

Biomarker Discovery in Precision Medicine

Patipark Kueanjinda, Ph.D.

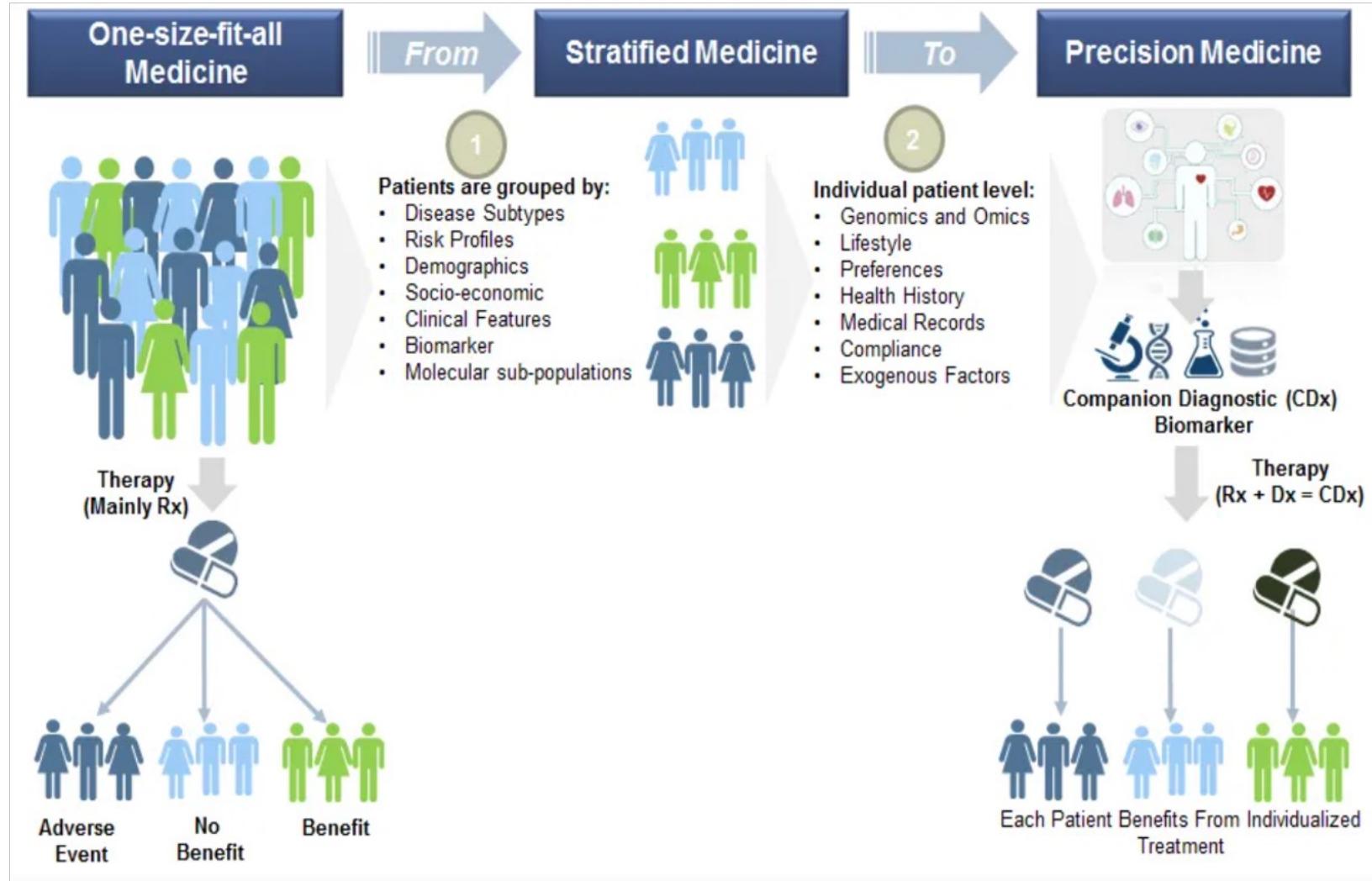
SIPM610 Systems Pharmacology Lab I

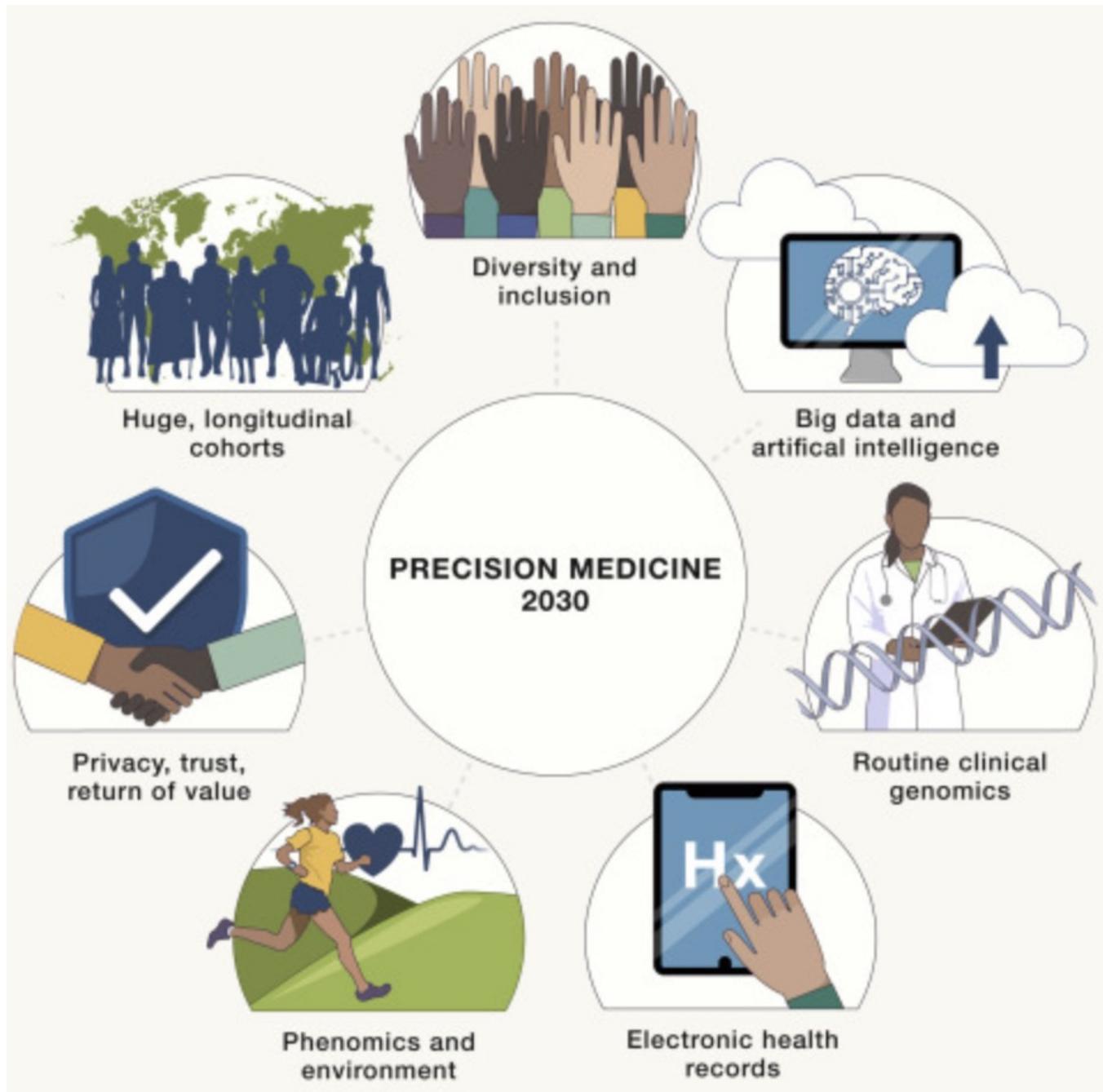
Outlines

1. Precision medicine
2. Introduction to biomarkers
3. Common biomarkers in cancer
4. Uses of biomarkers in cancer
5. Biomarker development
6. Ethical, legal, and social issues with biomarkers

1. Precision Medicine

Precision medicine delivers right treatment at the right time

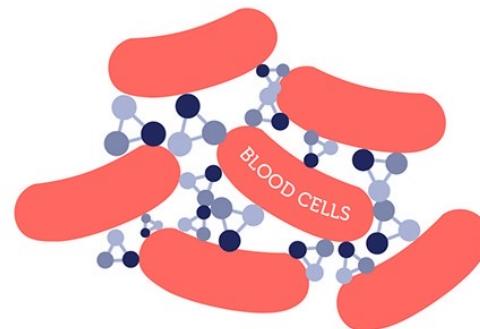




How warfarin reduces blood clots



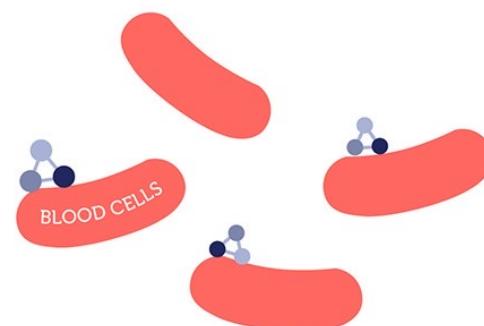
Vitamin K, produced by the body, helps form *blood-clotting proteins*



Blood-clotting proteins hold blood cells together to form clots



Warfarin reduces the body's ability to make Vitamin K which interferes with protein creation

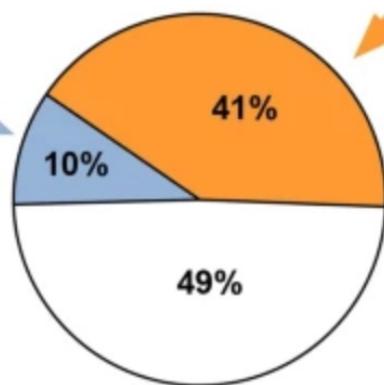


Lower levels of clotting protein makes blood cells less likely to clot

Factors contributing to warfarin response

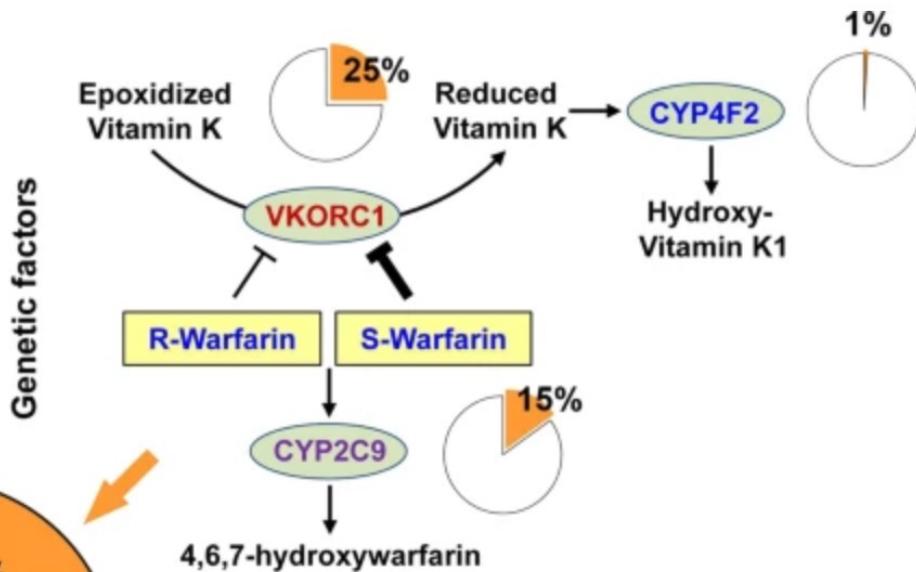
Clinical and environment factors

Age
Height
Weight
Race
Gender
Interaction medicine
Diet
Smoking
Alcohol



New factors

New common and rare genetic variants
Pharmacomicrobiomic factors
Pharmacoepigenetic factors
⋮



Warfarin therapy optimization

VKORC1	CYP2C9					
	*1/*1	*1/*2	*1/*3	*2/*2	*2/*3	*3/*3
G/G	Normal Responder (5-7 mg/d)	Normal Responder (5-7 mg/d)	Sensitive Responder (3-4 mg/d)	Sensitive Responder (3-4mg/d)	Sensitive Responder (3-4mg/d)	Highly Sensitive Responder (0.5-2 mg/d)
G/A	Normal Responder (5-7 mg/d)	Sensitive Responder (3-4 mg/d)	Sensitive Responder (3-4mg/d)	Sensitive Responder (3-4mg/d)	Highly Sensitive Responder (0.5-2 mg/d)	Highly Sensitive Responder (0.5-2 mg/d)
A/A	Sensitive Responder (3-4mg/d)	Sensitive Responder (3-4mg/d)	Highly Sensitive Responder (0.5-2 mg/d)			

2. Introduction to Biomarkers

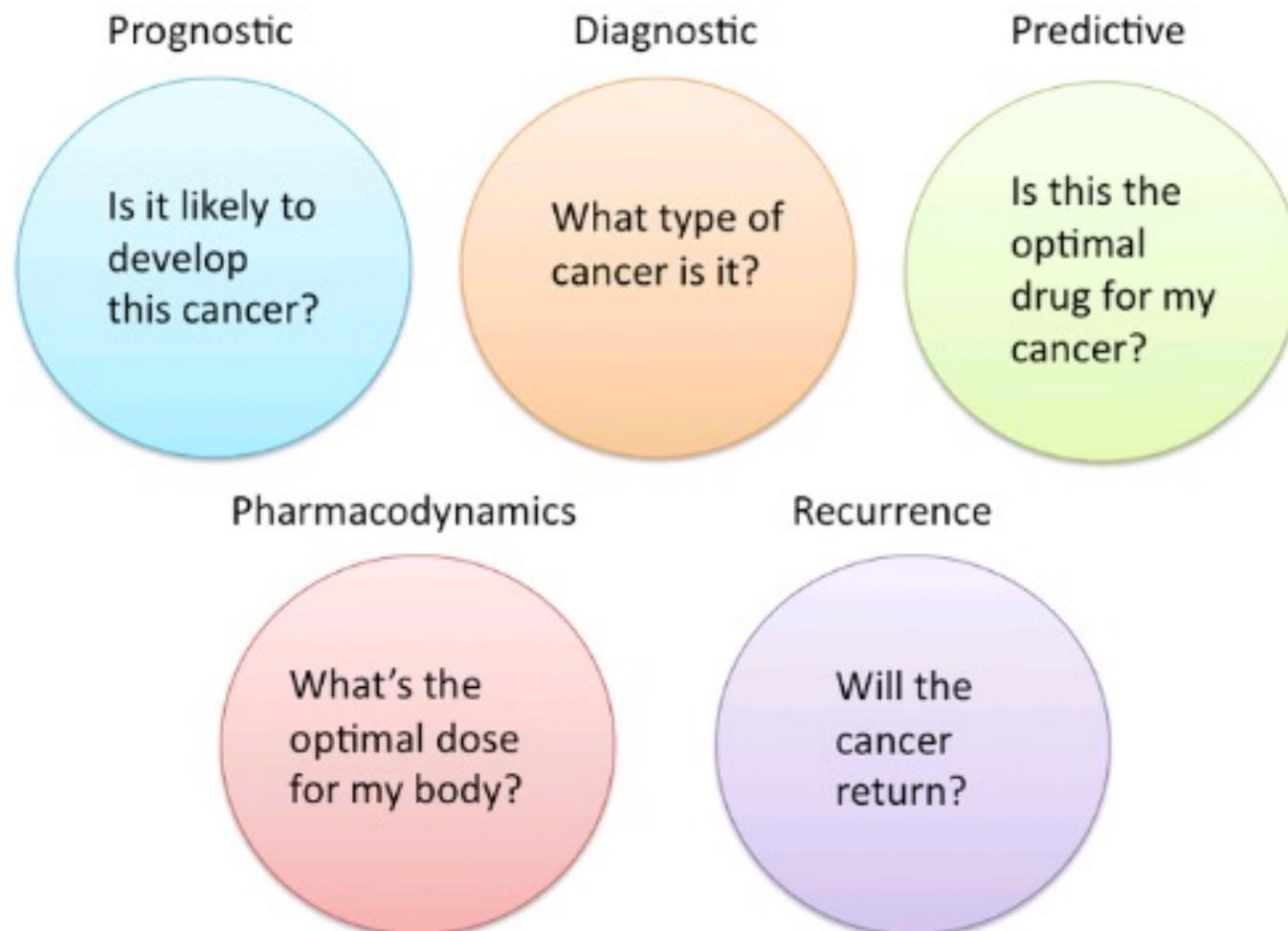
Biomarker definitions

Source	Definition
National Cancer Institute (NCI)	A biological molecule in blood or other body fluids, or tissues that is a sign of a normal or abnormal process or of a condition of disease.
Center for Biomarkers in Imaging (MGH)	Anatomic, physiologic, biochemical, or molecular parameters associated with the presence and severity of specific disease states
Biomarkers Consortium (NIH)	Characteristics that are objectively measured and evaluated as indicators of normal biological processes, pathogenic processes, or pharmacological responses to therapeutic intervention

What are biomarkers?

- Biological/biochemical molecules
- Anatomical and physiological parameters
 - Blood pressure, cholesterol level
- Mostly proteins identified and measurable in blood or urine
- Surrogate measures of the biology of the cancer
 - Provide insight into the clinical behavior of the disease
 - Change overtime as new technologies have been developed

Questions answered by cancer biomarkers



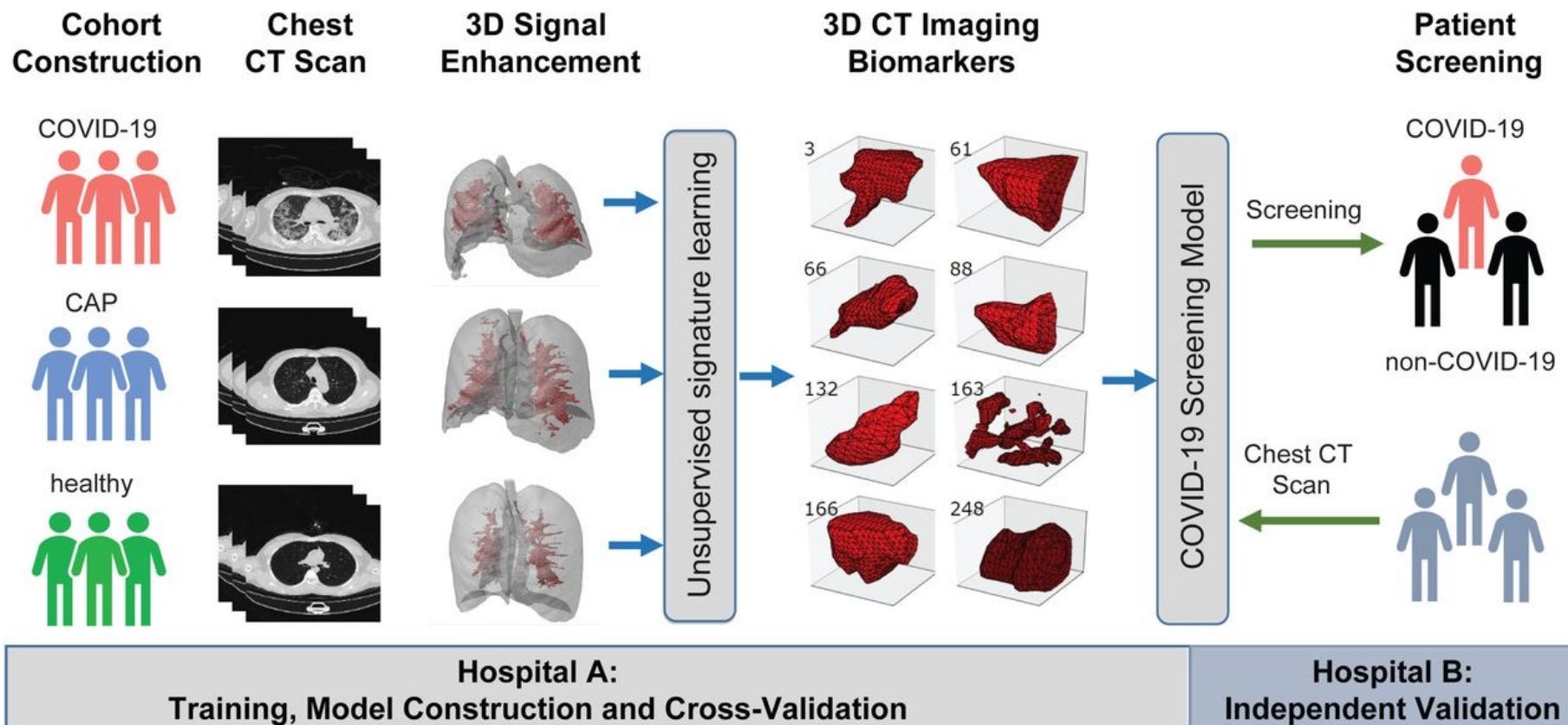
Types of biomarkers

- Type 0: Natural history marker
 - Correlate with known clinical symptoms over the course of the disease
 - e.g., serum c-reactive protein (CRP) => inflammation
- Type 1: Biological activity marker
 - Capture the effects of an intervention or treatment
 - e.g., serum HbA1c => blood sugar level
- Type 2: Single or multiple markers of therapeutic efficacy
 - Surrogate endpoints that predict clinical benefit
 - e.g., serum LDL-c => blood cholesterol

Safety biomarkers

- Common lab biomarkers for monitoring vital organ function
 - Liver safety: AST, ALT, ALP, bilirubin
 - Renal safety: BUN, creatine, GFR
 - Hematology safety: complete blood count
 - Bone safety: calcium, inorganic phosphate
 - Basic metabolic safety: glucose, cholesterol, uric acid

Imaging as biomarkers

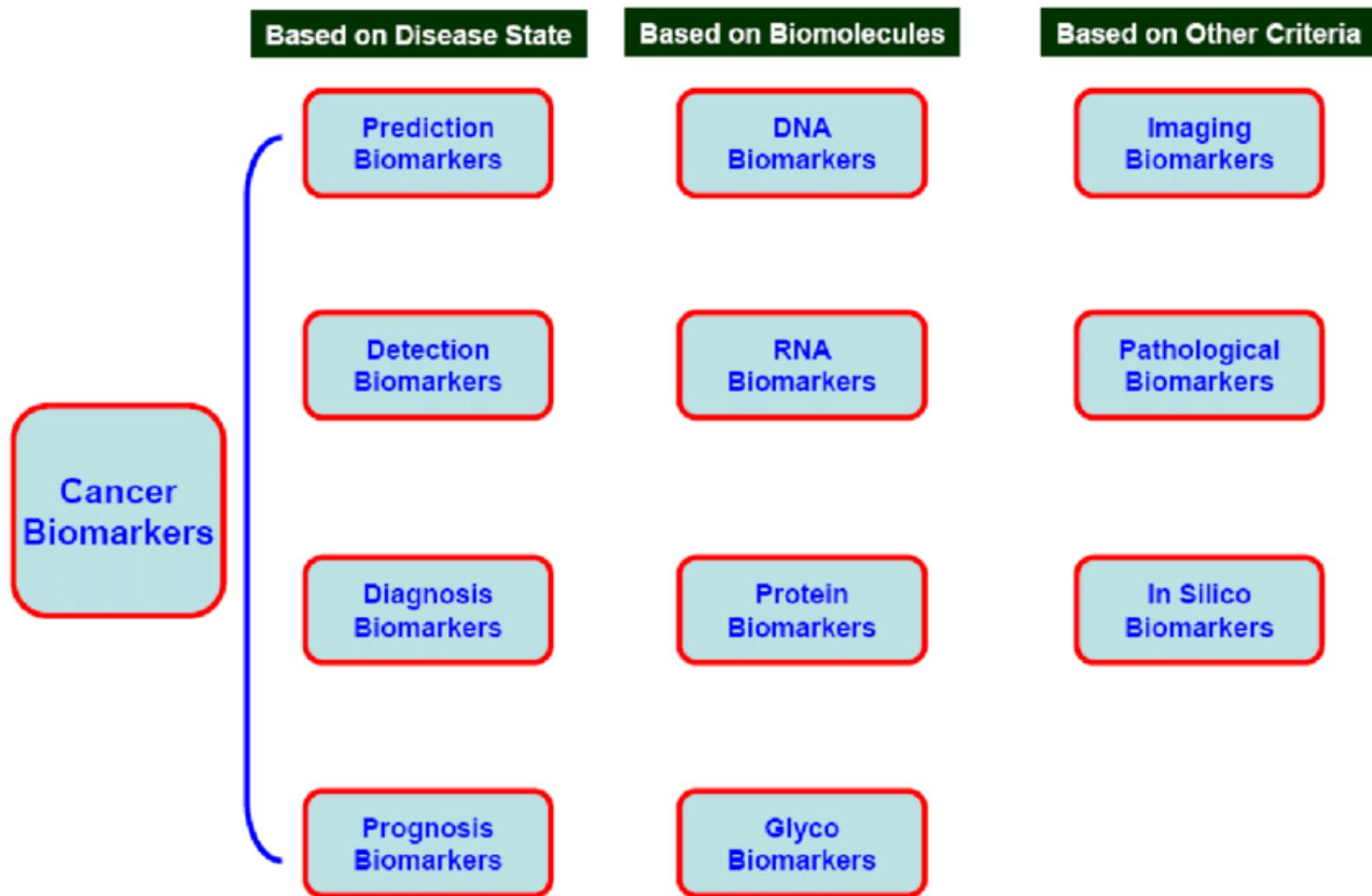


Surrogate endpoints

- Lab or physical measurement used in clinical trials to indicate drug's response
- Tell the benefit or harm from therapy

Disease	Surrogate endpoints	Clinical endpoints
Hypertension	Blood pressure	Stroke
Dyslipidemia	Cholesterol, LDL	Coronary artery disease
Diabetes	Glycosylated hemoglobin (HbA1c)	Retinopathy, nephropathy, neuropathy, heart disease
Glaucoma	Intraocular pressure	Loss of vision
Cancer	Biomarkers, tumor shrinkage, response rate	Progression-free survival, overall survival

Types of cancer biomarkers



Are biomarkers perfect predictors or prognosticators?

- No
- Accuracy of biomarkers varies greatly depending on several factors:
 - Specificity for the disease
 - Measurement accuracy

False positive and false negative

	Test says you don't have it	Test says you do have it
You really don't have it	True Negative (TN)	False Positive (FP)
You really do have it	False Negative (FN)	True Positive (TP)

Consequences for patients

False negative – misdiagnosis, delayed treatment

False positive – psychological, economic impacts

Sensitivity and specificity

$$\text{Sensitivity} = \text{TP} / (\text{TP} + \text{FN})$$

Note: ideal = 95-98%

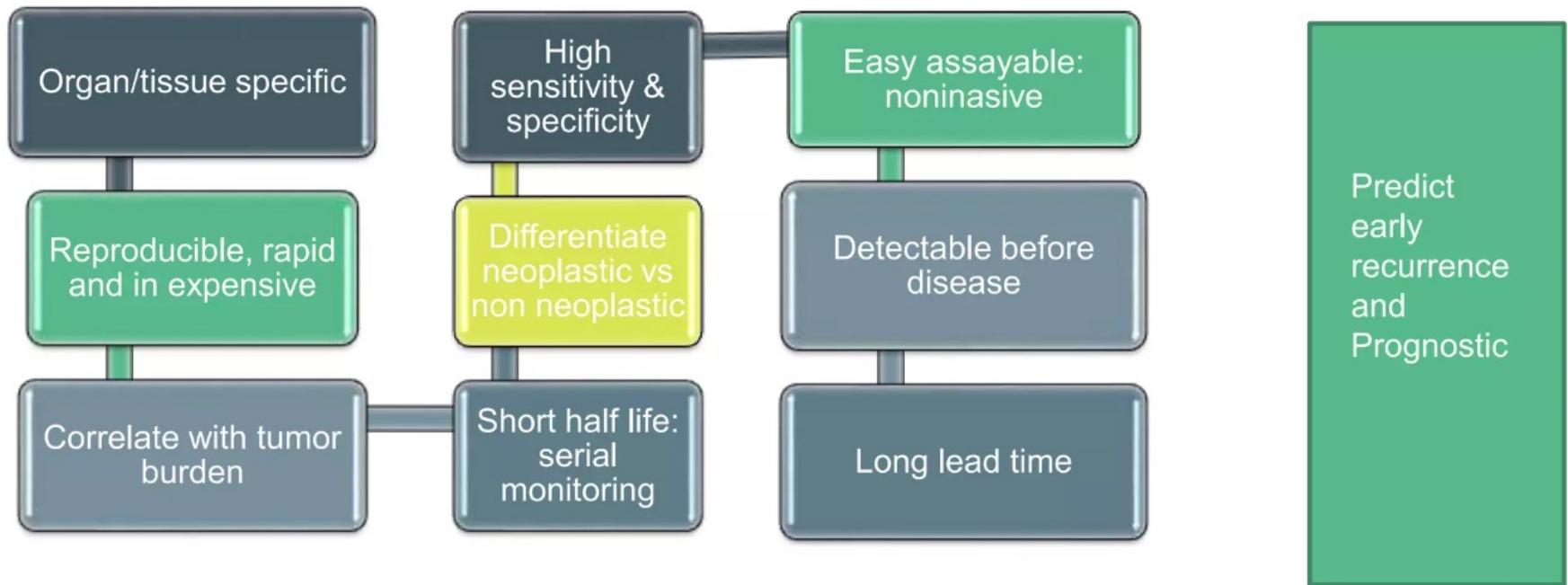
$$\text{Specificity} = \text{TN} / (\text{TN} + \text{FP})$$

Note: ideal = 90-95%

- Poorly sensitive tests will result in many false negatives
- Poorly specific tests will result in many false positives
- Tissue/tumor specific markers are rare

Ideal biomarkers

- 100% of the people who have the disease
- 0% of the people who do not have the disease



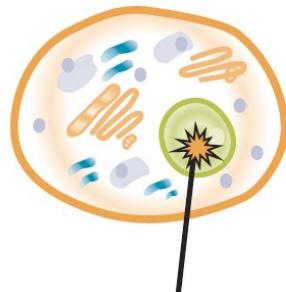
3. Common Biomarkers in Cancer

Multi-hit theory of cancer

- Accumulation of multiple gene alterations in genes lead to abnormal cell growth
- Cancerous cells compete with normal cells in the tissues and may escape into the blood stream

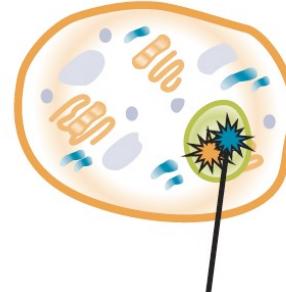
Multiple Gene Alterations in Critical Cell Pathways Can Lead to Cancer

Normal Cell



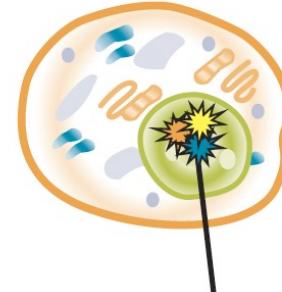
May contain single mutation – not enough to cause cancer in humans

Normal Cell



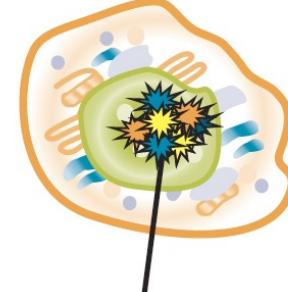
Cell may acquire another mutation over time

Cell beginning to transform



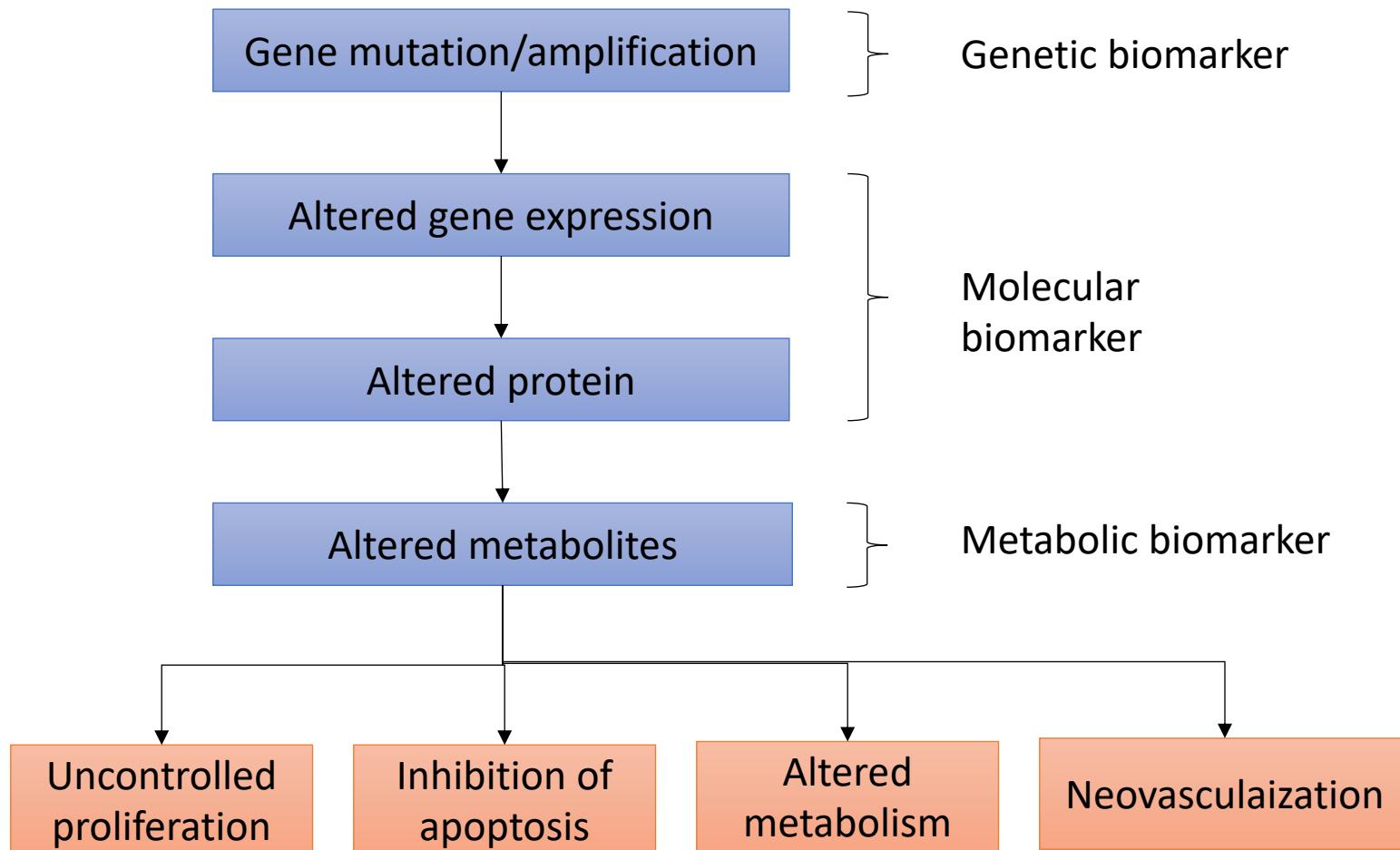
Multiple mutations accumulate over time

Cancerous Cell



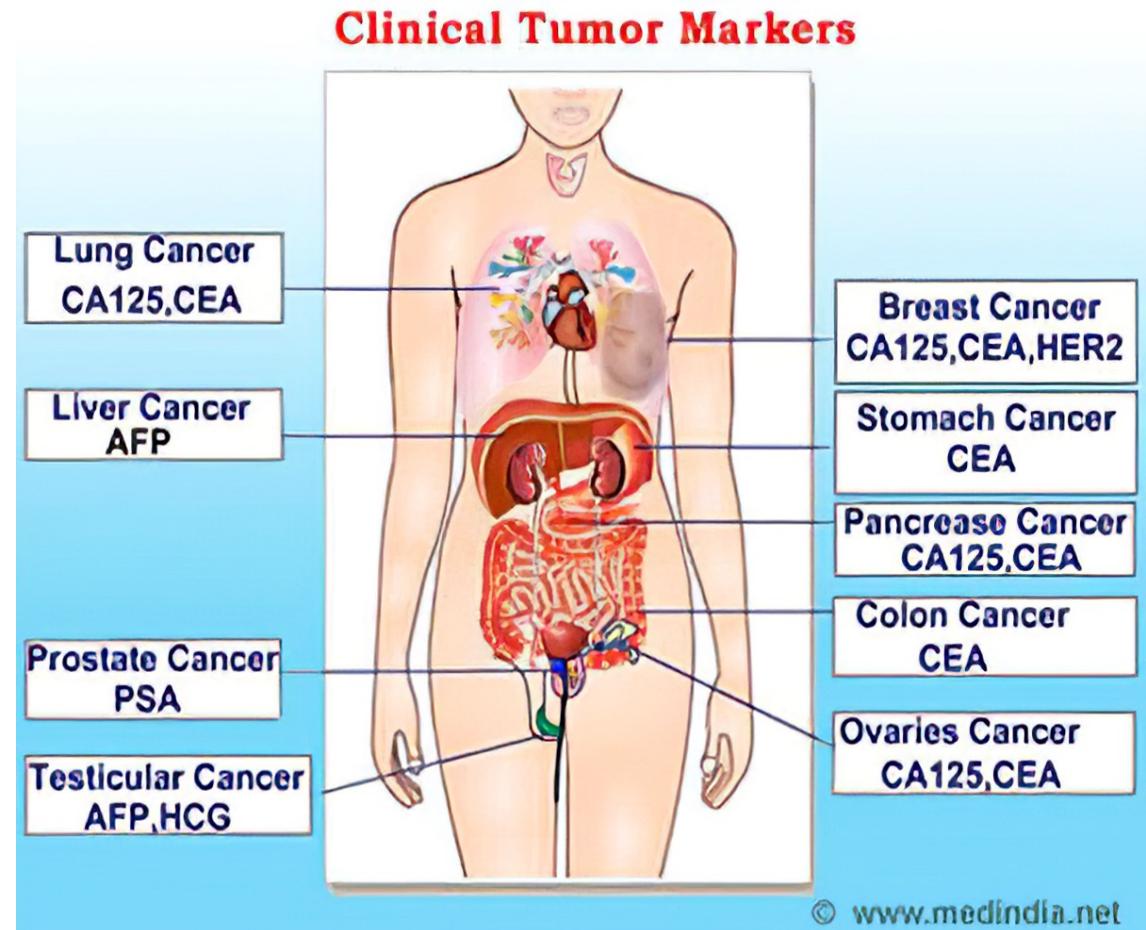
Too many mutations – cell no longer responsive to normal regulatory signals

Mutations give rise to cancerous cells



Classical cancer biomarkers

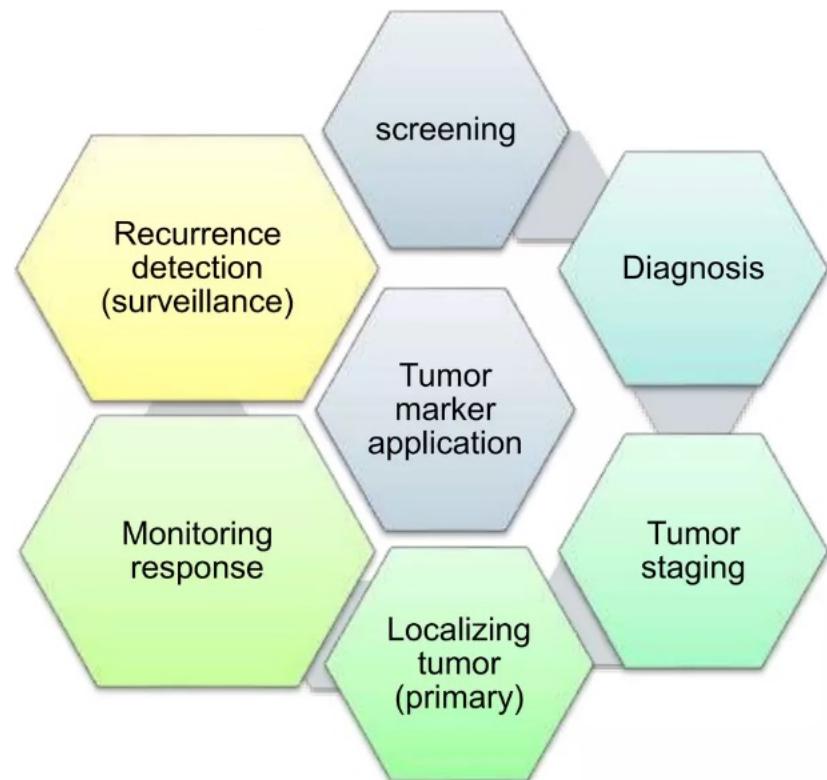
- Measurable biological indicators
- Frequently found in bodily fluids or tissues
- Products of cellular, biochemical, molecular, or genetic alterations
- Correlates with a normal or pathogenic physiological state



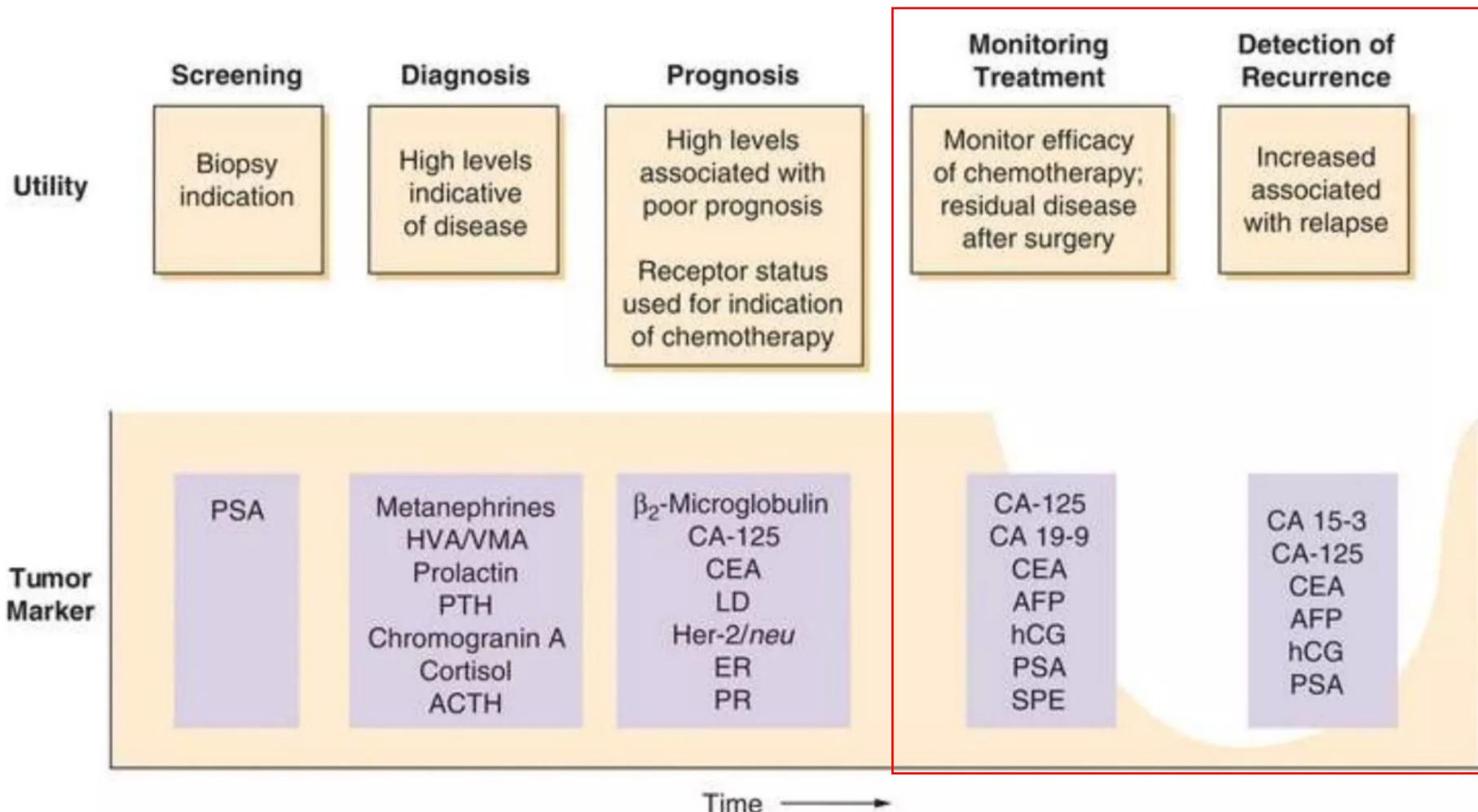
4. Uses of Biomarkers in Cancer

Tumor biomarkers often correlate with tumor burden

Role of Biomarker	Description of Use
Diagnostic	To help diagnose a cancer, distinguish benign vs. malignant disease
Prognostic	To forecast how aggressive the disease process is and/or how a patient can expect to fare in the absence of therapy
Predictive	To help identify which patients will respond to which drugs



Clinical application



Cancer biomarkers in clinical use

Names	Alternative Names	Cancer Type	Clinical Use
HER2	ErbB2, NEU, CD340	Breast cancer	Select patients for trastuzumab therapy
PSA	Prostate-specific antigen, Kallikrein 3, KLK3	Prostate cancer	Screening, diagnosis (with digital rectal examination)
Alfa-fetoprotein	AFP	Germ-cell cancer, hepatoma cancer	Diagnosis, differential diagnosis, staging, detection, recurrence, monitoring therapy
Human chorionic gonadotropin-beta	Beta-hCG	Testicular cancer	Diagnosis, staging, detection, recurrence, monitoring therapy

Cancer biomarkers in clinical use

Names	Alternative Names	Cancer Type	Clinical Use
Calcitonin	Thyrocalcitonin	Medullary thyroid cancer	Diagnosis, monitoring therapy
CA125	Mucin 16, MUC16	Ovarian cancer	Prognosis, detecting recurrence, monitoring therapy
CA 15-3	Carcinoma antigen 15-3	Breast cancer	Monitoring therapy
CA 19-9	Cancer antigen 19-9, sialylated Lewis (a) antigen	Pancreatic cancer	Monitoring therapy
Cacinoembryonic antigen	CEA	Colon cancer	Monitoring therapy, prognosis, detecting recurrence, screening for metastasis

Cancer biomarkers in clinical use

Names	Alternative Names	Cancer Type	Clinical Use
ER	Estrogen receptor	Breast cancer	Select patients for endocrine therapy
PgR	Progesterone receptor	Breast cancer	Select patients for endocrine therapy
Lactate dehydrogenase	LDH, LD	Germ cell cancer	Diagnosis, prognosis, detecting recurrence, monitoring therapy
Thyroglobulin	Tg	Thyroid cancer	Monitoring

Roles of biomarkers and their limitations

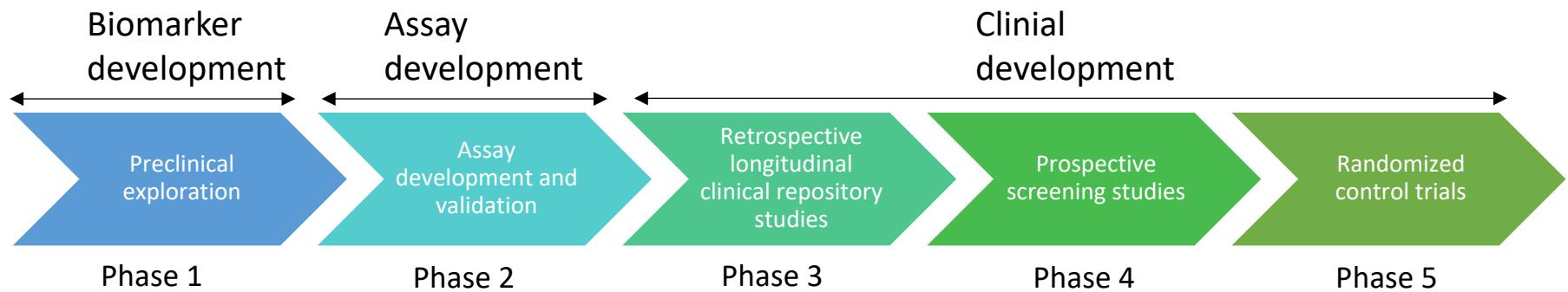
Role	Current usefulness	Comments
Population screening	Limited	<p>Screening test must have high sensitivity and specificity to avoid false positives in low cancer prevalence populations</p> <p>Most biomarkers suffer from low sensitivity and specificity as screening markers</p>
Diagnosis	Limited	Most biomarkers suffer from low sensitivity and specificity as diagnostic markers
Prognosis	Limited	Most markers have some prognostic value
Prediction of therapeutic response	High	Most provide information that aids therapy selection

Roles of biomarkers and their limitations

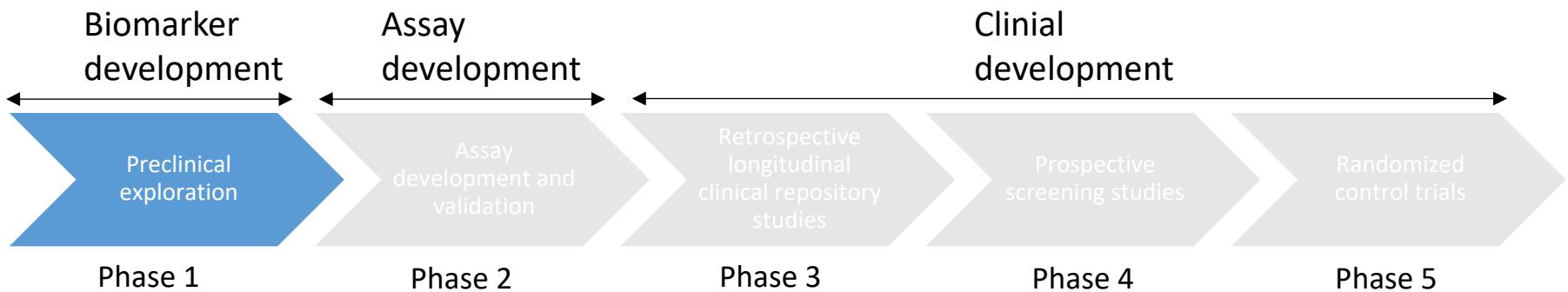
Role	Current usefulness	Comments
Tumor staging	Limited	Besides AFP and HCG, the accuracy of the markers in determining tumor stage is poor
Detecting early tumor recurrence	Controversial	Relapses may occur without marker elevation Marker elevation can be non-specific
Monitoring effectiveness of cancer therapy	High	Makers provide information on therapeutic response (effective vs. non-effective)

5. Biomarker Development

Biomarker development phases



Biomarker development phases

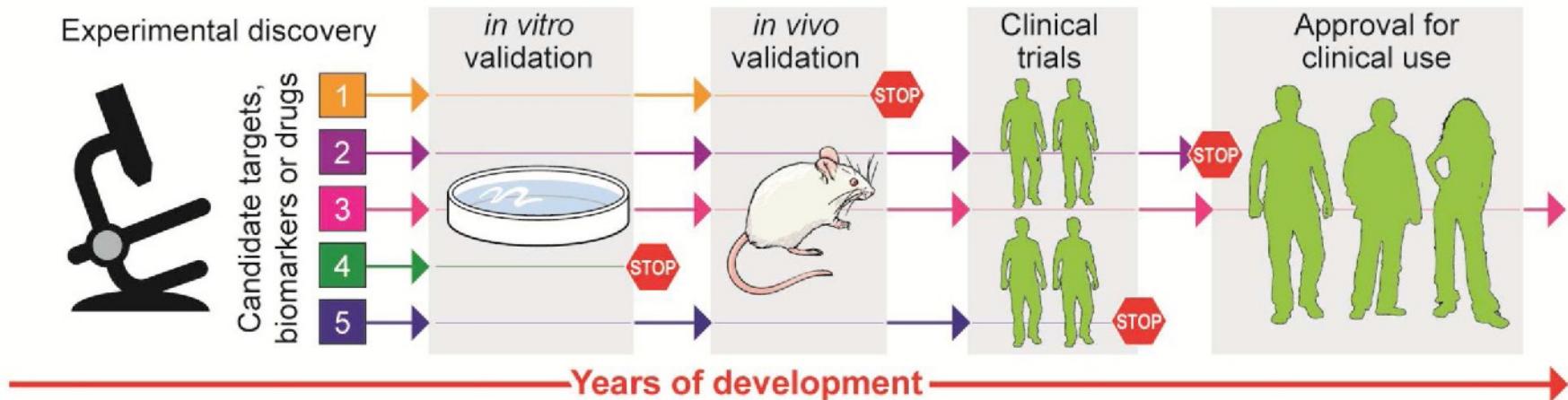


- Exploratory phase
- Prioritized based on diagnostic/prognostic/therapeutic value
- Most are pharmacodynamic biomarkers that are targets of drug inhibition
- Provide evidence that the agent reaches or modulates the putative target
- Can be conducted by analysis of samples and/or images obtained prior to and after treatment or by comparison to an untreated control

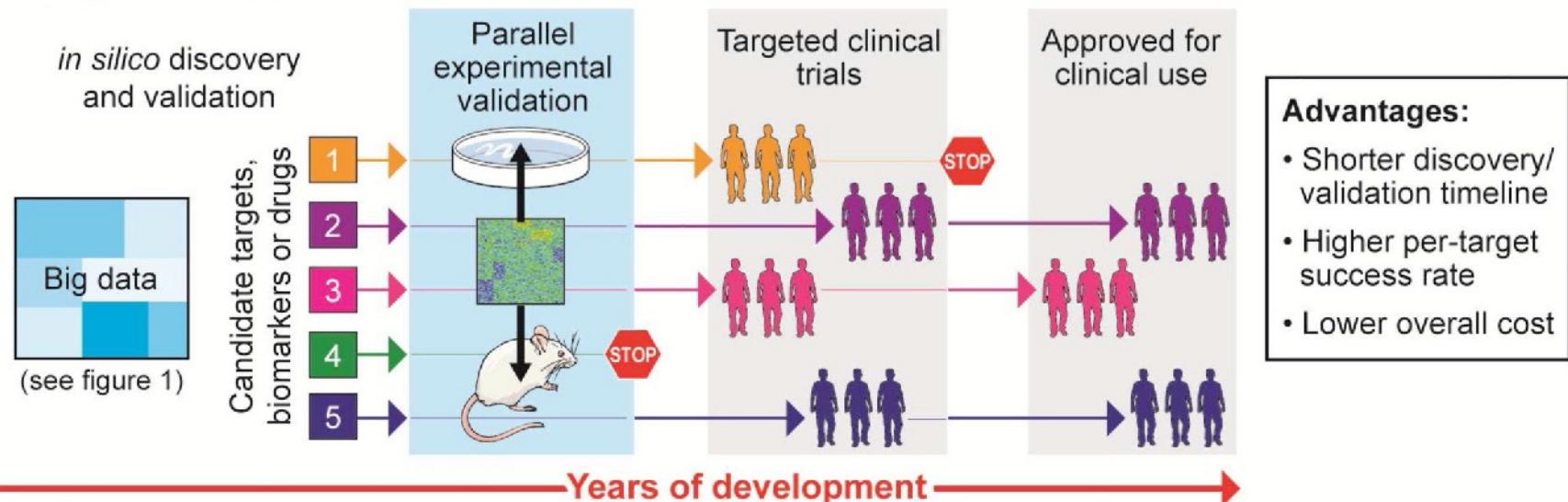
Different approaches to uncover biomarkers

- Hypothesis-driven approach
 - determine which genes to examine based on the pre-existing scientific literature
 - e.g., compare the expression of cell growth-related genes in cancer vs. healthy group
- Comprehensive approach
 - analyze the complete set of DNA and try to relate a pattern of gene expression to some feature of the cancer
 - e.g., GWAS in cancer vs. healthy group

A. Traditional biological hypothesis-based approach



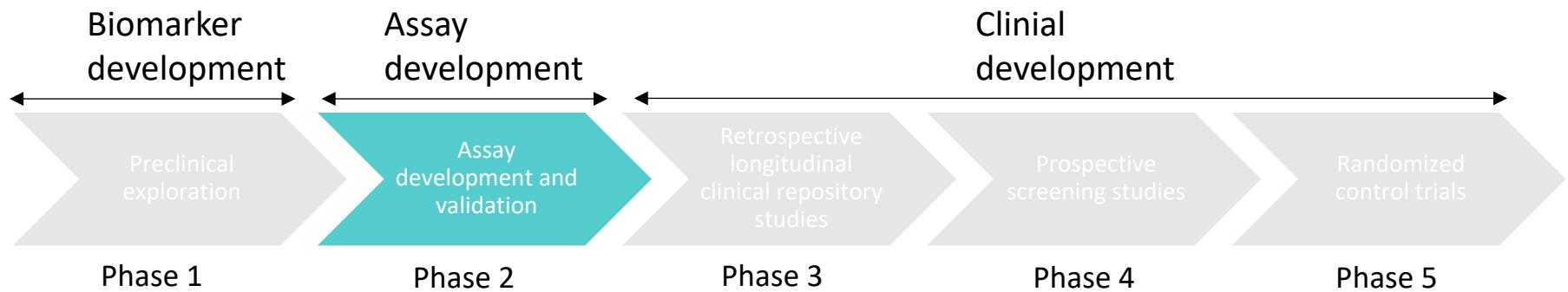
B. Big Data approach



Advantages:

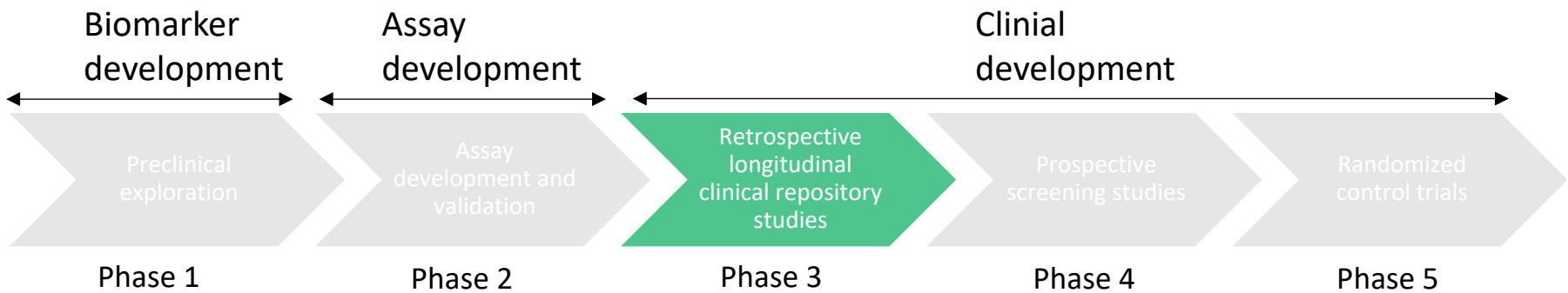
- Shorter discovery/validation timeline
- Higher per-target success rate
- Lower overall cost

Biomarker development phases



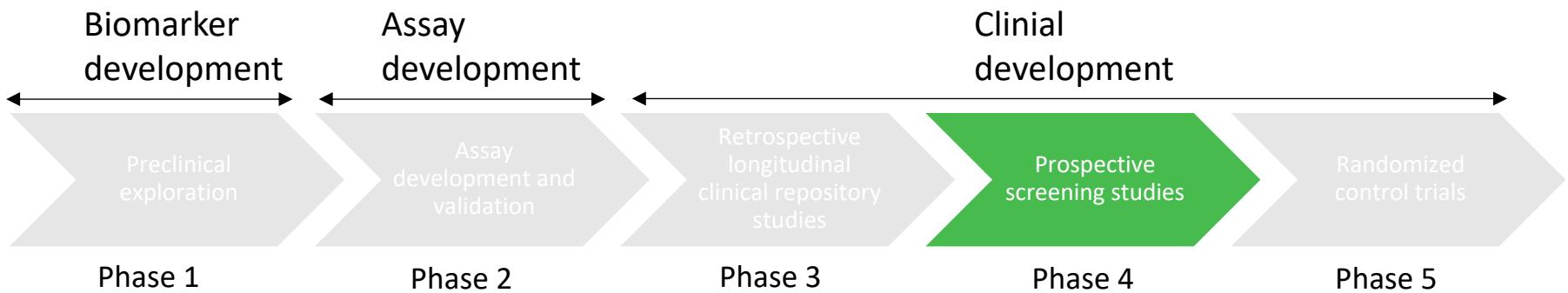
- Establish a clear association between the biomarker and clinical outcome
- Determine the dose-response relationship of a pharmacodynamic marker across a narrow set of dose cohorts and more homogenous patient population
- Assay required reproducibility validation across laboratories

Biomarker development phases



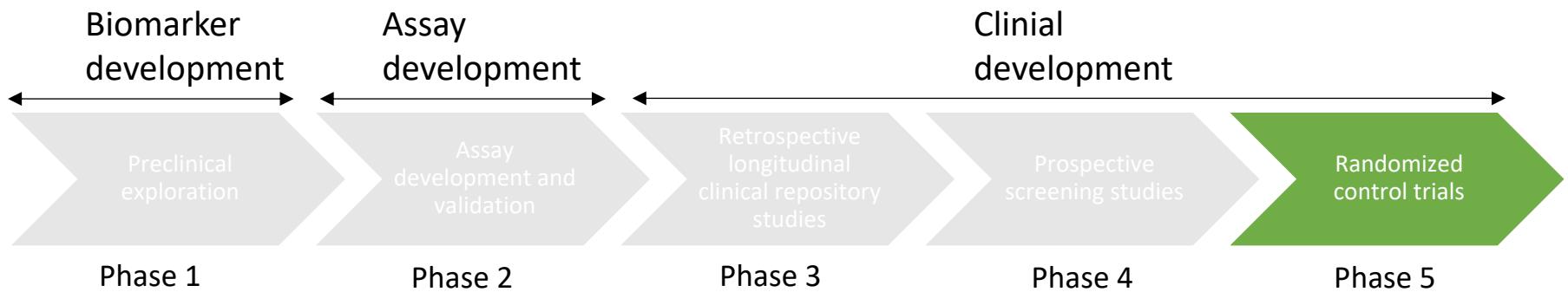
- Use specimens from patients under disease monitoring
- Evaluates the sensitivity and specificity of the assay for detection of the disease
- Demonstrate capacity of the biomarker to detect early stage of the disease (optional)

Biomarker development phases



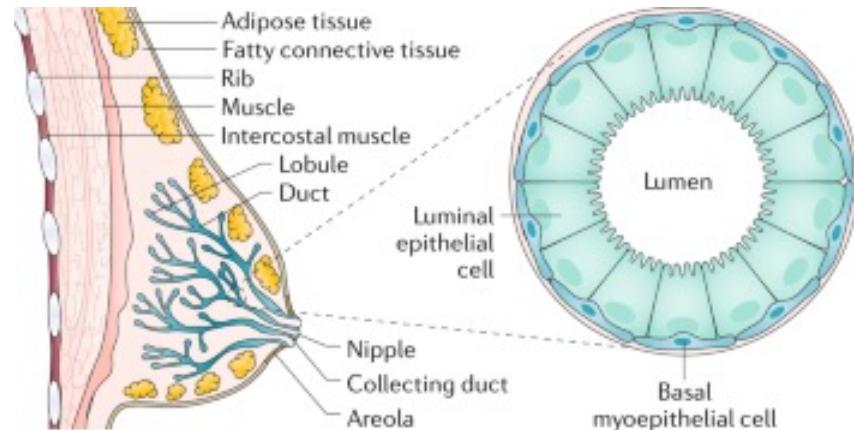
- Evaluate the sensitivity and specificity of the assay on a prospective cohort
- Individuals are screened with the biomarker assay
- Positive individuals are given diagnosis
- Biomarkers are monitored as disease progresses

Biomarker development phases



- Evaluate the overall benefits and risks of the assay on the screening population
- Determine whether the screening reduces the burden of cancer in the screening population

Breast cancer



Breast Cancer Subtypes

Luminal A

Luminal B

HER2

Basal

ER+ / PR+ / HER2-

ER+ / PR+ / HER2+

ER- / PR- / HER2+

ER- / PR- / HER2-

Chemotherapy

HER2 targeted therapy

Endocrine Therapy

Prognosis

Better

Worse

Differentiated

Cellular Morphology

Dedifferentiated

Table 1 Comparison of Oncotype DX, PAM50, and MammaPrint Multigene Tests

	Oncotype DX	PAM50	MammaPrint
Number of genes	21	50 (+5 control genes)	70
Sample requirements	Formalin-fixed, paraffin-embedded tissue	Formalin-fixed, paraffin-embedded tissue	Fresh-frozen tissue
Technique	Quantitative PCR	Quantitative PCR and nCounter technology	DNA microarray
Study population used to develop the test	Patients with ER+, node-negative, breast cancer	Patients with stage I-III breast cancer	Women <61 years, with T1-T2, N0 disease
Features	Recurrence score predicts likelihood of recurrence at 10 years Identify low-risk patients who can be spared from adjuvant chemotherapy	Provides the intrinsic subtype classification Predicts distant relapse-free survival and likelihood of recurrence at 10 years in the setting of ER+ breast cancer treated with tamoxifen Identifies patients who would benefit from neoadjuvant endocrine therapy or chemotherapy	Stratifies patients into good or poor prognosis signatures
Guidelines	NCCN and ASCO CLIA assay, no formal regulatory approval	FDA and European approvals pending	FDA approved
Prospective clinical trials	TAILORx and RxPONDER	RxPONDER trial will compare PAM50 and Oncotype DX scores	MINDACT

Abbreviations: CLIA, Clinical Laboratory Improvement Amendments; ER, estrogen receptor; PCR, polymerase chain reaction.

Oncotype DX

PROLIFERATION
Ki-67
STK15
Survivin
Cyclin B1
MYBL2

ESTROGEN
ER
PR
Bcl2
SCUBE2

BAG1

GSTM1

CD68

INVASION
Stromelysin 3
Cathepsin L2

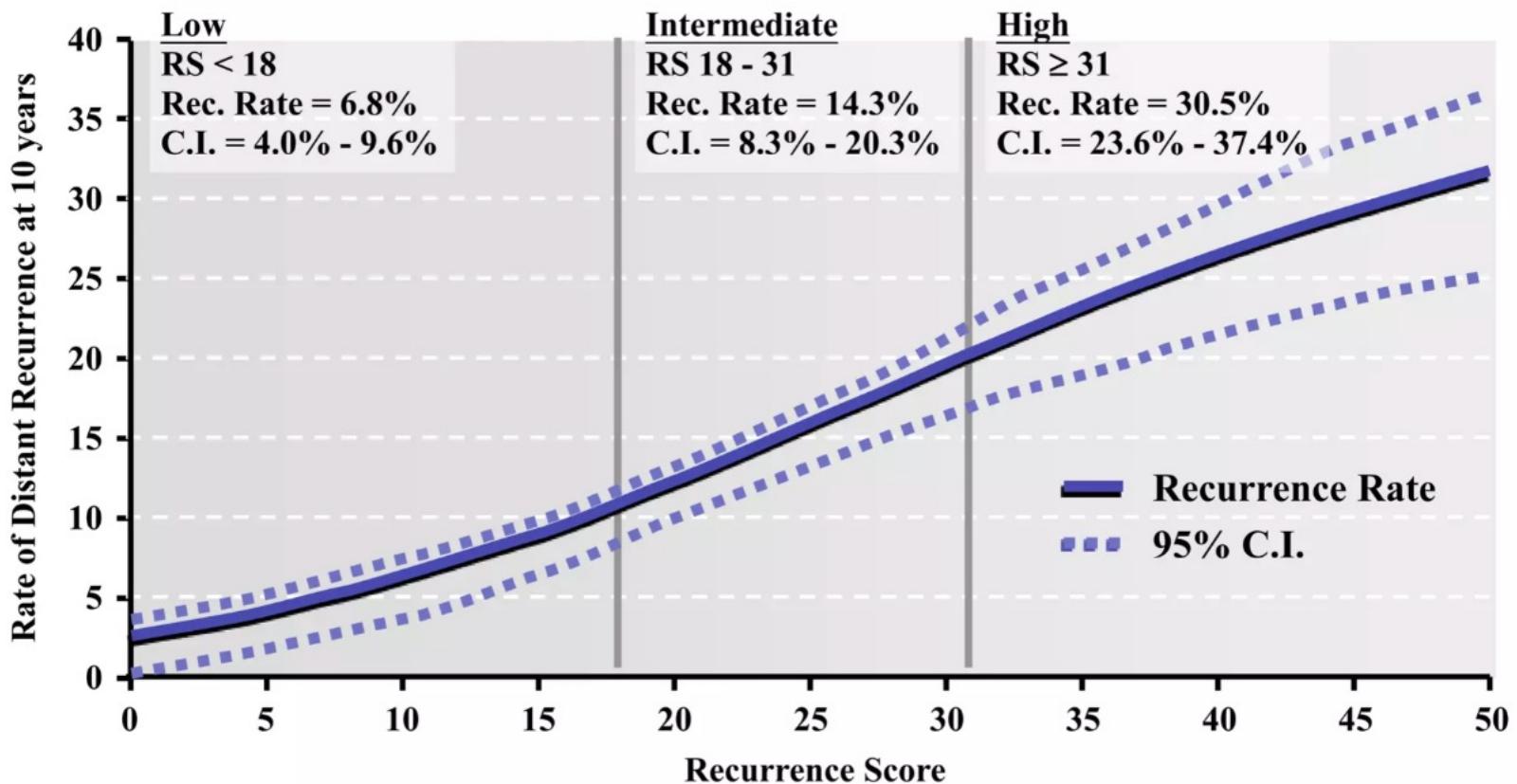
HER2
GRB7
HER2

REFERENCE
Beta-actin
GAPDH
RPLPO
GUS
TFRC

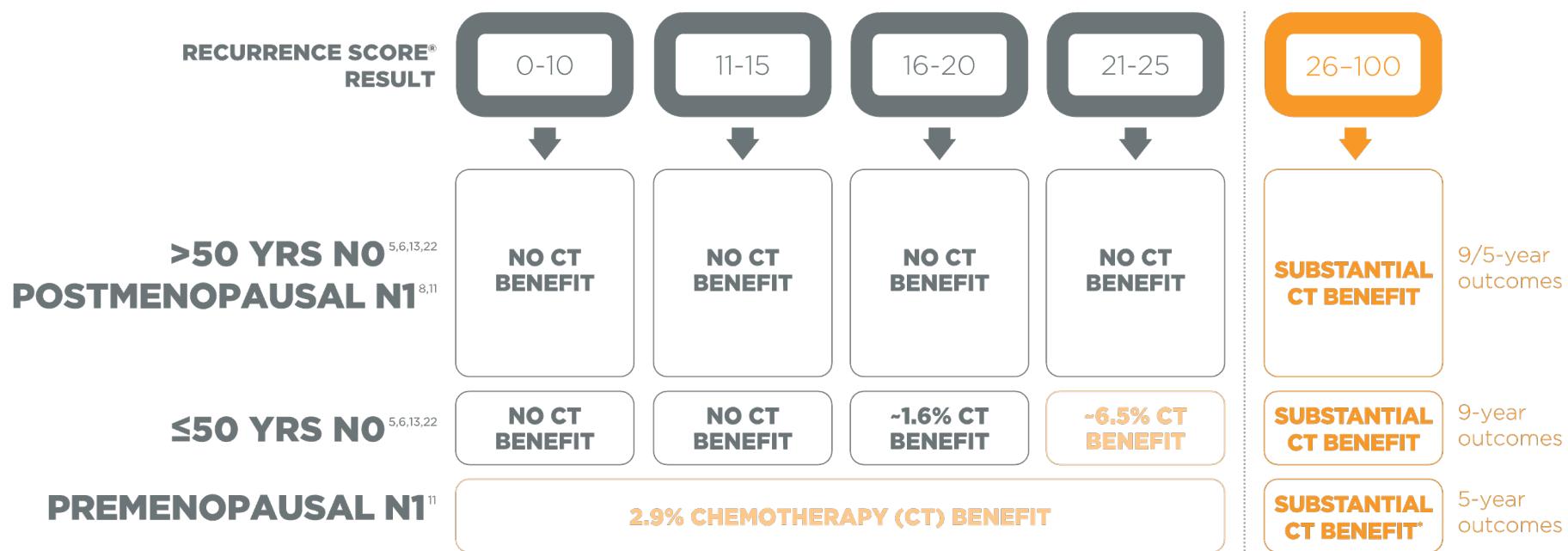
$$\text{RS} = + 0.47 \times \text{HER2 Group Score} \\ - 0.34 \times \text{ER Group Score} \\ + 1.04 \times \text{Proliferation Group Score} \\ + 0.10 \times \text{Invasion Group Score} \\ + 0.05 \times \text{CD68} \\ - 0.08 \times \text{GSTM1} \\ - 0.07 \times \text{BAG1}$$

Paik *et al.* *N Engl J Med.* 2004;351:2817-26.

Oncotype DX recurrence score



Guided treatment with Oncotype DX test



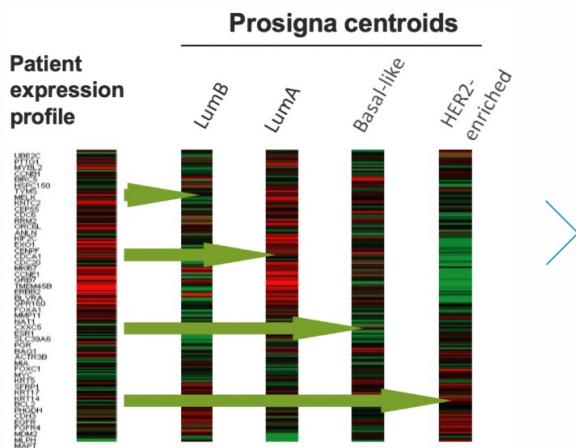
CT: chemotherapy

N0 (node-negative): patients with no cancer in axillary lymph nodes

N1 (node-positive): patients with cancer in axillary lymph nodes

Prosigna

Determine Intrinsic (Molecular) Subtype Through Pearson's Correlation to Centroids



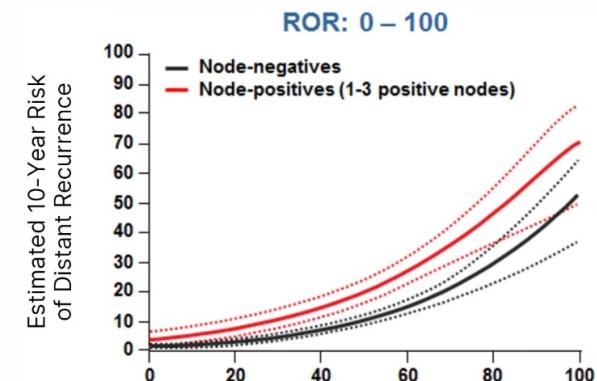
Calculate Prosigna Score

$$\text{Prosigna Score} = \frac{aR_{\text{LumA}} + bR_{\text{LumB}} + cR_{\text{Her2e}} + dR_{\text{Basal}}}{eP + fT}$$

Where:

- aR_{LumA} , bR_{LumB} , cR_{Her2e} , dR_{Basal} represent Pearson's correlation to centroids.
- eP and fT represent Proliferation score and Tumor size.

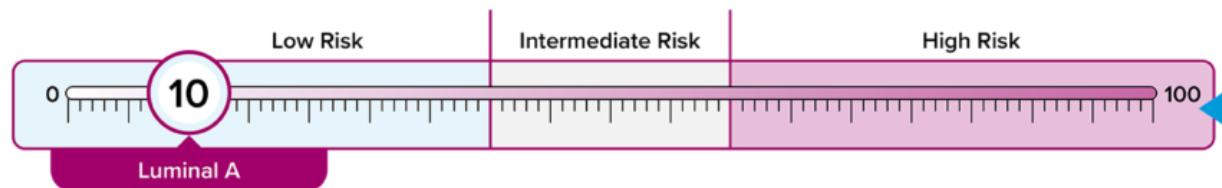
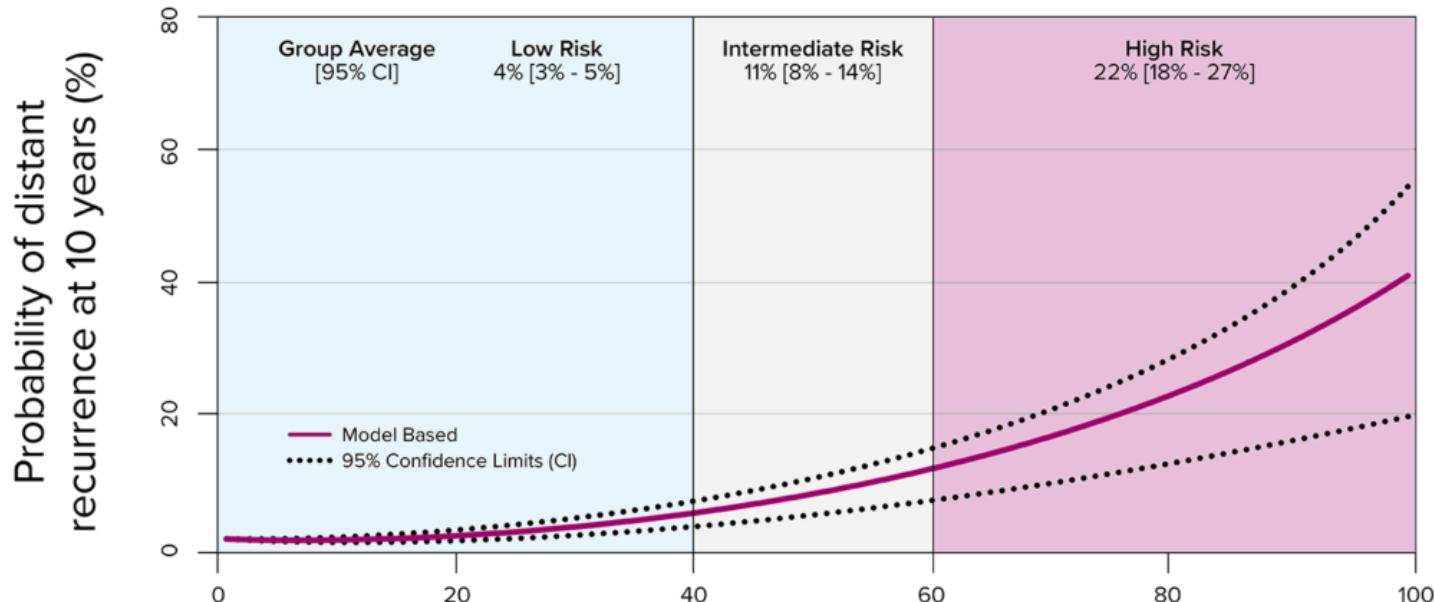
Estimate 10-Year Risk of Distant Recurrence on Endocrine Therapy Alone



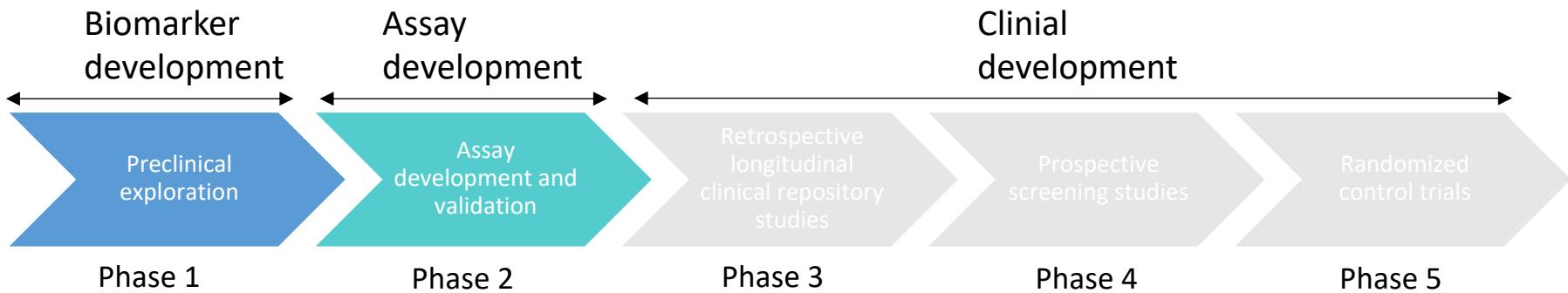
- There is a discordance between IHC-based subtype and Prosigna molecular subtypes
- IHC test is recommended to improve accuracy

Prosigna

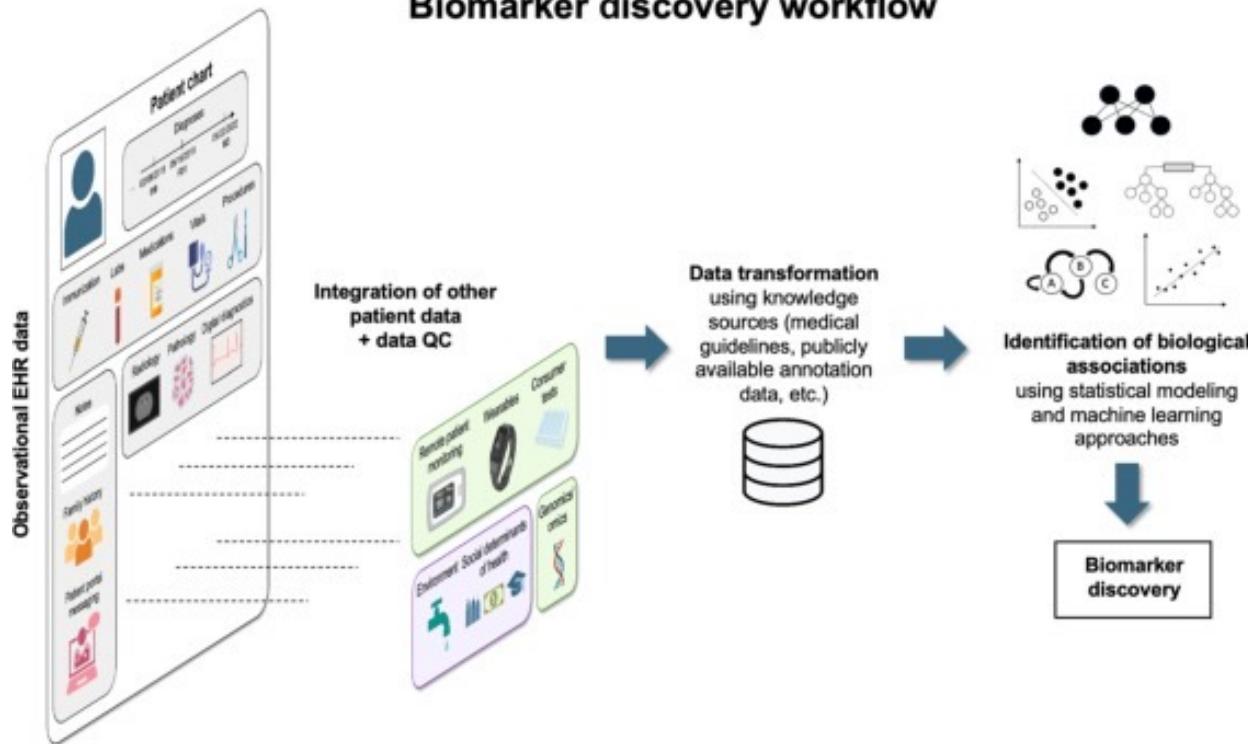
Clinical Trial Results: Probability of Distant Recurrence *(based on combined analysis of 2,400 women from the ABCSG-8 and TransATAC cohorts)*



Biomarker development phase I-II



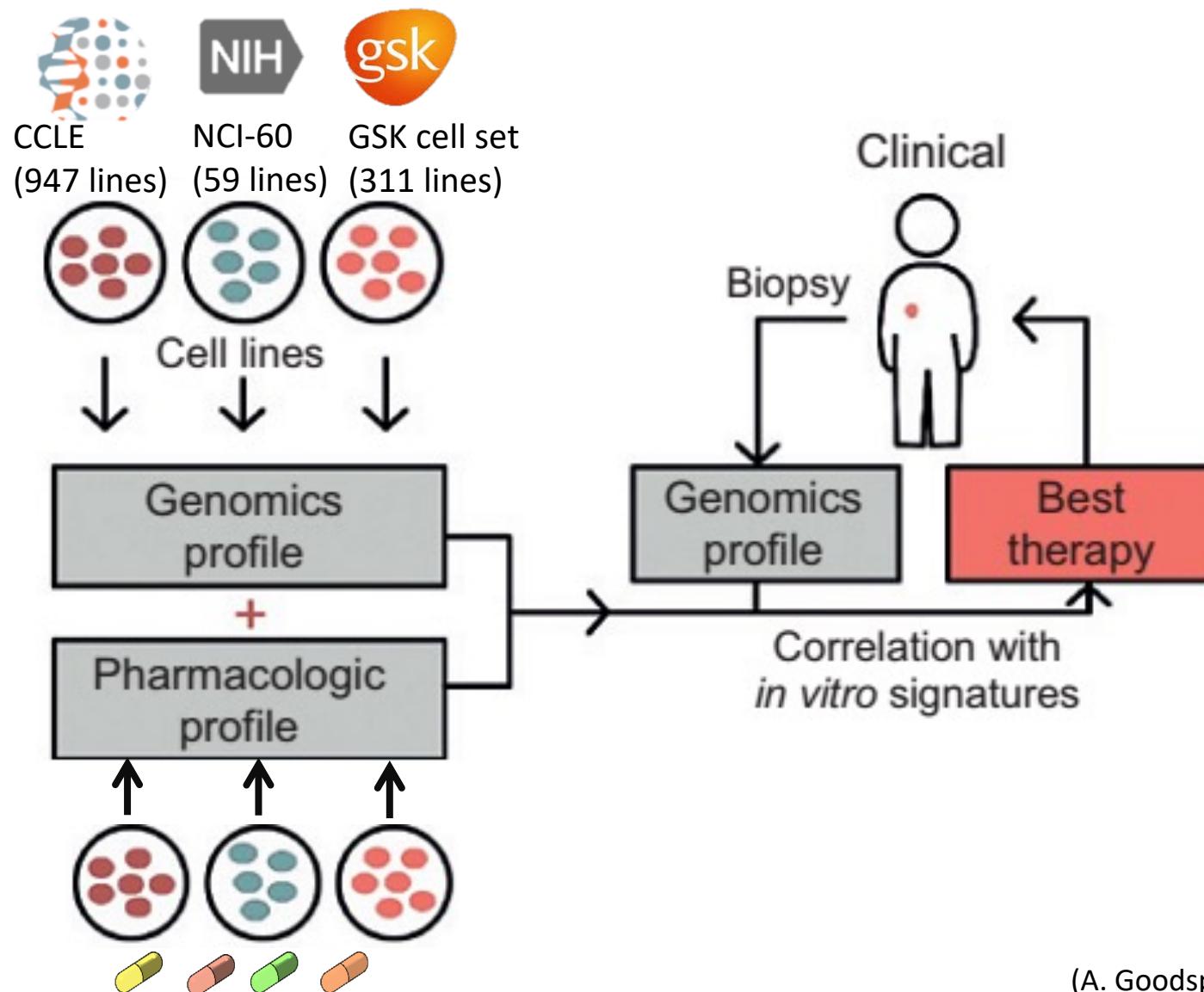
Biomarker discovery workflow



List of cancer genomics databases

#	Cancer genomic database name	Cancer alteration types	Organisms
1	The Cancer Genome Atlas (TCGA)	Copy number, mutation, methylation, gene expression, miRNA expression	Human
2	The International Cancer Genome Consortium (ICGC)	Mutation	Human
3	Catalog of Somatic Mutations in Cancer (COSMIC)	Mutation	Human
4	cBio Cancer Genomics Portal	Copy number, mutation, methylation, gene expression, miRNA expression, protein, phosphorylation	Human
5	MethyCancer	Methylation	Human
6	MutaGene	Mutation	Human
7	Moonshot project	Copy number, gene expression	Human
8	Integrative Oncogenomics Cancer Browser (IntOGen)	Copy number, mutation, gene expression	Human
9	Mouse Retrovirus Tagged Cancer Gene Database	Mutation	Mouse
10	Mouse Tumor Biology Database	Copy number, mutation, methylation, gene expression	Mouse
11	OncoDB.HCC	Copy number, gene expression, QTL	Human, mouse, rat
12	UCSC Cancer Genomics Browser	Copy number, mutation, gene expression, miRNA	Human, mouse, rat

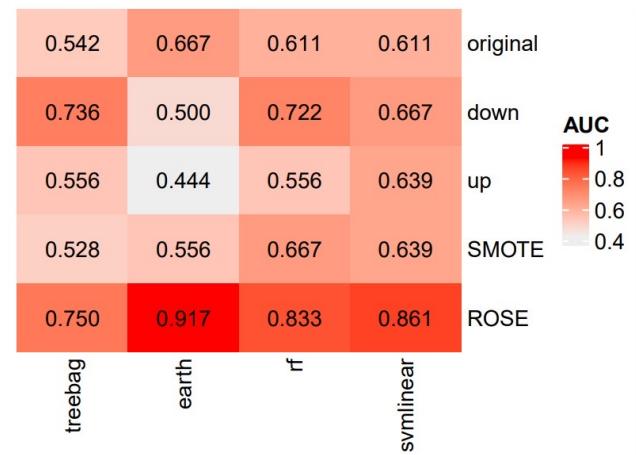
The power of cell lines for cancer treatment



Drug resistant biomarkers in NSCLC



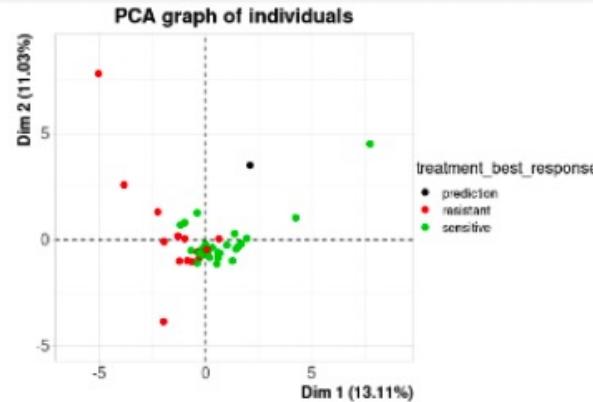
- Early stage NSCLC accounts for 20% of lung cancer in Thailand (2015)
- TCGA cohort
 - NSCLC dataset
 - Responders vs. non-responders to carboplatin/cisplatin
- Gene expression data
 - CPM normalized counts
- Features selection
 - Differentially expressed genes
 - Gene set enrichment analysis (GSEA) leading edge genes
 - Principal component analysis (PCA)
- Build classifier models
 - SVM
- Evaluate model accuracy
- Create scoring method using coefficients from SVM



Predictive Biomarker for Treatment of NSCLC with Carboplatin

Patient is

Carboplatin resistant

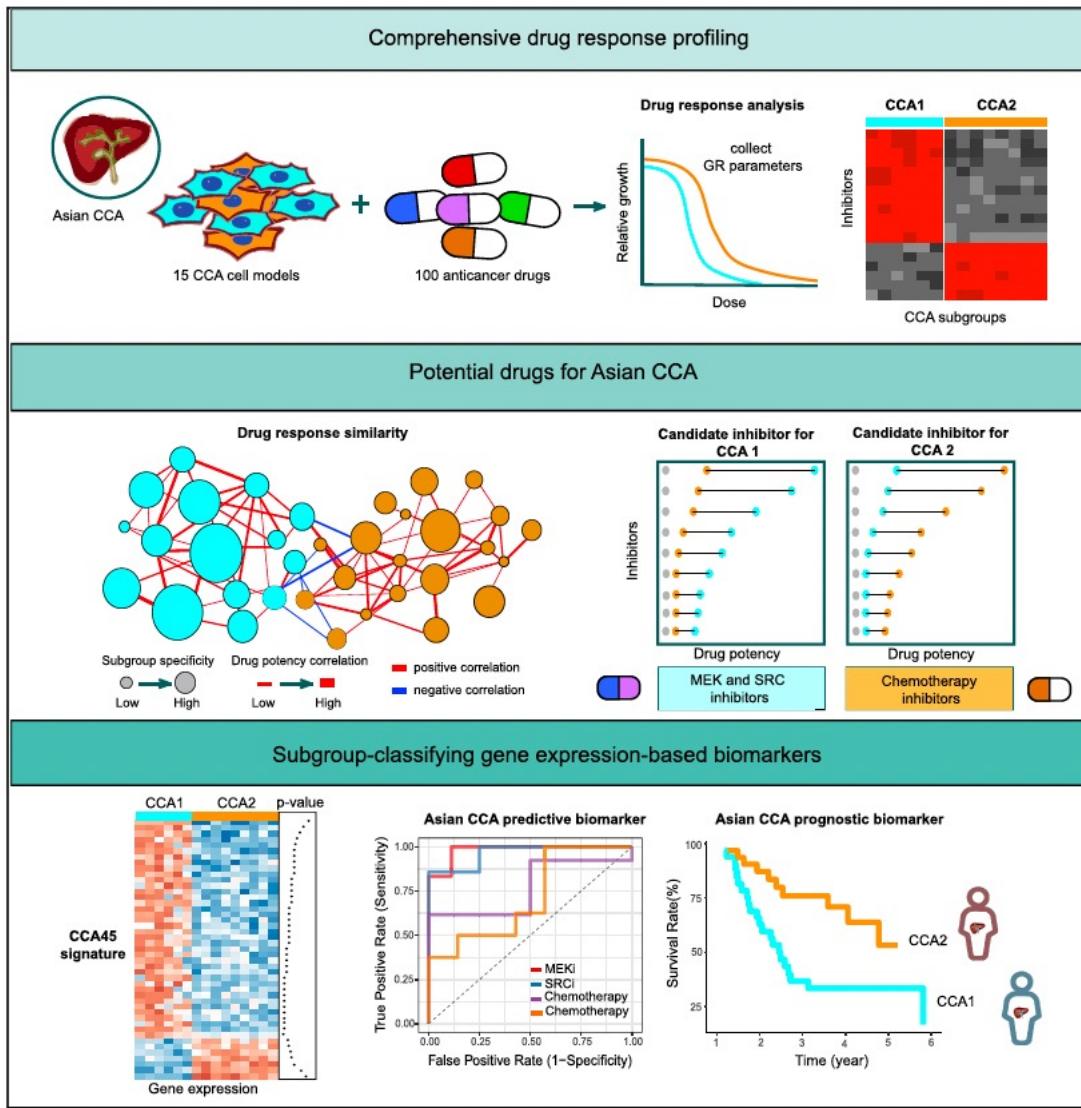


Input count number of each genes:

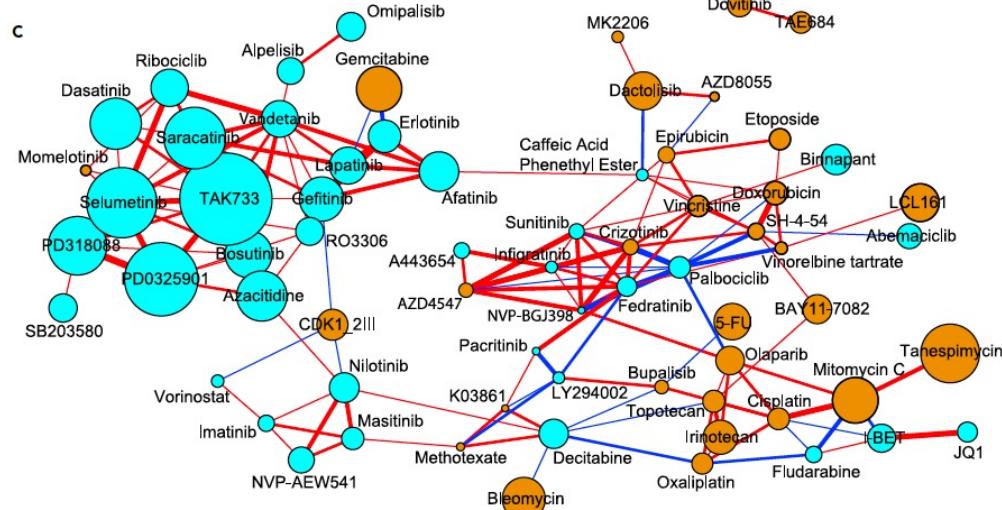
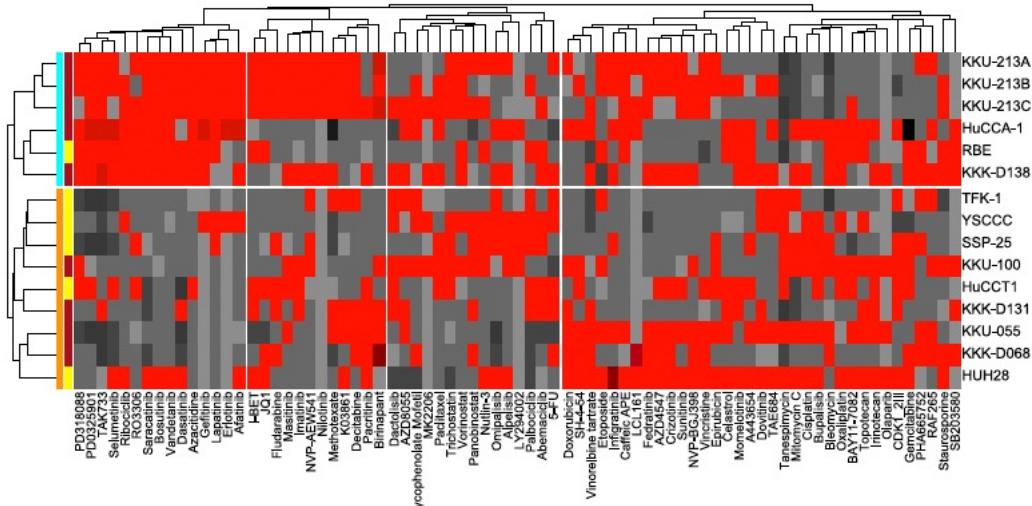
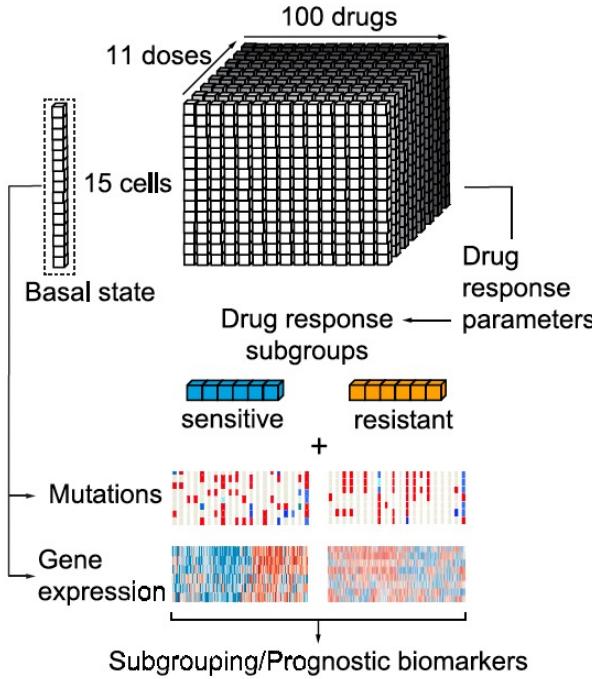
ISL1 count no.	NECAB2 count no.	H19 count no.	NTRK2 count no.	CALB2 count no.	LERFS count no.
243.742694731641	230.656405210617	-1456.88879352423	-10687.6693399111	-1210.0765146949	40.4184723721402
CPS1 count no.	CTCFL count no.	MYCN count no.	PLAAT5 count no.	OXTR count no.	RTL3 count no.
-3300.40595507413	188.24716969953	63.7861717277916	385.054711263266	160.940567796449	-22.9987076935667
COL19A1 count no.	MAGEA10 count no.	SCN2A count no.	SYT8 count no.	RPL10P6 count no.	LINC00942 count no.
248.522984056408	1142.52553169698	1008.14648701112	13.7653750522678	-73.2380875591228	996.262106824751
LAMB4 count no.	SYT5 count no.	COL2A1 count no.	LGALS4 count no.	ERVH48-1 count no.	KCNJ18 count no.
171.957949090985	219.656960733834	-1829.10891941775	-2508.90142855885	38.4254122349371	8.54881154272926

- Accuracy test in other validation data set
- Presence of the biomarkers in carboplatin-resistant NSCLC cell lines or patients

Drug response gene panel in CCA



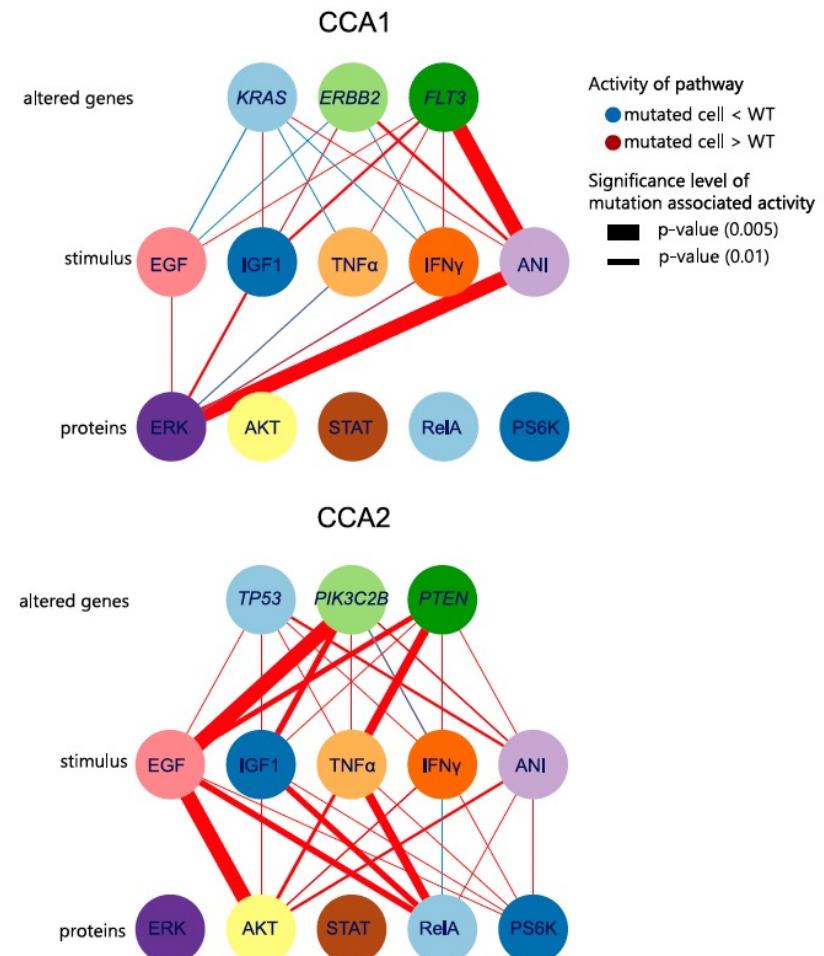
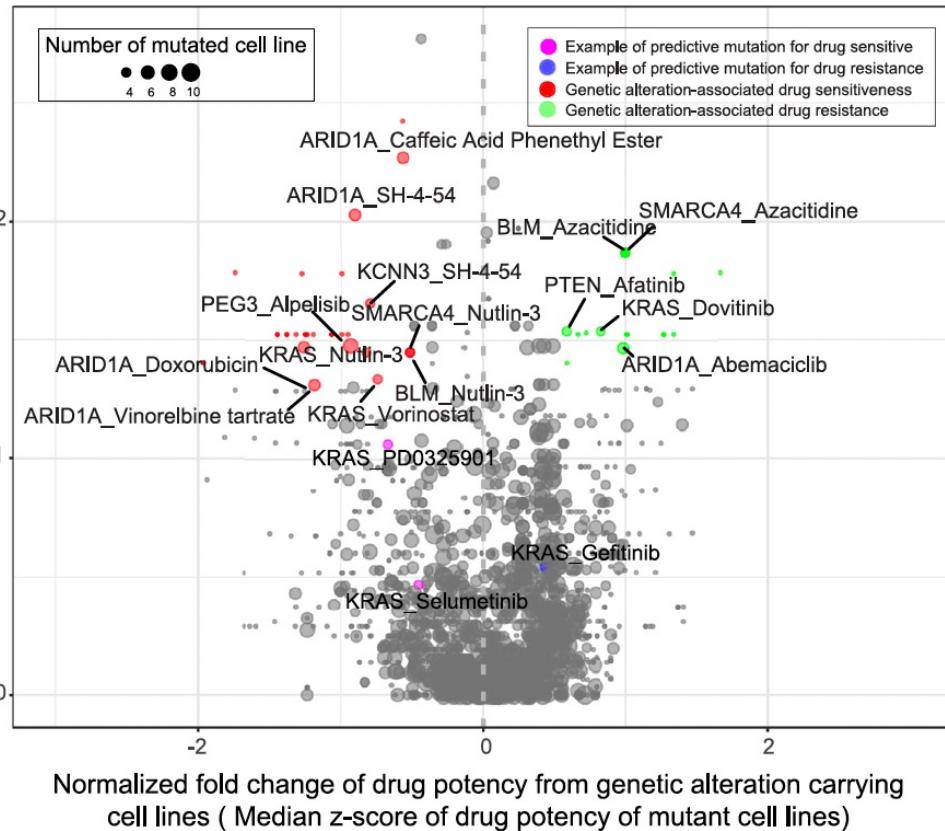
Subgroup classification based on drug response profile



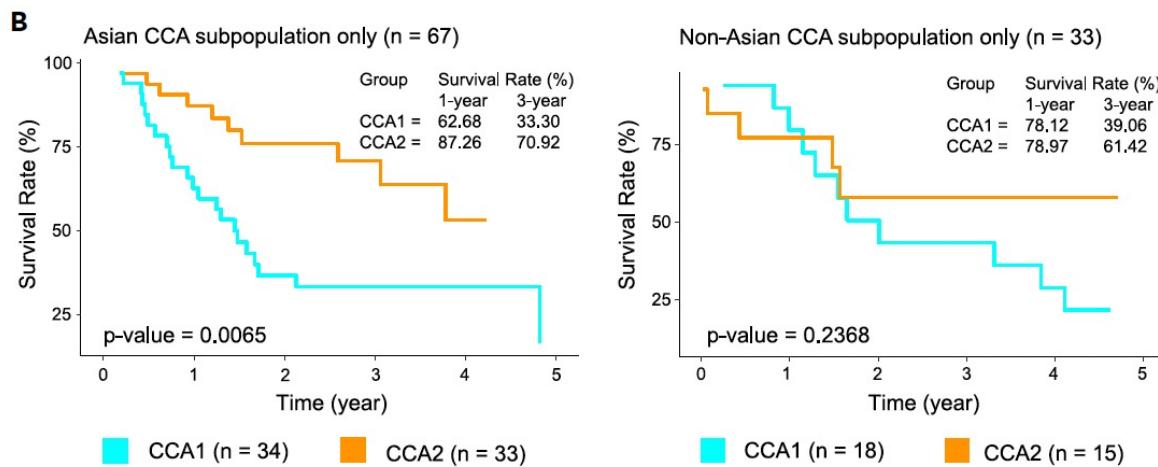
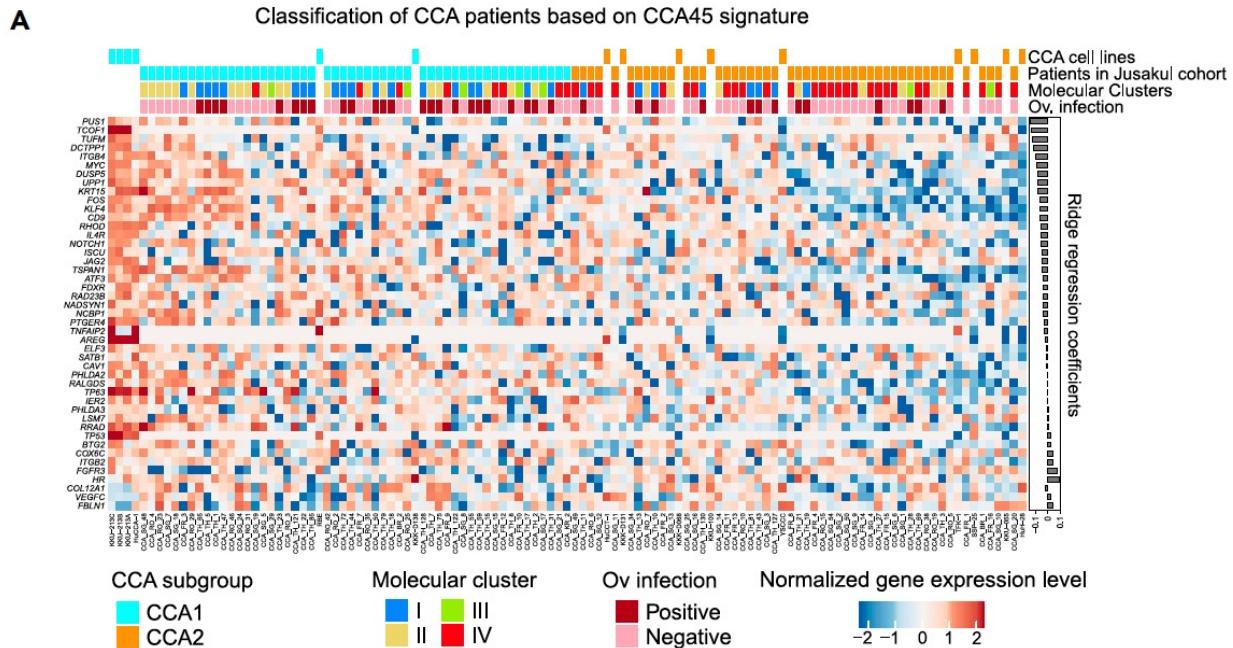
Association between drug responses and omic profiles

Association between genetic alterations and drug response

Association significance(-log₁₀(p-value))



Gene signatures predict drug response



Challenges in biomarker selection

(i) Input data selection and pre-processing

- Identifying Target variables:

Are Cancer cells

Are **not** Cancer cells (Normal/ control)

Q: Are variables mutually exclusive?

- Sample pre-processing

-Normalization

-Noise filtration

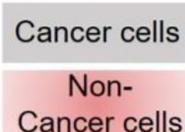
-Feature selection

Q: Is the data of good quality?

- Creating training data set:

Q: Are the no. of samples enough for the test?

Q: Are the annotation consistent?



(ii) Selection of algorithm/ prediction model and data integration

- Prepare data matrix:

Q: Are the features relevant of the task?
Q: Are the no. of samples enough for the classification?

Q: How many missing values ?

- Find proper algorithm for clustering:

Data classifier
-Supervised
-Un-supervised

Cancer cells
Non-
Cancer cells

Q: What is the no. of false negative and false positive?
Q: Can the model handle missing values?
Q: How many training data is needed?

(iii) Testing the prediction models

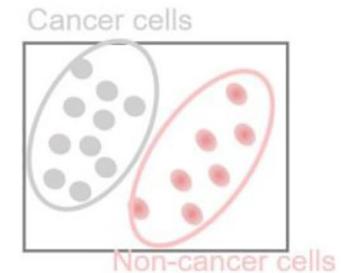
- Independent evaluation:

Q: Does the model fits data beyond the current predictive model?

Q: Are the results as expected or meaningful?

Q: Is the model clinically translatable?

- Multi-omics data integration



Specificity and sensitivity of biomarkers

- The biomarker does not have a good specificity
 - In cancer, we often find that biomarker is associated not only with cancer, but also with other diseases or conditions
- The biomarker does not have a good sensitivity
 - People who have cancer may not always have a candidate biomarker

Tissue sample accessibility

- Bodily fluids such as blood can be collected with minimal invasiveness
- Solid tumors, on the other hand, are more challenging because tissue samples need to be removed by surgery, needle biopsy or endoscopy
- Collection of multiple tissue samples are not typically included in the clinical trial's protocol

Tests for detecting biomarkers

- Accurate biomarker test
 - No need to take the test again
 - Different lab should get the same result
- Test validity
 - Specificity – give a positive result if the biomarker is present and negative result if the biomarker is no present
 - Sensitivity – give a positive result every time the biomarker is present
 - Clinical validity – biomarker correlates with a clinical outcome
- Test reliability
 - Standardized and repeatable results

Clinical relevance of biomarker research

1. Patient-centered
2. Substantially improve patient outcomes
3. Conduct in an openly shared environment
4. Encourage innovation that improves patient lives
5. All stakeholders must agreed upon standards and guidelines for conducting research, reporting results, and the clinical use of biomarkers

6. Ethical, Legal, and Social Issues with Cancer Biomarkers

Basic principles of medical ethics

Principle	Definition	Explanation
Beneficence	Duty to do more good than harm	Who will benefit and in what ways?
Non-malfeasance	Duty not to cause harm	Who might be harmed? How might this information be misused?
Individual rights?	Respect for an individual's right to be his/her own person and choose his/her own course of action	Are rights of all individuals considered and respected?
Privacy	Control over one's body and personal information; freedom from interference with personal choices	Protection of confidentiality; are there limits to this? What information is needed to save another person's life?
Justice/equity	Fair, equitable treatment for all	Are the interests of all the community considered and is potential discrimination prevented? Are resources allocated fairly?

Thank you
Q&A

patipark.k@chula.ac.th