Ontology for Heart Rate Turbulence Domain Applying the Conceptual Model of SNOMED-CT

Cristina Soguero-Ruiz[1], Luis Lechuga[2], Inmaculada Mora-Jiménez[1], Javier Ramos-López[1]
[1] Universidad Rey Juan Carlos, Fuenlabrada, Spain
[2] Hospital Universitario de Fuenlabrada, Fuenlabrada, Spain

***Abstract*—Although cardiovascular risk stratification (CVRS) based on ECG-derived indices has been deeply studied, many current findings are not being widely used in the clinical practice. We hypothesized that, in addition to the necessary scientific evidence, also a clear and standardized connection among the current knowledge in the scientific literature, its availability for the cardiologist, and the actual patient data, is necessary for the practical implementation and refinement of these indices.**

For this purpose, we implemented an standardized framework for CVRS based on ECG-derived indices, focused on the actual knowledge of Heart Rate Turbulence (HRT) indices (with concise guidelines and clear procedures to parameter calculations). An ontology for HRT was built according to a set of logical and relational rules, yielding the class hierarchy model and its corresponding inferred model (Protégé-OWL, 4.1) for completeness. Different from other biomedical ontologies, ours was based on the international standard SNOMED-CT. The model of SNOMED-CT not only considers terminology, but also properties and relationships, what guaranteed the standardization and compatibility with current and emerging Electronic Health Records. Our HRT ontology consisted of 308 concepts (289 from SNOMED-CT, and 19 a local extension to model the main concepts of the HRT-domain). As an application example, a database of 27 instances of patients with HRT from 24-Holter monitoring recordings was considered, with basic HRT indices and also conventional and emergent signal processing calculations. A consistence of 86% and 77% was achieved between averaged procedure for HRT index calculations given in the guidelines and with a filtering procedure.

***Index Terms*—Cardiovascular Risk Stratification, Heart Rate Turbulence, ontology and SNOMED-CT.**

I. Introduction

CARDIOVASCULAR Risk Stratification (CVRS) based on indices derived from electrocardiogram (ECG) has received wide attention in the clinical, scientific and technical literature [1], however, many current findings in this area are not being currently used in the clinical practice. Maybe not only the required scientific evidence for some findings to be actually established into routine cardiac medicine is necessary, but also a clear and standardized connection among the current scientific knowledge, its availability for the cardiologist, and the actual patient data bases, are necessary for the clinical routine use and refinement of these indices.

Recent developments in Semantic Web applications have made the ontologies a key tool for logical and consistent reasoning and for domain description support. Technical ontology has been defined [2] as a *formal, explicit specification of shared conceptualization*. An ontology typically consists of a finite set of terms (classes of objects of the domain), relationships among such terms, properties of relationships, value restrictions, disjointness between terms, and specification of logical relationships between objects [3]. Biomedical domain ontologies have been mostly applied to date for semantically driven information integration, by using different biomedical terminologies, modeling, and reasoning about complex entities [4]. An ECG ontology was first proposed in [5], and it was subsequently extended for inference capabilities given by OWL DL and SWRL [6]. An ECG-based decision making system was proposed in [7] using an ontological model for the ECG and for the heart rate variability measurements. In [8], a framework was proposed for predicting types of arrhythmia from ECG signal in a body sensor network node. Previous work on ECG ontologies has aimed to give a principled approach to this knowledge domain, but still their standardization with current and emerging Electronic Health Recordings (EHR) is not warrantied. However, the Systematized Nomenclature of Medicine - Clinical Terms (SNOMED-CT) is becoming the standard for semantic interoperability both in EHR and in Hospital Information Systems (HIS), and it is the most comprehensive, multilingual clinical healthcare terminology worldwide [9].

Accordingly, we addressed an standardized framework for CVRS based on ECG-derived indices, by making an ontology completely based on SNOMED-CT conceptual model, hence yielding advanced interoperability capabilities. Our first focus was the cardiac indices obtained from HRT, for which concise guidelines and clear procedures for parameter calculations can be found [1].

The paper is organized as follows. Section 2 summarizes the state of art about ontologies, SNOMED-CT conceptual model, and HRT fundamentals. Section 3 presents our proposed HRT ontology for CVRS. A simple case study on a patient data base of ECG 24h-Holter is given in Section 4. Conclusions are finally stated in Section 5.

II. Background

Figure 1 shows a basic block diagram for the interoperability among the EHR, the inference information system, the shared ontology, and the data presentation to the clinical user. The role of the underlying ontology consists of: (a) yielding a logical and actual structure to the relationships among the clinical concepts; (b) providing with a standardized information architecture; and (c) allowing the logical and

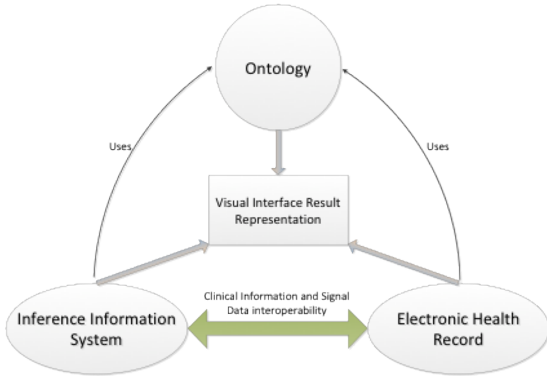

Fig. 1. *Schema of the elements for interoperability.*

conceptual inference to generate complete relationships among concepts from the current basic knowledge, and for detecting inconsistencies. A number of biomedical and cardiovascular ontologies have limited interoperability despite of their widespread use. In [10], a system was able to perform risk classification of patients using medical image analysis and semantically structured information of patient data, by means of MeSH thesaurus. Improved interoperability was achieved in [11] with an ontology for chronic patients built on the SNOMED-CT reference terminology.

SNOMED-CT Conceptual Model. Far from being just a terminology or thesaurus, the SNOMED-CT system is a multi-language linguistic tool. SNOMED-CT probably represents the most complete classification for clinical use, its main advantage being the reference to terminologies of most of different health professions, hence facilitating the communication within diversified health teams and in searching health information.

The concepts in SNOMED-CT are structured in 19 top-level hierarchies (axes). Although SNOMED-CT includes more than 311,000 concepts since 2011, this impressive number can be not enough for representing many clinical expert domains. Hence, local and national extensions can be created in SNOMED-CT, so that contents (such as subsets of concepts) may be delivered locally by a specific clinical expert group and can move to a national extension and to the international core. Most of current EHR and HIS are adopting SNOMED-CT as their standard for interoperability, and many authors agree that its conceptual model is in fact an ontology.

Basic Concepts on HRT. HRT is the phenomenon of short-term fluctuation in sinus cycle length over about 20 beats following a Ventricular Premature Complex (VPC) [1]. HRT is usually assessed from 24-h ECG signals. From such recordings, a VPC tachogram is constructed by aligning and averaging the R-R interval sequences around isolated VPCs, according to the guidelines. HRT is quantified by two parameters: turbulence onset (TO) and turbulence slope (TS). TO reflects the amount of sinus acceleration following a VPC. TS reflects the rate os sinus decelerations that follows sinus acceleration. In this study, these parameters are calculated from the tachogram averaged from all VPC tachograms on a 24-h ECG record. TO is the shortening of the two sinus beat RR interval average after the compensatory pause, expressed as a percentage of the two sinus beat RR interval average preceding the VPC.

III. HRT ONTOLOGY FOR CVRS

HRT Ontology. The HRT ontology defined in this study has been represented in the Ontology Web Language (OWL) [12]. According to the general methodology proposed in [3], the process of constructing the HRT ontology consisted of the following set of steps.

(1) Determine the domain. The domain was HRT.

(2) Enumerate the relevant concepts of the constructed ontology, to get a well-defined domain. First, a set of clinical, anatomical, electrophysiological, and pharmacological features, were identified in order to represent the HRT domain. This first set was mostly based on SNOMED-CT concepts (among more than 311,000 available concepts), however, a subset of concepts which were highly relevant for defining HRT could not be taken from SNOMED-CT. Accordingly, up to 19 over 300 concepts defined in our ontology had to be extended using the *Fuenlabrada Hospital* namespace. In Figure 4, an schema of the ontology is shown, where extended concepts are in a rectangle.

(3) Define the concepts and their hierarchy (down-top development process). After the concepts were enumerated, a hierarchy development was necessary. HRT concepts mapped directly in SNOMED-CT terminology were just included by following the SNOMED-CT hierarchy. For the remaining concepts, we first decided in which of 19 SNOMED-CT axes they were included, i.e., in a *body structure*, in a *clinical finding*, or in an *observable entity*. After, we assigned the concept to its parent class. The SNOMED CT July 2006 release was systematically examined using the CliniClue terminology browser [13] to determine whether the concepts defined are represented in the terminology. In Figure 2 we present an example of a defined concept, the ECG feature concept, its hierarchy, status and identificator. This identificator is used to complete the name of ECG feature concept, what we propose to be used in the future to standardize the ontology concepts so that they will be used in any HIS. In Figure 3, a screen snapshot of Protégé shows the ECG feature concept inside our HRT ontology hierarchy.

(4) Define the properties of the concepts. Data Properties define features associated with one concept, whereas Object Properties define a relationship between two concepts. SNOMED-CT defines some properties, but new properties can be generated for the most conceptual aspects of an specific domain, as shown in Table I for HRT.

(5) Define the non-hierarchical relationships (Table I).

(6) Use inference mechanism, in our case, to get a risk stratification criterion in patients with cardiac disease, which is described next.

Inference on CVRS with HRT. The HRT evaluation has been found useful for risk stratification after acute myocardial infarction and monitoring of disease progression in heart failure. HRT values are usually classified into: *Category 0,* when both TO and TS are normal; *Category 1,* when either TO or TS is abnormal; and *Category 2,* when both TO and TS

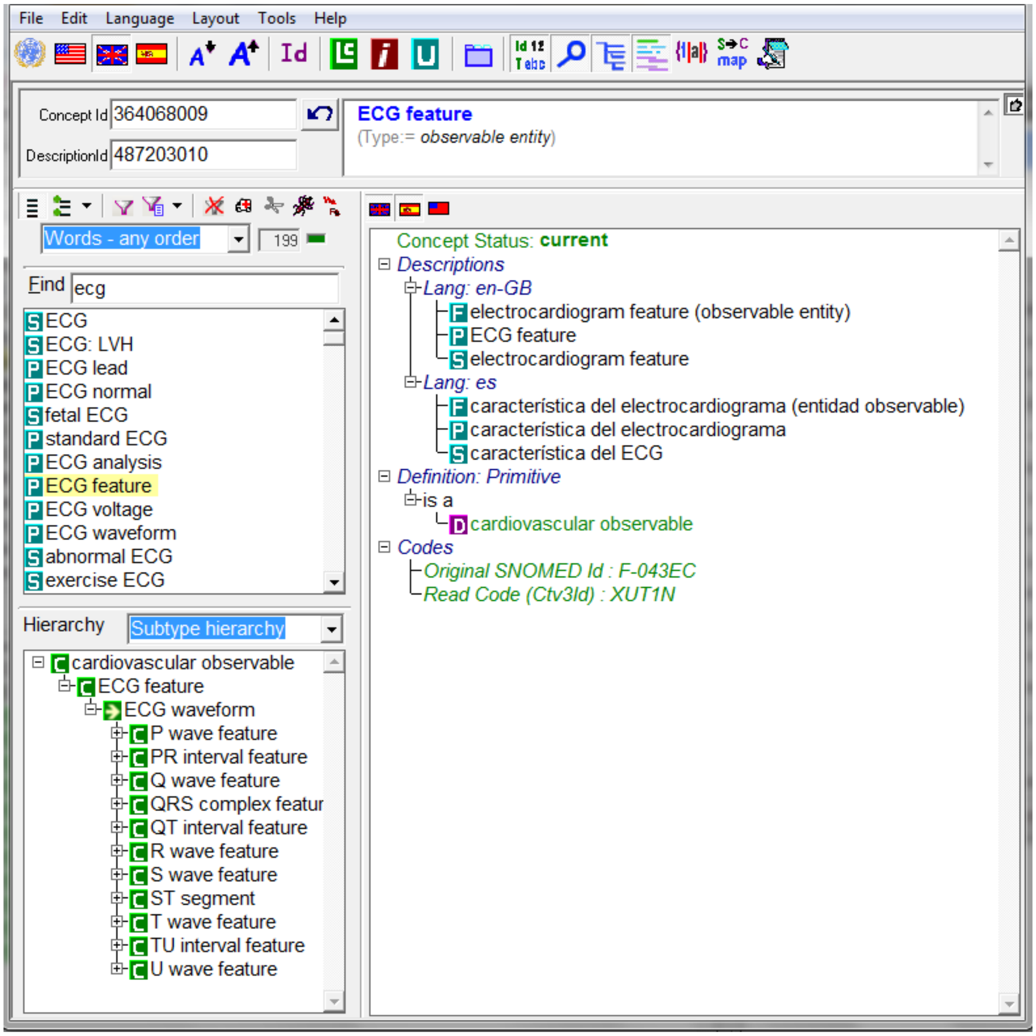

Fig. 2. *CliniClue Xplore screen snapshot selecting the ECG feature concept and showing its hierarchy, the concept status and id.*

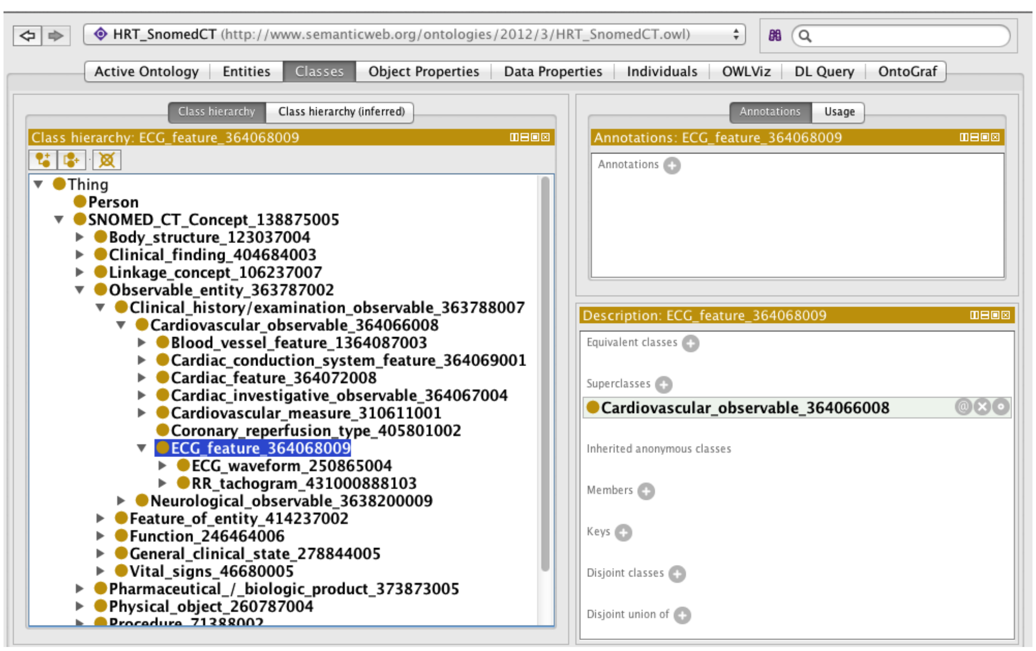

Fig. 3. *Protégé screen snapshot remarking one concept and its corresponding SNOMED-CT concept identificator.*

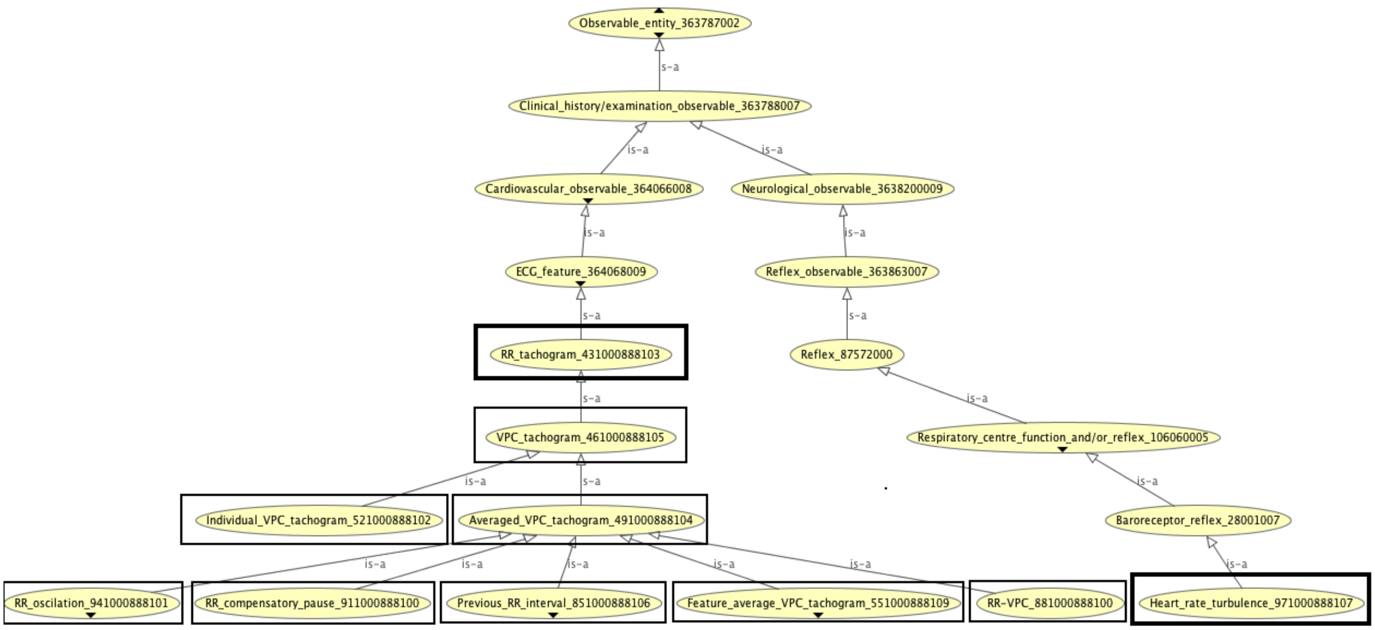

Fig. 4. *HRT ontology. Highlighted rectangles represent local extended concepts.*

TABLE I
SOME SNOMED-CT PROPERTIES IN HRT ONTOLOGY. LOCAL EXTENDED CONCEPTS ARE IN ITALIC.

Individual	Property	Individual
Normal heart rate_78663003	**Interprets_363714003**	Cardiac conducting system structure_24964005
Disorder of cardiac function_105981003	**Finding site_363698007**	Heart structure_80891009
Disorder of cardiac ventricle_415991003	**Finding site_363698007**	Cardiac ventricular structure_21814001
24 hour ECG_252417001	**Method_260686004**	Monitoring-action_360152008
24 hour ECG_252417001	**Using device_424226004**	Electrocardiographic monitor and recorder_74108008
HRT_971000888107	***is a measurement of_161000888104***	Baroreceptor_reflex_28001007
TO_641000888108	***is influenced by_191000888105***	Left_ventricular_ejection_fraction_250908004
TS_731000888107	***is influenced by_191000888105***	Left_ventricular_ejection_fraction_250908004
TS_731000888107	***is calculated in_101000888100***	*Averaged tachogram_491000888104*

are abnormal. Up to four studies have reported that mean TO ranged from -2.7% to -2.3% and mean TS ranged from 11.0 to 19.2 ms/RR in healthy subjects [1]. By considering these limits, we created the instances of classes TO and TS. OWL can deal with classification problems using Description Logic, but many applications for risk stratification (as it was our case) need augmented OWL with Semantic Web Rule Language (SWRL) [14] (see Table II for several examples).

Some authors claimed that the averaging process in the VPC-tachogram calculation could be masking physiological effects, in addition to the expected denoising [15]. Accordingly, several advanced signal processing procedures have been proposed for denoising individual (rather than averaged) VPC-tachograms, including Support Vector Machine (SVM) nonlinear regression. These methods have been proven to be effective for denoising from a technical and physiological point of view, however, their impact on the risk stratification capabilities has not been benchmarked with the averaged tachogram signal processing.

As a case study, we used our HRT ontology to compare the output of the HIS both for the conventional and emergent signal processing methods for calculating HRT indices. We used a database of 24h Holter recordings from Congestive Heart Failure (CHF) patients, collected in the Arrhythmia Unit of Hospital Universitario Virgen de la Arrixaca (Spain). RR-tachograms were analyzed to identify reliable isolate VPC tachograms, according to the criteria proposed in [1]. Up to 27 of 60 recordings were useful for this analysis (14 recordings with only one isolated VPC tachogram). Table III shows that conventional averaged calculated indices yielded 0, 5 and 22 patients on Category 0, 1, and 2, respectively, whereas the single-VPC procedure based on SVM yielded 0, 4, and 23 patients, hence showing that where will be differences in the inference depending on the underlying signal processing calculation algorithm. For comparison purposes, the patient classification according to individual VPC without averaging and without denoising, was 0, 4, and 20, where three patients could not be classified by majority.

IV. CONCLUSION

An ontology for CVRS has been proposed on the HRT domain. Interoperability with EHR and HIS has been reinforced by using the SNOMED-CT conceptual model, and extension

TABLE II
SWRL RULES.

Category 0:
$TO(p?) \wedge swrlb : greaterThan(?p, -2.7)$
$TO(p?) \wedge swrlb : greaterThan(?p, -2.7)$ $\wedge\, TO(p?) \wedge swrlb : lessThan(?p, -2.3)$ $\wedge\, TS(p?) \wedge swrlb : greaterThan(?p, 11.0)$ $\wedge\, TS(p?) \wedge swrlb : lessThan(?p, 19.0) \Rightarrow$ $swrlb : equal(Category(?p, 0))$
Category 1:
$TO(p?) \wedge swrlb : greaterThan(?p, -2.7)$ $\wedge\, TO(p?) \wedge swrlb : lessThan(?p, -2.3)$ $\wedge\, TS(p?) \wedge swrlb : lessThan(?p, 11.0) \Rightarrow$ $swrlb : equal(Category(?p, 1))$
$TO(p?) \wedge swrlb : greaterThan(?p, -2.7)$ $\wedge\, TO(p?) \wedge swrlb : lessThan(?p, -2.3)$ $\wedge\, TS(p?) \wedge swrlb : greaterThan(?p, 19.2) \Rightarrow$ $swrlb : equal(Category(?p, 1))$
$TS(p?) \wedge swrlb : greaterThan(?p, 11.0)$ $\wedge\, TS(p?) \wedge swrlb : lessThan(?p, 19.0)$ $\wedge\, TO(p?) \wedge swrlb : lessThan(?p, -2.7) \Rightarrow$ $swrlb : equal(Category(?p, 1))$
$TS(p?) \wedge swrlb : greaterThan(?p, 11.0)$ $\wedge\, TS(p?) \wedge swrlb : lessThan(?p, 19.0)$ $\wedge\, TO(p?) \wedge swrlb : greaterThan(?p, -2.3) \Rightarrow$ $swrlb : equal(Category(?p, 1))$
Category 2:
$TO(p?) \wedge swrlb : lessThan(?p, -2.7)$ $\wedge\, TS(p?) \wedge swrlb : lessThan(?p, 11) \Rightarrow$ $swrlb : equal(Category(?p, 2))$
$TO(p?) \wedge swrlb : lessThan(?p, -2.7)$ $\wedge\, TS(p?) \wedge swrlb : greaterThan(?p, 19.2) \Rightarrow$ $swrlb : equal(Category(?p, 2))$
$TO(p?) \wedge swrlb : greaterThan(?p, -2.3)$ $\wedge\, TS(p?) \wedge swrlb : lessThan(?p, 11.0) \Rightarrow$ $swrlb : equal(Category(?p, 2))$
$TO(p?) \wedge swrlb : greaterThan(?p, -2.3)$ $\wedge\, TS(p?) \wedge swrlb : greaterThan(?p, 19.2) \Rightarrow$ $swrlb : equal(Category(?p, 2))$

TABLE III
CARDIOVASCULAR RISK STRATIFICATION RESULTS.

	Avg. tachogram			**Individual tachogram**			**Ind. tachogram SVM**			
Patient Number	C0	C1	C2	C0(%)	C1(%)	C2(%)	C0(%)	C1(%)	C2(%)	Total no. Ind. tachograms
1	-	-	x	0.0	0.0	**100.0**	0.0	0.0	**100.0**	37
2	-	-	x	0.0	0.0	**100.0**	0.0	0.0	**100.0**	1
3	-	-	x	0.0	40.0	**60.0**	0.0	10.0	**90.0**	40
4	-	-	x	0.0	17.4	**82.6**	0.0	16.3	**83.7**	86
5	-	-	x	0.0	0.0	**100.0**	0.0	**100.0**	0.0	1
6	-	-	x	0.0	0.0	**100.0**	0.0	0.0	**100.0**	1
7	-	x	-	0.0	11.0	**88.9**	0.0	22.2	**77.8**	9
8	-	x	-	0.0	**100.0**	0.0	0.0	**100.0**	0.0	1
9	-	-	x	0.0	0.0	**100.0**	0.0	0.0	**100.0**	1
10	-	-	x	0.0	0.0	**100.0**	0.0	0.0	**100.0**	2
11	-	-	x	0.0	13.0	**87.0**	0.0	13.0	**87.0**	23
12	-	-	x	0.0	0.0	**100.0**	0.0	**100.0**	0.0	1
13	-	-	x	0.0	0.0	**100.0**	0.0	0.0	**100.0**	1
14	-	-	x	0.0	**50.0**	**50.0**	0.0	0.0	**100.0**	2
15	-	-	x	0.0	0.0	**100.0**	0.0	0.0	**100.0**	1
16	-	-	x	1.8	35.7	**62.5**	3.6	28.6	**67.9**	56
17	-	x	-	0.0	**100.0**	0.0	0.0	0.0	**100.0**	1
18	-	x	-	0.0	**100.0**	0.0	0.0	0.0	**100.0**	1
19	-	-	x	0.0	0.0	**100.0**	0.0	0.0	**100.0**	1
20	-	-	x	0.0	0.0	**100.0**	0.0	0.0	**100.0**	2
21	-	-	x	0.0	0.0	**100.0**	0.0	0.0	**100.0**	4
22	-	-	x	0.0	0.0	**100.0**	0.0	0.0	**100.0**	1
23	-	-	x	**33.3**	**33.3**	**33.3**	0.0	**66.7**	33.3	3
24	-	x	-	0.0	**100.0**	0.0	0.0	0.0	**100.0**	1
25	-	-	x	0.0	0.0	**100.0**	0.0	0.0	**100.0**	1
26	-	-	x	0.0	**50.0**	**50.0**	0.0	0.0	**100.0**	2
27	-	-	x	1.4	11.6	**87.0**	0.0	2.9	**97.1**	69

of the concepts and rules has been achieved. A simple application of risk categorization has been used for straightforward comparison of conventional and emergent criteria, and its impact in the classification of a patient data base.

Oncoming work is devoted to extend the methodology to other relevant ECG-derived indices for CVRS, and to include new properties (for instance, effect of drugs, effect of age, and others). An extension of this work is in preparation to be submitted to the technical journal *IEEE Transactions on Information Technology in Biomedicine*, with an index factor of 1.960 (second quartile) in the category Computer Science, Information Systems (Scopus).

ACKNOWLEDGEMENTS

The author wish to thank Óscar Barquero-Pérez, Arcadi García-Alberola and José Luis Rojo-Álvarez for suggestions, comments, help with experiments, and data. A reduced version of this work was submitted to Rosanna Degani Young Investigator Award competition in the international conference *Computing in Cardiology 2012*. This work has been partly supported by Research Projects TEC2009-12098, TEC2010-19263 and TSI-020100-2009-332 from Spanish Government. Additionally, it has also been partially funded by the Ministerio de Sanidad y Política Social (Spanish Government, 2010 strategy - CISNS of date 02/10/2010), as support to the implementation of Information Systems.

REFERENCES

[1] A. Bauer, M. Malik, G. Schmidt, P. Barthel, Bonnemeier, and et. al, "Heart rate turbulence: Standards of measurement, physiological interpretation, and clinical use: (ishne consensus)," *J Am Coll Cardiol*, vol. 52, no. 17, pp. 1353–1365, 2008.

[2] T. R. Gruber, "A translation approach to portable ontology specifications," *Knowledge Acquisition*, vol. 5, no. 2, pp. 199–220, 1993.

[3] N. F. Noy and D. L. McGuinness, "Ontology development 101: A guide to creating your first ontology," 2001.

[4] A. Burgun, "Desiderata for domain reference ontologies in biomedicine." *J. of Biom. Informatics*, vol. 39, no. 3, pp. 307–313, 2006.

[5] B. Gonçalves, G. Guizzardi, and J. G. P. Filho, "An electrocardiogram (ecg) domain ontology," in *Workshop on Ontologies and Metamodels for Software and Data Engineering*, 2007, pp. 68–81.

[6] B. Gonçalves, V. Zamborlini, and et. al, "Using a lightweight ontology of heart electrophysiology in an interactive web application," in *Proc. XIV Brazilian Symposium on Multimedia and the Web*, 2008, pp. 77–80.

[7] G. Acampora, C.-S. Lee, A. Vitiello, and M.-H. Wang, *Evaluating cardiac health through semantic soft computing techniques*. Springer Berlin / Heidelberg, 2011.

[8] T. Tanantong, E. Nantajeewarawat, and S. Thiemjarus, "Towards continuous electrocardiogram monitoring based on rules and ontologies," in *Bioinformatics and Bioengineering, IEEE 11th Int Conf on*, oct. 2011, pp. 327 –330.

[9] (2008) Snomed-ct. [Online]. Available: http://www.ihtsdo.org

[10] C. Doulaverakis, M. Papadogiorgaki, A. B. Vasileios Mezaris, and et. al, "Ivus image processing and semantic analysis for cardiovascular diseases risk prediction," *Int.l J. Biomedical Engineering and Technology*, vol. 3, pp. 349–374, 2010.

[11] T. Sampalli, M. Shepherd, and J. Duffy, "A patient profile ontology in the heterogeneous domain of complex and chronic health conditions," in *Int Conf on System Sciences*, 2011, pp. 1–10.

[12] G. Antoniou and F. van Harmelen, "Web ontology language: Owl," in *Handbook on Ontologies in Information Systems*. Springer-Verlag, 2003, pp. 76–92.

[13] "Cliniclue 2006. clinical terminology services from the clinical information consultancy. available at www.cliniclue.com."

[14] I. Horrocks, P. F. Patel-Schneider, and et. al, "SWRL. A Semantic Web Rule Language Combining OWL and RuleML," 2004.

[15] J. L. Rojo-Álvarez, O. Barquero-Pérez, I. Mora-Jiménez, and et. al, "Heart rate turbulence denoising using support vector machines," *IEEE Trans Biomedical Engineering*, vol. 56, no. 3, pp. 310–319, 2006.