Measuring Effectiveness and Implementation Outcomes

Angelo Volandes, MD, MPH
Professor and Vice Chair of Research, Department of Medicine
Dartmouth Health and Geisel School of Medicine at Dartmouth

NIH PRAGMATIC TRIALS COLLABORATORY
Rethinking Clinical Trials®

Disclosures

- Dr. Angelo Volandes has a financial interest in ACP Decisions, a nonprofit organization developing advance care planning video decision support tools. Dr. Volandes' interests were reviewed and are managed by Dartmouth in accordance with their conflict-of-interest policies. No other disclosures to report.

Learning goals

- Describe methods for measuring outcomes using real-world data sources such as electronic health records (EHRs) and patient-reported outcomes (PROs)
- Identify outcomes commonly used in implementation research
- Learn about key issues in measurement of implementation outcomes

NIH PRAGMATIC TRIALS COLLABORATORY
Rethinking Clinical Trials®

Outcome, measure, endpoint

- **Outcome** usually refers to a variable of interest or a meaningful aspect of health (such as oxygen volume or fatigue)

- **Measure** usually refers to a specific, standardized process to obtain information on an outcome
  - Includes instructions, administration materials, content, formatting, and scoring rules

Types of measures

Patient-reported outcome measure (PROM)	Observer-reported outcome measure (ObsRO)
Clinician-reported outcome measure (ClinRO)	Performance outcome measure (PerfO)

NIH PRAGMATIC TRIALS COLLABORATORY
Rethinking Clinical Trials®

Outcome, measure, endpoint

- **Endpoint** usually refers to a precisely defined variable that is statistically analyzed to address a particular research question

Examples:

- Change from baseline at 6 weeks in mean PROMIS Fatigue score
- Mean difference in PROMIS Fatigue score between patients in the intervention and usual care groups, after controlling for baseline status

NIH PRAGMATIC TRIALS COLLABORATORY
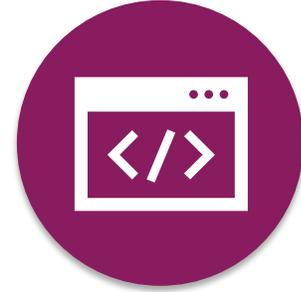Rethinking Clinical Trials®

Important things to know

Outcomes and their related endpoints should be **meaningful** to providers and patients

Outcomes and related measures should be relatively **easy** to collect (ie, pragmatic)

Researchers do not control the design or data collected in EHR systems

NIH PRAGMATIC TRIALS COLLABORATORY
Rethinking Clinical Trials®

Choosing and specifying endpoints in ePCTs

Outcomes and related endpoints should be available as part of routine care

Easy

- Acute myocardial infarction
- Broken bone
- Hospitalization

Hard

- Suicide attempts
- Gout flares
- Silent myocardial infarction
- Early miscarriage

NIH PRAGMATIC TRIALS COLLABORATORY
Rethinking Clinical Trials®

Key questions for choosing endpoints

Is the outcome medically significant
such that a patient would seek care?

Does it require hospitalization?

Is treatment provided in inpatient or outpatient settings?

Will the event be medically attended?
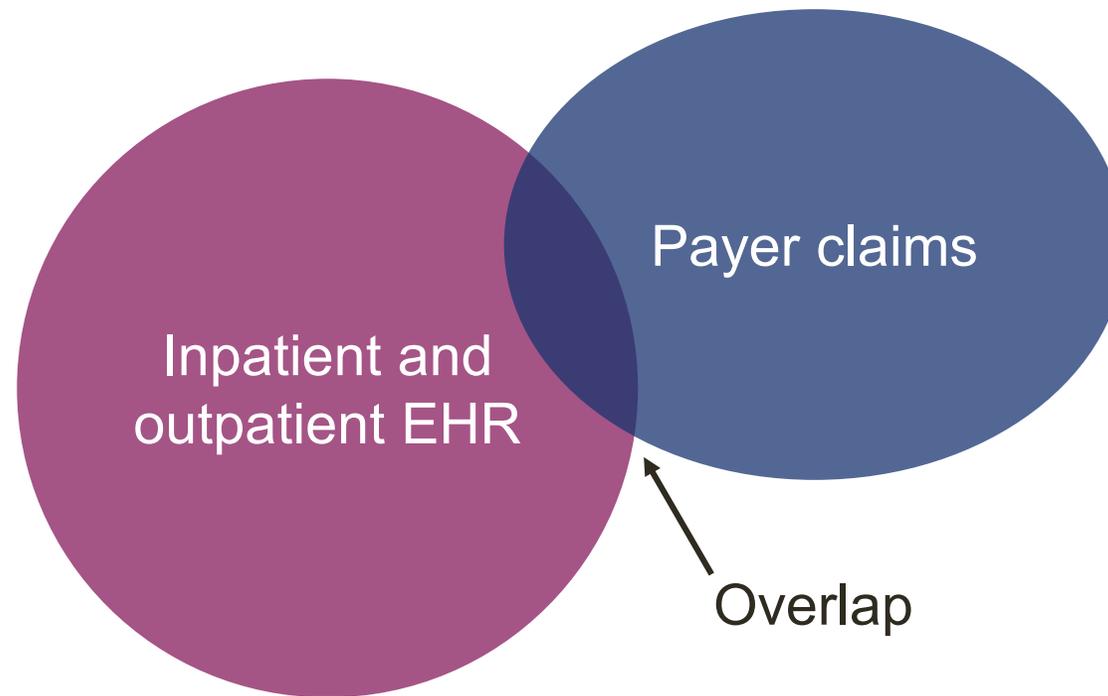
Data sources for endpoints in ePCTs

> *" The first challenge in using big biomedical data effectively is to identify what the potential sources of healthcare information are and to determine the value of linking these together."*
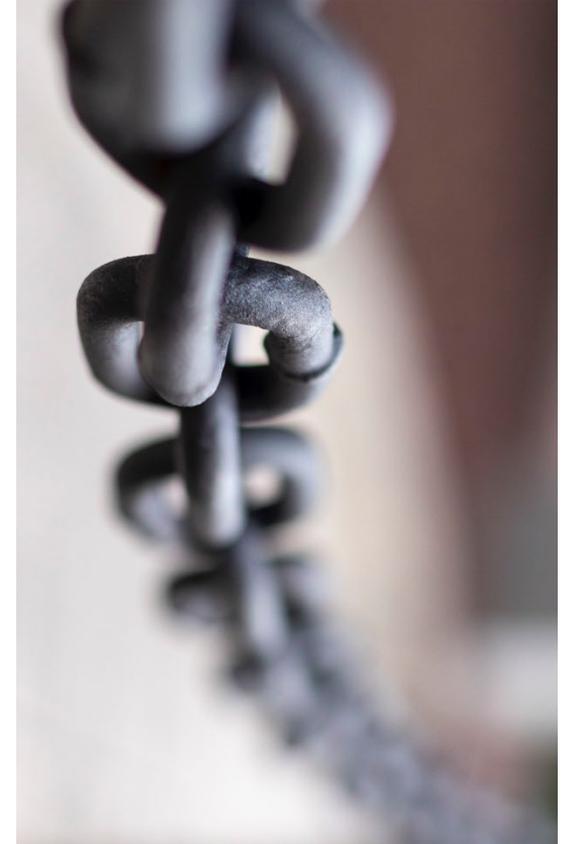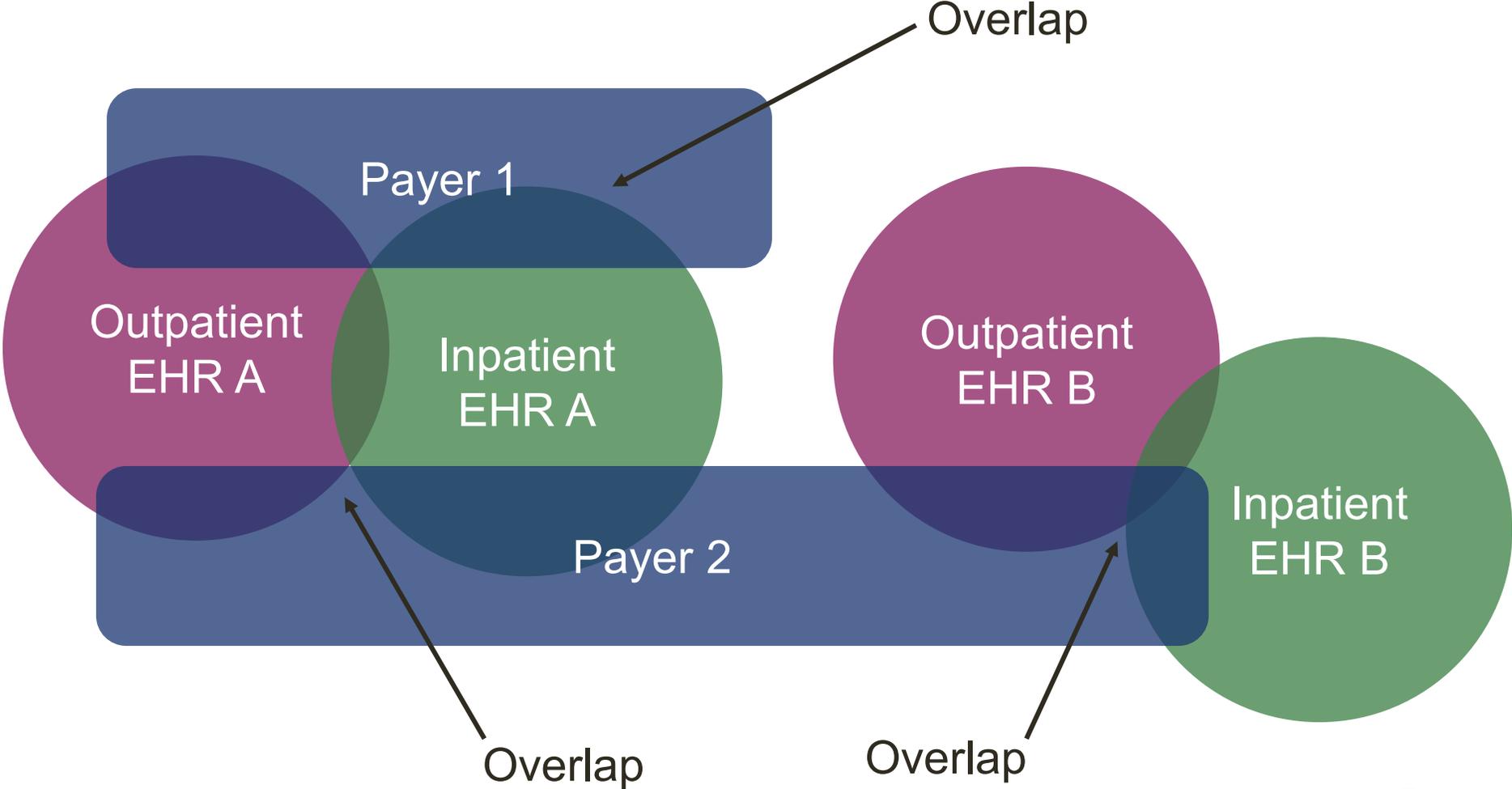
Weber GM, Mandl KD, Kohane IS. Finding the missing link for big biomedical data. *JAMA*. 2014 Jun 25;311(24):2479-80. doi: 10.1001/jama.2014.4228.

NIH PRAGMATIC TRIALS COLLABORATORY
Rethinking Clinical Trials®

Where is the signal?

- EHR (laboratory values, treatments, etc.)
- Claims data (Does the event generate a bill?)

Inpatient and outpatient EHR

Payer claims

Overlap

NIH PRAGMATIC TRIALS COLLABORATORY
Rethinking Clinical Trials®

Reality is not straightforward

Overlap

Payer 1

Outpatient EHR A

Inpatient EHR A

Outpatient EHR B

Inpatient EHR B

Payer 2

Overlap

Overlap

NIH PRAGMATIC TRIALS COLLABORATORY
Rethinking Clinical Trials®

Longitudinal data linkage

- To fully capture all care—complete longitudinal data—linking research and insurance claims data is often necessary

- Without explicit consent, getting longitudinal data from an insurance carrier can be an insurmountable hurdle, both technically and legally

Data is a surrogate for clinical phenomena

Error Impact on Trials
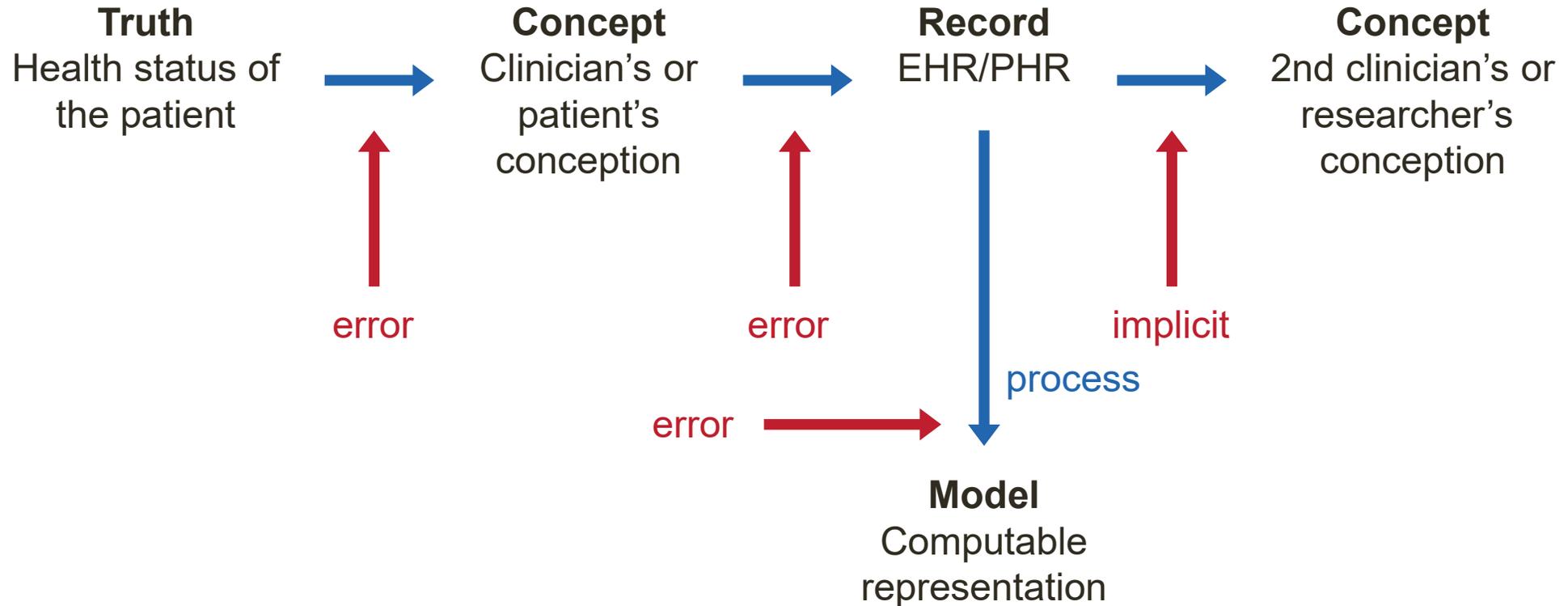
Adapted from Hripcsak et al. J Am Med Inform Assoc. 2009 Mar-Apr;16(2):220-7.

NIH PRAGMATIC TRIALS COLLABORATORY
Rethinking Clinical Trials®

Data sources for endpoints in ePCTs

Traditional

- EHR or ancillary health information systems

Complementary

- Other types of health data not routinely collected outside standard clinical practice, such as PROs

It's a balancing act

Relevance to real-world decision-making
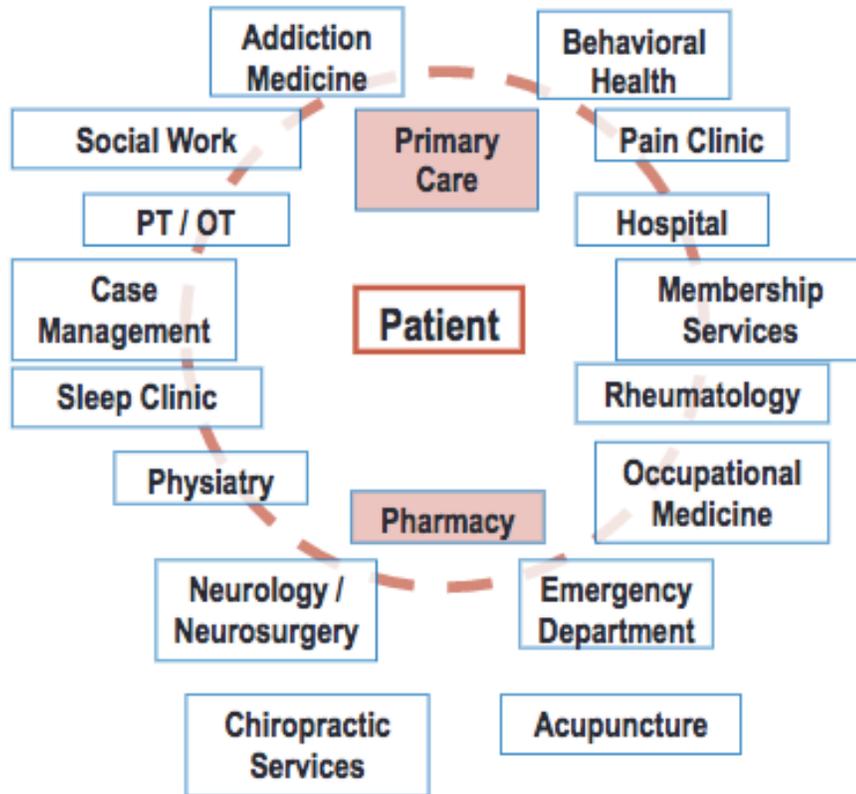may come at the expense of efficiency

A trial measuring outcomes that matter most to patients and healthcare systems may not be able to rely exclusively on information from the EHR, and instead may need to assess patient-reported outcomes, which is more expensive and less efficient

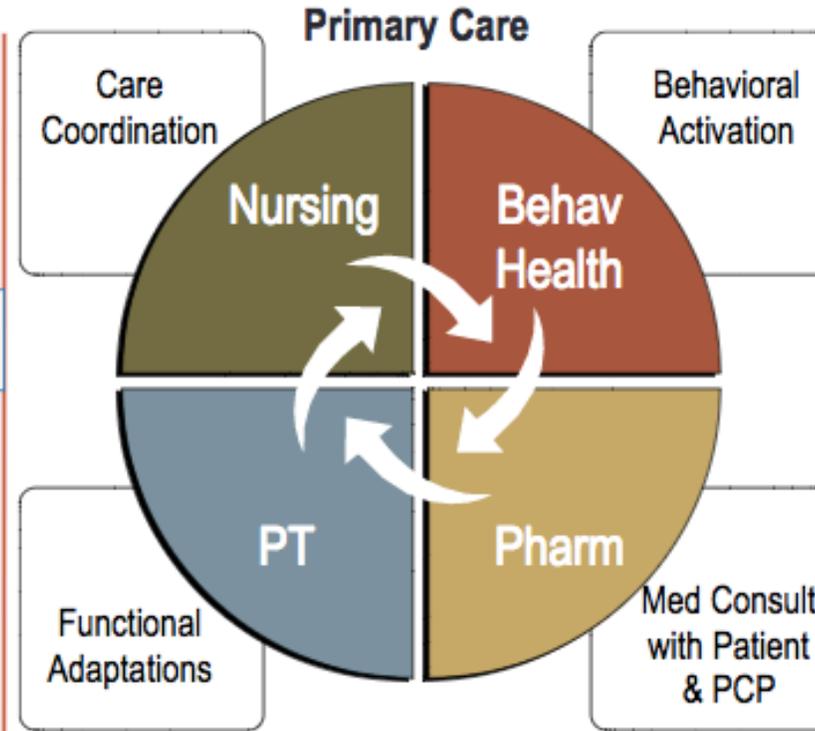# Outcomes measured via direct patient report

- PROs are the best way to measure quality of life and often the best way to measure how patients are feeling and functioning

- Challenges
  - Not routinely or consistently used in clinical care
  - Not regularly recorded in EHR

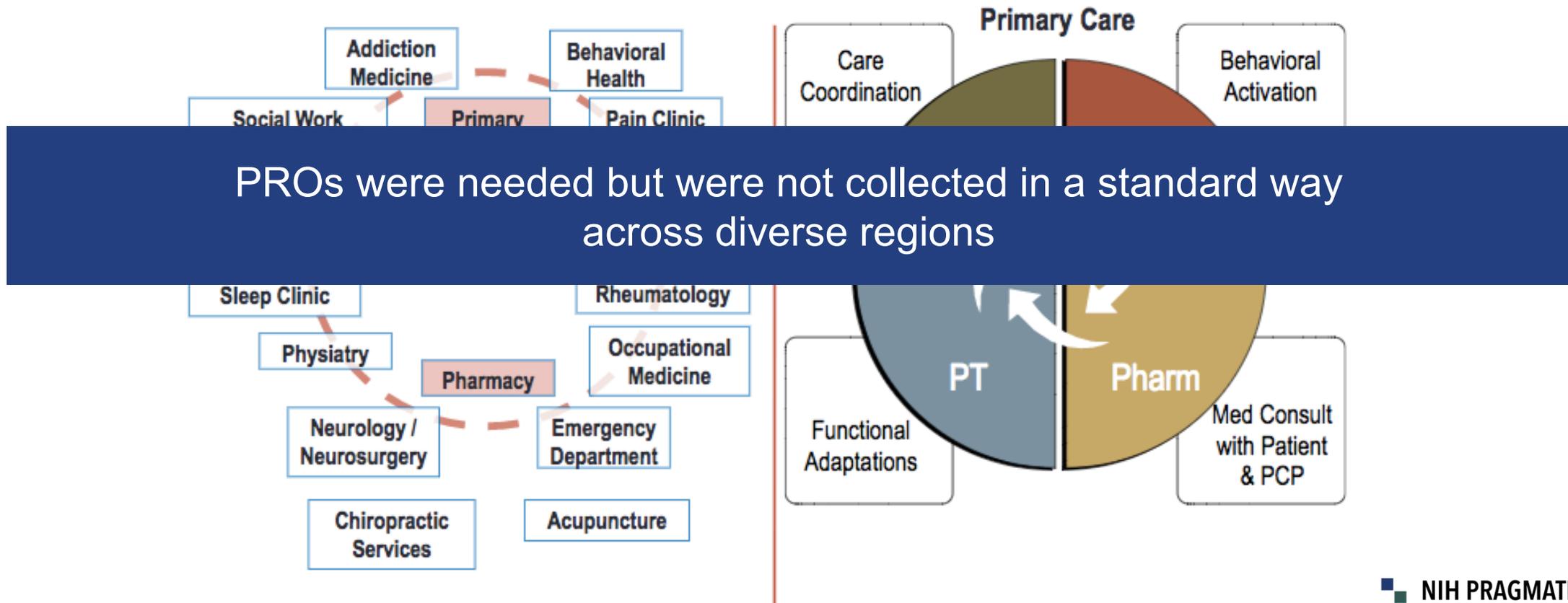- Need a mechanism to collect PROs

Example: PPACT Trial

Source: Lynn DeBar, Kaiser Permanente Washington Health Research Institute

Example: PPACT Trial

Pain Management: Usual Care

- Addiction Medicine
- Behavioral Health
- Social Work
- Primary
- Pain Clinic
- Sleep Clinic
- Rheumatology
- Physiatry
- Pharmacy
- Occupational Medicine
- Neurology / Neurosurgery
- Emergency Department
- Chiropractic Services
- Acupuncture

Interdisciplinary Management Embedded in Primary Care

Primary Care

- Care Coordination
- Behavioral Activation
- PT
- Pharm
- Functional Adaptations
- Med Consult with Patient & PCP

PROs were needed but were not collected in a standard way across diverse regions

Source: Lynn DeBar, Kaiser Permanente Washington Health Research Institute

NIH PRAGMATIC TRIALS COLLABORATORY
Rethinking Clinical Trials®

Example: PPACT Trial

- Project leaders worked with national Kaiser Permanente to create buy-in for a common instrument

- Local IT staff built it within each region

- A multitiered approach supplemented PROM data collected in clinics at 3, 6, 9, and 12 months

- A follow-up phone call by research staff was necessary to maximize data collection at each time point

NIH PRAGMATIC TRIALS COLLABORATORY

Rethinking Clinical Trials®

Enabling pragmatic research

E-screening, e-enrollment, and e-follow-up
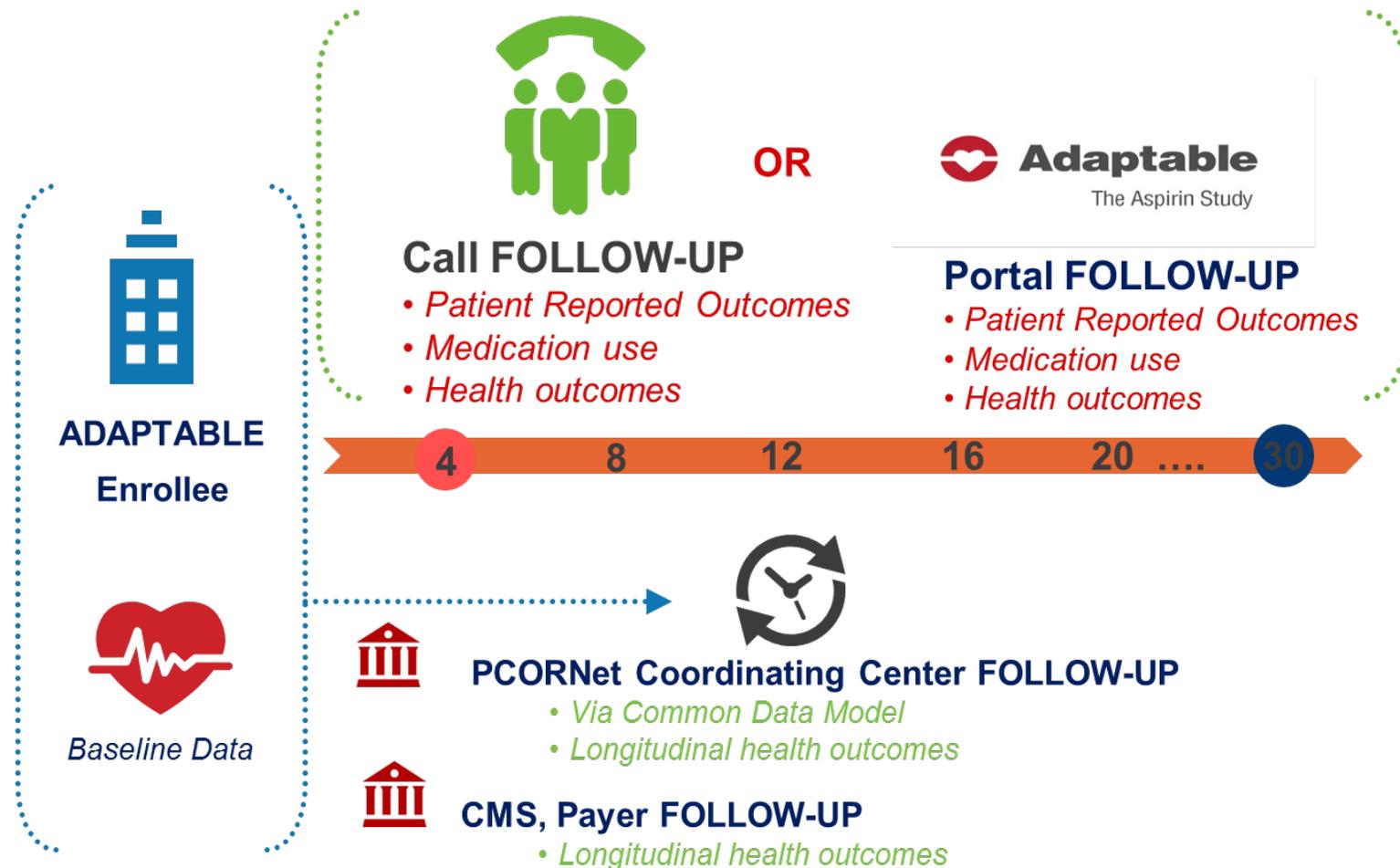
Mobile devices for outcome measurement

- Smartphones, tablet computers, and portable, implantable, or wearable medical devices (mHealth)
  - Some mHealth devices transmit data to a data warehouse every night
  - Largely considered imperfect measures

Data quality assessment

- Identify variation between populations at different sites or in different study groups
- Recommend formal assessment of accuracy, completeness, and consistency for key data
- Data quality should be described, reported, and informed by workflows

Important things to do

- Ask questions that the data will support
- Design trials to minimize new data collection
- Talk to patients and stakeholders when identifying outcomes
- Engage EHR and data experts when defining endpoints
- Budget for data and systems experts at each site (…then double it)
- Carefully consider bias and take steps to promote generalizability
- Develop a data quality assessment plan to improve the value of the data and detect and address data issues early

NIH PRAGMATIC TRIALS
COLLABORATORY
Rethinking Clinical Trials®

Implementation science in simpler terms…

- The intervention, practice, or innovation is **the thing**
- Effectiveness research looks at whether **the thing** works
- Implementation research looks at how best to help people/places **do the thing**
- Implementation strategies are the stuff we do to help people/places **do the thing**
- Main implementation outcomes are **how much** and **how well** they **do the thing**

Curran, *Implementation Science Communications*, 2020

Implementation outcomes

- "The effects of deliberate and purposive actions to implement new treatments, practices, and services."

Proctor, *Adm Policy Ment Health*, 2011

Implementation outcomes

- "The effects of deliberate and purposive actions to implement new treatments, practices, and services." (implementation strategies)

Proctor, *Adm Policy Ment Health*, 2011

Proctor taxonomy of implementation outcomes

Implementation Outcomes

Acceptability
Adoption
Appropriateness
Costs
Feasibility
Fidelity
Penetration
Sustainability

Service Outcomes

Efficiency
Safety
Effectiveness
Equity
Patient-centeredness
Timeliness

Client Outcomes

Satisfaction
Function
Symptomatology

Adapted from Proctor, *Adm Policy Ment Health*, 2011

Proctor taxonomy of implementation outcomes

Implementation Outcomes	Service Outcomes	Client Outcomes
Acceptability	Efficiency	Satisfaction
Adoption	Safety	Function
Appropriateness	Effectiveness	Symptomatology
Costs	Equity	
Feasibility	Patient-centeredness	
Fidelity	Timeliness	
Penetration		
Sustainability		

Adapted from Proctor, *Adm Policy Ment Health*, 2011

NIH PRAGMATIC TRIALS COLLABORATORY
Rethinking Clinical Trials®

Definitions

- **Adoption**: "The intention, initial decision, or action to try or employ an innovation or evidence-based practice."

- **Fidelity**: "The degree to which an intervention was implemented as it was prescribed in the original protocol or as it was intended by the program developers."

- **Penetration**: "Integration of a practice within a service setting and its subsystems. Penetration…can be calculated in terms of the number of providers who deliver a given service or treatment, divided by the total number of providers trained in or expected to deliver the service."

- **Sustainability***: "The extent to which a newly implemented treatment is maintained or institutionalized within a service setting's ongoing, stable operations."

Proctor taxonomy of implementation outcomes

Outcome	Level	Theory	Other terms	Stage	Measurement
Acceptability	Individual provider, patient	"Complexity" from Rogers	Satisfaction	Early	Surveys, interviews
Adoption	Individual provider, org	"Adoption" from RE-AIM	Uptake; utilization; intention to try	Early to mid	Observation, surveys, interviews
Appropriateness	Individual provider, patient, org	"Compatibility" from Rogers	Perceived fit; relevance	Early	Surveys, interviews, focus groups
Feasibility	Individual provider, org	"Compatibility" and "trialability" from Rogers	Practicability	Early	Survey; admin data
Fidelity	Individual provider	"Implementation" from RE-AIM	Delivered as intended; adherence, integrity	Early to mid	Observation; checklist; self-report, admin data
Implementation cost	Provider or org	"Costs" and "resources" from TCU Program Change Model	Marginal cost, cost-effectiveness	Early	Admin data
Penetration	Org	"Reach" from RE-AIM	Spread	Mid to late	Case audit; checklist
Sustainability	Org	"Maintenance" from RE-AIM	Maintenance; integration; institutionalization	Late	Surveys, interviews, checklists

Proctor, *Adm Policy Ment Health*, 2011

NIH PRAGMATIC TRIALS COLLABORATORY
Rethinking Clinical Trials®

Proctor taxonomy of implementation outcomes

Outcome	Level	Theory	Other terms	Stage	Measurement
Acceptability	Individual provider, patient	"Complexity" from Rogers	Satisfaction	Early	Surveys, interviews
Adoption	Individual provider, org	"Adoption" from RE-AIM	Uptake; utilization; intention to try	Early to mid	Observation, surveys, interviews
Appropriateness	Individual provider, patient, org	"Compatibility" from Rogers	Perceived fit; relevance	Early	Surveys, interviews, focus groups
Feasibility	Individual provider, org	"Compatibility" and "trialability" from Rogers	Practicability	Early	Survey; admin data
Fidelity	Individual provider	"Implementation" from RE-AIM	Delivered as intended; adherence, integrity	Early to mid	Observation; checklist; self-report, admin data
Implementation cost	Provider or org	"Costs" and "resources" from TCU Program Change Model	Marginal cost, cost-effectiveness	Early	Admin data
Penetration	Org	"Reach" from RE-AIM	Spread	Mid to late	Case audit; checklist
Sustainability	Org	"Maintenance" from RE-AIM	Maintenance; integration; institutionalization	Late	Surveys, interviews, checklists

Proctor, *Adm Policy Ment Health*, 2011

NIH PRAGMATIC TRIALS COLLABORATORY
Rethinking Clinical Trials®

IMPLEMENTATION SCIENCE

SYSTEMATIC REVIEW

Outcomes for implementation science: an enhanced systematic review of instruments using evidence-based rating criteria

Cara C. Lewis[1,2]*, Sarah Fischer[1], Bryan J. Weiner[3], Cameo Stanick[4], Mimi Kim[5,6] and Ruben G. Martinez[7]

Evidence-Based Assessment (EBA) rating criteria

- **Reliability (or Internal Consistency)**: Extent to which an instrument <u>consistently</u> measures what it is meant to measure

- **Structural (or Construct) Validity**: Extent to which instrument <u>accurately</u> measures what it is meant to measure

- **Predictive (or Criterion) Validity**: Extent to which a measure is related to an outcome

- **Norms**: An estimate of the position of the tested individual in a predefined population, with respect to construct being measured

- **Responsiveness**: Ability of instrument to detect change over time

- **Usability**: Ease with which instrument can be administered

Lewis, *Implementation Science*, 2015

NIH PRAGMATIC TRIALS COLLABORATORY
Rethinking Clinical Trials®

Fidelity measures

- Extent to which an <u>intervention</u> is delivered as intended
  - Core intervention components or functions
- Checklists
- Direct observation
- Questionnaires

Lewis, *Implementation Science*, 2015

Cost

- Costs to deliver an intervention and implementation strategy (direct and/or indirect)

- Cost-effectiveness of an intervention: costs relative to a health outcome (eg, quality-adjusted life years)

- Return on investment: value for money

- Specific algorithms (eg, substance use treatment programs)*

Penetration

- % of "eligible" or target clinicians who deliver the intervention
- Questionnaires: for example, level of institutionalization scale for health promotion programs

Instrumentation issues revealed

- Minimal psychometric testing, reporting, and strength
- Non-use of theories and frameworks
- Frequent use of home-grown, use-once measures
- Redundant development across teams
- Lack of attention to pragmatic relevance (usability)
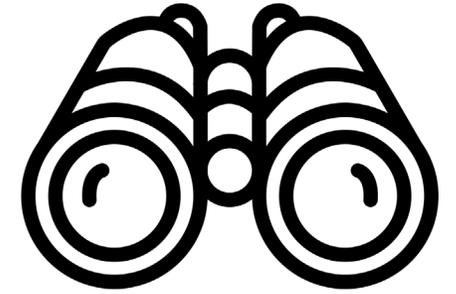
Important things to do

- Look to implementation science protocol papers to see which measures are commonly used

- Increase availability of instruments with promising psychometric properties to further establish their quality

- Work to establish psychometrics of underdeveloped measures

- Engage key partners

NIH PRAGMATIC TRIALS COLLABORATORY
Rethinking Clinical Trials®

Q&A

NIH PRAGMATIC TRIALS COLLABORATORY
Rethinking Clinical Trials®

Knowledge checkpoint

- Getting longitudinal data from an insurance carrier does not require explicit consent from the patient.
  - True
  - False

Knowledge checkpoint

- Suicide attempts would be a good endpoint to use in pragmatic clinical trials since they are routinely documented.
  - True
  - False

Knowledge checkpoint

- Mobile devices are promising means of obtaining data although presently largely imperfect measures.
    - True
    - False