

Assessing the feasibility of using an electronic records database system in use in a group of private chiropractic clinics for practice-based research

Edward F. Owens, Jr., MS, DC¹

Joseph Esposito, DC²

Ronald S. Hosek, DC, MPH, PhD¹

Stephanie G.B. Sullivan, DC, PhD¹

Objective: To explore the feasibility of collecting aggregated patient data from the electronic records of a group of private practices and to determine the suitability of the data for comparative effectiveness or other practice-based research (PBR).

Methods: Assess the type and quality of health-related variables contained in a commercially available electronic records system (Vitalogics) in use in consenting chiropractor's offices. Descriptively analyze baseline patient records to identify demographic variables, vital signs, case types and diagnoses.

Évaluation des possibilités d'utilisation d'une base de données sur les dossiers électroniques par un groupe de cliniques de chiropratique privées aux fins de recherches fondées sur la pratique

Objectif : Examiner la possibilité de recueillir des macrodonnées de patients dans les dossiers électroniques d'un groupe de cliniques privées et déterminer leur utilité pour mener une recherche comparative d'efficacité ou mener d'autres recherches fondées sur la pratique.

Méthodologie : Évaluer le type et la qualité des variables de santé contenus dans le logiciel de gestion de dossiers médicaux informatisés de Vitalogics en usage dans les cabinets des chiropraticiens consentants. Effectuer une analyse descriptive de dossiers de patients de référence pour trouver des variables démographiques, des signes vitaux, des cas types et des diagnostics.

¹ Dr. Sid E. Williams Center for Chiropractic Research, Life University

² AlignLife

Corresponding author: Edward F. Owens, Jr., Dr. Sid E. Williams Center for Chiropractic Research, Life University. Marietta GA, USA
E-mail: edward.owens@life.edu

© JCCA 2021

The study was supported in part with funds from the Health Missions foundation and approved by the Institutional Review Board of Life University in Marietta, GA, USA. The authors have no other disclaimers, competing interests, or sources of support or funding to report in the preparation of this manuscript.

Results: *Of the 46,000 individual patient records that were analyzed, only a fraction had usable demographic data while more than half contained ICD-9 codes, including records for non-insurance case-types. None contained outcome variables.*

Conclusion: *We did not find that the electronic record system we examined had the types of demographic and outcomes variables that would be useful for comparative effectiveness research. The addition of special fields to code chief complaint and health status indicators not typically included in an electronic records system would be most useful.*

(JCCA. 2021;65(1):94-104)

KEY WORDS: chiropractic, database, effectiveness, health record, practice-based research

Introduction

Health-services researchers and policymakers are continuously looking for ways to reduce costs of health-care in the USA while still making sure everyone has access to quality care. Chiropractic care is considered a healthcare option with similar efficacy for neck pain and low back pain as physical therapy or standard medical care. Current European and North American guidelines recommend a trial of chiropractic care, which includes spinal manipulative therapy (SMT) as well as other modalities such as strengthening and flexibility exercises, education and advice on self-management strategies; all within the scope of practice of chiropractors for both of these afflictions.¹⁻⁵

Current guidelines are mostly based on systematic reviews of randomized clinical trials (RCTs) of previous guidelines. While RCTs are generally considered to be the gold standard for studying the efficacy of health care interventions, many such studies are carried out in institutional settings which are unlike the typical chiropractor's office. Hence, clinical trial results don't necessarily reflect a true measure of the effectiveness of chiropractic

Résultats : *Des 46 000 dossiers de patients examinés, seule une partie contenaient des données démographiques utiles alors que moins de la moitié contenaient les codes de la classification ICD-9, et des données sur des cas types non assurés. Aucun ne contenait de variables de résultats.*

Conclusion : *À notre avis, le système de gestion de dossiers électroniques examinés ne fournit pas des types de variables démographiques et de variables de résultats qui seraient utiles pour mener une recherche comparative d'efficacité. L'ajout de champs spéciaux permettant de saisir le code du symptôme principal et des indicateurs de l'état de santé, ce qui d'habitude ne se trouve pas dans un logiciel de gestion de dossiers électroniques, s'avérerait très utile.*

(JACC. 2021;65(1):94-104)

MOTS CLÉS : chiropratique, base de données, efficacité, dossier médical, recherche fondée sur la pratique

care nor do they provide a true picture of real-world outcomes and costs.^{6,7} Blanchette *et al.*, found only a limited number of studies of effectiveness or cost-effectiveness for chiropractic care for low back pain when they focused on 'pragmatic' studies (i.e. carried out in close to real-world settings) in 2016.⁷ They found conflicting evidence for the cost/benefit of chiropractic care compared to other types of non-invasive care, and indicated a need for more high-quality pragmatic studies.

Practice-based research (PBR) offers an alternative approach to studying the effectiveness of chiropractic care for a range of health conditions. By moving the locus of research to the offices of one or more chiropractors practicing in the field, it may be possible to amass very large volumes of data relating to problems, therapeutic modalities, costs and outcomes. PBR works well for what has been called Comparative Effectiveness Research (CER)^{8,9}; CER seeks to compare outcomes across many practices, with differing clinical approaches, to see what works best for what clinical presentations. The power of CER is that the results are immediately generalizable to the real-world practice, since data are gathered from that

milieu. Also, the sample size can be expansive and geographically diverse if enough practitioners are involved.

However, the collection of data directly from doctors' offices involves its own set of challenges, especially in terms of network development, staff training and buy-in and quality control.¹⁰ A recent article by Bussieres *et al.* points out the challenges of developing a PBR network (PBRN) in chiropractic, which stem from a lack of research infrastructure and research training by chiropractors.¹¹

Currently there are two PBRNs listed with the U.S. Agency for Healthcare Research and Quality (AHRQ): the International Chiropractic Pediatric Association (ICPA) PBRN and the ResearchLink Chiropractic Learning Healthcare Community.¹² The ICPA PBRN is perhaps the most productive of the currently operational PBRNs in chiropractic, with five recent publications.¹³⁻¹⁷

Traditional PBR projects have depended on paper forms for data collection, especially of health outcomes in the form of patient questionnaires or doctors' assessments.¹⁰ More recently, researchers are using online resources, including emails to contact patients, and online questionnaires.^{15,16,18}

It is also possible to collect PBR data from private and public databases such as those containing medical claims data. These databases contain claims data from a mixture of healthcare practitioners, including chiropractors, and have gone far to describe the economics of complementary and alternative medicine (CAM) healthcare consumption¹⁹⁻²³ as well as the distribution of Medicare utilization²⁴⁻²⁶.

Public and Insurance databases in the United States contain only claims-level data and some information about the patient's date of birth, and date of injury or disability as seen on the insurance claim. While diagnostic information is present, claims databases do not typically have information about patient health assessments or response to care.

Other government-sponsored health information databases, such as the Medical Expenditures Panel Survey (MEPS), combine health expenditure information, with employers' records and health interviews with representative samples of adult consumers. Analyses of these data have provided information about outcomes of care and comparisons of expenses between CAM practitioners and standard medical care.²⁷⁻³⁰

The present study approaches PBR from a relatively new and untested direction, namely using the practitioner's own Electronic Record System as the source of data. Since such systems are now mandated, the computerized data management systems presumably installed in all chiropractic offices in the US may be a rich source of information about patients, their progress under chiropractic care and the cost-effectiveness of the care. It may be possible to collect patient data retrospectively from the doctor's electronic files. The main benefit would be that no extra work is required from doctors or their staff to enter study-specific information.

It should be noted that the electronic records systems used in a particular doctor's office may be one of two types: an electronic medical record (EMR) or an electronic health record (EHR). An EMR is a single practice's digital version of a patient's chart. It contains the patient's medical history, diagnoses and treatments by a particular physician, nurse practitioner, specialist, dentist, surgeon or clinic. An EHR is also a digital version of a patient chart, but it is a more inclusive snapshot of the patient's medical history. EHRs are designed to be shared with other providers, so authorized users may instantly access a patient's data across different providers. A USF Medical School web page discusses the differences.³¹

Mior *et al.* used data from an electronic billing system to study the economics of chiropractors in private practice in Ontario, Canada.³² The study assessed summary data pertaining to numbers of patients and visits seen over time but did not look at patient outcomes or demographics. The use of EHR as a data source for retrospective studies of chiropractic patients has been carried out in institutional settings using college or hospital EHR systems. Peterson *et al.* used a hospital EHR in Switzerland to study the relative effectiveness of two treatments for low back pain.³³ Kaeser *et al.* used the EMR of a chiropractic college in the US to compare demographics and diagnoses of patients in the care of student interns to patients seen in private practice.³⁴ They determined that there were distinct differences between patients seen in the school clinics and those reported by practitioner surveys. Hence, there is a need for more information from general practitioners in the chiropractic field.

The present report describes our initial venture into the use of an office-based EMR for possible use in PBR studies. In what follows, we explore the issues encountered in

practitioner recruitment and buy-in, database access, data transmission, personal health information (PHI) de-identification and data analysis. In addition, we briefly discuss institutional review board approval as it relates to this type of work.

Methods

Study design

A retrospective cross-sectional descriptive analysis of a selected set of variables extracted from aggregated data samples contained in cloud-based backups of database files from practitioners using the Vitalogics records system.

Data source

We elected to utilize the commercial electronic records system marketed by Vitalogics Wellness Technologies [www.vitalogics.com, Peoria, IL, USA], largely because the owner of the company is a chiropractor who offered us access to his software engineers and to assist in recruiting clinics in return for our help in improving the software. The software, which was in use in hundreds of chiropractic offices, is an EMR system specifically designed for Chiropractic offices. Besides functionality specifically for practice management and billing, it features modules for healthcare management, outcome assessment, records management and SOAP notes. It is these latter features that we hoped to exploit for comparative effectiveness research.

The actual extraction of data and assembly of working files was performed by programmers employed by Vitalogics. Two cross-sectional data samples were extracted from records of consenting offices, one for all active patient with visits occurring in July 2014 and the other for January 2015. The records extracted contained only baseline data from the patient registration database and did not include information about services performed.

Participants

Participants were of two types: the doctors or practices providing the data and the patients whose records were part of the data sample. As regards the former, any doctor or practice using the Vitalogics software was eligible to participate. Doctors were required to sign a consent form agreeing to the data extraction. Our recruitment plan for

doctors using the software called for the software owner to make an initial request either through email or personal contact. As a further recruiting effort, the research team made a presentation to a regional meeting of software users to familiarize them with the project and enlist them in the study. Practitioners were able to voice their concerns about the project goals and possible outcomes. Practitioners were required to provide written consent before any data could be accessed. The study consent form described the goals of the research and the HIPAA compliant security and de-identification methods that would be used to guarantee data safety and anonymity.

Since our patient participants were 'virtual' and could not consent to participation in person, it was necessary to collaborate with our Institutional Review Board (IRB) to develop a study-specific ethics protocol appropriate to the parameters of this study. Following recent ethics discussions³⁵ and the U.S. federal regulations on human research protections³⁶, it was decided that patient consent would not be required for this project for two reasons: (1) the data were retrospective, implying that in many cases the care plan has been completed and it would be difficult to contact patients to gain consent, and (2) any personal health information (PHI) would be transformed using HIPAA safe harbor plus statistical de-identification methodologies.

Variables

We learned from the software engineers that the software contained several thousand variables, many of which were related to internal software processes. The software engineers provided an initial list of 62 variables focusing on patient demographics, vital statistics, the visit/service record, diagnosis codes and outcomes. Upon inspection, we found that many of these variables were numbers that could be used to identify individual patients, and we could not use them if confidentiality was to be maintained. Also, we had determined early on in the study plan not to include visit-specific data, but to focus on patient characteristics and health status.

We also developed a list of desirable outcome variables specific to comparative effectiveness research, including psychosocial factors, such as marital status, number of children and income level; and educational level which are known to contribute to health and availability of care.^{37,38} The variables we suggested are typically utilized in health services research databases.³⁹

The final list consisted of 34 variables, with 17 variables for record identification and patient demographics, two diagnosis variables, 13 health indicators and two related to referral source or the presence of outcomes (Table 1). The patient’s method of payment, whether cash or some form of third party payment was included as the variable ‘Case Type’. The software engineers matched our list to their data dictionary and identified 27 variables in the database that best matched our requested list. For the final data capture, the programmers developed specific database queries to amass a working dataset.

Data handling and analysis

According to pre-established protocols, the resulting datasets were conveyed through a secure server to a single researcher at our institution who performed initial data cleaning, including removal or transformation of any residual Protected Health Information (PHI). To facilitate this process, we developed protocols for data blinding and transfer based on best practices as well as federal research and HIPAA guidelines.⁴⁰ The Life University institutional review board (IRB) approved the data handling plan and the collection of data directly from the EMR database.

Using Microsoft Excel, a descriptive analysis was performed; to find the rate of utilization of each variable, we first simply counted the number of instances for each variable where data were found to determine the percentages present. Next, we evaluated the range of each numeric variable, such as age, to determine validity. Invalid values (e.g. dates in the wrong century) were removed from the database.

To assess the frequency of utilization of the various diagnosis codes, we constructed a subsidiary database whose records contained the diagnosis codes, case types and age variables. For grouping purposes, the dataset was augmented by the addition of a three-digit ICD ‘root’ code, and short descriptors matched to an ICD-9 code database downloaded from the US Centers for Medicare & Medicaid Services.⁴¹

Diagnosis codes were also compared to those presented in previous articles that used MEPS to study the use of chiropractic care in back pain patients. Traditionally, there is a limited set of ICD codes that are considered “Back Pain” by researchers.^{27,28} Smith used this same list, but augmented it with two codes (846 & 847) to include

Table 1.
Variable types and numbers requested, compared to the actual variable list received in both data slices combined.

Identification Codes	% present in sample
Clinic_ID	100.0%
Pt_ID	100.0%
Date at 1 st visit	100.0%
Creation Date	68.0%
Age at 1st visit	72.3%
Gender	65.8%
Marital Status	49.3%
Education level	Not in db
Number of children	Not in db
Family income	Not in db
Employment Status	32.7%
Occupation	14.0%
Case Type (Cash, Insurance, Medicare, WC)	63.5%
Residence ZIP Code (3 digits)	79.0%
Residence City	79.5%
Residence County	Not in db
Residence State	77.4%
Diagnosis	
ICD Code	51.1%
Date of Diagnosis	Not in db
Other	
Referral Source	17.4%
Outcomes (NDI, RMDQ, VAS, etc.)	Not in db
Vital Signs	
Height	Not in db
Weight	Not in db
Diastolic Pressure	Not in db
Systolic Pressure	Not in db
Pulse rate	Not in db
Activity level	Not in db
Family Medical History	Not in db
Reason for seeking care this case	Not in db
Duration of health issue	Not in db
Severity of health issue	Not in db
Fitness Rating	Not in db
Goal of care this case	Not in db
Status at last visit	Not in db

db = database

Sprains & Strains of the spine or sacroiliac region.^{29,30} We flagged the codes in our database that were used by Smith to indicate back pain as “used in MEPS” and marked them with an asterisk (*) in Table 2.

Results

Practitioners

The recruitment process yielded 40 clinics that signed consent forms and agreed to participate in the study. The clinics were located across the US with 15 located in the Midwest, 12 in the Pacific region, seven in the Southeast

Table 2.

Diagnosis Code breakdown by Case Type. Values are the number and percentage of patients of a certain case type that was given the diagnosis code type shown at each row. The diagnosis codes were reduced to just the 3-digit root code and the general description for each. Codes are sorted by frequency of use from most frequent at the top and only the most frequent 20 codes are shown. Counts and percentages of each case type are shown in the bottom row. ‘’ indicates a code recognized as ‘Back Pain’ in MEPS studies (n=22552)*

Diagnosis Code (3 digit) & Description	Major Medical	Cash	Medicare	Gratis	PI	Managed	Wellness	Workers Comp	Medicaid	Total number	Per-cent
739 Somatic dysfunction	57.4%	41.3%	50.8%	44.4%	45.9%	63.7%	50.7%	46.7%	46.1%	11614	51.5%
724 Other /unspecified disorders of back*	10.2%	17.1%	18.4%	7.6%	9.5%	20.0%	13.7%	19.4%	14.4%	2864	12.7%
723 Other disorders of cervical region*	7.3%	10.7%	9.4%	17.5%	9.4%	5.9%	6.9%	9.8%	6.7%	1985	8.8%
722 Intervertebral disk disorders*	9.8%	5.1%	3.3%	4.3%	2.4%	3.5%	1.1%	2.0%	0.4%	1646	7.3%
847 Sprains and strains of other and unspecified parts of back*	4.2%	4.4%	3.2%	4.3%	14.5%	0.6%	14.9%	6.5%	5.3%	1060	4.7%
839 Other, multiple, and ill-defined dislocations of spine*	1.4%	8.7%	4.2%	0.2%	9.4%	0.2%	0.2%	5.4%	2.5%	834	3.7%
729 Pain in Limb	2.1%	2.3%	2.5%	7.9%	1.3%	0.1%	4.5%	2.9%	12.0%	541	2.4%
728 Muscle weakness, spasm, ligament laxity	1.7%	1.3%	0.4%	11.3%	0.6%	0.1%	0.0%	1.3%	0.4%	406	1.8%
720 Ankylosing spondylitis and other inflammatory spondylopathies*	0.8%	2.0%	3.6%	0.3%	1.7%	0.1%	6.3%	0.7%	6.3%	316	1.4%
719 Joint Pain – upper limb, lower limb, Pelvis	1.2%	1.9%	1.3%	0.1%	0.9%	2.0%	0.3%	2.2%	2.8%	316	1.4%
784 Headache	0.6%	1.2%	0.8%	0.1%	1.3%	2.4%	0.3%	0.4%	1.4%	203	0.9%
840 Sprain of shoulder	1.1%	0.4%	0.0%	0.0%	0.9%	0.2%	0.0%	1.1%	0.0%	158	0.7%
721 Spondylosis and allied disorders*	0.6%	0.9%	0.6%	0.0%	0.5%	0.0%	0.0%	0.9%	0.0%	135	0.6%
737 Curvature of spine*	0.4%	0.7%	0.4%	0.2%	0.3%	0.0%	0.0%	0.0%	0.0%	113	0.5%
736 Unequal leg length	0.5%	0.1%	0.0%	1.1%	0.0%	0.0%	0.0%	0.0%	0.0%	68	0.3%
346 Migraine Headache	0.2%	0.5%	0.2%	0.0%	0.2%	0.6%	0.2%	0.0%	0.4%	68	0.3%
782 Edema	0.1%	0.5%	0.3%	0.0%	0.3%	0.0%	0.2%	0.2%	1.1%	68	0.3%
781 Abnormality of gait/posture	0.1%	0.6%	0.1%	0.1%	0.3%	0.1%	0.6%	0.2%	0.0%	45	0.2%
307 Tension Headache	0.2%	0.3%	0.2%	0.0%	0.3%	0.7%	0.0%	0.0%	0.4%	45	0.2%
846 Sprains and strains of sacroiliac region*	0.2%	0.1%	0.2%	0.5%	0.2%	0.0%	0.3%	0.2%	0.0%	45	0.2%
Total N	11998	6495	1105	970	925	496	361	113	90	22552	
% of patients by Case Type	53.2%	28.8%	4.9%	4.3%	4.1%	2.2%	1.6%	0.5%	0.4%		100%

Number of Patients within SCF

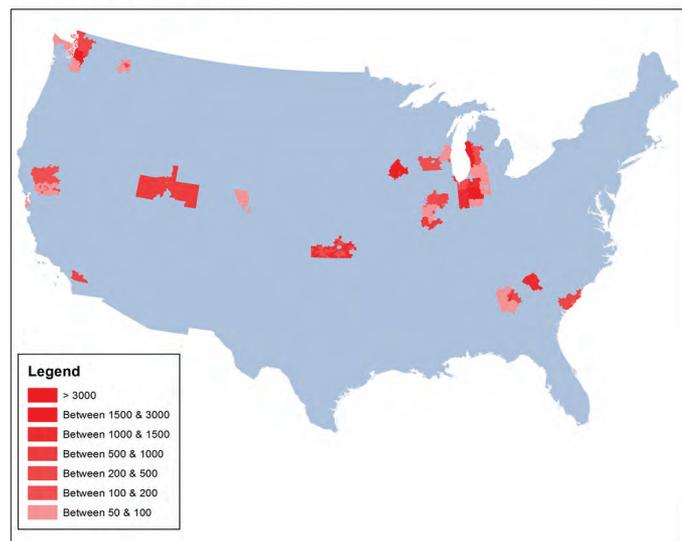


Figure 1.

Geographic distribution of patients in EMR sample.

and three each in the Mountain and Northeast regions (Figure 1).

Patient data

The July 2014 data consisted of 31,000 patient records and the January 2015 slice included 44,500 records. After removing duplicate records, there were 46,692 records. Table 1 shows the list of variables requested and the percent of records that contained usable information for each variable. Of 34 variables requested, 15 were found in the dataset, with several being present in less than half of the records. Age at first visit was present and valid in 72.3% of the records and ranged from 0 years to 93 (Mean 38.7 (SD 18.5)). Gender was present in 65.8% of records (F=60.4%, M=39.6%). Figure 2 shows the frequency distribution of ages, broken out by gender, when both variables were present for the same patient.

Marital Status was found in 49.3% of records — the most frequent being Married (53.2%) and Single (37.5%). There were no data available for education level, number of children or family income. Employment status and occupation were found in less than half of the records.

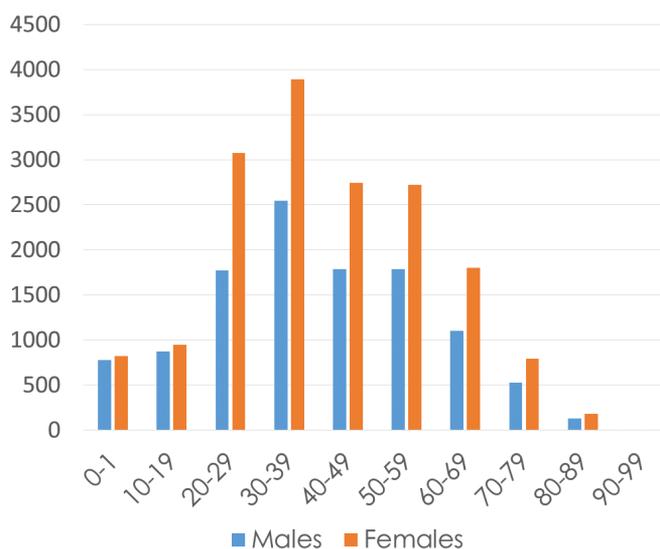


Figure 2.

Frequency distribution of ages, broken out by gender, when both variables were present for the same patient. Vertical axis is number of patients, horizontal axis is age in decades at the time of 1st visit.

No data were available on vital signs, medical history, or several other health- and fitness-related variables.

The Case Type variable was present in 63.5% of patient records. The variable was originally entered as an open-ended text field and yielded 80 different values. It was cleaned and grouped into 9 different case types, including: Major Medical, Cash, Medicare, Gratis, Personal Injury, Managed, Wellness, Workers Comp, & Medicaid. Major Medical and Cash were the most frequent with 53.2% and 28.8% respectively.

Diagnosis codes

Diagnosis codes were assigned by doctors at or near the first visit. The dataset included up to four ICD-9 diagnosis codes for each patient and 23,854 patients (51%) had at least one diagnosis. Of patients with diagnoses present, 82% had all four diagnosis codes assigned. In the original database, we found 88,900 codes across all patients, which were one of 366 unique ICD codes. We found 120 unique 3-digit root codes. The most frequent group of codes was 739 “Somatic Dysfunction” having been assigned to 51.5% of patients; of these, 29% were

for cervical region, 26% thoracic and 22% lumbar. The remainder were limbs or pelvis.

Because of the relative completeness of the diagnosis code data, we decided to look at it more closely. While there were 120 unique code groups when looking at the three-digit root code, 99% of those codes fell into only 20 unique groups. Table 2 shows the breakdown of those top 20 codes by case type (e.g., major medical, cash or gratis, etc). The frequency and percentage of each case type appears in the bottom two rows of the table, respectively. Major Medical was most frequent, followed by Cash, etc. [Note that ‘*’ after code descriptions indicates that the code is “in MEPS.”].

Comparing the frequency of diagnosis codes across the case types suggests that Cash patients were assigned codes in the 739 family (Somatic Dysfunction) less often than Major Medical patients and were assigned codes in the 724 (Other/unspecified disorders of the back) family more often.

We also looked at diagnosis code breakdowns with respect to each clinic. One clinic was omitted because it only contributed one patient to the dataset. The usage by clinic is quite varied, but there appear to be 4 major groups that can account for 34 of the 39 practices:

- 15 clinics use Somatic Dysfunction for the majority of their diagnoses (97%);
- 10 use Somatic Dysfunction most of time, but for less than 50% of their diagnoses. They tend to also use “Other and unspecified disorders of back*” and “Other disorders of cervical region*”;
- Six use “Other and unspecified disorders of back*” most often;
- Three use “Other, multiple, and ill-defined dislocations of spine*” for most of their diagnoses.

Discussion

In this study, we investigated the feasibility of collecting data from an EMR system for potential use in research. We could glean little from our data samples about the patients themselves. The most often populated fields were age (reduced to decade), and gender. Marital and employment status were included in the database, but not often completed. We found very little useful data in the database regarding other demographics, vital signs or the outcomes of care.

The patients in this study were more often females, as also found in a recent scoping review by Beliveau *et al.*⁴² and the NBCE 2015 survey of chiropractic practitioners⁴³. The median age of patients in our study (38.7) was slightly younger than the median reported across several studies surveyed by Beliveau *et al.* (43.4) but was within the interquartile range.

Diagnosis coding

As a major component of insurance billing systems, the diagnosis codes were completed for more than half of the patients. In general, there was a preponderance of musculoskeletal diagnoses, particularly in the neck, mid- and low back. We found a higher occurrence of neck complaints at 29% than did Beliveau *et al.* (22%).⁴²

A unique feature of this dataset is that we have diagnosis codes for a full range of 9 different case types, including five that might be considered insurance of one form or another. The ‘Cash’ case type was the second-most frequent type found, representing a significant portion (28.4%) of the patients. An interesting finding is that non-insurance case types (e.g. Cash, Gratis) have somewhat similar diagnostic profiles to insurance patients as seen in Table 2, but there is less reliance on somatic dysfunction codes. Medicare and insurance claims data would not contain records of patients that do not have insurance, so this is perhaps a new finding.

Another remarkable finding of the examination of diagnoses was the reliance on the 739 family of codes, which is a somatic dysfunction of the neck, back and upper and lower extremities. The large majority of these were in the neck, spine or pelvis. It was the most frequently used code, no matter which way we looked at the data: by case type, age or clinic. It is remarkable because this is not a code that would have been picked up by previous back pain researchers following the standard methodology to identify chiropractic care.²¹⁻³⁰

The Somatic Dysfunction code (739.xx) was often used in combination with other codes. In total, 78.9 % of patients were assigned at least one diagnosis code that we flagged as ‘In MEPS’, meaning they would have been classified as back or neck pain patients in previous studies. That means that 21.1% would have gone undetected if researchers relied solely on doctor assigned-diagnosis codes. It suggests the possibility that a fair number of patients in previous studies of chiropractic claims data may

have been misclassified, or not included in analyses of ‘back pain’.

Feasibility of using EMR data for practice-based research

We are not the first to suggest that chiropractic research could be based on EMR data from practitioners — Johnson recommended this action in 2010.⁹ Peterson *et al.* have been successful using EMR data in a hospital setting to compare SMT to nerve root injections for patients with lumbar disc herniation.³³ Uniquely, however, this is the first report on actual patient data extraction from a commercial EMR written specifically for chiropractors. While research to document chiropractic’s impact on the population is clearly a possibility, our efforts to date had mixed results. We were successful in collecting information directly from practitioners’ databases, but our success with gathering variables beyond those typically found in claims databases was limited.

Kukaftka *et al.* suggested that EHR systems should be designed from the ground up in such a way as to serve not only clinical goals but also efforts in public health.⁴⁴ Similarly, EMR systems such as the one we tested need to be augmented with additional modules to enable better harvesting of data for health services research.

The EMR system that we used for our study is equipped with a free-form text field for chief complaint; it is not coded in any way and hence cannot be easily correlated with ICD codes. Thus, due to the complexity of data extraction and reduction to a limited set of chief complaints, such a correlation has not been accomplished to date.

In addition to codified presenting complaints, EMR systems need the ability to track outcomes.⁴⁵ Doctors most certainly track improvement in patient’s subjective symptoms and use this information to guide care and for billing. If recorded electronically in the EMR, the information is mostly buried in the case notes as some form of SOAP (Subjective, Objective, Assessment, and Plan). The EMR we used does in fact contain SOAP notes in an uncoded form; however, our software engineers could not justify the effort it would have taken to extract the information. EMR systems could include modules to record outcome measures typically used in clinical trials (SF-12, Roland-Morris Disability Questionnaire, etc.). Indeed, the designers of the software we tested developed a module

for the SF-12, but we were not able to obtain any records in that domain to ascertain its frequency of use.

Early on in the implementation of EMRs, doctors were offered financial incentives to purchase “certified” software systems that included certain public health-related fields and modules. Having the ability to enter enhanced variables is the first step. In the end, the modules will only be useful if doctors begin to make use of them.

A major unresolved issue is what incentive(s) software developers might have to include new features in their systems. Software developers are challenged and even hesitant about adding functionality to a software system for the purpose of research because it typically does not enhance the economic value of the product to the clinician. However, the data collected on patient services and outcomes may provide valuable insight on how the software is being utilized. This utilization insight may inspire programmers to improve the functionality of the software which may provide the enhanced economic value the software company is looking for. Further complicating the issue is the fact that clinicians must regularly use the software enhancements for the research benefit to be realized. The clinician’s concern and the threat of non-engagement is going to be based on the efficiency of documentation. However, objective outcome assessments that provide data to support third party reimbursement will provide high value to the clinician and enhance their potential utilization of additional software functionality.

Limitations

This study only included a limited set of variables in one specific commercial electronic medical record system. We did not collect data on visits and services over time, so there is no perspective on process of care, outcomes, or cost. Assessing the type and quality of health-related variables from other databases might produce different results.

Conclusion

We were successful in gathering a large volume of data (more than 46,000 individual patient records) from the EMR software of practicing clinicians with the cooperation of the software designers. While we have looked at initial patient records only, and not the visit records, the cleanest and most complete data we found were in the diagnosis codes. Other variables important in healthcare

research such as patient demographics, clinical information, and outcomes of care were not found in the database. As such, the software at this stage was not particularly useful for comparative effectiveness research. We recommend that additional data fields be implemented to indicate the overall presenting profile of the patient including demographics and health indicators.

References

1. Kjaer P, Kongsted A, Hartvigsen J, et al. National clinical guidelines for non-surgical treatment of patients with recent onset neck pain or cervical radiculopathy. *Eur Spine J*. 2017;26(9):2242-2257. doi:10.1007/s00586-017-5121-8
2. Côté P, Wong JJ, Sutton D, et al. Management of neck pain and associated disorders: A clinical practice guideline from the Ontario Protocol for Traffic Injury Management (OPTiMa) Collaboration. *Eur Spine J*. 2016;25(7):2000-2022. doi:10.1007/s00586-016-4467-7
3. Bussi eres AE, Stewart G, Al-Zoubi F, et al. The Treatment of neck pain-associated disorders and whiplash-associated disorders: a clinical practice guideline. *J Manipulative Physiol Ther*. 2016;39(8):523-564.e27. doi:10.1016/j.jmpt.2016.08.007
4. Wong JJ, C ot e P, Sutton DA, et al. Clinical practice guidelines for the noninvasive management of low back pain: a systematic review by the Ontario Protocol for Traffic Injury Management (OPTiMa) Collaboration. *Eur J Pain*. 2017;21(2):201-216. doi:10.1002/ejp.931
5. Stochkendahl MJ, Kjaer P, Hartvigsen J, et al. National Clinical Guidelines for non-surgical treatment of patients with recent onset low back pain or lumbar radiculopathy. *Eur Spine J*. 2018;27(1): 60-75. doi:10.1007/s00586-017-5099-2
6. Dagenais S, Brady O, Haldeman S, Manga P. A systematic review comparing the costs of chiropractic care to other interventions for spine pain in the United States. *BMC Health Serv Res*. 2015; 15:474.
7. Blanchette MA, Stochkendahl MJ, Borges Da Silva R, Boruff J, Harrison P, Bussi eres A. Effectiveness and economic evaluation of chiropractic care for the treatment of low back pain: a systematic review of pragmatic studies. *PLoS One*. 2016;11(8): e0160037. doi:10.1371/journal.pone.0160037
8. Sox HC, Greenfield S. Comparative effectiveness research: a report from the Institute of Medicine. *Ann Intern Med*. 2009;151(3): 203-205.
9. Johnson C. Comparative effectiveness research and the chiropractic profession. *J Manipulative Physiol Ther*. 2010;33: 243-250.
10. Ax en I, Leboeuf-Yde C. Conducting practice-based projects among chiropractors: a manual. *Chiropr Man Therap*. 2013; 21(1):8.
11. Bussi eres A, C ot e P, French S, Godwin M, et al. Creating a chiropractic practice-based research network (PBRN): enhancing the management of musculoskeletal care. *J Can Chiropr Assoc*. 2014; 58(1): 8-15.
12. Practice-Based Research Networks. Agency for Healthcare Quality and Research. <https://pbrn.ahrq.gov/search/chiropract> (Accessed 12/20/2019).
13. Alcantara J, Ohm J, Kunz D. The safety and effectiveness of pediatric chiropractic: a survey of chiropractors and parents in a practice-based research network. *Explore (NY)*. 2009; 5(5):290-295.
14. Alcantara J, Ohm J, Kunz D. The chiropractic care of children. *J Altern Complement Med*. 2010; 16(6):621-626.
15. Alcantara J, Lamont AE, Ohm J, Alcantara J. The quality of life of children under chiropractic care using promis-25: results from a practice-based research network. *J Altern Complement Med*. 2018; 24(4):378-384.
16. Alcantara J, Nazarenko AL, Ohm J, Alcantara J. The use of the patient reported outcomes measurement information system and the Rand VSQ9 to measure the quality of life and visit-specific satisfaction of pregnant patients under chiropractic care utilizing the Webster technique. *J Altern Complement Med*. 2018; 24(1): 90-98.
17. Alcantara J. The presenting complaints of pediatric patients for chiropractic care: Results from a practice-based research network. *Clinical Chiropr*. 2008; 11, 193-198.
18. Field J, McKernan P, Newell D. Comparative service evaluation of patients presenting for chiropractic care privately with those referred for care via the NHS. ECU Conference, Dublin, Ireland; May 2015.
19. Davis MA, Martin BI, Coulter ID, Weeks WB. US spending on complementary and alternative medicine during 2002-08 plateaued, suggesting role in reformed health system. *Health Aff (Millwood)*. 2013; 32(1): 45-52. doi:10.1377/hlthaff.2011.0321.
20. Grieves B, Menke JM, Pursel KJ. Cost minimization analysis of low back pain claims data for chiropractic vs medicine in a managed care organization. *J Manipulative Physiol Ther*. 2009; 32(9):734-739.
21. Weigel P, Hockenberry JM, Bentler SE, et al. A longitudinal study of chiropractic use among older adults in the United States. *Chiropr Osteopat*. 2010;18:34.
22. Weigel PA, Hockenberry JM, Bentler SE, Kaskie B, Wolinsky FD. Chiropractic episodes and the co-occurrence of chiropractic and health services use among older Medicare beneficiaries. *J Manipulative Physiol Ther*. 2012;35(3):168-175.
23. Weigel PA, Hockenberry JM, Bentler SE, Kaskie B, Wolinsky FD. Chiropractic use in the Medicare population: prevalence, patterns, and associations with 1-year changes in health and satisfaction with care. *J Manipulative Physiol Ther*. 2014;37: 542-551.
24. Whedon JM, Davis MA. Medicare part B claims for

- chiropractic spinal manipulation, 1998 to 2004. *J Manipulative Physiol Ther.* 2010;33: 558-561.
25. Whedon JM, Song Y, Davis MA, Lurie JD. Use of chiropractic spinal manipulation in older adults is strongly correlated with supply. *Spine.* 2012; 37(20):1771-1777.
 26. Whedon JM, Song Y. Geographic variations in availability and use of chiropractic under Medicare. *J Manipulative Physiol Ther.* 2012;35: 101-109.
 27. Martin BI, Deyo RA, Mirza SK, Turner JA, Comstock BA, Hollingworth W, Sullivan SD. Expenditures and health status among adults with back and neck problems. *JAMA.* 2008; 299(6):656-664.
 28. Martin BI, Gerkovich MM, Deyo RA, Sherman KJ, Cherkin DC, Lind BK, Goertz CM, Lafferty WE. The association of complementary and alternative medicine use and health care expenditures for back and neck problems. *Med Care.* 2012;50(12)
 29. Smith M, Davis MA, Stano M, Whedon JM. Aging baby boomers and the rising cost of chronic back pain: secular trend analysis of longitudinal medical expenditures panel survey data for years 2000 to 2007. *J Manipulative Physiol Ther.* 2013; 36(1): 2–11.
 30. Smith M. Identifying Episodes of Back Pain Using Medical Expenditures Panel Survey (MEPS) Data: Patient Experience, Use of Services, and Chronicity. *J Manipulative Physiol Ther.* 2010; 33(8): 562–575.
 31. EHR vs EMR | USF Health Online. <https://www.usfhealthonline.com/resources/key-concepts/ehr-vs-emr/> (accessed 6/22/2020).
 32. Mior SA, Waalen J. Personal and practice predictors associated with the income of Ontario chiropractors. *J Manipulative Physiol Ther.* 2008;31(4): 277-284.
 33. Peterson CK, Leeman S, Lechman M, Pfirrmann AW. Manipulative therapy or imaging-guided lumbar nerve root injections. *J Manipulative Physiol Ther.* 2013; 36: 218-225.
 34. Kaeser MA, Hawk C, Anderson M. Patient characteristics upon initial presentation to chiropractic teaching clinics: A descriptive study conducted at one university. *J Chiropr Educ.* 2014;28(2):146–151.
 35. Wolf LE1, Walden JF, Lo B. Human subjects issues and IRB review in practice-based research. *Ann Fam Med.* 2005;3(Suppl 1): S30-S37.
 36. 45 CFR §46 US Department of Health & human Services. <https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html> (accessed 6/10/2019).
 37. Marmot MG, Fuhrer R, Ettner SL, Marks NF, Bumpass LL, Ryff CD. Contribution of psychosocial factors to socioeconomic differences in health. *Milbank Q.* 1998;76(3):403-448, 305.
 38. Zimmerman E, Woolf SH. Understanding the relationship between education and health. Discussion Paper. Washington, DC: Institute of Medicine; 2014.
 39. MEPS codebook. https://meps.ahrq.gov/mepsweb/data_stats/download_data_files_codebook.jsp?PUFId=H181 (accessed 6/5/2018).
 40. Barth-Jones, D., Statistical de-identification: challenges and solutions, HHS workshop on HIPAA privacy rule's de-identification standard, 2010.
 41. Centers for Medicare and Medicaid Services (www.cms.gov).
 42. Beliveau PJH, Wong JJ, Sutton DA, et al. The chiropractic profession: a scoping review of utilization rates, reasons for seeking care, patient profiles, and care provided. *Chiropr ManTher.* 2017; 25:35.
 43. National Board of Chiropractic Examiners. Practice Analysis of Chiropractic 2015. NBCE. Greeley Colorado, 2015.
 44. Kukafka R, Ancker JS, Chan C, Chelico J et al. Redesigning electronic health record systems to support public health. *J BioMed Inform.* 2007; 40: 398–409.
 45. Wu AW, Kharrazi H, Boulware LE, Snyder CF. Measure once, cut twice – adding patient-reported outcome measures to the electronic health record for comparative effectiveness research. *J Clin Epidemiol.* 2013; 66(8 Suppl): S12-S20.