October 29, 2018

Data Considerations

Critical Dataset Features

Amy Abernethy, MD, PhD
Chief Medical Officer / Chief Scientific Officer & SVP - Oncology, Flatiron Health *(a member of the Roche Group)*
Adjunct Professor of Medicine, Duke University School of Medicine
@dramyabernethy ◊◊ amy@flatiron.com

flatiron

Disclosures

Amy Abernethy is Chief Medical Officer, Chief Scientific Officer and SVP, Oncology at Flatiron Health, a member of the Roche Group, and has stock ownership in Roche

Other disclosures include:

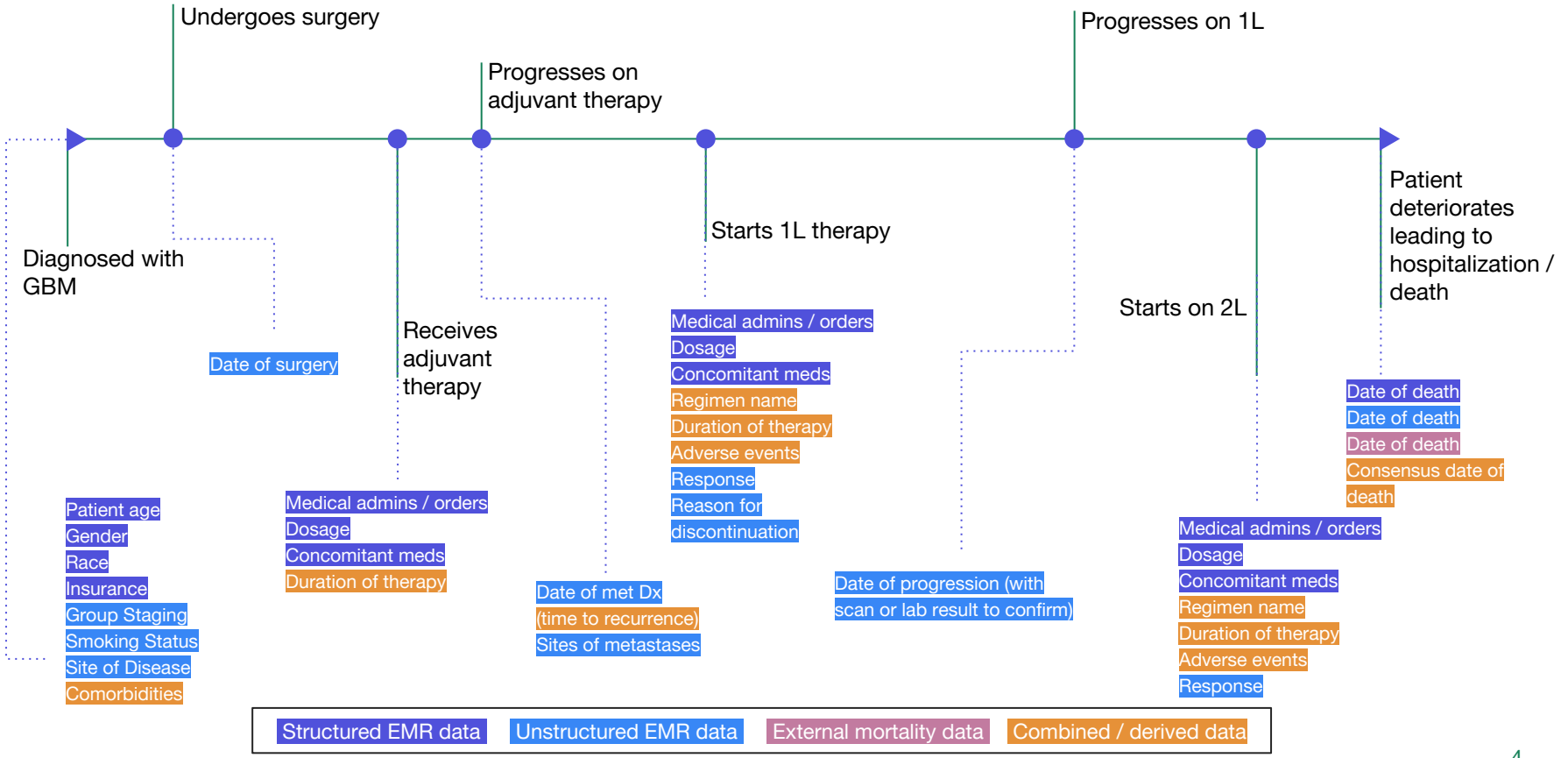
- Athenahealth (NASDAQ: ATHN), Board of Directors, Stock ownership
- CareDx (NASDAQ: CDNA), Board of Directors, Stock ownership
- Orange Leaf Associates, LLC, Owner
- Highlander Partners, Senior Advisor (http://highlander-partners.com/)
- SignalPath Research, Advisor (http://signalpath.com/)
- The One Health Company, Special Advisor (http://www.theonehealthcompany.com/)
- RobinCare, Advisor (https://getrobincare.com/)
- KelaHealth, Inc. - peri-surgical risk prediction, Advisor (kelahealth.com)
- Patent Pending: Current patent pending for a technology that facilitates the extraction of unstructured information from medical records.

1. Importance of having all of the aspects of the data you need to create algorithms, etc

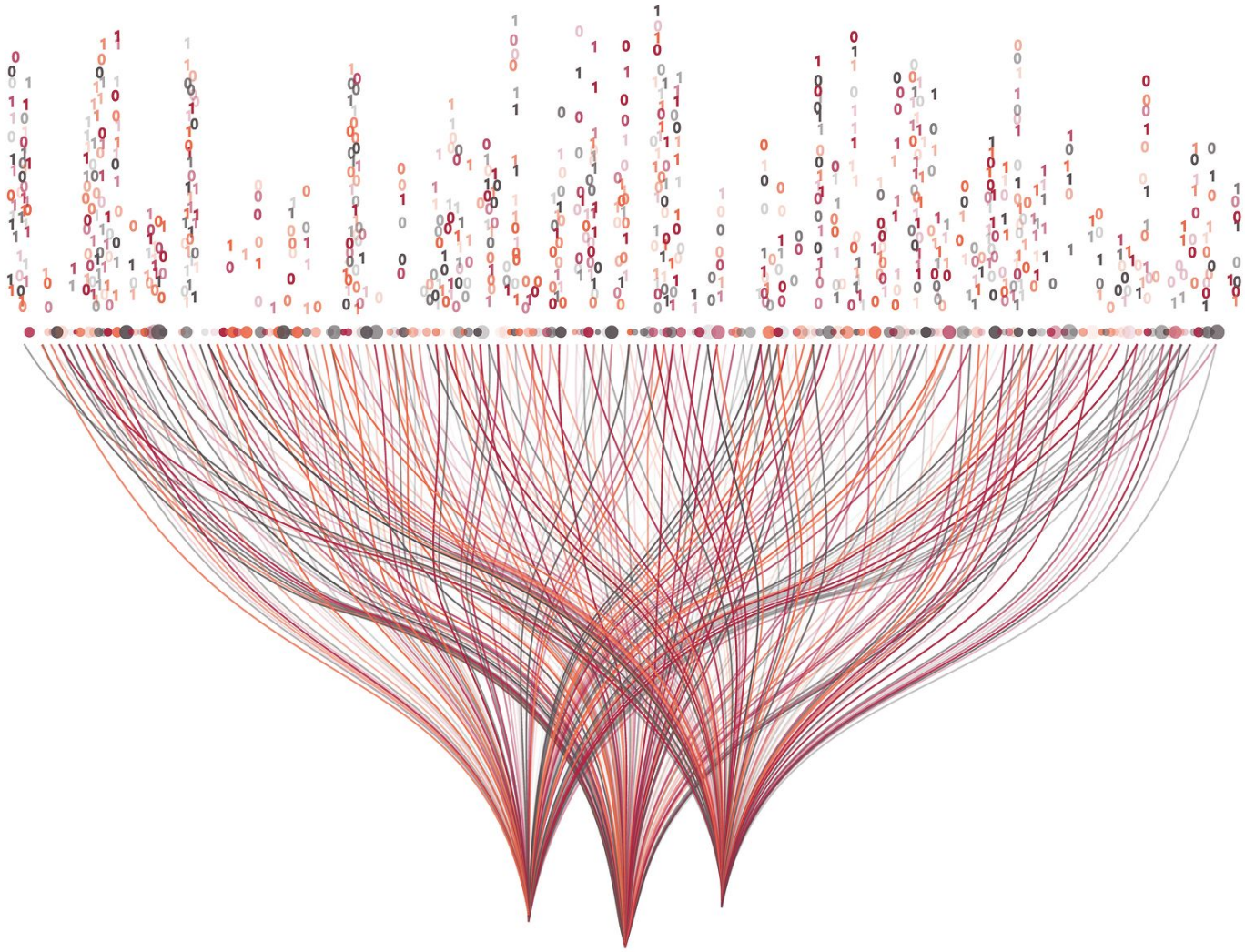
Reliable data produce reliable algorithms and reliable results

flatiron

A comprehensive view of the patient journey



*Relative timing not exact



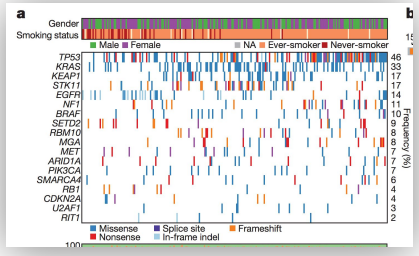
flatiron

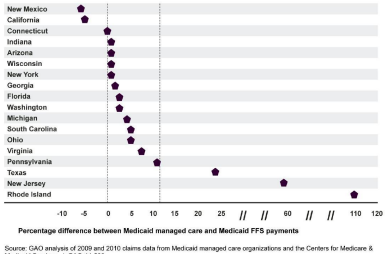
Evaluating relevancy and quality

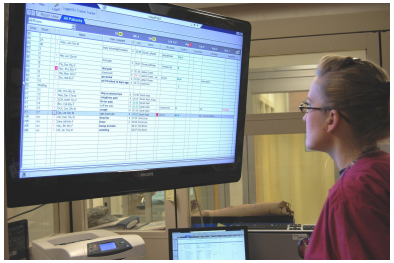
2. Different data types have different technical, contextual and governance features which influence reliability

flatiron

Different data types have different technical, contextual and governance features







Instrumentation data (e.g., genomics, imaging, immune profiling)

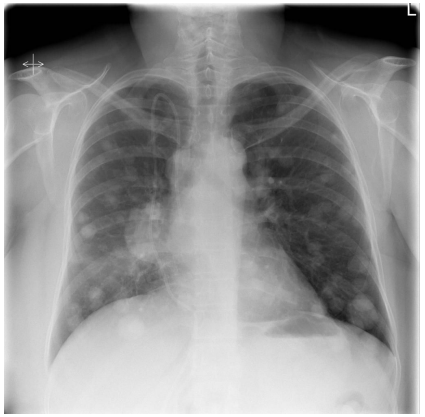
Administrative data (e.g., claims, mortality)

Longitudinal clinical data (e.g, electronic health records, registries)

Patient generated data (e.g., PROs, biosensors)

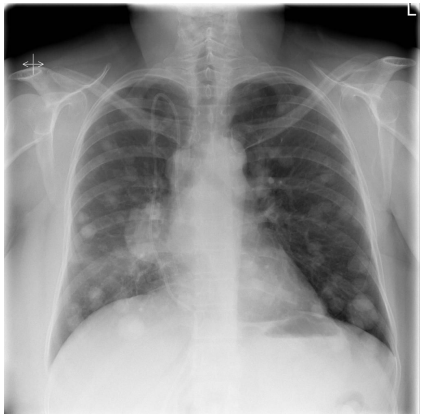
flatiron

Example: Imaging Data



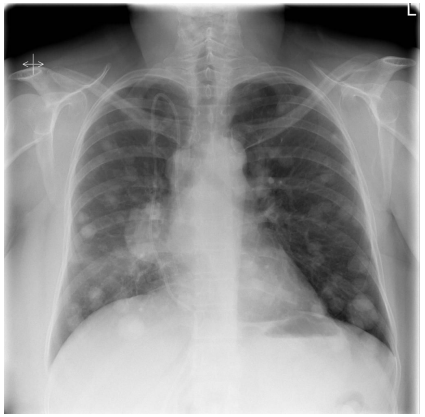
- Raw images are best if available, but require massive storage capabilities and are hard to de-identify
  - Pixalated data is not the same as an interpretation of findings
  - Information needed for subsequent analyses and algorithms is the interpretation
- An alternative is to focus on radiology reports in the EHR
  - Reports must be curated into a analytical dataset
  - Quality of the interpretation varies depending on reviewer
  - Findings in routine clinical practice vary from third party independent review for clincial trials
- Adequate interpretation of scans requires information about context
- Standardized systems for proxy data from scans (e.g., RECIST)
  - Rely on a series of features that may not be available outside of clinical trials (e.g., target lesions, prior scans)
  - May miss clinical features - must understand the pitfalls of the approach

Example: Imaging Data



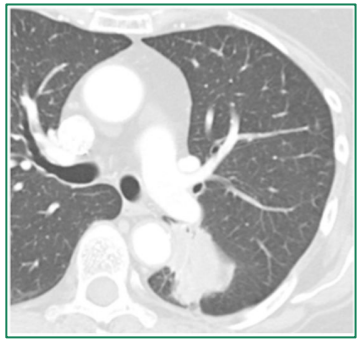
- Raw images are best if available, but require massive storage capabilities and are hard to de-identify
  - Pixalated data is not the same as an interpretation of findings
  - Information needed for subsequent analyses and algorithms is the interpretation
- An alternative is to focus on radiology reports in the EHR
  - Reports must be curated into a analytical dataset
  - Quality of the interpretation varies depending on reviewer
  - Findings in routine clinical practice vary from third party independent review for clincial trials
- Adequate interpretation of scans requires information about context
- Standardized systems for proxy data from scans (e.g., RECIST)
  - Rely on a series of features that may not be available outside of clinical trials (e.g., target lesions, prior scans)
  - May miss clinical features - must understand the pitfalls of the approach

Example: Imaging Data



- Raw images are best if available, but require massive storage capabilities and are hard to de-identify
  - Pixalated data is not the same as an interpretation of findings
  - Information needed for subsequent analyses and algorithms is the interpretation
- An alternative is to focus on radiology reports in the EHR
  - Reports must be curated into a analytical dataset
  - Quality of the interpretation varies depending on reviewer
  - Findings in routine clinical practice vary from third party independent review for clincial trials
- Adequate interpretation of scans requires information about context
- Standardized systems for proxy data from scans (e.g., RECIST)
  - Rely on a series of features that may not be available outside of clinical trials (e.g., target lesions, prior scans)
  - May miss clinical features - must understand the pitfalls of the approach

Curation of Unstructured Data



Gross Description
The specimen is received in formalin labeled with the patient's name. It consists of a 1 x 0.3 x 0.1 cm aggregate of pink-tan to red-pink soft tissue cores and fragments entirely submitted in one block.

Dictated by: GREGORY W SMITH PA
Entered: *02/06/12 - 1544 JAM*

Microscopic Description
The specimen consists of a well differentiated adenocarcinoma, favor lung primary. CK7 and TTF are positive. CK20 is negative. A colleague agrees with this malignant diagnosis.

Dictated by: THOMAS J GRIFONE MD
Entered: *02/07/12 - 1423 SML*

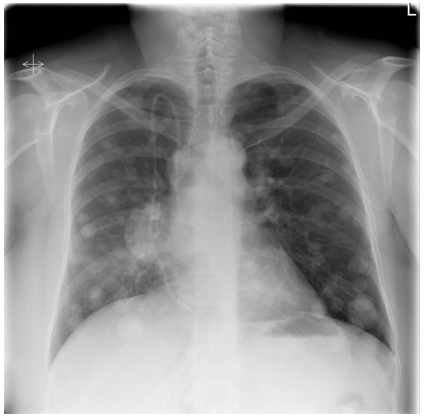
Diagnosis
SPECIMEN SUBMITTED AS TRUCUT BIOPSY LEFT LUNG NODULE:
- WELL-DIFFERENTIATED ADENOCARCINOMA, FAVOR LUNG PRIMARY.
- SEE ABOVE.

Impression:
1. Large left upper lobe bronchogenic carcinoma extending to the left hilum markedly increased in size since prior study
2. Mediastinal adenopathy increased in size particularly subcarinal space. The adenopathy in the AP window has undergone partial necrosis since previous exam.
3. Stable right apical nodularity possibly scar
4. Emphysema

6.2 → 4.7cm
US PET
121 [redacted]

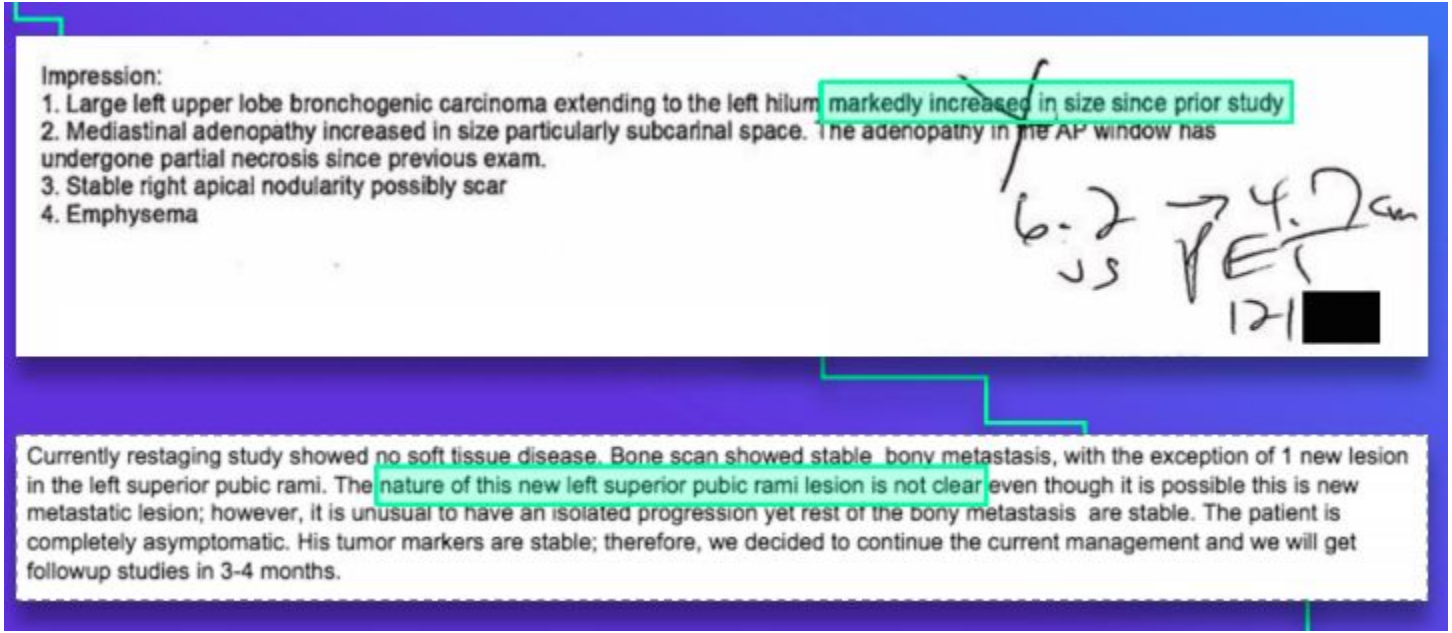
Currently restaging study showed no soft tissue disease. Bone scan showed stable bony metastasis, with the exception of 1 new lesion in the left superior pubic rami. The nature of this new left superior pubic rami lesion is not clear even though it is possible this is new metastatic lesion; however, it is unusual to have an isolated progression yet rest of the bony metastasis are stable. The patient is completely asymptomatic. His tumor markers are stable; therefore, we decided to continue the current management and we will get followup studies in 3-4 months.

Example: Imaging Data



- Raw images are best if available, but require massive storage capabilities and are hard to de-identify
  - Pixalated data is not the same as an interpretation of findings
  - Information needed for subsequent analyses and algorithms is the interpretation
- An alternative is to focus on radiology reports in the EHR
  - Reports must be curated into a analytical dataset
  - Quality of the interpretation varies depending on reviewer
  - Findings in routine clinical practice vary from third party independent review for clincial trials
- Adequate interpretation of scans requires information about context
- Standardized systems for proxy data from scans (e.g., RECIST)
  - Rely on a series of features that may not be available outside of clinical trials (e.g., target lesions, prior scans)
  - May miss clinical features - must understand the pitfalls of the approach

Context: Accounting for Changing Interpretations Over Time

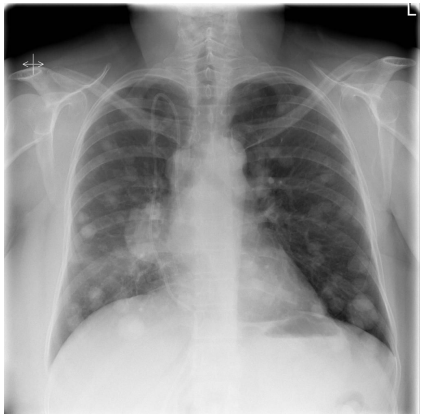


Impression:
1. Large left upper lobe bronchogenic carcinoma extending to the left hilum markedly increased in size since prior study
2. Mediastinal adenopathy increased in size particularly subcarinal space. The adenopathy in the AP window has undergone partial necrosis since previous exam.
3. Stable right apical nodularity possibly scar
4. Emphysema

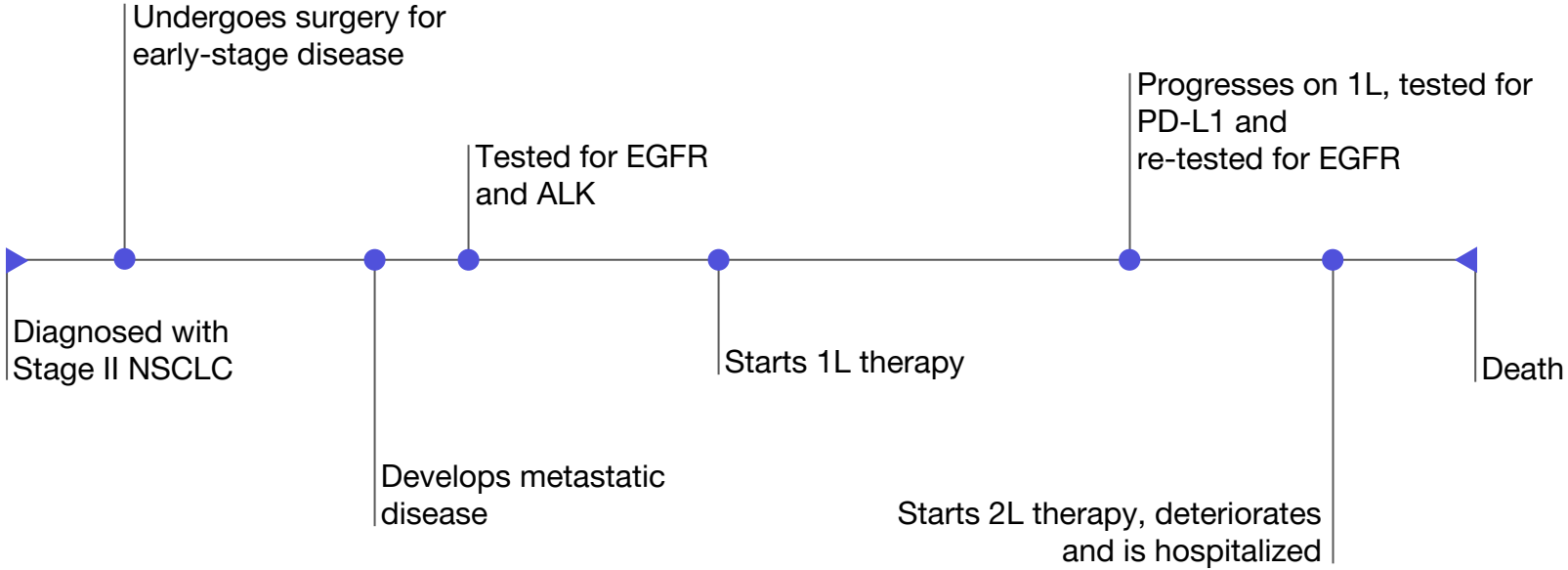
6.2 → 4.7 cm
VS PET
121 [redacted]

Currently restaging study showed no soft tissue disease. Bone scan showed stable bony metastasis, with the exception of 1 new lesion in the left superior pubic rami. The nature of this new left superior pubic rami lesion is not clear even though it is possible this is new metastatic lesion; however, it is unusual to have an isolated progression yet rest of the bony metastasis are stable. The patient is completely asymptomatic. His tumor markers are stable; therefore, we decided to continue the current management and we will get followup studies in 3-4 months.

Example: Imaging Data



- Raw images are best if available, but require massive storage capabilities and are hard to de-identify
  - Pixalated data is not the same as an interpretation of findings
  - Information needed for subsequent analyses and algorithms is the interpretation
- An alternative is to focus on radiology reports in the EHR
  - Reports must be curated into a analytical dataset
  - Quality of the interpretation varies depending on reviewer
  - Findings in routine clinical practice vary from third party independent review for clincial trials
- **Adequate interpretation of scans requires information about context**
- Standardized systems for proxy data from scans (e.g., RECIST)
  - Rely on a series of features that may not be available outside of clinical trials (e.g., target lesions, prior scans)
  - May miss clinical features - must understand the pitfalls of the approach



Clinical events are a combination of clinical, pathological, radiological, & biomarker data - *in context*

Undergoes surgery for early-stage disease

Tested for EGFR and ALK

Progresses on 1L, tested for PD-L1 and re-tested for EGFR

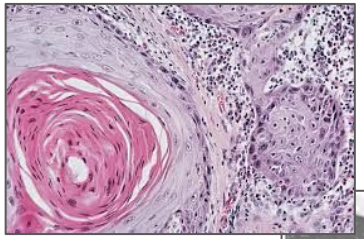
Diagnosed with Stage II NSCLC

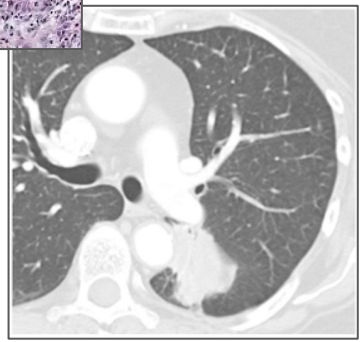
Starts 1L therapy

Death

Develops metastatic disease

Starts 2L therapy, deteriorates and is hospitalized





Gross Description
The specimen is received in formalin labeled with the patient's name. It consists of a 1 x 0.3 x 0.1 cm aggregate of pink-tan to red-pink soft tissue cores and fragments entirely submitted in one block.

Dictated by: GREGORY W SMITH PA
Entered: *02/06/12 - 1544 JAM*

Microscopic Description
The specimen consists of a well differentiated adenocarcinoma, favor lung primary. CK7 and TTF are positive. CK20 is negative. A colleague agrees with this malignant diagnosis.

Dictated by: THOMAS J GRIFONE MD
Entered: *02/07/12 - 1423 SML*

Diagnosis
SPECIMEN SUBMITTED AS TRUCUT BIOPSY LEFT LUNG NODULE:
- WELL-DIFFERENTIATED ADENOCARCINOMA, FAVOR LUNG PRIMARY.
- SEE ABOVE.

flatiron

Undergoes surgery for early-stage disease

Tested for EGFR and ALK

Progresses on 1L, tested for PD-L1 and re-tested for EGFR

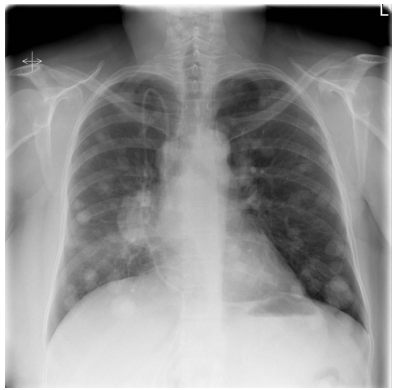
Diagnosed with Stage II NSCLC

Starts 1L therapy

Death

Develops metastatic disease

Starts 2L therapy, deteriorates and is hospitalized



Path?

flatiron

Undergoes surgery for early-stage disease

Tested for EGFR and ALK

Progresses on 1L, tested for PD-L1 and re-tested for EGFR

Diagnosed with Stage II NSCLC

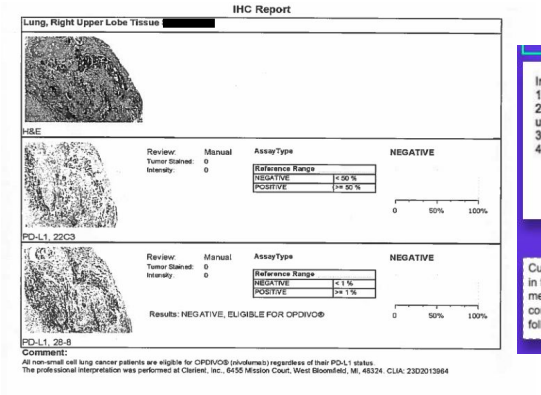
Develops metastatic disease

Starts 1L therapy

Starts 2L therapy, deteriorates and is hospitalized

Death

Time to progression is dependent on when patient is evaluated


IHC Report

Lung, Right Upper Lobe Tissue [redacted]

H&E

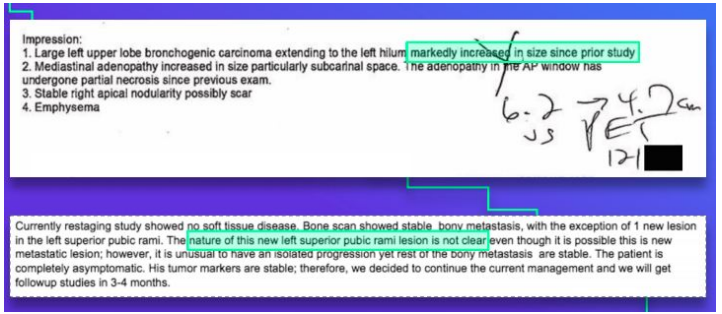
Review: Manual
Tumor Stained: 0
Intensity: 0
AssayType
NEGATIVE
Reference Range
NEGATIVE < 50 %
POSITIVE >= 50 %
0 50% 100%

PD-L1, 22C3

Review: Manual
Tumor Stained: 0
Intensity: 0
AssayType
NEGATIVE
Reference Range
NEGATIVE < 1 %
POSITIVE >= 1 %
Results: NEGATIVE, ELIGIBLE FOR OPDIVO®
0 50% 100%

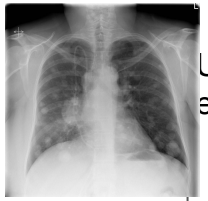
PD-L1, 28-8

Comment:
All non-small cell lung cancer patients are eligible for OPDIVO® (nivolumab) regardless of their PD-L1 status.
The professional interpretation was performed at Clarient, Inc., 6455 Mission Court, West Bloomfield, MI, 48324. CLIA: 23D2013964

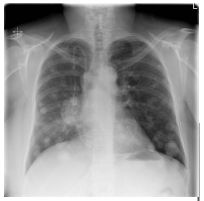

Impression:
1. Large left upper lobe bronchogenic carcinoma extending to the left hilum markedly increased in size since prior study
2. Mediastinal adenopathy increased in size particularly subcarinal space. The adenopathy in the AP window has undergone partial necrosis since previous exam.
3. Stable right apical nodularity possibly scar
4. Emphysema

6.2 → 4.7cm
JS PET
121 [redacted]

Currently restaging study showed no soft tissue disease. Bone scan showed stable bony metastasis, with the exception of 1 new lesion in the left superior pubic rami. The nature of this new left superior pubic rami lesion is not clear even though it is possible this is new metastatic lesion; however, it is unusual to have an isolated progression yet rest of the bony metastasis are stable. The patient is completely asymptomatic. His tumor markers are stable; therefore, we decided to continue the current management and we will get followup studies in 3-4 months.



Undergoes surgery for early-stage disease



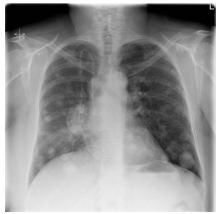
Progresses on 1L, tested for PD-L1 and re-tested for EGFR

Tested for EGFR and ALK

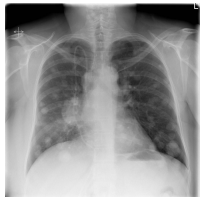
Diagnosed with Stage II NSCLC

Starts 1L therapy

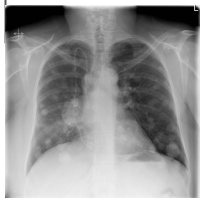
Death



Develops metastatic disease

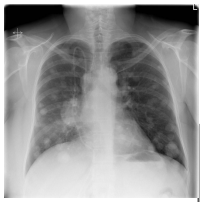


Starts 2L therapy, deteriorates and is hospitalized



Complete datasets increase reliability

Undergoes surgery for early-stage disease



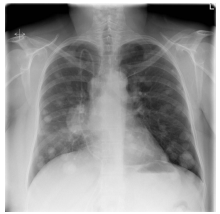
Progresses on 1L, tested for PD-L1 and re-tested for EGFR

Tested for EGFR and ALK

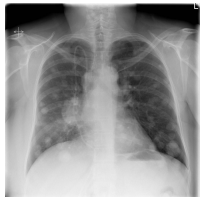
Diagnosed with Stage II NSCLC

Starts 1L therapy

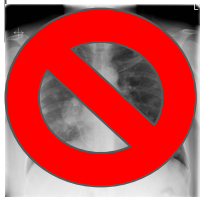
Death



Develops metastatic disease

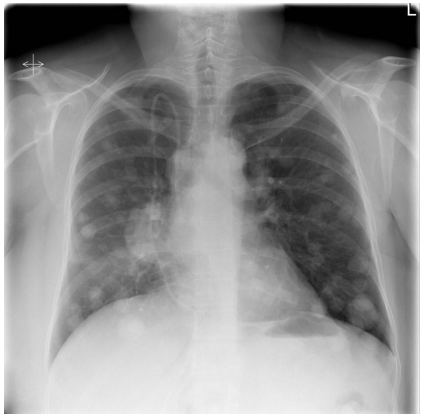


Starts 2L therapy, deteriorates and is hospitalized



Complete datasets increase reliability

Example: Imaging Data

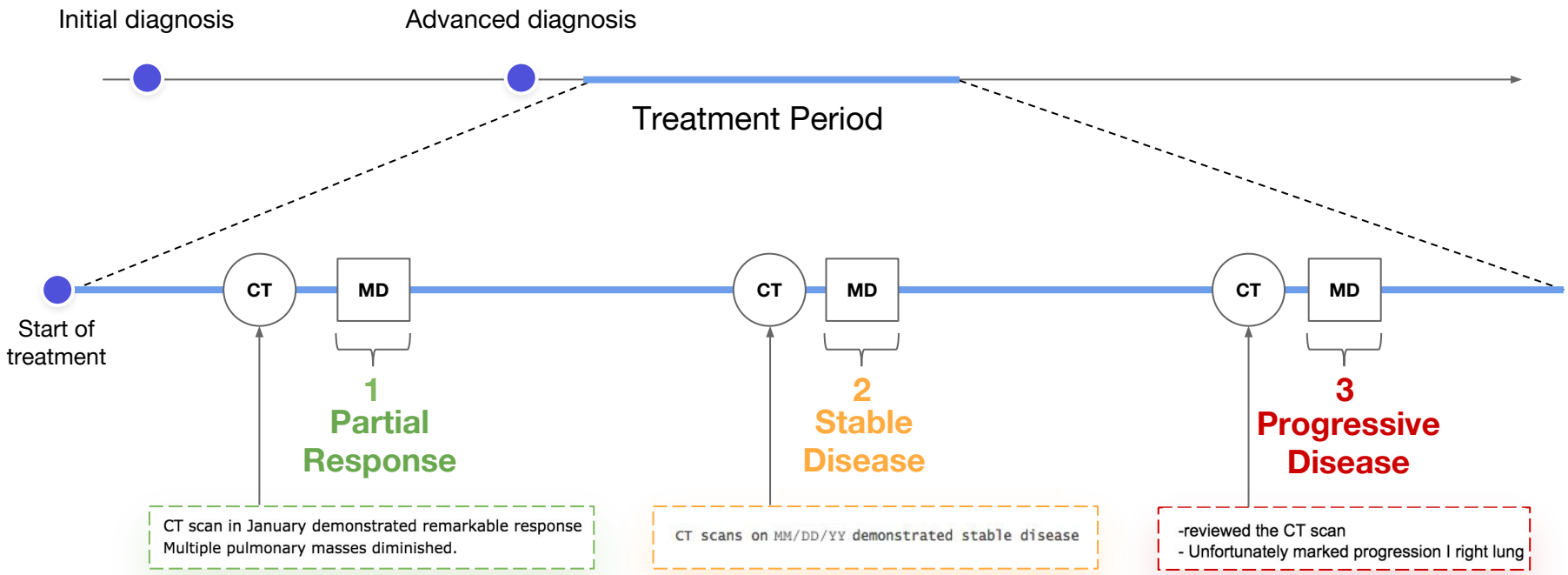


- Raw images are best if available, but require massive storage capabilities and are hard to de-identify
  - Pixalated data is not the same as an interpretation of findings
  - Information needed for subsequent analyses and algorithms is the interpretation
- An alternative is to focus on radiology reports in the EHR
  - Reports must be curated into a analytical dataset
  - Quality of the interpretation varies depending on reviewer
  - Findings in routine clinical practice vary from third party independent review for clincial trials
- Adequate interpretation of scans requires information about context
- Standardized systems for proxy data from scans (e.g., RECIST)
  - Rely on a series of features that may not be available outside of clinical trials (e.g., target lesions, prior scans)
  - May miss clinical features - must understand the pitfalls of the approach

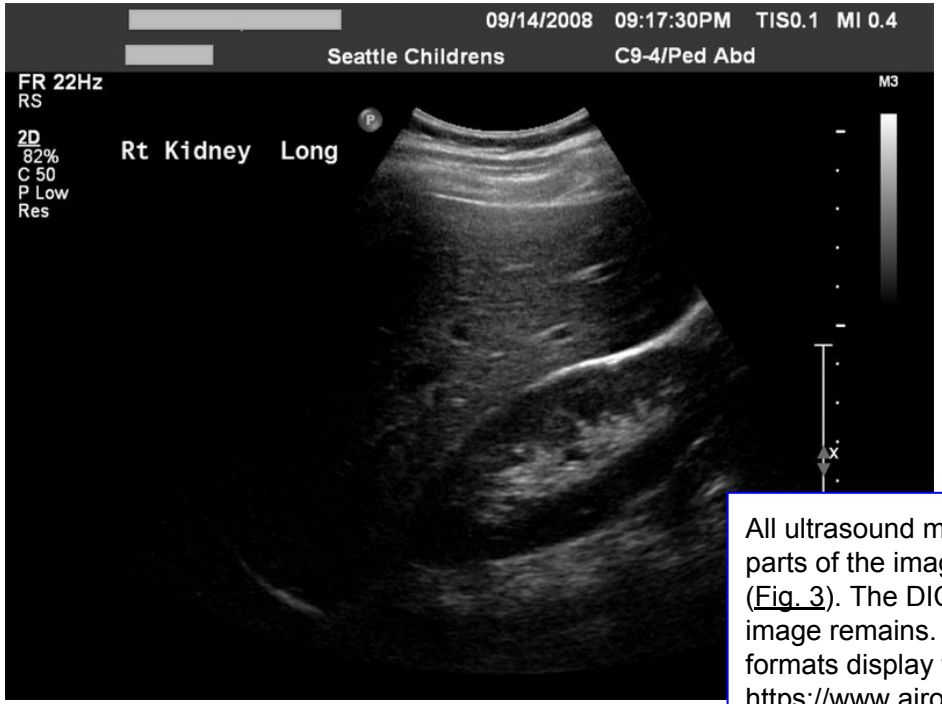
rwTR Assessment Categories

Category	Description
Complete Response (CR)	Complete resolution of all visible disease.
Partial Response (PR)	Partial reduction in size of visible disease in some or all areas without any areas of increase in visible disease. Captures a decrease in disease volume even though disease is still present.
Stable Disease (SD)	No change in overall size of visible disease. Also includes cases where some lesions increased in size and some lesions decreased in size.
Progressive Disease (PD)	Increase in visible disease and/or presence of any new lesions. Includes cases where the clinician indicates progressive disease, PD, or POD as the overall assessment.
Pseudoprogression	Clinician indicates pseudoprogression or related terminology (e.g., tumor flare).
Indeterminate response	Clinician specifically indicates that the response is “indeterminate” or “uncertain,” or if the clinician’s interpretation of the scan(s) cannot be mapped to one of the above assessment categories.
Not documented	Clinician’s note references this response imaging (e.g., “Patient had recent scan”) but does not mention any assessment of tumor response.

Response is tied to exposure to a therapy



Governance and De-ID

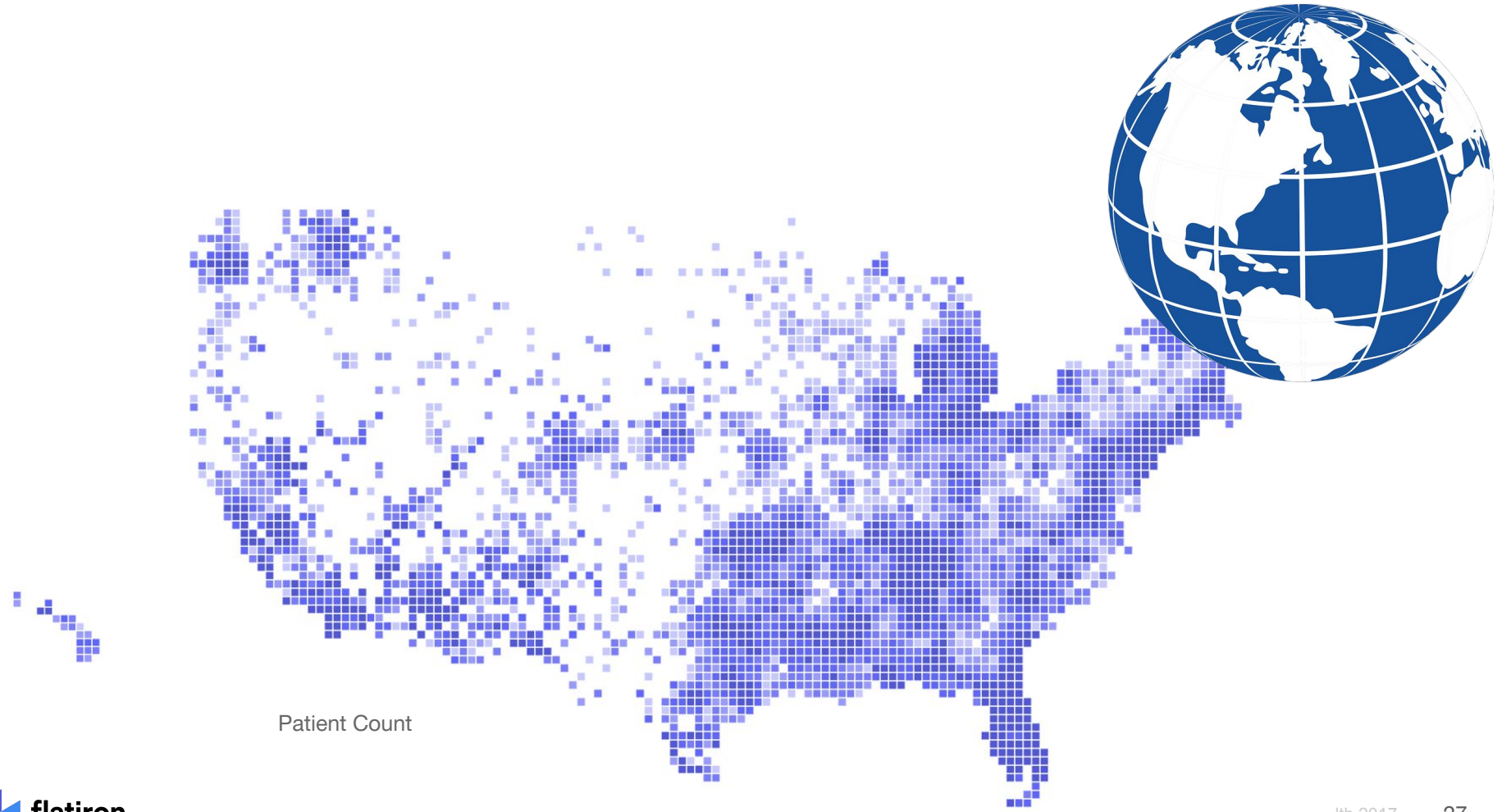


- Who approves access?
- What kind of access?
- How is access QA'ed?

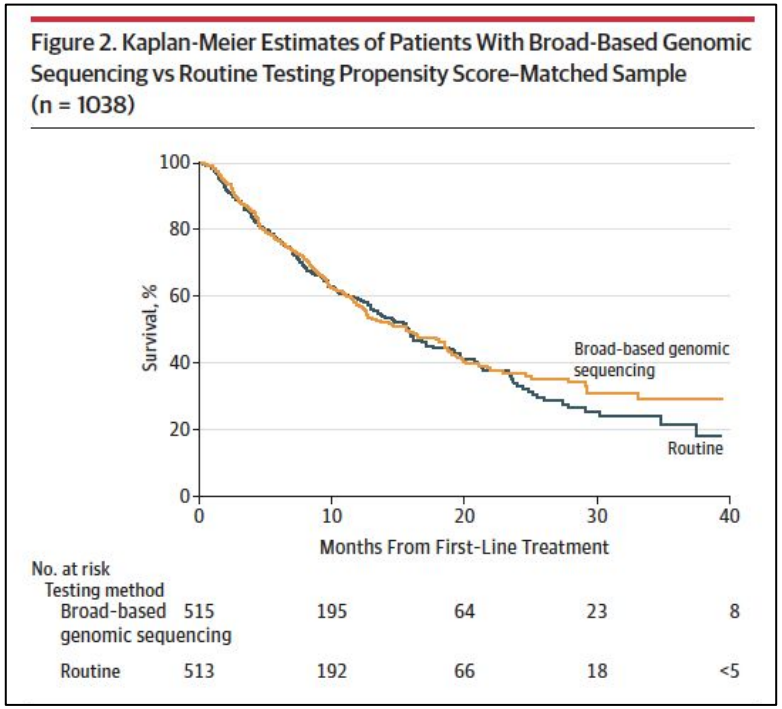
All ultrasound machines store the patient's name and MRN as intrinsic parts of the image and are displayed whenever the image is displayed (Fig. 3). The DICOM headers can be deidentified, but the PHI in the image remains. In addition to the patient's name and MRN, some formats display the birth date, scan date and time, or facility name. https://www.ajronline.org/doi/full/10.2214/AJR.13.11789

flatiron

3. Representativeness of datasets (genomic, geographic); risk of bias in the underlying data



Does genomic testing improve survival for lung cancer patients?

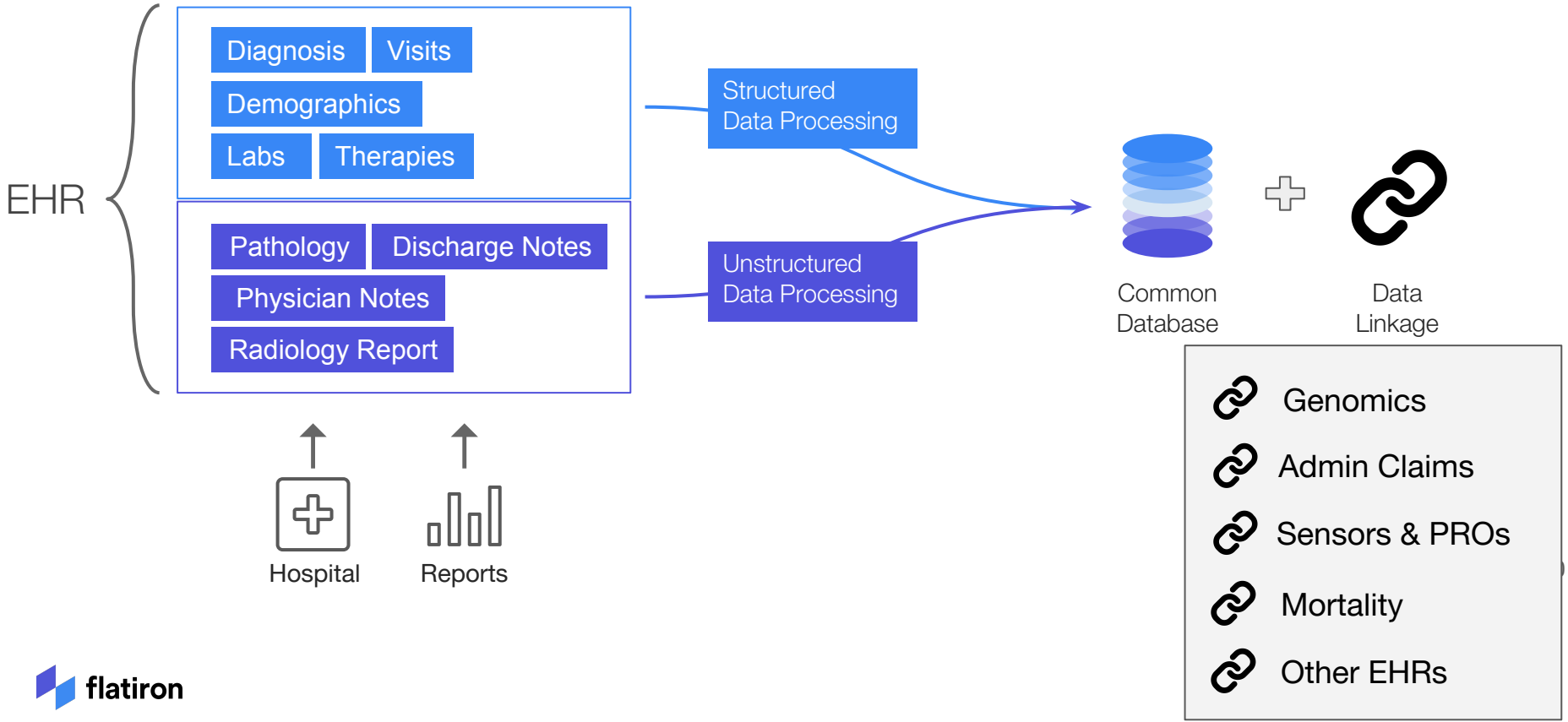


Total	No. (%) Broad-Based Genomic Sequencing (n=875)	Routine Testing (n=4813)	*P* Value[a]
Age at diagnosis, y			
≤45	41 (4.7)	89 (1.9)	
46-55	106 (12.1)	588 (12.2)	
56-65	277 (31.7)	1367 (28.4)	<.001
66-75	307 (35.1)	1642 (34.1)	
76-85	144 (16.5)	1127 (23.4)	
Sex			
Male	407 (46.5)	2247 (46.7)	.93
Female	468 (53.5)	2566 (53.3)	
Race/ethnicity			
Non-Hispanic white	594 (67.9)	3023 (62.8)	
Non-Hispanic black	49 (5.6)	380 (7.9)	
Hispanic or Latino	28 (3.2)	166 (3.5)	<.001
Asian	39 (4.5)	172 (3.6)	
Other	98 (11.2)	377 (7.8)	
Unknown	67 (7.7)	695 (14.4)	

Presley et al. *JAMA* 2018

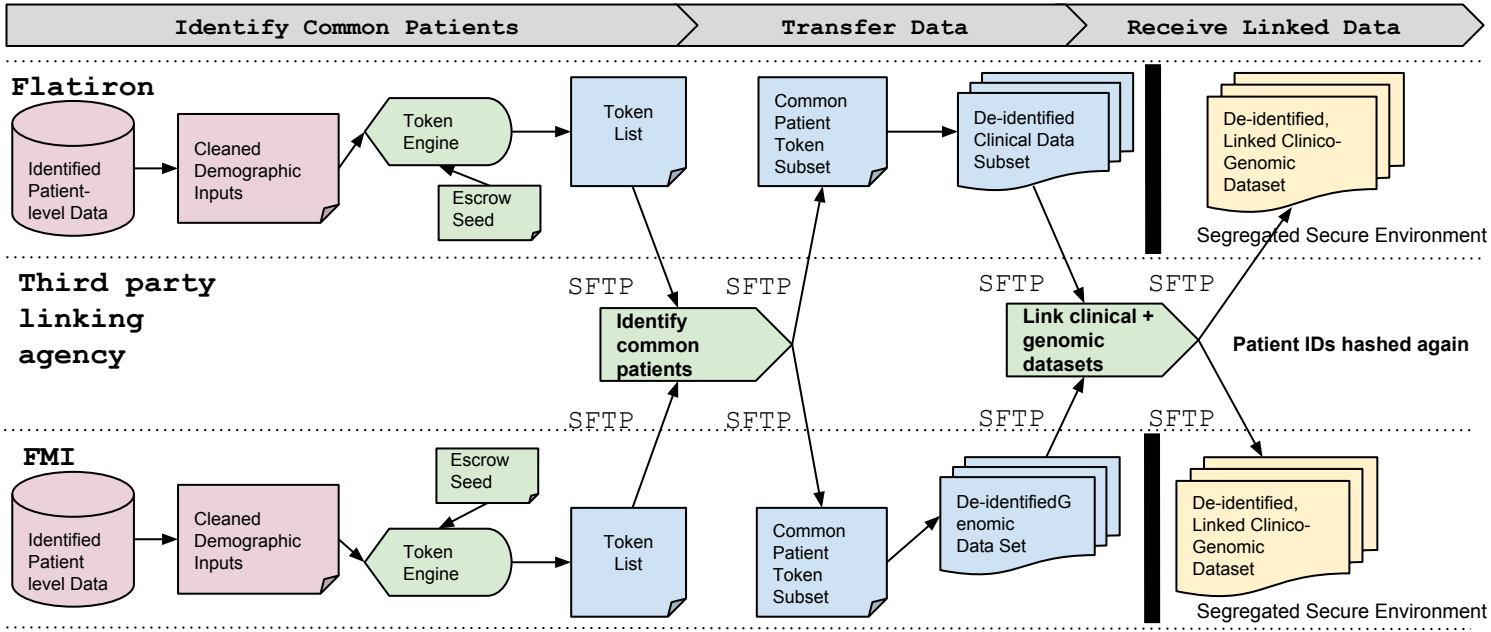
4. Methods to resolve data gaps

Data linkage to fill gaps



flatiron

Data linkage solutions



Red - Identified Data
Blue - De-identified Data
Yellow - De-identified Linked Data
Green - Third party linking agency's technology

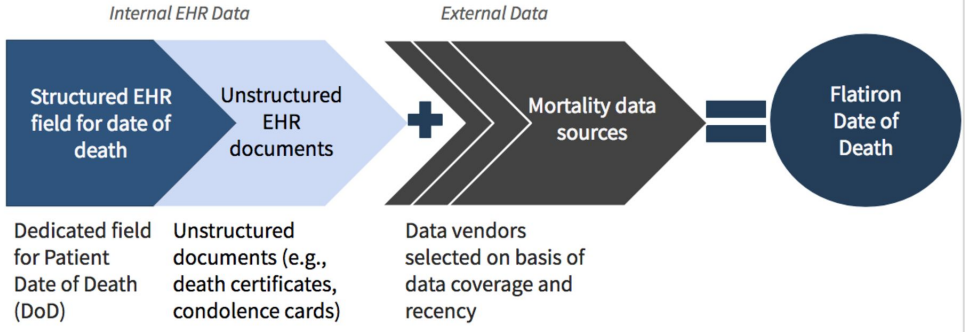
This entire process is repeated quarterly at FH-FMI for longitudinal refresh of the data (e.g., new chart abstraction)

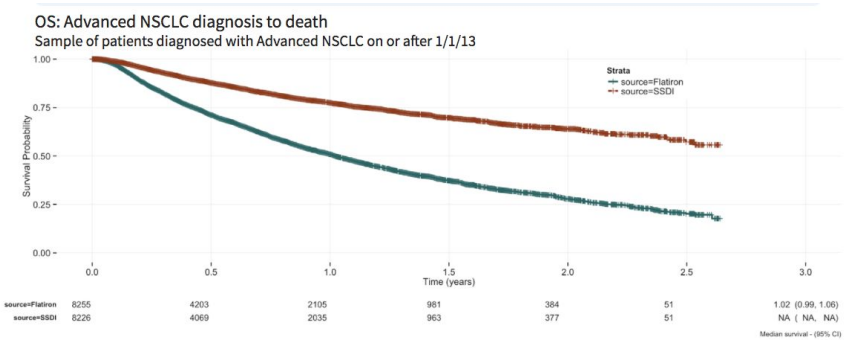
Final linked clinico-genomic dataset cannot be re-identified or linked to other FH or FMI datasets (unique patient ID)

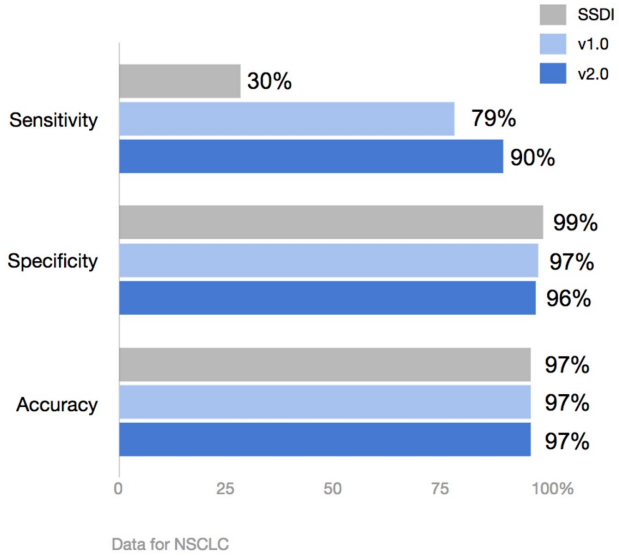
flatiron FOUNDATION MEDICINE®

Evaluate data against a reference standard

E.g., gold standard = National Death Index

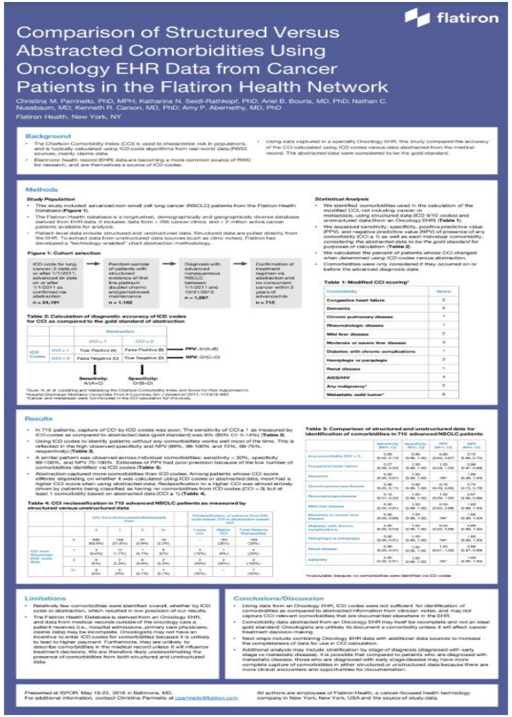






flatiron

Developing proxy and surrogate variables when needed



Comparison of structured vs. abstracted comorbidities

Results:

- Capture of CCI by ICD codes was poor. The sensitivity of CCI ≥ 1 as measured by ICD codes as compared to abstracted data (gold standard) was 9% (95% CI: 5-14%)

- Using ICD codes to identify patients without any comorbidities works well most of the time. This is reflected in the high observed specificity and NPV (99%, CI: 98-100% and 72%, CI: 69-75%, respectively)

- Abstraction captured more comorbidities than ICD codes. However, comorbidity data abstracted from an Oncology EHR may itself be incomplete and not an ideal gold standard. Oncologists are unlikely to document a comorbidity unless it will affect cancer treatment decision-making

5. Issues in Precision Medicine: Available biologic/biomarker data, small cohorts

flatiron

Deep Phenotype is Needed

In an era of abundant data, merging biological information with deep clinical phenotype is more important than ever

Nucleic Acids Research, 2013, 1–9
doi:10.1093/nar/gkt1026

The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data

DEEP PHENOTYPING

The details of disease

Precision medicine demands precise matching of deep genomic and phenotypic models — and the deeper you go, the more you know.

S14 | NATURE | VOL 527 | 5 NOVEMBER 2015

Warner *et al. Genome Medicine* (2016) 8:113
DOI 10.1186/s13073-016-0371-3

Genome Medicine

REVIEW Open Access

Integrating cancer genomic data into electronic health records

Jeremy L. Warner[1,2,3*], Sandeep K. Jain[2,4] and Mia A. Levy[1,2,3]

flatiron

Precision medicine often yields small cohorts

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Vemurafenib in Multiple Nonmelanoma Cancers with *BRAF* V600 Mutations

David M. Hyman, M.D., Igor Puzanov, M.D., Vivek Subbiah, M.D., Jason E. Faris, M.D., Ian Chau, M.D., Jean-Yves Blay, M.D., Ph.D., Jürgen Wolf, M.D., Ph.D., Noopur S. Raje, M.D., Eli L. Diamond, M.D., Antoine Hollebecque, M.D., Radj Gervais, M.D., Maria Elena Elez-Fernandez, M.D., Antoine Italiano, M.D., Ph.D., Ralf-Dieter Hofheinz, M.D., Manuel Hidalgo, M.D., Ph.D., Emily Chan, M.D., Ph.D., Martin Schuler, M.D., Susan Frances Lasserre, M.Sc., Martina Makrutzki, M.D., Florin Sirzen, M.D., Ph.D., Maria Luisa Veronese, M.D., Josep Tabernero, M.D., Ph.D., and José Baselga, M.D., Ph.D.

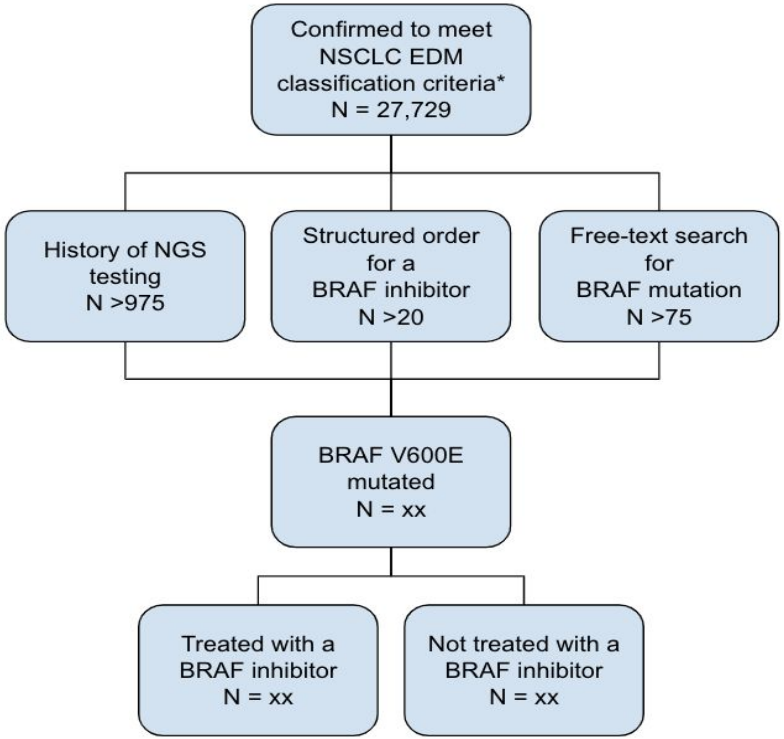
ABSTRACT

BACKGROUND

BRAF V600 mutations occur in various nonmelanoma cancers. We undertook a histology-independent phase 2 "basket" study of vemurafenib in *BRAF* V600 mutation–positive nonmelanoma cancers.

METHODS

We enrolled patients in six prespecified cancer cohorts: patients with all other



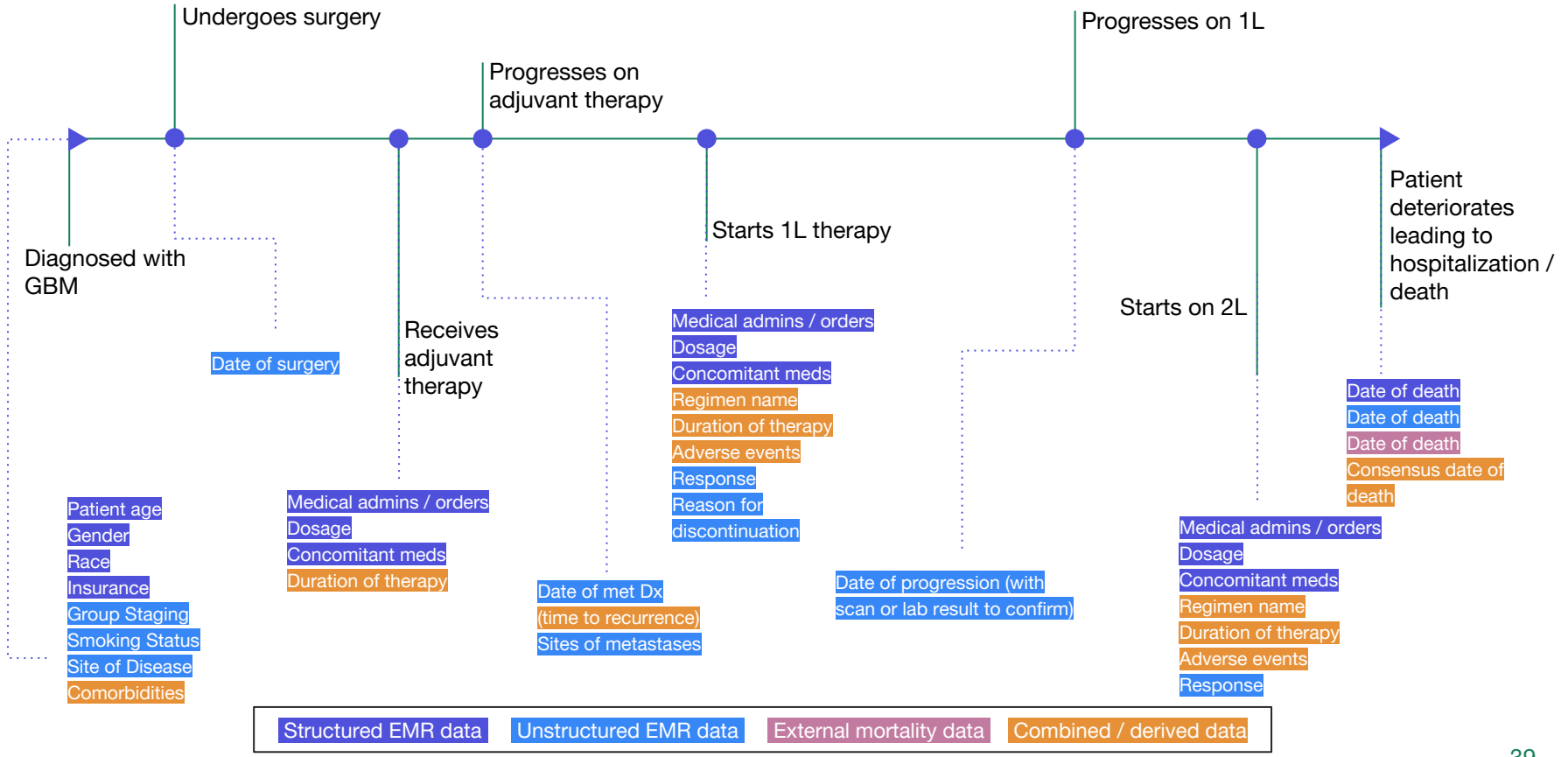
6. Data reliability and quality

flatiron

Reliable data produce reliable algorithms and reliable results

flatiron

Documentation of source, quality and provenance



*Relative timing not exact

Need a consistent approach to documenting completeness and quality

Appendix B: Flatiron Health PD-L1: Inter-rater agreement and kappas on abstracted variable

Project: **FDA** **PD-1 inhibitors in aNSCLC**

Note: For questions where a high percentage of patients have a common answer (e.g., PD-L1 testing status), kappa may be significantly lower than inter-rater agreement. In these cases, it may be more accurate to use inter-rater agreement to measure reliability of the data.

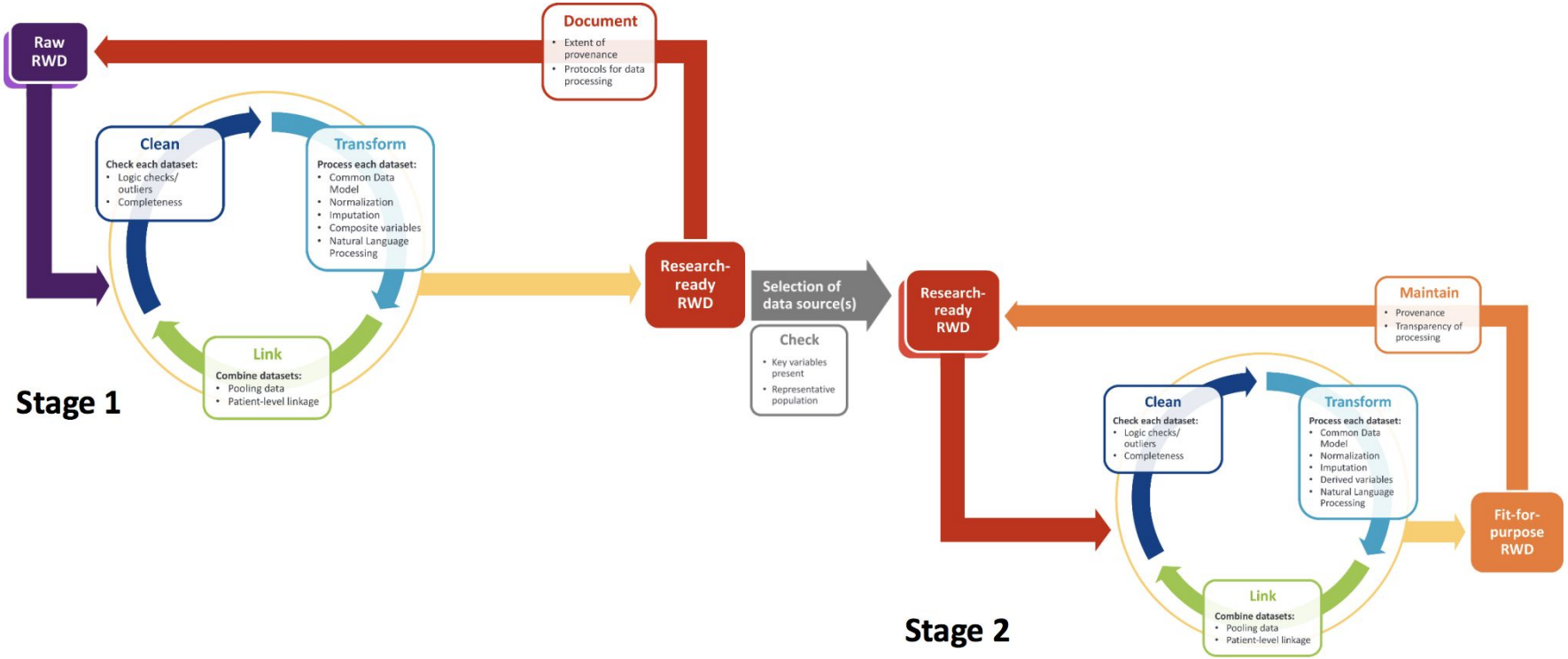
Kappas scale	
Almost perfect	0.8 to 1.0
Substantial	0.6 to 0.8
Moderate	0.4 to 0.6
Fair	0.2 to 0.4
Slight	0 to 0.2

Table: Enhanced_AdvancedNSCLC

Summary of variable inter-rater agreement and kappas

Variable	Description of variable	Corresponding question(s) on abstraction form	Question type	Inter-rater agreement (exact day for dates)	Kappa (exact agreement)	Kappa (30-day window for dates)
DiagnosisDate	Date of initial diagnosis	Enter the date of initial diagnosis	date	0.795	0.794	0.902
AdvancedDiagnosisDate	Date of diagnosis of advanced disease: first recurrence or metastasis	Enter the date of the first diagnosis of metastatic or advanced NSCLC	date	0.695	0.695	0.796
MetastaticDiagnosisDate	Date of diagnosis of metastatic disease	Enter the date of initial diagnosis [for ~55% of patients in the cohort who are diagnosed metastatic]	date	0.795	0.794	0.902
		Enter the date of distant metastatic diagnosis [for ~45% of patients in the cohort who are diagnosed non-metastatic]	date	0.527	0.476	0.557
Histology	Histology	Select the histology	drop down	0.947	0.894	
GroupStage	Group stage at time of initial diagnosis	Select the group stage	drop down	0.848	0.768	
SmokingStatus	Documented history of smoking	Smoking status	drop down	0.934	0.695	
EgfrTested	Indicator of whether the tumor was tested for a EGFR mutation	Was the tumor tested for a EGFR mutation?	boolean	0.927	0.84	
AlkTested	Indicator of whether the tumor was tested for an ALK rearrangement	Was the tumor tested for an ALK rearrangement?	boolean	0.901	0.791	
PdL1Tested	Indicator of whether the tumor was tested for PD-L1 expression	Was the tumor tested for PD-L1 expression?	boolean	0.901	0.547	
KrasTested	Indicator of whether the tumor was tested for a KRAS mutation	Was the tumor tested for a KRAS mutation?	boolean	0.894	0.728	
Ros1Tested	Indicator of whether the tumor was tested for a ROS-1 rearrangement	Was the tumor tested for a ROS-1 rearrangement?	boolean	0.881	0.725	

Research-ready databases



7. Recency and availability

flatiron

8. Privacy and security

flatiron

9. Policy Considerations

Reliable data produce reliable algorithms and reliable results

flatiron

Policy Considerations

- Require accurate complete datasets that incorporate many data types
- Details about the data matter
- Biased and unreliable data lead to biased and unreliable algorithms
- Accessible data at point of care
- While maintaining patient privacy
- Solving this requires input of many types of experts - clinical, analytic, software, hardware, privacy, etc
- Stimulate tech innovation that solves all of these problems - do not skimp
- Expect transparency about data reliability and quality - as well as accuracy of output from algorithms
- Create standards for documenting reliability, quality and accuracy at the data source, dataset and algorithm level
- Develop a the workforce

flatiron