Identifying SNOMED Concepts Relevant to CHA_2DS_2–VASc and HAS-BLED Scores

Peter L. Elkin, MD; Edwin Anand, MD; Chris Crowner, MS; Sina Erfani, BS; Grégoire Ficheur, PhD MD; Daniel R. Schlegel, PhD

University at Buffalo

Introduction

The CHA_2DS_2–VASc[1] and HAS-BLED[2] scores are used to assess the risk of stroke in patients with atrial fibrillation, and risk of major bleeding due to oral anticoagulants used to treat atrial fibrillation, respectively. In this study we identify a list of SNOMED concepts which are relevant to the generation of these two scores.

Eventual goal: build a reasoning model to calculate these scores automatically.

Methods

SCALES

CHA_2DS_2–VASc Scoring Criteria	
Age	+1 (65-74), +2 (75+)
Sex	+1 (Female)
History of Congestive Heart Failure	+1
History of Hypertension	+1
History of Diabetes Mellitus	+1
History of Stroke	+2
History of Vascular Disease	+1

HAS-BLED Scoring Criteria	
Age	+1 (65+)
History of Hypertension	+1
Renal Disease	+1
Liver Disease	+1
History of Stroke	+1
Prior Major Bleeding / Predisposition to Bleeding	+1
Labile INR	+1
Medication Use Predisposing to Bleeding	+1
History of Alcohol Use	+1

CONCEPT SELECTION

Two clinicians (GF, EA) familiar with the scales and SNOMED:

1. examined all SNOMED concepts related to CHA_2DS_2–VASc and HAS-BLED using the IHTSDO SNOMED CT Browser;
2. examined all descendants of relevant concepts to ensure their relevance;
3. formed a consensus with the help of an adjudicator (PLE)

This study is sponsored by Pfizer

STATISTICAL ANALYSIS

- The contingency tables between the two clinicians (GF, EA) were built at both the *top level*, selected by clinicians, and *including descendants*.
- The proportions of common concepts were computed for each of these two tables
- A Cohen's κ score could not be calculated because the (potentially very large) set of SNOMED concepts which both clinicians found to be irrelevant is unknown, and an estimate could yield an artificially high κ score.

Results

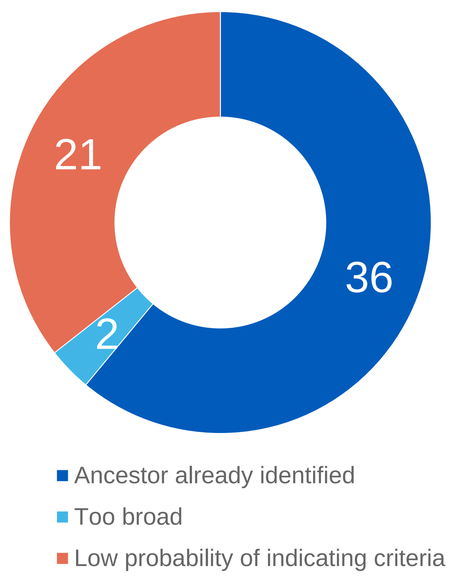
Top Level		EA +	EA -
GF	+	13	36
	-	57	

With Descendants		EA +	EA -
GF	+	5,321	42
	-	5,862	

Figure 1. Similarities and differences in the generation of the consensus (left), and also considering descendants (right).

- 106 unique concepts identified
- 13 (12%) of were common to both clinicians.
- Including the descendants, the intersection was just over 47%.
  - We believe this is a more useful measure of the intersection, since including a parent and child concept deserves much less of a penalty than including entirely incorrect concepts.
- GF selected 36 concepts which EA did not
  - including descendants only added 42 concepts to the total, while the 57 which EA selected that GF did not added 5,862.
- The final consensus included 47 concepts, indicating a rather large amount of culling during arbitration was performed.

Reasons for Exclusion of Top Level Concepts



Discussion

There were 59 total exclusions. Decisions about the correctness of concept selections were made along the following lines:

- ancestor concept already having been identified
  - e.g., *congestive heart failure* is a child of *heart failure* in SNOMED CT
- low probability of indicating the criteria
  - *Endocrinologist* for the renal disease category
  - *Respiratory crackles* for the congestive heart failure category
- The remaining concepts were too broad and required clinical decisions about whether to include them or not
  - e.g., include *Chronic liver disease* instead of *Disorder of liver*, since scales are concerned with long-term effects, rather than short-lived diseases - of which there are many under *Disorder of liver*.

Conclusion

- Agreement among SNOMED CT concepts identification remains sub optimal in our case.
- Previous studies have compared professional coders to study SNOMED CT, in which the inter rater agreement was only 33%. A study using UMLS found a similarly low inter-rater reliability.[3,4]
- Standardized NLP algorithms aiding physicians' coding for disease conditions would optimize coding and make such data more reliable.

References

1. Lip, G.Y., R. Nieuwlaat, R. Pisters, D.A. Lane, and H.J. Crijns, *Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation.* Chest Journal, 2010. **137**(2): p. 263-272.
2. Pisters, R., D.A. Lane, R. Nieuwlaat, C.B. de Vos, H.J. Crijns, and G.Y. Lip, *A novel user-friendly score (HAS-BLED) to assess 1-year risk of major bleeding in patients with atrial fibrillation: the Euro Heart Survey.* Chest Journal, 2010. **138**(5): p. 1093-1100.
3. http://browser.ihtsdotools.org/
4. Andrews, J.E., R.L. Richesson, and J. Krischer, Variation of SNOMED CT coding of clinical research concepts among coding experts. J Am Med Inform Assoc, 2007. 14(4): p. 497-506.
5. Rothschild, A.S., H.P. Lehmann, and G. Hripcsak, Inter-rater agreement in physician-coded problem lists. AMIA Annu Symp Proc, 2005: p. 644-8.