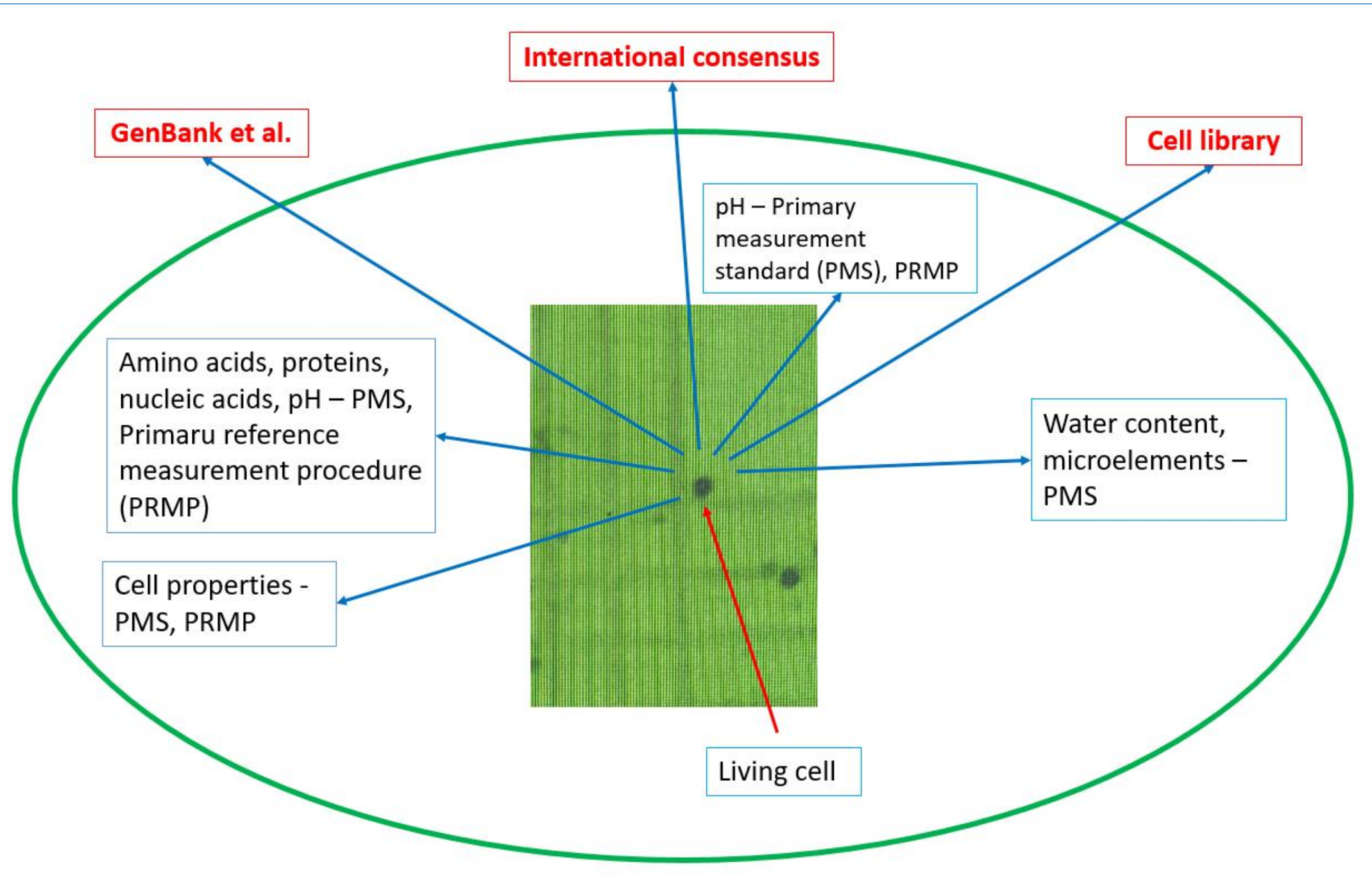


NEW BRANCH OF METROLOGY – METROLOGY OF NOMINAL PROPERTIES

Elena Kulyabina^{1*}, Valeria Morozova², Tatiana Kulyabina¹, Vitaly Morozov³
¹ Research Center for Applied Metrology - Rostest (RCAM-Rostest), Ozeraya str. 46, Moscow, Russian Federation, 119361
² California State University, 5151 State University Drive, Los Angeles, CA 90032, USA
³ Rusagro Group PJSC, Moscow, Russian Federation, 115054
* To whom correspondence should be addressed. Email: kuliabina@mail.ru

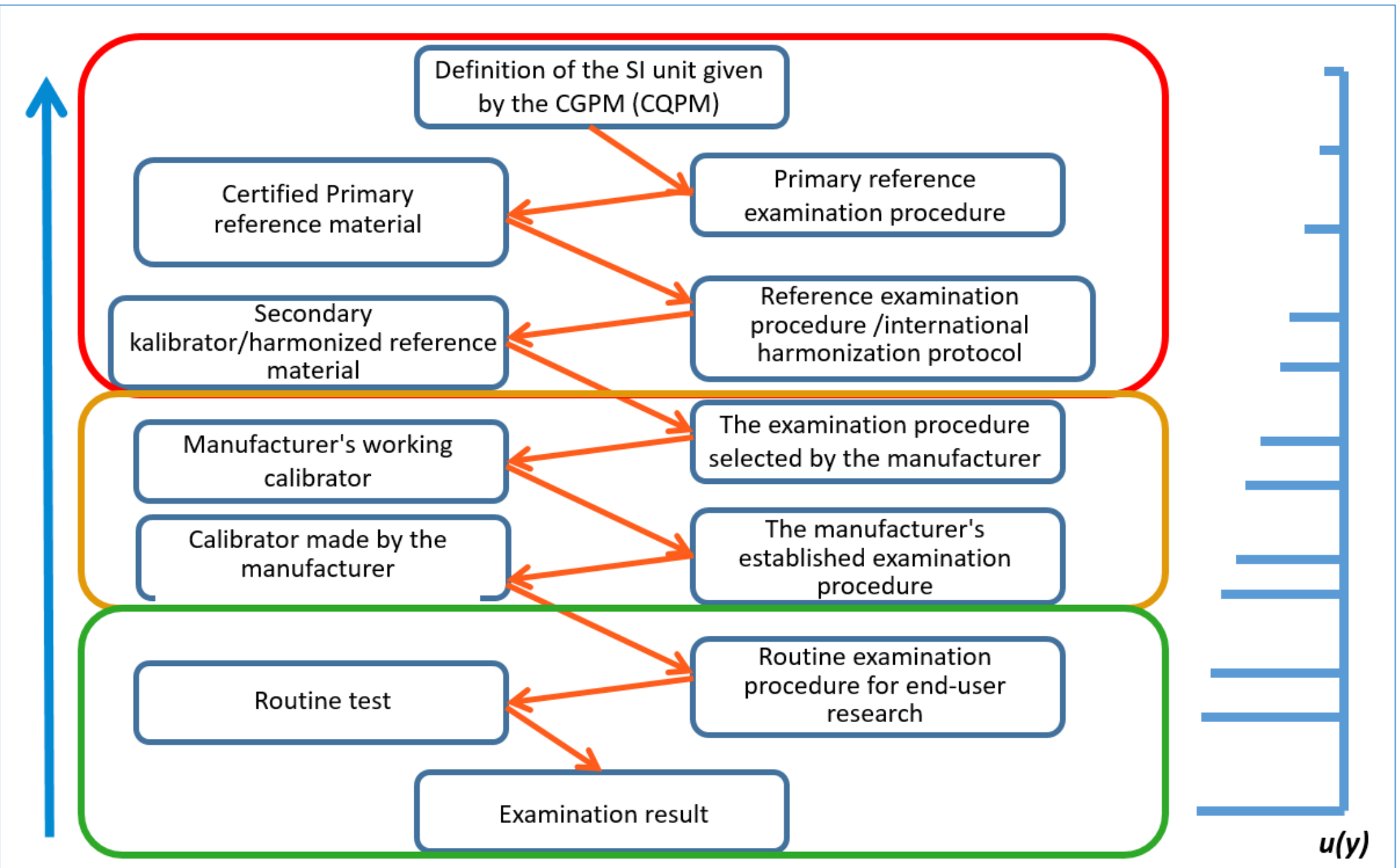
Quantitative properties describe only a part of the biological object properties required for diagnostics - the red blood cells concentration, and their state describes a nominal property, for example - shape. With comprehensive information, the probability of correct diagnosis reaches the level necessary for prescribing an adequate effect. The proposed approach is to create a metrological traceability chain of nominal properties, in the form of associated structure with the determined analyte with the metrological traceability chain of quantitative properties.



The living cell and its properties traceability – from the authors

Quantity is a property that is common in nominal terms for various objects, but different in quantitative terms. Quantities are mass, amount of substance, time interval, and other properties that can be measured and expressed as a number. Nominal properties are a separate class of quantities considered by metrology. These are biological properties, classification, identification... Nominal properties include the shape of a living cell (which differs depending on the cell functionality, on its state), the cell membrane structure, the sequence of amino acids or nucleic acids. Nominal properties cannot be measured, they can be examined, which is discussed by G. Nordin et al. Vocabulary on nominal property, examination, and related concepts for clinical laboratory sciences (IFCC-IUPAC Recommendations 2017), Pure Appl. Chem., 2018.

However, quantitative properties describe only a part of the biological object properties required for diagnostics - the red blood cells concentration, and their state describes a nominal property - shape. The use of all available methods of analysis of both measurable quantitative characteristics and nominal properties will obviously increase the probability of correct diagnosis. Adequate prescriptions and treatment methods will be based on comprehensive, complete information. The obviously measurable quantitative properties are traceable to primary standards or primary reference measurement procedures, and further to the definition of the SI unit given by the CQPM. At the head of the traceability chain of the studies results of nominal properties should be the primary reference research procedures, if it is impossible to fulfill this condition - reference methods, or international harmonization protocols, the scheme should include CRM of qualitative properties, routine research procedures up to the result of measurements of a sample of human biological material.



An example of metrological traceability chain of nominal properties – modified from ISO 17511

To assess the identification reliability, we will use a comparison method based on the use of databases. Thus, a substance will be identified if its characteristics (DNA nucleotide sequences) match with the characteristics of one substances from the database. Two data types are used for identification. The first type of data is information from the database obtained by sequencing - this is the basis for comparison - what is compared with. The second type of data is experimentally obtained DNA nucleotide sequences and the corresponding values of research precision - this is what is compared. To enable the application of statistical methods to the identification of the data second type, we will introduce the theory applicability limits. The assumption that allows us to assess the identification reliability is the values of the substances characteristics during examination/measurements behave like random variables, and their distributions are normal. However, we admit that the characteristic of the analyzed substance (sequence) cannot be a random variable for it. At the same time, the difference in sequences for two different substances will already be a random variable, and methods of mathematical statistics can be applied to it. Identification is considered effective if difference between the measured value of the characteristic and the one given in the database is no more than a certain criterion - a tolerance for this value, which is assumed in the range of (2-3) standard deviation values. When carrying out identification, statistical hypotheses H₀ (null hypothesis) and H₁ (alternative hypothesis) are formulated. In our case, the **null hypothesis H₀** is a statement about the absence of a substance (we consider it unknown) in the analyzed sample, i.e. we identify it as absent. **Hypothesis H₁** - the substance is present in the sample. To find out which of the two hypotheses is realized, it is proposed to use the Student's criterion. If the value of the quantile t(0,j) calculated from the experimental data is less than the tabulated one, i.e. the condition is met (1):

$$t(0,j) < t_{N-1,P} \quad (1)$$

then the corresponding hypothesis is rejected, the alternative one is accepted. If the test result showed the presence of a substance in the analyte, i.e. due to measurement error, its result satisfies the tolerance condition, and if on the basis of such a test a decision is made about the presence of a substance in the analyte, then an error of the 1st kind (false positive result) is made. In other words, in case of a type 1 error, hypothesis H₀ is rejected and hypothesis H₁ is accepted. The probability of a type 1 error is denoted by α . Thus, α is the probability of accepting the alternative hypothesis, which is called the significance level. The alternative situation is the validity of hypothesis H₁. If the test result showed the absence of the substance in the analyte, and if on the basis of such a measurement a decision is made about the absence of the substance in the analyte, then a type 2 error (false negative result) is committed. In other words, in case of a type 2 error, hypothesis H₁ is rejected and hypothesis H₀ is accepted. The probability of a type 2 error is denoted by β . Another value is associated with this probability – the power of the criterion M. It is calculated using the formula

$$M = 1 - \beta \quad (2)$$

In particular, the probability of a type 1 error α is a function of the power of M, calculated under the null hypothesis.

For clarity, the situations described above can be presented in the form of a table:

The real situation	Identification result	Nature of the error
is <u>absent</u> from the sample (H ₀)	Identification gives a <u>negative</u> result	<u>There is no</u> (hypothesis H ₀ is accepted)
is <u>absent</u> from the sample (H ₀)	Identification gives a <u>positive</u> result	<u>Type 1 error</u> (false positive result) (hypothesis H ₀ is rejected, hypothesis H ₁ is accepted)
is <u>present</u> in the sample (H ₁)	Identification gives a <u>negative</u> result	<u>Type 2 error</u> (false negative result) (hypothesis H ₁ is rejected, hypothesis H ₀ is accepted)
is <u>present</u> in the sample (H ₁)	Identification gives a <u>positive</u> result	<u>There is no</u> (hypothesis H ₁ is accepted)

The reliability of identification is determined by the function P, associated with the probabilities α and β by the relation: $P = 1 - \alpha - \beta$.

The function P cannot be completely identified with the probability of correct identification. In fact, it can take values from - 1 to + 1 depending on the values of α and β .

CONCLUSION

Thus, the possibility of constructing of metrological traceability chain for the results of examined nominal properties, interconnected through the object of research with the chain of traceability of quantitative results of quantitative properties measurements of a biological object, is shown. For example, the conformation or structure of a protein directly affects its activity. The shape of an erythrocyte, which is a qualitative property and traceable to examined procedure (and not to measurement procedures), tells us about its condition, and as a result, about the health or illness of a person. The concentration of hormones indicates the state of the thyroid gland. And all these properties can be measured or examined, and their traceability is ensured using the above approaches. The solution to a particular problem of determining the reliability of identification of a biological object by simple transformations using methods of mathematical statistics is also shown.