Modeling the Clinical Observation Space of Electronic Medical Record in Primary Care

Elmehdi Aitnouri
GPEMA R&D department
GPEMA Canada
Sherbrooke, Québec, Canada
elmehdi.aitnouri@gpema.com

Ali Abbassene
Networking and Information Systems Team
Centre de Développement des Technologies Avancées
Baba Hassen, Algiers, Algeria
a_abbassene@cdta.dz

***Abstract*—We present a new model for structuring clinical data in primary care, denoted by the observation space. This model is intended to be used within Electronic Medical Record (EMR), and is designed to meet the requirements of clinical contexts and nuances that characterize primary care practices. While these contexts are generally captured with free-text descriptions, structuring them makes the use of EMR very attractive, since it opens many possibilities such as clinical data exchange and the design of Clinical Decision Support Systems. In partnership with practitioners, the results presented in this paper are being used to build clinical patterns.**

Keywords-Electronic medical record; clinical data modeling;

I. INTRODUCTION

Although the Electronic Medical Record (EMR) is clearly identified as one of the tool that may help manage the many primary care challenges such as population aging, its use is far to be generalised [1]. Many authors noted that the use of EMR is not systematic since there is no clear consensus on how clinical aspects of EMR should be processed [2]. While practitioners see that coding clinical data will limit necessary nuances required for describing the clinical context and argue to use free-text data for this purpose, managers see that coding clinical data is important for information gathering, searching and comparing as well as the possibility to use upstream Clinical Decision Support Systems (CDSS) [3].

We have initiated a research project in order to design an intelligent clinical module to deal with clinical aspects in primary care. Such a module may be part of any EMR, and will have the ability to address the needs of practitioners in their diagnosis process. We have started with the needs to diagnose and monitor chronic diseases. The project has four phases as follow: Phase1 deals with the capture of clinical data in primary care. Phase 2 deals with the design of aggregation and fusion of that clinical data in order to extract patterns. Phase 3 deals with the build of the reasoning capability. Finally, phase 4 deals with medical knowledge management. This paper presents some results obtained in phase 1 of that project, and dealing with the structure of the observation space.

In primary care, clinical information, denoted in this article by Primary Care Clinical Data (PCCD), may be classified in nine categories, namely symptoms, signs, review of systems, medications, labs, personal history, family history, style of life and demographic data [4]. In their study of the applicability of medical data standards in clinical research, Richesson and Krischer [5] concluded that there is still much work to be done since many gaps have been identified, especially in the area of physical exam: "There is a conspicuous lack of medical standards for the structuring of questions and case report form, particularly in the areas of physical exam, medical history, family history, and eligibility criteria". Interestingly, the major gaps concern directly structuring and standardizing PCCDs.

The present work investigates on how to structure some of the PCCD, especially those describing the clinical observation. In addition to the first classification of the PCCD in nine classes, it is possible to consider a further categorization following the approach recommended by Scheuermann [6] for treating diseases and diagnosis within an ontological point of view. The approach suggests that diseases may be considered as predispositions rooted in physical disorders in the organism and realized in pathological process. This conceptualization of diseases helps to divide PCCD in two main categories. On one hand, the observation space groups PCCD that describe the manifestation of the disease phenomenon, namely symptoms, signs and review of the systems. On the other hand, the predisposition space groups the remaining PCCD and describes pre-clinical and pre-conditions that make the disease follows a given course. This article will focus on the structuring of the observation space and propose a global model of it.

This paper is organized as follow. Section 2 introduces the PCCD's and will focus on the description of the observation space. Section 3 discusses how it is possible to structure it in order to capture clinical nuances and contexts. Section 4 presents results of the application of the observation space model to a real clinical case.

II. PRIMARY CARE CLINICAL DATA (PCCD)

PCCDs are categorized following their functional characteristics, independently of patient's health and disease diagnosis. It is possible to further categorize PCCDs in order to understand the clinical data structure in primary care. For this purpose, it is important to return to the pathophysiological basis of diseases [7]. Contemporary classification of human disease derives from observational correlation between pathological analysis and clinical syndromes, which makes diseases primarily described by symptoms and signs they cause.

 ISBN: 978-1-61208-252-3

However, diseases are dispositions rooted in physical disorders in the organism and realized in pathological processes. The disorders are active and the organism as a complex system will use all required metabolic fluxes to recover the homeostasis. If a deregulation exceeds a given threshold or organisms fail to recover the homeostasis, a pathological process triggers and a disease course, which may vary widely between patients who have the same disease, starts [6].

A. Observation Space

The observation space is the clinical picture of a patient on a given point in time of a disease evolution.

The *observation space* contains the PCCDs symptoms, signs, and review of systems. It describes the condition of the patient's organism due to the effect of the disease. From this space, a practitioner should answer the question, *what is the clinical condition of the patient given these observations?* Theoretically, to construct the *observation space*, physicians are not required to know the diagnosis or the disease of the patient.

B. Predisposition space

The predisposition space represents all preclinical conditions of a patient that may variably influence the evolution of a disease course.

The predisposition space contains the PCCDs labs, demographic data, past medical history, present medical condition, family history, present medication, and style of life. It represents the preclinical conditions of a patient independently of any actual pathological process. The predispositions will condition any disease course. From this space, the practitioner should answer the question: *Given the information gathered from the predisposition space, what should be the clinical picture of the patient for a given pathology?* Many clinical researches concentrate on discovering relations between some predisposition space PCCD and disease severities.

This paper focuses on the structuring of the *observation space*.

III. Structuring Of The Observation Space

Due to the presence of symptoms, the *observation space* is subjective and contextual. It is, however, this space who introduces the major clinical nuances which require free-text description. The present work will focus on describing how data and information that constitute this space may be structured and modeled in taking into account appropriate medical knowledge and practices.

A. Symptom model

Symptoms are the kernel of the *observation space*. A symptom is a subjective evidence of a disease perceived by the patient. Psychological aspects have seldom evident effects on how patients perceive and report what they feel. It has been reported that patients that categorize themselves with a group of people that have a given illness will perceive concurrent symptoms relevant to that illness to be more severe [8]. This is a clear evidence of the subjectivity character of symptoms. They are also the main input of CDSS due to the fact that knowledge on diseases is described in terms of symptoms. Major CDSS, such as rule-based, model disease knowledge as causal relevance between symptoms and diseases, and symptoms are generally described by only their intensities [9]. Innocent [10], however, proposed a way to add temporal aspects to model the causal relevance between symptoms and diseases, and proposed rules such as: *Influenza always causes symptom fever in day one to day 3*.

Providing only few attributes to describe symptoms is far to be enough for practitioners who want to capture different clinical nuances and contexts. In a specialized literature, symptoms are described by many attributes.

Here are two examples of symptoms that need many attributes to be described:

"*Complains of intermittent severe pain in lower abdomen since the last three weeks".*

"*Severe pain in the upper abdomen for five days. Burning in nature especially occurs at night in bed".*

It is possible to categorize the clinical information of each symptom as illustrated in the following table:

TABLE I. Clinical Information Categories

Descriptors	Description1	Description2
Onset		-
Duration	3 weeks	5 days
Intensity	Severe	Severe
Location	Lower abdomen	Upper abdomen
Quality	-	Burning
Periodicity	Intermittent	-
Factors	-	Night in bed

We propose, in this work, to define symptoms with a set of attributes as follows:

Standard Symptom Qualifiers (SSQ) is defined to be a set of attributes that describe a symptom SSQ = {onset, duration, periodicity, intensity, quality, site, factors}.

The actual symptom model does not take into account influences other than those of the observed disease, such as associated symptoms, previous episodes, etc.

B. Sign model

A sign is an objective evidence of a disease perceptible to the physical examiner. Signs are findings resulting from the physical examination of the patient. This definition compared to that of symptoms doesn't allow to totally discriminating between these two concepts. The borders are fuzzy and there is an overlap. However, in order to understand the *observation space*, it is important to examine how signs information is present and what are their relations with symptoms. In the literature, signs are not always defined as a medical entity that may be analyzed independently of diseases. In this work, signs are classified in two main categories, which we may call respectively signs and symptom-signs.

A symptom-sign is a sign that is associated with a symptom and will fuse with it once evaluated.

 ISBN: 978-1-61208-252-3

The fever symptom is an example. A patient may report a fever as symptom, and the physician will measure objectively the value of the fever. We say that the value of the temperature (fever sign) will fuse with the fever symptom, since the fever symptom will remain, and the value of its intensity will be changed by the taken temperature.

Definition: A sign is not related to a symptom.

Blood pressure is a good example of a sign. Signs by this definition are different than vital signs. Signs have generally an explicit and direct link to pathologies. They constitute for this purpose an objective description of the *observation space*.

In this work, we also propose that signs are characterised by the SSQ model. Thus, they are identified by their onset, duration, periodicity, intensity, quality, site and factors. In doing so, fusing signs with symptoms will be very obvious. The degree of objectivity that a sign may bring to a symptom will be discussed in another article.

C. Review of Systems (ROS)

The review of systems is a head-to-toe survey to screen for additional symptoms not related to the patient's main complaint. The information obtained from a ROS has its own weight to the final diagnosis. A typical ROS is composed of a set of symptoms classified by body systems or physiological functions and are symptoms that have no SSQ.

D. Observation space model

Figure 1 shows a general structure of the *observation space* model.

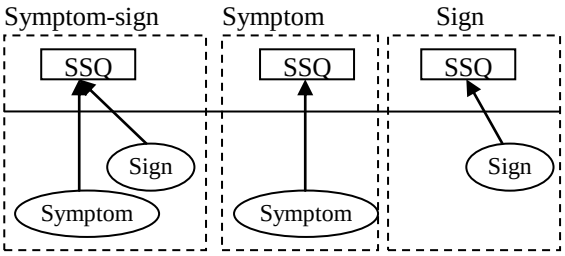

Figure 1. *Observation space* general structure

There are three main structures that form the *observation space*; Symptoms that do not possess signs, symptoms that possess signs (symptom-signs), and signs. Each structure has its SSQ.

IV. RESULTS

This section presents some results on the application of the *observation space* model to a real clinical situation. It will be used a clinical story of a 38 years old man with fatigue as chief complaint.

A. A clinical story example

This clinical sample story was extracted from [11].

1) Patient story

"*Mr. X, a 38-year-old homeless presented to a primary care clinic with severe, progressive fatigue of one month's duration; progressive weakness and shortness of breath while engaging in any kind of physical exertion and dizziness whenever he tried to get out of bed. Mr. X reported that he had not felt like eating for a month; whenever he ate something, he experienced severe abdominal pain. It usually occurred during the evening hours and was always triggered by eating.*

The patient had lost 20 pounds during the previous month, which he attributed to poor intake due to abdominal pain, anorexia, nausea, and fever. He reported having a mild, intermittent fever during the last month, particularly during the evening, which he attributed to fatigue".

2) Physical examination

Looking tired; Signs of pallor and jaundice. Brown coated tongue with no atrophy or inflammation.
Temperature: 100.4°F; blood pressure: 140/90
Cardiovascular system: bounding pulse, evidence of 3rd heart sound, no murmur.

Gastrointestinal system: abdomen: lower border is 1 inch down the costal margin; no liver tenderness. The spleen is 2 inches below the costal margin.

B. Discussion

The following table captures the clinical case described above using the *observation space* model. The observed symptoms and signs are listed in the rows. The SSQ attributes are listed in the columns. The ***Type*** columns in the table indicates which of the three types (symptom: Sym; Sign: Sig; Symptom-Sign: SS) the observation belongs to. The last column denoted symptom-sign gives the correspondent symptom or sign for a given symptom-sign observation. Indeed, in this example there are three symptom-signs, namely (fatigue-tired), (Pain-Abdomen) and (Fever, Jaundice).

The value of each SSQ attribute for symptoms and signs in the table are not dealt with in this article. For symptoms, the values such as described in the patient story are used. For signs, a +1 value in the intensity means that there is evidence that the sign is observed. Moreover, how symptom-sign SSQ values may be merged is not dealt with in this article.

 ISBN: 978-1-61208-252-3

TABLE II. CLINICAL DATA CAPTURED USING THE OBSERVATION SPACE MODEL

SSQ	Onset	Duration	Period.	Intensity	Quality	Site	Factors	Type	Symptom -sign
Symptom									
Fatigue		1 month		Severe	progressive			Symp	Tired
Weakness				Severe	progressive			Symp	
Shortness of breath				severe			While engaging in physical exertion	Symp	X
Dizziness							Get out of bed	Symp	X
Pain	When eating	1 month	Evening hours	severe		abdomen		Symp	Abdomen
Weight lost		Previous months		20 pounds				Symp	X
Fever		Last months		mild	intermittent		During evenings	Symp	Jaundice
Signs									
Tired				+1				SS	Fatigue
Pallor				+1				Sign	X
Jaundice				+1				SS	Fever
coated tongue				+1	Brown		no atrophy or inflammation	Sign	X
Temperature				100.4°F				Sign	X
Blood press.				140/90				Sign	X
Cardiac auscultation					Evidence	3rd	No murmur	Sign	X
Hepatomegaly				1 inch	Down		Lower	SS	Pain

V. CONCLUSION AND FUTURE WORK

This paper presented a pathway on how to structure *the observation space* of PCCD. Despite the existence of many standards for medical data, it has been shown previously that PCCD is poorly represented by them, and this contribution is primarily intended to fill some of the associated gaps.

However, the most innovative aspect in this article concerns signs and their characterizations. Signs are described only in a very specialized literature and are complex to structure. The use of the SSQ model will help represent them and formalize their relations with symptoms.

Finally, the next steps are to use the same methodology to structure *the predisposition space*, as well as to describe how it is possible to use aggregation and fusion technics to extract patterns from PCCD. Results will be presented in another contribution.

ACKNOWLEDGEMENT

The major parts of this contribution have been done when Dr. Aitnouri was working at the Université de Sherbrooke.

REFERENCES

[1] D. A. Ludwicka and J. Doucettea, "Adopting electronic medical records in primary care: Lessons learned from health information systems implementation experience in seven countries", International Journal of Medical Informatics, vol. 78, Jan. 2009, pp. 22-31, doi:10.1016/j.ijmedinf.2008.06.005.

[2] R. H. Miller and I. Sim, "Physicians' use of electronic medical records: Barriers and solutions", Health Affairs, vol. 23, Mar. 2004, pp. 116-126, doi:10.1377/hlthaff.23.2.116.

[3] S. Lusignan, S. Wells, N. Hague, and K. Thiru, "Managers see the problems associated with coding clinical data as a technical issue whilst clinicians also see cultural barriers", Methods of Information in Medicine, vol. 42, 2003, pp. 416-422.

[4] M. L. Tierney and C. M. Henderson, The patient history: evidence-based approach, New York: Lange Medical Books/McGraw-Hill, ed. 2005.

[5] R. Richesson and J. Krischer, "Data standards in clinical research: Gaps, overlaps, challenges and future directions", Journal of American Medical Informatics Association, vol. 14, Nov-Dec. 2007, pp. 687-96, doi:10.1197/jamia.M2470.

[6] H. Scheuermann, W. Ceusters, and B. Smith, "Toward an ontological treatment of disease and diagnosis", In Proc. of AMIA Summit on Translational Bioinformatics (TBI 2009), 15-17 Mar. 2009, pp. 116-120, San Francisco, CA, USA.

[7] J. Loscalzo, I. Kohane, and A. Barabasi, "Human disease classification in the postgenomic era: A complex systems approach to human pathobiology", Molecular systems biology, vol. 3, Jul. 2007, article number 124 (11 pages), doi:10.1038/msb4100163.

[8] L. St. Clair, A. Clift, and L. Dumbelton, "How do I know what I feel? Evidence for the role of self-categorization in symptom perceptions", European journal of social psychology, vol. 38, Jan. 2008, pp. 173-186, doi:10.1002/ejsp.417.

[9] M. Daniel, P. Hájek, and H. P. Nguyen, "CADIAG-2 and MYCIN-like systems", Artificial Intelligence in Medicine, vol. 9, Mar. 1997, pp. 241-259.

[10] P. R. Innocent and R. I. John, "Computer aided fuzzy medical diagnosis", Information Sciences, vol. 162, May. 2004, pp. 81-104, doi:10.1016/j.ins.2004.03.003.

[11] A. M. Butt, A. Mannan, A. Talib, R. Asim, and P. Post, "Decompensated Chronic Liver Disease with Comorbid Treatable Aplastic Anemia in a Homeless Adult Male", Homeless Health Care Case Report: Sharing Practice-Based Experience, vol. 4, Dec. 2008.

 ISBN: 978-1-61208-252-3