Gold standard method

Gold standard method in testing laboratories: Is it appropriate or a
misconception?

Samara Bin Salem and Premanandh Jagadeesan*

Central Testing Laboratories

Quality and Conformity Council

P.O Box 853, Abu Dhabi, United Arab Emirates

*Corresponding Author: jpanandh@yahoo.com

27 Abstract

28 Gold standard is a very common but most powerful terminology amongst
29 diagnostic and analytical laboratory scientists. The genesis dates back to 19th
30 century monetary system where the standard economic unit of account is based
31 on a fixed quantity of gold and has nothing to do with science. Nevertheless, this
32 term had taken an indispensable place in the testing arena ever since its
33 appearance in reputed scientific journals. Application of this term in analytical
34 science from both proponents and opponents perspective has been discussed in
35 detail. It's applicability with regards to bacteriological culture and pulsed field gel
36 electrophoresis for typing bacteria and clinical diagnosis as examples are
37 presented. In conclusion, diagnostic and analytical test methods should not be
38 described or presented in a misleading manner as it is likely to create an
39 erroneous impression on the use of such terminology. We need to revisit gold
40 standard terminology for its appropriation.

41

42 Key words: gold standard, culture, analytical, colony forming units, diagnostic

43

44 **Introduction**

45 Laboratory testing plays a pivotal role in the assessment of specific infection or
46 contaminants which may otherwise cause serious health issues. Hence analytical
47 testing methods are developed to study quantitative and/or qualitative presence
48 of harmful microbes/contaminants in food, water and clinical specimen. A typical
49 analytical test is expected to be fit for its intended purpose with sufficient
50 reliability that any decision based on it can be taken with confidence. Any new
51 method is validated against the existing reference/gold standard method. In
52 general, an analytical procedure sufficiently free of random or systematic error to
53 make it useful for validating proposed new analytic procedures for the same
54 analyte is defined as reference method/gold standard. Indeed, is not the perfect
55 method, but merely the best available one that has a standard with known result
56 (Cardoso *et al.*, 2014).

57 Gold standard is a very common but most powerful terminology amongst
58 diagnostic and analytical laboratory scientists. Although analytical and diagnostic
59 performances are two separate concepts, their relation lies in the fact that a poor
60 analytical performance is likely to have poor diagnostic performance (Flatland *et*
61 *al.*, 2014). In general, non-clinical laboratory testing such as food, water and
62 environment are considered as analytical testing while clinical lab tests are
63 referred as diagnostic tests. Nevertheless, it is very common that both clinical
64 and non-clinical labs use the term Gold standard as the reference method with
65 high degree of recognition.

Gold standard method

The genesis dates back to 19th century monetary system where the standard economic unit of account is based on a fixed quantity of gold. In other words, it was more of an economical term used as the definitive exemplar of quality and reliability. Nevertheless, application of this term was highly contested amongst the financial community leading to abandonment. On the contrary, this phrase had taken an indispensable place in the testing arena ever since its appearance in reputed scientific journals like Lancet and New England Journal of Medicine (Jones & Podolsky, 2015).

Appropriateness of the term seems valid based on the consensus that a given test in a given situation is the best and hence serves as gold standard to which newer tests can be compared (Versi, 1992). On the contrary, testing methods are constantly challenged and superseded with new ones. But who is omniscient to make such statement and decision as there is always something to supersede.

In the case of food and water microbial testing, bacteriological culture remains gold standard till date despite enigmatic culture anomalies. For instance, the ability of *Escherichia coli* to survive sub-lethal environmental stress through viable but non-cultivable cells (VNBC) state (Pienaar *et al.*, 2016) undermines their enumerations by existing culture methods. In addition oxidative stress influenced colony growth also contribute to underestimation of this technique.

A several orders of magnitude of discrepancy between viable plate count and total direct microscopic count of bacteria in aquatic habitats has already been reported, a phenomenon that is very well known for many generations but conveniently often discounted (Cundell, 2015). Nonetheless, culture is still

Gold standard method

regarded as Gold standard test for bacterial enumeration and detection (Tanaka *et al.*, 2014).

According to United States Pharmacopeia (USP, 2015), microorganisms existing in aggregates (e.g., clumps or chains) would normally grow into a single colony forming units (CFU), miscalculating the exact number of cells present. Therefore, it is not appropriate to consider only the unit of microbiological enumeration highlighting its inappropriateness as a gold standard. Similarly, cell culture is still regarded as gold standard for clinical diagnosis of *Chlamydia trachomatis*, a sexually transmittable infection despite its poor sensitivity over nucleic acid amplification tests (Janssen *et al.*, 2016).

Yet another typical example is pulsed-field gel electrophoresis (PFGE) which has been the laboratory gold standard for foodborne bacterial subtyping for almost two decades. Nevertheless, studies are mounting to demonstrate the discriminatory power of whole genome sequencing based approaches over the currently designated gold standard method (PFGE). Although the method served meaningful comparison of historical isolates for years (Rumore, *et al.*, 2018), its current affiliation is not appropriate especially when other well validated superior methods are available (Tang *et al.*, 2019).

From the analytical perspective, gold Standard is expected to be a perfect test without classification error. But in reality, a perfect gold standard is less often available than an imperfect gold standard which is measured with error according to Maxim and coworkers (2014). Therefore it is important to understand that gold

standard test does not mean ‘perfect test’. It is a term used to refer a benchmark method for other tests until other methods prove otherwise.

Scientific advancements results in better understanding and novel methods are continuously being developed that usually supersede existing techniques in the diagnostic arena. Thus improvisation is an ongoing phenomenon where today’s most preferred test method may become tomorrow’s least preferred ones.

Concluding remarks:

In conclusion, uncertainty is central to science and pervades almost all scientific activities (Jensen et al., 2017) and testing method is not an exception. In analytical science, existing diagnostic accuracies are constantly challenged or improvised by new developments. Thus diagnostic and analytical test methods should not be described or presented in a misleading manner as it is likely to create an erroneous impression on the use of such terminology. We need to revisit gold standard terminology for its appropriation.

Acknowledgments

The authors gratefully acknowledge the support and encouragement of the Abu Dhabi Quality and Conformity Council. Gratitude is also extended to the management of central testing laboratories for invaluable support.

References

Cardoso, J. R., Pereira, L. M., Iversen, M. D., & Ramos, A. L. 2014. What is gold standard and what is ground truth?. *Dental Press J Orthod.* 19(5), 27–30.

Gold standard method

- 132 Flatland, B., Freeman, K. P, Vap, L. M., and Harr, K. E. 2013. ASVCP
133 guidelines: quality assurance for point-of-care testing in veterinary medicine. *Vet*
134 *Clin Pathol.* 42: 405–423.
- 135 Jones, D. S, and Podolsky, S. H. 2015. The history and fate of the gold
136 standard. *Lancet.* 385:1502-1503.
- 137 Versi, E. “Gold standard” is an appropriate term. 1992. *BMJ.* 305:187.
- 138 Pienaar, J. A., Singh, A., and Barnard, T. G. 2016. The viable but
139 nonculturable state in pathogenic *Escherichia coli*: a general review. *Afr J Lab*
140 *Med.* 5: 9.
- 141 Cundell, T. 2015. “The limitations of the colony-forming unit in
142 microbiology,” *Eur Pharm Rev.* 20: (6) 11–13.
- 143 Tanaka, T., Kawasaki, K., Daimon, S., Kitagawa, W., Yamamoto, K.,
144 Tamaki, H., et al. 2014. A hidden pitfall in the preparation of agar media
145 undermines microorganism cultivability. *Appl Environ Microbiol.* 80: 7659–7666.
- 146 United States Pharmacopeia. 2015. <1223> Validation of Alternative
147 Microbiological Methods. United States Pharmacopeial Convention. USP
148 38/NF33:1439.
- 149 Rumore, J., Tschetter, L., Kearney, A., Kandar, R., McCormick, R., Walker
150 M, et al. 2018. Evaluation of whole-genome sequencing for outbreak detection of
151 verotoxigenic *Escherichia coli* O157:H7 from the Canadian perspective. *BMC*
152 *Genomics.* 19(1):870.

Gold standard method

153 Tang S, Orsi R. H, Luo H, Ge C, Zhang G, Baker RC, Stevenson A and
154 Wiedmann, M. 2019. Assessment and Comparison of Molecular Subtyping and
155 Characterization Methods for Salmonella. *Front Microbiol.* 10: 1591.

156 Janssen, K. J. H., Hoebe, C. J. P. A., Dukers-Muijrers, N. H. T. M.,
157 Eppings, L., Lucchesi, M., and Wolffs, P. F. G. 2016. Viability-PCR Shows That
158 NAAT Detects a High Proportion of DNA from Non-Viable Chlamydia
159 trachomatis. *PLoS ONE* 11(11): e0165920.

160 Maxim, L. D., Niebo, R., and Utell, M. J. 2014. Screening tests: a review
161 with examples. *Inhal Toxicol.* 26: 811–28.

162 Jensen, J. D., Pokharel, M., Scherr, C. L., King, A. J., Brown, N., and
163 Jones, C. 2017. Communicating uncertain science to the public: How amount
164 and source of uncertainty impact fatalism, backlash, and overload. *Risk Analysis.*
165 37: 40-51.

166