NIST/NIBIB Symposium on Medical Metrology and Standards for American Healthcare and Commerce

Impact of medical metrology and standards in the clinic

Daniel C. Sullivan, M.D.

Duke University Medical Center

NIST/NIBIB Symposium on Medical Metrology and Standards for American Healthcare and Commerce

Disclosures: None

Daniel C. Sullivan, M.D.

Duke University Medical Center

Outline

1. Framing the discussion
2. Basic Concepts
3. Clinical Priorities

Why Are We Here?

Clinical decision variability is a huge problem.



Example: wide variation is inherent in subjective, qualitative radiologist interpretations

One example: Recommendations for follow-up of pancreatic lesions for each radiologist

Radiologist subjective interpretations
reflect a combination of

1. his/her skill and

2. value judgements.

Both components are very resistant to standardization.

This observation is not unique to Radiology. It contributes to variability across all aspects of healthcare.

Editorial Comment:

- **“Individual patients, referring physicians, and society as a whole cannot possibly accept this degree of variability.”**

Macari & Megibow, Radiol, Apr 2011

ELSEVIER

CrossMark

THE SPINE JOURNAL

The Spine Journal 1 (2017) 54–561

Clinical Study

Variability in diagnostic error rates of 10 MRI centers performing lumbar spine MRI examinations on the same patient within a 3-week period

"… a patient should expect to receive the same radiological diagnosis regardless of which imaging center he or she visits, or which radiologist reviews the examination."

"Based on our extensive clinical experience, the authors believe that this assumption is not correct and that it can negatively impact patient care, outcomes, and costs."

STUDY DESIGN: This is a prospective observational study comparing the interpretive findings reported for one patient scanned at 10 different MRI centers over a period of 3 weeks to each other and to reference MRI examinations performed immediately preceding and following the 10 MRI examinations.

PATIENT SAMPLE: The sample is a 63-year-old woman with a history of low back pain and right L5 radicular symptoms.

Thought Experiment:

Deploy a suite of AI algorithms for spine, providing each radiologist with quantitative results for:

Need VERY LARGE training datasets of standardized, population-relevant , high-quality scans.

- Gray matter cross-sectional areas
- Alignment vectors
- Bone density values
- Facet joint alignment
- Foraminal sizes
- Etc.

Outline

1. Framing the discussion
2. **Basic Concepts**
3. Clinical Priorities

Setting the stage …

Precision Medicine …

Underlying assumption is that each patient has a unique combination of genotypic, epigenetic, and phenotypic profiles for his/her disease.

Precision medicine: tailoring diagnosis and treatment to the individual patient, using patient-specific biomarkers rather than population averages.

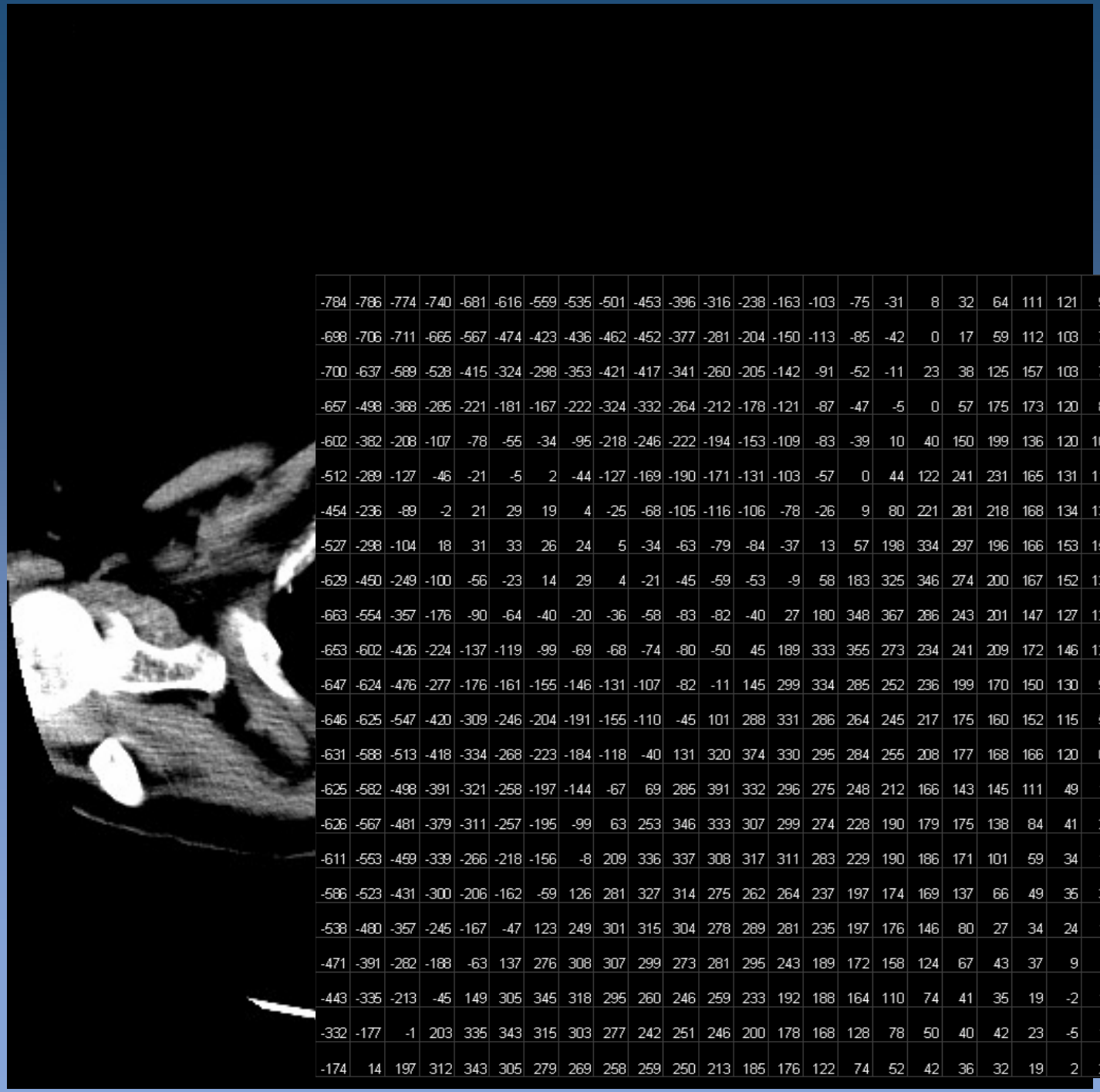
Basic concepts: Biomarkers …

- A characteristic that is objectively measured and evaluated as an indicator of normal biologic or pathogenic processes or pharmacological responses to a therapeutic intervention.

NIH Workshop definition (1999)

Basic concepts: What is an Image?

- an *n*-D data set, where n>1;
- each voxel value provides spatially- and/or temporally localized information about some property of matter with which the energy has interacted;
- "interpretation"/quantification requires extracting a sub-set of data ("segmentation");
- In biology, usually a noisy, blurry dataset.



-784	-786	-774	-740	-681	-616	-559	-535	-501	-453	-396	-316	-238	-163	-103	-75	-31	8	32	64	111	121	97
-698	-706	-711	-665	-567	-474	-423	-436	-462	-452	-377	-281	-204	-150	-113	-85	-42	0	17	59	112	103	74
-700	-637	-589	-528	-415	-324	-298	-353	-421	-417	-341	-260	-205	-142	-91	-52	-11	23	38	125	157	103	74
-657	-498	-368	-285	-221	-181	-167	-222	-324	-332	-264	-212	-178	-121	-87	-47	-5	0	57	175	173	120	88
-602	-382	-208	-107	-78	-55	-34	-95	-218	-246	-222	-194	-153	-109	-83	-39	10	40	150	199	136	120	107
-512	-289	-127	-46	-21	-5	2	-44	-127	-169	-190	-171	-131	-103	-57	0	44	122	241	231	165	131	115
-454	-236	-89	-2	21	29	19	4	-25	-68	-105	-116	-106	-78	-26	9	80	221	281	218	168	134	130
-527	-298	-104	18	31	33	26	24	5	-34	-63	-79	-84	-37	13	57	198	334	297	196	166	153	151
-629	-450	-249	-100	-56	-23	14	29	4	-21	-45	-59	-53	-9	58	183	325	346	274	200	167	152	139
-663	-554	-357	-176	-90	-64	-40	-20	-36	-58	-83	-82	-40	27	180	348	367	286	243	201	147	127	120
-653	-602	-426	-224	-137	-119	-99	-69	-68	-74	-80	-50	45	189	333	355	273	234	241	209	172	146	127
-647	-624	-476	-277	-176	-161	-155	-146	-131	-107	-82	-11	145	299	334	285	252	236	199	170	150	130	92
-646	-625	-547	-420	-309	-246	-204	-191	-155	-110	-45	101	288	331	286	264	245	217	175	160	152	115	57
-631	-588	-513	-418	-334	-268	-223	-184	-118	-40	131	320	374	330	295	284	255	208	177	168	166	120	67
-625	-582	-498	-391	-321	-258	-197	-144	-67	69	285	391	332	296	275	248	212	166	143	145	111	49	18
-626	-567	-481	-379	-311	-257	-195	-99	63	253	346	333	307	299	274	228	190	179	175	138	84	41	24
-611	-553	-459	-339	-266	-218	-156	-8	209	336	337	308	317	311	283	229	190	186	171	101	59	34	17
-586	-523	-431	-300	-206	-162	-59	126	281	327	314	275	262	264	237	197	174	169	137	66	49	35	37
-538	-480	-357	-245	-167	-47	123	249	301	315	304	278	289	281	235	197	176	146	80	27	34	24	14
-471	-391	-282	-188	-63	137	276	308	307	299	273	281	295	243	189	172	158	124	67	43	37	9	1
-443	-335	-213	-45	149	305	345	318	295	260	246	259	233	192	188	164	110	74	41	35	19	-2	14
-332	-177	-1	203	335	343	315	303	277	242	251	246	200	178	168	128	78	50	40	42	23	-5	10
-174	14	197	312	343	305	279	269	258	259	250	213	185	176	122	74	52	42	36	32	19	2	20

Basic concepts:
Definition of Quantitative Imaging

- Quantitative imaging is the development, standardization, and optimization of anatomical, functional, and molecular imaging acquisition protocols, data analyses, display methods, and reporting structures in order to permit the validation of accurately and precisely obtained image-derived metrics with physiologically relevant parameters, and the use of such metrics in clinical research and patient care.

Basic concepts:
Quantitative Imaging Biomarker…

- A characteristic derived from an vivo image and objectively measured as an indicator of normal biological processes, pathogenic processes or a response to a therapeutic intervention

- A quantitative imaging biomarker (QIB) is an ***imaging assay.***

Basic concepts:
QIB Technical Parameters…

- **Metrology** is the science of measurement:
- "A measurement result is complete only when it is accompanied by a statement of the associated uncertainty,."

- NIST Policy Statement

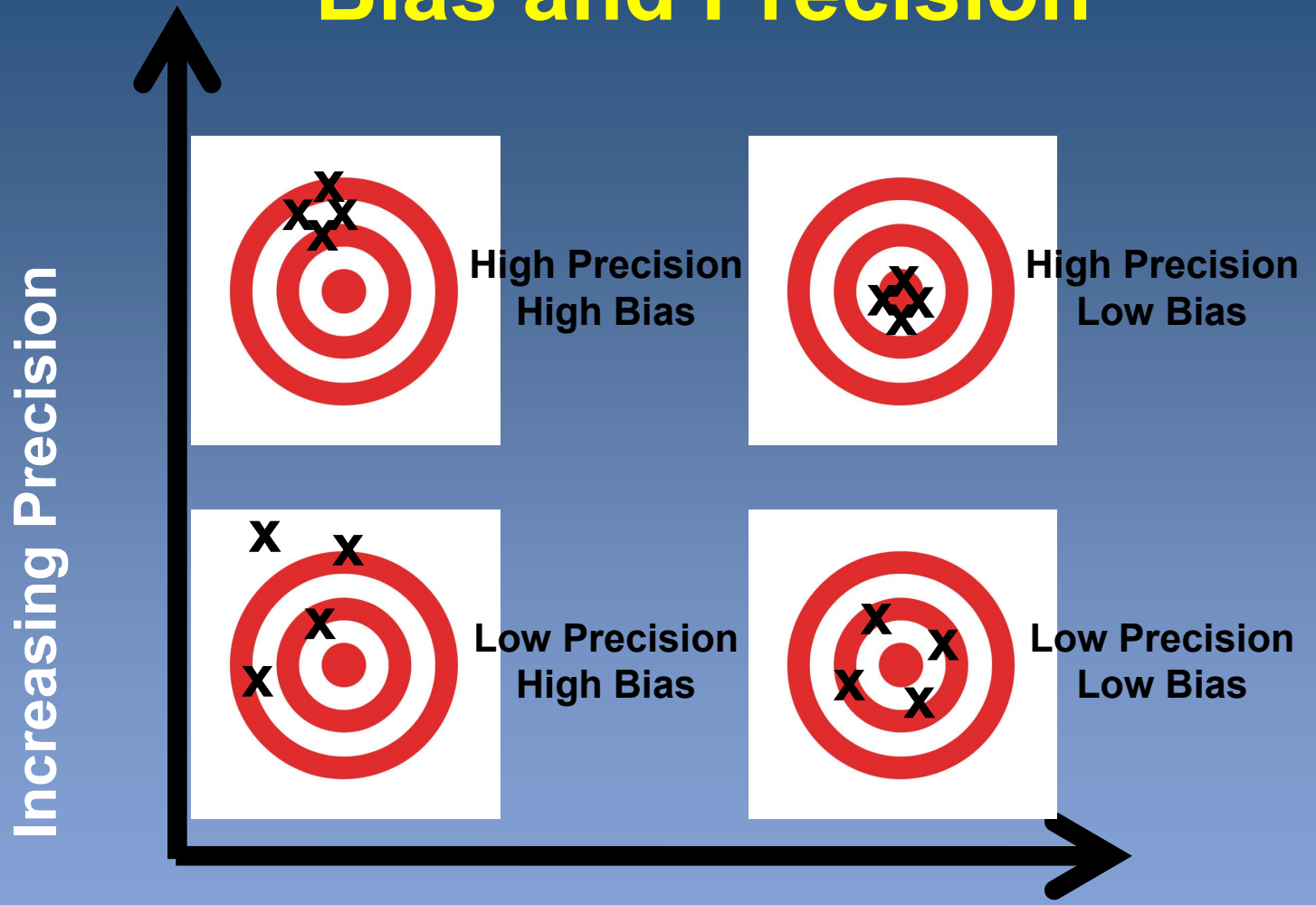
NIST NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY U.S. DEPARTMENT OF COMMERCE

- For QIBs, "Uncertainty" is strongly determined by:
  - Bias
  - Precision

Basic concepts:
Bias and Precision

Bias = Systematic measurement error;

Precision = Closeness of agreement.



High precision is good;

Low bias is good.

Basic concepts: Precision…

1. Repeatability
   - precision that occurs with identical or near-identical conditions.
2. Reproducibility
   - precision when location, operator, measuring system, or other factors differ.

Outline

1. Framing the discussion
2. Basic Concepts
3. **Clinical Priorities**

Setting Priorities for U.S. Healthcare Measurement

From this 2015 Federal Initiative on Precision Medicine –

Six major clinical areas of Precision Medicine (PM) interest emerged:

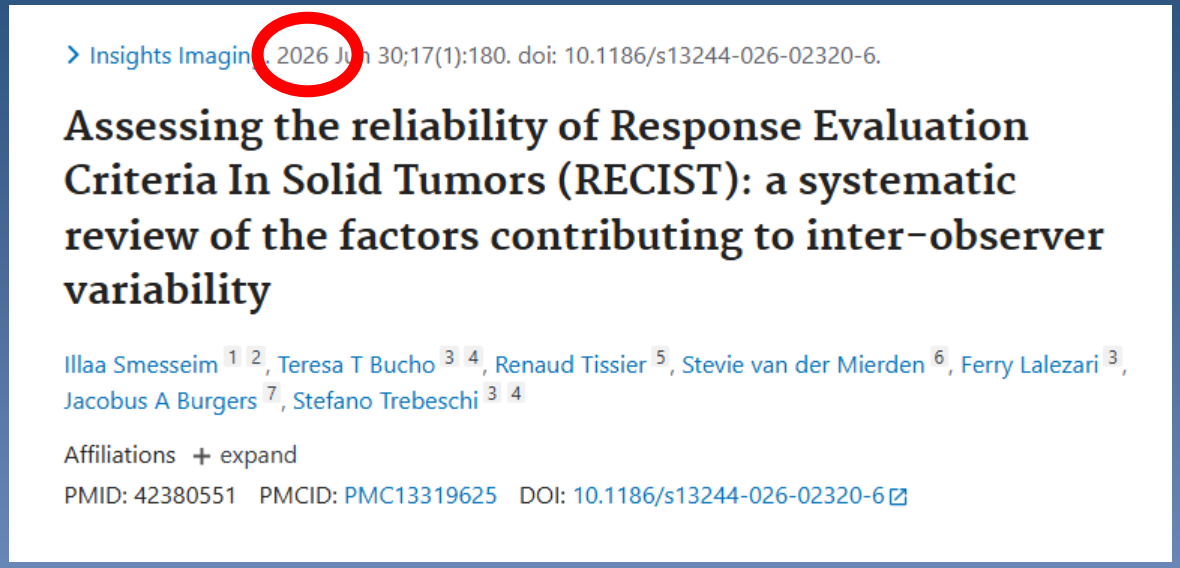
- oncology
- brain disorders
- cardiovascular disorders
- infections
- metabolic disorders, and
- other inflammatory disorders.

Precision Medicine

Six major clinical areas of PM interest:

- **oncology**
- **brain disorders**
- **cardiovascular disorders**
- **infections**
- **metabolic disorders, and**
- **other inflammatory disorders.**

Reproducibility of Tumor size …



> Insights Imagin. 2026 Jun 30;17(1):180. doi: 10.1186/s13244-026-02320-6.

Assessing the reliability of Response Evaluation Criteria In Solid Tumors (RECIST): a systematic review of the factors contributing to inter-observer variability

Illaa Smesseim [1 2], Teresa T Bucho [3 4], Renaud Tissier [5], Stevie van der Mierden [6], Ferry Lalezari [3], Jacobus A Burgers [7], Stefano Trebeschi [3 4]

Affiliations + expand

PMID: 42380551 PMCID: PMC13319625 DOI: 10.1186/s13244-026-02320-6

"Inter-observer variability (IOV) is primarily driven by 3 main factors:
measurement methods (manual vs. automated),
lesion selection (choosing target vs. non-target lesions), and
imaging quality."

"Standardizing these practices will improve the evaluation of tumor burdens, increase the reliability of clinical trial decisions, and optimize oncology patient outcomes."

Reproducibility of SUV & MTV …



JNM
The Journal of Nuclear Medicine

search

Advanced Se

Home | Content | Subscriptions | Authors | Info | About | More

Research Article | Clinical Investigation

International Benchmark for Total Metabolic Tumor Volume Measurement in Baseline ^{18}F-FDG PET/CT of Lymphoma Patients: A Milestone Toward Clinical Implementation

Ronald Boellaard, Irène Buvat, Christophe Nioche, Luca Ceriani, Anne-Ségolène Cottereau, Luca Guerra, Rodney J. Hicks, Salim Kanoun, Carsten Kobe, Annika Loft, Heiko Schöder, Annibale Versari, Conrad-Amadeus Voltin, Gerben J.C. Zwezerijnen, Josée M. Zijlstra, N. George Mikhaeel, Andrea Gallamini, Tarec C. El-Galaly, Christine Hanoun, Stephane Chauvie, Romain Ricci, Emanuele Zucca, Michel Meignan and Sally F. Barrington

Journal of Nuclear Medicine September 2024, 65 (9) 1343-1348; DOI: https://doi.org/10.2967/jnumed.124.267789

Precision Medicine

Six major clinical areas of PM interest:

- oncology
- **brain disorders**
- cardiovascular disorders
- infections
- metabolic disorders, and
- other inflammatory disorders.

Signs & Symptoms of Brain Disorders

- Abnormal -
  - Thoughts
  - Emotions
  - Behaviors

Global Burden:
Probably 1 billion people.
in 2019 ≈418M DALYs (16% of global DALYs) — ≈US$5 trillion economic burden.

by 2030, that figure is projected to go up to US$6.0 trillion.

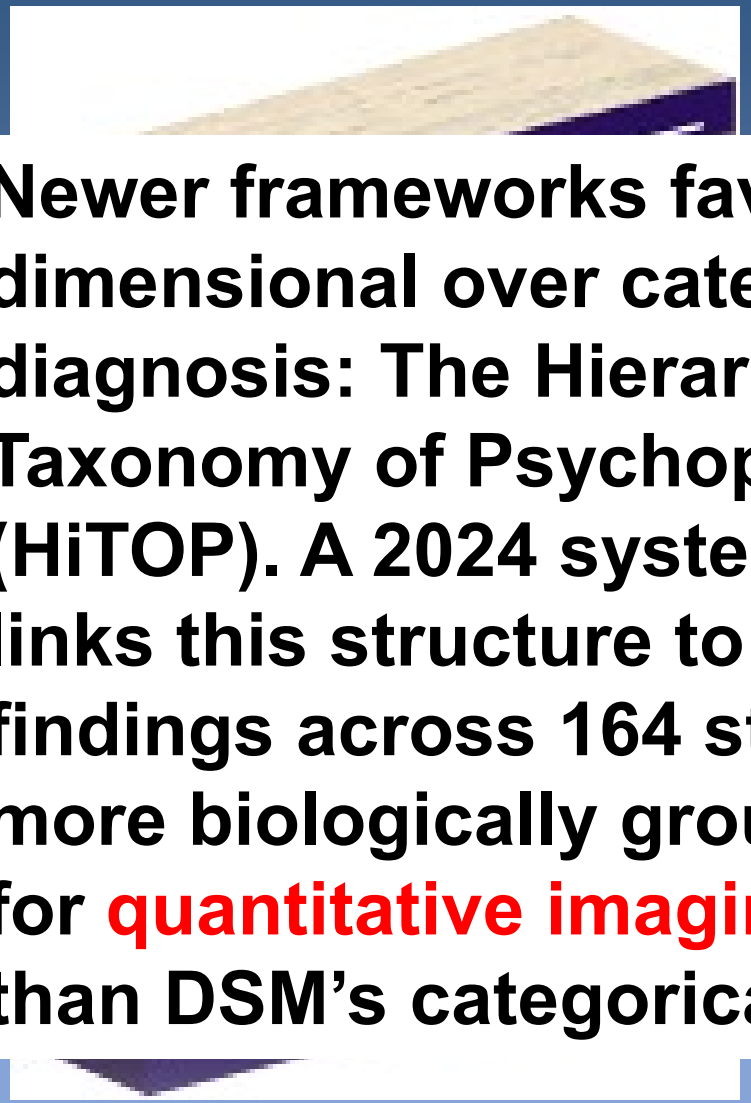
DALYs (disability-adjusted life years): years of healthy life lost to disease or premature death.

Signs & Symptoms of Brain Disorders

- Abnormal -
  - Thoughts
  - Emotions
  - Behaviors

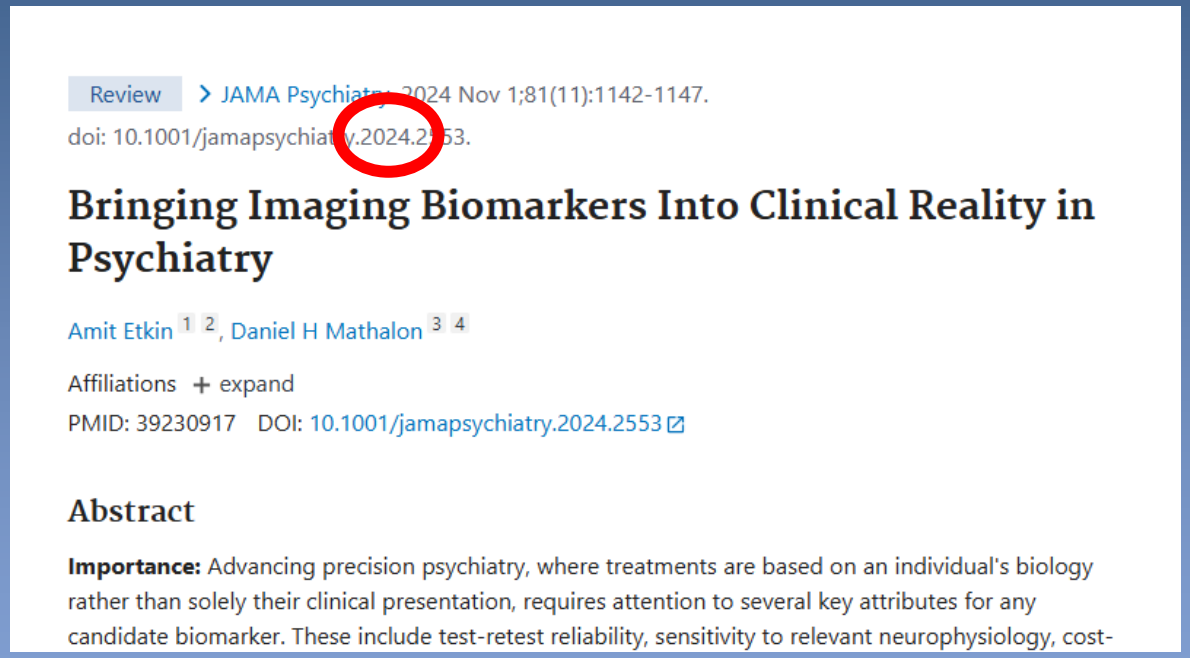
"DSM standardizes psychiatric diagnostic categories and criteria.

Critics, including the NIMH, argue the DSM represents an unscientific and subjective system.



Newer frameworks favor dimensional over categorical diagnosis: The Hierarchical Taxonomy of Psychopathology (HiTOP). A 2024 systematic review links this structure to neuroimaging findings across 164 studies — a more biologically grounded target for quantitative imaging biomarkers than DSM's categorical labels.

Huge, Unmet Need for Objective Biomarkers of Brain Disorders



Review > JAMA Psychiatry. 2024 Nov 1;81(11):1142-1147.
doi: 10.1001/jamapsychiatry.2024.2553.

Bringing Imaging Biomarkers Into Clinical Reality in Psychiatry

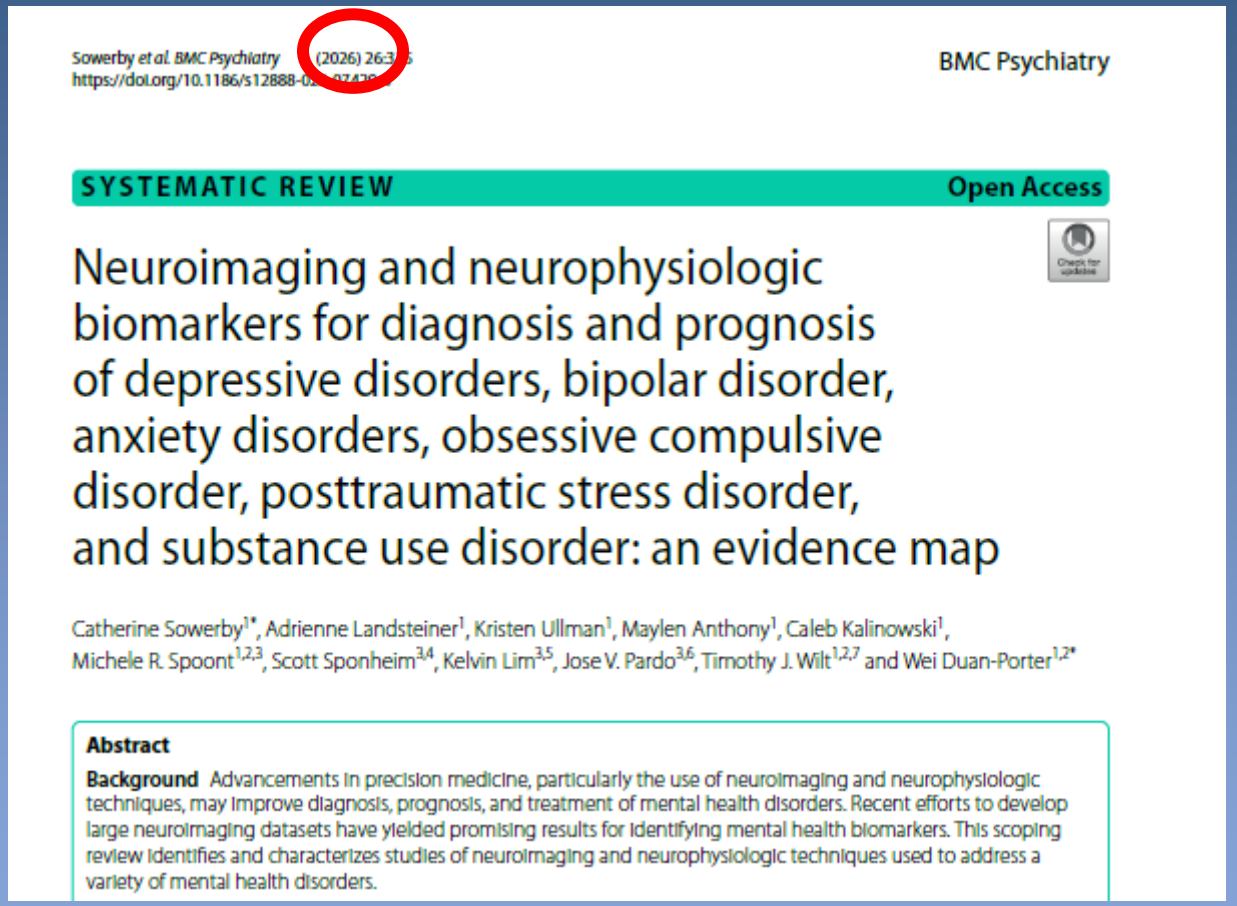
Amit Etkin [1 2], Daniel H Mathalon [3 4]

Affiliations + expand

PMID: 39230917 DOI: 10.1001/jamapsychiatry.2024.2553

Abstract

Importance: Advancing precision psychiatry, where treatments are based on an individual's biology rather than solely their clinical presentation, requires attention to several key attributes for any candidate biomarker. These include test-retest reliability, sensitivity to relevant neurophysiology, cost-



Sowerby et al. BMC Psychiatry (2026) 26:3
https://doi.org/10.1186/s12888-025-07439-1

BMC Psychiatry

SYSTEMATIC REVIEW Open Access

Neuroimaging and neurophysiologic biomarkers for diagnosis and prognosis of depressive disorders, bipolar disorder, anxiety disorders, obsessive compulsive disorder, posttraumatic stress disorder, and substance use disorder: an evidence map

Catherine Sowerby[1*], Adrienne Landsteiner[1], Kristen Ullman[1], Maylen Anthony[1], Caleb Kalinowski[1], Michele R. Spoont[1,2,3], Scott Sponheim[3,4], Kelvin Lim[3,5], Jose V. Pardo[3,6], Timothy J. Wilt[1,2,7] and Wei Duan-Porter[1,2*]

Abstract

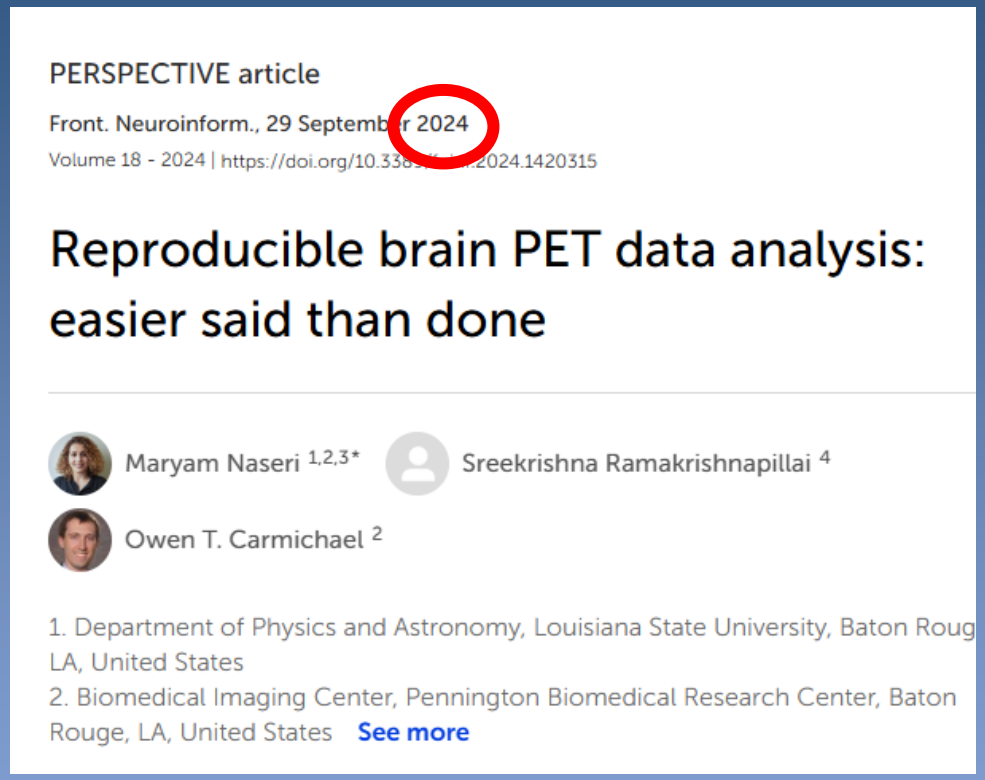
Background Advancements in precision medicine, particularly the use of neuroimaging and neurophysiologic techniques, may improve diagnosis, prognosis, and treatment of mental health disorders. Recent efforts to develop large neuroimaging datasets have yielded promising results for identifying mental health biomarkers. This scoping review identifies and characterizes studies of neuroimaging and neurophysiologic techniques used to address a variety of mental health disorders.

ADNI Biomarkers

Alzheimer's Disease Neuroimaging Initiative; 2004:

open-data model:
all imaging and clinical data were made available to researchers worldwide without embargo.

PET Reproducibility challenges:



PERSPECTIVE article

Front. Neuroinform., 29 September 2024

Volume 18 - 2024 | https://doi.org/10.3389/fninf.2024.1420315

Reproducible brain PET data analysis: easier said than done

Maryam Naseri [1,2,3*] Sreekrishna Ramakrishnapillai [4]

Owen T. Carmichael [2]

1. Department of Physics and Astronomy, Louisiana State University, Baton Rouge, LA, United States

2. Biomedical Imaging Center, Pennington Biomedical Research Center, Baton Rouge, LA, United States See more

Case study: "we tried to reproduce the processing pipeline of a large public dataset from the ADNI study."

They hit undocumented preprocessing steps, proprietary in-house software, and outdated tools they couldn't validate against.

"Promoting the development of standardized pipelines and tools should be a top priority."

Precision Medicine

Six major clinical areas of PM interest:

- oncology
- brain disorders
- **cardiovascular disorders**
- infections
- metabolic disorders, and
- other inflammatory disorders.

Quantitative Cardiovascular Imaging

- Cardiologists have extracted quantitative measures from clinical images for more than 50 years.
  - SPECT/PET
  - Echocardiography
  - CT (e.g., calcium scores; CT angiography
  - MR
- Assessing bias and precision in CV imaging has been a more recent focus.
- Thresholds between “normal” and “abnormal” periodically get revised because of the increased therapeutic options.

CV Imaging Standardization Efforts Have Advanced

Echocardiography

Guidelines for the Standardization of Adult Echocardiography Reporting

Taub, Stainback, et al. J Am Soc Echocardiogr. 2025.

Modernizes ASE's 2002 reporting standard; builds on the 2015 chamber-quantification guideline (Lang et al.) with standardized measurements, terminology, and structured reporting.

Coronary CT Angiography

Standards for Quantitative Assessments by Coronary CT Angiography (CCTA)

Nieman, Garcia-Garcia, et al. J Cardiovasc Comput Tomogr. 2024; 18(5):429-443.

SCCT expert consensus defining standards for quantitative (not just visual) CCTA measures of coronary disease - supersedes the 2016 acquisition guideline for quantitative use.

Nuclear Cardiology / PET

18F-Flurpiridaz PET Myocardial Perfusion Imaging and Blood Flow Quantitation

Packard, Maddahi, et al. (SNMMI/EANM/ASNC/ACNM). J Nucl Med, 2025; 66(10):1538-1554.

First multi-society procedure standard for quantitative PET myocardial blood flow - succeeds the 2015 EANM/EACVI reporting position paper.

Cardiac MRI (CMR)

Standardized Cardiovascular Magnetic Resonance Imaging (CMR) Protocols, 2020 Update

Kramer, Barkhausen, Bucciarelli-Ducci, et al. J Cardiovasc Magn Reson, 2020;22:17.

Direct successor to the 2013 SCMR protocol update; expands general technique guidance and adds device-patient imaging.

Precision Medicine

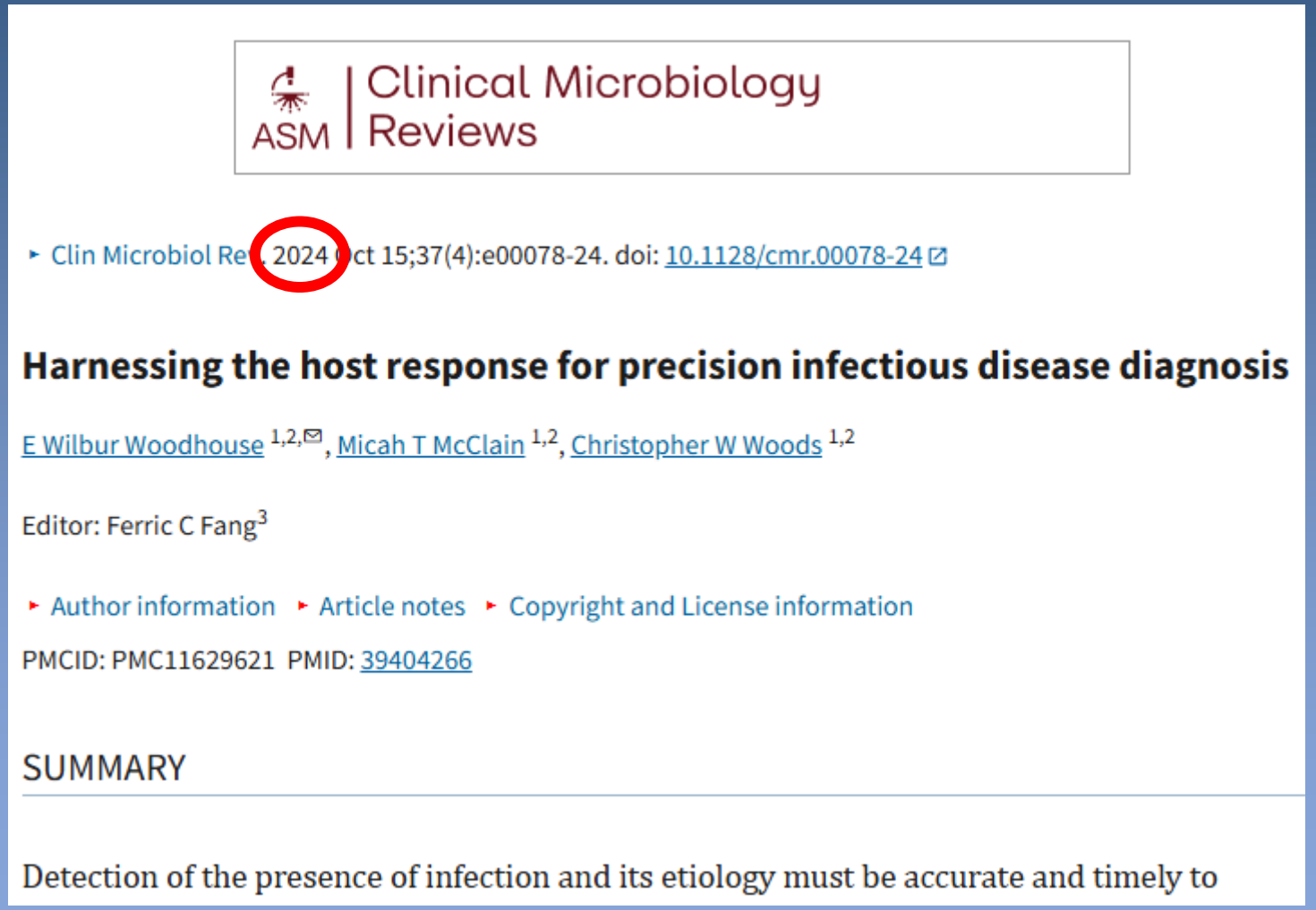
Six major clinical areas of PM interest:

- oncology
- brain disorders
- cardiovascular disorders
- **infections**
- metabolic disorders, and
- other inflammatory disorders.

QI in Infection

- **“Understanding the host’s response to infection — not just detecting the pathogen — gives more accurate, timely, and actionable diagnostic information.”**

Woodhouse, et.al., Clin Microbiol Rev, 2024;



ASM | Clinical Microbiology Reviews

▸ Clin Microbiol Rev. 2024 Oct 15;37(4):e00078-24. doi: 10.1128/cmr.00078-24

Harnessing the host response for precision infectious disease diagnosis

E Wilbur Woodhouse [1,2,✉], Micah T McClain [1,2], Christopher W Woods [1,2]

Editor: Ferric C Fang[3]

▸ Author information ▸ Article notes ▸ Copyright and License information

PMCID: PMC11629621 PMID: 39404266

SUMMARY

Detection of the presence of infection and its etiology must be accurate and timely to

QI in Infection

1. Imaging can be an alternative to biopsy
    - Tuberculosis
    - Hepatitis
    - Abdominal infections
    - Fungal infections
    - Osteomyelitis
    - Sepsis
    - Diabetic foot
2. Quantify drug distribution

- Tuberculosis remains the leading infectious cause of death worldwide, with 10.7 million cases and 1.23 million deaths reported in 2024.
- There is strong quantitative evidence behind PET/CT for detecting treatment response earlier than anatomic imaging and predicting relapse risk with high specificity.
- Yet, WHO has not endorsed routine PET/CT use for TB management, citing cost, access, and standardization gaps.

Precision Medicine

Six major clinical areas of PM interest:

- oncology
- brain disorders
- cardiovascular disorders
- infections
- **metabolic disorders**, and
- other inflammatory disorders.

Metabolic dysfunction-associated steatotic liver disease (MASLD): a chronic liver disease

Progressive form = Metabolic dysfunction-associated steatohepatitis (MASH)

- MASLD affects 1.3 billion people worldwide
- MASH estimated to affect 400 million worldwide
- Resmetirom FDA approval, March 14, 2024
- Many payers now require quantitative MRI-PDFF and MRE
- However, payer criteria for MRE/MRI-PDFF aren't uniform yet
- No generally accepted QA or calibration criteria.

This is a protype, precedent-setting example of *imaging as an assay*.

All of the concepts developed over decades for laboratory assays of fluid and specimen biomarkers must be adapted and applied to imaging assays.

Precision Medicine

Six major clinical areas of PM interest:

- oncology
- brain disorders
- cardiovascular disorders
- infections
- metabolic disorders, and
- other inflammatory disorders.

Inflammatory Conditions for which QIBs are Needed

- Rheumatoid arthritis
- Inflammatory bowel disease (IBS)
- Osteoarthritis
- Idiopathic pulmonary fibrosis
- Autoimmune disorders
- Neurodegenerative disorders
- Asthma

JCC JOURNAL of CROHN'S and COLITIS

4.1. Monitoring disease in UC

> Recommendation 8
>
> In patients with UC requiring treatment initiation or optimization, we recommend early (within 12 weeks) clinical (EL1), biochemical (EL1), and endoscopic (EL1) or ultrasonographic (EL2) response assessment. Results should be interpreted based on prior baseline assessment. (92% agreement)

2. Methodology of the guidelines update process

3. Initial diagnosis

Published: 31 July 2025 **Article history** ▾

What's Broken?

Variation in clinical practice results in poorer outcomes and higher costs.

That's the basic premise!

Treating physicians need quantitative results.

They want measurements instead of ambiguous qualitative descriptors.

Those of you participating or listening to this Symposium have the power to fix these problems.

Thank you.



