Lecture 11:

Precision Medicine and Biomarkers

- Personalized treatment era
- Success stories
- Patient impact

Introduction to Biomedical Data Science

Lecture Contents

Part 1/3:

Precision Medicine

- Conceptual framework
- Technology enablers
- Clinical applications

Personalized vs Precision Medicine

Terminology Evolution

- N-of-1 trials
- Population stratification
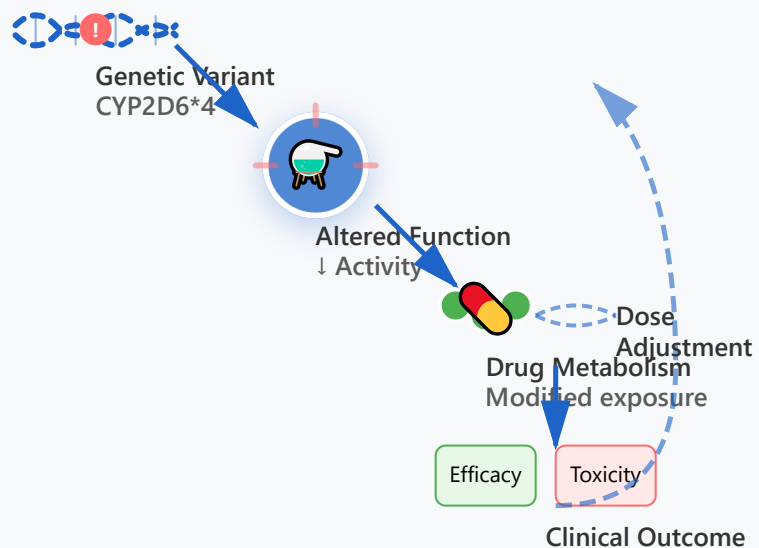- Individual variability

Clinical Approaches

- Targeted therapies
- Genomic profiling
- Treatment optimization

Healthcare Economics

Precision medicine balances cost-effectiveness with improved patient outcomes through targeted interventions, reducing trial-and-error approaches and minimizing adverse drug reactions.

Pharmacogenomics (PGx)

Genetic Variant
CYP2D6*4

Altered Function
↓ Activity

Drug Metabolism
Modified exposure

Dose Adjustment

Efficacy

Toxicity

Clinical Outcome

PGx Variants

Genetic variations affecting drug response, metabolism, and efficacy

Drug Metabolism

CYP450 enzymes and transporter polymorphisms impact drug levels
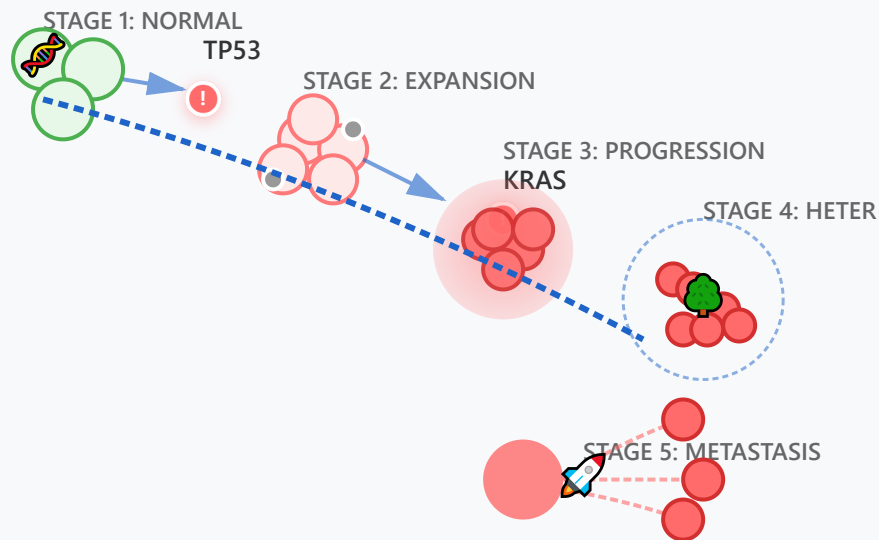
Dosing Algorithms

Genotype-guided dosing for warfarin, clopidogrel, and more

CPIC Guidelines

Clinical Pharmacogenetics Implementation Consortium standards

Implementation Barriers: Cost, infrastructure, education, and workflow integration remain challenges for widespread PGx adoption

Cancer Genomics

STAGE 1: NORMAL
TP53

STAGE 2: EXPANSION

STAGE 3: PROGRESSION
KRAS

STAGE 4: HETER

STAGE 5: METASTASIS

Driver Mutations
Mutations providing selective growth advantage to cancer cells

Passenger Mutations
Neutral mutations accumulated during tumor evolution

Clonal Evolution
Tumor progression through sequential genetic changes

Tumor Heterogeneity
Genetic diversity within and between tumors

Minimal Residual Disease: Ultra-sensitive detection of remaining cancer cells post-treatment

Tumor Profiling

NGS Panels

Next-generation sequencing for comprehensive mutation analysis

TMB Assessment

Tumor Mutational Burden predicts immunotherapy response

MSI Status

Microsatellite Instability as biomarker for treatment selection

HRD Testing

Homologous Recombination Deficiency guides PARP inhibitor use

PD-L1 Expression: Key biomarker for immune checkpoint inhibitor therapy eligibility

Liquid Biopsy

Blood Draw → Isolation → Analytes → NGS/PCR → Results

ATCG
TAGC
CGAT
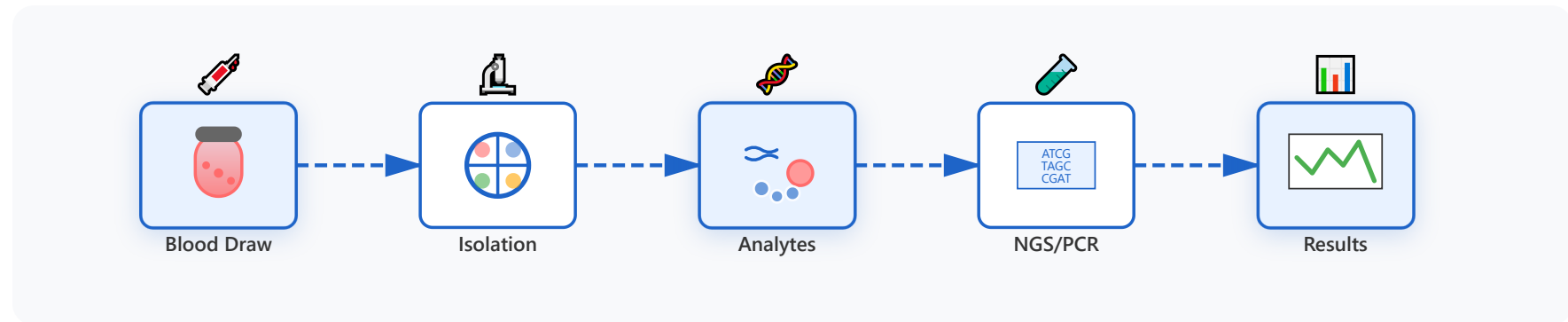
ctDNA Detection
Circulating tumor DNA analysis from blood samples

CTCs Analysis
Circulating Tumor Cells isolation and characterization

Exosomes
Extracellular vesicles carrying tumor-derived molecules

Early Detection
Non-invasive screening for cancer presence

Monitoring Response: Real-time treatment efficacy assessment through serial sampling

Companion Diagnostics

FDA Co-development
Simultaneous drug and diagnostic test development

Test Validation
Analytical and clinical validation requirements

Labeling Requirements
Regulatory guidelines for test-drug pairing

Market Examples
HER2 testing, EGFR mutations, ALK rearrangements

Future Trends: Multi-gene panels and liquid biopsy-based companion diagnostics

Part 2/3:

Biomarker Discovery

- Discovery strategies

- Validation frameworks

- Statistical considerations

Types of Biomarkers

Diagnostic   Identify presence or absence of disease

Prognostic   Predict disease outcome independent of treatment

Predictive   Forecast response to specific therapy

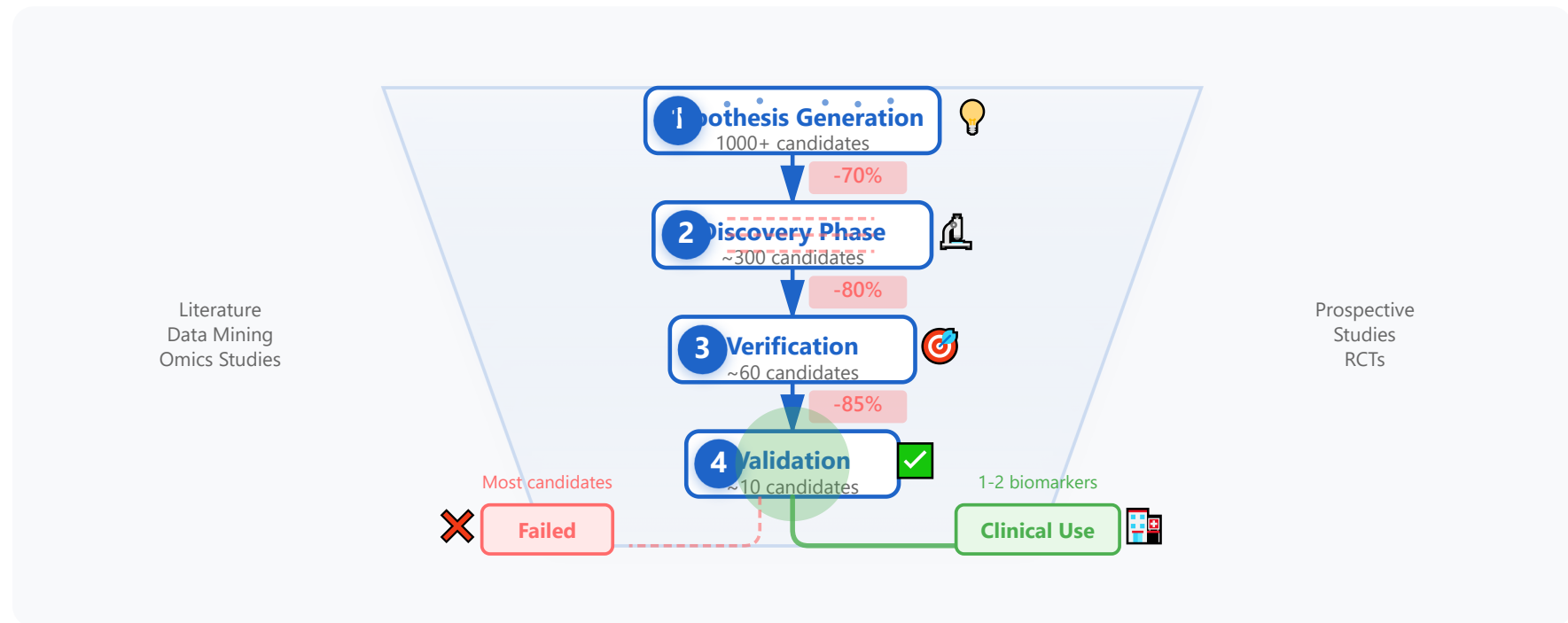**Pharmacodynamic**   Measure biological response to treatment

Safety   Detect potential adverse effects or toxicity

Each biomarker type serves distinct clinical purposes in precision medicine workflows

Discovery Pipeline

1 Hypothesis Generation
1000+ candidates

-70%

2 Discovery Phase
~300 candidates

-80%

3 Verification
~60 candidates

-85%

4 Validation
~10 candidates

Literature
Data Mining
Omics Studies

Prospective
Studies
RCTs

Most candidates

❌ Failed

1-2 biomarkers

Clinical Use

⚠️ Attrition rate is high - most candidates fail before reaching clinical implementation

Statistical Methods

Feature Selection   LASSO, elastic net, random forests for variable selection

Multiple Testing   FDR correction and permutation-based adjustments
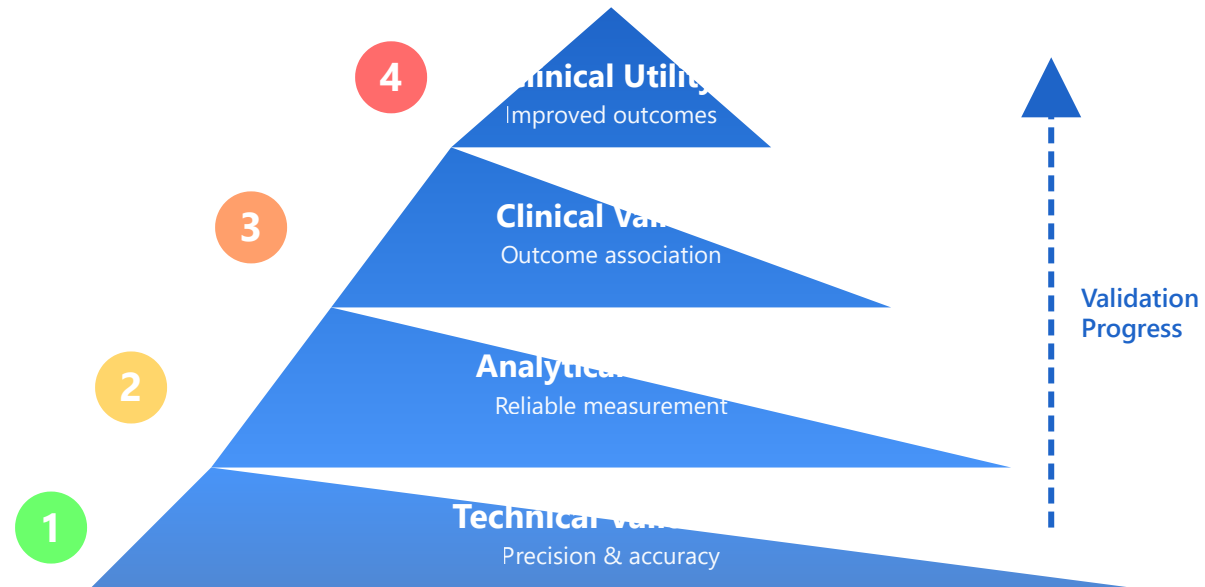
Cross-validation   K-fold CV and nested CV prevent overfitting

Bootstrap CI   Confidence intervals through resampling methods

Power Analysis: Essential for determining adequate sample size and detecting true effects

Validation Strategies

4 Technical Utility
Improved outcomes

3 Clinical Val...
Outcome association

2 Analytical...
Reliable measurement

1 Technical Val...
Precision & accuracy

Validation Progress

Technical
Assay precision, accuracy, and reproducibility in controlled settings

Analytical
Reliable and accurate measurement of biomarker across conditions

Clinical
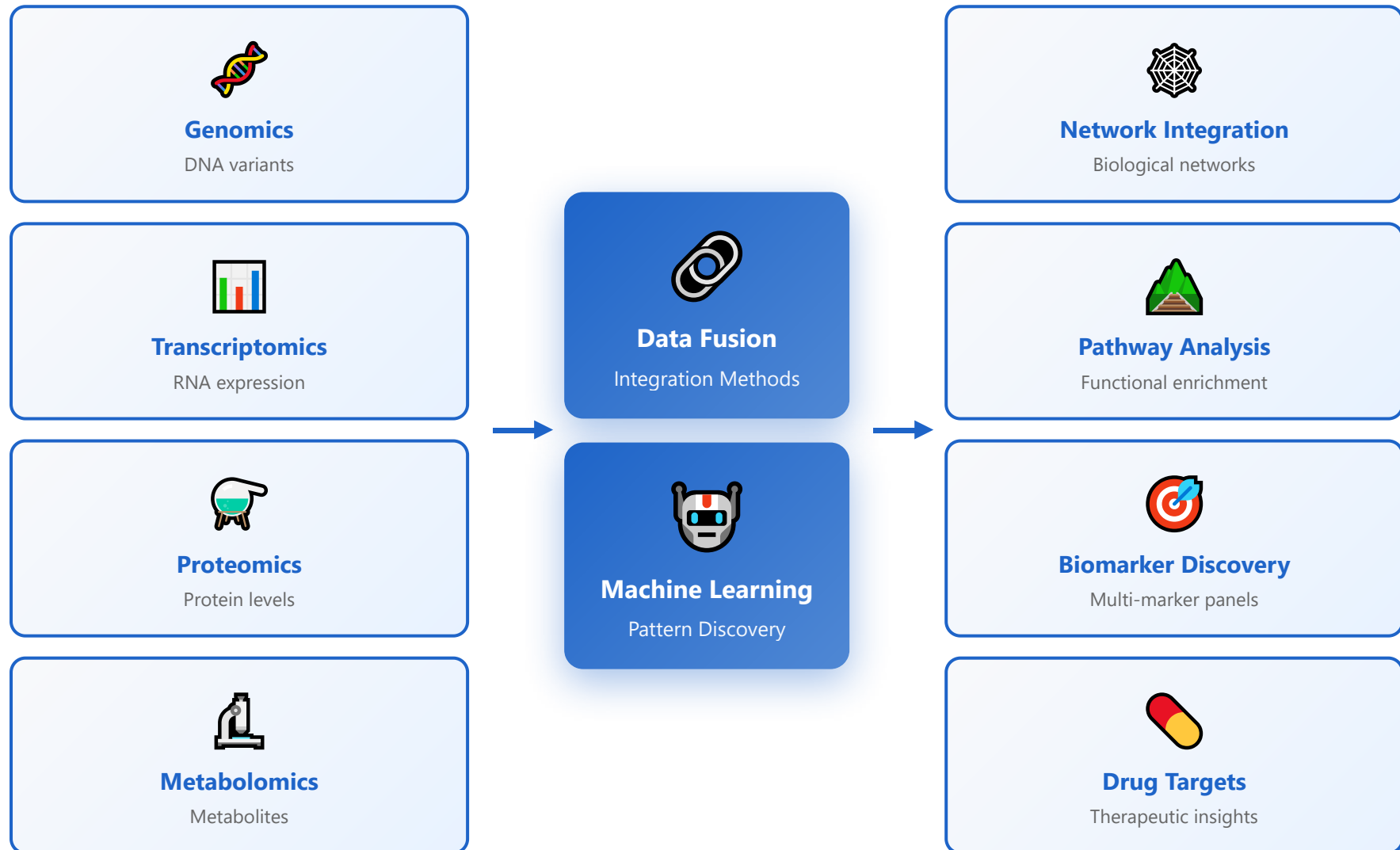Performance and association with outcome in patient population
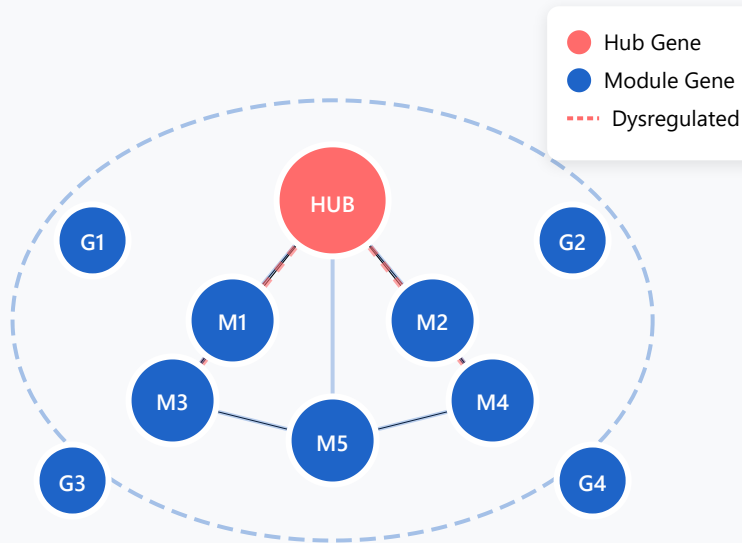
Utility
Improves decision-making and patient outcomes in practice

Clinical Utility: Biomarker must improve decision-making and patient outcomes in practice

Multi-omics Integration

Genomics
DNA variants

Transcriptomics
RNA expression

Proteomics
Protein levels

Metabolomics
Metabolites

Data Fusion
Integration Methods

Machine Learning
Pattern Discovery

Network Integration
Biological networks

Pathway Analysis
Functional enrichment

Biomarker Discovery
Multi-marker panels

Drug Targets
Therapeutic insights

Network Biomarkers

Module Identification
Detect functional gene/protein modules

Hub Genes
Central nodes in biological networks

Edge Biomarkers
Dysregulated interactions between molecules
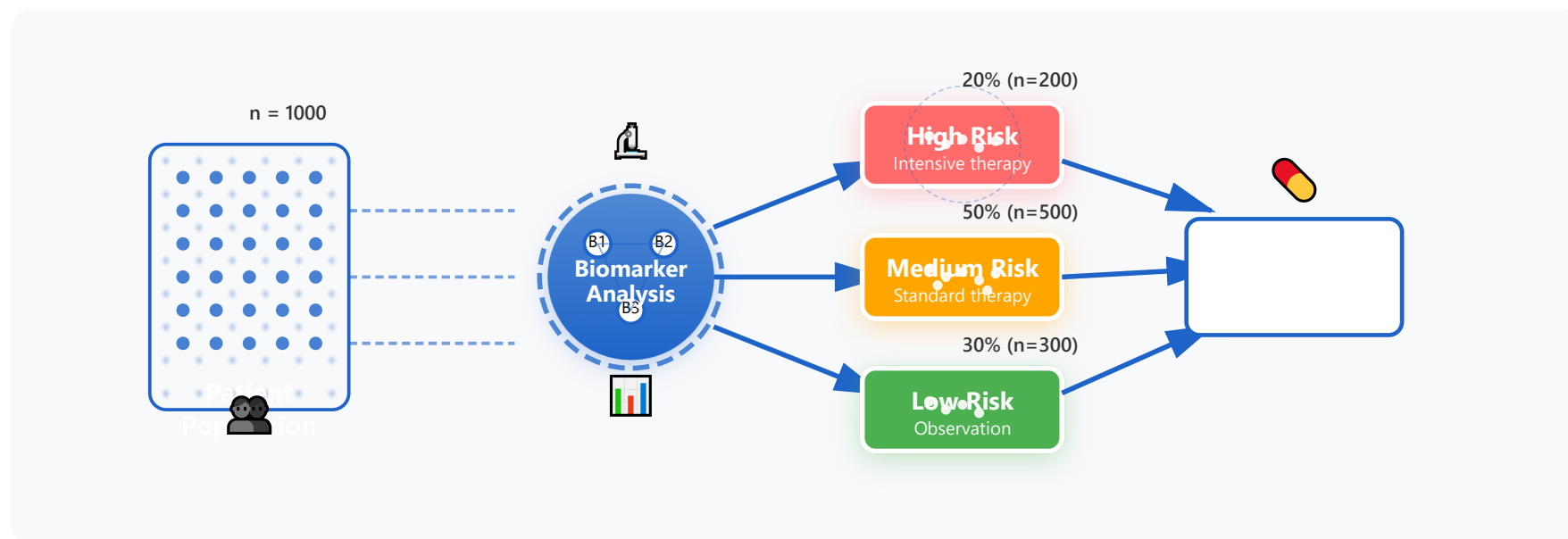
Dynamic Networks
Time-varying network states

Robustness Analysis: Network biomarkers often more stable than individual markers

Legend:
- Hub Gene
- Module Gene
- Dysregulated

Part 3/3:

Clinical Translation

- Implementation strategies

- Real-world evidence

- Healthcare integration

Patient Stratification

n = 1000

Patient Population

Biomarker Analysis

B1 B2 B3

20% (n=200)
High Risk
Intensive therapy

50% (n=500)
Medium Risk
Standard therapy

30% (n=300)
Low Risk
Observation

Subgroup Identification

Clustering and classification to define patient groups

Risk Groups

Stratify patients by disease risk or prognosis

Treatment Assignment

Match therapies to molecular profiles

Enrichment Strategies

Select patients likely to benefit from intervention

Treatment Selection

Decision Algorithms

Computational tools for treatment recommendations

Combination Therapies

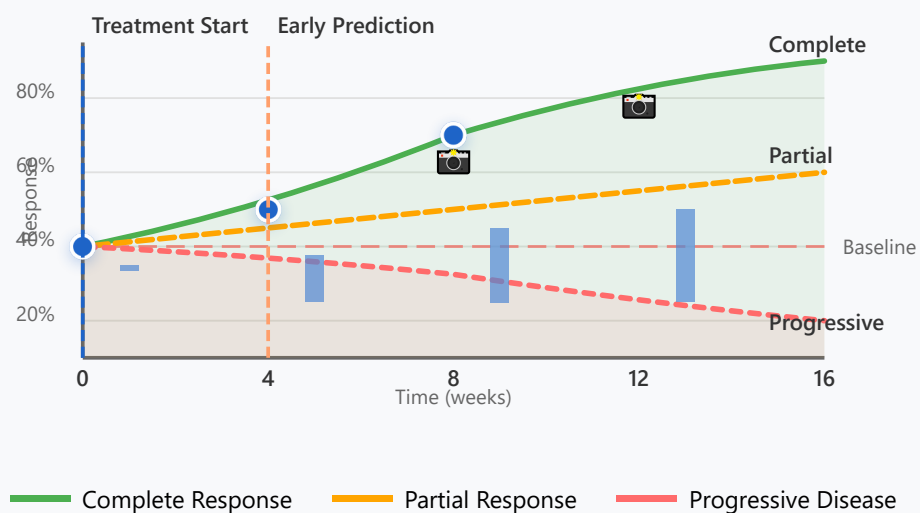Biomarker-guided multi-drug regimens

Sequential Treatments

Treatment sequencing based on molecular changes

Resistance Mechanisms

Anticipate and overcome therapy resistance

Cost-effectiveness: Balance improved outcomes with healthcare resource utilization

Response Prediction

Early Indicators

Biomarkers predicting response before clinical signs

Surrogate Endpoints

Intermediate markers replacing clinical outcomes
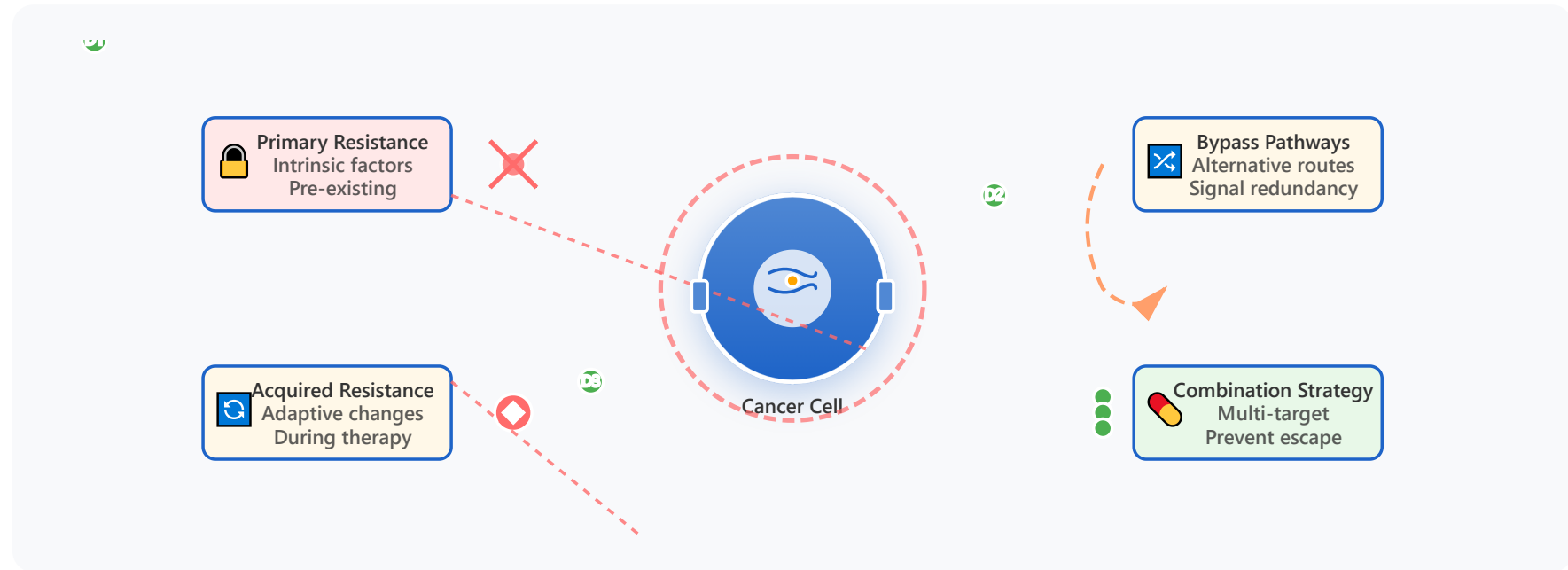
Imaging Biomarkers

Radiological features predicting treatment response

ctDNA Dynamics

Circulating DNA changes track treatment efficacy

Resistance Prediction: Anticipate resistance before clinical progression occurs

Resistance Mechanisms

Primary Resistance
Intrinsic factors
Pre-existing

Acquired Resistance
Adaptive changes
During therapy

Cancer Cell

Bypass Pathways
Alternative routes
Signal redundancy

Combination Strategy
Multi-target
Prevent escape

Monitoring Approaches: Serial biopsies and liquid biopsy track resistance evolution

Clinical Trial Design

Basket Trials

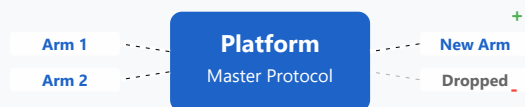Multiple tumor types with same molecular alteration

Lung — BRAF — Melanoma
Colon — BRAF — Thyroid

Umbrella Trials

Single disease with multiple biomarker-driven arms

NSCLC
EGFR+   ALK+   KRAS+

Platform Trials

Perpetual trials adding/dropping arms dynamically

Arm 1
Arm 2
Platform
Master Protocol
New Arm +
Dropped −

Biomarker-stratified

Randomization based on biomarker status

Patients — Test
Positive → Drug A
Negative → Drug B

Master Protocols: Efficient frameworks for precision oncology trials

Regulatory Approval

Biomarker Qualification

FDA/EMA process to approve biomarkers

Context of Use

Specific intended application of biomarker

Evidence Requirements

Standards for analytical and clinical validation

Global Harmonization

Align biomarker standards across regions

Expedited Pathways: Breakthrough designations accelerate biomarker-drug approval

Cost-effectiveness

QALY Analysis

Quality-Adjusted Life Years in economic evaluation

Budget Impact

Healthcare system financial implications

Value Frameworks

ASCO, ESMO, NCCN value assessment tools

Reimbursement

Insurance coverage and payment models

Access Issues: Balance innovation with affordability and equitable distribution

Implementation Barriers

Technical Challenges

Assay complexity and standardization needs

Clinical Workflow

Integration into existing healthcare systems

Education Needs

Training clinicians in precision medicine

Infrastructure

Lab capabilities and IT systems requirements

Equity Concerns: Ensure access across socioeconomic and geographic boundaries

🎯 Precision Medicine Success Stories

HER2+ Breast Cancer

Trastuzumab (Herceptin)

First companion diagnostic: **HER2 testing** identifies patients who benefit from targeted therapy. Dramatically improved survival in HER2+ patients.

CML Treatment Revolution

Imatinib (Gleevec)

Targeting **BCR-ABL fusion** transformed CML from fatal to manageable chronic disease. 5-year survival >90%.

MSI-High Cancers

Pembrolizumab (Keytruda)

First tissue-agnostic approval based on **MSI biomarker**. Immunotherapy success across multiple cancer types.

Melanoma Breakthrough

BRAF Inhibitors

BRAF V600E mutation testing enables combination therapy. Median survival improved from 9 to 25+ months.

CAR-T Cell Therapy

Tisagenlecleucel (Kymriah)

💡 Key Lessons

- Strong biomarker-target link
- Rigorous validation studies
- Drug-diagnostic co-development

CD19+ B-cell targeting in leukemia/lymphoma. Complete remission in 80%+ of refractory patients.

🚀 Future Directions in Precision Medicine

Single-cell Biomarkers
Tumor heterogeneity analysis at single-cell resolution for personalized therapy

Digital Biomarkers
Wearables and mobile health monitoring for continuous patient assessment

AI-discovered Biomarkers
Machine learning identifies novel patterns invisible to traditional analysis

Combination Panels
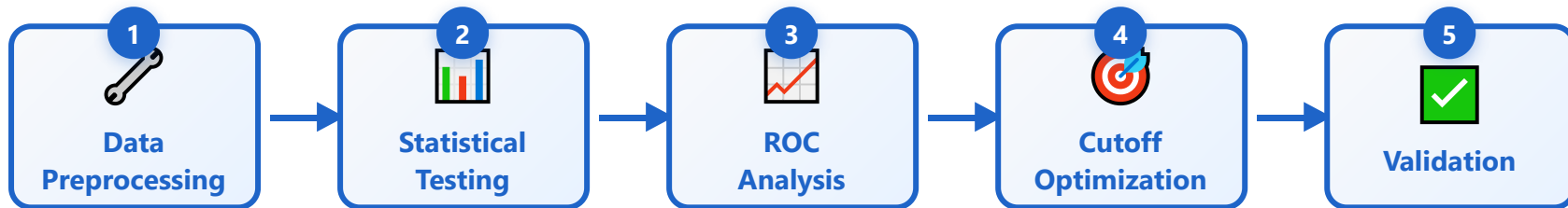Multi-biomarker signatures improve accuracy and clinical utility

⚡ Dynamic Monitoring
Real-time tracking of disease evolution and treatment response

Vision: Seamless integration of multi-modal biomarkers into routine clinical care

💻 Hands-on: Biomarker Analysis Pipeline

1 Data Preprocessing → **2** Statistical Testing → **3** ROC Analysis → **4** Cutoff Optimization → **5** Validation

🔑 Key Steps

- ‣ Normalize and quality control data
- ‣ Feature selection with LASSO/elastic net
- ‣ Statistical comparison (t-test, ANOVA)
- ‣ Multiple testing correction (FDR, Bonferroni)

📏 Performance Metrics

- ‣ AUC-ROC for discrimination ability
- ‣ Sensitivity, specificity, PPV, NPV
- ‣ Cross-validation for robustness
- ‣ Independent cohort validation

💻 Hands-on: Patient Stratification Methods

📊 Clustering Methods

K-means, hierarchical clustering, and consensus clustering to identify patient subgroups based on molecular profiles

📈 Risk Score Development

Cox regression and machine learning models to create prognostic scores for treatment decisions

⏱ Survival Analysis

Kaplan-Meier curves, log-rank tests, and Cox proportional hazards models for outcome prediction

🔄 Treatment Interaction

Subgroup analysis to identify biomarker-treatment interactions and optimal therapy assignment

📉 Visualization

✅ Cross-validation

Heatmaps, forest plots, and interactive dashboards for communicating stratification results

Internal validation with bootstrap and external validation in independent cohorts

🛠️ Tools & Packages

| scikit-learn | lifelines | survminer |
| ComplexHeatmap | ggplot2 | plotly |

Thank You

🎯 **Precision Health**

🛡️ **Prevention Focus**

🌍 **Global Implementation**

🤝 **Equitable Access**

Introduction to Biomedical Data Science