Opinion Paper

Madeleen Bosma* and Christa Cobbaert

Digital metrology in laboratory medicine: a call for bringing order to chaos to facilitate precision diagnostics

https://doi.org/10.1515/cclm-2025-0152
Received February 7, 2025; accepted April 2, 2025;
published online April 16, 2025

Abstract: Laboratory medicine is faced with rapid developments in data exchange, secondary use of data and artificial intelligence. Safe exchange of laboratory data requires a suitable terminology standard. NPU, LOINC and SNOMED CT are increasingly used for this purpose, but none of these terminology standards can currently accommodate safe exchange across the full spectrum of conventional laboratory data. Furthermore, rapid technological advances in, amongst others, the 'omics' area will enforce a shift towards precision diagnostics. These emerging technologies demand an appropriate and future-proof terminology standard. Given the current and future challenges in laboratory terminologies, we here present a concept for digital metrology in laboratory medicine. Terminology standards used in laboratory medicine should be adjusted to the current state of science to allow safe data exchange and interpretation. Essential test information entails the full spectrum of pre-pre-analysis to post-post-analysis. Major improvements needed include sufficient coding detail for the molecular form of the measurand and information on metrological traceability. Furthermore, especially given the advances in precision diagnostics, it will become essential to indicate interrelationships between measurands. Herefore, integration with established taxonomies would allow improved identification of interrelationships between measurands and linkage with scientific information for multidisciplinary data science. Hence, laboratory data can further gain in specificity and value. The time has come to lay the basis for safe data exchange in the era of precision diagnostics, with a global focus. A consensus for digital metrology in laboratory medicine will be essential to move forward with health data exchange within Europe and beyond.

Keywords: metrology; laboratory terminology standards; emerging technologies; secondary use of data; artificial intelligence

Introduction

Laboratory medicine is faced with rapid developments in data exchange, secondary use of data and artificial intelligence (AI). Furthermore, data exchange moves from regional initiatives to global movements (Figure 1), for instance the European Health Dataspace (EHDS). Moreover, rapid technological advances in the 'omics' area will enforce a shift towards precision diagnostics, aiming to stratify patients in order to precisely diagnose and monitor patients. Current laboratory terminology systems are not ready to facilitate these widespread technological advances.

Correct data exchange and interpretation of data from different sources requires sufficient detail in the data being exchanged. Developments in precision medicine will result in growing availability of measurands with different levels of selectivity. Furthermore, despite significant efforts, standardization of medical tests is lagging behind, which implies availability of tests with different levels of standardization. Laboratory terminology standards and IT systems will need to develop in order to facilitate the correct identification of selectivity and standardization level of the measurands.

The European regulation for *in vitro* diagnostics (IVDR) demands safe and effective IVD's. To realize this, a new clinical evidence requirement has been incorporated in the IVDR, encompassing the triad of scientific validity, analytical performance and clinical performance, which is key for guaranteeing that a medical test is fit-for-clinical-purpose. Note that analytical performance, including metrological traceability of test results within allowable measurement uncertainty, and clinical performance are interdependent requirements which should be aligned. For enabling

***Corresponding author: Madeleen Bosma**, MSc PhD, Department of Clinical Chemistry and Laboratory Medicine, Leiden University Medical Centre, Leiden, The Netherlands, E-mail: m.de_geus-bosma@lumc.nl. https://orcid.org/0000-0003-0547-0100
Christa Cobbaert, Department of Clinical Chemistry and Laboratory Medicine, Leiden University Medical Center, Leiden, The Netherlands. https://orcid.org/0000-0003-3565-1404

Open Access. © 2025 the author(s), published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.

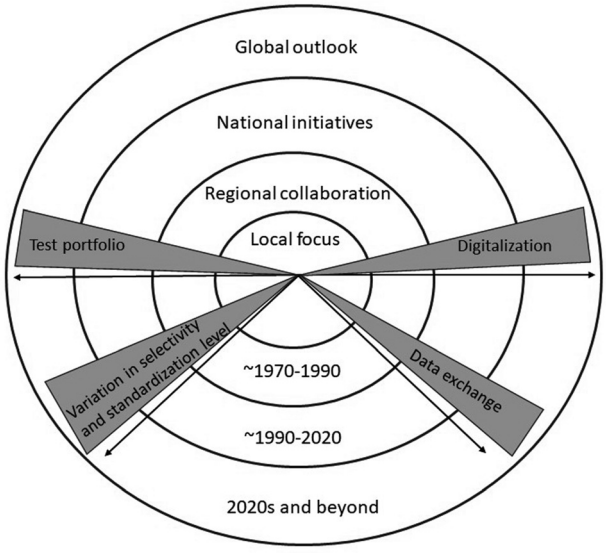

Figure 1: Developments in laboratory medicine and digitalization: from a local focus to a global outlook. The test portfolio is expanding, while standardization is lagging behind, resulting in variation in test selectivity and standardization level. On the other hand, there is an increasing demand for digitalization and data exchange. The combination of these challenges underscores the need for digital metrology in laboratory medicine.

exchangeability of medical test results, reference intervals and clinical decision limits in this digitally connected world, metrological traceability of test results has become increasingly important. This requires concurrent digitalization of this information to ensure proper interpretation.

Furthermore, with increasing data exchange between laboratories and organizations and increased secondary use of data it becomes essential that all available laboratory information, originating from the total test process, is coded uniformly and uniquely. This is in line with the FAIR guiding principles for scientific data management and stewardship: Fair, Accessible, Interoperable and Reusable [1]. Terminology standards currently used in medical laboratories, like LOINC, NPU and SNOMED CT do however not allow for sufficient coding detail, resulting in absent essential information when data exchange is based on current terminology standards.

When essential medical test information is lacking due to data exchange without sufficient detail, this could potentially harm patient care. Additionally, lack of sufficiently detailed information may affect conclusions of (retrospective) research studies (secondary use of data) and clinical decision support (CDS) tools. Developments in AI have accelerated big data analysis for research and clinical or epidemiological purposes. Data availability and AI model development are quickly expanding, but proper data coding is lagging behind. Next to primary data (laboratory test results), also properly coded clinical laboratory meta- and peridata are essential for effective AI model development and implementation [2]. Metadata is defined as data derived from the testing process that describe the characteristics and the requirements that are relevant for assessing the quality and the validity of the laboratory test results. Peridata is data facilitating the accurate interpretation of test results (for example, clinical data and reference intervals) [2]. For safe and effective CDS, standardized coding and structured data is indispensable, inadequately coded data harms the effectiveness of CDS and can potentially cause patient harm [3]. Hence, sufficient coding detail across the full spectrum of laboratory data (primary data, metadata and peridata) is of utmost importance for innovation and safe patient care. With the expected exponential growth in clinical data available per patient, exceeding the human cognitive capacity, in combination with a trend towards personalized medicine and holistic approaches, the adoption of AI and clinical decision support in healthcare will grow, thus further increasing the need for proper identification and coding of laboratory data. Here, we present a concept for digital metrology in laboratory medicine.

Current state of the art of terminology standards in laboratory medicine

The most frequently used terminology standards in medical laboratories for standardized reporting of tests and interpretation are LOINC, NPU and SNOMED CT. Additionally, regional terminology standards are in use. Current terminology standards are based on measurands conventionally used in diagnostic laboratories and do not currently facilitate exchange of, amongst others, proteoform information. Furthermore, terminology standards do not optimally facilitate listing of detailed method information, including the calibration hierarchy.

LOINC codes medical laboratory test results based on the system (sample type), component/analyte, method (optional, e.g. immunoassay, chromogenic activity test or molecular analysis), property (the quantity, e.g. mass or concentration, NB: not the specific unit), time (moment in time or time interval) and scale (the level of measurement, e.g. quantitative or ordinal) [4]. LOINC terms are organized in basic hierarchies. LOINC codes the component and system in combination with the measured property and optionally the method, not the result. Hence, with LOINC, gene variant

analysis for antithrombin (SERPINC1) can be coded (LOINC 93814-2 or 70107-8), but not the resulting gene variant. The number of LOINC codes per measurand increases rapidly, partly due to the desire of laboratory professionals to include method information. Only the Regenstrief Institute can assign new LOINC codes. LOINC codes are generated on request, a basis for sufficient method information per measurand is however lacking. Furthermore, a validated approach for safe grouping of LOINC codes, i.e. information on which LOINC codes are sufficiently similar to be grouped, does not exist. Hence, researchers may group test results with different LOINC codes without a metrological basis, which may result in incorrect conclusions. With technological advances an exponential increase in LOINC codes is expected, which without a proper grouping strategy will hamper data exchange and secondary use of data.

NPU codes laboratory test results based on system, component and unit. The system can be a patient (e.g. body of a patient), a distinct part of the patient (e.g. blood, cerebrospinal fluid), or a part of the patient's surroundings (e.g., house dust or drinking water). In contrast to LOINC, NPU includes the measurement unit (based on the IUPAC rules on nomenclature and units). NPU codes the property of a system and not the measurement technique. NPU is based on metrological concepts, for NPU codes with a WHO IU, the certified reference material (CRM) is stated (this does not apply to NPU codes with SI units). Like LOINC, NPU offers codes for gene variant analysis but currently does not offer the possibility to code protein variant analysis and protein variant quantification. Furthermore, NPU is a terminology system without a hierarchy and, like LOINC, no validated approach exists for identification and grouping of NPU terms that can safely be exchanged.

SNOMED CT is a multiaxial nomenclature system (taxonomy) covering the full spectrum of clinical medicine [4]. SNOMED CT is based on hierarchical relationships. SNOMED CT concepts are arranged in different hierarchies (19 currently). The hierarchy 'specimen' specifies the type of sample, the laboratory result is coded either as 'observable entity' when it regards a precoordinated concept or as 'substance' for so-called post-coordinated concepts, which can be further specified with attributes [4]. Additionally, also the hierarchies 'clinical finding' and 'procedure' may contain laboratory-derived information. For example, in the hierarchy 'procedure' phenotype determination can be coded (e.g. apolipoprotein E phenotype determination, 250692007) and in the hierarchy 'finding' the interpretation of a phenotype can be coded (e.g. finding of apolipoprotein E phenotype, 365833001). Notably, this phenotype finding is currently based on genetic testing, while phenotype testing may also be performed with mass spectrometry/proteomics. Furthermore, there is no possibility in SNOMED CT currently to code quantitative proteoform analysis. With SNOMED CT, preanalytical information (for example the blood draw location and type of sample) and postanalytical information (interpretative comments) can be coded.

LOINC and SNOMED CT cooperate, in order to achieve that LOINC data will be integrated in the SNOMED CT terminology. Unified Code for Units of Measure (UCUM) is the preferred method for unit expression in combination with LOINC within SNOMED CT [4]. This cooperation will enhance standardization of laboratory data. A shortcoming of UCUM is that it does not display scientific characters such as the Greek letter, μ, or integer powers. A mapping of LOINC and SNOMED CT with NPU is to be expected. The combination of NPU or LOINC with SNOMED CT allows coding from preanalysis to postanalysis. A major shortcoming of these terminologies however is the lack of a requirement for method specification (including measurand intended to be measured and the ISO 17511-proof calibration hierarchy), which hampers safe data exchange. Furthermore, these terminologies are not ready for emerging technologies in the 'omics' area.

In conclusion, NPU, LOINC, UCUM and SNOMED CT have improved standardized reporting of medical laboratory test information. However, critical test information is currently lacking and these terminologies are not yet ready for emerging technologies in laboratory medicine.

Critical test information

Essential test information for safe data exchange and secondary use of data now and in the future entails the full spectrum of pre-pre-analysis to post-post-analysis. A summary of key elements of test information is listed below.

Critical test information

- Preanalytical information, including the blood draw location, type of sample and the processing of the sample
- Time aspect
- Measurand, molecular form if applicable
- Test result
- Unit (SI system)
- Detailed method information
- Calibration hierarchy (including the certified reference material, primary reference material and secondary matrix-based reference material to which test results are traceable) and allowable measurement uncertainty (per intended use)

- Analyzer/application (if relevant)
- Type of operator (for POCT)
- Reference-interval (for non-standardized tests and when locally validated) or clinical decision limit (CDL)
- (Interpretative) comments (including what has not been excluded)

The combination of SNOMED CT with LOINC or NPU can cover the above mentioned elements partly. The main shortcomings are the lack of sufficient coding detail of the measurand or the molecular form (selectivity) and metrological traceability (calibration hierarchy). Furthermore, relationships between measurands can currently only be indicated at a very basic level. Other aspects could be addressed in other data fields as meta- or peridata in HL7 (a set of international standards created to facilitate the exchange of electronic health care data between healthcare providers).

The basis for precision diagnostics: nomenclature and taxonomy

The range of measurands in laboratory medicine spans from small molecules to cells and from gene variants to proteins and proteoforms. Across laboratory medicine, technologies enabling precision diagnostics are expanding and entering the routine clinical care. Hence, for all these types of elements terminology standards should facilitate correct identification of the measurand. We propose to integrate established nomenclature and taxonomy systems. Nomenclature for the name of the element, taxonomy for their relationships. Furthermore, a solution should be sought for combinations of conventional techniques with advanced, more selective methods; a taxonomy should describe their interrelationship and it should be interpretable which tests can safely be exchanged and which not.

We will illustrate the need for advances in digital metrology in laboratory medicine on the basis of a few examples.

Example 1: essential method information

Creatinine, the cornerstone of renal function monitoring, can be measured with different methods that differ significantly. The Jaffé method is a colorimetric method which is hampered by interference from pseudochromogens and other interferences such as bilirubin and glucose. Several variants of the Jaffé method exist. The Jaffé method is evidently inferior to the more selective enzymatic methods, and patients with kidney function loss may be misclassified with Jaffé methods [5–8]. Jaffé methods should have been abandoned especially in secondary and tertiary care centers and in pediatric patients because of potential patient harm. Nevertheless a significant number of Jaffé using laboratories remained because of economic reasons (much lower cost of the Jaffé method compared to the high cost of the enzymatic method). Next to enzymatic methods, LC-MS tests may be needed in case of heavily interfered enzymatic creatinine tests in e.g. icteric liver transplant patients. For creatinine analyzed with LC-MS a separate LOINC code exists (103616-9), but the chemical and enzymatic methods are currently not distinguished within LOINC. In NPU, a separate code exists for the Jaffé method (NPU01807) and the enzymatic methods (NPU04998), but no code exists for the LC-MS method. SNOMED CT offers a procedure code for ‘quantitative measurement of mass concentration of creatinine in serum or plasma specimen using colorimetric enzyme technique’ (444207000, see Supplemental Figure 1), which is linked to the group of creatinine measurements and contains attributes describing the scale type and specimen. An additional code for the Jaffé (colorimetric non-enzymatic) method is lacking. Yet, given the hierarchical orientation of SNOMED CT and the possibilities to add further details by attributes, this design does offer the possibility to distinguish between methods and potentially calibration information with the possibility to list interrelationships. For SNOMED CT, NPU and LOINC applies that codes are mainly created on request, which may explain the lack of completeness.

In case of highly heterogenous measurands, method information is also highly relevant. An intriguing example is apolipoprotein(a) in lipoprotein(a) (Lp(a)), the former being characterized by a KIV type 2 size polymorphism and a variation in molecular mass from 280 to 850 Da. These protein characteristics in combination with the use of polyclonal immunoassays made that Lp(a) test results were confounded and masked the association with coronary heart disease, stroke and peripheral arterial disease [9–11]. Due to overlooking this confounding factor, clinical researchers lost interest in Lp(a), a very relevant genetic pro-thrombotic, pro-atherogenic and pro-inflammatory risk factor, for almost 2 decades. Lp(a)’s renaissance came when GWAS and Mendelian randomization studies confirmed its causality with myocardial infarction, stroke and peripheral arterial disease.

A third example is related to the secondary use of test results from the electronic health record (EHR) of patients. If software tools like LiverPro are developed for detecting and/or monitoring liver fibrosis [12, 13] it is of utmost importance that the test results that are used in this software tool are either standardized and exchangeable among participating

hospitals, or method information is included so that tools can be adjusted to the specific method used locally. If this is not guaranteed, these software tools and scores will be misleading and harmful for patients and caregivers.

In summary, LOINC, NPU and SNOMED CT all lack completeness and need optimization when it comes to coding of essential method information on creatinine and Lp(a) results in order to judge equivalence.

Example 2: calibration hierarchy

Commercial assays for serum ferritin are calibrated against different WHO international standards. These differences in calibration hierarchies have introduced a significant bias [14]. Therefore, rather than following clinical decision limits from clinical guidelines, clinicians should base their interpretation of ferritin data on locally verified reference ranges and clinical decision limits. This is possible within hospitals, but when ferritin data is exchanged between hospitals or is included in research projects, it will be essential to include the information on metrological traceability. Within the NPU system, different codes are available for the different WHO international standards: IS 80/578 (NPU29748), IS 80/602 (NPU28969), and IS 94/572 (NPU28586). The possibility to include information on the metrological traceability, including the reference material, does currently not exist in LOINC and SNOMED CT.

Terminology standards should adapt to the current state of science and practice in metrology, ISO 17511 compliant. Evolution to matrix-based secondary reference materials that are commutable should be enabled. For non-standardized methods terminology standards should facilitate inclusion of calibration hierarchy information. For laboratory tests that are sufficiently standardized, a label could be generated that indicates that these tests are sufficiently comparable for safe exchange.

Example 3: the molecular form(s) of the measurand(s)

Developments in the area of quantitative proteomics have great potential to enhance precision diagnostics. Quantitative proteomics using LC-MS has excellent molecular selectivity and hence allows analysis of proteoforms; the range of different structures of a protein product of a single gene, including variations in amino acid sequence and post-translational modifications (Forgrave et al., 2022). Specific proteoforms may be linked to disease and risk for disease and/or disease activity. For example, specific antithrombin proteoforms are associated with lack of heparin-binding properties (Kruijt et al. 2023) and specific apolipoprotein E isoforms are associated with risk for cardiovascular disease and Alzheimer's disease (Reijnders et al., 2024). The driver for implementation of precision diagnostics is the clinical need/clinical gap in the care pathway. Only when proteoform analysis has clinical value and delivers actionable results, precision diagnostics has a place in patient care. Hence, conventional laboratory medicine tests will remain cornerstone tests of patient follow-up, supplemented with precision diagnostics.

Due to marked technological improvements in the field of proteomics, bottom-up and top-down approaches in analysis of proteoforms, the human proteoform database is growing [15]. In parallel, LC-MS techniques become better suitable for implementation in routine clinical laboratories. Fully automated analyzers have become available for small molecules, therapeutic drug monitoring (TDM) and for some proteoforms. Hence, an expansion in proteoform analysis is expected in the nearby future in clinical care. Furthermore, with increasing data availability and secondary use of data it becomes essential that protein measurands are coded uniformly and uniquely to guarantee proper data exchange. Current terminology standards (SNOMED CT, LOINC and NPU) do not facilitate proteoform coding but should be enabled to accommodate the evolution to more refined molecular measurands and selective, properly calibrated tests.

UniProtKB is the leading database for functional protein data. It contains information on protein structure, protein function, proteoforms and clinical relevance. Proteins have a unique protein identification number and proteoforms representing different protein isoforms coming from alternatively spliced transcripts have their own unique protein identifiers in UniProtKB. UniProtKB has adopted the FAIR principles for scientific data management and stewardship [16]. Furthermore, UniProtKB links to a range of other relevant databases with gene and protein information, like the NCBI snpDB, the dominant database for single-nucleotide polymorphisms. Additionally, UniProtKB identification is frequently used in other scientific disciplines. Hence, UniProtKB can serve as a bridge between clinical data and scientific information, thus facilitating multidisciplinary research, which is expected to grow since combined data has the potential to provide new insights. For example, environmental (exposome) data could be combined with medical data to obtain insights into environmental risk factors. In NPU, all proteins are defined and termed according to UniprotKB, protein variant data is however currently lacking. UniprotKB has the potential to enhance current laboratory terminologies to facilitate coding of proteomics data. To

Table 1: Clinically relevant type II mutations of antithrombin affecting antithrombin activity. The list is based on Ruhaak et al. 2017 [15] supplemented with newly added variants/proteoforms listed in UniProtKB. UniProtKB protein ID: P01008 · ANT3_HUMAN. Additional information on variant was copied from the UniProtKB database (field: variant description). All variants are associated with antithrombin deficiency (ATD) type II (exceptions where variants are also associated with type I are indicated). Additional variant description is listed in the Table. Date of data extraction: September 5th, 2024. NB: genetic variants do not always have a unique NCBI snpDB (e.g. 5a and 5b), yet they do have unique UniProtKB variant IDs.

No.	AA no.	Mutation	NCBI snpDB	UniProtKB	Variant description	Additional information on variant
1	39	I→R	rs121909558	VAR_007033	Rouen-3	Lack of heparin-binding properties
2	53	C→F	N.A.	VAR_071199	N.A.	N.A.
3	56	R→C	rs28929469	VAR_007035	Rouen-4	Lack of heparin-binding properties
4	73	P→L	rs121909551	VAR_007036	N.A.	Lacks heparin-binding ability.
5a	79	R→C	rs121909547	VAR_007037	Tours/Alger/Amiens/Toyama/Paris-1/ Paris-2/ Padua-2/Barcelona-2/Kumamoto/Omura/ Sasebo	Lacks heparin-binding ability.
5b	79	R→S	rs121909547	VAR_007039	Rouen-2	Lack of heparin-binding properties
5c	79	R→H	rs121909552	VAR_007038	Rouen-1/Padua-1/Bligny/Budapest-2	Lack of heparin-binding properties
6	112	P→S	N.A.	VAR_086227	N.A.	Severely decreased antithrombin activity
7	125	G→D	N.A.	VAR_071200	N.A.	N.A.
8a	131	L→F	rs121909567	VAR_007045	Budapest-3/Budapest-7	N.A.
8b	131	L→V	N.A.	VAR_007046	Southport	N.A.
9	146	K→E	rs1170430756	VAR_027456	Dreux	Complete loss af heparin binding
10	147	T→A	rs2227606	VAR_013085	N.A.	N.A.
11	148	S→P	rs121909569	VAR_007049	Nagasaki	Defective heparin binding associated with thrombosis
12	150	Q→P	rs765445413	VAR_007050	Vienna	N.A.
13	161	R→Q	rs121909563	VAR_007054	Geneva	N.A.
14	170	S→P	rs1657786518	VAR_071201	N.A.	N.A.
15	198	Y→C	rs1425532034	VAR_007056	In AT3D type-I and -II; Whitechapel	N.A.
16	218	I→N	N.A.	VAR_071202	N.A.	N.A.
17a	219	N→D	rs121909571	VAR_007059	Rouen-6	Increases affinity for heparin
17b	219	N→K	N.A.	VAR_007058	Glasgow-3	N.A.
18	248	V→G	N.A.	VAR_071203	N.A.	N.A.
19	269	E→K	rs758087836	VAR_007060	Truro	Increases affinity for heparin
20a	283	M→I	rs2102783093	VAR_007062	N.A.	N.A.
20b	283	M→V	N.A.	VAR_027468	N.A.	N.A.
21	293	R→P	N.A.	VAR_071204	N.A.	N.A.
22	316	I→N	N.A.	VAR_007064	Haslar/Whitechapel	N.A.
23	334	E→K	N.A.	VAR_007065	N.A.	N.A.
24	401	H→R	N.A.	VAR_071205	N.A.	N.A.
25	414	A→T	rs121909557	VAR_007069	Hamilton/Glasgow-2	Reduces interaction with thrombin by 90 %
26	416	A→S	rs121909548	VAR_007071	Cambridge-2	N.A.
27	424	G→D	rs121909566	VAR_007073	Stockholm	N.A.
28a	425	R→C	rs121909554	VAR_007075	N.A.	N.A.
28b	425	R→H	rs121909549	VAR_007074	Glasgow/Sheffield/Chicago/Avranches/ Kumamoto-2	Increases affinity for heparin; deprived of inhibitory activity.
28c	425	R→P	rs121909549	VAR_007076	Pescara	Deprived of inhibitory activity
29	426	S→L	rs121909550	VAR_007077	Denver/Milano-2	Deprived of inhibitory activity
30a	434	F→C	rs1572084546	VAR_007078	Rosny	N.A.
30b	434	F→L	N.A.	VAR_007080	Maisons-Laffite	N.A.
30c	434	F→S	N.A.	VAR_007079	Torino	N.A.
31	436	A→T	rs121909546	VAR_007081	Oslo/Paris-3	N.A.
32	437	N→K	rs1301351856	VAR_007082	La Rochelle	N.A.
33a	438	R→G	N.A.	VAR_009258	In AT3D type I and type-II	N.A.
33b	438	R→M	N.A.	VAR_007083	Kyoto	N.A.
34a	439	P→A	rs1487411568	VAR_071206	N.A.	N.A.
34b	439	P→L	rs121909555	VAR_007084	Utah	Deprived of inhibitory activity
34c	439	P→T	rs1487411568	VAR_007085	Budapest-5	N.A.

Table 1: (continued)

No.	AA no.	Mutation	NCBI snpDB	UniProtKB	Variant description	Additional information on variant
35	441	L→P	rs1188571702	VAR_027473	N.A.	N.A.
36	457	R→T	N.A.	VAR_007088	N.A.	N.A.
37	461	P→L	rs121909564	VAR_007091	Budapest	N.A.

N.A., not available.

illustrate the richness of UniProtKB data, Table 1 lists the UniProtKB protein variant identifiers related to antithrombin (UniProtKB protein ID: P01008 · ANT3_HUMAN), a protein for which a multiplex LC-MS/MS assay was developed in our group [17].

Given the taxonomy structure, broad scope and wide usage internationally, we argue that a promising design would be an extension of the current SNOMED CT ontology (including the integration of LOINC codes, possibly mapped with NPU) in combination with UniProtKB identifiers for protein and proteoform identification. Hence, protein variant data can be added and hierarchical relationships (i.e. "parent-child" relationships for proteins and proteoforms) can be integrated. This is in line with developments in the area of pathology, where a management plan for molecular pathology sequence data involving SNOMED CT, LOINC and Human Genome Variant Society nomenclature (HGVS) was presented [18, 19]. However, unlike genetic identification, quantification is also an element of the analysis in proteomics. Hence, for proteomics additional extension of the model would be necessary.

By creating different hierarchical levels, the level of specification of the measurand can be indicated. Hence, immunoassays in which a mixture of proteoforms is measured should receive a different code level as compared to assays (like mass spectrometry) where specific proteoforms are measured. Hence, the hierarchy 'substance' should be supplemented with a secondary hierarchy to be used to specify the specific proteoform.

While UniProtKB allows protein variant coding, post-translational modification (PTM) is listed in the UniProtKB database but is not coded (yet). UniProt states that assigning identifiers for PTM is complicated, since they may include crosslinks involving more than one residue either within or between proteins. Coding single residue PTMs however, like the clinically relevant glycosylation of antithrombin at position Asn-167 (β-antithrombin, in contrast to the more abundant α-antithrombin, lacks the glycan at position Asn-167 and is more active than α-antithrombin owing to its increased affinity for heparin) [20], does seem feasible and will hopefully be implemented in the UniProtKB database.

Where UniprotKB is a promising candidate taxonomy for protein and proteoform information, established scientific databases for other types of measurands, for example the NCBI gene database for gene variants, should be integrated for other measurands, where relevant. This will facilitate correct identification of the measurand and its interrelationships with other measurands, as well as links with scientific information, which will enhance secondary use of data.

Conclusions

Given the fast technological advances and the global demand for data exchange, secondary use of data and CDS, terminology standards used in laboratory medicine should be adjusted to allow unequivocal digital data exchange and interpretation. The major improvements needed concern sufficient coding detail regarding the molecular form(s) of the measurand and sufficient information on metrological traceability. Furthermore, with integrated diagnostics looming on the horizon, where laboratory test results are combined with imaging and pathology data coupled with advanced information technology, and the emerging technologies that enter routine clinical care, it will become essential to identify the interrelationships between measurands. Herefore, integration with established taxonomies like UniprotKB would allow improved specification of the measurand and its interrelationships and linkage with scientific information for multidisciplinary data science. The time is now to build the era of precision diagnostics and software tools, with a global rather than a local or national focus. A global and standardized approach for digital metrology in laboratory medicine will be essential to enable health data exchange within Europe and beyond.

Acknowledgments: The authors thank Dr. Renee Ruhaak and dr. Mirjam Kruijt for their contribution to the list of antitrombin variants.

Research ethics: Not applicable.

Informed consent: Not applicable.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Use of Large Language Models, AI and Machine Learning Tools: None declared.
Conflict of interest: The authors state no conflict of interest.
Research funding: None declared.
Data availability: Not applicable.

References

1. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. Comment: the FAIR Guiding Principles for scientific data management and stewardship. Sci Data 2016;3:160018.
2. Padoan A, Cadamuro J, Frans G, Cabitza F, Tolios A, De Bruyne S, et al. Data flow in clinical laboratories: could metadata and peridata bridge the gap to new AI-based applications? An investigation on behalf of the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) Working Group on Artificial Intelligence (WG-AI). Clin Chem Lab Med 2024;63:684–91.
3. Plebani M, Cadamuro J, Vermeersch P, Jovičić S, Ozben T, Trenti T, et al. A vision to the future: value-based laboratory medicine. Clin Chem Lab Med 2024;62:2373–87.
4. Bietenbeck A, Boeker M, Schulz S. NPU, LOINC, and SNOMED CT: a comparison of terminologies for laboratory results reveals individual advantages and a lack of possibilities to encode interpretive comments. J Lab Med 2018;42:267–75.
5. Drion I, Cobbaert C, Groenier KH, Weykamp C, Bilo HJ, Wetzels JF, et al. Clinical evaluation of analytical variations in serum creatinine measurements: why laboratories should abandon Jaffe techniques. BMC Nephrol 2012;13:133.
6. Delanghe JR, Cobbaert C, Harmoinen A, Jansen R, Laitinen P, Panteghini M. Focusing on the clinical impact of standardization of creatinine measurements: a report by the EFCC working group on creatinine standardization. Clin Chem Lab Med 2011;49:977–82.
7. Cobbaert CM, Baadenhuijsen H, Weykamp CW. Prime time for enzymatic creatinine methods in pediatrics. Clin Chem 2009;55:549–58.
8. Weykamp C, Kuypers A, Bakkeren D, Franck P, Loon DV, Gunnewiek JK, et al. Creatinine, Jaffe, and glucose: another inconvenient truth. Clin Chem Lab Med 2015;53:e347–9.
9. Ridker PM, Stampfer MJ, Rifai N. Novel risk factors for systemic atherosclerosis. JAMA 2001;285:2481–5.
10. Ridker PM, Stampfer MJ, Hennekens CH, Women's Hospital B. Plasma concentration of lipoprotein(a) and the risk of future stroke. JAMA 1995;273:1269–73.
11. Ridker PM, Hennekens CH, Stampfer MJ. A prospective study of lipoprotein(a) and the risk of myocardial infarction. JAMA 1993;270:2195–9.
12. Lindvig KP, Thorhauge KH, Hansen JK, Kjærgaard M, Hansen CD, Johansen S, et al. Development, validation, and prognostic evaluation of LiverPRO for the prediction of significant liver fibrosis in primary care: a prospective cohort study. Lancet Gastroenterol Hepatol 2025;10:55–67.
13. Åberg F, Männistö V. The promise of automated liver disease risk stratification in primary care. Lancet Gastroenterol Hepatol 2025;10:5–7.
14. Braga F, Pasqualetti S, Frusciante E, Borrillo F, Chibireva M, Panteghini M. Harmonization status of serum ferritin measurements and implications for use as marker of iron-related disorders. Clin Chem 2022;68:1202–10.
15. Smith LM, Agar JN, Chamot-Rooke J, Danis PO, Ge Y, Loo JA, et al. The human proteoform project: defining the human proteome. Sci Adv 2021;7:eabk0734.
16. Garcia L, Bolleman J, Gehant S, Redaschi N, Martin M, Bateman A, et al. FAIR adoption, assessment and challenges at UniProt. Sci Data 2019;6:175.
17. Ruhaak LR, Romijn FPHTM, Smit NPM, Van Der Laarse A, Pieterse MM, De Maat MPM, et al. Detecting molecular forms of antithrombin by LC-MRM-MS: defining the measurands. Clin Chem Lab Med 2018;56:1704–14.
18. Campbell WS, Karlsson D, Vreeman DJ, Lazenby AJ, Talmon GA, Campbell JR. A computable pathology report for precision medicine: extending an observables ontology unifying SNOMED CT and LOINC. J Am Med Inf Assoc 2018;25:259–66.
19. Campbell WS, Carter AB, Cushman-Vokoun AM, Greiner TC, Dash RC, Routbort M, et al. A model information management plan for molecular pathology sequence data using standards: from sequencer to electronic health record. J Mol Diagn 2019;21:408–17.
20. Kruijt M, Treep MM, Cobbaert CM, Ruhaak LR. Antithrombin diagnostics by mass spectrometry: development and analytical validation of a next-generation test. Res Pract Thromb Haemost 2023;7:100079.

Supplementary Material: This article contains supplementary material (https://doi.org/10.1515/cclm-2025-0152).