IEEE JOURNAL OF
BIOMEDICAL AND HEALTH INFORMATICS

Intelligent Mortality Reporting with FHIR

Ryan A. Hoffman, Hang Wu, Janani Venugopalan, Paula Braun, and May D. Wang

Abstract— **One pressing need in the area of public health is timely, accurate, and complete reporting of deaths and the diseases or conditions leading up to them. Fast Healthcare Interoperability Resources (FHIR) is a new HL7 interoperability standard for electronic health record (EHR), while Sustainable Medical Applications and Reusable Technologies (SMART)-on-FHIR enables third-party app development that can work "out of the box". This research demonstrates the feasibility of developing SMART-on-FHIR applications to enable medical professionals to perform timely and accurate death reporting within multiple different jurisdictions of US. We explored how the information on a standard certificate of death can be mapped to resources defined in the FHIR standard (DSTU 2) and common profiles. We also demonstrated analytics for potentially improving the accuracy and completeness of mortality reporting data.**

I. Introduction

Mortality is one of the most reliable sources of health-related data that is comparable across different geographical locations and is a large source of population-level health data, with approximately 56 million deaths per year world-wide [1]. In the United States of America (US) alone 2.6 million people die each year [2]. Timely and accurate mortality data is essential for formulating emergency response to epidemics and new disease threats, prevention of communicable diseases such as flu, determining vital statistics such as life expectancy, mortality trends, etc.

Accurate collection and aggregation of high-quality mortality data remains an ongoing challenge primarily due to issues such as the lack of practice for physicians to perform death certification (on the order of 1-2 times a year), inconsistent training in determining the cause of death information, complex data flow between the funeral home, the certifying physician and the registrar, and non-standard practices of data acquisition and transmission [3, 4]. In the US, the National Center for Health Statistics aggregates mortality data from the 57 reporting jurisdictions around the country. However, the precise regulations and local laws of each reporting jurisdiction differ [5].

A decision support system that can assist the physician to determine the appropriate cause of death and put on the death certificate in the requisite format can largely assist to mitigate these challenges of data accuracy. Additionally, current efforts towards mortality reporting standardization using technologies such as Clinical Document Architecture (CDA) [6, 7] and Health Level 7 (HL7) V2 [8] have some shortcomings, such as challenges in integrating with large-scale web services. Additionally, efforts at standardization such as Integrating the Healthcare Enterprise (IHE) have met with limited adoption. As it stands today, the current flow of information between the various providers and registrars is not optimal. Also, under the new meaningful use of electronic health records (EHR), the government requires healthcare institutions to show at least partial support of patient-facing application programming interfaces (APIs) to show potential for sharing data, interoperability and clinical decision support [9].

To overcome these challenges, we propose a framework that utilizes HL7's Fast Healthcare Interoperability Resources (FHIR) as both an application platform and a means of accessing EHR data. FHIR is a new emerging health standard that is aimed at streamlining and standardizing healthcare communication using a resource-centric approach (as opposed to document-centric) for specification of data elements. It is designed to allow simple implementation using existing technologies such as RESTful (REpresentational State Transfer) APIs, OAuth security, and XML/JSON data. [10]. FHIR was chosen for the application of death certificates, because it is vendor-neutral, scalable, and is positioned to emerge as a global standard. FHIR is designed to work within current EHR systems using APIs and can potentially be used to pre-populate sections of the death certificates and to provide information from the decedent's health history to aid physicians in determining the cause of death.. FHIR is also the standard which more easily supports the addition of analytics into the EHR systems [11, 12]. The ultimate goals of this project are to generate information that will aid in more complete physician reporting of the causes of death, and to provide valuable mortality information to registrars, public health departments, and other authorized parties in a timelier manner.

A preliminary version of this work was presented at the IEEE Biomedical and Health Informatics (BHI) conference in February 2017 [13].

II. Web Application Design

A web application was implemented in HTML (HyperText Markupb Language)and JavaScript, using the SMART-on-

This work was supported in part by grants from the National Center for Advancing Translational Sciences of the National Institutes of Health (NIH) under Award UL1TR000454 to Dr. May D. Wang, T32 GM105490 Traineeship to Dr. Greg Gibson of Georgia Tech for trainee R. A. Hoffman, National Science Foundation Award NSF1651360, and the US Department of Health and Human Services (HHS) Centers for Disease Control and Prevention (CDC) HHSD2002015F62550B to Dr. May D. Wang, and Microsoft Research and Hewlett Packard. This article does not reflect the official policy or opinions of the CDC, NSF, or the US Department of HHS and does not constitute an endorsement of the individuals or their programs.

R. A. Hoffman and J. Venugopalan: Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, GA 30332 USA.

H. Wu: School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA.

P. Braun: Centers for Disease Control and Prevention (CDC), Atlanta, GA 30329 USA (co-corresponding author, e-mail: pabraun@cdc.gov).

M. D. Wang is with the Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, GA 30332 USA (co-corresponding author, phone: 404-385-2954; e-mail: maywang@gatech.edu).

Digital Object Identifier 10.1109/JBHI.2017.2780891

FHIR JavaScript client library (https://github.com/smart-on-fhir/client-js). The application runs in the browser, securely accessing the FHIR server using OAuth2 authentication. The application was developed and tested using a virtual FHIR server. In addition to the SMART-on-FHIR compliant EHR server, an application server hosts simple RESTful interfaces to UMLS and data mining functionality. An outline of the proposed infrastructure for a SMART-on-FHIR-based mortality is shown in Fig. 1.

A. Application Features

This application was designed with the ultimate goal of enabling not just more timely, but also more accurate and complete data about the chain of diseases or conditions ultimately leading to death. As such, it is designed to allow the simultaneous visualization of a large portion of the decedent's health history. As mortality reporting in the United States has

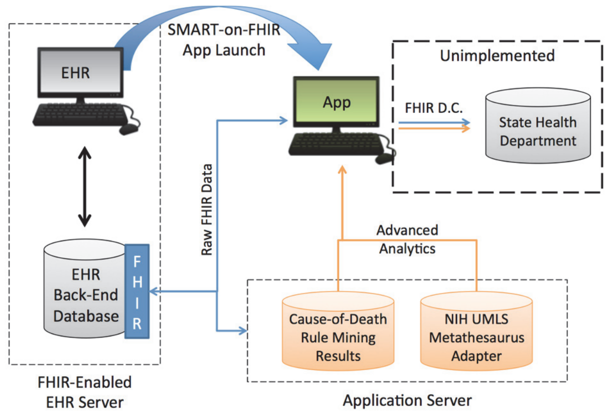

Fig. 1: Proposed Infrastructure for Death Reporting Application. A user's existing EHR system with compatible FHIR interface can directly launch the application with patient context. The application can contact an internal or external analytics server for decision support and other tools, before packaging the death certificate object and sending it.

adopted the ICD-10 standard since 1999, integration with the Unified Medical Language System (UMLS) Metathesaurus is necessary to enable crosswalk between medical event coding systems, ensuring compatibility between retrieved records and the results of rule mining.

B. Illustrative Synthetic Data and Files

To aid in interface prototyping and illustration of the application interface in Fig. 2, the application was tested using synthetic patient data available through the Cerner SMART-on-FHIR app development sandbox (https://code.cerner.com). To illustrate the application's function, the output of the application after initial processing of the Cerner sandbox test patient "Joe Smart" is included as a supplementary file.

C. Interface Design

The application's main interface is illustrated in Fig. 2. The application's interface is broken into sequential pages, which each address a section of the death certificate. The specific page illustrates uses and interactive visualization to assist in determining the causes of death.

This displays an interactive patient history timeline generated using the popular D3 and D3-tip visualization libraries (https://d3js.org, https://github.com/Caged/d3-tip).

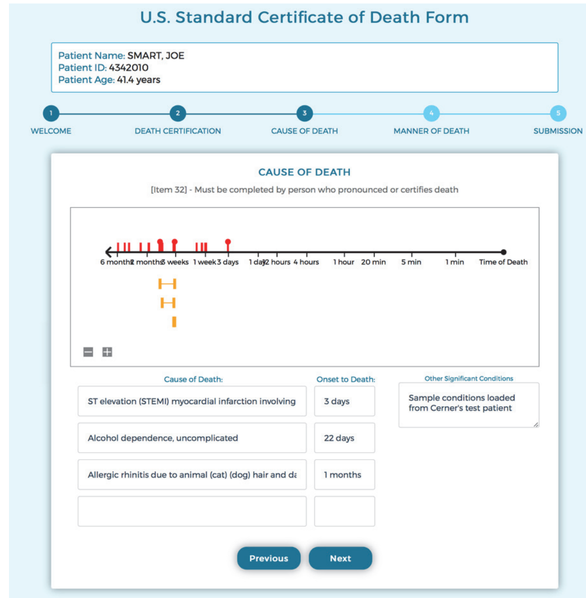

Fig. 2: Prototype Application User Interface. The web-based application runs in the user's browser. The app pulls patient information and notes from the EHR to provide context. Conditions are laid out on a timeline, alongside proposed sequential linkages between those events. The user can edit, add, or remove conditions as needed.

Events displayed on the timeline are spaced logarithmically, with the axis anchored at the time of death. This allows simultaneous visualization of events occurring around the time of death alongside relevant context from the patient's more distant history. Scrolling adjusts the scaling to allow focus on past and recent events. The events shown on this timeline are generated using Condition resources accessed from the FHIR server.

The bottommost section are designed so as to recreate the familiar-to-users appearance of the US Standard Certificate of Death's cause-of-death field layout, with a chain of one or more causes occurring as consequences of one another.

Buttons are provided to access additional pages, which contain such fields as injury information, the provider's information, and submission / download controls. The application can be downloaded from http://miblab.bme.gatech.edu/software.

III. Representing Death Certificate Data in FHIR

A significant milestone in developing a FHIR-based electronic death record is mapping the elements of a death certificate to FHIR Resources. Such a mapping must balance multiple design goals. Such a mapping should be expressive, in that it should be intuitively clear to application developers what various Resource elements are to be used for. It should also be modular and extensible, so that the rich variation in data elements required by the various state health agencies can be represented in a single, reusable Resource. To ensure that such a profile can be quickly adopted by potential users, it should be idiomatically correct and should not misuse the standard fields. It should also be designed with stakeholders' current data practices in mind, mirroring existing processes wherever feasible to limit administrative friction and encourage adoption. Here we propose a mapping of FHIR

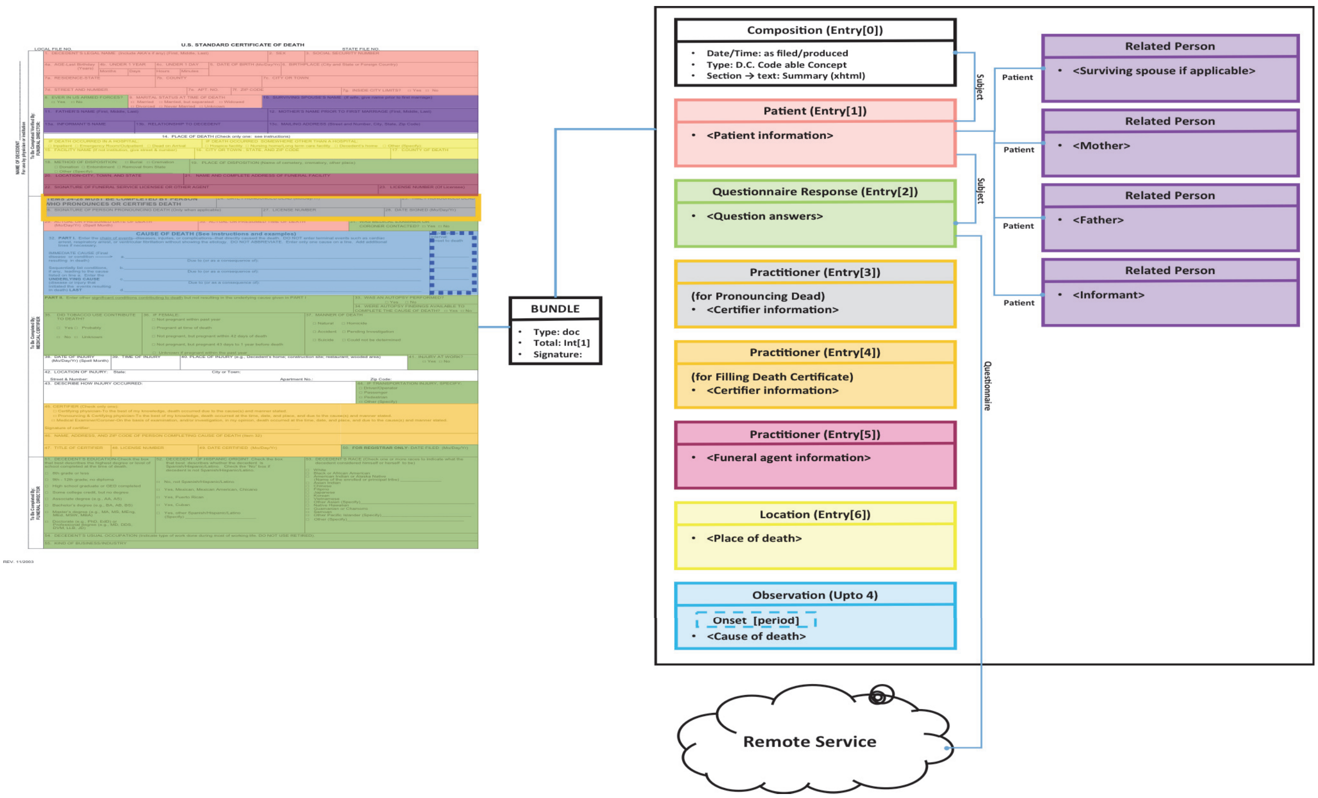

Figure 3: Resource-Level Mapping Overview. (Left) The U.S. Standard Certificate of Death is overlaid with colored blocks representing mapped FHIR resources, to be bundled into a death certificate document (Right). The exact number of Conditions will vary with the circumstances of death, and specific observations may be specified by public health reporting jurisdictions.

DSTU 2 resources and data elements to fully represent death certificate data. DSTU 2 was selected for this application and profiling effort given its much wider acceptance over the newer STU 3 version at the dime of development.

A. *Proposed Profiles*

We propose using the standard FHIR metaphor of a "document" to represent a death certificate object. Unlike the older CDA document standard, FHIR documents are modular compositions of full EHR Resources, which can be readily split apart and incorporated into another interoperable system. FHIR Documents are defined as a Bundle of resources, where the first entry in the Bundle is a Composition, which in turn contains a human readable summary of the Bundle's contents. Additional Resources are then added to the bundle, in support of the information contained in the document. Leveraging FHIR's built-in flexibility, we propose to profile a specific document type that contains a defined minimum set of resources, representing death certificate data elements, illustrated in Fig. 3.

B. *Mapping to Existing Data Standards*

Mapping the proposed profile to known standards for mortality data will maximize its usefulness and enable interoperability even with electronic death reporting systems (EDRS) which are not truly or fully FHIR-enabled.

To that end, we have completed a mapping of the HL7 Vital Records Domain Analysis Model (VR DAM) section on mortality reporting to FHIR resources. The complete mapping table is available as a supplementary file to this report.

C. *Challenges and Design*

Some death certificate data elements map very naturally and directly to FHIR Resources (e.g. the decedent name and address map directly to a FHIR Patient). Others require more careful thought and design, one example being the choice between FHIR Patient.contacts or RelatedPersons for representing the decedent's family members.

In choosing between contacts and RelatedPersons, as a specific example of making any such design decision, several factors should be considered. First, both representations are complete, and can represent all of the data elements from the U.S. Standard Certificate of Death and the HL7 VR DAM without extension. Both are intuitive enough mappings to communicate the needed information clearly to an implementer. The deciding factor in our implementation is idiomatic FHIR "correctness" – i.e., which choice most clearly follows from the originally intended uses. The contacts field of a FHIR Patient is intended primarily to enable contacting the patient or his/her decision-makers in a clinical setting, whereas RelatedPerson resources are stated to be sources of patient information with non-healthcare relationships to the patient. From this, it follows that RelatedPerson resources are most appropriate.

Another informative example is the choice between a collection of Observation resources or Questionnaire / QuestionnaireResponse resources to represent the demographics and medical history sections of the death certificate (shown in green in Fig. 3). The clean metaphor of a questionnaire representing a list of questions on a physical death certificate form is attractive, however we have opted to propose the use of Observations instead. The variations in lists of medical history questions and valid responses between reporting jurisdictions makes developing a single Questionnaire for national use infeasible. If each jurisdiction were using a unique, "flat" Questionnaire to represent their set of questions and answers, it would complicate the processes of integrating data sets from across the nation and identifying data elements present in multiple of those data sets.

IV. SEQUENTIAL PATTERN MINING ANALYTICS

This application illustrates how next-generation web services can be developed to aid in timely, accurate mortality reporting. To better understand the availability of this data and demonstrate this capability, we use sequential pattern mining on one year of public causes of death data to mine a list of rules that can be used directly in the application to propose common pathways of events that may have led to death.

A. Public Use Data

Through the National Vital Statistics System (NVSS), the National Center for Health Statistics aggregates the causes of death for all deaths occurring within the United States from 1959 to 2014 [14]. Each death certificate format in vital statistics offices of each state, the District of Columbia, and other special jurisdictions varies, but generally consists of the underlying, immediate, and contributing causes of death as recorded by physicians and other details such as the demographics, race and ethnicity. Thr available fields are shoen in Table 1.

B. SPM Background and Related Work

The temporal models commonly seen in the literature include models such as sequence analysis [15-19] and association rule mining [18, 20, 21]. Sequential Pattern Mining (SPM) is a data mining technique that seeks temporal relationships among events (in this case the underlying causes of death) [22] and has been extensively examined in the literature with applications in pattern mining [23] (AprioriAll [24], SPADE [25]) and database projections. (PrefixSpan [26], MEMISP [27]). This method has been applied successfully to clinical data [28, 29]. Recently, privacy preserving pattern mining [30-32] and distributed mining [30] have attracted considerable interests. In health care, SPM has applications in heart disease prediction [33] , healthcare auditing [34], and neurological diagnosis [35], violent death reporting [36] etc. The input data to an SPM is a set of sequences, which comprises a list of events ordered by temporal relations.

The goal of SPM is to discover all valid sequential patterns with pre-specified minimum support, where support of a candidate pattern is the proportion of sequences in the data that exhibit the pattern [37]. For example, in the NVSS Multiple Cause-of-Death public use data, each record contains a list of ordered conditions (up to 20 conditions) that could lead to a person's death. While SPM's output an ordered list of sequences which correlates with the target outcome and is able to find rules such as "Condition 1 -> Condition 2". This means that if we observe Condition 1, we can assert that Condition 2 will possibly follow Condition 1. This was introduced as an improvement over Association Rule Mining (ARM) [23], which doesn't take the temporal relations into consideration, so will only output rules like "[Condition 1, Condition 2]", meaning if we observe Condition 1, we are also likely to observe Condition 2 for some confidence. Hence SPMs form an ideal algorithm for the current task of discovering the most probable sequence of events that led to the cause of death to help certifying physicians fill the death certificates.

C. SPM Problem Formulation

As discussed above, the NVSS public use data sets consist of up to 20 underlying conditions $C = [C_1, C_2, ..., C_K]$ which lead to death, where $C = [C_1, C_2, ..., C_K]$ is the list of unique events/conditions. Using this data as the training set, our goal is to find the list of most frequent sequence conditions $S = <s_1, s_2, ..., s_T>$ which can occur given the outcome and comorbidities. The relative support of a rule R in the set of sequences D is defined as the percentage of sequences that contain this rule, i.e.,

$$rel\_support(R) = \frac{|\{S \mid S \in D \,\&\, R \subset S\}|}{|D|},$$

Where $|\cdot|$ is the cardinality of a set. SPM aims to discover sequential patterns that have support larger than a pre-specified minimum support. For example, we have a set of sequences and want to identify rules with a minimum support of 0.8.

D. SPM Methodology

In our experiments, we use the BIDE algorithm, short for BI-Directional-Extension-based frequent closed sequence mining, proposed in Wang et al. [38]. Conventional sequence mining algorithms adopt a candidate maintenance-and-test paradigm, in which they maintain a list of discovered closed rules and use the rules to prune the search space and determine whether new rules are promising to be closed. Such paradigm is accurate but lacks scalability with respect to the number of frequent closed rules, both in time and storage. On the other hand, the BIDE algorithm aims to

Table 1. Death Record Layout in Multiple Cause-of-Death Mortality Data from NCHS

Type	Information Included
Demographics	Age, Gender, Residence, Death Time, etc.
Underlying Cause of Death	Cause of death coded according to ICD and several other coding systems.
Conditions	A maximum of 20 conditions that correlate with the death
Race and Ethnicity	The reported race for States that are reporting single race or the bridged race for States that are reporting multiple race.

find all the frequent closed rules, without candidate maintenance.

E. Pattern Mining Results

We apply the aforementioned algorithm BIDE to Multiple Cause-of-Death Mortality Data from the NVSS public use data sets to find most frequent sequences of diseases or conditions before people's deaths. We picked Year 2012's Mortality Data, which contains 2,547,864 deaths. We set the minimum support to be 50 and identified a total of 65,915 frequent closed rules. We present the distribution of rules of different lengths in Table 2.

We present the top 20 rules of length-2 in Table 3 for illustration. The full set of rules was deployed as a lookup table service using a CGI script, integrating it into the death reporting prototype application.

Table 2. Frequent Rules Count of Different Lengths

Length	Count
1	961
2	19,160
3	33,179
4	11,508
5	1,081
6	26
>7	0

V. Conclusions and Future Work

This work demonstrates the feasibility of using the SMART-on-FHIR application framework to develop public health applications for mortality reporting, improving the timeliness and accessibility of such reports. Intelligent analytics have been show integrated with the prototype application, demonstrating future potential for improving the accuracy of death reporting. Future work may focus on using alternative data sets, as well as the more complete patient information made accessible though the interoperability of FHIR, to construct more personalized and precise analytics systems. Further development is ongoing to develop precise FHIR resource profiles to concisely, completely, and flexibly represent death certificate data.

Acknowledgments

The authors thank Dr. Mark Braunstein (Georgia Tech), Dr. Myung Choi (Georgia Tech Research Institute), and Charles Sirc (CDC) for their invaluable assistance and support in shaping this project. The authors especially thank Paula Braun, Entrepreneur-in-Residence with the Centers for Disease Control and Prevention, for her guidance and assistance throughout the course of this project.

Table 3. Rules from 2012 NCHS Mortality Data

Rule	Count	Percent
Mental and behavioral disorders due to use of tobacco -> Other chronic obstructive pulmonary disease	100,920	3.96%
Chronic ischemic heart disease -> Cardiac arrest	85,952	3.37%
Essential (primary) hypertension -> Chronic ischemic heart disease	77,249	3.03%
Mental and behavioral disorders due to use of tobacco -> Malignant neoplasm of bronchus and lung	73,212	2.87%
Chronic ischemic heart disease -> Heart failure	63,283	2.48%
Essential (primary) hypertension -> Cardiac arrest	59,771	2.35%
Mental and behavioral disorders due to use of tobacco -> Chronic ischemic heart disease	58,130	2.28%
Chronic ischemic heart disease -> Acute myocardial infarction	53,042	2.08%
Essential (primary) hypertension -> Heart failure	47,787	1.88%
Unspecified diabetes mellitus -> Chronic ischemic heart disease	47,322	1.86%
Mental and behavioral disorders due to use of tobacco -> Other chronic obstructive pulmonary disease -> Respiratory failure, not elsewhere classified	19,194	0.75%
Essential (primary) hypertension -> Chronic ischemic heart disease -> Cardiac arrest	17,150	0.67%
Mental and behavioral disorders due to use of tobacco -> Other chronic obstructive pulmonary disease -> Malignant neoplasm of bronchus and lung	13,081	0.51%
Disorders of lipoprotein metabolism and other lipidemias -> Essential (primary) hypertension -> Chronic ischemic heart disease	12,123	0.48%
Essential (primary) hypertension -> Chronic ischemic heart disease -> Acute myocardial infarction	11,692	0.46%
Mental and behavioral disorders due to use of tobacco -> Essential (primary) hypertension -> Chronic ischemic heart disease	11,539	0.45%
Mental and behavioral disorders due to use of tobacco -> Other chronic obstructive pulmonary disease -> Chronic ischemic heart disease	11,205	0.44%
Mental and behavioral disorders due to use of tobacco -> Essential (primary) hypertension -> Other chronic obstructive pulmonary disease	11,044	0.43%
Essential (primary) hypertension -> Chronic ischemic heart disease -> Heart failure	10,824	0.42%
Unspecified diabetes mellitus -> Chronic ischemic heart disease -> Cardiac arrest	9,568	0.38%

REFERENCES

[1] "Top 10 Causes of Death - Factsheet," W. H. O. (W.H.O.), Ed., ed, 2013.

[2] K. D. Kochanek, S. L. Murphy, J. Xu, and B. Tejada-Vera, "Deaths: Final Data for 2014.," *National vital statistics reports : from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System,* vol. 65, pp. 1-122, Jul 2016.

[3] B. Randall, "Death Certification: A Primer Part I-An Introduction to the Death Certificate," *South Dakota Medicine,* vol. 67, 2014.

[4] "Possible Solutions to Common Problems in Death Certification," N. C. f. H. Statistics, Ed., ed: Centers for Disease Control and Prevention 1997.

[5] D. C. Cowper, J. D. Kubal, C. Maynard, and D. M. Hynes, "A primer and comparative review of major US mortality databases," *Annals of epidemiology,* vol. 12, pp. 462-468, 2002.

[6] R. H. Dolin, L. Alschuler, S. Boyer, C. Beebe, F. M. Behlen, P. V. Biron*, et al.*, "HL7 clinical document architecture, release 2," *Journal of the American Medical Informatics Association,* vol. 13, pp. 30-39, 2006.

[7] R. H. Dolin, L. Alschuler, C. Beebe, P. V. Biron, S. L. Boyer, D. Essin*, et al.*, "The HL7 clinical document architecture," *Journal of the American Medical Informatics Association,* vol. 8, pp. 552-569, 2001.

[8] M. Eichelberg, T. Aden, J. Riesmeier, A. Dogac, and G. B. Laleci, "Electronic health record standards-a brief overview," in *Proceedings of the 4th IEEE International Conference on Information and Communications Technology (ICICT 2006)*, 2006.

[9] D. Blumenthal and M. Tavenner, "The "meaningful use" regulation for electronic health records," *New England Journal of Medicine,* vol. 363, pp. 501-504, 2010.

[10] D. Bender and K. Sartipi, "HL7 FHIR: An Agile and RESTful approach to healthcare information exchange," in *Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on*, 2013, pp. 326-331.

[11] G. Alterovitz, J. Warner, P. Zhang, Y. Chen, M. Ullman-Cullere, D. Kreda*, et al.*, "SMART on FHIR Genomics: Facilitating standardized clinico-genomic apps," *Journal of the American Medical Informatics Association,* vol. 22, pp. 1173-1178, 2015.

[12] S. N. Kasthurirathne, B. Mamlin, H. Kumara, G. Grieve, and P. Biondich, "Enabling Better Interoperability for HealthCare: Lessons in Developing a Standards Based Application Programing Interface for Electronic Medical Record Systems," *Journal of Medical Systems,* vol. 39, p. 182, October 07 2015.

[13] R. A. Hoffman, H. Wu, J. Venugopalan, P. Braun, and M. D. Wang, "Intelligent mortality reporting with FHIR," in *2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, 2017, pp. 181-184.

[14] N. C. f. H. Statistics, "Health, United States, 2014: With special feature on adults aged 55–64," 2015.

[15] C. Tao, K. Wongsuphasawat, K. Clark, C. Plaisant, B. Shneiderman, and C. G. Chute, "Towards event sequence representation, reasoning and visualization for EHR data," presented at the Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium, Miami, Florida, USA, 2012.

[16] T. D. Wang, C. Plaisant, A. J. Quinn, R. Stanchak, S. Murphy, and B. Shneiderman, "Aligning temporal data by sentinel events: discovering patterns in electronic health records," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2008, pp. 457-466.

[17] H. Syed and A. K. Das, "Identifying Chemotherapy Regimens in Electronic Health Record Data Using Interval-Encoded Sequence Alignment," in *Artificial Intelligence in Medicine*, ed: Springer, 2015, pp. 143-147.

[18] I. J. Casanova, M. Campos, J. M. Juarez, A. Fernandez-Fernandez-Arroyo, and J. A. Lorente, "Using Multivariate Sequential Patterns to Improve Survival Prediction in Intensive Care Burn Unit," in *Artificial Intelligence in Medicine: 15th Conference on Artificial Intelligence in Medicine, AIME 2015, Pavia, Italy, June 17-20, 2015. Proceedings*, H. J. Holmes, R. Bellazzi, L. Sacchi, and N. Peek, Eds., ed Cham: Springer International Publishing, 2015, pp. 277-286.

[19] I. Batal, H. Valizadegan, G. F. Cooper, and M. Hauskrecht, "A pattern mining approach for classifying multivariate temporal data," in *Bioinformatics and Biomedicine (BIBM), 2011 IEEE International Conference on*, 2011, pp. 358-365.

[20] H. Yang and C. C. Yang, "Using Health-Consumer-Contributed Data to Detect Adverse Drug Reactions by Association Mining with Temporal Analysis," *ACM Trans. Intell. Syst. Technol.,* vol. 6, pp. 1-27, 2015.

[21] R. Bellazzi, F. Ferrazzi, and L. Sacchi, "Predictive data mining in clinical medicine: a focus on selected methods and applications," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery,* vol. 1, pp. 416-430, 2011.

[22] Q. Zhao and S. S. Bhowmick, "Sequential pattern mining: A survey," *ITechnical Report CAIS Nayang Technological University Singapore,* pp. 1-26, 2003.

[23] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules," in *Proc. 20th int. conf. very large data bases, VLDB*, 1994, pp. 487-499.

[24] R. Agrawal and R. Srikant, "Mining sequential patterns," in *Data Engineering, 1995. Proceedings of the Eleventh International Conference on*, 1995, pp. 3-14.

[25] M. J. Zaki, "SPADE: An efficient algorithm for mining frequent sequences," *Machine learning,* vol. 42, pp. 31-60, 2001.

[26] J. Pei, J. Han, B. Mortazavi-Asl, J. Wang, H. Pinto, Q. Chen*, et al.*, "Mining sequential patterns by pattern-growth: The prefixspan approach," *Knowledge and Data Engineering, IEEE Transactions on,* vol. 16, pp. 1424-1440, 2004.

[27] M.-Y. Lin and S.-Y. Lee, "Fast discovery of sequential patterns by memory indexing," in *Data Warehousing and Knowledge Discovery*, ed: Springer, 2002, pp. 150-160.

[28] S. Boytcheva, G. Angelova, D. Tcharaktchiev, and Z. Angelov, "Big Data Analytics in Healthcare–Pattern Mining of Temporal Clinical Events," 2011.

[29] A. P. Wright, A. T. Wright, A. B. McCoy, and D. F. Sittig, "The use of sequential pattern mining to predict next prescribed medications," *Journal of biomedical informatics,* vol. 53, pp. 73-80, 2015.

[30] B. R. Mistry and A. Desai, "Privacy preserving heuristic approach for association rule mining in distributed database," in *Innovations in Information, Embedded and Communication Systems (ICIIECS), 2015 International Conference on*, 2015, pp. 1-7.

[31] A. W.-C. Fu, R. C.-W. Wong, and K. Wang, "Privacy-preserving frequent pattern mining across private databases," in *Data Mining, Fifth IEEE International Conference on*, 2005, p. 4 pp.

[32] K. Sathiyapriya and G. S. Sadasivam, "A survey on privacy preserving association rule mining," *International Journal of Data Mining & Knowledge Management Process,* vol. 3, p. 119, 2013.

[33] J. Soni, U. Ansari, D. Sharma, and S. Soni, "Predictive data mining for medical diagnosis: An overview of heart disease prediction," *International Journal of Computer Applications,* vol. 17, pp. 43-48, 2011.

[34] S. Concaro, L. Sacchi, C. Cerra, P. Fratino, and R. Bellazzi, "Mining health care administrative data with temporal association rules on hybrid events," *Methods Inf Med,* vol. 50, pp. 166-79, 2011.

[35] G. Dolce, M. Quintieri, S. Serra, V. Lagani, and L. Pignolo, "Clinical signs and early prognosis in vegetative state: a decisional tree, data-mining study," *Brain Injury,* vol. 22, pp. 617-623, 2008.

[36] M. R. McNally, C. L. Patton, and W. J. Fremouw, "Mining for Murder-Suicide: An Approach to Identifying Cases of Murder-Suicide in the National Violent Death Reporting System Restricted Access Database," *Journal of forensic sciences,* 2015.

[37] R. Srikant and R. Agrawal, *Mining sequential patterns: Generalizations and performance improvements*: Springer, 1996.

[38] J. Wang, J. Han, and C. Li, "Frequent closed sequence mining without candidate maintenance," *Knowledge and Data Engineering, IEEE Transactions on,* vol. 19, pp. 1042-1056, 2007.