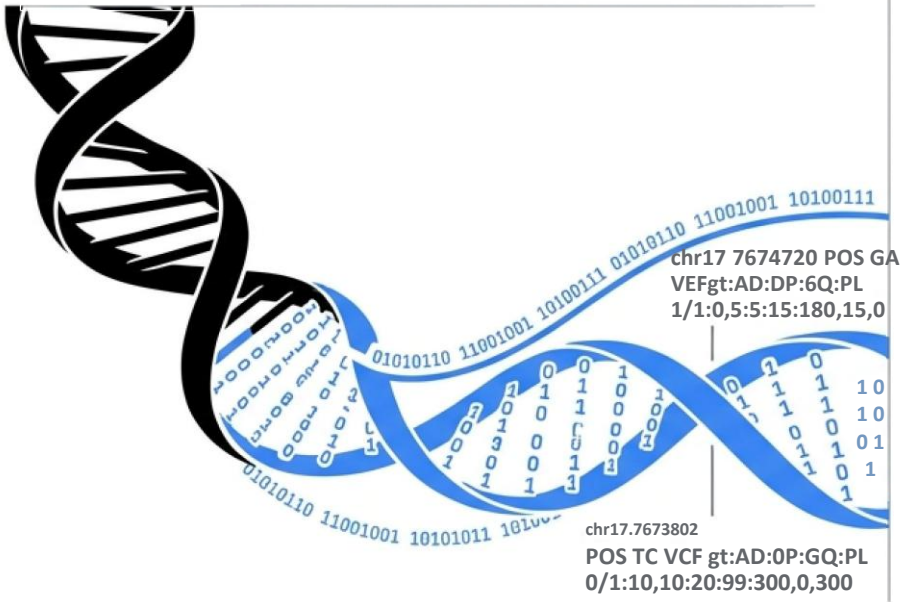
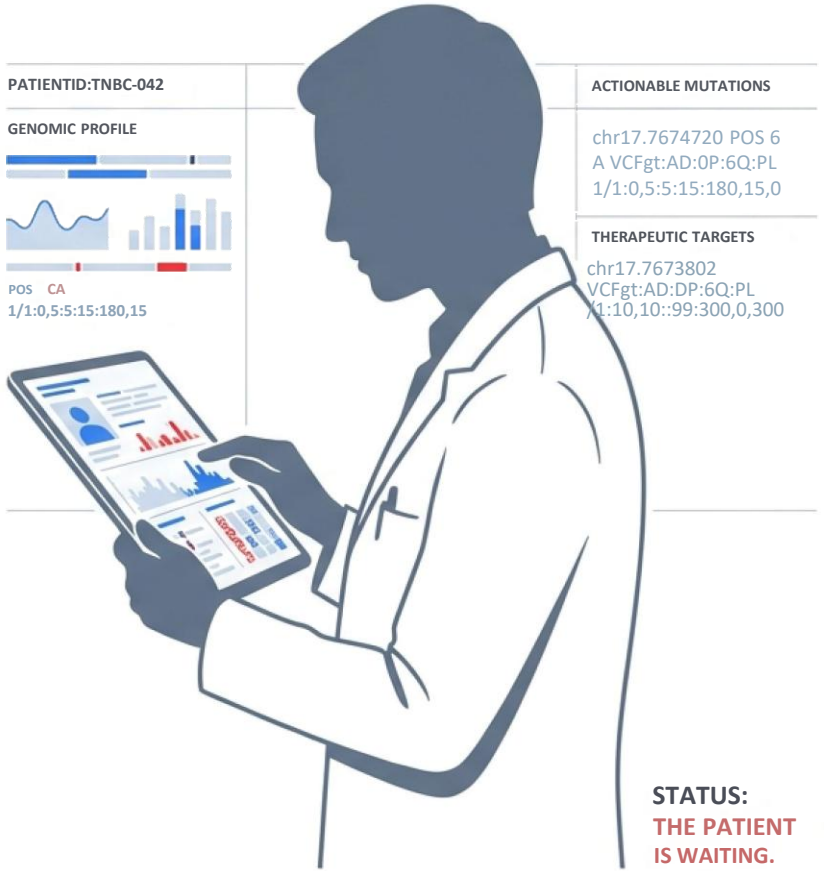
BIOMARKERS

From workbench to clinic

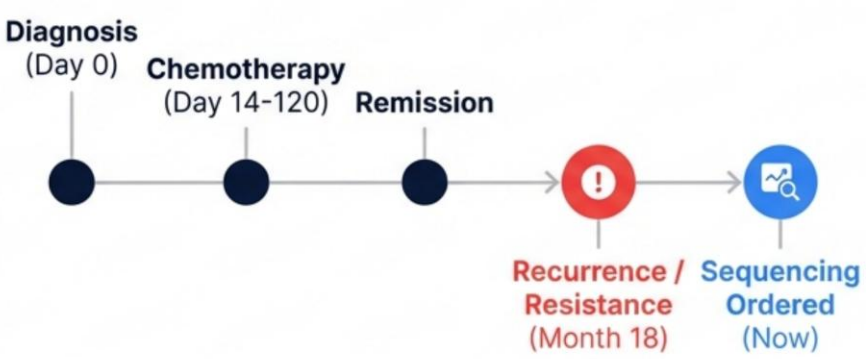
A Workshop on Biomarker Discovery & Analysis
In the setting of Triple Negative Breast Cancer





THE CASE: PATIENT PROFILE

PATIENT ID:	TNBC-43F
AGE:	43 Years
SEX:	Female
DIAGNOSIS:	Triple Negative Breast Cancer (TNBC)
HISTORY:	Initial response to chemotherapy (AC-T regimen). Recurrence detected at 18 months.
CURRENT STATUS:	Non-responsive / Resistant to Standard of Care.
INTERVENTION:	Whole Exome Sequencing (WES) Ordered.



Mission Statement

MISSION: **Analyze sequencing output to identify Biomarkers for targeted therapy.**

BIOLOGICAL SIGNALS TO MEASURABLE FEATURES

What is a Biomarker?

An objectively measured, quantifiable characteristic, such as proteins, DNA, or physiological scans that indicates normal biological processes, pathogenic processes, or responses to therapeutic interventions.

FDA Definition

"A defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention."

Biomarker vs. Clinical Endpoint

Biomarker
(The Signal)

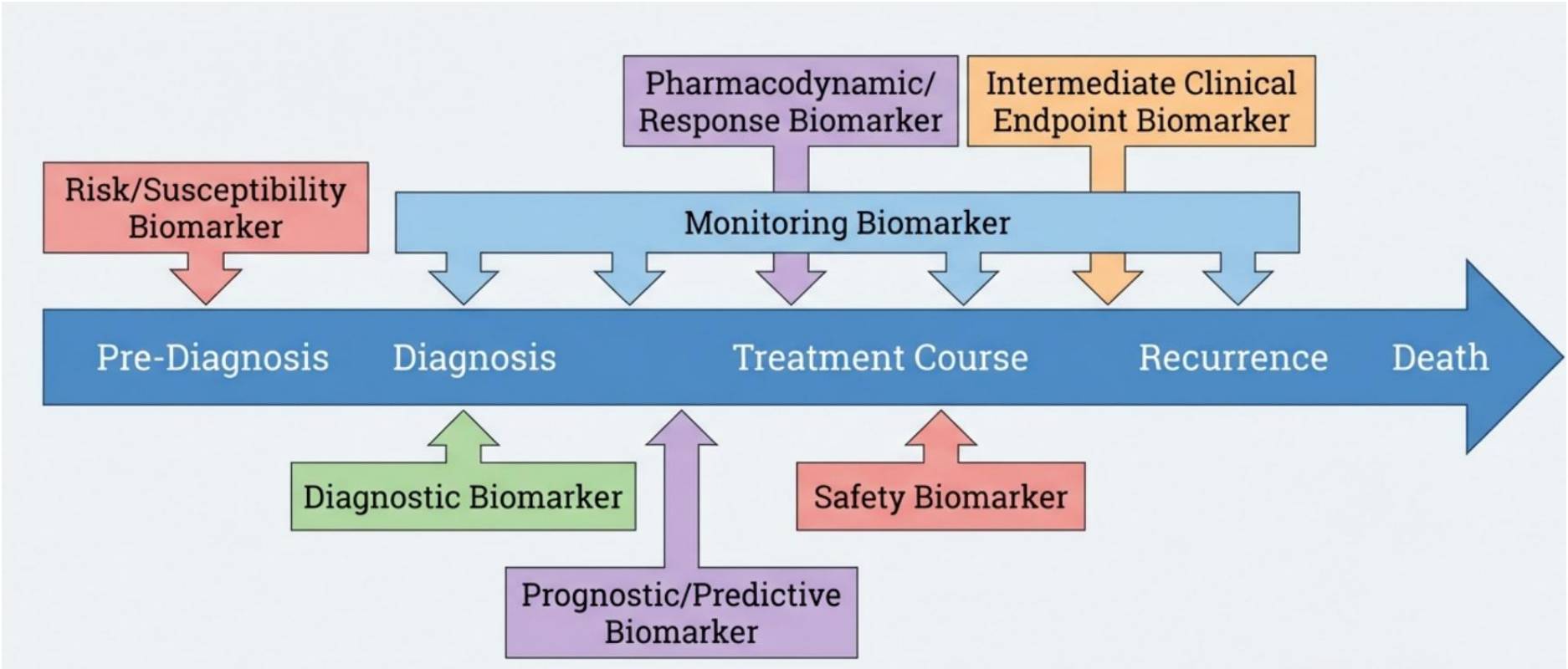
- Molecular alteration (e.g., **Mutation**)
- Physiological measurement (e.g., **Blood pressure**)
- Radiographic sign (e.g., **Tumor shrinkage**)

Biomarkers **PREDICT** or **MONITOR** Endpoints.

Clinical Endpoint
(The Outcome)

- How the patient feels
- Survival (**OS**, **PFS**)
- Functional capacity

BIOMARKERS OVER THE COURSE OF TREATMENT



Features of a Good Biomarker

Stability

Consistent over time.

Reproducibility

Same result in different labs.

Interpretability

Clear clinical meaning.

Cost & Scalability

Feasible for widespread use.

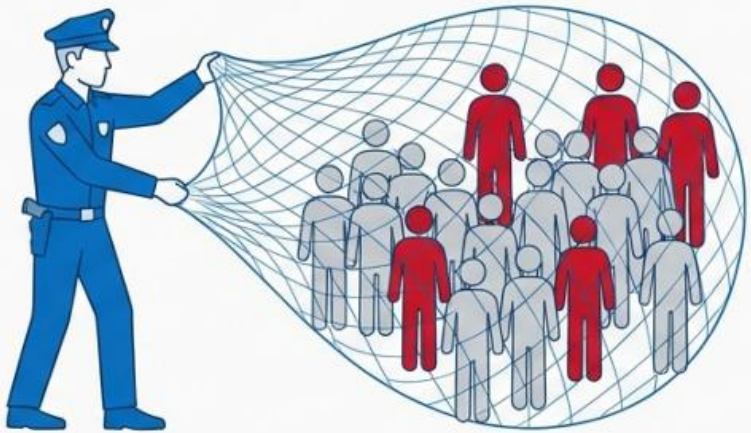
Clinical Utility

Improves patient outcomes.

A robust biomarker must fulfill these criteria to be a valuable tool in clinical decision-making and precision medicine.

Validating the Evidence: Sensitivity vs. Specificity

Sensitivity (The Net)



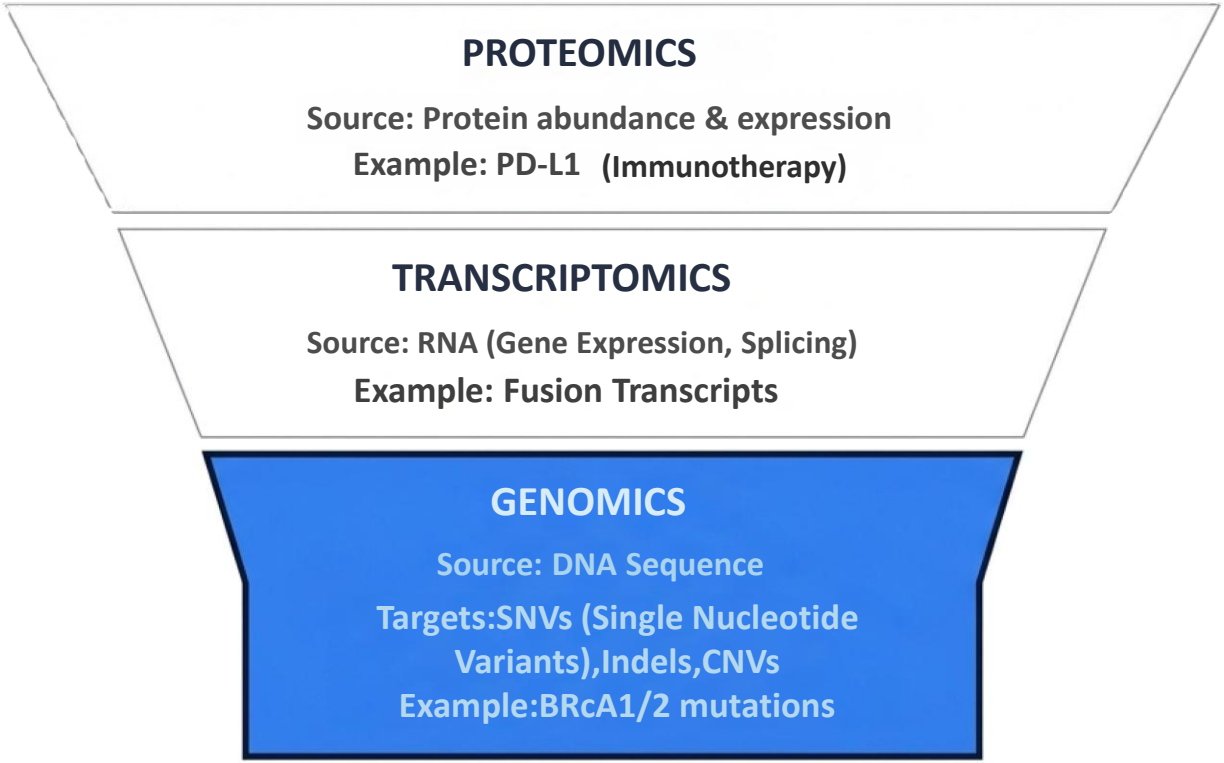
Catches all disease (True Positive), but flags healthy patients too (False Positive).

Specificity (The Filter)

Never flags a healthy patient (True Negative), but might miss some disease (False Negative).

In cancer, we need high sensitivity to detect the tumor, but high specificity to treat it accurately.

SOURCES OF BIOMARKERS: THE OMICS LANDSCAPE

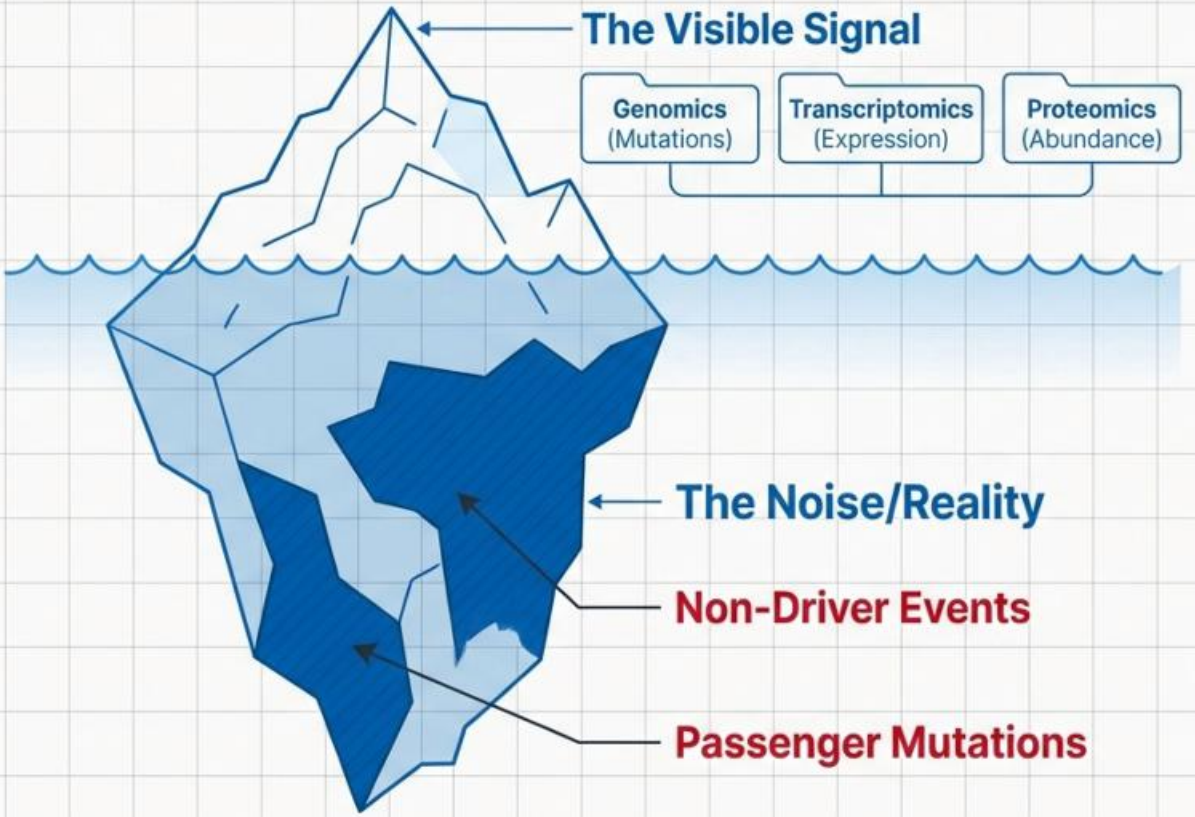


INDUSTRY EXAMPLES

Oncology:
BRCA1/2 (Genomic)

Immunotherapy:
PD-L1 (Proteomic)

Not All Signal is a Biomarker



The Signal Problem:

A tumor may have thousands of mutations. Most are "**passengers**" that **do not drive the cancer** or predict drug response.

Key Rule:

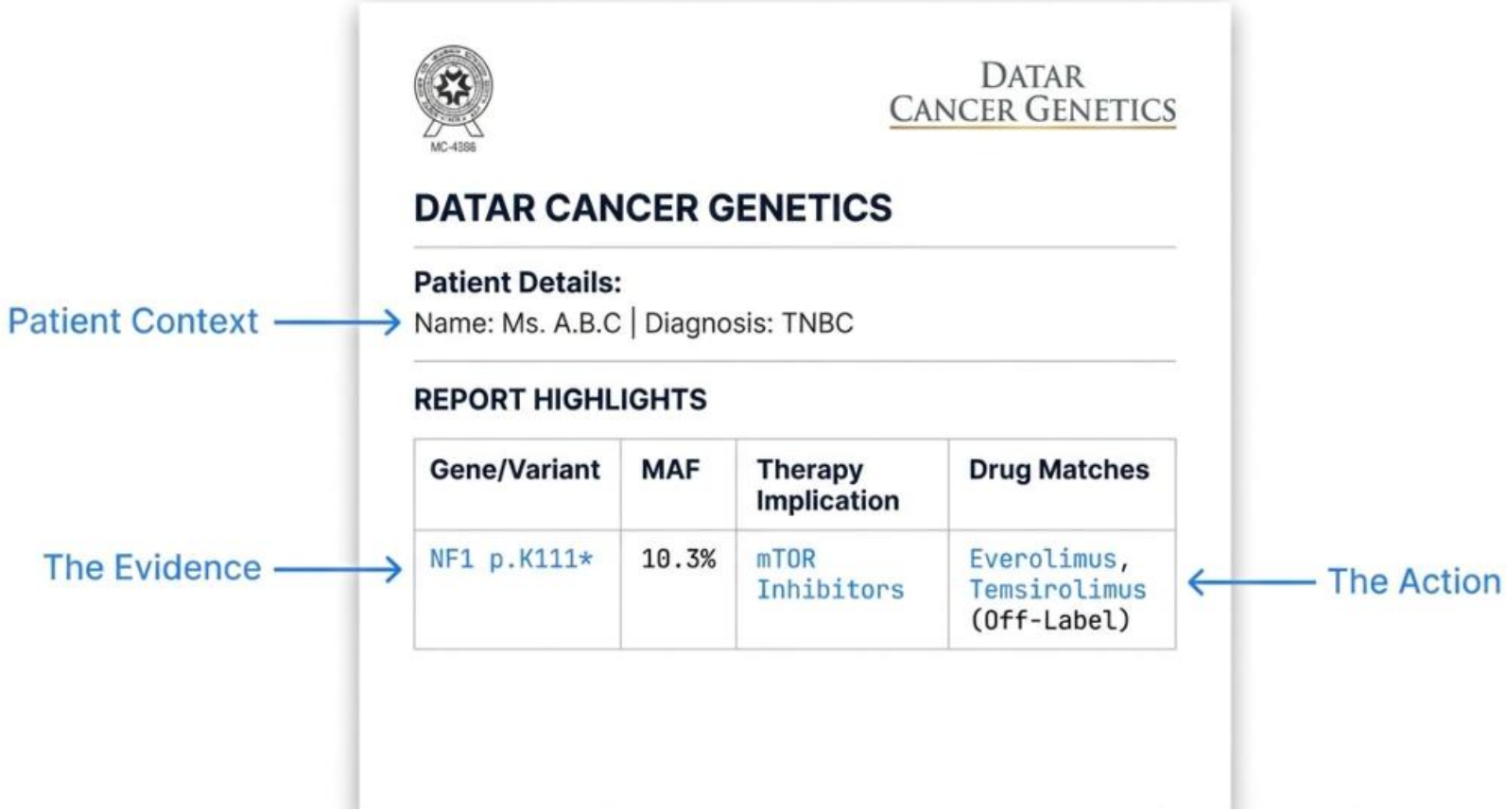
A biomarker must have an **established scientific framework** elucidating its clinical significance. Mere presence is not enough.

THE RULEBOOK: AMP/ASCO/CAP GUIDELINES

Joint Consensus Recommendation for Somatic Variant Interpretation

TIER I: STRONG CLINICAL SIGNIFICANCE	TIER II: POTENTIAL CLINICAL SIGNIFICANCE	TIER III: UNKNOWN SIGNIFICANCE	TIER IV: BENIGN / LIKELY BENIGN
Therapeutic, prognostic & diagnostic	Therapeutic, prognostic & diagnostic	**VUS:** Not observed in population, no convincing cancer association.	Observed at high frequency in general population.
Level A Evidence **FDA-approved therapy** Included in **professional guidelines**	**Level C Evidence** FDA-approved therapies for different tumor types or investigational therapies Multiple small published studies with some consensus	Not observed at a significant **allele frequency** in the general or specific specific subpopulation **databases**, or pan-cancer or tumor-specific variant databases No convincing published evidence of cancer association	Observed at significant **allele frequency** in the general or specific subpopulation **databases** No existing published evidence of cancer association
Level B Evidence Well-powered studies with consensus from experts in the field	**Level D Evidence** Preclinical trials or a few case reports without consensus		

THE DESTINATION: THE CLINICAL REPORT



THE EVIDENCE: RAW VCF FILES

```
##fileformat=VCFv4.2
##source=SelectVariants
##INFO=<ID=AF,Number=A,Type=Float,Description=\"Allele Frequency">
##INFO=<ID=DP,Number=1,Type=Integer,Description="Total Depth">
##INFO=<ID=GENE,Number=1,Type=String,Description="Gene Name">
#CHROM POS     ID    REF   ALT   QUAL   FILTER INFO
1      13417   .     G     A   .  .      PASS   AF=0.001;DP=150
1      13420   .     C     T   .  PASS   PASS   AF=0.005;DP=160
1      13435   .     A     C   .  PASS   PASS   AF=0.002;DP=155
17     7577120 .     C     T   .  PASS   PASS   AF=0.25;GENE=TP53
17     7578432 .     A     G   .  PASS   PASS   AF=0.30;GENE=TP53
17     7579890 .     T     C   .  PASS   PASS   AF=0.28;GENE=TP53
13     329144  .     T     G   .  PASS   PASS   AF=0.60;GENE=BRCA2
13     329340  .     G     C   .  PASS   PASS   AF=0.62;GENE=BRCA2
13     329567  .     A     T   .  PASS   PASS   AF=0.59;GENE=BRCA2
1      13500   .     G     A   .  PASS   AF=0.001;DP=150
17     7577200 .     C     T   .  PASS    AF=0.25;GENE=TP53
13     329200  .     T     G   .  PASS    AF=0.60;GENE=BRCA2
... (thousands more lines)
```

Meta-information

Header

Data Lines

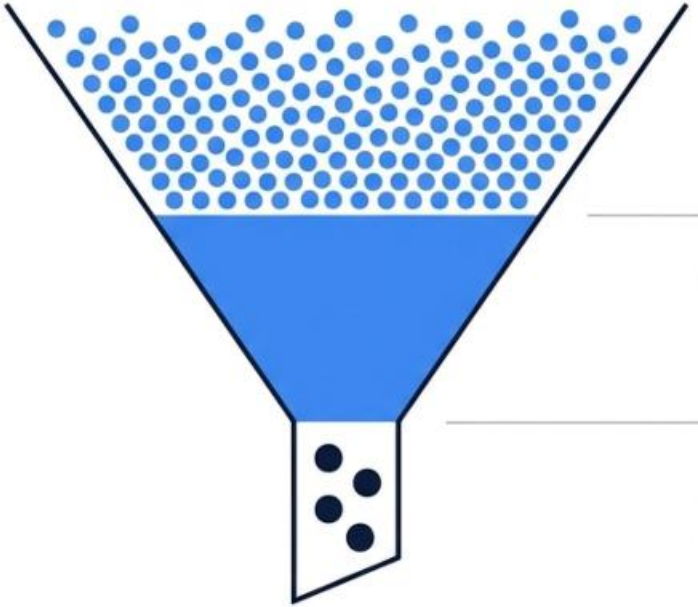
The Haystack: A single VCF contains thousands of variants.

Most are noise. Most are benign. We must filter to find the cure.

INVESTIGATION STEP 1: POPULATION FILTERING

Databases of Normality
Inter

1. 1000 Genomes Project
2. ExAC (Exome Aggregation Consortium)
3. dbSNP



Input: ~25,000 Variants

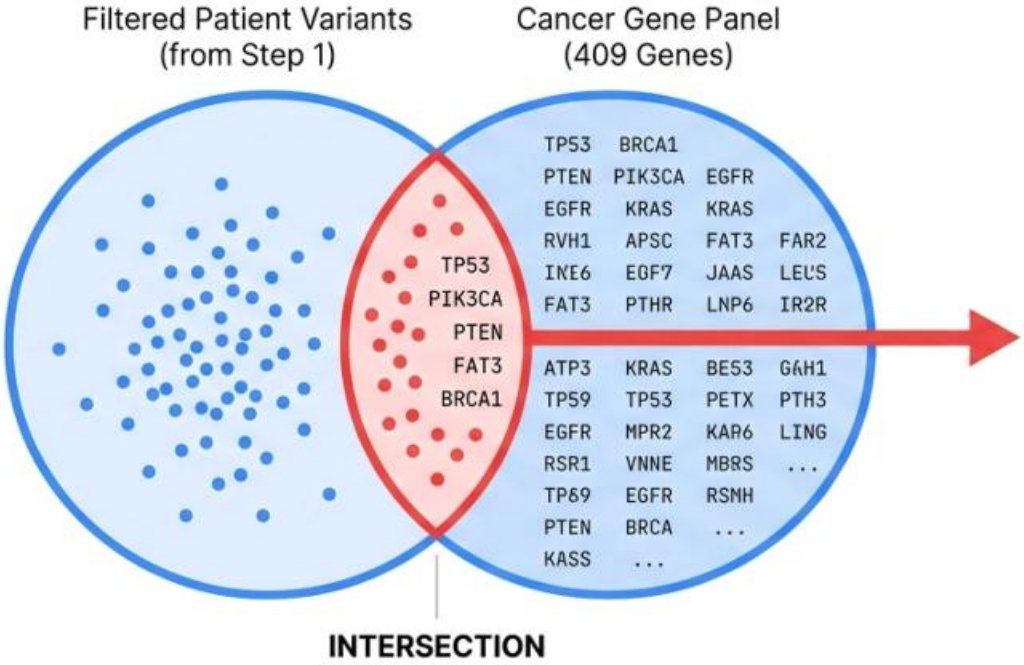
Filter: Common Polymorphisms (>1% Frequency)

Output: ~400 Variants

LOGIC: If a mutation is in everyone, it is not the cause of this rare cancer.

AMP Guidelines: Tier IV (Benign).

INVESTIGATION STEP 2: RELEVANCE FILTERING

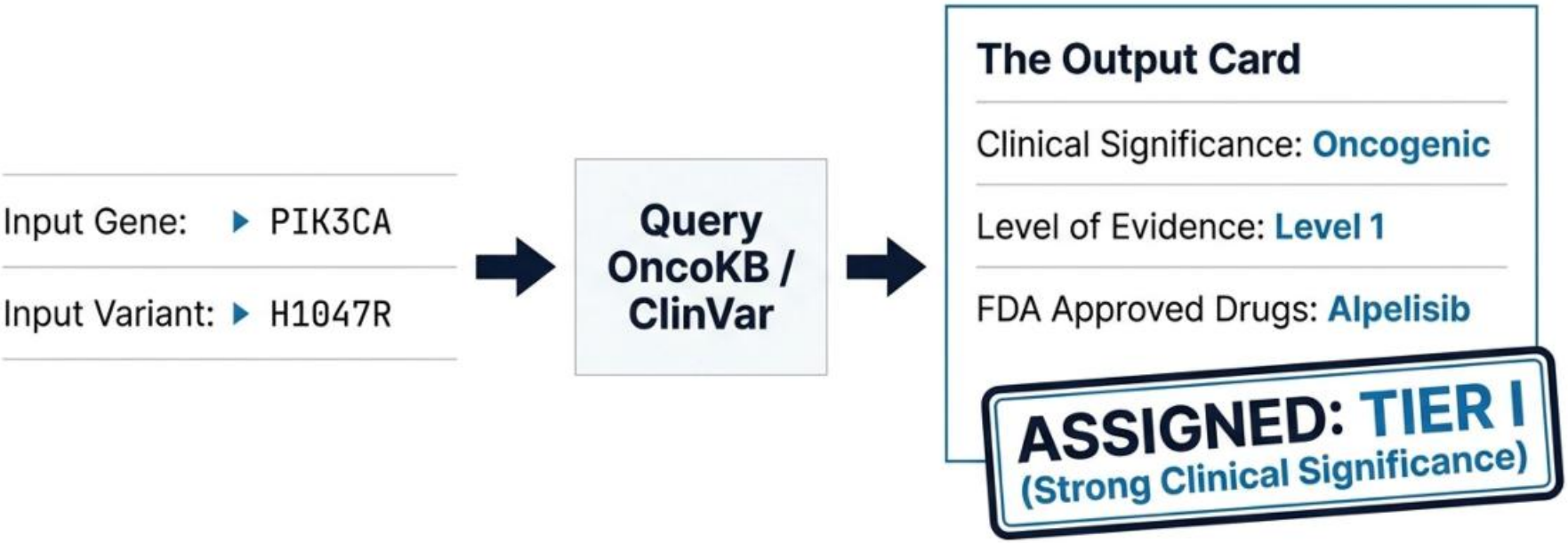


TARGET GENES:

- TP53 (Guardian of the Genome)
- PIK3CA (PI3K/AKT Pathway)
- PTEN (Tumor Suppressor)
- FAT3 (Prognostic Marker in TNBC)
- BRCA1/2

We are left with a handful of SOMATIC variants to investigate.

INVESTIGATION STEP 3: WEIGHING THE EVIDENCE



We repeat this for every remaining variant to build the treatment table.

AN EMERGING BIOMARKER: TMB

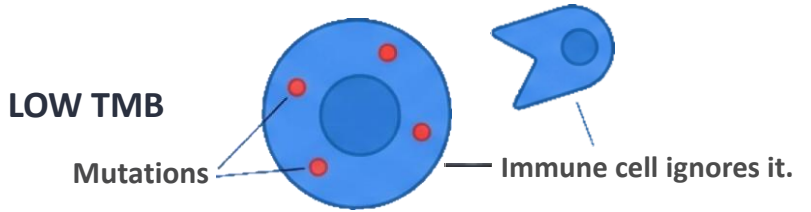
Tumor Mutational Burden

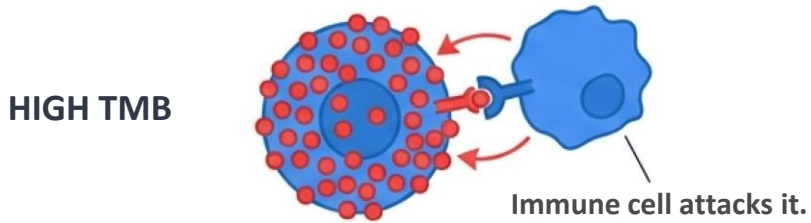
DEFINITION:

Total number of somatic coding mutations per megabase (muts/Mb) of tumor DNA.

THE LOGIC:

More Mutations = More Neoantigens = More visible to Immune System.

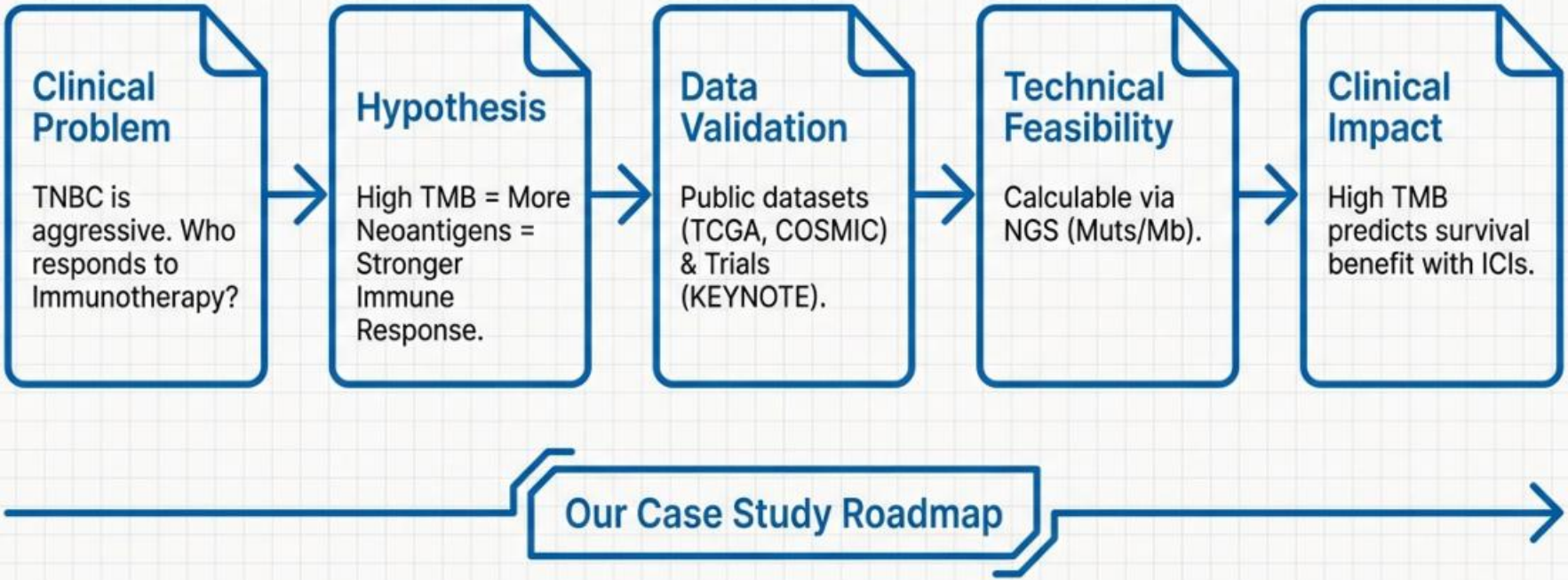




High TMB predicts response to Checkpoint Inhibitors (Immunotherapy).

DRUGS:Pembrolizumab,Nivolumab

Pipeline to Discovery: TMB in Triple Negative Breast Cancer



Calculating the Biomarker: Tumor Mutational Burden (TMB)

$$\text{TMB} = \frac{(\text{Total \# of Somatic, Non-Synonymous Mutations})}{(\text{Size of Sequenced Coding Area in Megabases})}$$

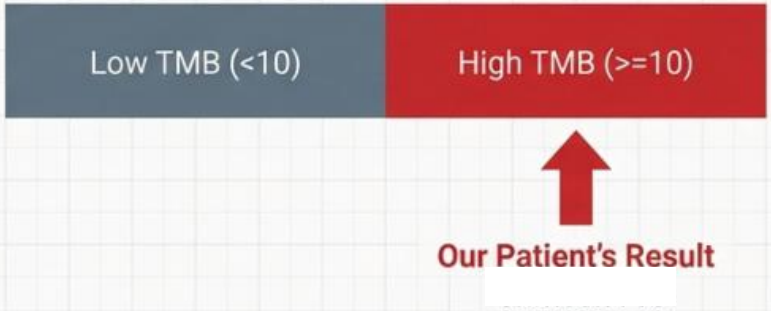
Calculation Example

Step 1
Count valid variants in the VCF (e.g., **450 mutations**).

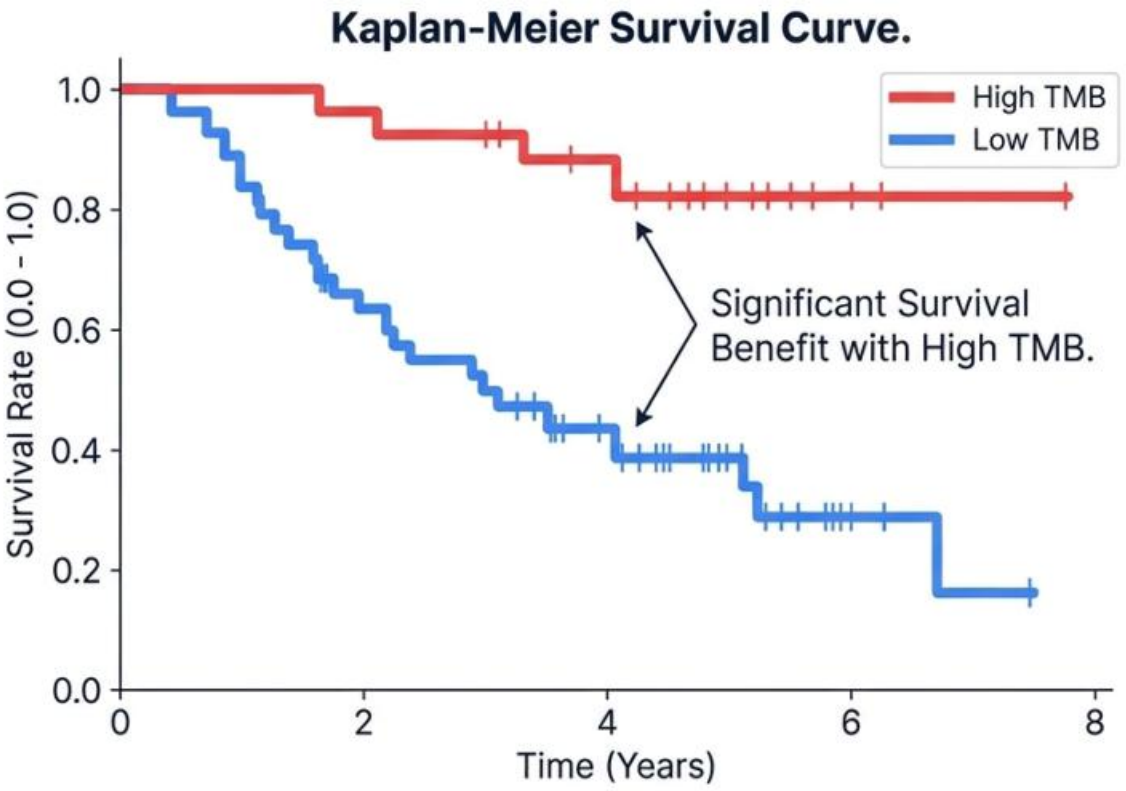
Step 2
Divide by Exome Size (e.g., 30 Mb).

Calculation
450 / 30 = 15 Mutations/Mb.

Result Scale



CASE INSIGHT: TMB & SURVIVAL IN TNBC



Gene Insight: FAT3

- FAT3 mutations are associated with altered immune infiltration.
- Presence of **FAT3 mutation + High TMB = Strong candidate** for **Immunotherapy**.