*MEDINFO 2019: Health and Wellbeing e-Networks for All*
*L. Ohno-Machado and B. Séroussi (Eds.)*

*doi:10.3233/SHTI190484*

# SNOMED CT Coding and Analytics of *in vitro* Diagnostics Observations

**Maël Le Gall[a,b], René Vachon[a], Raphaël Petit[c], Xavier Gansel[a]**

[a] *System Development – Knowledge Engineering, bioMérieux, Centre C. Mérieux,5 Rue des Berges, Grenoble, France*
[b] *Master BioInformatique Modélisation et Statistique, Université de Rouen Normandie, Normandie Université, France*
[c] *System Development – Software Development, bioMérieux, 3 route de Port Michaud, 38390 La Balme-les-Grottes, France*

**Abstract**

*This work investigates the capability of SNOMED CT to encode microbiology laboratory data with the goal of fully describing multidrug resistance and breakpoint assignment by specimen.*

***Keywords:***
Systematized Nomenclature of Medicine; Microbiology

## Introduction

Antimicrobial resistance (AMR) is transforming the treatment of common infectious diseases. There is a strong international call for actions [1; 2] which include improved data sharing to support secondary usage of medical data [1].

Clinical microbiology laboratory data plays a key role in the fight against AMR [3]. Recording lab data without systematized nomenclature, such as LOINC & SNOMED CT interoperability standards, creates a burden for primary and secondary users of the data [4]. Initiatives in several countries demonstrate the added value of nomenclature standards to support national AMR surveillance [5]. Adopting standards by *in vitro* diagnostic (IVD) systems manufacturers should help microbiology labs report data to local, national and international AMR surveillance systems.

bioMérieux's (IVD solutions provider for the diagnosis of infectious diseases) VITEK®2, VITEK® MS, ETEST® & VIDAS® systems, have [6] >99% LOINC coverage and 90% coverage of nominal SNOMED CT observation values and 60% of SNOMED CT ordinal values. We have identified actions to help close the remaining gaps.

The work presented here describes (1) how we plan to complete the SNOMED CT encoding of our microbiology laboratory data using analytics ; (2) analyze the capability to SNOMED CT to represent and support analysis of MultiDrug Resistant Organisms (MDRO) [7] and (3) initiate SNOMED CT mediated laboratory data analytics.

## Methods

### Biological specimen

Specimen used in bioMérieux MYLA®, ARGENE® CONNECT systems are grouped according to their implied semantic and business role. Those groups were manually mapped to SNOMED CT (sub-) hierarchies concepts (table 1). The mapping was guided by the HL7 SPM segments.

*Table 1– SNOMED CT hierarchies for each group & link to HL7 SPM segment*

| Group | SNOMED CT hierarchy or sub-hierarchy | HL7 (SPM) |
|---|---|---|
| What | *105590001 \|Substance (substance)\|* | 4, 5 |
| Where | *442083009 \|Anatomical or acquired body structure (body structure)\|* | 8, 9 |
| How | *71388002 \|Procedure (procedure)\|* | 7 |
| Why | Out of the scope | -- |
| Status | *362981000 \|Qualifier value (qualifier value)\|* | 24 |

### Antibiotic

Antibiotics supported by bioMérieux VITEK®2, ETEST® & ATB™ antibiotic susceptibility testing systems were semi-automatically mapped to SNOMED CT subtypes of *410942007 |Drug or medicament (substance)|* using the process described in figure 1.

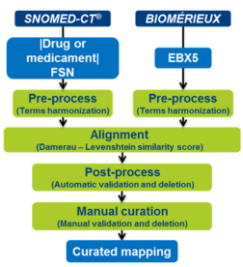


*Figure 1*

*Workflow of antibiotics mapping to SNOMED CT*

### SNOMED CT mediated labs data analytics

The specimen analytics were performed using the OWL version of SNOMED CT integrated in the Jena Fuseki triple store. The endpoint analysis, Figure2, used SPARQL queries.

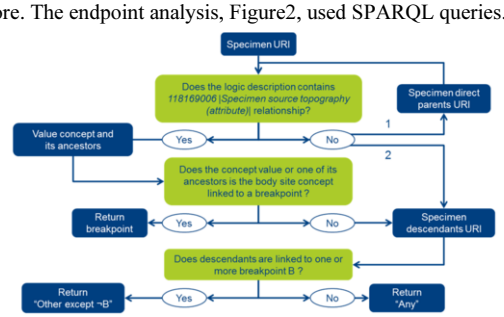


*Figure 2– Workflow to run the appropriate sample preparation workflow*

**Representing MDRO in SNOMED CT**

MRDO may be defined through rules in the form of <species name> and ((<susceptibility testing result> to <drug name>) or (<test result> to < Susceptibility test name >)). We used "methicillin resistant *Staphylococcus aureus*" (MRSA) as a test case to examine SNOMED CT's ability to described MDRO by:

1. Pre-coordination of the concept model.
2. Post-coordination of concepts and SNOMED CT compositional grammar expressivity.

**SNOMED CT**

We used the July 2018 International release of SNOMED CT for all our work.

## Results

A total of 107 internal specimen concepts were mapped to SNOMED CT with a 70% (75/107) coverage. 19 additional mappings are under validation by an expert biologist. We used 358 antibiotics form 3 antibiotic susceptibility testing systems; SNOMED CT mapping gave an 88% (316/358) coverage.

Our data analytics use case is based on drug clinical breakpoints [8; 9]. Those are key to interpret antibiotic susceptibility testing results. Breakpoints are applicable for microbial isolates originating, in the case of *Streptococcus pneumoniae* from 'non meningitis' (understood here as 'pneumonia') or 'meningitis' or 'neither pneumonia nor meningitis' specimens origin [9]. We defined 'pneumonia' breakpoint origin as "a specimen obtained from a body site subtype of *20139000 |Structure of respiratory system (body structure)|*" and 'meningitis' breakpoint as "a specimen obtained from a body site subtype of *1231004 |Meninges structure (body structure)|*". Under the working assumption that a lab may use other specimen than in the above mentioned mapping ; we considered all SNOMED CT subtype of *123038009 |Specimen (specimen)|*. Breakpoints were identifed through the *118169006 |Specimen source topography (attribute)|* relationship that tie specimen to body structure. All cases allowed to identify the breakpoints to be applied.

"Methicillin resistant *Staphylococcus aureus*" analysis, as an archetype of MDRO, shows that MDRO concepts themselves do exist in SNOMED CT. Nevertheless, they exist as primitives, and thus do not embark definition supporting future analytics. It appears that (1) some concepts needed to post-coordinate MDRO rule do also exist and (2) that compositional grammar syntax allows building some MDRO rules but not all. The limitation lies in both the SNOMED CT concept model (implemented in the MRCM) and in EL++ that do not support the exclusive disjunction 'XOR' as described in the SNOMED CT Editorial Guide.

## Conclusions

Although, SNOMED CT is able to encode biological specimen and antibiotics for the bioMérieux systems analyzed, it does not contain some of the higher-level microbiological concepts required for tracking antimicrobial resistance concepts

Our mediated analytics shows that one can use the SNOMED CT concept model to identify high level ARM concepts if a biological specimens origin is available so clinical breakpoints can be identified. We will continue to investigate this aspect based on additional relationships and usage of LOINC codes to refine and improve our analysis.

In the future we would like to model the microbial phenotypes MRSA and MDRO using pre or post-coordinated analysis of SNOMED CT concepts. Additionally, we would like to supplement our analysis by including the LOINC code mappings. Those do conflict with the SNOMED CT concept model and EL++ expressivity.

## Acknowledgements

This work as support by bioMérieux and Master BioInformatique Modélisation et Statistique, Université de Rouen Normandie, Normandie Université

## References

[1] J. Emberger, D. Tassone, M.P. Stevens, and J.D. Markley, The Current State of Antimicrobial Stewardship: Challenges, Successes, and Future Directions, *Current Infectious Disease Reports* **20** (2018), 31.
[2] WHO, Antimicrobial resistance: global report on surveillance, in, World Health Organization, 2014.
[3] P.-E. Fournier, M. Drancourt, P. Colson, J.-M. Rolain, B.L. Scola, and D. Raoult, Modern clinical microbiology: new challenges and solutions, *Nature Reviews Microbiology* **11** (2013), 574-585.
[4] R. Flatman, Terminology, units and reporting - how harmonized do we need to be?, *Clin Chem Lab Med* (2018).
[5] X. Gansel, M. Mary, and A. van Belkum, Semantic data interoperability, digital medicine, and e-health in infectious disease management: a review, *European Journal of Clinical Microbiology & Infectious Diseases* (2019).
[6] M.l. Gal, R. Vachon, B. Cellière, M. Dante, F. Holder, M. Mary, M. Olleon, and X. Gansel, Semantic laboratory data interoperability: Encoding microbiology tests and test results using LOINC ® and SNOMED-CT ® from an IVD manufacturer perspective, in: *ASM microbe 2018,* A.S.f. Microbiology, ed., American Society for Microbiology, Atlanta, 2018.
[7] A.-P. Magiorakos, A. Srinivasan, R.B. Carey, Y. Carmeli, M.E. Falagas, C.G. Giske, S. Harbarth, J.F. Hindler, G. Kahlmeter, B. Olsson-Liljequist, D.L. Paterson, L.B. Rice, J. Stelling, M.J. Struelens, A. Vatopoulos, J.T. Weber, and D.L. Monnet, Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance, *Clinical Microbiology and Infection* **18** (2012), 268-281.
[8] CLSI, M100 Performance Standards for Antimicrobial Susceptibility Testing - 28th Edition, in: *CLSI standards*, CLSI, USA, 2018.
[9] EUCAST, Clinical Breakpoints, in, European Society of Clinical Microbiology and Infectious Diseases, , 2018.

**Address for correspondence**

Xavier Gansel ; xavier.gansel@biomerieux.com ; +33 624 01 43 30
bioMérieux, Centre C. Mérieux - 5 rue des Berges - 38000 Grenble - FRANCE

ORCID https://orcid.org/0000-0001-7148-5461