Implementation and management of a biomedical observation dictionary in a large healthcare information system

Pierre-Yves Vandenbussche,[1,2] Sylvie Cormont,[3] Christophe André,[3] Christel Daniel,[3] Jean Delahousse,[1] Jean Charlet,[2] Eric Lepage[3,4]

[1]Mondeca, Paris, France
[2]Eq.20, INSERM UMRS872, Paris, France
[3]Assistance Publique—Hôpitaux de Paris, Centre de compétences et de services, Paris, France
[4]Département d'Information Hospitalier, Université Paris 12, CHU Henri Mondor, Créteil, France

Correspondence to
Dr Pierre-Yves Vandenbussche, Mondeca, 3 cité Nollez, Paris 75018, France; py.vandenbussche@gmail.com

ABSTRACT

Objective This study shows the evolution of a biomedical observation dictionary within the Assistance Publique Hôpitaux Paris (AP-HP), the largest European university hospital group. The different steps are detailed as follows: the dictionary creation, the mapping to logical observation identifier names and codes (LOINC), the integration into a multiterminological management platform and, finally, the implementation in the health information system.

Methods AP-HP decided to create a biomedical observation dictionary named AnaBio, to map it to LOINC and to maintain the mapping. A management platform based on methods used for knowledge engineering has been put in place. It aims at integrating AnaBio within the health information system and improving both the quality and stability of the dictionary.

Results This new management platform is now active in AP-HP. The AnaBio dictionary is shared by 120 laboratories and currently includes 50 000 codes. The mapping implementation to LOINC reaches 40% of the AnaBio entries and uses 26% of LOINC records. The results of our work validate the choice made to develop a local dictionary aligned with LOINC.

Discussion and Conclusions This work constitutes a first step towards a wider use of the platform. The next step will support the entire biomedical production chain, from the clinician prescription, through laboratory tests tracking in the laboratory information system to the communication of results and the use for decision support and biomedical research. In addition, the increase in the mapping implementation to LOINC ensures the interoperability allowing communication with other international health institutions.

INTRODUCTION

Context

The Assistance Publique Hôpitaux Paris (AP-HP) is the largest European university hospital with 12 hospital groups composed of 44 hospitals, 165 laboratories in and around Paris (23 000 beds, 1 000 000 inpatients per year and 4 000 000 outpatients per year).

The objective of AP-HP is to build an information system based on the concept of information sharing, integrating clinic, medico-technical, medico-economic and biomedical research domains. In 2005, AP-HP decided to develop a new laboratory information system (LIS) common to the 12 hospital groups and based on a single shared core application. Consequently, a biomedical observation dictionary was required to accommodate the entire production chain. This dictionary should ideally remain independent of any system constraints and be common to all laboratories.

The biomedical observation dictionary, named AnaBio, is based on previous works (C-NPU,[1] Euclides,[2] Names-Lab,[3] Cumul[4] and logical observation identifier names and codes (LOINC)).[5–7] It consists of all codes and labels for the definition of each observation (the laboratory domain of the international initiative 'integrating the healthcare enterprise', IHE LAB, defines a laboratory test result as an observation that is a measurement of a single variable or a single value derived logically and/or algebraically from other measured or derived values). It standardizes other elements, such as the display label, the edition label, the mnemonic codes and units. The dictionary must rely on a terminology shared by as many as possible to secure the communication of the test results to health practitioners, no matter where the practice is located. Considering that LOINC has now the widest international development,[8] AP-HP decided to map its biomedical observation dictionary to LOINC and ensures the maintenance of this mapping. Other hospitals showing a more or less complete mapping to LOINC have also made this choice.[9–11] In addition, the French translation of LOINC labels is part of a cooperative work between the French Society of Laboratory Computing (SFIL) and the AP-HP, promoted by the shared healthcare information system agency (ASIP-Santé: Agence des Systèmes d'Information Partagés de Santé.[12]

Objectives

This paper aims to illustrate the production, the implementation and the routine management of a biomedical observation dictionary shared by 120 out of 165 AP-HP laboratories. This management is enabled by the use of a knowledge management platform.

In the Methods section, we first introduce the AP-HP biomedical observation dictionary AnaBio and the differences and similarities with the LOINC dictionary. We also present the biomedical observation workflow based on AnaBio. Second, we describe the methodology used to represent the biomedical knowledge and the AnaBio management. In the Results section, we detail the implementation in the AP-HP institution and describe the AnaBio integration into the platform and the mapping to LOINC. We then discuss the impact of the dictionary integration and the advantages and disadvantages of using

AnaBio as a LIS dictionary. Finally, we conclude with the perspectives on relevant open areas for further research and improvements.

ANABIO AND LOINC
Need for a specific dictionary
In 1998, AP-HP introduced its first information server of biomedical observations. It identified the problems related to the diversity of expression modes and the need to rely on standard semantic.[13] In the absence of any adequate representation framework available at that time, the AP-HP set up a working group drawing on the expertise of its biology laboratories. This work was based on a highly formalized building process that would later articulate Names-Lab nomenclature. The 18 000 observation entries from Names-Lab were managed in a relational database (SIAM: Medical observation information server).[14] This telematics system was dropped in 2004. However, it created the basis for the development of the AP-HP biomedical observation dictionary AnaBio.

Since 2004, three approaches have been considered within the AP-HP:
- using the reference terminology LOINC codes as an unique AP-HP observatory entry identifier and translate into French,
- transcoding LOINC codes into a unique identifier specific to the institution,
- developing a biomedical observation dictionary from the observations already in use and map the dictionary codes to LOINC.

AP-HP chose to map the biomedical observation dictionary (AnaBio) to the reference terminology LOINC and to ensure the maintenance of this mapping.[15] Some inadequacies motivated the choice not to use LOINC directly, including:
- Incompleteness: LOINC provides records that are not necessarily useful for AP-HP laboratories and, conversely, does not meet all the needs. For example, a large number of the exploration tests used by the AP-HP endocrinology services are not yet referenced in LOINC. This example is illustrated in supplementary appendix 1 (available online only) in the 'incompleteness' sheet.
- Granularity: a finer detail level than the one provided by LOINC is sometimes required by AP-HP biologists, for whom some elements must be more specifically described in specialized disciplines of biology. Some test results described by a unique code in LOINC correspond to more than 10 codes in AP-HP (eg, a study of fusion transcript). This example is illustrated in supplementary appendix 1 (available online only) in the 'granularity' sheet. Conversely, LOINC codes of a more generic level of detail are sometimes needed, as for example for LOINC system in microbiology.
- Specific parameters: national regulatory authorities, such as the nomenclature of medical biology procedures (NABM), impose the inclusion of specific attributes to an observation entry for billing concerns.
- Language requirement: the data must be available in French.

Supplementary appendix 1 (available online only) presents a sample of the AnaBio database and the mapping between AnaBio observation entries and the LOINC record as well as some examples of inadequacies stated above.

Dictionary structure
An AnaBio observation entry consists of five core elements:
- Analyte: substance or body element to be analyzed (eg, chemical molecules, hormone, bacteria, virus, cellmark).
- Parameter: element specifying the conditions applying to an analyte (eg, challenge, formula, allergens, activator or inhibitor reaction).
- System: fluid, solution or substance for biological consideration (eg, serum, urine).
- Method: method used to carry out the biomedical test (eg, electrophoresis, culture, colorimetry).
- Unit: international units and common units used in laboratories (eg, mmol/l, g/l).

It has to be compared with LOINC, which is based on a pre-coordinated system of six core elements:[16] (1) component; (2) property; (3) time aspect; (4) system; (5) scale type; and (6) method. The two terminologies are very similar in structure. The LOINC component matches with the AnaBio analyte plus parameter; the system and method are identical. In AnaBio, we prefer using unit that is more precise than the LOINC property. The LOINC time aspect is included in the AnaBio system element. However, the main difference in each element is the level of description.

Each AnaBio observation entry is described by a label (AP-HP label) to which is assigned a five-character alphanumeric code (AP-HP index). This dictionary is also linked to related external data, such as the list of hospital units as well as their contacts. Each AnaBio observation entry is mapped to LOINC and NABM when possible.

AnaBio observation entry includes elements that are necessary to the LIS setup: display labels, edition labels, observation calling codes (mnemonic codes) and scale. Unlike LOINC, it does not take the panels into account.

The AnaBio dictionary offers the necessary managerial flexibility for daily use while maintaining the semantic interoperability with other international health institutions via the mapping to LOINC. The biomedical observation dictionary is an interface terminology mapped to a reference terminology.[17] This involves a daily mapping task between these two terminologies.

Dictionary use
The AP-HP competency and services center (figure 1) is in charge of the observation entries collection in the laboratories, the dictionary maintenance (creation, modification and deprecation: when removing a dictionary entry, it is not deleted as it may be currently in use. However, it is deprecated in favor of the most recent entry that will be used from now on) and its mapping to LOINC. The management workflow is composed of the following steps:
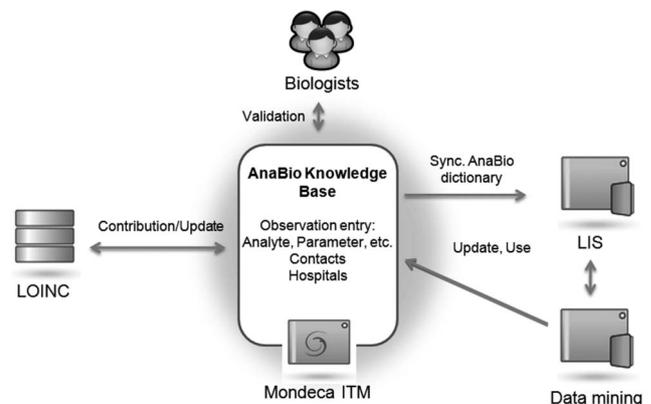
Figure 1 Data flow around the biology knowledge database. ITM, intelligent topic manager; LIS, laboratory information system; LOINC, logical observation identifier names and codes.

1. the biomedical tests selection performed by the biologists via a collection tool,
2. the request for observation entry creation or modification,
3. the request validation by the biologists experts,
4. the possible mapping to LOINC,
5. the creation or modification import in all LIS bases.

Only the competency center performs these imports. No creation or observation entry modification is directly performed in the LIS.

A PLATFORM FOR THE ANABIO DICTIONARY MANAGEMENT
System requirements and limitations
The AnaBio management faces some constraints regarding the day-to-day dictionary evolution, its effective distribution and its use. The main dictionary management and implementation requirements are:

1. to have a tool able to manage a data volume evolving continuously,
2. to allow the editing of the knowledge repository by several users simultaneously,
3. to have traceability of the items of the dictionary (eg, date of creation),
4. to implement a validation process of the observation entries used by either management team or domain experts,
5. to control the data integrity (eg, duplicate control),
6. to generate statistics about the knowledge repository content,
7. to consider the data related to the observation (eg, hospital units, contacts),
8. to ensure the links between related data and observation entry (eg, use),
9. to be able to provide exports of all or part of the knowledge database in standard formats,
10. to support mapping to other terminologies (eg, LOINC, SNOMED),
11. to support internationalization (eg, use of several languages to designate a dictionary entry element).

The current system used so far is a spreadsheet enhanced by several scripts that facilitate exports and imports from and to the dictionary. Whereas this solution partly fulfills requirements 3, 4, 5, 6, 9, it does not address the other requirements listed above. A spreadsheet-based knowledge representation has some limitations.[18] Practically, the Excel tool does not support a concurrent edition (req. 2) or traceability of the items of the dictionary (req. 3) and is not well suited for relation browsing and cross-modification updates (req. 7, 8, 10). As the file size increase, any modification is slower to perform, decreasing by the meantime the work experience comfort (req. 1). Translation relationships are also hard to manage with this kind of solution (req. 11).

This system requirements identification leads us to implement a solution based on the system management tool named intelligent topic manager (ITM) developed by Mondeca.[19]

Knowledge representation
ITM is a dynamic web user interface tool based on a domain model represented by a knowledge graph in a semantic web format. This model may be a posteriori changed on the fly to adapt to a domain change.

The literature describes several approaches to model terminologies and mappings:
1. The alignment implementation between distributed terminologies;[20–22]
2. The definition of a top-ontology or a top-thesaurus to link the knowledge organization systems;[23–25]
3. The design of a hub meta-model supervising these organization systems.[26 27]

Our method fits into this latter approach by defining a unique representation model. The advantage of this unique model is the integration of the different terminologies in a single server and thus to allow the editing of these repositories. The platform set up requires the transformation of semistructured data to a structured model representation and relies on a knowledge engineering method[28] summarized here in three steps: modeling, data migration, and integration/validation.

Modeling
The modeling task is conducted in close collaboration with the terminology maintenance unit. This collaboration aims to understand the usefulness of each domain element and apprehend the needs impacting the model. Practically, some status indicators, such as 'to be sent', 'validated by the experts', 'to be requested to LOINC' have been put in place to ensure a smooth follow-up and generate statistics based on different criteria.

The model designed and expressed in OWL (web ontology language)[29] is generic enough to represent any type of terminology (figure 2). It includes LOINC, AnaBio but also future resources that will be useful to improve the interoperability, such as SNOMED. This model remains extensible to consider the particularities of each terminology (eg, the NABM codes) but also to link the AnaBio dictionary to related knowledge such as hospital units, contacts and to map it to reference terminologies (such as LOINC). In the field of knowledge organization system representation, some norms and standards exist such as SKOS[30] (simple knowledge organization system) and BS8723[31] (British standard). Our approach does not pretend to define an ex nihilo model but wants to be a good practice paradigm for controlled vocabularies representation. Our method uses and extends parts of modeling in these standards.[32] Supplementary appendix 2 (available online only) presents the AnaBio abstract model in OWL format.

Data migration
The data migration task has been initiated in parallel to the model design process. It requires the transformation of the entire spreadsheet data to allow the integration and the conformity with the new formal model. Some information related to the dictionary is available separately in spreadsheet format and must be included in this model. During the migration process, some data inconsistencies are identified and cleaned (eg, duplicate controls, spelling errors, integrity control on some values). The data migration relies on a JAVA tool developed specifically to generate RDF/XML files from XLS files. The RDF/XML output format is consistent with the formal rules defined in the model and is ready for integration. This step allows the validation of the model by comparison with the data and the improvement of the data quality included in the dictionary while preserving the AP-HP identifiers and labels uniqueness.

Integration/validation
The modeling and data migration tasks allow an iterative refinement work. The model and migrated data are imported into the platform to be validated. During this step, the terminology
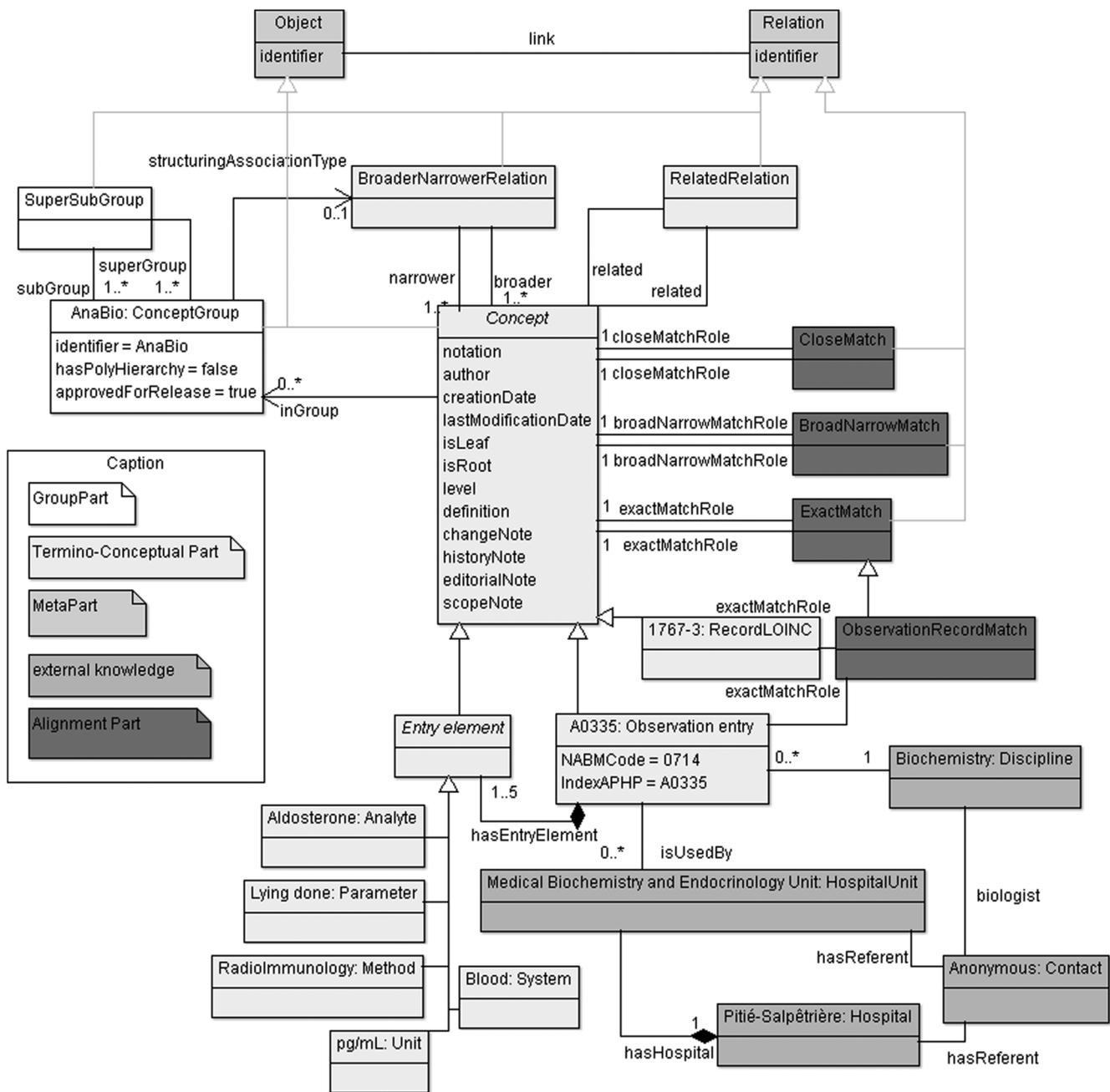
Figure 2 Simplified UML class model of the biology knowledge database and instantiation of one observation entry. LOINC, logical observation identifier names and codes; NABM, nomenclature of medical biology procedures.

maintenance team works in parallel with the spreadsheet and ITM tool. After this validation, corrections and improvements are made to the model and thus to the data migration. After the validation process, the platform is deployed in the production environment.

Platform implementation
Imports and exports
A transformation module is designed to transform AnaBio data representation format into LIS-specific formats. It satisfies the requirements of the knowledge import and export. In addition, ongoing work aims at exchanging information in accordance with the IHE LCSD profile.[33]

A specific module has been developed to ensure the update of the LOINC terminology. When removing a LOINC record, this element is not deleted as it could be currently in use. However, the record is deprecated in favor of the most recent element that will be used from now on. Mass import is used for introducing new LOINC codes resulting from the communication with the Regenstrief Institute.[34]

Integrity rules
A set of rules has been defined to enable a regular integrity checking of the knowledge base. The rules are executed via an inference engine included in the ITM tool. For example, a rule has been defined to verify the number of characters in the labels as per the requirements of the AP-HP common LIS. Another rule checks the uniqueness of the AP-HP index and the mnemonic within the dictionary.

Statistics

It is necessary to be able to provide and generate statistics regarding the content of the knowledge base. The statistical results are obtained using the SPARQL language[35] and are available upon request in report table. The support team can modify existing requests and add new ones at any time.

RESULTS

Dictionary implementation in the platform

The platform implementation phase lasted 1 year from August 2009 to July 2010. It required six iterations of the modeling, data transfer, integration and validation tasks.

The transition from semistructured data (spreadsheet) to structured data (formal model) has forced the cleaning of data considered as incoherent. This AnaBio data quality improvement is a major achievement of this project. We adopted a similar approach to the one used in LOINC. The observation entry elements have been progressively structured (analyte and parameter have been differentiated) and writing rules have been defined. For example, all acids should be written in the form of salts (lactate rather than lactic acid), microorganisms labels in Latin (*Streptococcus pneumoniae* rather than pneumococcus) and chemical molecules with their INN (paracetamol instead of acetaminophen). In addition, the data were cleaned when merging similar terms, such as IL-2 and interleukin-2, or synonyms, such as free fatty acids and fatty acids non-esterified.

As illustrated in figure 3, observation entry and observation entry element (analyte, parameter, system, method and unit) numbers increased, respectively, from 26 458 to 35 714 and from 103 758 to 144 126. However, this increase was not linear. When analyzing the figures related to the creation, depreciation and modification, we observed two important cleaning stages and one major dictionary enrichment stage.

Stage 1 (between August and October 2009): numerous observation entries and observation entry elements were depreciated; 96% of observation entry depreciations correspond to the discipline of microbiology subdivided into virology (47%), bacteriology (33%) and myco-parasitology (16%); 84% of observation entries creations correspond to the disciplines of bacteriology (48%) and myco-parasitology (36%). The aim is continuously to enrich the repository, clean the data and improve the quality (prevalent at this stage). During this stage, the biologists in charge of microbiology initiated a new conception work leading to a large depreciation number.

Stage 2 (between November 2009 and February 2010): a large number of observation entries were created (8719) and 83% of these creat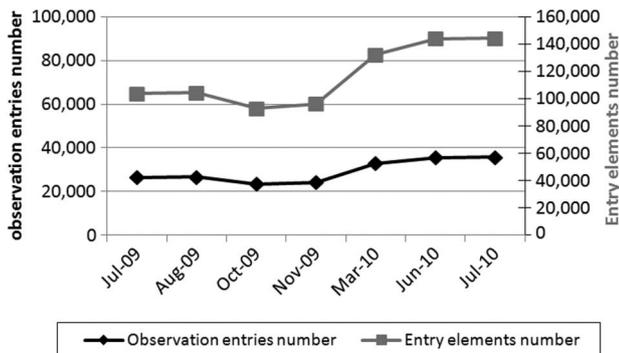ions were for virology. This increase is mainly due to the new observation integration and to the observation duplication based on the technique used (eg, GC-MS, LC-MS). The duplication constraint mainly applies to pharmacology, which enables clinicians to compare only comparable values.

Stage 3 (between March 2010 and July 2010): a moderate number of observation entries were created and depreciated; 85% of entries depreciated correspond to microbiology (59%) and pharmaco-toxicology (26%). The creations were shared between several disciplines. This stage tends to reinforce the dictionary by cleaning the data in view of the adequacy with the model and the integration into the tool.

The use of the platform within the information system

Currently, AnaBio implemented since 2006 at the AP-HP includes 51 281 entries used by 120 out of 165 laboratories covering all disciplines of biology. By using AnaBio, the number of codes transiting the AP-HP results servers dropped from 560 000 (when using different dictionaries between the laboratories) to 51 281 (when using the AnaBio common dictionary).

Due to the use of AnaBio as an execution terminology, 15% of the AnaBio entries (n=7932) are not supposed and will never match with LOINC codes. Indeed, these entries are used in the LIS to keep a trace of intermediary tests results (eg, during a microbiological diagnostic) and are out of LOINC scope. This figure applies only to LIS internal observation entries or to the one allowing biologists to send a biological comment to the clinicians. The remaining 84% of AnaBio entries (n=43 349) are supposed to match with LOINC codes: 40% of Anabio entries (n=20 407) are currently mapped to LOINC and due to resource and time constraints 45% of the AnaBio entries (n=22 942) are still to be mapped, as illustrated in figure 4. Figure 5 presents the percentage of the mapping done, the mapping to be done per discipline (with a maximum of 65% in pharmacology) and the reuse percentage of the same LOINC codes for different AnaBio entries (indicating a lack of granularity and incompleteness of LOINC).

It should be underlined that on average 54% of the mapped AnaBio codes (n=11 045) are linked to a LOINC code that is used more than once in the AnaBio database. This is particularly obvious for microbiology, pharmacology and molecular biology.
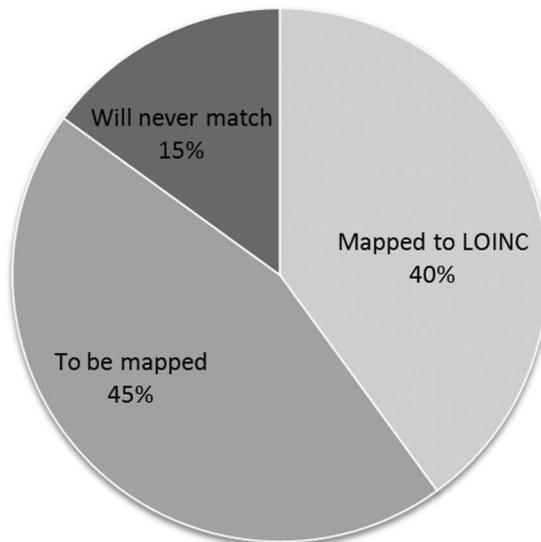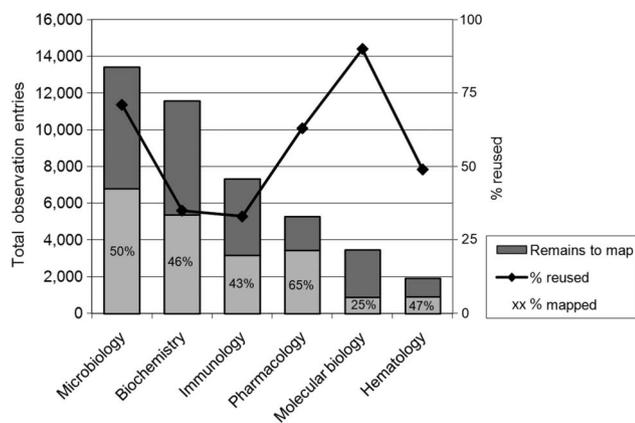

Figure 3 Evolution of observation entries/entry elements numbers during the project duration.

Figure 4 State of the mapping between AnaBio and logical observation identifier names and codes (LOINC).

Figure 5 AnaBio/logical observation identifier names and codes (LOINC) mapping implementation (40% mapped) of the observation entries where there is a potential correspondence and the reuse percentage of the same LOINC codes for different AnaBio observations.

With respect to molecular biology, the mapping rate is low (25%). However, we noticed for this particular discipline a high reuse of the same LOINC codes (90% of AnaBio codes reuse the same LOINC code). It shows the current lack of granularity of the LOINC records compared to the AP-HP needs. In addition, the duplication of the AnaBio codes depending on the techniques led us to use the same LOINC codes for microbiology (71%) and pharmacology (63%). Finally, the decision not to reuse LOINC directly is supported by the study on the number of separate LOINC codes mapped to AnaBio. In fact, the current mapping implementation uses only 26% of the LOINC codes dedicated to biology (n=11 020).

DISCUSSION

The implemented platform automates and integrates a large number of tasks (eg, automatic index creation or input control implemented according to constraints defined in the model). It releases the team in charge of the AnaBio dictionary from proceeding outside of their expertise. The methods and tools used by the platform allow a rigorous management of the dictionary and a better data consistency.

The transition from a free entry without real data validation to a knowledge database constrained by a formal model revealed data inconsistencies. More than 10% of the original data have been identified as inconsistent and have been cleaned mainly by merging identical information, depreciating incorrect information or adding the missing data.

The use of the platform as a multilingual multiterminology server and the experience gained with this project have been rewarded. We have been chosen by ASIP Santé to act as focal point for the LOINC French translation.

The use of AnaBio as a LIS dictionary presents both benefits and challenges. Practically, AnaBio, as an interface terminology, should support some LIS setting elements that should not exist in a reference terminology. This results in 15% of observation entries out of LOINC scope and can therefore not be mapped to LOINC. However, this configuration enables a centralized management of the dictionary and provides a greater visibility of the real use of the AP-HP biomedical observation entries.

Other projects in different application areas have benefited from the multiterminology model designed in this work. It includes the InterSTIS project[36] and the Eurovoc thesaurus editing.[37] Our model has been submitted to the research group

on the elaboration of the ISO 25964 on multilingual thesaurus representation. The specification of the future norm now includes our concept group primitive.

CONCLUSION AND PERSPECTIVES

The different steps of the implementation of a knowledge management platform aiming at the creation, maintenance and integration of a biomedical observation dictionary within the AP-HP have been described in this paper. It was decided to create an interface terminology dedicated to the AP-HP observation description and to maintain its mapping to LOINC. The migration process to the platform forced us to improve the quality of the dictionary data by formulating and enforcing new semantic and syntax rules.

The mapping to LOINC had to be increased to ensure the necessary interoperability allowing the transmission between health institutions. At present, the mapping implementation reaches 40% of the AnaBio codes. The two possible improvement options that have been identified and should be analyzed to enhance the mapping implementation are as follows:

1. to request the creation of all AnaBio codes in LOINC via the RELMA interface;[17]
2. to relax the mapping criteria.

Keeping in mind the specificity of our requirements and the fact that it is impossible for LOINC to create all our codes, we have to find the right balance between these two approaches. The work presented in this paper constitutes a first step towards a wider use of the platform. The next step is the modeling and the integration of the biomedical test prescription data to cover the entire biology information flow within the AP-HP in the context of both patient care coordination and biomedical research.

REFERENCES

1 Cornelis R, Fuentes-Arderiu X, Bruunshuus I, et al. International Union of Pure and Applied Chemistry and International Federation of Clinical Chemistry, Committee on Nomenclature. Properties and Units (C-NPU): properties and units in the clinical laboratories sciences. IX. Properties and units in trace elements (IFCC-IUPAC technical report 1997). *Clin Eur J Clin Chem Clin Biochem* 1997;35:833.

2 EUCLIDES. *Coding system Version 4.01.* Ghent, Belgium: EUCLIDES Foundation International, 1994.

3 Cormont S, Erman A, Burckel Y, et al. Names-lab: un modèle de normalisation des messages de biologie [names-lab : a model for the standardization of biology message exchanges]. *Ann Biol Clin (Paris)* 2002;60:173–81.

4 De Haan JB, et Deom A. [for a universal test directory: a French version of LOINC]. Journées Francophones d'Informatique Médicale, Lille 12–13 mai 2005. *Ann Biol Clin (Paris)* 1999;57:730–6.

5 Forrey AW, McDonald CJ, DeMoor G, et al. Logical observation identifier names and codes (LOINC) database: a public use set of codes and names for electronic reporting of clinical laboratory test results. *Clin Chem* 1996;42:81–90.

6 Huff SM, et al. LOINC (logical observation identifier names and codes. developpement of the logical observation identifier names and codes (LOINC) vocabulary. *JAMIA* 1998;5:276–92.

7 Dolin RH, Huff SM, Rocha RA, et al. Evaluation of a "lexically assign, logically refine" strategy for semi-automated integration of overlapping terminologies. *J Am Med Inform Assoc* 1998;5:203–13.

8 McDonald CJ, Huff SM, Suico JG, *et al*. LOINC, a universal standard for identifying laboratory observations: a 5-year update. *Clin Chem* 2003;49: 624–33.

9 Lin M, Vreeman D, McDonald C, *et al*. Characterization of local LOINC mapping for laboratory tests in three large institutions. *Methods Inf Med* 2011;50:105–14.

10 Baorto DM, Cimino JJ. Combining laboratory data sets from multiple institutions using the logical observation identifier names and codes (LOINC). International Journal of Medical Informatics. 1998;51:29–37.

11 Zunner C, Bürkle T, Prokosch HU. Ganslandt: mapping local laboratory interface terms to LOINC at a German university hospital using RELMA V.5: a semi-automated approach. *J Am Med Inform Assoc* 2013;20:293–7.

12 ASIP-Santé: Agence des Systèmes d'Information Partagés de Santé. See http:// esante.gouv.fr/ (accessed 24 Jan 2013).

13 Landais P, Jais JP, Frutiger P. *Sémantique des classifications et nomenclatures. Informatique et Santé 1*. Springer-Verlag, Informatique et Santé, 1989:211–22.

14 Carayon A, Noe E, Liebard G, *et al*. Informatisation d'un fichier d'analyses biologiques. *Rev Fr Lab Janv* 1992;232:69–76.

15 Cormont S, Buemi A, Horeau T, *et al*. Construction of a dictionary of laboratory tests mapped to LOINC at AP-HP. *AMIA Annual Symposium*, Washington. 2008;1200.

16 LOINC user's guide. June 2012. http://loinc.org/downloads/files/LOINCManual.pdf (accessed 25 Apr 2013).

17 Daniel C, Buemi A, Mazuel L, *et al*. Functional requirements of terminology services for coupling interface terminologies to reference terminologies. *Studies Health Technol Inform* 2009;150:205–9.

18 Kohlhase A, Kohlhase M. Compensating the Semantic Bias of Spreadsheets. In: Baumeister J, Atzmüller M, eds. *Wissens-und Erfahrungsmanagement LWA (Lernen, Wissensentdeckung und Adaptivität) Conference Proceedings*; 2008:448.

19 Mondeca. Intelligent topic manager. http://www.mondeca.com/ (accessed 24 Jan 2013).

20 Si LE, Obrien A, Probets S. Integration of distributed terminology resources to facilitate subject cross-browsing for library portal systems. In: Aslib Proceedings, vol. 62, Emerald Group Publishing Limited, 2010:415–27.

21 Isaac A, Schlobach S, Matthezing H, *et al*. Integrated access to cultural heritage resources through representation and alignment of controlled vocabularies. *Libr Rev* 2009;57:187–99.

22 Macgregor G, Joseph A, Nicholson D. A SKOS core approach to implementing an m2m terminology mapping server. *International Conference on Semantic Web and Digital Libraries (ICSD Proceedings)*; 2007:21–3.

23 Fung K, Bodenreider O. Utilizing the UMLS for semantic mapping between terminologies. *AMIA Annual Symposium Proceedings*, 2005:266.

24 Stenzhorn H, Beißwanger E, Schulz S. Towards a top-domain ontology for linking biomedical ontologies. *Stud Health Technol Inform* 1999;2007:1225.

25 Gangemi A, Guarino N, Masolo C, *et al*. *Sweetening ontologies with DOLCE. Lecture notes in computer science*. Berlin, Heidelberg: Springer, 2002:166–81.

26 Picca D, Gliozzo A, Gangemi A. LMM: an OWL-DL metamodel to represent heterogeneous lexical knowledge. *Proceedings of LREC*, Marrakech, Morocco; 2008.

27 Pathak J, Solbrig H, Buntrock J, *et al*. LexGrid: a framework for representing, storing, and querying biomedical terminologies from simple to sublime. *J Am Med Inform Assoc* 2009;16:305–15.

28 Fernández-López M, Gómez-Pérez A. Overview and analysis of methodologies for building ontologies. *The Knowledge Engineering Review* 2002;17:129–56.

29 OWL (web ontology language). See http://www.w3.org/TR/owl-features/ (accessed 24 Jan 2013).

30 Miles A. SKOS: requirements for standardization. *DC-2006: Proceedings of the International Conference on Dublin Core and Metadata Applications*; 2006:55–64.

31 BS8723. Structured vocabularies for information retrieval, Part 4: Interoperability between vocabularies. British Standards Institution, 2008.

32 Cormont S, Vandenbussche P-Y, Buemi A, *et al*. Implementation of a platform dedicated to the biomedical analysis terminologies management. *Am Med Inform Assoc (AMIA) Ann Symp* 2011;2011:1418–28.

33 Macary F, Spronk R. IHE Laboratory Profiles. 2008.

34 Regenstrief Institute. See http://www.regenstrief.org/ (accessed 24 Jan 2013).

35 Prud'hommeaux E, Seaborne A. SPARQL query language for RDF. W3C Recommendation 15 January 2008. World Wide Web Consortium. http://www.w3. org/TR/rdf-sparql-query/ (accessed 25 Apr 2013).

36 InterSTIS project. French multiterminologies server project. http://www.interstis.org/ (accessed 24 Jan 2013).

37 Eurovoc thesaurus editing. Multilingual, multidisciplinary thesaurus covering the activities of the EU. See http://europa.eu/eurovoc/ (accessed 24 Jan 2013).