Leveraging Electronic Dental Record Data for Clinical Research in the National Dental PBRN Practices

Thankam Paul Thyvalikakath[1,2,*]   William D. Duncan[3,*]   Zasim Siddiqui[1]   Michelle LaPradd[4]
George Eckert[4]   Titus Schleyer[2,5]   Donald Brad Rindal[6]   Mark Jurkovich[6]   Tracy Shea[6]
Gregg H. Gilbert[7]   and National Dental PBRN Collaborative Group[7,**]

[1] Dental Informatics Core, Department of Cariology, Operative Dentistry & Dental Public Health, Indiana University School of Dentistry, IUPUI, Indianapolis, Indiana, United States
[2] Center for Biomedical Informatics, Regenstrief Institute, Inc., Indianapolis, Indiana, United States
[3] Environmental Genomics and Systems Biology, Lawrence Berkeley National Laboratory, Berkeley, California, United States
[4] Department of Biostatistics, Indiana University School of Medicine, Indianapolis, Indiana, United States
[5] Department of Internal Medicine, Indiana University School of Medicine, Indianapolis, Indiana, United States
[6] HealthPartners Institute, Minneapolis, Minnesota, United States
[7] Department of Clinical and Community Sciences, University of Alabama at Birmingham, Birmingham, Alabama, United States

Address for correspondence   Thankam Paul Thyvalikakath, DMD, MDS, PhD, Dental Informatics Core, Indiana University School of Dentistry, IUPUI, Regenstrief Institute, Inc., 1050 Wishard Boulevard, R2206, Indianapolis, IN 46202, United States (e-mail: tpt@iu.edu).

Abstract

Objectives   The aim of this study is to determine the feasibility of conducting clinical research using electronic dental record (EDR) data from U.S. solo and small-group general dental practices in the National Dental Practice-Based Research Network (network) and evaluate the data completeness and correctness before performing survival analyses of root canal treatment (RCT) and posterior composite restorations (PCR).

Methods   Ninety-nine network general dentistry practices that used Dentrix or EagleSoft EDR shared de-identified data of patients who received PCR and/or RCT on permanent teeth through October 31, 2015. We evaluated the data completeness and correctness, summarized practice, and patient characteristics and summarized the two treatments by tooth type and arch location.

Results   Eighty-two percent of practitioners were male, with a mean age of 49 and 22.4 years of clinical experience. The final dataset comprised 217,887 patients and 11,289,594 observations, with the observation period ranging from 0 to 37 years. Most patients (73%) were 18 to 64 years old; 56% were female. The data were nearly 100% complete. Eight percent of observations had incorrect data, such as incorrect tooth number or surface, primary teeth, supernumerary teeth, and tooth ranges, indicating multitooth procedures instead of PCR or RCT. Seventy-three percent of patients had dental insurance information; 27% lacked any insurance information. While gender was documented for all patients, race/ethnicity was missing in the dataset.

Keywords
► dental record
► data quality
► electronic health record

* Dr. Thankam Paul Thyvalikakath and Dr. William D. Duncan are the first authors.
** http://www.nationaldentalpbrn.org/collaborative-group.php.

Conclusion This study established the feasibility of using EDR data integrated from multiple distinct solo and small-group network practices for longitudinal studies to assess treatment outcomes. The results laid the groundwork for a learning health system that enables practitioners to learn about their patients' outcomes by using data from their own practice.

Background and Significance

The increasing availability of electronic health record (EHR) data is enabling significant insights into the health profiles and treatment outcomes of diverse patient cohorts in real-world clinical settings.[1] Harnessing EHR data for research can increase efficiency,[2,3] lower costs,[2] include the study of all patients, and facilitate longitudinal studies that are not possible with traditional methods.[4] In the last decade, electronic dental record (EDR) data have been increasingly used for clinical research and quality improvement purposes in academic settings and larger health care systems.[5–10] Studies in Europe and Canada have also utilized longitudinal data from solo and group practices to assess treatment outcomes, such as longevity of composite versus amalgam restorations.[11–14] These authors highlighted the need for more studies from practice-based contexts to study more-diverse patient cohorts and restorative procedures not performed in well-controlled randomized controlled trials. To the best of our knowledge, no study has determined the feasibility of utilizing EDR data in the United States to assess treatment outcomes in routine community clinical settings where most people receive their dental care.

When reusing EHR and EDR data for research, completeness and accuracy of data can be important limitations, because these data are recorded for patient care, not research purposes. It is well-recognized that data gathered for a specific purpose may not be beneficial for another purpose.[15] Therefore, it is important to assess the appropriateness of reusing the data for research and understand their limitations. Moreover, with the application of machine learning algorithms using EHR data in health care, it is important to examine the underlying data for potential misclassification, and missing data that may influence the model. Studies in medicine have established several data quality measures to assess EHR data.[16–18] Weiskopf et al. identified completeness, correctness, and currency as fundamental measures of EHR data quality. According to them, completeness refers to "whether or not a truth about a patient was present in the EHR." Correctness is "closeness of agreement between a data value and the true value"[19] and currency indicates whether the data are "representative of a patient state at a desired time of interest." They also emphasized the need to define correctness and completeness of EHR-derived datasets based on the context of the study.[16]

A few studies have evaluated the completeness and correctness of EDR data in academic institutions to identify gaps and improve dental students' training for patient care documentation.[20–22] These studies indicate the need to establish a systematic process to evaluate the data quality of EDR-derived datasets and to promote fidelity and reproducibility of secondary data analysis.

Objectives

The primary objective of this study was to determine the feasibility of extracting and utilizing EDR data from U.S. solo and small-group dental practices in the National Dental Practice-Based Research Network[23] for clinical research. The second objective was to evaluate the completeness and correctness of the data required to perform survival analyses of posterior composite restorations (PCR) and root canal treatments (RCT) performed on permanent teeth in network practices. To inform researchers and clinicians interested in leveraging EDR data for research and quality improvement purposes, we also report in detail the process followed to generate the dataset.

Methods

Practice Recruitment

We recruited network practices, which consisted of small group and solo general dental practices that met the following inclusion criteria: used EDR Dentrix (Henry Schein One, American Fork, Utah, United States) or EagleSoft (Patterson Dental, St Paul, Minnesota, United States) for at least 5 years; maintained electronic clinical information of at least existing conditions and treatment performed; placed at least one PCR on a permanent tooth on at least 100 patients or performed at least one RCT on a permanent tooth on at least 50 patients; had follow-up electronic data available for at least 2 years and had performed either of these procedures or both procedures between January 1, 2000 and October 31, 2015.

As part of enrolling in the network, practices completed an *enrollment questionnaire* regarding characteristics of their practice. The network's Regional Coordinators identified eligible practices by reviewing data from the network's enrollment questionnaire and contacted via email only those who used one of the two specified EDRs for at least 5 years. Upon receiving a response indicating interest, the coordinators confirmed that interested eligible practices met the remaining study inclusion criteria. The practices were provided a *brochure* that described the study and the *process required* to share de-identified data with researchers. If a practice had additional questions or concerns that the regional coordinators could not address, the principal investigator (T.P.T.) responded via email or telephone. Once practices agreed to participate, they signed the network-specific consents and necessary data sharing agreements with Indiana University (IU).

Data Transfer

Once consented, practices contacted their respective EDR vendors (Dentrix or EagleSoft) to share de-identified data based on the study criteria with the research team (►Fig. 1). After obtaining appropriate permissions from the practice owner, the vendors securely accessed and extracted the practice's data and generated de-identified databases containing the study data. The databases included de-identified treatment records of patients who had a PCR or RCT performed on permanent teeth through October 31, 2015. ►Supplementary Appendix Table S1 (available in the online version) lists the American Dental Association (ADA) code on dental procedures and nomenclature (CDT codes) for PCR and RCT used to identify the study records.[24] We included only procedures that contained the last four digits of the codes. The PCR codes include only resin or resin-based restorations placed on posterior teeth (►Supplementary Appendix Table S1, available in the online version). While codes 2385 to 2388 are retired codes used for PCR, codes 2391 to 2394 are current codes used for PCR as well as for restorations that used glass ionomer or other resin-based restorative materials. The vendors de-identified the data, which included offsetting the dates and redacting all identifiers according to the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule and transferred the
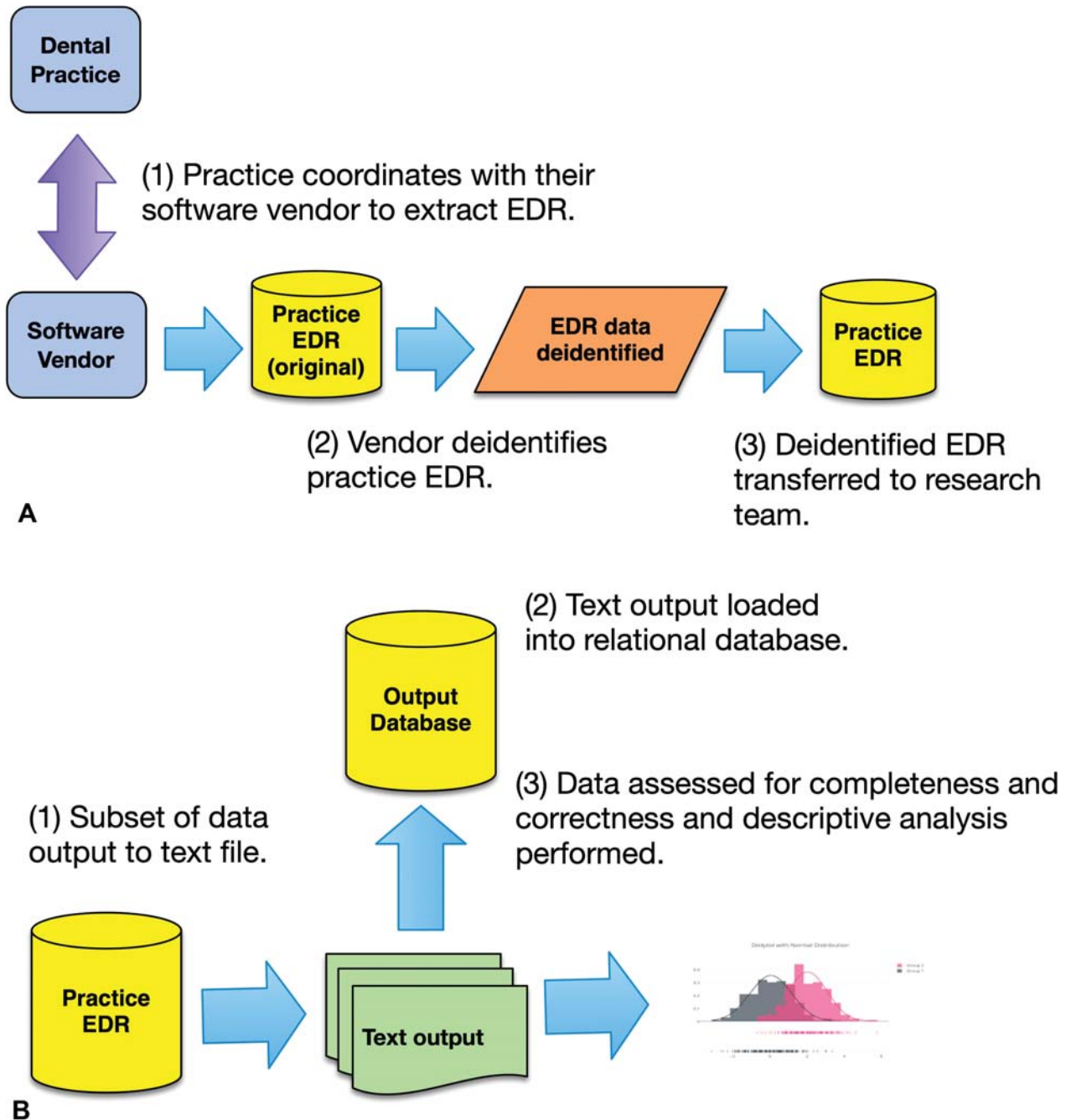
Fig. 1 Flow diagram demonstrating the steps involved in (A) extracting the electronic dental record data from 99 practices and transferring to the research team and (B) generating the final study dataset for analysis.

study data to researchers through an encrypted online portal maintained by IU.

Data Query and Creation of the Study Dataset

To aggregate data from Dentrix and EagleSoft databases, we first determined the tables and fields that contained the variables necessary to perform our research. These variables included demographic information, such as patients' gender, age (date of birth [DOB]), and the presence of dental insurance anytime during patients' dental care; provider information; and dental charting information regarding existing services, conditions, and completed CDT procedure codes at the tooth and tooth surface level. We wrote custom software scripts and structured query language queries to extract relevant study variables from the de-identified databases and stored the extracted data from each practice as a text file (►Fig. 1B). The individual data files were then loaded into a central repository (relational database) for use by the study team.

Data Quality Assessment

We used the work performed by Weiskopf et al[16,17] as a guide by which we defined two dimensions of data quality which we could assess: completeness and correctness of the EDR data variables needed to perform survival analysis of two treatment outcomes: RCT and PCR performed on permanent teeth.

For data completeness, we assessed the percentage of missing data for specific variables of interest: patient identification number, DOB, gender, tooth number, and date of procedure, and tooth surface for PCR procedures. Next, we removed all PCR and RCT procedures that were recorded as "existing services" because they were performed at another practice.

Although we could not directly compare the data to the actual patients, we defined the data to be reasonably correct if the distribution matched that of the overall population in previous studies.[25,26] We evaluated DOB, tooth number, tooth surface, and procedure codes for RCT and PCR. We first created frequency distributions of each of the variables and found that each of them were comparable to the population. In addition, we removed data which were not reasonable, using the

following metrics. A patients' DOB was assessed as incorrect if the patients' calculated age was less than 0 years or greater than 100 years. A tooth number was considered incorrect if it was outside the range 1 to 32, or tooth numbers were represented as tooth ranges (e.g., 2–8). Observations that represented tooth surfaces other than facial (F), buccal (B), mesial (M), distal (D), lingual (L), and occlusal (O) were considered incorrect. For procedure codes, we determined the correctness of PCR and RCT by considering the correctness of procedure codes entered for the respective tooth type. For example, a PCR procedure code was defined as incorrect if the procedure was performed on an anterior tooth. An RCT code was defined as incorrect if an RCT code was entered for the same tooth within 90 days, which is considered a follow-up procedure of the first RCT code for that tooth.

Data Analysis

We summarized the practice and patient characteristics using descriptive analysis. We performed a descriptive analysis of patients' age and gender based on their having received PCR, RCT, or both PCR and RCT treatments. For this analysis, we included only teeth 1 to 32 according to the Universal/National numbering system used in the United States.[27] Patient age was calculated as the difference between a patient's first PCR and/or RCT procedure and the patients' DOB. We calculated the observation time as number of years between a patient's first and last visit in a dental office. Patient characteristics were analyzed at the patient level, while PCR and RCT information were analyzed at the tooth level (tooth type, arch location [maxilla, mandible]) and PCR procedures, at the tooth surface level.

Results

Practices Recruited

We recruited 99 network practices that used Dentrix or EagleSoft EDR and who shared de-identified data of patients who received RCT and/or PCR on permanent teeth through October 31, 2015. Fifty-seven practices used Dentrix; the remaining 42 practices used EagleSoft. ►Table 1 displays the number of practices that shared their data across the six

Table 1 Number of study practices, number of teeth with posterior composite restorations (PCR) and root canal treatment (RCT) procedures across the 6 network regions

Network regions (number of practices)	PCR by number of surfaces				RCT by tooth type			
	One-surface	Two-surfaces	Three or >surfaces	Total	Anterior	Bicuspid	Molar	Total
Western (16)	54,406	50,343	23,780	128,529	1,460	2,908	4,415	8,783
Midwest (13)	43,403	37,315	15,390	96,108	1,938	3,371	4,701	10,010
Southwest (23)	81,792	77,126	23,925	182,843	2,585	4,885	7,575	15,045
South Central (16)	88,510	70,631	28,585	187,726	3,188	5,102	7,344	15,634
South Atlantic (16)	51,926	46,216	17,493	115,635	3,089	4,689	5,371	13,149
Northeast (15)	51,900	43,639	22,737	118,276	1,880	2,752	4,005	8,637
Total (99)	371,937	325,270	131,910	829,117	14,140	23,707	33,411	71,258

Table 2 Characteristics of participating practices

Practice type	Number reported[a] (n = 119)
General practitioner	98
Pediatric dentist	4
Endodontist	3
Oral/Maxillofacial surgeon	1
Orthodontist	5
Periodontist	4
Prosthodontist	4
Race/ethnicity of practice provider	Number (n = 98)
White	83
African American	2
American Indian or Alaska Native	1
Asian	8
Other	4
Age distribution of patients in 96 practices	Percentage of patients
1–18 y	18.8
19–44 y	29.5
45–64 y	33.3
65 y or older	18.4
Patients' dental visit characteristics	Percentage of patients
Regular	67.9
Irregular	15.1
Emergency only	10
Only one visit	7

[a]Numbers do not add to 99 because some practices reported multiple specialties in addition to general dentistry.

Table 3 Table showing the number (%) of patients by the observation time between the first and last visit

Time in years	Number of patients (%) N = 217,887	
No follow-up	32,922	(15.1)
Up to 5 y	91,289	(41.9)
>5 and ≤10 y	48,195	(22.1)
>10 and ≤15 y	31,078	(14.3)
>15 and ≤20 y	11,989	(5.5)
>20 y	2,414	(1.1)

network regions and the number of RCT and PCR procedures present in each region. Given the geographic distance between practices, we determined it was highly unlikely that practices shared patients.

Practice Characteristics

The distribution of practice types, as reported in the *enrollment questionnaire*, is shown in ►Table 2. Only two providers reported being Hispanic, and the gender breakdown was predominately male (81 males). The mean age of the providers was 49.4 years with a mean of 22.4 years of clinical experience, based on the number of years since graduation from dental school. A total of 96 out of 99 practices provided information on the enrollment questionnaire regarding the age distribution and insurance coverage of their patients (►Table 3).

The reported race/ethnicity of patients, based on enrollment questionnaire data, was mostly white (71.9%), followed by African American (12.3%), Asian (7.5%), American Indian (1.7%), and Hawaiian/Pacific Islander (0.9%). A total of 11.94% of patients were reported as being Hispanic. Gender percen-

tages of patients were not reported in the enrollment questionnaire. They reported 65% having private insurance, 9% having public insurance, 24% having no insurance, and 2% receiving a reduced fee. Finally, these practices reported the visit characteristics of their patients (►Table 2).

Data Quality and the Final Dataset

►Fig. 2 demonstrates the steps involved in assessing the data quality and generating the final dataset to perform descriptive and survival analysis. As shown in the figure, the final dataset consisted of 217,887 patients and 11,289,594 procedures. These patients' data included 0 to 37 years of observation time. As ►Table 3 shows, approximately 15% of the patients did not return after their first PCR or RCT. Forty-two percent of the patients had up to 5 years of observation time, and approximately 1% of the patients had more than 20 years of observation time after their first PCR or RCT.

Data Completeness

Gender was documented for all patients; race/ethnicity was missing in the EDR data from both systems. The completeness of variables needed to perform survival analyses of PCR was 99.73%, with 2,480 observations missing for tooth number, tooth surfaces, or date of treatment. The completeness of variables needed to perform survival analyses of RCT was 99.61%, with 284 observations missing for tooth number or date of treatment. Of the 217,887 patients, 159,028 patients (73%) had documented insurance information at least once during their dental care history; the remaining patients (27%) lacked documentation of any insurance information. About 57,685 (27%) of the 217,887 patients did not have any insurance coverage during their dental care.

Data Correctness

Missing Teeth, Primary Teeth, and Supernumerary Teeth for PCR

We considered the tooth number field incorrect for PCR procedures if the following were present: tooth number was missing, tooth number indicated primary teeth, supernumerary teeth (indicated as larger than 32), and tooth numbers indicated as anterior teeth. ►Fig. 2 and ►Supplementary Appendix Table S2 (available in the online version) list the number of observations with missing tooth number, primary
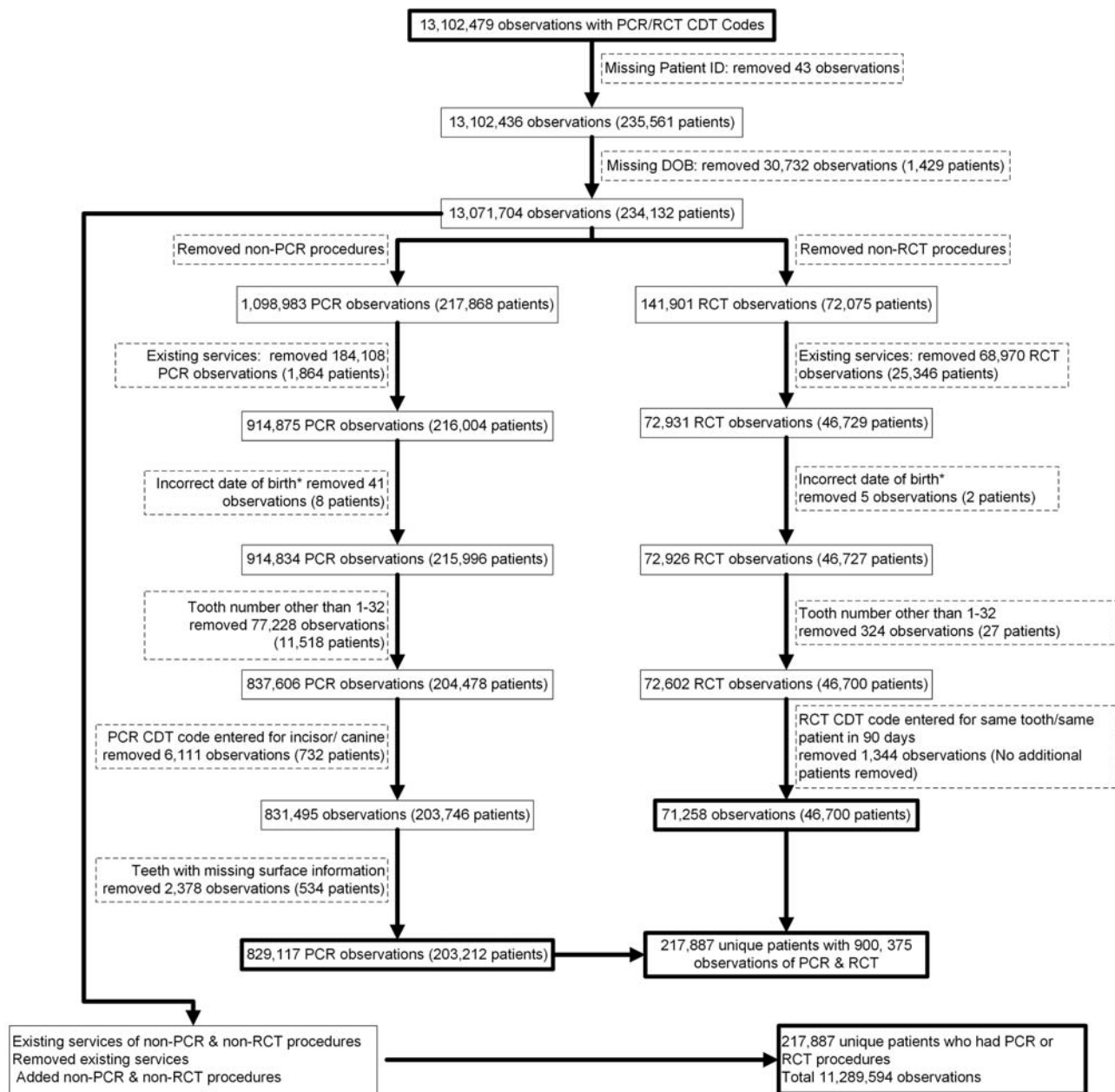
Fig. 2 Generating the final dataset. CDT code: code on dental procedures and nomenclature; DOB, date of birth; PCR, posterior composite restorations; RCT, root canal treatment.

teeth, and supernumerary teeth. The primary teeth and supernumerary teeth could have been present in the treatment history of patients less than 18 years who had a PCR and/or RCT on a permanent tooth. ►**Fig. 2** shows the number of observations where PCR code was incorrectly applied for anterior teeth restorations and observations with missing tooth surface information. The final dataset included 829,117 PCR observations in 203,212 patients, as shown in ►**Fig. 2**.

Missing Teeth, Primary Teeth, Supernumerary Teeth, and Tooth Ranges for RCT
We found instances of missing tooth number, presence of primary teeth numbers (B, E, F, K, O, and T: one each), tooth ranges (►**Fig. 2** and ►**Supplementary Appendix Table S3**, available in the online version) typically indicating multitooth

procedures and supernumerary teeth. We removed these 324 observations. As ►**Fig. 2** shows, 1,344 RCT observations entered for the same tooth and the same patient in 90 days, which we considered the continuation of the same procedure on the same tooth. After removing these observations, the final dataset included 71,258 RCT observations in 46,700 patients.

We also found 949 observations where the current version of CDT codes for RCT was coded incorrectly by tooth type. ►**Supplementary Appendix Table S4** (available in the online version) lists the number of RCT observations coded for the wrong tooth type. This error may have occurred due to changes in the CDT codes for RCT. The present RCT codes are indicated by code 3310 for anterior tooth, 3320 for bicuspid, and 3330 for molar teeth. Previously, RCT CDT codes were based on the number of canals, with 3310 for a tooth with one canal, 3320

for a tooth with two canals, and 3330 for a tooth with 3 or more canals. We fixed this error by recoding the RCT procedure for each tooth type according to the tooth number listed in the dataset, which is 100% reliable (►Supplementary Appendix Table S4, available in the online version). For example, tooth number 7, which is the upper right lateral incisor is matched to the code 3310 for anterior tooth.

In our examination of the CDT codes, we discovered that individual practices could enter the CDT codes as free text, and this created variations among the use of CDT codes. These differences included the use of letters either before or after the code (►Supplementary Appendix Table S5, available in the online version). In some cases, these altered codes still indicated the same procedure, while in others, the use of letters indicated a separate but related procedure. We determined that using all CDT codes that ended in the proper four-digit numerical code would be equivalent to the four-digit code alone because some EDR systems coded CDT codes using a "D" or "0" at the beginning. Because the data spanned many years, we had to account for multiple versions of CDT codes.

Final Dataset

The final dataset consisted of 217,887 unique patients with 900,375 observations for PCR and RCT procedures. These 217,887 patients had a total of 11,289,594 observations after including all dental visits. ►Fig. 2 and ►Supplementary Appendix Table S6 (available in the online version) show the number of patients by age and gender who received PCR, RCT, and both PCR and RCT. Nineteen percent of patients were less than 18 years old, 46% of patients were ages 18 to 44 years, 27% of patients were ages 45 to 64 years, and 8% of patients were 65 years or older.

PCR ($N = 829,117$) were performed almost equally on the maxillary and mandibular teeth (►Supplementary Appendix Table S7, available in the online version). Sixty-six percent of the PCR were performed on molars, 52% of which were mandibular molars. Among the PCR performed on bicuspids, 55% were performed on the maxillary teeth. For the RCT ($N = 71,258$), 55% were performed on the maxillary teeth; 47% of the RCT was performed on the molar teeth, followed by bicuspids and anterior teeth; 57% of the molar RCT was performed on the mandibular molar teeth, and 59% of the bicuspid RCT and 76% of anterior RCT were performed on the maxillary bicuspids and anterior maxillary teeth, respectively (►Supplementary Appendix Table S7, available in the online version).

Discussion

This study demonstrated the feasibility of utilizing EDR data integrated from multiple distinct solo and small-group network practices for longitudinal studies to assess treatment outcomes. We established a process to extract de-identified data from practices that used two different EDR systems and assessed the data completeness and correctness to perform survival analysis of RCT and PCR. Major findings of this investigation include: the near 100% completeness and high percentage of correct data; characterization of the incorrect data that may occur for these data types; presence of longer observation times of patients who received a PCR or RCT; unavailability of race/ethnicity data and the ability to study insured and uninsured populations of patients who sought treatment in these practices. Through this study, we intended to highlight the importance of understanding the data before analysis to identify the biases that could occur due to the health care recording process.

As ►Fig. 2 demonstrates, less than 1% of the patients had a missing DOB, and less than 0.1% had incorrect DOB. Approximately 8% of PCR observations had incorrect tooth numbers (►Supplementary Appendix Table S2, available in the online version), which were mostly primary or supernumerary tooth numbers. They were considered incorrect because we intended to limit survival analysis to permanent teeth. The rates of missing tooth numbers, incorrect/missing tooth surface information, and incorrect entry of PCR/RCT code (►Fig. 2) were low, at less than 1%. We also observed less variation in these data across practices. These results demonstrate that EDR data from solo and small-group practices in the United States can be utilized to study patient populations in real-world settings and to assess the longevity of dental procedures at the practice level. Moreover, we can analyze differences in outcomes across different geographic regions.

A major concern regarding the use of EDR data for research is the loss of patients following the initial patient visit to the dental office. Our results indicated that 85% of patient records had at least 5 years of observation time, out of which, 22% had 5 to 10 years, 14% 10 to 15 years, and 5% had 15 to 20 years of observation time (►Table 3). Having access to information over such long time enables researchers to perform longitudinal studies such as survival analysis at much lower costs compared with prospective clinical studies. Only 15% of patient records did not have observations following the initial date of performing a PCR or RCT. This finding is comparable to reports that private practices experience approximately 17% of patient attrition annually[28] due to patients changing dental practices. Our results indicate the potential of using EDR data for researchers and practitioners to study long-term outcomes of various treatments, which has not been possible previously.

Contrasted to claims data, this dataset provided access to both insured and noninsured populations, permitting us to study treatment outcomes in patients without dental insurance coverage. In this study, we retrieved 72% of the patients' dental insurance information; only 27% of the patient records lacked insurance documentation. Also, 27% of the patients did not have insurance coverage (coincidentally the same percentage as those who did not have any information about coverage), which is consistent with existing reports of dental insurance coverage among U.S. adults. We believe insurance information was not available specific to the date of treatment because they were stored as a patient characteristic in the EDR systems. Therefore, obtaining de-identified data without patient identifiers may have prevented us from obtaining treatment-specific insurance coverage. Previous studies utilizing EDR data have worked with limited datasets that included actual dates of treatment and birth.[8,10] Further work is needed to obtain treatment-specific insurance coverage because finances play a major role in the patient's decision to receive dental treatment.

As with any study, we encountered some limitations. First, race and ethnicity were not available in these EDR data. The vendors confirmed that their EDR systems were not designed to capture race and ethnicity. This information was reported in EDR data-based studies from academic and large health care systems. While race and ethnicity may not be essential information for patient care purposes, future studies would surely want to examine differences among different demographic groups to evaluate health disparities and health outcomes for subpopulations.[29] Second, we did not assess the decayed missing filled teeth and caries risk status because these data may be recorded as structured data or in clinical notes requiring more data processing. Future work should consider retrieving information such as patient and practitioner characteristics such as oral hygiene, frequency, and history of performing certain procedure, which can be derived from EDR data. Also, a better understanding of documentation practices is warranted because the potential for variations in documenting diagnosis and findings is high across practices. We also wanted to focus on the retrieval and completeness of data essential to perform survival analyses of treatments.

Finally, network members are not recruited randomly, so factors associated with network participation (e.g., an interest in clinical research) may make network dentists unrepresentative of dentists at large. While we cannot assert that network dentists are entirely representative, we can state that they have much in common with dentists at large, while also offering substantial diversity in these characteristics.[30,31] This assertion is warranted because: (1) substantial percentages of network general dentists are represented in the various response categories of the characteristics in the enrollment questionnaire; (2) findings from several network studies document that network general dentists report patterns of diagnosis and treatment are similar to patterns determined from non-network general dentists,[32–35] and the similarity of network dentists to non-network dentists based on characteristics was reported in the 2010 ADA Survey of Dental Practice.[36]

Conclusion

Despite these drawbacks, this study demonstrated the feasibility of leveraging EDR data to establish a learning health system for practitioners to gain insights about their patients' treatment outcomes. We also established a process for solo- and small group practices to share their data for research, and to learn about treatment outcomes. Thus, this work has laid the groundwork to establish a clinical data warehouse of solo and small-group practices data similar to the Big Mouth data[37] and virtual data warehouse repositories[38] that include dental data from academic and health care system settings respectively in the United States.

Clinical Relevance Statement

This study established a patient cohort using electronic dental record (EDR) data from multiple community practices who use different EDR systems to assess the longevity of two commonly performed dental treatments. We also describe methods to assess the completeness and correctness of the data. The results from this study laid the groundwork to establish a learning health system that enables practitioners to learn about their patients' outcomes by utilizing data from their own practice.

Multiple Choice Questions

1. What are the limitations of reusing Electronic Health Record (EHR) and Electronic Dental Record (EDR) data for research?
   a. Difficulty accessing data.
   b. Lack of structured data.
   c. Questionable completeness and accuracy of data.
   d. Absence of data quality measures.

   Correct Answer: The correct answer is option c, questionable completeness and accuracy of data. The data recorded are for patient care and not for research. Therefore, the data recorded for one purpose may not fit to use for another purpose. Thus, reuse of EHR and EDR data has certain limitations.

2. What is the main advantage of using electronic dental record data from community practices for clinical research?
   a. Practitioners learn about different treatment outcomes.
   b. Identify possible reasons for treatment failures.
   c. Enable study of diverse patient cohorts undergoing different treatments.
   d. Learn patient's adherence to treatment.

   Correct Answer: The correct answer is option c, enables study of diverse patient cohorts undergoing different treatments. Studies using electronic dental data from community practices offer an opportunity to include diverse patient cohorts and treatments provided or performed in real world settings.

Note

An internet site devoted to details about the nation's network is located at http://NationalDentalPBRN.org. We are grateful to the network's regional node coordinators, who function as the "face" of the network in recruiting, training, and interacting with the network's practitioners (Midwest Region: T.S., RDH, BSDH; Western Region: Stephanie Hodge, MA; Northeast Region: Christine O'Brien, RDH; South Atlantic Region: Hanna Knopf, BA, and Deborah McEdward, RDH, BS, CCRP; South Central Region: Shermetria Massengale, MPH, CHES, and Ellen Sowell, BA; Southwest Region: Stephanie Reyes, BA, Meredith Buchberg, MPH, and Colleen Dolan, MPH), and the network's program manager (Andrea Mathews, BS, RDH), and program coordinator (Terri Jones).

Protection of Human and Animal Subjects

Our study was determined to be nonhuman subject research by the Indiana University Institutional Review Board (IRB). We also received IRB approvals or exemptions separately from the six Network regions: the Midwest,

South Central, Southwest, Northeast, Western, and South Atlantic regions. The informed consent of all human subjects who participated in this investigation was obtained after the nature of the procedures had been explained fully.

Conflict of Interest

None declared.

References

1 Singh G, Schulthess D, Hughes N, Vannieuwenhuyse B, Kalra D. Real world big data for clinical research and drug development. Drug Discov Today 2018;23(03):652–660

2 Kohane IS. Using electronic health records to drive discovery in disease genomics. Nat Rev Genet 2011;12(06):417–428

3 Medicine Io. Roundtable on evidence-based medicine. In: Olsen LA, Aisner D, McGinnis JM, eds. The Learning Healthcare System: Workshop Summary. Washington, DC: National Academies Press (U.S.); 2007

4 Brownstein JS, Sordo M, Kohane IS, Mandl KD. The tell-tale heart: population-based surveillance reveals an association of rofecoxib and celecoxib with myocardial infarction. PLoS One 2007;2(09):e840

5 Maupomé G, Peters D, Rush WA, Rindal DB, White BA. The relationship between cardiovascular xerogenic medication intake and the incidence of crown/root restorations. J Public Health Dent 2006;66(01):49–56

6 Acharya A, VanWormer JJ, Waring SC, Miller AW, Fuehrer JT, Nycz GR. Regional epidemiologic assessment of prevalent periodontitis using an electronic health record system. Am J Epidemiol 2013;177(07):700–707

7 Rindal DB, Flottemesch TJ, Durand EU, Godlevsky OV, Schmidt AM, Gilbert GH; National Dental PBRN Collaborative Group. Practice change toward better adherence to evidence-based treatment of early dental decay in the National Dental PBRN. Implement Sci 2014;9:177

8 Chatzopoulos GS, Wolff LF. Implant and endodontic treatment selection are influenced by patients' demographic characteristics, insurance status, and medical history: a retrospective cohort study. Quintessence Int 2017;48(09):753–764

9 Siddiqui Z, Wang Y, Makkad P, Thyvalikakath T. Characterizing restorative dental treatments of Sjögren's syndrome patients using electronic dental records data. Stud Health Technol Inform 2017;245:1166–1169

10 Kumar SV, Bangar S, Neumann A, et al. Assessing the validity of existing dental sealant quality measures. J Am Dent Assoc 2018;149(09):756–764

11 Laske M, Opdam NJM, Bronkhorst EM, Braspenning JCC, Huysmans MCDNJM. Longevity of direct restorations in Dutch dental practices. Descriptive study out of a practice based research network. J Dent 2016;46:12–17

12 Opdam NJM, Bronkhorst EM, Loomans BAC, Huysmans MCDNJM. 12-year survival of composite vs. amalgam restorations. J Dent Res 2010;89(10):1063–1067

13 Palotie U, Eronen AK, Vehkalahti K, Vehkalahti MM. Longevity of 2- and 3-surface restorations in posterior teeth of 25- to 30-year-olds attending Public Dental Service-A 13-year observation. J Dent 2017;62:13–17

14 Wu X, Al-Abedalla K, Abi-Nader S, Daniel NG, Nicolau B, Tamimi F. Proton pump inhibitors and the risk of osseointegrated dental implant failure: a cohort study. Clin Implant Dent Relat Res 2017;19(02):222–232

15 Verheij RA, Curcin V, Delaney BC, McGilchrist MM. Possible sources of bias in primary care electronic health record data use and reuse. J Med Internet Res 2018;20(05):e185

16 Weiskopf NG, Bakken S, Hripcsak G, Weng C. A data quality assessment guideline for electronic health record data reuse. EGEMS 2017;5(01):14

17 Weiskopf NG, Hripcsak G, Swaminathan S, Weng C. Defining and measuring completeness of electronic health records for secondary use. J Biomed Inform 2013;46(05):830–836

18 Wright A, McCoy AB, Hickman TT, et al. Problem list completeness in electronic health records: a multi-site study and assessment of success factors. Int J Med Inform 2015;84(10):784–790

19 Zozus MN, Hammond WE, Green BB, et al. Assessing Data Quality for Healthcare Systems Data Used in Clinical Research (Version 1.0). NIH Collaboratory; 2014

20 Burcham WK, Romito LM, Moser EA, Gitter BD. Analyzing medication documentation in electronic health records: dental students' self-reported behaviors and charting practices. J Dent Educ 2019;83(06):687–696

21 Levitin SA, Grbic JT, Finkelstein J. Completeness of electronic dental records in a student clinic: retrospective analysis. JMIR Med Inform 2019;7(01):e13008

22 Barnett GO, Winickoff RN. Quality assurance and computer-based patient records. Am J Public Health 1990;80(05):527–528

23 Gilbert GH, Williams OD, Korelitz JJ, et al; National Dental PBRN Collaborative Group. Purpose, structure, and function of the United States National Dental Practice-Based Research Network. J Dent 2013;41(11):1051–1059

24 American Dental Association. CDT 2018: Dental Procedure Codes. American Dental Association, Chicago, Illinois, United States; 2017

25 Gilbert GH, Tilashalski KR, Litaker MS, McNeal SF, Boykin MJ, Kessler AW; DPBRN Collaborative Group. Outcomes of root canal treatment in Dental Practice-Based Research Network practices. Gen Dent 2010;58(01):28–36

26 Gordan VV, Riley JL III, Rindal DB, et al; National Dental Practice-Based Research Network Collaborative Group. Repair or replacement of restorations: a prospective cohort study by dentists in The National Dental Practice-Based Research Network. J Am Dent Assoc 2015;146(12):895–903

27 American Dental Association. Current Dental Terminology Third Edition (CDT-3). Chicago, IL: American Dental Association; 1999

28 Wells B Patient attrition: 3 steps to finding a hidden gold mine. Endeavor Business Media, Fort Atkinson, Wisconsin. Available at: https://www.dentaleconomics.com/practice/article/16386274/patient-attrition-3-steps-to-finding-a-hidden-gold-mine. Accessed November 24, 2019

29 Dye BA, Duran DG, Murray DM, et al. The importance of evaluating health disparities research. Am J Public Health 2019;109(S1): S34–S40

30 Makhija SK, Gilbert GH, Rindal DB, et al; DPBRN Collaborative Group. Practices participating in a dental PBRN have substantial and advantageous diversity even though as a group they have much in common with dentists at large. BMC Oral Health 2009;9:26

31 Makhija SK, Gilbert GH, Rindal DB, Benjamin PL, Richman JS, Pihlstrom DJ; DPBRN Collaborative Group. Dentists in practice-based research networks have much in common with dentists at large: evidence from the Dental Practice-Based Research Network. Gen Dent 2009;57(03):270–275

32 Gordan VV, Garvan CW, Richman JS, et al; DPBRN Collaborative Group. How dentists diagnose and treat defective restorations: evidence from the dental practice-based research network. Oper Dent 2009;34(06):664–673

33 Norton WE, Funkhouser E, Makhija SK, et al; National Dental Practice-Based Research Network Collaborative Group. Concor-dance between clinical practice and published evidence: findings from The National Dental Practice-Based Research Network. J Am Dent Assoc 2014;145(01):22–31

34 Gilbert GH, Riley JL, Eleazer PD, Benjamin PL, Funkhouser E; National Dental PBRN Collaborative Group. Discordance between presumed standard of care and actual clinical practice: the example of rubber dam use during root canal treatment in the National Dental Practice-Based Research Network. BMJ Open 2015;5(12):e009779

35 Gordan VV, Garvan CW, Heft MW, et al; DPBRN Collaborative Group. Restorative treatment thresholds for interproximal pri-mary caries based on radiographic images: findings from the Dental Practice-Based Research Network. Gen Dent 2009;57 (06):654–663

36 American Dental Association. 2010 Survey of Dental Practice: Characteristics of Dentists in Private Practice and Their Patients. Chicago, IL: American Dental Association; 2012

37 Walji MF, Kalenderian E, Stark PC, et al. BigMouth: a multi-institutional dental data repository. J Am Med Inform Assoc 2014;21(06):1136–1140

38 Fellows JL, Rindal DB, Barasch A, et al; DPBRN Collaborative Group. ONJ in two dental practice-based research network regions. J Dent Res 2011;90(04):433–438