

# MMCI Practical Data Science

## lecture 02

erich s. huang  
erich.huang@duke.edu

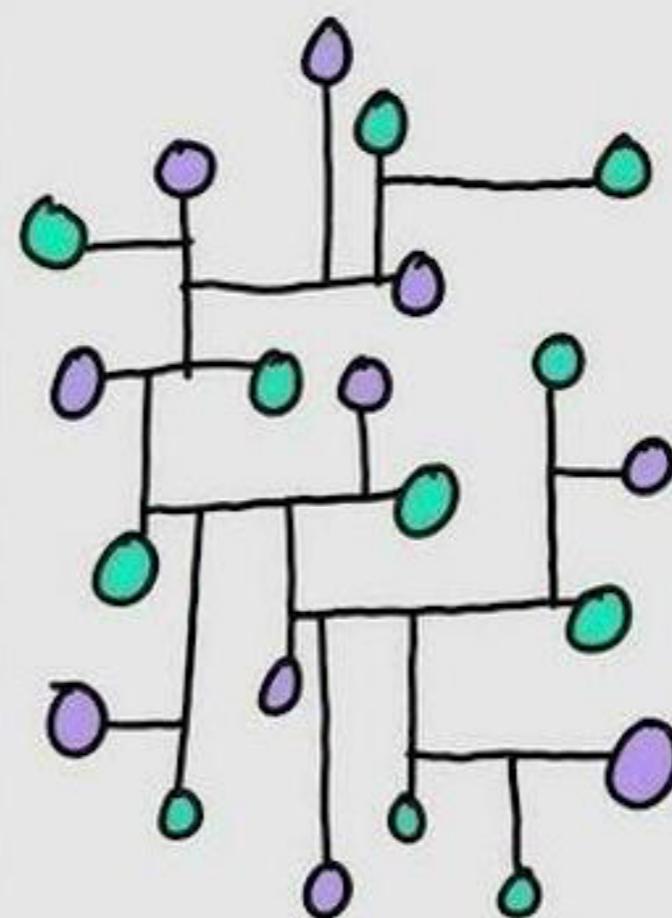
Scene from "The Knick"

How is knowledge acquired?

## **Knowledge**



## **Experience**



# The NEW ENGLAND JOURNAL of MEDICINE

VOL. 366 NO. 2

ESTABLISHED IN 1812

JANUARY 12, 2012

NEJM.ORG



200 NEJM

108 THIS WEEK AT NEJM.ORG

#### PERSPECTIVE

- 101 The Politics of Emergency Contraception A.J.J. Wood, J.M. Drazen, and M.F. Greene  
103 The Constitutionality of the ACA's Medicaid Expansion Mandate I.G. Cohen and J.F. Blumstein  
105 Expanding Eligibility, Cutting Costs — A Medicaid Update J.K. Iglehart  
e4 Achieving Accountable Care — "It's Not About the Bike" J. Walker and A. McKethan

#### ORIGINAL ARTICLES

- 109 Pertuzumab plus Trastuzumab plus Docetaxel for Metastatic Breast Cancer J. Baselga and Others  
120 Subclinical Atrial Fibrillation and the Risk of Stroke J.S. Healey and Others  
130 Cardiac Arrest during Long-Distance Running Races J.H. Kim and Others  
141 Germline Mutations in HOXB13 and Prostate-Cancer Risk C.M. Ewing and Others  
**SPECIAL ARTICLE**  
150 Fitness Memberships and Favorable Selection in Medicare Advantage Plans A.L. Cooper and A.N. Trivedi

#### CLINICAL PRACTICE

- 158 Hidradenitis Suppurativa G.B.E. Jemec  
165 Paragonimiasis M.A. Barrientos and A.U. Carrasco  
e5 Jugular Venous C-V Wave in Severe Tricuspid Regurgitation N.B. Senguttuvan and G. Karthikeyan

#### CASE RECORDS OF THE MASSACHUSETTS GENERAL HOSPITAL

- 166 A Man with Persistent Ulcers on the Hands D. Kroshinsky, M.P. Hoang, and R.P. Hasserjian

#### EDITORIALS

- 176 HER2 Therapy — An Abundance of Riches W.J. Gradishar

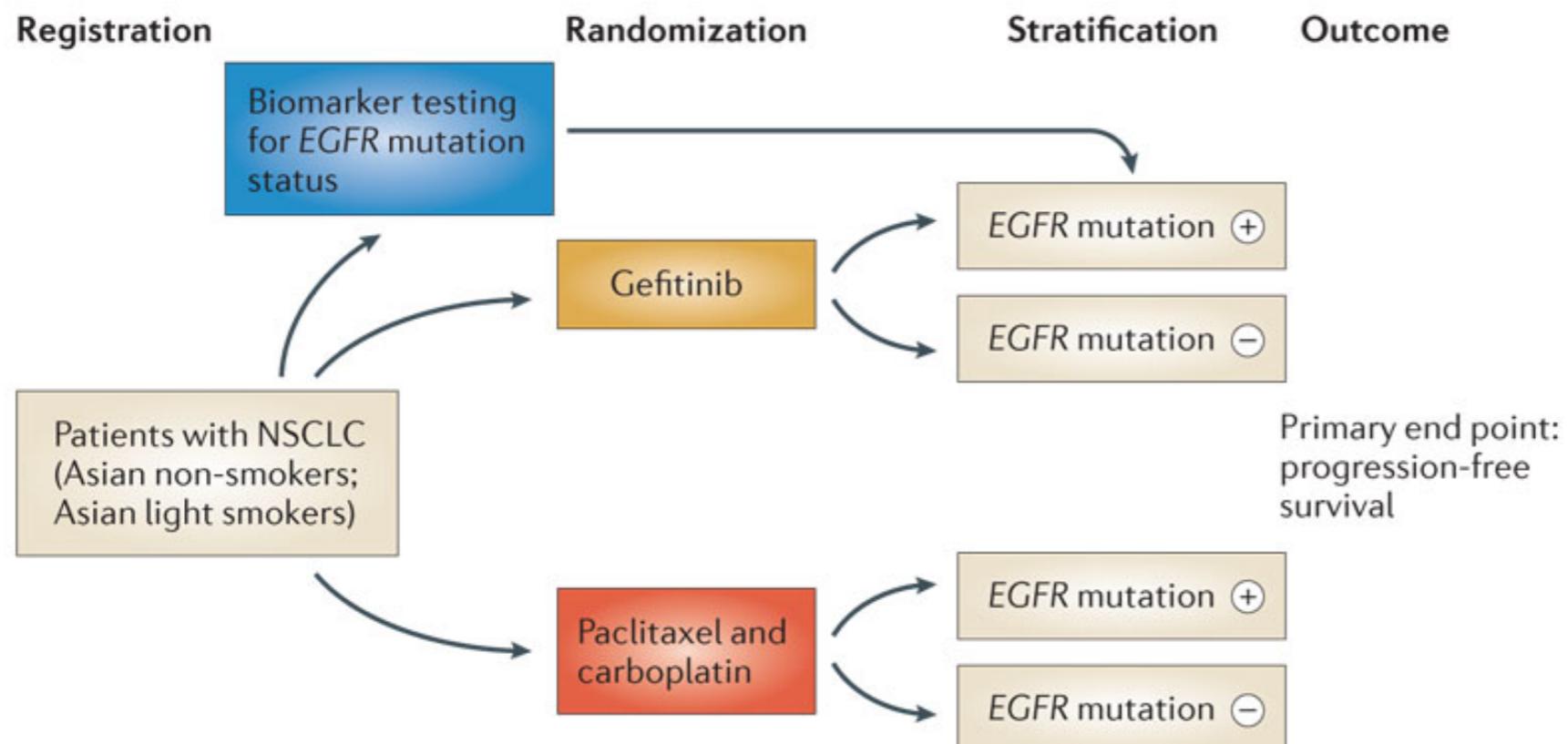
#### CLINICAL IMPLICATIONS OF BASIC RESEARCH

- 181 Diet and Intestinal Immunity H. Tilg  
184 CORRESPONDENCE  
Acute Coronary Thrombosis in Boston Marathon Runners  
An Impedance Threshold Device in Out-of-Hospital Cardiac Arrest  
Breast-Cancer Adjuvant Therapy with Zoledronic Acid  
Inflammatory Bowel Disease and ADAM17 Deletion  
Breast-Cancer Screening  
Generalizing Lung-Cancer Screening Results

#### NOTICES

#### CONTINUING MEDICAL EDUCATION

Owned & published by the MASSACHUSETTS MEDICAL SOCIETY © 2012.  
All rights reserved. ISSN 0028-4793.



Nature Reviews | Drug Discovery

# The NEW ENGLAND JOURNAL of MEDICINE

VOL. 366 NO. 2

ESTABLISHED IN 1812

JANUARY 12, 2012

NEJM.ORG

## 108 THIS WEEK AT NEJM.ORG

### PERSPECTIVE

- 101 The Politics of Emergency Contraception A.J.J. Wood, J.M. Drazen, and M.F. Greene
- 103 The Constitutionality of the ACA's Medicaid-Expansion Mandate I.G. Cohen and J.F. Blumstein
- 105 Expanding Eligibility, Cutting Costs — A Medicaid Update J.K. Iglesias
- e4 Achieving Accountable Care — "It's Not About the Bike" J. Walker and A. McKethan

- 158 Hidradenitis Suppurativa G.B.E. Jemec

### IMAGES IN CLINICAL MEDICINE

- 165 Paragonimiasis M.A. Barrientos and A.U. Carrasco
- e5 Jugular Venous C-V Wave in Severe Tricuspid Regurgitation N.B. Senguttuvan and G. Karthikeyan

### CASE RECORDS OF THE MASSACHUSETTS GENERAL HOSPITAL

- 166 A Man with Persistent Ulcers on the Hands D. Kroshinsky, M.P. Hoang, and R.P. Hasserjian

### EDITORIALS

- 176 HER2 Therapy — An Abundance of Riches W.J. Gradishar

- 178 How Much Atrial Fibrillation Is Too Much Atrial

- 120 Subclinical Atrial Fibrillation and the Risk of Stroke J.S. Healey and Others

- 130 Cardiac Arrest during Long-Distance Running Races



IMS (FOR, FamilyPractice) (Patient: CAFFREY, JENNIE )

Action View Setup Activities Billing Reports Utilities Windows Help

Check Out - CAFFREY, JENNIE (15401)

Date: 04/23/11 Office: 0001 Pt. All

Appointments (04/23/2011)

Provider	Time	Reason
Goodman, George	08:00 A	EADE, JOHN (10731)
AARON, JOHN (9851)	09:00 A	
BAARE, John (14033)	10:00 A	
CAFFREY, JENNIE (15401)	11:00 A	
OAKS, Merr (17110)	11:30 A	
SABADOD, John (17902)	12:30 P	
IBARGUEN, Mike (7896)	12:45 P	

Verify: system Primary Doctor: HOLLOWAY, Christina Date: Race:

Referral Doctor: Case: GENERAL 02 Auth. No.: \_\_\_\_\_

Address: 456 E Groop Street Rogersville MO 65742 Phone: (W) (510) 555-4101

DOB: 10/22/1970 40 Yr 6 Mo (H) Click here to RE-VERIFY patient detail

BC/BS OF MINNESOTA LT 534534545 Primary Click here to EDIT PATIENT detail

Click here to EDIT CASE detail

Insurance ID Priority Start Date End Date Group No Copay S I

P BC/BS OF MINNESOTA LT 534534545 Primary 25 M

Send Inquiry Eligibility History P= Patient Ins. C= Case Ins. S= Insurance Card I= Insured By M= Missing V= View

Today's SuperBill: Copay (Collected \$25.00) Bal. 25.00 CR

CPT	Amount	Allow. Amt	Copay	Note
1. 93000 EKG ROUTINE	41.00	41.00	25.00	
	41.00	41.00	25.00	

Follow Up Note: Give Patient Education Handout - On the Printer Options:

Print VN 1 letter(s), 2 careplan(s)  
Fax VN  
Prescription 2 Careplan added  
Follow Up Reminder  
Copy (Collected \$25.00)  
Rx Dispense Super Bill  
Letter (Visit Note)  
Lab Test  
Print Label

Print Appointment Card: S Date Time Provider

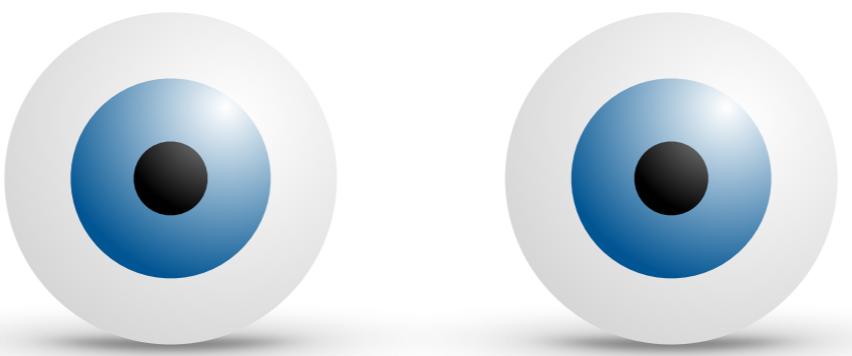
S	Date	Time	Provider
1.	05/07/2011	08:00 AM	Goodman, George

Select All Deselect All

Click here to print appointment card

Ready Desktop 4/23/11 21:40:13

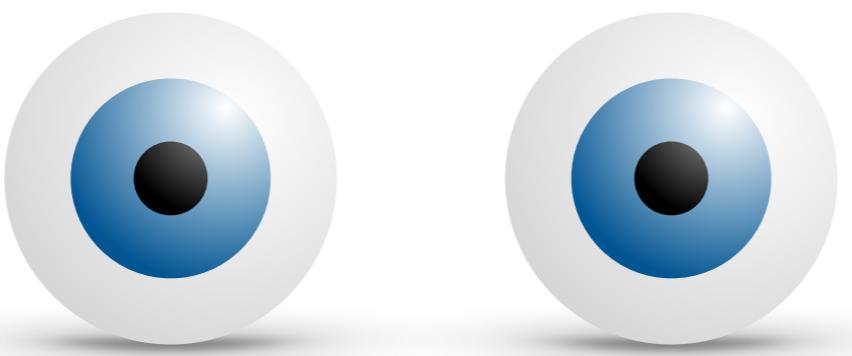




Generalizable

&

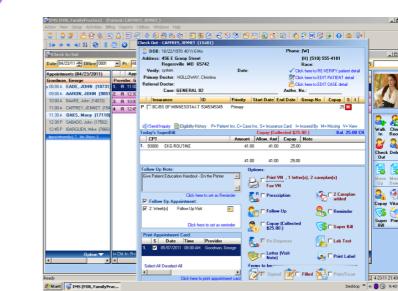
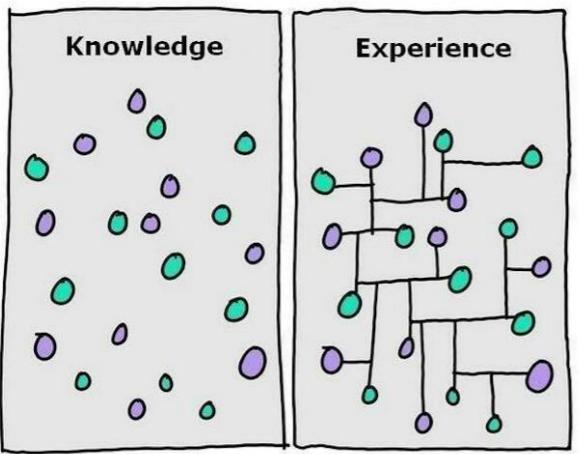
Scalable





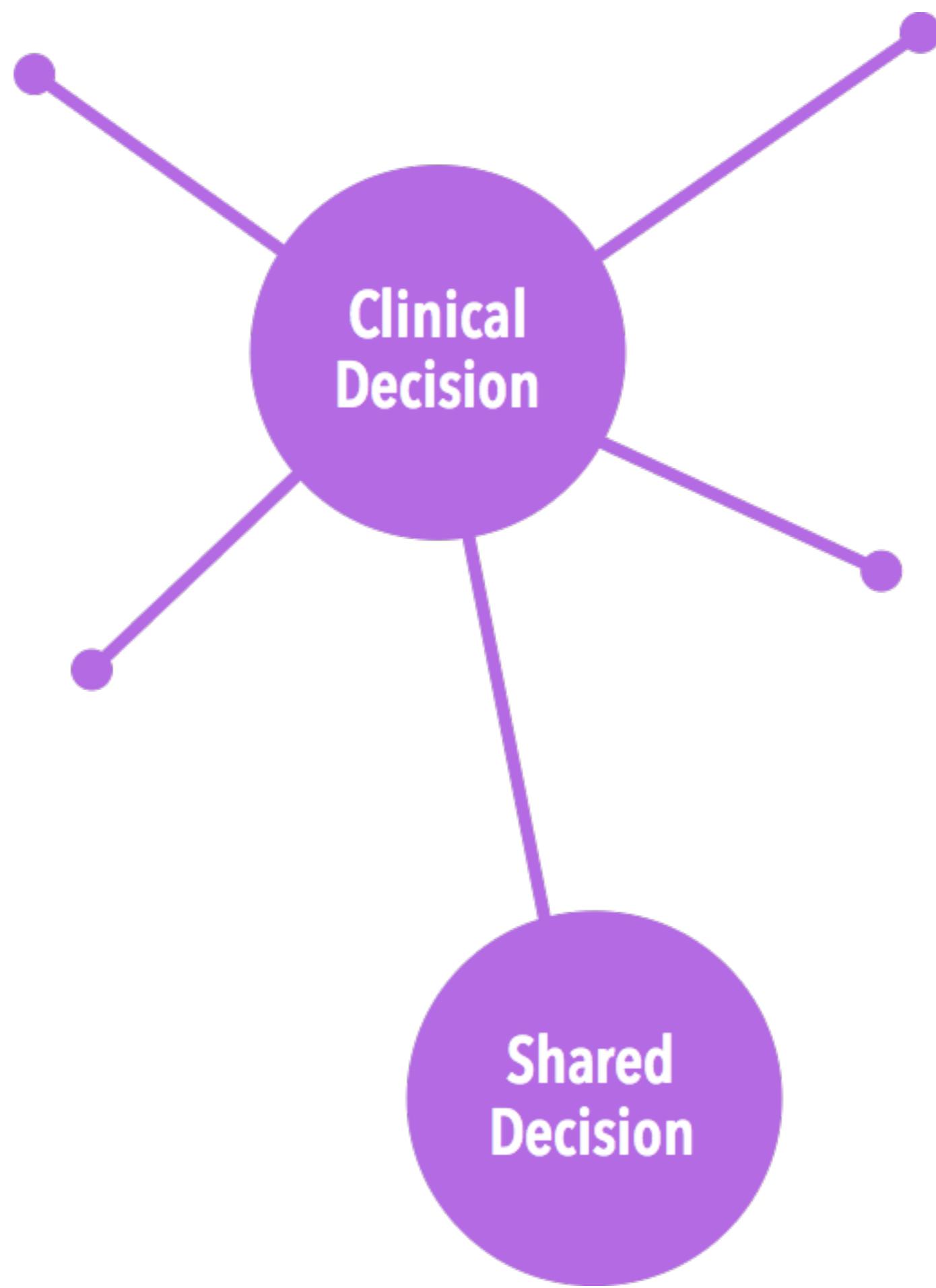


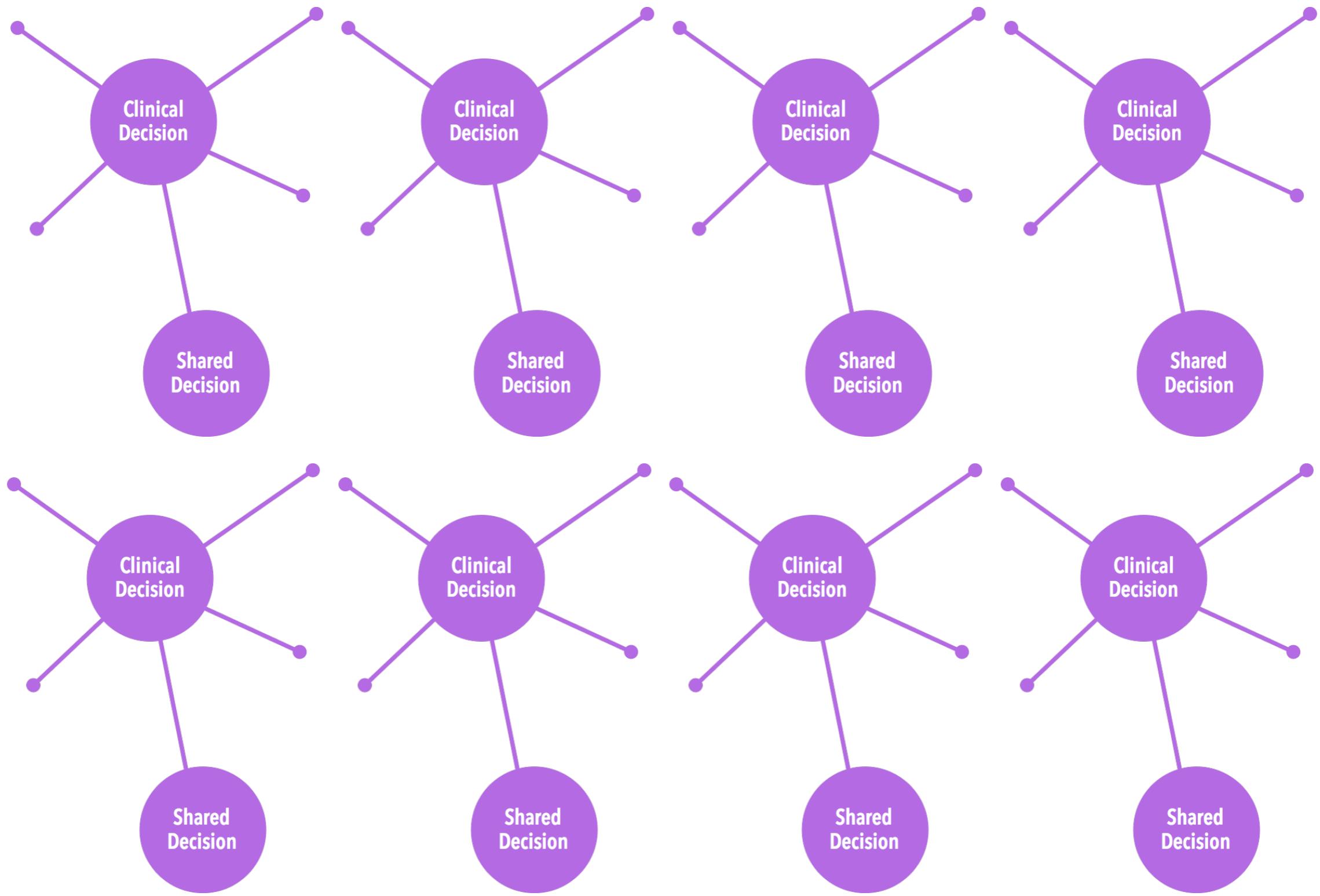
# Clinical Decision



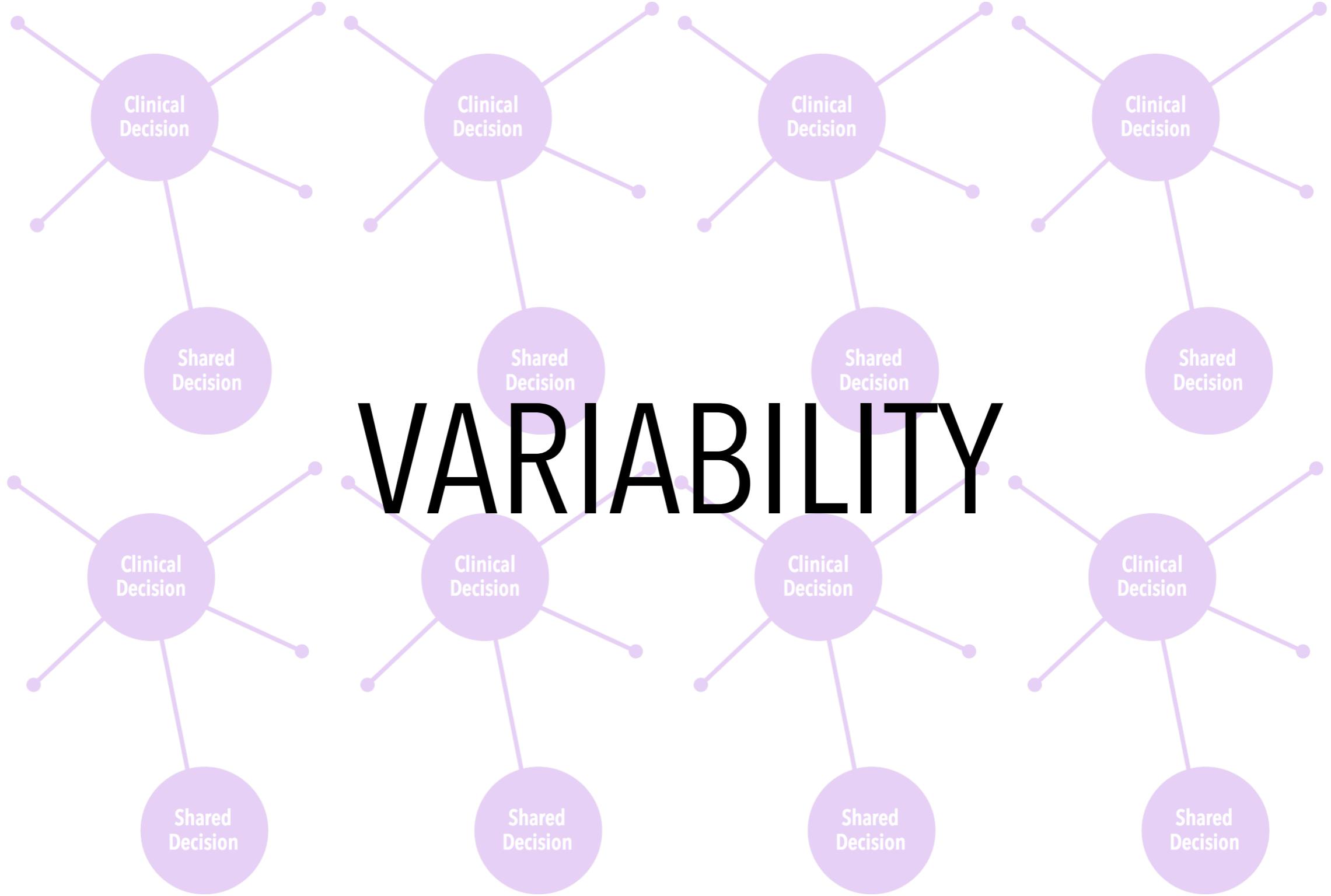


## Shared Decision



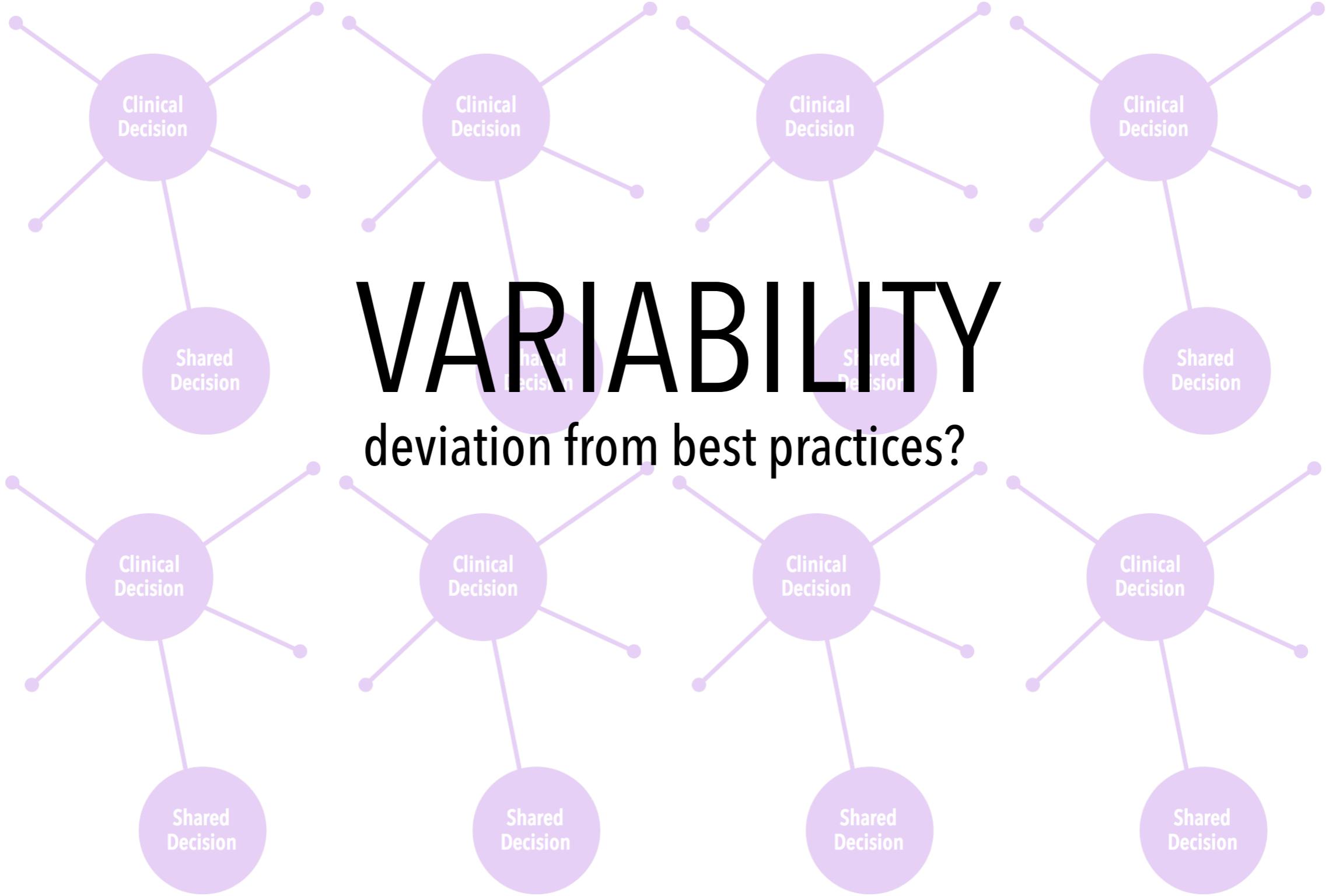


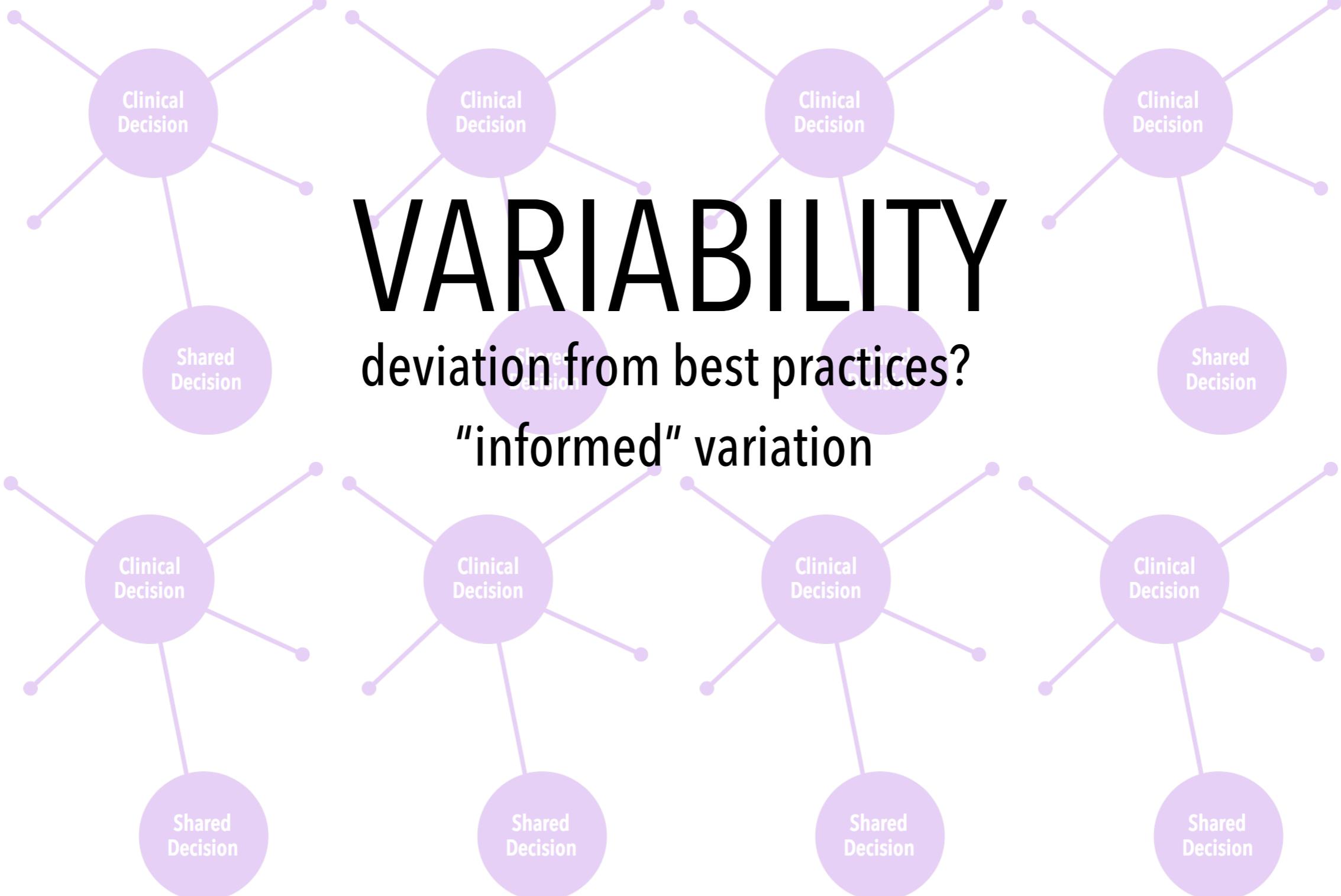
# VARIABILITY



# VARIABILITY

deviation from best practices?

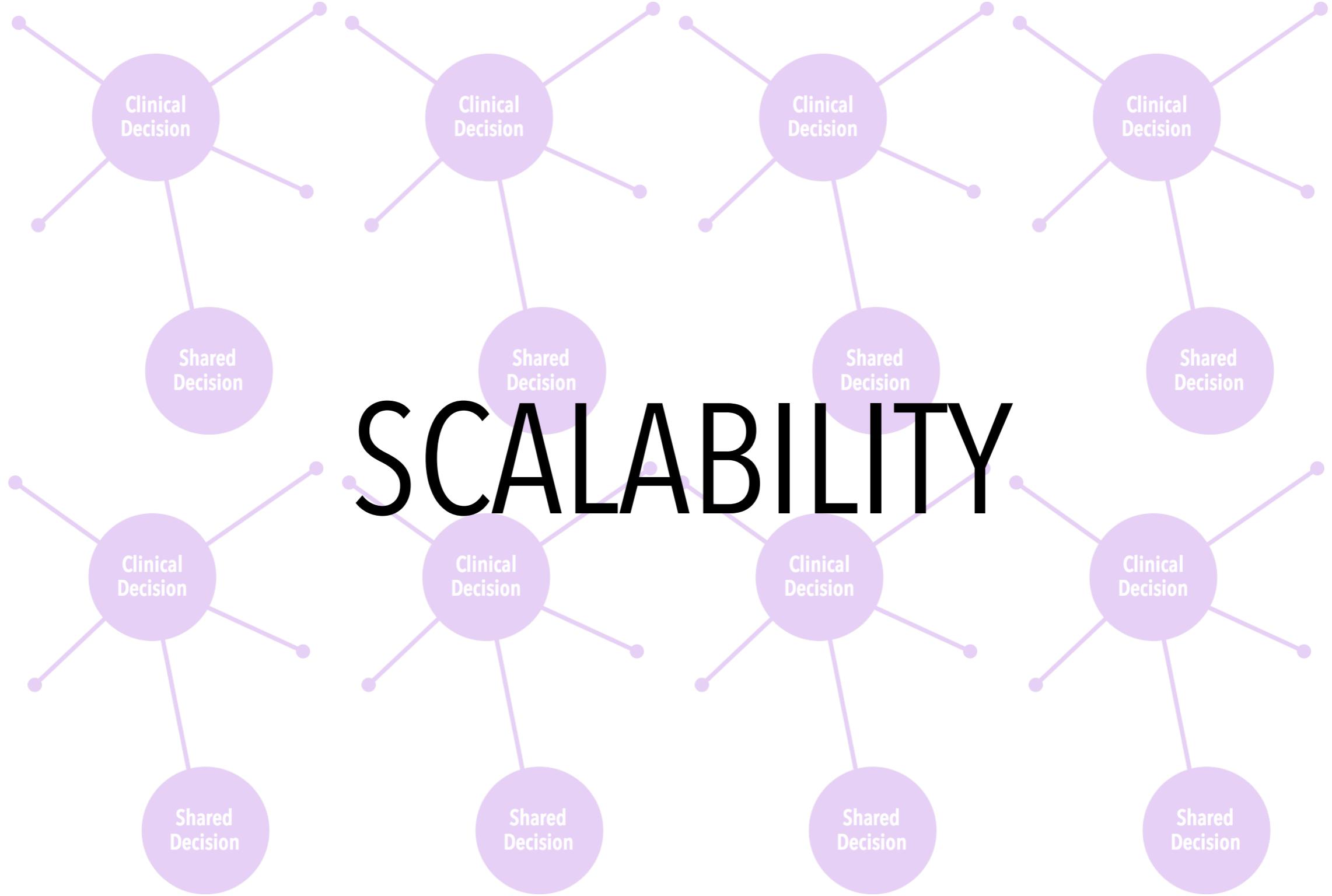




# VARIABILITY

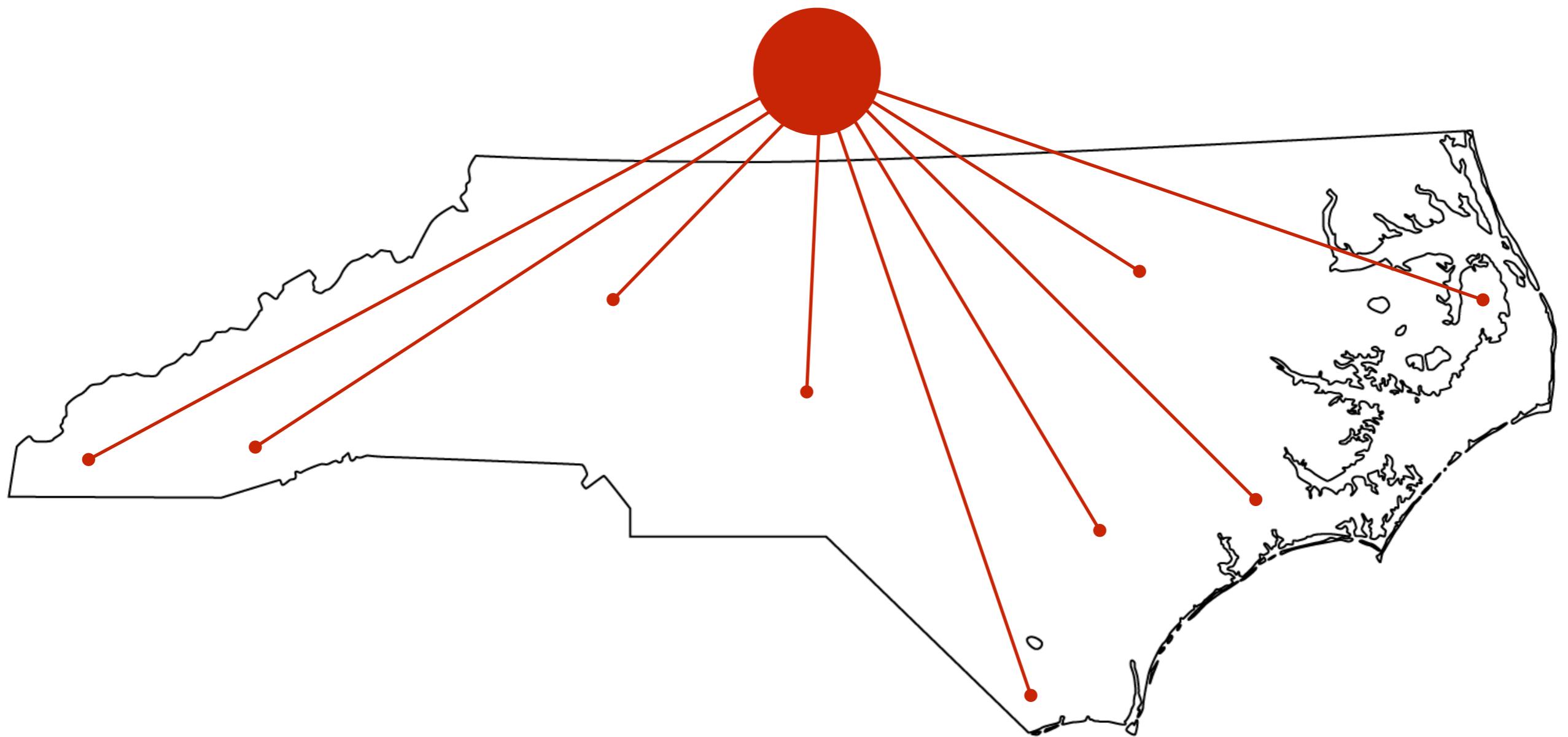
deviation from best practices?  
“informed” variation

# SCALABILITY





# Duke Cancer Institute



# Vinod Khosla says technology will replace 80 percent of doctors – sparks indignation

MATT MARSHALL   SEPTEMBER 2, 2012 9:12 PM



[BRIEFING ROOM](#)[ISSUES](#)[THE ADMINISTRATION](#)[PARTICIPATE](#)[1600 PENN](#)

Search



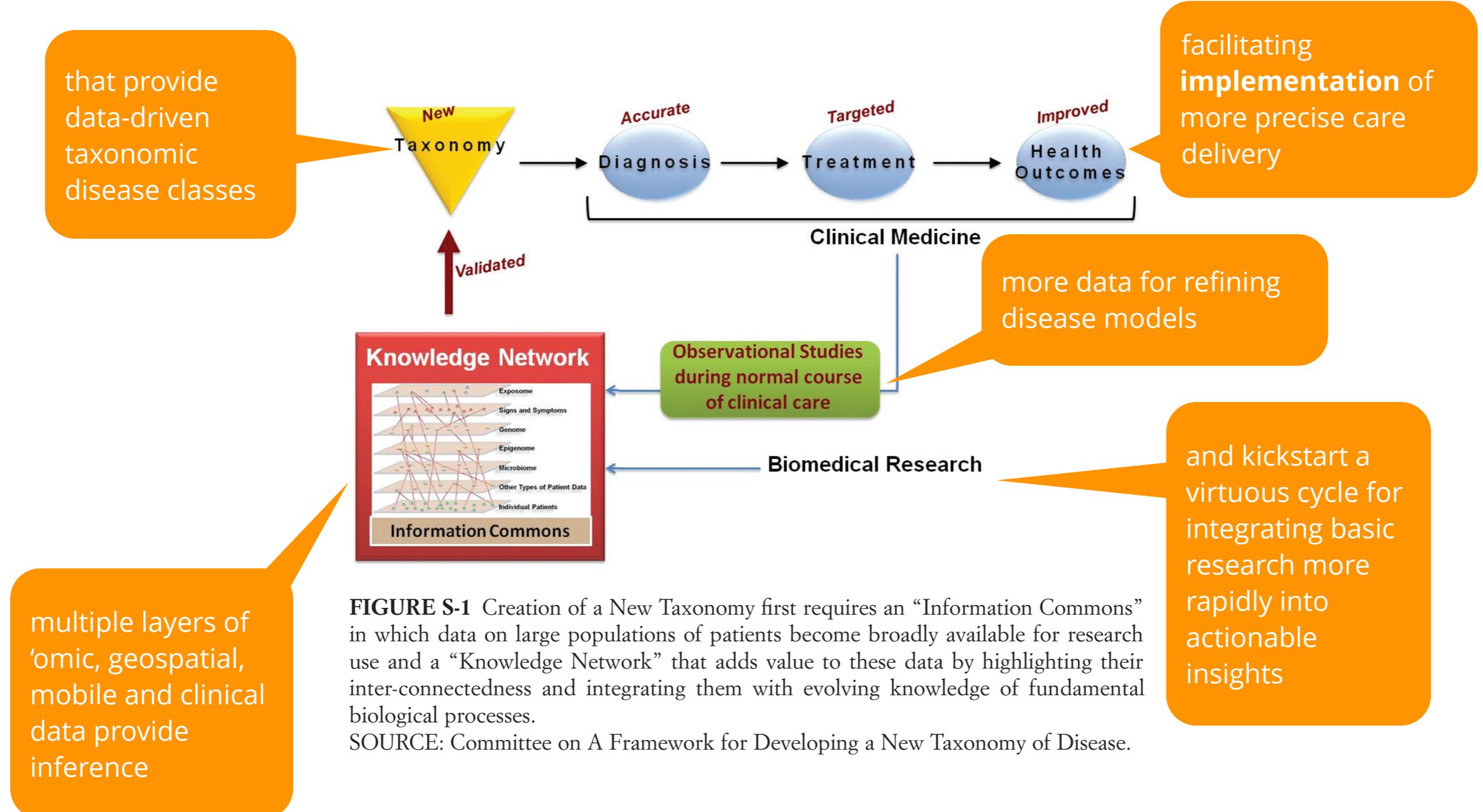
# THE PRECISION MEDICINE INITIATIVE

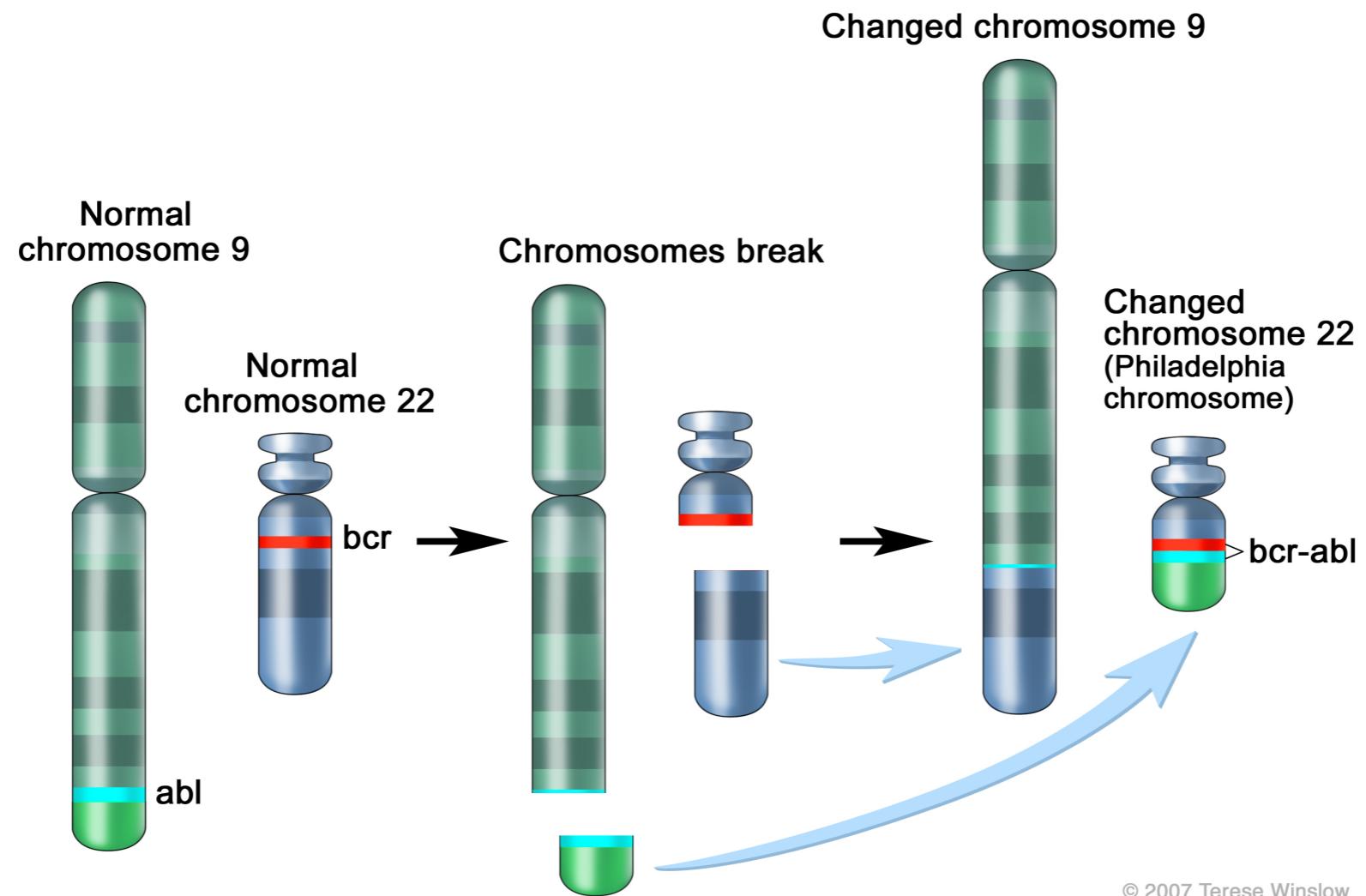
[PRECISION MEDICINE](#)[THE INITIATIVE](#)[PRINCIPLES](#)[STORIES](#)

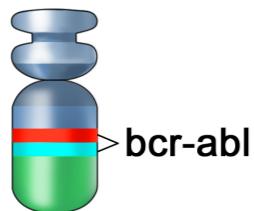
GO TO TOP

*"Doctors have always recognized that every patient is unique, and doctors have always tried to tailor their treatments as*

# The National Academy's Report on **Precision Medicine**



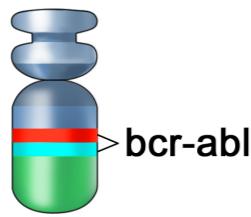




seen in ~95% of patients with  
**chronic myeloid leukemia** (CML)

encodes a novel fusion protein that  
is "always on" and induces  
dysregulated cell proliferation

in 2000, the death rate for CML in the first  
year was 25 to 50 percent

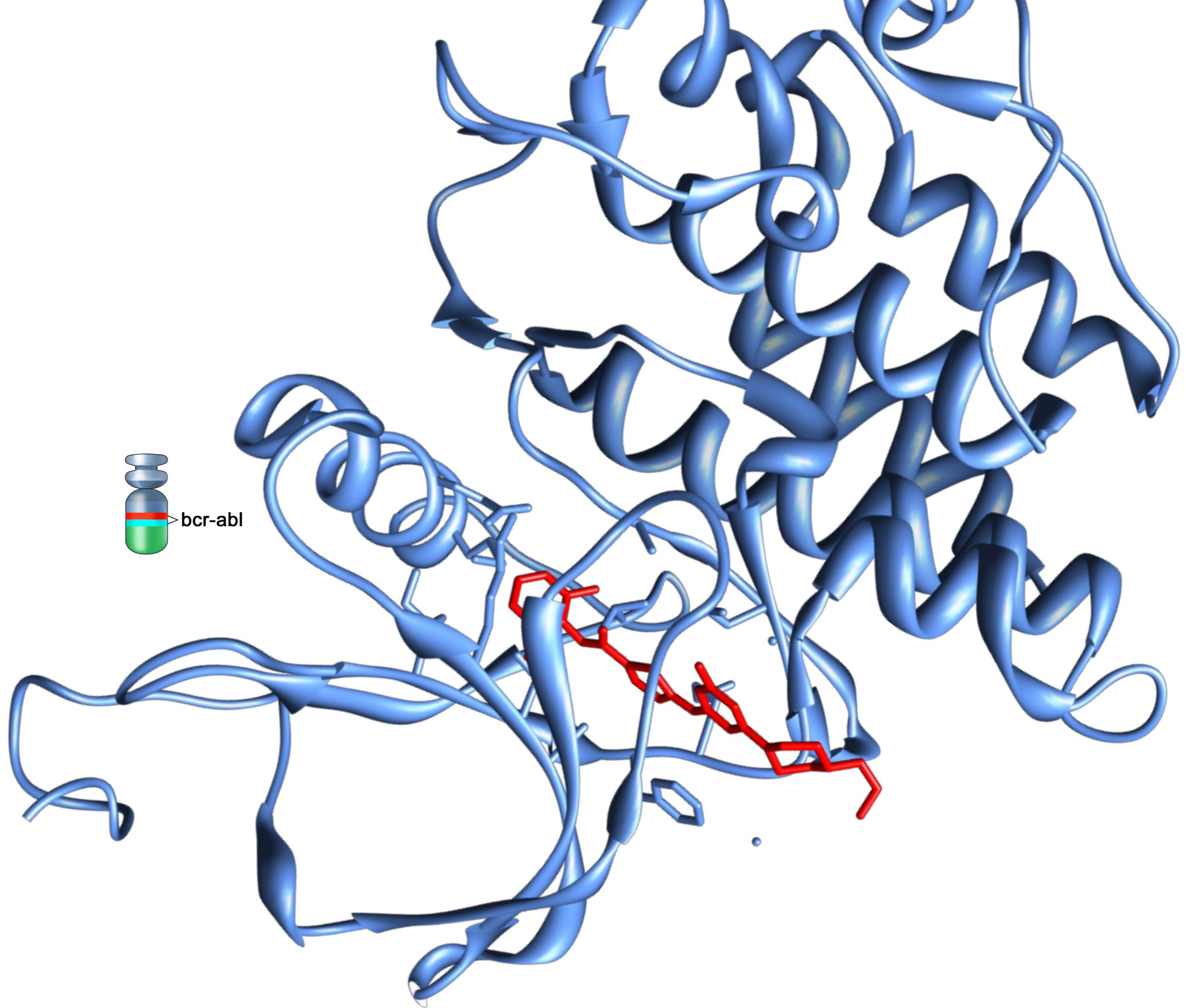


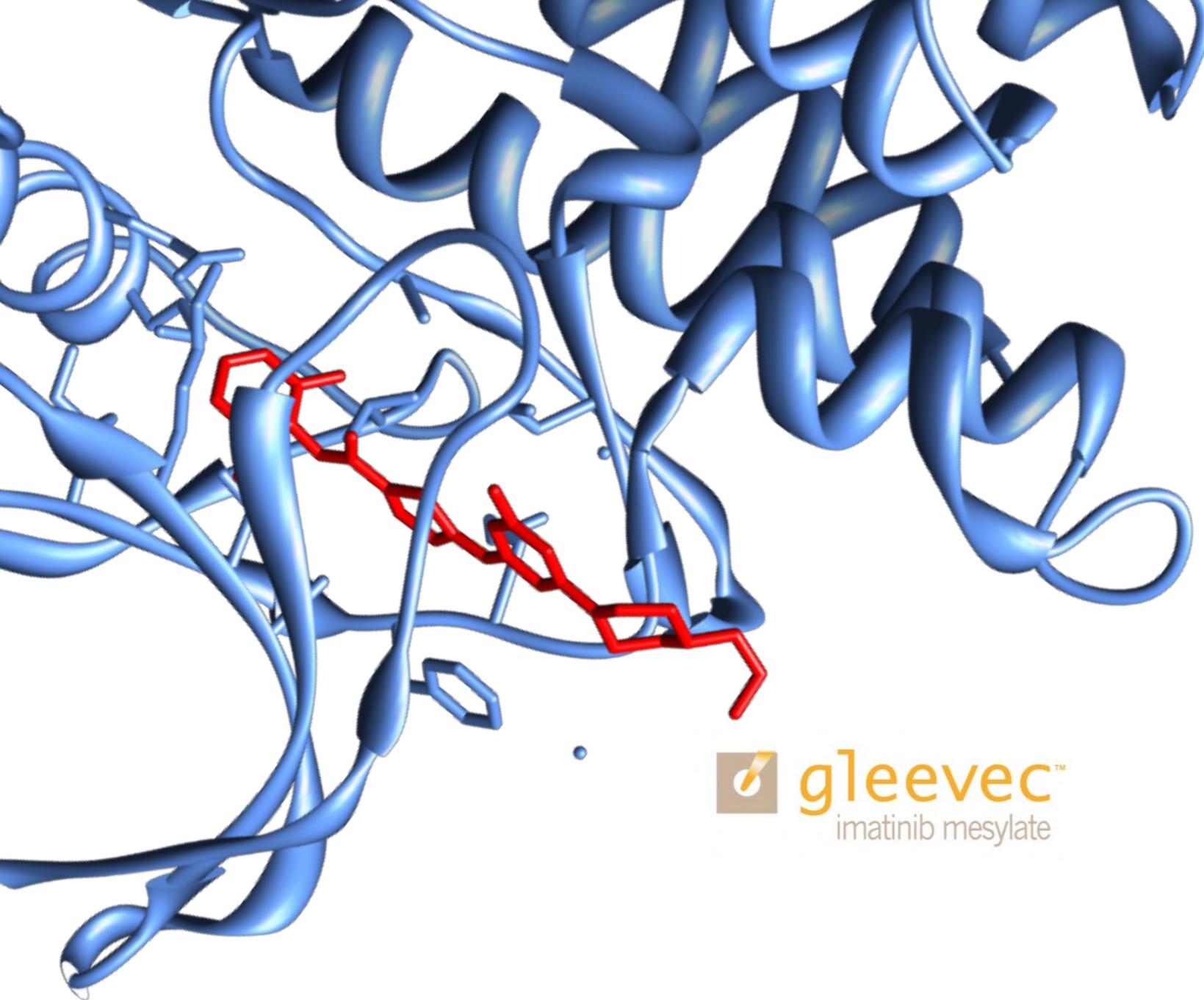
“

The way I'd been trained, cancer was seen as something like a light switch that was stuck in an "on" position. You were given a baseball bat, which was chemotherapy, and told to knock the light out with the bat. I thought, "Why don't we just try to figure out *why* the light is stuck on, then we can fix it without breaking everything."

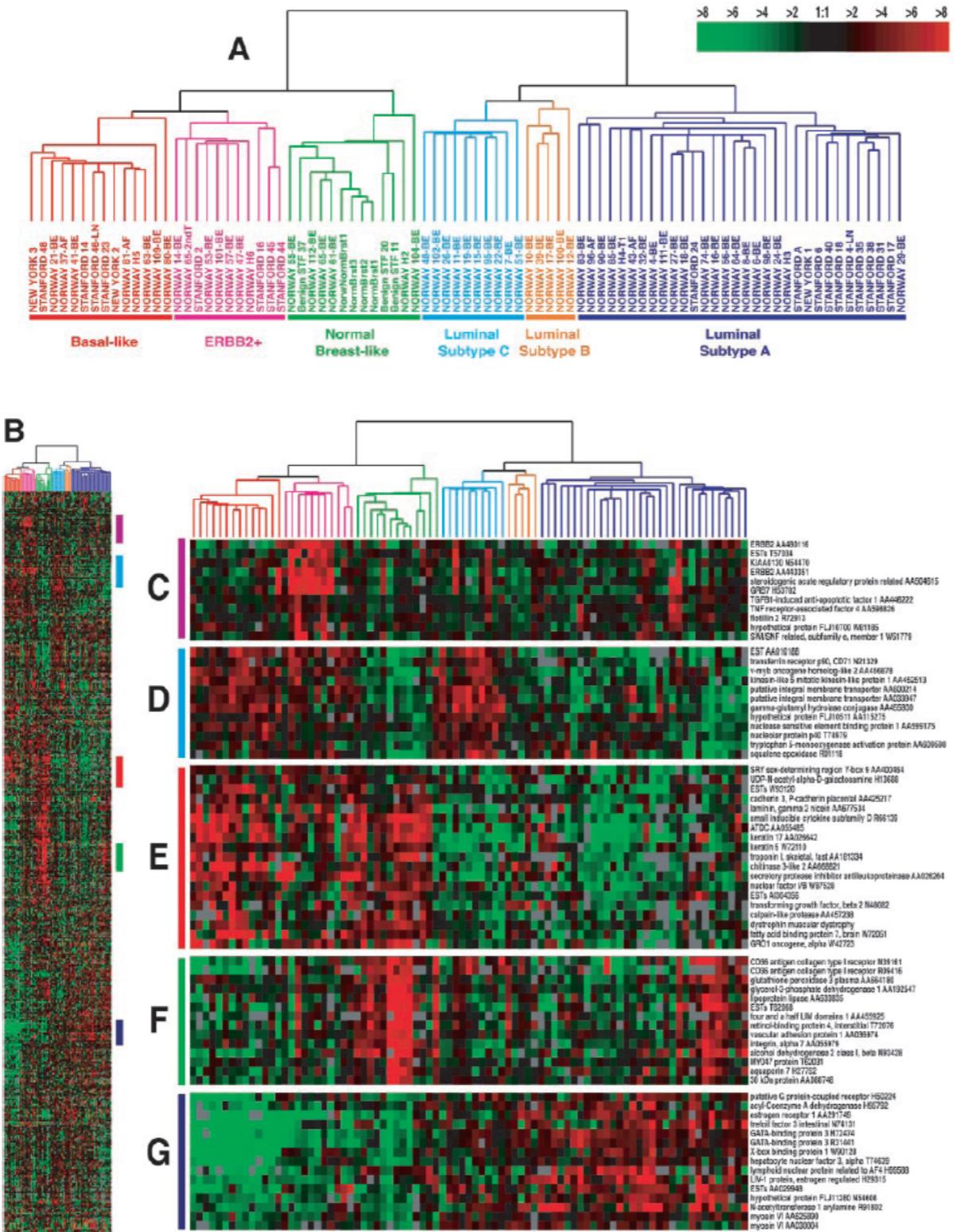
”

– Brian Druker



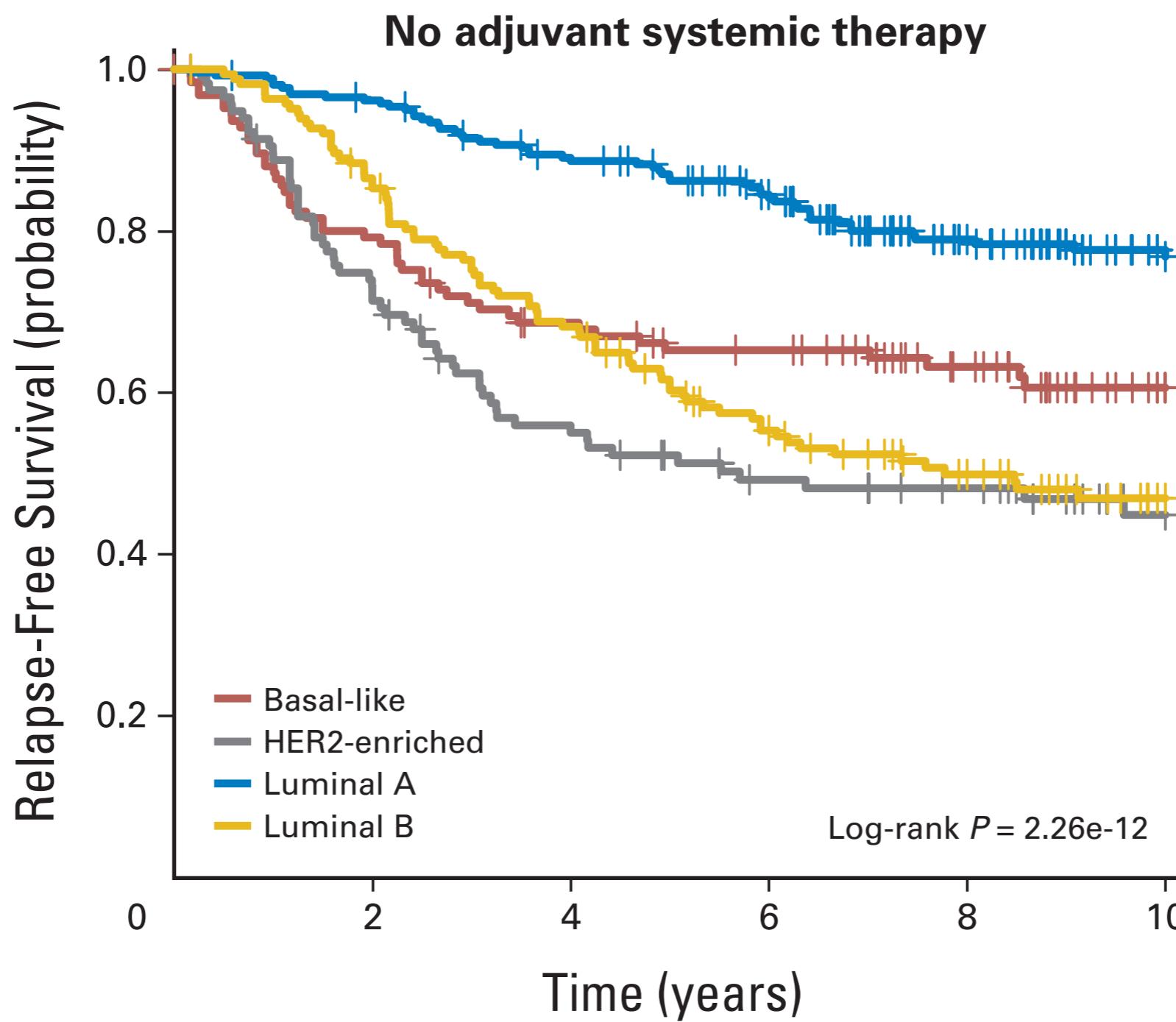


> 95% of patients with stable cytogenetic response are alive at 8 years



# breast cancer class discovery

*PNAS* 2001  
hierarchical clustering in 78  
Norwegian patients (not  
model-based class  
discovery)

**A**

A prediction needs...

# Discussion

## Discussion Questions

1. As Harlan Krumholz describes, over the last decade other industries have been taking advantage of 'big data' and 'machine learning'. Why hasn't this happened in healthcare? What are the main barriers? And how do we surmount them?
2. Obermeyer and Emanuel's *NEJM* editorial paints a sanguine picture of what machine learning will bring to clinical medicine. What fallacies do they point out? Importantly, what do they fail to point out? They point to three areas of disruption; what potential issues are they glossing over?



# The NEW ENGLAND JOURNAL of MEDICINE



HOME ARTICLES & MULTIMEDIA ISSUES SPECIALTIES & TOPICS FOR AUTHORS CME >



## Perspective

### Uncertainty in the Era of Precision Medicine

David J. Hunter, M.B., B.S., Sc.D.

N Engl J Med 2016; 375:711-713 | August 25, 2016 | DOI: 10.1056/NEJMmp1608282

Share: [f](#) [t](#) [g+](#) [in](#) [+](#)

Article References Metrics

A National Research Council report on “precision medicine” explains that the term “refers to the tailoring of medical treatment to the individual characteristics of each patient.” The report goes on to say, “It should be emphasized that in ‘precision medicine’ the word ‘precision’ is being used in a colloquial sense, to mean both ‘accurate’ and ‘precise.’”<sup>1</sup> In the colloquial sense, “precision” also implies a high degree of certainty of an outcome, as in “precision-guided missile” or “at what precise time will you arrive?” So will precision medicine usher in an age of diagnostic and prognostic certainty?

In fact, the opposite will probably result. The new tools for tailoring treatment will demand a greater tolerance of uncertainty and greater facility for calculating and interpreting probabilities than we have been used to as physicians and patients.

Oncology has been called “the clear choice for enhancing the near-term impact of precision medicine.”<sup>2</sup> New tools extract information from cancer genomes that include both the mutations that occur somatically (cancer genome sequencing) and the functional changes that result from both

[Article](#)[References](#)[Metrics](#)

A National Research Council report on “precision medicine” explains that the term “refers to the tailoring of medical treatment to the individual characteristics of each patient.” The report goes on to say, “It should be emphasized that in ‘precision medicine’ the word ‘precision’ is being used in a colloquial sense, to mean both ‘accurate’ and ‘precise.’”<sup>1</sup> In the colloquial sense, “precision” also implies a high degree of certainty of an outcome, as in “precision-guided missile” or “at what precise time will you arrive?” So will precision medicine usher in an age of diagnostic and prognostic certainty?

In fact, the opposite will probably result. The new tools for tailoring treatment will demand a greater tolerance of uncertainty and greater facility for calculating and interpreting probabilities than we have been used to as physicians and patients.

Oncology has been called “the clear choice for enhancing the near-term impact of precision medicine.”<sup>2</sup> New tools extract information from cancer genomes that include both the mutations that occur somatically (cancer genome sequencing) and the functional changes that result from both these mutations and epigenetic events (gene-expression alterations in tumors). For instance, by

## Discussion Questions

1. Lazer *et. al.* talks about "big data hubris". When is data being "bigger" not necessarily better? What about the issue of transparency and reproducibility?
2. Darcy *et. al.* raises the issue of privacy. Are privacy and machine learning reconcilable?

# Microsoft Azure Lab