

Identifying SNOMED Concepts Relevant to CHA₂DS₂-VASc and HAS-BLED Scores¹

Peter L. Elkin, MD, Edwin Anand, MD, Chris Crowner, MS, Sina Erfani, BS,
Grégoire Ficheur, PhD, MD, and Daniel R. Schlegel, PhD
Department of Biomedical Informatics, University at Buffalo, Buffalo, NY, USA

Introduction

The CHA₂DS₂-VASc¹ and HAS-BLED² 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. Our eventual goal is to build a reasoning model to calculate these scores automatically. In this study we identify a list of SNOMED concepts which are relevant to the generation of these two scores.

Methods

Two clinicians, EA and GF, familiar with CHA₂DS₂-VASc and HAS-BLED used the IHTSDO SNOMED CT Browser to find concepts relevant to the clinical use of those scales. The clinicians examined all SNOMED concepts which were found when searching for terms the clinicians thought were relevant to each criteria in the two scales. They examined all descendants of relevant concepts to ensure they are also relevant, and noted if they were not. A consensus² was formed with the help of an adjudicator (PLE). 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. Instead we report the further three variables in the contingency table. We computed these three values both at the top level (the one selected by the clinicians) and including descendants.

Results

Of the 106 concepts identified by the clinicians, only 13 (12%) of them were common to both. Including the descendants the intersection was just over 47%. The variability in depth of the tree under selected terms can be seen in Figure 1 – GF selected 36 concepts which EA did not, but 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.

Top Level	EA		With Descendants	EA			
	+	-		+	-		
GF	+	13	36	GF	+	5,321	42
	-	57			-	5,862	

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

Discussion

The overwhelmingly majority of issues resolved during arbitration was the removal of concepts for which an ancestor had already been identified. This accounts for 36 of the 59 concepts excluded during arbitration. In other cases a category was chosen which was too broad – this happened twice and accounts for many of the descendants EA captured but GF did not. Other issues resolved during arbitration had to do with whether to include items which have low probability of indicating the criteria, e.g., *Endocrinologist* for renal disease or *Respiratory crackles* for congestive heart failure. It was decided that these should be excluded. Finally, some clinical decisions had to be made. For example, it was decided during arbitration to include *Chronic liver disease* instead of *Disorder of liver*, which one clinician had identified, since the scales are about longer term effects, rather than short-lived diseases.

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.

¹ This study is sponsored by Pfizer.

² The final concept list is available on Google Drive at <http://bit.ly/1U4kmpc>