PANEL 3: DEVELOPMENT OF REAL WORLD ENDPOINTS

Amy Abernethy, Flatiron Health
Eric Rubin, Merck Research Labs
Richard Pazdur, US FDA

FLATIRON

Panel III:
Development of Real-World Endpoints

June 16, 2016

FRIENDS of CANCER RESEARCH

ALEXANDRIA SUMMIT®

Reprise to Panels 1 & 2

1. Assessment of reliable valid endpoints is a critical aspect of research that includes real-world data
   - E.g., Effectiveness of BRAF-inhibitors for BRAF+ NSCLC
   - E.g., Ramucirumab / PD-1 treatment sequence
2. What endpoints?
   - Effectiveness, Safety, PROs, Composite, Other
3. Need a reliable way to develop and assess endpoints
4. For what purposes can we use these endpoints?
   - Retrospective analyses? Publications ? Post-approval regulatory submissions? Payer discussions? Prospective pragmatic trial?
   - Other?

Panel 3 - Agenda

1. Approach to developing and evaluating real-world endpoints
2. Development and evaluation of the real-world progression (rwP) endpoint
3. Expansion to other endpoints including patient-generated health data
4. Proposed collaborative prospective real-world data study

Approach to developing and evaluating real-world endpoints

Objectives for real-world endpoints

Deliver endpoints for patients in the real world that are:

- Based on existing data captured routinely from the chart in the real world
  - Take advantage of data presented in electronic health records (EHRs) whenever possible
- Tied to source evidence (e.g., radiographic, laboratory, pathologic, clinical assessment)
- Shown to be a meaningful endpoint based on a predefined experimental validation framework
- Accepted by oncologists, researchers, regulatory bodies, and industry, with guidance around suitable applications

FLATIRON

© 2015 Flatiron Health, Inc. Proprietary and confidential.

Flatiron aggregates & processes EHR data at scale

240+

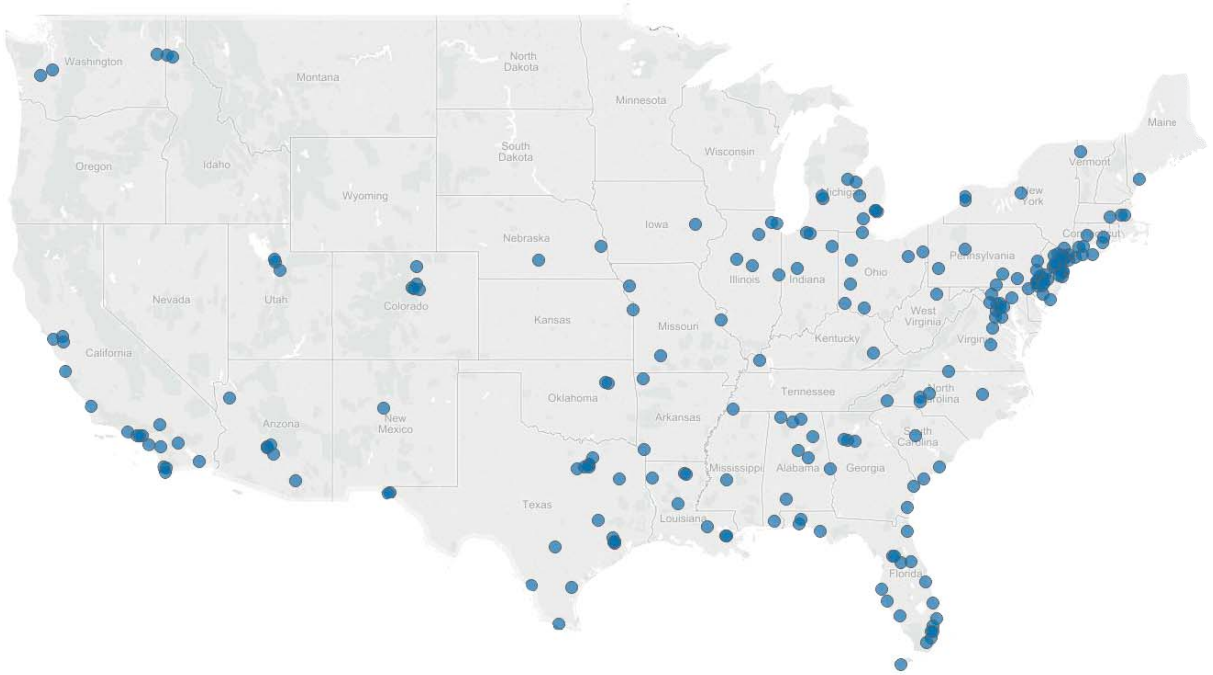
Cancer Clinics

2,000+

Clinicians

1,300,000+

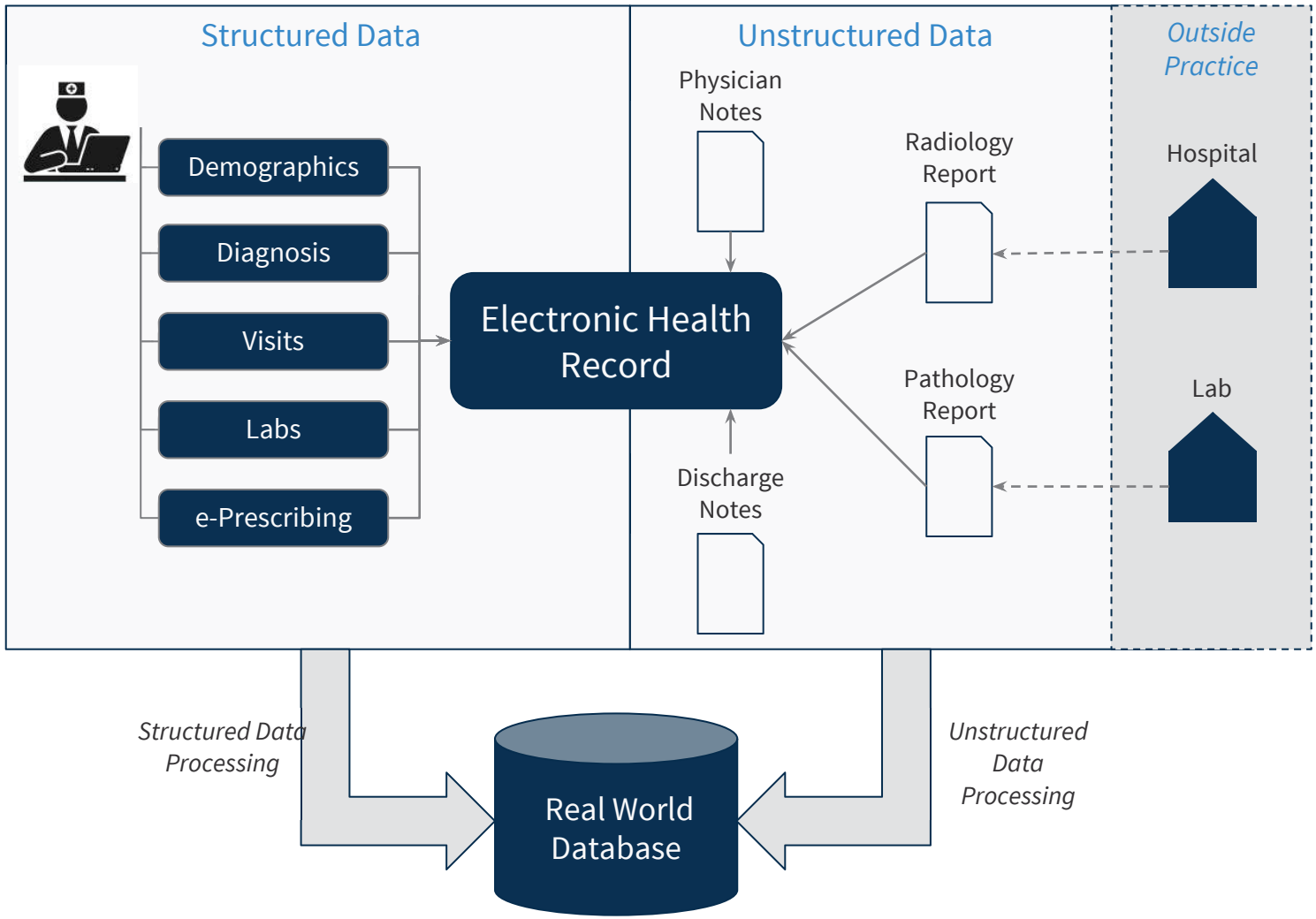
Active Cancer Patients



FLATIRON

© 2016 Flatiron Health, Inc. Proprietary and confidential.

The electronic health record comprises structured and unstructured chart data



FLATIRON

Linking high-value datasets: Real-world mortality endpoint example

Processed structured & unstructured EHR data

- Genomic data
- Closed claims data
- Prospective data capture
- Patient-reported outcomes
- **Mortality data**

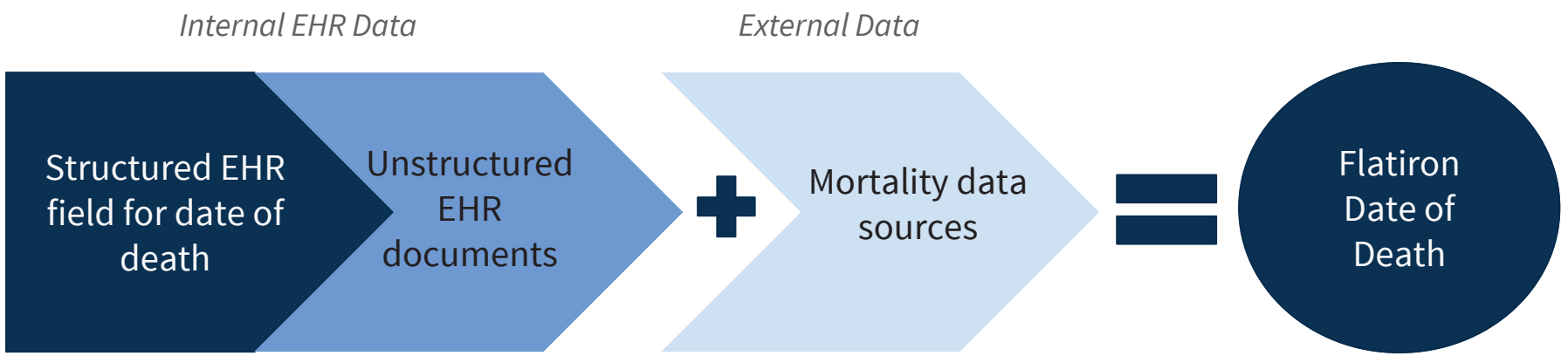
FLATIRON

© 2016 Flatiron Health, Inc. Proprietary and confidential.

Linking high-value datasets: Real-world mortality endpoint example

Processed structured & unstructured EHR data

- Genomic data
- Closed claims data
- Prospective data capture
- Patient-reported outcomes
- **Mortality data**

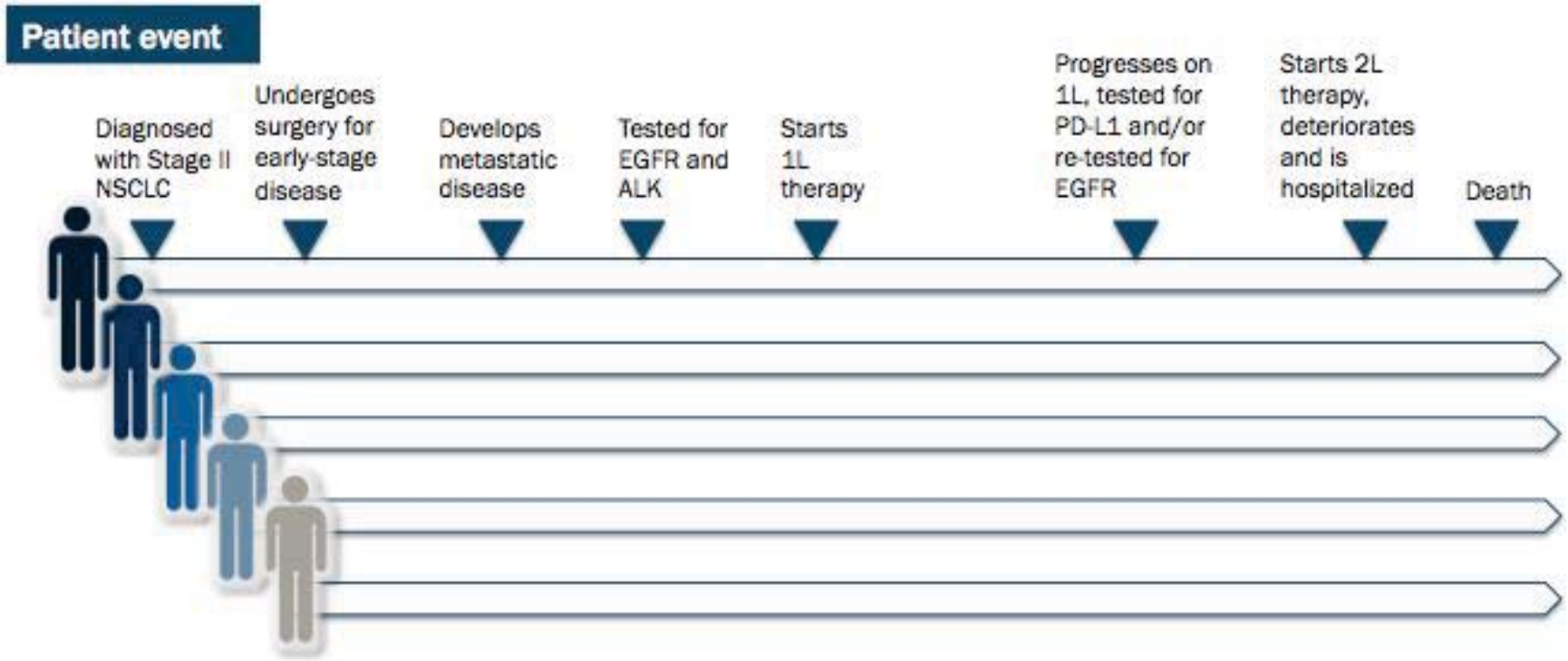


Combining data sources enables development of a high-quality consensus date of death for patients across the database

FLATIRON

© 2016 Flatiron Health, Inc. Proprietary and confidential.

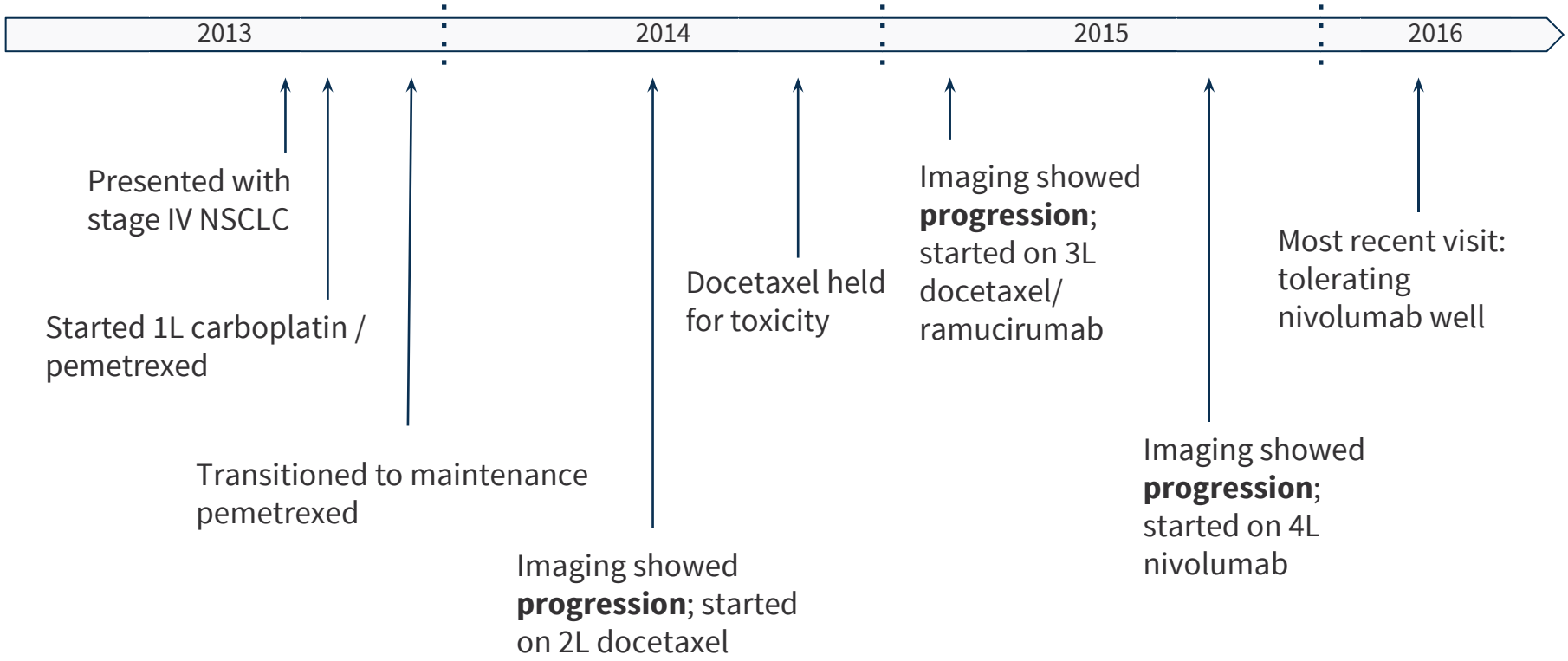
Developing a definition that works across many diverse patient stories



- *Objective is to aggregate a cohort of patient stories for research purposes, using replicable processes*
- *Need to develop a solution that works across many different patients, clinicians, documentation habits, EHRs, health systems, and diseases settings*

FLATIRON

Patient Journey: Female in Her Late 60s with NSCLC

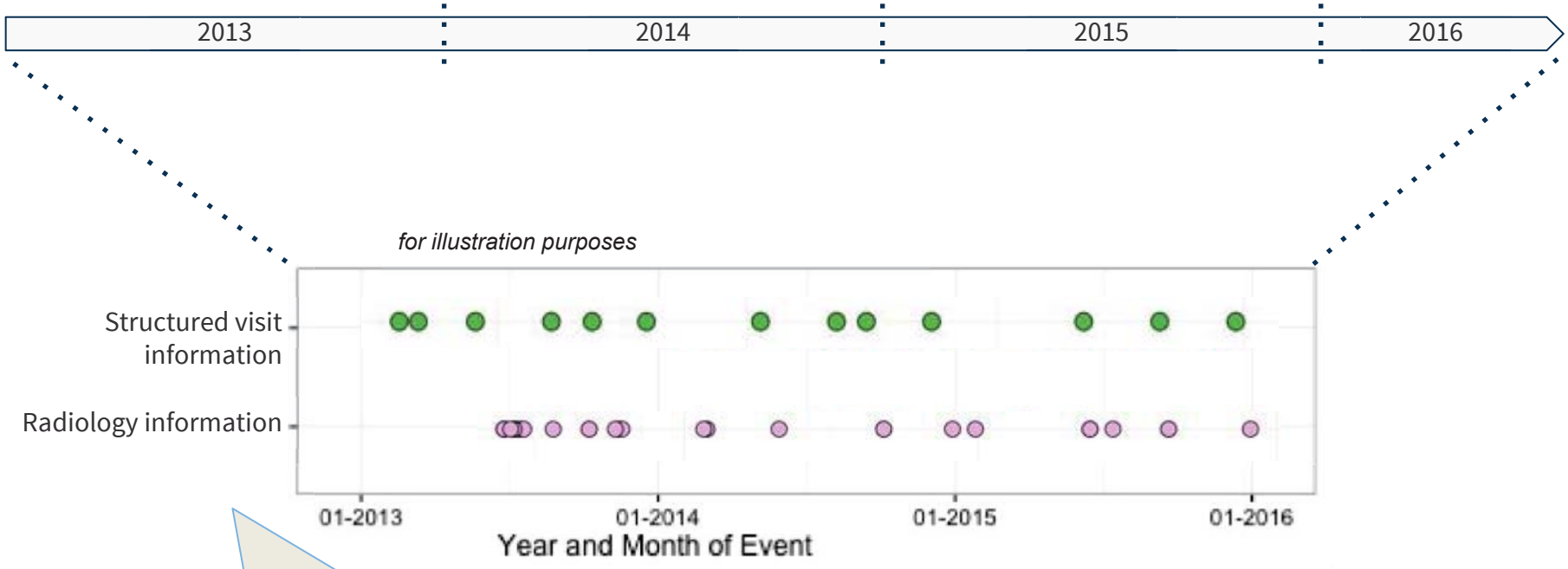


Over the next few slides we will show you how outcomes are documented in the patient chart, what kinds of source evidence are available, and some of the complexities of capturing this information at scale

FLATIRON

© 2016 Flatiron Health, Inc. Proprietary and confidential.

Progression Information is Captured in a Series of Clinic Notes and Radiology Reports



Both of these have associated unstructured documents in medical record that can be reviewed

FLATIRON

© 2016 Flatiron Health, Inc. Proprietary and confidential.

Imaging Showed Progression; Started on 2L Docetaxel

Started 1L carboplatin / pemetrexed

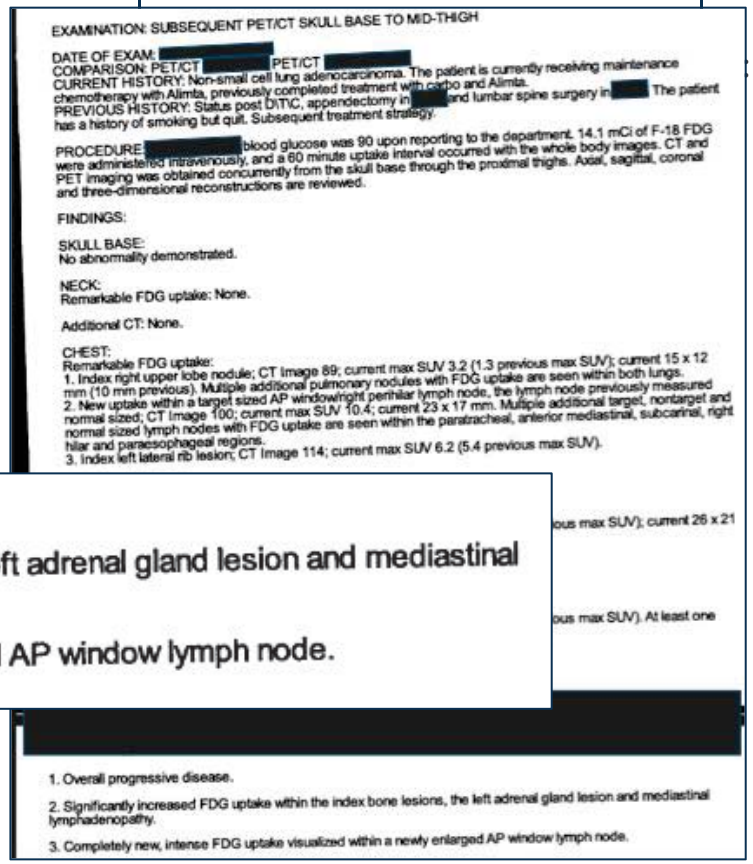
Imaging showed progression; started on 2L docetaxel

Imaging showed progression; started on 4L nivolumab

Presented with stage IV NSCLC

Transitioned to maintenance pemetrexed

Docetaxel held for toxicity



EXAMINATION: SUBSEQUENT PET/CT SKULL BASE TO MID-THIGH

DATE OF EXAM: [redacted]
COMPARISON: PET/CT [redacted] PET/CT [redacted]
CURRENT HISTORY: Non-small cell lung adenocarcinoma. The patient is currently receiving maintenance chemotherapy with Alimta, previously completed treatment with carbo and Alimta.
PREVIOUS HISTORY: Status post DI/TC, appendectomy in [redacted] and lumbar spine surgery in [redacted]. The patient has a history of smoking but quit. Subsequent treatment strategy.

PROCEDURE: [redacted] blood glucose was 90 upon reporting to the department. 14.1 mCi of F-18 FDG were administered intravenously, and a 60 minute uptake interval occurred with the whole body images. CT and PET imaging was obtained concurrently from the skull base through the proximal thighs. Axial, sagittal, coronal and three-dimensional reconstructions are reviewed.

FINDINGS:

SKULL BASE:
No abnormality demonstrated.

NECK:
Remarkable FDG uptake: None.

Additional CT: None.

CHEST:
Remarkable FDG uptake:
1. Index right upper lobe nodule; CT Image 89; current max SUV 3.2 (1.3 previous max SUV); current 15 x 12 mm (10 mm previous). Multiple additional pulmonary nodules with FDG uptake are seen within both lungs.
2. New uptake within a target sized AP window/right perihilar lymph node, the lymph node previously measured normal sized; CT Image 100; current max SUV 10.4; current 23 x 17 mm. Multiple additional target, nontarget and normal sized lymph nodes with FDG uptake are seen within the paratracheal, anterior mediastinal, subcarinal, right hilar and paraesophageal regions.
3. Index left lateral rib lesion; CT Image 114; current max SUV 6.2 (5.4 previous max SUV).

[redacted]

1. Overall progressive disease.

2. Significantly increased FDG uptake within the index bone lesions, the left adrenal gland lesion and mediastinal lymphadenopathy.

3. Completely new, intense FDG uptake visualized within a newly enlarged AP window lymph node.

1. Overall progressive disease.

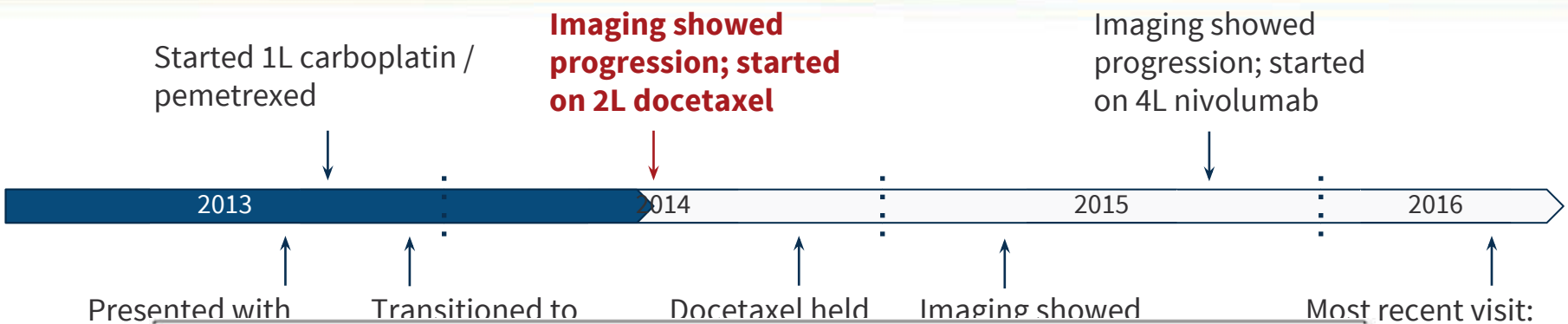
2. Significantly increased FDG uptake within the index bone lesions, the left adrenal gland lesion and mediastinal lymphadenopathy.

3. Completely new, intense FDG uptake visualized within a newly enlarged AP window lymph node.

FLATIRON

© 2016 Flatiron Health, Inc. Proprietary and confidential.

Imaging Showed Progression; Started on 2L Docetaxel



Assessment
Pet Ct evidence of Progressive Non Small cell lung cancer Status post alimta/carboplatin induction therpy followed by 12 cycles of maintenance of Alimta

Histological subtype Adenocarcinoma

History of Tobacco abuse.

Good performance status.

Normocytic Anemia
Disease Status: Progression of disease.

Recommendation/Plan
1. Discussed pet/ct results and the fact that she has evidence of disease progression.Pros and cons of further treatment options were discussed.

2. Incurable nature of disease was emphasized

3. Given her good performance status and the fact that she wants to pursue with further therapy the game plan is to proceed with salvage therapy utilising single agent taxotere at a dose of 60mg/m2 along with neulasta support.

4 Also would continue her on Exjeva which we would give her every 6 weeks.

5. Restage with PET CT after 3 cycles.

6 RTC on [redacted] for initiation of chemotherapy.

FLATIRON

© 2016 Flatiron Health, Inc. Proprietary and confidential.

Complexities of capturing endpoints at scale in real-world data based on the EHR

Complexities of real-world data

- Room for interpretation in radiology reports
- Variable time-points for disease assessments
- Radiologists may not access prior scans
- Several sources of evidence per event
- Clinical nuances (e.g., pseudoprogression, mixed response)
- Potential for missing data
- Length and complexity of abstraction, creating potential for errors

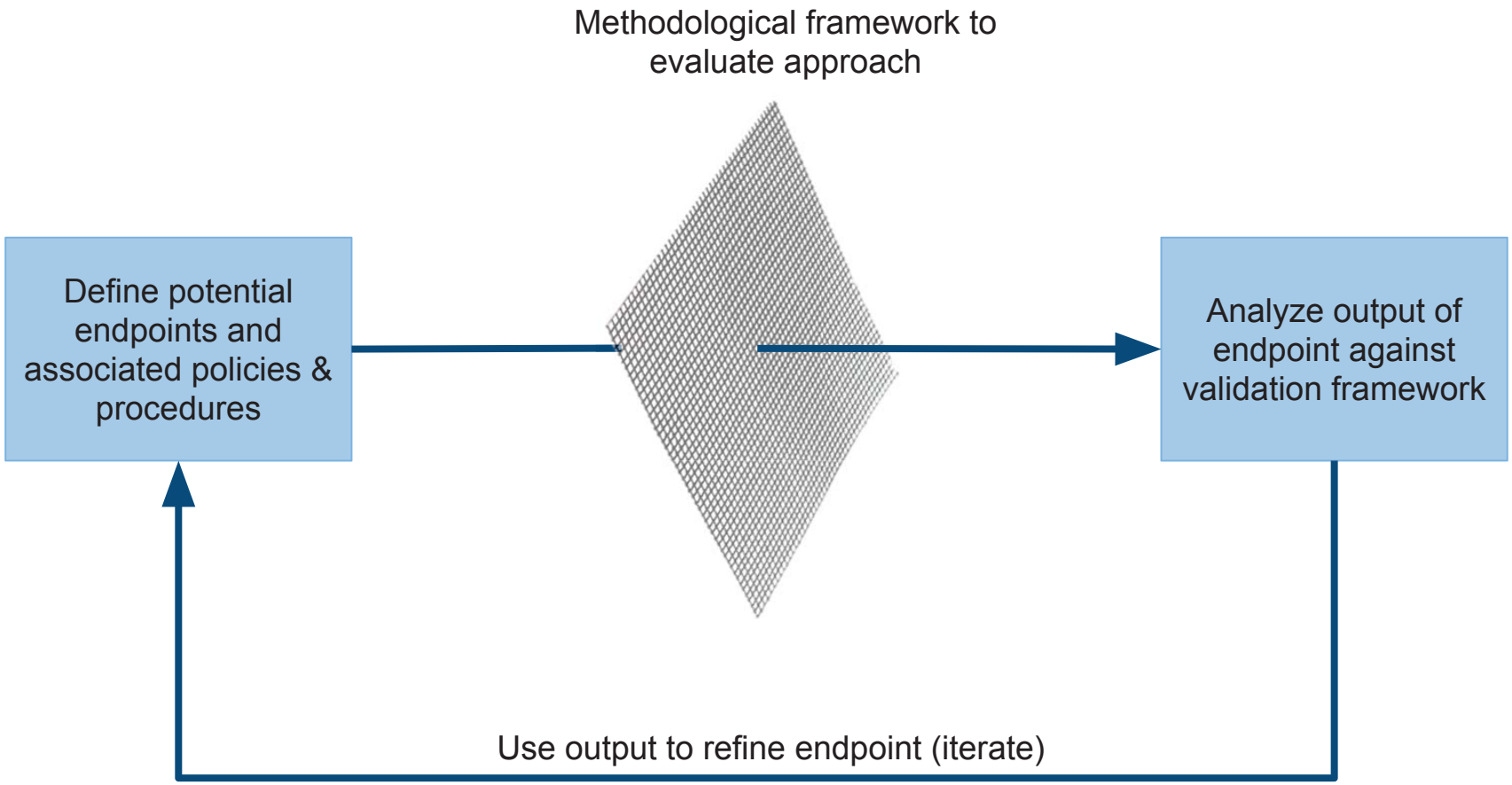
Need an approach that is:

- Able to account for complexities of the real world
- Scalable
- Replicable across abstractors
- Portable to multiple clinical settings and EHRs

FLATIRON

© 2015 Flatiron Health, Inc. Proprietary and confidential.

Overarching approach



FLATIRON

How do we evaluate our approach?

Performance criteria for real-world endpoints can be summarized using a "data quality and validation" framework, as proposed below:

Face validity	1. Oncologist agreement with definition & approach
	2. Regulator and other stakeholder agreement with definition & approach
Feasibility and quality of abstraction	3. Completeness of collected data
	4. Inter-rater agreement on progression dates for duplicate abstracted patients
	5. Qualitative feedback from abstractors reviewing the medical records
Validity of outputs	6. Likelihood of predicting downstream events (e.g., treatment change)
	7. Association between OS and PFS/TTP a. Patient-level correlation b. Responsiveness of endpoint to treatment effects

FLATIRON

Discussion points

1. What advantages and limitations of real-world data should be considered in the development real-world endpoints?

2. Is the data quality and validation framework a suitable approach to evaluating the performance of real-world endpoints?

3. What lessons from clinical trials endpoints should be applied to this framework?

Up next: Example application of framework to a proposed real-world endpoint → real-world progression (rwP)

Development and evaluation of the real-world progression (rwP) endpoint

Comparison of potential endpoints

	RECIST-based progression	Real-world progression (rwP)		Treatment failure (duration / time to next treatment)
		Radiologist interpretation	Clinician interpretation	
Definition	Change in tumor size based on radiologic evidence.	Change in tumor burden based on radiologist's interpretation of scans.	Improvement or worsening of disease based on clinician's interpretation of the entire patient chart.	Changes in treatment as a signal of the physician's interpretation of worsening disease*
Eligible source evidence	Radiographic imaging (or clinical decline)	Radiology is main source evidence, with laboratory, pathology and clinical assessments as confirmatory documentation	Clinician assessment is main source evidence, with radiology, laboratory, and pathology as confirmatory documentation	Longitudinal treatment data

FLATIRON

*Can include non-disease related events such as financial considerations, need to modify treatment schedule to accommodate homelife, etc.

Comparison of potential endpoints

	RECIST-based progression	Real-world progression (rwP)		Treatment failure (duration / time to next treatment)
		Radiologist interpretation	Clinician interpretation	
Definition	Change in tumor size based on radiologic evidence.	Change in tumor burden based on radiologist's interpretation of scans.	Improvement or worsening of disease based on clinician's interpretation of the entire patient chart.	Changes in treatment as a signal of the physician's interpretation of worsening disease*
Eligible source evidence	Radiographic imaging (or clinical decline)	Radiology is main source evidence, with laboratory, pathology and clinical assessments as confirmatory documentation	Clinician assessment is main source evidence, with radiology, laboratory, and pathology as confirmatory documentation	Longitudinal treatment data

FLATIRON

*Can include non-disease related events such as financial considerations, need to modify treatment schedule to accommodate homelife, etc.

Should real-world disease burden be assessed using RECIST criteria?

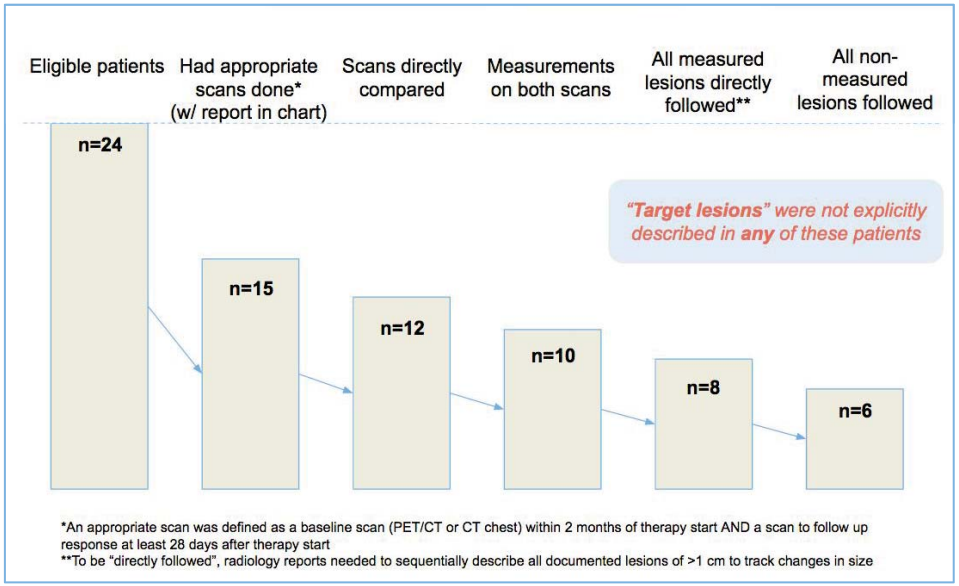
Experimental approach

We reviewed the charts of 24 advanced NSCLC patients to determine whether an existing framework for evaluating disease burden like RECIST can be applied to reliably define progression and tumor response in real-world EHR data.

In real world data, only charts are available for review; radiology scans are not available for central review

Summary of findings

- Using a strict definition of RECIST (require the radiologist to identify a target lesion), 0% of the patients were able to be assessed by RECIST.
- Relaxing the definition of RECIST (allow a trained abstractor to select the target lesion), 25% of the patients were able to be assessed by RECIST.
- For the other 75% of patients, the endpoint would have to be treated as missing.



Comparison of potential endpoints

	RECIST-based progression	Real-world progression (rwP)		Treatment failure (duration / time to next treatment)
		Radiologist interpretation	Clinician interpretation	
Definition	Change in tumor size based on radiologic evidence.	Change in tumor burden based on radiologist's interpretation of scans.	Improvement or worsening of disease based on clinician's interpretation of the entire patient chart.	Changes in treatment as a signal of the physician's interpretation of worsening disease*
Eligible source evidence	Radiographic imaging (or clinical decline)	Radiology is main source evidence, with laboratory, pathology and clinical assessments as confirmatory documentation	Clinician assessment is main source evidence, with radiology, laboratory, and pathology as confirmatory documentation	Longitudinal treatment data

FLATIRON

*Can include non-disease related events such as financial considerations, need to modify treatment schedule to accommodate homelife, etc.

What is the appropriate definition of real-world progression (rwP)?

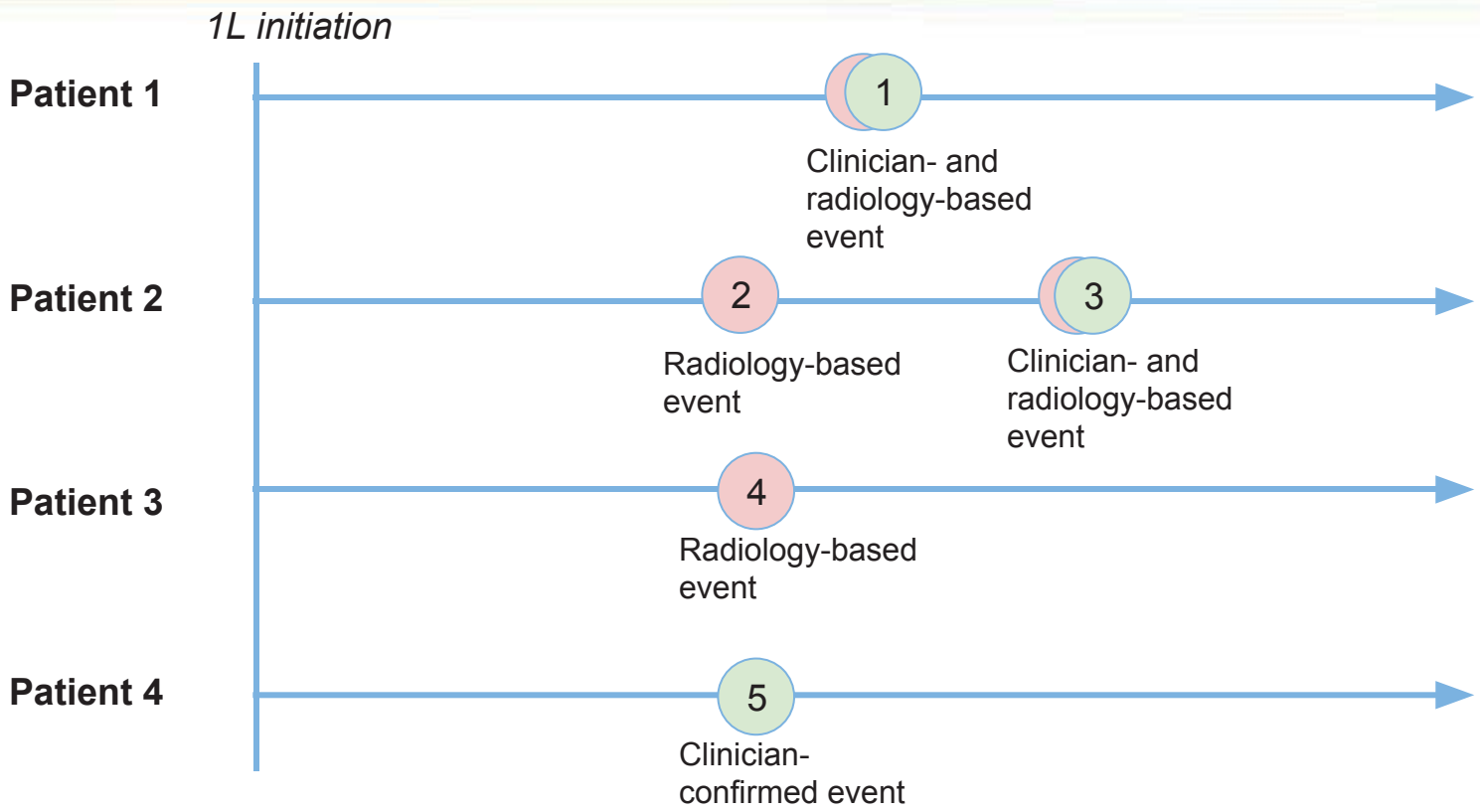
Background/rationale

- There are numerous potential ways to define progression based on available source data. The proposed rwP endpoint is a consensus endpoint that summarizes different inputs.
- It is critical to define the hierarchy of each piece of information and how to incorporate it into the endpoint, including when different information conflicts (e.g., radiologist vs clinician interpretation)

Key question:

Should we base the real-world progression endpoint on clinician assessment, radiologist assessment, or a hybrid approach?

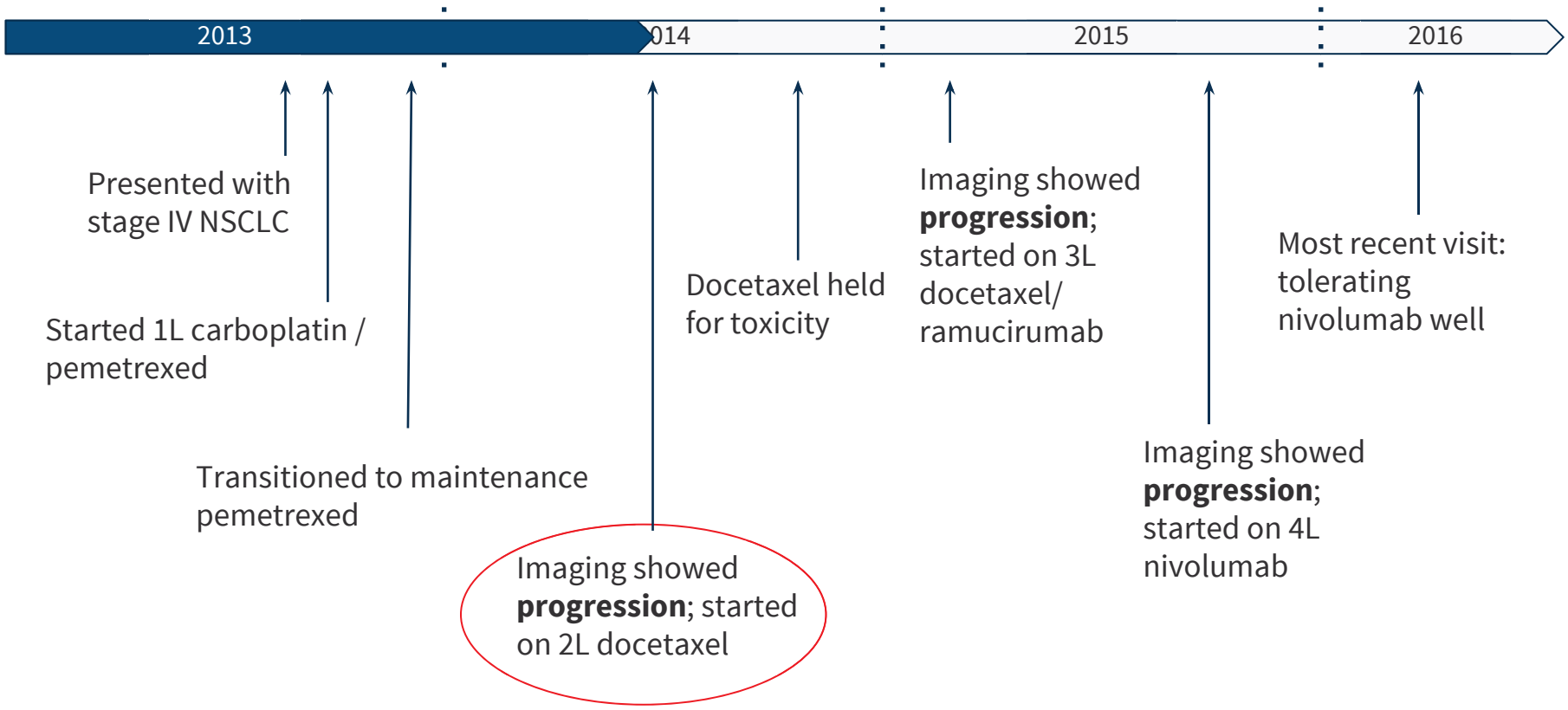
Example illustration of potential hierarchies



I. **First clinician-confirmed progression** *(green events only)*
II. **First radiology-based progression** *(red events only)*
III. **First clinician- or radiology-based progression** *(earlier of red or green for each patient)*

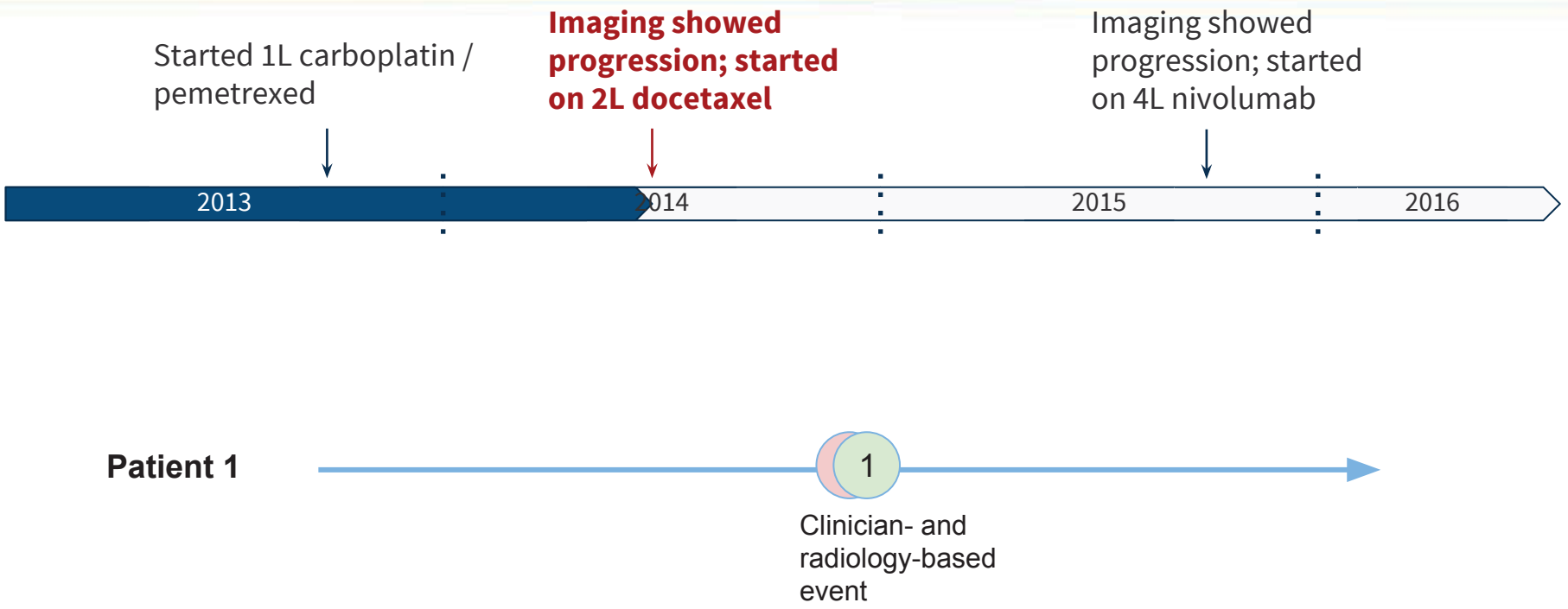
FLATIRON

Patient Journey: Female in Her Late 60s with NSCLC



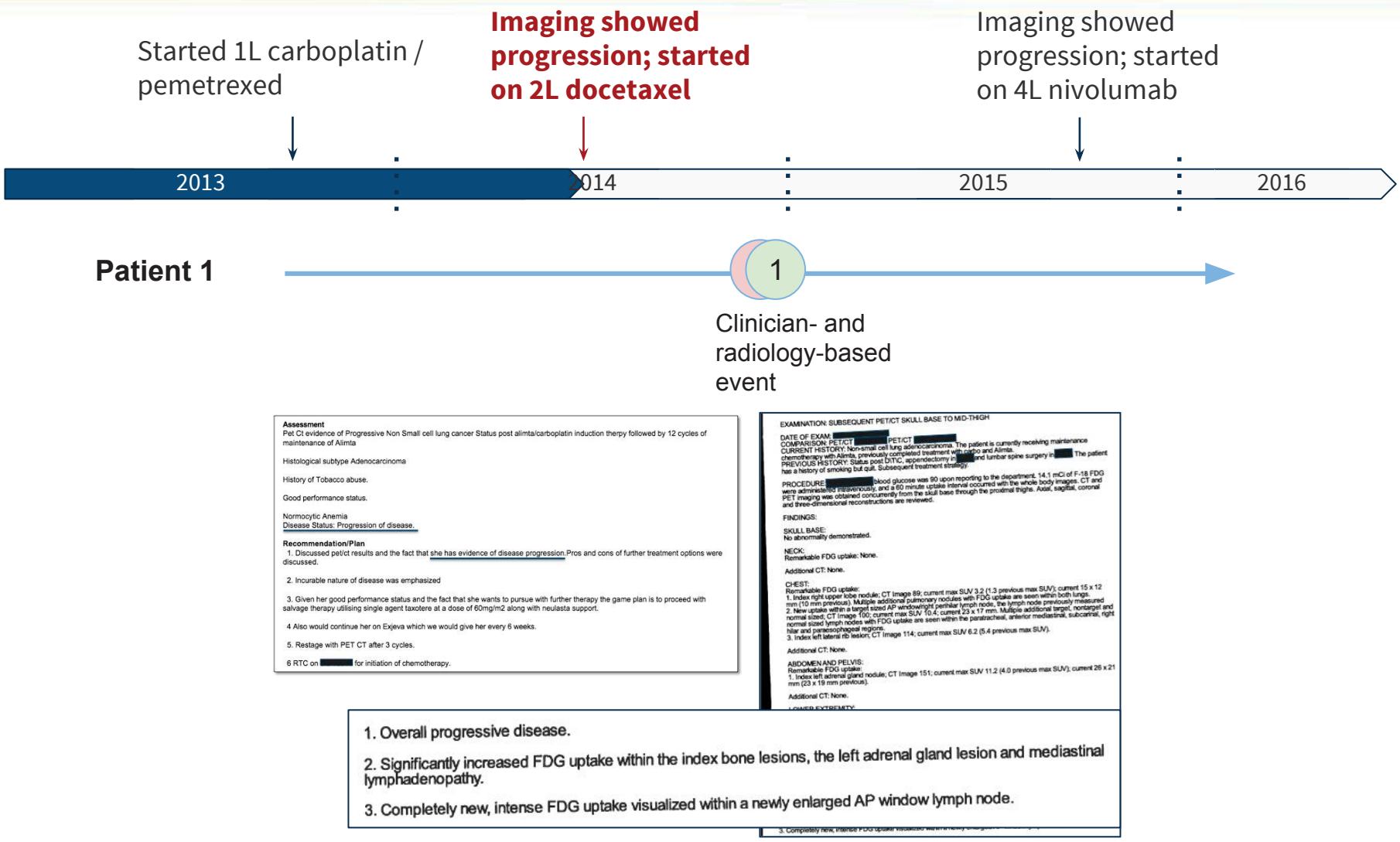
© 2016 Flatiron Health, Inc. Proprietary and confidential.

rwP as a consensus endpoint



© 2016 Flatiron Health, Inc. Proprietary and confidential.

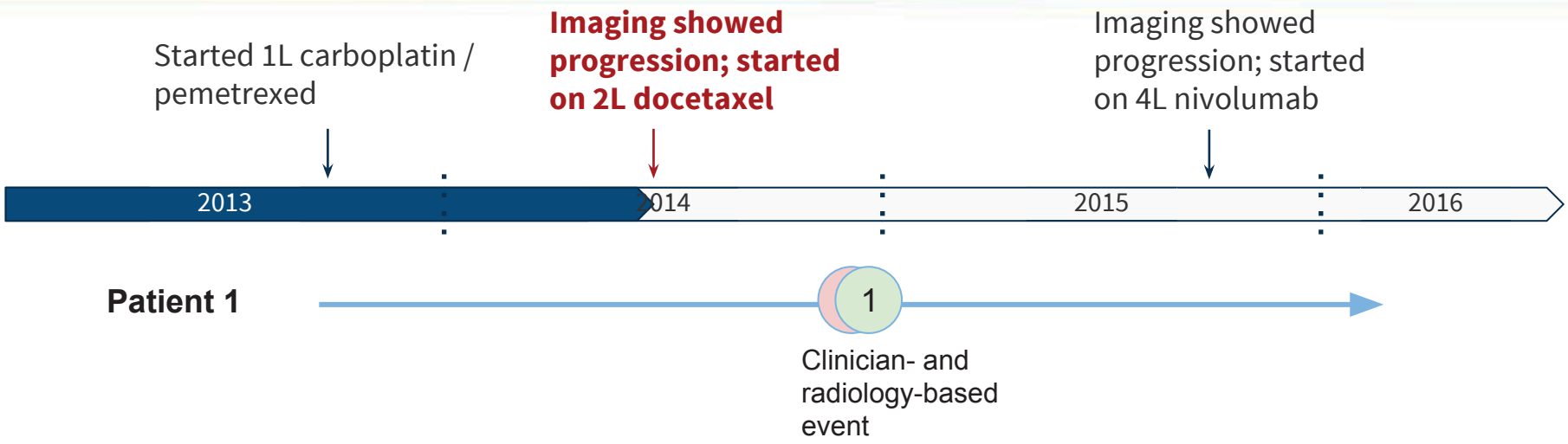
rwP as a consensus endpoint



FLATIRON

© 2016 Flatiron Health, Inc. Proprietary and confidential.

rwP as a consensus endpoint



Maintain all underlying component information

	Event?	Date?
Clinician note	X	03/11/2015
Radiology report	X	03/05/2015
Pathology report		

Consensus:
Did a progression event occur? YES

Associated date:
Clinician-confirmed: 03/05/2015
Radiology-reported: 03/05/2015
Either: 03/05/2015

Confidence:
2 of 3 potential elements; consistent evidence of progression; source dates within a month; no pathology available (e.g., score = 7/10)

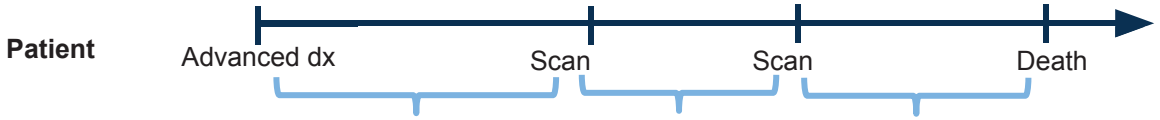
FLATIRON

© 2016 Flatiron Health, Inc. Proprietary and confidential.

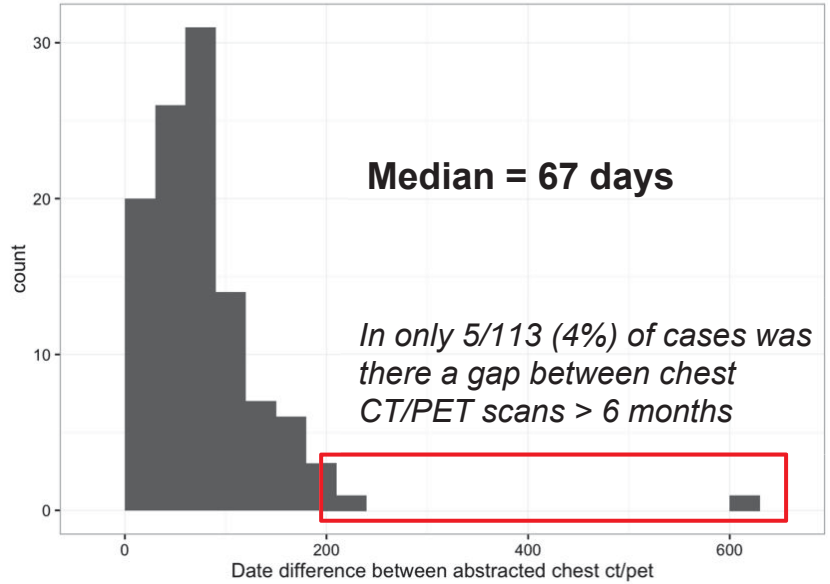
Scan frequency in real world practice

- The frequency of CT involving the chest (i.e., chest, chest/abdomen, pelvis, chest/abdomen) or whole-body PET scans, including PET/CTs was compared in a sample of 30 advanced NSCLC patients

- All gaps are described; For example, if a patient had a single CT scan to assess progression, that patient would have 2 gaps: the time between advanced diagnosis and the chest CT and the time between the chest CT and either last activity or death

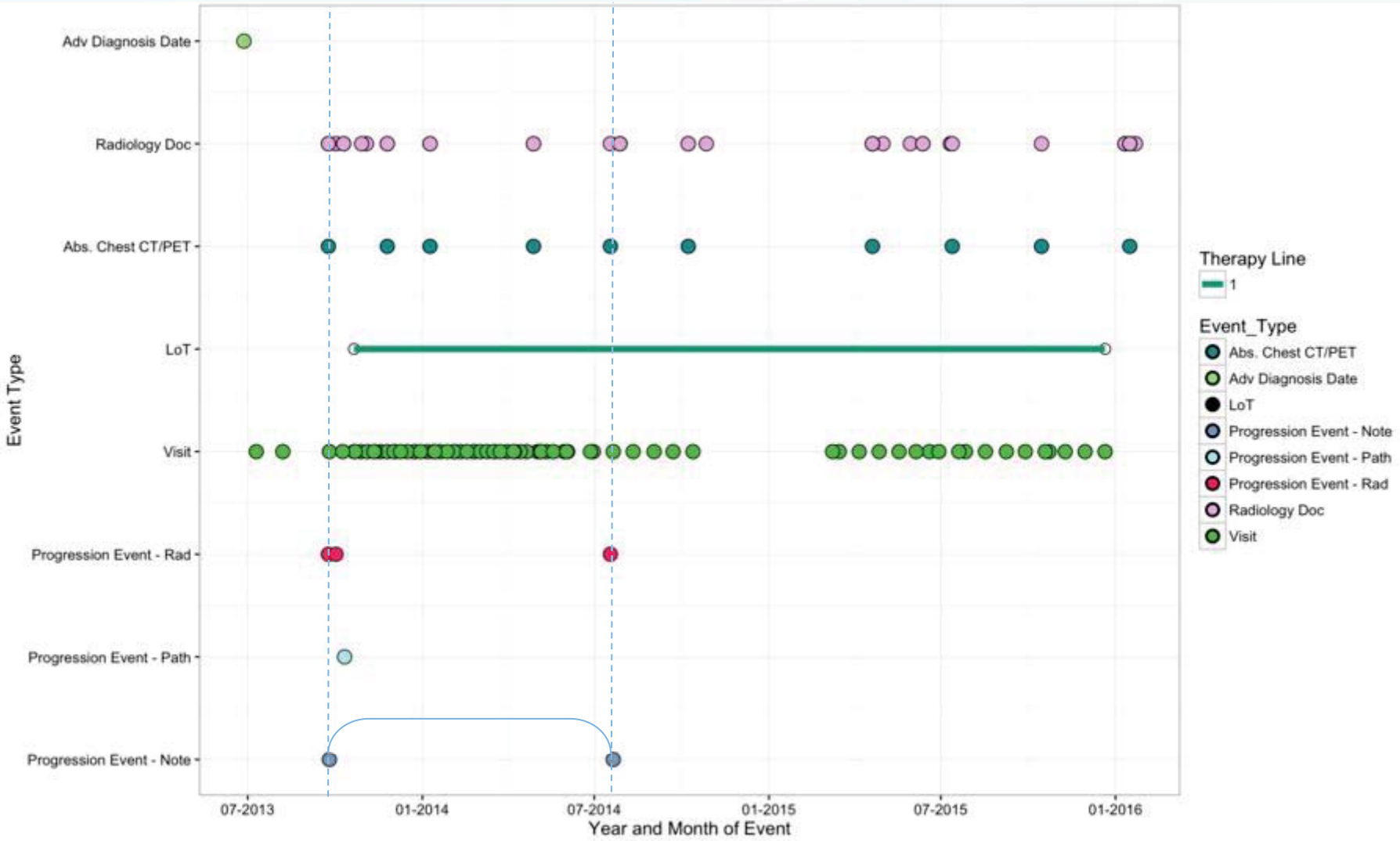


Frequency of chest CT or PET scans (abstracted)



FLATIRON

Multiple indicators of care during progression-free intervals: Patient example



FLATIRON

Notes:

1. *Date of death is generalized to last day of month*
2. *Abs. Chest CT/PET includes all abstracted CT scans involving the chest (i.e., chest, chest/abdomen, pelvis, chest/abdomen) or whole-body PET scans, including PET/CTs*

Summary of experimental approach

200 Advanced NSCLC patients were selected based on the following inclusion criteria:

- Confirmed diagnosis of Advanced NSCLC on or after January 1, 2011
- Completed at least one line of therapy and initiated at least one subsequent line

For these patients, abstractors were instructed to review all documents in the chart for evidence of potentially eligible progression events

- Events could be documented in clinician notes, radiology reports, and/or pathology reports
- Documented events contributed to the development of derived rwP variables
  - Evidence of a progression event
  - Date

Summary of experimental approach

Tested 3 different approaches using the first progression event after initiation of first line therapy:

- First clinician-confirmed progression[1]
- First radiology-based progression[2]
- First clinician- or radiology-based progression

Evaluated using data quality and validation parameters:

- Data completeness
- Association with treatment change
- Correlation with OS
- Data reliability, quality and qualitative feedback from abstractors

A Priori Hypothesis:

Clinician-confirmed progression yields the highest-quality endpoint because treating clinicians have the appropriate context to assess the patient's overall disease burden.

FLATIRON

1 Date of progression is indicated as date of radiology if available, or date of clinician note
2 One event was identified only via pathology; this was included in the radiology-based progression category

Several data slides have been removed and will be published at a later time.

Please contact Amy Abernethy, at amy@flatiron.com, for more information on the detailed slides from this section

Summary: What is the appropriate criteria for rwP?

Summary of findings

- Application of a RECIST approach is not a practical solution using real-world EHR data.
- Multiple approaches to defining real-world progression are possible, including clinician-confirmed events, radiology-based events, or both.
    - Overall, agreement is high between the approaches
    - Based on our validation of the resulting endpoints, clinician-confirmed progression events appear to be the best approach to rwP

Recommendation

Focus on clinician-confirmed progression as the primary definition and approach for the rwP variable. Use radiology and pathology data as supplemental confirmatory data.

Create consensus variable based upon underlying source information; maintain component information in the real-world dataset. Consider "confidence score" for each progression event.

Comparison of potential endpoints

	RECIST-based progression	Real-world progression (rwP)		Treatment failure (duration / time to next treatment)
		Radiologist interpretation	Clinician interpretation	
Definition	Change in tumor size based on radiologic evidence.	Change in tumor burden based on radiologist's interpretation of scans.	Improvement or worsening of disease based on clinician's interpretation of the entire patient chart.	Changes in treatment as a signal of the physician's interpretation of worsening disease*
Eligible source evidence	Radiographic imaging (or clinical decline)	Radiology is main source evidence, with laboratory, pathology and clinical assessments as confirmatory documentation	Clinician assessment is main source evidence, with radiology, laboratory, and pathology as confirmatory documentation	Longitudinal treatment data

FLATIRON

*Can include non-disease related events such as financial considerations, need to modify treatment schedule to accommodate homelife, etc.

Summary of experimental approach

727 Advanced NSCLC patients were selected based on the following inclusion criteria:

- Confirmed diagnosis of Advanced NSCLC on or after January 1, 2011
- Initiated at least one line of therapy
- Tested via next-generation sequencing (Foundation Medicine)

Abstractors were instructed to review all documents for evidence of clinician-confirmed progression events (consistent with results of prior experiment)

- All clinician-confirmed events were documented, as well as corresponding source evidence (e.g., radiographic, pathologic, clinical assessment)
- rwP was therefore defined using the "clinician assessment" approach

Real-world endpoints (rwP, time to next treatment) were tested using data quality and validation parameters, specifically correlation with OS and sensitivity analyses in order to test performance of the variable under different conditions

- The first progression event after initiation of 1L therapy was used for the analysis

Summary: focus on clinician-confirmed rwP

Summary of findings for advanced NSCLC

- rwP can be captured reliably at scale across a broad sample of Advanced NSCLC patients
- rwPFS correlates well with OS at the patient level, and is minimally impacted by various cohort selection factors
- rwP is more robust to cohort selection factors than time to treatment failure / time to next treatment

Proposed definition of rwP: All distinct episodes in which the *treating clinician concludes that there has been overall growth or worsening* of the disease of interest

- Distinct episodes are disease-specific time intervals in which the patient is assessed for progression
- Type of information considered includes radiology, laboratory evidence, pathology, clinical assessment

FLATIRON

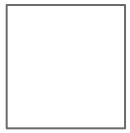
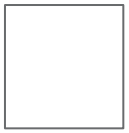
Proposed Next Steps

1. Analyze agreement between treatment effects on rwPFS and OS, adjusting for treatment selection in real-world population ("trial-level" analysis)
   - Cancer centers as the unit of measure, analogous to "trials"
   - Propensity score matching to approximate treatment randomization
2. Replicate a clinical trial cohort and analysis using real-world data and compare endpoints
3. Accommodate irregularities in intervals between visits and scans (e.g., interval censoring)
4. Develop a confidence score for each progression event (e.g., based on source evidence)
5. Extend rwP methodology to larger patient samples, other tumor types, and various treatment settings (e.g., adjuvant); compare results
6. Consider rwPFS alone and as part of composite endpoints; use the same framework to evaluate (e.g., PFS vs. PFS + QOL vs PFS + QOL + opiate utilization)

FLATIRON

© 2015 Flatiron Health, Inc. Proprietary and confidential.

Where we are in real-world endpoint development

Q3 2015	*Q4 2015 - Q2 2016*	*June 6*	*June 16*	*Q3-Q4*
Initiation of endpoint work	• Experimentation and analysis • Joint working sessions with researchers, life science partners, and clinicians	Research advisory board at ASCO with oncologists and academic researchers	Presentation of results at FOCR meeting and FDA input	Results & publication
				

FLATIRON

Discussion points

1. Additions to the data quality and validation framework?
2. What is the role and value of real-world endpoints of different types?
    - Real-world progression, tumor response, time to next treatment, composite endpoints
    - Patient-generated health data (e.g., PROs, sensors)
    - Endpoint “packages”?
    - Value of underlying component information?
3. For what purposes can we use these endpoints?
    - Retrospective analyses? Publications ? Post-approval regulatory submissions? Payer discussions? Prospective pragmatic trial?
    - Other?

Collaborative prospective real-world data study

FLATIRON

© 2015 Flatiron Health, Inc. Proprietary and confidential.

Context on mutational burden / MSI

- Immunotherapy is hypothesized to have activity in tumors with high mutational burden
- Deficiency in DNA mismatch repair (MMR) pathways is a cause of high mutational burden. MMR deficiency leads to genetic hypermutability, resulting in microsatellite instability (MSI)
- Evidence for the role of PD-1 inhibitors in MMR-deficient tumors is growing
  - A phase II study demonstrated that response rates to pembrolizumab is far higher in MMR-deficient colorectal cancer (CRC) than in other CRC patients (Le et al, NEJM 2015)
    - 11/2015: breakthrough therapy designation for pembrolizumab in MSI-high CRC
  - Preliminary evidence suggests that other GI / GU tumors with MSI may also respond to PD-1 inhibition
- Adoption of MSI testing in non-CRC cancers remains very low vs moderate in CRC

Opportunity to investigate MSI and PD-1 inhibitor treatment patterns in real-world dataset

Patient case example

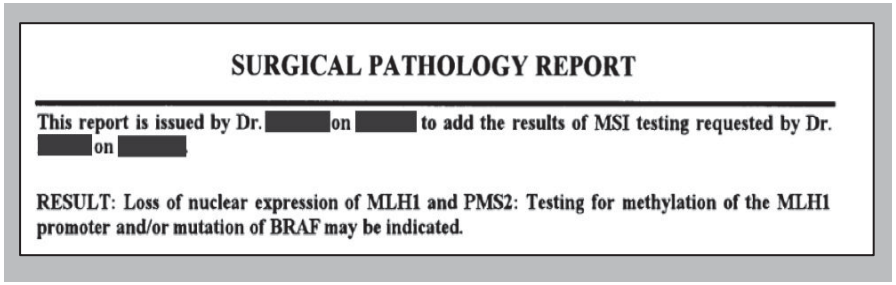
Case History

Late 70s male who presented in 2014 with epigastric pain and anemia, and work-up revealed gastric adenocarcinoma

- Underwent subtotal gastrectomy with path showing positive nodes and direct extension to serosal surface and omentum
- Radiation with concurrent capecitabine; 3 months later: peritoneal recurrence
- Carboplatin / paclitaxel x 3 cycles, did not tolerate well so switched to paclitaxel / ramucirumab
- Developed malignant ascites

Data

- NGS testing showed large number of genomic alterations
- MSI testing showed mismatch repair deficiency



SURGICAL PATHOLOGY REPORT

This report is issued by Dr. [redacted] on [redacted] to add the results of MSI testing requested by Dr. [redacted] on [redacted].

RESULT: Loss of nuclear expression of MLH1 and PMS2: Testing for methylation of the MLH1 promoter and/or mutation of BRAF may be indicated.

Decision-making

- Physician initiated treatment with pembrolizumab, citing 2015 *NEJM* study (see right)
- Treatment ongoing at the time of review

"Pembrolizumab has been used successfully in mismatch repair- deficient tumors in a phase 2 study (Le et al, NEJM 372;26, 2015)…. While only 1 gastric cancer patient was included, the study provides compelling evidence that MMR- deficiency may predict clinical benefit to PD1 inhibition."

FLATIRON

Proposed pilot study

Proposed pilot study of approved PD-1 inhibitors in patients with highly mutated tumors

- Initial feasibility assessment using retrospective databases, including testing and treatment patterns
- Prospectively designed pilot study to assess feasibility of using RWE to support regulatory decisions
  - Collaborative study
  - Validate data collection efforts
  - Test novel endpoints
  - Document efficacy and safety

Conclusions and future directions

- Real world datasets provide a valuable opportunity to explore early, untested clinic hypotheses such as PD-1 inhibitors in non-CRC MSI tumors
  - Early published evidence drives clinical decision-making, leading to real world treatment ahead of guideline changes and expansion of approved indications
  - Identifying patients of interest and conducting analysis of real world experience can result in rapid generation of insights
- Broad set of proposed applications for RWE, including:
  - Expanding the safety profile of a therapy
  - Identifying populations with enhanced benefit/risk for an approved therapy
  - Piloting studies to determine the correlation between real-world endpoints and clinical trial endpoints
  - Building evidence for a supplemental package to expand the indication profile for a therapeutic
  - Supporting efficacy results observed in clinical trial, to serve as post-market confirmation of benefit

Real World Endpoints
A Pharma Perspective

- Utility well established for OS endpoint obtained from EHR/real world information as well as other variables such as prevalence of a certain patient characteristic or biomarker, and treatment patterns
- Analyses presented by Flatiron suggest that a progression endpoint can be defined using EHR information that is reproducible and correlates well with overall survival in NSCLC
- Unclear if “RW PFS” would be similar to RECIST 1.1-defined PFS obtained in a clinical trial (even if evaluated in EHRs from patients enrolled in a clinical trial)
- Similar to PFS obtained from non-blinded clinical trials, RW PFS may be biased by physician beliefs and is dependent on frequency of imaging assessments
- Similar to RW OS, RW PFS has the potential to be useful in supporting results obtained from a prospective clinical trial
- Pre-specification of hypotheses is important to avoid “data dredging” and multiplicity errors