Representation of Functional Status Concepts from Clinical Documents and Social Media Sources by Standard Terminologies

Jinqiu Kuang, M.S. [1], April F. Mohanty, Ph.D.[2], M.P.H., Rashmi V. H. , M.Tech. [1], Charlene R. Weir, Ph.D., R.N. [1,2], Bruce E. Bray, M.D. [1], Qing Zeng-Treitler, Ph.D.[1,2]

[1]Biomedical Informatics, School of Medicine, University of Utah, Salt Lake City, UT, USA
[2]Informatics Decision-Enhancement and Analytic Sciences (IDEAS) Center, George E. Wahlen Department of Veterans Affairs Medical Center, Salt Lake City, UT, USA

Abstract

Patient-reported functional status is widely recognized as an important patient-centered outcome that adds value to medical care, research, and quality improvement. Functional status outcomes are, however, not routinely or uniformly collected in the medical record, except in certain small patient populations (e.g. geriatrics, nursing home residents). To utilize patient reported functional status for clinical research and practice, we manually collected 2,763 terms from clinical records and social media sites and modeled them on the widely used Short Form-36 Health Survey. We then examined the coverage of the Unified Medical Language System (UMLS) for these functional status terms through automated mapping. Most terms (85.9%) did not have exact matches in the UMLS. The partial matches were prevalent, however, they typically did not capture the terms' exact semantics. Our study suggests that there is a need to extend existing standard terminologies to incorporate functional status terms used by patients and clinicians.

Introduction

Patient reported functional status is widely recognized as an important patient-centered outcome that adds value to medical care, research, and quality improvement[1]. For instance, during the late 1980s researchers observed that patients' functional status was correlated to their compliance with antihypertensive treatment, even when treatment led to successful blood pressure control[2, 3]. Reduction in functional status is frequently the first sign of declining health for patients with chronic conditions and is related to the severity of acute illness and intensity of resource use. Functional status is a prognostic predictor of future risk of hospitalization readmission, morbidity and mortality[4, 5]. Functional status information is central to many healthcare decisions, including end-of-life care, living arrangements, and patient-tailored treatment. Some experts assert that functional status should be included as the "sixth vital sign"[6].

In the context of atrial fibrillation and heart failure, alteration to functional status is one of the most commonly reported symptoms, including reductions in exercise tolerance of 15-20% among patients diagnosed with atrial fibrillation[7, 8]. Measurement of functional status can help to inform diagnosis, prognosis, and can guide care/treatment strategies related to these conditions, in addition to facilitating examinations of treatment response[9].

While instruments such as the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36)[10] have been developed to capture patient reported functional status, functional status is not routinely or uniformly collected in the medical record. At the same time, clinical records as well as social media data do contain a large number of descriptions of patient functional status.

There has been much interest in and effort toward development of methods to extract functional status terms, using existing frameworks, such as the International Classification of Functioning, Disability and Health (ICF), especially for healthcare delivery for nursing home and rehabilitation patients[11-14]. The ICF was introduced 2001 by the World Health Organization as "a unified and standard language and framework for the description of health and health-related states"[15]. This standard terminology is currently included in the Unified Medical Language System (UMLS), a compendium of over 100 national and international vocabularies and classifications.

As these and other standard terminologies have been developed, it is important to update findings related to their coverage of functional status concepts. Also, since prior work on functional status concepts has been limited to the rheumatology care setting, there is a need to extend these findings to other settings including other medical care settings and consumer health.

In this study, we specifically examined functional status documented in the electronic medical record (EMR) and social media for patients with cardiovascular diagnoses such as atrial fibrillation. The terms were categorized and mapped to standard clinical terminologies. We also analyzed the causes of non-matches and partial matches.

Methods

Data Collection

We defined functional status terms as a span of text that expresses an idea or description corresponding to functional status. To collect functional status terms, we used clinical records and social media, both of which have been increasingly utilized by biomedical researchers. In social media, patients voluntarily describe functional status in their own terms. In the EMR, patient reports are often paraphrased and sometimes clinicians document functional status using a variety of instruments.

EMR data were obtained from the Veterans Administration Informatics and Computing Infrastructure (VINCI) database[17]. The Institutional Review Board approved the use of electronic medical data from VINCI and no human subjects were contacted for this study. We randomly selected a set of 800 clinical documents of patients who were diagnosed with atrial fibrillation or atrial flutter for review. Clinical documents were free text documents or progress notes such as cardiology visit notes, primary care nursing/physician notes, telephone encounter notes, palliative/geriatric medicine notes, and discharge summaries. Generally, these clinical documents do not use standardized terminologies to convey information regarding functional status.

Social media data were obtained from three different online discussion forums for cardiovascular diseases and atrial fibrillation. These sources were: (1) Atrial Fibrillation Support Group on dailystrength.org, (2) Heart & Cardiovascular Disease forum on healingwell.com, and (3) Cardiovascular Disease Prevention Expert Forum on medhelp.org. Questions, replies and comments posted by patients were extracted from the discussion forums using Web Scraper, a Chrome browser extension used for extraction of data from web pages. From the extracted content, 150 latest posts from source 1 and 100 latest posts each from source 2 and source 3 were chosen.

For both datasets, one investigator then manually reviewed clinical documents or posts containing word/phrases that indicated functional status of the patient according to the SF-36 Health Survey, with the intent to identify all potential candidate terms. Then a second investigator reviewed all the extracted terms. When a consensus could not be reached, a third adjudicator was included in the extraction of functional status terms.

We excluded duplicate words/phrases and lists of unique functional status terms from VINCI and social media data sources were created. For the VINCI data, we collected 1,050 functional status terms and removed duplicate terms. A total of 974 unique terms were mapped to UMLS. For the social media data, a total of 1,623 functional status terms were extracted out of which 980 terms were unique.

Classification of Functional Status Terms Based on SF-36 Health Survey Scales

To understand the types functional status terms being used in medical records and social media, we categorized them according to the SF-36 Health Survey scales. The SF-36 Health Survey includes eight scales: Physical Functioning, Role-Physical, Bodily Pain, General Health, Vitality, Social Functioning, Role-Emotional and Mental Health (http://www.sf-35.org/tools/SF36.shtml). Scores for each of these scales are determined by all but one (self-reported health transition) of the 36 items of the SF-36 Measurement Model. There are two summary measures of the SF-36 Health Survey that the eight scales aggregate to. The Physical Health summary measure includes Physical Functioning, Role-Physical, Bodily Pain, and General Health. The Mental Health summary measure includes Vitality, Social Functioning, Role-Emotional, and Mental Health.

Using the 35 items listed for the eight scales as a reference, three investigators jointly reviewed a subset of VINCI and social media terms to develop a classification guideline for assigning each term to one of the eight SF-36 Health Survey scales. The three investigators divided the remaining terms (two investigators reviewed the VINCI terms and one reviewed the social media terms) to independently review and assign them to one of the eight SF-36 Health Survey scales and based on the classification guideline.

Classification of Type of Match by Standard Terminologies

For the purpose of mapping functional status terms to the UMLS Metathesaurus, we used the MetaMap system developed by the National Library of Medicine (NLM)[18]. Unique functional status terms from VINCI and social media were parsed using MetaMap to determine the corresponding concepts in UMLS. The parsed terms, mapped concepts (along with concept unique identifiers [CUI]) and respective UMLS source vocabularies were recorded. Three investigators divided the functional status terms to independently review and categorize them as one of three types of matches: a 'complete match', a 'partial match', or a 'no match'. Terms where the entire word/phrase mapped to a single concept in UMLS were categorized as 'complete match'. Compound terms consisting of two or more words where only a few words mapped to UMLS concepts were categorized as 'partial match'. For example a 'partial match' could include "jogging (6 mph)", where "jogging" mapped to the UMLS but the qualifier "(6 mph)" did not map. Terms that did not map to any UMLS concepts were categorized as 'no match', an example of a term that did not map to any UMLS concepts was "debilitated". We further assessed whether complete matches were correct or incorrect. We defined complete incorrect matches as terms where the entire word/phrase mapped to a single concept in UMLS and did not capture the correct semantic meaning. An example of a complete incorrect match would be the term "fine", where the incorrect semantic meaning would be a financial fine versus feeling fine.

Review of Classifications

Once all of the functional status terms were manually assigned to one of the eight SF-36 Health Survey scales and were categorized as a complete-, partial-, or no-match, the three investigators exchanged lists of classified terms to reach a consensus on the final SF-36 Health Survey scale assignment and type of match for each term. Where there was a disagreement between two investigators, the final assignment was reached by including a third adjudicator.

Data Analysis

We examined the overlap of functional status terms and matched concepts across the two sources (i.e. VINCI or social media). For each source of functional status terms we examined the frequency distributions of functional status terms across SF-36 Health Survey scales. Percentages of partially matched terms, completely matched terms and terms that yielded no match under each of the SF-36 Health Survey scale were determined. We also calculated the percent coverage of functional status concepts by various vocabulary sources of the UMLS Metathesaurus.

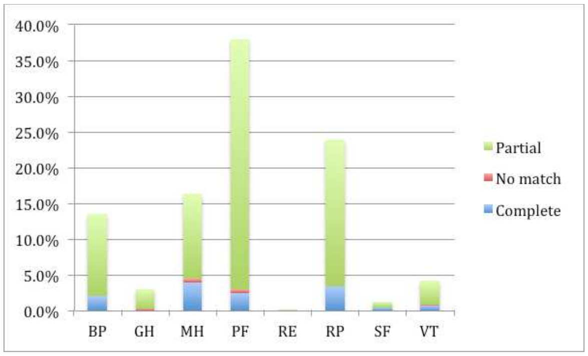
Results

Description of Functional Status Concepts in VHA and Social Media Sources

Overall, there was an overlap of 14 terms (exact term match) across both functional status term sources (i.e. VINCI and social media data). The overlap between mapped UMLS concepts from VINCI and social media data was much larger - 44.7% (or n = 378 concepts) of the mapped concepts from social media data were also found in VINCI data.

The most common SF-36 Health Survey scales that the extracted 974 VHA terms were classified to were Physical Functioning (37.9%), Role-Physical (23.9%), and Mental Health (16.3%), Figure 1a. The top three SF-36 Health Survey scales for the social media terms were Physical Functioning (31.9%), Mental Health (23.5%), and Role-Physical (13.5%), Figure 1b. Fewer than 2.0% of terms were classified to Role-Emotional or Social Functioning scales, regardless of source of functional status term.

a. Data Source: VINCI



b. Data Source: Social Media

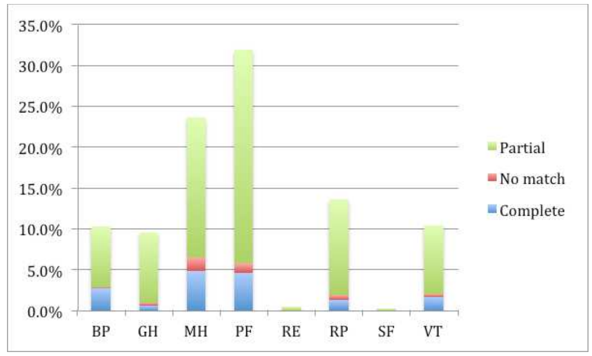


Figure 1. The Distribution of Three Types of Match by the Eight SF-36 Health Survey Scales
Abbreviations: VINCI, Veterans Affairs Informatics and Computing Infrastructure; BP, Bodily Pain; GH, General Health; MH, Mental Health; PF, Physical Functioning; RE, Role-Emotional; SF, Social Functioning; VT, Vitality

Coverage of Functional Status Concepts by Standard Terminologies

The comparison of the number of functional status terms that fell into 'complete match', 'no match' and 'partial match' categories in VINCI and social media data is shown in Figure 1 and Table 1. Overall, 14.1% of functional status terms were completely matched and 83.2% were partially matched to UMLS. Social media data contained a larger number of 'complete match' and 'no match' terms when compared to VINCI data. Most complete matches were in the Mental Health scale SF-36 Health Survey regardless of source of functional status term (Figure 1). Most match failures ('partial-' and 'no-match' combined) corresponded to the Physical Functioning SF-36 Health Survey scale, regardless of data source.

Table 1. Overall Match Percent for Terms from VINCI and Social Media Data Sources

Match Type	VINCI N (%)	Social Media N (%)
Complete	123 (12.6)	153 (15.6)
No match	12 (1.2)	40 (4.1)
Partial	839 (86.1)	787 (80.3)
Abbreviations: VINCI, Veterans Affairs Informatics and Computing Infrastructure		

The UMLS Metathesaurus is made up of more than 100 national and international source vocabularies and classifications. A concept can be found in one or multiple vocabularies. Mapped concepts from VINCI and social media data were spread across 43 different source vocabularies in UMLS. Table 2 lists the top 10 source vocabularies with the most mapped concepts. The top four vocabularies that had the most coverage were MTH, SNOMEDCT_US, National Cancer Institute (NCI) and Consumer Health Vocabulary (CHV) for both VINCI and social media data. Combination of these four vocabularies covered 96.2% of the mapped concepts from VINCI data and 97.3% of the mapped concepts from social media data.

Table 2. Top 10 UMLS Source Vocabularies for VINCI and Social Media Data Sources

Source Vocabulary (% Coverage)	
VINCI	Social Media
MTH (78.2)	MTH (81.9)
SNOMEDCT_US (48.6)	SNOMEDCT_US (51.9)
NCI (40.2)	NCI (47.6)
CHV (34.8)	CHV (36.4)
MSH (18.2)	MSH (18.3)
NLMSubSyn (17.0)	NLMSubSyn (15.2)
SNMI (6.8)	AOD (11.8)
LNC (6.2)	SNMI (10.7)
AOD (5.6)	LNC (4.6)
HL7V3.0 (3.1)	ICF (4.1)
Abbreviations: VINCI, Veterans Affairs Informatics and Computing Infrastructure Source vocabularies described at: http://www.nlm.nih.gov/research/umls/knowledge_sources/metathesaurus/release/source_vocabularies.html	

Some functional status terms from VINCI and social media data yielded complete matches, most of which were correctly matched. There were 123 complete matched terms from VINCI out of which 120 were correctly mapped and were found in 26 UMLS source vocabularies. Social media data had 153 complete matches out of which 141 mapped to correct concepts. These concepts were found in 24 different vocabulary sources of the UMLS Metathesaurus. Percentages of coverage of completely matched correct concepts by 10 source vocabularies that had maximum coverage for VINCI data and social media data are reported in Table 2b. SNOMEDCT_US, MTH, CHV and MSH were the top 4 source vocabularies with respect to coverage of completely matched terms that were mapped to correct concepts for both VINCI data and social media data. Combination of these top 4 source vocabularies contained 96.7% and 93.61% of the correctly mapped concepts from VINCI data and social media data, respectively.

Table 3. Top 10 UMLS Source Vocabularies for Complete and Correct Matches

Source Vocabulary (% Coverage)	
VINCI	Social Media
SNOMEDCT_US (65.8)	SNOMEDCT_US (66.0)
MTH (63.3)	MTH (65.3)
CHV (50.8)	CHV (55.3)
MSH (41.7)	MSH (46.1)
NLMSubSyn (38.3)	NCI (35.5)

NCI (27.5)	NLMSubSyn (33.3)
SNM (20.8)	SNMI (27.7)
AOD (17.5)	AOD (21.3)
CST (17.5)	CST (16.3)
ICD10CM (12.5)	OMIM (10.6)
Abbreviations: VINCI, Veterans Affairs Informatics and Computing Infrastructure Source vocabularies described at: http://www.nlm.nih.gov/research/umls/knowledge_sources/metathesaurus/release/source_vocabularies.html	

Most of the partial matches did not capture the exact semantics of the terms. We did not attempt to assess the correctness of the partial matches because frequently, a part that matched could be considered to be correct while another part could not. For instance, for the term "back to normal" the term "back" mapped to a number of concepts none of which were correct and the term "normal" mapped to the concept "Normal" which was correct. Together, they also failed to represent the complete semantic meaning.

We explored reasons for failures for a subset of non-matches and partial matches, these examples are presented in Table 4.

Table 4. Examples of Type of Match to UMLS and Reasons For Failures

Type of Match	Examples	Reason for Failure
No Match	Too zonked Debilitated Annoying Cowering Unfocused Go to gym	Concept does not exist in UMLS
	Doing OK	Use of colloquial term
Partial Match	Jogging (6 mph) Extreme fatigue	No match for qualifier/modifier
	Have problems tinkering around in the garage and outside Has difficulty taking arm to the side and away from the body Exercises daily on treadmill alternating with walking 3-4 miles	Too many descriptive details
	Feels lousy Drives me nuts	Use of colloquial term
	PAIN BEST: 4/10 FIM Score Discharge:[6] transfers [1] amb [1] stairs - Total [8]	Use of templates (specific to VINCI data)
	Feel like someone's choking me Feels like death taking over In a fog	Use of metaphors (specific to social media data)
Abbreviations: UMLS, Unified Medical Language System		

Discussion

With the goal of utilizing patient-reported functional status in research and clinical practice, we collected and analyzed 1,954 unique functional status terms extracted from VINCI EMR data among patients diagnosed with atrial fibrillation or atrial flutter and from social media posts of three cardiovascular disease-related online discussion forums. Overlap of the exact term across the two sources of functional status terms (i.e. EMR clinical documents containing free text data or social media posts) was small although, further examination of the overlap of mapped

UMLS functional status concepts across sources was much larger. The most common types of terms according the eight scales of the SF-36 Health Survey were Physical Functioning (34.9% of all terms) and Mental Health (20% of all terms). The least common types of terms were Role-Emotional (<0.1% of all terms) and Social Functioning (1.0% of all terms).

Overall, the coverage of functional status terms by current standard clinical terminologies was low: only 12.6% of terms extracted from VHA and 15.6% of terms extracted from social media were completely mapped to UMLS concept names. Some complete matches were not correct. Examples include the term "work out" (as in exercise) was mapped to "Unemployment" and the term "cardio" (again, as in exercise) was mapped to "Heart". Partial matches can be found for most of the remaining terms. They, however, do not reflect the complete semantic meaning of the terms. In many cases, they do not capture the core semantic meaning either. Natural Language Processing (NLP) frequently uses standards for automated data retrieval and there is a need for mapping functional status descriptions to standard terminologies. Although standard terminologies may contain many of the important and useful concepts related to functional status, our study suggests that the use of NLP for functional status data extraction requires mapping ontologies to more descriptive terms and concepts such as the ones we identified.

While the percent coverage of functional status terms by standard clinical terminologies was similar comparing VINCI EMR to social media sources, there was a lack of overlap in the terms used across sources. This lack of exact overlap of terms between the VINCI EMR and social media online discussion forums reflects the differences in the context and purpose of the writing as well as their authors. Aside from paraphrasing patients' reported functional status (e.g. "ambulate with assistance") about their activities, EMRs tend to use medical language as well as templates for functional assessment instruments (e.g. Activities of Daily Living (ADLs), Functional Independence Measure (FIM)). Patients who describe their functional status on a social media website may tend to use their own language which can be more informal, erratic, expressive or ambiguous. It is worth noting that the mapping rates of the two sources are fairly similar. In other words, in this domain, the standard vocabularies represent neither the clinician nor the patient terminology well. In addition, no single vocabulary stood out as a particularly good source for functional status terms. The mapped terms were spread across a number of different terminologies.

We found several reasons for failures for the subset of non-matches and partial matches that we investigated (Table 4). While some concepts did not exist in the UMLS, many more failed to match completely due to use of a colloquial term, no match for a qualifier/modifier or too many descriptive terms. In the VINCI data some terms failed to match due to use of templates. For the social media data some terms were not completely matched due to use of metaphors.

Increasingly, researchers are examining standard terminologies in terms of their support of interoperability and coverage of concepts existing in the biomedical domain. For instance, Bodenreider et al.[19] evaluated coverage by the UMLS for bioinformatics concepts from LocusLink and the Gene Ontology database. Others, including Frost et al.[20] are developing sophisticated methods such as the Markov Chain Ontology Analysis to improve analytical applications for biomedical ontologies such as enrichment analysis, which quantifies the importance of ontology classes relative to a dataset.

However, to the best of our knowledge, we are aware of only one other study by Ruggieri et al.[16] that examined the representation by standard terminologies of functional status concepts. The authors of this study found that the UMLS was superior to SNOMED-RT in matching functional status concepts from the Clinical Health Assessment Questionnaire (CLINHAQ) and the Modified Health Assessment Questionnaire (MHAQ). The CLINHAQ and MHAQ questionnaires are used to evaluate functional status of patients diagnosed with rheumatoid arthritis. Consistent with our study Ruggieri et al. reported that neither terminology had complete coverage of functional status terms and SNOMED-RT coverage was especially poor for concepts in the "activities" semantic class (terms found in the Physical Function and Role-Physical scales in our study).

There are a number of ongoing efforts to standardize functional status terminologies in industry. For example, standard representation of functional status has been actively studied in health information exchange and quality measures. Various groups including the Office of National Coordinator of Health Information Technology (ONC), Centers for Medicare and Medicaid Services (CMS) and Integrating the Healthcare Enterprise (IHE) are proposing functional status standards. The Consolidated-Clinical Document Architecture (C-CDA) specification defined the functional status section and specified vocabulary constraints for different types of functional status.

There has been increasing interest in automated extraction of functional status based on patient notes[11-14], including the use of the ICF standard terminology. One preliminary study, for example, used the ICF to extend an existing NLP system to encode functional status information noted in patient rehabilitation summaries[11]. In our study ICF was among the UMLS terminologies that were examined. The coverage of functional status terms by ICF was low, 1.9% for the VINCI data and 4.1% for social media data. We note that terms in the ICF are generally noun phrases, however as Ruggieri et al.[21] observed with functional-status language in questionnaires, we found that functional status information was mainly expressed with verbal phrases or sentences and included fewer medical terms. This may partly account for the low coverage.

One limitation of our study is that we focused on data from patients with atrial fibrillation or other cardiovascular diseases. Focusing on these patients may have limited the number of terms assigned to the Role-Emotional and Social Functioning SF-36 Health Survey scales. Future studies are needed to extend the research of functional status representation for other clinical domains. A second limitation is that we reviewed documents for a limited number of patients.

To address the gap in vocabulary coverage, we plan to update the Consumer Health Vocabulary (CHV), which is a UMLS source vocabulary developed by our research group, with terms identified in this study. Other possibilities include creating an ontology to represent the variety of functional status concepts and exploring the use of post-coordination. In anticipation of these efforts, one of the challenges we will face is to correctly map or link the terms, which are subjective by nature, to existing medical concepts. For example, it is not clear whether terms like “can’t get out of bed” or “feels like I’m going to die” refer to a physical or mental issue. At the same time they do indicate that the patients are not functioning well.

Conclusions

Individuals describe their physical status in a much more diverse way than current standard vocabularies capture. We collected of a large number (2,763) of patient-reported functional status terms from both clinician and consumer sources of data, including terms from social media. With the increasing emphasis on patient-reported outcomes and the application of automated data retrieval methods, it is important to capture both patients’ own reports as well as those rephrased by clinicians. We further examined the coverage by current standard terminologies of functional status terms extracted from the VINCI EMR and social media data sources, sources that to our knowledge have not been previously examined. Overall, we found that standard terminologies do not sufficiently provide coverage of patient-reported functional status, and could be enhanced by studying these existing textual sources from clinicians and patients.

Acknowledgements

This work is funded by the US Department of Veterans Affairs, Office of Research and Development, Health Services Research and Development grants CHIR HIR 08-374, HIR 08-204, CRE 12-315 and the CREATE: A VHA NLP Software Ecosystem for Collaborative Development and Integration. Dr. Mohanty is supported by the VA Advanced Fellowship Program in Medical Informatics of the Office of Academic Affiliations, Department of Veterans Affairs. We would also like to acknowledge the staff, resources and facilities of the VA Salt Lake City IDEAS Center.

References

1. Laine C and Davidoff F. Patient-centered medicine. A professional evolution. *Jama*. 1996; 275: 152-6.
2. Croog SH, Levine S, Testa MA, et al. The effects of antihypertensive therapy on the quality of life. *The New England journal of medicine*. 1986; 314: 1657-64.
3. Williams GH. Beyond blood pressure control. Effect of antihypertensive therapy on quality of life. *American journal of hypertension*. 1988; 1: 363S-5S.
4. de la Fuente SG, Bennett KM and Scarborough JE. Functional status determines postoperative outcomes in elderly patients undergoing hepatic resections. *J Surg Oncol*. 2013; 107: 865-70.
5. Holland R, Rechel B, Stepien K, Harvey I and Brooksby I. Patients' self-assessed functional status in heart failure by New York Heart Association class: a prognostic predictor of hospitalizations, quality of life and death. *Journal of cardiac failure*. 2010; 16: 150-6.
6. Bierman AS. Functional status: the six vital sign. *Journal of general internal medicine*. 2001; 16: 785-6.
7. Rienstra M, Lubitz SA, Mahida S, et al. Symptoms and functional status of patients with atrial fibrillation: state of the art and future research opportunities. *Circulation*. 2012; 125: 2933-43.
8. Ueshima K, Myers J, Graettinger WF, et al. Exercise and morphologic comparison of chronic atrial fibrillation and normal sinus rhythm. *Am Heart J*. 1993; 126: 260-1.
9. Fleg JL, Pina IL, Balady GJ, et al. Assessment of functional capacity in clinical and research applications: An advisory from the Committee on Exercise, Rehabilitation, and Prevention, Council on Clinical Cardiology, American Heart Association. *Circulation*. 2000; 102: 1591-7.
10. Ware JE, Jr. and Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical care*. 1992; 30: 473-83.
11. Kukafka R, Bales ME, Burkhardt A and Friedman C. Human and automated coding of rehabilitation discharge summaries according to the International Classification of Functioning, Disability, and Health. *J Am Med Inform Assoc*. 2006; 13: 508-15.
12. Bales M, Kukafka R, Burkhardt A and Friedman C. Extending a medical language processing system to the functional status domain. *AMIA Annu Symp Proc*. 2005: 888.
13. Giang P, Williams A and Argyros L. Automated extraction of the Barthel Index from clinical texts. *AMIA Annu Symp Proc*. 2013; 2013: 486-95.
14. Harris MR, Savova GK, Johnson TM and Chute CG. A term extraction tool for expanding content in the domain of functioning, disability, and health: proof of concept. *J Biomed Inform*. 2003; 36: 250-9.
15. Allan CM, Campbell WN, Guptill CA, Stephenson FF and Campbell KE. A conceptual model for interprofessional education: the international classification of functioning, disability and health (ICF). *J Interprof Care*. 2006; 20: 235-45.
16. Ruggieri AP, Elkin P and Chute CG. Representation by standard terminologies of health status concepts contained in two health status assessment instruments used in rheumatic disease management. *Proc AMIA Symp*. 2000: 734-8.
17. VA Informatics and Computing Infrastructure (VINCI)2015: (2012).(Available from:http://www.hsrd.research.va.gov/for_researchers/vinci/).
18. Divita G, Zeng QT, Gundlapalli AV, Duvall S, Nebeker J and Samore MH. Sophia: A Expedient UMLS Concept Extraction Annotator. *Proc AMIA Symp*. 2014: 467-76.
19. Bodenreider O, Mitchell JA and McCray AT. Evaluation of the UMLS as a terminology and knowledge resource for biomedical informatics. *Proc AMIA Symp*. 2002: 61-5.
20. Frost HR and McCray AT. Markov Chain Ontology Analysis (MCOA). *BMC Bioinformatics*. 2012; 13: 23.
21. Ruggieri AP, Pakhomov SV and Chute CG. A corpus driven approach applying the "frame semantic" method for modeling functional status terminology. *Stud Health Technol Inform*. 2004; 107: 434-8.