Quality of Life through Quality of Information
J. Mantas et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-101-4-477

Elicitating Patient Patterns of Physician Non-Compliance with Breast Cancer Guidelines Using Formal Concept Analysis

Jacques BOUAUD [a,1], Nizar MESSAI [b], Cédric LAOUÉNAN [c], France MENTRÉ [c], and Brigitte SÉROUSSI [d]

[a] *AP-HP, STIM, Paris, France ; INSERM UMR_S 872, éq. 20, CRC, Paris, France*
[b] *Université François Rabelais Tours, Laboratoire Informatique, UFR Sciences et Techniques, Tours, France*
[c] *Université Paris Diderot, Sorbonne Paris Cité, UMR738, Paris, France; INSERM, UMR738, Paris, France; AP-HP, Hôpital Bichat, Service de Biostatistique, Paris, France*
[d] *UPMC, UFR de Médecine, Paris, France; AP-HP, Hôpital Tenon, DSP, Paris, France ; Université Paris 13, UFR SMBH, LIM&BIO, Bobigny, France; APREC, Paris, France*

Abstract. Because they provide patient-specific guideline-based recommendations, clinical decision support systems (CDSSs) are expected to promote the implementation of clinical practice guidelines (CPGs). OncoDoc2 is a CDSS applied to the management of breast cancer. However, despite it was routinely used during weekly multidisciplinary staff meetings (MSMs) at the Tenon Hospital (Paris, France), the compliance rate of MSMs' decisions with CPGs did not reach 100%. Formal Concept Analysis (FCA) has been applied to elicit formal concepts related to non-compliance. A statistical pre-treatment of attributes has been proposed to leverage FCA and discriminate between compliant and non-compliant decisions. Among the 1,889 decisions made over a 3 year-period, 199 decisions of recommended re-excisions have been considered for analysis. In this sample, non-compliance was explained by uncommon clinical profiles and specific patient-centred clinical criteria.

Keywords. Computer decision support systems, Formal Concept Analysis, Clinical Practice Guidelines, Guideline adherence, Evaluation

Introduction

Clinical practice guidelines (CPGs) are currently widely developed to improve the quality of clinical care and promote cost-effective practices. Clinical decision support systems (CDSSs) providing patient-specific guideline-based recommendations may be efficient tools to promote CPGs implementation. OncoDoc2 [1] is a CDSS on breast cancer management based on local reference guidelines (CancerEst). It has been routinely used in multidisciplinary staff meetings (MSMs) the Tenon Hospital (Paris,

[1] Corresponding Author: J. Bouaud, INSERM U872 eq. 20, CRC, 15 rue de l'école de médecine, 75005 Paris, France; E-mail: jacques.bouaud@sap.aphp.fr.

France) for nearly 3 years. However, clinicians did not systematically follow the system's recommendations and compliance has been measured at 91% [2].

A great variety of barriers to physician compliance with CPGs has been identified [3]. The actual implementation of computerized CPGs relies indeed on a particular alchemy involving the CPGs, the CDSS embedding the CPGs, the decision maker(s), the clinical setting and the organizational process of decisions, as well as the patient herself. However, research currently conducted essentially explores the CDSS effect to elicit the "winning" technical features, *e.g.* design, implementation, and level of description, that could predict CDSS effectiveness to increase physician compliance with CPGs. Tu and Musen [4] have previously analysed the guideline effect, making the difference between "consultation guidelines", *i.e.* one-shot simple guidelines, and "management guidelines" used for the more complex management of patients with, for instance, chronic diseases. As for the patient effect, Messai *et al.* [5] suggested in a previous study that non-compliance was mostly observed for complex patients and explained by physicians' lack of familiarity with CPGs when not using OncoDoc2.

Formal Concept Analysis (FCA) [6] is a symbolic data analysis and classification method allowing to derive implicit relationships within a set of objects described by a set of attributes. The richness of lattice structures (derived by FCA) as well as their well-established formal properties have motivated a wide range of successful applications of FCA in many fields, particularly in medicine and psychology [7]. The objective of this work is to use FCA together with a statistical pre-treatment to study the "patient effect" on residual non-compliance with CPGs when a CDSS, reminding patient-specific recommendations, was used in MSMs.

1. Material and Methods

1.1. MSM decision data set

OncoDoc2 has been routinely used during weekly MSMs to support CPG-compliant decisions. Using the system consists in answering patient criteria while navigating in the knowledge base (KB) represented as a decision tree until a leaf is reached where treatment recommendations are issued. The navigation process selects a path in the KB, made of a set of criteria. They constitute the patient profile and are related to tumor properties (*e.g.* presence of microinvasion), patient data (*e.g.* menopausal status), and therapeutic history (*e.g.* prior neoadjuvant treatment). For each MSM decision, the profile was recorded along with the decision compliance status to form the data set.

1.2. Processing data for FCA

In the FCA theory, objects are grouped into units called "formal concepts" based on attribute sharing between objects. In its basic setting, FCA takes as input binary data represented as a formal context, a triplet (G, M, I), which materializes a set G of objects, a set M of attributes, and a binary relation I between the two sets indicating whether an object has or not an attribute [8]. A formal context is commonly represented as a binary table where rows correspond to objects, columns correspond to attributes, and each table cell (i, j) contains or not a cross (×) depending on whether the corresponding object (line i) has or not the corresponding attribute (column j). *Formal concepts* are then defined as maximal sets of objects having in common maximal sets of attributes.

They are computed based on a *Galois connection* which associates for each set of objects $A \subseteq G$ the maximal set of attributes common to all objects in A and, dually, for each set of attributes $B \subseteq M$ the maximal set of objects having all the attributes in B. Formal concepts of a context are partially ordered based on the inclusion between their sets of objects (and, dually, of attributes) and form a lattice.

In our application, objects are decisions made during MSMs and attributes are the recorded criteria. However, due to the tree structure of the KB, each profile uses a different set of criteria and a conceptual scaling has been performed (plain scaling) to get a valid formal context. Concept lattices were computed with Conexp (http://conexp.sourceforge.net/).

1.3. Formal context pre-treatment with respect to conformity

When run on the original formal context, FCA algorithm produced a high number of formal concepts and “bushy” lattices where potential relationships with (non-) compliance were hidden due to attribute dispersion among decisions. Thus, we extended the method used in [5] and developed a two-step simplification method to emphasize attributes which distribution among compliant and non-compliant decisions was significantly different. We first computed the observed proportions of each attribute in the subgroup of objects with CONFORMITY:NO and the complementary subgroup with CONFORMITY:YES. Then, we normalised these proportions to take into account the unevenness between non-compliant and compliant decisions and compared the two resulting ratios. We used the statistical Fisher exact test, where the risk level measures the probability of erroneously concluding that the proportion is different from what was expected. In a first step, attributes non-significantly differently distributed among the two subgroups, and thus indistinctly related to a subgroup, were totally removed. In a second step, we selectively pruned remaining “unbalanced” attributes. We assumed that values in the subgroup where they were less represented are “noise”, and thus, we removed them to emphasize the attribute relationship with the other subgroup. Classical FCA was then performed on the simplified data set. Figure 1 illustrates these 2 steps on a dummy context where attribute C1 is evenly distributed in the 2 subgroups while C2 is unequally balanced in the subgroups.

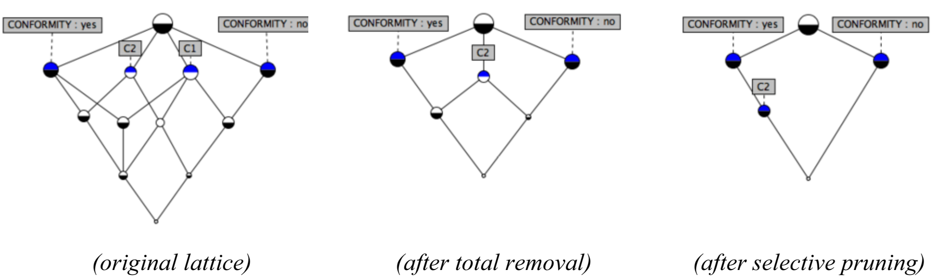

Figure 1. Context simplification on an example lattice (conformity and 2 additional attributes C1, C2).

1.4. Statistical data analysis

Univariate analysis was performed with Fisher exact tests on the data set. Then, variables associated with non-compliance (with $p < 0.20$) were entered in a multivariate adjusted logistic regression model with a backward selection procedure and a significance level of $p = 0.05$. Odds ratios (OR) and 95% confidence intervals were also estimated. SAS software was used (SAS Institute, Cary, NC, version 9.2).

2. Results

Between February 2007 and September 2009, 1,889 decisions for breast cancer patients have been recorded. Among them, 199 were included in the re-excision group which is only considered here for the analysis. The number of non-compliant decisions in this group was 28 (14.1%). The theoretical rate used for attribute selection is then 85.9%. Decisions in the re-excision group used 56 criteria from OncoDoc2's KB, mostly binary (*e.g.* multifocality = yes or no), sometimes ternary (*e.g.* SBR = 1, 2, or 3). After scaling, 102 attributes were created for the 199 "objects" (100 clinical criteria and the 2 attributes for conformity status). The identification of unevenly distributed attributes was performed for 3 risk levels: 0.05, the usual type I error risk, 0.10, and 0.20, similarly to the threshold used in the classical multivariate analysis.

We used the 20% risk level threshold for attribute selection and found 84 attributes equally distributed within both subgroups. They were totally removed during the first step of simplification. As a result, 16 attributes were selected for the analysis. With lower risk levels, we got less remaining attributes, 8 and 4, respectively, for the risks of 10% and 5%. They are however always included in the 16. In the second step of the simplification method, 7 attributes were selectively pruned in favour of compliance, and 9 in favour of non-compliance. Figure 2 shows the corresponding obtained lattice.

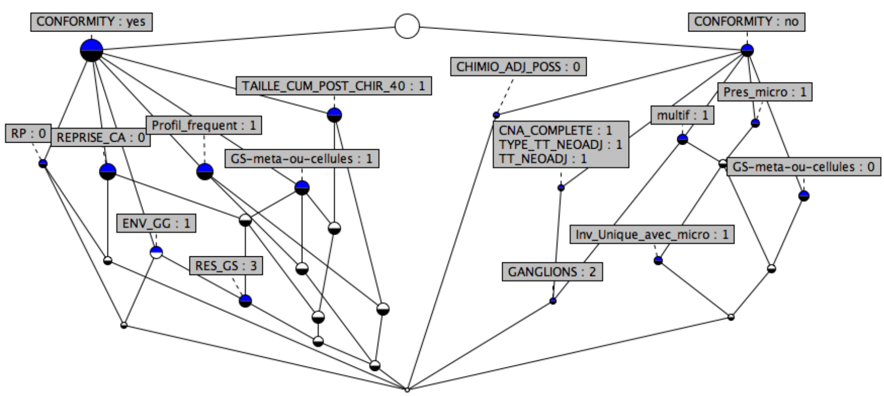

Figure 2. Lattice obtained after simplification (using a 20% risk).

Formal concepts leading to non-compliance are multifocal tumor, presence of microinvasion, unique invasive tumor with microinvasion, prior neoadjuvant treatment by complete chemotherapy, contra-indications to adjuvant chemotherapy and negative sentinel node procedure. Formal concepts leading to compliance are common patient profile, positive sentinel node procedure, negative progesterone receptors, no indication for lymph node dissection, and a tumor larger than 40 mm sized after surgery.

In the classic statistical approach, univariate analysis selected 6 patient-centered variables associated with non-compliance ($p < 0.20$): multifocal tumor, microinvasion, unique invasive tumor with microinvasion, prior neoadjuvant treatment, uncommon patient profile, and negative sentinel node procedure. In the final multivariate model, two variables were independently and significantly associated with non-compliance: uncommon patient profile (OR=3.4; 95% CI=[1.1-10.0]; $p = 0.03$) and invasive tumor with microinvasion (OR=4.9; 95% CI=[1.4-16.2]; $p = 0.01$).

3. Discussion and Conclusion

When using FCA on objects from "real life" such as breast cancer management decisions, formal contexts are complex and attributes are spread over compliant and non-compliant decisions to represent the singularity of actual patients. We have proposed a two-step simplification method as a pre-treatment of attributes and then used classical FCA. Patient-specific patterns of non-compliance such as unique invasive tumor with microinvasion, prior neoadjuvant treatment by complete chemotherapy, negative sentinel node procedure and uncommon clinical profile have been identified. When analysed, these patterns seem to correspond to clinical situations in which physicians do not agree with guidelines. This would explain why they do not comply despite the use of the CDSS OncoDoc2. These patterns could bring clinicians new insights for situations at risk of non-compliance, although it is sometimes legitimate not to comply with CPGs. These first results, obtained on a small data set ($n = 199$), should be confirmed for the other groups of decisions (pre-surgery, adjuvant treatment) and by using more complex data mining methods although they have been partially corroborated by multivariate analyses of the data.

Acknowledgments

We thank Prof. S. Uzan and all the participants of the senology MSM of the Tenon Hospital for their participation to the study and their support in using the OncoDoc2.

References

[1] Séroussi B, Bouaud J, and Antoine EC. OncoDoc, a successful experiment of computer-supported guideline development and implementation in the treatment of breast cancer. *Artif Intell Med* 2001;22(1):43– 64.

[2] Bouaud J, Séroussi B, Gligorov J, Daraï É, Lotz JP, Rouzier R, *et al*. Improving adherence with clinical practice guidelines when using OncoDoc2, a decision support system for breast cancer management: what is the next step? In: International Forum on Quality and Safety in Health Care, Nice, France. BMJ Group, April 20-23 2010.

[3] Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PAC, *et al*. Why don't physicians follow clinical practice guidelines? a framework for improvement. *JAMA* 1999;282(15):1458–65.

[4] Tu SW and Musen MA. A flexible approach to guideline modeling. *J Am Med Inform Assoc* 1999;6(suppl):420–4.

[5] Messai N, Bouaud J, Aufaure MA, Zelek L, and Séroussi B. Using formal concept analysis to discover patterns of non-compliance with clinical practice guidelines: a case study in the management of breast cancer. In: Peleg M, Combi C, Abu-Hanna A, and Andreassen S, eds, AIME 2011, Lecture Notes in Computer Science. Springer, 2011:119–28.

[6] Ganter B, Stumme G, and Wille R, eds. *Formal Concept Analysis - Foundations and Applications*, (vol3626) of *Lecture Notes in Computer Science*. Springer, 2005.

[7] Schnabel M. Representing and processing medical knowledge using formal concept analysis. *Meth Inform Med* 2002;41(2):160–7.

[8] Ganter B and Wille R. *Formal Concept Analysis*. Mathematical Foundations. Springer, 1999.