False Precision and Estimating the Reliability of Effects with the Traditional Evidence Generating Process

Robert M Califf MD

Vice Chancellor for Health Data Science

Duke University

Advisor, Verily Life Sciences

NAM September 20th, 2017

Preface

- What follows is partially a spoof to be consistent with the aura of the meeting
- I do believe that the “traditional” system of evidence generation has delivered medical products and delivery that has had a major impact on life expectancy, physical function and ability to enjoy life
- However, I also believe that the traditional evidence generation system has become bloated and burdened with practices that are damaging because they massively increase cost without improving quality and in many cases making quality worse because the SOPs have become more important than the science of clinical investigation.
- One of the damaging parts of the effort is the “asymptotic” effort to record each data item more precisely in the mistaken notion that a more reliable estimate of treatment effect will be derived for outcomes that matter
- **By a combination of refocusing on quality by design and use of automation and analytical methods, we can dramatically accelerate the generation of evidence while also improving the quality of the result**

We Believe…

- We are in an explosive phase of science that is making many new approaches to prevention, diagnosis and treatment possible
- The proliferation of technologies is fueling an increase in cost that will demand demonstration of value for new therapies
- The overall rising cost of health care requires an assessment of the value of old therapies already in use, particularly when there is a choice of a more expensive new therapy
- Due to increasing transparency people are more aware of evidence gaps
- The sum result is a dramatic need to answer more questions about therapeutics
- The current evidence generation system has unsustainable costs

The Old Doctrine..

- Was built at a time when automation was not possible because few electronic data existed
- Doctors wrote things on paper and then as time evolved this information was transferred to a computer
- In order to check that the information worked nurses and others had to fly around on airplanes and check that the information matched
- The regulators divised a doctrine (GCP) to enable these (now) arcane practices to be done in a consistent fashion
- Local experts took the doctrine and amplified it in voluminous SOP’s

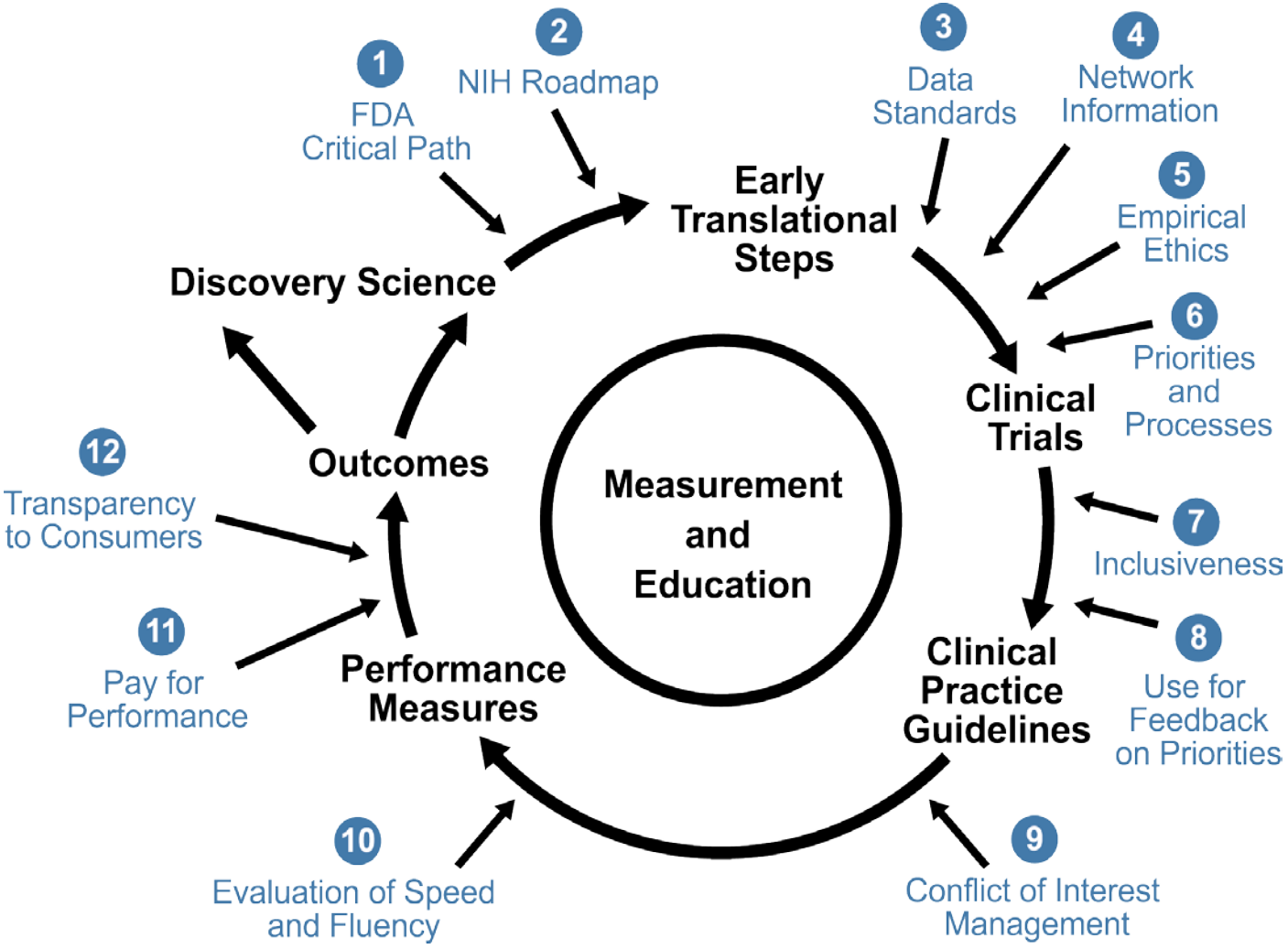
Thesis

- In a mistaken understanding of the theory and purpose of clinical trials, the regulated clinical trials industry has diverted enormous resources to an effort to increase precision
- The academic/NIH driven clinical research industry has adopted some of this thinking through the proliferation of “GCP”
- Tearing down the structure would be counter-productive–people need structure to conduct these complex human experiments
- To produce reliable clinical trial results that inform patients, carers, doctors (providers), health systems, payors and policy makers, we need to focus on reliable results

Common Definitions

- Precision
  - the quality, condition, or fact of being exact and accurate.
  - refinement in a measurement, calculation, or specification, especially as represented by the number of digits given.
- Reliable
  - consistently good in quality or performance; able to be trusted.
  - Synonyms: dependable, good, well founded, authentic, valid, genuine, sound, true

Generating Evidence to Inform Decisions



Our National Clinical Research System is Well-intentioned But Flawed

- High percentage of decisions not supported by evidence*
- Health outcomes and disparities are not improving
- Current system is great **except**:
  - Too slow, too expensive, and not reliable
  - Doesn't answer questions that matter most to patients
  - Unattractive to clinicians & administrators

We are not generating the evidence we need to support the healthcare decisions that patients and their doctors have to make every day.

Tricoci P et al. JAMA 2009;301:831-41

Which Treatment is Best for Whom?

High-Quality Evidence is Scarce

< 15% of Guideline Recommendations Supported by High Quality Evidence

ORIGINAL CONTRIBUTION

Scientific Evidence Underlying the ACC/AHA Clinical Practice Guidelines

Pierluigi Tricoci, MD, MHS, PhD
Joseph M. Allen, MA
Judith M. Kramer, MD, MS
Robert M. Califf, MD
Sidney C. Smith Jr, MD

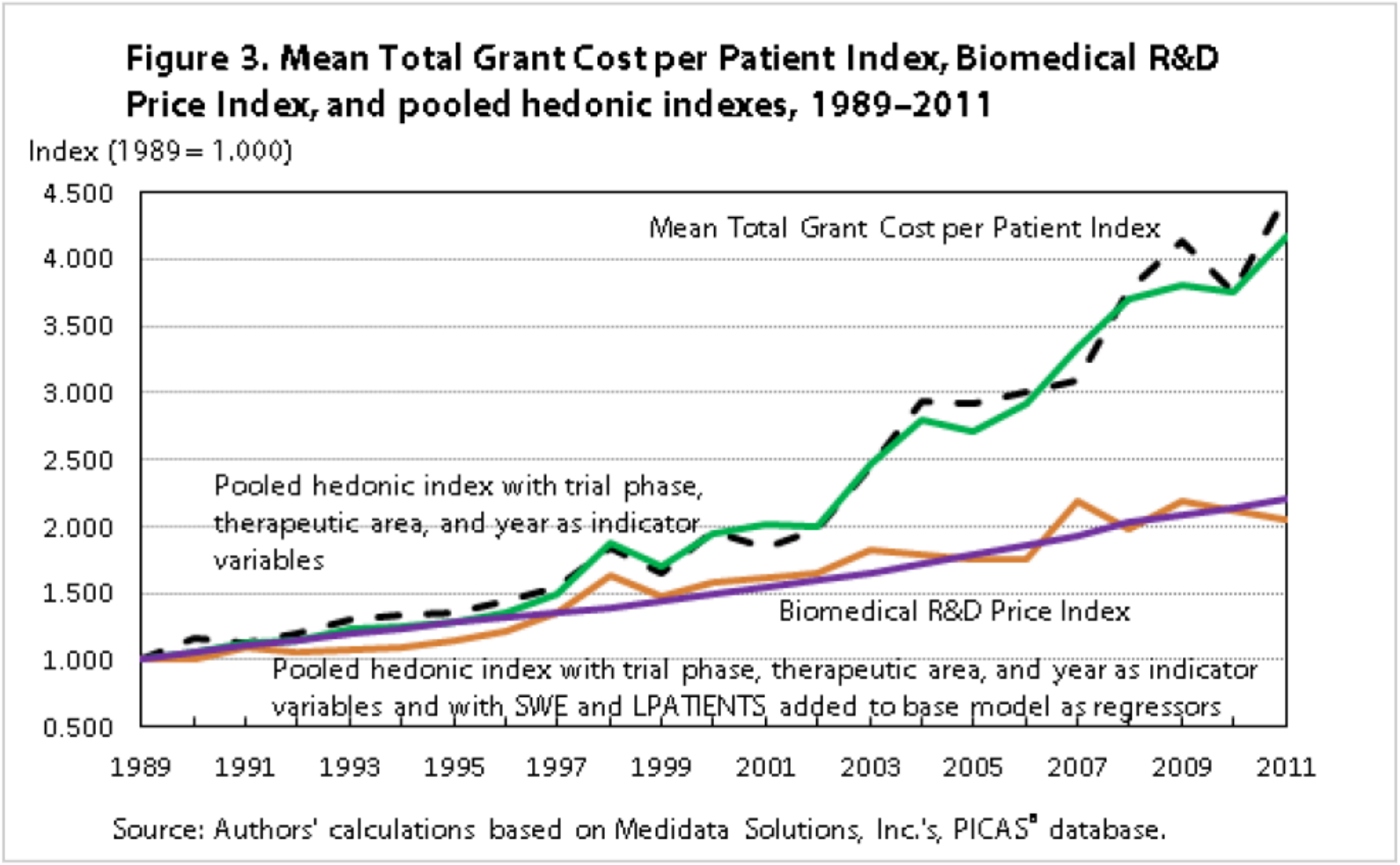
CLINICAL PRACTICE GUIDELINES are systematically developed statements to assist practitioners with decisions about appropriate health care for spe-

Context The joint cardiovascular practice guidelines of the American College of Cardiology (ACC) and the American Heart Association (AHA) have become important documents for guiding cardiology practice and establishing benchmarks for quality of care.

Objective To describe the evolution of recommendations in ACC/AHA cardiovascular guidelines and the distribution of recommendations across classes of recommendations and levels of evidence.

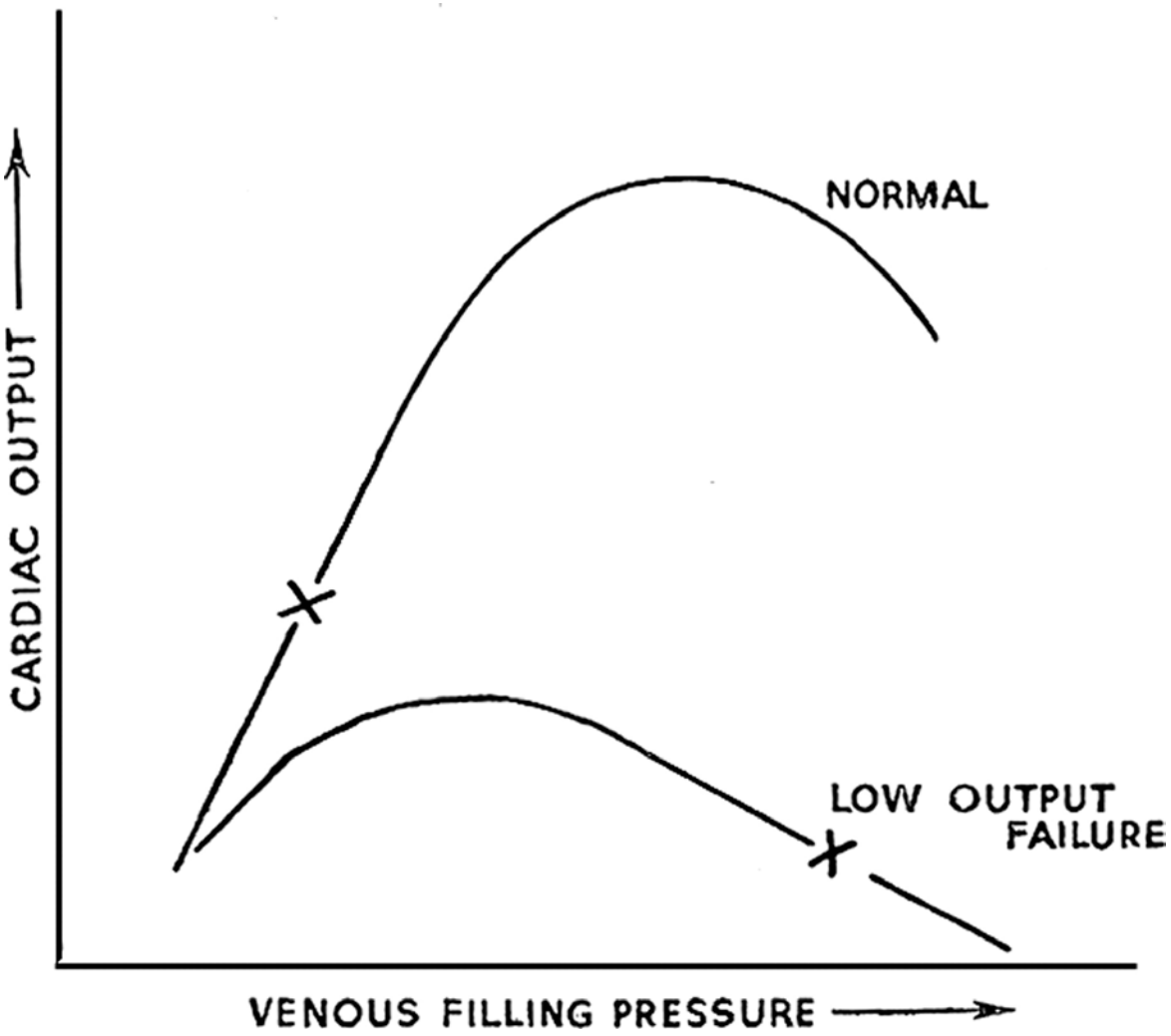
Data Sources and Study Selection Data from all ACC/AHA practice guidelines issued from 1984 to September 2008 were abstracted by personnel in the ACC Science and Quality Division. Fifty-three guidelines on 22 topics, including a total of 7196 recommendations, were abstracted.

Trial Hyperinflation



Berndt E, Cockburn I. Monthly Labor Review, June 2014

Figure 3. Diagram illustrating the failing heart operating on the descending limb of the Starling curve.



Arnold M. Katz Circ Heart Fail. 2008;1:63-71

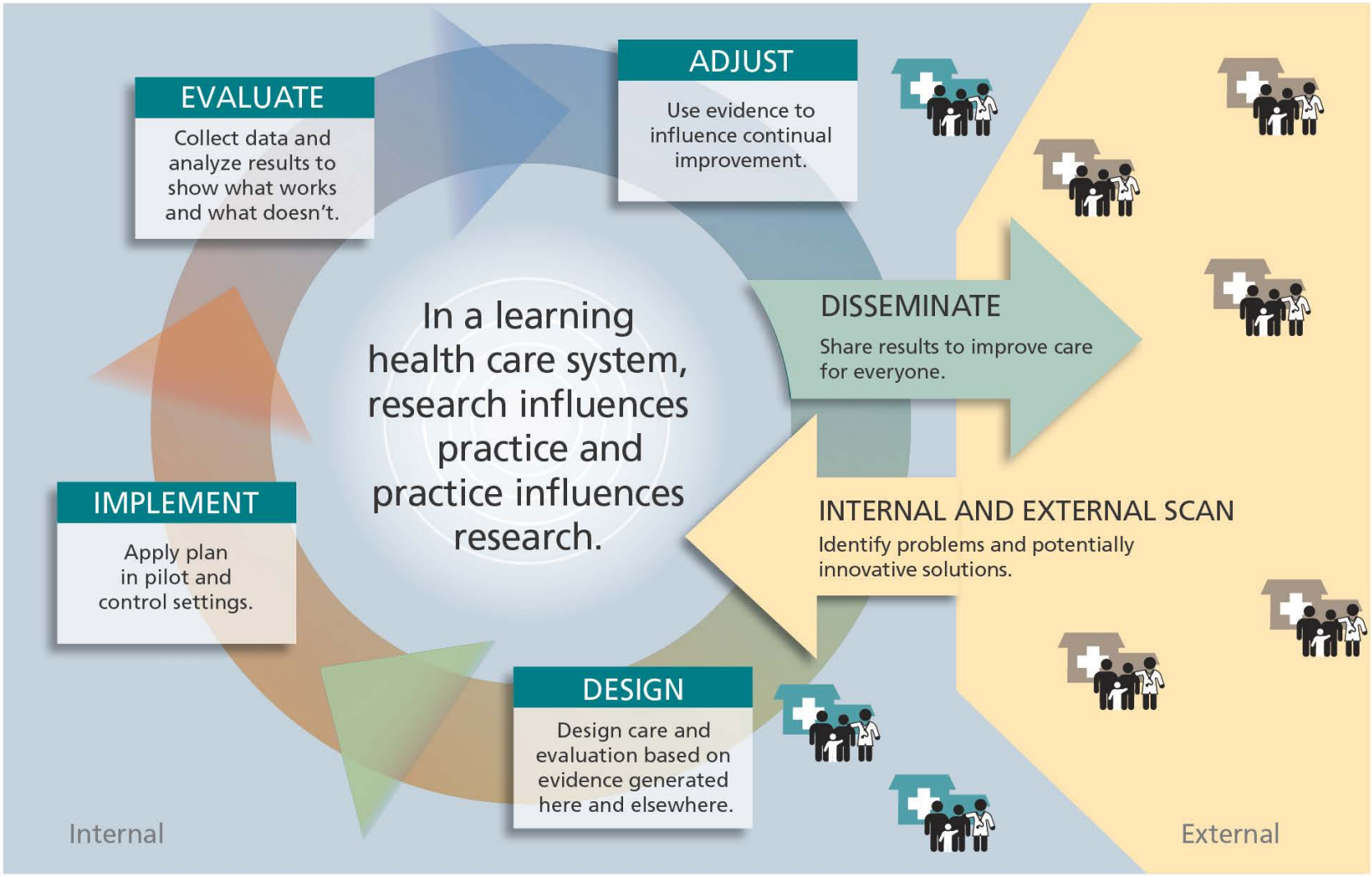
Definition of “Real World Evidence”—The Law

- **§355g. Utilizing real world evidence**
- **(a) In general**
- The Secretary shall establish a program to evaluate the potential use of real world evidence-
- (1) to help to support the approval of a new indication for a drug approved under section 355(c) of this title; and
- (2) to help to support or satisfy postapproval study requirements.
- **(b) Real world evidence defined**
- In this section, the term "real world evidence" means data regarding the usage, or the potential benefits or risks, of a drug derived from sources other than traditional clinical trials.

Key Elements to Get a New Start

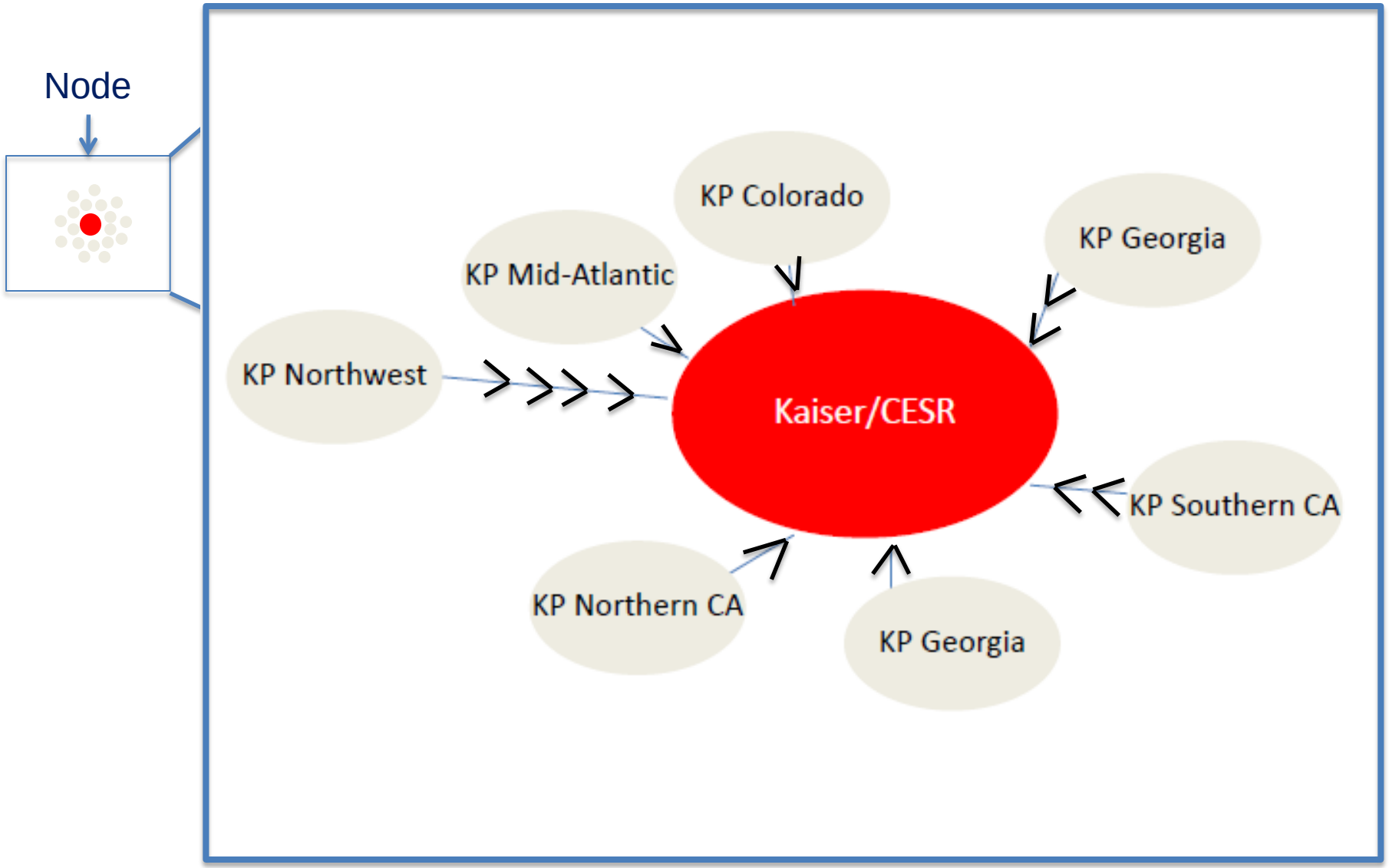
- **Build a reusable system embedded in practice**
- Use quality by design
- Use automation for repetitive tasks
- Operate from basic principles

Learning health care systems

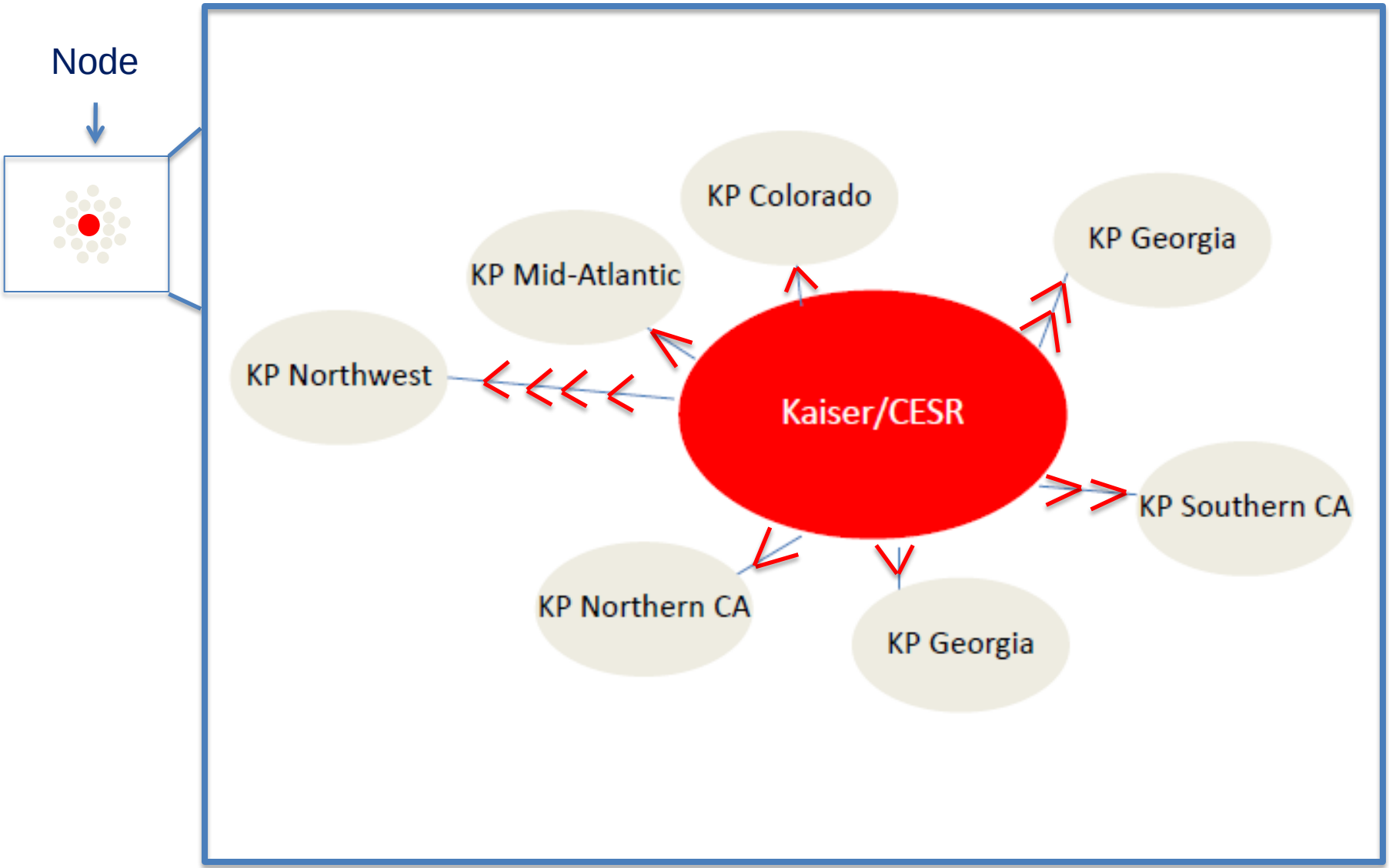


Public private partnerships are developing that could generate reliable evidence rapidly

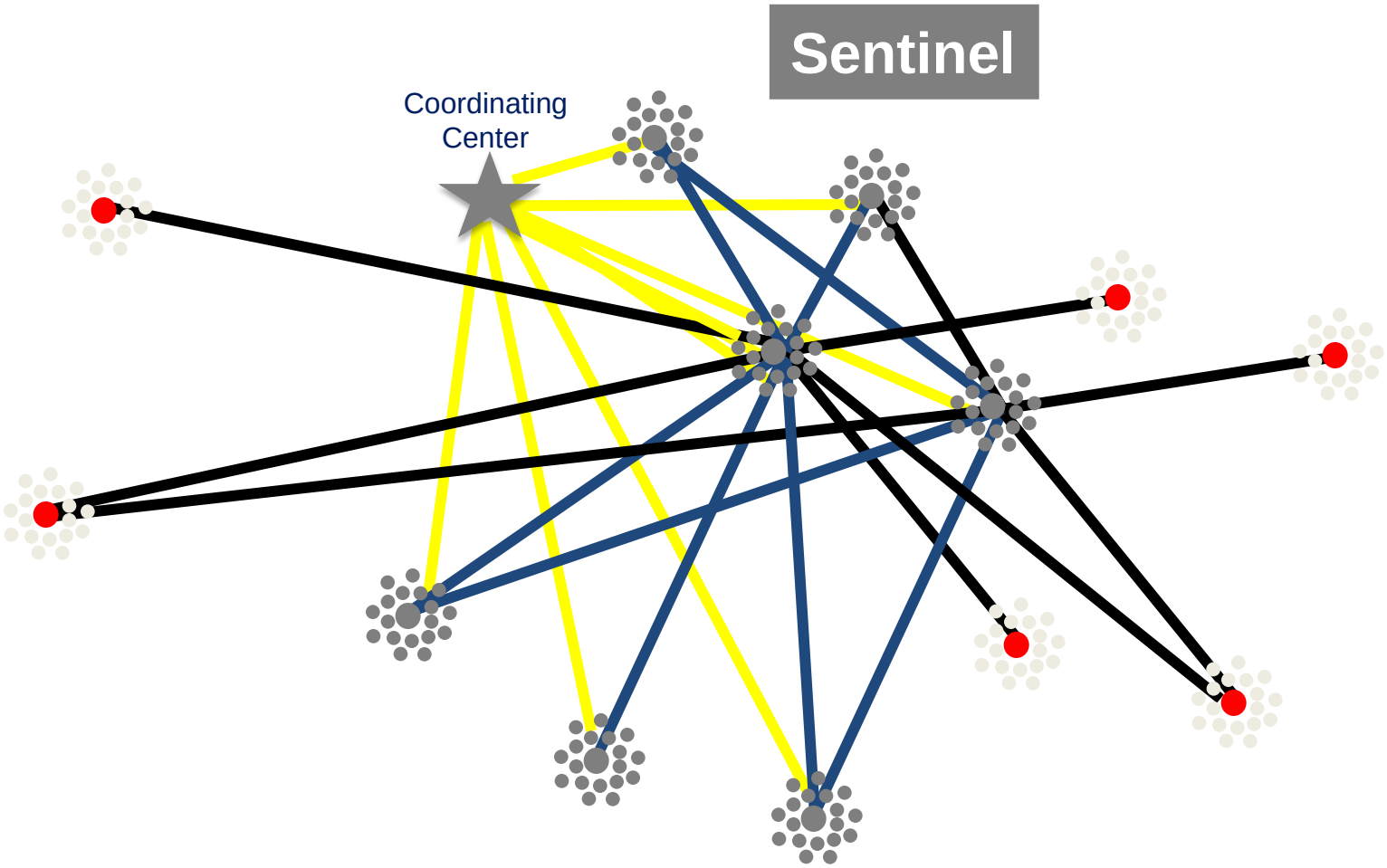
Previously Independent Sites now part of large integrated health systems
increasingly sophisticated data warehouses



Nodes are Operational Clusters Using Common Data

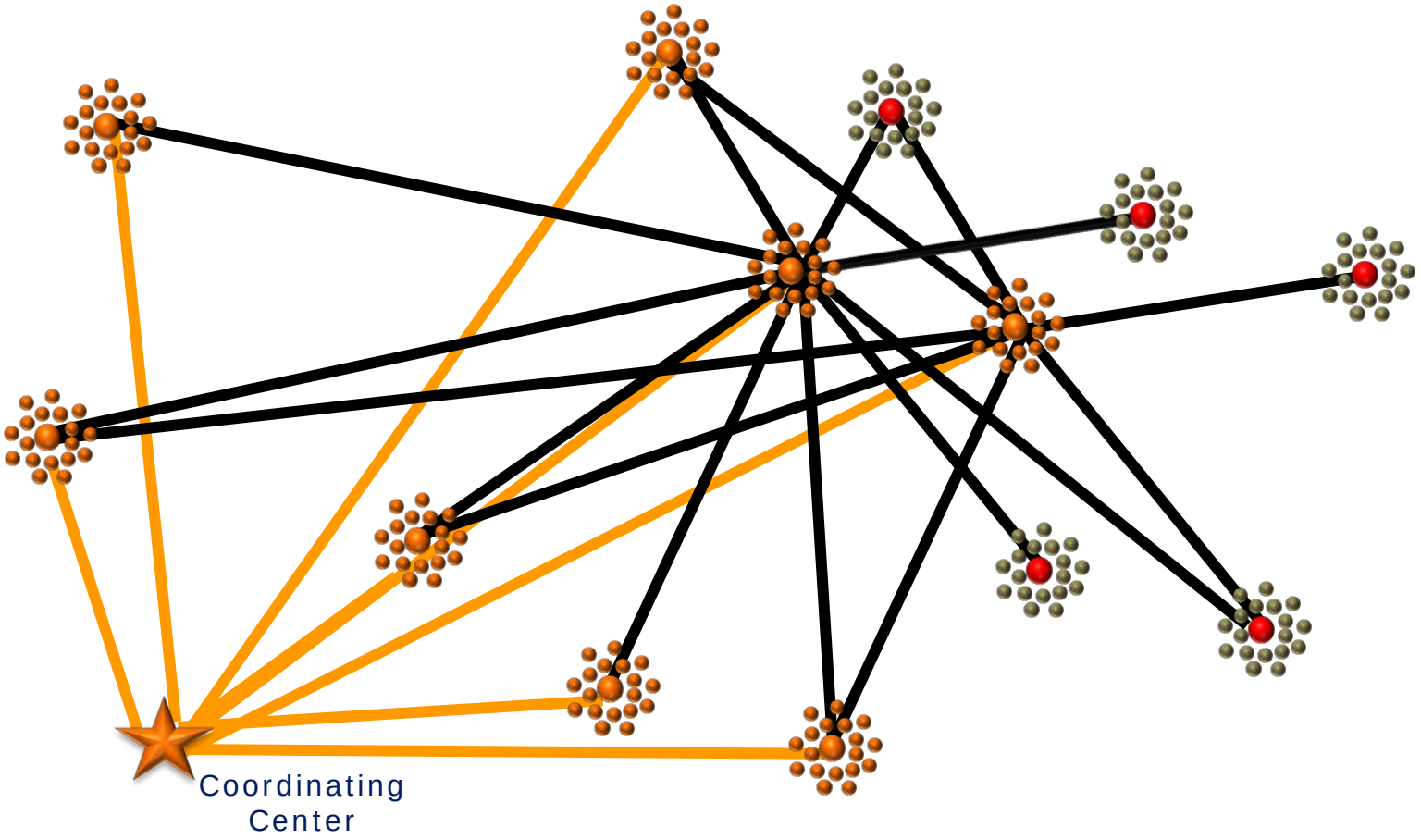


Drug Surveillance and Trials



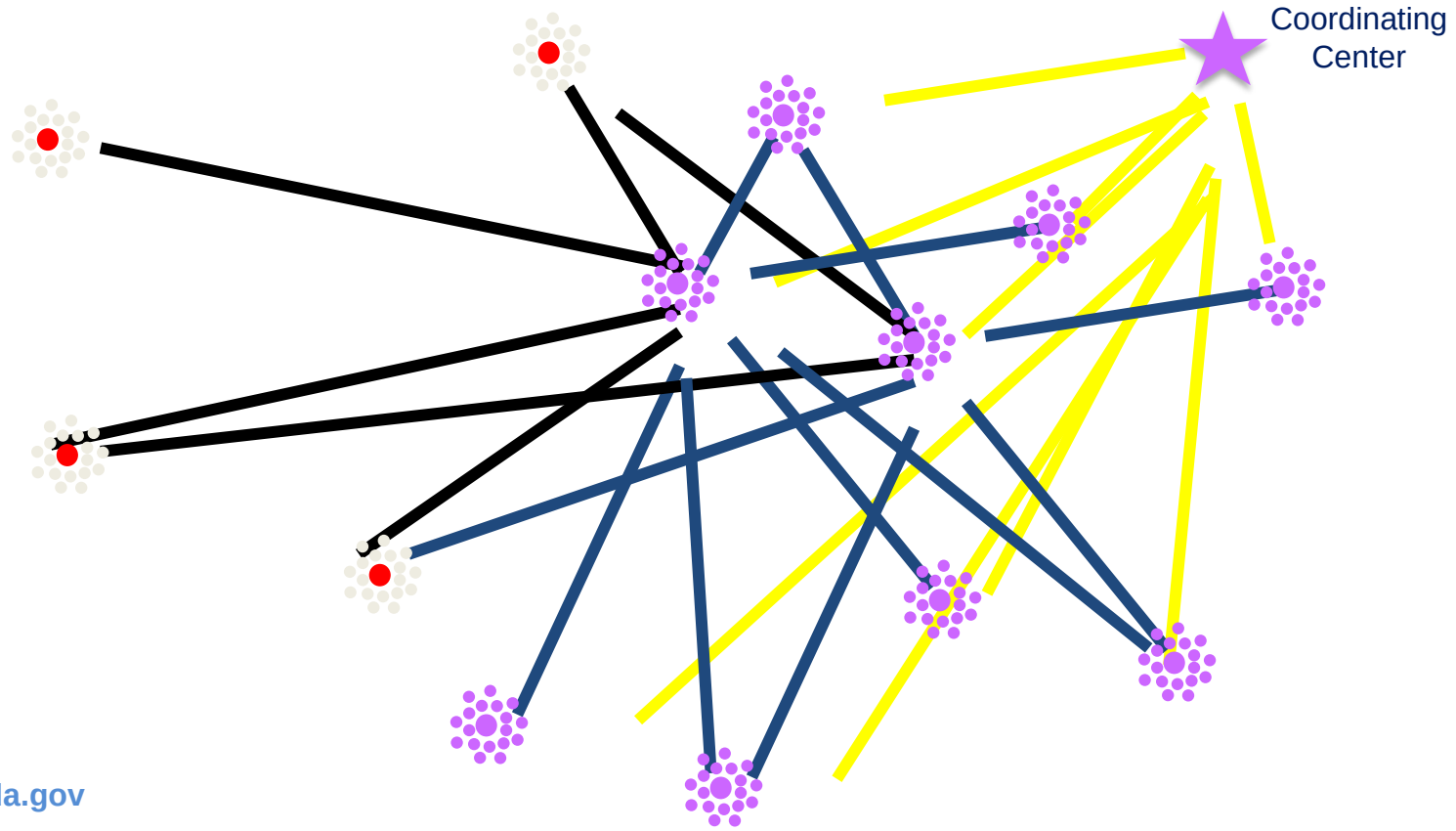
Device Surveillance and Trials

NEST

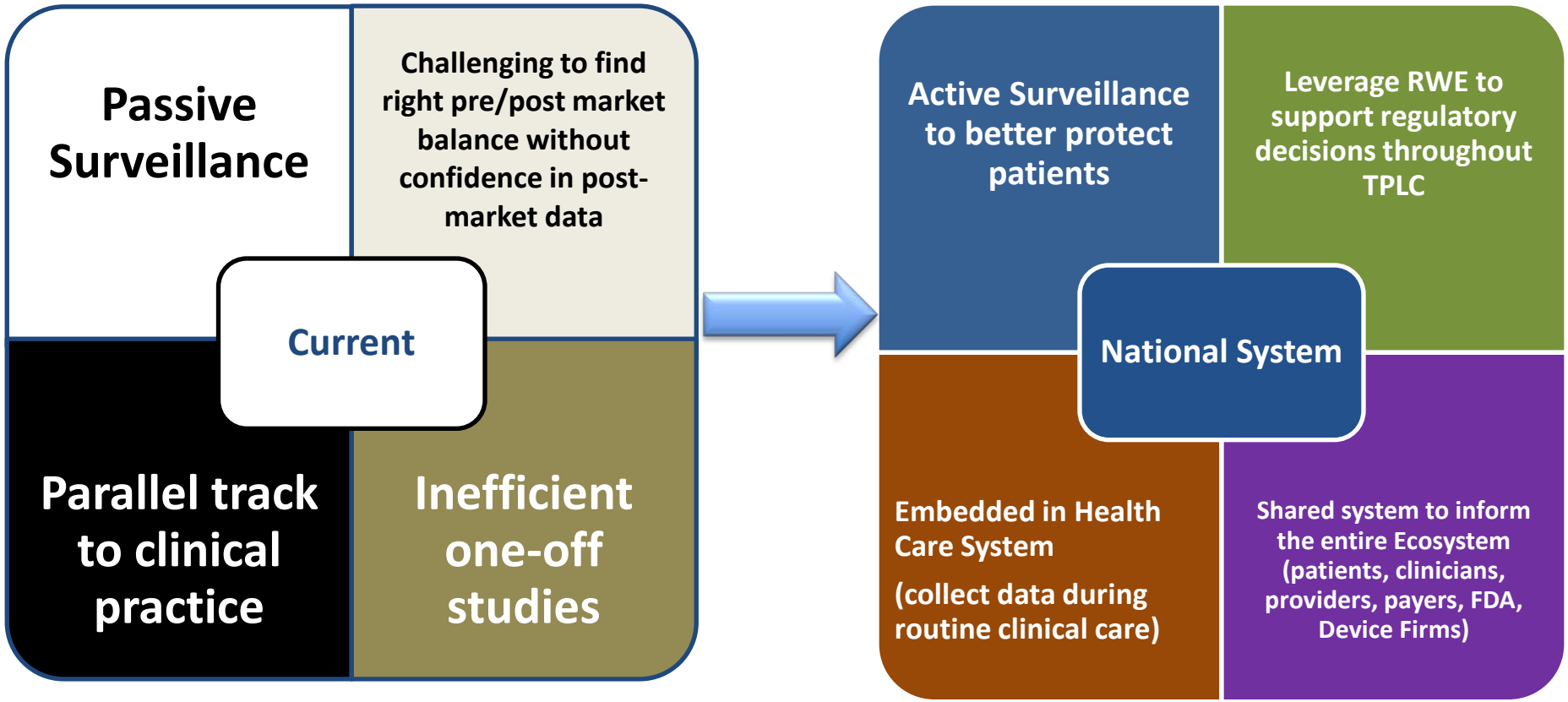


Post Market Studies, including comparative effectiveness

PCORnet



National System Paradigm Shift



Demonstration Project Overview-NIH Healthcare Systems Research Collaboratory

10 Demonstration Projects spanning 12 NIH institutes and centers

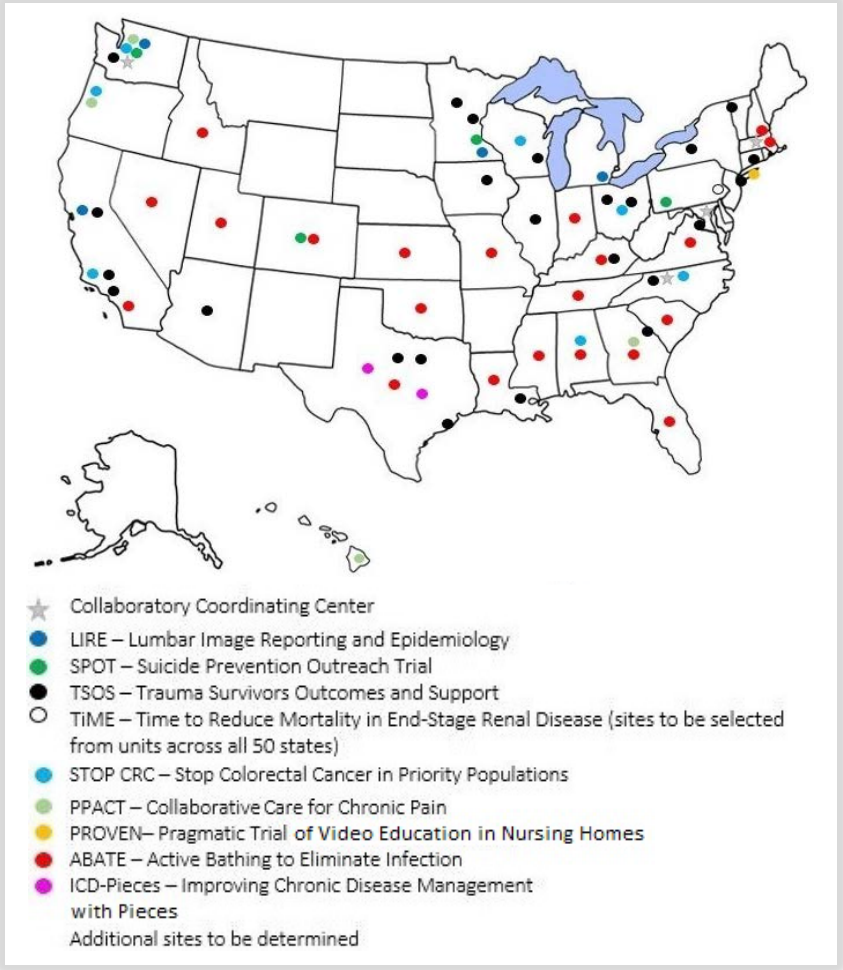
Major clinical outcome trials

1-year planning phase (UH2)

Implementation phase (UH3)

Using EHRs and minimal additional data collection

Log order reduction in cost



Key Elements to Get a New Start

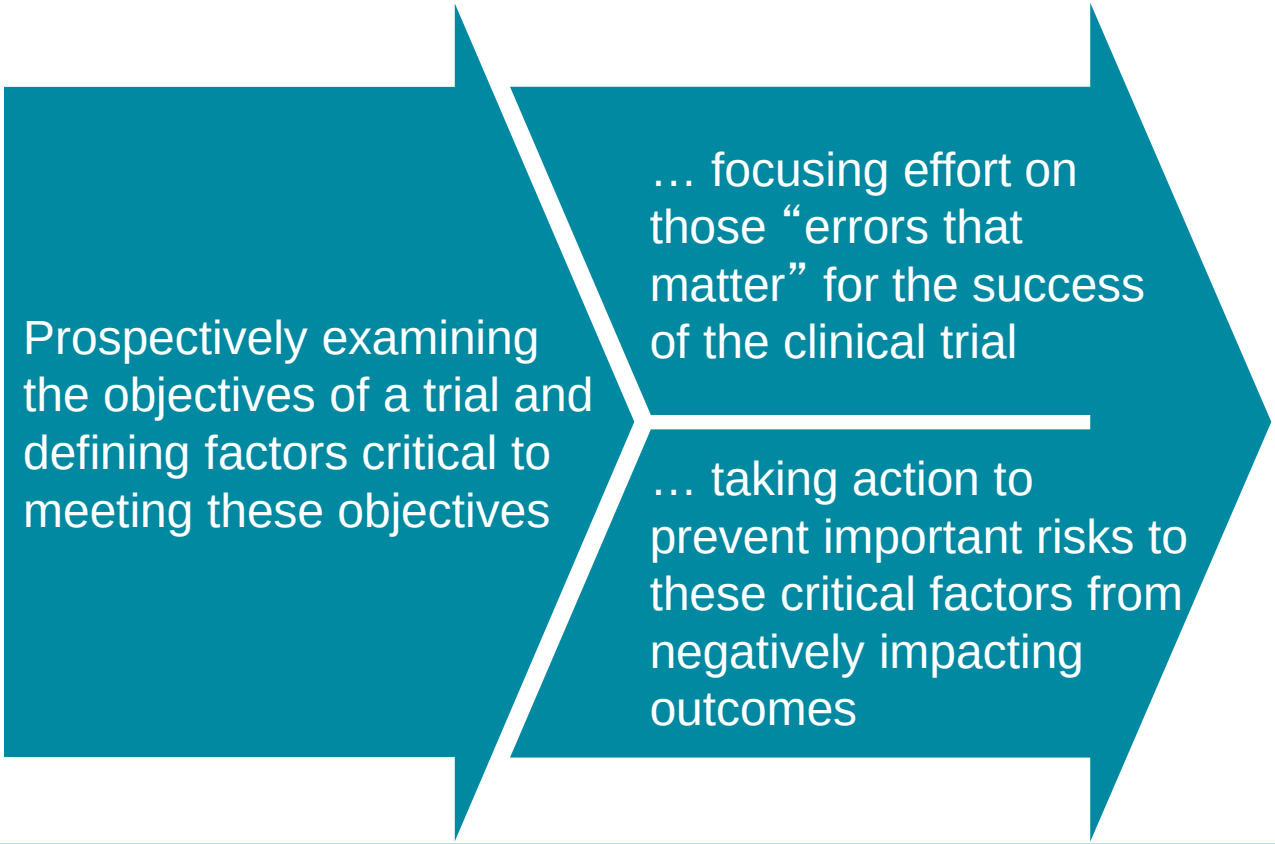
- Build a reusable system embedded in practice
- **Use quality by design**
- Use automation for repetitive tasks
- Operate from basic principles

What Is Quality by Design?

CLINICAL
TRIALS
TRANSFORMATION
INITIATIVE

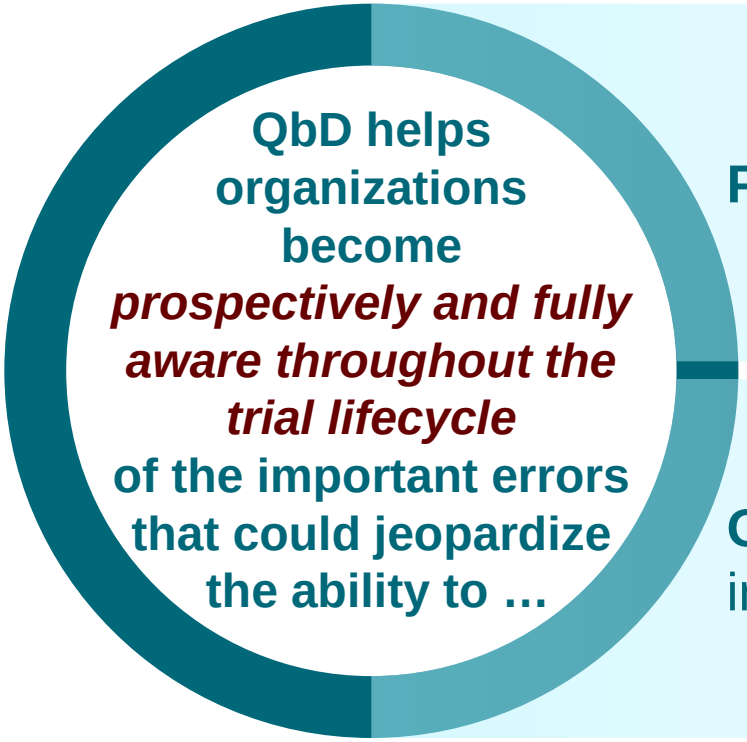
Quality by Design: QbD Defined

"Quality" in clinical trials is defined as the absence of errors that matter



CTTI

How QbD Improves Clinical Trials



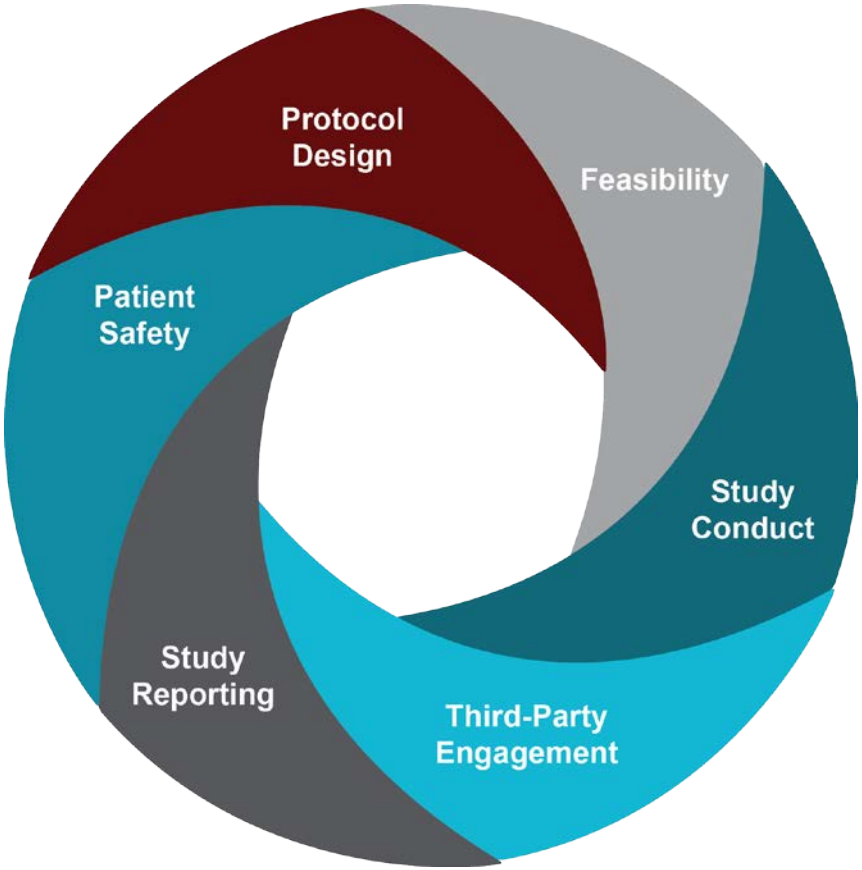
Protect patients during the trial

Obtain reliable results and meaningful information from the trial

CTTI

QbD Step 1

Identify “critical to quality” factors (CTQs) for your specific trial



CTTI

QbD Step 2

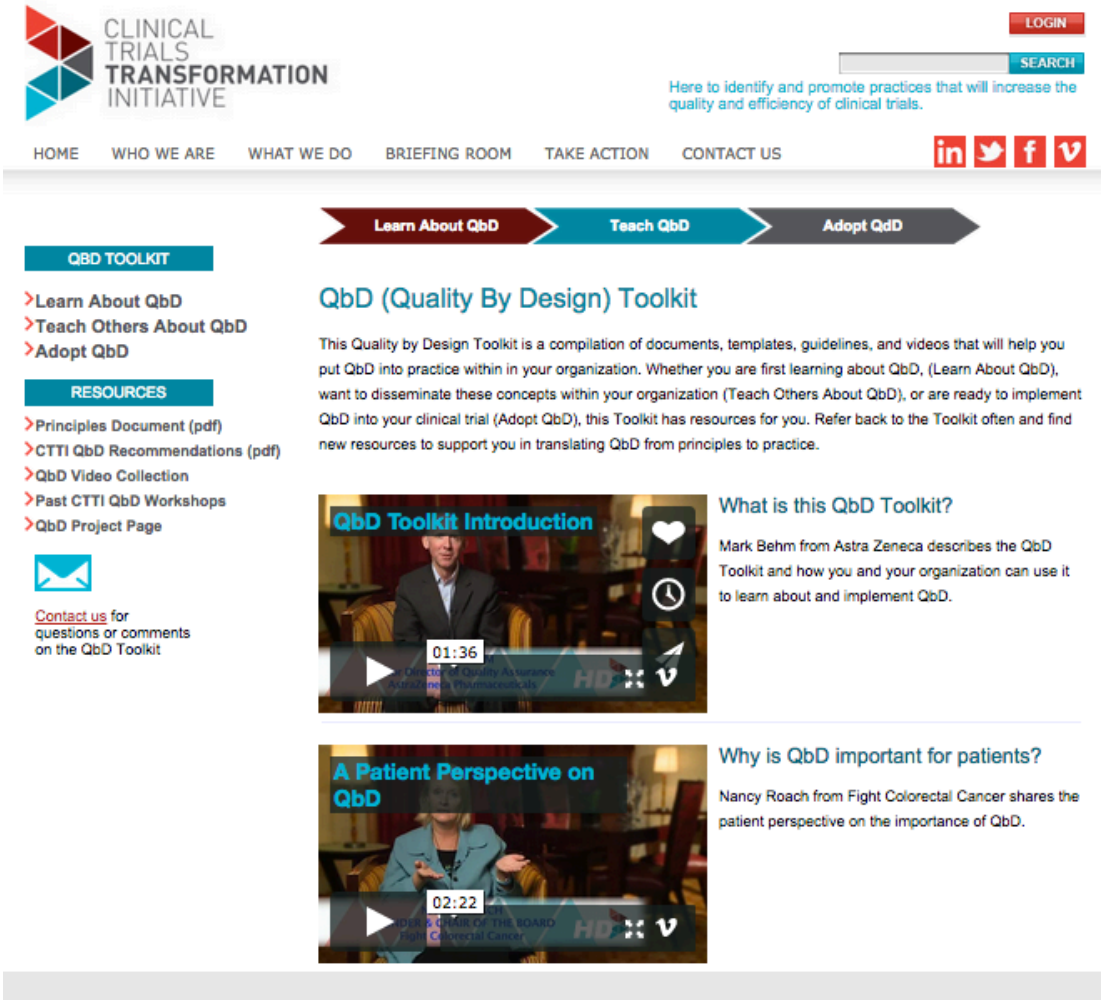
Discuss potential risks related to each CTQ identified that impact study quality (i.e., participant safety or credibility of results)

QbD Step 3

Mitigate those risks that will likely lead to errors that matter and determine how to rapidly identify and react when there is an issue

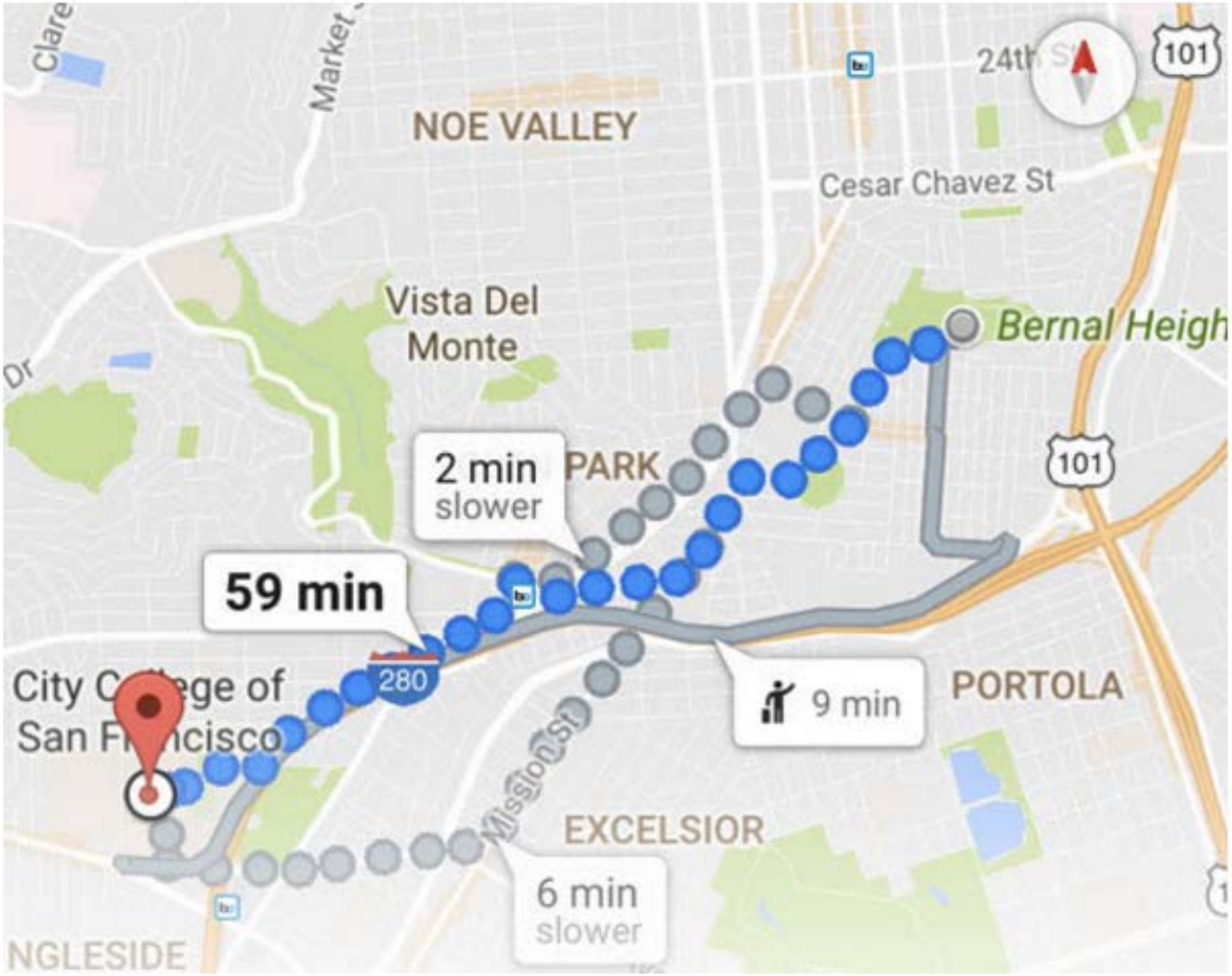
Use the QbD Toolkit

http://www.ctti-clinicaltrials.org/toolkit/QbD

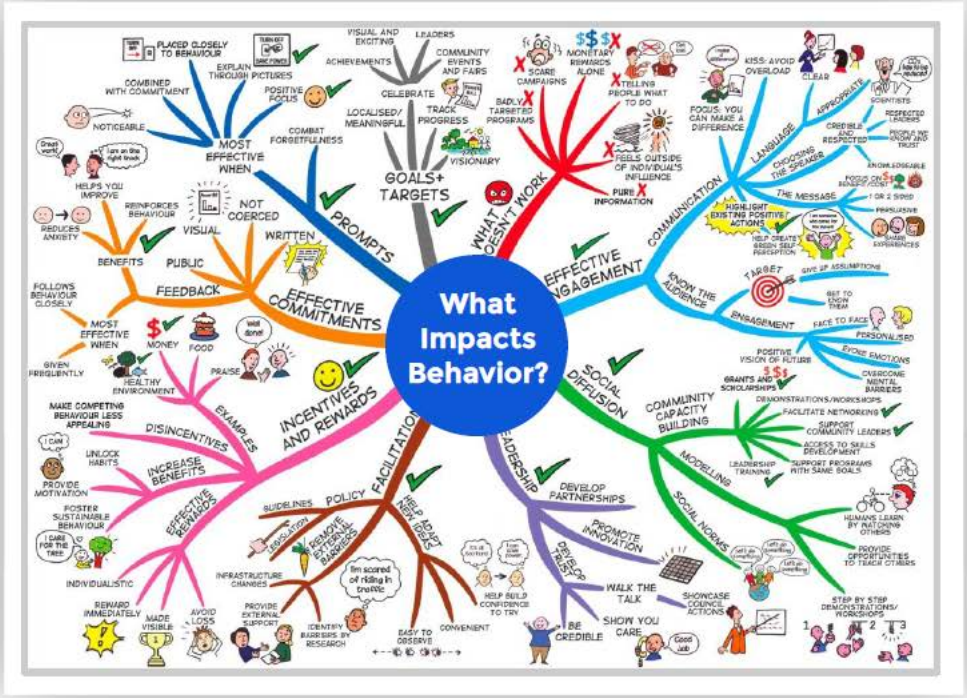


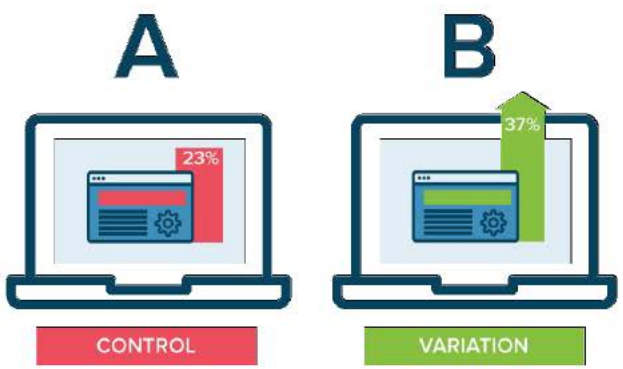
Key Elements to Get a New Start

- Build a reusable system embedded in practice
- Use quality by design
- **Use automation for repetitive tasks**
- Operate from basic principles


NOE VALLEY
Cesar Chavez St
Vista Del Monte
Bernal Heigh
2 min slower
PARK
59 min
101
280
9 min
PORTOLA
EXCELSIOR
6 min slower
Mission St

Data Activation and Testing Outcomes





Verily |

Key Elements to Get a New Start

- Build a reusable system embedded in practice
- Use quality by design
- Use automation for repetitive tasks
- **Operate from basic principles**

Examples of Basic Principles

- Errors that matter
  - Distinguishing random error from systematic error
- Enroll participants likely to inform the question
- Randomization
- Masking
- Measure outcomes in manner that is fit for purpose
- Strengths of different designs for different purposes
- Designing operations that yield an answer to the question in an efficient manner (“the answer is in the question”)

Levy's case

- Trial to determine if Drug X lowers the risk of death and stroke in patients with atrial fibrillation
- Sample size of 18000 and follow-up time of multiple years needed to answer this question
- Thousands of patients had already been studied to determine dosing regimen
- What would it cost to precisely record the time of ingestion of each dose of the twice a day drug?
  - For what purpose?

McCollister-Slipp Case

- People with diabetes live a long time on medicines
- Comparative effect of medicines on CV outcomes (”macrovascular”), including death, were not known
  - Using traditional GCP doctrine, cost per trial is hundreds of millions
- Comparative effect of medicines on neuropathy, retinopathy, nephropathy (“microvascular”) unknown
  - Using traditional methods of assessment
- Value of new, continuous measures of glucose and other biological parameters not known
- Suggested approach:
  - For CV outcomes streamline trial and use health system data to reduce cost dramatically
  - For microvascular outcomes, automate measurements in populations over time
  - Automate natural history studies to sort out which continuous measures of glucose and metabolism work

Twitter's C.E.O., Dick Costolo, Is Set to Exit, Feeling Heat of Criticism

ROBOTICA EPISODE 5
Sex Dolls That Talk Back

STATE OF THE ART
For Twitter, Future Means Here and Now

Sidewalk Labs, a Start-Up Created by Google, Has Bold Aims to Improve City Living

TECHNOLOGY

For Big-Data Scientists, 'Janitor Work' Is Key Hurdle to Insights

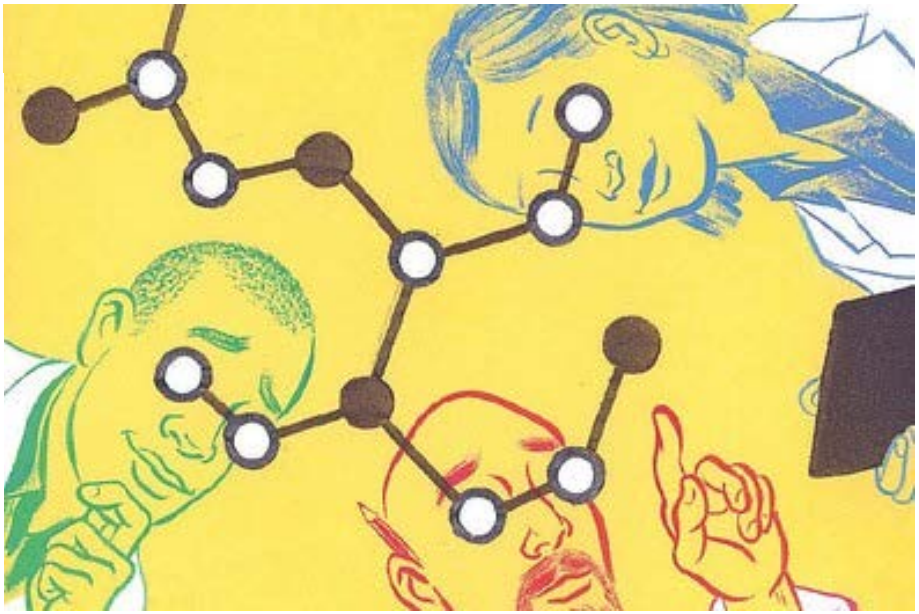
By STEVE LOHR AUG. 17, 2014

The New Einsteins Will Be Scientists Who Share

From cancer to cosmology, researchers could race ahead by working together—online and in the open

By MICHAEL NIELSEN

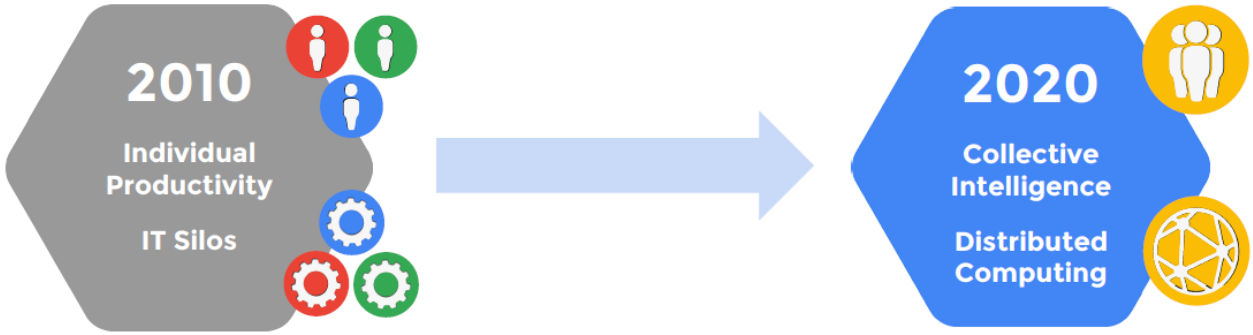
In January 2009, a mathematician at Cambridge University named Tim Gowers decided to use his blog to run an unusual social experiment. He picked out a difficult mathematical problem and tried to solve it completely in the open, using his blog to post ideas and partial progress. He issued an open invitation for others to contribute their own ideas, hoping that many minds would be more powerful than one. He dubbed the experiment the Polymath Project.



Several hours after Mr. Gowers opened up his blog for discussion, a Canadian-Hungarian mathematician posted a comment. Fifteen minutes later, an Arizona high-school math teacher chimed in. Three minutes after that, the UCLA mathematician Terence Tao commented. The discussion ignited, and in just six weeks, the mathematical problem had been solved.

THE WALL STREET JOURNAL.

Digital Transformation



- Data on premise, hard to access, analyze and use
- Productivity tools built for individual, local usage
- IT focusing on **where** it computes

- Data stored in cloud, simple to query
- Collaborative, cloud based productivity applications
- Machine learning drives deep, actionable insights
- IT changing **how** it computes

Verily

Recognizing the risk of creating another bureaucratic entanglement, we need to work across the ecosystem to codify the key principles and operating methods for evidence generation using real world data to produce real world evidence

The effort should embrace principles that enable quality design to produce reliable results