

RESEARCH ARTICLE

Predictors of tooth loss: A machine learning approach

Hawazin W. Elani^{1,2*}, André F. M. Batista³, W. Murray Thomson⁴, Ichiro Kawachi⁵, Alexandre D. P. Chiavegatto Filho³

1 Department of Oral Health Policy and Epidemiology, Harvard School of Dental Medicine, Boston, Massachusetts, United States of America, **2** Department of Health Policy and Management, Harvard T. H. Chan School of Public Health, Boston, Massachusetts, United States of America, **3** Department of Epidemiology, University of Sao Paulo Public Health School, Sao Paulo, Brazil, **4** Faculty of Dentistry, The University of Otago, Dunedin, New Zealand, **5** Department of Social and Behavioral Sciences, Harvard T. H. Chan School of Public Health, Boston, Massachusetts, United States of America

* hawazin_elani@hsdm.harvard.edu



OPEN ACCESS

Citation: Elani HW, Batista AFM, Thomson WM, Kawachi I, Chiavegatto Filho ADP (2021) Predictors of tooth loss: A machine learning approach. PLOS ONE 16(6): e0252873. <https://doi.org/10.1371/journal.pone.0252873>

Editor: Khanh N.Q. Le, Taipei Medical University, TAIWAN

Received: October 25, 2020

Accepted: May 24, 2021

Published: June 18, 2021

Copyright: © 2021 Elani et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: We used publicly available data from Centers for Disease Control and Prevention and did not have any special access or privileges that others would not have. National Health and Nutrition Examination Survey. <https://www.cdc.gov/nchs/hhanes/index.htm>.

Funding: Research reported in this publication was supported by the National Institute On Minority Health And Health Disparities of the National Institutes of Health (under Award Number K99MD012253 to HE). The content is solely the responsibility of the authors and does not

Abstract

Introduction

Little is understood about the socioeconomic predictors of tooth loss, a condition that can negatively impact individual's quality of life. The goal of this study is to develop a machine-learning algorithm to predict complete and incremental tooth loss among adults and to compare the predictive performance of these models.

Methods

We used data from the National Health and Nutrition Examination Survey from 2011 to 2014. We developed multiple machine-learning algorithms and assessed their predictive performances by examining the area under the receiver operating characteristic curve (AUC), accuracy, sensitivity, specificity, and positive and negative predictive values.

Results

The extreme gradient boosting trees presented the highest performance in the prediction of edentulism (AUC = 88.7%; 95%CI: 87.1, 90.2), the absence of a functional dentition (AUC = 88.3% 95%CI: 87.3, 89.3) and for predicting missing any tooth (AUC = 83.2%; 95%CI, 82.0, 84.4). Although, as expected, age and routine dental care emerged as strong predictors of tooth loss, the machine learning approach identified additional predictors, including socioeconomic conditions. Indeed, the performance of models incorporating socioeconomic characteristics was better at predicting tooth loss than those relying on clinical dental indicators alone.

Conclusions

Future application of machine-learning algorithm, with longitudinal cohorts, for identification of individuals at risk for tooth loss could assist clinicians to prioritize interventions directed toward the prevention of tooth loss.