FRIENDS
of CANCER
RESEARCH

ALEXANDRIA SUMMIT®

Blueprint for Breakthrough:
Exploring the Utility of Real World Evidence (RWE)

Table of Contents

Goals

Friends of Cancer Research (*Friends*) and Alexandria Summit will convene a multi-stakeholder meeting to include industry, real world evidence (RWE) vendors, FDA, academics, researchers, patients, advocacy organizations, and other vested stakeholders with the expressed intent of developing consensus toward the potential use of RWE in the regulatory setting. The following report, developed by a work group aims to build consensus in the following three areas which, will be presented on June 16th, 2016 to solicit public input and further refine content:

1) Identify disease and drug candidates in oncology as potential case studies,
2) Develop regulatory strategies for optimal use of RWE in oncology, and
3) Outline potential pilots in oncology that could be used for clinical evidence generation to support regulatory decisions.

Background

With bringing innovative medicines to patients in a timelier manner, there is great interest in designing clinical trials with adaptive features that make studies more efficient, more likely to demonstrate an effect of the drug, more informative, and better able to capture the totality of clinical evidence.[1] Advancements in our ability to build learning systems and improve data collection, so as to link clinical priorities and measurable outcomes, are beginning to inform clinical studies with respect to how patients are treated, should be treated, and wish to be treated with these new therapies in the real world. Thus, an opportunity exists to learn from current efforts to collect, interpret, and apply real world evidence to drug development.

Real World Evidence

RWE refers to evidence generated from data collected outside the traditional clinical trial setting, such as electronic health records (EHRs) including patients treated on- and off-label, pragmatic clinical trials, patient registries, patients treated through expanded access, administrative claims, surveys, and mobile health-generated data (e.g., smartphones, wearables, internet and social media). RWE is thought to better reflect the general population and the care they receive, given that enrollment in clinical trials is often limited to patients with specific baseline characteristics, with often restricted eligibility. Therefore, high quality RWE can provide different, and at times broader, estimates of the safety and effectiveness of therapies than certain traditional clinical trials with narrow eligibility criteria.

While randomized controlled trials (RCTs) are the gold standard for minimizing sources of potential bias, RWE may have utility in certain scenarios. Particularly, in the case of a drug where the effect size is likely to be significantly larger than any confounding factors that might occur and where confidence in the original efficacy data is relatively high,[2] such as for a therapy designated as a Breakthrough Therapy. While there may be concerns regarding data quality, owing to factors such as missing information and non-systematic data collection, information gathered from EHRs can allow for data to be collected on more patients, in an unselected patient population and more rapidly than traditional phase 4 trials designed to meet post-market requirement and commitments. Thus, in cases of transformative treatments, the question of whether it would be feasible and sufficient to confirm clinical benefit in the real world setting warrants serious consideration.

Breakthrough Therapy Designation

The Breakthrough Therapy Designation (BTD) was created in 2012 as a way to expedite the development of drugs for serious conditions with a large unmet medical need. Breakthrough Therapy designation decisions are based on preliminary clinical evidence demonstrating a substantial improvement over existing therapies. The designation and subsequent actions are primarily designed to maximize efficiency in the clinical development regulatory review processes for potentially transformative new medicines. An additional goal of the designation is to minimize the number of patients that are exposed to a less efficacious treatment throughout the entire development process, including the post-market setting. Drugs intended to treat a serious or life-threatening condition for which there is a large unmet medical need can receive accelerated approval by demonstrating a large effect on a surrogate or biomarker endpoint, e.g., tumor response, reasonably likely to predict a clinical benefit. Drugs approved using a surrogate endpoint generally

---

[1] Guidance for Industry Adaptive Design Clinical Trials for Drugs and Biologics. 2010 (version 1).
[2] Simon R et al, Clin Pharmacol Ther. 2015. 97(5):502-7.

require further evaluation to confirm clinical benefit where there is uncertainty in the relationship between the surrogate endpoint and the clinical benefit, or the observed clinical benefit and ultimate outcome. In the case of an expected substantial improvement in overall survival, as with drugs grants BTD, there may be loss of equipoise for conducting a randomized trial with a less effective therapy for confirmation of clinical benefit following accelerated approval. This raises the need for alternative approaches such as the use of RWE for confirmation of clinical benefit for highly active anticancer therapies such as those granted BTD.

Pragmatic Trials

In addition to traditional models for clinical trial designs, which may not always be practical, adaptive clinical studies that more closely reflect routine medical care represent opportunities to better understanding novel therapies in the real world setting. Pragmatic clinical trials (PCTs), which leverage existing clinical infrastructure, are designed to test interventions in everyday clinical settings to maximize therapeutic applicability and generalizability. Although PCTs are defined by FDA as being randomized studies, there are notable examples of nonrandomized pragmatic trials as well. For example, the Targeted Agent and Profiling Utilization Registry (TAPUR) Study,[3] launched in March 2016 by ASCO, is a nonrandomized pragmatic trial intended to collect data on the safety and efficacy of approved therapies in other disease settings. However, even with careful planning, such as with validation studies examining sensitivity, specificity, positive and negative predictive values to ensure the data needed to measure safety and effectiveness are captured reliably, it may still be challenging to identify confounding factors that impact study results and generalizability.

Challenges and Considerations

Technological advancements and the growing use of EHRs have facilitated collection of patient data outside of clinical trial settings, and hold potential to further inform patient care, supplement current clinical trial methodologies, and speed drug development. Although real world data (RWD) collection, like other data collection efforts, has challenges with variable data quality, heterogeneity in collection mechanisms, and privacy concerns, among others, a lot of progress has been made in this area recently. Several outstanding challenges that limit full implementation of RWE are highlighted in Table 1.

TABLE 1: Challenges that limit full implementation of RWE.

Data sources	EHRs (i.e., labs data, claims/billing codes, etc.) from academic, hospital, and community oncology sources, and registries, that may include patient reported outcomes (PROs), from patient advocate organizations and others.
Potential variables needed	The minimum information base includes: Diagnosis (International Classification of Disease (ICD) codes; dates of initial and advanced diagnoses; Disease staging; Dates and sites of metastases); Histology; Radiology and pathology reports; Treatment dates (start/stop, for prior and subsequent treatment), CPT codes; Labs (test date, result date, test name, result with units, and normal ranges); Demographics (e.g., smoking status); Biomarker status; Gene sequencing; Performance status; Medication and administration (date, drug, dose, routine, and units); Adverse events report/collection (Grade 3+ or serious adverse events only); Outcomes (e.g., date of death, other endpoints). Additional unique variables will need to be considered depending on the questions.
Data standards	Regulatory trials have adopted Clinical Data Interchange Standards Consortium (CDISC) data standards, however, universal data standards are still being developed within the EHR infrastructure. Similarly, electronic data capture (EDC) systems have become standardized, with fewer proprietary systems requiring extensive validation.[4] Among various hospital settings, there may still be variability in software for electronic records management, though a trend towards commercial systems, e.g., EPIC, has been observed.

---

[3] https://clinicaltrials.gov/ct2/show/NCT02693535

[4] Guidance for Industry - Part 11, Electronic Records; Electronic Signatures, establishes the United States Food and Drug Administration (FDA) regulations on electronic records and electronic signatures (ERES). Title 21 CFR Part 11 Section 11.1 (a).

Processes for data merging vs. established third party aggregators	While there is no established internal process to merge RWD sources, a number of third party organizations have begun doing this in oncology (e.g., Flatiron, USO) and beyond oncology (e.g., Humedica). The Flatiron dataset, for example, comprises longitudinal, patient level EHR data and incorporates data from both structured and unstructured EHR data streams.
Challenges/ considerations related to data merging	Variability of data sources, different EHR platforms, different site adaptations of platforms (e.g. custom workflows and proprietary logic, structured and unstructured data collection), high level of ambiguity and complexity in the data concepts (e.g. clinical, financial), lack of interoperability, and complexities associated with data merging, mapping, normalizing, while not insurmountable, requires significant programming and informatics work to prepare research ready datasets. Linking or tracking the same individual across different EHR and disparate RWD platforms is also an important consideration to fully capture the care of a single patient who may appear in more than one RWD source.
Structured, unstructured and missing data challenges	There can be significant variation across EHR systems/providers especially with respect to data collected via structured vs. unstructured parts of an EHR (e.g., some capture disease stage, others do not). It is estimated that half of critical variables for oncology-focused RWE are in unstructured documents[5] thereby requiring technology-enabled[6] or manual chart review (although Natural Language Processing is also under evaluation). Structured data include data points that are organized in a predefined manner, such as dropdown fields, and unstructured data may include free text from a physician note or a scanned pathology report. Derived variables such a "lines of therapy" and "real-world progression" can supplement RWE datasets; these are generated by combining structured and unstructured data elements using pre-defined business rules. Since each variable comes from a different potential source, the reliability and validity of variables should be described. In addition, missing data, an issue encountered in clinical trial reports as well, will need to be addressed regardless of structured or unstructured data collection.
Challenges/ considerations related to variable extraction processes	Data quality, missing variables, methods applied to extract unstructured content; audit trails are variable depending on EHR system. Hospitals have different extraction systems, heterogeneity in IT capabilities, and data captured by clinicians. Working with an oncology EHR data aggregator would be the simplest mechanism to leverage existing processes that map data from additional providers and present the data in different formats. This would ensure less variability and easier extraction. However, issues such as limited access to broad populations limit the ability to extrapolate results broadly. Ideally, EHR fields could be adjusted for specific data collection as per provider.
Challenges/ considerations related to defining endpoints	Each clinical endpoint, including hard endpoints (e.g., OS), surrogate endpoints (e.g., PFS), and other clinically meaningful endpoints (e.g., RR) have different challenges with respect to collection, reliability, and recording precision in EHR. Discussion and agreement with the appropriate regulatory agency is necessary. Can we quantify tumor shrinkage by mining radiology reports? Will tumor shrinkage require qualification with the same precision as the standard RECIST criteria? Should we integrate clinician assessment of the patient with radiologist assessment of scans into summary variables reflecting tumor burden? Would time to next therapy or time to treatment failure as assigned by the treating physician be reasonable proxies for efficacy? Treatment decisions may be hard to capture (e.g., endpoints used for treatment modification, choice of therapy, etc.). What additional work with EHR providers may be necessary to adjust the records if those are to be used for purposes other than primary billing? How will we determine if endpoints assessed through real-world data are reliable, valid, and meaningful?
HIPPA/ informed consent issues	Informed consent will likely be required due to the need for identifiable information for data linkages, AE reporting, etc. As such, patient level and physician data would be available for auditing. Information from EHRs does not currently have patient consent and would likely involve new processes and policy changes. While it may be possible to obtain consent for registries, these studies are monitored and may not be totally

---

[5] Berger,ML et al, Future Oncology, 2016. 12(10):1261-74.
[6] Liede A et al, Clin Epidemiol. 2015. 7:441-8.

| | reflective of RWE. While broad consent will be difficult to obtain, future trials using EHR may require prospective consenting with a clear explanation of the information gathering intent. |

Data Quality and Utility

Significant challenges remain with combining, organizing, and analyzing data from various information sources, including EHRs, insurance claims, biosensors, genomics datasets and patient-reports. Yet interest in using RWE in the assessment of drugs and other clinical interventions remains high. Programs like Sentinel and PCORNet, which aggregate multiple data sources, are relying on claims and EHR-based information to collect large amounts of health data to drive research, including comparative effectiveness. The value of EHR is further evidenced by the American Society of Clinical Oncology (ASCO) launch of the CancerLinQ™ system designed to improve patients' outcomes and quality of life based on EHR data. To date, however, it remains unclear whether such data could be suitable for regulatory purposes. Provided proper standards and methods of collecting, validating, and analyzing real-word data exists, RWE may support a number of activities that impact drug development and delivery.

Questions regarding RWD quality:
- What data quality elements need to be considered and should they differ by data source (e.g., EHRs)?
- How should various databases (community versus academic) with respect to extractability of relevant fields be considered?
- How should quality be reported (e.g., data completeness, variable reliability, variable validity, sources of variables, data provenance) and presented for review? What thresholds need to be reasonably considered for these categories?
- What details should be captured with respect to cohort selection when generating RWE?
- What additional analyses need to be done to generate RWE (i.e., sensitivity analyses)?

Questions regarding RWE use:
- What aspects of efficacy need to be captured with RWE in addition to addressing safety concerns? How best to consider the most appropriate endpoints and outcome measures for various intended uses?
- What requirements are needed for safety reporting based on RWD collected in a trial using EHR? Any specific regulatory advice considering that the drug is marketed? What additional reporting is needed in EHR beyond physicians' reports in their daily activities, according to Good Clinical Practice? Why and under what circumstances?
- How do data requirements change for differing regulatory use cases (e.g., post-market commitments, label expansion, improving dose selection, and defining safety in broader populations)?
- How do we accommodate changing data characteristics and needs for RWD over time?
- What specific adjustments (if any) need to be made to the EHR recordings to allow data transfer to the FDA?

Whereas multiple comprehensive EHR platforms collecting health information (i.e., EPIC) already exist, the individual modules vary by disease specialty and are frequently proprietary and not interoperable. Thus, it is pertinent to initially determine whether any of the collected information could be tested to meet a regulatory threshold.
- Do data fields vary with study design? If so, could these be grouped?
- What data fields are needed to address a study question?
  - Demographics
  - Diagnosis (data, test, treatment and length of treatment)
  - Efficacy outcome(s); therapy changes; subsequent treatments
  - Co-morbidities
  - Toxicities and side-effects
  - Other
- Could the above questions be tested in advance of a pilot to determine feasibility/data extraction, such as designing a simple randomized non-interventional study?

Case Studies

Taking the above challenges into consideration, the work group reviewed scenarios where RWD has been collected and identified opportunities to apply this evidence towards answering specific clinical questions in routine clinical care. The following case studies, while broad in scope, are intended to illustrate possible uses for RWE collection.

Safety (Ceritinib) – In 2014, the FDA approved ceritinib (Zykadia) for the treatment of patients with anaplastic lymphoma kinase-positive (ALK⁺) metastatic non-small cell lung cancer (NSCLC) who have progressed on or are intolerant to crizotinib. Anecdotal patient reports suggested that taking ceritinib with food may improve gastrointestinal (GI) tolerability but may lead to increased systemic exposure of the drug. The observed safety data led to a post-market commitment[7] to evaluate a lower dose of ceritinib taken with a meal that potentially improves GI tolerability. Furthermore, during FDA's review of ceritinib, the FDA noted signs and symptoms of pancreatitis (pancreatic enzyme elevations in addition to gastrointestinal symptoms) in several cases, but there was only one case of investigator-reported pancreatitis occurring in a supportive clinical trial.[8] Exploration of RWD following approval of ceritinib could have provided additional information on the safety profile of the drug and its association with pancreatitis.

Furthermore, RWD collection on ceritinib use including: dose interruptions, dose modifications (with and without food), concomitant medications, GI toxicity, diarrhea, therapy duration, other adverse events, may contribute to the enhanced evaluation of an appropriate dose of the drug in the post-market setting.

Treatment Sequencing (Ramucirumab) – Docetaxel has been one of the standards of care for the treatment of second line metastatic NSCLC regardless of histology. In December 2014, ramucirumab (Cyramza) was approved in combination with docetaxel for patients who have progressed on a platinum based combination therapy and have received an EGFR or ALK based therapy if indicated. In October 2015, two new immunotherapies, nivolumab and pembrolizumab were approved in the second line (or higher) NSCLC setting; they have already showed significant clinical benefit. These agents may be used in different lines of treatment for these patients; indeed, they are even sometimes being utilized prior to chemotherapy.

While there is no rationale that the safety of ramucirumab plus docetaxel will be affected by prior treatment of an immune checkpoint inhibitor, approval was based on clinical trials conducted before immune checkpoint inhibitors entered the market. Formal clinical trials can test the impact of sequencing on the safety and efficacy profiles of these therapies, yet such trials can be time and cost prohibitive and further complicated by the fact that new treatments will potentially be approved during the course of the study. Overall, it is difficult to study all of the permutations of treatment sequence, especially in a landscape where the available options are changing yearly. RWE can potentially be a practical solution. Thus, collecting RWD on patient characteristics, safety, and mortality in patients with advanced NSCLC receiving treatment with ramucirumab plus docetaxel as well as PD-1 inhibitor, in any treatment setting using even a limited patient pool of approximately 100 patients that meet the criteria of the study population, may facilitate determination of appropriate treatment sequencing.

Orphan Drug Application (Denosumab) – On September 12, 2013, FDA granted orphan drug status for the use of denosumab (XGEVA) for the treatment of hypercalcemia of malignancy (HCM) based on results of a study conducted with EHR data from oncology clinics.[9] This represents the first-time ruling by the FDA for orphan drug designation based primarily on RWE. The RWE was provided because the published medical literature on HCM ranged from <1% to 30% depending on tumor type and studies were mainly from single institutions or focused on a single tumor type.[10]

The RWE featured in the orphan drug application (ODA) was based on analysis of the Oncology Services Comprehensive Electronic Records (OSCER) database (established by Amgen for observational research, today powered by Flatiron Health), which captures outpatient data for a representative sample of more than 569,000 cancer patients treated at 52 community- and hospital- affiliated oncology practices (565 clinics) from 2004 forward. The widespread adoption of EHR by community oncology practices makes this a valuable tool for observational studies in oncology. Specifically, EHR, which captures routine laboratory results (i.e., serum calcium and albumin values), were used to estimate the

---

[7] Leong R et al, J Clin Oncol 33, 2015 (suppl; abstr 2574).

[8] Khozin S et al, 2015. 21(11):2436-9.

[9] Gastanaga VM et al, Cancer Medicine (in press) 2016. 51(13):1704-13.

[10] Basso U et al, Curr Med Chem. 201.

prevalence of HCM by tumor type (confirming findings in the literature) and grade, and describe trends over a recent time period (2009–2013) including the use of bone resorptive therapies (intravenous bisphosphonates [pamidronate and zoledronic acid] and denosumab). Additionally, EHR analyses also provided described renal impairment among patients with HCM, and survival for a subset of patients with vital status via external data linkage.

Indication Expansion (Vemurafenib) – On August 17, 2011, the FDA approved vemurafenib (Zelboraf) tablets for the treatment of patients with unresectable or metastatic melanoma with the BRAF^{V600E} mutation as detected by an FDA-approved test. The major efficacy outcome measures of the trial were overall survival (OS) and investigator-assessed progression-free survival (PFS). While less common in NSCLC, BRAF mutations do make up between 1 and 3% of patients predominantly in adenocarcinoma with a history of smoking. A histology-independent phase 2 basket study observed vemurafenib activity in NSCLC.[11] Determining whether additional data on BRAF patients in NSCLC can be extracted from EHRs may help build the case for an expanded indication without the need to confirm through a traditional clinical trial.

Indeed, supplementing patients' data from the Basket trial with about 40-50 patients with data on real world response rate, duration of therapy, prior treatment, and safety may be sufficient describe patient responses in BRAF V600E Mutation-Positive NSCLC.

Confirmatory Studies (Crizotinib) – In August 2011, FDA granted accelerated approval to crizotinib (Xalkori) for the treatment of patients with locally advanced or metastatic NSCLC that is ALK-positive. Full approval was contingent on the completion of two phase 3 randomized clinical studies in treatment-naïve (N=343) and in previously treated (N=347) ALK$^+$ NSCLC. Full approval was granted November 2013, based upon PFS results from the trial in treatment-naïve patients. A sNDA label update for the second phase 3 study in previously treated patients was approved September 2015. In 2016, crizotinib received BTD for the ROS1-positive development program and the sNDA application for patients with ROS1-positive disease was granted Priority Review and received approval in March 2016.

In retrospect, and in the context of breakthrough activity of crizotinib in the selected patient population, would a RWE study have been appropriate as a confirmatory study for ALK$^+$ NSCLC? Following crizotinib approval, a retrospective real world cohort study was conducted in the United States and Canada utilizing medical record review of 212 ALK$^+$ NSCLC patients who initiated crizotinib as first or later line therapy[12] This study provided further information on crizotinib use and outcomes of patients and was supportive of the phase 3 clinical studies. For example, response rates seen in the real world cohort study (66% overall; 69% in first line and 60% in second or later line) were similar to the response rates seen in treatment-naïve patients (74%)[13] and previously treated patients (65%)[14] in the clinical studies. In the real world study, one-year survival rates in first-line patients (85%) from the real world chart review was also similar to the one-year survival rate seen in the clinical study of treatment-naïve patients (84%).[15] These real world data provide support for the benefits of crizotinib in patients with ALK$^+$ NSCLC and are in line with data previously reported in clinical studies.

Following the early phase results, could RWD supplement, or replace, traditional requirements for post-market commitments in future development programs? A better understanding of whether real world studies are able to confirm clinical trial results would lend credence to this idea. Ultimately, observations confirming clinical trial findings may provide opportunities for novel trial designs that incorporate real world evidence earlier into development.[16]

Pilot Studies
Building on the above examples, this work group considered opportunities for designing prospective pilot studies to assess the feasibility of using RWE to support regulatory decisions. The key goal of this exercise would be to test and validate data collection efforts and identify novel endpoints that correlate with clinical benefit and reflect correlations between clinical practice and trial settings. Possible approaches for developing a pilot study are captured below.

---

[11]Hyman DM et al, NEJM. 2015.
[12] Davis KL et al, Presented at 16th World Conference on Lung Cancer, September 6-9, 2015; Denver, CO.
[13] Solomon B et al, NEJM. 2014. 371(23):2167-77.
[14] Shaw AT et al, NEJM. 2013. 368(25):2385-94.
[15] Solomon B et al, NEJM. 2014. 370(11):988-90.
[16] Koehler M et al, Ann Oncol. 2016. In print.

Immune checkpoint Inhibitors in tumors with high rates of mutational burden

Among various tumor types, melanoma and NSCLC patients have been the best responders to immunotherapies, possibly due to the number of somatic mutations present. Indeed, early studies point to an association between tumor mutational load and efficacy. For example, patients with high rates of microsatellite instability (MSI-high), a marker of defective mismatch repair mechanisms, have been observed to respond remarkably well to PD-1 and other immune checkpoint inhibitors. While most fully characterized in colorectal cancer (CRC), clinical reports in other gastrointestinal malignancies, and gynecological cancers among others are increasingly available. A recent report suggested that the objective response rate for the PD-1 inhibitor, pembrolizumb, in CRC was 40% when there was evidence of MSI-high (versus 0%) in patients with proficient mismatch repair. Additionally, 5 of the 7 non-CRC patients studied also responded to treatment.[17] Studies evaluating the link between MSI-high and immune checkpoint blockade are already ongoing and there may be additional biomarkers of mutational burden to better identify responders including, quantifying mutational load or identifying mutations in other DNA repair proteins.

To date, studies have not shown that immune checkpoint inhibitors are safe and efficacious for widespread use in highly mutated tumors. Hence, the clinical rationale for prescribing PD-1 inhibitors for people with evidence of mutational burden, such as microsatellite instability, exists but is insufficient. Based on this rationale, a prospective study of currently approved PD-1 inhibitors in patients with highly mutated tumors, could be addressed in the real world setting.

Determining the feasibility of this pilot would initially require a retrospective analysis of existing databases (e.g., the Flatiron Health dataset plus targeted chart abstraction and linked claims data) to address outstanding questions, such as, testing patterns (i.e., timing, disease state, test type, etc.) and treatment patterns (including toxicities and observed outcomes) in cancers with evidence of mutational burden. Once all the necessary components are identified, building a prospective trial would depend on scoping (i.e., cancer type, study size, etc.), optimizing the biomarker and test use, and ultimately defining mechanisms to measure efficacy.

Proposals for RWE Applications

With numerous advantages to collecting RWD, ranging from supplementing post-market data collection, decreasing costs and development timelines, defining novel outcomes, and minimizing the number of patients exposed to a less efficacious therapy, this working group proposes utilizing RWD with the intent of answering specific clinical questions and, when appropriate, informing product labels, in the following areas,

1. Expanding the safety profiles of a therapeutic,
2. Identifying populations with enhanced benefit/risk for an already approved therapy to inform clinical practice,
3. Piloting studies to determine the potential correlation between feasible real world measures (i.e., time to treatment switching) and more traditional clinical trial endpoints (i.e., time to progression),
4. Building evidence for a supplemental package to expand the indication profile for a therapeutic,
5. Supporting efficacy results observed in clinical trial setting, particularly in areas of unmet medical need, when a new drug shows substantial clinical benefit. Real world studies that are able to support the preliminary magnitude of effectiveness in a larger cohort may be sufficient to serve as post-market confirmation of clinical benefit.

These proposals are intended to guide developers in considering RWD collection during drug development; however, careful consideration and discussions with regulatory agencies will be needed in order to account for any observed outcomes, such as loss of efficacy, within RWD.

---

[17] Le DT et al, NEJM; 2015. 372(26):2509-20.

Appendix A: Detailed outline for the Ramucirumab case study

Clinical questions to be addressed utilizing RWE	The objectives of this study are to describe patient characteristics, safety and mortality in patients with advanced NSCLC receiving treatment with ramucirumab plus docetaxel (R/D) as well as a PD-1 inhibitor (PD-1) in any sequence. Included patients will be those who received the R/D in any treatment setting as well as a PD-1.
	Key questions include:
	• What are the demographic, clinical and treatment characteristics for the cohort?
	    ○ Demographics: age; gender; race
	    ○ Clinical characteristics: smoking status; ECOG performance status; stage at diagnosis; biomarker status (EGFR; ALK; ROS1; PD-L1; KRAS) time between diagnosis and advanced disease; histology
	    ○ Treatment history: use of systemic and targeted therapies
	    ○ Reimbursement: insurance status
	• What is the treatment sequencing of R/D, PD-1 and other therapies in this population?
	• What outcomes are observed when R/D precedes PD-1 vs comes after PD-1?
	    ○ What OS is observed?[18]
	    ○ What safety events are observed?[19]
	Proposed Milestones
	1. Feasibility, scoping and study design, presented at the 6/16/2016 Friends/Alexandria Summit meeting on RWE (intention is to present high-level study scope and design)
	2. Full analysis and report once the data matures
	3. Potential labeling updates based on the outcomes of the study

Included data sources	In order to address these questions, we plan to leverage RWD and specifically EHR data. The proposed data source is the Flatiron dataset, which comprises longitudinal, patient-level EHR data. The Flatiron dataset incorporates the data from both structured and unstructured EHR data streams. For unstructured sources in particular (e.g., physician notes, pathology reports, etc.), Flatiron uses technology-enabled abstraction to curate data points at scale. This enables Flatiron to consistently and accurately pull some of the most elusive clinical details out of the EHR. Each patient record passes through the technology-enabled process, with human review of each data element collected to confirm patient information (such as demographics, diagnosis, stage, histology, etc.). This abstracted information is often missing and/or inaccurate if relying on the structured EHR fields alone. By combining the processed structured and unstructured data, Flatiron can create a longitudinal view of each patient, tracking key events, interactions, and therapies over time. Flatiron's EHR dataset is already annotated with mortality data; No additional linkages are needed to add these variables. The mortality variable represents an amalgamation of internal and external data sources to represent the best understanding of a patient's vital status and date of death. Data is sourced from the EHR as well as obituaries, funeral homes, and other sources. No other data sources are needed, based upon the current scope of the research questions. If needed, the Flatiron EHR data can be linked to healthcare claims data of other sources.
Data merging processes and challenges	Data linkage is not necessary in order to accomplish this project, but would be doable if needed. Flatiron retains the underlying patient identifiers (in a HIPAA-approved manner), and can link directly to external data sources. Because Flatiron maintains access to the patient identifiers required to link patient-level data to external data sources, Flatiron must take extra precautions to ensure linking occurs in a de-identified manner. Flatiron works with a third party de-identification expert to oversee the linking process and confirm that all linked data is certified as de-identified.
Variables needed to address clinical questions	Flatiron will abstract the data variables included in the data model listed below to support the key research objectives. As noted above, Flatiron processes both structured data (i.e., data points that are organized in a predefined manner, such as dropdown fields) and unstructured data (e.g., free text from a physician note or a scanned pathology report). Together, these patient-level data provide a complete view of each patient with resolution at the indication, testing and treatment level.

---

[18] Since the potential follow up time for PD1 inhibitor therapy is shorter overall, it is likely necessary to restrict the cohort to people receiving therapy for metastatic NSCLC after the date of first approval of PD1 inhibitors in lung cancer in order to reduce bias.
[19] Exact safety events to be confirmed prior to study initiation

Structured and unstructured data variables	Preliminary data model to be confirmed upon formal study scoping with Lilly: **Structured Data:** • Demographics (to be determined based on de-identification requirements) • Diagnosis (ICD9 codes, ICD10 codes and dates) • Visits (date and type) • Labs (test date, result date, test name, result with units, and normal ranges) • Medication Administration (date, drug, dose and units) • Medication Orders (date, drug, dose and units, route) • Insurance • Performance Status (ECOG) **Unstructured Data:** • Date of initial diagnosis • Date of diagnosis of advanced disease: first recurrence or metastasis • Group stage at time of initial diagnosis • Documented history of smoking • Biomarker status: EGFR, ALK, PDL1, KRAS, ROS1 (including testing status, test result, and test date) • Safety events (specific events to be defined in collaboration with Lilly) **Derived Data Elements** (combine structured & unstructured data elements using clinician-defined business rules)**:** • Lines of therapy • Mortality
Challenges/ considerations with regard to consistent/ reproducible variable extraction process	Flatiron maintains robust QA/QC records to support data credibility and provenance requirements for RWE use cases. Documentation includes: • Cohort selection criteria • Overview of the source data, including completeness, inter-abstractor agreement, and kappas for each structured and unstructured variable • Analytic notes • Business rules used to develop each derived variable, including an audit trail of changes Specific data provenance initiatives include, but are not limited to: • Fulfillment of HIPAA certification requirements • Secure HiTrust certification • Data freeze and retention processes Flatiron also maintains internal QA/QC processes. For enhanced data captured by Flatiron abstractors from unstructured fields, in particular, Flatiron has developed multiple tools to monitor and measure quality. All Flatiron Health data abstractors are experienced oncology nurses, clinical research associates, or trained tumor registrars and continuous quality monitoring governs the abstraction process. Abstractor reliability scores, designated by kappas, are required to remain within a defined range (specific to the diagnosis and data model for each module) during both initial training and ongoing quality assurance (QA) checks. Furthermore, initial training and testing is conducted on actual EHR records against "gold standard" outputs (as defined by our internal oncologists) to ensure abstractor comprehension of specified data points and policies and procedures prior to initiating "live" abstraction on a set of tasks. Data elements are duplicate abstracted at the start of each new module to confirm agreement and ensure consistency of variable collection. The quality of each variable is monitored on an ongoing basis and medical outliers or edge cases are escalated via a "Review Panel" for adjudication by our Abstraction team leads, QA specialists, and/or medical oncologists.
Endpoints to be defined and associated challenges	Surrogate endpoints are not planned. This study will look at OS and safety. Specific safety events will be defined and scoped in collaboration with Lilly. Flatiron will abstract safety events as documented by the physician in the underlying chart. While Flatiron will be able to provide details around specific safety events that are recorded in the chart, visibility into specificity or severity of each event is sometimes limited. Flatiron has found that physicians typically do not systematically document grading in the chart in the real world; though where available, Flatiron would collect this information. Depending on the specific safety event, Flatiron can collaborate with Lilly to develop proxies for grading based on lab values that are in the structured data. As part of the development process, Flatiron plans to build a quality validation plan in order to assess the quality of the abstracted safety events. This plan will be developed collaboratively with Lilly and/or other key

	stakeholders in order to ensure that the resulting data represent a robust endpoint to support this research question.
Collection of patient level RWD	As of January 31, 2016, we have approximately 80-90 patients in the Flatiron dataset that generally meet the criteria for the study population. Since ramucirumab and PD-1 inhibitors were fairly recently approved, we anticipate that the use of these drugs will increase quickly over time as clinicians become more familiar with them and patients receive successive lines of therapy (e.g., 2nd line and beyond). We are seeing new ramucirumab patients added to our advanced NSCLC dataset at a rate of approximately 10-20 patients per month; New PD-1 patients are being added to our dataset at a rate of approximately 200 per month. Further, we anticipate that time will be needed for safety and survival outcomes to accrue with the population; A minimum of 6 months median follow up is likely needed.
	Based upon the above, we anticipate that the sample size available for review at a June meeting with Friends of Cancer Research (*Friends*) is a cohort of approximately 100 patients. Maturing of the dataset beyond June will add patients and longitudinal safety + outcomes data. The exact timeline and implications for sample size will be determined collaboratively with Lilly based upon review of the initial cohort and projections. Upon delivery of the dataset, Flatiron will provide full documentation of data completeness.
HIPPA/informed consent issues, regarding submission of patient level data to regulatory agencies and auditing of individual data. What new processes/policy changes need to be in place?	Deidentified patient level data can be used for this purpose without patient level consent. Flatiron partners with oncology care providers through several software products, in aggregate referred to as OncologyCloud™. Flatiron executes Business Associate Agreements (BAAs) with every cancer specialist using these software products thereby allowing Flatiron to pull a copy of the patient medical record into a central repository for data processing under the TPO exemption of HIPAA. Structured and unstructured data processing renders the main dataset, which is then deidentified and stored separately for secondary research purposes. Flatiron maintains the full patient chart for each patient in Flatiron's network, enabling an adaptive data model and the ability to conduct supplemental abstraction to support studies at any point in time. The Flatiron data repository, data processing approach, and approach to retrospective research is described in a protocol approved by the New England IRB.
	Flatiron maintains full documentation of the data abstraction processes and quality metrics (as described above), allowing for audits should the need arise. Individual data can be submitted to regulatory agencies (assuming deidentification requirements are met).

Appendix B: Glossary of Terms. The following terms, while still evolving, have defined for the purposes of this white paper here.

Breakthrough Therapy Designation (BTD) - Created in 2012, as part of the Food and Drug Administration Safety and Innovation Act (FDASIA) and intended to: 1. treat a serious or life threatening disease or condition and 2. preliminary clinical evidence indicates that the drug may demonstrate substantial improvement over existing therapies on one or more clinically significant endpoints, such as substantial treatment effects observed early in clinical development.

Electronic Health Records (EHRs) - A digital version of a paper chart that contains all of a patient's medical and clinical history throughout the patient's care and used by providers for diagnosis and treatment.

Pragmatic clinical trials (PCTs) - Are prospective randomized intervention studies that leverage the existing clinical infrastructure and are designed to test interventions in everyday clinical settings to maximize therapeutic applicability and generalizability.

Real World Data (RWD) - Data collected from sources outside of conventional randomized controlled trials – for example, from electronic systems used in health care delivery and to track patient experience with care – are commonly referred to as real-world data

Real World Evidence (RWE) - Evidence derived from the use, benefits and risks of medicines that fall outside the bounds of the classic clinical trial settings, including use of data that is routinely collected in the daily practice of medicine, and thus reflective of the heterogeneous patients seen in real world practice settings.