Reducing Uncertainty and Increasing Confidence in Research Results

Anne Plant
NIST Fellow
National Institute of Standards and Technology
anne.plant@nist.gov

National Institute of Standards and Technology (NIST)

- National Metrology Institute for the United States
- Mission:

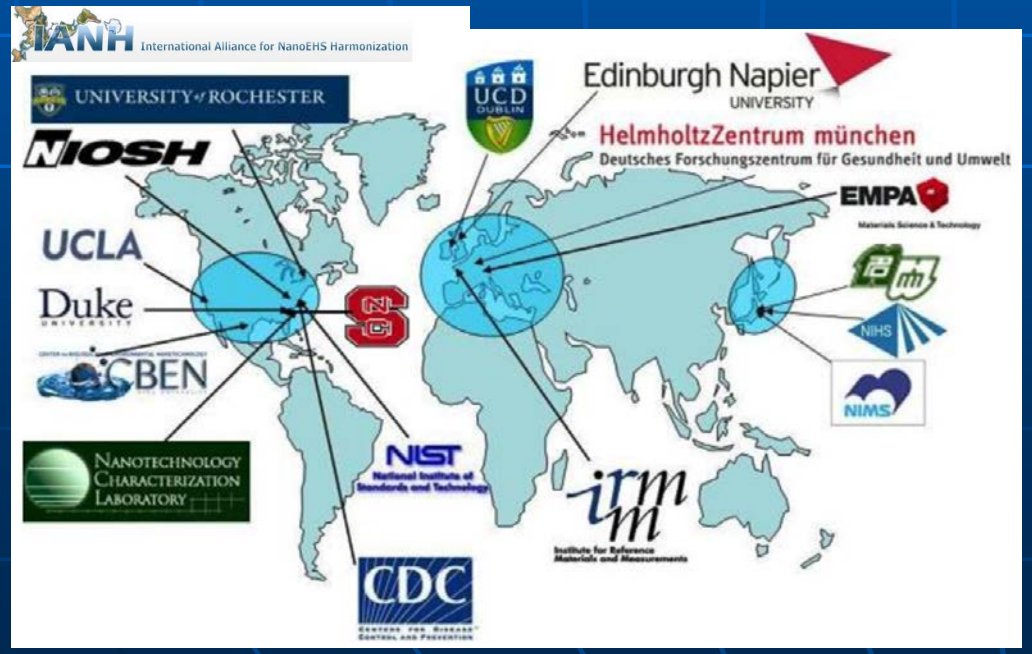
  To promote U.S. innovation and industrial competitiveness by advancing **measurement science, standards**, and **technology** in ways that enhance economic security and improve our quality of life

Measurement science in international policy and commerce

International standards and trade agreements

International Measurement Laboratories (BIPM)

International standards and harmonization groups



BIPM

JCTLM
Accurate results for patient care

IMDRF International Medical Device Regulators Forum

ASTM INTERNATIONAL

ISO

FDA

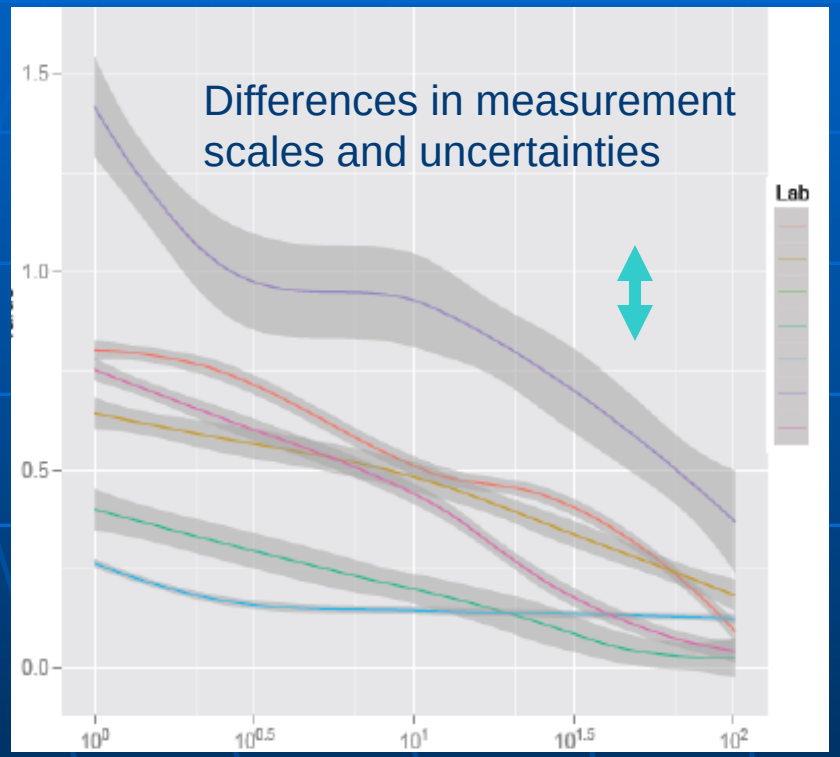
ICH
harmonisation for better health

IABS

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

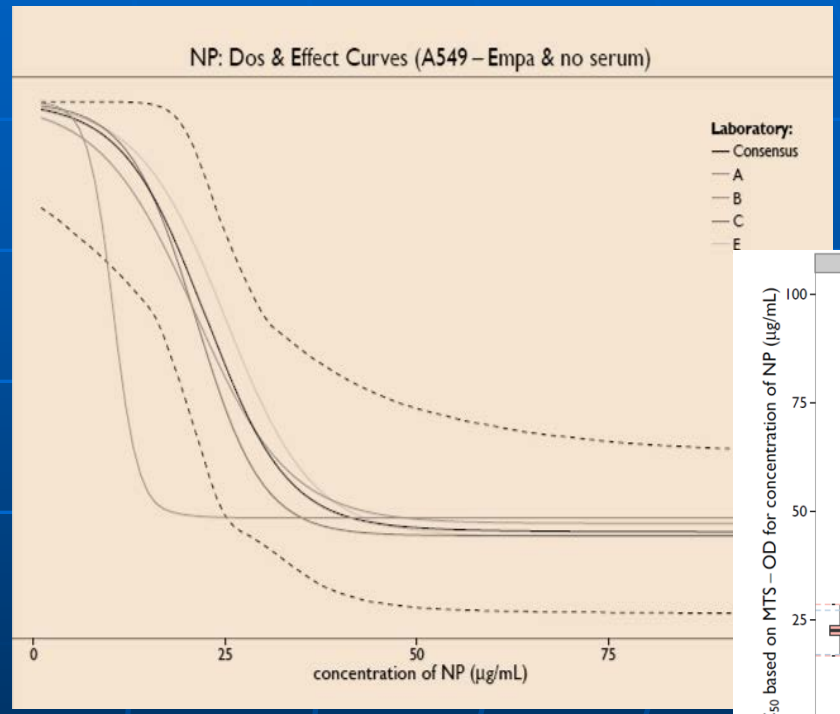
US interests in international policy and commerce

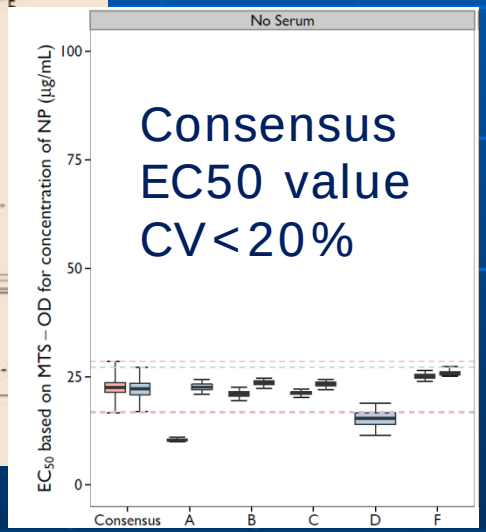
Improving Nanoparticle Toxicity Assay: Interlaboratory Comparison



Design Of Experiment

Identify/ control sources of uncertainty

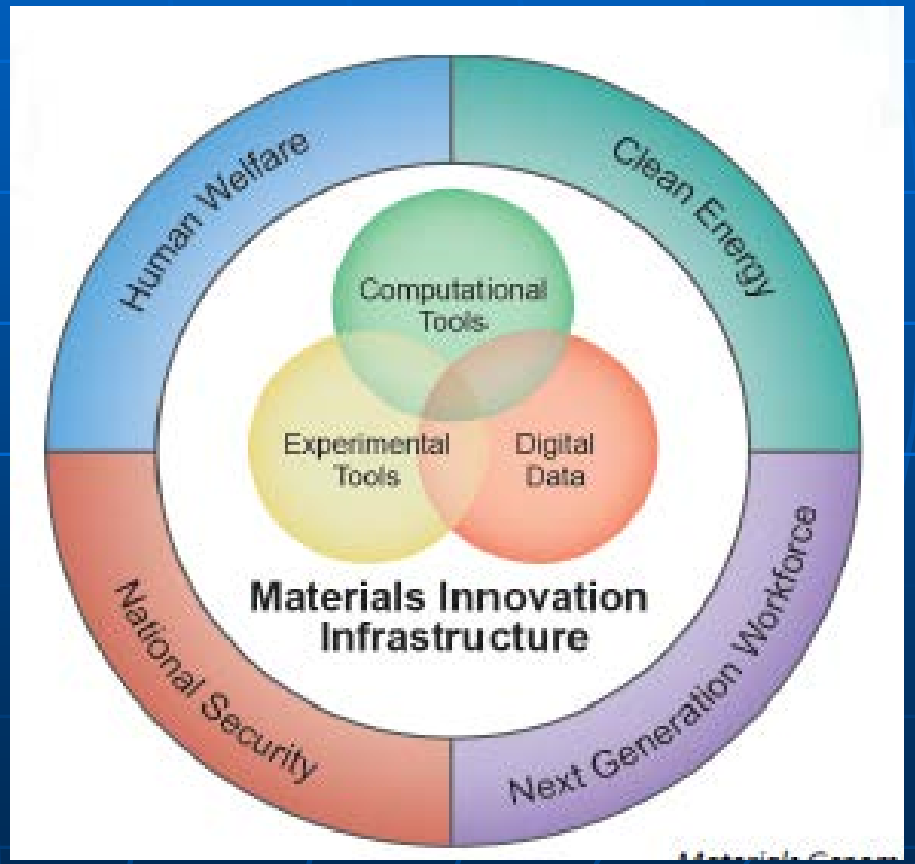




The Materials Genome Initiative (MGI)

Goal: to decrease the time to market for new materials by 50%

Data ⟺ **Models**



Jim Warren,Technical Program Director, MGI, NIST

The Materials Genome Initiative (MGI)

Goal: to decrease the time to market for new materials by 50%

CoRR

Assess & Improve ***Quality***

- Validate Experiments and Models
- Verify Model accuracy
- Quantify Uncertainty
- Quantify Sensitivity
- Define Next Generation of Experiments and Models

Cloud of Reproducible Records.
An important part of MGI at NIST is infrastructure and tools to enable reproducible research in computational and experimental materials science

Jim Warren,Technical Program Director, MGI, NIST

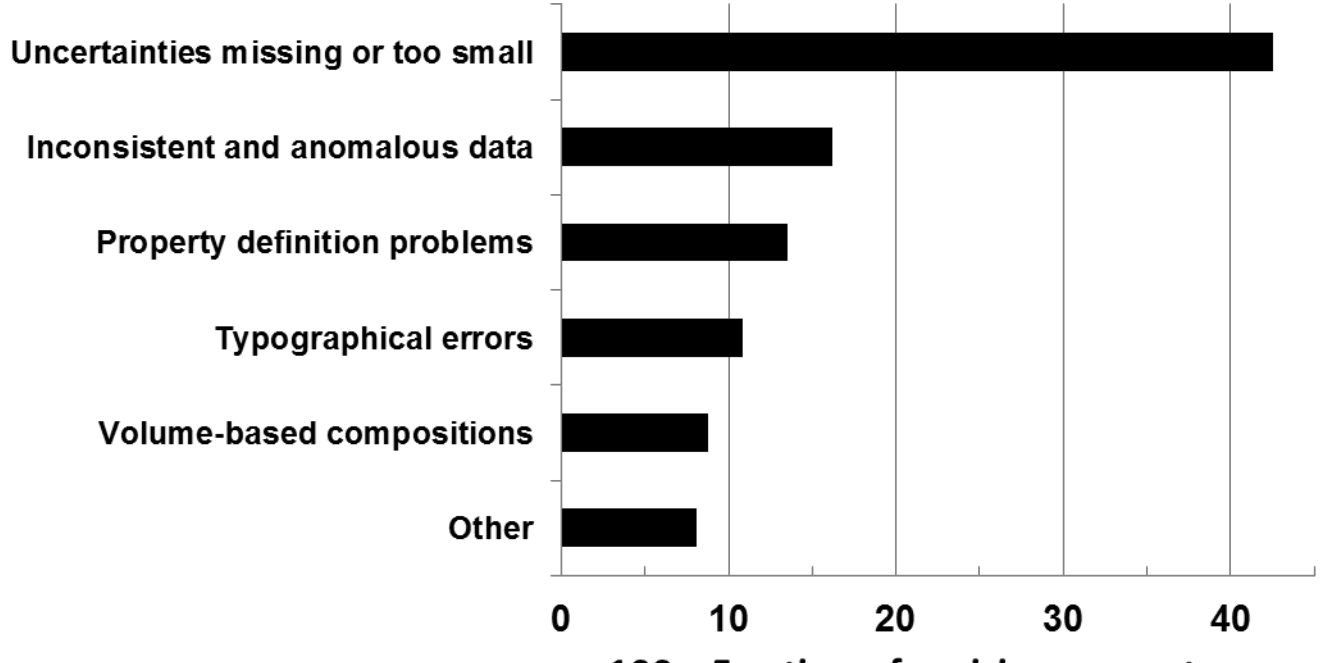
Journal of chemical & engineering data

Improvement of Quality in Publication of Experime
Thermophysical Property Data: Challenges, Assessr
Global Implementation, and Online Support

Robert D. Chirico,[a,*] Michael Frenkel,[a] Joseph W. M
Andrei F. Kazakov,[a] Kenneth Kroenlein,[a] Ilmutdin Ab
Joan F. Brenneke,[c] Paul L. Brown,[d] Peter T. Cummi
Anthony R. H. Goodwin,[h] Lee D. Hansen,[i] William M
Kenneth N. Marsh,[l] Paul M. Mathias,[m] Clare McCabe,
Christoph Schick,[q] J. P. Martin Trusler,[r] Sergey Vyaz

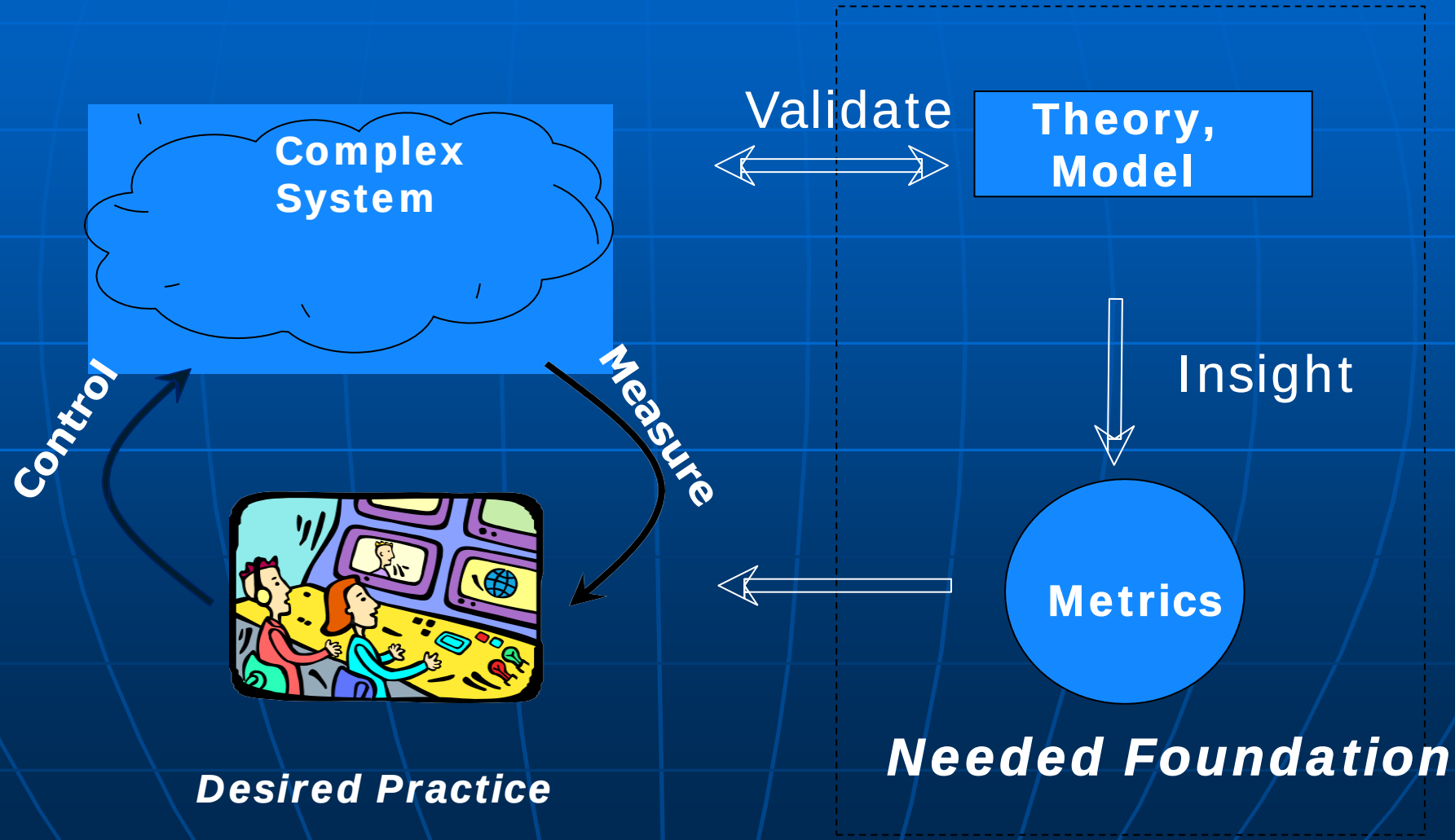
[a]Applied Chemicals and Materials Division, National Institute of Stand
[b]Department of Chemistry, University of North Texas, NT Station
[c]Department of Chemical and Biomolecular Engineering, Universit
[d]Rio Tinto Technology and Innovation, 1 Research Avenue, Bund
[e]Department of Chemical and Biomolecular Engineering, Vanderbi
[f]Department of Process and Energy, Engineering Thermodynamics
2628 CA Delft, Netherlands
[g]National Institute of Standards and Technology, Boulder, Colorad
[h]Schlumberger Technology Corporation, 150 Gillingham Lane, MI
[i]Department of Chemistry and Biochemistry, Brigham Young Univ
[j]Department of Science Education, Graduate School of Education,
Higashi-Hiroshima City, Japan
[k]Department of Mechanical and Industrial Engineering, University of T
[l]School of Mechanical and Chemical Engineering, University of W

Approximately ⅓ of articles that reach the “approve” stage are found to contain significant problems that require further revision



Foundations of Measurement Science for Information Systems

Experiment– Model – Understand– Measure – Improve



Ron Boisvert, Chief, Mathematical & Computational Sciences Division, NIST

Addressing Reproducibility via Publication Policies and Incentives within the Association for Computing Machinery (ACM)

- SIGMOD: establishing a reproducibility review of papers at the SIGMOD conference since 2008.
- The Artifact Evaluation for Software Conferences: formal evaluations of artifacts (such as software and data) for 11 major conferences since 2011.
- *ACM Transactions on Mathematical Software,* The *ACM Transactions on Modeling and Computer Simulation,* and the *Journal of Experimental Algorithmics*: initiatives for replication of computational results by a third party before acceptance.
- ACM Publications Board: policy on Result and Artifact Review and Badging to highlight papers that have undergone independent verification.

CRITERIA: documented, consistent, complete, exercisable, and include appropriate evidence of verification and validation.

Ronald Boisvert, the head of the Applied and Computational Mathematics Division at NIST and chair of the ACM Publications Board's Digital Library Committee. Communications of the ACM, Vol. 2016

NIST Products and Services

Measurement Research

- ~ 2,200 publications per year

Standard Reference Data

- ~ 100 different types
- ~ 6,000 units sold per year
- ~ 226 million data downloads per year

Standard Reference Materials

- ~ 1,300 products available
- ~ 30,000 units sold per year

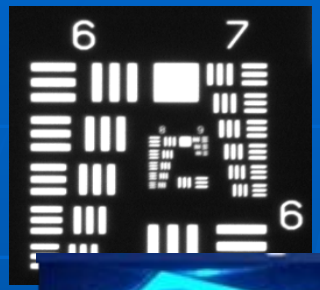
Calibration Tests

- ~ 18,000 tests per year

Laboratory Accreditation

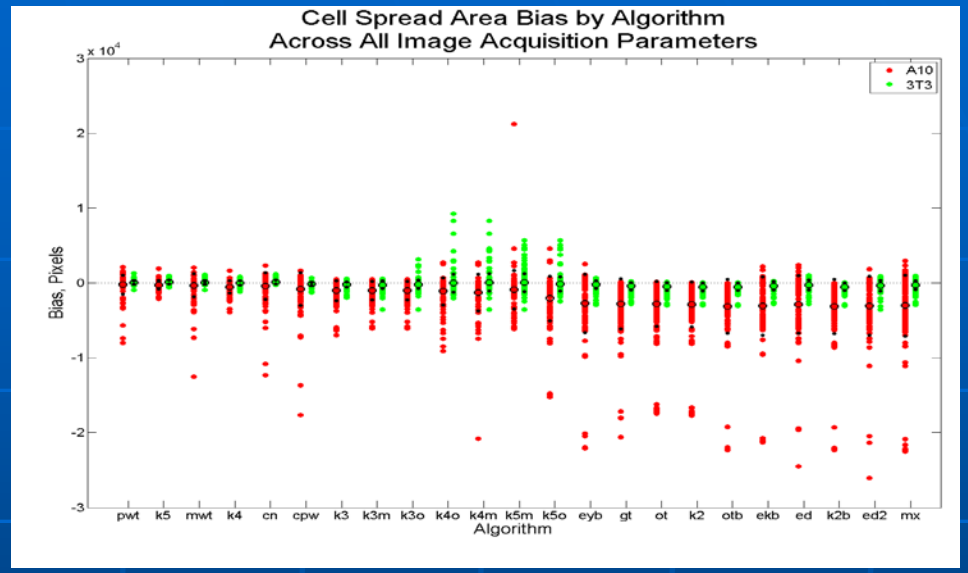
- ~ 800 accreditations of testing and calibration laboratories

Measurement assurance : Method validation

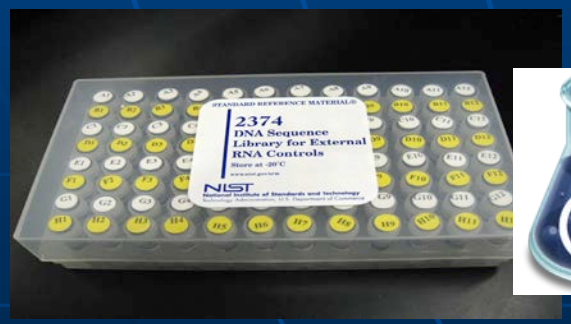


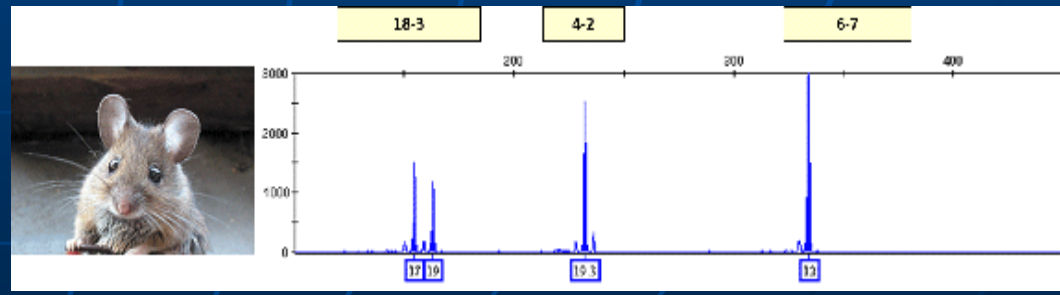
Halter et al Cytometry A 2014

- Benchmarking microscope performance.
- Measurement assurance of NA sequence
- Confirmation of cell line identity.
- Accuracy of image analysis algorithms.



Dima et al Cytometry A 2011





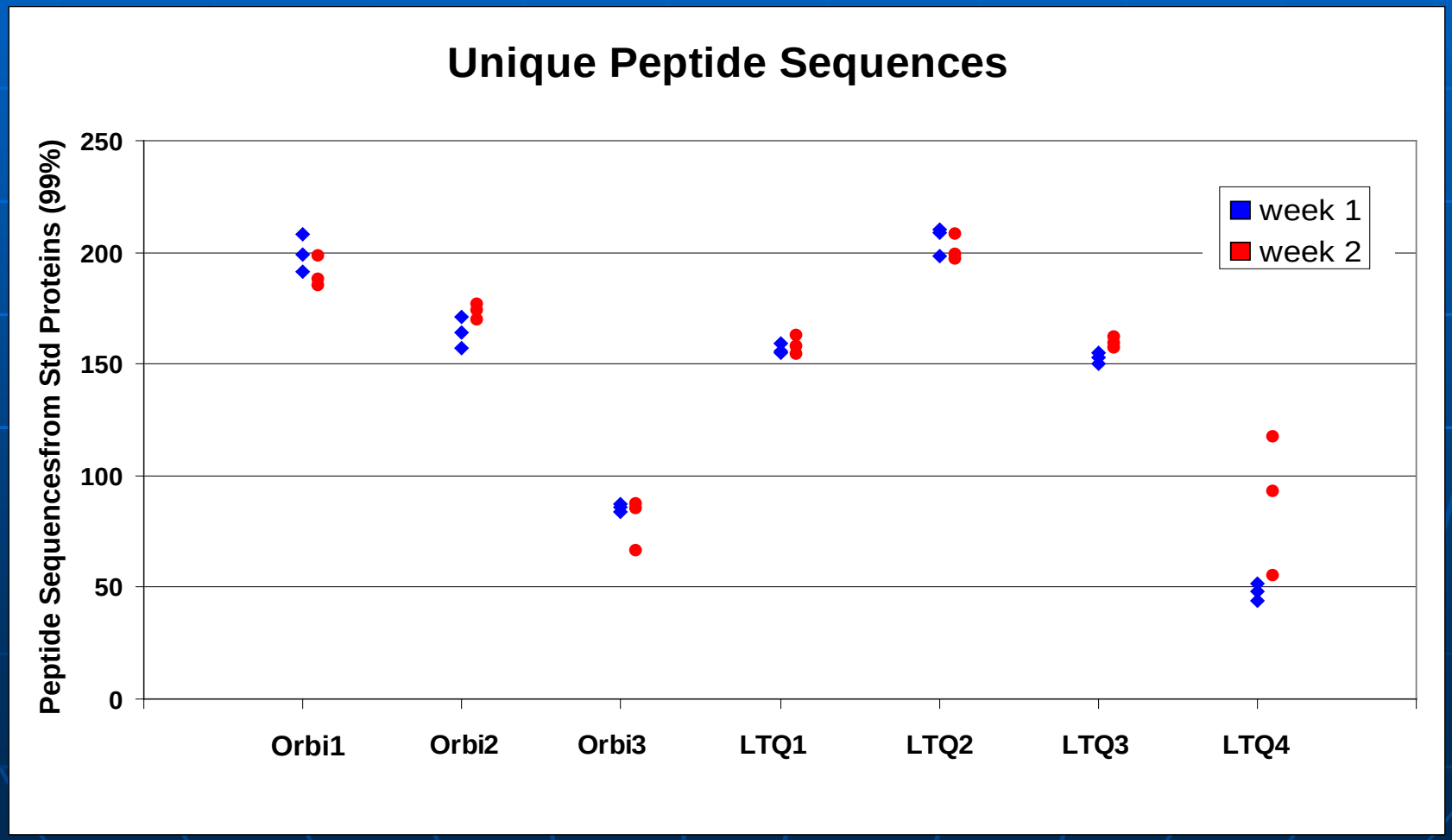
Almeida et al Cytotechnology 2013

Many laboratories don't think a lot about the quality of their measurements

Often hear: We don't have problems in our lab.
And: Our results are very reproducible.
And: It works for us.

Mass Spec Proteomics: Interlab Comparison

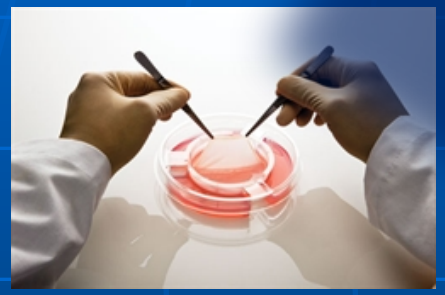
CPTAC Study 2 (SOP)



Transferring laboratory research to applications in health care: What does it take?

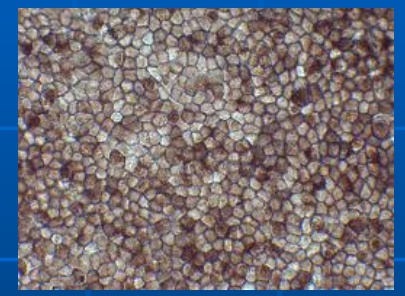
Challenge to commercialization of RM/CT products: Measurements

Comparability



I. Characterization of **product**

Assays to measure critical **quality attributes**

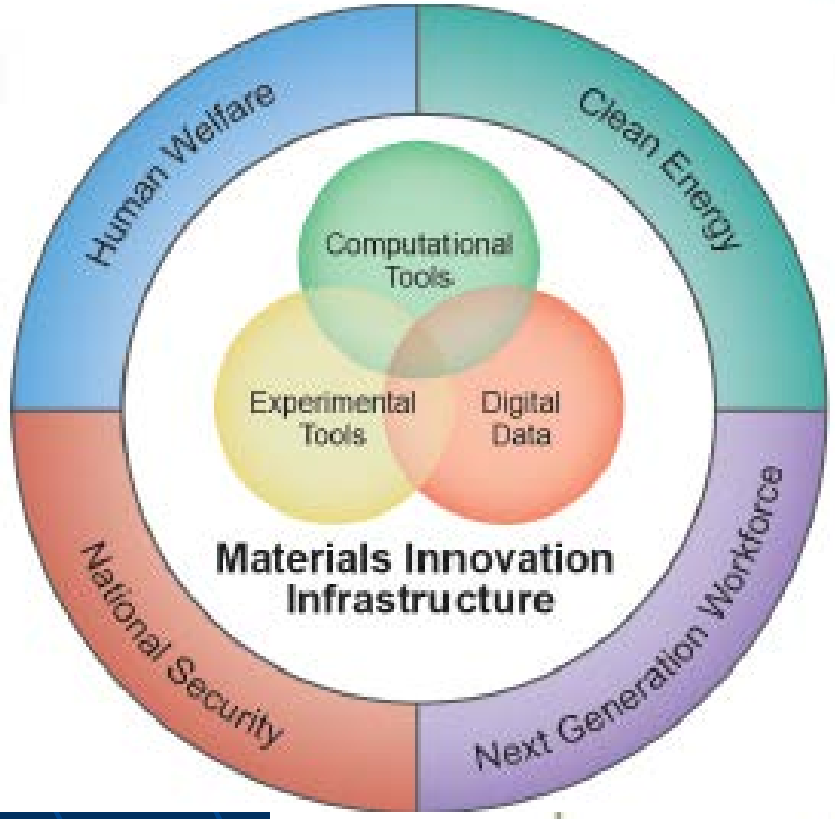


II. Control of the manufacturing process

Assays that assure consistency of product during

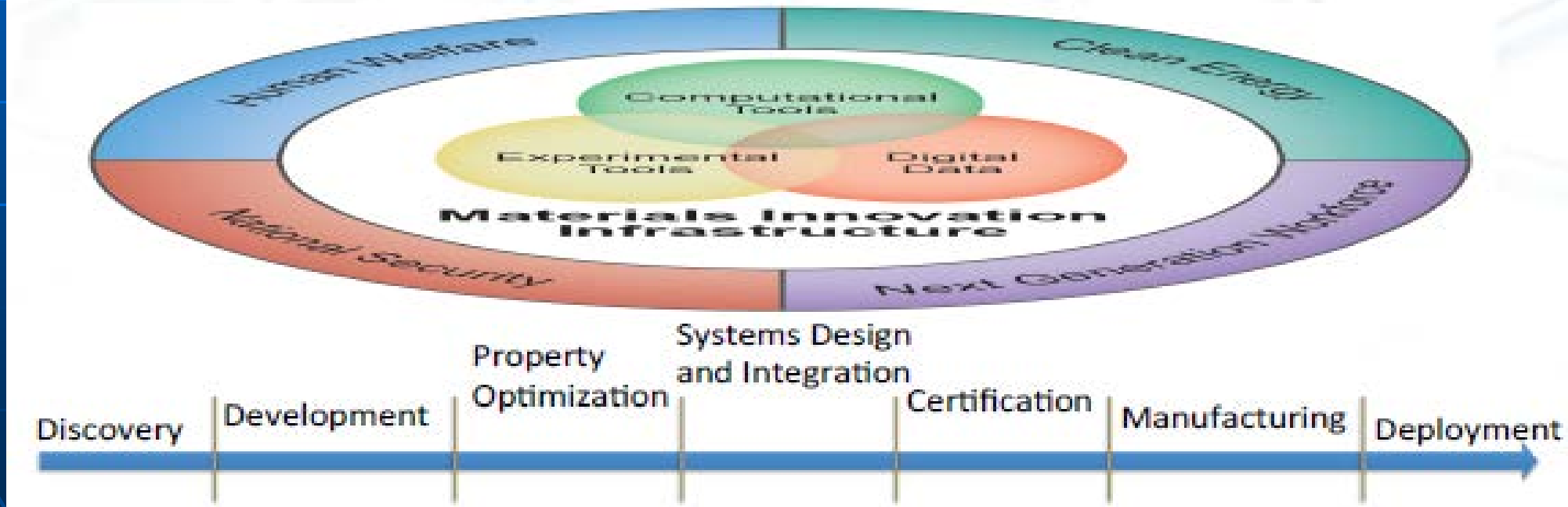
- Scale Up
- Change in personnel, processes, location
- Storage
- New analytical and culture equipment
- Improved assays
- Changes in raw materials

"If you can not measure it, you can not improve it." Lord Kelvin



IGNING MATERIALS TODAY

DESIGNING NEW MATERIALS IN THE FUTURE



How important is reproducibility

Establishing Confidence in Measurements : reproducibility is one element

COMMENTARY

Improved reproducibility by assuring confidence in measurements in biomedical research

Anne L Plant, Laurie E Locascio, Willie E May & Patrick D Gallagher

'Irreproducibility' is symptomatic of a broader challenge in measurement in biomedical research. From the US National Institute of Standards and Technology (NIST) perspective of rigorous metrology, reproducibility is only one aspect of establishing confidence in measurements. Appropriate controls, reference materials, statistics and informatics are required for a robust measurement process. Research is required to establish these tools for biological measurements, which will lead to greater confidence in research results.

A number of recent reports in the peer-reviewed literature[1,2] and in the popular press[3,4] have discussed irreproducibility of results in biomedical research. Some of these articles suggest that the inability of independent research laboratories to replicate published results has a negative impact on the development of, and confidence in, the biomedical research enterprise. One outcome of these concerns is the effort to test the reproducibility of

opportunities to recognize the underlying sources of irreproducibility, which may be inherently important for understanding the biological process and how to measure it.

It is also important to appreciate that reproducibility does not guarantee accuracy. A good measurement should meet a number of criteria: accuracy, precision, robustness, specificity and a well-defined detection limit and response function (Table 1). Although attention to

Replicate measurements quantify the dispersion in analytical values, but other aspects of measurement (Table 1) are also critical for establishing confidence in the data. The accuracy of an assay, for example, can be tested by comparing different protocols and by employing orthogonal measurement methods because different techniques are likely to be biased in different ways. A relevant study examined different fixation techniques,

Accuracy: Orthogonal method

Precision: Repeatability (replicates);

Reproducibility (different days, locations, technicians

Ruggedness: sensitivity to assay parameters

Dynamic range and Response function: Instrument benchmarking. +/- controls. Calibration curve. Limit of detection

Specificity: sensitivity to matrix effects

Identifying and mitigating the sources of uncertainty in a research study

- **Planning** (goals, generalizability, experimental design, preliminary data, assumptions)
- **Identifying systemic sources of bias and uncertainty** (characterize reagents, equipment response)
- **Characterizing the quality of experimental data** (precision, accuracy, dynamic range, robustness)
- **Minimizing bias in data reduction and interpretation of results** (basis of statistical tests, evaluate algorithms, compare results with previous data, identify uncontrolled sources of uncertainty, consider alternative interpretations)
- **Reporting and dissemination** (document data and code, protocols and process,