. Niraj LK, Patthi B, Singla A, et al. MRI in dentistry- A future towards radiation free imaging–systematic review. *J Clin Diagnostic Res*. 2016;10(10):ZE14-ZE19. https://doi.org/10.7860/JCDR/2016/19435.8658
50. Zhang R, Lee H, Zhao X, et al. Bone-selective MRI as a non-radiative alternative to CT for craniofacial imaging. *Acad Radiol*. 2020;27(11):1515-1522. https://doi.org/10.1016/j.acra.2020.03.001
51. De Felice F, Di Carlo G, Saccucci M, Tombolini V, Polimeni A. Dental Cone beam computed tomography in children: clinical effectiveness and cancer risk due to radiation exposure. *Oncol*. 2019;96(4):173-178. https://doi.org/10.1159/000497059
52. Fathi Kazerooni A, Pozo JM, McCloskey EV, Saligheh Rad H, Frangi AF. Diffusion MRI for assessment of bone quality; a review of findings in healthy aging and osteoporosis. *J Magn Reson Imaging*. 2020;51(4):975-992. https://doi.org/10.1002/jmri.26973
53. Sanz-Requena R, Martínez-Arnau FM, Pablos-Monzó A, et al. The role of imaging biomarkers in the assessment of sarcopenia. *Diagnostics*. 2020;10(8):534. https://doi.org/10.3390/diagnostics10080534
54. Kaidonis JA, Ranjitkar S, Townsend GC, Koh KSB, Toh VKL, Brook AH. The mature stomatognathic system is a complex adaptive system. *Int J Des Nat Ecodynamics*. 2016;11:670-675. https://doi.org/10.2495/DNE-V11-N4-670-675
55. Di Pietro G, Scimeca M, Iundusi R, et al. Differences between muscle from osteoporotic and osteoarthritic subjects: in vitro study by diffusion-tensor MRI and histological findings. *Aging Clin Exp Res*. 2020;32(12):2489-2499. https://doi.org/10.1007/s40520-020-01483-6
56. Gotti C, Sensini A, Zucchelli A, Carloni R, Focarete ML. Hierarchical fibrous structures for muscle-inspired soft-actuators: a review. *Appl Mater Today*. 2020;20:100772. https://doi.org/10.1016/j.apmt.2020.100772
57. Castelvecchi D. Can we open the black box of AI? *Nature*. 2016;538(7623):20-23. https://doi.org/10.1038/538020a
58. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019;25(1):44-56. https://doi.org/10.1038/s41591-018-0300-7
59. Wood R, Baxter P, Belpaeme T. A review of long-term memory in natural and synthetic systems. *Adapt Behav*. 2012;20(2):81-103. https://doi.org/10.1177/1059712311421219
60. Handelman GS, Kok HK, Chandra RV, Razavi AH, Lee MJ, Asadi H. eDoctor: machine learning and the future of medicine. *J Intern Med*. 2018;284(6):603-619. https://doi.org/10.1111/joim.12822
61. Lecun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436-444. https://doi.org/10.1038/nature14539
62. Schmidhuber J. Deep learning in neural networks: an overview. *Neural Networks*. 2015;61:85-117. https://doi.org/10.1016/j.neunet.2014.09.003
63. Hand DJ. Classifier technology and the illusion of progress. *Stat Sci*. 2006;21(1):1-14. https://doi.org/10.1214/088342306000000060
64. Cimini G, Squartini T, Saracco F, Garlaschelli D, Gabrielli A, Caldarelli G. The statistical physics of real-world networks. *Nat Rev Phys*. 2019;1(1):58-71. https://doi.org/10.1038/s42254-018-0002-6
65. Asiri SN, Tadlock LP, Schneiderman E, Buschang PH. Applications of artificial intelligence and machine learning in orthodontics. *APOS Trends Orthod*. 2020;10:17-24. https://doi.org/10.25259/apos_117_2019
66. Li P, Kong D, Tang T, et al. Orthodontic treatment planning based on artificial neural networks. *Sci Rep*. 2019;9(1):2037. https://doi.org/10.1038/s41598-018-38439-w
67. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med*. 2019;380(14):1347-1358. https://doi.org/10.1056/nejmra1814259
68. Obermeyer Z, Lee TH. Lost in thought—the limits of the human mind and the future of medicine. *N Engl J Med*. 2017;377(13):1209-1211. https://doi.org/10.1056/nejmp1705348
69. Bellazzi R, Ferrazzi F, Sacchi L. Predictive data mining in clinical medicine: a focus on selected methods and applications. *Wiley Interdiscip Rev Data Min Knowl Discov*. 2011;1(5):416-430. https://doi.org/10.1002/widm.23
70. Barelli E, Ottaviani E, Auconi P, et al. Exploiting the interplay between cross-sectional and longitudinal data in class III malocclusion patients. *Sci Rep*. 2019;9(1):6189. https://doi.org/10.1038/s41598-019-42384-7
71. Chapelle O, Scholkopf B, Zien, A (Eds). *Semi-Supervised Learning* [Book reviews] (2006). IEEE Trans Neural Networks; 2009. https://doi.org/10.1109/tnn.2009.2015974
72. Mousavi SS, Schukat M, Howley E. Deep reinforcement learning: an overview. In: *Lecture Notes in Networks and Systems*; 2018. https://doi.org/10.1007/978-3-319-56991-8_32
73. Zitnik M, Nguyen F, Wang B, Leskovec J, Goldenberg A, Hoffman MM. Machine learning for integrating data in biology and medicine: principles, practice, and opportunities. *Inf Fusion*. 2019;50:71-91. https://doi.org/10.1016/j.inffus.2018.09.012
74. Barabasi AL, Barabási A-L. *Network Science*. Cambridge University Press; 2016.
75. Cheng F, Jia P, Wang Q, Lin CC, Li WH, Zhao Z. Studying tumorigenesis through network evolution and somatic mutational perturbations in the cancer interactome. *Mol Biol Evol*. 2014;31(8):2156-2169. https://doi.org/10.1093/molbev/msu167
76. Cheng F, Liu C, Lin CC, et al. A gene gravity model for the evolution of cancer genomes: a study of 3,000 cancer genomes across 9 cancer types. *PLoS Comput Biol*. 2015;11(9):e1004497. https://doi.org/10.1371/journal.pcbi.1004497
77. Di Carlo G, Gili T, Caldarelli G, Polimeni A, Cattaneo PM. A Community detection analysis of malocclusion classes from orthodontics and upper airway data. *Orthod Craniofacial Res*. 2021. Online ahead of print.
78. Scarselli F, Gori M, Tsoi AC, Hagenbuchner M, Monfardini G. The graph neural network model. *IEEE Trans Neural Networks*. 2009;20(1):61-80. https://doi.org/10.1109/TNN.2008.2005605
79. Gori M, Ingegneria D, Monfardini G. A New model for learning in graph domains. *Ijcnn*. 2005;2:729-734.
80. Zhou J, Cui G, Zhang Z, et al. Graph neural networks: A review of methods and applications. *arXiv*. 2018;1:57-81. https://doi.org/10.1016/j.aiopen.2021.01.001
81. Goyal P, Ferrara E. Graph embedding techniques, applications, and performance: a survey. *Knowledge-Based Syst*. 2018;151:78-94. https://doi.org/10.1016/j.knosys.2018.03.022
82. Peng Y, Choi B, Xu J. Graph embedding for combinatorial optimization: a survey. *arXiv*. 2020;1-39.
83. Féraud R, Clérot F. A methodology to explain neural network classification. *Neural Networks*. 2002;15(2):237-246. https://doi.org/10.1016/S0893-6080(01)00127-7
84. Zere E, Chaudhari PK, Sharan J, Dhingra K, Tiwari N. Developing class III malocclusions: challenges and solutions. *Clin Cosmet Investig Dent*. 2018;10:99-116. https://doi.org/10.2147/CCIDE.S134303

How to cite this article: Gili T, Di Carlo G, Capuani S, Auconi P, Caldarelli G, Polimeni A. Complexity and data mining in dental research: A network medicine perspective on interceptive orthodontics. *Orthod Craniofac Res*. 2021;24(Suppl. 2):16–25. https://doi.org/10.1111/ocr.12520

16016343, 2021, S2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ocr.12520 by CochraneItalia, Wiley Online Library on [21/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License