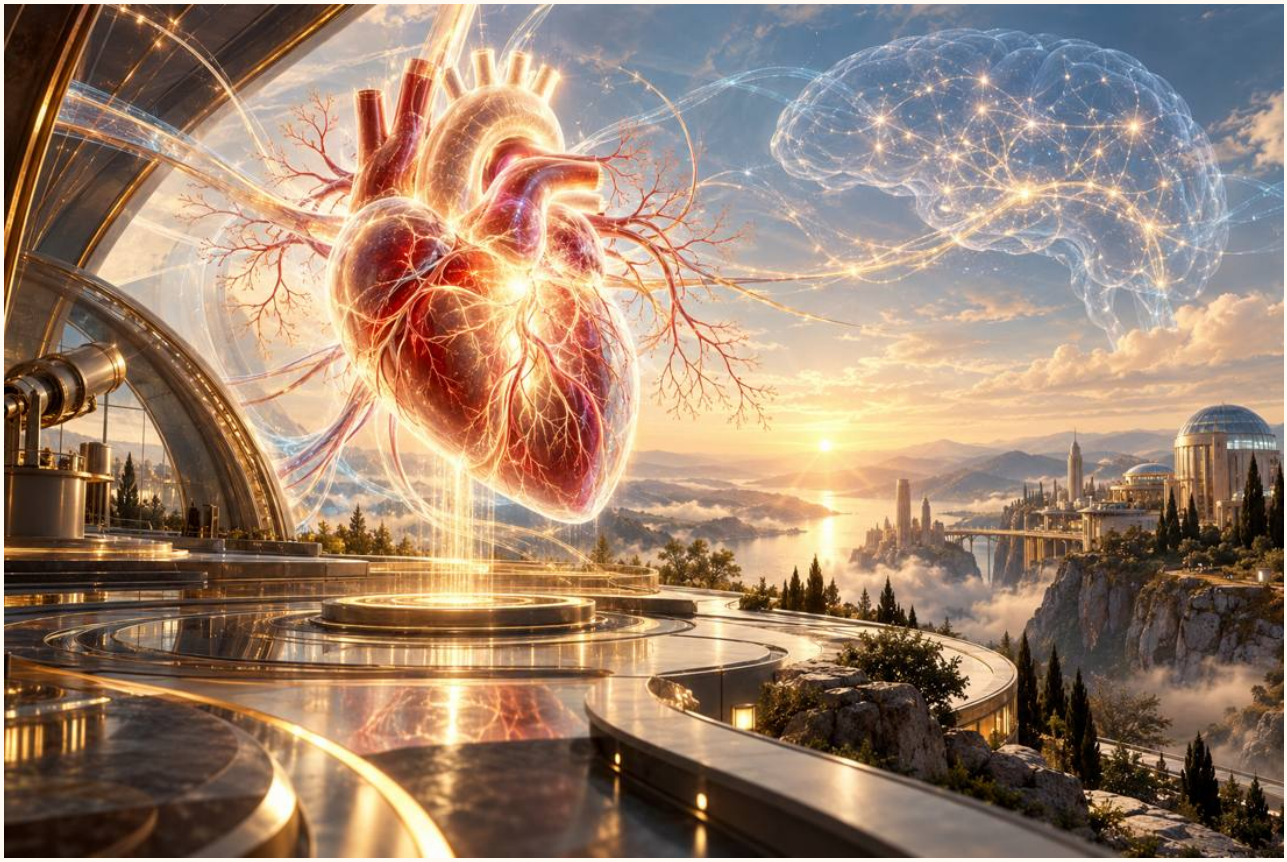
What Can We Measure That Truly Matters?

November 1, 2026 · John Griffin, MD



Conceptual editorial illustration for this edition of The Arising.

The Morning View

The most persuasive medical results this morning are not necessarily the most exotic. They are the ones tied to outcomes patients and families recognize: a serious bleed avoided, a mood episode prevented, a weakened hand used again. BRIGHT-4, a randomized study involving more than six thousand people having emergency treatment for a heart attack, found a 1.5-percentage-point reduction in death or major bleeding at one year under its specific anticoagulation protocol. That is consequential precisely because the comparison and the outcome are clear.

In parallel, tools that look like the future are arriving without all the answers. A compact multimodal cognitive assessment discriminated mild cognitive impairment in a research sample. A ninety-second magnetic cardiac recording offered information beyond ECG in a

separate test campus. Neither system has yet demonstrated that using it changes patient outcomes. They deserve serious development, not premature marketing.

Two pieces of negative evidence are equally useful. A randomized-trial synthesis found no convincing metabolic benefit from fecal microbiota transplantation. A meta-analysis of sham-controlled trials did not identify robust cognitive improvement from generic noninvasive brain stimulation in neurodevelopmental conditions. The research questions remain alive, but the clinical claims must be earned. The Arising can celebrate rigorous testing without requiring every exciting hypothesis to succeed.

1. In a 6,016-person heart attack trial, a modest absolute benefit proves substantial

Cardiology advances are often reported as hazard ratios. For the person arriving with a major heart attack, the more useful question is what happens in the year after the artery is opened. The BRIGHT-4 trial followed more than six thousand patients undergoing primary coronary intervention, and compared two anticoagulation strategies under a precisely defined procedural protocol.

The comparison matters: bivalirudin here was not simply given instead of heparin. It was followed by a high-dose infusion for two to four hours, and the patients largely underwent radial-access PCI. Those details constrain where the result applies. The composite reduction was supported by less major bleeding and stent thrombosis; death by itself was not conclusively lower.

For clinical practice this is a specialist protocol decision, requiring review of contemporary guideline recommendations, patient bleeding risk, cost and catheterization-lab workflow. It is not an invitation to generalize anticoagulation choices outside the trial setting.

The evidence: Open-label, multicenter RCT; 6,016 people with STEMI receiving radial-access primary PCI; 3,009 assigned bivalirudin plus full-dose post-PCI infusion and 3,007 heparin alone. At 12 months, death or BARC grade 3–5 bleeding occurred in 4.1% versus 5.6% (HR 0.73; 95% CI 0.58–0.92): an absolute 1.5-percentage-point difference, roughly 15 fewer events per 1,000 treated. Major bleeding was 0.6% versus 1.2%; stent thrombosis 0.8% versus 1.3%. The all-cause mortality difference alone (3.8% versus 4.8%) was not statistically significant (P=0.08).

Meaning for care: Clinically relevant now for interventional-cardiology review; not an unsupervised treatment change.

Read the original study

2. Twelve sessions of culturally adapted education may keep bipolar illness in remission longer

One of the most consequential discoveries in medicine is that treatment effectiveness can depend on whether people can actually use it. This trial did not introduce an exotic new molecule. It adapted structured psychoeducation to the language and circumstances of adults receiving bipolar-disorder care in Pakistan.

Participants met individually with trained psychologists over twelve weeks. Recurrence was measured across 48 weeks, not merely whether the sessions were popular. A hazard ratio near one-half is encouraging, although the absolute percentage experiencing relapse must be retrieved from the full report before publishing a number-needed-to-treat.

The intervention was delivered with usual psychiatric treatment. It did not replace medication, emergency services, or clinician follow-up. Successful implementation elsewhere would require cultural adaptation, clinician capacity and evaluation with appropriate comparators.

The evidence: 303 adults with bipolar I or II in remission at seven Pakistani centers; 155 randomized to 12 weekly individual culturally adapted psychoeducation sessions plus treatment as usual, 148 to usual care alone; 48-week recurrence follow-up. Time to recurrence of any mood episode favored the intervention (adjusted HR 0.54; 95% CI 0.42–0.69). Twelve-month data were available for 284/303 (93.7%). Depressive symptoms, medication adherence and patient-reported recovery also improved; the provided abstract does not give absolute recurrence risks.

Meaning for care: Potential clinical service-design implication for maintenance psychiatric care; verify protocol, absolute recurrence and safety before replication.

Read the original study

3. Less transfusion in leukemia care, but no demonstrated survival gain

Blood products are a crucial part of aggressive leukemia treatment, and reducing transfusion requirements could matter to patients and oncology units. The EPAG-2015 trial tested whether eltrombopag, introduced after chemotherapy, might help older adults through the difficult period of low platelet counts.

Its central outcome was not improved: the study did not establish a survival benefit, and the hazard-ratio confidence interval was wide. The reduction of two median platelet transfusions is potentially useful but secondary. It should not be rhetorically promoted into a successful mortality trial.

The clinical question is therefore narrower and practical: could a safely timed platelet-support strategy ease care in a defined setting without compromising cancer treatment? This requires the full protocol, adverse-event tables, blood-product use and local transfusion practices.

The evidence: 110 adults aged 60 and older receiving induction chemotherapy for newly diagnosed acute myeloid leukemia in 18 French hospitals; 55 eltrombopag, 55 placebo; median follow-up among survivors 20.9 months. The prespecified overall-survival objective was not met (HR 1.59; 95% CI 0.65–3.89; P=0.31). Complete response or response with incomplete recovery occurred in 45 people in each group. Median platelet transfusions were 8 (IQR 6–14) versus 10 (7–14); no increase in grade 3–4 adverse events was reported.

Meaning for care: Specialist research finding; not routine addition to AML induction therapy on this evidence.

Read the original study

4. A 14-minute digital assessment recognizes mild cognitive impairment—but what happens after the score?

A future cognitive-health clinic may measure how someone speaks, moves and solves problems, rather than rely only on a short paper questionnaire. The NeuroMarker study offers a promising look at that possibility. Across an adjudicated research cohort, a combination of tasks distinguished mild cognitive impairment from unimpaired cognition reasonably well.

AUC measures discrimination across thresholds, not whether a positive result means a particular person has Alzheimer’s disease. Clinical utility requires realistic disease prevalence, prespecified thresholds, confirmatory evaluation, false-positive burden, accessibility and longitudinal outcomes. The modest classifier agreement on repeated measurements also deserves careful examination.

This belongs directly in The Arising’s dementia monthly project. The valuable question is whether such tools help patients receive better evaluation earlier and improve subsequent decisions—not whether a dashboard can draw a persuasive curve.

The evidence: 638 adults aged 50 or older in three countries, 441 cognitively unimpaired and 197 with adjudicated mild cognitive impairment; self-administered multimodal tablet tasks involving motor, speech and augmented reality; independent cohort n=75. Median completion time 13.8 minutes. ROC AUC was 0.90 (95% CI 0.88–0.92), and 0.87 (0.80–0.93) in the 75-person independent sample. Test-retest classifier agreement κ=0.60; composite/domain ICCs ranged 0.55–0.80. Funded in part by Altoida Inc., a platform developer.

Meaning for care: Promising screening research, not a stand-alone diagnosis or dementia-prevention claim.

Read the original study

5. A 90-second magnetic recording offers a new view of the injured heart

The coronary catheterization lab is full of technology, yet identifying the artery responsible for a non-ST-elevation heart attack before angiography can still be difficult. This group asked whether optically pumped magnetometers could provide useful additional information in a recording lasting only ninety seconds.

A geographically separate testing campus improves credibility over a single retrospective training set. Still, the locked algorithm was correct on about two-thirds of three-vessel classifications and missed many left circumflex culprit vessels. A high ROC AUC for one artery must not obscure that clinical asymmetry.

The genuine breakthrough would not be accuracy alone. It would be a prospective demonstration that the information improves triage, prevents delays, changes management safely, or reduces patient-important harm.

The evidence: 611 hemodynamically stable NSTEMI patients already selected for angiography; XGBoost model on 65 prespecified magnetocardiography features trained in campus 1 (n=493), locked before campus 2 test (n=118). In external campus test, overall vessel identification accuracy was 63.6%; matched ECG-only model 51.7%; integrated standard care 55.9%. The exploratory addition of magnetocardiography probabilities to standard care achieved 67.8% accuracy; sensitivity was unequal across arteries (LAD 76.6%, LCx 34.5%, RCA 69.0%).

Meaning for care: Technical diagnostic research; not ready to choose invasive procedures.

Read the original study

6. After stroke, fifty hours of intensive practice changed how people used a weakened arm

Recovery is easier to measure in a laboratory than in a kitchen. For someone living with a stroke, the meaningful question is whether an affected hand begins helping with tasks again. This trial tested an intensive, task-specific approach involving the arms, legs and trunk, with outcomes that included how people used the affected upper limb.

The intervention group improved more than the usual-care group on the primary assisting-hand outcome, and gains persisted to three months. But the contrast was stark: fifty hours of concentrated practice versus under five hours of usual rehabilitation. We therefore cannot attribute the whole difference to the particular brand of therapy rather than dose, attention, intensity or environment.

The next test should hold total therapy hours and clinician contact more comparable, include meaningful activities at home, and establish durability. Even with those limits, it is a more useful healthspan story than another measurement of muscle size alone.

The evidence: 46 adults with chronic hemiparetic stroke randomized to a 50-hour, two-week camp-style HABIT-ILE program or usual neurorehabilitation averaging 4 hours 36 minutes

over the same period. The Assisting Hand Assessment Stroke measure favored intensive therapy, with reported intervention improvements of +6.24% of logits at two weeks and +7.05% at three months. Several arm-function and self-perceived occupational-performance endpoints also favored the intervention. No significant between-group improvement was demonstrated on the measured lower-extremity outcomes.

Meaning for care: Rehabilitation specialist interest; intensity and patient selection require prospective implementation appraisal.

Read the original study

7. A gut-microbe transplant has not yet become a metabolic treatment

The microbiome is one of the most exciting frontiers in metabolic medicine. The hypothesis is plausible: altering gut microbial communities might change insulin sensitivity. But the treatment question is simpler than the theory. Did giving fecal microbial material to patients reliably improve their glucose metabolism?

Across randomized trials, the pooled answer was inconclusive. Donor characteristics, preparation, background diets and comparators varied, and even the primary synthesis included just 120 people. That is too small and heterogeneous to turn an interesting biological mechanism into a clinical metabolic protocol.

For the nutrition and GLP-1 booklets, this is a useful negative result. It keeps the door open for better-designed trials while preventing an unproven intervention from appearing beside established treatments as though the evidence were comparable.

The evidence: 18 eligible randomized FMT studies in adults with metabolic disorders; primary meta-analysis drew on only six trials and 120 participants for insulin resistance measured by HOMA-IR. Between-group HOMA-IR mean difference -1.29 (95% CI -3.04 to 0.47; I^2 72%): no conclusive improvement. HbA1c data were too sparse to pool across different comparators. Evidence certainty was very low for HOMA-IR and HbA1c.

Meaning for care: No established FMT role for routine treatment of diabetes or insulin resistance based on these trials.

Read the original study

8. Thirty sham-controlled trials challenge claims about brain stimulation and cognition

A device can change neural activity without improving the thinking and daily functioning we care about. In neurodevelopmental conditions, noninvasive stimulation has attracted attention because it seems able to influence brain networks directly.

This meta-analysis pooled sham-controlled trials and treated multiple related outcomes as statistically dependent, an important methodological safeguard. Despite the number of tested cognitive domains, no specific domain showed a dependable gain over sham. That does not

mean individualized future protocols cannot work; it means the currently heterogeneous approaches have not earned a broad therapeutic claim.

The result also belongs in our dementia and cognitive-health coverage as a warning against extrapolation: a stimulation study in one diagnosis is not evidence that a device prevents Alzheimer's or reverses aging.

The evidence: 30 randomized trials across autism, ADHD and intellectual disability; 178 effects spanning seven cognitive domains, modeled with dependence-aware multilevel methods. No individual cognitive domain showed a statistically reliable advantage over sham. For working memory, Hedges g -0.16 (95% CI -0.37 to 0.05); for cognitive flexibility g -0.12 (-0.38 to 0.13). The small cross-domain pooled estimate g -0.18 (P=0.035) had a wide 95% prediction interval (-1.01 to 0.66) and no robust domain-specific signal.

Meaning for care: Do not present generic noninvasive stimulation as a proven cognitive enhancer for neurodevelopmental disorders.

Read the original study

9. AI-written therapy feedback performed similarly in one trial; that does not establish equivalence

This is a rare useful AI trial because it asks not just whether a language model can sound therapeutic, but whether people improve when it supplies supervised feedback during a structured therapy program. It also manipulated what participants believed the source to be.

The participants improved in active treatment, and no statistically detected difference emerged between AI and human feedback. However, the investigators did not establish equivalence or noninferiority. The feedback was also psychologist-reviewed. These are important boundaries for any proposal to replace a qualified therapist with an unsupervised companion.

The thoughtful next experiment is an appropriately powered noninferiority design with harms, treatment completion, alliance, human-review burden and longer-term remission. For Relationships, this should not be converted into an assertion that AI companions produce durable relationship benefit.

The evidence: 263 adults with social anxiety disorder randomized among psychologist-reviewed AI-generated feedback, human feedback, unguided self-help and waitlist, with experimental manipulation of stated feedback source. Active guided therapy reduced social anxiety relative to waitlist (reported d -1.23). AI versus human feedback did not differ statistically, but the trial explicitly did not establish noninferiority. Initial expectations favored a human source, whereas actual source/belief did not detectably change symptom improvement or therapeutic alliance.

Meaning for care: Promising supervised clinical-workflow research; no justification for unsupervised autonomous psychotherapy.

Read the original study

10. AI-designed balance exercises improved walking measures, but the human instructor still mattered

The most interesting finding here is not that a chatbot “trained” older adults. It is that a human instructor delivered an exercise plan generated with an AI model, and the small trial compared it against two other structured approaches.

Technology-assisted groups showed some larger gait gains. But sixteen people per arm cannot settle safety, fall prevention, durability or how much of the benefit comes from program design, instructor enthusiasm or the amount of practice. Every group improved during repeated exercise sessions, which is itself worth noticing.

The participants trained under supervision. Improvements on walking measures do not establish that an autonomous exercise program would be equally effective or safe.

The evidence: 48 retirement-home residents aged 60–75 randomized equally to human-delivered AI-generated exercise, virtual-reality exergaming or traditional balance training; 60-minute sessions three times weekly for eight weeks, with two-month follow-up. All groups improved balance and gait measures. AI- and VR-supported programs improved gait speed more than traditional exercise (both $P<0.001$ in reported comparisons), without a reported AI–VR difference. Virtual reality had the largest cadence gain and AI the largest gait-stability ratio reduction; no adverse events were recorded.

Meaning for care: Controlled educational pilot candidate; not evidence that AI exercise plans prevent falls or extend life.

Read the original study

Go deeper with The Arising

Read the illustrated November 1 edition on The Arising website.

Reading medical evidence · Sleep and rhythm · Strength and function · GLP-1 medicines · Browse the website · Explore the public Atlas

Alive. Awake. Aware.