



HHS Public Access

Author manuscript

Comput Biol Med. Author manuscript; available in PMC 2016 February 17.

Published in final edited form as:

Comput Biol Med. 2014 October ; 53: 203–205. doi:10.1016/j.combiomed.2014.07.022.

Informatics can identify systemic sclerosis (SSc) patients at risk for scleroderma renal crisis

Doug Redd^{1,2,3}, Tracy M. Frech^{1,2}, Maureen A. Murtaugh^{1,2}, Julia Rhiannon⁴, and Qing T. Zeng^{1,2,3}

¹Veterans Affairs Medical Center Salt Lake City Health Care System, Salt Lake City, Utah

²Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, Utah

³Department of Biomedical Informatics, University of Utah School of Medicine, Salt Lake City, Utah

⁴Veterans Affairs Medical Center Denver Health Care System, Denver, Colorado

Abstract

Background—Electronic medical records (EMR) provide an ideal opportunity for the detection, diagnosis, and management of systemic sclerosis (SSc) patients within the Veterans Health Administration (VHA). The *objective* of this project was to use informatics to identify potential SSc patients in the VHA that were on prednisone, in order to inform an outreach project to prevent scleroderma renal crisis (SRC).

Methods—The electronic medical data for this study came from Veterans Informatics and Computing Infrastructure (VINCI). For natural language processing (NLP) analysis, a set of retrieval criteria was developed for documents expected to have a high correlation to SSc. The two annotators reviewed the ratings to assemble a single adjudicated set of ratings, from which a support vector machine (SVM) based document classifier was trained. Any patient having at least one document positively classified for SSc was considered positive for SSc and the use of prednisone ≥ 10 mg in the clinical document was reviewed to determine whether it was an active medication on the prescription list.

Results—In the VHA, there were 4,272 patients that have a diagnosis of SSc determined by the presence of an ICD-9 code. From these patients, 1,118 patients (21%) had the use of prednisone ≥ 10 mg. Of these patients, 26 had a concurrent diagnosis of hypertension, thus these patients should not be on prednisone. By the use of natural language processing (NLP) an additional 16,522 patients were identified as possible SSc, highlighting that cases of SSc in the VHA may

Corresponding Author: Tracy M. Frech MD, MS, 4B200 SOM 30 N 1900 E, Salt Lake City, UT 84108, Fax: 801 581 6069, Telephone: 801 581 4334, tracy.frech@hsc.utah.edu.

The authors have no conflicts of interests.

All of our authors Doug Redd, Tracy M. Frech, Maureen A. Murtaugh, Julia Rhiannon, and Qing T. Zeng have declared no conflicts of interest.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

exist that are unidentified by ICD-9. A 10-fold cross validation of the classifier resulted in a precision (positive predictive value) of 0.814, recall (sensitivity) of 0.973, and f-measure of 0.873.

Conclusions—Our study demonstrated that current clinical practice in the VHA includes the potentially dangerous use of prednisone for veterans with SSc. This present study also suggests there may be many undetected cases of SSc and NLP can successfully identify these patients.

Keywords

Informatics; systemic sclerosis; scleroderma; renal crisis; prednisone; steroid; blood pressure; hypertension; natural language processing; management

Introduction

Systemic sclerosis (SSc; scleroderma) is a rare, complex autoimmune disease in which a poor prognosis is most closely related to organ fibrosis and/or hypertensive crisis. Hypertension is the key factor in the development of sudden kidney failure in SSc, which is called scleroderma renal crisis (SRC).¹ SRC is characterized by malignant hypertension and oliguric/anuric acute renal failure, and occurs in 2% to 5% of patients with systemic sclerosis (SSc).² If SRC occurs there is a 5-year survival rate of 65%, thus this condition is important to identify and prevent in SSc patients.

Several retrospective studies have found that a significant, but perhaps not widely recognized risk factor for SRC is recent use of prednisone.^{2,3} The use of prednisone for any indication in SSc remains controversial, but should only be used at the lowest possible dose (ideally <10 mg) and reserved for myositis, arthritis, interstitial lung disease, and inflammatory skin disease.⁴ SSc patients should be educated to monitor their blood pressure and to take the new onset on hypertension seriously.⁵ It is critical to detect SRC in its earliest stages because prompt treatment with the BP-lowering class of drugs called angiotensin-converting enzyme inhibitors (ACE-inhibitors) can help to prevent progression to serious kidney failure.⁶

Unfortunately, over half of cases of SRC have a delay in diagnosis, require dialysis and long-term mortality remains significant.⁷ Use of prednisone in a SSc patient and/or a delay in diagnosis of SRC unfortunately can result in high morbidity and mortality due to unnecessary delays in the referral process.⁸

Electronic medical records (EMR) provide an ideal opportunity for the detection, diagnosis, and management of SSc patients within the Veterans Health Administration (VHA). The VHA has one of the largest integrated healthcare systems in the United States with 8.6 million total enrollees in 2012. The VA Health Services Research and Development (HSR&D) office funded the Veterans Informatics and Computing Infrastructure (VINCI), which began operations in June 2008. VINCI is collaboration between the Office of Information and Technology (OI&T) and the Office of Research and Development (OR&D). VINCI was created to serve the data and Information Technology (IT) needs of the VA research community. The VINCI database is an excellent example of big data, a massive volume of both structured and unstructured data that is so large that it's difficult to

process using traditional database and software techniques. At the time of this study, VINCI provides access to structured and unstructured electronic medical data on 17,543,172 unique patients. VINCI and Consortium for Healthcare Informatics Research (CHIR) have shown that the EMR can be effectively utilized in a de-identified manner for patient safety and quality measurement.⁹

The objective of this project was to use informatics to identify SSc patients in the Veterans Health Administration (VHA) that were prescribed prednisone in order to inform an outreach project to prevent SRC. The investigators had two goals with this informatics project: (1) to identify current SSc patients that may be inappropriate therapy, and (2) to identify if there are potential SSc patients that are not identified by ICD-9 code in order to better understand the potential impact of an outreach project.

Methods

The electronic medical data for this study came from VINCI and was approved for use by the Institutional Review Board. No human subjects were contacted during this research. The data consisted of structured (i.e. problem lists, medication lists, lab reports, demographics) as well as unstructured (i.e. clinical notes) data. For natural language processing (NLP) analysis, a set of retrieval criteria was developed for documents expected to have a high correlation to SSc. These criteria were: patient had at least one systemic sclerosis diagnosis by a rheumatologist or at least two diagnoses by a primary care provider and the document written by that provider contained the text “systemic sclerosis” or “scleroderma.” Snippets containing “systemic sclerosis” or “scleroderma” were extracted from these documents, manually reviewed by two annotators, and assigned ratings of “Yes”, “No”, or “Uncertain” for the positive indication of SSc. In rating the snippets, additional terms that are strong indicators of SSc based on the classification criteria for this condition¹⁰: “skin thickening of fingers”, “digital tip ulcers”, “fingertip pitting scars”, “telangiectasia”, “abnormal nailfold capillaries”, “pulmonary arterial hypertension”, “interstitial lung disease”, “Raynaud's phenomenon”, “anticentromere”, “anti-topoisomerase”, “anti-RNA polymerase III”, or “scleroderma-related autoantibodies” were used. The antinuclear antibody (ANA) status was not reviewed.

The two annotators reviewed the ratings to assemble a single adjudicated set of ratings, from which a support vector machine (SVM) based document classifier was trained. Any patient having at least one document positively classified for SSc was considered positive for SSc. Once a patient was confirmed as definite SSc, the use of prednisone 10 mg in the clinical document was manually reviewed to determine whether it was an active medication on the prescription list.

Results

In the VHA, there were 4,272 patients that have a diagnosis of SSc determined by the presence of an ICD-9 code; all of these patients records were available for review. From a search of these patients, 1,118 patients (21%) had the use of prednisone \geq 10 mg documented in the EMR. Of these 1,118 SSc patients on steroid, 26 had a concurrent

diagnosis of hypertension and no clear plan educating the patient to monitor their blood pressure confirmed by manual review. Thus, these 26 patients were potentially being managed inappropriately.

From this manual review, the average age of the sample was 63 years. In this population 63% were confirmed to be on prednisone. Three patients were prescribed this therapy for gout; all others were prescribed this medication “for SSc” per the medical record. Only in 11 cases was indication specified as lung disease, tenosynovitis, or arthritis. No cases of myositis were identified. Prednisone doses as high as 60 mg were recorded in the EMR. The indication for the use of high rather than low dose prednisone was not clear. In the medical records of 37% of the SSc patients that were not on prednisone, phrases such as “allergy to prednisone,” “localized scleroderma,” “patient advised not to use prednisone,” or “past use of prednisone” were documented. Document types including anesthesia notes and disability forms were universally not accurate due to negation terms used, such as “this patient does not have scleroderma.”

There are limitations to our approach. We did not specifically look at ANA status because it is not a part of systemic sclerosis classification criteria, however this would have been helpful for understanding potential SSc cases. For the manual review, if scleroderma specific antibody information did not appear in the note, we could not confirm this data. Thus, low dose prednisone use in an anti-centromeric antibody SSc patient with long standing disease may not be a dangerous practice pattern, but is not captured by our study. Additionally, the study design does not adequately capture the prevalence of hypertension in this SSc population since we did not adjust for age, gender, and race/ethnicity. Patients that had both SSc (ICD-9 710.1) and localized scleroderma patients (ICD-9 701.0) coded were not excluded until NLP was applied. In this stage, only 3 patients were identified as having localized scleroderma. Thus this present study which suggests there may be many undetected cases of SSc and NLP can successfully identify these patients does not adequately exclude localized SSc. However, we demonstrated that NLP does have the ability to distinguish between these two diseases by negation terminology.

By the use of NLP an additional 16,522 patients were identified as possible SSc from all VHA records all multiple centers. From these 16,522 patients, the two annotators reviewed and rated 244 snippets. Snippets were selected by chronological date and were from multiple VAMC throughout the United States. A 10-fold cross validation of the classifier resulted in a precision (positive predictive value) of 0.814, recall (sensitivity) of 0.973, and f-measure of 0.873. Using this classifier the entire set of documents meeting the criteria was classified.

Conclusions

Current rheumatology guidelines emphasize early detection and effective management of SRC and highlight the risk of prednisone for SSc patients.^{3,11} Our study demonstrated that current clinical practice in the VHA includes the potentially dangerous use of prednisone for veterans with SSc, including the use of high dose steroid as well as unclear indications for its use. Medical plans from providers managing SSc patients with hypertension and on prednisone, did not document that patients were informed to track their blood pressure. Our

study also identified many additional patients by NLP that may have a diagnosis of SSc, but were not identified by ICD-9 coding. While the absence of serology, skin score information, and patient reported outcomes in the standard note structure is a limitation to this study, the snippets of diagnostic information, did allow for identification of SSc patients. This study suggests that there may be an unusually high prevalence of SSc among veterans, which warrants further investigation.

Advances in informatics allow identification of SSc patients potentially at risk for SRC and provides the opportunity to improve quality of care in these patients through education to clinical providers. This present study also suggests there may be many undetected cases of SSc and NLP can successfully identify these patients. Manual review of cases can help providers restrict the search terms to train a classifier which will recognize phrases, such as “patient advised not to use prednisone”, in order to implement alerts appropriately. NLP may allow better identification of possible SSc patients and aid providers in the care of US veterans.

Acknowledgments

Funding: Dr. Frech was supported by the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant 8UL1TR000105 (formerly UL1RR025764). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

References

1. Denton CP, Lapadula G, Mouthon L, Muller-Ladner U. Renal complications and scleroderma renal crisis. *Rheumatology (Oxford)*. 2009; 48(Suppl 3):iii32–5. [PubMed: 19487221]
2. Bussone G, Berezne A, Pestre V, Guillevin L, Mouthon L. The scleroderma kidney: progress in risk factors, therapy, and prevention. *Current rheumatology reports*. 2011; 13:37–43. [PubMed: 21061100]
3. Guillevin L, Berezne A, Seror R, et al. Scleroderma renal crisis: a retrospective multicentre study on 91 patients and 427 controls. *Rheumatology (Oxford)*. 2012; 51:460–7. [PubMed: 22087012]
4. Perez Campos D, Estevez Del Toro M, Pena Casanovas A, Gonzalez Rojas PP, Morales Sanchez L, Gutierrez Rojas AR. Are high doses of prednisone necessary for treatment of interstitial lung disease in systemic sclerosis? *Reumatologia clinica*. 2012; 8:58–62. [PubMed: 22317852]
5. Frech TM, Penrod J, Battistone MJ, Sawitzke AD, Stults BM. The prevalence and clinical correlates of an auscultatory gap in systemic sclerosis patients. *International journal of rheumatology*. 2012; 2012:590845. [PubMed: 22518163]
6. Shanmugam VK, Steen VD. Renal disease in scleroderma: an update on evaluation, risk stratification, pathogenesis and management. *Current opinion in rheumatology*. 2012; 24:669–76. [PubMed: 22955019]
7. Penn H, Denton CP. Diagnosis, management and prevention of scleroderma renal disease. *Current opinion in rheumatology*. 2008; 20:692–6. [PubMed: 18946330]
8. Farber HW, Simms RW, Lafyatis R. Care of patients with scleroderma in the intensive care setting. *Journal of intensive care medicine*. 2010; 25:247–58. [PubMed: 20542965]
9. Mayer J, Shen S, South BR, et al. Inductive creation of an annotation schema and a reference standard for de-identification of VA electronic clinical notes. *AMIA Annual Symposium proceedings / AMIA Symposium*. 2009; 2009:416–20. [PubMed: 20351891]
10. van den Hoogen F, Khanna D, Fransen J, et al. 2013 classification criteria for systemic sclerosis: an American college of rheumatology/European league against rheumatism collaborative initiative. *Annals of the rheumatic diseases*. 2013; 72:1747–55. [PubMed: 24092682]

11. Trang G, Steele R, Baron M, Hudson M. Corticosteroids and the risk of scleroderma renal crisis: a systematic review. *Rheumatology international*. 2012; 32:645–53. [PubMed: 21132302]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript