

UNIVERSITÀ DELLA CALABRIA



UNIVERSITA' DELLA CALABRIA

Dipartimento di Ingegneria Informatica, Modellistica, Elettronica e Sistemistica

Dottorato di Ricerca in

**Information and Communication Engineering for Pervasive Intelligent
Environments**

CICLO

XXIX

**Un support à la validation des codes LOINC® : une application de
TAL pour vérifier l'exactitude des données**

Settore Scientifico Disciplinare M-STO/08

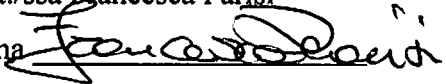
Coordinatore: Ch.mo Prof. Felice Crupi

Firma  _____

Supervisore/Tutor: Ch.mo Prof. Roberto Guarasci

Firma  _____

Dottoranda: Dott./ssa Francesca Parisi

Firma  _____

Index

Index	2
Liste des figures.....	4
Liste des tableaux	6
Tableau des sigles.....	7
Abstract.....	12
Introduction	14
Chapitre I.....	22
1.1 Les TIC au service de la santé	22
1.1.1 Le contexte Européen	25
1.1.2 Le patient européen	29
1.1.3 Le contexte italien	33
1.2 Interopérabilité des données sanitaires : les terminologies de référence ..	38
Chapitre II.....	41
2.1 Les systèmes de noms et de codes : LOINC.....	41
2.1.1 Le premier attribut : « Le Composant »	44
2.1.2 Le deuxième attribut : « La Grandeur »	46
2.1.3 Le troisième attribut : « Le Temps »	47
2.1.4 Le quatrième attribut : « Le milieu biologique ».....	48
2.1.5 Le cinquième attribut : « L'Échelle ».....	49
2.1.6 Le sixième attribut : « La technique »	51
2.2 Le processus de traduction : « LOINC Italia »	52
2.3 L'outil pour le mappage de codes : RELMA.....	55
Chapitre III	58
3.1 L'analyse du contexte et la problématique traitée	58
3.2 Le système des règles linguistiques : l'application NooJ	62
3.2.1 Les applications de TALN.....	69
3.3 Les logiciels de traitement automatique des langues.....	72

3.3.1	Le choix du logiciel NooJ.....	75
Chapitre IV	79
4.1	Le corpus.....	79
4.2	Les dictionnaires	82
4.2.1	Le dictionnaire électronique : « Les Analytes ».....	83
4.2.2	Le dictionnaire électronique : « Les attributs LOINC ».....	103
4.3	Les Grammaires flexionnelles	112
4.4	Grammaires hors contexte : « LOINC.nog ».....	116
4.5	Les contraintes contextuelles : les niveaux d'exactitude.....	124
4.5.1	Les contraintes d'aide aux experts : les niveaux « HELP »	131
Chapitre V	135
5.1	Les résultats obtenus : la formalisation de la connaissance.....	135
5.1.1	Annotation sémantique : la validation des codes	137
5.1.2	Annotation sémantique : les niveaux d'exactitude.....	142
5.1.3	Annotation sémantique : la ressource XML.....	146
5.2	L'extraction d'informations	148
5.2.1	L'extraction d'informations : un système questions - réponses	148
5.2.2	L'extraction d'informations : les niveaux d'exactitude.....	154
5.3	Résumé des résultats	159
Conclusions	163
Bibliographie	168
Annexe 1 – Les dictionnaire « Analiti.dic ».....		171
Annexe 2 – Dictionnaire « Elements_LOINC.dic ».....		188
Annexe 3 – Les grammaires d'extraction d'informations		204

Liste des figures

Figure 1 - Écran de recherche du logiciel RELMA.....	56
Figure 2 - Écran de support au mappage des examens locaux	57
Figure 3 - Le corpus	81
Figure 4 - Le dictionnaire « Analiti.dic » : les ALU simples.....	85
Figure 5 - Le dictionnaire « Analiti.dic » : les ALU composées.....	88
Figure 6 - Les morphèmes : suffixe "emia".....	92
Figure 7 - Les "Acides"	95
Figure 8 - Les acronymes	97
Figure 9 - Les synonymes.....	98
Figure 10 - La catégorie des globules blancs	100
Figure 11 - La catégorie des globules rouges	101
Figure 12 - La catégorie des plaquettes	101
Figure 13 - Les acronymes pour la catégorie des globules blancs	101
Figure 14 - Les acronymes pour la catégorie des globules rouges.....	101
Figure 15 - Les acronymes pour la catégorie des plaquettes.....	102
Figure 16 - Les expressions temporelles	102
Figure 17 - la catégorie des milieux biologiques.....	104
Figure 18 - La catégorie d'échelles de mesure.....	106
Figure 19 - La catégorie des méthodes.....	107
Figure 20 - La catégorie des unités de mesure	108
Figure 21 - Les adjectifs	109
Figure 22 - Les codes	110
Figure 23 - Les axes LOINC décrits en RELMA.....	111
Figure 24 - La propriété morphologique "SUFF"	113
Figure 25 - Les grammaires flexionnelles pour les morphèmes "emia" et "uria" .	113
Figure 26 - La grammaire "LOINC.nog"	117
Figure 27 - Structure de la grammaire "LOINC.nog"	117
Figure 28 - La grammaire "LOINC_ELEMENTS"	118
Figure 29 - La grammaire "COMPONENT"	119
Figure 30 - La grammaire "SYSTEM_AGG"	119
Figure 31 - La grammaire "PROPERTY"	120
Figure 32 - La grammaire "METHOD"	121
Figure 33 - La grammaire "SYSTEM"	122
Figure 34 - La grammaire "SCALE"	123

Figure 35 - Exemple d'annotation	124
Figure 36 - Les contraintes pour la grammaire « LOINC »	125
Figure 37 - Les contraintes "HELP"	133
Figure 38 - Résultats de l'annotation sémantique: les niveaux d'exactitude.....	138
Figure 39 - Résultats de l'annotation sémantique: les contraintes HELP.....	138
Figure 40 - La fonction de recherche dans RELMA: la chaîne de termes locales	139
Figure 41 - La fonction de recherche dans RELMA: le résultat de l'annotation ..	140
Figure 42 - La fonction de recherche dans RELMA: les expressions temporelles locales	140
Figure 43 - Ensemble des codes obtenu en utilisant les formes locales.....	141
Figure 44 - Ensemble des codes obtenu en utilisant la chaîne de termes annotée	141
Figure 45 - Les résultats obtenus par les contraintes contextuelles.....	143
Figure 46 - L'utilité des contraintes HELP.....	145
Figure 47 - La ressource XML	146
Figure 48 - L'élément XML: LOINC	147
Figure 49 - Les concordances des analytes	150
Figure 50 - Les concordances des acronymes	150
Figure 51 - Les concordances des synonymes	150
Figure 52 - Les concordances des morphèmes.....	151
Figure 53 - Les examens décrits seulement par le composant	152
Figure 54 - Les concordances pour la méthode.....	152
Figure 55 - Les concordance des adjectifs	153
Figure 56 - Les examens appartenant au niveau 1 d'exactitude	155
Figure 57 - La grammaire de sélection des examens qui s'arrêtent au niveau 1 ...	155
Figure 58 - Les examens qui s'arrêtent au niveau 1 d'exactitude	156
Figure 59 - La grammaire de sélection des examens appartenant au niveau 3 HELP	156
Figure 60 - Le niveau d'exactitude 3 help	157
Figure 61 - Examens du niveau 2 récupérés par la contrainte HELP.....	158
Figure 62 - L'examen appartenant au niveau 4 HELP.....	161

Liste des tableaux

Tableau 1 - Exemple de test codifié et structuré avec les six axes LOINC.....	43
Tableau 2 - Valeurs associés à l'attribut « Temp »	48
Tableau 3 – Valeurs associés à l'attribut « Milieu Biologique »	49
Tableau 4 - Exemple de valeurs associé à l'attribut « Technique »	51
Tableau 5 - Flexions invariantes.....	114
Tableau 6 - Flexions des mots composées (2 termes)	115
Tableau 7 - Flexions des mots composées (3 termes)	115
Tableau 8 - Flexions génériques.....	115
Tableau 9 - Flexions morphèmes en "emia" et "uria"	116
Tableau 10 - Les variables de la grammaire "LOINC"	125
Tableau 11 - Les variables pour les contraintes "HELP"	132
Tableau 12 - Synthèse résultats: catégories identifiées	159
Tableau 13 - Synthèse résultats: les niveaux d'exactitude.....	161

Tableau des sigles

Sigles et Acronymes	Définition
AgID	Agenzia per l'Italia digitale
AIC	Autorizzazione Immissione in Commercio
ALU	Atomic Linguistic Unit
ASIP	Agence des Systèmes d'Information Partagés de Santé
ATC	Anatomical Therapeutic Chemical classification system
AVM	Attribute Value Matrix
CDA	Clinical Document Architecture
CEF	Connecting Europe Facility
CI-SIS	Cadre d'interopérabilité des Systèmes d'Information de Santé
CNR	Consiglio Nazionale delle Ricerche
CREOLE	Collection of REusable Objects for Language Engineering

DL	Décret Législatif
DMP	Dossier Médical Personnel
DPCM	Decreto del Presidente del Consiglio dei Ministri
EEE	Islande, Liechtenstein et Norvège
epSOS	Smart Open Services for European Patients
eSens	Electronic Simple European Networked Services
EXPAND	Expanding Health Data Interoperability Services
FSE	Fascicolo Sanitario Elettronico
GATE	General Architecture for Text Engineering
HCP	Health Consumer Powerhouse
HL7	Health Level Seven
HPSG	Head-driven Phrase Structure Grammar
HTML	Hyper Text Markup Language

ICAR	Interoperabilità e Cooperazione Applicativa in rete tra le Regioni
ICD-9-CM	International Classification of Diseases – 9th Revision, Clinical Modification
IHTSDO	International Health Terminology Standards Development Organization
IIT	Istituto di Informatica e Telematica
InFSE	Infrastruttura tecnologica del Fascicolo Sanitario Elettronico
INPC	Indiana Network for Patient Care
INSERM	Institut national de la santé et de la recherche médicale
IRDES	Institut de Recherche et Documentation en Économie de la Santé
LFG	Lexical Functional Grammar
LOINC	Logical Observation Identifiers Names and Codes
NHS-CRS	National Health Service Care Record Service
NLTK	Natural Language Toolkit
OMS	Organisation Mondiale de la Santé

PR	Processing Ressources
RELMA	Regenstrief LOINC Mapping Assistant
RI	Regenstrief Institut
RTF	Rich Text Format
SFIL	Société Française d'Informatique de Laboratoire
SGML	Standard Generalized Markup Language
SIS	Systèmes d'Information de Santé
SNOMED CT	Systematized Nomenclature of Medicine – Clinical Terms
TALN	Traitement automatique des langues naturelles
TIC	Technologies de l'Information et de la Communication
UE	Union Européenne
UHC	United Healthcare
WHO	World Health Organization

XFST	Xerox Finite-State Tools
XML	eXtensible Markup Language

Abstract

Data interoperability in healthcare processes is necessary to exchange meaningful information among healthcare professionals and Institutions. Classification and coding systems such as the LOINC® (Logical observation identifiers names and codes) international standard for names and codes allows for information exchange while maintaining the intrinsic semantic value conveyed by the code unaltered. The methodology presented aims to give a support for data semantic interoperability in the e-Health domain.

The analyzed corpus contains laboratory databases of clinical and biological tests expressed in local linguistic forms and their mappings to the official LOINC codes (Italian version). In particular, the work presented concerns the construction of NooJ syntactical grammars for recognitions of local linguistic forms and detection of data correctness level for tests. All clinical laboratories use their local linguistic forms to identify specific clinical observations. Each biological or clinical test is described using terms, codes or acronyms valid only in the local context thus creating issues in data exchange and needing particular attention during the mapping operations to the official international standard LOINC definitions and codes.

The automatic recognition of local linguistic forms would represent an important step in helping biological and clinical laboratories to use the standard. The aim of this work is building a NooJ grammars system to allow the recognition of the local linguistic forms and detect specific correctness levels. The LOINC international standard describes clinical and biological tests through the combination of 6 parameters where each one of these has different linguistic variability. This data structure makes the tests identification clear and unambiguous. However, in local databases, clinical tests and observations are described with a combination of elements that often are not sufficient to uniquely identify the test and expressed in local linguistic forms (e.g.: names, suffixes, acronyms etc.) hard to be handled

during the mapping process. When the experts map the clinical observations to LOINC, they use their local forms causing errors and noise in the data. Sometimes local forms are so idiosyncratic that they cannot be retrieved in the LOINC international database. Therefore, the association (local tests-LOINC codes) resulting from mapping operations, has to be controlled and validated by the LOINC experts that give help during this important process. The realized set of hierarchical grammars aims to support LOINC experts during the validation process of LOINC codes associated to the clinical and biological tests expressed in local linguistic forms. The realized syntactical grammars can recognize the local linguistic forms as well as determine data accuracy.

In particular, for each analyzed biological or clinical test a matching between local linguistic forms and the corresponding property associated to the LOINC code is effected using the NooJ contextual constraints that identify the threshold of correctness. The created grammars allow the automatic collection of linguistic forms stored in local databases and give the possibility to extract important information regarding data correctness.

Introduction

L'objectif du travail de thèse présenté concerne la construction d'un système de règles linguistiques en support de la validation des opérations de mappage conduites dans le cadre d'interopérabilité des données sanitaires au niveau national et international. L'application de TALN (Traitement automatique des langues naturelles) développée a vu la réalisation d'un ensemble de règles d'annotation sémantique et de grammaires qui, de façon automatique, permettent de vérifier l'exactitude des données appartenant au corpus analysé et l'extraction des informations pertinentes aux objectifs fixés. En ne se basant que sur des règles linguistiques, l'application réalisée est capable de comprendre la sémantique liée aux formes linguistiques caractérisant un domaine extrêmement spécifique et de vérifier le niveau d'exactitude des associations de ces formes avec leur code identifiant.

Le travail suivant s'inscrit dans le domaine de la santé numérique et propose une contribution pour améliorer l'interopérabilité des données concernant les examens biologiques-médicaux effectués par les patients pendant leurs parcours de soins. Un ensemble de règles capables de vérifier l'exactitude des opérations de mappage entre les tests biologiques-médicaux et le standard international de noms et de codes LOINC® (*Logical Observation Identifiers Names and Codes*)¹ ont été réalisées pour permettre la reconnaissance automatique des formes linguistiques locales utilisées par les laboratoires et supporter les experts pendant les opérations de validation des codes associés. Le codage des tests médicaux se déroule par les laboratoires biologiques opérants dans le secteur public ou privé en vue de rendre leurs informations interopérables au niveau national et international. Encadré dans le domaine de l'e-santé, ce travail de thèse a vu la définition de règles d'annotation sémantique pour la reconnaissance de formes linguistiques locales utilisées pour

¹ LOINC: <loinc.org>. Date de la dernière consultation 18/12/2016.

décrire les caractéristiques des examens et leur association avec les termes et les propriétés propres au standard international LOINC.

Une importante contribution dans la réalisation d'un système automatique d'aide à la validation des opérations de mappage a été donnée par la définition de ces règles qui, à partir des résultats obtenus par l'annotation sémantique et par une comparaison entre ceux-ci et les propriétés associées à chaque code, déterminent pour chaque examen le niveau d'exactitude du code associé. Les données de laboratoire sont rendues interopérables par cette opération de mappage, qui permet à chaque patient de bénéficier d'un code univoque pour ses tests cliniques pendant tout son parcours médical. Le code associé aux examens permettra de garder les informations concernant la typologie d'observation durant tout le processus de transmission ou intégration des données collectées dans le dossier médical de chaque patient.

Le développement de la technologie au cours des dernières années a contribué à l'évolution du processus de gestion et planification des actions dans le domaine de la santé en obtenant des améliorations importantes au niveau du management des dépenses publiques ainsi que sur l'échange d'informations entre le système de santé² et les patients. A cet égard, la plupart des Pays Européens ont adoptés un véritable changement de leurs systèmes sanitaires en donnant aux citoyens la disponibilité des informations concernant leur parcours médical à tout moment en supprimant les difficultés liées surtout à leur récupération. On a donc constaté que le concept de santé publique a été transformé par une conception de santé qui, grâce à l'adoption de solutions technologiques capables d'intégrer et de résumer les données médicales des patients, offre aux citoyens des services de santé plus efficaces. Le but de ce travail d'intégration est d'améliorer la transmission et le partage des données autant entre les différentes Institutions qu'entre les Institutions

² Selon l'Organisation Mondiale de la Santé un système de santé comprend toutes les organisations, institutions et ressources dont le but principal est d'améliorer la santé. Le système de soins de santé est l'ensemble des institutions, des personnes et des ressources qui participent à la prestation de soins de santé. Une institution sanitaire est une organisation à caractère médicale et sociale dont la fonction consiste à assurer à la population des soins médicaux «complets», curatifs et préventifs et il peut être aussi un centre d'enseignement de la médecine et des recherches biomédicales.

et les patients. Le support donné par la technologie aux services de santé publique a contribué à la diffusion de cette nouvelle conception de santé à laquelle a été attribué le terme de « *e-santé* » et représente aujourd'hui un des défis les plus importants du contexte sanitaire européen. Le terme de « e-santé » (e-health en anglais) désigne les domaines où les technologies de l'information et de la communication (TIC) sont mises au service de la santé. Selon l'Organisation mondiale de la santé (OMS)³, la e-santé se définit comme « les services du numérique au service du bien-être de la personne » c'est-à-dire comment les TIC sont appliquées au domaine de la santé et du bien-être. La réduction des dépenses de santé et l'amélioration de la prise en charge des patients représentent des enjeux particulièrement considérés. La mise à disposition d'informations fiables pour les professionnels et les patients, l'utilisation du potentiel des nouvelles technologies pour améliorer les services publics et la surveillance ainsi que la promotion des politiques publiques en matière de santé, représentent les points principaux sur lesquels l'e-santé fonde ses propres bases.

Ceci étant, l'interopérabilité de données autant entre les systèmes qu'entre les acteurs opérants dans le processus d'échange d'informations ainsi que la transmission de celles-ci en utilisant un protocole informatique commun sont devenues indispensables pour atteindre ces objectifs.

Le concept d'interopérabilité est représenté par la capacité qu'un produit ou un système informatique possède pour communiquer avec d'autres systèmes ou produits en permettant de fonctionner ensemble⁴. Pour obtenir cette communication, il est nécessaire d'utiliser un langage commun qui peut permettre la correcte transmission des données, parfois de nature strictement confidentielle, entre différents opérateurs en s'assurant de garder la même valeur sémantique pendant tout le processus de transmission. À la lumière de ces éléments,

³ OMS: <www.who.int>. Date de la dernière consultation 18/12/2016.

⁴ IRDES – INSTITUT DE RECHERCHE ET DOCUMENTATION EN ECONOMIE DE LA SANTE: «*E-santé: télésanté, santé numérique et santé connectée*», Septembre 2016, disponible au lien suivant: <<http://www.irdes.fr/documentation/syntheses/e-sante.pdf>>. Date de la dernière consultation 18/12/2017

vocabulaires spécialisés, lexicques, systèmes de classification, taxonomies et standards ont été réalisés par ceux qui travaillent dans les secteurs spécialisés, comme celui biologique-médical, pour rendre possible le partage des données au niveau national et international. Dans cette catégorie de secteurs, l'interopérabilité sémantique concerne toutes les opérations d'accord sur les mots et les termes caractérisant une activité donnée. L'objet de l'interopérabilité sémantique dans le domaine de l'e-santé est de définir le vocabulaire commun qui sera employé dans les systèmes pour nommer telle ou telle maladie ainsi qu'un test clinique spécialisé ou une observation de laboratoire spécifique. L'interopérabilité sémantique permet, par exemple, d'avoir un identifiant univoque pour le même concept au-delà des différences linguistiques. De cette façon, la même information peut être échangée entre systèmes qui utilisent différents formats ainsi que différentes langues.

De manière analogue au langage naturel, lequel permet aux humains de communiquer en s'appuyant sur un vocabulaire et une syntaxe commune, le langage utilisé pour les échanges de données informatisées s'appuie sur des structures d'informations qui en constituent la syntaxe et sur des vocabulaires contraints qui fournissent les mots à placer dans ces structures pour composer des phrases⁵. Les systèmes de santé qui gère l'échange des informations sur le suivi des patients et leurs données privées doivent utiliser un langage commun pour transmettre efficacement le message et pour réaliser des traitements fiables à partir de l'information échangée.

Ce phénomène de transmission purement télématique a conduit vers un processus de dématérialisation des documents sanitaires (Ex.: diagnostics par les images, rapports médicaux, résultats des examens, etc.) en considérant tout particulièrement leur organisation et conservation. Pour suivre un patient correctement, les documents concernant son parcours médical doivent être consultables en tout lieu et à tout moment en assurant un égal accès au même contenu informatif.

⁵ Agence française de la santé numérique: <esante.gouv.fr>. Date de la dernière consultation 20/12/2017

Dans le contexte italien, le «Fascicolo Sanitario Elettronico» (FSE)⁶, l'équivalent du Dossier Médical Personnel (DMP)⁷ français, représente l'ensemble des documents sanitaires qui donnent une synthèse globale des étapes les plus significatives suivies pendant le parcours clinique-médical du patient. Le FSE représente le moyen le plus important adopté pour échanger les informations en utilisant des standards spécifiques qui permettent une communication claire et dépourvue d'ambiguïtés. Malgré les caractéristiques précieuses et les objectifs admirables de cette initiative, beaucoup d'obstacles n'ont pas encore été dépassés.

Dans le travail de thèse présenté, nous nous référons principalement aux problématiques liées au mappage vers le standard LOINC des observations de laboratoire ainsi qu'aux limitations existantes pour rendre ces informations interopérables. Comme précédemment indiqué, l'intégration et la communication des données provenant de systèmes d'informations différents sont des opérations très ardues. Le mappage entre la terminologie locale et les vocabulaires standardisés garantit l'interopérabilité sémantique nécessaire pour rendre les informations des patients prêtes au partage et à l'intégration étant identifiables efficacement. Actuellement en Italie, les laboratoires biologiques médicaux utilisent leurs propres systèmes de codes et leurs formes linguistiques locales pour nommer les observations effectuées. C'est pour cela que l'échange d'informations se manifeste très difficilement au sein des mêmes laboratoires qui doivent nécessairement communiquer au niveau national et international pour garantir une réutilisation possible de leurs données. Le mappage des termes et des codes utilisés au niveau local avec ceux qui appartiennent aux standards spécifiques est, donc, nécessairement demandé pour atteindre ces objectifs.

A ce jour, le langage spécialisé qui s'adapte le mieux aux exigences exposées est le standard des noms et des codes LOINC, qui grâce aux descriptions très détaillées

⁶ Decreto Legge 18 ottobre 2012, n. 179 recante "Ulteriori misure urgenti per la crescita del Paese" (c.d. Decreto Crescita 2.0), coordinato con la legge di conversione 17 dicembre 2012, n. 221 e pubblicato in Gazzetta Ufficiale 18 dicembre 2012, n. 294.

⁷ Le *Dossier Médical Personnel* a été institué par la Loi n° 2004-810 du 13 août 2004 relative à l'assurance maladie.

des examens et de leur usage au niveau international représente en absolu la solution la plus adaptée.

Le standard LOINC, débuté en 1994, est un système de codage pour les noms des analyses biologiques développé au sein du Regenstrief Institut (RI), situé à Indianapolis et devenu rapidement leader dans le secteur.

Une description d'analyse biologique se décompose selon LOINC en six axes: l'analyte (*Component*) qui représente la substance mesurée; la propriété (*Property*) mesurée (Ex. :concentration massique); la temporalité (*Timing*) de l'observation, pour indiquer à quel point dans le temps l'analyse a été effectuée; le milieu (*System*) pour indiquer le lieu de détermination de l'élément à doser (ex.: urines, sang, plasma etc.); l'échelle (*Scale*) pour préciser dans quelle mesure le test a été exécuté (quantitative, ordinale, nominale, etc.); la technique (*Method*) identifie la technique avec laquelle les résultats ont été obtenus. L'ensemble des six champs est associé à un et un seul code décrivant le test effectué et la nature de l'analyse.

Le niveau de détail qui caractérise le standard permet la communication précise et l'interconnexion entre les systèmes distincts en facilitant la récupération des documents concernant les examens extrêmement importants dans le suivi du patient. La granularité de sa structure permet l'utilisation des données codifiées dépourvues d'ambiguïté tant dans les documents cliniques structurés selon le standard CDA (Clinical Document Architecture) que dans l'échange des messages en format HL7⁸.

Néanmoins, la nomenclature minutieuse et l'élevé niveau de détail du standard LOINC si d'une part ils représentent un avantage pour la description des examens d'une autre part ils demandent aux experts un grand effort en terme d'association des informations locales, souvent exprimées en forme implicite, avec les paramètres et les codes du standard officiel.

La standardisation de formes locales et leur insertion dans l'ensemble des formes reconnues par le logiciel de mappage reste encore une étape copieuse si on

⁸ HL7 est l'acronyme de Health Level Seven.: <www.hl7.org/implement/standards/>. Date de la dernière consultation 20/12/2016.

considère la structure détaillée autant que variée des axes LOINC. Les axes Propriété, Temporalité et Échelle admettent un ensemble limité de valeurs normalisées. Les axes Analyte, Milieu et Technique, au contraire, présentent un ensemble de valeurs beaucoup plus larges, qui différencient plus en détail l'analyse décrite. L'association de chaque examen est effectuée par le logiciel officiel RELMA®⁹ qui gère les informations du standard et permet la recherche des codes à associer aux analyses de façon automatique partant de bases de données locales. Le mappage des analyses devient difficile à partir du moment où le logiciel n'est pas capable de reconnaître les formes linguistiques utilisées dans la base de données locale et, par conséquent, aucun résultat pertinent aux recherches effectuées ne sera pas produit pendant cette opération. En Italie, le groupe « LOINC Italia »¹⁰ du Conseil National des Recherches (CNR – Consiglio Nazionale delle Ricerche)¹¹ s'occupe de supporter le mappage des formes locales avec les codes du standard et de valider les associations effectuées par les experts de laboratoire. Le fait que le logiciel ne reconnaisse pas les formes linguistiques appartenant au lexique normalement utilisé par les experts de laboratoire, provoque une importante quantité d'erreurs possibles, en créant nombreux retards durant les opérations de validation. Si d'un côté le logiciel donne un support important pour le mappage des analyses vers la nomenclature LOINC, d'un autre côté son usage, inévitablement lié à la structure du standard, est parfois difficile et conduit les experts vers l'association de codes que ne décrivent pas vraiment les caractéristiques des analyses considérées. L'association aux codes corrects est très complexe étant donné que LOINC différencie les analyses par la combinaison univoque de ses six parties. L'opération de validation et de contrôle sur l'exactitude des codes assignés devient, donc, une phase très importante et cruciale dans le processus considéré.

Le développement de l'e-santé en Italie ainsi que les initiatives et l'impact des TIC dans ce contexte sont présentés dans le premier chapitre. Une attention

⁹ RELMA: <loinc.org/relma>. Date de la dernière consultation: 20/12/2016.

¹⁰ LOINC ITALIA : <www.labdoc.it/loincitalia>. Date de la dernière consultation: 20/12/2016.

¹¹ CNR: <www.cnr.it>. Date de la dernière consultation: 20/12/2016.

particulière sur les terminologies et les classifications utilisées dans le secteur biomédical a été donnée en soulignant leurs fonctionnalités dans le FSE.

Le deuxième chapitre est entièrement dédié au standard LOINC qui décrit ses caractéristiques et ses fonctionnalités plus en détail. La version italienne du standard LOINC est le résultat des activités de traduction et collecte de formes locales ainsi que d'autres éléments linguistiques, exécutées par le CNR pour soutenir les opérations de mappage. Les concepts d'associations et de codage des examens ainsi que les difficultés liées à ce genre d'opérations sont décrites dans la dernière section de ce chapitre.

L'interopérabilité des données sanitaires et la reconnaissance des formes linguistiques locales sont l'objet du troisième chapitre. Un état de l'art sur les travaux linguistiques effectués pour la standardisation de la terminologie locale avec la terminologie LOINC est présenté pour donner une vision d'ensemble sur la problématique à résoudre. Les motivations qui ont conduit vers le choix du logiciel de traitement automatique de langue NooJ¹² sont analysées pour mieux comprendre les avantages dérivants de cette approche.

Le chapitre 4 montre les étapes les plus importantes du travail linguistique réalisé. En particulier, les données, les dictionnaires et la structure des grammaires utilisées pour la construction des règles d'annotation sémantique ainsi que les contraintes contextuelles appliquées aux formes linguistiques objet de l'analyse, sont expliqués pour mieux comprendre les résultats obtenus.

Les résultats obtenus en terme d'informations extraites et concernant les associations « formes locales-codes LOINC » considérées dans le corpus analysé, les perspectives liées au travail réalisé et les limitations rencontrées sont présentées dans la dernière partie.

¹² Nooj Association : <www.nooj-association.org>. Date de la dernière consultation 20/12/2016

Chapitre I

1.1 Les TIC au service de la santé

Le concept de santé a évolué durant ces dernières années. L'application de TIC (Technologies de l'information et de la Communication) aux services sanitaires offerts aux citoyens, a sûrement représenté un des enjeux principaux pour obtenir cette transformation. Par ailleurs, les initiatives adoptées dans le secteur privé ont été souvent plus efficaces étant les processus de décision plus rapides et en ayant plus de possibilités d'investissement.

Le concept de « e-santé » se définit comme l'application des TIC à l'ensemble des activités en rapport avec la santé. Selon la commission européenne le terme « e-santé » désigne l'ensemble des technologies et des services pour les soins médicaux basés sur la santé¹³. Avec le nom d'e-santé, nous nous référons tous les processus d'innovation, objet de nombreux investissements, mises en œuvre afin d'améliorer le parcours des soins des patients grâce au support d'instruments technologiques.

L'introduction des TIC dans le domaine médical suscite de nombreuses attentes. Face au défi démographique, la possibilité d'assurer un meilleur suivi des patients, en particulier des personnes âgées atteintes de maladies chroniques ainsi que de renforcer leur sécurité en assurant aux personnels de santé l'accès à l'information concernant les patients à tout moment et en tout lieu. Offrir des outils de support à la médecine préventive et proactive, faciliter la personnalisation des traitements et donc, d'améliorer leur impact, contribuer à donner au patient un rôle d'acteur dans

¹³ « Les dossiers européens : La e-santé en Europe », Juin 2009. Disponible au site <ec.europa.eu >. Date de la dernière consultation 04/01/2017.

la gestion de sa santé représentent, par ailleurs, des avantages qui l'utilisation de TIC peut apporter au domaine sanitaire.

Il n'y a aucun doute que l'application de TIC apporte au domaine de la santé des avantages multiples mais, cette vision applicative d'ensemble tant au niveau national qu'au niveau international, reste encore un objectif difficile à réaliser.

En effet, les avantages de l'application de TIC au domaine de la santé ne sont pas pleinement exploités, étant encore une approche pas uniforme.

Une importante quantité d'informations est aujourd'hui utilisée et gérée par les systèmes de support aux activités dans le domaine sanitaire. Afin d'ajouter une valeur stratégique et donner la capacité de supporter la recherche de solutions aux problèmes cliniques qui se présentent, les informations et les données à caractère sanitaire ont la nécessité d'être recueillies et interprétées par des outils ad-hoc. La possibilité de ne disposer pas de ce genre d'informations au moment et au lieu demandé, reste encore un risque en croissance, en considérant la multitude de systèmes utilisés pour suivre les patients pendant leur parcours médical. La communication correcte entre les différents formats utilisés et une interprétation adéquate et capable de résumer les étapes les plus importantes de soins du patient reste encore le défi central. Un système de stockage automatique des données dérivées à partir de différentes sources pour être exhaustif doit fournir au personnel de santé les instruments nécessaires pour déduire des connaissances et donner un support au niveau décisionnel. Les institutions sanitaires stockent leurs données le plus souvent en format pas structuré en utilisant des codes et des noms utilisés au niveau local et beaucoup des fois imposée par le gouvernement pour effectuer déclarations financières. Les codes et les noms utilisés sont, donc, valide seulement pour une structure ou ensemble des structures particulier en créant des gros obstacles pour la réalisation d'un système sanitaire que soit globalement interopérable. Dans ce sens, l'application de TIC est très importante mais ce n'est pas le seul aspect qui compte. En effet, ce qui est nécessaire premièrement est une représentation des informations dans une forme qui soit compréhensible aux machines et réutilisable pour des élaborations futures et surtout avec une

sémantique commune et partagée. Pour cette motivation, ontologies, vocabulaires et terminologies spécialisées sont devenus particulièrement importants pour adresser la réutilisation des données puisqu'ils permettent la représentation des informations en utilisant des termes contrôlés [6]. Dans cette phase de passage au numérique les systèmes de codage ainsi que de classification et l'utilisation de terminologies contrôlées jouent un rôle d'importance fondamentale. Les systèmes d'information de santé (SIS) sont complexes, hétérogènes et sont rarement interopérables au niveau sémantique.

L'interopérabilité sémantique permet le partage et l'exploitation du sens des données des SIS - faits cliniques, décisions, activités - dès lors qu'elles sont représentées dans un format interprétable par la machine. Dans les domaines spécifiques, tel que le domaine de la santé, l'interopérabilité sémantique se fonde sur des référentiels partagés sous la forme de modèle d'information et des terminologies, développés par des organismes de standard internationaux et, ces derniers dépendent strictement de la capacité du SIS à intégrer ces référentiels pour permettre l'adoption des solutions d'information clinique qui facilitent l'interaction homme-machine [4].

Les structures et les systèmes de codage d'information clinique utilisés par les SIS, ne sont mis en œuvre qu'au niveau spécifique et local, sans un réel partage informatif même avec les établissements qui utilisent des SIS du même éditeur. Cette condition d'isolation, dépend de la capacité d'une structure déterminée d'utiliser les systèmes de codage et de classification de l'information afin de partager leur contenu sémantique au-delà de leur même organisation.

En général, l'utilisation et surtout la correcte utilisation de systèmes de codage et de classification des données de santé représentent aujourd'hui l'obstacle le plus difficile à résoudre. Due au niveau du détail de l'information, qui varie en basant de la profession de santé, de la spécialité ou encore du mode d'application, et en considérant que l'usage de ces systèmes n'est pas encore complètement obligatoire, l'application au niveau globale reste encore le défi le plus important. Il faut considérer les patients comme des patients globaux faisant partie d'un système de

santé qui ne s'arrête pas à la structure spécifique mais qu'est capable de partager les informations de soins avec une complexité des structures au niveau national ainsi que au niveau international. Ce concept de partage, rend l'échange d'informations plus facile autant au niveau géographique (en considérant les grosses distances) qu'au niveau démographique (en considérant les personnes les plus âgées souvent atteintes de maladies chroniques).

L'e-santé, donc, améliore la qualité des soins fournis par les professionnels de la santé puisque elle contribue à rendre plus facile l'accessibilité aux informations en tout lieu et à tout moment. Par ailleurs, la possibilité de partager les données sanitaires permet d'avoir une synthèse qui puisse produire des informations précises et pas redondantes, en ayant, donc, une amélioration en termes de temps et de frais d'échange.

1.1.1 Le contexte Européen

Les stratégies d'e-Gouvernement de l'Union Européenne au cours de ces 15 dernières années ont été mises en œuvre pour donner la possibilité aux États membres de définir leur offre de services technologiques dans différents domaines, comme par exemple le domaine sanitaire. Les États membres de l'Union Européenne sont responsables de la définition de leurs politiques de santé et leur offre de services de soins ; le rôle de l'Union Européenne est d'encourager la coopération entre eux sur les différents sujets touchés par les initiatives proposées.

En 2000, l'initiative « eEurope 2002¹⁴ » a été approuvée avec l'objectif d'utiliser les avantages des nouvelles TIC pour créer une Europe plus compétitive et dynamique au niveau mondiale. Elle proposait d'améliorer l'accès au réseau

¹⁴ « Piano d'azione eEurope 2002 », disponible sur le site <www.europafacile.net/Formulari%5CPOLITICHE%5CSociet%C3%A0Info%5CeEurope%5CPianoEurope.pdf>. Date de la dernière visite : 04/01/2017.

internet et utiliser au maximum son potentiel pour encourager l'innovation et l'offre de services numériques.

En 2002, la Commission Européenne a établi le plan d'action « eEurope 2005 »¹⁵ qui a succédé au plan d'action 2002 qui était surtout axé sur l'extension de la connectivité internet en Europe. Le nouveau plan d'action, approuvé par le Conseil européen de Séville en juin 2002, vise à traduire cette connectivité par un accroissement de la productivité économique et une amélioration de la qualité et de l'accessibilité des services au profit de l'ensemble des citoyens européens, en s'appuyant sur une infrastructure large bande sécurisée et disponible au plus grand nombre. La disponibilité en ligne des services publics, l'accessibilité la plus facile pour les utilisateurs, surtout pour les personnes avec besoins spécifiques (Ex.: personnes avec un handicap particulier), ont représenté les buts les plus importants des actions successives.

En 2004 la Commission Européenne a établi le premier plan d'action spécifique pour régler les opérations concernant les services de santé dans l'ensemble des États membres. Le « Plan d'action européen 2004-2010 en faveur de l'environnement et de la santé »¹⁶ vise à fournir à l'Union Européenne (UE) des informations fiables sur cet impact et à renforcer la coopération entre les différents acteurs intervenant dans les domaines de l'environnement, de la santé et de la recherche. À cette occasion, le terme « e-santé » (en anglais e-Health) a été définie comme toutes les TIC au service de la santé. Il se réfère, en effet, aux logiciels et aux services technologiques qui peuvent améliorer les diagnostics médicaux, les traitements de soin ainsi que le contrôle et la gestion des informations. Le but le plus important du plan d'action sur l'e-santé concerne l'utilisation du potentiel des TIC et des applications en ligne pour répondre aux nécessités du secteur sanitaire, construire une structure de données capable de préserver la vie privée des patients, standard pour l'identification univoque des données sanitaires et pour l'échange de

¹⁵ Description du plan d'action disponible sur le site : <<http://eur-lex.europa.eu/legal-content/FR/TXT/?uri=uriserv:l24226>>. Date de la dernière visite : 04/01/2017.

¹⁶ Description disponible sur le site : <eur-lex.europa.eu/legal-content/FR/TXT/?uri=uriserv%3A128145>. Date de la dernière consultation : 04/01/2017.

messages électroniques contenant des informations strictement confidentiels, afin de garantir leur interopérabilités tel au niveau technologique que sémantique.

En Mars 2010, le Plan stratégique Europe 2020 a été lancé en succession à la stratégie de Lisbonne¹⁷ mise en place durant la décennie 2000. La Stratégie Europe 2020 entend mieux s'adapter aux réalités de l'espace européen, et vise à mettre en place une croissance « intelligente, durable et inclusive ». L'objectif de cette initiative est d'achever l'Espace européen de la recherche, de développer un agenda stratégique de recherche axé sur des défis tels que la sécurité énergétique, les transports, le changement climatique et l'efficacité des ressources, la santé et le vieillissement, des méthodes de production et un aménagement du territoire respectueux de l'environnement, et de renforcer la programmation conjointe avec les États membres. Dans ce cadre, ce nouveau plan d'action se propose aussi d'augmenter les avantages liés à l'e-santé et d'améliorer l'interopérabilité.

L'interopérabilité et les avantages liés à la possibilité de rendre facilement utilisables les informations concernant les parcours de soins de patient, ont toujours été les objets les plus importants de projets et d'initiatives initiés, au niveau européen, au cours de ces dernières années¹⁸.

Le projet epSOS¹⁹ (2008 – 2014) était un projet européen auquel 25 pays²⁰ ont participé et qui visait à adresser les problématiques d'interopérabilité des systèmes qui rendent l'accès aux données du patient très difficiles entre les Etats membres.

¹⁷ La stratégie de Lisbonne est un programme de réforme en réponse aux enjeux de la mondialisation, du vieillissement de la population, des changements climatiques et de l'élargissement de l'Union Européenne. Son objectif est de faire de l'économie de l'Union Européenne l'économie la plus compétitive du monde. <http://www.europe-centre.eu/fr/80/Les_strategies_de_Lisbonne_et_de_Goteborg.html> (Date de la dernière consultation : 14/01/2017).

¹⁸ Direttiva 2011/24/UE del Parlamento Europeo e del Consiglio del 9 marzo 2011 concernente l'applicazione dei diritti dei pazienti relativi all'assistenza sanitaria transfrontaliera. La direttiva stabilisce norme volte ad agevolare l'accesso a un'assistenza sanitaria transfrontaliera sicura e di qualità e promuove la cooperazione tra gli Stati membri in materia di assistenza sanitaria, nel pieno rispetto delle competenze nazionali relative all'organizzazione e alla prestazione dell'assistenza sanitaria.

¹⁹ Le projet appelé epSOS (Smart Open Services for European Patients) a démarré en juillet 2008. Présentation du projet disponible sur le site <<http://www.epsos.eu/france.html>> (Dernière consultation : 04/01/2017).

²⁰ Autriche, Belgique, Croatie, République tchèque, Danemark, Estonie, Finlande, France, Allemagne, Grèce, Hongrie, Italie, Luxembourg, Malte, Norvège, Pologne, Portugal, Slovénie, Slovaquie, Espagne, Suède, Suisse, Pays-Bas, Turquie, Royaume-Uni.

Il avait plus spécifiquement pour objectif la mise en place d'une opération pilote à grande échelle sur deux services: le résumé patient (Patient Summary)²¹ et la e-prescription (ePrescription)²².

Le « résumé patient », le « dossier urgence », le « dossier médicament » ou la « prescription électronique » sont des services qui visent à renforcer la sécurité des patients en mettant à disposition des professionnels de santé des données médicales indispensables pour une meilleure prise en charge. Ce projet peut être, donc, considéré comme le précurseur pour la mise en opération de ces services qui se concrétisera sous le programme CEF (Connecting Europe Facility).

Le programme CEF²³, initié en 2014, cadre dans la stratégie générale européenne Horizon 2020²⁴ et couvre toutes les infrastructures de l'économie numérique dans l'UE²⁵. L'objectif final de ce programme est la création d'une infrastructure de services numériques en Europe, permettant l'échange de données entre les Etats membres via un réseau uniforme.

En ligne avec le projet epSOS, l'initiative Trillium Bridge²⁶ (2013 - 2015) avait le but d'améliorer le rapport entre l'Europe et les États Unis sur le thème de la santé en particulier, il visait à créer un système interopérable pour favoriser l'échange d'informations cliniques pour les citoyens en soin à l'étranger. Ce projet a vu une deuxième édition, Trillium Bridge II, avec l'objectif de faciliter l'interopérabilité des dossiers de santé électroniques et de rendre possible l'échange électronique du

²¹ Le « Patient Summary » est un résumé du dossier médical du patient spécialement créé pour l'opération pilote epSOS France avec l'accord préalable du patient. Lors de son séjour dans un pays européen, le patient peut autoriser un médecin étranger à consulter son « Patient Summary » traduit dans la langue locale afin de faciliter le diagnostic et le traitement. Définition consultable sur le glossaire disponible sur le site : <www.epsos.eu/france/glossaire.html> (Dernière consultation : 06/01/2017).

²² Il permet de la création par le médecin de la prescription de médicaments via un logiciel et sa transmission électronique au pharmacien. Définition consultable sur le glossaire disponible sur le site : <www.epsos.eu/france/glossaire.html> (Dernière consultation : 06/01/2017).

²³ <<https://ec.europa.eu/digital-single-market/en/connecting-europe-facility>> (Dernière consultation : 06/01/2017).

²⁴ Description disponible sur le site : <ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/H2020_FR_KI0213413FRN.pdf>. (Dernière consultation : 04/01/2017).

²⁵ Du budget total (1,14 milliards €) alloué à cette infrastructure, 7,5 millions € sont destinés aux services de santé en ligne. Rapport « Newsletter Thématique n°3 », Agence eSanté Luxembourg, disponible sur le site <www.esante.lu> (Dernière consultation : 04/01/2017).

²⁶ Bridging Patient Summaries across the Atlantic. <<http://www.trilliumbridge.eu/>> (Date de la dernière visite 04/01/2017).

résumé patient entre les pays européens et d'autres pays à travers le monde, dont plus particulièrement les Etats Unis. Trillium Bridge II constitue par conséquent un programme d'extension de l'échange de résumés patients testé en phase pilote sous epSOS et vise à développer un seul Résumé Patient International standard.

Le projet européen EXPAND²⁷ (2013 - 2015) visait à intégrer les résultats obtenus des précédents projets européens en e-santé, plus particulièrement du projet epSOS, et à établir une voie de développement durable vers des services de santé transfrontaliers. L'objectif de ce projet était d'évoluer vers un environnement de services e-santé transfrontaliers établis sur le plan européen à travers la CEF, et sur le plan national à travers le déploiement d'infrastructures et services nationaux appropriés. Le projet se propose de développer et tester les outils techniques réalisés dans le projet epSOS en vue de l'intégration au CEF.

L'objectif du projet eSens²⁸ (2013 – 2016) est d'améliorer l'accès frontalier aux services publics dans l'UE, et de supporter le développement d'un marché numérique unique. Ce projet se propose de développer une infrastructure pour des services publics interopérables en Europe, en utilisant les résultats de projets pilotes, dont les projets epSOS. Les résultats de ce projet seront utilisés pour soutenir l'innovation des administrations publiques en ligne avec le « European Interoperability Framework »²⁹.

1.1.2 Le patient européen

Malgré les diverses situations que connaît le secteur de la santé au niveau de l'UE, il y a la nécessité qu'un certain nombre de données, qu'on peut considérer basiques, accompagnent les personnes durant leurs voyages au sein de l'UE. Il est toujours possible que des Européens tombent malades ou soient victimes

²⁷ Le projet nommé EXPAND est l'acronyme de: Expanding Health Data Interoperability Services.

²⁸ Disponible sur le site <www.esens.eu>. Date de la dernière consultation : 04/01/2017.

²⁹ Disponible sur le site <ec.europa.eu/idabc/en/document/2319/5644.html>. Date de la dernière consultation : 04/01/2017.

d'accidents en dehors de leur pays. Ce genre de situations est l'exemple idéal pour comprendre pleinement le besoin d'avoir la disponibilité des données médicales (capable de sauver la vie de patients en certains cas) en tout lieu et à tout moment. L'accès par les professionnels de santé locaux aux dossiers médicaux peut apporter une contribution vitale aux soins des patients au lieu de ces situations, ou durant le moment cruciale qui suit un accident ou un évènement particulier.

En général, l'ensemble des informations au minimum utiles dans ce contexte, est très limité et concerne :

- l'identification et données de détails (par exemple adresse, nationalité, etc.) ;
- les données anthropomorphiques (dossiers médicaux de nature différente) ;
- les allergies, réaction aux antibiotiques ;
- les maladies chroniques et traitements/prises de médicaments en cours ou les passés³⁰ les plus importantes.

Pour acter, donc, un échange de ce type de données en toute liberté et sécurité une solution de partage doit être intégrée aux pratiques des médecins pour qu'elle puisse être adoptée facilement et comporter des procédures opérationnelles simples.

La Directive n° 2011/24/UE du 9 mars 2011 relative à l'application des droits des patients en matière de soins de santé transfrontaliers a pour objectifs de faciliter l'accès aux soins de santé transfrontaliers sûrs et de qualité élevée et d'encourager la coopération en matière de soins de santé entre les Etats membres et ceux de l'EEE (Islande, Liechtenstein et Norvège) dans le plein respect des compétences nationales en matière d'organisation et de prestation des soins de santé³¹.

À cet égard, les Pays européens ont acté différentes initiatives pour régler la construction et l'actuation de dossiers médicaux capables de partager les informations contenues autant sur le territoire national qu'international.

Selon deux enquêtes menées en 2013, en Europe, dans des hôpitaux pratiquant des interventions médicales ou chirurgicales et des soins de courte durée, ainsi

³⁰ <<https://www.orange.com/fr/content/download/13559/271676/version/3/file/Acces+electroniques+dossiers+medicaux+Orange+juin+2013.pdf>>. Date de la dernière visite : 29/01/2017.

³¹ <<http://www.cleiss.fr/docs/directive.html>> (Date de la dernière consultation : 29/01/2017).

qu'auprès des médecins généralistes, la santé en ligne prend son essor. Les pays les plus actifs dans la pratique de la santé en ligne au sein du système hospitalier sont le Danemark (66 % des structures sanitaires), l'Estonie (63 %), la Suède et la Finlande (62 % chacune). En ce qui concerne la numérisation des dossiers médicaux personnels, les Pays-Bas se placent en tête avec un taux de 83,2 %; le Danemark arrive en deuxième position (80,6 %), juste devant le Royaume-Uni (80,5 %). Par contre, seuls 9 % des hôpitaux en Europe permettent aux patients d'avoir accès à leur dossier médical en ligne et, dans la plupart des cas, cet accès n'est que partiel. Lorsqu'ils adoptent des pratiques de santé en ligne, les hôpitaux et les médecins généralistes rencontrent de nombreux obstacles, que ce soit le manque d'interopérabilité, l'absence de cadre réglementaire ou encore la pénurie de ressources³². En ce qui concerne l'échange d'informations sur la santé, entre les pays les plus actifs on retrouve encore le Danemark, l'Estonie, le Luxembourg, les Pays-Bas et la Suède, par contre, en ce qui concerne l'usage de l'ordonnance électronique et des interactions avec les patients par courrier électronique, les 3 principaux utilisateurs de l'ordonnance électronique sont l'Estonie (100 %), la Croatie (99 %) et la Suède (97 %), tandis que, pour l'utilisation du courrier électronique, le Danemark (100 %), l'Estonie (70 %) et l'Italie (62 %) donnent l'exemple. Ces données analysées jusqu'à 2013 peuvent donner une première image de la situation européenne en rapport avec ces thèmes.

Selon le rapport sur l'état des systèmes de santé européens (2014 – 2015) et les bénéfices qu'en tirent les patients élaborés par l'entreprise suédoise Health Consumer Powerhouse (HCP) sur la base de rapport à l'indice européen des consommateurs de soins et de santé, les Pays-Bas finissent en première position, loin devant les 35 autres pays analysés. En se fondant sur 48 indicateurs, allant de l'information aux patients, de leurs droits, à l'accessibilité aux soins autant que la prévention ou les résultats obtenus, Health Consumer Powerhouse établit une note

³² Toutes les informations et les données sont publiées sur "Communiqué de presse" de la Commission Européenne, 24 mars 2014. Disponible au site: <europa.eu/rapid/press-release_IP-14-302_fr.htm?locale=FR> (Date de la dernière consultation : 30/01/2017).

sur 1000. Plus les pays se rapprochent du score maximum, mieux le patient devrait être soigné.³³

Aujourd'hui l'attention s'adresse plutôt vers les droits des malades, parmi lesquelles l'accès direct au dossier médical, ne s'effectue que sur la structure informatique. L'accès au dossier médical est à la base de l'utilisation mais surtout de l'interopérabilité des données sanitaires dans un cadre d'utilisation globale. Il est nécessaire, donc, une structure des systèmes de santé capables de gérer autant les informations afin de donner de prestations efficaces, que l'accès direct aux mêmes données. L'instrument le plus efficace dans ce contexte est le dossier médical qui malgré les différences entre les normes en vigueur et les moyens d'actuation, usage et diffusion, devrait être aligné dans l'ensemble des Pays européens.

En France le Dossier Médical Personnel (DMP), institué avec la Loi n. 2004-810 du 13 août 2004 concernant l'assurance maladie, est réalisée par l'Agence des Systèmes d'Information Partagés de Santé (ASIP Santé). Le DMP se propose de garantir la continuité des soins des patients et d'améliorer la qualité des services sanitaires. Il permet aux professionnels de santé autorisés d'accéder aux informations utiles à leur prise en charge et de partager avec d'autres professionnels de santé des informations médicales concernant : antécédents, allergies éventuelles, les médicaments en dotation, le compte-rendu d'hospitalisation et de consultation, les résultats d'examen (radios, analyses biologiques, etc).³⁴

Une autre expérience significative a lieu en Angleterre, avec le *National Health Service Care Record Service* (NHS-CRS) qui se propose d'améliorer l'échange des informations entre patients et professionnels de soins et de garantir aux patients l'accès aux données sanitaires pertinentes. Le NHS-CRS est composé par: le « *Summary Care Records* » qui résume les informations cliniques essentielles des patients pour être utilisées dans le cas d'urgence surtout ; le « *Detailed Care*

³³ Les résultats de l'enquête sont montrés sur le site <<http://www.touteurope.eu/actualite/les-systemes-de-sante-europeens.html>> (Date de la dernière consultation : 30/01/2017).

³⁴ <<http://www.dmp.gouv.fr/>> (Date de la dernière consultation : 30/01/2017).

Records» qui concerne plutôt l'état de santé du patient enregistré au niveau local et qui peuvent être consultées par les opérateurs ainsi dans des territoires différents.

La Danemark représente aujourd'hui un modèle d'exemple de gestion des données sanitaires et fournitures des services d'e-santé. Le Danemark est particulièrement en avance en matière de développement des services d'e-santé. Il s'agit d'un pays peu vaste et peu peuplé, qui possède un système de santé centralisé: les décisions publiques en termes de santé sont donc applicables sur le terrain plus rapidement. La population est habituée à la transmission numérique sécurisée de données personnelles. Depuis plus de quarante ans, tous les citoyens Danois reçoivent un numéro d'identité personnel unique à la naissance utile pour la conservation, l'hébergement et la transmission des informations de santé de chacun. Le pays a adopté le suivi médical dématérialisé et les technologies de l'information appliquées à la santé depuis plus de dix ans. Aujourd'hui, la quasi-totalité des généralistes et près de la moitié des hôpitaux utilisent le dossier électronique. Même les ambulanciers ont accès aux dossiers médicaux informatisés, si bien qu'ils peuvent les mettre à jour pour informer les médecins pendant que le patient est conduit aux urgences, ce qui permet un gain de temps précieux.

Dans ce cadre, on peut donc comprendre comment l'adoption des initiatives en ligne avec tous les Pays européen, peut être importante pour permettre aux patients d'avoir les justes soins et être assuré en cas d'évènements accidentels dans un autre pays.

1.1.3 Le contexte italien

Le FSE (Fascicolo Sanitario Elettronico) représente l'ensemble des documents numériques sanitaires dérivant d'évènements cliniques présents ou passés concernant l'ensemble du parcours de soin des patients et, il est alimenté par les opérateurs qui sont impliqués dans le processus de soins et qui suivent les patients

et par les systèmes d'information utilisés par les structures. Le FSE est géré par les Régions et les Provinces Autonomes italiennes qui, au niveau local, s'engagent de son développement et sa diffusion pour en favoriser l'usage.

Le décret législatif (DL) 179/2012, modifié par le 69/2013 a défini formellement le FSE et a introduit certaines mesures à respecter pour son application au niveau de Régions et Provinces Autonomes. L'article 12 comma 15-bis du décret, établit qu'entre le 30 juin 2014 les Régions et les provinces autonomes devaient présenter à Agenzia per l'Italia digitale (AgID) et au Ministère de la Santé, le plan d'Action pour l'actuation du FSE en basant sur les lignes guides déterminés pour eux-mêmes. Par ailleurs, l'article 12 comma 15-ter, du même décret établit que AgID, en considérant les exigences des Régions et des Provinces Autonomes, détaillées dans les plans d'actions présentés, doit s'occuper, avec le support technique du CNR (Consiglio Nazionale delle Ricerche) et en accord avec le Ministère de Santé de la planification et la réalisation de la structure nationale nécessaire pour garantir l'interopérabilité des FSE et à terminer pour le 31 décembre 2015.

Le décret du Président du Conseil des Ministres du 29 Septembre 2015, n° 178³⁵, identifie les caractéristiques à considérer pendant la réalisation de la structure du FSE (finalité, contenus, traitement des données, etc.), en détaillant à l'article 28 les services à assurer dans une première phase. En particulier, on se réfère aux services d'accès pour le patient, aux services pour la connexion et l'accès du personnel médical, la disponibilité des services à support de l'interopérabilité du FSE et pour la gestion des rapports médicaux, les services de gestion du résumé sanitaire synthétique.

L'objectif stratégique de cette structure de FSE, concerne l'offre d'un ensemble des services au niveau central à utiliser pour l'interopérabilité des différents systèmes régionaux de FSE. La structure centrale réalisée fournira services supplémentaires aux Régions qui n'ont pas encore une structure de FSE propre en couvrant le manque de leurs systèmes de FSE déjà réalisés au niveau régional.

³⁵ Disponible au site : < <http://www.gazzettaufficiale.it/eli/id/2015/11/11/15G00192/sg>> (Date de la dernière consultation : 30/01/2017).

Dans le contexte italien, caractérisé d'une large variété de Régions, différentes initiatives ont été mise en œuvre pour atteindre les buts concernant l'interopérabilité ainsi que l'utilisation des données sanitaires pour être en ligne avec les directives européennes. Au-delà des opérations actées au niveau national, chaque région a mise en place ses propres opérations qui visaient à rendre disponible le FSE pour les citoyens du territoire. Malgré les initiatives mises en opération au niveau locale, il est possible d'identifier un ensemble d'objectifs à atteindre au niveau global, en particulier : l'amélioration de services pour les citoyens par la réduction des distances entre ceux-ci et les institution sanitaires ; donner une continuité au processus de soin, surtout en termes d'informations partagées et qui vont former le parcours de soins pendant la vie des patients ; rendre plus efficaces les processus du système de santé ; la construction d'un réseau entre les opérateurs impliqué dans le processus de soins capable d'échanger et partager les informations ; la réduction déplacements demandés aux patients ; l'introduction de documents standardisés pour les opérateurs sanitaires ; une politique de traitement de données personnels plus fiable en assurant une meilleure sécurité dans la gestion des informations ; la structuration des documents en considérant les standards internationaux (par exemple : HL7-CDA).

Le thème de l'interopérabilité de données sanitaires et d'informations personnelles des patients a toujours été considéré comme le fondement pour la réalisation d'une structure commune de FSE et d'un système de santé national capable de partager tous les types d'informations. À ce propos, beaucoup de projets et initiatives nationales ont été initiés pour atteindre cet objectif.

Le projet ICAR³⁶ (2006 - 2009) avait l'objectif de rendre interopérables les systèmes informatifs de santé des Régions pour permettre une correcte transmission des données au niveau national. Ce projet posait ses bases sur un travail de coopération entre l'État, les Régions et les administrations locales pour le

³⁶ « Interoperabilità e Cooperazione Applicativa in rete tra le Regioni ».

développement d'une structure d'e-gouvernement qui puisse assurer au mieux les services aux citoyens.

Le projet IPSE³⁷ institué dans le 2009, vise à développer un système diffusé d'interopérabilité des FSE régionaux, en particulier en ce qui concerne deux des documents les plus importants (« Patient Summary » et « ePrescription »), pour améliorer la continuité des soins et soutenir la coopération et l'efficacité des services pour optimiser les temps, la rationalisation des coûts et la réduction des erreurs médicales. La réalisation de ce projet a été mise en opération en considérant aussi les résultats obtenus par les projets « Mattoni SSN » et « ICAR », déjà expliqués ci-dessus. Le résultat le plus important concerne la réalisation d'une solution d'e-santé au niveau national, en ligne avec les opportunités européennes dans ce domaine, en faveur de l'interopérabilité européenne et de l'amélioration des rapports de coopération entre les États membres. Le projet se propose d'harmoniser les résultats obtenus au niveau européen avec le projet epSOS.

Il est donc possible, de vérifier que tous les projets nationaux et les initiatives proposées par les Régions ont comme objectif de base le développement de la structure du FSE, l'interopérabilité des données sanitaires sur le territoire national et international, et l'amélioration conséquente des services offerts par les institutions de santé. Comme cela a déjà été expliqué, le FSE a été développé, à aujourd'hui, au niveau régional en considérant les opérations mises en œuvre par chaque région, cela veut dire que chaque administration locale a pu opérer de façon autonome en ligne avec les directives nationales. Néanmoins, en conformité avec la loi qui en règle l'actuation et pour aligner toutes les initiatives déjà actées au niveau locale qu'ont montré beaucoup de divergence en termes d'innovation de processus et adoptions de technologies les plus avancées, le FSE vise à avoir une structure au plus haut niveau, qui soit la même sur tout le territoire national.

³⁷ « Sperimentazione di un sistema per l'interoperabilità Europea e Nazionale delle soluzioni di Fascicolo Sanitario Elettronico: Componenti Patient Summary e ePrescription ».Le projet représente l'initiative de participation au projet européen « epSOS ».

Initialement le modèle de FSE concernait l'interopérabilité directe entre les Régions alors que la nouvelle structure vise à un échange d'information possible et contrôlé aussi au niveau central.

En particulier, les nouvelles initiatives visent à atteindre différents buts : la création d'un FSE central pour chaque code identifiant des patients sur tout le territoire national ; la gestion univoque du consensus du patient sur le stockage et le traitement des données ainsi que sur l'utilisation de la structure ; la gestion centrale du résumé des patients formé par les données personnelles (nom, prénom, région d'appartenance, etc.) et les données sanitaires stockées par l'administration locale (région) ; gestion centrale de systèmes de codage ; interopérabilité des FSE régionaux avec un index univoque de référence ; alignement des données stockées au niveau locale vers la structure nationale et, au contraire, permettre aussi de garder une structure de FSE régional en partageant la mise à jour.

Il est évident, donc, que dans ce processus de passage d'une structure de FSE régional vers une structure nationale, l'interopérabilité est extrêmement importante tel au niveau technologique que sémantique ; l'information doit être échangée en utilisant les technologies les plus avancées mais surtout il faut s'assurer que le contenu sémantique ne soit pas altéré. En effet, nombreux sont les services à support de l'interopérabilité prévus pour réaliser ce changement autant concernant la récupération et la recherche des documents que l'identification univoque des données personnelles des patients. Dans ce cadre d'interopérabilité globale, les terminologies cliniques de références pour l'identification de la typologie de maladie ainsi que les résultats des examens de laboratoire jouent un rôle fondamental.

1.2 Interopérabilité des données sanitaires : les terminologies de référence

Les systèmes d'information dans le domaine de la santé doivent être capables de communiquer pour favoriser la coopération entre les professionnels impliqués dans le cadre des parcours de soin et les patients (les usagers), et pour aider à prendre des décisions au cours des processus médicaux. Les initiatives destinées à améliorer l'interopérabilité des SIS sont réalisées en ligne avec les règles d'informatique qui permettant cette communication en garantissant l'échange et le partage des données de santé dans le respect des exigences de sécurité et de confidentialité des données personnelles de santé (interopérabilité technique), et leur traitement et leur compréhension en s'appuyant sur un langage commun (interopérabilité des contenus métiers).

Le plan d'action pour l'interopérabilité de SIS s'appuie sur des normes et des standards internationaux et se construit en concertation avec les exigences requises pour optimiser les processus de soins en termes de partage et d'échange de données pertinentes. Les SIS sont utilisés à des fins non seulement de coordination du parcours de soins du patient mais aussi d'aide à la décision médicale au service des professionnels de santé, d'évaluation de leurs pratiques, d'épidémiologie et de santé publique, ainsi que de recherche clinique. En outre, la mise en opération de systèmes interopérables a pour corollaire une augmentation de la dématérialisation des données de santé pour en faciliter l'échange et le partage ainsi que le traitement.

Cette dématérialisation des données nécessite la définition de langages communs aux systèmes d'information pour éviter la définition de nouveaux langages et donc de nouveaux développements à chaque fois que deux systèmes d'information veulent échanger ou partager des données³⁸.

³⁸ Cadre d'interopérabilité des Systèmes d'Information de Santé (CI-SIS), disponible sur le site <<http://esante.gouv.fr/services/referentiels/referentiels-d-interoperabilite/cadre-d-interoperabilite-des-systemes-d>> (Date de la dernière consultation : 10/01/2017).

Le FSE, comme indiqué plus haut, représente l'ensemble des documents numériques sanitaires qui décrivent le parcours complet de soins des patients, et représente le moyen le plus important qu'amène à rendre interopérables données et des informations sanitaires. Ses contenus, comme prévu dans la normative, sont représentés par un nombre minimum des données et documents, ainsi que un certain nombre de données et de documents supplémentaires qui permettent de l'enrichir. Le socle minimal des documents et données à insérer dans le FSE est le même pour toutes les structures instituées par les Régions et Provinces Autonomes, et il est représenté par : les données identifiants du patient ; les rapports médicaux ; les comptes rendus de soins urgents ; le « Patient Summary » ; le dossier pharmaceutique ; le consensus ou le refus sur les dons d'organes.

Les informations contenues dans les documents du FSE sont stockées en utilisant terminologies de référence et systèmes de codage, qui en assurent l'interopérabilité sémantique tant dans le contexte régional que national et européen. Chaque région ou province autonome propose des services à support de l'interopérabilité du FSE qui doivent garantir au minimum la recherche et la récupération des documents insérés.

Conformément à l'article 24 du décret qui règle l'actuation et la réalisation du FSE, pour assurer le échange correct d'informations sans en altérer leur contenu sémantique, le standard HL7 doit être utilisé pour définir les données à échanger et un ensemble de systèmes de codage doit être adopté dans les documents. En particulier, les documents insérés dans le FSE doivent contenir le standard ICD 9-CM (International Classification of Diseases 9th revision – Clinical Modification) et LOINC (Logical Observation Identifiers Names and Codes) ainsi que les autorisations AIC³⁹ (Autorizzazione all'immissione in commercio di un farmaco) et la classification ATC⁴⁰ (Classificazione Anatomica Terapeutica Chimica).

³⁹ AIC, <<http://www.agenziafarmaco.gov.it/it/content/1%E2%80%99autorizzazione-all%E2%80%99immissione-commercio>> (Date de la dernière consultation 10/01/2017).

⁴⁰ Système de classification ATC, <<http://www.aifa.gov.it/sites/default/files/bif020659.pdf>> (Date de la dernière consultation 10/01/2017).

Le standard ICD 9-CM, développé par la World Health Organisation (WHO), représente le système de classification international capable d'organiser les maladies et les traumatismes en groupes en basant des critères bien définis qui permettent de définir avec précision le diagnostic associé aux patients. La traduction en italien a été réalisée par le centre de collaboration italien de la WHO et a été utilisée dans le FSE qui est la version du 2007.

Le standard LOINC est un système de noms et de codes universels qui identifient de façon univoque les observations cliniques et de laboratoire pour en faciliter le partage et l'échange des résultats de diagnostics provenant de systèmes appartenant à des structures sanitaires différents.

Les autorisations AIC représentent les autorisations à commercialiser un médicament dans le territoire national et permettent d'identifier avec précision les médicaments distribués et d'assurer l'assistance pharmaceutique, l'accès aux médicaments innovants pour les maladies les plus rares.

Le système de classification ATC concerne la subdivision des médicaments en groupe en basant les organes ou les appareils sur lesquelles agissent leurs propriétés chimiques, pharmacologiques et thérapeutiques ; contrairement à l'AIC, l'ATC identifie plutôt le principe actif des médicaments.

Pour l'implémentation de ces systèmes de classification et de codage, les Régions doivent acter un certain nombre de procédures concernant la formation du personnel sur l'utilisation et la prédisposition des systèmes informatiques.

Chapitre II

2.1 Les systèmes de noms et de codes : LOINC

Acronyme de « Logical Observation Identifiers Names and Codes », le standard LOINC est une terminologie de référence internationale pour le codage des observations biologiques-médicales publié par l'Institut de Recherche de l'Indiana University, le « Regenstrief Institute » dans sa version originale en anglais [17].

Le développement des services numériques et de l'usage des technologies innovatives et avancées ont conduit les professionnels de santé vers la nécessité d'échanger chaque jour un grand nombre de données par voie informatique. La transmission de résultats des analyses biologiques-médicales nécessite un encodage structuré, une nomenclature commune disponible pour tous les acteurs impliqués dans le processus de transmission, et qui puisse permettre de maintenir le contenu sémantique intact pendant le passage de données. Il est possible, donc, de transmettre telles informations dépourvues d'ambiguïté en utilisant des systèmes de classification et de codage opportuns, qui permettent de garder la même sémantique sans aucune altération des contenus.

Un des systèmes de classification et de codage qui se prête bien à ces buts est le système LOINC, qui avec sa nomenclature structurée et standard permet d'identifier le même test clinique tel au niveau local, étant fréquente l'utilisation de codes différents par les laboratoires appartenant à la même structure sanitaire, que au niveau international entre différents Pays, en surmontant les obstacles et les barrières linguistiques. La base de données LOINC représente une base de données « universel » dans laquelle il est possible de retrouver noms et codes capables de

décrire et identifier tests biologiques effectués par les laboratoires. Le jeu de valeurs LOINC pour les résultats de biologie est l'un des contenus sémantiques les plus importants à partager pour obtenir l'interopérabilité des données sanitaires sur le territoire national ainsi que l'international. L'objectif du système de classification et de codage LOINC consiste dans la création d'un identifiant qui puisse permettre l'échange de données entre les laboratoires cliniques-médicaux stockés dans systèmes informatiques différents, de façon univoque. Pour faciliter ce processus d'échange d'informations, chaque identifiant est formé de façon standard avec une structure déterminée. Chaque test est en effet, décrit par un ensemble de 5 ou 6 éléments, appelés axes, qui en décrivent la nature et les caractéristiques cliniques. Ce jeu de valeurs permet de coder une analyse élémentaire de biologie et de la différencier en se basant sur les résultats que l'on peut obtenir. Le standard LOINC représente un langage défini, qui facilite la communication autant entre les machines (systèmes caractérisés par différents environnement informatique) qu'entre les experts de laboratoire puisqu'il permet l'identification univoque des tests, en gardant la même sémantique (et donc le même code) dans tout le Pays qui utilisent le standard. Le système LOINC est adopté, à ce jour, par plus de 47500 utilisateurs dans 176 Pays⁴¹, pour lesquels ont été déroulé un travail de traduction à partir de la version originale en anglais pour permettre sa diffusion et son utilisation. Comme déjà énoncé, il est doté d'une nomenclature composée par 5 ou 6 attributs (Tableau 1), que vont préciser toutes les caractéristiques des analyses.

⁴¹ <loinc.org>. (Date de la dernière consultation 23/01/2017)

LOINC code	<i>5792-7</i>
Component	<i>Glucose</i>
Property	<i>MCnc</i>
Time	<i>Pt</i>
System	<i>Urine</i>
Scale	<i>Qn</i>
Method	<i>Test strip</i>

Tableau 1 - Exemple de test codifié et structuré avec les six axes LOINC

Chaque examen inséré dans la base de données officielle LOINC, en effet, est associée à une description (« Fully specified Name » en anglais) composée par les attributs LOINC de façon structurée. Cette description est très importante puisqu'elle représente la structure utilisée pour décrire un test dans la base de données officielles et détermine, en conséquence, le niveau de détail pour la recherche et le mappage des codes. L'opération de mappage effectuée par les laboratoires biologiques médicaux, dépende du niveau de détail dans la laquelle les examens sont décrits dans la base de données locale, ça veut dire, si les attributs composants de la description LOINC sont effectivement présents, ainsi que par la capacité d'utilisation du système LOINC et du logiciel de support RELMA. La présence dans la base de données locale, d'un nombre significatif des attributs correspondant au standard LOINC détermine le niveau de détail et d'exactitude de la recherche de tests pertinents. Pendant les opérations de mappage, les problèmes ne résident pas seulement dans la présence mais plutôt dans la forme linguistique dans laquelle les attributs sont exprimés. Chaque partie de la description, et donc, chaque partie appartenant au nom complexe associé aux examens décrits en LOINC, doit retrouver son correspondant dans la description du test local. La difficulté de cette opération dépend aussi de règles strictes du standard qui prévoit pour chaque attribut un ensemble de forme linguistique bien déterminé.

2.1.1 Le premier attribut : « Le Composant »

L'attribut « Analyse » appelé aussi « Composant » ou « Analyte » (dans la version anglaise « Component ») contient le nom principal de l'analyse avec si besoin des sous-classes (autres éléments linguistiques) qui apportent des informations supplémentaires par rapport aux informations de base.

L'attribut « Composant » est sûrement le plus varié dans ses formes linguistiques qui doivent suivre certaines règles définies, par exemple :

- les abréviations ou fractions de mots ne sont pas utilisées dans le standard sauf celles permises et opportunément identifiées⁴² ;
- l'ordre des éléments définit une sémantique précise, le premier élément représente, en effet, la partie centrale du test effectué, la substance qu'on veut mesurer, le composant principal de l'examen ;
- le nom taxonomique est préférable pour la description des organismes et le nom de virus pour décrire la présence de maladies ;
- il est préférable que les mots « direct », « indirect », « conjugué » « non conjugué » ne soient pas utilisés pour la recherche et le mot « total » est utilisé seulement pour identifier une fraction ;
- les vitamines sont décrites par le nom chimique : « thiamine » qui est préférée à « Vitamine B1 » qui peut être ajouté comme synonyme ;
- les alcools doivent être exprimé par un seul mot : « éthanol » qui est la forme utilisée pour exprimer « alcool éthylique » ;
- le nom de l'anion est préféré au nom de l'acide : « urate » qui est utilisé pour l'« acide urique » ;
- pour le symbole OH (« Hydrogène ») associé au composant ainsi que pour les lettres grecques, il faut utiliser la dénomination entier (par exemple : OH-progestérone identifie le composant « Hydroxyprogestérone »).

⁴² C.J. McDonald, S.Huff, J.Deckard, S.Armson, S. Abhyankar, D.J. Vreeman, *LOINC® Users' Guide*, Décembre 2016, p. 12.

Les règles décrites ci-dessus, représentent un ensemble des celles concernant les formes linguistique et parties de mots utilisés par le standard. Toutefois, diverses sont aussi les règles associées à la ponctuation ; les points ainsi que les virgules ont une sémantique précise qui beaucoup de fois suggèrent la position des éléments analysés dans leur composition chimique ou identifient l'exacte catégorie des composants. Pour indiquer les nombres sont utilisé plutôt la numérogie romaine que celle arabes. Toutefois, quand les nombres sont utilisés pour indiquer un exposant ils doivent être précédé par le mot « esp » à ne pas confondre avec l'indication avec les symboles « * » et « ^ », qui ont déjà une autre sémantique pour LOINC. Le symbole « ^ » agit pour séparer les parties du composant : la première partie est représentée par le nom principal (par exemple : de l'analyte ou la mesure à analyser), la deuxième peut indiquer la typologie de provocation avec l'indication du temps si nécessaire (par exemple : le composant « Glucose post prandial » qui devient « Glucosio^post pasto » dans la version italienne) , la substance de provocation ainsi que la quantité administrée, dans la troisième partie les méthodes de standardisation et de correction des valeurs mesurées sont indiquées (par exemple : « Leucocytes corrigés par érythroblastes nucléés » qui devient « Leucociti^corretti per eritrociti nucleati » dans la version italienne).

La variété linguistique du composant représente un des problèmes les plus traités dans la recherche d'une méthodologie de support pour les laboratoires qui reconnaissent les formes linguistiques locales et qui le puissent associer aux formes complexes du standard telles qu'elles sont décrites.

2.1.2 Le deuxième attribut : « La Grandeur »

Le deuxième attribut présent dans le nom LOINC définit la dimension (masse, volume, temps, etc.) du résultat des analyses concernées. L'attribut « Grandeur » (en anglais « Property »), se présente sous forme de nombreuses typologies, chacune avec une variété linguistique (on parle plutôt des sigles) différente. À ce propos, il dessine les attributs les plus difficiles à comprendre surtout pour la variété sigles et acronymes utilisés et qui n'explicitent pas directement le contenu à représenter. Néanmoins, dans la base de données locale cette typologie d'attribut n'est pas décrite directement, mais déduite à partir de l'unité de mesure associée à chaque test. Il est possible de distinguer en particulier :

- la catégorie qu'indique une *masse*⁴³ (milligrammes, grammes etc.), caractérisés d'avoir comme numérateur de l'unité de mesure une masse (mg, g, etc.). On peut avoir, donc, une « concentration de masse » (« Mass concentration » en anglais) ou un « contenu de masse » (« Mass content » en anglais)⁴⁴ ;
- la catégorie qu'indique une *substance*. LOINC préconise que toutes les unités de mesure ayant pour dimension une expression dont le numérateur est « mole » ou « mEq » (« milliéquivalents »), ils ont une grandeur qui commence par le mot « substance » ;
- la catégorie qu'indique une *activité catalytique*, se présente quand la grandeur des analyses commence par le terme « catalytique ». On peut avoir, en effet, une « concentration catalytique » ou un « contenu catalytique » ;
- la catégorie qu'indique une unité *arbitraire* se caractérise par les unités de mesures ayant comme numérateur le mot « arbitraire ». Toutes les analyses exprimées en U/L ou UI/L ont une grandeur "arbitraire".;

⁴³ Ibid. p. 19.

⁴⁴ L'axe LOINC qu'identifié la grandeur n'a pas été traduit pour garder la spécificité propre aux acronymes fournis par le système.

- la catégorie qu'indique un *nombre* est associée aux observations avec grandeur qu'identifie un comptage ou un nombre.

On a pu vérifier que l'attribut « grandeur » est strictement lié à l'attribut « unité de mesure ». Ce lien est très important parce que cela permet de déduire la valeur de la grandeur à partir de l'unité de mesure (par exemple : l'unité de mesure « mg/dl » exprime un rapport masse/volume, identifié par la grandeur *Concentration de Masse* « MCnc »). La déduction qu'il est possible d'effectuer et les liens très stricts qui existent entre les valeurs des attributs du standard, sont très importants à considérer pendant les opérations de mappage puisqu'ils permettent de déduire deux valeurs exprimées en forme standard à partir d'un seul.

Le système LOINC a introduit l'attribut « grandeur » pour réduire la variété des codes en correspondance d'examens avec unités de mesure qui peuvent être associées à la même grandeur (par exemple les unités de mesure « mg/dl » et « g/l » se réfèrent à la même valeur de grandeur « MCnc »).

2.1.3 Le troisième attribut : « Le Temps »

La troisième partie du nom complet LOINC est représentée par l'attribut « Temps » (en anglais « Timing »)⁴⁵ qui exprime le temps de réalisation de l'analyse. Il est possible d'effectuer un examen dans un moment précis ainsi que dans un laps de temps spécifique. Les quantités analysées durant un laps de temps défini sont exprimées souvent par un indice de masse (par exemple par la grandeur « MRat » (« Mass Rate ») et l'unité de mesure « g/24h ») ou par indice de substance (par exemple par la grandeur « SRat » (« Substance Rate ») et l'unité de mesure « mol/24h »). Les examens effectués dans un laps de temps déterminé sont appliqués surtout en considérant le milieu biologique « Urines » et « Fèces » (par

⁴⁵ Ibid. p. 26.

exemple : calcul de concentration d'urines en 24 heures) et sont associé à des valeurs déterminées dans la nomenclature du standard. En particulier, on utilise le sigle « Pt » pour les examens effectués ponctuellement donc dans un moment déterminé, « 24h » pour les observations déroulée dans le 24 heures suivant l'exécution du test ainsi qu'un chiffre suivi par l'indication concernant les minutes, les semaines, les jours ou dans certains cas le moins d'observations (Tableau 2).

Attribut « Temps »

<i>Abréviations</i>	<i>Nom complet</i>
15M	15Minutes
2H	2 hours
1D	1 days
2W	2 weeks

Tableau 2 - Valeurs associés à l'attribut « Temp »

Aussi pour l'attribut « Temps », il est possible de vérifier les liens possibles avec les autres valeurs. En particulier, l'information sur le temps peut être déduite (comme dans le cas de la grandeur) par l'unité de mesure quand elle exprime une valeur qui est mesurée dans un temps précis (par exemple mg/24h). Encore une fois le réseau de déduction créée peut être d'une importance fondamentale pour surmonter le manque d'information éventuelle dans la base de données locale. Le temps d'application des examens peut, donc, être déduit même s'il n'est pas complètement explicité au niveau local.

2.1.4 Le quatrième attribut : « Le milieu biologique »

La quatrième partie du nom complet LOINC est représenté par le milieu biologique (en anglais « System »), l'attribut qui identifie le type de système sur lequel l'analyse est pratiquée. Le standard LOINC offre un ensemble très détaillé des valeurs possibles du milieu biologique (par exemple il décrit toutes les

typologies de sang : sang, sang artériel, sang veineux, sérum, plasma, etc.)⁴⁶. Malgré que les laboratoires soient très restrictifs dans leur variété linguistique pour cet attribut. Les laboratoires, en effet, utilisent un petit ensemble de valeurs par rapport à celles fournis par le standard sans expliciter une telle différence dans leur base de données locale. Dans le Tableau 3 ci-dessus, un sous ensemble des formes linguistiques standard du milieu biologique les plus utilisées.

Attribut « Milieu Biologique »

<i>Abréviations</i>	<i>Nom complet</i>
Bld	Blood
Urine	Urine
BldA	Blood Arterial
BldC	Blood capillary

Tableau 3 – Valeurs associés à l'attribut « Milieu Biologique »

2.1.5 Le cinquième attribut : « L'Échelle »

La cinquième partie du nom complet LOINC concerne l'échelle de mesure (en anglais « Scale ») associée au résultat des tests⁴⁷. Cette valeur indiquée par des sigles, exprime la typologie de résultat qu'il est possible d'obtenir en effectuant les examens concernés. En particulier, on peut distinguer :

- les valeurs numériques, identifiées par le sigle « **Qn** », quand les résultats appartiennent à une échelle numérique et doivent être exprimés par un nombre entier, un rapport, un nombre réel ou un range (par exemple exprimé avec les symboles « > », « < », etc.) ;
- les valeurs ordinaux, par le sigle « **Ord** », quand on traite de valeurs ordonnées par catégories (positifs, négatifs, réactif, indéterminé, etc.) ;

⁴⁶ Ibid, p. 28.

⁴⁷ Ibid, p.. 31.

- les valeurs nominaux, avec le sigle « **Nom** », dans le cas de résultat sans un ordre naturel spécifique. C'est le cas des noms des éléments (par exemple dans le cas d'analyse de bactéries) ou des catégories concernant l'aspect (il y a des cas dans lequel le couleur est très importante) ;
- les valeurs indiquant un texte narratif, comme une description d'une partie microscopique analysée ou d'un élément concerné, les commentaires sur un résultat déterminé sont identifiées par le sigle « **Nar** ».

La valeur d'échelle est, comme les autres, liées aux valeurs des autres attributs. En particulier, à partir de l'unité de mesure et en considérant aussi la grandeur associée, il est possible de déduire la valeur d'échelle. Par exemple si l'unité de mesure est définie, dans la majorité des cas l'échelle sera inévitablement une valeur numérique. L'ensemble des éléments (grandeur, unité de mesure, milieu biologique, échelle) sont liés entre eux puisqu'ils peuvent être déduits les uns des autres, et représentent la majorité des éléments standards à identifier pendant les opérations de mappage. Clairement, il y a des cas dans lesquels cette déduction n'est pas si immédiate, étant donnée la variété linguistique fournie par le standard.

2.1.6 Le sixième attribut : « La technique »

La dernière partie du nom complet LOINC, concerne la typologie de technique adoptée pour effectuer les tests. La technique (en anglais « Method »)⁴⁸ est le seul paramètre qui ne doit pas être explicité obligatoirement, donc, sa présence éventuelle est demandée pour différencier cliniquement les examens. En effet, la méthode utilisée est explicitement indiquée surtout quand, au pair des autres éléments, la méthode donne une relevance clinique différente aux tests effectués (par exemple le test du glucose dans le sang effectué par une méthode déterminée doit être différencié par le dernier attribut). La liste des valeurs possibles (Tableau 4) pour les méthodes est très variée puisque la nomenclature officielle permet d'utiliser la définition complète pour la grande partie et des sigles seulement pour un sous ensemble.

Attribut « Technique »

<i>Abréviations</i>	<i>Nom complet</i>
Aggl	Agglutination
Coag	Coagulation Assay
CDC	Complement-dependent Cytotoxicity
IA	Immunoassay
IF	Immune Fluorescence

Tableau 4 - Exemple de valeurs associé à l'attribut « Technique »

La valeur de la méthodologie utilisée pour effectuer le test n'est pas obligatoire et son rôle dévient fondamental quand sa spécification différencie significativement les examens. En particulier, les opérations de mappage devraient être effectuées en tenant compte des cinq éléments obligatoires et en considérant la méthode déclarée seulement si elle est présente et si elle peut différencier les examens.

⁴⁸ Ibid. p. 33.

2.2 Le processus de traduction : « LOINC Italia »

La traduction de terminologies demande des étapes spécifiques qui doivent considérer autant les différences linguistiques que celles liées aux contextes d'usage des termes. Les premiers principes explicités par *Peter Newmark* dans la *theorie de quatre niveaux* peuvent être facilement appliquée aux terminologies médicales contrôlées [21]. À cet égard, le « *International Health Terminology Standards Development Organization* » (IHTSDO), a publié deux manuels pour désigner les lignes guides pour le travail de traduction de la terminologie SNOMED : « *Guidelines for Translation of SNOMED CT* » [15] et « *Guidelines for the Management of Translations of SNOMED CT* » [16], pour ceux qui doivent s'approcher pour la première fois à ce genre de travail. Les principes délinés dans ces deux manuels ont soulignés les nécessités d'assurer une cohérence linguistique ainsi que de réaliser une traduction basée sur les concepts et non pas seulement sur les mots.

Les RI considère nécessaire la traduction de LOINC pour sa diffusion et son adoption et, à cet propos, il a développé ces dernières années différents outils pour supporter l'important travail des traducteurs. Le développement de ces outils a été mené en tenant compte, en effet, de la nécessaire contribution des traducteurs humains experts du domaine, indispensable pour une bonne exécution du travail [25]. L'approche à la traduction LOINC suggérée par le RI [19], fonde ses bases sur l'utilisation d'un algorithme qui crée automatiquement la traduction de tous les termes de la base de donnée en partant des différentes parties du nom complet LOINC. Malgré la variété des règles grammaticales et des structures syntactiques propres de différentes langues dans laquelle on retrouve la nomenclature LOINC, la traduction basée sur les parties permet d'avoir une approche plus flexible dans la détermination de l'ordre des mots et du sens sémantique utilisé dans le contexte de la langue de destination. Par ailleurs, étant les termes subdivisés en classe de

spécialités, les décisions concernant les possibles règles de traduction, prises à partir d'un terme, peuvent être valides pour toute la classe d'appartenance.

Dans le cadre du projet « Infrastruttura tecnologica del Fascicolo Sanitario Elettronico » (InFSE)⁴⁹, le groupe de travail « LOINC Italia » au sein de l'Institut d'Informatique et Telematique (IIT – Istituto Informatica e Telematica) du Conseil National de la Recherche (CNR- Consiglio Nazionale delle Ricerche), ont abordé à la traduction de LOINC en langue italienne et testé son efficacité pendant les opérations de mappage en considérant un ensemble de bases de données locales appartenant à un échantillon de laboratoires considérés. La collaboration et le rapport de partenariat avec le RI, a porté à une profonde évaluation des capacités d'adaptation du système LOINC dans le contexte italien, pour rendre le travail de traduction plutôt en ligne avec les contextes d'utilisation que avec les seuls termes.

L'approche à la traduction de LOINC a suivi la méthode suggérée par le RI, qui prévoyait la traduction des seules parties pour la création suivante du nom complet LOINC et qui peut être subdivisé en deux phases. La première (approche « top-down ») où les parties sont traduites manuellement par les experts du secteur et du domaine spécifique et la deuxième (approche « bottom up ») qui par les opérations de mappage permet l'utilisation de la variabilité linguistique locale pour enrichir l'ensemble des synonymes associés aux noms LOINC dans la version italienne officielle. Une collecte générale de formes linguistiques utilisées par les laboratoires biologiques-médicales a été effectuée en voie préliminaire au travail de traduction. En particulier, à un échantillon des laboratoires⁵⁰ a été demandé de structurer leur examens selon la structure en axes LOINC pour pouvoir mesurer les possibles similarités et différences des variantes linguistiques tel entre les différents laboratoires que par rapport aux définitions officielles LOINC. Cette étude importante a permis de vérifier comme le même examen est nommé en quatre

⁴⁹ Le projet a été conçu par un accord entre le « Ministero per la Pubblica Amministrazione e l'Innovazione » et le « Consiglio Nazionale delle Ricerche » pour réaliser un modèle et une structure nationale pour le FSE (Fascicolo Sanitario Elettronico).

⁵⁰ La collaboration avec « l'Azienda Ospedaliera Universitaria Molinette di Torino » a permis d'identifier 8 laboratoires appartenant à la Région « Piemonte ».

modalités différentes sur un échantillon de huit laboratoires, en déterminant les premières lignes d'actions pour enrichir les définitions standards, avec synonymes (formes linguistiques locales retrouvées) qui puissent faciliter la recherche et les opérations de mappage effectuées par les laboratoires.

Les opérations d'enrichissement de la base de données LOINC ainsi que la méthodologie de traduction, ont été supportées aussi dans la comparaison avec les versions LOINC développées pour d'autres langues latines, le français et l'espagnol, grammaticalement et syntaxiquement similaires à l'italien et qui pouvaient donner une contribution importante pour améliorer les solutions adoptées.

La *Société Française d'Informatique de Laboratoire* (SFIL) [7] a été particulièrement importante à cet égard. Selon le RI les opérations de mappage ne veulent pas substituer les codes LOINC aux codes locaux, mais les intégrer dans les procédures d'échange informatif pour les faire devenir des identifiants univoques et accéder aux résultats des analyses de laboratoire au niveau international. L'activité de traduction n'a pas été déroulée, donc, pour réaliser une exacte correspondance entre les langues mais pour assurer la possibilité d'effectuer correctement les opérations de mappage vers le code LOINC correct et en obtenant une utilisation facile indépendamment de la langue.

Une première version de la traduction italienne de LOINC a été distribuée avec la « LOINC release » version 2.34 et RELMA 5.0 en décembre 2010, et était composée par 43152 codes⁵¹. Le résultat obtenu a été régulièrement conforme aux attentes, en effet, toutes les aspects critiques dérivait des différences syntaxiques et grammaticales entre la forme officielle en anglais et la version traduite en italien. Différents cycles de révision et modification des parties traduites ont été conduits pour évaluer les termes traduits à chaque fois, en utilisant une terminologie de secteur plus spécifique comme terme de comparaison et déterminant ultérieurs règles de traduction basée sur des évaluations linguistiques.

⁵¹ Le 43152 ont été obtenus par la traduction des 11748 parties générées automatiquement par l'algorithme de traduction des termes LOINC développée par le RI.

En particulier, grâce à la collaboration avec le « *LOINC Development Team*⁵² » a été possible de comprendre la logique du système de codage, concevoir des règles de traduction plus efficaces, trouver des solutions linguistiques pour les problématiques liées à la traduction de parties et avoir une mise à jour de parties LOINC plus organisée et cohérente.

2.3 L’outil pour le mappage de codes : RELMA

La base de données LOINC est distribuée gratuitement sur le site officiel <www.loinc.org> sur laquelle il est possible de télécharger aussi les tutoriels en différentes langues et le logiciel de support au mappage RELMA [20] (*REgenstrief LOINC Mapping Assistant*).

L’outil RELMA a été projeté par le RI et permet de surfer facilement dans la base de données LOINC, en donnant la possibilité de changer automatiquement la langue sur laquelle on veut travailler. Cet outil a été construit pour permettre aux utilisateurs d’être supportés pendant les opérations de mappage, puisqu’il consent de mapper les bases de données locales dans une section dédiée à ce genre d’opération ainsi que la possibilité de rechercher les codes pertinents pour certains types d’analyses. Par ailleurs, le logiciel prévoit certaines fonctions d’importation et de contrôle de la terminologie utilisée pour exprimer les examens au niveau local en permettant une éventuelle normalisation des termes locaux avec la terminologie LOINC. Les abréviations et les termes utilisés localement sont supprimés puisqu’ils ne sont pas reconnus du standard pour la recherche du code LOINC correct et ne peuvent pas être considérés comme synonymes à associer aux noms principaux. Les termes locaux sont, pour la majorité de fois, enrichis de caractéristiques propres de la langue dans laquelle sont exprimés et qui rendent, donc, difficile tel la recherche

⁵² LOINC groupe : <<https://loinc.org/adopters/regenstrief-institute-inc.html>> (Dernière consultation : 29/01/2017).

que une leur possible insertion dans la structure du standard. Le système LOINC a établi les noms officiels en basant des règles spécifiques qui doivent être respectées, aussi en considérant les variétés morphologiques qui chaque langue peut présenter.

Les fonctionnalités les plus importantes du logiciel RELMA peuvent être résumées ci-dessous :

- recherche des termes en utilisant une syntaxe de recherche similaire à celle utilisée pour les recherches avec Google (opérateurs booléens, caractère « jolly », etc.). En effet, les résultats de la recherche consistent dans la correspondance exacte des chaînes de recherches insérées (Figure 1) ;

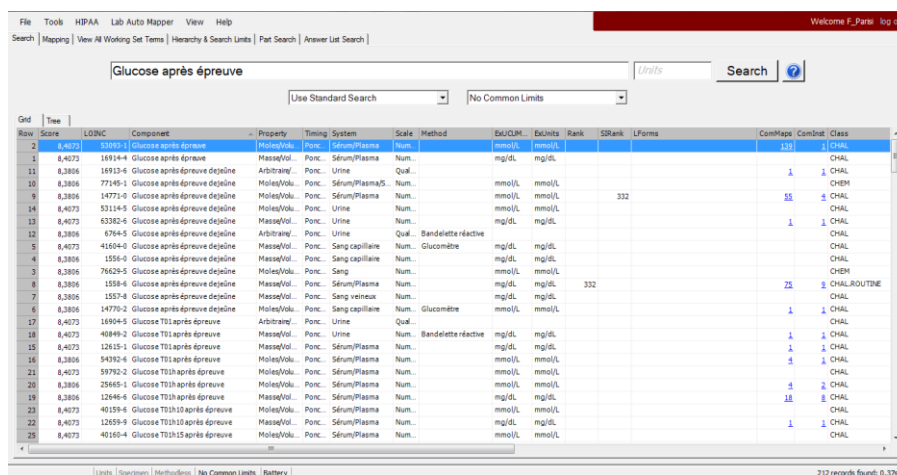


Figure 1 - Écran de recherche du logiciel RELMA

- visualisation de la page explicative de l'analyse cherchée, avec une description détaillée, une liste de termes corrélés et la traduction en différentes langues du terme candidat à être mappé ;
- effectuer des recherches détaillées et limitées, par exemple en considérant seulement un des attributs LOINC qui composent le nom complet ;

- focaliser la recherche seulement sur le *most common LOINC result*, c'est-à-dire, sur la base d'une liste de tests considérés comme étant les plus communs qui représentent 98% de ceux effectués [8] habituellement et les plus utilisés par trois organisations américaines (INPC, *Partners Healthcare System of Boston* e *United Healthcare*, UHC) ;
- mapper les codes locaux avec les codes LOINC, avec la possibilité d'ajouter des commentaires ou exporter en format Excel une sélection de codes candidats pour le mappage, pour les partager et échanger des opinions.

Avec le nombre croissant de codes et de termes LOINC, le RI a jugé nécessaire d'avoir un outil qu'en poussant consentir une gestion facile pour faciliter aussi l'éventuelle intégration ou modification de la base de données officiel. RELMA représente aussi, une très confortable solution pour les laboratoires qui ne sont pas trop experts sur la structure du standard, étant donné son approche plus immédiate par rapport à une opération de mappage manuelle (Figure 2).

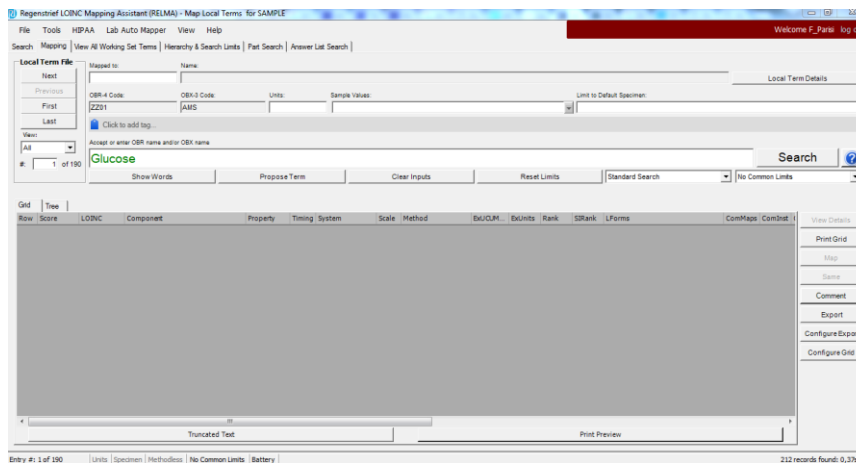


Figure 2 - Écran de support au mappage des examens locaux

Chapitre III

3.1 L'analyse du contexte et la problématique traitée

Le travail de thèse présenté répond à une exigence spécifique ressentie pendant la mise en œuvre du processus de validation des opérations de codage effectuées par les laboratoires biologiques-médicaux.

L'objet d'étude répond au besoin spécifique d'un support terminologique pour dérouler une opération importante et copieuse, même pour les experts du domaine.

Les laboratoires biologiques-médicaux, soient-ils publiques ou privés, effectuent des tests spécifiques et stockent les données dans leur système d'information. Les laboratoires d'analyses médicales possèdent ses systèmes de gestion fonctionnant sur le référentiel local du laboratoire. Les référentiels de chaque laboratoire en utilisant des terminologies différentes rendent les échanges de données difficiles entre deux structures et encore plus au niveau international ; chaque laboratoire effectue une large variété d'analyses identifiées par des formes linguistiques strictement liées à la langue d'appartenance. Les données médicales et notamment biologiques nécessitent d'une interopérabilité sémantique qui puissent les rendre utilisables en toutes les langues et en différents contextes, tel locaux que internationaux. À cet égard, la terminologie de référence LOINC (Logical Observation Identifiers Names and Codes) s'est imposée, et il est utilisé à aujourd'hui en plus de 150 Pays dans le monde. Pour contribuer à l'utilisation du standard les laboratoires biologiques médicaux doivent associer leurs propres examens aux codes officiels LOINC pour les rendre interopérables au niveau national et international.

Le processus de mappage effectué par les laboratoires médicaux concerne l'association de chaque analyse à un et un seul code officiel LOINC. Les codes

LOINC sont associés par définition à un ensemble d'attributs structurés en six parties qui identifient de façon univoque une observation déterminée. La combinaison des six paramètres LOINC est associée à un code spécifique qui le représente univoquement, c'est à dire que avec une seule valeur changée pour cette même combinaison le code LOINC associé sera différent.

Le standard LOINC permet l'identification d'un examen spécifique avec un code univoque qui reste le même indépendamment de la langue dans laquelle les paramètres sont exprimés. Par ailleurs, étant donné que le code reste invariable pour la même typologie des examens indépendamment de la langue utilisée, le codage des examens permet de surmonter tant les différences linguistiques liées aux Pays dans lesquels le standard est utilisé, que celles liées à la variété des formes locales utilisées par les laboratoires biologiques-médicaux. Pour ceux qui ne connaissent pas dans les détails les règles du standard LOINC, les valeurs des axes se présentent en forme complexe et difficile à interpréter (la majorité des paramètres se présentent sous forme d'acronymes, sigle ou conventions particulières) ainsi que la recherche des codes doit suivre des règles déterminées (par exemple en considérant la présence des symboles qui ne sont pas acceptés pour la recherche). Les laboratoires stockent l'ensemble de leurs tests effectués en utilisant des codes et des noms valides qu'au niveau local et pour laquelle l'association aux paramètres LOINC correspondants résulte très difficile.

Le niveau de détail et la précision dans laquelle LOINC différencie les examens si en décrivant les tests dans leurs totalité, ils causent problèmes aux non experts qui s'approchent pour la première fois à l'utilisation du standard et qui doivent donc, dérouler une opération si minutieuse.

L'opération d'association et de recherche des codes corrects est, donc, très difficile, car les examens de laboratoire sont exprimés en utilisant des formes linguistiques locales qui résultent souvent moins complètes par rapport aux informations requises pour identifier correctement le code LOINC. Généralement les laboratoires n'explicitent pas dans leur base de données toutes les informations correspondant aux 6 paramètres LOINC, en créant, de cette façon, des difficultés et

des problèmes pendant le mapping. En effet, la réduction des possibles codes à choisir pendant la recherche dépend fortement du nombre de paramètres insérés pour effectuer le codage ; c'est-à-dire que pour améliorer la recherche du code correct, il faut connaître un nombre significatif de paramètres LOINC, pour permettre la réduction de l'ensemble des possibles codes à choisir.

Le travail de thèse concerne la construction d'un système de grammaires formelles capable de reconnaître les formes locales utilisées par les laboratoires, de les annoter sémantiquement en considérant le langage du standard LOINC et de déterminer leur exactitude par rapport aux valeurs correspondantes aux codes associés.

Les experts du standard LOINC, en particulier les experts du groupe « LOINC ITALIA » pour les examens à coder et exprimés en langue italienne, sont appelés au contrôle des associations vers les codes officiels effectuées par les laboratoires, avant de les rendre définitives. Pour valider le mapping de la base de données locale vers les codes LOINC, il faut vérifier que chaque information liée aux analyses soit associée au paramètre standard correspondant. Si les laboratoires n'explicitent pas au niveau local toutes les caractéristiques de leurs analyses, le mapping et sa validation conséquente deviennent très compliqués.

Le manque d'informations fait partie des problématiques les plus importantes à considérer, sans oublier les problèmes liés aux aspects linguistiques. En effet, les formes linguistiques utilisées au niveau local gardent les caractéristiques morphologiques (par exemple : préfixes, suffixes, etc.) des termes liés à la langue d'appartenance (dans les cas objet d'étude, on parle que des formes en italien), en rendant très difficile la recherche du code candidat pour le mappage. Cette complexité linguistique rend difficile aux laboratoires, dans un premier temps, d'effectuer correctement le mapping de leurs analyses et d'identifier le juste code sans aucun type de support. La difficulté réside, donc, dans la variété linguistique tel des attributs LOINC que des forme linguistiques locales à associer puisque dotées de caractéristiques morphologiques différents par rapport à celles contenues dans le standard. Les différences syntaxiques et morphologiques existantes entre

anglais et italien, augmentent ces différences. L'utilisation de ces formes linguistiques locales entrave l'interopérabilité des données concernant les parcours de soins des patients qui représentent un argument très exploré au niveau national et international. Comme on a approfondi dans le premier chapitre, l'utilisation de systèmes standardisés est fondamentale pour la création d'un dossier médical qui soit interopérable au niveau national et international, afin que les informations soient compréhensibles dans différents environnements.

Le projet de thèse présenté propose une aide à la reconnaissance des formes linguistiques locales utilisées par les laboratoires, à leur collecte automatique, à leur association aux formes linguistiques standard ainsi que un support dans l'extraction des informations concernant la validation des associations effectuées. Le système de règles linguistiques considère les liens entre les attributs LOINC les plus importants pour déduire et contrôler de façon automatique leurs valeurs. Dans la majorité de cas, le lien entre les différents attributs LOINC (surtout ceux entre l'unité de mesure, la Grandeur, le Temps et l'Échelle) et les possibles déductions qui peuvent être faite à partir des valeurs déterminées, permettent d'arriver au code correct malgré le manque d'informations.

Le travail présenté propose, donc, un prototype de règles linguistiques pour surmonter les problématiques liées au manque d'information et à la variété des formes linguistiques locales utilisées par le laboratoire. Les grammaires syntaxiques construites, vérifient l'exactitude des données par rapport au domaine et au but spécifique qu'on s'adresse de réaliser.

Cette prémisse est faite pour introduire le contexte et surtout l'objectif du travail présenté, qui vise à donner une contribution pour rendre interopérables ce genre d'informations, notamment au niveau linguistique.

3.2 Le système des règles linguistiques : l'application NooJ

En considérant les problématiques présentées au paragraphe ci-dessus, ce travail de thèse, en effet, concerne la création d'un support à la validation des codes assignés aux examens par les laboratoires biologiques-médicaux par l'application de techniques de TALN et vise à tester les possibles utilités qui peuvent s'obtenir par la réalisation d'un support linguistique capable de reconnaître les formes linguistiques utilisées par les laboratoires.

L'association entre les formes linguistiques locales et les codes LOINC est une opération effectuée tant par les laboratoires pendant les opérations de mappage que par les experts LOINC pendant la validation. Les experts impliqués dans la validation doivent reconstruire le processus de mappage déroulé par les laboratoires et contrôler si l'association résultante a été effectuée correctement. Les difficultés linguistiques liées aux deux types d'activités peuvent être partagées aussi si elles nécessitent d'un support différent. Pour les opérations de validation il est suffisant de partir des formes linguistiques locales et de vérifier leur exactitude par rapport au code LOINC associé, alors que pour les opérations de mappage une interaction directe avec la base des données officielle est nécessaire afin de vérifier l'efficacité du support réalisé.

Le système LOINC n'impose pas l'utilisation des codes et des noms officiels en substitution de ceux utilisés dans les contextes locaux mais, au contraire, il les accompagne sous forme de traduction à partir d'un langage plus standard ; les opérations de mappage rendent possible l'association entre les examens locaux et les codes LOINC en maintenant le même contenu clinique.

Pour cela, l'exactitude des associations entre les codes LOINC et les examens de laboratoire est fondamentale pour contribuer à la fiabilité des données concernant la vie clinique du patient.

Avec le système des règles créé, la reconnaissance de ces formes est déroulée automatiquement en utilisant le logiciel NooJ qui a permis tant une modélisation de la connaissance liée à chacune entrée linguistique locale que l'annotation sémantique ainsi que le contrôle d'exactitude pour tous les examens décrits dans le corpus considéré [18]. Ce contrôle a été effectué par l'application des grammaires hors-contexte formalisées, capables de vérifier la structure des données et d'en déterminer l'exactitude.

En résumé, en utilisant uniquement des règles linguistiques et syntaxiques, le système réalisé est capable de reconnaître les formes linguistiques locales, de les annoter sémantiquement et de vérifier le niveau d'exactitude pour chaque examen. Cette identification permet un contrôle automatique des associations entre les formes linguistiques locales et les codes officiels LOINC et l'extraction d'informations spécifiques qui répondent par exemple à la question « combien d'examens ont seulement un élément correct par rapport au code associé ? ».

Par ailleurs, le système fonctionne comme collecteur automatique des formes linguistiques locales associées aux formes linguistiques officielles du système LOINC. Cette fonctionnalité peut donner une importante contribution au travail déroulé par le groupe LOINC Italia (décrit en paragraphe 2.2) qui donne support aux opérateurs sanitaires par la validation des codes et s'occupe de réaliser la version LOINC italienne par la traduction des parties et par l'association des synonymes aux formes officielles. L'association des synonymes locales avec les parties LOINC officielles est effectuée par le logiciel RELMA (Regenstrief LOINC Mapping Assistant). Cette opération de collecte ainsi que d'association porte importants avantages pour la recherche des codes corrects, étant que le système sera capable d'utiliser différents critères de recherche et reconnaître plusieurs paramètres exprimés dans leurs formes locales. La collecte des synonymes est très importante aussi pour faciliter l'approche au système de codage pour ceux qui ne connaissaient pas les règles syntactiques. Ces considérations soulèvent la question sur les possibles avantages qui peuvent porter la construction d'une base de données contenant tous les mappages effectués par les laboratoires

présents sur un territoire donné. La construction d'une base des données partagée et utilisée que au niveau locale entre les laboratoires, est d'une coté une bonne solution pour vérifier plus facilement les mappages effectués par les autres mais pourrait porter nombreux erreurs d'autre part. La tendance des experts de laboratoire pourrait être d'associer les analyses en prenant comme exemple le travail effectué par les autres, sans vérifier l'effective validité des opérations effectuées et l'effective correspondance avec le code LOINC choisi pour identifier l'examen, en créant dans cette façon beaucoup d'information pas fiable. L'utilisation des codes LOINC est obligatoire pour tous les laboratoires qui demandent, donc, support à ce propos.

Un système automatique de validation pourrait être une importante contribution pour améliorer les temps et l'efficacité de ce genre d'opérations.

Les phases les plus importantes suivies pour réaliser le système objet du travail présenté dans cette thèse, sont :

- *la préparation des données*. En particulier, le corpus considéré présente une structure spécifique en basant les règles de présentation des données demandées aux laboratoires ;
- *la construction des dictionnaires électroniques*. Dans la plateforme NooJ, chaque forme linguistique locale est décrite en considérant la morphologie et la connaissance intrinsèque liée au domaine d'appartenance ; cette phase agisse aussi comme collecteur des formes locales qui peuvent être insérées dans la mise à jour du groupe LOINC Italia de la traduction et synonymie de la langue italienne.
- *la gestion des ambiguïtés*. Les ambiguïtés décrites dans les dictionnaires ne sont pas liées à la morphologie des mots mais à la morphologie rapportée dans le domaine spécifique ;
- *identification des règles d'annotation sémantique* pour transformer les formes linguistiques locales en forme linguistiques acceptées par le standard LOINC ;

- *la construction de la grammaire hors-contexte* « LOINC » qui, divisée en différents niveaux d'analyses, reconnaît les éléments dans leurs différentes catégories ;
- *la définition des contraintes*, pour déterminer les niveaux d'exactitude en effectuant des contrôles spécifiques ;
- *la construction de grammaires*, pour l'extraction d'informations particulières, par exemple pour extraire les examens ayant seulement 1 élément correct ;
- *l'analyse des résultats obtenus*.

L'ensemble d'étapes expliquées ci-dessus portent à la réalisation d'un système de règles linguistiques à utiliser pour atteindre le besoin spécifique considéré dans cette thèse.

Les formes linguistiques considérées, même si dotées d'une large variété, doivent se rapporter à une terminologie standard déjà définie et détaillée. Un des aspects les plus importants de ce travail est représenté par l'utilisation d'une application de TALN (Traitement Automatique des Langues Naturelles)⁵³ réalisée à partir d'un corpus de textes structurés, techniques et contextualisés dans un domaine spécifique. La problématique liée à l'absence du contexte textuel a été surmontée par la réalisation de cette application qui, en ne considérant que de règles de nature linguistique, arrive à reconnaître et à évaluer l'exactitude des données présentées dans une structure déterminée et à donner des informations très importantes sur les opérations de mappage effectuées par les laboratoires. Le système réalisé obtient de meilleurs résultats si appliqué dans le cas où la sémantique des données détermine leur qualité, la spécificité du contexte est élevée et la validation sur l'exactitude des données demande une connaissance profonde du domaine. Le système des règles réalisé dans le cadre du suivant projet de thèse a été appliqué au corpus des examens biologiques exprimés en langue italienne.

⁵³Le terme TALN provient de la traduction de NLP (Natural Language Processing).

3.2 Le traitement automatique des langues

On regroupe sous le nom de « traitement automatique des langues » l'ensemble des recherches et développements visant à modéliser, à l'aide de machines, la capacité humaine à produire et à comprendre des énoncés linguistiques dans des buts de communication [26].

C'est vrai que l'être humain est capable d'analyser un discours ou un texte et manipuler son contenu, mais, ces dernières années, la quantité d'informations qui nous envahi notamment avec les avancées de la technologie nécessitent, pour leurs traitement et exploitation de nouvelles méthodes.

Le traitement automatique de la Langue Naturelle est une discipline scientifique assez récente. Elle ne concerne pas directement la langue, mais il porte sur les données linguistiques, les textes, codés dans une langue particulière [2]. Elle regroupe la linguistique, l'informatique et l'intelligence artificielle et vise la conception de systèmes et techniques informatiques permettant de manipuler le langage humain dans tous ses aspects.

On peut donc dire qu'une application de TALN est une manipulation d'une certaine quantité des données linguistiques d'un corpus, considérés comme objet d'entrée, dans l'objectif de les transformer en objets de sortie, autrement dit, la programmation d'un ensemble d'actions qui seront effectuées par l'ordinateur dans le but d'avoir comme résultat un système d'information enrichis par rapport au texte du départ. Ainsi le résultat peut être des propriétés syntaxiques et sémantiques d'un mot, propriétés syntaxique d'une phrase, ensemble de mots-clés, une formule logique, etc.[13].

Cette discipline est née aux États-Unis vers 1949⁵⁴ et dédiée à la conception de méthodes et d'outils informatiques pour analyser le langage naturel (écrit et/ou

⁵⁴Le TALN est né pendant la guerre froide et pour longtemps s'est concentré sur la traduction automatique avec évidemment un grand intérêt pour le passage du russe à l'anglais.

parlé) utilisé par les êtres humains. Le domaine a eu, depuis le début, diverses directions de recherches en suscitant une large variété d'intérêts pour les communautés scientifiques.

Les premiers travaux sur le domaine du TALN ont été déroulé pour la traduction automatique, avec la mise au point du premier traducteur automatique (très basilaire), dès 1954. C'est dans le contexte de la guerre froide qu'une course vers un outil capable de traduire la langue de l'autre est lancée entre américains et russes. Les projets consistaient à construire des dictionnaires et quelques règles structurelles seulement puisqu'il ne s'agit que d'une traduction mot à mot.

Depuis la publication *Structures syntaxiques* de N. Chomsky [5], publié en 1957, une autre aire linguistique s'ouvre avec l'idée de traduire formellement la langue. Plusieurs linguistes se penchent sur la question et beaucoup de travaux portant sur des grammaires génératives. Dans la même période, des travaux s'inscrivant dans la démarche des grammaires transformationnelles se sont développés notamment par Z. Harris dans, entre autres, *Structures mathématiques du langage* [14] et par Maurice Gross avec ces grammaire transformationnelles du français consacrées au verbe [10] et à la syntaxe du nom [11].

Néanmoins, dans cette période et jusqu'aux années soixante-dix, les grammaires de Chomsky sont considérées comme trop simples pour modéliser correctement l'ensemble des phénomènes de la langue. Ce constat a encouragé le développement de travaux consacrés beaucoup plus à la syntaxe et aux nouveaux formalismes de description grammaticale gérable par des grammaires algébrique.

Actuellement des travaux visent aussi à analyser et la formalisation des mécanismes d'acquisition et d'extraction automatique des connaissances, concernant le lexique, les règles grammaticales et la sémantique, en s'appuyant sur des corpus.

C'est dans la même période aussi qu'on peut également positionner la naissance de l'intelligence artificielle qui a posé la conception que tout aspect de l'intelligence humaine décrite d'une façon suffisamment précise, peut être simulé par une machine. On y discutait des possibilités de créer des programmes

d'ordinateurs qui se comportent intelligemment, et en particulier qui soient capables d'utiliser le langage. Puis, la recherche a porté vers des réflexions importantes ; la conférence mondiale de recherche en Intelligence Artificielle « IJCAI2011 » a regroupé plusieurs domaines à savoir, le traitement du langage naturel, la représentation des connaissances, le Web, planification automatique, robotique et vision, apprentissage automatique, etc.[1].

Dans le domaine du TALN, on peut distinguer des approches syntaxiques (linguistiques) liées à la théorie des langages formels et des approches numériques s'appuyant sur des probabilités et des statistiques. L'approche linguistique, en gros, prend la phrase comme unité fondamentale d'analyse et de traitement. Elle cherche surtout de simuler le processus de compréhension en encadrant la phrase dans un ensemble plus grand et cohérent (le contexte). Toutefois, la seconde approche est plus proche des mathématiques, et en particulier des probabilités. Plutôt que de construire des structures syntaxiques dans cette approche, on cherche à calculer les probabilités des cooccurrences entre mots ou expressions sans trop rentrer dans l'étude du sens.

Pour l'analyse de textes les techniques de TALN peuvent agir sous différents niveaux [2] :

- le niveau lexical/morphologique. Il considère les propriétés de mots et ses comportements au niveau morphologique. En particulier, le niveau lexical analyse leurs flexions, dérivations et leur association avec des autres mots (c'est le cas de mots composés) ainsi que leur cadre syntaxique. Ces informations sont contenues dans le dictionnaire électronique, principal outils pour adresser les buts déterminés dans ce niveau ;
- le niveau syntaxique (grammaires formelles). Il concerne les informations sur le positionnement des mots dans la phrase comment ils se combinent pour former une syntaxe correcte. L'objectif le plus important consiste en découper la phrase en syntagme pour pouvoir reconnaître les termes et repérer les dépendances syntaxiques présentes. Cette tâche est déroulée par

des grammaires formelles, un ensemble des règles explicites qui transforme une entrée en une sortie particulier ;

- le niveau sémantique. Pour analyser la sémantique d'un texte on considère la sémantique des mots afin de dégager ses différentes significations et la manière dont les sens des mots se combinent pour construire le sens global de la phrase ;
- le niveau pragmatique. Ce niveau est parfaitement dissocié du niveau sémantique ; la sémantique se préoccupe de comprendre les sens des énoncés ainsi que le pragmatique porte sur les attitudes (vérité, désirabilité, probabilité) que les locuteurs adoptent. On traite, donc, de interpréter la phrase en fonction des informations extralinguistiques, relatives à la connaissance qui n'est pas vraiment explicitée mais qui est contenu dans des informations de contexte.

3.2.1 Les applications de TALN

La discipline du TALN recouvre différents champs de recherche. Pour décrire brièvement ces domaines d'application on peut citer :

- la traduction automatique. Comme déjà expliqué, cette application a été le premier objectif souhaité au début de la guerre froide. Même si une traduction parfaite n'a pas encore vu le jour mais ce domaine reste important vu son utilité. Actuellement de nombreux logiciels, principalement commerciaux, sont disponibles (Systran, Logos, Metal, Aleth-Trad, etc.), ainsi que des sites de traduction comme *Systran* ou *Google* ;
- la correction orthographique. Des applications qui offrent les différentes règles orthographiques et grammaticales et qui permettent le respect des règles de la langue ;

- la recherche de documents et l'indexation de textes. Des millions de requêtes se font quotidiennement, parmi les systèmes les plus connus: *Google, Yahoo!, Amazon, etc.* ;
- l'apprentissage assisté par ordinateur ;
- la génération automatique de documents à partir de spécifications formelles.

Dernier point mais non des moindres, les applications de TALN permettent la recherche d'informations spécifiques à partir des textes structurés ou non structurés. Cet important aspect est, aujourd'hui, une des applications les plus importantes pour le traitement automatique des langues puisque il donne la possibilité d'explicitier une connaissance intrinsèque contenue dans les textes et strictement liée au domaine spécifique à laquelle elle se réfère.

Les plus importants avantages qu'offre le TALN au domaine médical sont, l'indexation automatique, la recherche d'informations et l'analyse de la parole. Cette utilité a permis au domaine de bénéficier de plusieurs projets importants, dont certains ont été appliqués à domaine spécifiques comme cela de la santé.

Parmi les projets les plus importants, la construction d'un système de détection des infections associées aux soins à partir de l'analyse de comptes rendus d'hospitalisations a été le projet décrit-en [12]. Ce système a pour objectif l'extraction d'informations à partir d'un ensemble de comptes rendus de médecins pour recueillir les informations relatives aux infections nosocomiales. Par ailleurs, un système d'acquisition automatique de connaissances morphologiques sur le vocabulaire médical a été lancé dans le projet décrit par N. Grabar et P. Zweigenbaum [9]: le projet a été construit à partir du glossaire SNOMED (Systematized Nomenclature of Human Medicine). Son objectif est d'extraire les informations liées à la morphologie des mots en les décomposant en racine, en préfixe et en affixe, et en les rongeant dans des familles de mots morphologiquement liées.

L'analyse de la dynamique phonétique causé par la maladie de Parkinson [22] : l'objectif de ce travail est d'identifier les difficultés de contrôle du mouvement chez les personnes atteintes de cette maladie à partir de l'étude de leur comportements

dans la production de sons de parole nécessitant des mouvements continus des articulateurs supra laryngés.

D'autres applications existent dans le domaine médical, Jean Charlet et *al.* décrivent trois d'entre elles construites à l'INSERM (Institut national de la santé et de la recherche médicale) [3]. La première a été conçue dans le projet *OntoPneumo*, elle vise à construire une ontologie de la pneumologie pour l'aide au codage des données médicales. La seconde est un projet de recherche sur les déterminants du raisonnement médical qui a abouti à la construction d'une première ontologie. Cette dernière a déjà comme résultat la proposition d'une mise à jour des formulaires d'entrées de données cliniques dans le domaine de l'hypertension artérielle. Par ailleurs, la troisième application est une application de pharmaco-vigilance fondée sur une ontologie construite spécifiquement pour augmenter les signaux de détections de causalité possibles entre un médicament et un effet indésirable. L'objectif principal de ces applications est de capter automatiquement et d'une façon précise le maximum d'informations liées aux pratiques des médecins à travers leurs vocabulaires, aux caractéristiques des patients, aux spécificités de leurs problèmes et la détection des effets indésirables par la définition formelle de chaque terme, etc.

Le projet *EPIDIMIA* construit avec le logiciel NooJ [23] s'inscrit dans une démarche qui consiste à modéliser des épidémies dans le monde. La tâche de NooJ est le dépouillement automatique des corpus susceptibles d'avoir des informations utiles comme *maladie*, *virus*, *bactérie*, etc. Le travail a été fait par étapes : construction de corpus, repérage de vocabulaire spécifique et son typage en différentes classes sémantiques, analyse des propriétés de ce vocabulaire (ex. les propriétés syntaxiques et sémantiques d'un verbe ou d'un nom), construction de grammaires permettant de reconnaître les éléments de la phrases, de les isoler et de les annoter, et enfin construction d'un langage standardisé à l'aide de vocabulaire isolés dans l'étape précédente.

Comme on peut remarquer, les utilités des applications de TALN peuvent être les plus variées, appliquées dans le cadre de différents domaines aussi spécifiques qu'ils soient utilisant des outils plus adaptés.

3.3 Les logiciels de traitement automatique des langues

Les travaux de constructions de logiciels du TALN s'appuient sur :

- *des connaissances en linguistique*, pour décrire les diverses connaissances relatives à la langue ;
- *l'aspect formel des données* pour exprimer ces connaissances dans un formalisme qui convient à un traitement automatique ;
- *l'informatique* qui utilise cette description formelle des connaissances dans une application informatique concrète.

Parmi les logiciels de TALN nous citons un sous ensemble des plus connus et des plus utilisés :

Le logiciel open-source GATE⁵⁵ a été développé en JAVA à partir du 1995. Ce système est composé par différentes parties qui permettant son application à différents types des ressources textuelles. En particulier, il est formé par différents éléments :

- des ressources linguistiques « *Language Ressource (LR)* », composées par lexiques, taxonomies, ontologies, corpus et autres ressources (support de plusieurs formats XML, RTF, HTML, SGML, etc.) ;
- des ressources pour dérouler les processus « *Processing Ressources (PR)* » composées par algorithmes et des composants qui effectuent les traitements ayant pour but d'ajouter ou de transformer des annotations au corpus (sous forme d'attributs/valeurs) ;

⁵⁵General Architecture for Text Engineering, <<https://gate.ac.uk/>> (Date de la dernière visite : 06/02/2017).

- une ressource visuelle « *Visual Ressource* » qui permet la présentation des résultats sous forme visuelle, à l'intérieur de l'environnement de développement GATE.

Ces fonctionnalités sont gérées par une collection d'objets réutilisables appelée CREOLE (*Collection of REusable Objects for Language Engineering*). CREOLE, qui représente aussi un large inventaire de ressources, donne la possibilité de définir le propre environnement de travail et les instruments nécessaires pour dérouler l'analyse.

Une autre application, NLTK⁵⁶ développée en Python pour effectuer le traitement en TAL. Elle représente un important support pour la recherche en linguistique computationnelle. Cet outil comprend une large quantité d'algorithmes pour effectuer les fonctions les plus importantes : identification des tokens, POS-tagging, clustérisations et classification. Il est très efficace surtout sur la langue anglaise mais ça marche aussi sur des autres comme par exemple l'italien. On retrouve aussi différentes librairies open-source qui sont liées à cette application comme TextBlob⁵⁷ qui utilise le répertoire de *Pattern*⁵⁸ qui contient les fonctions de web mining, TALNet machine learning. *Pattern* est un outil organisé en sections séparées différenciées par les utilités qu'on veut utiliser (par exemple : *pattern.web*, *pattern.en*, *pattern.search*, *pattern.vector*, *pattern.it*). NLTK et *Pattern* appartiennent aux librairies écrites en Python, mais il existe aussi une série de librairies développées en Java.

« OpenNLP »⁵⁹ est une librairie écrite en Java et conçue pour effectuer des activités de traitement automatique de langues. Il permet d'effectuer les fonctionnalités de base comme la tokenisation, la division en phrases, le POS-

⁵⁶ Bird, Steven. "NLTK: the natural language toolkit." Proceedings of the COLING/ACL on Interactive presentation sessions. Association for Computational Linguistics, 2006.

⁵⁷ Steven Loria, *TextBlob Documentation, Release 0.12.0.dev0*, février 2017. Disponible à l'adresse : <<https://media.readthedocs.org/pdf/textblob/dev/textblob.pdf>> .

⁵⁸ Smedt and Daelemans, 2012. Disponible au site <http://www.jmlr.org/papers/volume13/desmedt12a/desmedt12a.pdf>.

⁵⁹ Wilcock, Graham. "Text annotation with OpenNLP and Uima." (2009).

tagging et l'extraction d'informations linguistiques comme la recherche des entités nommées.

Un logiciel qui permet d'effectuer des analyses de sentiments (sentiment analysis), parsing basé sur théories de probabilités et la reconnaissance des entités nommées comme « personnes », « organisation », et « lieux », est représenté par le « Stanford CoreNLP⁶⁰ ». Cette application est développée pour reconnaître des textes en langue anglaise, chinoise, française, allemande, et espagnole, et pour l'italien il dispose d'un ensemble d'outils appelé TextPro⁶¹.

A la différence de ce type d'analyses, le logiciel XFST⁶² (Xerox Finite-State Tools) a été développé pour répondre d'une manière efficace aux problèmes morphologiques. Il est une interface qui donne accès à des opérations d'états finis (algorithmes tels que l'union, la concaténation, l'itération, l'intersection, la composition, etc.), il inclut un compilateur d'expressions régulières et il est bidirectionnel. L'interface comprend : une opération de recherche (appliquer) et une opération de génération (lookdown). Il est utilisé par exemple pour le remplacement des éléments d'un mot comme dans le passage du singulier au pluriel ou du masculin au féminin⁶³.

LFG (Lexical Functional Grammar)⁶⁴ est un logiciel adapté au traitement syntaxique de la langue. Son approche s'appuie sur la formalisation de la structure lexicale, de la structure constitutive (c-structure) et de la structure fonctionnelle (f-structure), avec possibilité de communication entre les trois niveaux. La première structure consiste à associer à l'entrée lexicale des informations de genre « *VERBE* », « *OBJET* », « *SUJET* », etc. La structure constitutive représente l'ordre des mots et leurs imbrication pour la constitution de phrases comme dans la déconstruction de d'une phrase en *SUJET-VERBE-COMPLEMENT*. En fin la

⁶⁰C. D.Manning et al., *The Stanford CoreNLP Natural Language Processing Toolkit*, 2014. Disponible à l'adresse : <<http://nlp.stanford.edu/pubs/StanfordCoreNlp2014.pdf>>.

⁶¹Pianta et al., *Teacher-Student Relationships and Engagement: Conceptualizing, Measuring, and Improving the Capacity of Classroom Interactions*, 2008. Disponible sur le site : <<https://pdfs.semanticscholar.org/530d/ec0c538bef5671ae51c152a7ebe5388cc2cf.pdf>>.

⁶²Mike Rosner, Dept ICS, <<http://cs.haifa.ac.il/~shuly/teaching/06/nlp/xfst-tutorial.pdf>>.

⁶³Pour plus d'informations voir M. Silberstein : 2015 : 42.

⁶⁴Voir l'article de Carol Neidle sur le site : <<http://www.bu.edu/asllrp/neidle-lfg.pdf>>.

structure fonctionnelle, qui consiste à unifier les précédentes informations ; ce niveau est représenté hiérarchiquement par les matrices AVM (*Attribute Value Matrix*) qui contiennent les informations sur le mot, sur les constituants de la phrase et sur la phrase elle-même.

HPSG (Head-driven Phrase Structure Grammar) est une plateforme qui tente, comme celle de LFG, de traduire formellement la syntaxe de n'importe quelle phrase. Elle est descendante directe de la grammaire syntagmatique généralisée (GPSG) et indirectement, de la grammaire fonctionnelle d'unification (FUG). HPSG présente l'avantage d'avoir des structures de traits comme seule structure de données contenant chaque aspect de la description linguistique (sémantique, syntaxe, phonologie, etc.) et présentant leurs interactions⁶⁵. Très proche aussi de la démarche formalisée de LFG, la méthode de HPSG repose sur des niveaux de représentations distincts à savoir, le niveau lexical qui s'occupe de l'organisation des propriétés phonologiques, syntaxiques et sémantiques de l'entrée, le niveau phrastique qui s'occupe de l'organisation des éléments dans leur constitution en phrases et le niveau des dépendances dédié à la description et à la formalisation des relations entre les constituants de la phrase. Et cela se fait d'une manière qui permet à chaque niveau d'hériter les informations contenues dans les autres niveaux⁶⁶.

3.3.1 Le choix du logiciel NooJ

Dans cette section, les caractéristiques du logiciel NooJ ainsi que les fonctionnalités possédées et qu'ont nous conduit à le choisir pour adresser les objectifs de cet travail de thèse ont été décrites

NooJ est un environnement de traitement automatique des langues naturelles conçu par Max Silberztein, à partir de 2002 pour surpasser les insuffisances de

⁶⁵Anne ABEILLE, *Les grammaires d'unification*, Hermès Lavoisier, 2007, 342 pages, ISBN 978-2-7462-1251-0, p. 158.

⁶⁶Plus des détails voir l'article de Robert D. Levine et W. Detmar Meurers sur l'adresse : <http://lingo.stanford.edu/sag/papers/meur-lev-hpsg.pdf>.

l'outil INTEX⁶⁷. L'objectif du logiciel NooJ est de formaliser les différentes spécificités des langues d'une manière simple et efficace, et dans le respect des principes linguistiques. Le projet de construction du logiciel NooJ s'inscrit au cœur de la linguistique : il a pour objectif de décrire les langues avec exhaustivité et une précision absolue l'ensemble des phrases d'une langue susceptibles d'apparaître dans les textes écrits. Le projet répond surtout à deux besoins : il s'agit d'une part de fournir aux linguistes des outils pour les aider à décrire les langues d'une façon exhaustive, et d'autre part de construire des logiciels capables de traiter automatiquement des textes rédigés en langue naturelle [24].

La plate-forme NooJ est développée pour atteindre différents buts et pour agir d'une façon plus efficace par rapport aux autres systèmes. Parmi ses spécificités [24]:

- son adaptation aux différents phénomènes de la langue. Il offre un analyseur spécial pour chaque phénomène, par exemple la morphologie avec des outils conçus spécialement pour la morphologie et la dérivation avec des outils conçus pour les phénomènes de dérivation, etc. ;
- sa capacité à associer les différents résultats obtenus. Chaque analyseur peut communiquer ses propres résultats à un autre analyseur ; les données morphologiques par exemple peuvent être exploitées facilement et sans aucune autre intervention par l'analyseur syntaxique ;
- son traitement du corpus en cascade : le traitement des données d'un texte se fait par niveau. L'analyseur morphologique s'occupe des informations liées à la morphologie, celui de la syntaxe s'occupe des informations syntaxiques, etc.

Dans son ouvrage *La formalisation des langues, l'approche NooJ* [24], l'auteur a tenu aussi à clarifier la différence de cette approche par rapport aux méthodes statistiques qui consistent étiqueter les mots en les comparant à un corpus de

⁶⁷Max SELBERZTEIN et Agnès TUTIN, « *NooJ, un outil TAL pour l'enseignement des langues. Application pour l'étude de la morphologie lexicale en FLE* », ALSIC [En ligne], vol.8, No. 2/2005.

référence. L'auteur a résumé les principaux problèmes et ambiguïtés de cette méthode dans les points suivants :

- si l'étiqueteur ne trouve pas dans le corpus de référence le mot à étiqueter ce dernier peut être associé à n'importe quelle autre étiquette en se basant sur des critères *incertains* comme la première lettre en majuscule d'un mot qu'on qualifiera d'un nom propre ou d'un nom commun ; à force d'apparaître dans le corpus de référence avec la même étiquette, cette dernière sera associée automatiquement au mot en question ;
- le contexte d'apparition du mot à étiqueter devient la principale référence lorsque différentes étiquettes conviennent au mot à étiqueter. Ainsi si le mot *place* apparaît entre *le* et *de* il sera considéré comme nom, comme dans l'expression « *la place de la république* », mais ce sera une ambiguïté dans des cas comme dans l'expression « *il la place de son côté* » [24].

Cette méthode est donc purement machinale et vide de tout traitement morphologique, syntaxique ou sémantique, autrement dit, loin de tout traitement linguistique de la langue. Étant nos corpus très spécifiques et composés par des phrases spécifiques et structurées, et étant le langage de référence très spécifique et standard, l'utilisation de l'approche statistique pourrait nous apporter différentes typologies de problèmes ; en premier, l'absence d'un corpus de référence ou d'une structure ontologique rend impossible la reconnaissance des entités retrouvées dans le corpus local analysé appartenant à un langage très spécifique. L'utilisation d'un corpus de référence appartenant au même domaine d'application provoque nombreuses cas d'ambiguïté que dans le contexte considéré dans ce travail de thèse et en considérant les objectifs à adresser, ça pourrait représenter un problème important. Le logiciel NooJ a été choisi pour dérouler le travail de thèse présenté. NooJ permet de personnaliser sa propre analyse textuelle et d'adapter parfaitement les ressources linguistiques aux domaines dans lesquelles ces-mêmes analyses sont déroulées et en apportant une valeur ajoutée à la qualité des résultats. Les motivations qui nous ont conduits vers l'utilisation de cet outil reposent surtout sur le besoin de traiter des données appartenant aux domaines hautement

spécifiques et sur la nécessité d'adopter une approche qui puisse garantir une haute précision concernant les objectifs.

Le fait que le contexte à analyser dans cette thèse est fortement spécifique, il y avait la nécessité d'utiliser un outil capable de modéliser les connaissances liées aux formes linguistiques, la nécessité d'associer des informations aux entrées lexicales, et celle d'obtenir exactement le résultat requis par une formalisation précise et en permettant l'extraction d'informations spécifiques. Des tâches que NooJ nous a permis de réaliser avec des grammaires formelles spécifiques d'annotations et d'extractions d'informations comme nous l'avons souhaité. Les grammaires réalisées permettent de transformer une langage commun aux autrement dit locale à un contexte spécifique dans un langage standard ayant une structure et une syntaxe spécifique.

La particularité de NooJ repose aussi dans la capacité de modéliser les entités linguistiques en faisant des deductions sur le domaine spécifique d'appartenance, en associant de classes déduites par leur valeur. Par exemple, il est possible d'associer une unité de mesure à une propriété spécifique ainsi que à un typologie de résultat (en faisant l'association avec les axes « property » et « scale » appartenent au standard LOINC). Ces capacité de modélisation ainsi que de deduction permet de surmonter les difficultés liées à l'absence d'une structure ontologique pour le standard officiel LOINC qui puisse décrire les propriétés entre les attributs appartenants aux classes principales.

Le choix liés à l'utilisation du logiciel NooJ sont très liés aux avantages que cette utilisation pourrait nous apporter en terme de modélisations des entités locales avec les entité standard, et des règles sémantiques capables de valider l'association avec les codes LOINC officiels ou suggérer le code correcte en partant des entités locales.

Chapitre IV

4.1 Le corpus

Le corpus de textes techniques-médicaux considère les données de 8 laboratoires différents, pour un total de 375 associations entre les données locales et le code LOINC (Logical Observation Identifiers Names and Codes). Le travail d'annotation et d'évaluation des données a été effectué en considérant les associations aux codes LOINC effectuées par les laboratoires et pour lesquelles les experts doivent déterminer le niveau d'exactitude. Chaque association représente le résultat des opérations de mappage des données locales vers un code LOINC effectué par les laboratoires.

Les 375 examens considérés ont été obtenus par une sélection de ceux qui appartiennent aux catégories choisies pour tester le système. En particulier, les catégories des examens considérées représentent celles avec une variété linguistique la plus élevée et pour laquelle le mappage avec les codes LOINC est effectué plus souvent. Les tests faisant partie du corpus ont été sélectionnés dans les bases de données locales des laboratoires en considérant l'appartenance à une catégorie déterminée.

Les 33 catégories d'examens considérées ont été définies par les experts les plus fréquentes et pour lesquelles le système de contrôle réalisé pouvait apporter des avantages importants.

On considère particulièrement les catégories génériques suivantes : « Acido Urico », « Acido Vanilmanelico », « Alalina », « Albumina », « Azoto », « Bilirubina », « Calcio », « Cloro », « Colesterolo », « Creatinina », « Emoglobina », « Epinefrina », « Estradiolo », « Estrone », « Follitropina », « Fosfatasi alcalina », « Fosfato », « Glucosio », « Aspartato transaminasi »,

« Alalina transaminasi », « Fattore di crescita insulino simile », « Insulina »,
« Lattico deidrogenasi », « Potassio », « Pregnenolone »,
« Progesterone », « Prolattina », « Proteina », « Sodio »,
« Trigliceridi », « Urobilinogeno », « Vitamina D ».

Ces catégories représentent le point de départ pour la réalisation des dictionnaires électroniques utilisés. Les entrées lexicales seront formalisées, comme on approfondira après, en considérant toutes les possibles flexions ainsi que la connaissance de domaine liée à chaque forme linguistique. Aux 33 catégories, on ajoute celles qui se réfèrent au groupe d'examens effectués dans l'hémogramme.

En particulier : « Neutrofili », « Linfociti », « Monociti », « Eosinofili »,
« Basofili », « Grandi cellule non colorate », « Emoglobina », « Ematocrito »,
« Volume globuli medio », « Contenuto medio emoglobina », « Concentrazione media emoglobina », « Distribuzione volume globuli rossi », « Distribuzione volume emoglobina media », « Aggregazione piastrinica », « Conta piastrinica »,
« Distribuzione piastrinica ampiezza », « Piastrine volume medio »,
« Piastrinocrito ».

La sélection des examens faisant partie du corpus a été effectuée en considérant ces catégories, exprimées en forme basilaire et en considérant toutes les flexions possibles. C'est-à-dire que, après avoir sélectionné tous les codes LOINC appartenant à chaque catégorie (grâce au logiciel RELMA - Regenstrief LOINC Mapping Assistant), le corpus a été obtenu par une comparaison croisée entre les résultats de cette sélection et les examens contenus dans les bases de données locales considérées. Cette opération a été effectuée aussi pour associer les formes linguistiques officielles à chaque code LOINC retrouvé dans les bases de données locales.

Le corpus résultant a été construit en considérant une structure de données précise, correspondant à celle qui est demandée aux laboratoires pour effectuer les opérations de codage et dérivée par la structure des codes LOINC dans les 6 axes pour identifier les examens. Les associations décrites en utilisant cette structure représentent les phrases sur lesquelles l'analyse sémantique a été effectuée. On

utilise les descriptions des examens comme phrases de laquelle on connaît déjà la structure exacte, l'ordre précis des éléments constituants.

La structure des phrases du corpus reportent les formes linguistiques avec laquelle les laboratoires expriment les valeurs des attributs LOINC.

Précisément, la phrase sera ainsi composée :

« Composant » - « unité de mesure » - « méthode d'analyse » - « milieu biologique » - « échelle de mesure » - « code LOINC associé ».

La Figure 3 montre la structure du corpus obtenu.

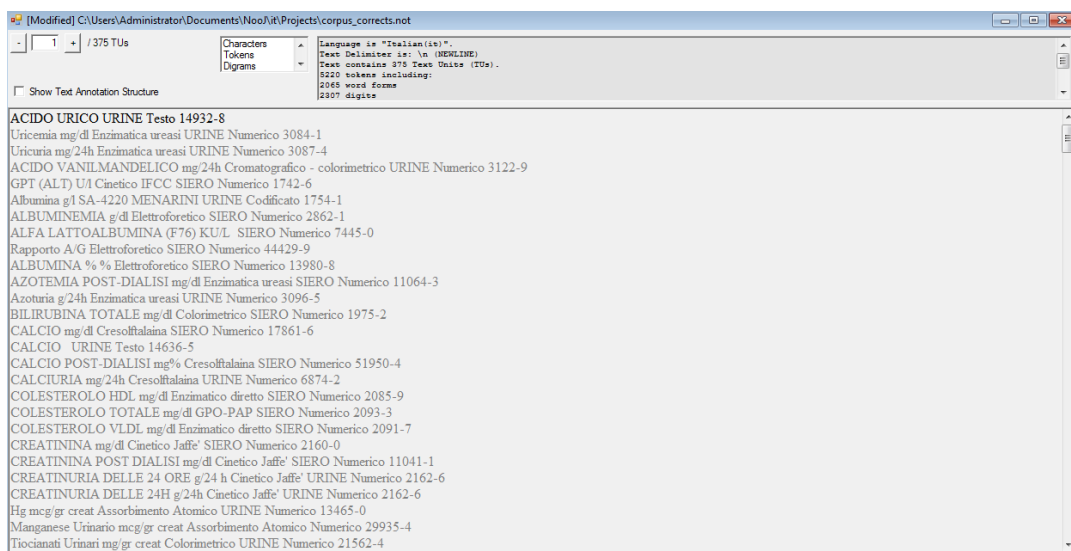


Figure 3 - Le corpus

Ce passage est très important, puisque le manque de contexte textuel pouvait provoquer beaucoup d'ambiguïtés morphologiques, sans considérer les difficultés pour la reconnaissance des mots. En ayant une structure des phrases déterminée *a priori* et la possibilité d'identifier les règles d'annotation précises, la structure du corpus obtenue a permis d'éviter ce genre de problèmes dans notre analyse.

4.2 Les dictionnaires

Le logiciel linguistique NooJ donne la possibilité de créer les dictionnaires électroniques nécessaires à la reconnaissance des mots, simples et composés.

NooJ exploite ses dictionnaires dans l'analyse des mots et des phrases, en décrivant chaque entrée lexicale d'une manière précise et exhaustive. Les entrées du dictionnaire NooJ contiennent, dans leur description, propriétés morphologiques, syntaxiques et sémantiques.

Les propriétés morphologiques représentent les différentes formes conjuguées sous lesquelles un mot peut apparaître (par exemple le mot « *écrire* » peut passer à « *écrit, écrivons, écrira* ») ainsi que ses possibles formes dérivées (par exemple « *écrire* » donnera *écriture* et *écrivain*) (appelée : morphologie dérivationnelle).

Les *propriétés syntaxiques* représentent des informations liées au mot dans sa combinaison avec d'autres mots pour la formation de la phrase. Un mot peut être par exemple un *Nom*, un *Verbe*, un *Adjectif*, un *Adverbe*, une *Préposition* et avoir un rôle spécifique s'il est inséré dans une structure (combinaison de mots) déterminée.

Les *propriétés sémantiques* expriment les différentes significations que contient le mot. Le mot « *cheval* », par exemple, a les propriétés suivantes : *animal, mammifère, il a quatre pattes, moyen de transport*, ou bien les propriétés *fruit ou aliment* pour le nom d'un fruit.

L'entrée du dictionnaire est appelée ALU [24] (Atomic Linguistique Unit) et il peut avoir différentes formes. Il peut être :

- un *mot simple*, donc, une séquence de lettres entre deux blancs. Par exemple la phrase *il regarde la télévision*, contient quatre mots simples (il, regarde, la, télévision) ;
- un *morphème* représenté par les préfixes, les racines, les suffixes ; des parties de mots pas forcément délimitées par des séparateurs. Par exemple la phrase *la réorganisation de la maison*, contient quatre mots simples (*la, de, la, maison*) et trois morphèmes (*ré-, -organis-, et -ation*) ;

- un mot composé, donc, une seule unité qui apparaisse avec des séparateurs. Par exemple la phrase *il mange une pomme de terre*, contient trois mots simple (*il, mange, une*) et un mot composé (*pomme de terre*) ;
- une expression composée par des unités qui apparaissent sous séquences discontinues de lettres et de séparateurs. Par exemple la phrase *il ne parle pas*, contient deux mots simples (*il, parle*) et une expression (*ne---pas*).

Les dictionnaires construits ici avec le logiciel NooJ, ont été réalisés en suivant une syntaxe particulière et en décrivant les entrées lexicales dans toute leurs caractéristiques morphologiques, sémantiques et syntaxiques ainsi que leurs flexions.

La valeur ajoutée donnée par le logiciel NooJ dans ce travail, provient de la possibilité de décrire les ALU avec leurs caractéristiques morphologiques ainsi que par des propriétés qui explicitent la connaissance liée au domaine spécifique d'appartenance. On pense par exemple à l'importance de la sémantique de certains termes pendant des opérations déterminées déroulée dans des contextes spécifiques où la connaissance des experts est fondamentale. La possibilité, donc, de personnaliser les ressources en les adaptant au domaine spécifique considéré, a été une des motivations les plus importantes pour le choix du logiciel à utiliser.

4.2.1 Le dictionnaire électronique : « Les Analytes »

Le dictionnaire « Analiti.dic » décrit toute les formes linguistiques locales utilisées par les laboratoires pour exprimer le paramètre LOINC « Composant ». Dans ce dictionnaire on peut distinguer différents typologies d'ALU : les mots simples, les mots composés, les formes locales accompagnées des symboles qui suggèrent une sémantique précise, les acronymes, les synonymes, les formes ayant suffixes déterminés et qui peuvent créer des ambiguïtés sémantiques dans le contexte spécifique et enfin les formes locales liées à différentes formes officielles.

Il contient 685 entrées lexicales décrites en relation à leur nature morphologique mais surtout en basant leur sémantique par rapport au système LOINC.

Pour bien comprendre de quelle manière la connaissance liée aux ALU a été décrite, il faut se référer à leurs propriétés divisées en : propriétés liées à la morphologie des mots, propriétés liées au domaine spécifique.

Le dictionnaire des analytes, étant capable de décrire la connaissance morphologique et spécifique du domaine, se présente en forme complexe et très varié.

Les définitions d'analyses contenues dans le dictionnaire « Analiti.dic » sont très variées et se différencient par : la morphologie et la nature des mots, et les catégories identifiées en considérant le domaine spécifique.

Pour la morphologie des ALU on différencie les mots simples, les mots composés et les morphèmes. À la catégorie générique de noms (le symbole « N » au début de chaque formalisation dans le dictionnaire) nous avons ajouté des autres catégories sémantiques en liaison avec le domaine d'application et, clairement, avec la langue d'appartenance. En particulier, nous avons distingués la catégorie « Acidi » (propriété « +AC »), les acronymes (propriété « +Acronimi »), les synonymes (propriété « +SIN »).

Les *ALU simples* (Figure 4), en général, les analytes qui se présentent comme un mot unique. Elles sont identifiées par les propriétés « +Analita » et « +CORE ».

La propriété « +Analita » a été utilisée comme pour toutes les unités faisant partie de ce dictionnaire, en vue d'une possible identification et reconnaissance par les grammaires développées.

« +CORE » identifie les entrées, mots simples et composés, qui représentent l'élément défini « *Core Component* », c'est-à-dire l'élément objet de l'analyse (par exemple : « glucose », « azote », « calcium », etc.), et laquelle définition est privée d'autres possibles morphèmes. Ce genre d'entrées, par contre, peuvent créer aussi des ambiguïtés en référence au domaine spécifique ; par exemple une ALU simple peut être assignée à différentes valeurs pour le paramètre composant.

Si on regarde l'image ci-dessus, on peut vérifier comment la description de chaque ALU a été effectuée en considérant leur nature morphologique et leur sémantique dans le contexte spécifique. Cette dernière est considérée en référence aux possibles valeurs assumées de paramètres du système de codage LOINC.

```

#####ALU simples

Albumina,N+Analita+CORE+Component=Albumina+FLX=F22
Aptoglobina,N+Analita+CORE+Component=Aptoglobina+System=Siero/Plasma+FLX=F22
Conalbumina,N+Analita+CORE+Component=Conalbumina Ab.IgE+FLX=F22
Azoto,N+Analita+CORE+Component=Azoto+FLX=F112
Azoto,N+Analita+CORE+Component=Azoto ureico+FLX=F112
Calcio,N+Analita+CORE+Component=Calcio+FLX=F1
Calcio,N+Analita+CORE+Component=Calcio.ionizzato+FLX=F1
Cloro,N+Analita+CORE+Component=Cloro+FLX=F1
Colecalciferolo,N+Analita+CORE+Component=Colecalciferolo+FLX=F7
Alanina,N+Analita+CORE+Component=Alanina+FLX=F6
Bilirubina,N+Analita+CORE+Component=Bilirubina+FLX=F2
Carbossiemoglobina,N+Analita+CORE+Component=Carbossiemoglobina+FLX=F22
Carbossiemoglobina,N+Analita+CORE+Component=Carbossiemoglobina/Carbossiemoglobina.totale+FLX=F22
Colesterolo,N+Analita+CORE+Component=Colesterolo+FLX=F7
Cortisolo,N+Analita+CORE+Component=CortisoloComponent+FLX=F1
Creatinina,N+Analita+CORE+Component=Creatinina+FLX=F2
Deossiemoglobina,N+Analita+CORE+Component=Deossiemoglobina+FLX=F22
Deossiemoglobina,N+Analita+CORE+Component=Deossiemoglobina/Emoglobina.totale+FLX=F22
Emoglobina,N+Analita+CORE+Component=Emoglobina+FLX=F22
Estradiolo,N+Analita+CORE+Component=Estradiolo+FLX=F7
    
```

Figure 4 - Le dictionnaire « Analiti.dic » : les ALU simples

En général, chaque ALU contenue dans le dictionnaire « Analiti.dic », indépendamment de la typologie, est décrite par la propriété « +Component » qui prend comme valeur la forme linguistique officielle correspondant au paramètre composant. La valeur de cette propriété représente la forme linguistique officielle pour l'attribut « Component » auquel la définition de l'examen local pourrait être associé.

Les ALU simples contenues dans le dictionnaire « Analiti.dic », représentent tous les analytes exprimés en forme simple par un mot. Si on les réfère au domaine spécifique analysé, les ALU simples peuvent créer des ambiguïtés qui nécessitent d'être considérées afin de les traiter et résoudre. Les ALU simples se différencient en basant leur sémantique dans le domaine considéré en identifiant pour chacune les différents cas à traiter.

On considère l'exemple 1) d'ALU comme décrite ci-dessus :

1) « *Albumina,N+Analita+CORE+Component=Albumina* »

Cette description simple indique que l'ALU « Albumina » est un analyte (propriété « +Analita »), qui représente l'élément testé, qu'il est, donc, le « core » de l'examen (propriété « +CORE ») considéré, et qu'il peut être associé à tous les codes du système LOINC ayant « Albumina » comme valeur officiel du composant (propriété « +Component=Albumina »).

Ceci est un exemple d'ALU simple associée à une valeur pour le paramètre « Component » parfaitement égal même à l'ALU décrite sans créer aucun type d'ambiguïté, et pour lequel le choix du code LOINC est déterminé surtout en considérant les autres attributs.

2) « *Conalbumina,N+Analita+CORE+Component=Conalbumina Ab.IgE* »

Le cas 2) représente une ALU simple associée à une valeur officielle du composant exprimé en forme composée.

Dans ce cas on a décrit, que l'unité simple « Conalbumina » doit être forcément associée aux codes ayant comme valeur « Conalbumina Ab.IgE » du « Composant ». Le sigle exprime la typologie de l'élément mais cette unité n'est clairement utilisée que dans la définition prévue par le système LOINC et rarement dans la forme linguistique locale.

Cet exemple est une petite démonstration pour montrer comment une forme locale (dans ce cas « Conalbumina ») peut être associée à la forme officielle exprimée par des règles spécifiques. Cette liaison (ALU simple – définition officielle) semble être banale, mais c'est une partie de connaissance du domaine fondamentale pour comprendre la complexité du système LOINC.

Dans cette façon, le lien « forme locale- définition officielle » est formalisé et reconnu par le système, qui « comprend » les valeurs de certaines propriétés et effectue ses analyses en considérant ces informations.

- 3) « *Azoto,N+Analita+CORE+Component=Azoto* »
« *Azoto,N+Analita+CORE+Component=Azoto ureico* »
« *Calcio,N+Analita+CORE+Component=Calcio* »
« *Calcio,N+Analita+CORE+Component=Calcio.ionizzato* »

Les exemples décrits dans le cas 3), montrent les possibles cas d'ambiguïtés créés à partir des ALU simples. En analysant le corpus des examens exprimés en formes linguistiques locales, on a eu la possibilité de vérifier que parfois la même définition simple peut être associée aux différentes valeurs du composant.

Dans le cas 3) l'ALU « Calcio » peut être associé aux codes avec définition égal à « Calcio » ou « Calcio.ionizzato » sans aucune règle de distinction définie. Ces considérations viennent de l'expérience des experts LOINC, qui tiennent en compte de la possible présence d'informations en forme tacite et partagée localement sans être explicitée complètement dans la base de données. Les experts de laboratoire considèrent certaines informations comme allant de soi, alors que pour qui doit trouver le code LOINC correct en se basant que sur les informations explicites, il y a plus des difficultés. L'exemple 3) montre comme peut être difficile trouver le code correct ainsi que valider le choix éventuel effectué par les laboratoires, sans considérer la connaissance et l'expérience des experts.

Les *ALU composées* (Figure 5), décrites dans le dictionnaire sont toutes les définitions des analyses exprimées par un ou plus mots. La majorité des examens décrits dans le dictionnaire se présentent en forme composée, étant aussi la variété des acronymes et des symboles utilisés pour accompagner le « core » de l'analyse.

```

#####ALU coposées
Ab Insulina,N+Analita+CORE+Component=Insulina Ab+FLX=F8
Ab anti insulina,Ab Insulina,N+Analita+CORE+Component=Insulina Ab+FLX=F8
17 beta estradiolo,Estradiolo,N+Analita+CORE+Component=Estradiolo+FLX=F7
17 Idrossi progesterone,17-Idrossiprogestosterone,N+Analita+CORE+Component=17-Idrossiprogestosterone+FLX=F7
17-beta estradiolo,Estradiolo,N+Analita+CORE+Component=Estradiolo+FLX=F7
Alanina aminotransferasi,N+Analita+CORE+Component=Alanina aminotransferasi+FLX=F6
Albumina/Globulina,N+Analita+CORE+Component=Albumina/Globulina+FLX=F6
Albumina/Proteina.totale,N+Analita+CORE+Component=Albumina/Proteina.totale+FLX=F22
Albumina %,Albumina/Proteina.totale,N+Analita+CORE+Component=Albumina/Proteina.totale+FLX=F22
Complemento C1 esterasi inibitore,N+Analita+CORE+Component=Complemento C1 esterasi inibitore+FLX=F7
Complemento C2,N+Analita+CORE+Component=Complemento C2+FLX=F7
Complemento C3,N+Analita+CORE+Component=Complemento C3+FLX=F7
Complemento C4,N+Analita+CORE+Component=Complemento C4+FLX=F7
C1 inibitore,Complemento C1 esterasi inibitore,N+Analita+CORE+Component=Complemento C1 esterasi inibitore+FLX=F7
C2,Complemento C2,N+Analita+CORE+Component=Complemento C2+FLX=F7
C3,Complemento C3,N+Analita+CORE+Component=Complemento C3+FLX=F7
C4,Complemento C4,N+Analita+CORE+Component=Complemento C4+FLX=F7
Emoglobina A,N+Analita+CORE+Component=Emoglobina A+FLX=F22
Emoglobina A1,N+Analita+CORE+Component=Emoglobina A1/Emoglobina.totale+FLX=F22
Emoglobina A1c,N+Analita+CORE+EMO+GR+Component=Emoglobina A1c+FLX=F22
Emoglobina A1c,N+Analita+CORE+EMO+GR+Component=Emoglobina A1c/Emoglobina.totale+FLX=F22
Emoglobina A2,N+Analita+CORE+Component=Emoglobina A2+FLX=F22
Emoglobina F,N+Analita+CORE+Component=Emoglobina F+FLX=F22
    
```

Figure 5 - Le dictionnaire « Analiti.dic » : les ALU composées

Comme pour les ALU simples, il est possible pour les ALU composées d'identifier différents cas à traiter.

- 1) « Ab Insulina,N+Analita+CORE+Component=Insulina Ab »
 « Ab anti insulina,Ab Insulina,N+Analita+CORE+Component=Insulina Ab »

Dans l'exemple d'ALU montrées en 1), on peut vérifier comment des entrées composées peuvent être associées à la valeur du paramètre « component » correspondent. L'entrée « Ab Insulina », par exemple, forme linguistique plus utilisée dans les bases de données locales, est associée à la valeur LOINC « Insulina Ab » puisque le système LOINC demande que le « Core Component » soit en tête à la description.

Par ailleurs, les règles de formalisation de dictionnaires NooJ permettent d'associer une forme linguistique à sa forme dérivée ou associée ; ce concept a été réutilisé pour décrire les ALU en référence au domaine spécifique, c'est-à-dire pour décrire une forme linguistique qui peut être utilisée en substitution d'une autre, sans être synonyme.

Par exemple, la forme linguistique « Ab anti insulina » est juste une façon locale d'exprimer « Ab Insulina » sans changer les propriétés principales. En effet, les

deux sont associées à la même valeur du paramètre « Component » dans le système LOINC, qui ne reconnaît que la forme linguistique officielle ou celle qui se rapproche le plus possible. Cet exemple montre comment une entrée peut être associée à sa forme linguistique substitutive qui la ressource linguistique nécessite de reconnaître puisqu'elle est plus utilisée au niveau locale et donc plus probable de retrouver dans le corpus.

L'association entre deux formes linguistiques est formalisée par la virgule : la forme linguistique décrite est associée à l'élément qui compare après la virgule, par exemple « Ab anti insulina, Ab Insulina » décrit que la définition « Ab anti insulina » provient et/ou elle est associée à « Ab Insulina ».

- 2) « Emoglobina A,N+Analita+CORE+Component=Emoglobina A »
« Emoglobina A1,N+Analita+CORE+Component=Emoglobina A1/Emoglobina.totale »
« Emoglobina A1c,N+Analita+CORE+EMO+GR+Component=Emoglobina A1c »
« Emoglobina A1c,N+Analita+CORE+EMO+GR+Component=Emoglobina
A1c/Emoglobina.totale »

Dans l'exemple 2), il est possible de vérifier les formes d'ALU composées par une ALU simple et un sigle qui en identifie la nature. L'élément « Emoglobina » suivi par les sigles « A, A1, A1c, A2, F » assume une sémantique différente et il est associé distinctement à l'attribut « Component ».

Il est possible de vérifier par exemple que la forme linguistique « Emoglobina A1 » est associée à la seule possible valeur pour l'attribut « Component » égal à « Emoglobina A1/Emoglobina.totale ». Ce lien est un clair exemple de connaissance formalisée, tenu compte que l'association entre les deux formes n'est pas du tout évident aux non experts.

Même chose aussi pour la forme « Emoglobina A1c » peuvent se créer des ambiguïtés ; la forme linguistique « Emoglobina A1c » peut être associée aux codes LOINC ayants comme valeur de l'attribut « Component » la forme « Emoglobina A1c » ou la forme « Emoglobina A1c/Emoglobina.totale ».

Parfois, on ne retrouve pas un symbole (par exemple « % ») ou d'autres éléments qui peuvent suggérer un rapport comme valeur du composant, et les ambiguïtés restent difficiles à résoudre. L'ambiguïté donnée par la présence des symboles est montrée dans l'exemple ci-dessus.

- 3) Albumina %,Albumina/Proteina.totale,N+Analita+CORE+Component=Albumina/Proteina.totale
Alfa2 %,Alfa 2 globulina,N+Analita+CORE+Component=Alfa 2 globulina/Proteina.totale

Dans l'exemple 3), il est possible de vérifier comment un symbole présent dans la définition des examens peut déterminer une valeur différente de l'attribut « Component ». Le symbole « % », utilisé dans les formes linguistiques locales n'est pas reconnu par le logiciel qui permet d'effectuer les opérations de mappage, malgré son importante sémantique. Ce symbole, en effet, identifie l'unité de mesure mais peut se refléter aussi dans le composant et dans la grandeur. C'est-à-dire que sa présence peut permettre de déduire la valeur d'attributs LOINC différents. Par exemple pour le composant (attribut LOINC « Component ») peut signifier d'avoir un rapport.

Partant, la propriété « +Component » pour la forme linguistique « Albumina % », différemment que pour la forme « Albumina », prend la valeur « Albumina/Proteina.totale » pour identifier le composant. Cet exemple montre clairement comment et pourquoi une reconnaissance des formes linguistiques locales accompagnées par toutes les possibles parties adjonctives est nécessaire pour l'amélioration des opérations de mappage et de validation des associations effectuées. Même chose pour l'ALU « Alfa 2 % » associée à la valeur «Alfa 2 globulina/Proteina.totale».

- 4) *BCR-ABL1 e1a1,N+Analita+CORE+Component=BCR-ABL1 e1a1 proteina di fusione+System=Sangue/Tess*
BCR-ABL1 e1a2,N+Analita+CORE+Component=BCR-ABL1 e1a2 proteina di fusione+System=Sangue/Tess

Le cas montré dans 4) est un exemple particulier d'entrée composée. La forme « BCR-ABL1 e1a1 » est une forme linguistique composée par des sigles, des symboles et des nombres, mais qui n'est pas considérée comme un acronyme. Cette forme, en effet, représente elle-même la définition de l'examen et la valeur du composant associé.

Les *morphèmes* considérés dans ce dictionnaire se réfèrent aux suffixes « *emia* » et « *uria* » qui déterminent des formes linguistiques utilisées pour décrire les examens et pour laquelle l'analyse sémantique est nécessaire.

Les formes obtenues par les morphèmes considérés, en effet, représentent des formes linguistiques très utilisées dans le langage commun mais qui ne sont pas reconnues comme définitions LOINC étant une traduction à partir de sa forme originelle en anglais pour comme le système a été conçu.

L'aspect le plus intéressant lié à la reconnaissance de ces formes, repose sur les ambiguïtés créées en référence au domaine spécifique. En particulier, les définitions composées par le morphème « *emia* », permettent d'identifier un ensemble possible des valeurs pour le paramètre LOINC « System » (milieu biologique) sans en identifier le seule exact, en créant des résultats qui peuvent être ambiguës.

Par la formalisation (Figure 6) et la suivante annotation sémantique des formes linguistiques locales effectuées avec NooJ, il est possible de décrire et résoudre les ambiguïtés dérivées par les morphèmes, en identifiant la valeur qui peut être considérée correcte (l'analyse de ces aspects est renvoyée aux prochains paragraphes).

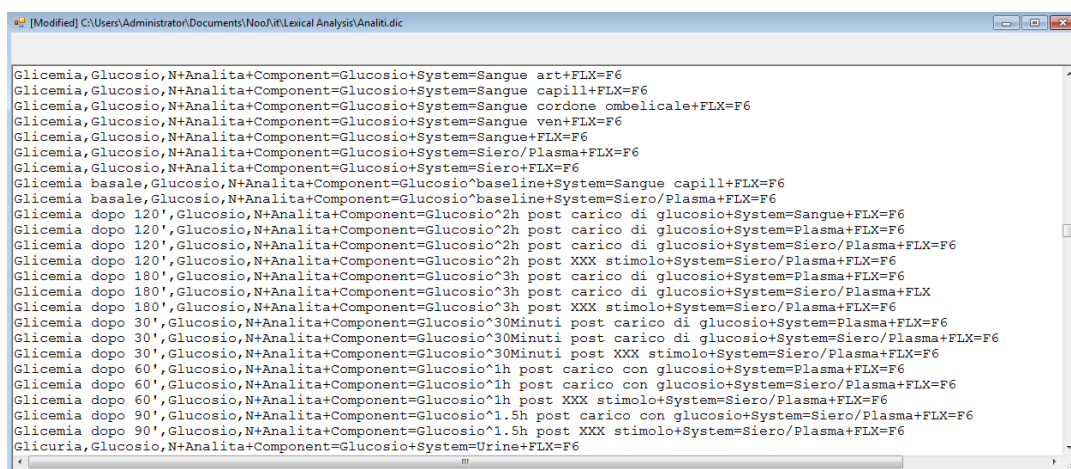


Figure 6 - Les morphèmes : suffixe "emia"

- 1) Glicemia,Glucosio,N+Analita+Component=Glucosio+System=Sangue art
 Glicemia,Glucosio,N+Analita+Component=Glucosio+System=Sangue capill
 Glicemia,Glucosio,N+Analita+Component=Glucosio+System=Sangue cordone ombelicale
 Glicemia,Glucosio,N+Analita+Component=Glucosio+System=Sangue ven
 Glicemia,Glucosio,N+Analita+Component=Glucosio+System=Sangue
 Glicemia,Glucosio,N+Analita+Component=Glucosio+System=Siero/Plasma
 Glicemia,Glucosio,N+Analita+Component=Glucosio+System=Siero

Le premier cas, montré dans 1) est un exemple de morphème en « emia » avec l'ambiguïté créée dans le contexte spécifique considéré.

La définition « Glicemia » (en français « Glycémie ») représente le concept « taux de glucose dans le sang⁶⁸ ». Sa composition provient du grec en particulier par –glico (« douce ») et –emia (« Sangue »)⁶⁹ et permet de garder une sémantique déterminée dans un domaine spécifique comme la biologie, la médecine etc.

Le système des noms et des codes LOINC, très spécifique et précis dans l'identification de la combinaison des attributs, considère différentes valeurs pour le milieu biologique, qui peuvent être déduit par le suffixe « emia ».

⁶⁸ <<http://www.larousse.fr/dictionnaires/francais/glyc%C3%A9mie/37279>> (data dell'ultima consultazione).

⁶⁹ Dans la version française « Glycémie » est composée par -glyco (« sucré ») et par -emie (« sang »).

Dans le travail présenté nous avons considéré les valeurs les plus communes qui peuvent être associées à une définition exprimée par le morphème « emia » ; en particulier : « Sangue », « Sangue art », « Sangue ven », « Sangue capill », « Sangue cordone ombelicale », « Siero/Plasma », « Siero ».

Les valeurs considérés sont le résultat d'une analyse des formes linguistiques les plus utilisées par les laboratoires pour les associations qu'ils effectuent le plus souvent, et d'une interaction directe avec les experts du système LOINC qui s'occupent de valider les codes identifiés.

La formalisation dans le dictionnaire NooJ, permet d'associer les morphèmes « emia » à l'élément testé (en utilisant la virgule « Glicemia, Glucosio ») et au correspondent paramètre « Component » du système LOINC par la propriété « +Component » qui, dans ce cas, prendra la valeur « Glucosio ».

Pour l'attribut LOINC « System », différentes formalisations ont été prévues pour différencier la sémantique des entrées ; en correspondance d'une valeur différente pour la propriété « +System » l'entrée lexicale assume une sémantique différentes. Cette différenciation n'est pas liée aux propriétés grammaticales et morphologiques mais uniquement au domaine spécifique considéré.

Il est très important, donc, de comprendre le lien entre une caractéristique morphologique de la définition locale et l'identification de la correspondante valeur LOINC à associer. L'ambiguïté considérée pour ce genre d'ALU n'est pas de nature linguistique mais elle dépend du domaine spécifique considéré.

Le travail présenté dans cette thèse répond à un besoin déterminé qui justifie la prise en compte des aspects strictement liés au domaine spécifique considéré et à la langue utilisée pour décrire les examens.

- 2) Albuminuria,Albumina,N+Analita+Component=Albumina+System=Urine
- Azoturia,Azoto,N+Analita+Component=Azoto+System=Urine
- Calciuria,Calcio,N+Analita+Component=Calcio+System=Urine
- Creatinuria,Creatinina,N+Analita+Component=Creatinina+System=Urine

Le cas 2) montre les formalisations pour les définitions composées par le suffixe « uria ». Dans ce cas, le morphème correspondant au suffixe « *uria* » ne crée pas d'ambiguïtés en référence au domaine puisque on considère qu'une seule valeur possible pour le milieu biologique. En vérifiant les associations « forme locale – code LOINC », nous avons remarqué que l'usage des formes avec les suffixes « uria » détermine, dans la majorité des cas, la valeur « urine » pour le milieu biologique.

La composition de la définition « Azoturia » (en français « Azoturie ») est composée par le mot « azoto » (en français « azote ») et le suffixe « uria » (en français « urie ») ; elle représente le concept de « quantité d'azote présente dans les urines ». Le suffixe « uria » identifie, donc, la quantité d'un élément (déterminé dans la racine du mot) présente dans les urines.

Comme pour les morphèmes ayant « emia », les définitions ayant le suffixe « uria », sont formalisées par les propriétés sémantiques « +Component » et « +System » qui prennent respectivement les valeurs « Azoto » et « urine », si on considère la forme « Azoturia ».

Dans l'exemple 2) il est possible de vérifier la formalisation aussi pour les définitons « Calciuria » (en français « Calciurie »), « Albuminuria » (en français « Albuminurie ») et « Creatinuria » (en français « Créatinurie »).

Les définitions montrée dans 1) et dans 2) sont qu'un sous ensemble des morphèmes décrites dans le dictionnaire des analytes.

Les entrées définies dans la catégorie « Acidi » (« Acides ») par la propriété sémantique « +AC » sont montrées dans l'image ci-dessus (Figure 7) :

```

#####Acidi
Citrato,N+Analita+AC+Component=Citrato+FLX=F78
Acido citrico,Citrato,N+Analita+AC+Component=Citrato+FLX=F11
Folato,N+Analita+AC+Component=Folato+FLX=F78
Acido folico,Folato,N+Analita+AC+Component=Folato+FLX=F11
Fosfato,N+Analita+AC+Component=Fosfato+FLX=F78
Acido fosforico,Fosfato,N+Analita+AC+Component=Fosfato+FLX=F11
Ippurato,N+Analita+AC+Component=Ippurato+FLX=F78
Acido ippurico,Ippurato,N+Analita+AC+Component=Ippurato+FLX=F11
Lattato,N+Analita+AC+Component=Lattato+FLX=F78
Acido lattico,Lattato,N+Analita+AC+Component=Lattato+FLX=F11
Metilippurato,N+Analita+AC+Component=Metilippurato+FLX=F78
Metilippurato,N+Analita+AC+Component=Metilippurato/Creatinina+FLX=F78
Acido metilippurico,Metilippurato,N+Analita+AC+Component=Metilippurato+FLX=F11
Acido metilippurico,Metilippurato,N+Analita+AC+Component=Metilippurato/Creatinina+FLX=F11
Ac metilippurico,Metilippurato,N+Analita+AC+Component=Metilippurato+FLX=F11
Ac metilippurico,Metilippurato,N+Analita+AC+Component=Metilippurato/Creatinina+FLX=F11
Omovanillato,N+Analita+AC+Component=Omovanillato+FLX=F78
Acido omovanillico,Omovanillato,N+Analita+AC+Component=Omovanillato+FLX=F11
Acido para aminoippurico,Para aminoippurato,N+Analita+AC+Component=Para aminoippurato+FLX=F11bis
Para aminoippurato,N+Analita+AC+Component=Para aminoippurato+FLX=F78
Acido solforico,Solfato,N+Analita+AC+Component=Solfato+FLX=F11
Solfato,N+Analita+AC+Component=Solfato+FLX=F78
Urato,N+Analita+AC+Component=Urato+FLX=F78
Acido urico,Urato,N+Analita+AC+Component=Urato+FLX=F11
Ac. urico,Urato,N+Analita+AC+Component=Urato+FLX=F7
Acido vanilmandelico,Vanillilmandelato,N+Analita+AC+Component=Vanillilmandelato+FLX=F11
Vanillilmandelato,N+Analita+AC+Component=Vanillilmandelato+FLX=F78
    
```

Figure 7 - Les "Acides"

Le système des noms LOINC, exige que la définition des acides soit dans leur forme avec le suffixe « ato ». Par exemple « Acido urico » (en français « Acide urique ») est une forme qui n'est pas reconnue dans le système LOINC qui accepte, au contraire, seulement la forme en -ato « Urato » (en français « Urate »). Aussi pour ces catégories, il est possible de différencier les cas traités.

- 1) *Metilippurato,N+Analita+AC+Component=Metilippurato*
- Metilippurato,N+Analita+AC+Component=Metilippurato/Creatinina*
- Acido metilippurico,Metilippurato,N+Analita+AC+Component=Metilippurato*
- Acido metilippurico,Metilippurato,N+Analita+AC+Component=Metilippurato/Creatinina*
- Ac metilippurico,Metilippurato,N+Analita+AC+Component=Metilippurato*
- Ac metilippurico,Metilippurato,N+Analita+AC+Component=Metilippurato/Creatinina*

L'exemple montré en 1), représente un cas d'ambiguïté. L'entrée « Acido metilpurrico » (en français « Acide méthylène-hippurique ») en particulier dans la forme en -ato « Metilippurato » (en français « Méthylhippurate »), peut être associé aux codes LOINC avec les valeurs du composant (propriété « +Component ») différentes : « Metilippurato » ou « Metilippurato/Creatinina ». Nous avons vérifié avec l'analyse de corpus que les laboratoires n'utilisent aucun symbole pour différencier les deux valeurs pour le paramètre « Component ». Il est seulement signalé parfois dans l'unité de mesure avec le sigle « creat ».

- 2) *Uricemia,Urato,N+Analita+AC+Component=Urato+System=Sangue art*
Uricemia,Urato,N+Analita+AC+Component=Urato+System=Sangue capill
Uricemia,Urato,N+Analita+AC+Component=Urato+System=Sangue cordone ombelicale
Uricemia,Urato,N+Analita+AC+Component=Urato+System=Sangue ven
Uricemia,Urato,N+Analita+AC+Component=Urato+System=Siero/Plasma
Uricemia,Urato,N+Analita+AC+Component=Urato+System=Siero
Uricuria,Urato,N+Analita+AC+Component=Urato+System=Urine

Dans l'exemple montré en 2), il est possible d'analyser le seul cas de morphèmes appariant pour la catégorie « Acidi » (« Acides »). Les mêmes conditions d'ambiguïtés, pour le suffixe « emia » et pour le suffixe « uria », ont été prises en considération pour formaliser sémantiquement les entrées. On retrouve, en effet, les différentes valeurs pour l'attribut « System » déterminées dans le suffixe « emia », et la seule valeur « urine » pour l'attribut « System » en correspondance des entrées avec le morphème « uria ».

Ces différentes valeurs sont formalisées par la propriété sémantique « +System ». Les entrées lexicales « Uricemia » et « Uricuria » se réfèrent au composant « Urato » qui est décrit par la propriété sémantique « +Component ».

La catégorie des acides et le lien entre les formes en « ico » et en « ato », sont décrits par la grammaire flexionnelle correspondante, qui doit tenir compte aussi des formes plus complexes comme l'exemple des mots composés montrés en 3) ci-dessous.

- 3) *Acido para aminoippurico,Para aminoippurato,N+Analita+AC+Component=Para aminoippurato*
Para aminoippurato,N+Analita+AC+Component=Para aminoippurato

Les entrées dans la catégorie « Acronymes » (Figure 8) sont identifiées par la propriété « +Acronimo ».

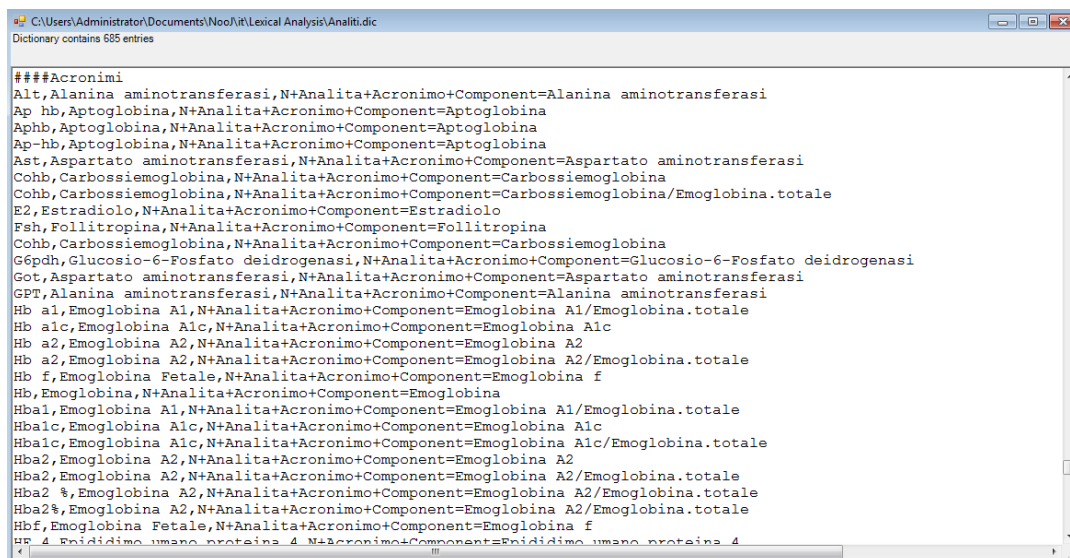


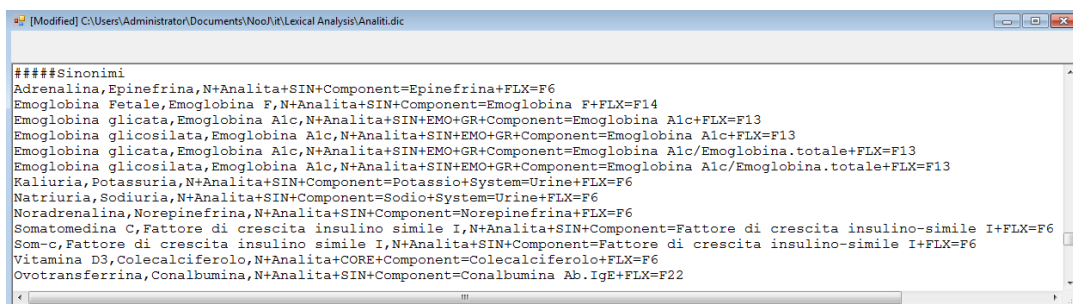
Figure 8 - Les acronymes

Dans le dictionnaire des analytes nous avons 100 acronymes décrits. Ils sont formalisés de la même façon que les autres entrées, et en maintenant aussi, les mêmes caractéristiques et ambiguïtés. Chaque acronyme est formalisé et associé à sa définition. Dans l'exemple ci-dessous la définition entière des acronymes est disposée après la virgule.

- *Alt,Alanina aminotransferasi,N+Analita+Acronimo+Component=Alanina aminotransferasi*
- Ast,Aspartato aminotransferasi,N+Analita+Acronimo+Component=Aspartato aminotransferasi*
- Fsh,Follitropina,N+Analita+Acronimo+Component=Follitropina*
- Got,Aspartato aminotransferasi,N+Analita+Acronimo+Component=Aspartato aminotransferasi*
- Hb a1,Emoglobina A1,N+Analita+Acronimo+Component=Emoglobina A1/Emoglobina.totale*
- Hba1c,Emoglobina A1c,N+Analita+Acronimo+Component=Emoglobina A1c*
- Hba1c,Emoglobina A1c,N+Analita+Acronimo+Component=Emoglobina A1c/Emoglobina.totale*

Les ambiguïtés qu'on a vu dans la définition « Emoglobina A1c », on peut le retrouver aussi pour son acronyme « HbA1c », décrites par la propriété sémantique « +Component ». La forme « N+Analiti+Acronimo » définit les analytes exprimés par des acronymes en les identifiant comme sous catégories sémantiques de la catégorie plus générique des noms (« N »).

Dans le dictionnaire « Analiti.dic », on identifie aussi la catégorie des « synonymes » (Figure 9). Les synonymes identifiés ne sont pas beaucoup, mais ils représentent un ensemble exhaustif pour l'analyse conduite. Il est nécessaire de rappeler que la formalisation effectuée est faite à partir des tests contenus dans le corpus analysé et en relation avec les catégories de tests choisies initialement. Partant les catégories comme cela, peuvent s'enrichir en effectuant des analyses sur corpus différents.



```
#####Sinonimi
Adrenalina,Epinefrina,N+Analita+SIN+Component=Epinefrina+FLX=F6
Emoglobina Fetale,Emoglobina F,N+Analita+SIN+Component=Emoglobina F+FLX=F14
Emoglobina glicata,Emoglobina A1c,N+Analita+SIN+EMO+GR+Component=Emoglobina A1c+FLX=F13
Emoglobina glicosilata,Emoglobina A1c,N+Analita+SIN+EMO+GR+Component=Emoglobina A1c+FLX=F13
Emoglobina glicata,Emoglobina A1c,N+Analita+SIN+EMO+GR+Component=Emoglobina A1c/Emoglobina.totale+FLX=F13
Emoglobina glicosilata,Emoglobina A1c,N+Analita+SIN+EMO+GR+Component=Emoglobina A1c/Emoglobina.totale+FLX=F13
Kaliuria,Potasuria,N+Analita+SIN+Component=Potassio+System=Urine+FLX=F6
Natriuria,Sodiuria,N+Analita+SIN+Component=Sodio+System=Urine+FLX=F6
Noradrenalina,Norepinefrina,N+Analita+SIN+Component=Norepinefrina+FLX=F6
Somatomedina C,Fattore di crescita insulino simile I,N+Analita+SIN+Component=Fattore di crescita insulino-simile I+FLX=F6
Som-c,Fattore di crescita insulino simile I,N+Analita+SIN+Component=Fattore di crescita insulino-simile I+FLX=F6
Vitamina D3,Colecalciferolo,N+Analita+CORE+Component=Colecalciferolo+FLX=F6
Ovotransferrina,Conalbumina,N+Analita+SIN+Component=Conalbumina Ab.IgE+FLX=F22
```

Figure 9 - Les synonymes

Comme pour les acronymes, aussi les synonymes sont formalisés en association avec leurs formes dérivées et ils sont identifiés par la propriété « +SIN ». Aussi pour cette catégorie sémantique, on peut avoir : les formalisations standards 1), les ambiguïtés de la définition principale reportées aussi pour son synonyme 2), les formes linguistiques locales pour exprimer le même synonyme 3).

1) *Adrenalina, Epinefrina, N+Analita+SIN+Component=Epinefrina*

La définition « Adrenalina » (en français « Adrénaline ») est le synonyme de « Epinefrina » (en français « Épinéphrine ») définition officielle utilisée en LOINC. L'entrée « Epinefrina » est formalisée par la propriété « Analita », son synonyme correspondant est « Analiti+SIN », et les deux sont associées à la valeur du composant « Epinefrina » par la propriété « +Component ».

2) *Emoglobina glicata, Emoglobina A1c, N+Analita+SIN+Component=Emoglobina A1c*
Emoglobina glicata, Emoglobina A1c, N+Analita+SIN+Component=Emoglobina
A1c/Emoglobina.totale

De la même façon que pour les acronymes, les ambiguïtés liées à la définition principale sont reportées aussi dans la formalisation des synonymes. La forme « Emoglobina glicata » est synonyme de « Emoglobina A1c » et maintient les mêmes ambiguïtés en référence au domaine spécifique et par rapport à la valeur de l'attribut composant décrit dans la propriété « +Component ».

3) *Somatomedina C, Fattore di crescita insulino simile I, N+Analita+SIN+Component=Fattore di*
crescita insulino-simile I
Som-c, Fattore di crescita insulino simile I, N+Analita+SIN+Component=Fattore di crescita
insulino-simile I

Les entrées montrées dans 3), sont un clair exemple de formes linguistiques locales associées à la même valeur pour l'attribut « Component ».

Le domaine de la médecine, comme celui de la biologie ou de la chimie, est plein d'acronymes, de morphèmes et de synonymes utilisés dans le langage commun (et pour certaines expressions aussi dans le langage plus technique). Le domaine LOINC, comme on a pu vérifier, est particulier en ce sens, puisqu'il refuse l'usage de certaines formes linguistiques étant le résultat d'un travail de traduction à partir de sa forme originale en anglais. LOINC n'accepte pas les formes

linguistiques propres des langues latines mais permet aux correspondants groupes de recherche, d'améliorer chaque année la version dans sa propre langue en ajoutant synonymes acronymes et formes utiles. Pour cette raison l'identification d'un ensemble consistant de catégories linguistiques utilisées au niveau locale a été un travail intéressant aussi pour le développement de la version italienne de LOINC.

La dernière catégorie sémantique considérée se réfère aux tests faisant partie de l'examen « *Emocromo* » (en français « Hémogramme »). Elle est représentée par la propriété « +EMO » et par les propriétés complémentaires « +GB » (pour identifier les tests sur les globules blancs), « +GR » (pour identifier les tests effectués sur les globules rouges), « +PIA » (pour identifier les tests effectués sur les plaquettes). Comme pour les autres catégories sémantiques, les entrées décrivant les tests contenus dans l'hémogramme ont été décrits suivant les mêmes règles et en divisant les entrées en : analytes simples, acronymes, synonymes. Dans les images suivantes (Figure 10, Figure 11, Figure 12, Figure 13, Figure 14, Figure 15), il est possible de voir comment les entrées lexicales dans la catégorie « *Emocromo* » ont été décrites dans le dictionnaire.

```

#####EMOCROMO EMO
#####Globuli bianchi GB
Basofili %,Basofili,N+Analita+EMO+GB+Component=Basofili/100 leucociti+FLX=F788
Basofili%,Basofili,N+Analita+EMO+GB+Component=Basofili/100 leucociti+FLX=F788
Basofilo,N+Analita+EMO+GB+Component=Basofili+FLX=F78
Baso,Basofili,N+Analita+EMO+GB+Component=Basofili+FLX=F78
Emocromo,N+Analita+EMO+Component=Emocromo+FLX=F7
Eosinofilo,N+Analita+EMO+GB+Component=Eosinofili+FLX=F78
% Eosinofili,Eosinofili,N+Analita+EMO+GB+Component=Eosinofili/100 leucociti+FLX=F78
%Eosinofili,Eosinofili,N+Analita+EMO+GB+Component=Eosinofili/100 leucociti+FLX=F78
Eosinofili%,Eosinofili,N+Analita+EMO+GB+Component=Eosinofili/100 leucociti+FLX=F78
Eosino%,Eosinofili,N+Analita+EMO+GB+Component=Eosinofili/100 leucociti+FLX=F78
Eosi%,Eosinofili,N+Analita+EMO+GB+Component=Eosinofili+FLX=F7
Eosi%,Eosinofili,N+Analita+EMO+GB+Component=Eosinofili/100 leucociti+FLX=F7
Eosi %,Eosinofili,N+Analita+EMO+GB+Component=Eosinofili/100 leucociti+FLX=F7
Globulo bianco,Leucocita,N+Analita+SIN+EMO+GB+Component=Leucociti+FLX=F15
Grandi cellule non colorate,N+Analita+EMO+GB+Component=Grandi cellule non colorate+FLX=F12
Granulocita basofilo,Basofilo,N+Analita+SIN+EMO+GB+Component=Basofili+FLX=F11
Granulocita eosinofilo,Eosinofilo,N+Analita+SIN+EMO+GB+Component=Eosinofili+FLX=F11
Granulocita neutrofilo,Neutrofilo,N+Analita+SIN+EMO+GB+Component=Neutrofilo+FLX=F11
Granulocita,N+Analita+EMO+GB+Component=Granulociti+FLX=F10
Leucocita,N+Analita+EMO+GB+Component=Leucociti+FLX=F78
Linfocita,N+Analita+EMO+GB+Component=Linfociti+FLX=F78
Linfociti%,Linfociti,N+Analita+EMO+GB+Component=Linfociti/100 leucociti+FLX=F78
Linfociti %,Linfociti,N+Analita+EMO+GB+Component=Linfociti/100 leucociti+FLX=F78
Linf%,Linfociti,N+Analita+EMO+GB+Component=Linfociti/100 leucociti+FLX=F78
    
```

Figure 10 - La catégorie des globules blancs

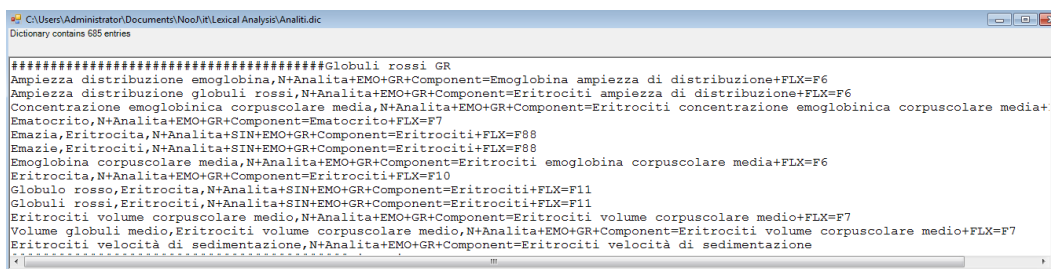


Figure 11 - La catégorie des globules rouges

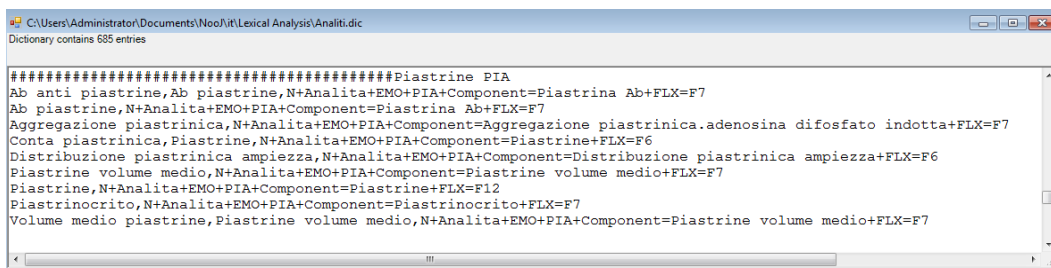


Figure 12 - La categorie des plaquettes

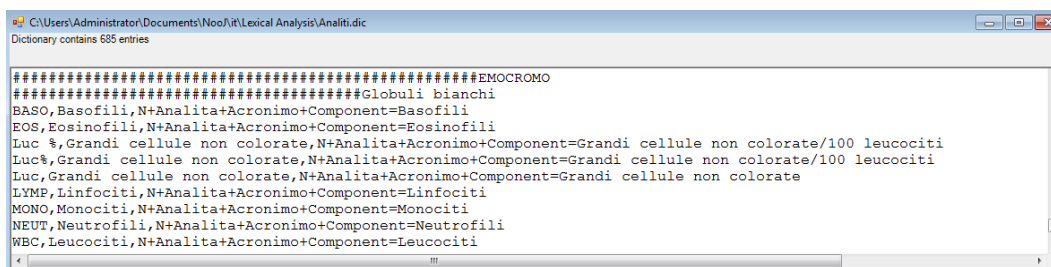


Figure 13 - Les acronymes pour la catégorie des globules blancs

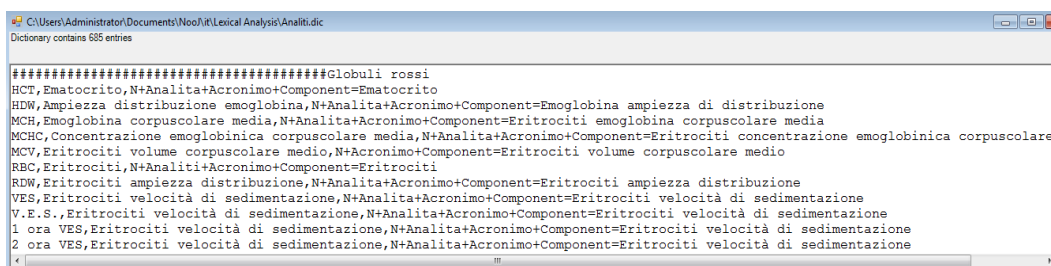


Figure 14 - Les acronymes pour la categorie des globules rouges

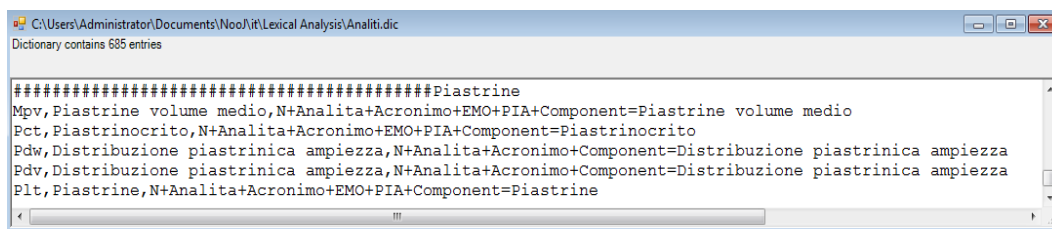


Figure 15 - Les acronymes pour la categorie des plaquettes

Pour cette typologie d'entrée la présence de symboles est très fréquente. En particulier, le symbole « % » identifie une valeur du composant exprimée en pourcentage, alors que le symbole « # » une valeur du composant égale à la quantité exprimée par un nombre.

Par ailleurs, il y a certaines formes linguistiques sous formes d'expressions composées par acronymes, symboles et surtout expressions temporelles, qu'en LOINC ont une sémantique particulière (Figure 16).

Quelques exemples de formes locales, pour faire mieux comprendre le travail de collection si minutieux, sont montrés dans l'image suivante :

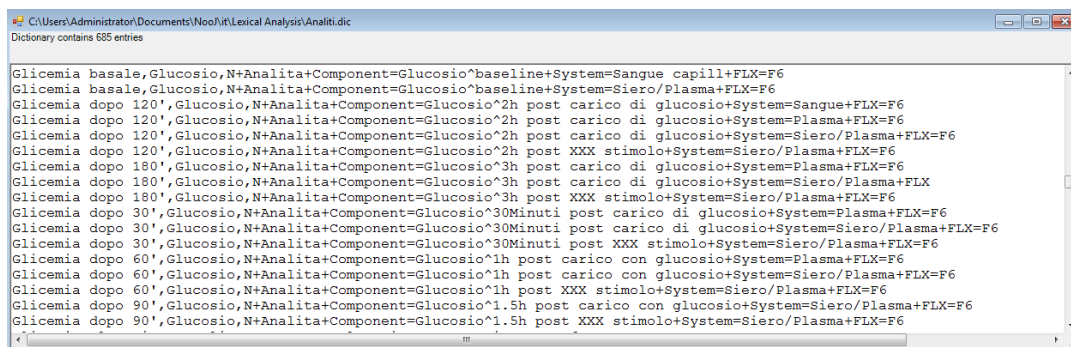


Figure 16 - Les expressions temporelles

Le dictionnaire « Analiti.dic » est montré dans sa version complète en Annexe 1 – Les dictionnaire « Analiti.dic ».

4.2.2 Le dictionnaire électronique : « Les attributs LOINC »

Le dictionnaire « Elements_LOINC » décrit les formes linguistiques locales utilisées en référence avec les autres paramètres LOINC. En particulier, les formes linguistiques locales utilisées pour exprimer les éléments concernant les milieux biologiques, les méthodes, les échelles et les unités de mesure.

Dans ce dictionnaire sont décrits aussi les adjectifs utilisés pour accompagner le « composant » et l'ensemble des codes LOINC retrouvés dans le corpus. Ce dictionnaire a permis donc la formalisation des liens existants entre les paramètres LOINC, pour déduire à partir d'une valeur déterminée et associée à un paramètre spécifique, les possibles valeurs des autres (par exemple à partir de l'unité de mesure on peut déduire la valeur de « l'échelle » ou du « Temps »).

Le paramètre « Grandeur » par exemple, est déduit à partir de l'unité de mesure ; c'est-à-dire que la valeur de la grandeur est formalisée comme valeur d'une propriété sémantique associée aux entrées lexicales décrivant les unités de mesure. Il contient 385 entrées lexicales et décrit les formes linguistiques associées aux attributs LOINC (milieu biologique, échelle, méthode, unité de mesure) et les autres éléments utiles pour le travail d'annotation et de vérification d'exactitude des données (les adjectifs associés aux définitions et les codes).

Comme on a vu pour le dictionnaire des analytes, NooJ donne la possibilité de formaliser les liens existants entre les différents attributs en associant par des propriétés sémantiques les valeurs des attributs LOINC correspondants dans leurs formes officielles. Les définitions des examens se réfèrent surtout à la valeur de l'attribut « Component » qui décrit le composant de l'examen ainsi que aux autres attributs (en particulier l'attribut « System ») quand la morphologie de mots crée d'ambiguïtés en suggérant les valeurs correspondants aux autres attributs.

Dans le dictionnaire « Elements_LOINC.dic », pour formaliser les formes linguistiques et les liens entre les différents attributs LOINC considérés, des propriétés spécifiques ainsi que de sous catégories de noms ont été prévues pour en décrire la variété.

Pour l'attribut concernant le milieu biologique (en anglais « System »), les formes linguistiques locales retrouvées dans le corpus ont été formalisées dans le dictionnaire (Figure 17), par la propriété « +Campione ».

```

#System
Sangue, N+Campione+System=Sangue+FLX=F7
Siero, N+Campione+System=Siero+FLX=F7
Urina, N+Campione+System=Urine+FLX=F88
Urina di fine turno, N+Campione+System=Urine+FLX=F882
Urina delle 24 ore, N+Campione+System=Urine+Timing=24h+FLX=F881
Urina delle 24, N+Campione+System=Urine+Timing=24h+FLX=F884
Urine 24 ore, N+Campione+System=Urine+Timing=24h+FLX=F884
Urina 24ore, N+Campione+System=Urine+Timing=24h+FLX=F884
Urina 24h, N+Campione+System=Urine+Timing=24h+FLX=F884
Urina fresca, N+Campione+System=Urine+FLX=F883
Urina fresc, N+Campione+System=Urine+FLX=F884
Na-k urinari, N+Campione+System=Urine+FLX=F7
Plasma, N+Campione+System=Plasma+FLX=F7
Feci, N+Campione+System=Feci+FX=F6
Sangue intero, N+Campione+System=Sangue+FLX=F7
Siero plasma, N+Campione+System=Siero/Plasma+FLX=F7
Siero o plasma, Siero plasma, N+Campione+System=Siero/Plasma+FLX=F7
Ematologia, N+Campione+System=Sangue+FLX=F6
Coagulazione, N+Campione+System=PPP+FLX=F6
    
```

Figure 17 - la categorie des milieux biologiques

Chaque forme linguistique locale retrouvée pour indiquer le milieu biologique est identifiée dans cette catégorie et associée à l'attribut LOINC « System » par la propriété sémantique « +System ». Aussi pour ces formes, il est possible d'identifier différents cas à traiter : les mots simples associés aux valeurs officielles correspondantes 1), les mots composés 2), les expressions locales qui suggèrent aussi la valeur pour les paramètres « temps » 3).

- 1) *Sangue, N+Campione+System=Sangue*
Coagulazione, N+Campione+System=PPP

Dans 1) est montré un exemple de forme linguistique simple (« Sangue » en français « Sang ») associée à la valeur du milieu biologique, par la propriété « +System » (« +System=Sangue »). Ce qui est intéressant de vérifier, est

l'association, à partir d'un mot simple, à un sigle spécifique ; par exemple l'entrée lexicale qui représente le milieu biologique « Coagulazione », (en français « Coagulation ») est associé à la valeur « PPP » pour la propriété « +System ». Cela représente un petit exemple de la complexité du système de noms et de codes LOINC et de comment l'association aux justes codes peut paraître très difficile pour les débutants.

2) *Siero plasma,N+Campione+System=Siero/Plasma*

Dans 2) est décrite la formalisation d'un mot composé (« Siero Plasma » en français « Sierum Plasma ») avec sa correspondante valeur dans la nomenclature LOINC par l'association « +System=Siero/Plasma ». Les mots composés décrits dans ce dictionnaire sont plutôt des ensembles des valeurs ou des concepts qui sont utilisés conjointement pour se référer à une valeur officielle spécifique. Par exemple « Siero Plasma » ou la forme « Siero e Plasma » (« Sierum et Plasma ») ou encore « Siero o Plasma » (« Sierum ou Plasma »), représentent l'ensemble des formes qui se réfèrent à la même valeur LOINC « Siero/Plasma ».

L'exemple montré dans 3) décrit les formes linguistiques qui peuvent suggérer aussi une valeur pour l'attribut « Temps ». Ce genre de formes est très important puisqu'il représente la possibilité d'identifier deux valeurs LOINC à partir d'une forme linguistique utilisée au niveau locale.

3) *Urina delle 24 ore,N+Campione+System=Urine+Timing=24h*
Urina delle 24,N+Campione+System=Urine+Timing=24h
Urine 24 ore,N+Campione+System=Urine+Timing=24h
Urine 24ore,N+Campione+System=Urine+Timing=24h
Urina 24h,N+Campione+System=Urine+Timing=24h

L'exemple ci-dessus, montre la possibilité de formaliser le lien entre l'attribut « System » (par la propriété « +System ») et l'attribut « Timing » par la propriété « +Timing ». La formalisation de dictionnaire NooJ permet de décrire le lien l'association entre une forme linguistique locales (par exemple « Urine delle 24

ore » en français « Urine des 24 heures ») et les deux valeurs correspondants aux paramètres LOINC, « System » et « Timing » exprimés dans leur forme officielle (« System=Urine+Timing=24h »). L'analyse de ces liens est très importante, puisque on peut vérifier que les formes linguistiques locales étant plus explicites, rendent la recherche du code candidat plus difficile si on ne connaît pas les règles de mappage et la structure du système LOINC.

Il est, donc, nécessaire de donner à ces formes locales une connaissance intrinsèque que puisse conduire les experts dans la recherche et dans la validation.

D'autre part, la possibilité de déduire avec une seule expression locale un ou plusieurs valeurs des attributs officiels, représente un des avantages les plus importants du travail présenté. L'importance de ce travail repose dans la formalisation de formes linguistiques locales en utilisant la connaissance des experts de laboratoire sur leurs tests effectués et celle des experts du système de noms LOINC impliqués dans les opérations de validation.

Les éléments qui expriment l'échelle de mesure des tests ont été décrits par la propriété « +Risultato », (Figure 18).

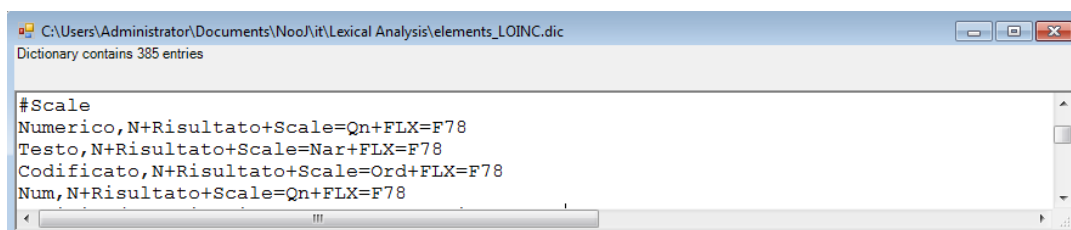


Figure 18 - La catégorie d'échelles de mesure

Ces formes linguistiques sont associées à l'attribut « Scale » (« échelle de mesure ») par la propriété sémantique « +Scale ». Dans le dictionnaire « Elements_LOINC.dic » les formes linguistiques les plus utilisées ont été prise en considération et formalisées dans la catégorie « N+Risultato ».

- *Numerico,N+Risultato+Scale=Qn*
Num,N+Risultato+Scale=Qn

Pour l'ensemble d'échelles de mesure, aussi si très minimal, on peut vérifier aussi la recueille automatique de formes locales à associer à la même valeur officiel : les formes « Numerico » et « Num » sont associées à la même valeur « Qn » de l'attribut « Scale (« +Scale=Qn »).

Les formes linguistiques qui expriment la méthode utilisée pour effectuer un test ont été formalisées par la propriété « +Metodo » et associées à la valeur de l'attribut « Method » du système LOINC par la propriété sémantique « +Method » (Figure 19).

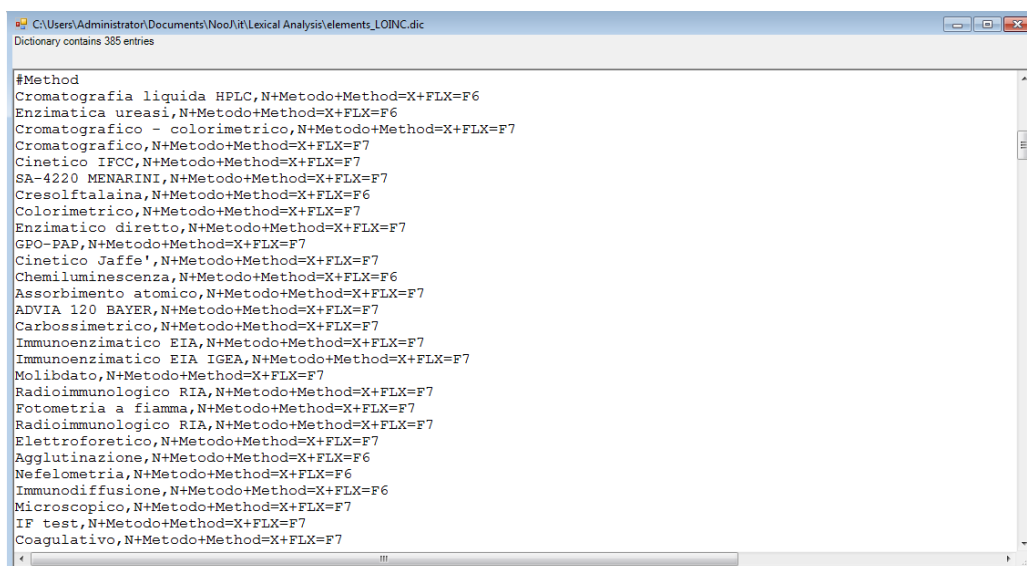


Figure 19 - La categorie des méthodes

La propriété « +Method » prévoit l'indication de présence («+Method= X ») ou l'absence («+Method=Absent»), sans l'association de la forme linguistique à la valeur correspondante en LOINC. L'attribut qu'indique la méthode utilisée pour obtenir les résultats de tests, détermine le choix du code surtout en évaluant l'information clinique associée à l'examen effectué avec une telle ou telle méthode. Les valeurs pour cet attribut ne sont pas exprimées avec leur forme précise, mais

par définition renferment un ensemble de définitions exprimant la même méthode d'analyse. La méthode est le seul attribut optionnel pour la description officiel des tests demandé en LOINC étant explicité seulement si l'adoption d'une certaine méthode change la nature clinique du test décrit. Pour ces raisons, nous avons choisi d'annoter pour cet attribut avec l'information sur sa présence ou son absence.

Les unités de mesure ont été formalisées (Figure 20) pour déduire les attributs « Property ».

Les formes linguistiques, sous formes de rapports ou formules qui représentent les unités de mesure, ont une structure particulière qui maintient beaucoup de connaissances en considérant leur nature et les liens avec les autres attributs.

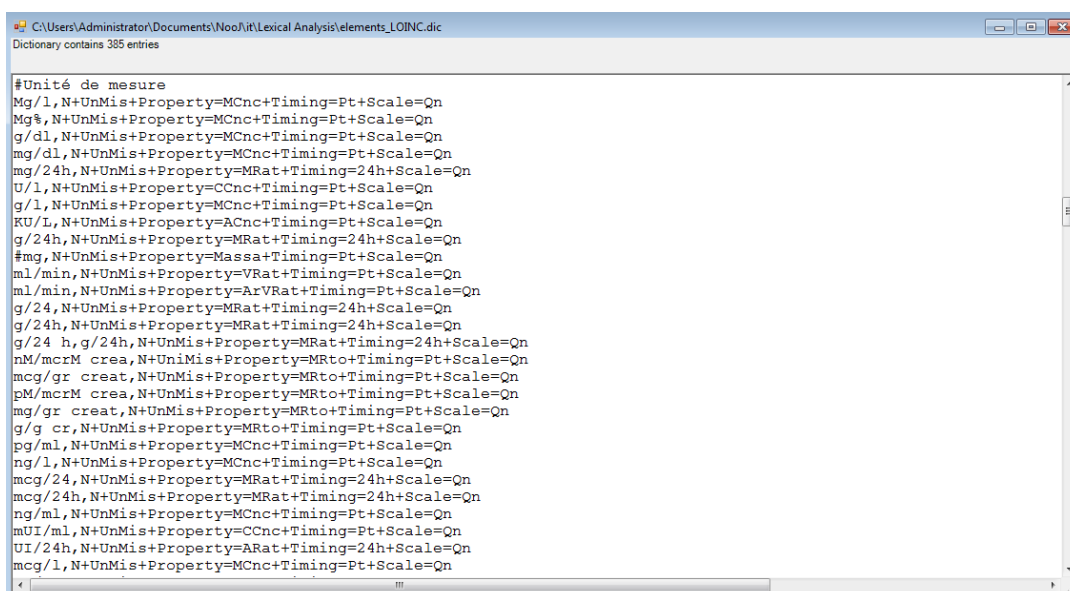


Figure 20 - La catégorie des unités de mesure

Les unités de mesure formalisées par la propriété « +UnMis » sont très importantes pour la déduction de la valeur de l'attribut « Property », et dans certains cas aussi des attributs « Scale » et « Timing ».

Par l'unité de mesure, on peut déduire d'autres paramètres et formaliser leurs valeurs par des propriétés spécifiques. En particulier, la valeur de l'attribut

« Grandeur » par la propriété « +Property », de l'attribut « temps » par la propriété « +Timing » et surtout de l'attribut « échelle » par la propriété « +Scale ».

L'unité de mesure détermine une valeur du paramètre « échelle » en représentant au même temps la nature de cette même mesure. Parfois le temps pris pour effectuer l'analyse est indiqué dans l'unité de mesure, surtout quand les examens doivent être effectués dans un laps de temps déterminé (24 heures, 12 heures etc.).

Si le temps n'est pas indiqué ça veut dire que l'examen a été effectué dans un temps ponctuel. La valeur de l'unité de mesure correspond à une valeur déterminée du paramètre « grandeur » dans le système LOINC.

- $Mg/l, N+UnMis+Property=MCnc+Timing=Pt+Scale=Qn$
 $g/24h, N+UnMis+Property=MRat+Timing=24h+Scale=Qn$

Ce lien entre les différents attributs est très important pour les analyses sémantiques et les contrôles sur l'exactitude des données conduites ensuite.

Dans le dictionnaire « Elements_LOINC.dic » d'autres entrées importantes ont été décrites ; on retrouve les éléments correspondants aux attributs LOINC et les autres entrées qui peuvent suggérer des valeurs. En particulier, les adjectifs utilisés pour expliciter des aspects déterminés dans la définition des examens (par exemple « Azoto urinario », en français « Azote urinaire ») en accompagnant l'élément principal, ont été identifiés par la propriété « +AgLoinc » (Figure 21).

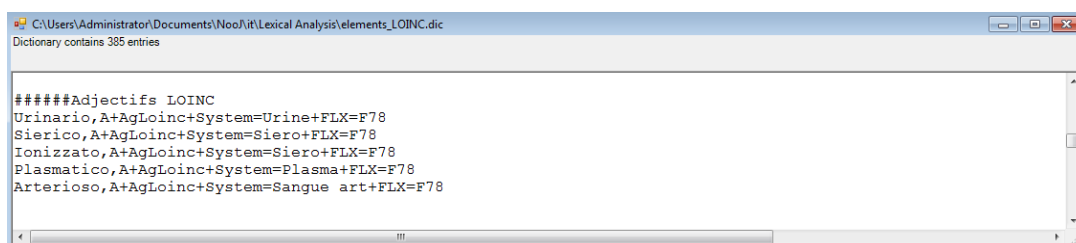


Figure 21 - Les adjectifs

Les adjectifs retrouvés dans le corpus, peuvent suggérer la valeur du paramètre « System » en indiquant le milieu biologique de l'examen ; l'adjectif « urinaire », par exemple, pour indiquer le milieu biologique « urine ».

La valeur ajoutée de ce dictionnaire est la formalisation de codes LOINC comme entrées lexicales (Figure 22).

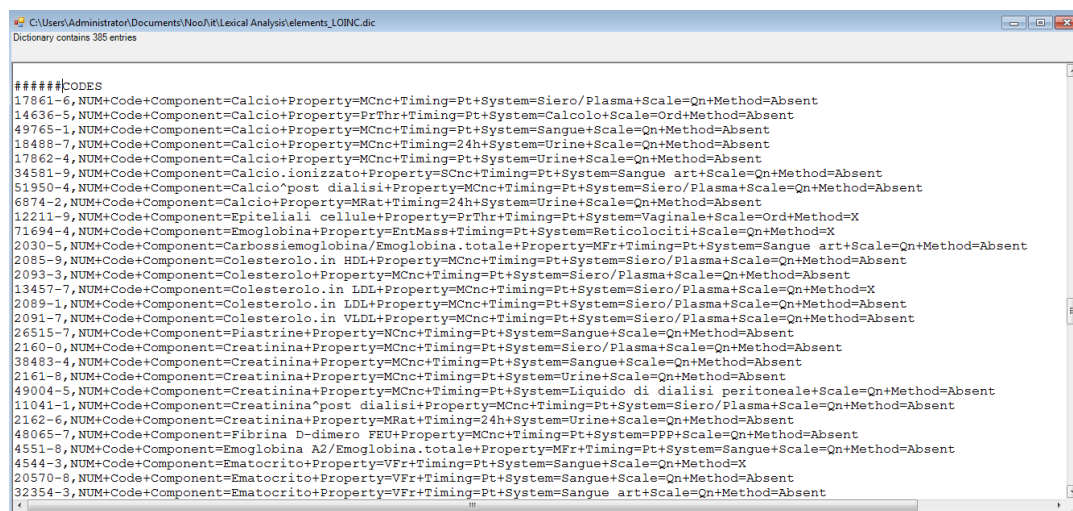


Figure 22 - Les codes

Chaque code LOINC a été formalisé par la propriété « +CODE » ajoutée comme sous-catégorie de nombres « NUM+CODE ». La particularité de ces entrées repose sur leur formalisation qui prévoit la présence de toutes les propriétés sémantiques qui gardent les valeurs des attributs LOINC. Les codes décrits dans le dictionnaire sont formalisés avec des propriétés sémantiques qui lui associent les valeurs correspondant exprimées dans leur forme officielles. Pour mieux comprendre, on peut regarder l'exemple de formalisation par rapport au code pour comme montré en Figure 23.

- 21305-8, NUM+Code
+Component=Glucosio+Property=MCnc+Timing=24h+System=Urine+Scale=Qn+Method=Absent

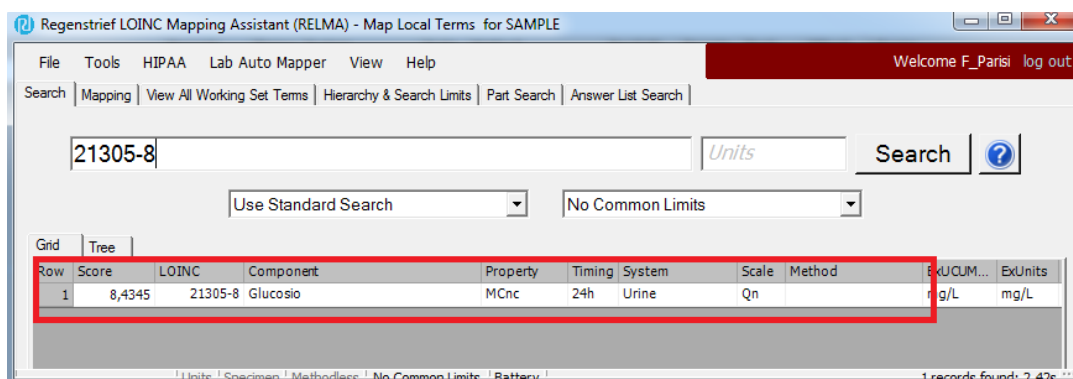


Figure 23 - Les axes LOINC décrits en RELMA

La formalisation des codes LOINC représente une étape fondamentale pour le travail présenté. Il permet la vérification directe et uniquement par des règles linguistiques, de l'association entre les formes linguistiques locales et les formes linguistiques officielles correspondant aux paramètres LOINC du code associé à chaque test. Par ailleurs, on peut dire que le code associé à chaque phrase (paragraphe 4.1) permet de déterminer le niveau d'exactitude de l'association effectuée et des résultats de l'annotation sémantique effectuée pour chaque élément et produites par le grammaire hors-contexte considérées. L'annotation sémantique et le contrôle d'exactitude sont renvoyés aux prochains paragraphes.

Le dictionnaire « Elements_LOINC.dic » est montré dans sa version complete en Annexe 2 – Dictionnaire « Elements_LOINC.dic ».

4.3 Les Grammaires flexionnelles

Les grammaires flexionnelles décrivent les différentes formes sous lesquelles une entrée du dictionnaire peut apparaître, comme les formes conjuguées d'un verbe, la mise au pluriel d'un nom ou d'un adjectif, etc.

Les grammaires flexionnelles dans NooJ, suivent une syntaxe particulière en utilisant des opérateurs morphologiques spécifiques [24]:

- pour effacer le caractère avant le curseur ;
- <C> pour changer la casse du caractère courant ;
- <D> pour dupliquer le caractère courant ;
- <L> pour déplacer le curseur vers la gauche ;
- <R> pour déplacer le curseur vers la droite ;
- <S> pour effacer le caractère après le curseur ;
- <N> pour déplacer le caractère au début du mot suivant ;
- <P> pour déplacer le curseur à la fin du mot précédent.

Le paradigme flexionnel « *COUSIN* », par exemple, est défini avec la forme « *COUSIN* = <E>/m+s | e/f+s | s/m+p | es/f+p; ». Ce paradigme décrit et reconnaît les formes suivantes du mot « *cousin* » :

- la forme du masculin singulier avec le suffixe <E> (vide), autrement dit, le mot est au masculin singulier si on ne lui rajoute rien ;
- la forme du féminin singulier si on ajoute le suffixe « *e* » ;
- la forme du pluriel masculin si on lui ajoute le suffixe « *s* » ;
- la forme du pluriel féminin si on lui rajoute le suffixe « *es* ».

Dans le travail suivant les grammaires flexionnelles construites se réfèrent à chaque forme linguistique considérée et ont été développées en considérant le domaine spécifique d'application. Cette considération est très importante puisque, malgré l'existence de ressources linguistique pour la langue italienne déjà disponible et exhaustive, l'application de TALN développée a été personnalisée

pour répondre à un besoin spécifique et construite ad-hoc sur les caractéristiques du domaine.

À cet égard, l'analyse des formes linguistiques qui représentent les morphèmes en « emia » et en « uria », est fondamentale à cause de leur connaissance intrinsèque très importante pour les analyses à effectuer. En considérant cet aspect, et pour permettre au logiciel de reconnaître les formes linguistiques ayant le suffixe « emia » et « uria », la propriété morphologique « SUFF » a été ajoutée (Figure 24) aux propriétés déjà disponibles pour la reconnaissance des formes en langue italienne, qui a permis la construction des grammaires flexionnelle pour les morphèmes considérés (Figure 25).

```

[Modified] C:\Users\Administrator\Documents\Noo\it\Lexical Analysis\_properties.def

# NooJ V1
# Dictionary properties' definition
#
# Language is: it
#
# Special Characters: '=' '+' '#' ' '
#
# List categories and properties associated with features
# Example: N_Number = m + f;
# Special KEYWORD: INFLECTION lists all inflectional features (used by variables $xF)
# Example: INFLECTION = m + f + Present + Futur;

INFLECTION = m + f + s + p + 1 + 2 + 3 + PR + IM + PA + F + S + IS + C + IMP + INF + L + G + PPR + PP + emia + uria ;

SUFF = emia + uria;

PRON_Caso = ACC + DAT + LOC;|

V_Synt = t + i;
V_Aux = a + e;

A_Genere = m + f;
A_Numero = s + p;
    
```

Figure 24 - La propriété morphologique "SUFF"

```

C:\Users\Administrator\Documents\Noo\it\Lexical Analysis\Analiti.nof

#Albubina, Emoglobina
F22=<E>/f+s|<B>e/f+p|<B>emia/emia|<B>uria/uria;

#Bilirubina, Creatinina
F2=<E>/f+s|<B>emia/emia|<B>uria/uria|<B3>uria/uria;

#Glucosio
F3=<E>/m+s|<B6>icemia/emia|<B6>icosuria/uria;

#Lipidi
F5=<E>/m+p|<B>emia/emia|<B3>emia/emia;
    
```

Figure 25 - Les grammaires flexionnelles pour les morphèmes "emia" et "uria"

Pour dérouler l'analyse linguistique présentée dans cette thèse, les grammaires flexionnelles « Analiti.nof » et « Elements_LOINC.nof » ont été construites pour décrire les formes linguistiques contenues dans les respectifs dictionnaires « Analiti.dic » et « Elements_LOINC.dic ». Pour chaque entrée à l'intérieur du dictionnaire a été formalisée sa flexion morphologique.

La flexion d'un mot est associée par la propriété « +FLX » qui prend la valeur correspondant au nom de la grammaire correspondant, par exemple « +FLX=F6 ». Il est possible pour chaque grammaire morphologique d'identifier différents catégories de flexions, en s'appuyant sur les entrées lexicales décrites dans le dictionnaire.

Un *sous ensemble* des flexions pour chaque catégorie est montrée dans les tableaux suivants.

- Flexion de formes « invariantes » ; ça veut dire les flexions morphologiques des mots qui ne changent pas le numéro, donc pour lesquelles n'existe que du pluriel ou du singulier. Un sous ensemble de flexions décrit dans les grammaires flexionnelles est montré en Tableau 5.

Flexions Invariantes
<E>/f+s
<E>/f+p
<E>/m+s
<E>/m+p

Tableau 5 - Flexions invariantes

- Flexions de mots composée par 2 termes ou plus ; les mots composés ont besoin, en général, d'une description particulière puisqu'ils sont différents les opérateurs morphologiques utilisée. Le Tableau 6 montre les flexions pour les mots composées par 2 termes et le Tableau 7 ceux composés par 3 termes ou plus.

Flexions noms composés (2 termes)
<E>/m+s i<P>i/m+p
<E>/m+s o/m+s i/m+p
<E>/m+s hi<P>i/m+p
<E>/f+s e<P>e/f+p
<E>/f+s e<P>i/f+p

Tableau 6 - Flexions des mots composés (2 termes)

Flexions noms composés (3 ou plus termes)
<E>/m+s i<P4>i/m+p
<E>/f+s <P3>e/f+p
<E>/f+s he<P>e/f+p
<E>/f+s <P>e/f+p

Tableau 7 - Flexions des mots composés (3 termes)

- Flexions génériques : associés aux mots qui prennent le masculin et le féminin (Tableau 8).

Flexions génériques
F78=<E>/m+s i/m+p
F88=<E>/f+s e/f+p

Tableau 8 - Flexions génériques

- Les flexions pour les suffixes « emia » et « uria » (Tableau 9) ; ces flexions peuvent utiliser la flexion en « emia » et en « uria » grâce à la propriété « SUFF » ajoutée dans les propriétés morphologiques. La définition de cette propriété a permis, comme on verra ensuite, la reconnaissance des morphèmes (« N+Analiti+emia » ou « N+Analiti+uria »), l'analyser de leurs concordances et la possibilité de leurs réutilisations pour autres

objectif (par exemple dans les grammaires hors-contexte développées pour l'annotation sémantique et le contrôle d'exactitude).

Flexions pour les suffixes EMIA - URIA
<E>/m+s emia/emia uria/uria
<E>/f+s emia/emia uria/uria <B3>uria/uria
<E>/m+s <B6>icemia/emia <B6>icosuria/uria
<E>/m+p emia/emia <B3>emia/emia

Tableau 9 - Flexions morphèmes en "emia" et "uria"

4.4 Grammaires hors contexte : « LOINC.nog »

Les grammaires hors-contexte (ou génératives de type 2) de NooJ sont des expressions rationnelles dans lesquelles peuvent figurer des références à d'autres expressions rationnelles, en permettant d'utiliser une grammaire plus puissante par rapport aux grammaires rationnelles (de type 3) ; c'est-à-dire, les grammaires hors contexte peuvent décrire des langages que les grammaires rationnelles ne peuvent pas décrire.

La grammaire hors-contexte LOINC a été réalisée pour effectuer le travail d'annotation sémantique du texte considéré. La grammaire « LOINC.nog » (Figure 26, Figure 27) est composée par la grammaire hiérarchique « LOINC_ELEMENTS » dont les entrées ayant catégories « NUM+Code » sont contenues dans la variable « C » avec un ensemble de contraintes contextuelles, qu'on va approfondir dans le prochain paragraphe.

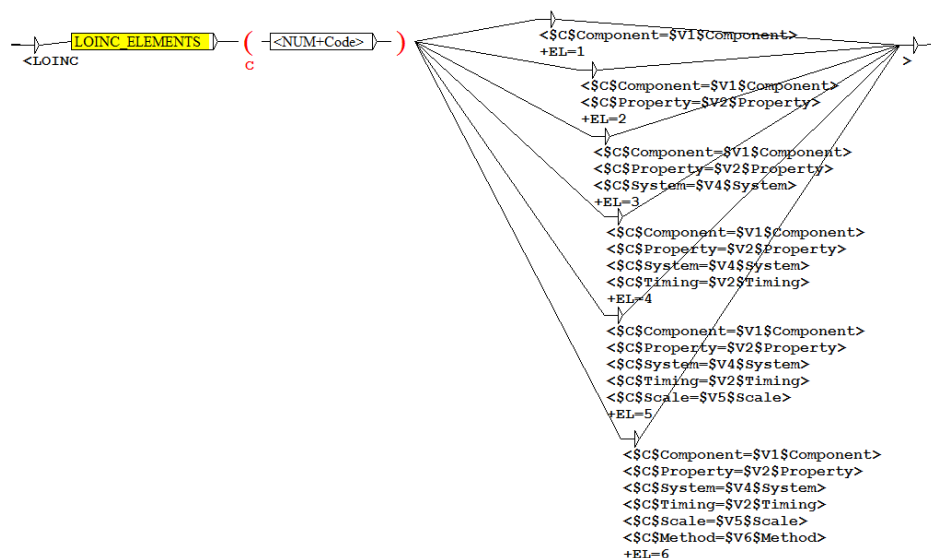


Figure 26 - La grammaire "LOINC.nog"

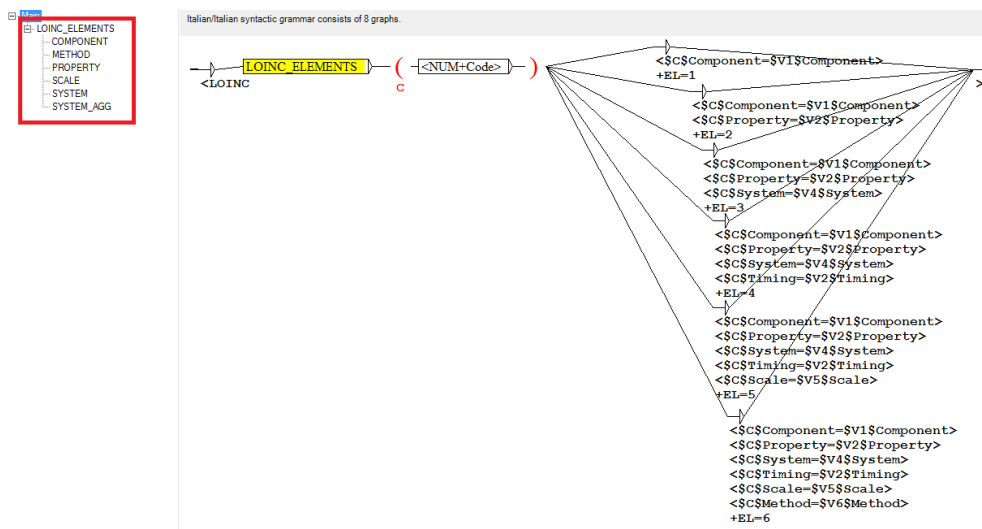


Figure 27 - Structure de la grammaire "LOINC.nog"

La grammaire « LOINC ELEMENTS » est composée de six sous-grammaires organisées en section différentes. Cette grammaire, reconnaît les éléments les plus importants pour déterminer la chaîne des mots résultant par l'annotation sémantique qu'on veut effectuer. La structure des phrases contenues dans notre texte

(paragraphe 4.1) est connue à priori et reste la même pour tout le texte. Chaque phrase du corpus considéré correspond à la description locale de test et des éléments qui le décrivent. La grammaire construite agisse, donc, dans un ordre précis formalisé dans le graphe régulier montré en Figure 28.

LOINC_ELEMENTS

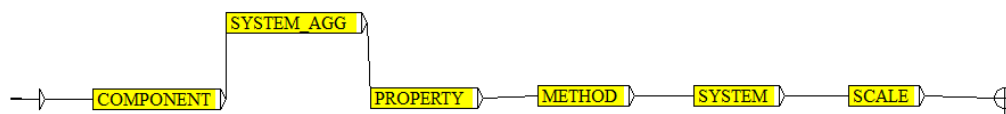


Figure 28 - La grammaire "LOINC_ELEMENTS"

Les premiers éléments reconnus par la sous-grammaire « COMPONENT » (Figure 29), sont les entrées correspondants à la catégorie « N+Analita ». La valeur de l'élément reconnu est contenu dans la variable « V1 », et il est annoté avec la valeur associée à la propriété « +Component ». La catégorie « N+Analita » représente les composants, en prenant en considération toute leurs formes et flexions possibles, des analyses décrits dans les phrases et formalisés dans le dictionnaire « Analiti.dic ». Les entrées décrites ont comme propriété sémantique « +Component » qui associe la valeur officielle du paramètre « Component » en LOINC à la forme locale.

Par la formalisation « +COMPONENT=\$V1\$Component », on annote la forme locale retrouvée sous la forme « N+Analiti » contenue dans la variable « V1 », avec la valeur de la propriété « +Component ».

Par exemple, si on considère la forme « Albuminemia » suivante :

Albuminemia,Albumina,N+Analita+Component=Albumina+System=Sangue art

la grammaire hors-contexte « LOINC » va créer l'annotation dans la forme « LOINC+COMPONENT=Albumina ».

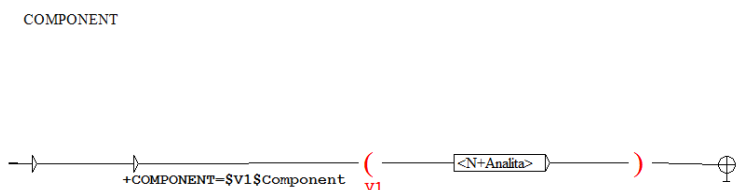


Figure 29 - La grammaire "COMPONENT"

Après avoir annoté l'analyse de l'analyse considérée, la grammaire vérifie la présence des entrées qui peuvent accompagner la définition en suggèrent d'autres valeurs d'attributs de LOINC (Figure 30). En particulier, il vérifie la présence de mots avec la catégorie « A+AgLOINC », une sous-catégorie des adjectifs considérés comme importants pour cette analyse spécifique. Pour cette catégorie d'entrées n'est prévue aucune annotation, étant que les possibles valeurs suggérées sont considérées de « support » et d'aide à déterminer l'exactitude des données par les contraintes qui considèrent les valeurs déduite dans une deuxième phase.

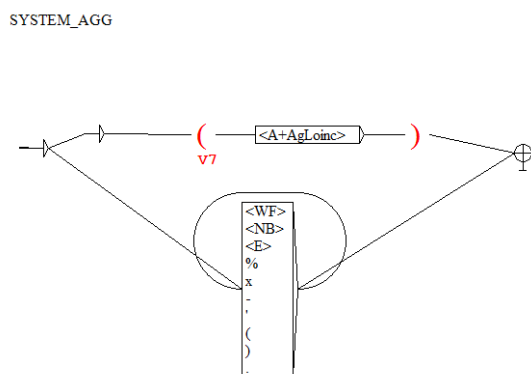


Figure 30 - La grammaire "SYSTEM_AGG"

Pour chaque élément retrouvé dans la phrase, il faut considerer sa presence mais aussi la possibilité de manque. Cet aspect c'était un des plus difficiles à traiter car, il arrive souvent que les laboratoires n'insèrent pas toutes les informations requises dans leur base de données. Le principe est que l'annotation sémantique et la

verification des informations doit continuer aussi si un élément manque dans sa place; par exemple, si l'unité de mesure n'est pas décrite en correspondance d'une certaine analyse, la grammaire doit continuer son travail de reconnaissance (annotation et contrôle). Pour chaque grammaire a été, donc, prévue (par la ramification vide) la possibilité de manque d'une quelque forme. On rappelle que le manque d'information ensemble à la variété des formes linguistiques locales est un des problemes les plus importants qui ont été affrontés dans ce travail.

L'unité de mesure determine la valeur pour l'attributs LOINC «Property» et « Timing ». Comme pour le cas du composant, la sous-grammaire « PROPERTY » (Figure 31) reconnais les formes appartenant à la catégorie « N+UnMis » en enregistrant leur valeurs dans la variable « V2 », et elle les annote en considérant les attributs associés par les propriétés « +Property » et « +Timing » décrite dans le dictionnaire.

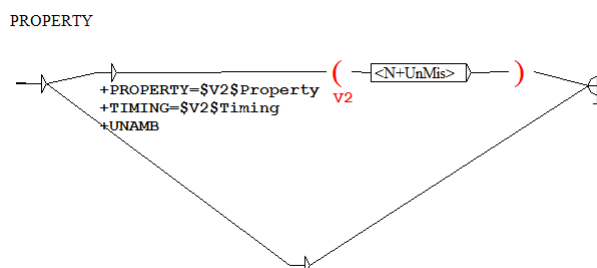


Figure 31 - La grammaire "PROPERTY"

Par exemple si on considère l'entrée « g/dl », formalisée comme ci-dessus :

$$g/dl, N+UnMis+Property=MCnc+Timing=Pt+Scale=Qn$$

la correspondant annotation sémantique obtenue sera «+PROPERTY=MCnc+TIMING=Pt ».

L'annotation sémantique obtenue par la grammaire hors-contexte LOINC est un exemple qui montre comment, à partir d'un certain langage, on peut créer par des

règles précises, un autre langage dérivant du précédent. On dit pour ce genre de grammaire qu'elles ne dépendent pas du contexte puisque leurs règles suivent une structure précise, et la reconnaissance des entrées ne dépend pas du contexte dans laquelle elles sont utilisées.

Pour tous les éléments, comme nous avons déjà expliqué, est prévue aussi la possibilité de manque ainsi que leur présence n'est pas ambiguë. L'opérateur « +UNAMB » utilisé en cas de présence d'un certain élément, permet d'enlever l'ambiguïté par rapport à son possible absence. Les grammaires à états-finis donnent les solutions tout en prenant en considération la formalisation graphique globale. « +UNAMB » permet à la grammaire de comprendre le fait qu'il ne doit pas considérer comme «solution» possible l'absence de l'élément puisque cet élément est présent. C'est-à-dire que, si la forme correspondante à une certaine catégorie est présente et reconnue par la grammaire, elle n'est pas ambiguë.

Les formes linguistiques des méthodes d'analyse sont reconnues par la grammaire « METHOD » qui travaille ensuite pour la grammaire qui, à son tours, annote les valeurs du composant, de la grandeur et du temps (Figure 32).

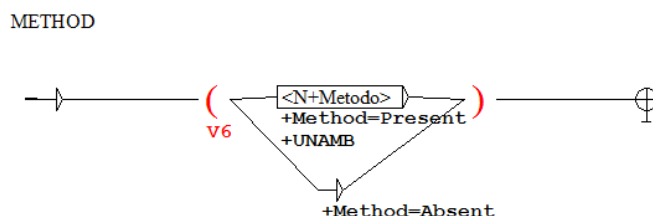


Figure 32 - La grammaire "METHOD"

Cette grammaire reconnaît les formes formalisées dans la catégorie « N+Metodo » en enregistrant leurs valeurs dans la variable « V6 » et en annotant sa présence ou son absence.

Pour les objectifs du travail présenté, les formes linguistiques qui décrivent une méthode d'analyse sont associées à deux seules valeurs considérées dans leurs en association à l'attribut officiel. Par exemple, on considère la forme :

Elettroforetico,N+Metodo+Method=X

«*Method =X* » indique la présence de la forme linguistique qui identifie la méthode adoptée, en annotant « +Method=Present » ou « +Method=Absent ». L'annotation obtenue sera « +Method=Present » si la forme est reconnue, ou « +Method=Absent » si la méthode n'est pas explicité.

L'élément qui identifie le milieu biologique est reconnu par la grammaire «SYSTEM» (Figure 33). Cette grammaire reconnaît les formes « N+Campione » en gardant sa valeur dans la variable « V4 » et annote les formes reconnues avec la valeur de la propriété sémantique « +System » associée à chaque entrée dans le dictionnaire.

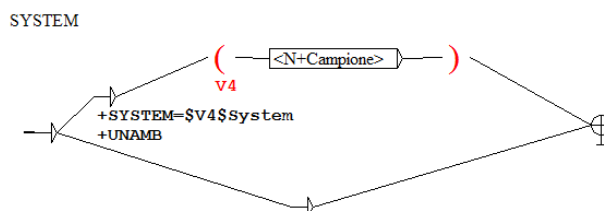


Figure 33 - La grammaire "SYSTEM"

Par exemple, pour la forme linguistique « Sanguo intero » décrite comme suit : « *Sanguo intero,N+Campione+System=Sanguo* » l'annotation produira « +SYSTEM=Sanguo ».

Pour la valeur de l'attribut qui représente l'échelle de mesure, la grammaire « SCALE » a été construite et posée après la grammaire qui identifie le milieu

biologique, et en suivant l'ordre des éléments pour comme décrits dans les phrases considérées (Figure 34).

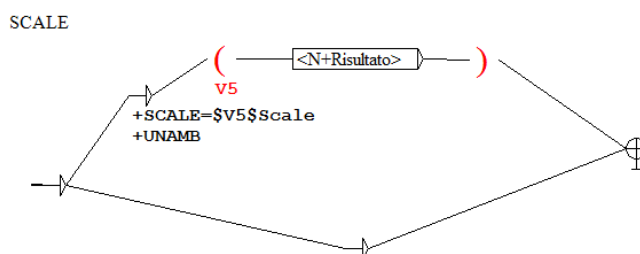


Figure 34 - La grammaire "SCALE"

Cette grammaire reconnaît les formes linguistiques appartenant à la catégorie « N+Risultato ». Elle garde la valeur retrouvée dans la variable « V5 » et annote la forme linguistique locale avec la forme décrite par la propriété « +Scale » associée aux entrées décrites par le dictionnaire. En considérant la forme linguistique « Numerico » formalisée dans « Numerico,N+Risultato+Scale=Qn+ », l'annotation sémantique correspondant produira comme résultat « +Scale=Qn ».

On a vu donc, comment la grammaire hors-contexte « LOINC » peut transformer une chaîne de termes exprimés dans une forme déterminée, en une chaîne d'autres termes exprimés en forme différente. On traite, donc, le passage d'un langage à un autre, par des règles précises et en considérant les objectifs à atteindre.

Dans l'image ci-dessus (Figure 35) on peut voir comment les valeurs insérées dans la base des données locales pour identifier le test et le code à lui associé « Albumina g/dl Elettroforetico Siero Numerico 2862-1 » ont été annotées dans la forme :

« LOINC+COMPONENT=Albumina+PROPERTY=MCnc+TIMING=Pt+Method=Present+SY STEM=Siero+SCALE=Qn ».

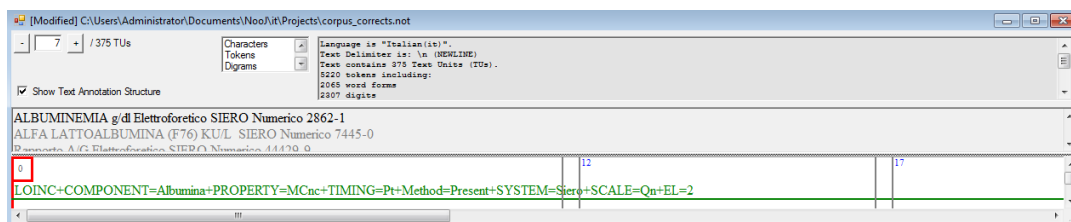


Figure 35 - Exemple d'annotation

4.5 Les contraintes contextuelles : les niveaux d'exactitude

Les contraintes contextuelles considérées dans la grammaire « LOINC » donnent la possibilité de vérifier l'exactitude du code associé à chaque examen par la comparaison entre les formes résultants du travail d'annotation.

On prend, donc, en considération la deuxième partie de la grammaire ; celle qui vient après l'annotation effectuée par la grammaire hiérarchique « ELEMENTS_LOINC » et analysée par le paragraphe précédent.

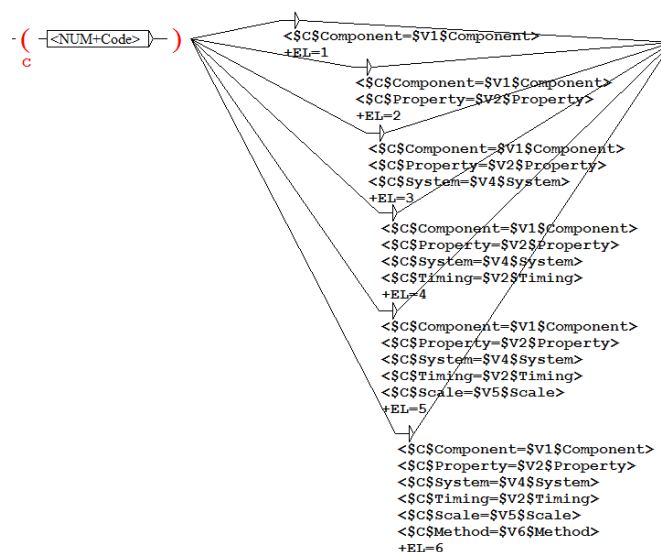


Figure 36 - Les contraintes pour la grammaire « LOINC »

Après la composition de la chaîne des termes annotés, la grammaire identifie les formes appartenant à la catégorie « NUM+Code » en gardant la valeur dans la variable « C ». En résumé, les valeurs correspondants gardées dans chaque variable considérée jusqu'à maintenant sont montrées dans le Tableau 10 ci-dessous.

Variable	Catégories d'entrés	Attributs LOINC suggérées
V1	N+Analiti (Composant)	Component <u>System</u>
V2	N+UnMis (Unité de mesure)	Property Timing <u>Scale</u>
V4	N+Campione (Milieu Biologique)	System <u>Timing</u>
V5	N+Risultato (Échelle)	Scale
V6	N+Metodo (Méthode d'analyse)	Method
V7	A+AgLoinc (Adjectif du composant)	<u>System</u>

Tableau 10 - Les variables de la grammaire "LOINC"

La variable « C » conserve la valeur (et donc le code même) des formes « N+Code » qui figurant dans la phrase analysée. La reconnaissance des codes est fondamentale puisque elle représente, pour chaque phrase analysée (voir paragraphe 4.1), le terme de comparaison entre les formes linguistiques résultant par l'annotation sémantique et les formes linguistiques officielles liées aux codes associés.

Pour effectuer cette comparaison, les contraintes contextuelles mise à disposition par logiciel NooJ ont été utilisés (Figure 36). Les contraintes considérées ont été divisées en groupe, chacun représentant un niveau d'exactitude différent. Les niveaux d'exactitude se différencient par le nombre d'éléments pour lesquels s'effectue le contrôle sur la comparaison des formes linguistiques.

- Le niveau 1, annoté par l'expression « +EL=1 » donne un résultat si le contrainte est satisfaite. La contrainte pour déterminer le niveau 1 d'exactitude se formalise dans la façon suivante :

« \$C\$Component=\$V1\$Component » et effectue une comparaison sur la propriété « +Component » des variables « C » et « V1 ». La variable « V1 » garde la valeur des formes « N+Analiti » (composant de l'examen) ayant propriété « +Component » qui prend la valeur officielle LOINC qui correspond. La variable « C » garde la forme des codes ayant la propriété « +Component » qui prend la correspondante valeur du composant associé au même code dans le système LOINC. Si le système vérifie cette contraintes, la chaine des termes annotés terminera avec l'expression « +EL=1 », comme dans l'exemple suivant :

On considère l'entrée « Albumina » décrite comme ci-dessus

« *Albumina,N+Analita+CORE+Component=Albumina+FLX=F22* »

et l'entrée « 2862-1 » décrite dans la forme

« *2862-1,*

NUM+Code+Component=Albumina+Property=MCnc+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=X ».

En considérant, par exemple, la phrase :

« *Albuminemia g/dl Elettroforetico Siero Numerico 2862-1* »

l'annotation sémantique produira le suivant résultat, obtenu par la comparaison satisfaisante (Component=Albumina) :

« *LOINC+COMPONENT=Albumina+PROPERTY=MCnc+TIMING=Pt+Method=Present+SYSTEM=Siero+SCALE=Qn+EL=1* »

- Le niveau 2, il fonctionne comme le niveau 1. En considérant les contrôles effectués sur deux valeurs associées aux propriétés intéressées, les contraintes formalisées pour ce niveau vérifient l'exactitude de la grandeur et du composant. Les contraintes « $\$C\$Component=\$V1\$Component$ » et « $\$C\$Property=\$V2\$Property$ » en particulier, vont comparer, respectivement, la valeur de la propriété « +Component » associée au code LOINC (variable « C ») avec la valeur de la propriété « +Component » du composant (variable « V1 »), et la valeur de la propriété « +Property » associée aux codes (variable « C ») avec la valeur de la propriété « +Property » associée à l'unité de mesure (variable « V2 »). Si les contraintes sont vérifiées, la chaîne de termes prend « +EL=2 » à la fin pour identifier que 2 représente les valeurs exactes pour l'association « formes locales –codes LOINC » considérée. Un possible exemple d'annotations en partant des entrées considérées est montré ci-dessous :

$V1 = \text{« Albumina,N+Analita+CORE+Component=Albumina+FLX=F22 »}$

$V2 = \text{« g/dl,N+UnMis+Property=MCnc+Timing=Pt+Scale=Qn »}$

$C = \text{« 2862-1, »}$

$NUM+Code+Component=Albumina+Property=MCnc+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=X$ »

$PHRASE = \text{« Albuminemia g/dl Elettroforetico Siero Numerico 2862-1 »}$

$ANNOTATION=LOINC+COMPONENT=Albumina+PROPERTY=MCnc+TIMING=Pt+Method=Present+SYSTEM=Siero+SCALE=Qn+EL=2$

- Le niveau 3, ajoute aux contraintes déjà considérées le contrôle sur la valeur de la propriété « +System » associée aux formes appartenant à la catégorie « N+Campione » (milieu biologique) et gardé dans la variable « V4 ». La contrainte ajoutée, en effet, se présente dans la forme suivante « \$C\$System=\$V4\$System » et effectue une comparaison entre la valeur de la propriété « +System » associée au code (variable « C ») et la valeur de la propriété « +System » associée au milieu biologique (variable « V4 »). Si les trois contraintes se vérifie au même temps l'annotation ajoutera « +EL=3 » à la fin de la chaîne des termes déjà annotés. Exemple :

VI= « Proteinuria,Proteina,N+Analita+Component=Proteina+System=Urine »

V2= « mg/24h,N+UnMis+Property=MRat+Timing=24h+Scale=Qn »

V4= « Urina,N+Campione+System=Urine »

C= « 2889-4,

NUM+Code+Component=Proteina+Property=MRat+Timing=24h+System=Urine+Scale=Qn+Method=Absent »

PHRASE= « Proteinuria delle 24h mg/24h Urine Numerico 2889-4 »

ANNOTATION=« LOINC+Component=Proteina+Property=MRat+Timing=24h+Method=Absent+System=Urine+Scale=Qn+EL=3

- Le niveau 4, prévoit le contrôle sur la valeur de la propriété « Timing » déduite par l'unité de mesure. La variable « V2 » garde la valeur de l'unité de mesure dans laquelle on peut déduire aussi la grandeur, et l'échelle de mesure concernant l'examen considéré.

Dans ce niveau on considère la comparaison de la valeur de la propriété « +Timing » associée au code (variable « C ») avec la valeur de la propriété « +Timing » associée à l'unité de mesure (variable « V2 »). La contrainte ajoutée à ces décrits pour déterminer le niveau 3, se présente dans la forme suivante : « \$C\$Timing=\$V2\$Timing » ; si la contrainte est vérifiée, l'annotation ajoute à la fin de la chaîne des termes annotés « +EL=4 ». On considérant le même exemple montré pour le niveau 3, on obtient :

VI= « *Proteinuria,Proteina,N+Analita+Component=Proteina+System=Urine* »

V2= « *mg/24h,N+UnMis+Property=MRat+Timing=24h+Scale=Qn* »

V4= « *Urina,N+Campione+System=Urine* »

C= « 2889-4,

NUM+Code+Component=Proteina+Property=MRat+Timing=24h+System=Urine+Scale=Qn+Method=Absent »

PHRASE= « *Proteinuria delle 24h mg/24h Urine Numerico 2889-4* »

ANNOTATION=« *LOINC+Component=Proteina+Property=MRat+Timing=24h+Method=Absent+System=Urine+Scale=Qn+EL=4*

- Le niveau 5, pareillement, ajoute au groupe de contraintes du niveau 4 le contrôle sur la propriété « +Scale » identifiée par la variable « V5 ». la contrainte « $\$C\$Scale=\$V5\$Scale$ » vérifie, en particulier, la comparaison entre la valeur de la propriété « +Scale » associée au code (variable « C ») et la valeur de la propriété « +Scale » associée aux formes reconnues et appartenant à la catégorie « N+Risultato ». Si les contraintes sont vérifiées conjointement, l'annotation va ajouter « +EL=5 » à la fin de la chaîne. Ce concept est clairement valide pour tous les niveaux qui se réalisent seulement si les comparaisons prévues se vérifient au même temps. On ajoute les considérations sur la variable V5 à l'exemple montré pour les précédents niveaux :

VI= « *Proteinuria,Proteina,N+Analita+Component=Proteina+System=Urine* »

V2= « *mg/24h,N+UnMis+Property=MRat+Timing=24h+Scale=Qn* »

V4= « *Urina,N+Campione+System=Urine* »

V5=*Numerico,N+Risultato+Scale=Qn+FLX=F78*

C= « 2889-4,

NUM+Code+Component=Proteina+Property=MRat+Timing=24h+System=Urine+Scale=Qn+Method=Absent »

PHRASE= « *Proteinuria delle 24h mg/24h Urine Numerico 2889-4* »

*ANNOTATION=« LOINC+Component=Proteina+Property=MRat+Timing=24h+Method
=Absent+System=Urine+Scale=Qn+EL=5*

- Le niveau 6 il est le dernier. Les contraintes considérées à ce niveau font référence à toutes les valeurs des propriétés vues dans les niveaux précédents en ajoutant le contrôle sur la méthode d'analyse. La contrainte « $\$C\$Method=V6\$Method$ » se réfère à la comparaison entre la valeur de la propriété « +Method » associée au code (variable « C ») et la valeur de la propriété « +Method » associée aux formes linguistiques appartenant à la catégorie « N+Metodo » (variable « V6 »). Dans le corpus analysé on retrouve des examens qui appartiennent à ce niveau d'exactitude, malgré les laboratoires difficilement arrivent à formaliser si en détail leur base de données. Si les contraintes sont vérifiées, la grammaire ajoute à la fin de la chaîne annotée la propriété « +EL=6 ». En considérant le même cas vu dans le niveau 5, on vérifie l'absence de la méthode dans la phrase et la conséquente assignation au niveau 6. Dans la grammaire « METHOD » (Figure 32), il est possible de vérifier qu'en cas d'absence de la forme linguistique, la variable « V6 » prend comme valeur de la propriété « +Method=Absent ».

V1= « Proteinuria,Proteina,N+Analita+Component=Proteina+System=Urine »

V2= « mg/24h,N+UnMis+Property=MRat+Timing=24h+Scale=Qn »

V4= « Urina,N+Campione+System=Urine »

V5= Numerico,N+Risultato+Scale=Qn+FLX=F78

V6 = +Method=Absent

C= « 2889-4,

NUM+Code+Component=Proteina+Property=MRat+Timing=24h+System=Urine+Scale=Qn+Method=Absent »

PHRASE= « Proteinuria delle 24h mg/24h Urine Numerico 2889-4 »

*ANNOTATION=« LOINC+Component=Proteina+Property=MRat+Timing=24h+Method
=Absent+System=Urine+Scale=Qn+EL=6*

Les contraintes appartenant à chaque niveau ont été déterminées sur l'expérience des experts appartenant au groupe « LOINC Italia » (voir paragraphe 2.2) et impliqués dans la validation des codes et les associations les plus fréquents vérifiées dans les bases de données locales.

Particulièrement, et pour tester le système, on a choisi de procéder en assignant en manier incrémentale les contraintes pour chaque niveau en considérant l'ordre de la structure LOINC et les résultats les plus significatifs.

On pense par exemple, aux couples « Composant - Propriété » et « Échelle - Timing ». Le premier a, sûrement, une valeur supérieure étant que le contrôle sur le composant résulte plus difficile et copieux, étant la variété des formes locales. Il faut clarifier, en tout cas, que, pour le système LOINC, le mappage des attributs maintient la même importance pour tous sans distinction. L'évaluation sur le choix des contraintes a été effectuée en basant les nécessités et les considérations référées à l'objectif spécifique qu'on vise à atteindre.

4.5.1 Les contraintes d'aide aux experts : les niveaux « HELP »

On présente ici les contraintes contextuelles définies (HELP). Les niveaux d'exactitude créés par ces contraintes ont été obtenus par la vérification croisée des valeurs correspondant aux propriétés déjà considérées pour déterminer le niveau d'exactitude par les contraintes «standard», mais déduites par des autres éléments. La valeur ajoutée par ces contraintes, repose dans la déduction de certaines valeurs des attributs LOINC, en considérant des entrées lexicales différentes et qui ne sont pas directement associées aux valeurs objet de la comparaison. Ce type de contrainte agit particulièrement, en assignant l'examen à un certain niveau, là où le contrôle prévu par les contraintes standard n'est pas vérifié.

L'objectif de ces contraintes est de vérifier si le niveau d'exactitude associé à un examen est effectivement le niveau correct ou si certaines associations peuvent être déduites par d'autres formes linguistiques.

Si on considère par exemple un examen appartenant au niveau 2 d'exactitude ; l'association au niveau 3 doit venir par la comparaison de la valeur de la propriété «+System» du code et par la valeur de la propriété «+System» des formes linguistiques appartenant à la catégorie «N+Campione». Si cette comparaison n'est pas vérifiée, les contraintes «System_HELP» peuvent donner aux examens une autre possibilité pour passer au niveau 3.

Dans le Tableau 11, sont montrées les variables sur lesquelles on focalise l'attention pour mieux expliquer les mécanismes de contraintes «HELP».

Variable	Catégories d'entrés	Attributs LOINC suggérées
V1	N+Analiti (Composant)	Component <u>System</u>
V2	N+UnMis (Unité de mesure)	Property Timing <u>Scale</u>
V4	N+Campione (Milieu Biologique)	System <u>Timing</u>
V7	A+AgLoinc (Adjectif du composant)	<u>System</u>

Tableau 11 - Les variables pour les contraintes "HELP"

Pour effectuer les contrôles de support, on considère, particulièrement, les possibles déductions suivantes : celle de la valeur associée à la propriété «+System» à partir de la variable «V1» (le composant), celle de la valeur associée à la propriété «+Scale» à partir de la variable «V2» (unité de mesure), celle de la valeur associée à la propriété «+Timing» à partir de la variable «V4» (milieu biologique) et celle de la valeur associée à la propriété «+System» à partir de la variable «V7» (adjectif explicité dans le composant). On voit, donc, les déductions croisées des valeurs des propriétés : composant – milieu biologique, unité de mesure – échelle de mesure, milieu biologique – temps, adjectif du composant – milieu biologique.

Les niveaux d'exactitude déterminés par ces types de contraintes identifient le seuil à partir duquel la vérification de l'expert est nécessaire pour l'effective assignation à un certain niveau. C'est-à-dire que, les examens qui arrivent au niveau 3 d'exactitude, satisfaits par les contraintes « HELP », ne peuvent pas passer au 4 seulement ensuite au contrôle des experts. Pour cette motivation, ce type de contraintes prend en considération les passages d'un niveau à un autre et en particulier ceux concernant le passage du niveau 2 au niveau 3, du niveau 3 au niveau 4 et du niveau 4 au niveau 5.

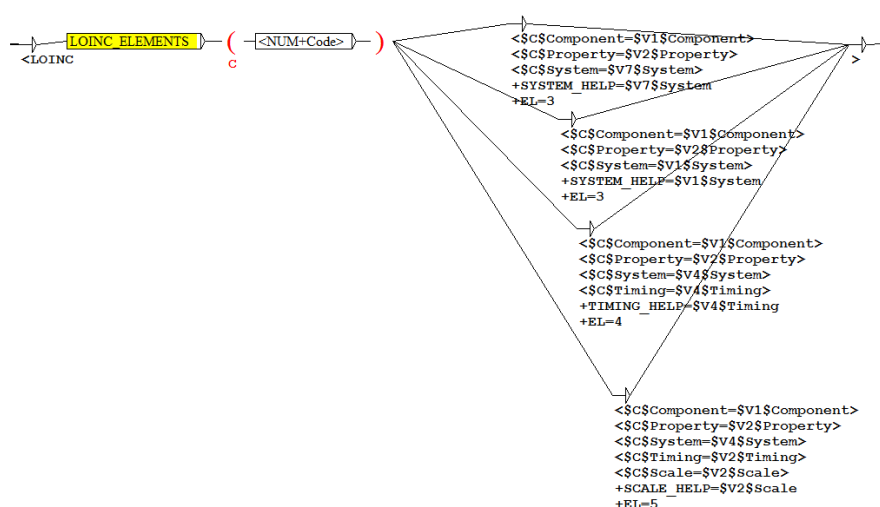


Figure 37 - Les contraintes "HELP"

Pour aider au passage du niveau 2 au niveau 3, deux contraintes ont été considérés pour déterminer la valeur du milieu biologique à partir des entrées différentes :

- la contrainte « \$C\$System=\$V7\$System », comme dans le cas des contraintes standards, vérifient la comparaison entre la valeur de la propriété « +System » du code (variable « C ») avec la valeur de la propriété « +System » associée aux adjectifs (variable « V7 »). Cette contrainte est très importante si on prend en considération le fait que les laboratoires

n'expriment pas toujours le milieu biologique avec le même niveau de détail requis par LOINC;

- la contrainte « $\$C\$System=\$V1\$System$ » agit sur le composant qui peut suggérer la valeur de la propriété « +System ». On considère, par exemple, que les formes linguistiques en « emia » ou « uria », gardent la sémantique par rapport à la définition du composant et au milieu biologique utilisé.

Le passage du niveau 3 au niveau 4 est déterminé par la contrainte « $\$C\$Timing=\$V4\$Timing$ ». Cette contrainte vérifie la valeur de la propriété « +Timing » du code avec la valeur de la propriété « +Timing » associée au milieu biologique. On considère par exemple que l'entrée « Urina delle 24 ore » (« Urine de 24 heure »), appartenant à la catégorie « N+Risultato » (milieu biologique), peut donner une suggestion sur la valeur du temps si, dans l'unité de mesure, (contrainte standard) cette valeur n'est pas indiqué.

Pareillement, pour passer du niveau 4 au niveau 5 on considère la contrainte « $\$C\$Scale=\$V2\$Scale$ ». Cette contrainte se réfère à la possibilité de déduire, à partir de l'unité de mesure, la valeur de la propriété « +Scale » suggérée par cette l'unité de mesure.

Les résultats obtenus par la formalisation de ces contraintes sont très intéressants puisque les formes linguistiques locales sont parfois non exhaustives et non conformes aux règles imposées par le standard LOINC.

Par exemple, beaucoup de fois en correspondance du milieu biologique, les laboratoires ne sont pas précis comme le forme linguistique demandé par LOINC. Ce qui rend la vérification d'exactitude tâche difficile. Les contrôles croisés sont nécessaires si on considère la complexité et la variété des morphèmes et des expressions utilisées dans le langage commun aux laboratoires. La variété des formes et leur rapport avec la langue dans laquelle sont exprimées rendent difficile leurs associations à une terminologie plus standard et privé des aspects linguistiques possibles.

Chapitre V

5.1 Les résultats obtenus : la formalisation de la connaissance

Dans ce chapitre on analyse les résultats obtenus et les utilités produites par l'application à textes techniques médicaux à domaine spécifique du système de grammaires réalisé.

La contribution du système de règles présenté dans cette thèse se situe dans l'opération de validation des codes, effectuée par les experts du système LOINC (Logical Observation Identifiers Names and Codes) . En particulier, une fois que les laboratoires ont associé à leurs tests un et un seul code LOINC, ces associations doivent être évaluées afin de réaliser un contrôle qualitatif et rendre effective leur intégration dans les rapports de laboratoires en format électronique.

Cette opération de validation demande aux experts beaucoup de temps étant nécessaire, pour chaque examen analysé, la reconstruction du processus de mapping en considérant en détail toutes les étapes qu'ont porté vers un choix déterminé. Pendant cette reconstruction l'expert comprend le niveau d'exactitude (totalement erroné – presque correct – correct) de chaque examen en se basant sur les informations contenues dans les catalogues d'examens locaux rendus accessibles par les laboratoires et sur sa connaissance ainsi que sur son expérience acquise au cours des années pour effectuer ce genre d'opérations.

Par conséquent, la capacité de donner une validation aux associations, et donc la capacité de vérifier l'exactitude du code assigné, dépend de la connaissance du domaine mais, surtout, de la connaissance profonde des règles adoptées par le système de codage LOINC pour décrire les examens.

La connaissance des experts du système LOINC permet de lier facilement les données présentes dans la base de données locale, si toutes les informations demandées ont été explicitées pour décrire un examen, à la terminologie officielle LOINC.

Ce processus, à ce jour, repose sur la reconnaissance des formes linguistiques locales qui se déroulent sans aucun support automatique et il est principalement le résultat de la capacité de validateurs.

Grâce à la formalisation des formes linguistiques avec leurs significations intrinsèques liées à la connaissance des experts du domaine a été possible d'effectuer une analyse linguistique qui va au-delà de la reconnaissance et l'association des formes linguistiques à une catégorie donnée, en analysant syntaxiquement les phrases (paragraphe 4.1) par la connaissance formalisée.

La réalisation de la ressource utilisée pour l'analyse linguistique objet de ce travail a permis, grâce à la définition détaillée des formes locales, la formalisation de la connaissance et de l'expérience des experts impliqués dans les opérations de validation de codes LOINC.

La formalisation obtenue par la réalisation de dictionnaires, ainsi que les grammaires qu'identifient le niveau d'exactitude assigné à chaque code, représentent les premières étapes vers l'automatisation de l'entier processus de mapping comme de validation à support des usagers LOINC. Les dictionnaires électroniques et leur formalisation contenant la connaissance intrinsèque aux formes linguistiques du domaine ont permis tant la reconnaissance des formes locales que l'extraction d'informations spécifiques contenues dans la base de données analysée.

Les informations extraites de façon automatique grâce à l'application de la ressource linguistique créée sont très détaillées, et leur détection aurait demandé une analyse plus détaillée et copieuse en termes de temps et de précision, si effectuée manuellement. L'application de TALN conçue a permis de réduire au maximum les temps utilisés pour analyser les bases des données locales composées par les associations entre les descriptions locales et les codes LOINC.

5.1.1 Annotation sémantique : la validation des codes

Le corpus analysé est composé par 375 examens sélectionnés sur la base des catégories de tests choisies pour dérouler ce travail. L'analyse linguistique effectuée avec NooJ prend en considération les deux dictionnaires réalisés « Analiti.dic » et « Elements_LOINC.dic », pour l'analyse lexicale, et les deux grammaires « LOINC » et « LOINC_HELP », pour l'analyse syntactique.

La technique d'annotation sémantique utilisée reconnaît les formes locales, comme on a vu dans le paragraphe 4.4, et produit à partir d'un langage local un autre langage spécifique et déterminé par règles linguistiques demandés par le système LOINC dans ce cas, qui répond aux objectifs fixés.

On passe donc d'une chaîne de termes, dérivante de la description locale d'un examen, à une autre chaîne de termes résultat de l'annotation des informations de départ et prête à être comparée avec les formes officielles associées aux codes LOINC. Cette transformation porte nombreux avantages tant pour la validation que pour la recherche de codes. En effet, les chaînes de termes résultants par l'annotation sémantique peuvent être utilisées comme point de départ pour effectuer une transformation automatique des termes de recherche utilisées pour retrouver un code déterminé.

On considère les exemples montrés ci-dessus dans les Figure 38 et Figure 39.

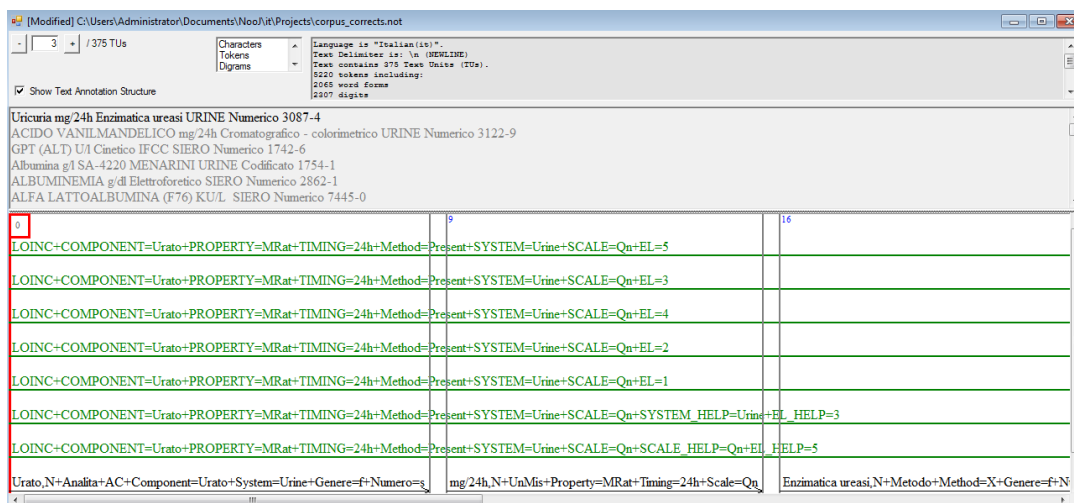


Figure 38 - Résultats de l'annotation sémantique: les niveaux d'exactitude

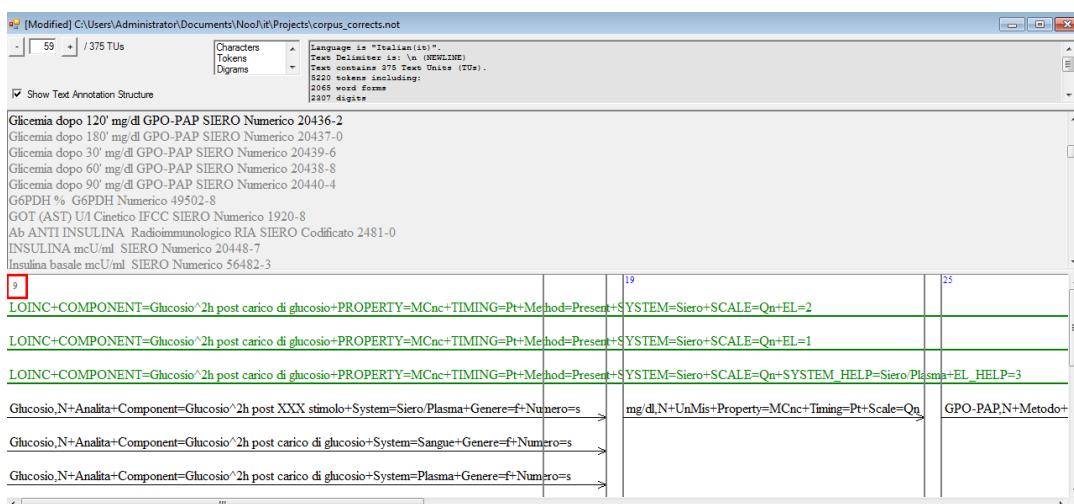


Figure 39 - Résultats de l'annotation sémantique: les contraintes HELP

Les experts du système LOINC appartenant au groupe LOINC Italia utilisent la fonction de recherche fournie par le logiciel RELMA (Regenstrief LOINC Mapping Assistant) pour effectuer les opérations de validation à partir des catalogues d'examen locaux. En correspondance de la chaîne de recherche dans le logiciel RELMA, après avoir choisi la langue italienne, la description « Uricuria mg/24h Enzimatica ureasi URINE Numerico » a été insérée

sans obtenir résultats ; c'est le cas évident que le langage local trouve ces limitations en se rapprochant à la terminologie standard (Figure 40).

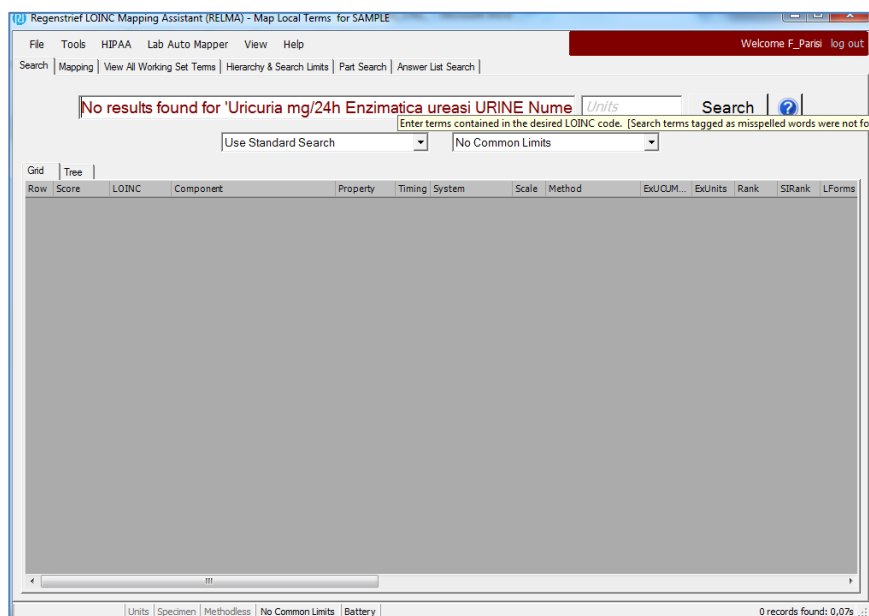


Figure 40 - La fonction de recherche dans RELMA: la chaîne de termes locales

Au contraire, la chaîne de termes obtenue par l'annotation sémantique effectuée (« Urato MRat 24h Urine Qn ») identifie l'examen de façon univoque en donnant un résultat précis dépourvu d'ambiguïté (Figure 41). La chaîne de mots obtenue reconnaît le morphème en « uria » en identifiant tant le composant que la valeur du milieu biologique associé. Par ailleurs, l'unité de mesure est annotée par la valeur de la propriété « +Property » correspondant, et l'échelle de mesure passe de la forme « Numerico » à le sigle « Qn ».

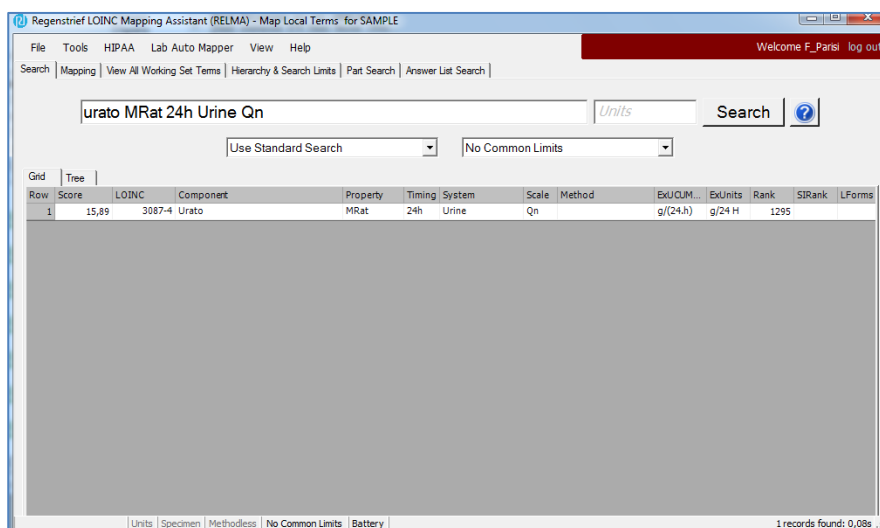


Figure 41 - La fonction de recherche dans RELMA: le resultat de l'annotation

La même chose peut être vérifiée pour l'exemple montré dans la Figure 42. La forme linguistique locale « Glicemia dopo 120' » n'est pas reconnue par RELMA qui demande, pour ce genre d'expressions temporelles associées au composant, une forme déterminée (« Glucosio, 2h post XXX »).

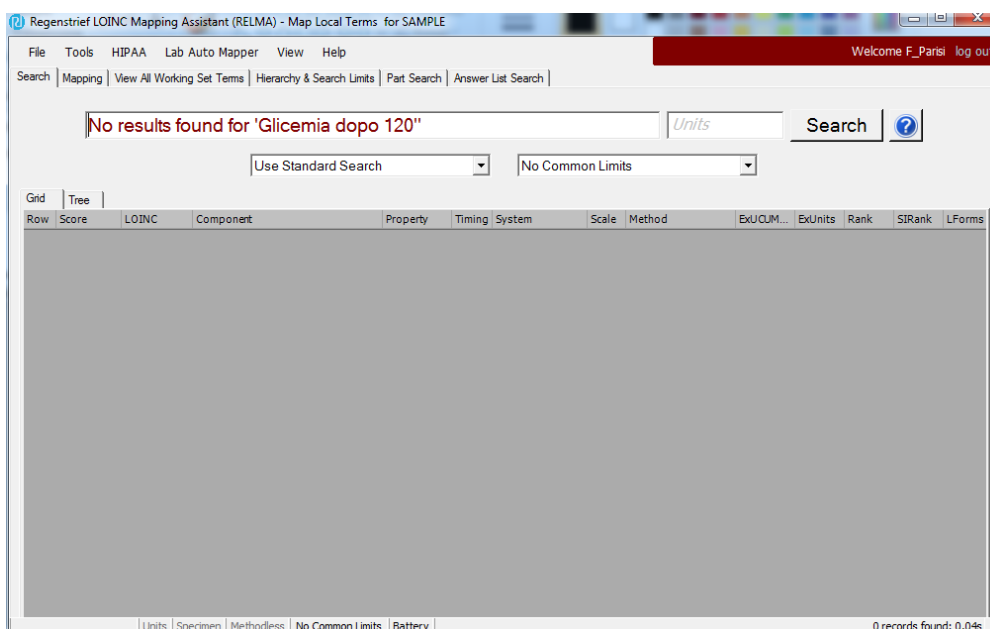


Figure 42 - La fonction de recherche dans RELMA: les expressions temporelles locales

Cette forme linguistique locale annotée, au contraire, avec seulement la définition du composant, identifie un ensemble de 49 codes LOINC possibles (Figure 43), qui se réduit à 5 si on ajoute les autres formes reconnues par l'annotation des autres attributs (Figure 44).

The screenshot shows the Regenstrief LOINC Mapping Assistant (RELMA) interface. The search bar contains 'Glucosio 2h post'. The results table below shows 49 records found. The table has columns: Row, Score, LOINC, Component, Property, Timing, System, Scale, Method, ExUQM..., ExUnits, Rank, SRank, LFor.

Row	Score	LOINC	Component	Property	Timing	System	Scale	Method	ExUQM...	ExUnits	Rank	SRank	LFor
1	13,572	58686-7	C peptide^2h post carico di glucosio	MCnc	Pt	Siero/Plasma	Qn		ug/dL	ug/L			
2	13,5406	47588-9	C peptide^2h post carico di glucosio	SCnc	Pt	Siero/Plasma	Qn		nmol/L	nmol/L			
3	13,572	25706-3	Chetoni^2h post carico di glucosio	MCnc	Pt	Urine	Qn	Test strip	mg/dL	mg/dL			
4	13,572	58574-5	Corticotropina^2h post carico di ...	MCnc	Pt	Plasma	Qn		pg/mL	pg/mL			
5	13,572	58575-2	Corticotropina^2h post carico di ...	SCnc	Pt	Plasma	Qn		pmol/L	pmol/L			
6	11,9922	30345-3	Glucosio^2h post 1,2 g/kg lattosio...	MCnc	Pt	Siero/Plasma	Qn		mg/dL	mg/dL			
7	14,7287	1514-9	Glucosio^2h post 100 g glucosio...	MCnc	Pt	Siero/Plasma	Qn		mg/dL	mg/dL	896		
8	14,7287	14757-9	Glucosio^2h post 100 g glucosio...	SCnc	Pt	Siero/Plasma	Qn		mmol/L	mmol/L		896	
9	14,7287	1516-4	Glucosio^2h post 100 g glucosio...	MCnc	Pt	Urine	Qn		mg/dL	mg/dL			
10	15,2788	41024-1	Glucosio^2h post 50 g glucosio per...	MCnc	Pt	Siero/Plasma	Qn		mg/dL	mg/dL			
11	11,9922	1517-2	Glucosio^2h post 50 g lattosio per os	MCnc	Pt	Siero/Plasma	Qn		mg/dL	mg/dL			
12	11,9922	14758-7	Glucosio^2h post 50 g lattosio per os	SCnc	Pt	Siero/Plasma	Qn		mmol/L	mmol/L			
13	11,9922	6750-4	Glucosio^2h post 50 g lattosio per os	PrThr	Pt	Urine	Ord	Test strip					
14	14,7287	1518-0	Glucosio^2h post 75 g glucosio per...	MCnc	Pt	Siero/Plasma	Qn		mg/dL	mg/dL	835		
15	14,7287	14995-5	Glucosio^2h post 75 g glucosio per...	SCnc	Pt	Siero/Plasma	Qn		mmol/L	mmol/L		835	
16	14,7287	1520-6	Glucosio^2h post 75 g glucosio per...	MCnc	Pt	Urine	Qn		mg/dL	mg/dL			
17	14,7287	6751-2	Glucosio^2h post 75 g glucosio per...	PrThr	Pt	Urine	Ord	Test strip					
18	12,0236	51769-8	Glucosio^2h post carico con...	MCnc	Pt	Siero/Plasma	Qn		mg/dL	mg/dL			
19	11,9922	26780-7	Glucosio^2h post carico con lattosio...	MCnc	Pt	Siero/Plasma	Qn		mg/dL	mg/dL			
20	12,0236	40279-2	Glucosio^2h post carico con lattosio...	SCnc	Pt	Siero/Plasma	Qn		mmol/L	mmol/L			
21	17,0736	42613-0	Glucosio^2h post carico con lattosio...	MCnc	Pt	Urine	On		mo/dL	mo/dL			

Figure 43 - Ensemble des codes obtenu en utilisant les formes locales

The screenshot shows the Regenstrief LOINC Mapping Assistant (RELMA) interface. The search bar contains 'Glucosio 2h post carico di glucosio MCnc Pt Siero/Plasma'. The results table below shows 5 records found. The table has columns: Row, Score, LOINC, Component, Property, Timing, System, Scale, Method, ExUQM..., ExUnits, Rank, SRank, LForms.

Row	Score	LOINC	Component	Property	Timing	System	Scale	Method	ExUQM...	ExUnits	Rank	SRank	LForms
1	25,4812	58686-7	C peptide^2h post carico di glucosio	MCnc	Pt	Siero/Plasma	Qn		ug/dL	ug/L			
2	22,8089	51769-8	Glucosio^2h post carico con fruttosio per os	MCnc	Pt	Siero/Plasma	Qn		mg/dL	mg/dL			
3	22,7833	26780-7	Glucosio^2h post carico con lattosio per os	MCnc	Pt	Siero/Plasma	Qn		mg/dL	mg/dL			
4	27,593	20436-2	Glucosio^2h post carico di glucosio	MCnc	Pt	Siero/Plasma	Qn		mg/dL	mg/dL	884		
5	25,4812	55507-8	Somatotropina^2h post carico di glucosio	MCnc	Pt	Siero/Plasma	Qn		ng/mL	ug/L, n...			

Figure 44 - Ensemble des codes obtenu en utilisant la chaîne de termes annotée

La forme linguistique annotée arrive à un niveau de détail supérieur en utilisant la forme « Glucosio^{2h} post carico di glucosio », qui identifie précisément le code cherché, à choisir dans un ensemble de possibilités réduit du 90% par rapport à la recherche initiale.

Les avantages de l'application du système d'annotation aux formes linguistiques locales utilisées dans la recherche de codes ont été vérifiés pour tous les examens faisant partie de notre corpus (375 examens), en obtenant pour chacune une amélioration, en terme de réduction de l'ensemble de possibles codes corrects de plus du 70% par rapport à ceux qui sont identifiés en utilisant les formes locales.

Par ailleurs, souvent les formes locales se présentent en forme totalement différente (acronymes ou synonymes) par rapport à la terminologie standard et ne produisent aucun résultat pendant la recherche. Par conséquent, l'identification d'un ensemble de possibles codes suite au processus d'annotation déroulé représente déjà un très bon résultat.

5.1.2 Annotation sémantique : les niveaux d'exactitude

Le travail d'annotation effectué assigne à chaque examen un niveau d'exactitude par rapport à la comparaison entre les formes locales et les formes officielles associées au code choisi pour identifier un test déterminé.

Les validations peuvent être organisées sur la base de leur niveau d'exactitude identifié par le système en indiquant la complexité des analyses. Grâce au système réalisé, les experts LOINC peuvent décider de commencer à valider les examens ayants le niveau d'exactitude le plus élevé qui demande un temps d'analyse mineur, et passer ensuite à la validation de ceux associés à un niveau inférieur. Ces derniers demandent une attention majeure et donc un temps d'analyse plus long, ainsi

qu'une éventuelle rencontre avec les experts de laboratoire qui ont effectué l'association aux codes des examens locaux.

Les contraintes contextuelles (paragraphe 4.5) effectuent un contrôle sur l'exactitude de chaque forme locale par la comparaison exacte avec les valeurs associées aux propriétés sémantiques identifiées pour décrire les codes. Chaque examen reçoit l'étiquette sémantique correspondant à chaque niveau d'exactitude d'appartenance différenciée par les grammaires d'extraction d'information décrites dans les paragraphes suivants (Figure 45).

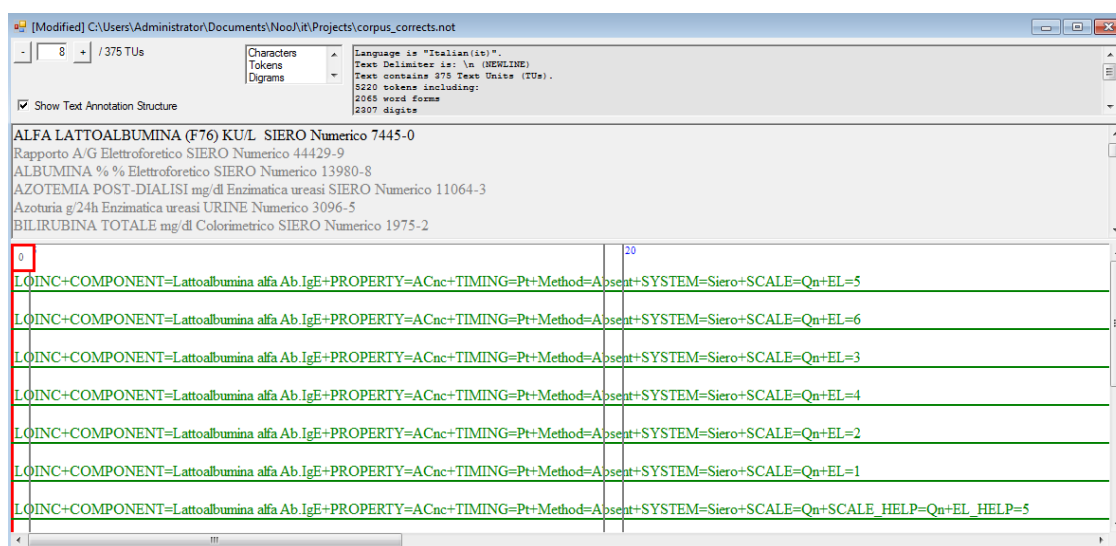


Figure 45 - Les résultats obtenus par les contraintes contextuelles

On a divisé les contraintes en deux typologie : « standard » quand elles effectuent la comparaison directe entre la valeur d'un attribut LOINC et l'information correspondant dans la chaîne de termes décrivant l'examen locale ; « HELP » quand elles effectuent une comparaison croisée entre des axes qui peuvent suggérer indirectement la valeur d'un attribut déterminé et leurs correspondants valeurs associées aux propriétés sémantiques décrites pour le code.

La différenciation de ces contraintes a porté nombreux bénéfices pour la vérification de l'exactitude des données.

Le développement de l'analyse linguistique présentée sur la base de la connaissance ainsi que de l'expérience des experts impliqués dans les opérations de validation nous a permis de souligner que dans la majorité des cas, les valeurs propres au système LOINC peuvent être déduites à partir des autres éléments appartenant à la description du test ; c'est-à-dire que beaucoup de fois l'exactitude d'une association « formes locales – code LOINC » est déterminée par la déduction des valeurs des attributs à partir des éléments qui ne représentent pas directement l'attribut en cause.

L'exemple montré dans la Figure 46 est significatif pour comprendre l'utilité de ce genre de contraintes.

La chaîne de termes locale « Uricemia mg/dl Enzimatica ureasi Urine Numerico » représente la description d'un test retrouvé dans le catalogue local.

La définition de l'examen qui prend le suffixe « emia » peut suggérer la valeur du milieu biologique (paragraphe 4.2.1). L'ensemble des possibles milieux biologiques fait référence au concept qui associe au suffixe « emia » le taux d'un quelque élément testé dans le « Sang ».

Au contraire, dans la chaîne de termes locale, on remarque en correspondance du milieu biologique la valeur « Urine », qui est complètement en opposition par rapport aux possibles valeurs suggérées par la définition du nom de l'examen.

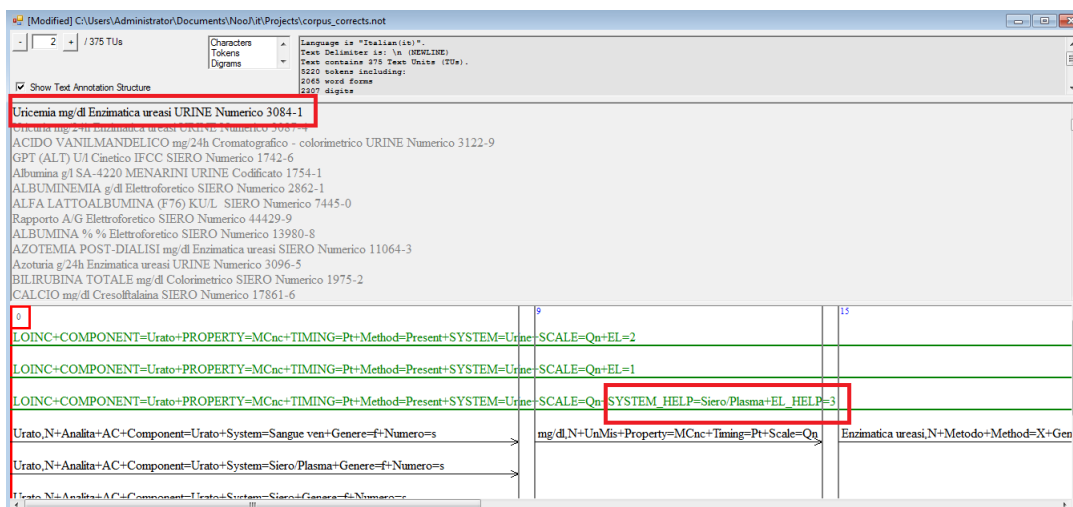


Figure 46 - L'utilité des contraintes HELP

Les contraintes « standard » associent l'examen au niveau 2 d'exactitude, alors que par les contraintes « HELP » l'examen passe au niveau 3, en résultant donc correct pour 3 élément.

Dans ce cas, l'ambiguïté liée à l'identification de la valeur pour le milieu biologique est résolue grâce à l'expérience des validateurs, permettant de présupposer que la valeur est gardée plutôt dans le composant étant l'attribut sur lequel les laboratoires focalisent plus d'attention. Par conséquent, l'association de la forme « Urine » comme milieu biologique représente une erreur qui le laboratoire a constaté pendant l'opération de mapping, en donnant plus d'importance à la définition de l'examen, mais pas modifié dans la base de données. L'aspect les plus importants est que notre système relève l'erreur et, sur la base de la connaissance formalisée, suit le processus d'association effectué par le laboratoire qui donne priorité à la signification intrinsèque du composant.

5.1.3 Annotation sémantique : la ressource XML

Les informations obtenues par l'annotation sémantique ont été décrites de façon automatique en format XML, suite à l'exportation des résultats possibles dans NooJ. La ressource réalisée permet de collecter les formes locales et les informations annotées dans un format réutilisable et capable d'être interrogé ou intégré avec d'autres outils.

La reconnaissance des formes locales et leur transformation suivant les formalismes linguistiques demandés par le système LOINC, représente une des utilités les plus importantes obtenues par ce travail, en considérant aussi la recueille automatique effectuée en enregistrant les données et les résultats obtenus en format XML.

Les chaînes de termes annotés sont enregistrées automatiquement dans la ressource XML fournie par le logiciel NooJ. Le stockage des données dans une ressource XML donne la possibilité de construire une base de connaissance qui peut s'alimenter et être manipulée automatiquement : un système automatique de recueille de formes locales associées aux définitions officielles.

```

<LOINC COMPONENT="Urato" Method="Absent" SYSTEM="Urine" SCALE="Nar" EL="1">
ACIDO URICO URINE Testo 14932-8
</LOINC>
<LOINC COMPONENT="Urato" PROPERTY="MCnc" TIMING="Pt" Method="Present" SYSTEM="Urine" SCALE="Qn" EL="1">
Uricemia mg/dl Enzimatica ureasi URINE Numerico 3084-1
</LOINC>
<LOINC COMPONENT="Urato" PROPERTY="MRat" TIMING="24h" Method="Present" SYSTEM="Urine" SCALE="Qn" EL="1">
Uricuria mg/24h Enzimatica ureasi URINE Numerico 3087-4
</LOINC>
<LOINC COMPONENT="Vanillilmandelato" PROPERTY="MRat" TIMING="24h" Method="Present" SYSTEM="Urine" SCALE="Qn" EL="1">
ACIDO VANILMANDELICO mg/24h Cromatografico - colorimetrico URINE Numerico 3122-9
</LOINC>
<LOINC COMPONENT="Alanina aminotransferasi" PROPERTY="CCnc" TIMING="Pt" Method="Present" SYSTEM="Siero" SCALE="Qn" EL="1">
GPT (ALT) U/l Cinetico IFCC SIERO Numerico 1742-6
</LOINC>
<LOINC COMPONENT="Albumina" PROPERTY="MCnc" TIMING="Pt" Method="Present" SYSTEM="Urine" SCALE="Ord" EL="1">
Albumina g/l SA-4220 MENARINI URINE Codificato 1754-1
</LOINC>
<LOINC COMPONENT="Albumina" PROPERTY="MCnc" TIMING="Pt" Method="Present" SYSTEM="Siero" SCALE="Qn" EL="1">
ALBUMINEMIA g/dl Elettroforetico SIERO Numerico 2862-1</LOINC>
<LOINC COMPONENT="Lattoalbumina alfa Ab.IgE" PROPERTY="ACnc" TIMING="Pt" Method="Absent" SYSTEM="Siero" SCALE="Qn" EL="1">
ALFA LATTOALBUMINA (F76) KU/L SIERO Numerico 7445-0
</LOINC>
<LOINC COMPONENT="Albumina/Globulina" Method="Present" SYSTEM="Siero" SCALE="Qn" EL="1">
Rapporto A/G Elettroforetico SIERO Numerico 44429-9
</LOINC>
<LOINC COMPONENT="Albumina/Proteina.totale" Method="Present" SYSTEM="Siero" SCALE="Qn" EL="1">
ALBUMINA % Elettroforetico SIERO Numerico 13980-8
</LOINC>
<LOINC COMPONENT="Azoto ureico*post dialisi" PROPERTY="MCnc" TIMING="Pt" Method="Present" SYSTEM="Siero" SCALE="Qn" EL="1">
AZOTEMIA POST-DIALISI mg/dl Enzimatica ureasi SIERO Numerico 11064-3
</LOINC>
<LOINC COMPONENT="Azoto ureico" PROPERTY="MRat" TIMING="24h" Method="Present" SYSTEM="Urine" SCALE="Qn" EL="1">

```

Figure 47 - La ressource XML

La ressource XML (Figure 47) montre comme à chaque examen contenu dans le catalogue local sont associés les correspondantes valeurs des attributs LOINC. Les informations concernant chaque examen sont contenues dans l'élément <LOINC> alors que les informations composant la description des examens trouvent leur correspondants valeurs des axes LOINC dans les attributs (Figure 48).

```
<LOINC COMPONENT="Glucosio*baseline" PROPERTY="MCnc" TIMING="Pt" Method="Present" SYSTEM="Siero" SCALE="Qn" EL="1">
Glicemia basale mg/dl GPO-PAP SIERO Numerico 1547-9
</LOINC>
<LOINC COMPONENT="Glucosio*2h post carico di glucosio" PROPERTY="MCnc" TIMING="Pt" Method="Present" SYSTEM="Siero" SCALE="Qn" EL="1">
Glicemia dopo 120' mg/dl GPO-PAP SIERO Numerico 20436-2
</LOINC>
```

Figure 48 - L'élément XML: LOINC

L'exemple de ressource montré a été obtenu par l'exportation en filtrant les examens avec un niveau d'exactitude minimum et donc avec l'attribut « EL=1 ». Le résultat XML obtenu garde tous les niveaux d'exactitude pour chaque examen, et va s'alimenter de façon automatique sur la base des opérations d'annotation effectuées. L'XML obtenu sera utilisé comme point de départ pour établir les règles linguistiques capables de reconnaître un ensemble plus large des formes linguistiques utilisées localement. La ressource créée agit comme base de connaissance pour la construction future d'un support automatique aux opérations de mappage ainsi que de validation traitées dans ce travail.

5.2 L'extraction d'informations

L'extraction d'informations à partir de données structurées ou non structurées est une opération qui se révèle difficile et complexe. L'information textuelle est un exemple de données non structurées pour laquelle l'extraction d'informations est conditionnée par différents facteurs, tel que le contexte linguistique et le domaine de référence.

Dans le cas objet d'étude, on traite des données structurées adaptées en forme textuelle semi-structurée (paragraphe 4.1). L'extraction d'information permet d'avoir une vision d'ensemble du corpus analysé plus détaillée et précise en analysant les catégories sémantiques identifiées et le niveau d'exactitude pour chaque examen. L'annotation sémantique, en effet, permet d'analyser les deux aspects en contribuant à l'amélioration et à la collecte des formes linguistiques à associer aux définitions LOINC et en supportant les experts pendant l'opération de validation. L'identification du seuil d'exactitude représente un aide important pour les contrôles des codes puisqu'il donne une indication importante sur les axes à contrôler. Actuellement, le RI ne dispose pas d'un système pour le contrôle automatique d'exactitude sur les codes qui est, en effet, effectué en considérant tous les six axes.

5.2.1 L'extraction d'informations : un système questions - réponses

Les formes linguistiques reconnues ont été associées à une catégorie sémantique déterminée en référence au domaine d'application ou à leur morphologie.

L'extraction de formes appartenant à une catégorie déterminée permet aux experts d'analyser le contenu de la base de données considérée et d'améliorer, en voyant plus clairement l'usage des formes, le système de reconnaissance

automatique. La vérification des formes utilisées et de la façon dont elles sont distribuées dans le corpus permet d'étendre les dictionnaires construits ou de découvrir éventuellement ultérieures règles linguistiques, en analysant la structure des données où elles se retrouvent.

On rappelle que le but du travail repose dans la nécessité de reconnaître certaines formes linguistiques pour supporter les opérations de validation qui, pour être effectuées, ont besoin d'une analyse des données profonde et d'une reconstruction détaillée du processus de mapping dérouillé par les laboratoires. L'extraction de certaines informations de façon automatique a porté à une considérable amélioration des temps et à une vision plus claire des données à analyser ainsi qu'à une importante réduction des contrôles manuels à effectuer, étant donné qu'un ensemble significatif d'attributs a été déjà contrôlé automatiquement.

Grâce à l'analyse des concordances, qui permet de vérifier le contexte dans lequel se retrouve une forme linguistique déterminée, l'extraction d'information a été possible même sur la structure dans laquelle se retrouve une forme linguistique spécifique ou si celle-ci est associée à un nombre limité d'attributs donc à une phrase avec beaucoup de manques.

Dans l'exemple montré dans la Figure 49, les concordances pour la catégorie « N+Analita » ont été analysées, ainsi que pour les catégories « N+Acronimi » (Figure 50) et « N+SIN » (Figure 51). L'analyse des concordances nous a permis de vérifier l'usage de certaines formes en relation aux codes LOINC (par exemple des examens définis par des formes différentes sont associés au même code) ; étant le texte composé par des phrases ayant la même structure (définition de l'analyse – unité de mesure - méthode - milieu biologique – échelle de mesure – code associé) on considère surtout le contexte droit si on analyse les définitions (les premiers éléments des phrases), gauche et droit si on veut analyser les autres attributs.

Text	Before	Seq.	After
URICO URINE Testo 14932-8	URICO URINE	ACIDO URICO	URINE Testo 14932-8 Uricemia mg/dl
ureasi URINE Numerico 3084-1	ureasi URINE	Uricemia	mg/dl Enzimatica ureasi URINE
ureasi URINE Numerico 3087-4	ureasi URINE	ACIDO VANILMANDELICO	mg/24h Enzimatica ureasi URINE
colorimetrico URINE Numerico 3122-9	colorimetrico URINE	GPT (ALT)	mg/24h Cromatografico - colorimetrico URINE
IFCC SIERO Numerico 1742-6	IFCC SIERO	Albumina	U1 Cinetico IFCC SIERO
MENARINI URINE Codificato 1754-1	MENARINI URINE	ALBUMINEMIA	g/l SA-4220 MENARINI URINE
Elettroforetico SIERO Numerico 2862-1	Elettroforetico SIERO	ALFA LATTOALBUMINA	g/dl Elettroforetico SIERO Numerico
Numerico 2862-1 ALFA LATTOALBUMINA (Numerico 2862-1 ALFA LATTOALBUMINA (F76	(F76) KUL SIERO Numerico
IL SIERO Numerico 7445-0	IL SIERO Numerico 7445-0	Rapporto A/G) KUL SIERO Numerico 7445-0 Rapporto
Elettroforetico SIERO Numerico 4429-9	Elettroforetico SIERO Numerico 4429-9	ALBUMINA %	Elettroforetico SIERO Numerico 4429-9 ALBUMINA % % Elettr
Elettroforetico SIERO Numerico 13980-8	Elettroforetico SIERO Numerico 13980-8	AZOTEMIA POST-DIALISI	% Elettroforetico SIERO Numerico 13980-8 AZOTEMIA POST
ureasi SIERO Numerico 11064-3	ureasi SIERO Numerico 11064-3	Azoturia	mg/dl Enzimatica ureasi SIERO
ureasi URINE Numerico 3096-5	ureasi URINE Numerico 3096-5	BILIRUBINA TOTALE	g/24h Enzimatica ureasi URINE
Colorimetrico SIERO Numerico 1975-2	Colorimetrico SIERO Numerico 1975-2	CALCIO	mg/dl Colorimetrico SIERO Numerico
Cresolfalina SIERO Numerico 17861-6	Cresolfalina SIERO Numerico 17861-6	CALCIO	mg/dl Cresolfalina SIERO Numerico
CALCIO URINE Testo 14636-5	CALCIO URINE Testo 14636-5	CALCIO POST-DIALISI	URINE Testo 14636-5 CALCIO POST-DIALISI
Cresolfalina SIERO Numerico 51950-4	Cresolfalina SIERO Numerico 51950-4	CALCIURIA	mg% Cresolfalina SIERO Numerico 51950-4 CALCIURIA
Cresolfalina URINE Numerico 6874-2	Cresolfalina URINE Numerico 6874-2	COLESTEROLO HDL	mg/24h Cresolfalina URINE Numerico
diretto SIERO Numerico 2085-9	diretto SIERO Numerico 2085-9	COLESTEROLO TOTALE	mg/dl Enzimatico diretto SIERO
-PAP SIERO Numerico 2093-3	-PAP SIERO Numerico 2093-3	COLESTEROLO VLDL	mg/dl GPO-PAP SIERO
diretto SIERO Numerico 2091-7	diretto SIERO Numerico 2091-7	CREATININA	mg/dl Enzimatico diretto SIERO
Jaffe' SIERO Numerico 2160-0	Jaffe' SIERO Numerico 2160-0	CREATININA POST DIALISI	mg/dl Cinetico Jaffe' SIERO
Jaffe' SIERO Numerico 11041-1	Jaffe' SIERO Numerico 11041-1	CREATINURIA	mg/dl Cinetico Jaffe' SIERO
Jaffe' URINE Numerico 2162-6	Jaffe' URINE Numerico 2162-6	CREATINURIA	DELLE 24 ORE g/24 h Cinetico
Jaffe' URINE Numerico 2162-6	Jaffe' URINE Numerico 2162-6	He	DELLE 24H g/24h Cinetico
			mcg/cre creat Assorbimento Atomico

Figure 49 - Les concordances des analytes

Text	Before	Seq.	After
colorimetrico URINE Numerico 3122-9	colorimetrico URINE Numerico 3122-9	GPT (ALT)	U1 Cinetico IFCC SIERO
Numerico 2862-1 ALFA LATTOALBUMINA (Numerico 2862-1 ALFA LATTOALBUMINA (F76) KUL SIERO Numerico 7445-0 Rapporto
Jaffe' URINE Numerico 2162-6	Jaffe' URINE Numerico 2162-6	Hg	mcg/g creat Assorbimento Atomico
URINE Numerico 13774-5 DOSAGGIO	URINE Numerico 13774-5 DOSAGGIO	Hb A2	% PLASMA Numerico 4551-8 Emoglobina mg/dl
URINE Codificato 50559-4 EMOGLOBINA (URINE Codificato 50559-4 EMOGLOBINA (HGB) g/dl ADVIA 120 BAYER EMATOLOGIA
Numerico 718-7 EMOGLOBINA GLICOSILATA -	Numerico 718-7 EMOGLOBINA GLICOSILATA -	HbA1c	% EMOGLOBINA Numerico 4548-4 HBA1 % Elettroforetico SANGUE
1c % EMOGLOBINA Numerico 4548-4	1c % EMOGLOBINA Numerico 4548-4	HBA1	% Elettroforetico SANGUE Numerico 51196-4 HBA2 % Elettroforetico
Elettroforetico SANGUE Numerico 51196-4	Elettroforetico SANGUE Numerico 51196-4	HBA2 %	Elettroforetico SANGUE Numerico 4552-6 Solfoemoglobina % Hb
SANGUE Numerico 4552-6 Solfoemoglobina %	SANGUE Numerico 4552-6 Solfoemoglobina %	Hb	totale SIERO Numerico 4685-4 Noradrenalina mcg
ml SIERO Numerico 2484-4	ml SIERO Numerico 2484-4	FSH	URINARIO U1/24h URINE Numerico
-PAP SIERO Numerico 20440-4	-PAP SIERO Numerico 20440-4	G6PDH	% G6PDH Numerico 49502-8 GOT (AST
Numerico 20440-4 G6PDH %	Numerico 20440-4 G6PDH %	G6PDH	Numerico 49502-8 GOT (AST) U1
G6PDH Numerico 49502-8	G6PDH Numerico 49502-8	GOT (AST)	U1 Cinetico IFCC SIERO
ml SIERO Numerico 47862-8	ml SIERO Numerico 47862-8	LDH 1	% SIERO Numerico 2536-1 LDH 2 % SIERO Numerico
LDH 1 % SIERO Numerico 2536-1	LDH 1 % SIERO Numerico 2536-1	LDH 2	% SIERO Numerico 2539-5 LDH 3 % SIERO Numerico
LDH 2 % SIERO Numerico 2539-5	LDH 2 % SIERO Numerico 2539-5	LDH 3	% SIERO Numerico 2542-9 LDH 4 % SIERO Numerico
LDH 3 % SIERO Numerico 2542-9	LDH 3 % SIERO Numerico 2542-9	LDH 4	% SIERO Numerico 2545-2 LDH 5 % SIERO Numerico
LDH 4 % SIERO Numerico 2545-2	LDH 4 % SIERO Numerico 2545-2	LDH 5	% SIERO Numerico 2548-6 LH mIU/ml
LDH 5 % SIERO Numerico 2548-6	LDH 5 % SIERO Numerico 2548-6	LH	mIU/ml Immunoenzimatico ELA SIERO
EIA SIERO Numerico 10501-5	EIA SIERO Numerico 10501-5	LH	URINARIO U1/24h URINE Numerico
Fotometria a fiamma	Fotometria a fiamma	NA	-K URINARI Numerico 2829-0 17 OH PROGESTERONE
a fiamma NA -	a fiamma NA -	K	URINARI Numerico 2829-0 17 OH PROGESTERONE ng
-K URINARI Numerico 2829-0	-K URINARI Numerico 2829-0	17 OH	PROGESTERONE ng/ml Radioimmunologico RIA
RIA SIERO Numerico 1469-3	RIA SIERO Numerico 1469-3	17 OH	PROGESTERONE Numerico 2829-0 SIERO

Figure 50 - Les concordances des acronymes

Text	Before	Seq.	After
BAYER EMATOLOGIA Numerico 718-7	BAYER EMATOLOGIA Numerico 718-7	EMOGLOBINA GLICOSILATA	- HbA1c % EMOGLOBINA Numerico 4548-4 HBA
totale SIERO Numerico 4685-4	totale SIERO Numerico 4685-4	Noradrenalina	mcg/24h Cromatografico - colorimetrico URINE Numerico
colorimetrico URINE Numerico 2668-2	colorimetrico URINE Numerico 2668-2	Adrenalina	mcg/24h Cromatografico - colorimetrico URINE
ml SIERO Numerico 2258-2	ml SIERO Numerico 2258-2	SOMATOMEDINA C	ng/ml SIERO Numerico 2484-4 FSH
-PAP SIERO Numerico 2571-8	-PAP SIERO Numerico 2571-8	VITAMINA D3	(25 OH) ng/ml SIERO Numerico
BAYER EMATOLOGIA Numerico 713-8	BAYER EMATOLOGIA Numerico 713-8	GLOBULI ROSSI	(RBC) 10 ⁶ /µl ADVIA 120 BAYER EMATOLOGIA
BAYER EMATOLOGIA Numerico 788-0	BAYER EMATOLOGIA Numerico 788-0	Emazie	COLTURE Codificato 48051-7 1 ora VES Numerico
BAYER EMATOLOGIA Numerico 17788-1	BAYER EMATOLOGIA Numerico 17788-1	GLOBULI BIANCHI	(WBC) 10 ³ /µl ADVIA 120 BAYER EMATOLOGIA
-LATTOALBUMINA U1/ml 7445-0	-LATTOALBUMINA U1/ml 7445-0	ADRENALINA	URINARIA mcg/24h HPLC Urina
Urina delle 24 ore 2232-7	Urina delle 24 ore 2232-7	NORADRENALINA	URINARIA mcg/24h HPLC Urina
HPLC Urina delle 24 2668-2	HPLC Urina delle 24 2668-2	ADRENALINA	pg/ml HPLC Sanguine intero
Sanguine intero Numerico 2230-1	Sanguine intero Numerico 2230-1	NORADRENALINA	pg/ml HPLC Sanguine intero
FLUSSO SANGUE INTERO 713-8	FLUSSO SANGUE INTERO 713-8	GLOBULI ROSSI	10 ⁶ /ml CITOMETRIA A FLUSSO SANGUE
Sanguine NUM 7556-4 F323	Sanguine NUM 7556-4 F323	OVOTRANSFERRINA	kUA1 Sanguine NUM 25615-6 F
chr pg NUM 71694-4	chr pg NUM 71694-4	Emoglobina glicata	NUM 41995-2 Emoglobina g/dl NUM
/dl Sanguine NUM 2258-2	/dl Sanguine NUM 2258-2	SOM-C	ng/ml Sanguine NUM 2484-4 FSH
S. mm Siero 30341-2	S. mm Siero 30341-2	Globuli Rossi	26453-1 MCV 30428-7 MCH 28539-5 MCHC 28540-3 Eosinofili % 26450-7 FOSFATASI
UL Siero 14805-6	UL Siero 14805-6	Globuli Bianchi	26464-8 Linfociti % 731-0 Linfociti % 26478-8 Monociti % 742-7 Monociti % 26485-3 Neutrofili

Figure 51 - Les concordances des synonymes

Text	Before	Seq	After
MENARINI URINE Codificato 1754-1	ALBUMINEMIA		g/dl Elettroforetico SIERO Numerico
Elettroforetico SIERO Numerico 13980-8	AZOTEMIA POST-DIALISI		mg/dl Enzimatica ureasi SIERO
ureasi SIERO Numerico 11064-3	Azoturia		g/24h Enzimatica ureasi URINE
Cresolitalaina SIERO Numerico 51950-4	CALCIURIA		mg/24h Cresolitalaina URINE Numerico
Jaffe' SIERO Numerico 11041-1	CREATTINURIA		DELLE 24 ORE g/24 h Cinetico
Jaffe' URINE Numerico 2162-6	CREATTINURIA		DELLE 24H g/24h Cinetico
Molibdato URINE Numerico 2779-7	GLICEMIA		mg/dl GPO-PAP SIERO
-PAP SIERO Numerico 2345-7	Glicemia		basale mg/dl GPO-PAP
-PAP SIERO Numerico 1547-9	Glicemia		dopo 120' mg/dl GPO-PAP
-PAP SIERO Numerico 20436-2	Glicemia		dopo 180' mg/dl GPO-PAP
-PAP SIERO Numerico 20437-0	Glicemia		dopo 30' mg/dl GPO-PAP
-PAP SIERO Numerico 20439-6	Glicemia		dopo 60' mg/dl GPO-PAP
-PAP SIERO Numerico 20438-8	Glicemia		dopo 90' mg/dl GPO-PAP
/dl SIERO Numerico 2885-2	PROTEINURIA		DELLE 24 h mg/24h URINE
A/G % NUM 44429-9	Azotemia		mg/dl Sangue NUM 3094-0 Azoto
/dl Urine NUM 2778-9	Glicemia		mg/dl Sangue NUM 2345-7 G
g/dl, Sangue 786-4	Glicemia		mg/dl Siero Plasma 2345-7 AST
Siero 1742-6 Basofili 704-7 Basofili % 30180-4	BILIRUBINEMIA		TOTALE mg/dl Siero 1975-2 BILIRUBINEMIA
mg/dl Siero 1975-2	BILIRUBINEMIA		DIRETTA mg/dl Siero 1968-7 COLESTEROLO

Figure 52 - Les concordances des morphèmes

On a retrouvé donc 371 formes définies comme analytes (catégorie générique) ; en plus 110 sont les acronymes utilisés et 18 les synonymes, 19 les formes en « emia » et « uria » dans la Figure 52 et 191 ont été formalisées comme « core » de l'analyse. Ces informations sont très importantes pour avoir une vision plus détaillée des données qui, au contraire, ne sont pas formalisées sous ces aspects et donc difficiles à extraire automatiquement.

À cet égard, étant chaque forme associée à une catégorie définie, il est possible aussi d'extraire des informations plus particulières qui répondent à des requêtes plus précises et détaillées. Le système réalisée, en effet, peut être utilisé aussi comme base pour la construction d'un système « questions – réponses », suite une transformation en langage naturelle du langage d'interrogation pour effectuer une requête en NooJ. On pense par exemple, aux requêtes suivantes :

- 1) « Quels sont les examens décrits seulement par un acronyme ? »
- 2) « Combien de synonymes se retrouvent dans une phrase en forme complète ? »
- 3) « Combien de fois y a-t-il un adjectif après l'analyte ? »

La requête 1) effectuée dans NooJ sera le résultat de la forme <N+Acronimo><NUM+Code> et donnera le contexte pour toutes les couples d'acronymes et des codes présents dans le corpus, c'est-à-dire tous les examens décrits par le seul acronyme et associés au code LOINC (Figure 53).

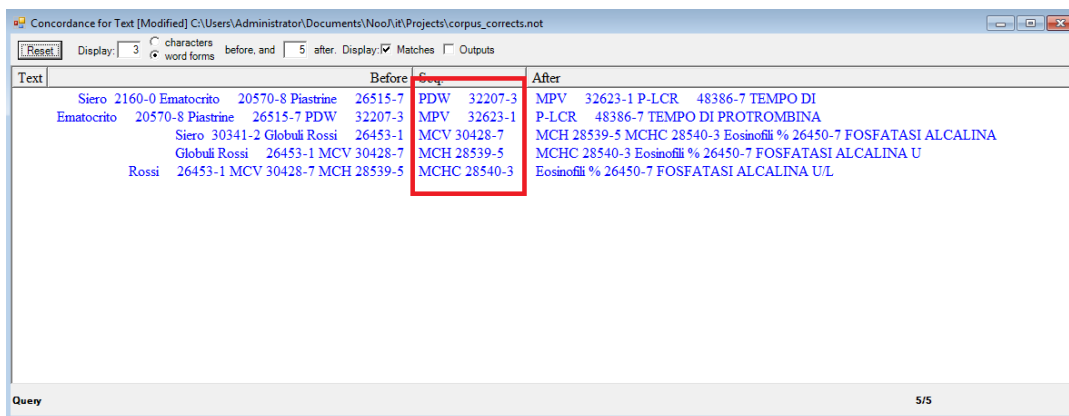


Figure 53 - Les examens décrits seulement par le composant

La requête 2) est les résultats des expressions suivantes:

- <N+Analita+SIN><N+UnMis><N+Metodo><N+Campione><N+Risultato><N+UM+Code>
- <N+Analita+SIN><N+UnMis><N+Campione><N+Risultato><NUM+Code>

La Figure 54 montre en bleu les résultats obtenus en considérant la présence de la méthode, en rouge ceux qui ne l'explicitent pas.

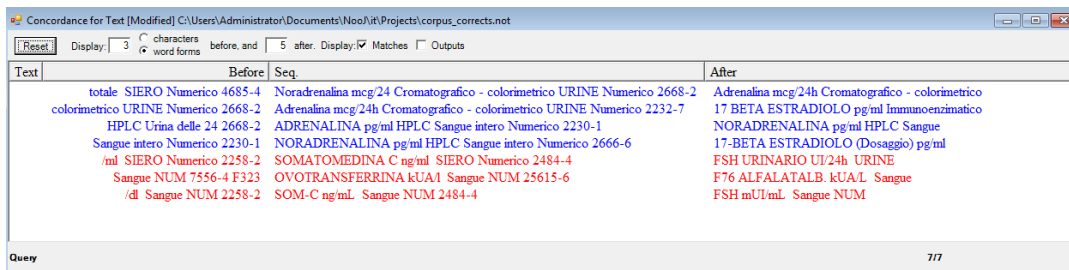


Figure 54 - Les concordances pour la méthode

De la même manière, la requête présentée dans la forme :

- $\langle N+Analita \rangle \langle A+AgLoinc \rangle$

identifie tous les composants accompagnés par un adjectif qui identifie le milieu biologique à associer (Figure 55).

Text	Before	Seq.	After
Atomico URINE Numerico 13465-0	Manganese Urinario		mcg/gr creat Assorbimento Atomico
Colorimetrico URINE Numerico 21562-4	Ac metilippurico Urinario		mg/gr creat Cromatografia liquida
/ml SIERO Numerico 2484-4	FSH URINARIO		UI/24h URINE Numerico 15068-0 FOSF
Molibdato SIERO Numerico 2777-1	FOSFORO URINARIO		mg/24h Molibdato URINE Numerico
EIA SIERO Numerico 10501-5	LH URINARIO		UI/24h URINE Numerico 15080-5 POTASSIO
fiamma SIERO Numerico 29349-8	POTASSIO URINARIO		mEq/24h Fotometria a fiamma
fiamma SIERO Numerico 2951-2	SODIO URINARIO		mEq/24h Fotometria a fiamma
Ratio COAGULAZIONE Numerico 13590-5	ACIDO IPPURICO URINARIO		Mg/l HPLC Urina di
fine turno Numerico 6709-0	ACIDO METILIPPURICO URINARIO		Mg/l HPLC Urina di
delle 24 ore Numerico 3122-9	ALBUMINA URINARIA		mg/dl Colorimetrico con rilevazione
-LATTOALBUMINA UI/ml 7445-0	ADRENALINA URINARIA		mcg/24h HPLC Urina delle
Urina delle 24 ore 2232-7	NORADRENALINA URINARIA		mcg/24h HPLC Urina delle
/dl Sangue NUM 3084-1	Acido urico Urinario		mg/dl Urine NUM 3086-6 GPT
/dl Sangue NUM 3094-0	Azoto Urinario		mg/dl NUM 3095-7 Baso% % NUM
/dl Sangue NUM 38483-4	Creatinina Urinaria		mg/dl NUM 2161-8 Ematocrito % NUM
/dl Sangue NUM 2777-1	Fosforo Urinario		mg/dl Urine NUM 2778-9 Glicemia
/l Sangue NUM 2823-3	Potassio Urinario		mg/dl Urine NUM 2828-2 17 OH
URINE FRESCHE NUM 2955-3	CALCIO IONIZZATO		Mmol/L Sangue NUM 34581-9 EMATOCRITO
/L Sangue NUM 34581-9	EMATOCRITO ARTERIOSO		Sangue NUM 32354-3 COHb % Sangue NUM

Figure 55 - Les concordance des adjectifs

Les cas montrés sont pris comme exemple pour exposer les capacités du système de donner de façon automatique toute une série d'informations utiles et liées aux connaissances qui, au contraire, ne seraient pas formalisées en aucune manière. Il est clair donc que les requêtes montrées ne sont qu'un sous ensemble de celles possibles, et que suite à une formalisation plus « user-friendly » du langage utilisé elles pourraient représenter un point de départ pour la construction d'un système questions réponses qu' à partir d'informations techniques médicales conservées dans des structures des données déterminées puisse donner des informations utiles pendant l'arrangement des procédures spécifiques et très délicates.

5.2.2 L'extraction d'informations : les niveaux d'exactitude

Dans ce paragraphe on focalise l'attention sur le niveau d'exactitude que l'annotation sémantique nous a permis d'identifier.

Pour chaque examen on retrouve une étiquette qui représente le nombre des éléments qui résultent appropriés suite aux contrôles spécifiés dans les contraintes formalisées. Une fois que l'annotation sémantique associe à chaque examen son niveau (en forme incrémentale), ce niveau d'appartenance est enregistré par la propriété « +EL » en devenant « attribut » de l'élément « LOINC » qui décrit chaque examen dans la ressource XML obtenue. L'extraction d'information, donc, peut être effectuée par la requête simple « LOINC+EL=1 » si on veut tous les examens associés au niveau 1, même chose mais avec une valeur différente pour la propriété « +EL » pour les autres niveaux.

Le fait que les niveaux d'exactitude soient associés de façon incrémentale signifie que si un examen appartient au niveau 3 automatiquement cela veut dire qu'il appartient aussi au niveau 2 et au niveau 1. L'annotation sémantique pour sa nature associe aux mots, qu'elle retrouve et reconnaît, une certaine étiquette morphologique ou sémantique en associant une forme à telle ou telle catégorie.

Sur la base des règles décrites dans notre système l'annotation sémantique, pour comme géré dans la grammaire « LOINC.nog », associe à chaque examen son niveau d'appartenance en laissant tous les choix possibles qui respectent les contraintes. L'identification des quantités est nécessaire pour l'analyse des résultats obtenus par l'application du système des règles linguistiques présenté. Pour surmonter les limitations dues à la construction des niveaux de façon incrémentale (les codes appartenant au niveau 3 appartiennent aussi au niveau 2 et au niveau 1 par conséquence), des grammaires pour l'extraction d'informations spécifiques ont été construites.

Comme déjà explique ci-dessus, l'extraction des examens appartenant à un niveau spécifique peut être effectuée de la même façon des requêtes montrées dans le paragraphe 4.5, en particulier, « LOINC+EL=1 » permet d'extraire les 346

examens appartenant au niveau 1, le niveau minimum pour considérer que l'annotation sémantique a été effectuée (Figure 56).

Before	Seq	After
URICO URINE Testo 14932-8	ACIDO URICO URINE Testo 14932-8	Uricemia mg/dl Enzimatica ureasi
ureasi URINE Numerico 3084-1	Uricemia mg/dl Enzimatica ureasi URINE Numerico 3084-1	Uricuria mg/24h Enzimatica ureasi
ureasi URINE Numerico 3087-4	Uricuria mg/24h Enzimatica ureasi URINE Numerico 3087-4	ACIDO VANILMANDELICO mg/24h Crona
colorimetrico URINE Numerico 3122-9	ACIDO VANILMANDELICO mg/24h Cromatografico - colorimetrico URINE Numerico 3122-9	GPT (ALT) UI Cinetico
IFCC SIERO Numerico 1742-6	GPT (ALT) UI Cinetico IFCC SIERO Numerico 1742-6	Albumina g/l SA-4220 MENARINI
MENARINI URINE Codificato 1754-1	Albumina g/l SA-4220 MENARINI URINE Codificato 1754-1	ALBUMINEMIA g/dl Elettroforetico SIERO
Elettroforetico SIERO Numerico 2862-1	ALBUMINEMIA g/dl Elettroforetico SIERO Numerico 2862-1	ALFA LATTOALBUMINA (F76) KU/L
IL SIERO Numerico 7445-0	ALFA LATTOALBUMINA (F76) KU/L SIERO Numerico 7445-0	Rapporto A/G Elettroforetico SIERO
Elettroforetico SIERO Numerico 44429-9	Rapporto A/G Elettroforetico SIERO Numerico 44429-9	ALBUMINA % Elettroforetico SIERO Num
Elettroforetico SIERO Numerico 13980-8	ALBUMINA % Elettroforetico SIERO Numerico 13980-8	AZOTEMIA POST-DIALISI mg/dl
ureasi URINE Numerico 3096-5	AZOTEMIA POST-DIALISI mg/dl Enzimatica ureasi SIERO Numerico 11064-3	Azoturia g/24h Enzimatica ureasi
Colorimetrico SIERO Numerico 1975-2	Azoturia g/24h Enzimatica ureasi URINE Numerico 3096-5	BILIRUBINA TOTALE mg/dl Colorimetrico
Cresolftalaina SIERO Numerico 17861-6	BILIRUBINA TOTALE mg/dl Colorimetrico SIERO Numerico 1975-2	CALCIO mg/dl Cresolftalaina SIERO
CALCIO URINE Testo 14636-5	CALCIO mg/dl Cresolftalaina SIERO Numerico 17861-6	CALCIO URINE Testo 14636-5
Cresolftalaina SIERO Numerico 51950-4	CALCIO URINE Testo 14636-5	CALCIO POST-DIALISI mg% Cresolftalaina SIERO Numerico 51950-4
Cresolftalaina SIERO Numerico 6874-2	CALCIO POST-DIALISI mg% Cresolftalaina SIERO Numerico 51950-4	CALCIURIA mg/24h Cresolftalaina URINE Numerico 6874-2
Cresolftalaina URINE Numerico 6874-2	CALCIURIA mg/24h Cresolftalaina URINE Numerico 6874-2	COLESTEROLO HDL mg/dl Enzimatico diretto SIERO Numerico 2085-9
diretto SIERO Numerico 2085-9	COLESTEROLO HDL mg/dl Enzimatico diretto SIERO Numerico 2085-9	COLESTEROLO TOTALE mg/dl GPO-PAP SIERO Numerico 2093-3
-PAP SIERO Numerico 2093-3	COLESTEROLO TOTALE mg/dl GPO-PAP SIERO Numerico 2093-3	COLESTEROLO VLDL mg/dl Enzimatico
	COLESTEROLO VLDL mg/dl Enzimatico diretto SIERO Numerico 2091-7	CREATININA mg/dl Cinetico Jaffe

Figure 56 - Les examens appartenant au niveau 1 d'exactitude

On rappelle que les examens qui ont l'étiquette « EL=1 » sont 346, c'est-à-dire que l'annotation du composant a été effectuée correctement puisque elle a permis d'obtenir des résultats malgré sa variété linguistique.

Toutefois, ceci n'exclut pas la présence des autres étiquettes et donc l'appartenance aussi aux niveaux les plus élevés. Si on veut exactement le nombre d'examens qui s'arrêtent au niveau 1, sans passer aux niveaux les plus élevés, il faut effectuer la requête en appliquant la grammaire montrée dans la Figure 57.

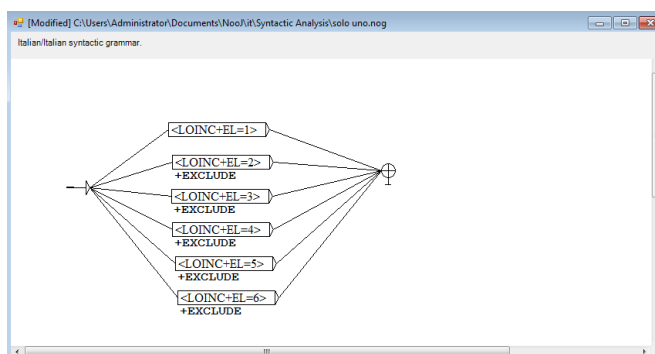


Figure 57 - La grammaire de sélection des examens qui s'arrêtent au niveau 1

Avec l'opérateur « +EXCLUDE », la grammaire NooJ nous permet de sélectionner seulement le niveau d'exactitude qu'on veut analyser en excluant les autres. Des 346 appartenant au niveau 1, 146 s'arrêtent au niveau 1 sans passer aux niveaux les plus élevés (Figure 58).

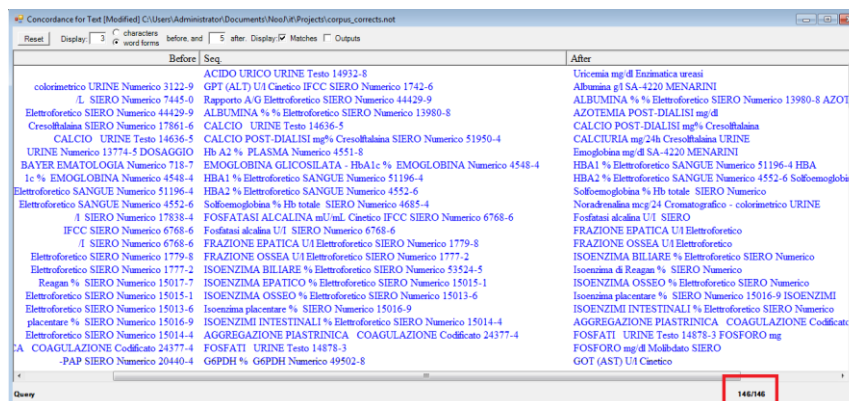


Figure 58 - Les examens qui s'arrêtent au niveau 1 d'exactitude

Ce genre de grammaires a été construit pour la sélection univoque des examens appartenant aux autres niveaux ainsi que pour les niveaux déterminés par les contraintes « HELP ». Par exemple, pour extraire les examens déterminés seulement par les contraintes « HELP » pour passer au niveau 3, la grammaire dans la Figure 59 a été appliquée pour la sélection des informations.

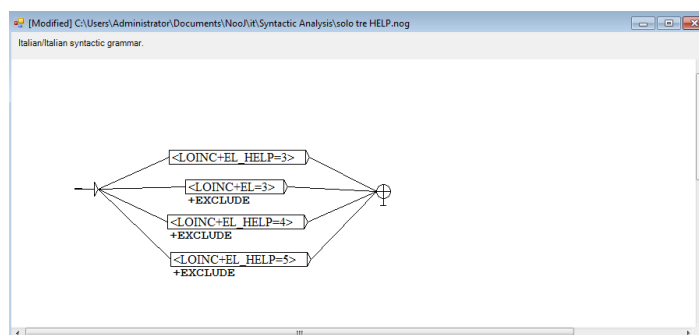


Figure 59 - La grammaire de sélection des examens appartenant au niveau 3 HELP

Les résultats obtenus nous soulignent que les examens qui à partir du niveau 2 passent au niveau 3 par le contraintes « HELP » sont 17 ; pour ces tests, la valeur

de l'attribut « milieu biologique » est déterminée par le composant sous forme de morphème (avec les suffixes -emia et -uria) ou par l'adjectif qui l'accompagne.

Text	Before	Seq.	After
	URICO URINE Testo 14932-8	URICEMIA mg/dl Enzimatica ureasi URINE Numerico 3084-1	Uricuria mg/24h Enzimatica ureasi
	MENARINI URINE Codificato 1754-1	ALBUMINEMIA g/dl Elettroforetico SIERO Numerico 2862-1	ALFA LATTOALBUMINA (F76) KUL
	Elettroforetico SIERO Numerico 13980-8	AZOTEMIA POST-DIALISI mg/dl Enzimatica ureasi SIERO Numerico 11064-3	Azoturia g/24h Enzimatica ureasi
	Atomico URINE Numerico 13465-0	Manganeso Urinario mg/g creat Assorbimento Atomico Numerico 29935-4	Tiocianati Urinari mg/g creat
	Molbdato URINE Numerico 2779-7	GLICEMIA mg/dl GPO-PAP SIERO Numerico 2345-7	Glicemia basale mg/dl GPO
	-PAP SIERO Numerico 2345-7	Glicemia basale mg/dl GPO-PAP SIERO Numerico 1547-9	Glicemia dopo 120' mg/dl GPO
	-PAP SIERO Numerico 1547-9	Glicemia dopo 120' mg/dl GPO-PAP SIERO Numerico 20436-2	Glicemia dopo 180' mg/dl GPO
	-PAP SIERO Numerico 20437-0	Glicemia dopo 30' mg/dl GPO-PAP SIERO Numerico 20439-6	Glicemia dopo 60' mg/dl GPO
	-PAP SIERO Numerico 20439-6	Glicemia dopo 60' mg/dl GPO-PAP SIERO Numerico 20438-8	Glicemia dopo 90' mg/dl GPO
	-PAP SIERO Numerico 20438-8	Glicemia dopo 90' mg/dl GPO-PAP SIERO Numerico 20440-4	G6PDH % G6PDH Numerico
	BAYER EMATOLOGIA Numerico 4544-3	APTOGLOBINA g/l Nefelometria SIERO Numerico 46127-7	C1 INIBITORE mg/dl Immunodiffusione
	A/G % NUM 44429-9	Azotemia mg/dl Sangue NUM 3094-0	Azoto Urinario mg/dl NUM
	/dl Sangue NUM 3094-0	Azoto Urinario mg/dl NUM 3095-7	Baso% % NUM 704-7 Bilirubina Diretta mg
	/dl Sangue NUM 38483-4	Creatinina Urinaria mg/dl NUM 2161-8	Ematocrito % NUM 4544-3 Aptoglobina g/l
	NUM 2161-8 Ematocrito % NUM 4544-3	Aptoglobina g/l Sangue NUM 4542-7	C1 Inibitore mg/dl Sangue
	/dl Urine NUM 2778-9	Glicemia mg/dl Sangue NUM 2345-7	G6PDH mU/ml Sangue
	Siero 1742-6 Basofili 704-7 Basofili % 30180-4	BILIRUBINEMIA TOTALE mg/dl Siero 1975-2	BILIRUBINEMIA DIRETTA mg/dl Siero

Figure 60 - Le niveau d'exactitude 3 help

Ce résultat est très important puisque on récupère 17 tests laissés au niveau 2 pour les passer à leur effectif niveau d'appartenance plus élevé (Figure 60). Les grammaires de récupérations des informations ont été réalisées aussi pour les autres contraintes « HELP ». Les grammaires pour sélectionner les examens appartenant à chaque niveau d'exactitude déterminé par les contraintes standard et/ou « Help » sont montrés en Annexe 3 – Les grammaires d'extraction d'informations.

L'extraction de l'information concernant certains niveaux d'exactitude peut répondre à des questions spécifiques référées à un analyte déterminé. Par exemple si on veut savoir combien d'examens ayant comme composant l'analyte « Glucosio » (« Glucose » en français) sont associés au niveau 2 et récupérés par les contraintes « HELP » pour passer au niveau 3, il est possible d'effectuer les requêtes suivantes:

- <LOINC+EL=2>

- <LOINC+COMPONENT=Glucosio+EL_HELP=3>

Les résultats sont montrés dans la Figure 61: en bleu les examens appartenant au niveau 2, en rouge ceux qui ont été évalués appropriés pour passer au niveau 3 par les contraintes « HELP ».

Text	Before	Seq	After
Molbdata URINE Numerico 2779-7	GLICEMIA mg/dl GPO-PAP SIERO Numerico 2345-7		Glicemia basale mg/dl GPO
-PAP SIERO Numerico 2345-7	Glicemia basale mg/dl GPO-PAP SIERO Numerico 1547-9		Glicemia dopo 120' mg/dl GPO
-PAP SIERO Numerico 1547-9	Glicemia dopo 120' mg/dl GPO-PAP SIERO Numerico 20436-2		Glicemia dopo 180' mg/dl GPO
-PAP SIERO Numerico 20436-2	Glicemia dopo 180' mg/dl GPO-PAP SIERO Numerico 20437-0		Glicemia dopo 30' mg/dl GPO
-PAP SIERO Numerico 20437-0	Glicemia dopo 30' mg/dl GPO-PAP SIERO Numerico 20439-6		Glicemia dopo 60' mg/dl GPO
-PAP SIERO Numerico 20439-6	Glicemia dopo 60' mg/dl GPO-PAP SIERO Numerico 20438-8		Glicemia dopo 90' mg/dl GPO
-PAP SIERO Numerico 20438-8	Glicemia dopo 90' mg/dl GPO-PAP SIERO Numerico 20440-4		G6PDH % G6PDH Numerico
/d/ Urine NUM 2778-9	Glicemia mg/dl Sangue NUM 2345-7		G6PDH mU/ml Sangue
g/dL Sangue 786-4	Glicemia mg/dl Siero Plasma 2345-7		AST/GOT U/L Siero
URINE 24h NUM 21305-8	GLUCOSIO mg/dl Ematico URINE FRESC NUM 2350-7		POTASSIO mmol/L Potenzimetrico Urine
Sangue % Sangue NUM 19225-2	GLUCOSIO GAS mmol/L Sangue NUM 39481-7		GLUCOSIO GAS Mmol/L Sangue
/L Sangue NUM 39481-7	GLUCOSIO GAS Mmol/L Sangue NUM 39480-9		POTASSIO EMOGAS Mmol/L Sangue
Molbdata URINE Numerico 2779-7	GLICEMIA mg/dl GPO-PAP SIERO Numerico 2345-7		Glicemia basale mg/dl GPO
-PAP SIERO Numerico 2345-7	Glicemia basale mg/dl GPO-PAP SIERO Numerico 1547-9		Glicemia dopo 120' mg/dl GPO
-PAP SIERO Numerico 1547-9	Glicemia dopo 120' mg/dl GPO-PAP SIERO Numerico 20436-2		Glicemia dopo 180' mg/dl GPO
-PAP SIERO Numerico 20437-0	Glicemia dopo 30' mg/dl GPO-PAP SIERO Numerico 20439-6		Glicemia dopo 60' mg/dl GPO
-PAP SIERO Numerico 20439-6	Glicemia dopo 60' mg/dl GPO-PAP SIERO Numerico 20438-8		Glicemia dopo 90' mg/dl GPO
-PAP SIERO Numerico 20438-8	Glicemia dopo 90' mg/dl GPO-PAP SIERO Numerico 20440-4		G6PDH % G6PDH Numerico
/d/ Urine NUM 2778-9	Glicemia mg/dl Sangue NUM 2345-7		G6PDH mU/ml Sangue
g/dL Sangue 786-4	Glicemia mg/dl Siero Plasma 2345-7		AST/GOT U/L Siero

Figure 61 - Examens du niveau 2 récupérés par la contrainte HELP

5.3 Résumé des résultats

La formalisation de la connaissance par la réalisation des dictionnaires électroniques décrit en détail les formes linguistiques et permet, comme on a vu dans les sections précédentes, l'extraction des informations concernant les catégories reconnues ainsi qu'obtenues en réponse à des questions déterminées.

Dans le Tableau 12 montré ci-dessus, on retrouve les quantités pour chaque catégorie retrouvée dans le corpus analysé, avec leurs totaux agrégés.

Catégorie	Totale		
Analytes	371		
Analytes (Core Component)	191		
Morphemes « emia »	14	Morphèmes 19	
Morphemes « uria »	5		
Synonymes	10	Synonymes 18	
Synonymes Hemo	8		
Hémogramme (Globules Blancs)	10	Analytes Hémogramme 62	
Hémogramme (Globules Rouges)	44		
Hémogramme (Plaquettes)	8		
Acronymes Hemo	39	Acronymes 110	
Acronymes	71		
Acides	16		
			Hémogramme (Analytes- Synonymes- Acronymes) 101

Tableau 12 - Synthèse résultats: catégories identifiées

Les analytes définis « core component » (paragraphe 4.2.1), donc qui représentent l'élément testé, sont 191 et peuvent comprendre aussi les formes linguistiques (morphèmes, synonymes, acronymes, etc.) associées aux catégories de ce genre.

Les morphèmes avec les suffixes « emia » et « uria » sont seulement 19 tandis que les synonymes utilisés sont 18. Dans cette analyse, nous avons distingué les catégories retrouvées pour les éléments testés dans l'hémogramme. Cette distinction a été faite seulement pour mieux détailler les examens étant l'hémogramme un examen difficile à valider à cause de sa structure composée par différents niveaux d'analyse (examens concernant les globules blancs, les globules rouges ou les plaquettes).

Les acronymes, par exemple, ont été divisés en acronymes et acronymes des examens contenus dans l'hémogramme. Dans le total agrégé, on retrouve pour l'hémogramme 101 formes linguistiques dont 62 sont les analytes divisées en globules blancs (10), globules rouges (44), plaquettes (8), synonymes (8) et acronymes (39).

La vision d'ensemble des caractéristiques des examens contenus dans le corpus donne la possibilité aux experts de réduire les temps d'analyse et d'arriver plus préparés pour la validation des associations des codes. C'est-à-dire que connaître les catégories d'appartenance sémantiques des examens et comment elles vont se distribuer dans le corpus permet de savoir à priori la typologie d'examens que nous allons examiner. Ce genre d'analyses effectué sur le corpus permet d'associer des examens ayant des définitions identiques et de récupérer du temps pendant les opérations de validation.

En tout cas les résultats les plus importants concernent les niveaux d'exactitude identifiés pour chaque test analysé. En particulier, dans le Tableau 13 est montrée la quantité d'examens appartenant à chaque niveau d'exactitude en référence aussi aux niveaux déterminés par les contraintes « HELP ».

<i>Niveau</i>	<i>Totale</i>	<i>Reste au niveau</i>	<i>Passent au niveau plus élevé</i>	<i>Dimension du corpus</i>
1	346	146		375
2	200	101	17 (niveau 3 help)	
3	82	5	1 (niveau 4 help)	
4	76		23 (niveau 5 help)	
5	53	27		
6	26			

Tableau 13 - Synthèse résultats: les niveaux d'exactitude

Le niveau 1 est composé par 346 examens dont 146 s'arrêtent au niveau 1 et les restants passent aux niveaux les plus élevés. En effet, au niveau 2 on retrouve les 200 examens obtenus par la différence entre ceux qui sont au niveau 1 et ceux qu'y restent. Des 200 appartenant au niveau 2, 101 restent au niveau 2 et 17 sont passés au niveau 3 par les contraintes « HELP ».

Ce résultat est très important puisqu'il nous dit que pour 17 tests le milieu biologique peut être déterminé par le composant (dans la forme de morphème) ou d'un adjectif qui accompagne la définition de l'analyse.

Au niveau 3 sont assignés 82 test dont 5 restent au niveau 3 et 1 est passé au niveau 4 par les contraintes « HELP ». C'est-à-dire que seulement une fois l'attributs temps est spécifié dans le milieu biologique et pas dans l'unité de mesure ; on le considère donc un phénomène très rare puisque la majorité des fois les temps est indiqué dans l'unité de mesure (Figure 62).

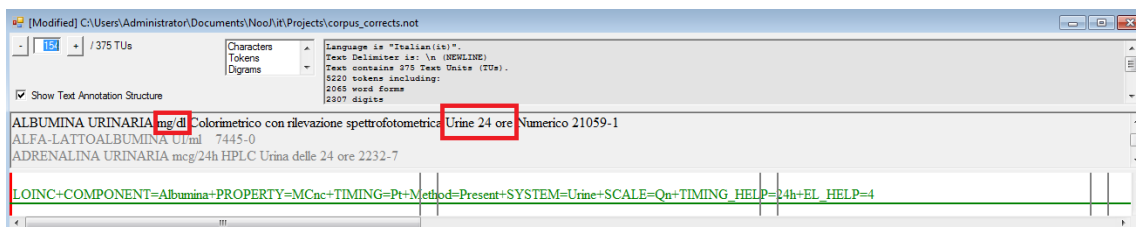


Figure 62 - L'examen appartenant au niveau 4 HELP

Le niveau 4 est composé par 76 examens, obtenus par la différence entre les 82 du niveau 3, le cinq qui restent au niveau 3 et le seul qui passe au 4. Des 76 appartenant au niveau 4 aucun examen reste à ce niveau et 23 sont passés au 5 par les contraintes « HELP ». Le résultat obtenu nous donne une information très importante. Tous les examens qui arrivent à avoir 4 éléments corrects passent aux niveaux les plus élevés ; c'est-à-dire qu'une fois que l'unité de mesure identifie correctement l'axe « Property », qui est le plus difficile à déterminer, les autres déductions pour les attributs « temps » et « échelle de mesure » (qui détermine le niveau 5) sont définies correctes. À cet égard, les examens qui arrivent au niveau 4 automatiquement passent au 5.

Pour les niveaux 5 on retrouve 53 examens dont 27 restent au niveau 5 et les restants passent aux derniers niveaux. Les examens définis totalement corrects sont ceux qui ont tous les éléments bien associés aux codes LOINC officiels.

L'extraction d'information nous a permis d'analyser le corpus et de vérifier que seulement 26 examens peuvent être identifiés comme totalement corrects de façon automatique, contre les 146 qui ne s'arrêtent qu'au niveau 1. Cette différence montre que la description des laboratoires sur les tests médicaux présente encore une large variété linguistique et elle très loin des formes linguistiques officielles, en démontrant l'utilité d'un système de support linguistique pendant les opérations de mapping et de validation.

Conclusions

L'objectif du travail présenté dans cette thèse a concerné la construction d'un système de support aux opérations de validation effectuées par les experts du système de codes et de noms LOINC qui appartiennent au groupe « LOINC Italia ». Ils doivent effectuer périodiquement des opérations de validations et de contrôle des bases des données locales gérées par les laboratoires biologiques-médicaux qui leur demandent de l'aide afin de rendre fiables et interopérables leurs examens.

En particulier, les experts doivent contrôler que les associations entre les définitions locales et les codes officiels LOINC soient effectuées de façon correcte et dépourvue de possibles ambiguïtés. Un support linguistique à ce propos a été nécessaire pour faciliter ces opérations qu'autrement demandent une analyse longue et profonde.

Le contexte d'application de cette thèse offre des importantes lignes de recherche à suivre, et le travail déroulé a permis d'effectuer des évaluations significatives sur les étapes futures à réaliser pour améliorer la situation du contexte actuelle.

Le travail réalisé s'encadre au milieu entre les avantages qui pourraient être apportées pendant l'opération de mapping et les conséquentes facilitations en termes de temps et de précision qui peuvent y être pendant les opérations de validation. En effet, l'application de TALN réalisée sous forme des règles linguistiques et grammaires formelles, a été appliquée en considérant le processus de validation des associations entre les formes linguistiques locales et les codes LOINC déjà effectuées par les laboratoires. La prise en considération du processus de validation a permis d'obtenir des résultats significatifs qui pourront être utilisés aussi pour l'amélioration des stratégies de support au processus de mappage effectué par les laboratoires. En effet, pour appliquer cette typologie d'application de TALN pendant le processus de mapping et contribuer donc à l'amélioration en

termes de temps et de performances à la recherche du code correct, il est demandé d'agir sur les entités officielles LOINC en relation aux possibles codes LOINC que chaque entité pourrait suggérer pendant les opérations de recherche. Cet aspect est très important, et il nous porte vers une réflexion significative. L'absence d'une structure ontologique pour les codes et les entités LOINC crée des difficultés supérieures pendant le processus de recherche. Le niveau de détail et la syntaxe du standard LOINC et la divergence très marquée avec le langage locale utilisé par les laboratoires représente la problématique la plus important à régler.

En considérant une communication entre des experts ayant pour objet un examen de laboratoire, un niveau de détail si minutieux n'est pas nécessaire puisque ils possèdent des connaissances et des compétences suffisantes pour que l'échange d'information ne soit pas ambiguë. Au contraire, la communication entre machines nécessite d'identifier exactement les données traitées au fin de garantir l'interopérabilité.

En effet, une des possibilités à approfondir pour améliorer cette communication technique (effectuée donc entre machines) peut être la construction d'un niveau d'abstraction ontologique partagé qui puisse rendre plus fiable et facile la communication ainsi que la recherche des codes. La construction d'une ontologie de référence des entités officielles et des codes LOINC décrits et associés en classes sémantiques spécifiques liées par des propriétés et composée par des attributs pourrait permettre d'identifier le code LOINC par les liens logiques décrits sous forme de relation dans l'ontologie. Cela permettrait une réduction plus importante des possibles codes à choisir pendant la recherche d'un examen spécifique.

L'application de TALN réalisée a permis de vérifier les divergences entre le langage utilisé par les laboratoires et les formes linguistiques officielles demandées par le système LOINC. La spécificité du logiciel NooJ et la possibilité de décrire les entités par l'association aux classes sémantiques les plus appropriées nous a permis de faire des associations et des déductions entre les éléments locales en référence aux paramètres du standard LOINC (par exemple il est possible de décrire le lien entre l'unité de mesure et la typologie de résultat qui celle-là suggère). Ce

niveau de détail et les déductions existantes entre les différentes entités pourraient être la base pour la construction d'une structure ontologique associée au standard LOINC. Il y aurait, donc, d'une côté l'amélioration de la recherche des codes avec une réduction plus significative de l'ensemble de possibles choix et, d'autre part, la possibilité d'utiliser les classes sémantiques identifiées et décrites dans l'ontologie pour le travail d'annotation sémantique effectué dans ce projet de thèse.

L'avantage obtenu par l'application des règles d'annotation sémantique au corpus analysé concerne la reconnaissance des formes linguistiques locales avec leur association à catégories sémantiques déterminées mais surtout, l'identification du niveau d'exactitude pour chaque examen présent dans le catalogue local analysé. L'identification du niveau d'exactitude, obtenu par l'application des contraintes contextuelles de différente nature, a permis d'améliorer l'efficacité du processus de validation et de réduire fortement le temps d'analyse. L'expert de validation peut commencer en analysant les tests du catalogue local que le système réalisé présente comme ayant un niveau d'exactitude élevé et se focaliser ensuite sur ceux qui présentent un niveau d'exactitude inférieur qui demandent une analyse plus complexe.

Les règles et les informations obtenues par l'application du système de règles réalisé ont alimenté la ressource créée de façon automatique et en format XML. La base de donnée XML représente le collecteur des règles d'association entre les formes locales et les formes linguistiques officielles LOINC, à utiliser comme point de départ pour la construction d'un système de support tant à la validation qu'aux opérations de mappage. Un support à la validation des opérations de mappage demande une intégration de la base de données LOINC ainsi que une formalisation de la connaissance des experts de laboratoire aussi en forme de structure ontologique qui pourra déterminer un niveau de détail supérieur par rapport à celui déjà formalisé. La construction d'un système de support autant pour la validation que pour les opérations de mappage effectuées par les experts de laboratoire représente la prospective la plus importante pour ce travail. Premièrement, l'insertion des règles linguistiques pour la forme en Italien ainsi que l'association

des synonymies à l'intérieur du système de support au codage RELMA peut apporter des avantages pour la recherche plus ponctuelle des codes corrects associés aux examens locaux. En plus, la formalisation de ces règles permet d'avoir déjà une base de connaissance disponible tel pour la construction d'un niveau d'abstraction ontologique que pour la réalisation des ressources linguistiques rendues disponibles pour des utilisations différentes, comme par exemple pour l'intégration des services terminologique standardisés.

En outre, le format XML rend usable la structure des règles identifiées, aussi dans d'autres domaines qui peuvent partager ce genre d'hétérogénéité linguistique. La divergence entre les formes linguistiques utilisées au niveau local et la terminologie standard se vérifie aussi dans d'autres contextes, par exemple pour le codage des diagnostics médicaux. Les diagnostics médicaux se présentent sous forme de commentaire exprimé en langage naturel et viennent associées au code correspondant par le médecin de base que, avec les éléments disponible commence le processus de recherche. Les règles linguistiques et les grammaires formelles réalisées dans ce projet de thèse pourraient être appliquées aussi pour cette typologie de langage en analysant les récursivités et les structures que peuvent se présenter dans les textes, au fin de supporter le spécialiste dans le choix du code le plus spécifique pour le diagnostic médicale effectuée.

Les perspectives de recherche présentées montrent l'importance de ce travail de thèse comme première étape dans le processus de formalisation de mécanismes associatifs qui permettent la correcte utilisation d'un standard. Cette formalisation devient nécessaire dans le contexte actuel, caractérisé, spécialement dans le domaine biomédical, par la prolifération de standards utilisés pour codifier tous les aspects du processus sanitaire, mais pas par un correspondant développement d'outils de support à leur utilisation. L'application de techniques de TALN, suivant des règles formelles comme celles définies dans ce travail de thèse, peut représenter le trait d'union entre le langage naturel et flexible des spécialistes du domaine et le langage structuré des systèmes de codage standard.

Bibliographie

- [1] Agence Nationale de Recherche (ANR), L'intelligence artificielle (AI) et ses enjeux, in « Intelligence artificielle et robotique, les cahiers de l'ANR », Mars 2012, pp. 9-19.
- [2] Bouillon P., Traitement automatique des langues naturelles, Duculot, Belgique, 1998.
- [3] Charlet J., Baneyx A., Steichen O., et al., Utiliser et construire des ontologies en médecine : le primat de la terminologie, in « Revue RSTI - TSI- 28/2009, Web sémantique », pp. 147. Article consulté sur <http://www.medialab.sciences-po.fr/publications/TSI2009_Baneyx.pdf> (Date de la dernière consultation 29/01/2017).
- [4] Chniti A., Griffon N., Traore L., et al., Terminologies et référentiels d'interopérabilité sémantique en Santé, in «JFIM», 2014, pp. 44-58.
- [5] Chomsky N., Structures syntaxiques, Le seuil, Paris, 1957.
- [6] Cimino J.J., Bright T.J., Li J., Medication reconciliation using natural language processing and controlled terminologies, in « Studies in Health Technology and Informatics », vol. 129, 2007, pp. 679-683.
- [7] Domas G., Marchand M., Priorité au contenu, in «Spectra Biologie», n. 158, 2007.
- [8] Frassica, Joseph J., Frequency of laboratory test utilization in the intensive care unit and its implications for large-scale data collection efforts, in «Journal of the American Medical Informatics Association», vol. 12, n. 2, 2005, pp. 229-233.
- [9] Grabar N., Zweigenbaum P., Acquisition automatique de connaissances morphologiques sur le vocabulaire médical, in "Conférence TALN 1999", Cargèse, 12-17 juillet 1999.
- [10] Gross M., Grammaires transformationnelle du français, 1 : le verbe, Larousse, Paris, 1968.
- [11] Gross M., Grammaires transformationnelle du français, 2 : syntaxe du nom, Larousse, Paris, 1977.
- [12] Hagège C., Proux D., Gicquel Q., et al., Développement d'un système de détection des infections associées aux soins à partir de l'analyse de

- comptes-rendus d'hospitalisation, in "TALN 2011", Montpellier, 27 juin - 1er juillet 2011.
- [13] Harris Z., *Methods in Structural Linguistics*, Université de Chicago Press, 1951.
- [14] Harris Z., *Structures mathématiques du langage*, Dunod, Paris, 1973.
- [15] International health terminology standards development organization, *Guidelines for translation of SNOMED CT®*, 2012, <www.snomed.org/resource/resource/110> (Date de la dernière consultation 29/01/2017).
- [16] International health terminology standards development organization, *Guidelines for management of translation of SNOMED CT®*, 2012, <www.snomed.org/resource/resource/110> (Date de la dernière consultation 29/01/2017).
- [17] McDonald C.J., Huff S., Deckard J., et al., *LOINC® Users' Guide*, Décembre 2016.
- [18] Parisi F., *Clinical Term Recognition: From Local to LOINC Terminology. An Application for Italian Language*, in Barone et al. (Eds.) *Automatic Processing of Natural-Language Electronic Texts with NooJ*, Springer International Publishing: NooJ 2016, pp. 1–9, 2017. DOI: 10.1007/978-3-319-55002-2_10. https://link.springer.com/chapter/10.1007/978-3-319-55002-2_10.
- [19] Regenstrief Institute, *Logical Observation Identifiers Names and Codes (LOINC®): translation users' guide*, Regenstrief Institute, 2009.
- [20] Regenstrief Institute, *RELMA® Users' Manual Version 6.17*, 2016, <<https://loinc.org/downloads/files/RELMAManual.pdf>> (Date de la dernière consultation 29/01/2017).
- [21] Reynoso G.A., March A.D., Berra C.M., et al., *Development of the Spanish version of the systematized nomenclature of medicine: methodology and main issues*, in "Proceedings AMIA annual symposium", 2000, pp. 694-698.
- [22] Roland V., Delvaux V., Huet K., et al., *Dynamique phonétique et contrôle moteur dans la maladie de Parkinson: analyse du contrôle de la production des glides*, in "Actes de la conférence conjointe JEP-TALN-RECITAL" 2016, pp. 211-218.
- [23] Roux M., El Zant M., Royauté J., *Projet EPIDEMIA Intervention des transducteurs NooJ*, in "Actes des 9èmes journées scientifiques

INTEX/NooJ, Belgrade, 1-3 Juin 2006, Article consulté sur :
<http://NooJ.matf.bg.ac.rs/pptpdf/18%20Michel%20Roux%20-%20EPIDEMIA_def.pdf> (Date de la dernière consultation 29/01/2017).

- [24] Silberztein M., La formalisation des langues, l'approche de NooJ, ISTE, London, 2015.
- [25] Vreeman D.J., Chiaravalloti M.T., Hook J., et al., Enabling international adoption of LOINC through translation, in «Journal of Biomedical Informatics», 2012.
- [26] Yvon F., Une petite introduction au traitement automatique des langues naturelles, in “Conference on Knowledge discovery and data mining”, 2010, pp. 27-36.

Annexe 1 – Les dictionnaire « Analiti.dic »

Albumina,N+Analita+CORE+Component=Albumina+FLX=F22
Aptoglobina,N+Analita+CORE+Component=Aptoglobina+System=Siero/Plasma+FLX=F22
Conalbumina,N+Analita+CORE+Component=Conalbumina Ab.IgE+FLX=F22
Azoto,N+Analita+CORE+Component=Azoto+FLX=F112
Azoto,N+Analita+CORE+Component=Azoto ureico+FLX=F112
Calcio,N+Analita+CORE+Component=Calcio+FLX=F1
Calcio,N+Analita+CORE+Component=Calcio.ionizzato+FLX=F1
Cloro,N+Analita+CORE+Component=Cloro+FLX=F1
Colecalciferolo,N+Analita+CORE+Component=Colecalciferolo+FLX=F7
Alanina,N+Analita+CORE+Component=Alanina+FLX=F6
Bilirubina,N+Analita+CORE+Component=Bilirubina+FLX=F2
Carbossiemoglobina,N+Analita+CORE+Component=Carbossiemoglobina+FLX=F22
Carbossiemoglobina,N+Analita+CORE+Component=Carbossiemoglobina/Carbossiemoglobina.totale+FLX=F22
Colesterolo,N+Analita+CORE+Component=Colesterolo+FLX=F7
Cortisolo,N+Analita+CORE+Component=CortisoloComponent+FLX=F1
Creatinina,N+Analita+CORE+Component=Creatinina+FLX=F2
Deossiemoglobina,N+Analita+CORE+Component=Deossiemoglobina+FLX=F22
Deossiemoglobina,N+Analita+CORE+Component=Deossiemoglobina/Emoglobina.totale+FLX=F22
Emoglobina,N+Analita+CORE+Component=Emoglobina+FLX=F22
Estradiolo,N+Analita+CORE+Component=Estradiolo+FLX=F7
Ab Insulina,N+Analita+CORE+Component=Insulina Ab+FLX=F8
Ab anti insulina,Ab Insulina,N+Analita+CORE+Component=Insulina Ab+FLX=F8
17 beta estradiolo,Estradiolo,N+Analita+CORE+Component=Estradiolo+FLX=F7
17 Idrossi progesterone,17-Idrossiprogestosterone,N+Analita+CORE+Component=17-Idrossiprogestosterone+FLX=F7
17-beta estradiolo,Estradiolo,N+Analita+CORE+Component=Estradiolo+FLX=F7
Alanina aminotransferasi,N+Analita+CORE+Component=Alanina aminotransferasi+FLX=F6
Albumina/Globulina,N+Analita+CORE+Component=Albumina/Globulina+FLX=F6
Albumina/Proteina.totale,N+Analita+CORE+Component=Albumina/Proteina.totale+FLX=F22
Albumina %,Albumina/Proteina.totale,N+Analita+CORE+Component=Albumina/Proteina.totale+FLX=F22
Complemento C1 esterasi inibitore,N+Analita+CORE+Component=Complemento C1 esterasi inibitore+FLX=F7
Complemento C2,N+Analita+CORE+Component=Complemento C2+FLX=F7
Complemento C3,N+Analita+CORE+Component=Complemento C3+FLX=F7
Complemento C4,N+Analita+CORE+Component=Complemento C4+FLX=F7
C1 inibitore,Complemento C1 esterasi inibitore,N+Analita+CORE+Component=Complemento C1 esterasi inibitore+FLX=F7
C2,Complemento C2,N+Analita+CORE+Component=Complemento C2+FLX=F7
C3,Complemento C3,N+Analita+CORE+Component=Complemento C3+FLX=F7

C4, Complemento C4, N+Analita+CORE+Component=Complemento C4+FLX=F7
 Emoglobina A, N+Analita+CORE+Component=Emoglobina A+FLX=F22
 Emoglobina A1, N+Analita+CORE+Component=Emoglobina A1/Emoglobina.totale+FLX=F22
 Emoglobina A1c, N+Analita+CORE+EMO+GR+Component=Emoglobina A1c+FLX=F22
 Emoglobina A1c, N+Analita+CORE+EMO+GR+Component=Emoglobina A1c/Emoglobina.totale+FLX=F22
 Emoglobina A2, N+Analita+CORE+Component=Emoglobina A2+FLX=F22
 Emoglobina F, N+Analita+CORE+Component=Emoglobina F+FLX=F22
 17 Idrossiprogesterone, 17-Idrossiprogesterone, N+Analita+CORE+Component=17-Idrossiprogesterone+FLX=F7
 17 OH progesterone, 17-Idrossiprogesterone, N+Analita+CORE+Component=17-Idrossiprogesterone+FLX=F7
 17 OH-progesterone, 17-Idrossiprogesterone, N+Analita+CORE+Component=17-Idrossiprogesterone+FLX=F7
 17 beta-estradiolo, Estradiolo, N+Analita+CORE+Component=Estradiolo+FLX=F7
 17-Idrossiprogesterone, N+Analita+CORE+Component=17-Idrossiprogesterone+FLX=F7
 17-OH progesterone, 17-Idrossiprogesterone, N+Analita+CORE+Component=17-Idrossiprogesterone+FLX=F7
 17-OH-progesterone, 17-Idrossiprogesterone, N+Analita+CORE+Component=17-Idrossiprogesterone+FLX=F7
 17 OH progesterone tempo 30', N+Analita+CORE+Component=17-Idrossiprogesterone^30Minuti post XXX stimolo+FLX=F7
 17 OH progesterone tempo 60', N+Analita+CORE+Component=17-Idrossiprogesterone^1h post XXX stimolo+FLX=F7
 17 OH progesterone tempo 90', N+Analita+CORE+Component=17-Idrossiprogesterone^1.5h post XXX stimolo+FLX=F7
 17 OH progesterone tempo 0', N+Analita+CORE+Component=17-Idrossiprogesterone^baseline+FLX=F7
 Adenosina monofosfato, N+Analita+CORE+Component=Adenosina monofosfato.ciclico+FLX=6
 Albuminemia, Albumina, N+Analita+Component=Albumina+System=Sangue art+FLX=F6
 Albuminemia, Albumina, N+Analita+Component=Albumina+System=Sangue capill+FLX=F6
 Albuminemia, Albumina, N+Analita+Component=Albumina+System=Sangue cordone ombelicale+FLX=F6
 Albuminemia, Albumina, N+Analita+Component=Albumina+System=Sangue ven+FLX=F6
 Albuminemia, Albumina, N+Analita+Component=Albumina+System=Sangue+FLX=F6
 Albuminemia, Albumina, N+Analita+Component=Albumina+System=Siero/Plasma+FLX=F6
 Albuminemia, Albumina, N+Analita+Component=Albumina+System=Siero+FLX=F6
 Albuminuria, Albumina, N+Analita+Component=Albumina+System=Urine+FLX=F6
 Alfa 1 %, Alfa 1 globulina, N+Analita+CORE+Component=Alfa 1 globulina/Proteina.totale+FLX=F6
 Alfa 1 globulina, N+Analita+CORE+Component=Alfa 1 globulina+FLX=F88
 Alfa 1, Alfa 1 globulina, N+Analita+CORE+Component=Alfa 1 globulina+FLX=F6
 Alfa 2 %, Alfa 2 globulina, N+Analita+CORE+Component=Alfa 2 globulina/Proteina.totale+FLX=F6
 Alfa 2 globulina, N+Analita+CORE+Component=Alfa 2 globulina+FLX=F88
 Alfa 2, Alfa 2 globulina, N+Analita+CORE+Component=Alfa 2 globulina+FLX=F6
 Alfa 1 %, Alfa 1 globulina, N+Analita+CORE+Component=Alfa 1 globulina/Proteina.totale+FLX=F6
 Alfa1, Alfa 1 globulina, N+Analita+CORE+Component=Alfa 1 globulina+FLX=F6
 Alfa2 %, Alfa 2 globulina, N+Analita+CORE+Component=Alfa 2 globulina/Proteina.totale+FLX=F6
 Alfa2, Alfa 2 globulina, N+Analita+CORE+Component=Alfa 2 globulina+FLX=F6
 Alfa-lattoalbumina, Lattoalbumina alfa, N+Analita+CORE+Component=Lattoalbumina alfa Ab.IgE+FLX=F22
 Alfa lattoalbumina, Lattoalbumina alfa, N+Analita+CORE+Component=Lattoalbumina alfa Ab.IgE+FLX=F22
 Alfa-lattoalbumina, Lattoalbumina alfa, N+Analita+CORE+Component=Lattoalbumina alfa Ab.IgA+FLX=F22
 Alfa lattoalbumina, Lattoalbumina alfa, N+Analita+CORE+Component=Lattoalbumina alfa Ab.IgA+FLX=F22
 Alfa-lattoalbumina, Lattoalbumina alfa, N+Analita+CORE+Component=Lattoalbumina alfa Ab.IgG+FLX=F22
 Alfa lattoalbumina, Lattoalbumina alfa, N+Analita+CORE+Component=Lattoalbumina alfa Ab.IgG+FLX=F22
 Alfa-lattoalbumina, Lattoalbumina alfa, N+Analita+CORE+Component=Lattoalbumina alfa Ab.Ig4+FLX=F22

Alfa lattoalbumina,Lattoalbumina alfa,N+Analita+CORE+Component=Lattoalbumina alfa Ab.Ig4+FLX=F22
Alfa-lattoalbumina,Lattoalbumina alfa,N+Analita+CORE+Component=Lattoalbumina alfa Ag+FLX=F22
Alfa lattoalbumina,Lattoalbumina alfa,N+Analita+CORE+Component=Lattoalbumina alfa Ag+FLX=F22
Aspartato aminotransferasi,Aspartato transaminasi,N+Analita+CORE+Component=Aspartato aminotransferasi+FLX=F7
Aspartato aminotransferasi,N+Analita+CORE+Component=Aspartato aminotransferasi+FLX=F7
Azotemia post dialisi,Azoto ureico post dialisi,N+Analita+Component=Azoto ureico^post dialisi+System=Sangue art+FLX=F6
Azotemia post dialisi,Azoto ureico post dialisi,N+Analita+Component=Azoto ureico^post dialisi+System=Sangue capill+FLX=F6
Azotemia post dialisi,Azoto ureico post dialisi,N+Analita+Component=Azoto ureico^post dialisi+System=Sangue cordone ombelicale+FLX=F6
Azotemia post dialisi,Azoto ureico post dialisi,N+Analita+Component=Azoto ureico^post dialisi+System=Sangue ven+FLX=F6
Azotemia post dialisi,Azoto ureico post dialisi,N+Analita+Component=Azoto ureico^post dialisi+System=Sangue+FLX=F6
Azotemia post dialisi,Azoto ureico post dialisi,N+Analita+Component=Azoto ureico^post dialisi+System=Siero/Plasma+FLX=F6
Azotemia post dialisi,Azoto ureico post dialisi,N+Analita+Component=Azoto ureico^post dialisi+System=Siero+FLX=F6
Azotemia post-dialisi,Azoto ureico post dialisi,N+Analita+Component=Azoto ureico^post dialisi+System=Sangue art+FLX=F6
Azotemia post-dialisi,Azoto ureico post dialisi,N+Analita+Component=Azoto ureico^post dialisi+System=Sangue capill+FLX=F6
Azotemia post-dialisi,Azoto ureico post dialisi,N+Analita+Component=Azoto ureico^post dialisi+System=Sangue cordone ombelicale+FLX=F6
Azotemia post-dialisi,Azoto ureico post dialisi,N+Analita+Component=Azoto ureico^post dialisi+System=Sangue ven+FLX=F6
Azotemia post-dialisi,Azoto ureico post dialisi,N+Analita+Component=Azoto ureico^post dialisi+System=Sangue+FLX=F6
Azotemia post-dialisi,Azoto ureico post dialisi,N+Analita+Component=Azoto ureico^post dialisi+System=Siero/Plasma+FLX=F6
Azotemia post-dialisi,Azoto ureico post dialisi,N+Analita+Component=Azoto ureico^post dialisi+System=Siero+FLX=F6
Azotemia,Azoto ureico,N+Analita+Component=Azoto ureico+System=Sangue art+FLX=F6
Azotemia,Azoto ureico,N+Analita+Component=Azoto ureico+System=Sangue capill+FLX=F6
Azotemia,Azoto ureico,N+Analita+Component=Azoto ureico+System=Sangue cordone ombelicale+FLX=F6
Azotemia,Azoto ureico,N+Analita+Component=Azoto ureico+System=Sangue ven+FLX=F6
Azotemia,Azoto ureico,N+Analita+Component=Azoto ureico+System=Sangue+FLX=F6
Azotemia,Azoto ureico,N+Analita+Component=Azoto ureico+System=Siero/Plasma+FLX=F6
Azotemia,Azoto ureico,N+Analita+Component=Azoto ureico+System=Siero+FLX=F6
Azoto ureico post dialisi,N+Analita+CORE+Component=Azoto ureico^post dialisi+FLX=F113
Azoto ureico post-dialisi,Azoto ureico post dialisi,N+Analita+CORE+Component=Azoto ureico^post dialisi+FLX=F113
Azoto ureico,N+Analita+CORE+Component=Azoto ureico+FLX=F111
Azoturia,Azoto,N+Analita+Component=Azoto+System=Urine+FLX=F6
Beta 1 %,Beta 1 globulina,N+Analita+CORE+Component=Beta 1 globulina/Proteina.totale+FLX=F6
Beta 1 globulina,N+Analita+CORE+Component=Beta 1 globulina+FLX=F88
Beta 1%,Beta 1 globulina,N+Analita+CORE+Component=Beta 1 globulina/Proteina.totale+FLX=F6
Beta 1,Beta 1 globulina,N+Analita+CORE+Component=Beta 1 globulina+FLX=F6
Beta 2 %,Beta 2 globulina,N+Analita+CORE+Component=Beta 2 globulina/Proteina.totale+FLX=F6

Beta 2 globulina,N+Analita+CORE+Component=Beta 2 globulina+FLX=F88
 Beta 2%,Beta 2 globulina,N+Analita+CORE+Component=Beta 2 globulina/Proteina.totale+FLX=F6
 Beta 2,Beta 2 globulina,N+Analita+CORE+Component=Beta 2 globulina+FLX=F6
 Beta1 %,Beta 1 globulina,N+Analita+CORE+Component=Beta 1 globulina/Proteina.totale+FLX=F6
 Beta1%,Beta 1 globulina,N+Analita+CORE+Component=Beta 1 globulina/Proteina.totale+FLX=F6
 Beta1,Beta 1 globulina,N+Analita+CORE+Component=Beta 1 globulina+FLX=F6
 Beta2 %,Beta 2 globulina,N+Analita+CORE+Component=Beta 2 globulina/Proteina.totale+FLX=F6
 Beta2%,Beta 2 globulina,N+Analita+CORE+Component=Beta 2 globulina/Proteina.totale+FLX=F6
 Beta2,Beta 2 globulina,N+Analita+CORE+Component=Beta 2 globulina+FLX=F6
 Beta-lattoalbumina,Lattoalbumina beta,N+Analita+CORE+Component=Lattoalbumina beta+FLX=F22
 Bilirubina diretta,Bilirubina,N+Analita+CORE+Component=Bilirubina+FLX=F2
 Bilirubina indiretta,Bilirubina,N+Analita+CORE+Component=Bilirubina+FLX=F2
 Bilirubina totale,Bilirubina,N+Analita+CORE+Component=Bilirubina+FLX=F2
 Bilirubinemia,Bilirubina,N+Analita+Component=Bilirubina+System=Sangue art+FLX=F6
 Bilirubinemia,Bilirubina,N+Analita+Component=Bilirubina+System=Sangue capill+FLX=F6
 Bilirubinemia,Bilirubina,N+Analita+Component=Bilirubina+System=Sangue cordone ombelicale+FLX=F6
 Bilirubinemia,Bilirubina,N+Analita+Component=Bilirubina+System=Sangue ven+FLX=F6
 Bilirubinemia,Bilirubina,N+Analita+Component=Bilirubina+System=Sangue+FLX=F6
 Bilirubinemia,Bilirubina,N+Analita+Component=Bilirubina+System=Siero/Plasma+FLX=F6
 Bilirubinemia,Bilirubina,N+Analita+Component=Bilirubina+System=Siero+FLX=F6
 Calcemia,Calcio,N+Analita+Component=Calcio+System=Sangue art+FLX=F6
 Calcemia,Calcio,N+Analita+Component=Calcio+System=Sangue capill+FLX=F6
 Calcemia,Calcio,N+Analita+Component=Calcio+System=Sangue cordone ombelicale+FLX=F6
 Calcemia,Calcio,N+Analita+Component=Calcio+System=Sangue ven+FLX=F6
 Calcemia,Calcio,N+Analita+Component=Calcio+System=Siero/Plasma+FLX=F6
 Calcemia,Calcio,N+Analita+Component=Calcio+System=Siero+FLX=F6
 Calcio post dialisi,N+Analita+CORE+Component=Calcio^post dialisi+FLX=F115
 Calcio post-dialisi,N+Analita+CORE+Component=Calcio^post dialisi+FLX=F115
 Calciuria,Calcio,N+Analita+Component=Calcio+System=Urine+FLX=F6
 Cloremia,Cloro,N+Analita+Component=Cloro+System=Sangue art+FLX=F6
 Cloremia,Cloro,N+Analita+Component=Cloro+System=Sangue capill+FLX=F6
 Cloremia,Cloro,N+Analita+Component=Cloro+System=Sangue cordone ombelicale+FLX=F6
 Cloremia,Cloro,N+Analita+Component=Cloro+System=Sangue ven+FLX=F6
 Cloremia,Cloro,N+Analita+Component=Cloro+System=Sangue+FLX=F6
 Cloremia,Cloro,N+Analita+Component=Cloro+System=Siero/Plasma+FLX=F6
 Cloremia,Cloro,N+Analita+Component=Cloro+System=Siero+FLX=F6
 Colesterolo HDL,N+Analita+CORE+Component=Colesterolo.in HDL+FLX=F7
 Colesterolo LDL,N+Analita+CORE+Component=Colesterolo.in LDL+FLX=F7
 Colesterolo totale,Colesterolo,N+Analita+CORE+Component=Colesterolo+FLX=F7
 Colesterolo VLDL,N+Analita+CORE+Component=Colesterolo.in VLDL+FLX=F7
 Creatinina post dialisi,N+Analita+CORE+Component=Creatinina^post dialisi+FLX=F114
 Creatinina post-dialisi,Creatinina post dialisi,N+Analita+CORE+Component=Creatinina^post dialisi+FLX=F114
 Creatininemia post dialisi,Creatinina post dialisi,N+Analita+Component=Creatinina^post dialisi+System=Sangue art+FLX=F6

Creatininemia post dialisi,Creatinina post dialisi,N+Analita+Component=Creatinina^post dialisi+System=Sangue capill+FLX=F6
 Creatininemia post dialisi,Creatinina post dialisi,N+Analita+Component=Creatinina^post dialisi+System=Sangue cordone ombelicale+FLX=F6
 Creatininemia post dialisi,Creatinina post dialisi,N+Analita+Component=Creatinina^post dialisi+System=Sangue ven+FLX=F6
 Creatininemia post dialisi,Creatinina post dialisi,N+Analita+Component=Creatinina^post dialisi+System=Sangue+FLX=F6
 Creatininemia post dialisi,Creatinina post dialisi,N+Analita+Component=Creatinina^post dialisi+System=Siero/Plasma+FLX=F6
 Creatininemia post dialisi,Creatinina post dialisi,N+Analita+Component=Creatinina^post dialisi+System=Siero+FLX=F6
 Creatininemia post-dialisi,Creatinina post dialisi,N+Analita+Component=Creatinina^post dialisi+System=Sangue art+FLX=F6
 Creatininemia post-dialisi,Creatinina post dialisi,N+Analita+Component=Creatinina^post dialisi+System=Sangue capill+FLX=F6
 Creatininemia post-dialisi,Creatinina post dialisi,N+Analita+Component=Creatinina^post dialisi+System=Sangue cordone ombelicale+FLX=F6
 Creatininemia post-dialisi,Creatinina post dialisi,N+Analita+Component=Creatinina^post dialisi+System=Sangue ven+FLX=F6
 Creatininemia post-dialisi,Creatinina post dialisi,N+Analita+Component=Creatinina^post dialisi+System=Sangue+FLX=F6
 Creatininemia post-dialisi,Creatinina post dialisi,N+Analita+Component=Creatinina^post dialisi+System=Siero/Plasma+FLX=F6
 Creatininemia post-dialisi,Creatinina post dialisi,N+Analita+Component=Creatinina^post dialisi+System=Siero+FLX=F6
 Creatininemia,Creatinina,N+Analita+Component=Creatinina+System=Sangue art+FLX=F6
 Creatininemia,Creatinina,N+Analita+Component=Creatinina+System=Sangue capill+FLX=F6
 Creatininemia,Creatinina,N+Analita+Component=Creatinina+System=Sangue cordone ombelicale+FLX=F6
 Creatininemia,Creatinina,N+Analita+Component=Creatinina+System=Sangue ven+FLX=F6
 Creatininemia,Creatinina,N+Analita+Component=Creatinina+System=Sangue+FLX=F6
 Creatininemia,Creatinina,N+Analita+Component=Creatinina+System=Siero/Plasma+FLX=F6
 Creatininemia,Creatinina,N+Analita+Component=Creatinina+System=Siero+FLX=F6
 Creatinuria,Creatinina,N+Analita+Component=Creatinina+System=Urine+FLX=F6
 Emoglobinemia,Emoglobina,N+Analita+Component=Emoglobina+System=Sangue art+FLX=F6
 Emoglobinemia,Emoglobina,N+Analita+Component=Emoglobina+System=Sangue capill+FLX=F6
 Emoglobinemia,Emoglobina,N+Analita+Component=Emoglobina+System=Sangue cordone ombelicale+FLX=F6
 Emoglobinemia,Emoglobina,N+Analita+Component=Emoglobina+System=Sangue ven+FLX=F6
 Emoglobinemia,Emoglobina,N+Analita+Component=Emoglobina+System=Sangue+FLX=F6
 Emoglobinemia,Emoglobina,N+Analita+Component=Emoglobina+System=Siero/Plasma+FLX=F6
 Emoglobinemia,Emoglobina,N+Analita+Component=Emoglobina+System=Siero+FLX=F6
 Emoglobinuria,Emoglobina,N+Analita+Component=Emoglobina+System=Urine+FLX=F6
 Epididimo umano proteina 4,N+Analita+CORE+Component=Epididimo umano proteina 4+FLX=F7
 Epinefrina,N+Analita+CORE+Component=Epinefrina+FLX=F6
 Esame striscio periferico,Striscio di sangue,N+Analita+CORE+Component=Striscio di sangue osservazione+FLX=F7
 Estradiolo 1° campione,N+Analita+CORE+Component=Estradiolo^1° campione post XXX stimolo+FLX=F7
 Estradiolo 1 camp,Estradiolo 1° campione,N+Analita+CORE+Component=Estradiolo^1° campione post XXX stimolo+FLX=F7
 Estradiolo 2° campione,N+Analita+CORE+Component=Estradiolo^2° campione post XXX stimolo+FLX=F7

Estradiolo 120',Estradiolo 2° campione,N+Analita+CORE+Component=Estradiolo^2° campione post XXX stimolo+FLX=F7
Estrone,N+Analita+CORE+Component=Estrone+FLX=F7
Fattore di coagulazione II,N+Analita+CORE+Component=Fattore di coagulazione II inibitore+FLX=F7
Fattore di coagulazione IX,N+Analita+CORE+Component=Fattore di coagulazione IX Ag+FLX=F7
Fattore di coagulazione IX,N+Analita+CORE+Component=Fattore di coagulazione IX attivazione+FLX=F7
Fattore di coagulazione IX,N+Analita+CORE+Component=Fattore di coagulazione IX attività effettivo/normale+FLX=F7
Fattore di coagulazione IX,N+Analita+CORE+Component=Fattore di coagulazione IX attività+FLX=F7
Fattore di coagulazione IX,N+Analita+CORE+Component=Fattore di coagulazione IX inibitore+FLX=F7
Fattore di coagulazione V,N+Analita+CORE+Component=Fattore di coagulazione V Ag+FLX=F7
Fattore di coagulazione V,N+Analita+CORE+Component=Fattore di coagulazione V attivazione+FLX=F7
Fattore di coagulazione V,N+Analita+CORE+Component=Fattore di coagulazione V attività effettivo/normale+FLX=F7
Fattore di coagulazione V,N+Analita+CORE+Component=Fattore di coagulazione V inibitore+FLX=F7
Fattore di coagulazione VI,N+Analita+CORE+Component=Fattore di coagulazione VI Ag+FLX=F7
Fattore di coagulazione VII,N+Analita+CORE+Component=Fattore di coagulazione VII Ag+FLX=F7
Fattore di coagulazione VII,N+Analita+CORE+Component=Fattore di coagulazione VII attivazione+FLX=F7
Fattore di coagulazione VII,N+Analita+CORE+Component=Fattore di coagulazione VII attività effettivo/normale+FLX=F7
Fattore di coagulazione VII,N+Analita+CORE+Component=Fattore di coagulazione VII attività+FLX=F7
Fattore di coagulazione VII,N+Analita+CORE+Component=Fattore di coagulazione VII inibitore+FLX=F7
Fattore di coagulazione VIII,N+Analita+CORE+Component=Fattore di coagulazione VIII Ab+FLX=F7
Fattore di coagulazione VIII,N+Analita+CORE+Component=Fattore di coagulazione VIII Ag+FLX=F7
Fattore di coagulazione VIII,N+Analita+CORE+Component=Fattore di coagulazione VIII attivazione+FLX=F7
Fattore di coagulazione VIII,N+Analita+CORE+Component=Fattore di coagulazione VIII attività
effettivo/normale+FLX=F7
Fattore di coagulazione VIII,N+Analita+CORE+Component=Fattore di coagulazione VIII attività+FLX=F7
Fattore di coagulazione VIII,N+Analita+CORE+Component=Fattore di coagulazione VIII inibitore+FLX=F7
Fattore di coagulazione X,N+Analita+CORE+Component=Fattore di coagulazione X Ag+FLX=F7
Fattore di coagulazione X,N+Analita+CORE+Component=Fattore di coagulazione X attivazione+FLX=F7
Fattore di coagulazione X,N+Analita+CORE+Component=Fattore di coagulazione X attività effettivo/normale+FLX=F7
Fattore di coagulazione X,N+Analita+CORE+Component=Fattore di coagulazione X attività+FLX=F7
Fattore di coagulazione X,N+Analita+CORE+Component=Fattore di coagulazione X inibitore+FLX=F7
Fattore di coagulazione XI,N+Analita+CORE+Component=Fattore di coagulazione XI Ag+FLX=F7
Fattore di coagulazione XI,N+Analita+CORE+Component=Fattore di coagulazione XI attivazione+FLX=F7
Fattore di coagulazione XI,N+Analita+CORE+Component=Fattore di coagulazione XI attività effettivo/normale+FLX=F7
Fattore di coagulazione XI,N+Analita+CORE+Component=Fattore di coagulazione XI attività+FLX=F7
Fattore di coagulazione XI,N+Analita+CORE+Component=Fattore di coagulazione XI inibitore+FLX=F7
Fattore di coagulazione XII,N+Analita+CORE+Component=Fattore di coagulazione XII Ag+FLX=F7
Fattore di coagulazione XII,N+Analita+CORE+Component=Fattore di coagulazione XII attivazione+FLX=F7
Fattore di coagulazione XII,N+Analita+CORE+Component=Fattore di coagulazione XII attività effettivo/normale+FLX=F7
Fattore di coagulazione XII,N+Analita+CORE+Component=Fattore di coagulazione XII attività+FLX=F7
Fattore di coagulazione XII,N+Analita+CORE+Component=Fattore di coagulazione XII inibitore+FLX=F7
Fattore di coagulazione XIII,N+Analita+CORE+Component=Fattore di coagulazione XIII Ag+FLX=F7
Fattore di coagulazione XIII,N+Analita+CORE+Component=Fattore di coagulazione XIII attivazione+FLX=F7
Fattore di coagulazione XIII,N+Analita+CORE+Component=Fattore di coagulazione XIII attività
effettivo/normale+FLX=F7
Fattore di coagulazione XIII,N+Analita+CORE+Component=Fattore di coagulazione XIII inibitore+FLX=F7

Fattore di crescita insulino simile I,N+Analita+CORE+Component=Fattore di crescita insulino-simile I+FLX=F9

Fattore V,Fattore di coagulazione V,N+Analita+CORE+Component=Fattore di coagulazione V attività effettivo/normale+FLX=F7

Fattore VII,Fattore di coagulazione VII,N+Analita+CORE+Component=Fattore di coagulazione VII attività effettivo/normale+FLX=F7

Fattore VIII,Fattore di coagulazione VIII,N+Analita+CORE+Component=Fattore di coagulazione VIII attività effettivo/normale+FLX=F7

Fibrinogeno,N+Analita+CORE+Component=Fibrinogeno+FLX=F78

Follitropina,N+Analita+CORE+Component=Follitropina+FLX=F6

Follitropina 1 campione,N+Analita+CORE+Component=Follitropina^1° campione+FLX=F6

Follitropina 2 campione,N+Analita+CORE+Component=Follitropina^2° campione+FLX=F6

Follitropina 3 campione,N+Analita+CORE+Component=Follitropina^3° campione+FLX=F6

Follitropina 4 campione,N+Analita+CORE+Component=Follitropina^4° campione+FLX=F6

Follitropina 5 campione,N+Analita+CORE+Component=Follitropina^5° campione+FLX=F6

Follitropina 6 campione,N+Analita+CORE+Component=Follitropina^6° campione+FLX=F6

FSH 1 campione,Follitropina 1 campione,N+Analita+Acronimo+Component=Follitropina^1° campione+FLX=F6

FSH 2 campione,Follitropina 2 campione,N+Analita+Acronimo+Component=Follitropina^2° campione+FLX=F6

FSH 3 campione,Follitropina 3 campione,N+Analita+Acronimo+Component=Follitropina^3° campione+FLX=F6

FSH 4 campione,Follitropina 4 campione,N+Analita+Acronimo+Component=Follitropina^4° campione+FLX=F6

FSH 5 campione,Follitropina 5 campione,N+Analita+Acronimo+Component=Follitropina^5° campione+FLX=F6

FSH 6 campione,Follitropina 6 campione,N+Analita+Acronimo+Component=Follitropina^6° campione+FLX=F6

Fosfatasi alcalina biliare,N+Analita+CORE+Component=Fosfatasi alcalina.bile+FLX=F6

Fosfatasi alcalina di regan,N+Analita+CORE+Component=Fosfatasi alcalina.regan+FLX=F6

Fosfatasi alcalina epatica,Fosfatasi alcalina fegato,N+Analita+CORE+Component=Fosfatasi alcalina.fegato+FLX=F6

Fosfatasi alcalina fegato,N+Analita+CORE+Component=Fosfatasi alcalina.fegato+FLX=F6

Fosfatasi alcalina intestinale,N+Analita+CORE+Component=Fosfatasi alcalina.intestinale+FLX=F6

Fosf alcalina iso osseo,Fosfatasi alcalina osso,N+Analita+CORE+Component=Fosfatasi alcalina.osso+FLX=F6

Fosfatasi alcalina ossea,Fosfatasi alcalina osso,N+Analita+CORE+Component=Fosfatasi alcalina.osso+FLX=F6

Fosfatasi alcalina osso,N+Analita+CORE+Component=Fosfatasi alcalina.osso+FLX=F6

Fosfatasi alcalina placentare,N+Analita+CORE+Component=Fosfatasi alcalina.placentale+FLX=F6

Fosfatasi alcalina,N+Analita+CORE+Component=Fosfatasi alcalina+FLX=F6

Fosforo,Fosfato,N+Analita+CORE+Component=Fosfato+FLX=F7

Frazione epatica,Fosfatasi alcalina fegato,N+Analita+CORE+Component=Fosfatasi alcalina.fegato+FLX=F6

Frazione ossea,Fosfatasi alcalina osso,N+Analita+CORE+Component=Fosfatasi alcalina.osso/Fosfatasi alcalina.totale+FLX=F6

Frazione ossea,Fosfatasi alcalina osso,N+Analita+CORE+Component=Fosfatasi alcalina.osso+FLX=F6

Gamma %,Gamma globulina,N+Analita+CORE+Component=Gamma globulina/Proteina.totale+FLX=F6

Gamma globulina,N+Analita+CORE+Component=Gamma globulina+FLX=F88

Gamma,Gamma globulina,N+Analita+CORE+Component=Gamma globulina+FLX=F6

Glicemia,Glucosio,N+Analita+Component=Glucosio+System=Sangue art+FLX=F6

Glicemia,Glucosio,N+Analita+Component=Glucosio+System=Sangue capill+FLX=F6

Glicemia,Glucosio,N+Analita+Component=Glucosio+System=Sangue cordone ombelicale+FLX=F6

Glicemia,Glucosio,N+Analita+Component=Glucosio+System=Sangue ven+FLX=F6

Glicemia,Glucosio,N+Analita+Component=Glucosio+System=Sangue+FLX=F6

Glicemia,Glucosio,N+Analita+Component=Glucosio+System=Siero/Plasma+FLX=F6

Glicemia,Glucosio,N+Analita+Component=Glucosio+System=Siero+FLX=F6
 Glicemia basale,Glucosio,N+Analita+Component=Glucosio^baseline+System=Sangue capill+FLX=F6
 Glicemia basale,Glucosio,N+Analita+Component=Glucosio^baseline+System=Siero/Plasma+FLX=F6
 Glicemia dopo 120',Glucosio,N+Analita+Component=Glucosio^2h post carico di glucosio+System=Sangue+FLX=F6
 Glicemia dopo 120',Glucosio,N+Analita+Component=Glucosio^2h post carico di glucosio+System=Plasma+FLX=F6
 Glicemia dopo 120',Glucosio,N+Analita+Component=Glucosio^2h post carico di glucosio+System=Siero/Plasma+FLX=F6
 Glicemia dopo 120',Glucosio,N+Analita+Component=Glucosio^2h post XXX stimolo+System=Siero/Plasma+FLX=F6
 Glicemia dopo 180',Glucosio,N+Analita+Component=Glucosio^3h post carico di glucosio+System=Plasma+FLX=F6
 Glicemia dopo 180',Glucosio,N+Analita+Component=Glucosio^3h post carico di glucosio+System=Siero/Plasma+FLX=F6
 Glicemia dopo 180',Glucosio,N+Analita+Component=Glucosio^3h post XXX stimolo+System=Siero/Plasma+FLX=F6
 Glicemia dopo 30',Glucosio,N+Analita+Component=Glucosio^30Minuti post carico di glucosio+System=Plasma+FLX=F6
 Glicemia dopo 30',Glucosio,N+Analita+Component=Glucosio^30Minuti post carico di glucosio+System=Siero/Plasma+FLX=F6
 Glicemia dopo 30',Glucosio,N+Analita+Component=Glucosio^30Minuti post XXX stimolo+System=Siero/Plasma+FLX=F6
 Glicemia dopo 60',Glucosio,N+Analita+Component=Glucosio^1h post carico con glucosio+System=Plasma+FLX=F6
 Glicemia dopo 60',Glucosio,N+Analita+Component=Glucosio^1h post carico con glucosio+System=Siero/Plasma+FLX=F6
 Glicemia dopo 60',Glucosio,N+Analita+Component=Glucosio^1h post XXX stimolo+System=Siero/Plasma+FLX=F6
 Glicemia dopo 90',Glucosio,N+Analita+Component=Glucosio^1.5h post carico con glucosio+System=Siero/Plasma+FLX=F6
 Glicemia dopo 90',Glucosio,N+Analita+Component=Glucosio^1.5h post XXX stimolo+System=Siero/Plasma+FLX=F6
 Glicuria,Glucosio,N+Analita+Component=Glucosio+System=Urine+FLX=F6
 Glucosio post cena,Glucosio post pasto,N+Analita+CORE+Component=Glucosio^post pasto+FLX=F116
 Glucosio post colazione,Glucosio post pasto,N+Analita+CORE+Component=Glucosio^post pasto+FLX=F116
 Glucosio post pasto,N+Analita+CORE+Component=Glucosio^post pasto+FLX=F11
 Glucosio post prandiale,Glucosio post pasto,N+Analita+CORE+Component=Glucosio^post pasto+FLX=F116
 Glucosio post pranzo,Glucosio post pasto,N+Analita+CORE+Component=Glucosio^post pasto+FLX=F116
 Insulina basale,Insulina baseline,N+Analita+CORE+Component=Insulina^baseline+FLX=F8
 Insulina post cena,Insulina post pasto,N+Analita+CORE+Component=Insulina^post pasto+FLX=F6
 Insulina post colazione,Insulina post pasto,N+Analita+CORE+Component=Insulina^post pasto+FLX=F6
 Insulina post pasto,N+Analita+CORE+Component=Insulina^post pasto+FLX=F6
 Insulina post prandiale,Insulina post pasto,N+Analita+CORE+Component=Insulina^post pasto+FLX=F6
 Insulina post pranzo,Insulina post pasto,N+Analita+CORE+Component=Insulina^post pasto+FLX=F6
 Glucosio,N+Analita+CORE+Component=Glucosio+FLX=F3
 Glucosio-6-Fosfato deidrogenasi,N+Analita+CORE+Component=Glucosio-6-Fosfato deidrogenasi+FLX=F7
 Idrossido,N+Analita+CORE+Component=Idrossido+FLX=F78
 Insulina,N+Analita+CORE+Component=Insulina+FLX=F8
 Insulina 1 campione,N+Analita+CORE+Component=Insulina^1° campione+FLX=F8
 Insulina 2 campione,N+Analita+CORE+Component=Insulina^2° campione+FLX=F8
 Insulina 3 campione,N+Analita+CORE+Component=Insulina^3° campione+FLX=F8
 Insulina 4 campione,N+Analita+CORE+Component=Insulina^4° campione+FLX=F8
 Insulina 5 campione,N+Analita+CORE+Component=Insulina^5° campione+FLX=F8
 Insulina 6 campione,N+Analita+CORE+Component=Insulina^6° campione+FLX=F8
 Insulina 7 campione,N+Analita+CORE+Component=Insulina^7° campione+FLX=F8
 Insulina 1 camp.,Insulina 1 campione,N+Analita+CORE+Component=Insulina^1° campione+FLX=F8
 Insulina 2 camp.,Insulina 2 campione,N+Analita+CORE+Component=Insulina^2° campione+FLX=F8

Insulina 3 camp.,Insulina 3 campione,N+Analita+CORE+Component=Insulina^3° campione+FLX=F8
 Insulina 4 camp.,Insulina 4 campione,N+Analita+CORE+Component=Insulina^4° campione+FLX=F8
 Insulina 5 camp.,Insulina 5 campione,N+Analita+CORE+Component=Insulina^5° campione+FLX=F8
 Insulina 6 camp.,Insulina 6 campione,N+Analita+CORE+Component=Insulina^6° campione+FLX=F8
 Insulina 7 camp.,Insulina 7 campione,N+Analita+CORE+Component=Insulina^7° campione+FLX=F8
 Insulinemia,Insulina,N+Analita+Component=Insulina+System=Sangue art+FLX=F6
 Insulinemia,Insulina,N+Analita+Component=Insulina+System=Sangue capill+FLX=F6
 Insulinemia,Insulina,N+Analita+Component=Insulina+System=Sangue cordone ombelicale+FLX=F6
 Insulinemia,Insulina,N+Analita+Component=Insulina+System=Sangue ven+FLX=F6
 Insulinemia,Insulina,N+Analita+Component=Insulina+System=Sangue+FLX=F6
 Insulinemia,Insulina,N+Analita+Component=Insulina+System=Siero/Plasma+FLX=F6
 Insulinemia,Insulina,N+Analita+Component=Insulina+System=Siero+FLX=F6
 Isoenzima biliare,Fosfatasi alcalina biliare,N+Analita+CORE+Component=Fosfatasi alcalina.bile/Fosfatasi
 alcalina.totale+FLX=F11
 Isoenzima biliare,Fosfatasi alcalina biliare,N+Analita+CORE+Component=Fosfatasi alcalina.bile+FLX=F11
 Isoenzima di regan,Fosfatasi alcalina di regan,N+Analita+CORE+Component=Fosfatasi alcalina.regan/Fosfatasi
 alcalina.totale+FLX=F16
 Isoenzima di regan,Fosfatasi alcalina di regan,N+Analita+CORE+Component=Fosfatasi alcalina.regan+FLX=F16
 Isoenzima epatico,Fosfatasi alcalina fegato,N+Analita+CORE+Component=Fosfatasi alcalina.fegato/Fosfatasi
 alcalina.totale+FLX=F11
 Isoenzima intestinale,Fosfatasi alcalina intestinale,N+Analita+CORE+Component=Fosfatasi alcalina.intestinale/Fosfatasi
 alcalina.totale+FLX=F11
 Isoenzima intestinale,Fosfatasi alcalina intestinale,N+Analita+CORE+Component=Fosfatasi alcalina.intestinale+FLX=F11
 Isoenzima osseo,Fosfatasi alcalina osso,N+Analita+CORE+Component=Fosfatasi alcalina.osso/Fosfatasi
 alcalina.totale+FLX=F11
 Isoenzima osseo,Fosfatasi alcalina osso,N+Analita+CORE+Component=Fosfatasi alcalina.osso+FLX=F11
 Isoenzima placentare,Fosfatasi alcalina placentare,N+Analita+CORE+Component=Fosfatasi alcalina.placentale/Fosfatasi
 alcalina.totale+FLX=F11
 Isoenzima placentare,Fosfatasi alcalina placentare,N+Analita+CORE+Component=Fosfatasi alcalina.placentale+FLX=F11
 Isoenzima,Fosfatasi alcalina,N+Analita+CORE+Component=Fosfatasi alacalina+FLX=F78
 Lattato deidrogenasi 1,N+Analita+CORE+Component=Lattato deidrogenasi 1/Lattato deidrogenasi 2+FLX=F7
 Lattato deidrogenasi 1,N+Analita+CORE+Component=Lattato deidrogenasi 1/Lattato deidrogenasi.totale+FLX=F7
 Lattato deidrogenasi 1,N+Analita+CORE+Component=Lattato deidrogenasi 1+FLX=F7
 Lattato deidrogenasi 2,N+Analita+CORE+Component=Lattato deidrogenasi 2/Lattato deidrogenasi.totale+FLX=F7
 Lattato deidrogenasi 2,N+Analita+CORE+Component=Lattato deidrogenasi 2+FLX=F7
 Lattato deidrogenasi 3,N+Analita+CORE+Component=Lattato deidrogenasi 3/Lattato deidrogenasi.totale+FLX=F7
 Lattato deidrogenasi 3,N+Analita+CORE+Component=Lattato deidrogenasi 3+FLX=F7
 Lattato deidrogenasi 4,N+Analita+CORE+Component=Lattato deidrogenasi 4/Lattato deidrogenasi.totale+FLX=F7
 Lattato deidrogenasi 4,N+Analita+CORE+Component=Lattato deidrogenasi 4+FLX=F7
 Lattato deidrogenasi 5,N+Analita+CORE+Component=Lattato deidrogenasi 5/Lattato deidrogenasi.totale+FLX=F7
 Lattato deidrogenasi 5,N+Analita+CORE+Component=Lattato deidrogenasi 5+FLX=F7
 Lattato deidrogenasi,N+Analita+CORE+Component=Lattato deidrogenasi+FLX=F7
 Lattato deidrogenasi isoenzimi,N+Analita+CORE+Component=Lattato deidrogenasi isoenzimi+FLX=F7
 Lattoalbumina,N+Analita+CORE+Component=Lattoalbumina Ab.IgE+FLX=F22
 Lattoalbumina,N+Analita+CORE+Component=Lattoalbumina Ab.IgG+FLX=F22

Lattoalbumina alfa,N+Analita+CORE+Component=Lattoalbumina alfa Ab.IgE+FLX=F22
 Lattoalbumina alfa,N+Analita+CORE+Component=Lattoalbumina alfa Ab.IgA+FLX=F22
 Lattoalbumina alfa,N+Analita+CORE+Component=Lattoalbumina alfa Ab.IgG+FLX=F22
 Lattoalbumina alfa,N+Analita+CORE+Component=Lattoalbumina alfa Ab.Ig4+FLX=F22
 Lattoalbumina alfa,N+Analita+CORE+Component=Lattoalbumina alfa Ag+FLX=F22
 Lipidemia,Lipidi,N+Analita+Component=Lipidi+System=Sangue art+FLX=F6
 Lipidemia,Lipidi,N+Analita+Component=Lipidi+System=Sangue capill+FLX=F6
 Lipidemia,Lipidi,N+Analita+Component=Lipidi+System=Sangue cordone ombelicale+FLX=F6
 Lipidemia,Lipidi,N+Analita+Component=Lipidi+System=Sangue ven+FLX=F6
 Lipidemia,Lipidi,N+Analita+Component=Lipidi+System=Sangue+FLX=F6
 Lipidemia,Lipidi,N+Analita+Component=Lipidi+System=Siero/Plasma+FLX=F6
 Lipidemia,Lipidi,N+Analita+Component=Lipidi+System=Siero+FLX=F6
 Lipidi,N+Analita+CORE+Component=Lipidi+FLX=F5
 Lutropina,N+Analita+CORE+Component=Lutropina+FLX=F6
 Lutropina 1 campione,N+Analita+CORE+Component=Lutropina^1° campione+FLX=F6
 Lutropina 2 campione,N+Analita+CORE+Component=Lutropina^2° campione+FLX=F6
 Lutropina 3 campione,N+Analita+CORE+Component=Lutropina^3° campione+FLX=F6
 Lutropina 4 campione,N+Analita+CORE+Component=Lutropina^4° campione+FLX=F6
 Lutropina 5 campione,N+Analita+CORE+Component=Lutropina^5° campione+FLX=F6
 Lutropina 6 campione,N+Analita+CORE+Component=Lutropina^6° campione+FLX=F6
 LH 1 camp.,Lutropina 1 campione,N+Analita+CORE+Component=Lutropina^1° campione+FLX=F6
 LH 2 camp.,Lutropina 2 campione,N+Analita+CORE+Component=Lutropina^2° campione+FLX=F6
 LH 3 camp.,Lutropina 3 campione,N+Analita+CORE+Component=Lutropina^3° campione+FLX=F6
 LH 4 camp.,Lutropina 4 campione,N+Analita+CORE+Component=Lutropina^4° campione+FLX=F6
 LH 5 camp.,Lutropina 5 campione,N+Analita+CORE+Component=Lutropina^5° campione+FLX=F6
 LH 6 camp.,Lutropina 6 campione,N+Analita+CORE+Component=Lutropina^6° campione+FLX=F6
 Manganese,N+Analita+CORE+Component=Manganese+FLX=F1
 Manganese,N+Analita+CORE+Component=Manganese/Creatinina+FLX=F1
 Mercurio,N+Analita+CORE+Component=Mercurio+FLX=F1
 Mercurio,N+Analita+CORE+Component=Mercurio/Creatinina+FLX=F1
 Metaemoglobina,N+Analita+CORE+Component=Metaemoglobina+FLX=F22
 Metaemoglobina,N+Analita+CORE+Component=Metaemoglobina/Emoglobina.totale+FLX=F22
 Metaemoglobinemia,Metaemoglobina,N+Analita+Component=Metaemoglobina+System=Sangue art+FLX=F6
 Metaemoglobinemia,Metaemoglobina,N+Analita+Component=Metaemoglobina+System=Sangue capill+FLX=F6
 Metaemoglobinemia,Metaemoglobina,N+Analita+Component=Metaemoglobina+System=Sangue cordone
 ombelicale+FLX=F6
 Metaemoglobinemia,Metaemoglobina,N+Analita+Component=Metaemoglobina+System=Sangue ven+FLX=F6
 Metaemoglobinemia,Metaemoglobina,N+Analita+Component=Metaemoglobina+System=Sangue+FLX=F6
 Metaemoglobinemia,Metaemoglobina,N+Analita+Component=Metaemoglobina+System=Siero/Plasma+FLX=F6
 Norepinefrina,N+Analita+CORE+Component=Norepinefrina+FLX=F6
 OH progesterone,17-Idrossiprogestosterone,N+Analita+CORE+Component=17-Idrossiprogestosterone+FLX=F7
 OH-progesterone,17-Idrossiprogestosterone,N+Analita+CORE+Component=17-Idrossiprogestosterone+FLX=F7
 Ossiemoglobina,N+Analita+CORE+Component=Ossiemoglobina+FLX=F22
 Ovoalbumina,N+Analita+CORE+Component=Ovoalbumina Ab.IgE+FLX=F22
 Plasminogeno,N+Analita+CORE+Component=Plasminogeno Ag+FLX=F7

Plasminogeno,N+Analita+CORE+Component=Plasminogeno effettivo/normale+FLX=F7
 Plasminogeno,N+Analita+CORE+Component=Plasminogeno+FLX=F7
 Potassiemia,Potassio,N+Analita+Component=Potassio+System=Sangue art+FLX=F6
 Potassiemia,Potassio,N+Analita+Component=Potassio+System=Sangue capill+FLX=F6
 Potassiemia,Potassio,N+Analita+Component=Potassio+System=Sangue cordone ombelicale+FLX=F6
 Potassiemia,Potassio,N+Analita+Component=Potassio+System=Sangue ven+FLX=F6
 Potassiemia,Potassio,N+Analita+Component=Potassio+System=Siero/Plasma+FLX=F6
 Potassiemia,Potassio,N+Analita+Component=Potassio+System=Siero+FLX=F6
 Potassio post dialisi,N+Analita+CORE+Component=Potassio^post dialisi+FLX=F115
 Potassio post-dialisi,Potassio post dialisi,N+Analita+CORE+Component=Potassio^post dialisi+FLX=F115
 Potassio,N+Analita+CORE+Component=Potassio+FLX=F1
 Potassuria,Potassio,N+Analita+Component=Potassio+System=Urine+FLX=F6
 Pregnenolone,N+Analita+CORE+Component=Pregnenolone+FLX=F7
 Progesterone,N+Analita+CORE+Component=Progesterone+FLX=F7
 Prolattina,N+Analita+CORE+Component=Prolattina+FLX=F6
 Prolattina tempo 10',N+Analita+CORE+Component=Prolattina^10Minuti post XXX stimolo+FLX=F6
 Prolattina tempo 15',N+Analita+CORE+Component=Prolattina^15Minuti post XXX stimolo+FLX=F6
 Prolattina tempo 20',N+Analita+CORE+Component=Prolattina^20Minuti post XXX stimolo+FLX=F6
 Prolattina tempo 40',N+Analita+CORE+Component=Prolattina^40Minuti post XXX stimolo+FLX=F6
 Prolattina tempo 40,N+Analita+CORE+Component=Prolattina^40Minuti post XXX stimolo+FLX=F6
 Prolattina campione 1,N+Analita+CORE+Component=Prolattina^1° campione+FLX=F6
 Prolattina campione 2,N+Analita+CORE+Component=Prolattina^2° campione+FLX=F6
 Prolattina campione 3,N+Analita+CORE+Component=Prolattina^3° campione+FLX=F6
 Prolattina campione 4,N+Analita+CORE+Component=Prolattina^4° campione+FLX=F6
 Prolattina campione 5,N+Analita+CORE+Component=Prolattina^5° campione+FLX=F6
 Prolattina 1 camp,Prolattina campione 1,N+Analita+CORE+Component=Prolattina^1° campione+FLX=F6
 Prolattina 2 camp,Prolattina campione 2,N+Analita+CORE+Component=Prolattina^2° campione+FLX=F6
 Prolattina 3 camp,Prolattina campione 3,N+Analita+CORE+Component=Prolattina^3° campione+FLX=F6
 Prolattina 4 camp,Prolattina campione 4,N+Analita+CORE+Component=Prolattina^4° campione+FLX=F6
 Prolattina 5 camp,Prolattina campione 5,N+Analita+CORE+Component=Prolattina^5° campione+FLX=F6
 Proteina C reattiva,N+Analita+CORE+Component=Proteina C reattiva+FLX=F6
 Proteina,N+Analita+CORE+Component=Proteina+FLX=F22
 Proteine tot.,Proteina,N+Analita+CORE+Component=Proteina+FLX=F12
 Proteina S100,N+Analita+CORE+Component=Proteina S100 B calcio legante+FLX=F884
 Proteinemia,Proteina,N+Analita+Component=Proteina+System=Sangue art+FLX=F6
 Proteinemia,Proteina,N+Analita+Component=Proteina+System=Sangue capill+FLX=F6
 Proteinemia,Proteina,N+Analita+Component=Proteina+System=Sangue cordone ombelicale+FLX=F6
 Proteinemia,Proteina,N+Analita+Component=Proteina+System=Sangue ven+FLX=F6
 Proteinemia,Proteina,N+Analita+Component=Proteina+System=Sangue+FLX=F6
 Proteinemia,Proteina,N+Analita+Component=Proteina+System=Siero/Plasma+FLX=F6
 Proteinemia,Proteina,N+Analita+Component=Proteina+System=Siero+FLX=F6
 Proteinuria,Proteina,N+Analita+Component=Proteina+System=Urine+FLX=F6
 Rapporto A/G,Albumina/Globulina,N+Analita+CORE+Component=Albumina/Globulina+FLX=F7
 Sodiemia,Sodio,N+Analita+Component=Sodio+System=Sangue art+FLX=F6
 Sodiemia,Sodio,N+Analita+Component=Sodio+System=Sangue capill+FLX=F6

Sodiemia,Sodio,N+Analita+Component=Sodio+System=Sangue cordone ombelicale+FLX=F6
 Sodiemia,Sodio,N+Analita+Component=Sodio+System=Sangue ven+FLX=F6
 Sodiemia,Sodio,N+Analita+Component=Sodio+System=Siero/Plasma+FLX=F6
 Sodiemia,Sodio,N+Analita+Component=Sodio+System=Siero+FLX=F6
 Sodio,N+Analita+CORE+Component=Sodio+FLX=F1
 Sodiuria,Sodio,N+Analita+Component=Sodio+System=Urine+FLX=F6
 Solfoemoglobina % Hb
 totale,Solfoemoglobina,N+Analita+CORE+Component=Solfoemoglobina/Emoglobina.totale+FLX=F22
 Solfoemoglobina,N+Analita+CORE+Component=Solfoemoglobina+FLX=F22
 Solfoemoglobinemia,Solfoemoglobina,N+Analita+Component=Solfoemoglobina+System=Sangue art+FLX=F6
 Solfoemoglobinemia,Solfoemoglobina,N+Analita+Component=Solfoemoglobina+System=Sangue capill+FLX=F6
 Solfoemoglobinemia,Solfoemoglobina,N+Analita+Component=Solfoemoglobina+System=Sangue cordone
 ombelicale+FLX=F6
 Solfoemoglobinemia,Solfoemoglobina,N+Analita+Component=Solfoemoglobina+System=Sangue ven+FLX=F6
 Solfoemoglobinemia,Solfoemoglobina,N+Analita+Component=Solfoemoglobina+System=Sangue+FLX=F6
 Solfoemoglobinemia,Solfoemoglobina,N+Analita+Component=Solfoemoglobina+System=Siero/Plasma+FLX=F6
 Striscio di sangue,N+Analita+CORE+Component=Striscio di sangue osservazione+FLX=F7
 Tiocianato,N+Analita+CORE+Component=Tiocianato+FLX=F78
 Tiocianato,N+Analita+CORE+Component=Tiocianato/Creatinina+FLX=F78
 Trigliceridi,N+Analita+CORE+Component=Trigliceridi+FLX=F77
 Urobilinogeno,N+Analita+CORE+Component=Urobilinogeno+FLX=F7
 Vitamina D,N+Analita+CORE+Component=Vitamina D+FLX=F6
 Coagulazione indotta del fattore tissutale,N+Analita+CORE+Component=Coagulazione indotta del fattore
 tissutale.INR+FLX=F6
 BCR-ABL1 e1a1,N+Analita+CORE+Component=BCR-ABL1 e1a1 proteina di fusione+System=Sangue/Tess+FLX=F6
 BCR-ABL1 e1a2,N+Analita+CORE+Component=BCR-ABL1 e1a2 proteina di fusione+System=Sangue/Tess+FLX=F6
 BCR/ABL1 b2a2,N+Analita+CORE+Component=BCR-ABL1 b2a2 proteina di fusione+System=Sangue/Tess+FLX=F6
 BCR-ABL1 b3a2,N+Analita+CORE+Component=BCR-ABL1 b3a2 proteina di fusione+System=Sangue/Tess+FLX=F6
 BCR/ABL e1a2,BCR-ABL1 e1a2,N+Analita+CORE+Component=BCR-ABL1 e1a2 proteina di
 fusione+System=Sangue/Tess+FLX=F6
 BCR/ABL b2a2,BCR/ABL1 b2a2,N+Analita+CORE+Component=BCR/ABL1 b2a2 proteina di
 fusione+System=Sangue/Tess+FLX=F6
 Basofili %,Basofili,N+Analita+EMO+GB+Component=Basofili/100 leucociti+FLX=F788
 Basofili%,Basofili,N+Analita+EMO+GB+Component=Basofili/100 leucociti+FLX=F788
 Basofilo,N+Analita+EMO+GB+Component=Basofili+FLX=F78
 Baso,Basofili,N+Analita+EMO+GB+Component=Basofili+FLX=F78
 Emocromo,N+Analita+EMO+Component=Emocromo+FLX=F7
 Eosinofilo,N+Analita+EMO+GB+Component=Eosinofili+FLX=F78
 % Eosinofili,Eosinofili,N+Analita+EMO+GB+Component=Eosinofili/100 leucociti+FLX=F78
 % Eosinofili,Eosinofili,N+Analita+EMO+GB+Component=Eosinofili/100 leucociti+FLX=F78
 Eosinofili%,Eosinofili,N+Analita+EMO+GB+Component=Eosinofili/100 leucociti+FLX=F78
 Eosinofili %,Eosinofili,N+Analita+EMO+GB+Component=Eosinofili/100 leucociti+FLX=F78
 Eosi,Eosinofili,N+Analita+EMO+GB+Component=Eosinofili+FLX=F7
 Eosi%,Eosinofili,N+Analita+EMO+GB+Component=Eosinofili/100 leucociti+FLX=F7
 Eosi %,Eosinofili,N+Analita+EMO+GB+Component=Eosinofili/100 leucociti+FLX=F7

Globulo bianco,Leucocita,N+Analita+SIN+EMO+GB+Component=Leucociti+FLX=F15
 Grandi cellule non colorate,N+Analita+EMO+GB+Component=Grandi cellule non colorate+FLX=F12
 Granulocita basofilo,Basofilo,N+Analita+SIN+EMO+GB+Component=Basofili+FLX=F11
 Granulocita eosinofilo,Eosinofilo,N+Analita+SIN+EMO+GB+Component=Eosinofili+FLX=F11
 Granulocita neutrofilo,Neutrofilo,N+Analita+SIN+EMO+GB+Component=Neutrofilo+FLX=F11
 Granulocita,N+Analita+EMO+GB+Component=Granulociti+FLX=F10
 Leucocita,N+Analita+EMO+GB+Component=Leucociti+FLX=F10
 Linfocita,N+Analita+EMO+GB+Component=Linfociti+FLX=F78
 Linfociti%,Linfociti,N+Analita+EMO+GB+Component=Linfociti/100 leucociti+FLX=F78
 Linfociti %,Linfociti,N+Analita+EMO+GB+Component=Linfociti/100 leucociti+FLX=F78
 Linf%,Linfociti,N+Analita+EMO+GB+Component=Linfociti/100 leucociti+FLX=F78
 Linf %,Linfociti,N+Analita+EMO+GB+Component=Linfociti/100 leucociti+FLX=F78
 Linf,Linfociti,N+Analita+EMO+GB+Component=Linfociti+FLX=F78
 #Globuli bianchi,Leucociti,N+Analita+SIN+EMO+GB+Component=Leucociti+FLX=F15
 Linfociti totali %,Linfociti,N+Analita+EMO+GB+Component=Linfociti/100 leucociti+FLX=F78
 Linfociti totali%,Linfociti,N+Analita+EMO+GB+Component=Linfociti/100 leucociti+FLX=F78
 Monocita,N+Analita+EMO+GB+Component=Monociti+FLX=F78
 % Monociti,Monociti,N+Analita+EMO+GB+Component=Monociti/100 leucociti+FLX=F78
 % Monociti,Monociti,N+Analita+EMO+GB+Component=Monociti/100 leucociti+FLX=F78
 Monociti%,Monociti,N+Analita+EMO+GB+Component=Monociti/100 leucociti+FLX=F78
 Monociti %,Monociti,N+Analita+EMO+GB+Component=Monociti/100 leucociti+FLX=F78
 Neutrofilo,N+Analita+EMO+GB+Component=Neutrofilo+FLX=F78
 % Neutrofilo,Neutrofilo,N+Analita+EMO+GB+Component=Neutrofilo/100 leucociti+FLX=F78
 % Neutrofilo,Neutrofilo,N+Analita+EMO+GB+Component=Neutrofilo/100 leucociti+FLX=F78
 Neutrofilo%,Neutrofilo,N+Analita+EMO+GB+Component=Neutrofilo/100 leucociti+FLX=F78
 Neutrofilo %,Neutrofilo,N+Analita+EMO+GB+Component=Neutrofilo/100 leucociti+FLX=F78
 Neut%,Neutrofilo,N+Analita+EMO+GB+Component=Neutrofilo/100 leucociti+FLX=F78
 Neut %,Neutrofilo,N+Analita+EMO+GB+Component=Neutrofilo/100 leucociti+FLX=F78
 Neut,Neutrofilo,N+Analita+EMO+GB+Component=Neutrofilo+FLX=F78
 Ampiezza distribuzione emoglobina,N+Analita+EMO+GR+Component=Emoglobina ampiezza di distribuzione+FLX=F6
 Ampiezza distribuzione globuli rossi,N+Analita+EMO+GR+Component=Eritrociti ampiezza di distribuzione+FLX=F6
 Concentrazione emoglobinica corpuscolare media,N+Analita+EMO+GR+Component=Eritrociti concentrazione emoglobinica corpuscolare media+FLX=F6
 Ematocrito,N+Analita+EMO+GR+Component=Ematocrito+FLX=F7
 Emazia,Eritrocita,N+Analita+SIN+EMO+GR+Component=Eritrociti+FLX=F88
 Emazie,Eritrociti,N+Analita+SIN+EMO+GR+Component=Eritrociti+FLX=F88
 Emoglobina corpuscolare media,N+Analita+EMO+GR+Component=Eritrociti emoglobina corpuscolare media+FLX=F6
 Eritrocita,N+Analita+EMO+GR+Component=Eritrociti+FLX=F10
 Globulo rosso,Eritrocita,N+Analita+SIN+EMO+GR+Component=Eritrociti+FLX=F11
 Globuli rossi,Eritrociti,N+Analita+SIN+EMO+GR+Component=Eritrociti+FLX=F11
 Eritrociti volume corpuscolare medio,N+Analita+EMO+GR+Component=Eritrociti volume corpuscolare medio+FLX=F7
 Volume globuli medio,Eritrociti volume corpuscolare medio,N+Analita+EMO+GR+Component=Eritrociti volume corpuscolare medio+FLX=F7
 Eritrociti velocità di sedimentazione,N+Analita+EMO+GR+Component=Eritrociti velocità di sedimentazione
 Ab anti piastrine,Ab piastrine,N+Analita+EMO+PIA+Component=Piastrina Ab+FLX=F7

Ab piastrine,N+Analita+EMO+PIA+Component=Piastrina Ab+FLX=F7
 Aggregazione piastrinica,N+Analita+EMO+PIA+Component=Aggregazione piastrinica.adenosina difosfato indotta+FLX=F7
 Conta piastrinica,Piastrine,N+Analita+EMO+PIA+Component=Piastrine+FLX=F6
 Distribuzione piastrinica ampiezza,N+Analita+EMO+PIA+Component=Distribuzione piastrinica ampiezza+FLX=F6
 Piastrine volume medio,N+Analita+EMO+PIA+Component=Piastrine volume medio+FLX=F7
 Piastrine,N+Analita+EMO+PIA+Component=Piastrine+FLX=F12
 Piastrinocrito,N+Analita+EMO+PIA+Component=Piastrinocrito+FLX=F7
 Volume medio piastrine,Piastrine volume medio,N+Analita+EMO+PIA+Component=Piastrine volume medio+FLX=F7
 Citrato,N+Analita+AC+Component=Citrato+FLX=F78
 Acido citrico,Citrato,N+Analita+AC+Component=Citrato+FLX=F11
 Folato,N+Analita+AC+Component=Folato+FLX=F78
 Acido folico,Folato,N+Analita+AC+Component=Folato+FLX=F11
 Fosfato,N+Analita+AC+Component=Fosfato+FLX=F78
 Acido fosforico,Fosfato,N+Analita+AC+Component=Fosfato+FLX=F11
 Ippurato,N+Analita+AC+Component=Ippurato+FLX=F78
 Acido ippurico,Ippurato,N+Analita+AC+Component=Ippurato+FLX=F11
 Lattato,N+Analita+AC+Component=Lattato+FLX=F78
 Acido lattico,Lattato,N+Analita+AC+Component=Lattato+FLX=F11
 Metilippurato,N+Analita+AC+Component=Metilippurato+FLX=F78
 Metilippurato,N+Analita+AC+Component=Metilippurato/Creatinina+FLX=F78
 Acido metilippurico,Metilippurato,N+Analita+AC+Component=Metilippurato+FLX=F11
 Acido metilippurico,Metilippurato,N+Analita+AC+Component=Metilippurato/Creatinina+FLX=F11
 Ac metilippurico,Metilippurato,N+Analita+AC+Component=Metilippurato+FLX=F11
 Ac metilippurico,Metilippurato,N+Analita+AC+Component=Metilippurato/Creatinina+FLX=F11
 Omovanillato,N+Analita+AC+Component=Omovanillato+FLX=F78
 Acido omovanillico,Omovanillato,N+Analita+AC+Component=Omovanillato+FLX=F11
 Acido para aminoippurico,Para aminoippurato,N+Analita+AC+Component=Para aminoippurato+FLX=F11bis
 Para aminoippurato,N+Analita+AC+Component=Para aminoippurato+FLX=F78
 Acido solforico,Solfato,N+Analita+AC+Component=Solfato+FLX=F11
 Solfato,N+Analita+AC+Component=Solfato+FLX=F78
 Urato,N+Analita+AC+Component=Urato+FLX=F78
 Acido urico,Urato,N+Analita+AC+Component=Urato+FLX=F11
 Ac. urico,Urato,N+Analita+AC+Component=Urato+FLX=F7
 Acido vanilmandelico,Vanillilmandelato,N+Analita+AC+Component=Vanillilmandelato+FLX=F11
 Vanillilmandelato,N+Analita+AC+Component=Vanillilmandelato+FLX=F78
 Uricemia,Urato,N+Analita+AC+Component=Urato+System=Sangue art+FLX=F6
 Uricemia,Urato,N+Analita+AC+Component=Urato+System=Sangue capill+FLX=F6
 Uricemia,Urato,N+Analita+AC+Component=Urato+System=Sangue cordone ombelicale+FLX=F6
 Uricemia,Urato,N+Analita+AC+Component=Urato+System=Sangue ven+FLX=F6
 Uricemia,Urato,N+Analita+AC+Component=Urato+System=Siero/Plasma+FLX=F6
 Uricemia,Urato,N+Analita+AC+Component=Urato+System=Siero+FLX=F6
 Uricuria,Urato,N+Analita+AC+Component=Urato+System=Urine+FLX=F6
 Adrenalina,Epinefrina,N+Analita+SIN+COMP+Component=Epinefrina+FLX=F6
 Emoglobina Fetale,Emoglobina F,N+Analita+SIN+COMP+Component=Emoglobina F+FLX=F14
 Emoglobina glicata,Emoglobina A1c,N+Analita+SIN+EMO+GR+Component=Emoglobina A1c+FLX=F13

Emoglobina glicosilata,Emoglobina A1c,N+Analita+SIN+EMO+GR+Component=Emoglobina A1c+FLX=F13
 Emoglobina glicata,Emoglobina A1c,N+Analita+SIN+EMO+GR+Component=Emoglobina
 A1c/Emoglobina.totale+FLX=F13
 Emoglobina glicosilata,Emoglobina A1c,N+Analita+SIN+EMO+GR+Component=Emoglobina
 A1c/Emoglobina.totale+FLX=F13
 Kaliuria,Potassuria,N+Analita+SIN+COMP+Component=Potassio+System=Urine+FLX=F6
 Natriuria,Sodiuria,N+Analita+SIN+COMP+Component=Sodio+System=Urine+FLX=F6
 Noradrenalina,Norepinefrina,N+Analita+SIN+COMP+Component=Norepinefrina+FLX=F6
 Somatomedina C,Fattore di crescita insulino simile I,N+Analita+SIN+COMP+Component=Fattore di crescita insulino-simile
 I+FLX=F6
 Som-c,Fattore di crescita insulino simile I,N+Analita+SIN+COMP+Component=Fattore di crescita insulino-simile
 I+FLX=F6
 Vitamina D3,Colecalciferolo,N+Analita+SIN+COMP+Component=Colecalciferolo+FLX=F6
 Ovotransferrina,Conalbumina,N+Analita+SIN+COMP+Component=Conalbumina Ab.IgE+FLX=F22
 Alt,Alanina aminotransferasi,N+Analita+Acronimo+COMP+Component=Alanina aminotransferasi
 Ap hb,Aptoglobina,N+Analita+Acronimo+COMP+Component=Aptoglobina
 Apha,Aptoglobina,N+Analita+Acronimo+COMP+Component=Aptoglobina
 Ap-hb,Aptoglobina,N+Analita+Acronimo+COMP+Component=Aptoglobina
 Ast,Aspartato aminotransferasi,N+Analita+COMP+Acronimo+Component=Aspartato aminotransferasi
 Cohb,Carbossiemoglobina,N+Analita+Acronimo+COMP+Component=Carbossiemoglobina
 Cohb,Carbossiemoglobina,N+Analita+Acronimo+COMP+Component=Carbossiemoglobina/Emoglobina.totale
 E2,Estradiolo,N+Analita+Acronimo+COMP+Component=Estradiolo
 Fsh,Follitropina,N+Analita+Acronimo+COMP+Component=Follitropina
 Cohb,Carbossiemoglobina,N+Analita+Acronimo+COMP+Component=Carbossiemoglobina
 G6pdh,Glucosio-6-Fosfato deidrogenasi,N+Analita+Acronimo+COMP+Component=Glucosio-6-Fosfato deidrogenasi
 Got,Aspartato aminotransferasi,N+Analita+Acronimo+COMP+Component=Aspartato aminotransferasi
 GPT,Alanina aminotransferasi,N+Analita+Acronimo+COMP+Component=Alanina aminotransferasi
 Hb a1,Emoglobina A1,N+Analita+Acronimo+COMP+Component=Emoglobina A1/Emoglobina.totale
 Hb a1c,Emoglobina A1c,N+Analita+Acronimo+COMP+Component=Emoglobina A1c
 #Hb a1c,Emoglobina A1c,N+Analita+Acronimo+COMP+Component=Emoglobina A1c/Emoglobina.totale
 Hb a2,Emoglobina A2,N+Analita+Acronimo+COMP+Component=Emoglobina A2
 Hb a2,Emoglobina A2,N+Analita+Acronimo+COMP+Component=Emoglobina A2/Emoglobina.totale
 Hb f,Emoglobina Fetale,N+Analita+Acronimo+COMP+Component=Emoglobina f
 Hb,Emoglobina,N+Analita+Acronimo+COMP+Component=Emoglobina
 Hba1,Emoglobina A1,N+Analita+Acronimo+COMP+Component=Emoglobina A1/Emoglobina.totale
 Hba1c,Emoglobina A1c,N+Analita+Acronimo+COMP+Component=Emoglobina A1c
 Hba1c,Emoglobina A1c,N+Analita+Acronimo+COMP+Component=Emoglobina A1c/Emoglobina.totale
 Hba2,Emoglobina A2,N+Analita+Acronimo+COMP+Component=Emoglobina A2
 Hba2,Emoglobina A2,N+Analita+Acronimo+COMP+Component=Emoglobina A2/Emoglobina.totale
 Hba2 %,Emoglobina A2,N+Analita+Acronimo+COMP+Component=Emoglobina A2/Emoglobina.totale
 Hba2%,Emoglobina A2,N+Analita+Acronimo+COMP+Component=Emoglobina A2/Emoglobina.totale
 Hbf,Emoglobina Fetale,N+Analita+Acronimo+COMP+Component=Emoglobina f
 HE 4,Epididimo umano proteina 4,N+Acronimo+COMP+Component=Epididimo umano proteina 4
 HE4,Epididimo umano proteina 4,N+Acronimo+COMP+Component=Epididimo umano proteina 4
 HE-4,Epididimo umano proteina 4,N+Acronimo+COMP+Component=Epididimo umano proteina 4

HE-4 ovarian cancer marker,Epididimo umano proteina 4,N+Acronimo+COMP+Component=Epididimo umano proteina 4
Hg,Mercurio,N+Analita+Acronimo+COMP+Component=Mercurio
Hg,Mercurio,N+Analita+Acronimo+COMP+Component=Mercurio/Creatinina
Hgb,Emoglobina A1c,N+Analita+Acronimo+COMP+Component=Emoglobina A1c
Hhb,Deossioglobina,N+Analita+Acronimo+COMP+Component=Deossioglobina
Hhb,Deossioglobina,N+Analita+Acronimo+COMP+Component=Deossioglobina/Emoglobina.totale
K,Potassio,N+Analita+Acronimo+COMP+Component=Potassio
Ldh 1,Lattato deidrogenasi 1,N+Analita+Acronimo+COMP+Component=Lattato deidrogenasi 1
Ldh 1,Lattato deidrogenasi 1,N+Analita+Acronimo+COMP+Component=Lattato deidrogenasi 1/Lattato deidrogenasi 2
Ldh 1,Lattato deidrogenasi 1,N+Analita+Acronimo+COMP+Component=Lattato deidrogenasi 1/Lattato deidrogenasi.totale
Ldh 2,Lattato deidrogenasi 2,N+Analita+Acronimo+COMP+Component=Lattato deidrogenasi 2
Ldh 2,Lattato deidrogenasi 2,N+Analita+Acronimo+COMP+Component=Lattato deidrogenasi 2/Lattato deidrogenasi.totale
Ldh 3,Lattato deidrogenasi 3,N+Analita+Acronimo+COMP+Component=Lattato deidrogenasi 3
Ldh 3,Lattato deidrogenasi 3,N+Analita+Acronimo+COMP+Component=Lattato deidrogenasi 3/Lattato deidrogenasi.totale
Ldh 4,Lattato deidrogenasi 4,N+Analita+Acronimo+COMP+Component=Lattato deidrogenasi 4
Ldh 4,Lattato deidrogenasi 4,N+Analita+Acronimo+COMP+Component=Lattato deidrogenasi 4/Lattato deidrogenasi.totale
Ldh 5,Lattato deidrogenasi 5,N+Analita+Acronimo+COMP+Component=Lattato deidrogenasi 5
Ldh 5,Lattato deidrogenasi 5,N+Analita+Acronimo+COMP+Component=Lattato deidrogenasi 5/Lattato deidrogenasi.totale
Ldh,Lattato deidrogenasi,N+Analita+Acronimo+COMP+Component=Lattato deidrogenasi
Ldh isoenzimi,Lattato deidrogenasi isoenzimi,N+Analita+Acronimo+COMP+Component=Lattato deidrogenasi isoenzimi
Ldh1,Lattato deidrogenasi 1,N+Analita+Acronimo+COMP+Component=Lattato deidrogenasi 1
Ldh2,Lattato deidrogenasi 2,N+Analita+Acronimo+COMP+Component=Lattato deidrogenasi 2
Ldh3,Lattato deidrogenasi 3,N+Analita+Acronimo+COMP+Component=Lattato deidrogenasi 3
Ldh4,Lattato deidrogenasi 4,N+Analita+Acronimo+COMP+Component=Lattato deidrogenasi 4
Ldh5,Lattato deidrogenasi 5,N+Analita+Acronimo+COMP+Component=Lattato deidrogenasi 5
Lh,Lutropina,N+Analita+Acronimo+COMP+Component=Lutropina
Methb,Metaemoglobina,N+Analita+Acronimo+COMP+Component=Metaemoglobina
Methb,Metaemoglobina,N+Analita+Acronimo+COMP+Component=Metaemoglobina/Emoglobina.totale
Na,Sodio,N+Analita+Acronimo+COMP+Component=Sodio
Oh,Idrossido,N+Analita+Acronimo+COMP+Component=Idrossido
17 oh,17-Idrossiprogesterone,N+Analita+Acronimo+COMP+Component=17-Idrossiprogesterone
P,Fosfato,N+Analita+Acronimo+COMP+Component=Fosfato
Pt,Proteina,N+Analita+Acronimo+COMP+Component=Proteina
Shb,Solfoemoglobina,N+Analita+Acronimo+COMP+Component=Solfoemoglobina
Amp,Adenosina monofosfato,N+Analita+Acronimo+COMP+Component=Adenosina monofosfato.ciclico
Iaa,Ab Insulina,N+Analita+Acronimo+COMP+Component=Insulina Ab
S100,Proteina S100,N+Analita+Acronimo+COMP+Component=Proteina S100 B calcio legante
INR,Coagulazione indotta del fattore tissutale,N+Analita+Acronimo+COMP+Component=Coagulazione indotta del fattore
tissutale.INR
Gpt (alt),Alanina aminotransferasi,N+Analita+Acronimo+COMP+Component=Alanina aminotransferasi
Got (ast),Aspartato aminotransferasi,N+Analita+Acronimo+COMP+Component=Aspartato aminotransferasi
F76,Lattoalbumina alfa,N+Analita+Acronimo+COMP+Component=Lattoalbumina alfa Ab.IgE+FLX=F22
F76,Lattoalbumina alfa,N+Analita+Acronimo+COMP+Component=Lattoalbumina alfa Ab.IgA+FLX=F22
F76,Lattoalbumina alfa,N+Analita+Acronimo+COMP+Component=Lattoalbumina alfa Ab.IgG+FLX=F22
F76,Lattoalbumina alfa,N+Analita+Acronimo+COMP+Component=Lattoalbumina alfa Ab.Ig4+FLX=F22

F76,Lattoalbumina alfa,N+Analita+Acronimo+COMP+Component=Lattoalbumina alfa Ag+FLX=F22
BASO,Basofili,N+Analita+Acronimo+EMO+Component=Basofili
EOS,Eosinofili,N+Analita+Acronimo+EMO+Component=Eosinofili
Luc %,Grandi cellule non colorate,N+Analita+Acronimo+EMO+Component=Grandi cellule non colorate/100 leucociti
Luc%,Grandi cellule non colorate,N+Analita+Acronimo+EMO+Component=Grandi cellule non colorate/100 leucociti
Luc,Grandi cellule non colorate,N+Analita+Acronimo+EMO+Component=Grandi cellule non colorate
LYMP,Linfociti,N+Analita+Acronimo+EMO+Component=Linfociti
MONO,Monociti,N+Analita+Acronimo+EMO+Component=Monociti
NEUT,Neutrofilii,N+Analita+Acronimo+EMO+Component=Neutrofilii
WBC,Leucociti,N+Analita+Acronimo+EMO+Component=Leucociti
HCT,Ematocrito,N+Analita+Acronimo+EMO+Component=Ematocrito
HDW,Ampiezza distribuzione emoglobina,N+Analita+Acronimo+EMO+Component=Emoglobina ampiezza di distribuzione
MCH,Emoglobina corpuscolare media,N+Analita+Acronimo+EMO+Component=Eritrociti emoglobina corpuscolare media
MCHC,Concentrazione emoglobinica corpuscolare media,N+Analita+Acronimo+EMO+Component=Eritrociti
concentrazione emoglobinica corpuscolare media
MCV,Eritrociti volume corpuscolare medio,N+Acronimo+EMO+Component=Eritrociti volume corpuscolare medio
RBC,Eritrociti,N+Analita+Acronimo+EMO+Component=Eritrociti
RDW,Eritrociti ampiezza distribuzione,N+Analita+Acronimo+EMO+Component=Eritrociti ampiezza distribuzione
VES,Eritrociti velocità di sedimentazione,N+Analita+Acronimo+EMO+Component=Eritrociti velocità di sedimentazione
V.E.S.,Eritrociti velocità di sedimentazione,N+Analita+Acronimo+EMO+Component=Eritrociti velocità di sedimentazione
1 ora VES,Eritrociti velocità di sedimentazione,N+Analita+Acronimo+EMO+Component=Eritrociti velocità di
sedimentazione
2 ora VES,Eritrociti velocità di sedimentazione,N+Analita+Acronimo+EMO+Component=Eritrociti velocità di
sedimentazione
Mpv,Piastrine volume medio,N+Analita+Acronimo+EMO+Component=Piastrine volume medio
Pct,Piastrinocrito,N+Analita+Acronimo+EMO+Component=Piastrinocrito
Pdw,Distribuzione piastrinica ampiezza,N+Analita+Acronimo+EMO+Component=Distribuzione piastrinica ampiezza
Pdv,Distribuzione piastrinica ampiezza,N+Analita+Acronimo+EMO+Component=Distribuzione piastrinica ampiezza
Plt,Piastrine,N+Analita+Acronimo+EMO+Component=Piastrine

Annexe 2 – Dictionnaire « Elements_LOINC.dic »

Sangue,N+Campione+System=Urine+FLX=F7
Plasma,N+Campione+System=Plasma+FLX=F7
Feci,N+Campione+System=Feci+FX=F6
Sangue intero,N+Campione+System=Sangue+FLX=F7
Siero plasma,N+Campione+System=Siero/Plasma+FLX=F7
Siero o plasma,Siero plasma,N+Campione+System=Siero/Plasma+FLX=F7
Ematologia,N+Campione+System=Sangue+FLX=F6
Coagulazione,N+Campione+System=PPP+FLX=F6
Sangue midollare,N+Campione+FLX=F7
Cromatografia liquida HPLC,N+Metodo+Method=X+FLX=F6
Enzimatica ureasi,N+Metodo+Method=X+FLX=F6
Cromatografico - colorimetrico,N+Metodo+Method=X+FLX=F7
Cromatografico,N+Metodo+Method=X+FLX=F7
Cinetico IFCC,N+Metodo+Method=X+FLX=F7
SA-4220 MENARINI,N+Metodo+Method=X+FLX=F7
Cresolftalaina,N+Metodo+Method=X+FLX=F6
Colorimetrico,N+Metodo+Method=X+FLX=F7
Enzimatico diretto,N+Metodo+Method=X+FLX=F7
GPO-PAP,N+Metodo+Method=X+FLX=F7
Cinetico Jaffe',N+Metodo+Method=X+FLX=F7
Chemiluminescenza,N+Metodo+Method=X+FLX=F6
Assorbimento atomico,N+Metodo+Method=X+FLX=F7
ADVIA 120 BAYER,N+Metodo+Method=X+FLX=F7
Carbossimetrico,N+Metodo+Method=X+FLX=F7
Immunoenzimatico EIA,N+Metodo+Method=X+FLX=F7
Immunoenzimatico EIA IGEA,N+Metodo+Method=X+FLX=F7
Molibdato,N+Metodo+Method=X+FLX=F7
Radioimmunologico RIA,N+Metodo+Method=X+FLX=F7
Fotometria a fiamma,N+Metodo+Method=X+FLX=F7
Radioimmunologico RIA,N+Metodo+Method=X+FLX=F7
Elettroforetico,N+Metodo+Method=X+FLX=F7
Agglutinazione,N+Metodo+Method=X+FLX=F6
Nefelometria,N+Metodo+Method=X+FLX=F6
Immunodiffusione,N+Metodo+Method=X+FLX=F6

Microscopico,N+Metodo+Method=X+FLX=F7
 IF test,N+Metodo+Method=X+FLX=F7
 Coagulativo,N+Metodo+Method=X+FLX=F7
 HPLC,N+Metodo+Method=X+FLX=F7
 Colorimetrico con rilevazione spettrofotometrica,N+Metodo+Method=X+FLX=F7
 Nefelometrico,N+Metodo+Method=X+FLX=F7
 Turbidimetrico,N+Metodo+Method=X+FLX=F7
 Immunoenzimatico con rilevazione in chemiluminescenza,N+Metodo+Method=X+FLX=F7
 RIA,N+Metodo+Method=X+FLX=F7
 IRMA,N+Metodo+Method=X+FLX=F7
 CITOMETRIA A FLUSSO,N+Metodo+Method=X+FLX=F6
 Rt-PCR REAZIONE A CATENA DELLA POLIMERASI,N+Metodo+Method=X+FLX=F6
 Enzimatico colorimetrico,N+Metodo+Method=X+FLX=F7
 Enzimatico,N+Metodo+Method=X+FLX=F7
 Potenziometrico,N+Metodo+Method=X+FLX=F7
 Colture,N+Metodo+Method=X+FLX=F6
 Numerico,N+Risultato+Scale=Qn+FLX=F78
 Testo,N+Risultato+Scale=Nar+FLX=F78
 Codificato,N+Risultato+Scale=Ord+FLX=F78
 Num,N+Risultato+Scale=Qn+FLX=F78
 Positivo/negativo in percentuale,N+Risultato+FLX=F7
 Positivo/negativo,N+Risultato+FLX=F78
 Tec,N+Risultato
 Mg/l,N+UnMis+Property=MCnc+Timing=Pt+Scale=Qn
 Mg%,N+UnMis+Property=MCnc+Timing=Pt+Scale=Qn
 g/dl,N+UnMis+Property=MCnc+Timing=Pt+Scale=Qn
 mg/dl,N+UnMis+Property=MCnc+Timing=Pt+Scale=Qn
 mg/24h,N+UnMis+Property=MRat+Timing=24h+Scale=Qn
 U/l,N+UnMis+Property=CCnc+Timing=Pt+Scale=Qn
 g/l,N+UnMis+Property=MCnc+Timing=Pt+Scale=Qn
 KU/L,N+UnMis+Property=ACnc+Timing=Pt+Scale=Qn
 g/24h,N+UnMis+Property=MRat+Timing=24h+Scale=Qn
 #mg,N+UnMis+Property=Massa+Timing=Pt+Scale=Qn
 ml/min,N+UnMis+Property=VRat+Timing=Pt+Scale=Qn
 ml/min,N+UnMis+Property=ArVRat+Timing=Pt+Scale=Qn
 g/24,N+UnMis+Property=MRat+Timing=24h+Scale=Qn
 g/24h,N+UnMis+Property=MRat+Timing=24h+Scale=Qn
 g/24 h,g/24h,N+UnMis+Property=MRat+Timing=24h+Scale=Qn
 nM/mcrM crea,N+UnMis+Property=MRto+Timing=Pt+Scale=Qn
 mcg/gr creat,N+UnMis+Property=MRto+Timing=Pt+Scale=Qn

pM/mcrM crea,N+UnMis+Property=MRto+Timing=Pt+Scale=Qn
 mg/gr creat,N+UnMis+Property=MRto+Timing=Pt+Scale=Qn
 g/g cr,N+UnMis+Property=MRto+Timing=Pt+Scale=Qn
 pg/ml,N+UnMis+Property=MCnc+Timing=Pt+Scale=Qn
 ng/l,N+UnMis+Property=MCnc+Timing=Pt+Scale=Qn
 mcg/24,N+UnMis+Property=MRat+Timing=24h+Scale=Qn
 mcg/24h,N+UnMis+Property=MRat+Timing=24h+Scale=Qn
 ng/ml,N+UnMis+Property=MCnc+Timing=Pt+Scale=Qn
 mUI/ml,N+UnMis+Property=CCnc+Timing=Pt+Scale=Qn
 UI/24h,N+UnMis+Property=ARat+Timing=24h+Scale=Qn
 mcg/l,N+UnMis+Property=MCnc+Timing=Pt+Scale=Qn
 mU/mL,N+UnMis+Property=CCnc+Timing=Pt+Scale=Qn
 U/I,N+UnMis+Property=CCnc+Timing=Pt+Scale=Qn
 mcU/ml,N+UnMis+Property=ACnc+Timing=Pt+Scale=Qn
 mIU/ml,N+UnMis+Property=ACnc+Timing=Pt+Scale=Qn
 mEq/l,N+UnMis+Property=SCnc+Timing=Pt+Scale=Qn
 mEq/24h,N+UnMis+Property=SRat+Timing=24h+Scale=Qn
 Pmol/l,N+UnMis+Property=SCnc+Timing=Pt+Scale=Qn
 gr/dl,N+UnMis+Property=MCnc+Timing=Pt+Scale=Qn
 mg / dL,mg/dl,N+UnMis+Property=MCnc+Timing=Pt+Scale=Qn
 Ratio,N+UnMis+Property=Rapporto+Timing=Pt+Scale=Qn
 10^3/uL,N+UnMis+Property=NCnc+Timing=Pt+Scale=Qn
 10^3/ml,N+UnMis+Property=NCnc+Timing=Pt+Scale=Qn
 10^3/ul,N+UnMis+Property=NCnc+Timing=Pt+Scale=Qn
 10e3/uL,N+UnMis+Property=NCnc+Timing=Pt+Scale=Qn
 10^6/ml,N+UnMis+Property=NCnc+Timing=Pt+Scale=Qn
 10^6/ul,N+UnMis+Property=NCnc+Timing=Pt+Scale=Qn
 UI/ml,N+UnMis+Property=ACnc+Timing=Pt+Scale=Qn
 nmol/ml,N+UnMis+Property=SCnc+Timing=Pt+Scale=Qn
 pg,N+UnMis+Property=EntMass+Timing=Pt+Scale=Qn
 kUA/l,N+UnMis+Property=ACnc+Timing=Pt+Scale=Qn
 ug/dl,N+UnMis+Property=MCnc+Timing=Pt+Scale=Qn
 mU/mil,N+UnMis+Property=ACnc+Timing=Pt+Scale=Qn
 uUI/mL,N+UnMis+Property=ACnc+Timing=Pt+Scale=Qn
 ug/L,N+UnMis+Property=MCnc++Timing=Pt+Scale=Qn
 mm,N+UnMis+Property=Len+Timing=Pt+Scale=Qn
 mEq/l,N+UnMis+Property=SCnc+Timing=Pt+Scale=Qn
 mmol/l,N+UnMis+Property=SCnc+Timing=Pt+Scale=Qn
 Mmoli/l,N+UnMis+Property=SCnc+Timing=Pt+Scale=Qn
 Fl,N+UnMis+Property=EntVol+Timing=Pt+Scale=Qn

Urinario,A+AgLoinc+System=Urine+FLX=F78
Sierico,A+AgLoinc+System=Siero+FLX=F78
Ionizzato,A+AgLoinc+System=Siero+FLX=F78
Plasmatico,A+AgLoinc+System=Plasma+FLX=F78
Arterioso,A+AgLoinc+System=Sangue art+FLX=F78
704-7,NUM+Code+Component=Basofili+Property=NCnc+Timing=Pt+System=Sangue
+Scale=Qn+Method=X
711-2,NUM+Code+Component=Eosinofili+Property=NCnc+Timing=Pt+System=Sangue
+Scale=Qn+Method=X
30364-4,NUM+Code+Component=Linfociti+Property=NCnc+Timing=Pt+System=Sangue
+Scale=Qn+Method=X
742-7,NUM+Code+Component=Monociti+Property=NCnc+Timing=Pt+System=Sangue
+Scale=Qn+Method=X
751-8,NUM+Code+Component=Neutrofili+Property=NCnc+Timing=Pt+System=Sangue
+Scale=Qn+Method=X
713-8,NUM+Code+Component=Eosinofili/100 leucociti+Property=NFr+Timing=Pt+System=Sangue
+Scale=Qn+Method=X
30365-1,NUM+Code+Component=Linfociti/100 leucociti+Property=NFr+Timing=Pt+System=Sangue
+Scale=Qn+Method=X
5905-5,NUM+Code+Component=Monociti/100 leucociti+Property=NFr+Timing=Pt+System=Sangue
+Scale=Qn+Method=X
706-2,NUM+Code+Component=Basofili/100 leucociti+Property=NFr+Timing=Pt+System=Sangue
+Scale=Qn+Method=X
4537-7,NUM+Code+Component=Eritrociti velocità di sedimentazione+Property=Vel+Timing=Pt
+System=Sangue+Scale=Qn+Method=X
2243-4,NUM+Code+Component=Estradiolo+Property=MCnc+Timing=Pt+System=Siero/Plasma
+Scale=Qn+Method=Absent
1668-3,NUM+Code+Component=17-Idrossiprogesterone+Property=MCnc+Timing=Pt+System=Siero/Plasma
+Scale=Qn+Method=Absent
48340-4,NUM+Code+Component=17-Idrossiprogesterone^30Minuti post XXX stimolo+Property=MCnc
+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=Absent
48341-2,NUM+Code+Component=17-Idrossiprogesterone^1h post XXX stimolo+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
56616-6,NUM+Code+Component=17-Idrossiprogesterone^1.5h post XXX stimolo+Property=MCnc
+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=Absent
27996-8,NUM+Code+Component=17-Idrossiprogesterone^baseline+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
18184-2,NUM+Code+Component=Eritrociti velocità di sedimentazione+Property=Vel+Timing=Pt
+System=Sangue+Scale=Qn+Method=X
2481-0,NUM+Code+Component=Insulina Ab+Property=PrThr+Timing=Pt+System=Siero

+Scale=Ord+Method=Absent
57440-0,NUM+Code+Component=Piastrina Ab+Property=PrThr+Timing=Pt+System=Siero
+Scale=Ord+Method=X
13774-5,NUM+Code+Component=Metilippurato/Creatinina+Property=MRto+Timing=Pt+System=Urine
+Scale=Qn+Method=Absent
3086-6,NUM+Code+Component=Urato+Property=MCnc+Timing=Pt+System=Urine
+Scale=Qn+Method=Absent
6709-0,NUM+Code+Component=Ippurato+Property=MCnc+Timing=Pt+System=Urine
+Scale=Qn+Method=Absent
21387-6,NUM+Code+Component=Metilippurato+Property=MCnc+Timing=Pt+System=Urine
+Scale=Qn+Method=Absent
2724-3,NUM+Code+Component=Para aminoippurato+Property=MCnc+Timing=Pt+System=Urine
+Scale=Qn+Method=Absent
14932-8,NUM+Code+Component=Urato+Property=PrThr+Timing=Pt+System=Calcolo
+Scale=Ord+Method=Absent
3084-1,NUM+Code+Component=Urato+Property=MCnc+Timing=Pt+System=Siero/Plasma
+Scale=Qn+Method=Absent
3122-9,NUM+Code+Component=Vanillilmandelato+Property=MRat+Timing=24h+System=Urine
+Scale=Qn+Method=Absent
2232-7,NUM+Code+Component=Epinefrina+Property=MRat+Timing=24h+System=Urine
+Scale=Qn+Method=Absent
2230-1,NUM+Code+Component=Epinefrina+Property=MCnc+Timing=Pt+System=Plasma
+Scale=Qn+Method=Absent
24377-4,NUM+Code+Component=Aggregazione piastrinica.adenosina difosfato indotta+Property=PrThr
+Timing=Pt+System=PRP+Scale=Ord+Method=Absent
1754-1,NUM+Code+Component=Albumina+Property=MCnc+Timing=Pt+System=Urine
+Scale=Qn+Method=Absent
13980-8,NUM+Code+Component=Albumina/Proteina.totale+Property=MFr+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=X
21059-1,NUM+Code+Component=Albumina+Property=MCnc+Timing=24h+System=Urine
+Scale=Qn+Method=Absent
1751-7,NUM+Code+Component=Albumina+Property=MCnc+Timing=Pt+System=Siero/Plasma
+Scale=Qn+Method=Absent
2862-1,NUM+Code+Component=Albumina+Property=MCnc+Timing=Pt+System=Siero/Plasma
+Scale=Qn+Method=X
13978-2,NUM+Code+Component=Alfa 1 globulina/Proteina.totale+Property=MFr+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=X
13981-6,NUM+Code+Component=Alfa 2 globulina/Proteina.totale+Property=MFr+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=X
7445-0,NUM+Code+Component=Lattoalbumina alfa Ab.IgE+Property=ACnc+Timing=Pt+System=Siero

+Scale=Qn+Method=Absent
1742-6,NUM+Code+Component=Alanina aminotransferasi+Property=CCnC+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
22712-4,NUM+Code+Component=Adenosina monofosfato.ciclico+Property=SCnC+Timing=Pt+System=Urine
+Scale=Qn+Method=Absent
46127-7,NUM+Code+Component=Aptoglobina+Property=MCnC+Timing=Pt+System=Siero/Plasma
+Scale=Qn+Method=X
4542-7,NUM+Code+Component=Aptoglobina+Property=MCnC+Timing=Pt+System=Siero/Plasma
+Scale=Qn+Method=Absent
14979-9,NUM+Code+Component=Coagulazione superficiale indotta+Property=Tempo+Timing=Pt
+System=PPP+Scale=Qn+Method=X
1920-8,NUM+Code+Component=Aspartato aminotransferasi+Property=CCnC+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
3174-0,NUM+Code+Component=Antitrombina+Property=ACnC+Timing=Pt
+System=PPP+Scale=Qn+Method=X
3094-0,NUM+Code+Component=Azoto ureico+Property=MCnC+Timing=Pt+System=Siero/Plasma
+Scale=Qn+Method=Absent
11064-3,NUM+Code+Component=Azoto ureico^post dialisi+Property=MCnC+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
3095-7,NUM+Code+Component=Azoto ureico+Property=MCnC+Timing=Pt
+System=Urine+Scale=Qn+Method=Absent
3096-5,NUM+Code+Component=Azoto ureico+Property=MRat+Timing=24h
+System=Urine+Scale=Qn+Method=Absent
30180-4,NUM+Code+Component=Basofili/100 leucociti+Property=NFr+Timing=Pt
+System=Sangue+Scale=Qn+Method=Absent
49490-6,NUM+Code+Component=BCR/ABL1 b2a2 proteina di fusione+Property=PrThr+Timing=Pt
+System=Sangue/Tess+Scale=Ord+Method=X
55149-9,NUM+Code+Component=BCR-ABL1 e1a2 proteina di fusione+Property=PrThr+Timing=Pt
+System=Sangue/Tess+Scale=Ord+Method=X
32732-0,NUM+Code+Component=Beta 1 globulina/Proteina.totale+Property=MFr+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=X
32733-8,NUM+Code+Component=Beta 2 globulina/Proteina.totale+Property=MFr+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=X
15152-2,NUM+Code+Component=Bilirubina.glucuronidata+Property=MCnC+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
1975-2,NUM+Code+Component=Bilirubina+Property=MCnC+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
1968-7,NUM+Code+Component=Bilirubina.glucuronidata+Bbilirubina.albumina legata+Property=MCnC
+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=Absent
4477-6,NUM+Code+Component=Complemento C1 esterasi inibitore+Property=MCnC+Timing=Pt

+System=Siero/Plasma+Scale=Qn+Method=Absent
4485-9,NUM+Code+Component=Complemento C3+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
4498-2,NUM+Code+Component=Complemento C4+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
17861-6,NUM+Code+Component=Calcio+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
14636-5,NUM+Code+Component=Calcio+Property=PrThr+Timing=Pt
+System=Calcolo+Scale=Ord+Method=Absent
49765-1,NUM+Code+Component=Calcio+Property=MCnc+Timing=Pt
+System=Sangue+Scale=Qn+Method=Absent
18488-7,NUM+Code+Component=Calcio+Property=MCnc+Timing=24h
+System=Urine+Scale=Qn+Method=Absent
17862-4,NUM+Code+Component=Calcio+Property=MCnc+Timing=Pt
+System=Urine+Scale=Qn+Method=Absent
34581-9,NUM+Code+Component=Calcio.ionizzato+Property=SCnc+Timing=Pt
+System=Sangue art+Scale=Qn+Method=Absent
51950-4,NUM+Code+Component=Calcio^post dialisi+Property=MCnc+Timing=Pt+System=Siero/Plasma
+Scale=Qn+Method=Absent
6874-2,NUM+Code+Component=Calcio+Property=MRat+Timing=24h
+System=Urine+Scale=Qn+Method=Absent
12211-9,NUM+Code+Component=Epiteliali cellule+Property=PrThr+Timing=Pt+System=Vaginale
+Scale=Ord+Method=X
71694-4,NUM+Code+Component=Emoglobina+Property=EntMass+Timing=Pt
+System=Reticolociti+Scale=Qn+Method=X
2030-5,NUM+Code+Component=Carbossiemoglobina/Emoglobina.totale+Property=MFr
+Timing=Pt+System=Sangue art+Scale=Qn+Method=Absent
2085-9,NUM+Code+Component=Colesterolo.in HDL+Property=MCnc+Timing=Pt+System=Siero/Plasma
+Scale=Qn+Method=Absent
2093-3,NUM+Code+Component=Colesterolo+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
13457-7,NUM+Code+Component=Colesterolo.in LDL+Property=MCnc+Timing=Pt+System=Siero/Plasma
+Scale=Qn+Method=X
2089-1,NUM+Code+Component=Colesterolo.in LDL+Property=MCnc+Timing=Pt+System=Siero/Plasma
+Scale=Qn+Method=Absent
2091-7,NUM+Code+Component=Colesterolo.in VLDL+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
26515-7,NUM+Code+Component=Piastrine+Property=NCnc+Timing=Pt+System=Sangue
+Scale=Qn+Method=Absent
2160-0,NUM+Code+Component=Creatinina+Property=MCnc+Timing=Pt+System=Siero/Plasma

+Scale=Qn+Method=Absent
38483-4,NUM+Code+Component=Creatinina+Property=MCnc+Timing=Pt+System=Sangue
+Scale=Qn+Method=Absent
2161-8,NUM+Code+Component=Creatinina+Property=MCnc+Timing=Pt+System=Urine
+Scale=Qn+Method=Absent
49004-5,NUM+Code+Component=Creatinina+Property=MCnc+Timing=Pt
+System=Liquido di dialisi peritoneale+Scale=Qn+Method=Absent
11041-1,NUM+Code+Component=Creatinina^post dialisi+Property=MCnc+Timing=Pt+System=Siero/Plasma
+Scale=Qn+Method=Absent
2162-6,NUM+Code+Component=Creatinina+Property=MRat+Timing=24h
+System=Urine+Scale=Qn+Method=Absent
48065-7,NUM+Code+Component=Fibrina D-dimero FEU+Property=MCnc+Timing=Pt+System=PPP
+Scale=Qn+Method=Absent
4551-8,NUM+Code+Component=Emoglobina A2/Emoglobina.totale+Property=MFr+Timing=Pt
+System=Sangue+Scale=Qn+Method=Absent
4544-3,NUM+Code+Component=Ematocrito+Property=VFr+Timing=Pt
+System=Sangue+Scale=Qn+Method=X
20570-8,NUM+Code+Component=Ematocrito+Property=VFr+Timing=Pt
+System=Sangue+Scale=Qn+Method=Absent
32354-3,NUM+Code+Component=Ematocrito+Property=VFr+Timing=Pt
+System=Sangue art+Scale=Qn+Method=Absent
48051-7,NUM+Code+Component=Eritrociti+Property=PrThr+Timing=Pt
+System=Vaginale+Scale=Ord+Method=Absent
50559-4,NUM+Code+Component=Emoglobina+Property=MCnc+Timing=Pt
+System=Urine+Scale=Qn+Method=X
718-7,NUM+Code+Component=Emoglobina+Property=MCnc+Timing=Pt
+System=Sangue+Scale=Qn+Method=Absent
41995-2,NUM+Code+Component=Emoglobina A1c+Property=MCnc+Timing=Pt+System=Sangue
+Scale=Qn+Method=Absent
4548-4,NUM+Code+Component=Emoglobina A1c/Emoglobina.totale+Property=MFr+Timing=Pt
+System=Sangue+Scale=Qn+Method=Absent
26450-7,NUM+Code+Component=Eosinofili/100 leucociti+Property=NFr+Timing=Pt+System=Sangue
+Scale=Qn+Method=Absent
5909-7,NUM+Code+Component=Striscio di sangue osservazione+
Property=Prid+Timing=Pt+System=Sangue+Scale=Nom+Method=X
51499-2,NUM+Code+Component=Estradiolo^1° campione post XXX stimolo
+Property=MCnc+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=Absent
51501-5,NUM+Code+Component=Estradiolo^2° campione post XXX stimolo+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
2258-2,NUM+Code+Component=Estrone+Property=MCnc+Timing=Pt+System=Siero/Plasma

+Scale=Qn+Method=Absent
7556-4,NUM+Code+Component=Ovoalbumina Ab.IgE+Property=ACnc+Timing=Pt+System=Siero
+Scale=Qn+Method=Absent
25615-6,NUM+Code+Component=Conalbumina Ab.IgE+Property=ACnc+Timing=Pt+System=Siero
+Scale=Qn+Method=Absent
3193-0,NUM+Code+Component=Fattore di coagulazione V attività
effettivo/normale+Property=RelTime+Timing=Pt+System=PPP+Scale=Qn+Method=X
3198-9,NUM+Code+Component=Fattore di coagulazione VII attività effettivo/normale+Property=RelTime
+Timing=Pt+System=PPP+Scale=Qn+Method=X
3209-4,NUM+Code+Component=Fattore di coagulazione VIII attività effettivo/normale+Property=RelTime
+Timing=Pt+System=PPP+Scale=Qn+Method=X
3255-7,NUM+Code+Component=Fibrinogeno+Property=MCnc
+Timing=Pt+System=PPP+Scale=Qn+Method=X
17838-4,NUM+Code+Component=Fosfatasi alcalina.osso+Property=MCnc+Timing=Pt+System=Siero/Plasma
+Scale=Qn+Method=Absent
6768-6,NUM+Code+Component=Fosfatasi alcalina+Property=CCnC+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
14878-3,NUM+Code+Component=Fosfato+Property=PrThr+Timing=Pt
+System=Calcolo+Scale=Ord+Method=Absent
2777-1,NUM+Code+Component=Fosfato+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
21458-5,NUM+Code+Component=Fosfato+Property=MCnc+Timing=24h
+System=Urine+Scale=Qn+Method=Absent
2778-9,NUM+Code+Component=Fosfato+Property=MCnc+Timing=Pt
+System=Urine+Scale=Qn+Method=Absent
2779-7,NUM+Code+Component=Fosfato+Property=MRat+Timing=24h
+System=Urine+Scale=Qn+Method=Absent
1779-8,NUM+Code+Component=Fosfatasi alcalina.fegato+Property=CCnC+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
1777-2,NUM+Code+Component=Fosfatasi alcalina.osso+Property=CCnC+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
15067-2,NUM+Code+Component=Follitropina+Property=ACnc
+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=Absent
50477-9,NUM+Code+Component=Follitropina^1° campione+Property=ACnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
50478-7,NUM+Code+Component=Follitropina^2° campione+Property=ACnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
50479-5,NUM+Code+Component=Follitropina^3° campione+Property=ACnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
50480-3,NUM+Code+Component=Follitropina^4° campione+Property=ACnc+Timing=Pt+

System=Siero/Plasma+Scale=Qn+Method=Absent
50481-1,NUM+Code+Component=Follitropina^5° campione+Property=ACnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
50482-9,NUM+Code+Component=Follitropina^6° campione+Property=ACnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
15068-0,NUM+Code+Component=Follitropina+Property=ARat+Timing=24h
+System=Urine+Scale=Qn+Method=Absent
49502-8,NUM+Code+Component=Glucosio-6-Fosfato deidrogenasi+Property=EntCat+Timing=Pt
+System=Sangue+Scale=Qn+Method=Absent
2357-2,NUM+Code+Component=Glucosio-6-Fosfato deidrogenasi+Property=CCnC+Timing=Pt
+System=RBC+Scale=Qn+Method=Absent
13983-2,NUM+Code+Component=Gamma globulina/Proteina.totale+Property=MFr+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=X
26466-3,NUM+Code+Component=Leucociti+Property=NCnc
+Timing=Pt+System=Liquido corporeo+Scale=Qn+Method=Absent
2345-7,NUM+Code+Component=Glucosio+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
1547-9,NUM+Code+Component=Glucosio^baseline+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
20436-2,NUM+Code+Component=Glucosio^2h post carico di glucosio+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
20437-0,NUM+Code+Component=Glucosio^3h post carico di glucosio+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
20439-6,NUM+Code+Component=Glucosio^30Minuti post carico di glucosio+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
20438-8,NUM+Code+Component=Glucosio^1h post carico con glucosio+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
20440-4,NUM+Code+Component=Glucosio^1.5h post carico con glucosio+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
26464-8,NUM+Code+Component=Leucociti+Property=NCnc+Timing=Pt
+System=Sangue+Scale=Qn+Method=Absent
6690-2,NUM+Code+Component=Leucociti+Property=NCnc+Timing=Pt
+System=Sangue+Scale=Qn+Method=X
789-8,NUM+Code+Component=Eritrociti+Property=NCnc+Timing=Pt
+System=Sangue+Scale=Qn+Method=X
26453-1,NUM+Code+Component=Eritrociti+Property=NCnc+Timing=Pt
+System=Sangue+Scale=Qn+Method=Absent
21305-8,NUM+Code+Component=Glucosio+Property=MCnc+Timing=24h
+System=Urine+Scale=Qn+Method=Absent
2350-7,NUM+Code+Component=Glucosio+Property=MCnc+Timing=Pt

+System=Urine+Scale=Qn+Method=Absent
39481-7,NUM+Code+Component=Glucosio+Property=SCnc+Timing=Pt
+System=Sangue art+Scale=Qn+Method=Absent
39480-9,NUM+Code+Component=Glucosio+Property=SCnc+Timing=Pt
+System=Sangue ven+Scale=Qn+Method=Absent
51196-4,NUM+Code+Component=Emoglobina A1/Emoglobina.totale+Property=MFr+Timing=Pt
+System=Sangue+Scale=Qn+Method=X
4552-6,NUM+Code+Component=Emoglobina A2/Emoglobina.totale+Property=MFr+Timing=Pt
+System=Sangue+Scale=Qn+Method=X
46423-0,NUM+Code+Component=Emoglobina ampiezza di distribuzione+Property=MCnc+Timing=Pt
+System=Sangue+Scale=Qn+Method=Absent
55180-4,NUM+Code+Component=Epilidimo umano proteina 4+Property=SCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
13465-0,NUM+Code+Component=Mercurio/Creatinina+Property=MRto+Timing=Pt
+System=Urine+Scale=Qn+Method=Absent
19225-2,NUM+Code+Component=Deossioglobina/Emoglobina.totale+Property=MFr
+Timing=Pt+System=Sangue art+Scale=Qn+Method=Absent
8072-1,NUM+Code+Component=Insulina Ab+Property=ACnc+Timing=Pt+System=Siero
+Scale=Qn+Method=Absent
2483-6,NUM+Code+Component=Proteina 3 legante il fattore di crescita insulino-simile+Property=MCnc
+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=Absent
6301-6,NUM+Code+Component=Coagulazione indotta del fattore tissutale.INR+Property=RelTime
+Timing=Pt+System=PPP+Scale=Qn+Method=X
20448-7,NUM+Code+Component=Insulina+Property=ACnc
+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=Absent
50501-6,NUM+Code+Component=Insulina^1° campione+Property=ACnc
+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=Absent
50502-4,NUM+Code+Component=Insulina^2° campione+Property=ACnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
50503-2,NUM+Code+Component=Insulina^3° campione+Property=ACnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
50504-0,NUM+Code+Component=Insulina^4° campione+Property=ACnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
50505-7,NUM+Code+Component=Insulina^5° campione+Property=ACnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
50506-5,NUM+Code+Component=Insulina^6° campione+Property=ACnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
50507-3,NUM+Code+Component=Insulina^7° campione+Property=ACnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
56482-3,NUM+Code+Component=Insulina^baseline+Property=ACnc

+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=Absent
47862-8,NUM+Code+Component=Insulina^post pasto+Property=ACnc+Timing=Pt+System=Siero/Plasma
+Scale=Qn+Method=Absent
53524-5,NUM+Code+Component=Fosfatasi alcalina.bile/Fosfatasi alcalina.totale+Property=CFr+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
15017-7,NUM+Code+Component=Fosfatasi alcalina.regan/Fosfatasi alcalina.totale+Property=CFr+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
15015-1,NUM+Code+Component=Fosfatasi alcalina.fegato/Fosfatasi alcalina.totale+Property=CFr+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
15013-6,NUM+Code+Component=Fosfatasi alcalina.osso/Fosfatasi alcalina.totale+Property=CFr+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
15016-9,NUM+Code+Component=Fosfatasi alcalina.placentale/Fosfatasi alcalina.totale+Property=CFr
+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=Absent
15014-4,NUM+Code+Component=Fosfatasi alcalina.intestinale/Fosfatasi alcalina.totale+Property=CFr
+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=Absent
17096-9,NUM+Code+Component=Linfociti.kappa/100 linfociti+Property=NFr
+Timing=Pt+System=Sangue+Scale=Qn+Method=Absent
46237-4,NUM+Code+Component=Linfociti.kappa/100 linfociti+Property=NFr
+Timing=Pt+System=Mid oss+Scale=Qn+Method=Absent
17224-7,NUM+Code+Component=Linfociti.lambda/100 linfociti+Property=NFr+Timing=Pt
+System=Sangue+Scale=Qn+Method=Absent
14805-6,NUM+Code+Component=Lattato deidrogenasi+Property=CCnC+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=X
60021-3,NUM+Code+Component=Lattato deidrogenasi+Property=CCnC+Timing=Pt
+System=Liquido pleurico+Scale=Qn+Method=X
2536-1,NUM+Code+Component=Lattato deidrogenasi 1/Lattato deidrogenasi.totale+Property=CFr
+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=X
2539-5,NUM+Code+Component=Lattato deidrogenasi 2/Lattato deidrogenasi.totale+Property=CFr
+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=X
2542-9,NUM+Code+Component=Lattato deidrogenasi 3/Lattato deidrogenasi.totale+Property=CFr
+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=X
2545-2,NUM+Code+Component=Lattato deidrogenasi 4/Lattato deidrogenasi.totale+Property=CFr
+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=X
2548-6,NUM+Code+Component=Lattato deidrogenasi 5/Lattato deidrogenasi.totale+Property=CFr
+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=X
5910-5,NUM+Code+Component=Lattato deidrogenasi isoenzimi+Property=Imp+
Timing=Pt+System=Siero/Plasma+Scale=Nom+Method=Absent
53226-7,NUM+Code+Component=Leucociti+Property=Naric
+Timing=Pt+System=Uretra+Scale=Qn+Method=X
12226-7,NUM+Code+Component=Leucociti+Property=PrThr

+Timing=Pt+System=Vaginale+Scale=Ord+Method=X
10501-5,NUM+Code+Component=Lutropina+Property=ACnc
+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=Absent
50509-9,NUM+Code+Component=Lutropina^1° campione+Property=ACnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
50510-7,NUM+Code+Component=Lutropina^2° campione+Property=ACnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
50511-5,NUM+Code+Component=Lutropina^3° campione+Property=ACnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
50512-3,NUM+Code+Component=Lutropina^4° campione+Property=ACnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
50513-1,NUM+Code+Component=Lutropina^5° campione+Property=ACnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
50514-9,NUM+Code+Component=Lutropina^6° campione+Property=ACnc
+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=Absent
15080-5,NUM+Code+Component=Lutropina+Property=ARat
+Timing=24h+System=Urine+Scale=Qn+Method=Absent
736-9,NUM+Code+Component=Linfociti/100 leucociti+Property=NFr+Timing=Pt
+System=Sangue+Scale=Qn+Method=X
731-0,NUM+Code+Component=Linfociti+Property=NCnc
+Timing=Pt+System=Sangue+Scale=Qn+Method=X
26478-8,NUM+Code+Component=Linfociti/100
leucociti+Property=NFr+Timing=Pt+System=Sangue+Scale=Qn+Method=Absent
17789-9,NUM+Code+Component=Grandi cellule non colorate+Property=NCnc+Timing=Pt
+System=Sangue+Scale=Qn+Method=X
17788-1,NUM+Code+Component=Grandi cellule non colorate/100 leucociti+Property=NFr
+Timing=Pt+System=Sangue+Scale=Qn+Method=X
29935-4,NUM+Code+Component=Manganese/Creatinina+Property=MRto
+Timing=24h+System=Urine+Scale=Qn+Method=Absent
785-6,NUM+Code+Component=Eritrociti emoglobina corpuscolare media+Property=EntMass
+Timing=Pt+System=RBC+Scale=Qn+Method=X
28539-5,NUM+Code+Component=Eritrociti emoglobina corpuscolare media+Property=EntMass
+Timing=Pt+System=RBC+Scale=Qn+Method=Absent
786-4,NUM+Code+Component=Eritrociti concentrazione emoglobinica corpuscolare media+Property=MCnc
+Timing=Pt+System=RBC+Scale=Qn+Method=X
28540-3,NUM+Code+Component=Eritrociti concentrazione emoglobinica corpuscolare media
+Property=MCnc+Timing=Pt+System=RBC+Scale=Qn+Method=Absent
30428-7,NUM+Code+Component=Eritrociti volume corpuscolare medio+Property=EntVol
+Timing=Pt+System=RBC+Scale=Qn+Method=Absent
2615-3,NUM+Code+Component=Metaemoglobina/Emoglobina.totale

+Property=MFr+Timing=Pt+System=Sangue art+Scale=Qn+Method=Absent
26485-3,NUM+Code+Component=Monociti/100 leucociti+Property=NFr+Timing=Pt
+System=Sangue+Scale=Qn+Method=Absent
32623-1,NUM+Code+Component=Piastrine volume medio+Property=EntVol+Timing=Pt
+System=Sangue+Scale=Qn+Method=X
34419-2,NUM+Code+Component=Reticolociti.maturi/Reticolociti.totali+Property=NFr
+Timing=Pt+System=Sangue+Scale=Qn+Method=Absent
770-8,NUM+Code+Component=Neutrofili/100 leucociti+Property=NFr+Timing=Pt
+System=Sangue+Scale=Qn+Method=X
26511-6,NUM+Code+Component=Neutrofili/100 leucociti+Property=NFr+Timing=Pt
+System=Sangue+Scale=Qn+Method=Absent
2668-2,NUM+Code+Component=Norepinefrina+Property=MRat+Timing=24h
+System=Urine+Scale=Qn+Method=Absent
2666-6,NUM+Code+Component=Norepinefrina+Property=MCnc+Timing=Pt
+System=Plasma+Scale=Qn+Method=Absent
51637-7,NUM+Code+Component=Piastrinocrito+Property=VFr+Timing=Pt
+System=Sangue+Scale=Qn+Method=Absent
51631-0,NUM+Code+Component=Distribuzione piastrinica ampiezza+Property=Rapporto
+Timing=Pt+System=Sangue+Scale=Qn+Method=Absent
32207-3,NUM+Code+Component=Distribuzione piastrinica ampiezza+Property=EntVol
+Timing=Pt+System=Sangue+Scale=Qn+Method=X
777-3,NUM+Code+Component=Piastrine+Property=NCnc+Timing=Pt
+System=Sangue+Scale=Qn+Method=X
28660-9,NUM+Code+Component=Plasminogeno effettivo/normale+Property=RelCCnc
+Timing=Pt+System=PPP+Scale=Qn+Method=X
48386-7,NUM+Code+Component=Piastrine.grandi/Piastrine+Property=NFr
+Timing=Pt+System=Sangue+Scale=Qn+Method=X
2823-3,NUM+Code+Component=Potassio+Property=SCnc
+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=Absent
21476-7,NUM+Code+Component=Potassio+Property=SCnc+Timing=24h
+System=Urine+Scale=Qn+Method=Absent
2828-2,NUM+Code+Component=Potassio+Property=SCnc+Timing=Pt
+System=Urine+Scale=Qn+Method=Absent
32713-0,NUM+Code+Component=Potassio+Property=SCnc+Timing=Pt+System=Sangue art
+Scale=Qn+Method=Absent
29349-8,NUM+Code+Component=Potassio^post dialisi+Property=SCnc+Timing=Pt+System=Siero/Plasma
+Scale=Qn+Method=Absent
2829-0,NUM+Code+Component=Potassio+Property=SRat
+Timing=24h+System=Urine+Scale=Qn+Method=Absent
2839-9,NUM+Code+Component=Progesterone+Property=MCnc+Timing=Pt

+System=Siero/Plasma+Scale=Qn+Method=Absent
2842-3,NUM+Code+Component=Prolattina+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
20568-2,NUM+Code+Component=Prolattina+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=X
50517-2,NUM+Code+Component=Prolattina^1° campione+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
50518-0,NUM+Code+Component=Prolattina^2° campione+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
50519-8,NUM+Code+Component=Prolattina^3° campione+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
50520-6,NUM+Code+Component=Prolattina^4° campione+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
50521-4,NUM+Code+Component=Prolattina^5° campione+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
12832-2,NUM+Code+Component=Prolattina^10Minuti post XXX stimolo+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
12834-8,NUM+Code+Component=Prolattina^15Minuti post XXX stimolo+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
12833-0,NUM+Code+Component=Prolattina^20Minuti post XXX stimolo+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
17242-9,NUM+Code+Component=Prolattina^40Minuti post XXX stimolo+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
1988-5,NUM+Code+Component=Proteina C reattiva+Property=MCnc+Timing=Pt+System=Siero/Plasma
+Scale=Qn+Method=Absent
21482-5,NUM+Code+Component=Proteina+Property=MCnc+Timing=24h
+System=Urine+Scale=Qn+Method=Absent
2888-6,NUM+Code+Component=Proteina+Property=MCnc+Timing=Pt
+System=Urine+Scale=Qn+Method=Absent
2885-2,NUM+Code+Component=Proteina+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
2889-4,NUM+Code+Component=Proteina+Property=MRat+Timing=24h
+System=Urine+Scale=Qn+Method=Absent
44429-9,NUM+Code+Component=Albumina/Globulina+Property=MRto
+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=X
63561-5,NUM+Code+Component=Coagulazione superficiale indotta effettivo/normale+Property=RelTime
+Timing=Pt+System=PPP+Scale=Qn+Method=X
788-0,NUM+Code+Component=Eritrociti ampiezza distribuzione+Property=Rapporto+Timing=Pt
+System=RBC+Scale=Qn+Method=X
47275-3,NUM+Code+Component=Proteina S100 B calcio legante+Property=MCnc+Timing=Pt

+System=Siero+Scale=Qn+Method=Absent
2951-2,NUM+Code+Component=Sodio+Property=SCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
2947-0,NUM+Code+Component=Sodio+Property=SCnc+Timing=Pt+System=Sangue
+Scale=Qn+Method=Absent
21525-1,NUM+Code+Component=Sodio+Property=SCnc
+Timing=24h+System=Urine+Scale=Qn+Method=Absent
2955-3,NUM+Code+Component=Sodio+Property=SCnc+Timing=Pt
+System=Urine+Scale=Qn+Method=Absent
32717-1,NUM+Code+Component=Sodio+Property=SCnc+Timing=Pt
+System=Sangue art+Scale=Qn+Method=Absent
2956-1,NUM+Code+Component=Sodio+Property=SRat+Timing=24h+System=Urine
+Scale=Qn+Method=Absent
4685-4,NUM+Code+Component=Solfoemoglobina/Emoglobina.totale+Property=MFr
+Timing=Pt+System=Sangue+Scale=Qn+Method=Absent
2484-4,NUM+Code+Component=Fattore di crescita insulino-simile I+Property=MCnc
+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=Absent
5894-1,NUM+Code+Component=Coagulazione indotta del fattore tissutale effettivo/normale
+Property=RelTime+Timing=Pt+System=PPP+Scale=Qn+Method=X
5902-2,NUM+Code+Component=Coagulazione indotta del fattore tissutale+Property=Tempo
+Timing=Pt+System=PPP+Scale=Qn+Method=X
21562-4,NUM+Code+Component=Tiocianato/Creatinina+Property=MRto
+Timing=Pt+System=Urine+Scale=Qn+Method=Absent
2571-8,NUM+Code+Component=Trigliceridi+Property=MCnc
+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=Absent
3043-7,NUM+Code+Component=Trigliceridi+Property=MCnc+Timing=Pt
+System=Sangue+Scale=Qn+Method=Absent
3087-4,NUM+Code+Component=Urato+Property=MRat
+Timing=24h+System=Urine+Scale=Qn+Method=Absent
30341-2,NUM+Code+Component=Eritrociti velocità di sedimentazione+Property=Vel+Timing=Pt
+System=Sangue+Scale=Qn+Method=Absent
13590-5,NUM+Code+Component=Proteina C attivata resistenza+Property=TRto+Timing=Pt
+System=PPP+Scale=Qn+Method=X
35365-6,NUM+Code+Component=Vitamina D+Metaboliti+Property=MCnc+Timing=Pt
+System=Siero/Plasma+Scale=Qn+Method=Absent
1990-1,NUM+Code+Component=Colecalciferolo+Property=MCnc
+Timing=Pt+System=Siero/Plasma+Scale=Qn+Method=Absent

Annexe 3 – Les grammaires d'extraction d'informations

<p>Extraction des éléments appartenant au niveau 1, sans passer aux niveaux suivants.</p>	
<p>Extraction des éléments appartenant au niveau 2, sans passer aux niveaux suivants.</p>	
<p>Extraction des éléments appartenant au niveau 3, sans passer aux niveaux suivants.</p>	

<p>Extraction des éléments appartenant au niveau 3 déterminé par les contraintes HELP.</p>	
<p>Extraction des éléments appartenant au niveau 4, sans passer aux niveaux suivants.</p>	
<p>Extraction des éléments appartenant au niveau 4 déterminé par les contraintes HELP.</p>	
<p>Extraction des éléments appartenant au niveau 5, sans passer aux niveaux suivants.</p>	
<p>Extraction des éléments appartenant au niveau 5 déterminé par les contraintes HELP.</p>	