Optum

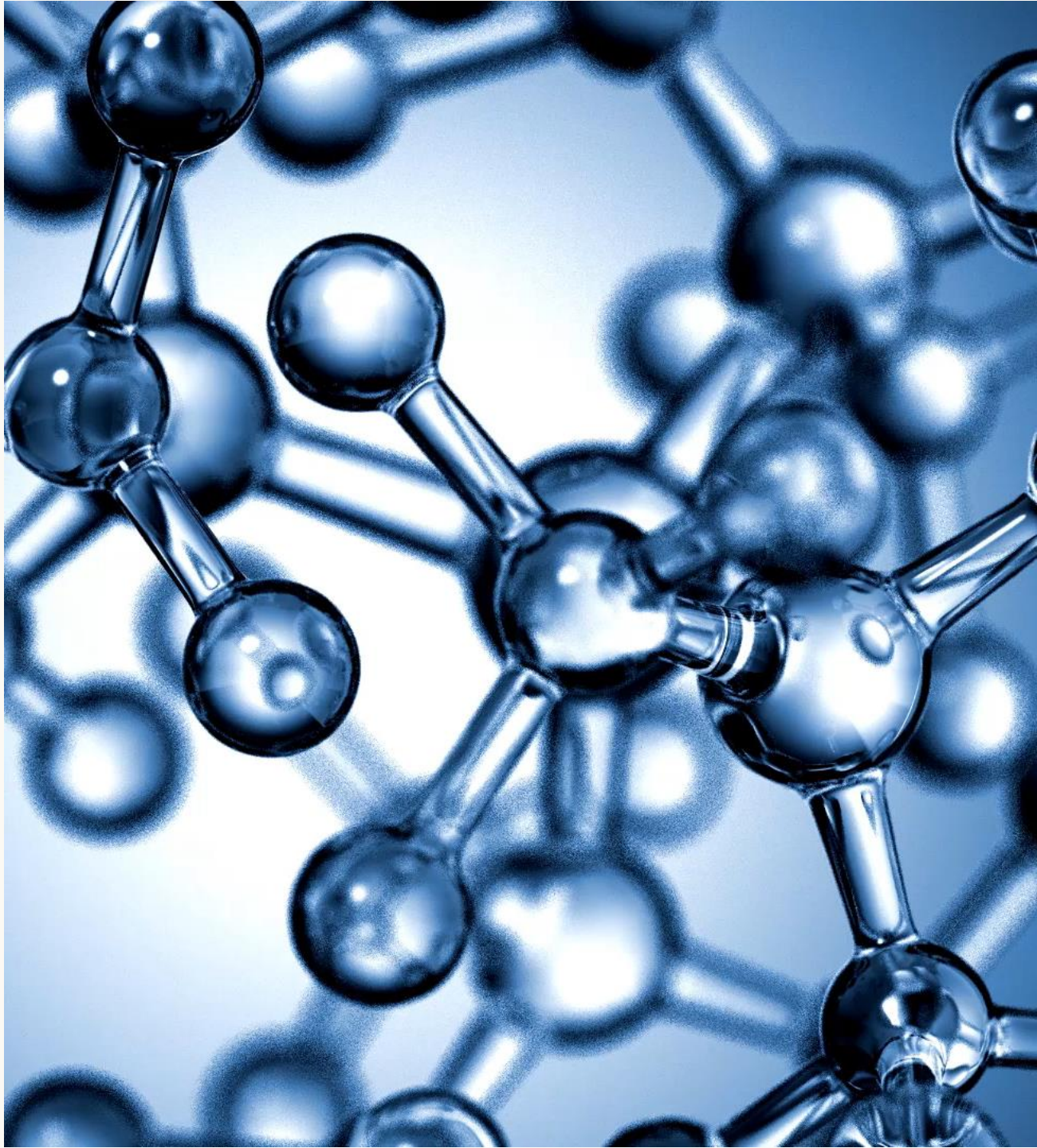
What Does the Future Look Like for Big Data and Precision Medicine?

Amy Sainski-Nguyen, PhD, Optum Life Sciences, Eden Prairie, MN, US

Ashley Brenton, PhD, Optum Life Sciences, Eden Prairie, MN, US

Jennifer Webster, MS, Pfizer, New York, NY, US

ISPOR Europe 2022
6-9 November 2022
Vienna, Austria and Virtual



Disclaimer

The views and opinions expressed in the following PowerPoint slides are those of the individual presenter and should not be attributed to Optum, Pfizer, ISPOR, or their directors, officers, employees, volunteers, members, chapters, councils, communities or affiliates, or any organization with which the presenter is employed or affiliated.

These PowerPoint slides are the intellectual property of the presenters and are protected under the copyright laws of the United States of America and other countries. Used by permission. All rights reserved. All trademarks are the property of their respective owners

© 2022 Optum, Inc. All rights reserved.

Agenda

History of precision medicine

Real-world genomic data sources

Genomic data in the drug discovery and development life cycle

RWD examples and best practices

Discussion

© 2022 Optum, Inc. All rights reserved.

Precision medicine, also known as "personalized medicine", is tailoring disease prevention and treatment to an individual's genes, environment, and lifestyle.[1]

[1] MedlinePlus [Internet]. Bethesda (MD): National Library of Medicine (US); Precision medicine; [updated 2022 May 17; cited 2022 Sept 22]; Available from: https://medlineplus.gov/genetics/understanding/precisionmedicine/

Optum

© 2022 Optum, Inc. All rights reserved.

Case Study: TMPT testing for 6-MP treatment in children with acute lymphoblastic leukemia

1970's: Treatment of pediatric ALL patients with thiopurine drugs results in adverse drug reactions including hematopoietic toxicity and life-threatening myelosuppression.[2]

1980's: TPMT is shown to be genetically regulated and poor activity results in cytoxicity.[2]

1990's: A blood test measuring TPMT activity it developed and 11% of individuals have intermediate levels and 1 in 300 have undetectable levels of activity.[3,4]

2002: First FDA change in label drawing attention to the genomics aspect[5,6]

2019: There are 64 anticancer therapies targeted against molecular alterations.[7]

- Of these, the number of targetable molecular alterations in 24
- In 19 of these the detection of the alteration was required to effectively indicate a specific prescription

[2] Weinshilboum RM, Sladek SL. Mercaptopurine pharmacogenetics: Monogenic inheritance of erythrocyte thiopurine methyltransferase activity. *Am J Hum Genet*. 1980;32:651-662.

[3] Szumlanski C, Otterness D, Her C et al. Thiopurine methyltransferase pharmacogenetics: Human gene cloning and characterization of a common genetic polymorphism. *DNA and Cell Biology*. 1997;15:17-30.

[4] Coulthard S, Howell C, Robson, J, et al. The relationship between thiopurine methyltransferase activity and genotype in blasts from patients with acute leukemia. *Blood.* 1998;92(8):2856-2862.

[5] FDA. Table of Pharmacogenomic Biomarkers in Drug Labeling with Labeling Text. Accessed 26 Sept 2022. https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm?event=overview.process&varApplNo=09053

[6] Lesko JL, Zineh I. DNA, drugs and chariots: On a decade of pharmacogenomics at the US FDA. *Pharmacogen. 2010;11(4):10-16.*

[7] Colomer R, Mondejar R, Romero-Laorden N, Alfranca A, Sanchez-Madrid F, Quintela-Fandino M. When should we order a next generation sequencing test in a patient with cancer? *E Clinical Med*. 2020;25:100487. doi: 10.1016/j.eclinm.2020.100487.

Optum

© 2022 Optum, Inc. All rights reserved.

Growth Contributions

Technology

Regulatory and advisory groups

Costs

Reimbursement policies

© 2022 Optum, Inc. All rights reserved.

Reasons for growth

Technology

- Immunohistochemistry, and fluorescence in situ hybridization (FISH), and polymerase chain reaction (PCR) are three conventional and commonly used molecular diagnostic techniques[8]
- Next Generation Sequencing (NGS), which can simultaneously analyze a broad spectrum of genomic alterations, is replacing these methods[9]
  - "Hot spot" panels that target identified gene alterations (e.g. 10-50 genes)
  - Broad NGS panels can characterize the tumor genetic profile and can include several hundred genes

Regulator and advisory consortiums recommendations and approvals

- The proportion of new drug approvals with pharmacogenomics labeling has increased from 10.3% in 2000 to 28.2% in 2020[10]
- This includes all clinical areas, but especially cancer therapies which comprise the largest proportion of biomarker-drug pairs that require pharmacogenomics testing[10]
- The US FDA maintains tables which outline important biomarker-drug interactions.[11]

[8] Malone ER, Oliva M, Sabatini PJB, Stockley TL, Siu LL. Molecular profiling for precision cancer therapies. Genome Med. 2020;12(1):8.

[9] Vanderwalde AM, Ma E, Yu E, et al. NGS testing patterns in advanced non-small cell lung cancer (aNSCLC) and metastatic breast cancer (mBC): OneOncology (OO) sites compared to Flatiron Health Nationwide (NAT). Journal of Clinical Oncology 2021;39:28_suppl, 288-288

[10] Kim JA, Ceccarelli R, Lu CY. Pharmacogenomic biomarkers in US FDA-approved drug labels (2000-2020). J Pers Med. 2021;11(3):179

[11] US FDA. Table of Pharmacogenetic Associations. Accessed 23 Sept 2022. https://www.fda.gov/medical-devices/precision-medicine/table-pharmacogenetic-associations#section2

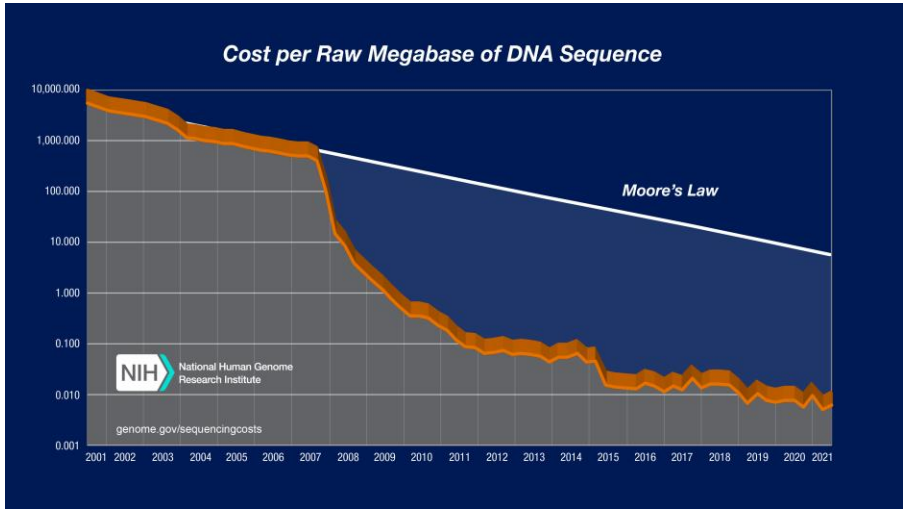
Optum

© 2022 Optum, Inc. All rights reserved.

Reasons for growth

Decreased costs

- In 2006 it cost $20-25 million[12,13] to sequence the genome, with NGS there are **hopes of the $1,000 genome**.
- A review[14] of economic evaluations of costs in 2018:
  - £382 ($555) to £3592 ($5169) for **exome sequencing**
  - £1312 ($1906) to £17,243 ($24,810) for **genome sequencing**



Changes in insurance reimbursement policies[15]

- Medicare and several commercial insurers cover NGS among certain groups of patients
  - Since 2015 there have been changes in coverage decision for sequencing with the largest increment being in 2018
- Variability exists in coverage

[12] National Human Genome Research Institute. The cost of sequencing a human genome. 2016. https://www.genome.gov/27565109/the-cost-of-sequencing-a-human-genome/. Accessed 23 Sept 2022.

[13] Schwarze K, Buchanan J, Fermont JM, et al. The complete costs of genome sequencing: A microcosting study in cancer and rare diseases from a single center in the United Kingdom. Genet Med. 2019;22:85-94.

[14] Schwarze K, Buchanan J, Taylor JC, Wordsworth S. Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature. Genet Med. 2018;20:1122–1130.

[15] Trosman JR, Douglas MP, Liang SY, et al. Insights from a temporal assessment of increases in US private payer coverage of tumor sequencing from 2015-2019. Value Health. 2020;23(5):551-558.

Optum

© 2022 Optum, Inc. All rights reserved.

Challenges

Knowledge

Costs

Time to report of results

Ethical and social challenges

Optum

© 2022 Optum, Inc. All rights reserved.

Challenges: Knowledge

Limitations exist in the knowledge of patients, providers, and payers

- Specific groups of oncologists that may be **more likely** to test patients
  - A 2017 study identified specific characteristics related to using tests to direct therapy: age (< 50 years) and having a faculty appointment, genomics training, seeing more than 50 unique patients monthly, and having access to a molecular tumor board.[16, 17]

- Although **85%** of cancer care is conducted in a community setting, NGS testing and sequencing-directed treatment are generally mostly available in academic centers which are more likely to have in-house testing and clinical trials.[18]

- Health care professionals who do not have genomics training **may** lack the confidence in interpreting results and making clinical decisions from these results.[19]

[16] Freedman AN, Klabunde CN, Wiant K, et al. Use of Next-Generation Sequencing Tests to Guide Cancer Treatment: Results From a Nationally Representative Survey of Oncologists in the United States. JCO Precis Oncol. 2018;2:1-13.

[17] Boehmer L, Roy UKB, Schrag J, et al. Identifying barriers to equitable biomarker testing in underserved patients with NSCLC: A mixed-methods study to inform quality improvement opportunities. Journal of Clinical Oncology 2021;39:28_suppl, 123-123

[18] Tucker TC, Charlton ME, Schroeder MC, et al. Improving the Quality of Cancer Care in Community Hospitals. Ann Surg Oncol. 2021;(2):632-638.

[19] Gray SW, Hicks-Courant K, Cronin A, Rollins BJ, Weeks JC. Physicians' attitudes about multiplex tumor genomic testing. J Clin Oncol. 2014;32(13):1317-23.

Optum

© 2022 Optum, Inc. All rights reserved.

Challenges: Costs

While the consumable costs have been decreasing with the advent of NGS, other costs are increasing or remain high.

- While there are hopes of a "$1000 genome", this is only the consumable component and **doesn't consider overall costs** like sequencing, bioinformatic processing, analysis, interpreting/reporting, and data storage.[13]
  - **"the $1000 genome and $100,000 analysis."**[20]

- Despite decreasing costs, **out-of-pocket costs** can be prohibitively high, even for those who are insured.[21]

[13] Schwarze K, Buchanan J, Fermont JM, et al. The complete costs of genome sequencing: A microcosting study in cancer and rare diseases from a single center in the United Kingdom. Genet Med. 2019;22:85-94.

[20] Mardis ER. The $1,000 genome, the $100,000 analysis? Genome Med. 2010;2:84.

[21] Yabroff KR, Zhao J, de Moor JS, Sineshaw HM, Freedman AN, Zheng Z, Han X, Rai A, Klabunde CN. Factors Associated With Oncologist Discussions of the Costs of Genomic Testing and Related Treatments. J Natl Cancer Inst. 2020;112(5):498-506

Optum

© 2022 Optum, Inc. All rights reserved.

Challenges: Time to report of results

The time it takes to receive results may cause anxiety in patients and providers

- Depending on the type of test and where it is done, it can take **4 or more weeks** to receive results.[22]
- There is always a possibility that actionable mutations **may not be found** or a re-biopsy could be needed.
- For patients with advanced and metastatic disease where extremely timely treatment is critical to maximize patient survival, physicians may elect to start patients on other non-targeted therapies **rather than wait** – especially if they believe that the actionable mutation is rare.[22]

22 Mileham KF, Roy UKB, Bruinooge SS, et al. Physician concern about delaying lung cancer treatment while awaiting biomarker testing: Results of a survey of U.S. oncologists. J Clin Oncol. 2021;39(suppl 15):9067. doi:10.1200/JCO.2021.39.15_sup-pl.9067

Optum

© 2022 Optum, Inc. All rights reserved.

Challenges: Ethical and Social

Providers may be less likely recommend biomarker testing in certain types of patients

- In a mixed-methods study of academic and community-based clinicians, respondents acknowledged offering biomarker testing less frequently to certain patients:
  - **57%** reported being extremely likely to order a test for a patient of high SES **vs. 32%** being extremely likely for a patient with a low SES or who was homeless.[23]

- Differential testing has been documented across other socioeconomic characteristics:
  - By **race**: lower rates in Black patients[23-25]
  - By **age**: lower rates in older patients[24]
  - By **insurance**: lower rates among uninsured patients[24]
- Patients who are not tested are **less likely to receive targeted treatment**.[24]

[23] Bruno DS, Hess LM, Li X, Su EW., Zhu YE, and Patel M. Racial disparities in biomarker testing and clinical trial enrollment in non-small cell lung cancer (NSCLC) Journal of Clinical Oncology 2021;39:15_suppl, 9005-9005

[24] Lamba N, Iorgulescu B. Disparities in microsatellite instability/mismatch repair biomarker testing for patients with advanced colorectal cancer. AACR; Cancer Epidemiol Biomarkers Prev 2020;29(12 Suppl): PO-091.

[25] Reid S, Cadiz S, Pal T. Disparities in Genetic Testing and Care among Black women with Hereditary Breast Cancer. Curr Breast Cancer Rep. 2020;12(3):125-131

Optum

© 2022 Optum, Inc. All rights reserved.

Genomic Data Sources

EMR

Molecular Testing

Direct to Consumer

© 2022 Optum, Inc. All rights reserved.

Genomic Data Sources: EMR

- Dependent on provider notation
- Often binary result--difficult to interpret
- Unstructured data
- Incomplete

© 2022 Optum, Inc. All rights reserved.

Genomic Data Sources: Molecular Testing Results

Broad vs Narrow Panel

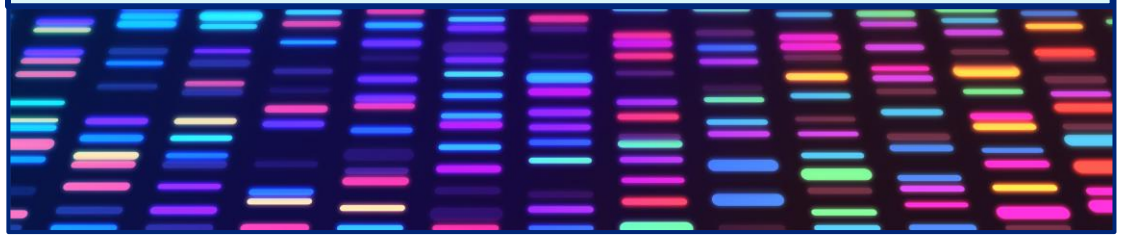
- Definition?

What test was ordered?

- Where is that information found? CPT? EMR note?
- Actual test results

Clinical test result vs RUO?

- Regulatory guidance



What does a negative result mean?

- Definition?

Integration into larger dataset

Consent for downstream uses?

Optum

© 2022 Optum, Inc. All rights reserved.

Genomic Data Sources: DTC

It is estimated that over 100 million individuals have undergone direct-to-consumer genetic testing by end of 2021[1]

How to Protect Your DNA Data Before and After Taking an at-Home Test
June 12, 2019

Your DNA Profile is Private? A Florida Judge Just Said Otherwise
Nov. 5, 2019

Sooner or Later Your Cousin's DNA Is Going to Solve a Murder
April 25, 2019

With a Simple DNA Test, Family Histories Are Rewritten
Aug. 28, 2017

Sigrid Johnson Was Black. A DNA Test Said She Wasn't.
Nov. 19, 2018

ethics

quality

depth

utility

consent

[1] AMA Board of Trustees, Nov 2021 Meeting

[2]New York Times

Optum

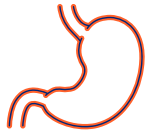
© 2022 Optum, Inc. All rights reserved.

Other -omics

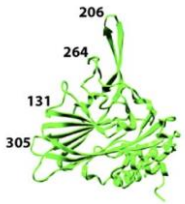
Transcriptomics

Epigenomics

Metabolomics



Microbiomics



Proteomics

Optum

© 2022 Optum, Inc. All rights reserved.

Other –omics: Transcriptomics

The full range of mRNA

ORIGINAL PAPER

Aging Cell WILEY

BiT age: A transcriptome-based aging clock near the theoretical limit of accuracy

David H. Meyer[1,2] | Björn Schumacher[1,2]

nature COMMUNICATIONS

ARTICLE

DOI: 10.1038/s41467-018-06500-x OPEN

DRUG-seq for miniaturized high-throughput transcriptome profiling in drug discovery

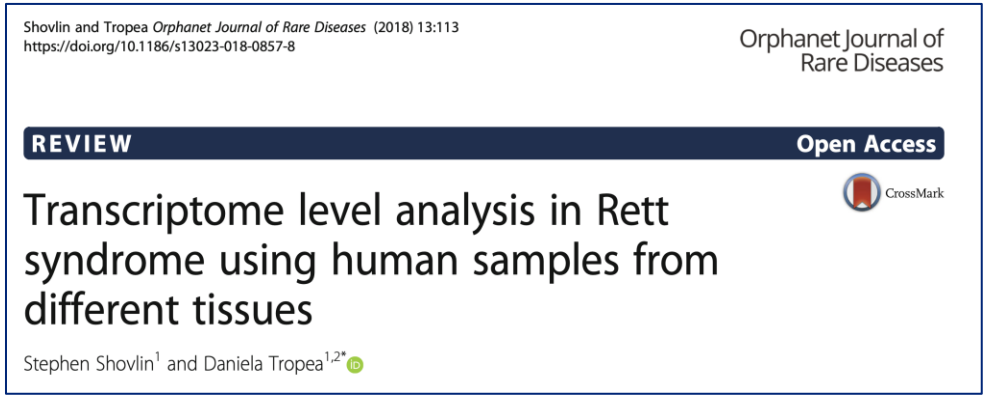
Chaoyang Ye[1,7], Daniel J. Ho[1], Marilisa Neri[2], Chian Yang[3], Tripti Kulkarni[4], Ranjit Randhawa[1], Martin Henault[3], Nadezda Mostacci[2], Pierre Farmer[2], Steffen Renner[2], Robert Ihry[1], Leandra Mansur[5], Caroline Gubser Keller[2], Gregory McAllister[6], Marc Hild[3], Jeremy Jenkins[6] & Ajamete Kaykas[1]

Annals of neurology

Author Manuscript HHS Public Access

A transcriptome-based drug discovery paradigm for neurodevelopmental disorders

Ryan S. Dhindsa, Ph.D, Anthony W. Zoghbi, M.D., [...], and David B. Goldstein, Ph.D



Shovlin and Tropea *Orphanet Journal of Rare Diseases* (2018) 13:113
https://doi.org/10.1186/s13023-018-0857-8

Orphanet Journal of Rare Diseases

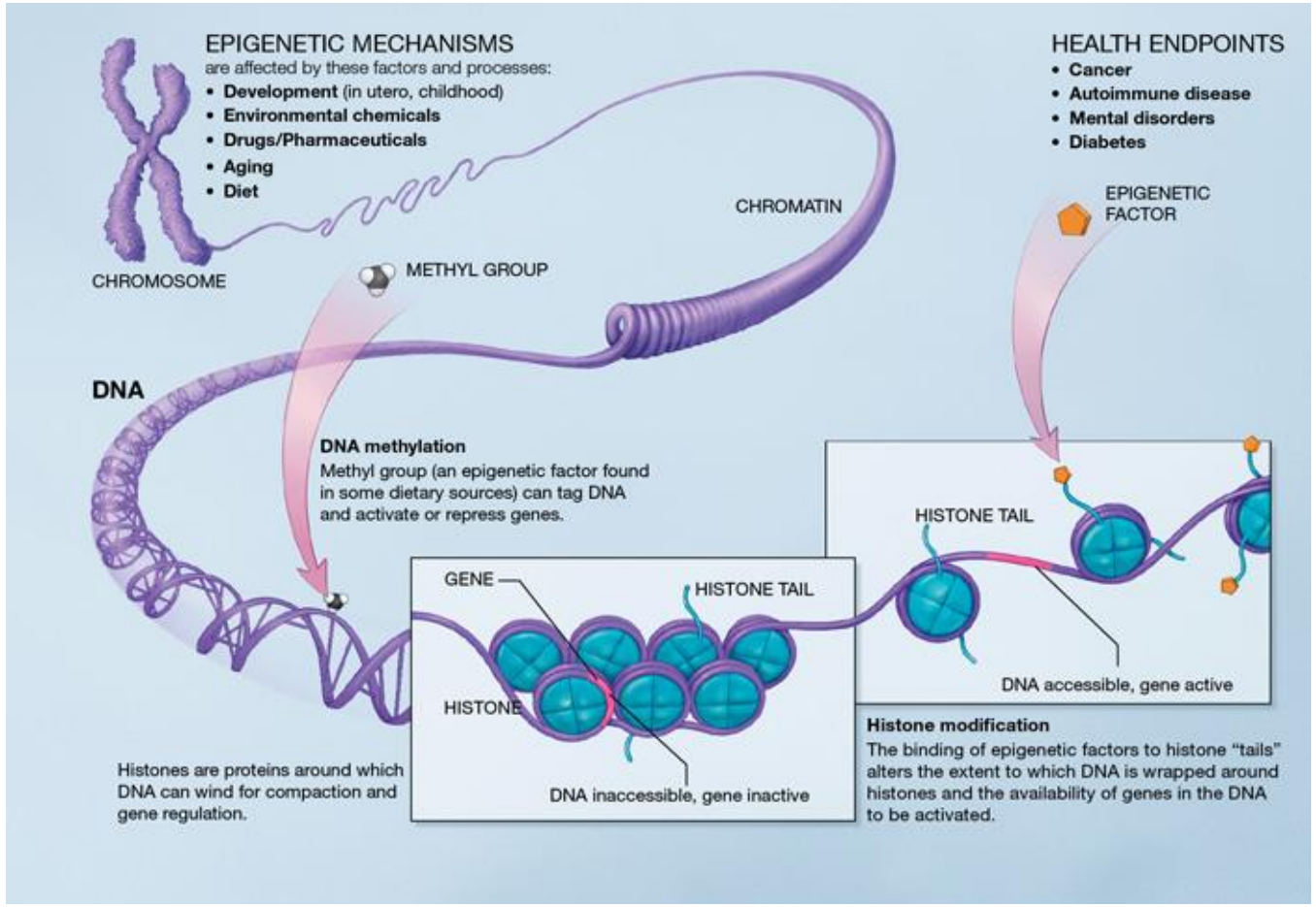
REVIEW Open Access

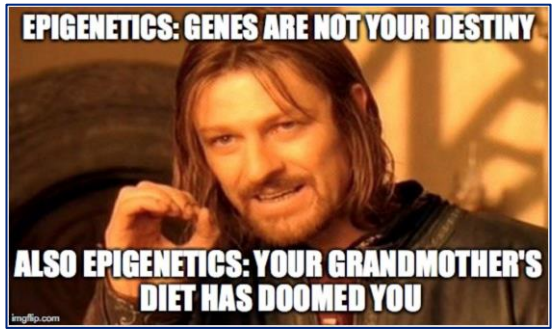
Transcriptome level analysis in Rett syndrome using human samples from different tissues

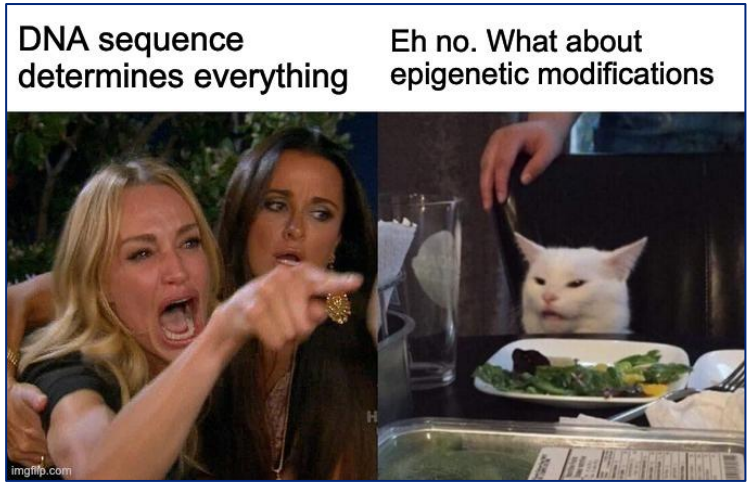
Stephen Shovlin[1] and Daniela Tropea[1,2*]

© 2022 Optum, Inc. All rights reserved.

Other –omics: Epigenomics

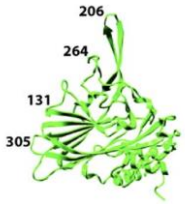


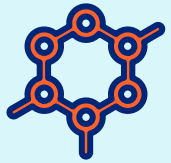




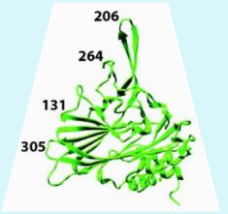
© 2022 Optum, Inc. All rights reserved.

Other –omics: Metabolomics and proteomics





Large-scale study of small molecules (metabolites)



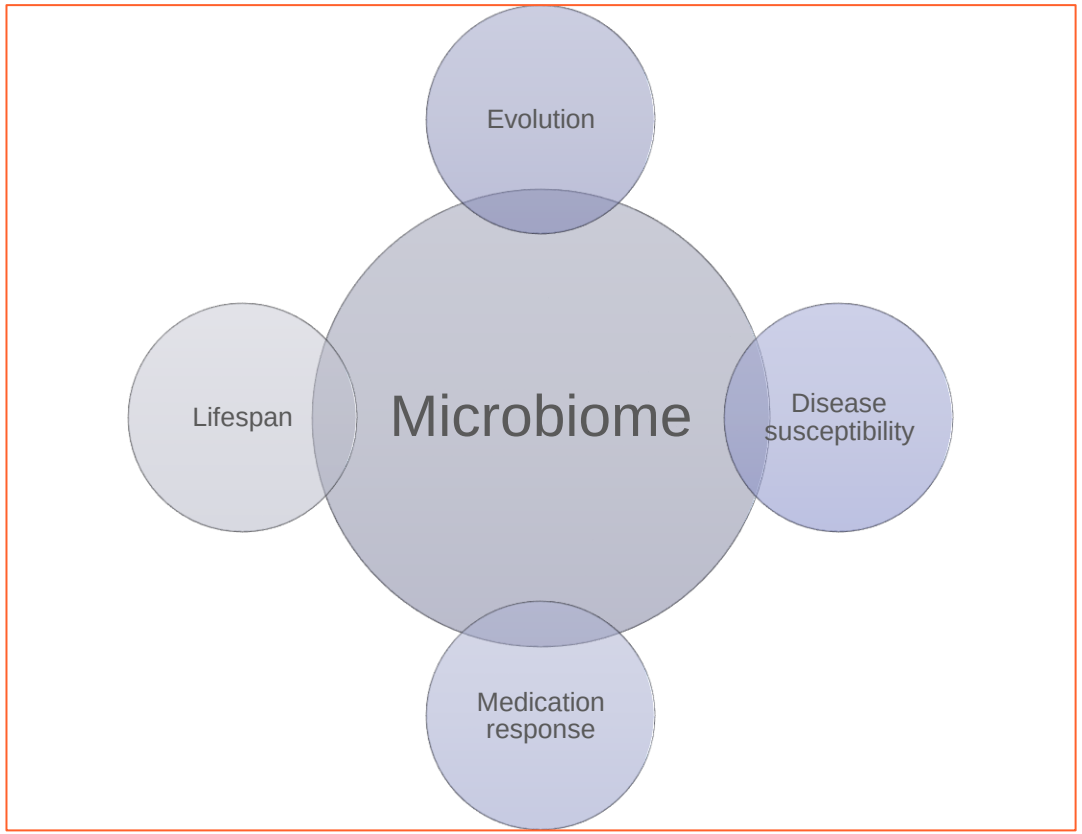
Large-scale study of proteome.
Proteome = set of proteins

Differs from cell to cell

Differs over time

© 2022 Optum, Inc. All rights reserved.

Other –omics: Microbiomics



Changes over time:

- Medication
- Birth
- Diet
- Illness

Challenges:

- Snapshot
- How to measure
- Frequency of measurements
- Interpretation

© 2022 Optum, Inc. All rights reserved.

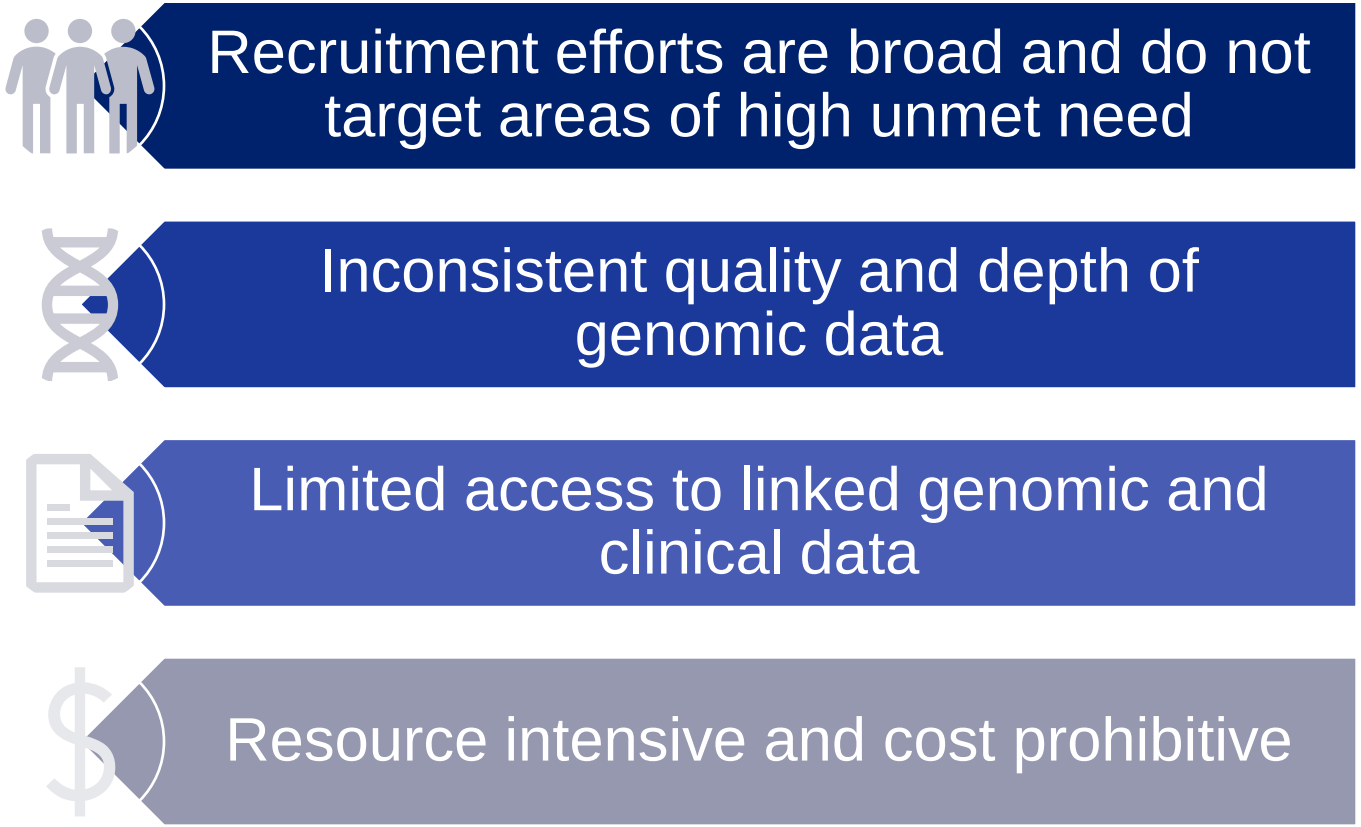
Challenges with Biobanks

- Availability
- Lack of Diversity
- Resource Intensive

Optum

© 2022 Optum, Inc. All rights reserved.

Challenges with Biobanks: Availability



Optum
© 2022 Optum, Inc. All rights reserved.

Challenges with Biobanks: Lack of Diversity

2009: Europeans represent 96% of all individuals in GWAS.[1]

2016: Europeans represent 81% of all individuals in GWAS.[2]

2019: Europeans represent 78% of all individuals in GWAS.[3]

- Participants from the US, UK and Iceland represent 71.8% of all individuals in GWAS[2].

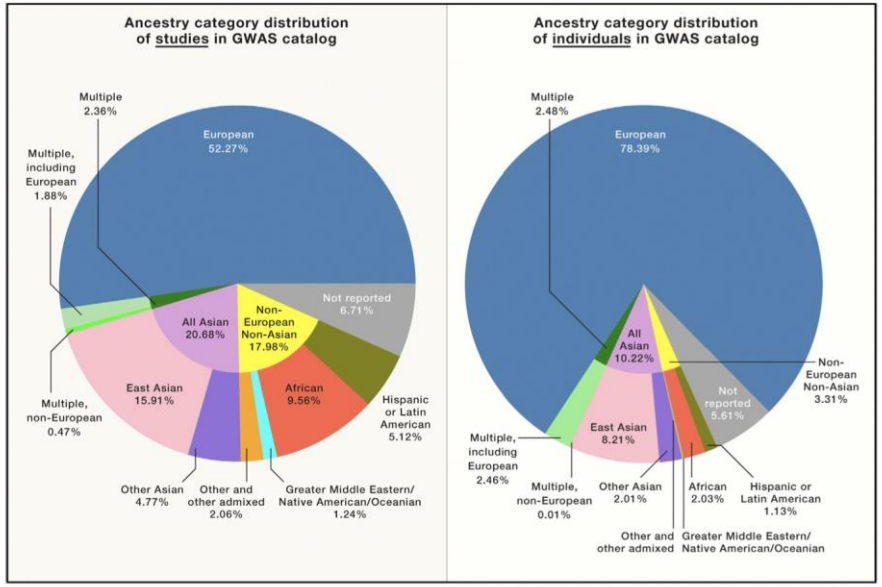


Figure 2. Summary of GWAS Studies by Ancestry for Studies in the GWAS Catalog through January 2019
We show the distribution of ancestry categories in percentages included in GWAS (https://www.ebi.ac.uk/gwas/home) based on the study (left) and based on the total number of individuals (right).

"The lack of ethnic diversity in human genomic studies means that our ability to translate genetic research into clinical practice or public health policy may be dangerously incomplete or, worse, mistaken."[3]

[1] Need AC, Goldstein. Next generation disparities in human genomics: concerns and remedies. Trends Genet. 2009;25(11):489-494.

[2] Mills MC, Rahal C. A scientometric review of genome-wide association studies. Commun Biol. 2019;2(1):9.

[3] Sirugo G, Williams SM, Tishkoff SA. The missing diversity in human genetic studies. Cell. 2019;177(1):26-31.

Optum

© 2022 Optum, Inc. All rights reserved.

Challenges with Biobanks: Resource Intensive

Baseline estimates for 5000-person cohort with common conditions

Cumulative effort

RECRUIT

$1-1.5M

- Patient identification using RWD
- Patient outreach
- Patient screening
- Consent
- Sample collection

SEQUENCE THE DATA

$5-40M

- Sample storage
- Whole genome sequencing
- Tumor sequencing
- Transcriptome RNA sequencing

SOURCE EHRs

$2.5M

- Productization and standardization
- Contracting
- Cleaning, curation, and NLP development

LINK DATA

$1.5M-2.4M

- Tokenization
- De-identification certification
- Data management
- Data linkage
- Data storage

DEVELOP PLATFORM

$10-15M /year

- Software engineering
- Data visualization
- Data science and analytics
- End user training
- Vendor solution
- In-house solution

Cumulative time

Optum

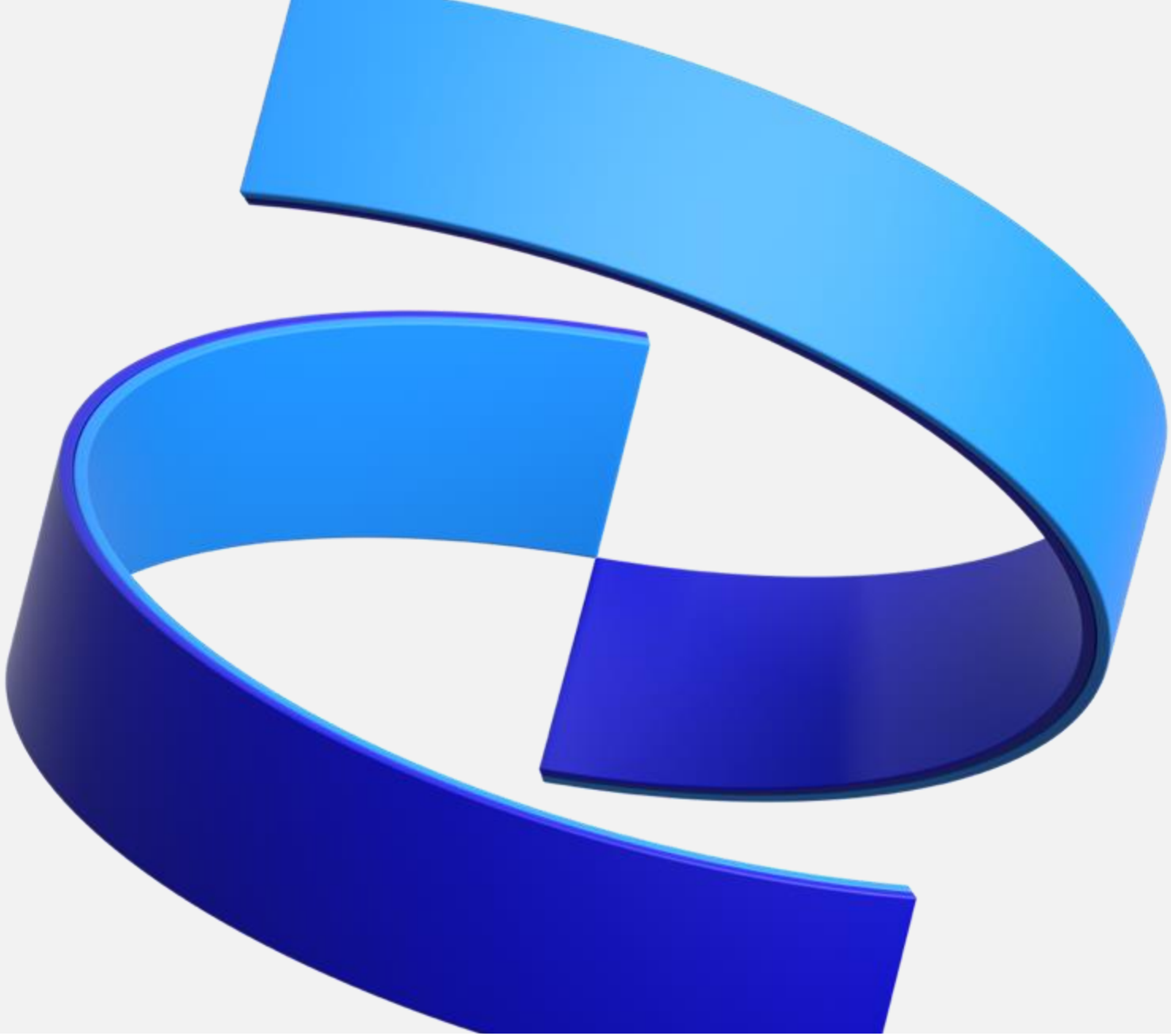
Estimates vary based on therapeutic area, cohort size, and other factors.

© 2022 Optum, Inc. All rights reserved.

https://www.genomicsengland.co.uk/news/journey-to-100000-genomes

https://piru.ac.uk/assets/files/100k%20genomes%20project-Final%20report.pdf

Big Data and Precision Medicine



ISPOR Europe 2022, November 8
Jennifer Webster, Precision Medicine RWE Lead

What do we mean when we say we can use real world data to drive precision medicine?

Precision medicine tells us that patients exposed to the same medication have heterogenous treatment effects (HTE)

And, increasingly, those heterogenous treatment effects are knowable and predictable

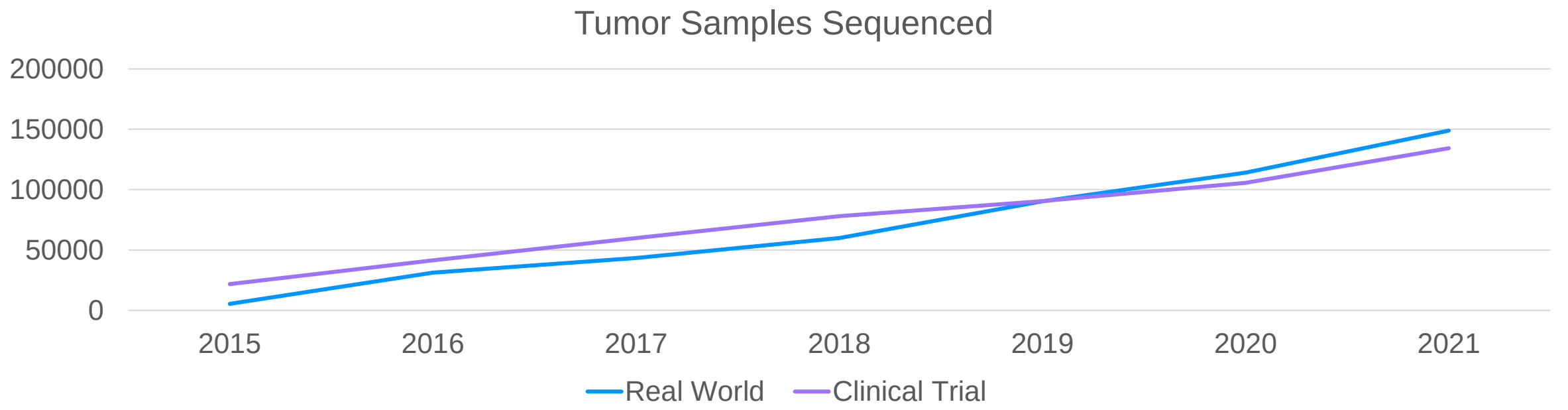
We focus our real-world data precision medicine work on five specific use cases.

- Incidence and prevalence of patient subpopulations (ex: disease x biomarker)
- Real world testing patterns and associated clinical utility
- Molecular Mechanisms of Response and Resistance
- Target Discovery & Validation
- Clinical Trial Design & Recruitment

Why now: Since 2020, more oncology patients have NGS testing in the real world than in clinical trials

Comparison of patients sequenced via industry sponsored interventional trials vs. real world standard clinical care.

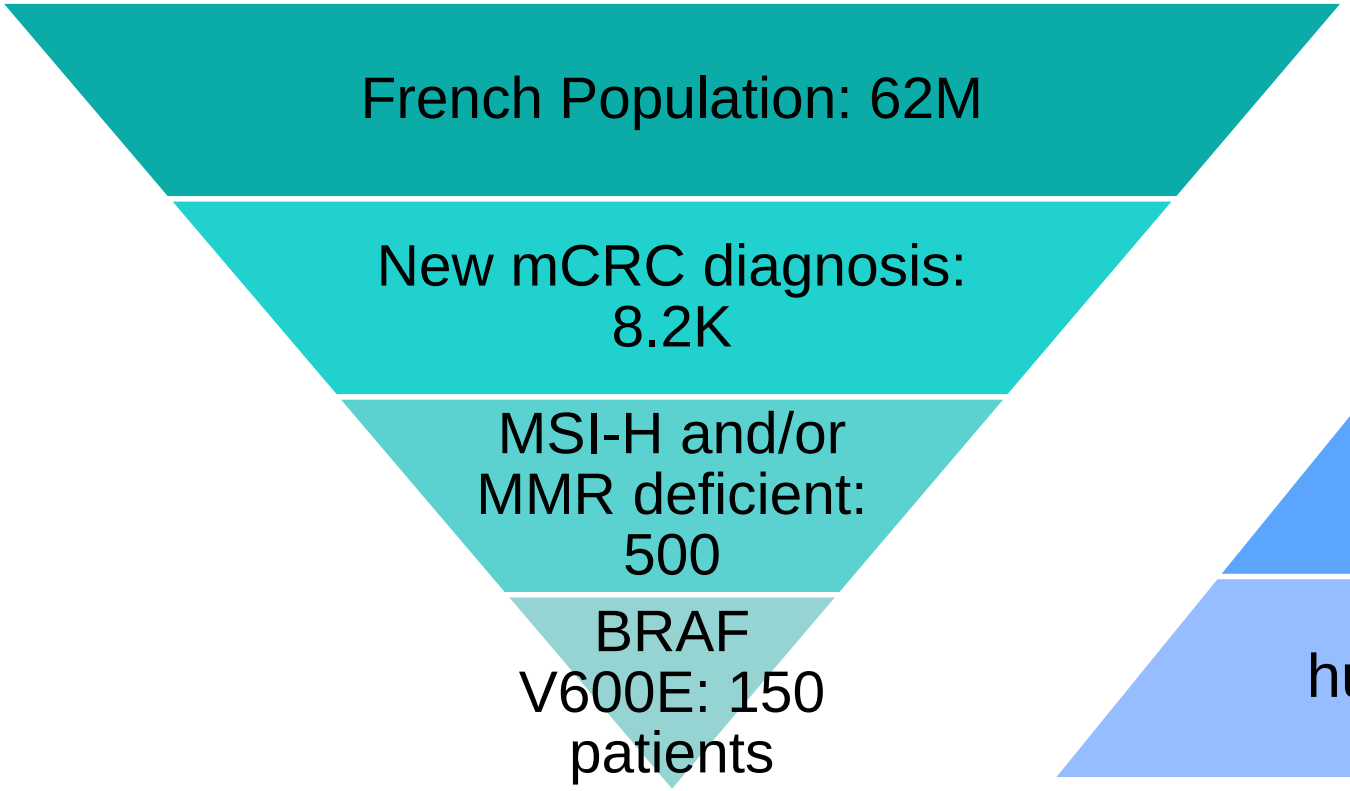
Real world data yields larger populations.



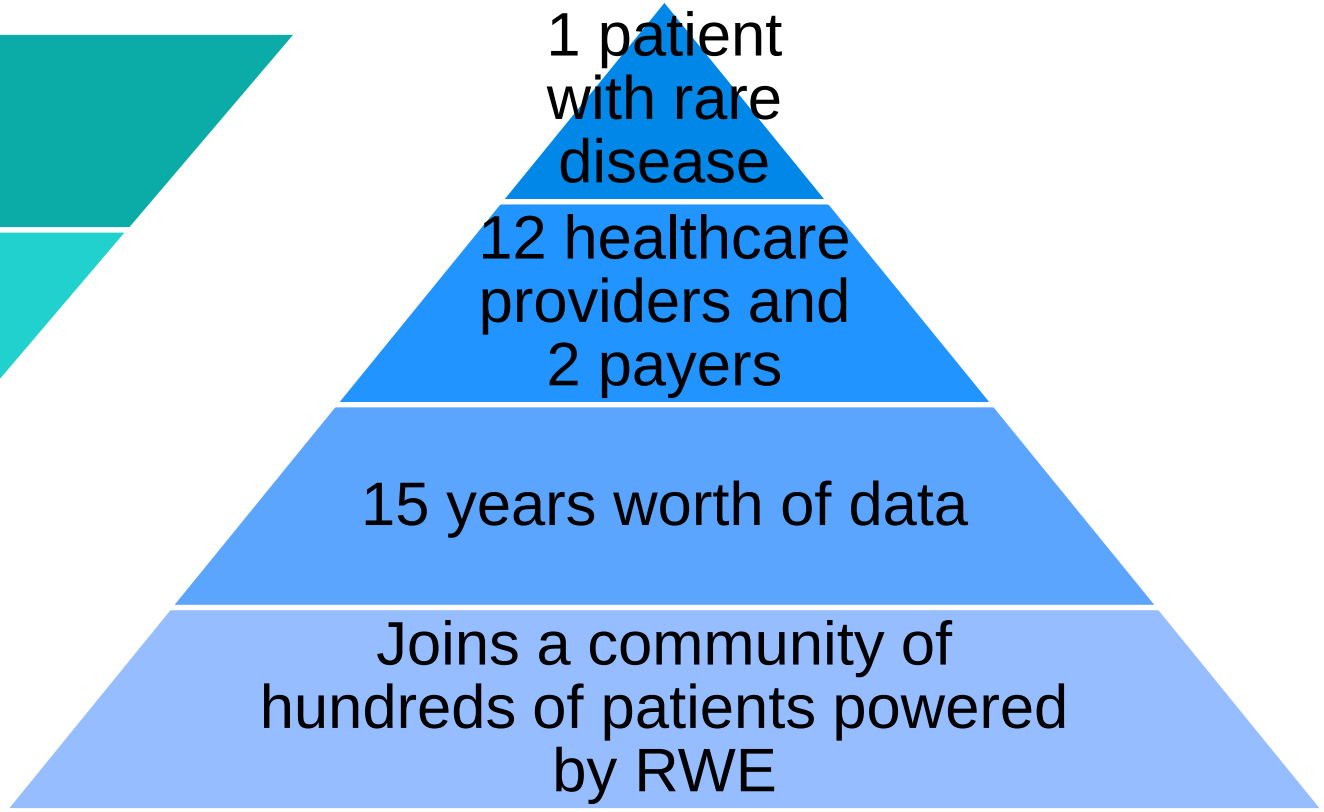
Sources: Clinicaltrials.gov, IQVIA Pharmetrics

In precision medicine we move between big data and small data

Example 1: Trial recruitment for 1L mCRC MSH-I/MMR and BRAF V600E

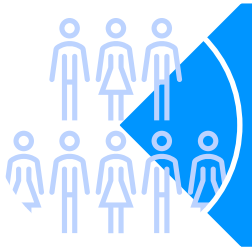


Example 2: Patient Directed Data



Precision medicine: How small is too small?

By use case, how many patients do we really need to generate evidence & insights



Incidence and prevalence of molecularly defined subpopulations:

20 – 30 pts

Mechanisms of response and resistance:

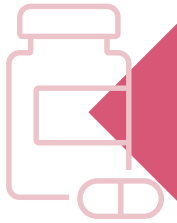
20 patients with paired samples

Understanding testing landscape:

50 – 100 patients

Target Validation:

As little as 1 sample with a novel finding



Clinical trial recruitment:

As little as 1 patient who meets I/E criteria

How big is too big? When does it stop being precision medicine?

Precision Medicine Approaches in context. Assume we have a therapy that is 6x more effective in the biomarker = yes population

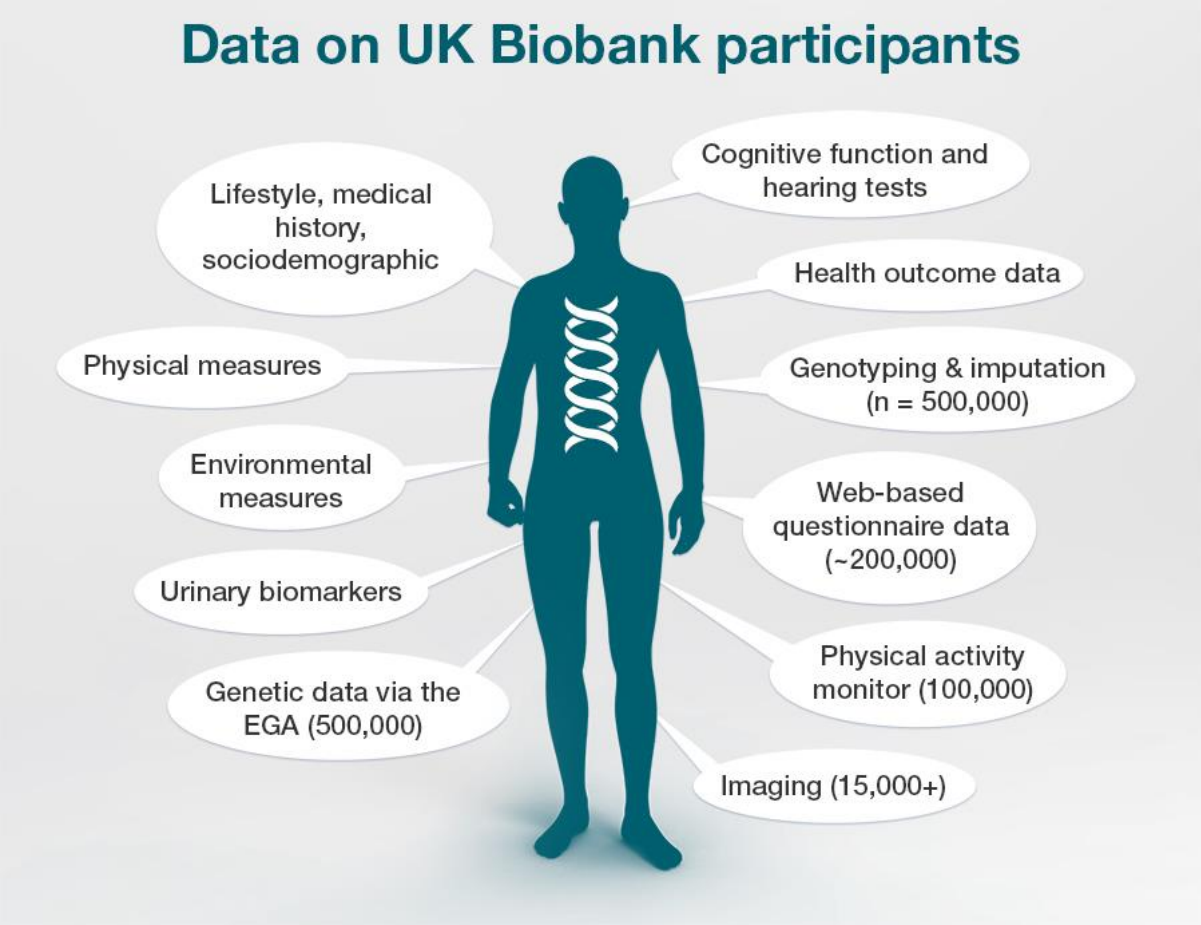
All-Comers Population 1L (100%)

Biomarker = yes (25%) **ORR = 75%**	**Biomarker = no (75%)** **ORR = 12.5%**

Approach	Impact
Approach 1: Target all-comers population with no biomarker subgroups	Possible for study to succeed based on strength of signal in biomarker=yes group. No evidence developed about biomarker or patient selection. Benefit accrues to biomarker = yes population and biomarker = no patients risk losing opportunity for more effective therapy.
Approach 2: Target biomarker = yes population	Likely successful trial (biomarker-driven trials have 2x PTRS). Evidence developed to support biomarker & patient selection. Benefit accrues to biomarker = yes population
Approach 3: Use combinations to make targeted tx effective for all comers	Lowest likelihood of success. Requires evidence generation for a biomarker plus at least two drugs. If successful, benefit accrues to entire all comers population.

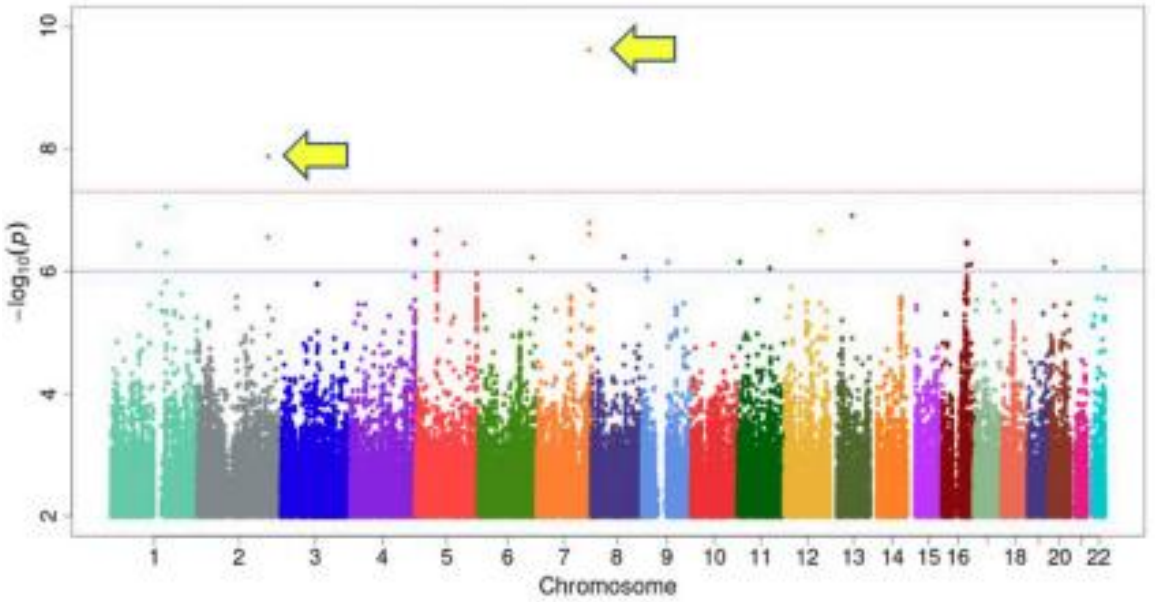
Examples

Identifying new genetic targets and biomarkers



RWI
- Identified 2 novel genetic variants associated with HCM
- Found new biometrics and biomarkers associated with HCM

Hypertrophic cardiomyopathy (HCM)
- Disorder of heart muscles associated with variants in 8 genes
- Compared genomes of 363 individuals with HCM to 7,260 controls matched for age, sex, and ancestry
- Examined comorbidities based on ICD diagnosis codes



Gene	rs Number	Chromosome	Position	Variant Type	OR	P-value	Minor Allele Frequency
KMT2C	rs78630626	7	152,056,039	Intronic	3.8	2.4×10^{-10}	1.6%
PARD3B	rs188937806	2	205,754,718	Intronic	3.8	1.3×10^{-8}	1.0%

References
Gyftopoulos A, et al. Identification of Novel Genetic Variants and Comorbidities Associated With ICD-10-Based Diagnosis of Hypertrophic Cardiomyopathy Using the UK Biobank Cohort. Front Genet. 2022;13:866042.

Understanding natural history to enrich clinical trial population

Background

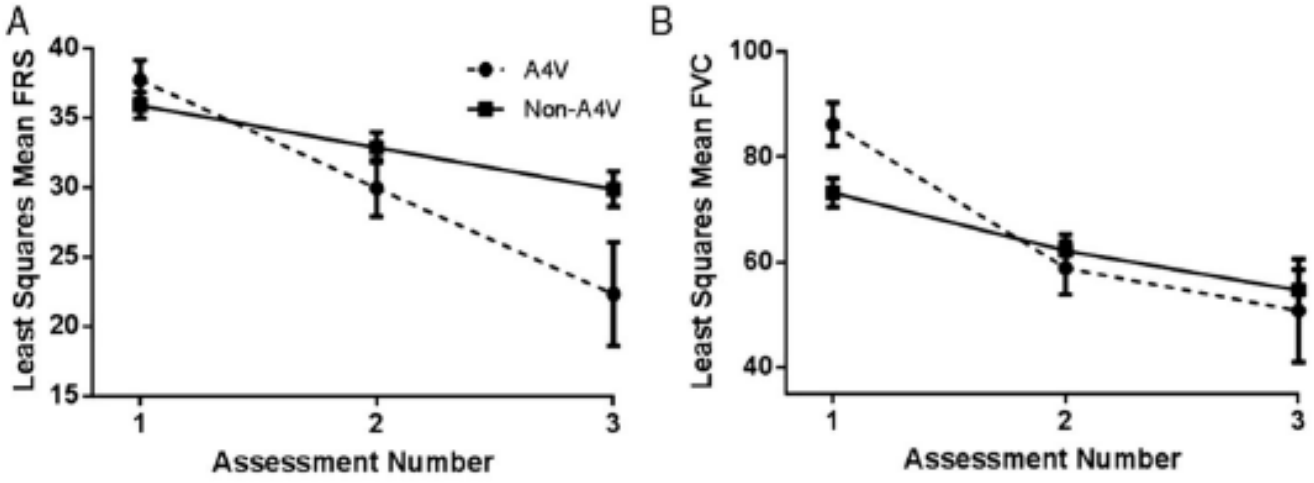
- Amyotrophic lateral sclerosis (ALS) is a fatal and progressive neurological disease with few therapies
- A subgroup of patients with familial ALS have mutations in the SOD1 gene
- Therapies aimed at SOD1 need to understand natural history of disease progression

Methods

- Consortium conducted retrospective chart review to identify 175 patients with ALS and SOD1 mutations
- Results were pooled to analyze changes in ALS-Functional Rating Scale (FRS) and forced vital capacity (FVC) over time
- Compared 2 subgroups of SOD1 mutations (A4V vs non-A4V)

RWI

- Significant differences were found in disease progression between A4V and non-A4V SOD1 mutations



- Outcomes within A4V subgroup were homogeneous
- Focusing on A4V subgroup could reduce sample size required by ~40%

Group	Sample size
SOD1 overall	N = 88
SOD1 A4V	N = 52

References

Bali T, Self W, Liu J, Siddique T, Wang LH, Bird TD, et al. Defining SOD1 ALS natural history to guide therapeutic clinical trial design. J Neurol Neurosurg Psychiatry. 2017;88(2):99-105.

Improving clinical trial feasibility and external validity

Background

- RCTs of novel therapies for multiple myeloma (MM) have reported significant increases in long-term survival
- Findings from RCTs where only a fraction of patients participate are difficult to extrapolate
- Overly strict study eligibility criteria are an important barrier to clinical trial participation

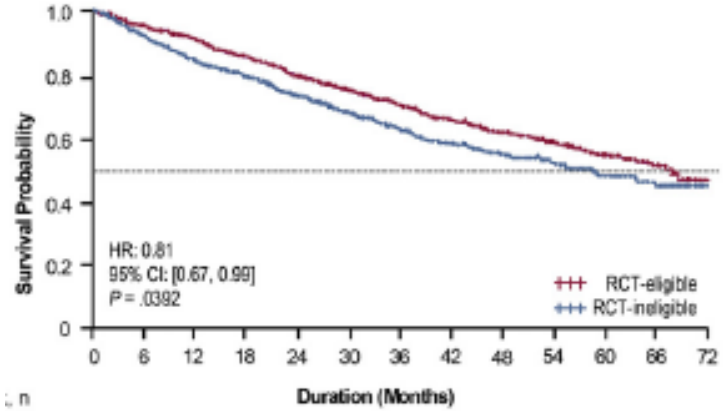
Methods

- Compared characteristics and survival for 1,406 patients in MM registry (ie, RWD) vs. participants in RCTs

RWI

- 40.0 to 56.8% of patients meeting minimal criteria for novel therapies would be ineligible for RCTs
- Patients ineligible for RCTs have more comorbidities, more advanced disease, and worse outcomes

RCT Eligibility Criteria[16,18-43]	n (%)
RCT Ineligible Not Considering Hemoglobin Level or M-Protein ≤1.0 g/dL	563 (40.0)
RCT Ineligible Including Hemoglobin Level ≤8 g/dL	618 (44.0)
RCT Ineligible Including Hemoglobin Level ≤8 g/dL and M-Protein ≤ 1.0 g/dL	799 (56.8)



- Modifying commonly used study eligibility criteria could increase external validity without jeopardizing trials

References

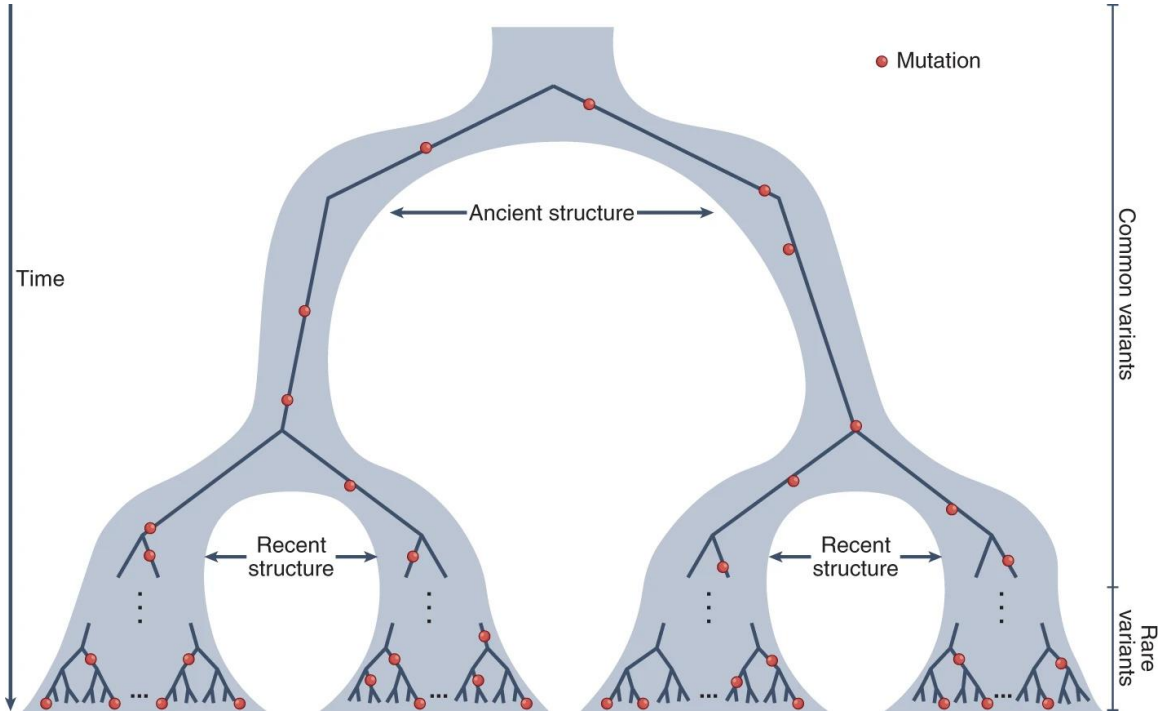
Shah JJ, et al. Analysis of Common Eligibility Criteria of Randomized Controlled Trials in Newly Diagnosed Multiple Myeloma Patients and Extrapolating Outcomes. Clin Lymphoma Myeloma Leuk. 2017;17(9):575-83 e2.

What’s Next?

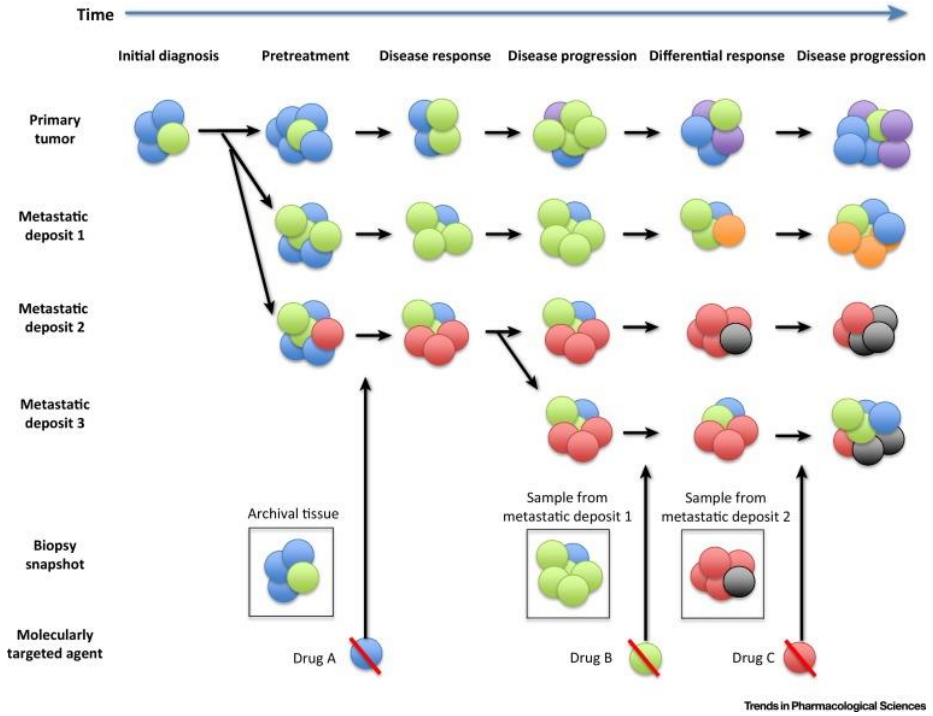
As patient subgroups get smaller and smaller, we must take the opportunity to learn from every patient, inform every patient

Will precision medicine lead us to practice n of (less than) 1 medicine?

Outsize impact of recent, rare variants drive disease[1]



Heterogeneity even within the same patient[2]



1. Young, A.I. Discovering missing heritability in whole-genome sequencing data. *Nat Genet* **54**, 224–226 (2022). https://doi.org/10.1038/s41588-022-01012-3
2. Towards Precision Medicine in the Clinic: From Biomarker Discovery to Novel Therapeutics. Collins, Dearbhaile C. et al. Trends in Pharmacological Sciences, Volume 38, Issue 1, 25 - 40

Reasons for data sparsity: structural

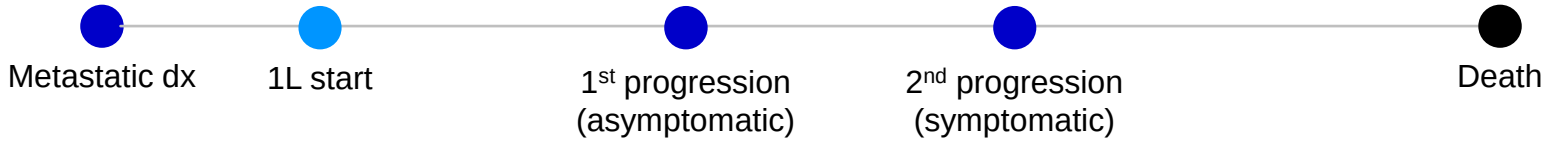
Patients receiving care in normal ways results in data that does not match our mental models

Data can be technically correct, but still appear erroneous

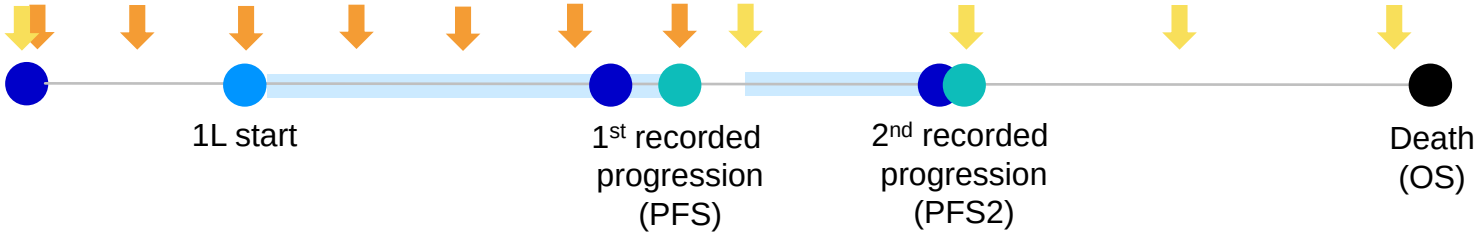
RW Care Scenario	Impact
Patient managed by community center, but sees an academic oncologist for a second opinion	Patient will have (at least) two records, neither of which is complete.
Patient is a “snowbird” – spending six months in one location and six months in another every year	Patient will have two records, an appear in each to only receive care for six months of the year followed by a six month “treatment holiday”
Patient participates in a clinical trial for first line therapy	Care received in clinical trial does not show in real world record – including therapy & study-sponsored biomarker testing
Patient receives therapy via patient assistance program	Treatment record spotty (impression of poor adherence whether true or not) or missing altogether
Patient experiences long delay between therapy order & administration due to insurance or other factors	Patients who experience these barriers & care delays are frequently excluded from RW studies.

Reasons for sparsity: Assessment patterns matter (maybe more than treatment patterns)

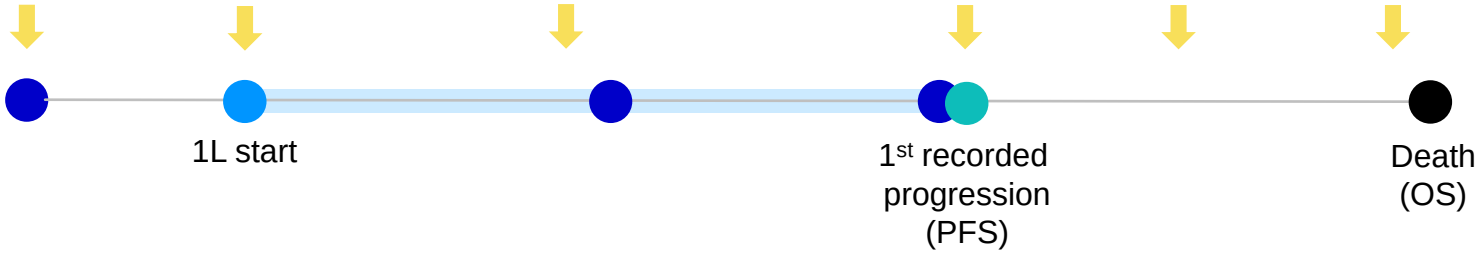
Disease course



Assessments when patient is enrolled in a clinical trial



Assessments if that same patient is treated via standard real-world care



Key

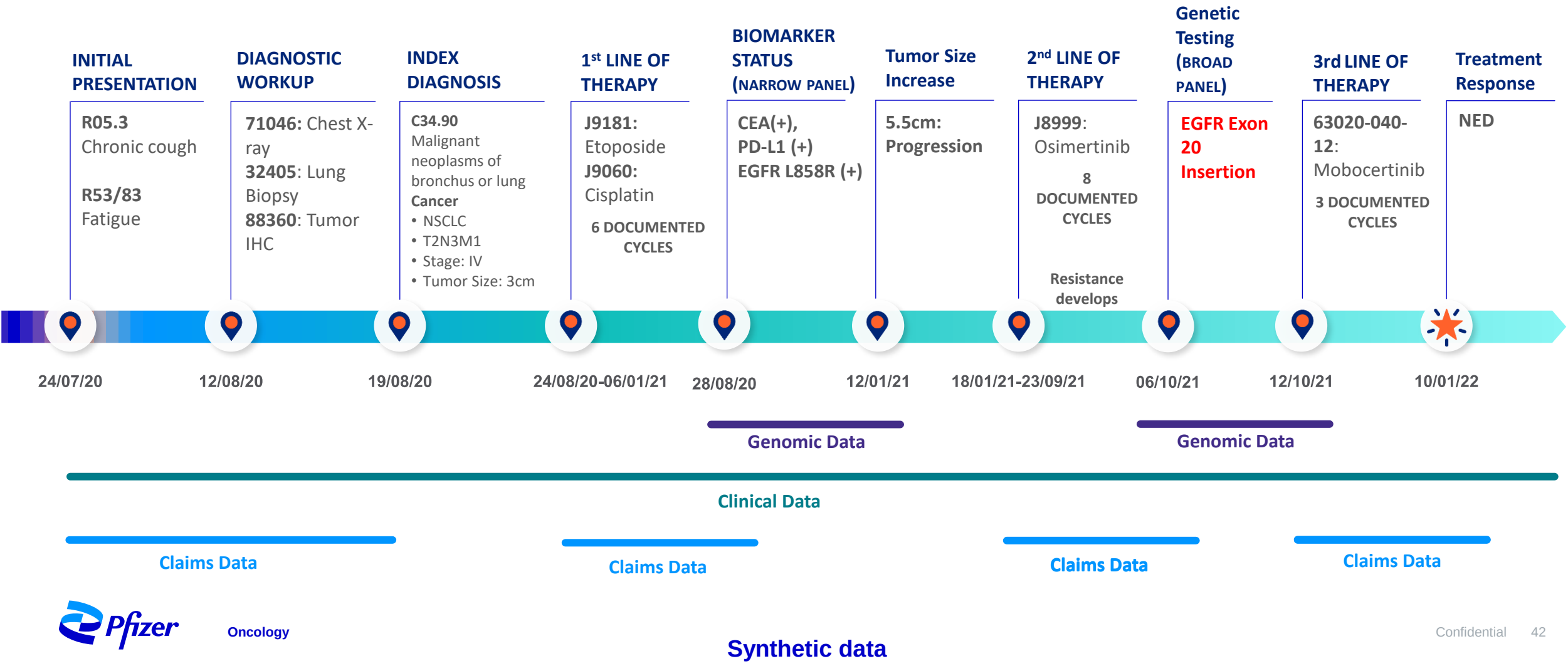
- Underlying biology
- Treatment
- Assessment (CT)
- Assessment (RW)
- Measured PFS/PFS2
- Treatment

Our current data and methodologies will not support this transition. And the data quality is reaching an asymptotic limit.

We need new methods

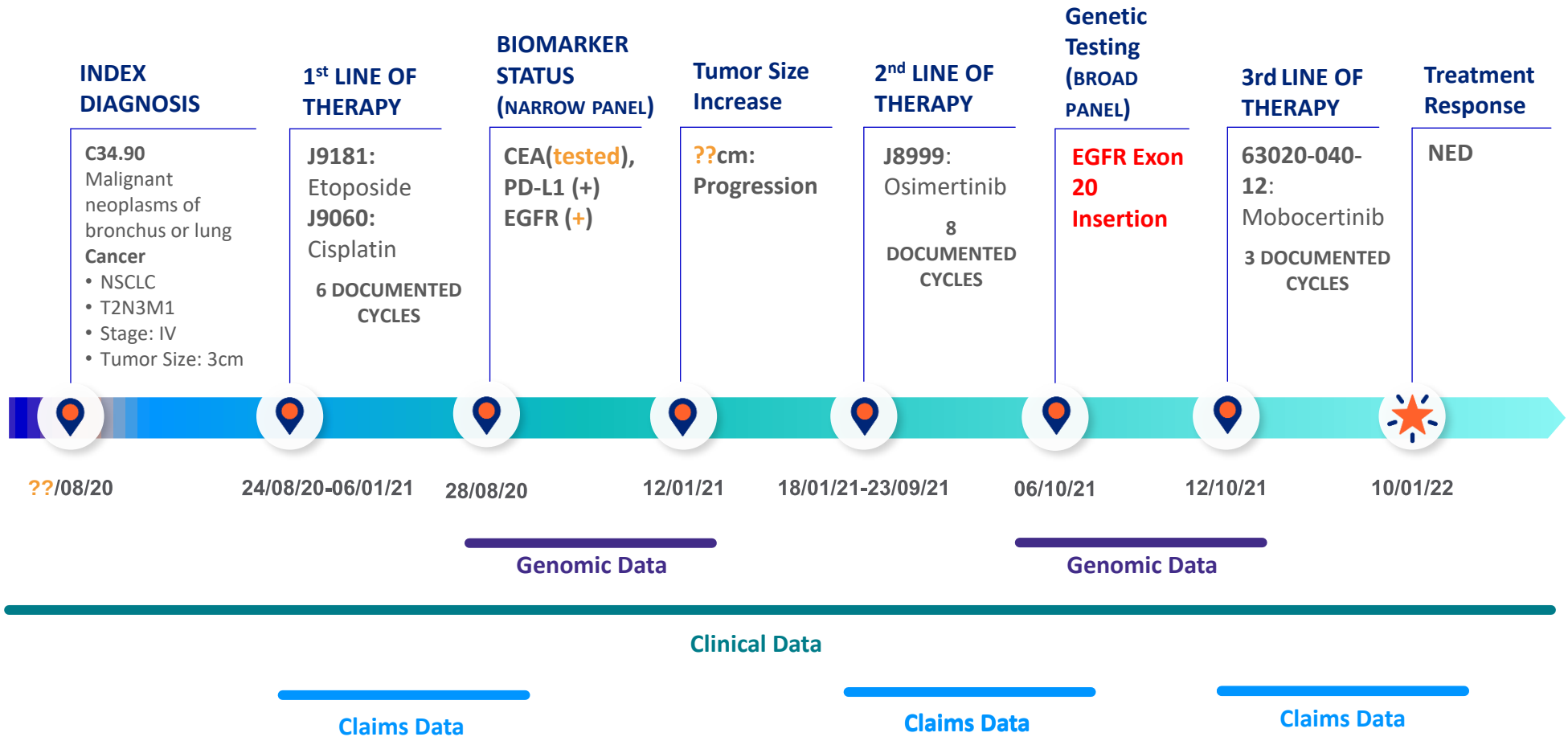
Every patient record in our datasets is informative (1 of 3: best case)

It is our job as scientists employing real world data to extract and use that information.



Every patient record in our datasets is informative (2 of 3: common RWE case)

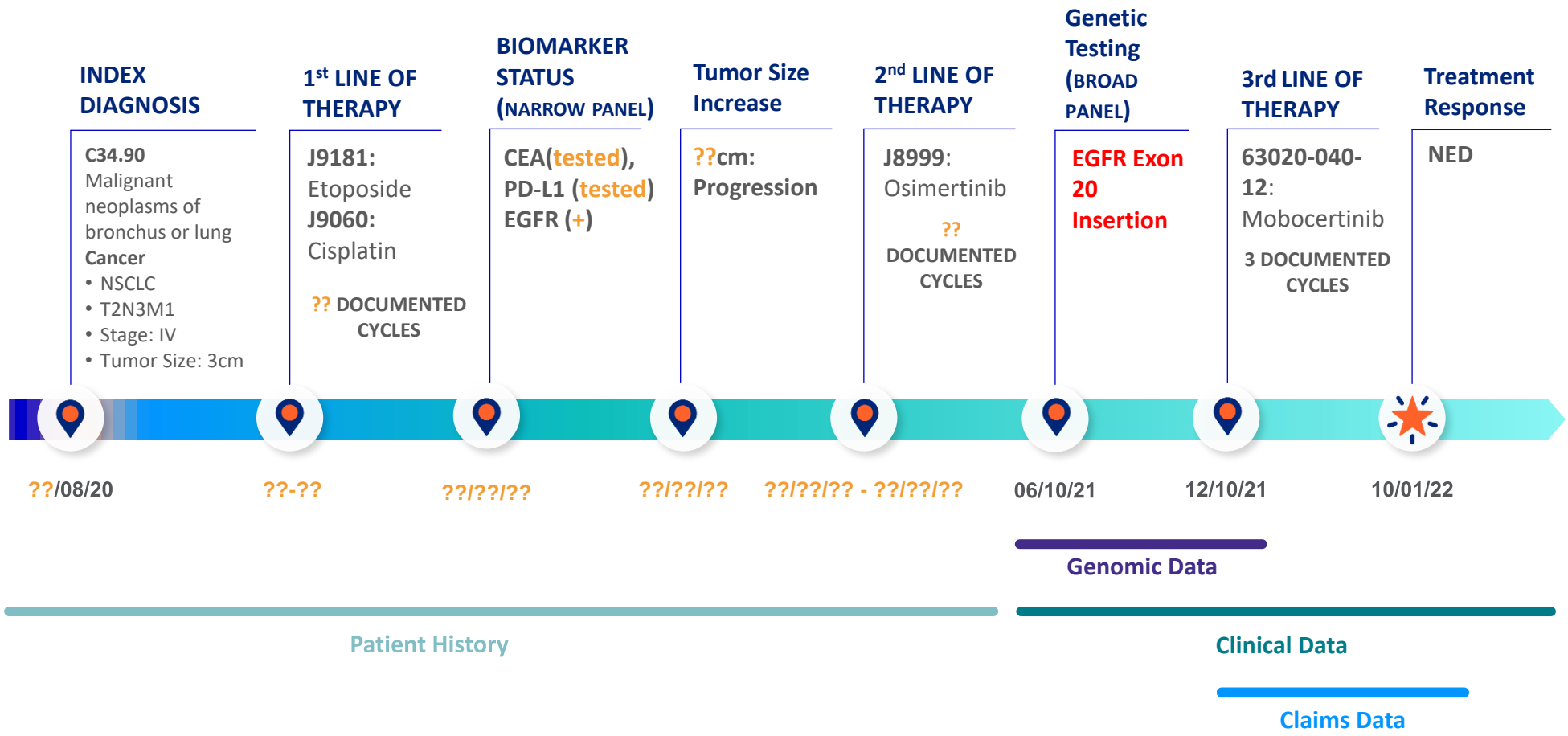
It is our job as scientists employing real world data to extract and use that information.



Synthetic data

Every patient record in our datasets is informative (3 of 3: still informative)

It is our job as scientists employing real world data to extract and use that information.



Synthetic data

Example RWD Genomic Inquires

1. **Identifying patients who underwent molecular testing.**
   - 1a Claims algorithms
2. **Identifying the types of test patients receive.**
   - 2a Claims algorithms – Code stacking and provider ID/Tax ID
3. **Identifying molecular testing results.**
   - 3a EHR – Structured and unstructured
   - 3b Lab Direct Results

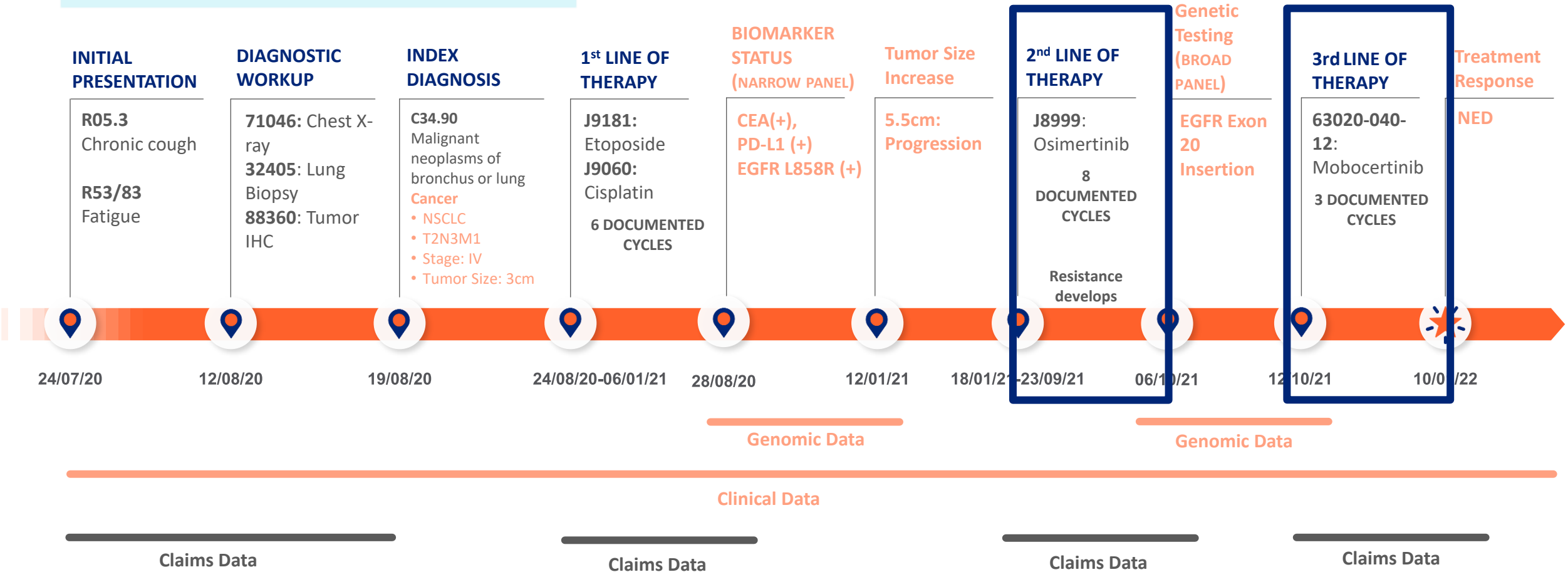
Optum

© 2022 Optum, Inc. All rights reserved.

Claims Data

PATIENT CHARACTERISTICS
- 60 YO Asian Female
- Activity Period: 03/2017-03/2022

Claims Only Data:
Limitations: Absence of molecular test results
Solutions: Presence of a targeted treatment implies the modification was identified even in the absence of claims for a molecular test.



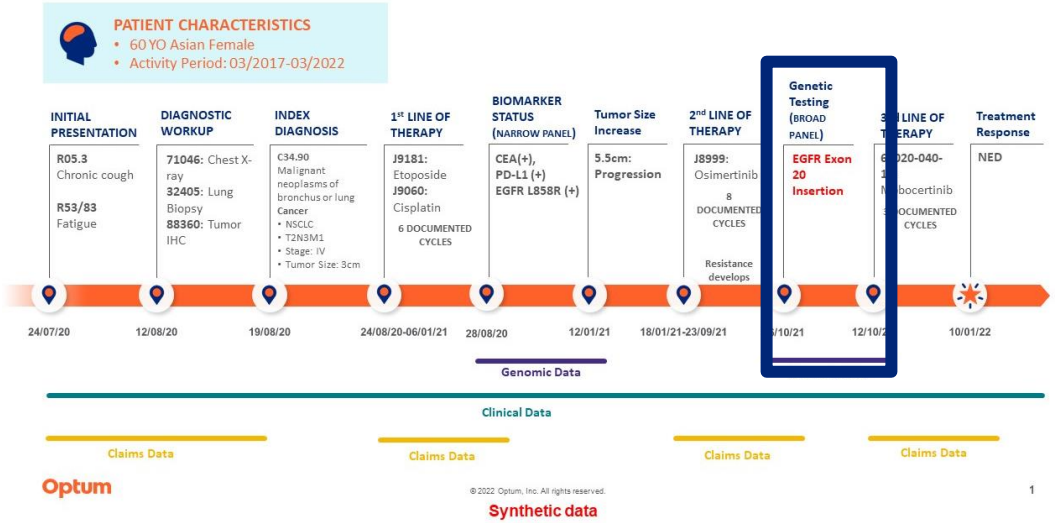
Optum

© 2022 Optum, Inc. All rights reserved.

Synthetic data

Claims Data

Claims Only Data: Documentation of type of test, methodology (NGS v PCR) and targets to be evaluated are limited. Code descriptions become key but due to the prevalence to unspecified codes, crosswalks between provider or tax ID and code stacking become necessary.



Ideal:

Test Date	Code (CPT)	Code Description
04/10/2021	81445	Targeted genomic sequence analysis panel, solid organ or hematolymphoid neoplasm, DNA analysis, and RNA analysis when performed, 51 or greater genes (eg, ALK, BRAF, CDKN2A, CEBPA, DNMT3A, EGFR, ERBB2, EZH2, FLT3, IDH1, IDH2, JAK2, KIT, KRAS, MLL, NPM1, NRAS, MET, NOTCH1, PDGFRA, PDGFRB, PGR, PIK3CA, PTEN, RET), interrogation for sequence variants and copy number variants or rearrangements, if performed

Common:

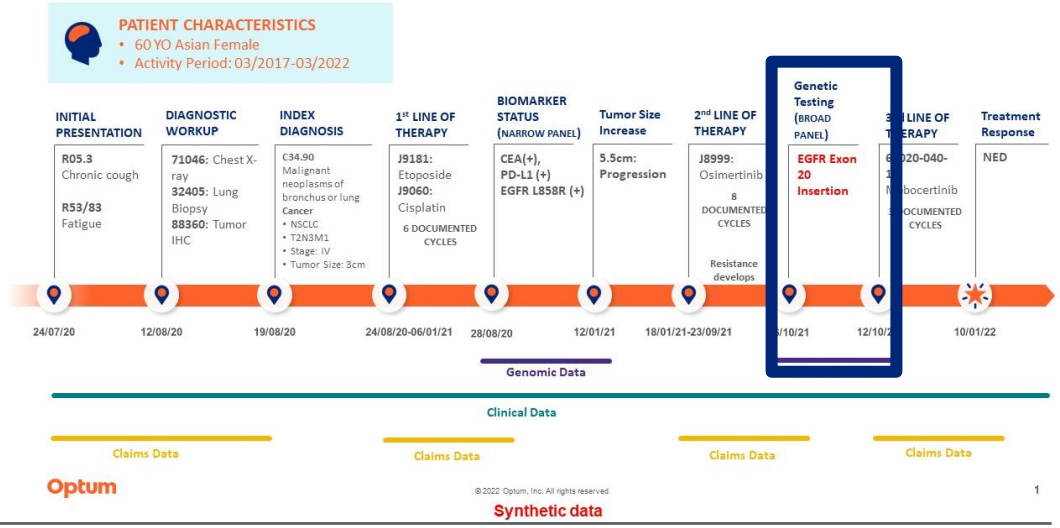
Test Date	Code (CPT)	Code Description
04/10/2021	81235	EGFR (epidermal growth factor receptor) (eg, non-small cell lung cancer) gene analysis, common variants
04/10/2021	81275	KRAS (Kirsten rat sarcoma viral oncogene homolog) (eg, carcinoma) gene analysis
04/10/2021	81400	MOLECULAR PATHOLOGY PROCEDURE LEVEL 1
04/10/2021	81401	MOLECULAR PATHOLOGY PROCEDURE LEVEL 2

© 2022 Optum, Inc. All rights reserved.

Synthetic data

Claims Data

Claims Only Data: Code stacking allows assumptions if broad panel testing were likely utilized. Identification of the targets or the utilization of narrow panel testing is still elusive. Provider or Tax ID allows assumptions of the type of molecular testing. For example, Guardant Health Inc only performs liquid biopsies using broad panel NGS.



Test Date	Code (CPT)	Code Description	Provider ID
04/10/2021	81235	EGFR (epidermal growth factor receptor) (eg, non-small cell lung cancer) gene analysis, common variants	Guardant Health Inc
04/10/2021	81275	KRAS (Kirsten rat sarcoma viral oncogene homolog) (eg, carcinoma) gene analysis	Guardant Health Inc
04/10/2021	81400	MOLECULAR PATHOLOGY PROCEDURE LEVEL 1	Guardant Health Inc
04/10/2021	81401	MOLECULAR PATHOLOGY PROCEDURE LEVEL 2	Guardant Health Inc
04/10/2021	81401	MOLECULAR PATHOLOGY PROCEDURE LEVEL 2	Guardant Health Inc
04/10/2021	81401	MOLECULAR PATHOLOGY PROCEDURE LEVEL 2	Guardant Health Inc
04/10/2021	81403	MOLECULAR PATHOLOGY PROCEDURE LEVEL 4	Guardant Health Inc
04/10/2021	81403	MOLECULAR PATHOLOGY PROCEDURE LEVEL 4	Guardant Health Inc
04/10/2021	81403	MOLECULAR PATHOLOGY PROCEDURE LEVEL 4	Guardant Health Inc

© 2022 Optum, Inc. All rights reserved.

Synthetic data

Unstructured EHR: Free Text

EHR provider free text: Continual lack of extensive, clear documentation of not only tested molecular targets but the subsequent results. Manual review or natural language processing (NLP) are key.

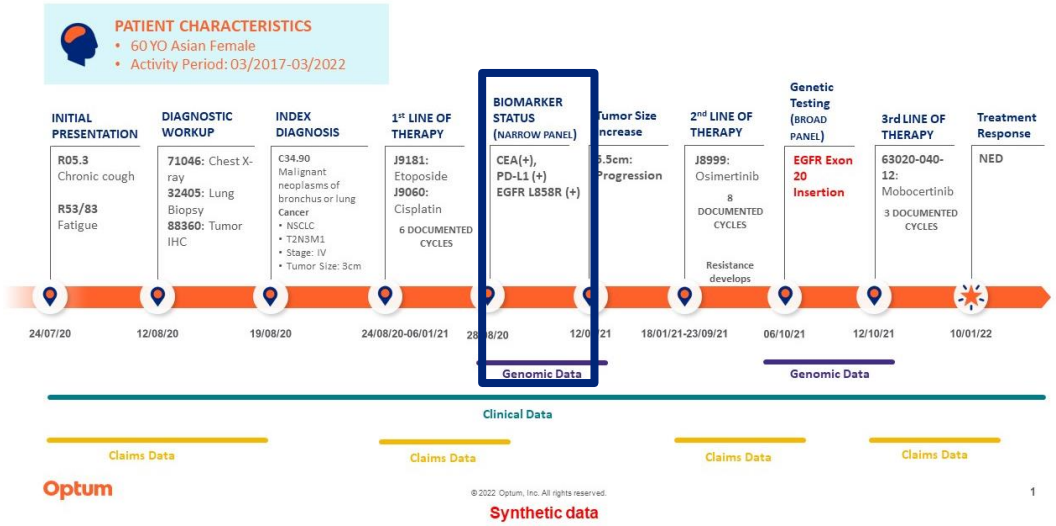
28/08/2020: Patient tested for CEA, PD-L1, EGFR, and KRAS. Patient tested positive for CEA, PD-L1 and has a EGFR L858R mutation. Patient was started on etoposide and cisplatin prior to results. Patient tolerating tx well and no adjustments will be made.

Ideal EHR Documentation

28/08/2020: Patient CEA, PD-L1, EGFR positive.

28/08/2020: Results: CEA, PD-L1, EGFR and KRAS.

Common EHR Documentation



© 2022 Optum, Inc. All rights reserved.

Synthetic data

Unstructured EHR: PDF lab results

Summary of Detected Somatic Alterations, Immunotherapy Biomarkers & Associated Treatment Options

KEY ✓ Approved in indication ~ Approved in other indication ✗ Lack of response

Detected Alteration(s) / Biomarker(s)	Associated FDA-approved therapies	Clinical trial availability (see page 3)	% cfDNA or Amplification
EGFR N771_P772insH (Exon 20 insertion)	✓ Amivantamab, Mobocertinib ✗ Afatinib, Erlotinib, Gefitinib	Yes	2.6%
TP53 G245A	None	Yes	1.0%

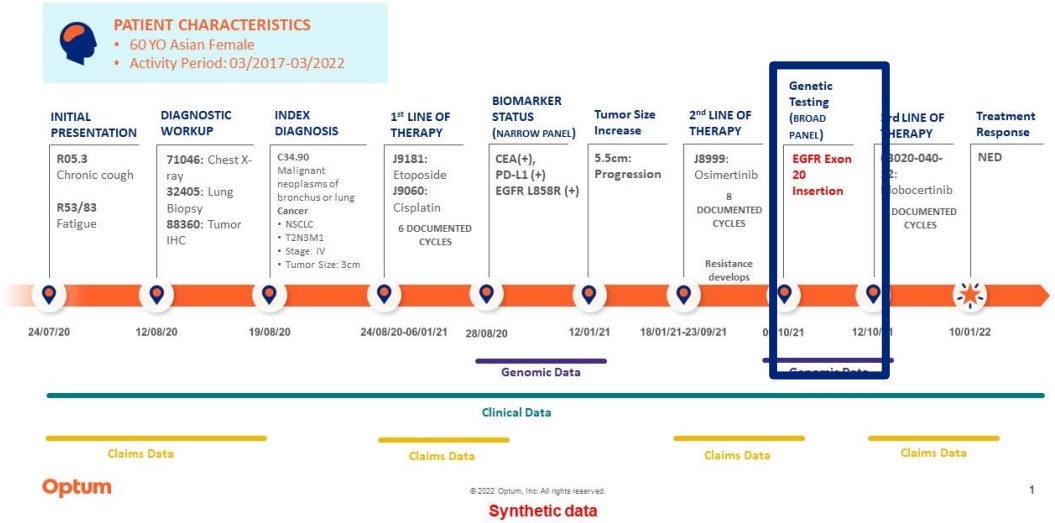
Additional Biomarkers

Biomarker	Additional Details
MSI-High	NOT DETECTED

Alterations or biomarkers that were "NOT DETECTED" have been excluded from the summary table above.

We evaluated this sample for 74 genes, including the following guideline-recommended genes for NSCLC

EGFR(T790M and others)	*ALK*	*ROS1*	*BRAF*	*MET*	*ERBB2*(HER2)	*RET*	*NTRK*

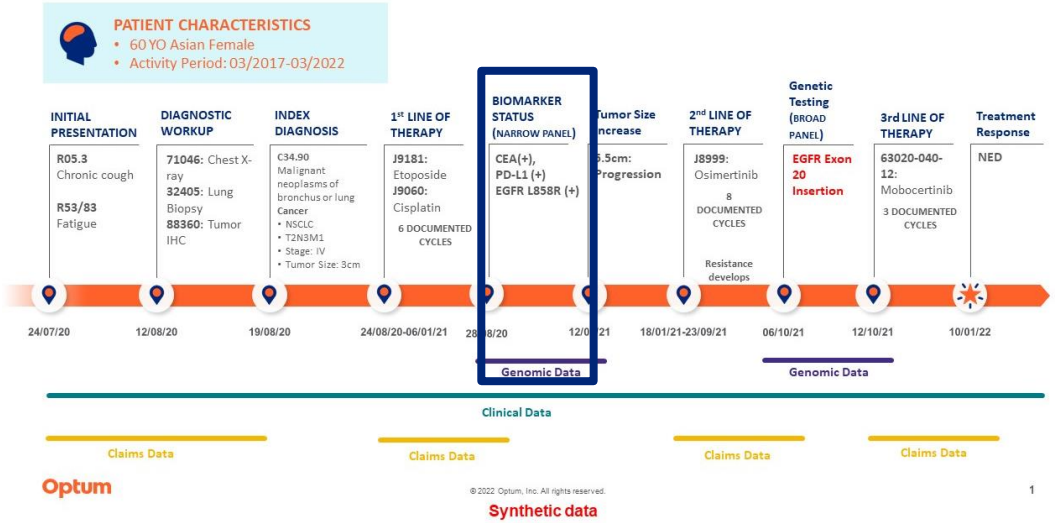


Unstructured EHR PDF lab results: Requires manual review to abstract the test, results, and implications.

© 2022 Optum, Inc. All rights reserved.
Synthetic data

Structured EHR Molecular Testing Examples

Structured EHR: Compliments unstructured EHR molecular target missingness and does not require manual review nor NLP.



Test Date	Biomarker	Narrative Result	Numeric Result	Test Name
21/08/2020	CD274 MOLECULE (CD274 OR PD-L1 OR PDL1)	Positive		IMMUNOHISTOCHEMISTRY (IHC)
21/08/2020	CD274 MOLECULE (CD274 OR PD-L1 OR PDL1)		2+	
11/01/2021	CD274 MOLECULE (CD274 OR PD-L1 OR PDL1)	Negative		
11/01/2021	CD274 MOLECULE (CD274 OR PD-L1 OR PDL1)	Positive		IMMUNOHISTOCHEMISTRY (IHC)
11/01/2021	CD274 MOLECULE (CD274 OR PD-L1 OR PDL1)	Positive	20%	IMMUNOHISTOCHEMISTRY (IHC)
18/03/2021	CD274 MOLECULE (CD274 OR PD-L1 OR PDL1)	Expressed	<1%	
18/03/2021	CD274 MOLECULE (CD274 OR PD-L1 OR PDL1)		2%	
21/03/2021	CD274 MOLECULE (CD274 OR PD-L1 OR PDL1)		20%	IMMUNOHISTOCHEMISTRY (IHC)

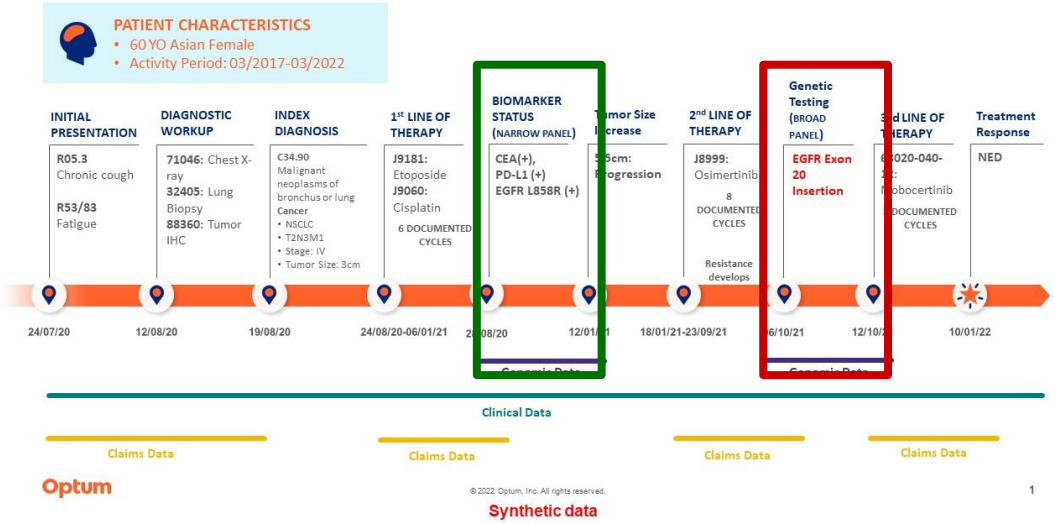
Optum

© 2022 Optum, Inc. All rights reserved.

Synthetic data

Direct from Lab Derived Data

Lab Only Derived Data: Allows for the capture of nonclinically relevant modifications.



Test Date	Test Type	Target	Modification	Result	Clinically Relevant at Time of Test
21/08/2020	RT-PCR	Epidermal Growth Factor Receptor (EGFR)	SNV	L858R	Yes
21/08/2020	RT-PCR	Kirsten rat sarcoma virus (KRAS)	None	-	No
04/10/2021	NGS	Epidermal Growth Factor Receptor (EGFR)	SNV	L858R	Yes
04/10/2021	NGS	Epidermal Growth Factor Receptor (EGFR)	Insertion	N771_P772insH	Yes
04/10/2021	NGS	Kirsten rat sarcoma virus (KRAS)	SNV	G12A	No
04/10/2021	NGS	Rearranged during transfection (RET)	SNV	F393L	No
04/10/2021	NGS	Anaplastic lymphoma kinase (ALK)	None	-	No

Optum

© 2022 Optum, Inc. All rights reserved.

Synthetic data

Discussion

© 2022 Optum, Inc. All rights reserved.

Presenters

Thank you for joining us today!

Amy Sainski-Nguyen

Optum LifeSciences

Director, Oncology

Director, Clinicogenomics

amy.nguyen1@optum.com

Jennifer Webster

Pfizer

Senior Director, RWE Oncology

jennifer.webster@pfizer.com

Ashley Brenton

Optum LifeSciences

VP, Clinicogenomics

ashley.brenton@optum.com

© 2022 Optum, Inc. All rights reserved.