Evaluating Deep Learning Algorithm Assessment of Digital Photographs for Dental Public Health Surveillance

New

Supporting Statement A

Program Official/Contact
Susan O Griffin, PhD
Economist, Division of Oral Health
National Center for Chronic Disease Prevention and Health Promotion
Centers for Disease Control and Prevention
P: 404-422-0867
F: 770-488-6080
Sig1@cdc.gov

8/25/2023

TABLE OF CONTENTS

ATTACHMENTS

1. Public Health Service Act [42 U.S.C. 247, 301]
2A. Data collection form part A (Paper clinical dental screening from)
2B. Data collection form part B (Sample photos of teeth)
2C. Screenshot of data collection application
2D. Invitation to schools to participate
2E. Consent form
2F. Notice of screening results
3A. 60-Day Federal Register Notice
3B. 60-day Federal Register Notice public comments and agency response
4. CDC Institutional Review Board determination

JUSTIFICATION SUMMARY

Goal of the project: The Centers for Disease Control and Prevention (CDC) is examining the validity and feasibility of using digital photos taken by non-dental professionals, which in turn would be analyzed by deep learning algorithms to assess youth's oral health status in lieu of examination by dental professionals. This deep learning assessment tool ultimately could be used by public health officials for dental public health surveillance at the local, state, and national levels. It is anticipated that obtaining information on dental conditions via deep learning algorithm assessment of digital images as opposed to human assessment will 1) be more cost-effective as it would not require dental personnel and 2) improve the accuracy of assessment due to minimal bias and less confounding factors associated with the examiner (e.g., subjective index and thresholding). This tool also would offer mobility, simplicity, and affordability for rapid and scalable adoption in community settings.

Intended use of the resulting data: The data collected from this request will be used to both 1) train the deep learning algorithms to assess tooth decay, sealant presence, and fluorosis severity and 2) test the accuracy of the algorithm assessments compared to the gold standard human examination.

Methods to be used to collect: CDC is funding the Colorado Department of Public Health and Environment to collect data for 1,000 students. The Colorado state health department will implement the collection by recruiting selected schools and dental examiners, gaining consent, arranging logistics, and collecting de-identified data from dental examination and photos taken by the dental examiners. CDC will provide dental examination protocols, and funding for researchers at Purdue University to test photo taking protocols and analyze data, and an expert in dental public health data collection to train the examiners. CDC met with cariologists and experts in fluorosis to review protocols for data collection. Data collected for each student will include 1) human assessment of fluorosis severity in the 6 upper anterior teeth and caries/sealant assessment of the occlusal surfaces of the 8 permanent molars (Attachments 2A) and 2) 9 digital photos of the upper anterior teeth acquired using an onboard smartphone camera and 24 digital photos of the occlusal surfaces of the 8 permanent molars acquired using an intraoral camera (Attachment 2B). The intra-oral camera will be linked to the smartphone via Wi-Fi. Digital photos of both the teeth and the completed paper screening form will be automatically saved and stored to the smartphone via a data collection app (Attachment 2C). After data have been collected for all students in the school, the data stored on the phone will be uploaded to a HIPAA compliant cloud storage that only can be accessed by examiners and designated CDC researchers with administrative rights.

The subpopulation to be studied: The study population will be comprised of middle-school Colorado students attending schools in areas with naturally

occurring fluoride in the tap water at or exceeding 1 part per million (ppm). Adolescents were chosen because some teeth at high risk for caries and fluorosis do not erupt until then. Colorado was chosen because it has several schools located in areas with high naturally occurring fluoride in tap water. In order to train the deep learning algorithms there will need to be a sufficient number of photos of moderate/severe fluorosis. Prevalence of fluorosis in its more severe forms is higher in areas with elevated tap water fluoride content.

How data will be analyzed: Two deep learning algorithms will be designed and trained to detect fluorosis severity and caries/sealant, respectively. The algorithm for fluorosis severity detection will consist of three modules of color correction, tooth segmentation, and classification. The algorithm for caries/sealant assessment will consist of two modules of tooth segmentation and classification. In both cases, separate training and validation datasets will be employed to strengthen the learning algorithms for training: 75% of the data points will be randomly selected as a training dataset and the remaining 25% will be blindly tested as a testing dataset. The performance of each application will be assessed by using a confusion matrix that can describe the performance of a multi-class classification model on the dataset where the ground true values are the assessments from the standardized human examiners. Performance criteria include precision (proportion of correct predictions among all predictions for a particular class), recall (proportion of positive class predictions made from all positive examples), and F1 score (Harmonic mean of precision and recall).

A. JUSTIFICATION

A1. Circumstances Making the Collection of Information Necessary

The Centers for Disease Control and Prevention (CDC) requests OMB approval for a new collection of data on dental caries, sealants, and fluorosis among 1,000 youth in Colorado over 1 year. These data would be used to train and test deep learning algorithms that will analyze digital phots of teeth to detect these dental conditions. CDC is authorized to collect the information under the Public Health Service Act, Title 42, Section 247b–14, Oral health promotion and disease prevention; and the Public Health Service Act, Title 42, Section 301 (Attachment 1).

The CDC Division of Oral Health (DOH) provides leadership to improve the nation’s oral health. One of its core functions is to monitor the burden of oral diseases and use of preventive measures especially among high-risk populations. To do this, DOH funds the National Centers for Health Statistics (NCHS) to collect data obtained during an examination by a dental professional in the National Health and Nutrition Examination Survey (NHANES: OMB No. 0920-0950) to estimate national prevalence of caries, sealants, and fluorosis among youth and funds 20

states to collect data on caries and sealant prevalence (OMB No. 0920-1346). Although dental caries is preventable it remains highly prevalent among US youth. National data on caries are used by DOH to track progress in meeting Healthy People goals and to inform and evaluate policy and resource allocation. Monitoring caries and fluorosis is also important as in 2015, the US Public Health Service noted the need for enhanced surveillance of fluorosis and caries to evaluate the impact of changing the recommended level of the optimal fluoride content of drinking water to prevent caries (US Department of Health and Human Services Federal Panel on Community Water Fluoridation 2015).

It is resource intensive to have dental personnel collect data and many states report difficulty in gaining access to schools since the COVID-19 pandemic. Having a tool that could collect this data via school personnel taking digital photographs would address these concerns. In addition, detecting fluorosis severity via human assessment is difficult. The level of fluorosis severity is based on the proportion of the tooth surface affected. An NCHS assessment on the quality of the fluorosis data in NHANES noted a biologically implausible increase in fluorosis prevalence from 2001–2004 to 2011–2014 in a synthetic cohort analysis and wide variation in prevalence across 2-year data collection cycles in 1999–2004 and 2011–2016. The NCHS data release warned users to strongly consider above issues when determining whether the data were appropriate to measure fluorosis prevalence and trends (National Center for Health Statistics and National Center for Chronic Disease Prevention and Health Promotion 2019). It is anticipated deep learning algorithms would improve the accuracy of fluorosis detection.

A2. Purpose and Use of the Information Collection

The purpose of this collection is to examine the validity and feasibility of using digital photos taken by non-dental professionals, which in turn would be analyzed by deep learning algorithms, to assess youth’s oral health status in lieu of examination by dental professionals.

The CDC Division of Oral Health is funding the Colorado Department of Public Health and Environment to collect data on dental caries, sealants and fluorosis for 1,000 Colorado middle-school students attending schools in areas with naturally occurring fluoride in the tap water at or exceeding 1 part per million (ppm). Colorado was chosen because it has several schools located in areas with high naturally occurring fluoride in tap water. Whereas caries and sealants are common, fluorosis in tis more severe forms is rare and thus limiting data collection to areas with higher fluoride content in tap water will ensure a sufficient number of fluorosis cases. The Colorado state health department will implement the collection by recruiting selected schools and dental examiners, gaining consent (Attachments 2D and 2E), arranging logistics, and collecting de-identified data from dental examination and photos taken by the dental examiners (Attachments 2A to 2C). CDC will provide dental examination protocols, and funding for researchers at Purdue University to test photo taking protocols and analyze data, and an expert in dental public health data collection to train the examiners.

Data collected for each student will include 1) human assessment of fluorosis severity in the 6 upper anterior teeth and caries/sealant assessment of the occlusal surfaces of the 8 permanent molars (Attachment 2A) and 2) 9 digital photos of the upper anterior teeth acquired using an onboard smartphone camera and 24 digital photos of the occlusal surfaces of the 8 permanent molars teeth acquired using an intraoral camera (Attachment 2B). The intra-oral camera will be linked to the smartphone via Wi-Fi. Digital photos of both the teeth and the completed paper screening form will be automatically saved and stored to the smartphone via a data collection app (Attachment 2C). After data have been collected for all students in the school, the data stored on the phone will be uploaded to a HIPAA compliant cloud storage that only can be accessed by examiners and designated CDC researchers with administrative rights. Upon completion of this project, the de-identified data will be stored on a CDC secured server.

The Division of Oral Health, CDC is funding Dr. Young Kim and research associates (School of Biomedical Engineering, Purdue University) to develop deep learning algorithm applications for dental public health surveillance. These deep learning algorithms will analyze digital images obtained with smart phones and intra-oral cameras. Dr. Kim and his team have developed similar applications to assess anemia, inflammation, and blood hemoglobin and are one of the first groups to extract spectroscopic information from red-green-blue (RGB) images for biomedical applications (Jeon et al. 2022; Ji et al. 2021; Kim et al. 2016a; Kim et al. 2016b; Kim et al. 2017; Kwak et al. 2021; Park et al. 2020). Deep learning technologies can offer advantages over visual assessment by human examiners, including minimal bias and less confounding factors associated with the examiner (e.g., subjective index and thresholding). These technologies also offer mobility, simplicity, and affordability for rapid and scalable adaptation in community-based settings.

The data collected from this request will be used to train and test the deep learning algorithms that will diagnose dental fluorosis severity, presence of dental caries, and presence of sealant using digital images. This technology ultimately can be used by public health officials to generate local, state, and national oral health surveillance data and enhance dental public health capacity to 1) evaluate the impact of the revised US Public Health Service standard for fluoride concentration in drinking water; 2) monitor children’s oral health status, trends, and disparities, 3) inform planning, implementation and evaluation of effective oral health interventions, programs, and policies; 4) measure progress toward Healthy People objectives; and 5) educate the public and policy makers regarding cross-cutting public health programs. It is anticipated that obtaining information on dental conditions via digital images as opposed to clinical assessment by human examiners will 1) be more cost-effective as it would not require dental personnel and 2) improve the accuracy of fluorosis severity assessment. The levels of fluorosis severity—normal, very mild, mild, moderate, and severe—are determined by the proportion of the tooth surface affected. For example, very

mild is present if less than ¼ of the surface is affected; mild, less than ½ of surface; and moderate, greater than ½ is affected. Quantifying the proportion of the tooth surface affected with a deep learning algorithm assessing a digital image would likely be more accurate than human judgement.

A3. Use of Improved Information Technology and Burden Reduction

The data collected in this request will be used to train and test deep learning algorithms to assess children's oral health status from digital images taken with a smart phone or intra-oral camera by non-dental professionals. The current method to collect these data requires dental professionals to travel and conduct oral assessments for each child in the school. A deep learning technology to assess oral health would likely reduce this burden as it only requires digital photos of select teeth taken by non-dental professionals, which could include onsite school employees. Because this technology is less resource intensive, it would allow public health entities to collect data at the community/school level, which would result in more effective and efficient prioritization of schools for school-based caries prevention programs. This technology also has the potential to increase assessment accuracy as it is less vulnerable to human examiner bias. Dr. Kim and his Purdue research team have developed tools to standardize digital images regardless of lighting and field conditions. These tools decrease the required sample size to train and test the deep learning algorithms.

A4. Efforts to Identify Duplication and Use of Similar Information

There are two surveys that use clinical examinations by trained dental personnel to assess children's oral health—the NCHS administered National Health and Nutrition Examination Survey, from which national estimates of disease prevalence can be obtained and the state department of health administered Basic Screening Survey, from which state estimates can be obtained. We are proposing a tool that could be used by either of these entities to collect the same data. The benefit of our proposed tool is that it is less resource intensive as it does not require dental personnel and potentially more accurate as it would not be subject to examiner bias. The Division of Oral Health has discussed this tool with NCHS, which has agreed to pilot the tool once it is validated and completed. Upon completion of pilot testing at the federal level CDC will next conduct piloting at the state and local levels. There are no similar tools to collect and analyze data for population oral health surveillance currently available.

A5. Impact on Small Businesses or Other Small Entities

The proposed collection does not include any small entities, only state governments.

A6. Consequences of Collecting the Information Less Frequently

This is a one-time data collection.

A7. Special Circumstances Relating to the Guidelines of 5 CRF 1320.5

This request fully complies with the regulation 5 CFR 1320.5.

A8. Comments in Response to the FRN and Efforts to Consult Outside the Agency

Part A: PUBLIC NOTICE

A 60-day Federal Register Notice was published in the Federal Register on June 5, 2023, Volume 88, Number 107, pages 36581–36583 to obtain comments from the public and affected agencies. CDC received two comments. See attachment 3B for comments and response.

Part B: CONSULTATION

CDC has consulted with experts in caries, dental sealant materials, fluorosis, and dental public health surveillance to review and provide input on protocols for data collection (Table A8).

Table A8: Consultations

Name	Title	Contact Information	Role
Dr. Margherita Fontana	Clifford Nelson Endowed Professor and Professor of Dentistry Director, Global Initiatives Program in Oral and Craniofacial Health, University of Michigan	Email: mfontan@umich.edu Phone: (734) 647-1225	Caries and sealant material consultant
Dr. Steven Levy	Wright-Bush-Shreves Endowed Professor of Research, University of Iowa	Email: Steven-levy@uiowa.edu Phone: (319) 335-7185	Caries, sealant material, and fluorosis consultant
Dr. John Warren	Graduate Program Director and Professor of Preventive & Community Dentistry, University of Iowa	Email: john-warren@uiowa.edu Phone: (319) 335-7205	Fluorosis and caries consultant
Dr. Jayanth Kumar	California State Dental Director	Email: Jayanth.Kumar@cdph.ca.gov Phone: (916) 324-1715	Caries, fluorosis, and surveillance consultant
Dr. Mark Moss	Division Director of Public Health Dentistry	Email: mossm17@ecu.edu Phone: (252) 737-7229	Caries, fluorosis, and

Name	Title	Contact Information	Role
	Associate Professor, Department of Foundational Sciences, Eastern Carolina University		surveillance consultant
Dr. Eugenio Beltran	Adjunct Professor Epidemiology and Health Promotion, New York University	E-mail: eba3@nyu.edu Phone: (212) 998-9800	Will train dental examiners to assess caries, fluorosis, and dental sealants

A9. Explanation of Any Payment or Gift to Respondents

Respondents do not receive an incentive.

A10. Protection of the Privacy and Confidentiality of Information Provided by Respondent

CDC has determined all collected data to be both de-identified and unlinkable (Attachment 4).
During data collection all data will be automatically saved and stored to the smartphone via a data collection application (Attachment 2C). After data have been collected for all students in a school, the data stored on the phone will be uploaded to a HIPAA compliant cloud storage that only can be accessed by examiners and designated CDC researchers. Upon completion of this project, the de-identified data will be stored on a secure CDC server. CDC will retain records in accordance with the applicable CDC records control schedule.

A11. Institutional Review Board (IRB) and Justification for Sensitive Questions

CDC determined that the data collected for this project are de-identified and unlinkable (Attachment 4). Although CDC review determined this project to be non-exempt human subjects research, it does not require review by the CDC Human Research Protection Office as no CDC employees are directly involved in data collection or analysis. Dr. Young Kim submitted the initial inquiry to Purdue University IRB on April 17, 2023 and submitted the official package to Purdue University IRB on August 18, 2023.

A12. Estimates of Annualized Burden Hours and Costs

Data will be collected for 1,000 students. Total estimated annualized burden hours are 827 and the corresponding total cost is $32,961 (Tables A12-1 and A12-2).

Based on feedback from CDC-funded states, it is estimated that it will take a parent/guardian 1 minute to complete a consent form, which translates to 17 hours and cost burden of $561 for 1,000 students.

Based on feedback from CDC-funded states It is estimated that it will take 16 minutes to acquire data per child (5 minutes for screening and 11 minutes for photo taking, infection control, and moving from class to screening area). This translates to 270 hours across all students. In addition to the 270 hours (45 hours per screener) spent with students, screeners will also require 192 hours (32 hours per screener) for training and calibration, 48 hours for travel to schools to screen students (8 hours per screener), and 30 hours of ongoing technical assistance (5 hours per screener). Total burden hours for screeners will be 540 and the corresponding total burden cost will be $32,400.

Average hourly wages are based on consultation with the Colorado state oral health program.

Table A12-1: Estimated Annualized Burden Hours

Type of Respondent	Form Name	Number of Respon-dents	Number of Responses per Resp-ondent	Avg Burden per Response (in hr)	Total Burden (in hr)
Parent or caretaker	Consent	1,000	1	1/60	17
Child	Screening/photo form	1,000	1	16/60	270
Screener	Screening/photo form*	6	1	90	540
Total					827

*Includes time for training, travel, screening and photos, and ongoing technical assistance

Table A12-2: Estimated Annual Burden Cost

Type of Respondent	Form Name	Number of Respon-dents	Total Burden (in hr)	Avg. Hourly Wage	Total Cost

Parent or caretaker	Consent	1,000	17	$33	$561
Child	Screening/ photo form	1,000	270	NA	NA
Screener	Screening/ photo form*	6	540	$60	$32,400
Total			827		$32,961

*Includes time for training, travel, screening and photos, and ongoing technical assistance

A13. Estimates of Other Total Annual Cost Burden to Respondents and Record Keepers

The total burden for other costs is $56,033 annualized for this one-time data collection, including other training related costs, oral health screening supplies, and costs for the Colorado state oral health program staff members' time involved (Table A13). These costs are based on input provided by the Colorado state oral health program. The collection is free for schools and families.

Table A13. Estimated Other Annualized Cost to the Respondents and Record Keepers

Item	**Cost per unit**	**Number of units**	**Total**
Training: Screener transportation/per diem	$1,400	6	$8,400
Training: volunteer lunch and training supplies	$81.4 per volunteer	30	$2,442
Disposable instruments	$1.53	1100	$1,683
Hygiene Kits	$1.5	1100	$1,650
Photocopies	$0.45	1000	$450
Translation	$500	6	$3,000
Project manager	$95,136 per FTE	0.1	$9,514
School coordinator	$64,816 per FTE	0.15	$9,722
Oral health consultant	$64,334 per	0.1	$6,433

	FTE		
Fluoridation engineer	$84,924 per FTE	0.15	$12,739
Total Other Costs			$56,033

A14. Annualized Cost to the Federal Government

The total estimated annualized cost to the federal government for this one-time data collection is $170,182 (Table A14). CDC is paying two full-time employees (epidemiologist and economist) for 20% time; a contractor oral health epidemiologist/examiner trainer for $20,400, and Dr. Young Kim and supporting staff for $82,172. These entities are developing the dental screening data collection protocols and photo taking protocols to ensure that digital images are comparable under different lighting and field conditions and that only the minimum number of digital images to train the deep learning algorithms. The contractor oral health epidemiologist/examiner trainer and Dr. Kim's team will provide examiners trainings on dental examinations, photo taking and data collection. In addition, Dr. Kim's team is developing a data collection application that links images from the intra-oral camera to the smartphone via Wi-Fi; and automatically saves and stores all digital photos of the teeth and the photos of the completed paper screening form to the smartphone. CDC will pay for the data collection application, smartphones and intraoral cameras for examiners at an estimated cost of $11,000.

Table A14. Estimated Annualized Cost to Federal Government

Type of Government Cost	Annualized Cost
Federal Staff	
Economist at 20% FTE	$30,262
Epidemiologist at 20% FTE	$26,348
Consultants	
Oral health epidemiologist/examiner trainer	$20,400
Dr. Kim and supporting staff	$82,172
Smartphones, intraoral cameras and data collection application	$11,000
TOTAL	$170,182

A15. Explanation for Program Changes or Adjustments

This request is a new data collection.

A16. Plans for Tabulation and Publication and Project Time Schedule

It is anticipated that this research (including both data collection and analysis) will be completed by January 2026. Upon completion, all de-identified data will be stored on a secured CDC server and be publicly available upon request. CDC, Dr.

Kim and research team and the oral health epidemiologist consultant plan on publishing articles related to protocols for taking digital images of teeth, the performance of the deep learning algorithms, and feasibility for use in public health surveillance in 2025 and 2026. The Division of Oral Health and National Center for Health Statistics are projecting piloting the deep learning algorithms in 2027–2028 NHANES data collection.

A17. Reason(s) Display of OMB Expiration Date is Inappropriate

The display of the OMB expiration date is appropriate.

A18. Exceptions to Certification for Paperwork Reduction Act Submission

There are no exceptions to the certification.

REFERENCES

Jeon HJ, Leem JW, Ji Y, Park SM, Park JW, Kim KY, Kim SW, L. KY. 2022. Cyber-physical watermarking with inkjet edible bioprinting. Advanced Functional Materials.

Ji Y, Kwak Y, Park SM, Kim YL. 2021. Compressive recovery of smartphone RGB spectral sensitivity functions. Optics Express 29:11947-11961

Kim T, Choi SH, Lambert-Cheatham N, Xu Z, Kritchevsky JE, Bertin FR, Kim YL. 2016a. Toward laboratory blood test-comparable photometric assessments for anemia in veterinary hematology. Journal of Biomedical Optics 21:107001.

Kim T, Kim JI, Visbal-Onufrak MA, Chapple C, Kim YL. 2016b. Nonspectroscopic imaging for quantitative chlorophyll sensing. Journal of Biomedical Optics 21:16008.

Kim T, Visbal-Onufrak MA, Konger RL, Kim YL. 2017. Data-driven imaging of tissue inflammation using RGB-based hyperspectral reconstruction toward personal monitoring of dermatologic health. Biomedical Optics Express 8:5282-5296

Kwak Y, Park SM, Ku Z, Urbas A, Kim YL. 2021. A pearl spectrometer. Nano Letters 21:921-930

National Center for Health Statistics, National Center for Chronic Disease Prevention and Health Promotion. 2019. Data quality evaluation of the dental fluorosis clinical assessment data from the National Health and Nutrition Examination Survey, 1999–2004 and 2011–2016. Vital Health Stat. 2(183).

Park SM, Visbal-Onufrak MA, Haque MM, Were MC, Naanyu V, Hasan MK, Kim YL. 2020. mHealth spectroscopy of blood hemoglobin with spectral super-resolution. Optica. 7:563-573.

US Department of Health and Human Services Federal Panel on Community Water Fluoridation. 2015. U.S. Public Health Service Recommendation for Fluoride Concentration in Drinking Water for the Prevention of Dental Caries. Public Health Rep. 130(4):318-331.