Connecting Medical Informatics and Bio-Informatics
R. Engelbrecht et al. (Eds.)
IOS Press, 2005

ICPCview: visualizing the International Classification of Primary Care

Pierre P Lévy[a,b], Laetitia Duché[a], Laszlo Darago[c], Yves Dorléans[b], Laurent Toubiana[b], Jean-François Vibert[b], Antoine Flahault[a,b]

[a] *Hôpital Tenon(Assistance Publique Hôpitaux de Paris), France*
[b] *INSERM U707, Paris, France*
[c] *University of Debrecen, Hungary*

Abstract

This paper proposes a method to visualize the semantic content of data bases where the medical information is coded with the International Classification of Primary Care. The main idea is the identification of a pixel with a code and the conversion of all the data associated with these into an image the ICPCview. The method proceeds in two step, defining the reference frame and using this reference frame to visualize data. The reference frame is built by using a sign/diagnosis binary criterion, a seventeen category nosological criterion and an age ordinal criterion .The results are visualization of the signs and diagnosis of the ICPC according to gender, age and time period of the year. A limitation of the method lies in the fact that the result depends on the chosen reference frame. Further work has to be done with various reference frames and data. However the main point is that, when both the reference set of the image and of the mind of the user are built, the method is powerful at extracting the hidden content of a very large amount of data.

Keywords:
ICPCview, Caseview; Information visualization; International Classification of Primary Care, General practitioner, Code. French Sentinelles network

1. Introduction - Context

A current problem in medicine and biology is the existing of huge data banks whose semantic content is difficult to grasp. A solution was proposed for the Diagnosis Related Group (DRG) data base used in the context of hospital tarrifing. This solution was the caseview method proposed as a generic method applicable to all DRG classifications in the world [1,2,3,4,5].

In the present paper a new step is taken: the method is generalised and applied to the International Classification of Primary Care (ICPC). The ICPC [6] is a classification originating from the International Classification of Diseases (ICD) . It was defined to account for the specific activity of general practitioners. Each contact of a practitioner with a patient is described with a chain of codes corresponding to the implied signs, diagnoses and processes. The codes are classified according to three axes: signs, diagnoses and processes. These three axes are used to browse seventeen socio-nosological categories (social, haematology, metabolic…).

In the present study, this coding was applied to the data transmitted by 1200 general practitioners belonging to the French Sentinelles© network [7,8,9].

The point of this paper is to show the various insights offered by the generalised caseview method applied to this huge ICPC coded medical data bank

2. Material and Methods

The main idea of the generalised caseview is to assimilate an informational entity to a pixel. Then the method proceeds in two steps. The first step is to define a reference frame by ordering these pseudo-pixels according to three criteria. Using this reference frame and defining a colour scale, the second step allows the visualization of the associated with these informational entities data.

The three criteria are a binary criterion which allows splitting of the reference frame into two parts, a nominal criterion which allows the definition of columns where the entities share a common property and an ordinal criterion which allows the arranging of pseudo-pixels symmetrically for the two parts inside each column.

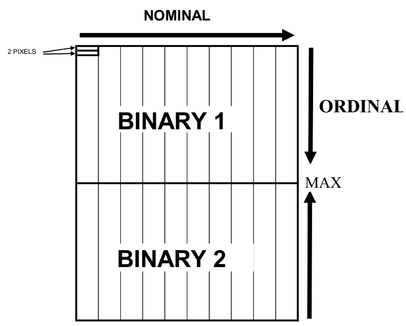

Figure 1- The three criteria of the generalized caseview method

Since August 1997 the data base was built applying a coding algorithm to the transmitted by the practitioners' data. These data correspond to the in-hospital set of their patients. A program converted these referrals into three codes which were allocated to three specific fields of the central data base. Indeed, this automatic coding represents a unique real time system allowing to monitor health at the national level [7,8].

3. Results

3.1 Defining the reference frame

The informational entities used are the diagnosis and signs of the ICPC. The processes are not included because the small number of data.

The binary criterion was signs versus diagnoses: the diagnoses are in the upper part of the reference frame and the signs are in the lower part.

The nominal criterion was the 17 categories of the ICPC: each column of the reference set corresponds to a category (Table 1).

Table 1 –Title of the columns of the reference frame

SOCIAL	HEMATO	METAB, ENDOC	GENERAL	SKIN	MUSCULO-SKELETAL
NEURO	MENTAL	RESPI	CARDIO	DIGESTIVE	KIDNEY
PREGNANCY	FEMALE SYST.	MALE SYST.	EYE	EAR	

The ordinal criterion was the ordering of the codes according to the average age associated with them in the data base.

The reference frame is shown on figure 2. In this figure each "pixel" contains the average age associated with the corresponding "pixel" code. Dark colour corresponds to "old" code (patient older than 65 years old) whereas very light grey corresponds to children (younger than fifteen years old).

The first result appears after analysing the reference frame: the proportion of codes concerned by the study is rather large. This pattern of codes is approximately symmetric, with a larger proportion of diagnoses compared to the signs. Moreover the area where the age is older than sixty five is asymmetric in that there are more "old" diagnoses than "old" signs. This is particularly prominent for the "circulatory" column.

The paediatric diagnoses and signs are marginal.

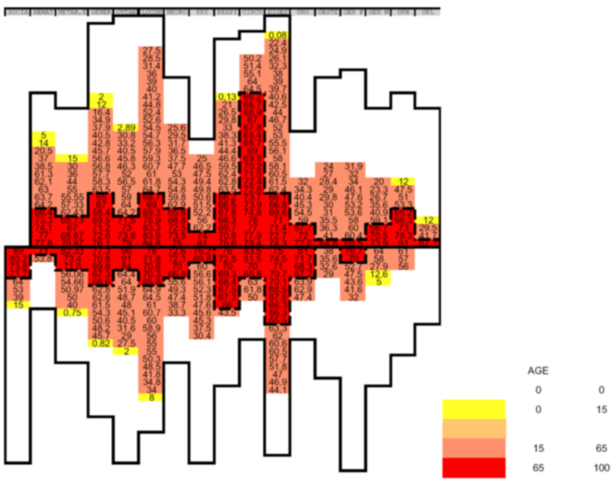

Figure 2 - The Reference frame: each pixel contains the average age associated with the code

3.2 Using the reference set to visualize data

Figure 3 shows in-patient counts corresponding to the three fields of the data base.

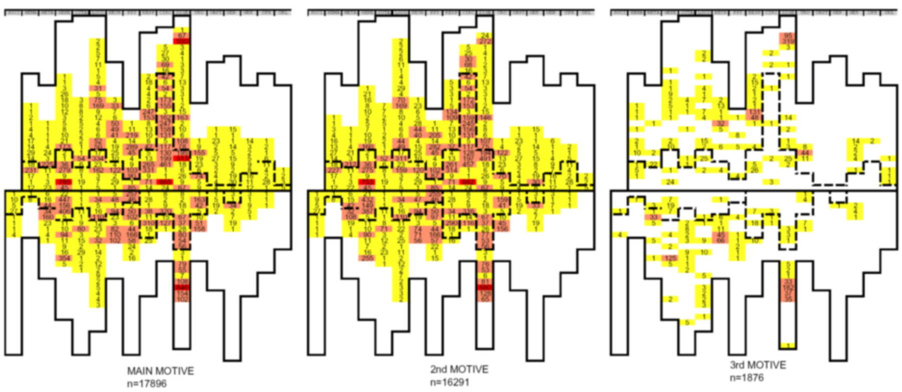

Figure 3 – In-patient count of the three motives

The image on the left is the main in-patient motive: there are many diagnoses for the circulatory, digestive and respiratory systems. The three motives have show similar patterns: only the total number of diagnoses decreases from the left to the right.

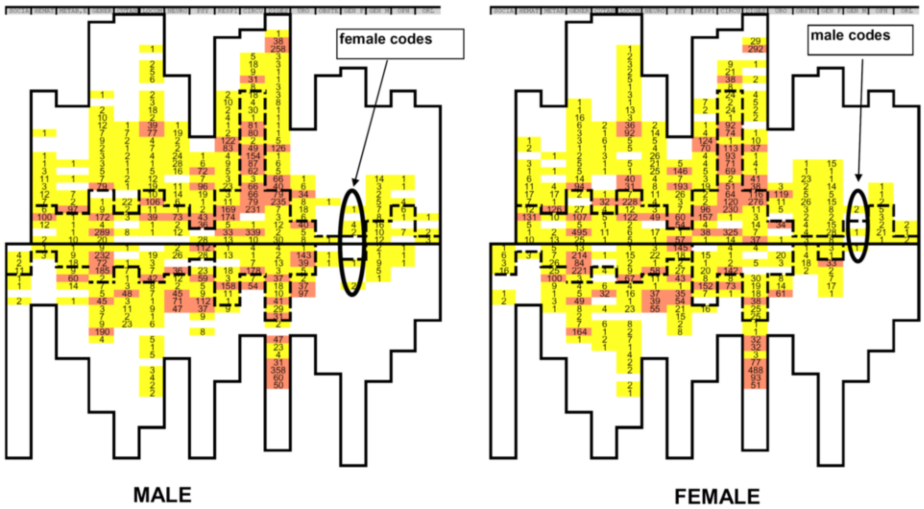

Figure 4 –in-patient count according to Gender

Figure 4 shows that the male and female patterns seem similar. However there are coding errors as it shown on the figure.

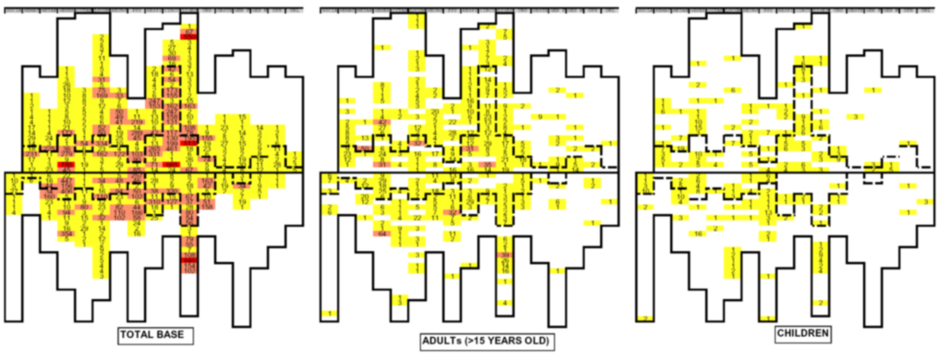

Figure 5 –in-patient count according to age.

Figure 5 shows that in general, the patterns of the total data base are similar to those corresponding to the “adult”. The pattern corresponding to children presents a low number of cases at the nucleus level.

To analyse the temporal data, the year was split into two semesters: a cold semester – October to March- and a hot one – April to September.

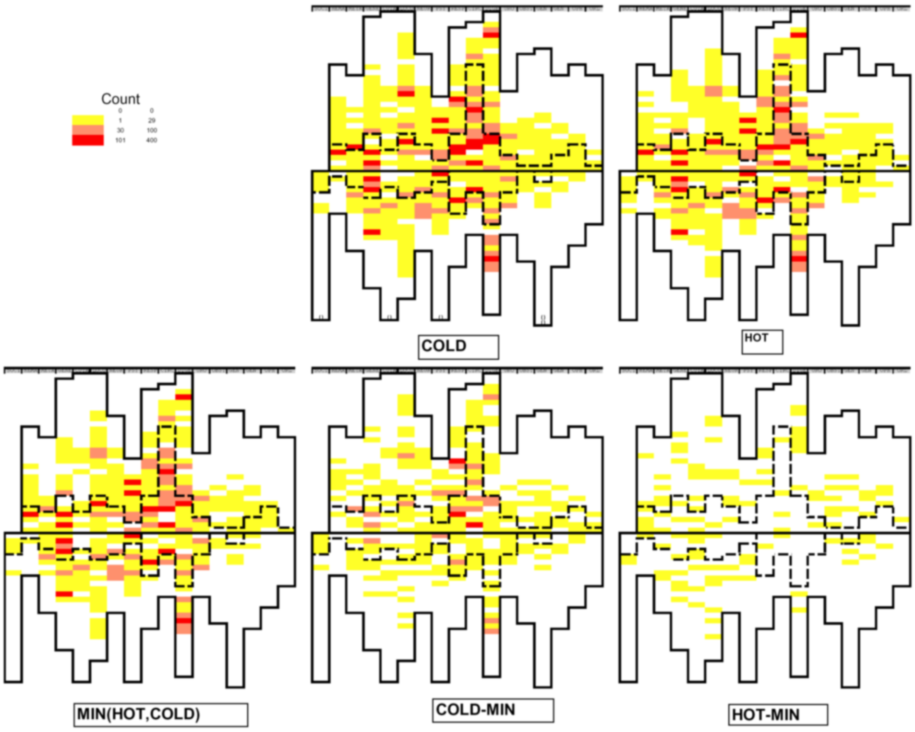

Figure 6 –In-patient count according to the time period of the year.

On the top of figure 6 we can see that the two periods seem similar. The intersection of the two images is shown on the lower left corner. The two images on the lower right show a higher number of diseases (respiratory, circulatory and general) for the "old" diagnoses of the cold period.

4. Discussion

The first benefit of the method lies in its capability to give a first look at the enclosed inside the data base information. However the nature of this information is limited to the point of view "window", in other words, the reference frame. In this way, the data base is viewed from the point of view of the sign/diagnosis, nosological and age triple axis. Only the meaning paned by these axes is visible. More precisely it is important to keep in mind that the ordinal age criterion corresponds to the codes of the first field used to encode pathology. This was visible in figure 2 . The information outside the set of codes present in the first field, is lacking. To summarise, the "window" is restricted to the first field and to the set of information existing in the actual data base. This is the limitation of the method. Other ordinal criteria have to be tested to improve the reference frame.

However the main point is that, when both the reference set of the image and of the mind of the user are built, the method is powerful at extracting the hidden content of a very large amount of data. [10].

5. Conclusion

The main benefit of the method is its capability to allow quick insight on the data. Furthermore it has to be used in combination with other methods.

Moreover, it is important to note that this method is not static: the images can be processed and compared by simple computing (intersection difference…) leading to new results and interpretations.

6. Acknowledgments

We are grateful to the GPs from the French Sentinelles system for transmitting their hospital referral reports.

Moreover, we are grateful to Ann Rocheford for correction of the English version of the paper.

7. References

[1] Lévy PP**.** Le case view une methode de visualisation du case mix*. Journal d'économie médicale mars 2002*; 20(2), 118-127.

[2] Lévy PP The case view a generic method of visualization of the case mix. *International journal of Medical Informatics 2004, 73*, 713-718.

[3] www.b3e.jussieu.fr/caseview/drgview

[4] www.medinfo.hu/darago/caseview.php

[5] László Daragó: Caseview_HUN: easy DRG overview. *Studies in health Technology and Informatics*. Volume 105. Transformation of Healthcare with information technologies. IOS Press. Editors: Marius Duplaga, Krzysztof Zielinski, David Ingram.pp182-189, 2004

[6] World organization of National Collges, Academies and Academic Associations of General Practitioners/Family Physicians (WONCA). *ICPC-2 International Classidication of primary Care; S*econd Edition*.* Oxford University Press, 1998

[7] Letrilliart L, Viboud C, Boelle PY, Flahault A. Automatic coding of reasons for hospital referral from general medicine free-text reports. *Proc AMIA Symp*. 2000; 487-91.

[8] www.sentiweb.org

[9] Valleron AJ, P Garnerin (1992). "Computer networking as a tool for public health surveillance: the French experiment." *MMWR Morb Mortal Wkly Rep 41 Suppl*: 101-10, 1992.

[10] Lévy PP. Caseview: building the reference set. In: *Studies in Health Technology and Informatics*. Volume 105. Transformation of Healthcare with information technologies. IOS Press. Editors: Marius Duplaga, Krzysztof Zielinski, David Ingram.pp172-181, 2004.

Address for correspondence

Pierre Lévy
Hôpital Tenon, Unité de Biostatistique et informatique Médicale
4, rue de la chine, 75970 Paris cedex 20, France
Email: pierre.levy@tnn.aphp.fr