NON-INTERVENTIONAL (NI) STUDY PROTOCOL

Study Information

Title	Comparative Effectiveness of Elranatamab (PF-06863135) in Clinical Study C1071003 Versus Standard of Care (SOC) in Real-World (RW) External Control Arms in Patients with Triple-Class Refractory (TCR) Multiple Myeloma (MM)
Protocol number	C1071024
Protocol version identifier	1.0
Date	15 August 2022
Active substance	Elranatamab (PF-06863135)
Research question and objectives	**Primary objective:** 1. To compare the objective response rate (ORR) among TCR MM patients treated with elranatamab in Study C1071003 with a comparable cohort of TCR MM patients receiving SOC therapy within external comparator arms. **Secondary objectives:** 1. To compare time to treatment response (TTR) in TCR MM patients treated with elranatamab in Study C1071003 with a comparable cohort of TCR MM patients receiving SOC therapy. 2. To compare the duration of response (DOR) in TCR MM patients treated with elranatamab in Study C1071003 with a comparable cohort of TCR MM patients receiving SOC therapy.

Author	PPD , PhD PPD

1. TABLE OF CONTENTS

2. LIST OF ABBREVIATIONS

Abbreviation	Term
ADC	Antibody-drug conjugate
AE	Adverse Event
Anti-CD38	Anti-CD38 monoclonal antibodies
AIPW	Augmented Inverse Probability Weights
ATE	Average Treatment Effect
AST	Aspartate Aminotransferase
ALT	Alanine Aminotransferase
BCMA	B-Cell Maturation Antigen
BM	Bone Marrow
CAR	Chimeric Antigen Receptor
CI	Confidence Interval
CR	Complete Response
DOR	Duration of Response
ECOG	Eastern Cooperative Oncology Group
EHR	Electronic Health Record
EMA	European Medicines Agency
FDA	Food and Drug Administration
FLC	Free Light Chain
GBS	Guillain-Barre Syndrome
GVHD	Graft Versus Host Disease
HBV	Hepatitis B Virus
HCV	Hepatitis C Virus
HIV	Human Immunodeficiency Virus
IMiD	Immunomodulatory Drug
IMWG	International Myeloma Working Group
IPTW	Inverse Probability of Treatment Weights
KarMMa-RW	Comparison of Idecabtagene Vicleucel with Real-World Outcomes in Relapsed and Refractory Multiple Myeloma
LVEF	Left Ventricular Ejection Fraction
LOT	Line of Therapy
MAMMOTH	Monoclonal Antibodies in Multiple Myeloma: Outcomes after Therapy Failure
MM	Multiple Myeloma
M-protein	Monoclonal immunoglobulin protein
MUGA	Multigated acquisition scan
NI	Non-Interventional
OR	Objective Response
ORR	Objective Response Rate
OS	Overall Survival
PD	Progressive Disease

Abbreviation	Term
PI	Proteasome Inhibitor
PFS	Progression-Free Survival
POEMS	Polyneuropathy, Organomegaly, Endocrinopathy, Monoclonal Gammopathy, and Skin Changes
PR	Partial Response
PS	Propensity Scores
QT	Time from the beginning of the QRS Complex to the End of the T Wave
QTcF	Corrected QT (Fridericia Method)
RW	Real-World
RWD	Real-World Data
RRMM	Relapsed/Refractory Multiple Myeloma
SAP	Statistical analysis plan
SARS-Cov2	Severe Acute Respiratory Syndrome- Coronavirus 2
sCR	Stringent Complete Response
SMD	Standardized Mean Difference
SOC	Standard of Care
SPEP	Serum Protein Electrophoresis
TCR	Triple-Class Refractory
TTR	Time to Treatment Response
ULN	Upper Limit of Normal
UPEP	Urine Protein Electrophoresis
US	United States
VGPR	Very Good Partial Response

3. RESPONSIBLE PARTIES

Principal Investigator(s) of the Protocol

Name, degree(s)	Job Title	Affiliation	Address
PPD PhD	PPD		
PPD, MPH	PPD		
PPD, MD, PhD	PPD		
PPD, PharmD, MBA	PPD		
PPD, PhD	PPD		
PPD, MPH	PPD		
PPD, MD, MPH	PPD		
PPD, MPH	PPD		

4. ABSTRACT

Not applicable

5. AMENDMENTS AND UPDATES

None

6. MILESTONES

Milestone	Planned date
Start of data collection	03 October 2022
End of data collection	16 October 2022
Final study report	11 December 2022

7. RATIONALE AND BACKGROUND

MM is a rare and progressive type of cancer that is characterized by the rapid accumulation of monoclonal plasma cells in the bone marrow (BM). This results in the production of monoclonal immunoglobulin protein (M-protein) and eventually leads to end-organ damage. The global incidence of MM is rising and is estimated at 160,000 in 2018, with 106,000 deaths attributed to MM annually. In the past 3 decades, from 1990 to 2016, the global incidence of MM has increased by 126%, with the highest incidence rates reported in Australia, Western Europe, and the United States (US)(Padala *et al.*, 2021).

Over the years, advancements have been made in the discovery and utilization of MM therapies, with current SOC including proteasome inhibitors (PIs) (bortezomib, carfilzomib, ixazomib), immunomodulatory drugs (IMiD) (thalidomide, lenalidomide, pomalidomide) and anti-CD38 monoclonal antibodies (anti-CD38: daratumumab and isatuximab)(Wallington-Beddoe and Pitson, 2017).

Despite these treatment advances, a major proportion of MM patients will fail all 3 classes of the current SOC regimens (including PI, IMIDs, and anti-CD38) and become triple class refractory (TCR). The treatment of TCR MM presents a therapeutic challenge because of the inherent clonal heterogeneity and genetic instability of MM tumor cells, influencing the eventual development of therapeutic resistance (Nooka *et al.*, 2015) . Therefore, as TCR patients progress and are exposed to an increasing number of therapies, the duration of response (DOR) decreases along with overall survival (OS) until the disease is ultimately fatal (Yong *et al.*, 2016).

Pfizer has developed elranatamab (PF-06863135), an investigational B-cell maturation antigen (BCMA) CD3-targeted bispecific antibody, for the treatment of patients with TCR MM. Belantamab mafodotin-blmf was the first BCMA therapy approved by the FDA, in August 2020. The preliminary results of the Phase 1 study (NCT03269136) demonstrated manageable safety of elranatamab and a response rate of 69% (ie, 9/13 patients) at the recommended dose of 1000 µg/kg (Sebag *et al.*, 2022). Additionally, elranatamab has been granted Fast Track Designation by the US Food and Drug Administration (FDA) and European Medicines Agency (EMA) to aid the rapid development and review of this novel agent and thereby address the unmet medical needs of TCR MM patients (*Pivotal Phase 2 MagnetisMM-3 Trial| Pfizer*, 2022). Pfizer is currently conducting a Phase 2 study to evaluate the clinical benefit of elranatamab in TCR MM patients (Study C1071003). Considering the ethical and practical challenges associated with conducting a randomized clinical trial within a particularly difficult-to-treat population of MM patients with unmet needs and no clear single SOC, a single arm design was adopted for the Phase 2 study (*First Data from Planned Interim Analysis of Pivotal Phase 2 MagnetisMM-3 Trial | Pfizer*, 2022). Thus, there is a need to generate evidence on the clinical outcomes of TCR patients who receive SOC to help contextualize the results from Study C1071003. This will be accomplished by creating an external control arm using Real-World Data (RWD) to support the ongoing evaluation of elranatamab.

Previous RW studies have provided data on clinical outcomes in patients with relapsed/refractory MM (RRMM) using RWD. In the MAMMOTH study (Gandhi *et al.*,

2019), 275 patients from 14 academic institutions were identified in the US with a diagnosis of active MM who were refractory to anti-CD38, daratumumab, or isatuximab, administered as monotherapy or in combination. Median OS from the time when patients were refractory to an anti-CD38 was 8.6 months (95% confidence interval, CI: 7.5 - 9.9). The ORR to the first regimen after patients became refractory to anti-CD38 was 31% and the median progression-free survival (PFS) and OS in patients in the same group were 3.4 and 9.3months, respectively.

The KarMMa-RW study was a global NI, retrospective study that assessed treatment patterns and outcomes in RW RRMM patients treated with currently available therapies, and whose characteristics were similar to the KarMMa study, a Phase 2 single-arm study of idecabtagene vicleucel in heavily pretreated RRMM patients who were triple-class exposed and refractory to the last regimen (Jagannath *et al.*, 2020). Approximately 43% of the KarMMa-RW population were TCR. The ORR of the 190 eligible RRMM patients in the RW setting was 32% (95% CI: 24.4 - 42.3), and 29% in the matched adjusted eligible RRMM subset. In the eligible RRMM cohorts, the median PFS was 3.5 months, and the median OS was 14.7 (95% CI: 14.0-15.4) months.

Regulatory Agencies, including FDA and EMA, have recognized the utility of external control arms derived from RWD in decision making, and have established guidance on conditions that should be satisfied before an external control approach is deemed appropriate, as well as key methodological considerations (EMA, 2001; FDA, 2018; FDA, 2019). According to these guidelines, an external control approach is considered appropriate under certain conditions, including(a) there is a high unmet need (eg, serious rare diseases), (b) there is a well-documented, highly predictable disease course that can be objectively measured and verified, and (c) there is an expected drug effect that is large, self-evident, and temporally closely associated with the intervention (CHMP, 2002; Jahanshahi *et al.*, 2021).

This study aims to compare the ORR, TTR, and the DOR in patients from Study C1071003 treated with elranatamab versus RW TCR MM patients treated with SOC. To reduce the potential for bias, external control arms will be constructed by selecting fit-for-purpose RW data (RWD) sources (ie, reliable and relevant)(FDA, 2018), appropriate comparative effectiveness methods, and robust statistical techniques (eg, inverse probability of treatment weighting [IPTW]) (Carroll, 2016; Ghadessi *et al.*, 2020; Chesnaye *et al.*, 2022).

8. RESEARCH QUESTION AND OBJECTIVES

This study aims to assess the comparative effectiveness of elranatamab versus SOC treatment in TCR MM patients using external control arms for the open-label, multicenter, non-randomized single-arm Phase 2 Study C1071003.

Primary objective:

To compare ORR among TCR MM patients treated with elranatamab in Study C1071003 with a comparable cohort of TCR MM patients receiving SOC therapy within external comparator arms.

Secondary objectives:

1. To compare TTR in TCR MM patients treated with elranatamab in Study C1071003 with a comparable cohort of TCR MM patients receiving SOC therapy within external comparator arms.
2. To compare DOR in TCR MM patients treated with elranatamab in Study C1071003 with a comparable cohort of TCR MM patients receiving SOC therapy within external comparator arms.

9. RESEARCH METHODS

9.1. Study design

This retrospective cohort study builds upon 2 previous studies, which identified 2 cohorts of RW TCR MM patients to use as external control arms for Study C1071003. The cohorts were identified from 2 of the US-based oncology electronic health record (EHR) databases, Flatiron Health, and COTA (see relating study protocols in ANNEX 1. LIST OF STAND ALONE DOCUMENTS). In the current study, 2 external arms will be constructed from these 2 cohorts of RW TCR MM patients to maximize comparability to patients from Study C1071003 (see Section 9.2.3 for more detail).

MM patients eligible for selection into external control arms will be those patients who are refractory to at least 1 PI, 1 IMiD, and 1 anti-CD38, and have started at least 1 new treatment since documentation of TCR status. Refractory is defined as having disease progression, according to International Myeloma Working Group (IMWG) criteria or clinical assessment, while on therapy or within 60 days of the last dose in any line of therapy (LOT), regardless of response. However, if a subsequent LOT is initiated and a progression is observed after at least 30 days of the LOT start, the patient will be considered refractory to the subsequent LOT even if the progression occurred within 60 days after the last dose of the preceding LOT. Details on the operational definitions of IMWG criteria for progression used in Flatiron and COTA, and a side-by-side comparison with the criteria used in Study C1071003 are provided in the relating study protocols in the APPENDIX.

In the RW setting, no single SOC currently exists for TCR MM patients, and combinations of treatments are frequently used in lieu of monotherapy (Mikhael, 2020). For simplicity, in this protocol, the term “SOC” refers to all standard treatment options available for TCR MM patients. Selection of TCR MM patients initiating a new treatment in the external control arm

enables comparability of patients at a similar stage in disease progression following TCR documentation.

The date of initiation of the first regimen after TCR MM eligibility will be defined as the index date. Only patients with an index date occurring between 16 November 2015, and 31 March 2022 will be selected (the first anti-CD38 therapy was approved by the FDA on 16 November 2015). The study period will be comprised of the baseline period (time preceding the index date) and the observational period (time following the index date). The observational period will span from the index date to the earliest of death, or the latest available patient's record, whichever comes first. Clinical outcomes of interest will be ORR, TTR, and DOR.

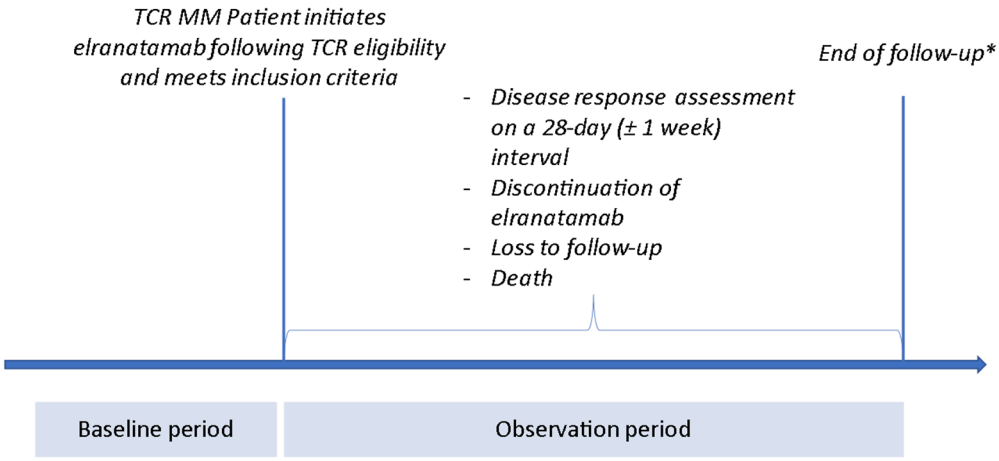
For the primary analyses, differences in baseline and key covariate characteristics including treatment history and disease-related characteristics at the index date between patients in Study C1071003 and each external control arm will be balanced using IPTW (see Section 9.7.2).

Sensitivity analyses will be conducted to evaluate the effect via doubly robust estimation using augmented inverse probability weights (AIPW), alternative inclusion/exclusion criteria, and, if feasible, to evaluate any differences in the magnitude of treatment effect in patients with and without prior exposure to BCMA-directed therapy (see in Section 9.7.3).

Finally, a quantitative bias analysis (nullification analysis) will be performed to evaluate the robustness of results in the presence of potential threats to internal validity see Section 9.7.3).

A schematic figure of Study C1071003 and external control arms is provided in Figure 1 and Figure 2, respectively.

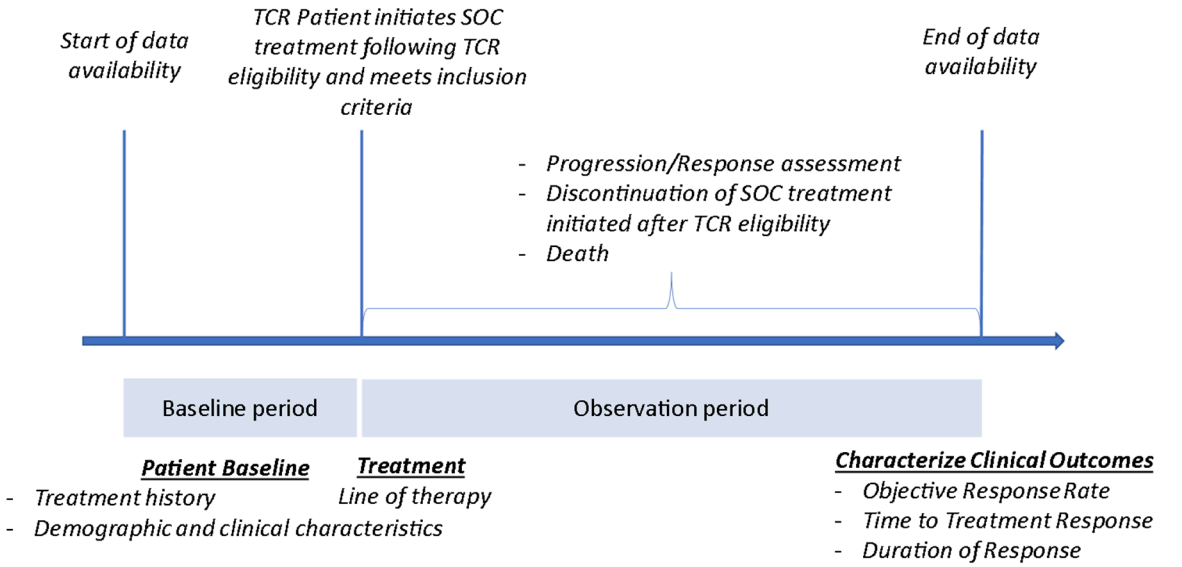
Figure 1. Baseline and observation periods in Study C1071003



** For this protocol, all participants will be followed for at least 9 months from the date of enrollment*

Abbreviations: MM=multiple myeloma; TCR=triple -class refractory.

Figure 2. Baseline and observation periods in external control arms



Abbreviations: MM=multiple myeloma; TCR=triple -class refractory; SOC=standard of care

9.2. Setting

9.2.1. Study C1071003

Study C1071003 is an open-label, multi-center, non-randomized Phase 2 study of elranatamab (PF-06863135) monotherapy, which was initiated in February 2021 (Pfizer,

2021). To determine the effects of prior BCMA-directed therapy on the response to elranatamab monotherapy, Study C1071003 enrolled 2 independent and parallel cohorts, 1 with patients who are naïve to BCMA-directed therapies (Cohort A; approximately 120 patients) and the other with patients previously exposed to BCMA-directed therapy (Cohort B; approximately 60 patients).

9.2.2. Real-world data sources

The cohorts of RW TCR MM patients for the external control arm were previously extracted from Flatiron Health and COTA databases that collect information on MM patients treated across the US. Flatiron and COTA have been selected according to data availability and as fit-for-purpose for fulfilling study objectives (the US FDA, 2018 Framework). The dates for the study periods for each RW database have been selected to align as closely as possible to each other and to Study C1071003 which will include at least 9 months of follow-up.

Flatiron Health

The Flatiron Health database is a longitudinal, demographically, and geographically diverse database derived from EHR data. Flatiron includes data from over 280 community cancer centers and academic institutions (~800 sites of care) representing more than 2.4 million active the US cancer patients available for analysis. The source population is the overall population reported in the EHR and includes patients managed in at least 1 of the US oncology centers taking part in the Flatiron Health network from 01 January 2011 onwards.

Flatiron Health databases consist of longitudinal data on MM patients, including normalized data from structured EHR, enhanced data on patients' characteristics abstracted from unstructured EHR, and derived data that is created based on Flatiron-specific business rules (Flatiron Health, 2021).

Structured data include information on patients' demographics, visit dates, diagnoses, vitals, medications, the Eastern Cooperative Oncology Group (ECOG) performance status, and laboratory tests. Unstructured data contain additional information on comorbidities, biomarker reports, and details of transplants. Flatiron Health also provides derived data such as progression and response variables where algorithms are developed following adapted IMWG criteria to identify these events throughout the course of a patients' journey.

Details on the identification of the external control arm using Flatiron Health data can be found in the related protocol as Document 001 in ANNEX 1. LIST OF STAND ALONE DOCUMENTS.

COTA

COTA maintains a multidisciplinary data curation approach. The COTA database is a longitudinal database derived from the EHR of healthcare provider sites including academic institutions, community centers, and hospital systems representing 500,000 patients from over 200 sites of care in the US. Data elements are standardized across sources and ontologies to create a single, structured dataset to cover the full longitudinal history of a patient's clinical care. For each patient, the datasets contain information from the time of initial MM diagnosis to the most recent documentation in the EHR. Patients with insufficient

documentation or those having significant gaps in their patient history in the EHR are excluded from COTA datasets (COTA,2022).

As with Flatiron Health, COTA collects information from structured and unstructured data sources. Derived data based on adapted IMWG criteria are also provided (eg, disease progression and treatment response). COTA differs from Flatiron Health in the provision of specific databases that are dedicated to reporting factors relevant to the experience of a patient with MM and includes datasets containing abstracted information on bone lesions, plasmacytomas, radiation therapy, and adverse events (AE).

Details on the identification of the external control arm using COTA data can be found in the related protocol as Document 002 in ANNEX 1. LIST OF STAND ALONE DOCUMENTS.

9.2.3. Inclusion and exclusion criteria for external control arms

Comparability of patients between Study C1071003 and the external control arms is 1 of the key considerations to minimize bias. Care will be taken to ensure that patients identified from each RW TCR MM cohort are similar to the Study C1071003 patients by satisfying pre-specified inclusion and exclusion criteria. Inclusion and exclusion criteria from Study C1071003 are provided in Table 1 and will be modified for each cohort of RW TCR patients based on data availability. Operational definitions of inclusion and exclusion criteria for each RWD source are summarized in the APPENDIX.

Due to the missingness present within RWD sources, there is an inherent tradeoff between analyzable sample size and the degree of comparability of patients from Study C1071003 with the external control arm sample. In other words, the more closely the inclusion/exclusion criteria for the external control arms are aligned to Study C1071003, the smaller the available sample size. From each cohort of RW TCR MM patients, we will establish 2 samples of external control arms (“critical eligibility criteria” sample and “expanded eligibility criteria” sample). Both samples will incorporate the inclusion/exclusion criteria from Study C10701003 which represent key clinical variables to define the cohort. However, 1 of the samples (“expanded eligibility criteria”) will also incorporate additional inclusion/exclusion criteria from Study C1071003 (eg, lab values) at the expense of a smaller sample size.

The critical eligibility criteria sample will constitute the primary analysis set. Sensitivity analyses will be conducted using the expanded eligibility criteria to assess the potential influence of the selection process on the observed effects (ie, primary analyses will be repeated using this sample). The number of patients included/excluded at each step of the study will be reported.

For all samples, patient selection into the TCR MM external control arm will include:

1. Having a diagnosis of MM (identified previously).
2. Those who are refractory to at least 1 PI, 1 IMiD, and 1 anti-CD38 treatment (TCR MM). Refractory is defined as having disease progression, IMWG criteria or clinical assessment, while on therapy or within 60 days of the last dose (if the duration of a

subsequent LOT is ≤ 30 days at the time of reported disease progression) in any line, regardless of response (identified previously).

3. Patients start at least 1 treatment following their TCR eligibility. The index date will be defined as the date of initiation of the first regimen post-TCR MM eligibility.
4. Those having measurable disease according to IMWG criteria (Rajkumar *et al.*, 2014), ECOG performance status ≤2, and aged 18 years and older at index date.

A tabular view summarizing the inclusion and exclusion criteria of Study C1071003 which are available and applied for RW patients is provided in Table 1 (signified with a checkmark symbol where the critical and/or expanded selection criteria definitions are applicable).

Table 1. Inclusion and Exclusion criteria for Study C1071003, Flatiron Health, and COTA

Patient criteria per Study C1071003	Implementation in Flatiron Health		Implementation in COTA	
	Critical	Expanded	Critical	Expanded
Inclusion Criteria				
Male or female patients age ≥18 years	ü	ü	ü	ü
Willing and able to comply with all scheduled visits, treatment plans, laboratory tests, lifestyle considerations, and other study procedures	X	X	X	X
Prior diagnosis of MM as defined according to IMWG criteria (Rajkumar et al, 2014)	ü	ü	ü	ü
Measurable disease, based on IMWG criteria as defined by at least 1 of the following a) Serum M-protein ≥0.5 g/dL by SPEP b) Urinary M-protein excretion ≥200 mg/24 hours by Urine Protein Electrophoresis (UPEP) c) Serum immunoglobulin Free Light Chain (FLC)≥10 mg/dL (≥100 mg/L) and abnormal serum immunoglobulin kappa to lambda FLC ratio (<0.26 or >1.65)	ü	ü	ü	ü
Patients are TCR defined as being refractory to all 3 of the following: a) Refractory to at least 1 IMiD b) Refractory to at least 1 PI c) Refractory to at least 1 anti-CD38 antibody Relapsed/refractory to last anti-MM regimen	ü	ü	ü	ü
Initiated at least 1 anti-MM systemic therapy after becoming TCR eligible. The first systemic treatment initiation after becoming TCR eligible must not comprise any study or investigational agent.*	ü	ü	ü	ü
ECOG performance status ≤2	ü	ü	ü	ü
Adequate hepatic function characterized by all of the following: a) Total bilirubin ≤2 x Upper Limit of Normal (ULN) (≤3 x ULN if documented Gilbert's syndrome) b) aspartate aminotransferase (AST) ≤2.5 x ULN c) alanine aminotransferase (ALT) ≤2.5 x ULN	X	ü	X	ü
Adequate renal function, defined by an estimated creatinine clearance ≥30 mL/min (according to the Cockcroft Gault formula, by 24-hour urine collection for creatinine clearance, or according to the local institutional standard method)	X	ü	X	ü

Table 1. Inclusion and Exclusion criteria for Study C1071003, Flatiron Health, and COTA

Patient criteria per Study C1071003	Implementation in Flatiron Health		Implementation in COTA	
	Critical	Expanded	Critical	Expanded
Adequate BM function characterized by all of the following a) Absolute neutrophil count ≥1.0 × 109/L (use of granulocyte-colony stimulating factors is permitted if completed at least 7 days before planned start of dosing) b) Platelets ≥25 × 109/L (transfusion support is permitted if completed at least 7 days before the planned start of dosing) c) Hemoglobin ≥8 g/dL (transfusion support is permitted if completed at least 7 days before the planned start of dosing)	X	ü	X	ü
Left ventricular ejection fraction ≥40% as determined by a multigated acquisition scan (MUGA) or echocardiogram	X	X	X	X
Resolved acute effects of any prior therapy to baseline severity or Common Terminology Criteria for AE Grade ≤1	X	X	X	X
Exclusion Criteria				
Active plasma cell leukemia	ü	ü	ü	ü
Amyloidosis	ü	ü	ü	ü
Previous treatment with an anti-BCMA bispecific antibody.	X	X	X	X
Previous administration with an investigational drug within 30 days (or as determined by the local requirement) or 5 half-lives preceding the first dose of study intervention used in this study (whichever is longer).	ü	ü	ü	ü
Smoldering MM	ü	ü	ü	ü
Stem cell transplant within 12 weeks before enrolment or active Graft Versus Host Disease (GVHD).	ü	ü	ü	ü
Any other active malignancy within 3 years before enrolment, except for adequately treated basal cell or squamous cell skin cancer, or carcinoma *in situ*.	ü	ü	ü	ü
POEMS syndrome	X	X	X	X
Impaired cardiovascular function or clinically significant cardiovascular diseases, defined based on the history of any of the following conditions within 6 months before enrolment: a) Acute myocardial infarction or acute coronary syndromes (eg, unstable angina, coronary artery bypass graft, coronary angioplasty or stenting, symptomatic pericardial effusion) b) Clinically significant cardiac arrhythmias (eg, uncontrolled atrial fibrillation or uncontrolled paroxysmal supraventricular tachycardia) c) Thromboembolic or cerebrovascular events (eg, transient ischemic attack, cerebrovascular accident, deep vein thrombosis, or pulmonary embolism) d) Prolonged Time from the beginning of the QRS Complex to the End of the T Wave (QT) syndrome (or triplicate average Corrected QT (Fridericia Method) (QTcF) >470 msec).	X	ü	X	ü
Active Hepatitis B Virus (HBV), hepatitis C virus Hepatitis C Virus (HCV), Severe Acute Respiratory Syndrome- Coronavirus 2 (SARS-CoV2), known Human Immunodeficiency Virus (HIV), or any active, uncontrolled bacterial, fungal, or viral infection. Active infections must be resolved at least 14 days before enrolment.	ü	ü	ü	ü

Table 1. Inclusion and Exclusion criteria for Study C1071003, Flatiron Health, and COTA

Patient criteria per Study C1071003	Implementation in Flatiron Health		Implementation in COTA	
	Critical	Expanded	Critical	Expanded
Other surgical (including major surgery within 14 days prior to enrolment), medical, or psychiatric conditions including recent (within the past year) or active suicidal ideation/behavior or laboratory abnormality that may increase the risk of study participation or, in the investigator's judgment, make the patient inappropriate for the study.	X	X	X	X
Ongoing Grade ≥2 peripheral sensory or motor neuropathy.	X	ü	X	X
History of any grade peripheral sensory or motor neuropathy with prior BCMA-directed therapy (Cohort B).	X	ü	X	ü
History of Guillain-Barre Syndrome (GBS) or GBS variants, or history of any Grade ≥3 peripheral motor neuropathy	X	ü	X	X
Investigator site staff or Pfizer employees directly involved in the conduct of the study, site staff otherwise supervised by the investigator, and their respective family members.	X	X	X	X
Known or suspected hypersensitivity to the study intervention or any of its excipients.	X	X	X	X
Live attenuated vaccine must not be administered within 4 weeks of the first dose of the study intervention.	X	X	X	X

Note:
* This criterion is not among the inclusion/exclusion criteria of Study C10710003 and is applied only to RW patients.

9.3. Variables

Exposure definition

In each comparison, patients will be classified into one of the two treatment groups according to the therapy received after TCR eligibility, as those treated with elranatamab and those treated with a SOC regimen (any standard treatment option available for RW TCR MM patients).

If feasible based on data availability, patients from elranatamab and SOC treatment groups will further be classified into those with and without exposure to BCMA therapy before the index date for the purposes of subgroup analyses.

Study outcomes

Consistency between outcome definitions across data sources is another key consideration to minimize bias. Comparative effectiveness endpoints will be evaluated based on availability within each RWD source. Definitions of the outcomes will be aligned, where possible, with Study C1071003 (Table 2). In Study C1071003, the capture of objective response (OR) is defined according to IMWG criteria.

Table 2. Definitions of comparative effectiveness outcomes in Study C1071003, Flatiron Health, and COTA

	Outcome	Study C1071003	Flatiron Health	COTA
1	Objective response rate (ORR)	The proportion of patients with an OR based on blinded independent central review per IMWG criteria. OR is defined as having a best overall response of confirmed Stringent Complete Response (sCR), CR, Very Good Partial Response (VGPR), and PR per IMWG criteria, from the date of the first dose until confirmed PD, death or start of new anticancer therapy, whichever occurs first.	The proportion of patients who achieved at least VGPR or PR Based on IMWG criteria*	The proportion of patients who achieved sCR, CR, VGPR, or PR Based on IMWG criteria*
2	TTR	For patients with an OR per IMWG criteria, the time from the first dose to first documentation of OR that is subsequently confirmed.	For subjects with an OR (based on PR or VGPR), time from the initiation of the first line after TCR to the first documented OR (first date at which at least VGPR or PR was documented).	For subjects with an OR, time from the initiation of the first line after TCR to the first documented OR.
3	Duration of response (DOR)	For patients with an OR per IMWG criteria, the time from the first documentation of OR that is subsequently confirmed, until confirmed PD per IMWG criteria, or death due to any cause, whichever occurs first.	Among patients who achieved an OR (based on PR or VGPR), the time from the first documentation of OR until progression or death due to any cause, whichever occurs first.	Among patients who achieved an OR, the time from the first documentation of OR until progression or death due to any cause, whichever occurs first.

Note:

*Details on the operational definitions of IMWG criteria used in Flatiron and COTA to define response, and a side-by-side comparison with the criteria used in Study C1071003 are provided in the relating study protocols as Document 001 and 002 in ANNEX 1. LIST OF STAND ALONE DOCUMENTS.

Other variables

Ensuring that baseline factors related to treatment assignment and outcomes are captured is of utmost importance for the validity of external control arms. The availability of these variables in selected RWD sources determines the feasibility of the external control arm.

Baseline covariates will be selected to compare patients from elranatamab and external control arms, and to perform further statistical adjustments to control for the baseline confounding (eg, via IPTW).

Baseline covariates to be captured will include those related to patient demographic characteristics, disease characteristics, comorbidity profile, laboratory measurements, and MM treatment patterns, including LOT. The exact variables and definitions will be dependent upon availability within each RWD source. Adaptations of variable definitions or the use of proxy variables may be necessary if the preferred variables are unavailable in each data source. Each variable will be taken on or before the index date; if before, the most recent measurement will be used. All variables listed below will be used to describe the study cohorts and empirically reviewed for data completeness. Since small sample sizes may limit the number of covariates that can be reasonably accounted for in the statistical adjustment (eg, IPTW), a systematic literature review was conducted to identify the variables most strongly and consistently correlated with outcomes in RWD studies conducted among RRMM patients. These variables would be prioritized for inclusion in the present study. Fifty-seven studies were extracted as part of our literature review which used either univariate (N=22) or multivariate (N=35) analyses to predict outcomes. Collectively, these studies indicated the variables with the strongest relationships with future outcomes in an RRMM population were age, sex, cytogenetic risk, number of prior lines of therapy, and ISS/R-ISS. Additionally, both hemoglobulin and albumin levels exhibited limited evidence as prognostic variables.

Priority covariates

- Age at index date
- Sex (male; female)
- ISS (International Staging System) stage (I; II; III) within 90 days before or on the index date, if feasible
- ECOG performance status within 90 days before or on the index date, if feasible
- Time since initial MM diagnosis (years)
- Number of pre-index treatment lines
- Quad-and Penta-refractory status at index date
- Cytogenic risk (high; standard) at index date, if feasible

Note that a high proportion of missingness is possible when using RWD and this may limit the feasibility to use the above variables (especially ISS and cytogenetic risk) as priority covariates in the statistical analysis.

Additional covariates

Additional covariates will be collected to characterize patients from two study arms and to assess for baseline confounding adjustment. Given the adequate sample size, some of these covariates may be included in the analyses:

- Race/ethnicity (White / Non-White)
- Body mass index (kg/m^2; identified from the most recent measure on or before the index date)
- Treatment setting (community-based; academic-based) at the index date
- Extramedullary disease identified within 12 months before or on the index date
- Presence of bone lesions within 12 months before or on the index date
- Duration of prior therapy (days)
- Stem cell transplantation (yes; no) within 12 months before or on the index date
- Radiation therapy (yes; no) within 12 months before or on the index date
- BCMA exposure during the baseline period
- History of any grade peripheral sensory or motor neuropathy for those with prior BCMA directed therapy
- Charlson Comorbidity Index within 12 months before or on the index date
- Impaired cardiovascular function within 12 months before or on the index date
- Basal cell or squamous cell skin cancer, or carcinoma *in situ* (yes; no) within 12 months on or before the index date
- Adequate hepatic function reported within 90 days before or on the index date
- Adequate renal function reported within 90 days before or on the index date
- Adequate BM function reported within 90 days before or on the index date
- Aspartate aminotransferase (U/L) within 90 days before or on the index date, if feasible
- Alanine aminotransferase (U/L) within 90 days before or on the index date, if feasible

- Creatinine clearance (mL/min) within 90 days before or on the index date, if feasible
- ANC ($\times 10^9$/L) within 90 days before or on the index date, if feasible
- Platelets ($\times 10^9$/L) within 90 days before or on the index date, if feasible
- Hemoglobin (g/L) within 90 days before or on the index date, if feasible
- Calcium in serum or plasma (mg/dL) reported within 90 days before or on the index date
- Lactate dehydrogenase (U/L) reported within 90 days before or on the index date
- Beta-2 microglobulin (mg/L) reported within 90 days before or on the index date
- Serum albumin (g/dL) reported within 90 days before or on the index date

9.4. Data sources

Data integration

Upon constructing the analytical file for the comparison between Study C1071003 and Flatiron Health and between Study C1071003 and COTA, a common data model with standardized data elements will be created and documented for transparency. The variable transformation will include the creation of the common variable type, format, and taxonomy. Consistent definitions will be applied to create derived variables for treatments, index date, outcomes, and comorbidities.

Combined study datasets

Combined study datasets will be created from the entirety of the Study C1071003 arm plus the selected external control arm based on data source and eligibility criteria.

A total of 4 combined study datasets will be created. Each will use the same patients from Study C1071003 along varying RW patients, depending upon the RWD source and inclusion/exclusion criteria applied to identify them, as indicated below.

For primary analyses:

- Study C1071003 arm plus the external control arm selected using critical eligibility criteria from Flatiron Health.
- Study C1071003 arm plus the external control arm selected using critical eligibility criteria from COTA.

For sensitivity analyses:

- Study C1071003 arm plus the external control arm selected using expanded eligibility criteria from Flatiron Health.

- Study C1071003 arm plus the external control arm selected using expanded eligibility criteria from COTA.

9.5. Study size

The study sample will be identified from the analysis of secondary data that have already been collected. Accordingly, the sample size will be limited by the duration of the observation window. All patients who meet the inclusion/exclusion criteria defined in Section 9.2.3 will be included in the analyses. Preliminary analysis of the Flatiron and COTA data identified 341 and 448 TCR MM patients before applying inclusions and exclusion criteria, respectively. No formal sample size estimations have been performed for this observational study.

9.6. Data management

This study will use structured databases from Study C1071003, an open-label, multi-center, non-randomized Phase 2 study of elranatamab (PF-06863135) monotherapy (Pfizer, 2021), Flatiron Health, and COTA (the latter two databases collect information on MM patients treated across the US). Flatiron and COTA have been selected according to data availability and as fit-for-purpose for fulfilling study objectives (the US FDA, 2018 Framework). The dates for the study periods for each RW database have been selected to align as closely as possible to each other and to Study C1071003 which will include at least 9 months of follow-up.

9.7. Data analyses

The detailed methods associated with the statistical analyses of the data collected in this study will be documented in the statistical analysis plan (SAP), which will be dated, filed, and maintained by the sponsor. Table 3 provides the summary of statistical analyses for the combined study cohorts.

Table 3. Summary of statistical analyses for combined study cohorts.

	Analyses based on the critical eligibility criteria RW sample	**Analyses based on the expanded eligibility criteria RW sample**
Primary analyses	• Descriptive statistics of elranatamab and SOC-treated patients Primary objective • Naïve (unadjusted) comparison of ORR in elranatamab versus SOC-treated patients • IPTW comparison of ORR in elranatamab versus SOC-treated patients Secondary objectives	

Table 3. Summary of statistical analyses for combined study cohorts.

	Analyses based on the critical eligibility criteria RW sample	Analyses based on the expanded eligibility criteria RW sample
	• Naïve (unadjusted) comparison of TTR, and DOR in elranatamab versus SOC-treated patients • IPTW comparison of TTR, and DOR in elranatamab versus SOC-treated patients	
Sensitivity analyses	• Doubly robust comparisons of ORR in elranatamab versus SOC-treated patients • Quantitative bias assessments (using raw, IPTW, doubly robust estimates) when comparing ORR, TTR, and DOR in elranatamab versus SOC-treated patients	• Descriptive statistics of elranatamab and SOC-treated patients • Naïve (unadjusted) comparison of ORR, TTR, and DOR in elranatamab versus SOC-treated patients • IPTW comparison of ORR, TTR, and DOR in elranatamab versus SOC-treated patients • Doubly robust comparisons of ORR in elranatamab versus SOC-treated patients
Subgroup analyses (if feasible)	• Descriptive statistics of elranatamab and SOC-treated patients stratified by prior BCMA exposure • Naïve (unadjusted), IPTW, and doubly robust comparisons of ORR in elranatamab versus SOC-treated patients with and without prior BCMA exposure • Naïve (unadjusted) and IPTW comparisons of TTR, and DOR in elranatamab versus SOC-treated patients with and without prior BCMA exposure	

In this study, measures taken to reduce potential bias include a pre-specified and detailed SAP. Additionally, the propensity score (PS) balancing steps will be conducted independently and prior to availability/knowledge of the outcomes (i.e., balancing will be conducted using only baseline characteristic data from an earlier data cut than the data cut which will contain the efficacy outcomes from Study C1071003).

9.7.1. Descriptive statistics

Tabular summaries of baseline patient demographics and clinical characteristics by treatment arm will be presented. Summary statistics will include frequencies and percentages (categorical variables) and mean/median, minimum/maximum with standard deviation/interquartile range (continuous variables).

9.7.2. Primary analyses

For comparative effectiveness, ORR, TTR, and DOR will be assessed using IPTW comparisons estimating the average treatment effect (ATE). The naïve (unadjusted) estimates for ORR, TTR and DOR will be reported.

Propensity scores and inverse probability of treatment weighting

In analyses of RWD, an important consideration in the identification of potentially causal effects is the control for confounding. IPTW is a well-established method for causal inference in nonrandomized studies. Estimating the PS is a form of dimensionality reduction, in which several individual characteristics relevant to treatment assignment and the outcome, or those related only to outcome development, are used to estimate the conditional probability that the patient is assigned to a given treatment. IPTW is a quasi-experimental approach used to create a pseudo-population in which the covariates are independent of the treatment assignment, thereby permitting an unbiased estimate of the ATE.

The PS will be estimated using logistic regression models, where the dependent variable is a binary indicator of the treatment arm (elranatamab versus SOC). Covariates in the logistic regression will include the priority covariates described in Section 9.3 and, optional, some additional covariates. Diagnostics of the estimated PS will be applied by examining the distribution of the PS in each treatment group, the area of PS overlap, or any outliers. The estimated PS will be used to generate stabilized IPTWs. A standardized mean difference (SMD) of ≥10% will be used to identify potential residual confounding and as an indicator of covariate imbalance between the treatment arms requiring further investigation(Austin, 2009). Grossly higher SMDs (over 25%) may lead to refining the logistic model for deriving the weights.

Analysis models for endpoints

The following analyses will be performed for endpoints.

Objective response

The dichotomous outcome ORR will be compared between treatment arms using a logistic regression model. Unweighted and IPTW estimates will be reported together with corresponding 95% CIs. Standard errors and CIs for the IPTW estimators will be obtained via a robust sandwich-type estimator of variance (Lin and Wei, 1989).

Time to event endpoints (TTR and DOR)

TTR and DOR will be compared between treatment arms using hazard ratios estimated from Cox proportional hazards models. The proportional hazards assumption will be checked, and in case of deviations, a restricted mean survival time model will be applied instead of the Cox proportional hazards regression model (Zhao and Tsiatis, 1999) Standard errors and CI for the IPTW estimator will be obtained via a robust sandwich-type estimator of variance (Lin and Wei, 1989).

Kaplan-Meier figures will be used for DOR to visually examine survival probabilities over the follow-up period. Results from both weighted and unweighted analyses will be provided.

9.7.3. Sensitivity analyses

To evaluate the robustness of results from the primary analyses, sensitivity analyses will be conducted.

Doubly robust comparison

The dichotomous outcome ORR will be compared between treatment arms using doubly robust methods. Analyses employing a doubly robust estimator require fitting 2 models: a model for treatment or exposure status, and a second model for the outcome of interest. As with other causal inference methods, valid and unbiased estimates require assumptions of no unmeasured confounding (exchangeability), positivity (the experimental treatment assumption), no interference, and consistency (Funk *et al.*, 2011). Assuming these assumptions are upheld, if at least 1 of these 2 models is correctly specified, resulting in a doubly robust estimate that may remain consistent and unbiased (Bang and Robins, 2005).

The AIPW approach will be used to estimate a PS model for weighting, and subsequently augment the inverse probability weights using predicted values from the outcome model. The general procedure for estimation of PS, and IPTW is provided in Section 9.7.2.

Diagnostics will be applied to assess whether the doubly robust model appears to be

well-specified (Austin and Stuart, 2015).

Alternative eligibility criteria

To assess the potential influence of population selection on the observed effect estimates, additional analysis will be performed identical to the primary analysis but using the combined study cohorts with the expanded eligibility criteria.

Subgroup analysis by prior BCMA exposure

A subgroup analysis by prior BCMA therapy (yes/no) was intended to be performed if at least

20 patients with prior BCMA exposure were available in at least one of the external comparator arms. However, based on preliminary assessment of the Flatiron and COTA data, the minimum sample size of BCMA-exposed patients is not met in either comparator arm; therefore, only a subgroup analysis of non-BCMA-exposed patients will be performed. This analysis will only be performed for the primary analysis and not for sensitivity analyses unless stated otherwise.

Nullification analysis

Nullification analysis will be applied to assess the potential influence of unmeasured confounding on the observed associations. A shortlist of suspected unmeasured confounders will be identified; these are variables that were observed in Study C1071003 but not available (or only with substantial missingness) in the RWD sources. The E-value will be computed and reflects the minimum strength of association the confounder would need to have with the exposure and outcome, conditional on the measured covariates, to fully explain away the observed treatment effect (van der Weele and Ding, 2017). This sensitivity analysis will be applied for the primary analysis.

E-values are relatively intuitive, and offer a comprehensible, easily communicated summary of the findings. The lowest possible E-value is 1, and as the value increases, so does the implied degree of bias required to explain the results (van der Weele and Ding, 2017).

E-values will be computed for measures of effect and adjusted for the same set of relevant covariates as in the primary analyses.

9.8. Quality control

This is a retrospective study, so issues of quality control at study sites, e.g., data queries, do not apply. Analyses are programmed according to the specifications in the protocol, and if applicable, the statistical analysis plan and documented in a programming plan. Final deliverables are reviewed and verified by a second, independent programmer who may also perform double programming. All quality checks are documented in the programming plan.

9.9. Limitations of the research methods

Unlike clinical trial settings with specific definitions of study outcomes and scheduled assessments described in the protocol, the assessment of TTR in RW clinical practice settings may not be made consistently across patients and physicians. Specifically, in real-world observational studies, especially those performed retrospectively, it is not possible to implement consistent monitoring and application of homogenous evaluation criteria (eg, IMWG) that are inherent to clinical trial design. As such, depending on how often clinical assessments are made, outcomes may be subject to surveillance bias, which occurs when the outcome is more likely to be captured among patients who are followed more closely (Haut and Pronovost, 2011). This bias could favor the RW control groups, considering that in real life, the patient is monitored less frequently for progression.

The use of 2 RWD sources has further limitations. First, missing data and the accuracy of recorded data on disease characteristics, lab results, and comorbidities in RWD may introduce an information bias and residual confounding. Second, patients without sufficient information in their EHRs were excluded from the datasets by the data providers to avoid information bias. However, this may introduce a selection bias. Third, the variability in populations and differences in a and/or variables/covariates can result in changes in evaluations.

To address and/or reduce the impact of potential bias and improve exchangeability between the trial and the external control arms (Pocock, 1976; Ghadessi *et al.*, 2020; Hatswell *et al.*, 2020) measures will be taken to align comparable populations, advanced adjustment methods will be employed, and a series of sensitivity analyses will be conducted to evaluate the impact of key assumptions and selection criteria. However, this analysis is still subject to unmeasured confounding. Quantitative bias analysis will attempt to estimate the potential bias from unmeasured confounders, but these estimates are likely to be conservative and may nullify a true treatment effect.

9.10. Other aspects

Not applicable.

10. PROTECTION OF HUMAN SUBJECTS

10.1. Patient information

This study involves data that exist in anonymized structured format and contain no patient personal information.

10.2. Patient consent

As this study involves anonymized structured data, which according to applicable legal requirements do not contain data subject to privacy laws, obtaining informed consent from patients by Pfizer is not required.

10.3. Patient withdrawal

Not applicable

10.4. Institutional review board (IRB)/ Independent ethics committee (IEC)

Approval from an institutional review board/independent ethics committee is not required for this study as only de-identified secondary data sources and anonymized medical record data from EHR will be used. Therefore, this study is considered exempt from the requirements for "human subjects research".

10.5. Ethical conduct of the study

The study will be conducted in accordance with legal and regulatory requirements, as well as with scientific purpose, value, and rigor and follow generally accepted research practices described in Guidelines of Good Pharmacoepidemiology Practices (GPP) issued by the International Society for Pharmacoepidemiology (ISPE), Good Practices for Outcomes Research issued by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR), International Ethical Guidelines for Epidemiological issued by the Council for International Organizations of Medical Sciences (CIOMS).

11. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

This study is retrospective in nature, it involves data that exist as structured data by the time of study start or a combination of existing structured data and unstructured data, which will be converted to structured form during the implementation of the protocol solely by a computer using automated/algorithmic methods, such as natural language processing.

In these data sources, individual patient data are not retrieved or validated, and it is not possible to link (ie, identify a potential association between) a particular product and medical event for any individual. Thus, the minimum criteria for reporting an AE (ie, identifiable patient, identifiable reporter, a suspect product, and event) cannot be met.

12. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

For all publications relating to the study, Pfizer will comply with recognized ethical standards concerning publications and authorship, including Section II - "Ethical

Considerations in the Conduct and Reporting of Research" of the Uniform Requirements for Manuscripts Submitted to Biomedical Journals, http://www.icmje.org/index.html#authorship, established by the International Committee of Medical Journal Editors.

In the event of any prohibition or restriction imposed (e.g., clinical hold) by an applicable competent authority in any area of the world, or if the investigator party responsible for collecting data from the participant is aware of any new information which might influence the evaluation of the benefits and risks of a Pfizer product, Pfizer should be informed immediately.

13. REFERENCES

Austin, P. (2009) 'Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples', *Statistics in medicine*, 28, pp. 3083–3107. Available at: https://doi.org/10.1002/sim.3697.

Austin, P.C. and Stuart, E.A. (2015) 'Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies', *Statistics in Medicine*, 34(28), pp. 3661–3679. Available at: https://doi.org/10.1002/sim.6607.

Bang, H. and Robins, J.M. (2005) 'Doubly Robust Estimation in Missing Data and Causal Inference Models', *Biometrics*, 61(4), pp. 962–973. Available at: https://doi.org/10.1111/j.1541-0420.2005.00377.x.

Carroll, NM. (2016) 'Application of Propensity Score Models in Observational Studies using VDW data', pp. 2522–2018.

Chesnaye, N.C. *et al.* (2022) 'An introduction to inverse probability of treatment weighting in observational research', *Clinical Kidney Journal*, 15(1), pp. 14–20. Available at: https://doi.org/10.1093/CKJ/SFAB158.

COTA. Multiple myeloma real-world evidence. Data manual. (2022).

Committee for Medicinal Products for Human Use (CHMP) guideline on clinical trials in small populations draft agreed by efficacy working party/ad hoc group on clinical trials in small populations adoption by CHMP for release for consultation. (2002). Available at: http://www.emea.eu.int (Accessed: 24 April 2022).

Flatiron Health, Inc. (2021) *Analytic guide for Flatiron Health Data. Triple-class refractory multiple myeloma. Version 1.*

'Framework for FDA's Real-World Evidence Program' (2018). Available at: www.fda.gov (Accessed: 24 May 2022).

Funk, M.J. *et al.* (2011) 'Doubly Robust Estimation of Causal Effects', *American Journal of Epidemiology*, 173(7), pp. 761–767. Available at: https://doi.org/10.1093/aje/kwq439.

Gandhi, U.H. *et al.* (2019) 'Outcomes of patients with multiple myeloma refractory to CD38-targeted monoclonal antibody therapy', *Leukemia*, 33(9), pp. 2266–2275. Available at: https://doi.org/10.1038/S41375-019-0435-7.

Ghadessi, M. *et al.* (2020) 'A roadmap to using historical controls in clinical trials - By Drug Information Association Adaptive Design Scientific Working Group (DIA-ADSWG)', *Orphanet Journal of Rare Diseases*, 15(1). Available at: https://doi.org/10.1186/S13023-020-1332-X.

Hatswell, A. *et al.* (2020) 'Summarising salient information on historical controls: A structured assessment of validity and comparability across studies', *Clinical Trials*, 17(6), pp. 607–616. Available at: https://doi.org/10.1177/1740774520944855.
Haut, E.R. and Pronovost, P.J. (2011) 'Surveillance bias in outcomes reporting', *JAMA - Journal of the American Medical Association*, 305(23), pp. 2462–2463. Available at: https://doi.org/10.1001/JAMA.2011.822.
ICH E10 Choice of control group in clinical trials | European Medicines Agency (2001). Available at: https://www.ema.europa.eu/en/ich-e10-choice-control-group-clinical-trials (Accessed: 24 May 2022).
Jagannath, S. *et al.* (2020) 'KarMMa-RW: A study of real-world treatment patterns in heavily pretreated patients with relapsed and refractory multiple myeloma (RRMM) and comparison of outcomes to KarMMa.', *Journal of Clinical Oncology*, 38(15_suppl), pp. 8525–8525. Available at: https://doi.org/10.1200/JCO.2020.38.15_SUPPL.8525.
Jahanshahi, M. *et al.* (2021) 'The Use of External Controls in FDA Regulatory Decision Making', *Therapeutic innovation & regulatory science*, 55(5), pp. 1019–1035. Available at: https://doi.org/10.1007/S43441-021-00302-Y.
Lin, D.Y. and Wei, L.J. (1989) 'The Robust Inference for the Cox Proportional Hazards Model', *Journal of the American Statistical Association*, 84(408), p. 1074. Available at: https://doi.org/10.2307/2290085.
Mikhael, J. (2020) 'Treatment Options for Triple-class Refractory Multiple Myeloma', *Clinical lymphoma, myeloma & leukemia*, 20(1), pp. 1–7. Available at: https://doi.org/10.1016/J.CLML.2019.09.621.
Nooka, A.K. *et al.* (2015) 'Treatment options for relapsed and refractory multiple myeloma', *Blood*, 125(20), pp. 3085–3099. Available at: https://doi.org/10.1182/BLOOD-2014-11-568923.
Padala, S.A. *et al.* (2021) 'Epidemiology, Staging, and Management of Multiple Myeloma', *Medical sciences (Basel, Switzerland)*, 9(1), p. 3. Available at: https://doi.org/10.3390/MEDSCI9010003.
Pfizer (2021) MagnetisMM-3, an open-label, multicenter, non-randomized phase 2 study of elranatamab (pf-06863135) monotherapy in participants with MM who are refractory to at least one proteasome inhibitor (PI), one immunomodulatory drug, and one anti-CD38 antibody. Final Protocol Amendment 7, 11 November 2021.
Pfizer Initiates Pivotal Phase 2 MagnetisMM-3 Trial of BCMA-CD3 Bispecific Antibody Elranatamab (PF-06863135) in Multiple Myeloma | Pfizer (2022). Available at: https://www.pfizer.com/news/press-release/press-release-detail/pfizer-initiates-pivotal-phase-2-magnetismm-3-trial-bcma (Accessed: 24 May 2022).
Pfizer Presents First Data from Planned Interim Analysis of Pivotal Phase 2 MagnetisMM-3 Trial of BCMA-CD3 Bispecific Antibody Elranatamab Under Investigation for Relapsed/Refractory Multiple Myeloma | Pfizer (2022). Available at: https://www.pfizer.com/news/press-release/press-release-detail/pfizer-presents-first-data-planned-interim-analysis-pivotal (Accessed: 26 June 2022).
Pocock, S.J. (1976) 'The combination of randomized and historical controls in clinical trials', *Journal of chronic diseases*, 29(3), pp. 175–188. Available at: https://doi.org/10.1016/0021-9681(76)90044-8.

Rajkumar, S.V. *et al.* (2014) 'International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma', *The Lancet. Oncology*, 15(12), pp. e538–e548. Available at: https://doi.org/10.1016/S1470-2045(14)70442-5.
Rare Diseases: Natural History Studies for Drug Development | FDA (2019). Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/rare-diseases-natural-history-studies-drug-development (Accessed: 24 May 2022).
Sebag, M. *et al.* (2022) 'Elranatamab (PF-06863135), a B-Cell Maturation Antigen (BCMA) Targeted CD3-Engaging Bispecific Molecule, for Patients with Relapsed or Refractory Multiple Myeloma: Results from Magnetismm-1'. Available at: https://ash.confex.com/ash/2021/webprogram/Paper150519.html (Accessed: 24 April 2022).
Wallington-Beddoe, C.T. and Pitson, S.M. (2017) 'Novel therapies for multiple myeloma', *Aging*, 9(8), pp. 1857–1858. Available at: https://doi.org/10.18632/AGING.101284.
van der Weele, T.J. and Ding, P. (2017) 'Sensitivity Analysis in Observational Research: Introducing the E-Value', *Annals of internal medicine*, 167(4), pp. 268–274. Available at: https://doi.org/10.7326/M16-2607.
Yong, K. *et al.* (2016) 'Multiple myeloma: patient outcomes in real-world practice', *British journal of haematology*, 175(2), pp. 252–264. Available at: https://doi.org/10.1111/BJH.14213.
Zhao, H. and Tsiatis, A.A. (1999) 'Efficient estimation of the distribution of quality-adjusted survival time', *Biometrics*, 55(4), pp. 1101–1107. Available at: https://doi.org/10.1111/J.0006-341X.1999.01101.X.

14. LIST OF TABLES

15. LIST OF FIGURES

16. APPENDIX

16.1. Appendix. Inclusion and exclusion criteria in Study C1071003, Flatiron Health and COTA.

Patient inclusion criteria per C1071003 Study	**Flatiron Spotlight RW alternative approach to be used in the external control arm**	*Implementation*	**COTA RW alternative approach to be used in the external control arm**	*Implementation*
Age and Sex:				
Male or female participants age ≥18 years.	Male or female participants age ≥18 years.	Critical and expanded criteria	Male or female participants age ≥18 years.	Critical and expanded criteria
A female participant is eligible to participate if she is not pregnant or breastfeeding.	Not available.	N/A	Not available.	N/A
Type of Participant and Disease Characteristics:				
Participants who are willing and able to comply with all scheduled visits, treatment plan, laboratory tests, lifestyle considerations, and other study procedures.	Not applicable	N/A	Not applicable	N/A
Prior diagnosis of MM as defined according to IMWG criteria (Rajkumar et al, 2014).	Perfect or close match Patients were extracted by Flatiron based on the presence of a diagnosis code for MM (ICD-9 203.0x or ICD-10 C90.0x) any time prior to the index date. Active MM diagnosis was confirmed by Flatiron via abstraction of unstructured data.	Critical and expanded criteria	Perfect or close match MM diagnosis confirmed through pathology reports or through clinical diagnosis and supporting documentation such as lab tests.	Critical and expanded criteria

Patient inclusion criteria per C1071003 Study	**Flatiron Spotlight RW alternative approach to be used in the external control arm**	*Implementation*	**COTA RW alternative approach to be used in the external control arm**	*Implementation*
Measurable disease based on IMWG criteria as defined by at least 1 of the following: a. Serum M-protein ≥0.5 g/dL by SPEP b. Urinary M-protein excretion ≥200 mg/24 hours by UPEP c. Serum immunoglobulin FLC ≥10 mg/dL (≥100 mg/L) AND abnormal serum immunoglobulin kappa to lambda FLC ratio (<0.26 or >1.65)	Perfect or close match. Measurable disease, defined within 90 days prior to or on the index date.	Critical and expanded criteria	Perfect or close match. Measurable disease, defined within 90 days before or on the index date.	Critical and expanded criteria
Refractory to at least one IMiD.	Perfect or close match. Refractory to at least one IMiD defined as disease progression, based on IMWG criteria or HCP assessment, while a patient is on therapy or within 60 days of the last therapy dose in any line, regardless of response.	Critical and expanded criteria	Perfect or close match. Refractory to at least one IMiD defined as disease progression, based on IMWG criteria or HCP assessment, while a patient is on therapy or within 60 days of the last therapy dose in any line, regardless of response.	Critical and expanded criteria
Refractory to at least one PI.	Perfect or close match. Refractory to at least one PI defined as disease progression, based on IMWG criteria or HCP assessment, while a patient is on therapy or within 60 days of the last therapy dose in any line, regardless of response.	Critical and expanded criteria	Perfect or close match. Refractory to at least one PI defined as disease progression, based on IMWG criteria or HCP assessment, while a patient is on therapy or within 60 days of the last therapy dose in any line, regardless of response.	Critical and expanded criteria

Patient inclusion criteria per C1071003 Study	**Flatiron Spotlight RW alternative approach to be used in the external control arm**	*Implementation*	**COTA RW alternative approach to be used in the external control arm**	*Implementation*
Refractory to at least one anti-CD38 antibody.	Perfect or close match. Refractory to at least one anti-CD38 defined as disease progression, based on IMWG criteria or HCP assessment, while a patient is on therapy or within 60 days of the last therapy dose in any line, regardless of response.	Critical and expanded criteria	Perfect or close match. Refractory to at least one anti-CD38 defined as disease progression, based on IMWG criteria or HCP assessment, while a patient is on therapy or within 60 days of the last therapy dose in any line, regardless of response.	Critical and expanded criteria
Relapsed/refractory to last anti-MM regimen. Note: Refractory is defined as having progressive disease while on therapy or within 60 days of the last dose in any line, regardless of response.	Perfect or close match By design as the index date is the first line after the TCR eligibility date, all patients will be refractory to the last anti-myeloma regimen.	Critical and expanded criteria	Perfect or close match By design as the index date is the first line after the TCR eligibility date, all patients will be refractory to the last anti-myeloma regimen.	Critical and expanded criteria
Cohort A: Has not received prior BCMA-directed therapy.	Information is collected as baseline covariates.	Possible inclusion criteria for subgroup analysis	Information is collected as baseline covariates.	Possible inclusion criteria for subgroup analysis
Cohort B: Has received prior BCMA-directed ADC or BCMA-directed CAR T-cell therapy, either approved or investigational.	Information is collected as baseline covariates.	Possible inclusion criteria for subgroup analysis	Information is collected as baseline covariates.	Possible inclusion criteria for subgroup analysis

Patient inclusion criteria per C1071003 Study	**Flatiron Spotlight RW alternative approach to be used in the external control arm**	*Implementation*	**COTA RW alternative approach to be used in the external control arm**	*Implementation*
ECOG performance status ≤2.	Perfect or close match. ECOG performance status ≤2 within 90 days prior to or on the index date	Critical and expanded criteria	Perfect or close match. ECOG performance status ≤2 within 90 days prior to or on the index date KPS score to be converted to ECOG if patient's performance status test closest to the index date is KPS	Critical and expanded criteria
LVEF ≥40% as determined by a MUGA scan or ECHO.	Not available.	N/A	Not available.	N/A
Adequate hepatic function characterized by the following: a. Total bilirubin ≤2 x ULN (≤3 x ULN if documented Gilbert's syndrome); b. AST ≤2.5 x ULN; and c. ALT ≤2.5 x ULN	Perfect or close match Adequate hepatic function within 30 days before study index date.	Baseline covariate, expanded criteria	Perfect or close match Adequate hepatic function within 30 days before study index date.	Baseline covariate, expanded criteria
Adequate renal function defined by an estimated creatinine clearance ≥30 mL/min (according to the Cockcroft Gault formula, by 24-hours urine collection for creatinine clearance, or according to local institutional standard method).	Perfect or close match Adequate renal function within 30 days before study index date.	Baseline covariate, expanded criteria	Perfect or close match Adequate renal function within 30 days before study index date.	Baseline covariate, expanded criteria

Patient inclusion criteria per C1071003 Study	**Flatiron Spotlight RW alternative approach to be used in the external control arm**	*Implementation*	**COTA RW alternative approach to be used in the external control arm**	*Implementation*
Adequate BM function characterized by the following: a. ANC ≥1.0 × 109 /L (use of granulocyte-colony stimulating factors is permitted if completed at least 7 days prior to planned start of dosing); b. Platelets ≥25 × 109 /L (transfusion support is permitted if completed at least 7 days prior to planned start of dosing); and c. Hemoglobin ≥8 g/dL (transfusion support is permitted if completed at least 7 days prior to planned start of dosing).	Perfect or close match Adequate BM function within 90 days before study index date.	Baseline covariate, expanded criteria	Perfect or close match Adequate BM function within 90 days before study index date.	Baseline covariate, expanded criteria
Resolved acute effects of any prior therapy to baseline severity or CTCAE Grade ≤1.	Not available.	N/A	Not available.	N/A
Informed Consent:				
Capable of giving signed informed consent, which includes compliance with the requirements and restrictions listed in the ICD and in this protocol.	Not applicable	N/A	N/A	N/A

Patient exclusion criteria per C1071003 Study	**Flatiron Spotlight RW alternative approach to be used in the external control arm**	*Implementation*	**COTA RW alternative approach to be used in the external control arm**	*Implementation*
Medical Conditions:				
Smoldering MM.	Not Available Not identifiable directly in Flatiron Spotlight data however patients are only included in the data if they have evidence of active MM. Since smoldering MM (SMM) can be a precursor to active MM, patients may have had SMM before progressing to active MM.	Not implemented directly	Perfect or close match Patient will be excluded if SMM is the type of MM listed for the patient closet to (but not after) the index date	Critical and expanded criteria
Active plasma cell leukemia.	Perfect or close match Diagnosis of plasma cell leukemia is identified based on the presence of ≥1 diagnosis code within the 365 days prior to or on the index date, or based on identification of plasma cell leukemia in enhanced data abstracted from unstructured sources.	Critical and expanded criteria	Perfect or close match Active plasma cell leukemia within the 365 days before or on the index date can be identified in the secondary disease diagnosis dataset. Patients are censored by COTA and no longer followed in the data after a determination of active plasma cell leukemia.	Critical and expanded criteria
Amyloidosis.	Perfect or close match Diagnosis of amyloidosis is identified based on the presence of ≥1 diagnosis code within the 365 days prior to or on the index date, or based on identification of amyloidosis in enhanced data abstracted from unstructured sources.	Critical and expanded criteria	Perfect or close match Amyloidosis within the 365 days before or on the index date can be identified in the secondary disease diagnosis dataset. Patients are censored by COTA and no longer followed in the data after a diagnosis of amyloidosis.	Critical and expanded criteria

Patient exclusion criteria per C1071003 Study	**Flatiron Spotlight RW alternative approach to be used in the external control arm**	*Implementation*	**COTA RW alternative approach to be used in the external control arm**	*Implementation*
POEMS syndrome	Not available	N/A	Not available	N/A
Stem cell transplant within 12 weeks prior to enrollment or active GVHD.	Perfect or close match Diagnosis of GVHD is identified based on the presence of ≥1 diagnosis code within the 12 weeks prior to or on the index date, but there are no specific codes for "active". The indicator of stem cell transplant and the date the patient received the stem cell transplant is provided.	Critical and expanded criteria	Partial match. Stem cell transplants within 12 weeks prior to index date.	Critical and expanded criteria
Impaired cardiovascular function or clinically significant cardiovascular diseases, defined as any of the following within 6 months prior to enrollment: a. Acute myocardial infarction or acute coronary syndromes (eg, unstable angina, coronary artery bypass graft, coronary angioplasty or stenting, symptomatic pericardial effusion); b. Clinically significant cardiac arrhythmias (eg, uncontrolled atrial fibrillation or uncontrolled paroxysmal supraventricular tachycardia);	Partial match Identifiable based on diagnosis codes (except for QT syndrome); enhanced data abstracted from unstructured sources can identify the following CV conditions: angina, atrial arrhythmia, cerebrovascular accident, congestive heart failure, DVT, MI, another arrhythmia, PE, TIA, ventricular arrhythmia. However, for arrhythmias there is no information as to their clinically significance (eg, if the arrhythmia is uncontrolled).	Baseline covariate, expanded criteria	Partial match Identifiable based on comorbid conditions that are included in the CCI. However, comorbid conditions included in the CCI are only captured once per patient. History of cardiac arrhythmia, QT syndrome, and other forms of acute coronary syndrome are not identifiable.	Baseline covariate, expanded criteria

Patient exclusion criteria per C1071003 Study	**Flatiron Spotlight RW alternative approach to be used in the external control arm**	*Implementation*	**COTA RW alternative approach to be used in the external control arm**	*Implementation*
c. Thromboembolic or cerebrovascular events (eg, transient ischemic attack, cerebrovascular accident, deep vein thrombosis [unless associated with a central venous access complication] or pulmonary embolism); d. Prolonged QT syndrome (or triplicate average QTcF >470 msec at screening).				
Ongoing Grade ≥2 peripheral sensory or motor neuropathy.	Partial match. Ongoing Grade ≥2 peripheral sensory or motor neuropathy cannot be identified in the data as there is no information about Grade. However, treatments used to manage neuropathic pain will be used as a proxy to identify presence of more severe neuropathy cases to describe patients	Baseline covariate, expanded criteria	Not available	N/A
History of any grade peripheral sensory or motor neuropathy with prior BCMA-directed therapy (Cohort B).	Perfect or close match Sensory and motor neuropathy may be identified based on the presence of ≥1 diagnosis code any time prior to or on the index date. Among patients with prior BCMA-directed therapy only.	Possible exclusion criteria for subgroup analysis	Perfect or close match May be identified based on AEs that resulted in BCMA treatment discontinuation	Possible exclusion criteria for subgroup analysis

Patient exclusion criteria per C1071003 Study	**Flatiron Spotlight RW alternative approach to be used in the external control arm**	*Implementation*	**COTA RW alternative approach to be used in the external control arm**	*Implementation*
History of GBS or GBS variants, or history of any Grade ≥3 peripheral motor polyneuropathy.	Partial match GBS could be identified based on diagnosis codes. There is no information about the grade in the data that allows us to identify Grade ≥3 peripheral motor polyneuropathy. Treatments used to manage neuropathic pain will be used as a proxy to identify more severe neuropathy cases.	Baseline covariate, expanded criteria	Not available	N/A
Active HBV, HCV, SARS-CoV2, HIV, or any active, uncontrolled bacterial, fungal, or viral infection. Active infections must be resolved at least 14 days prior to enrollment.	Partial match Identifiable based on lab tests as well as diagnosis codes. There are no specific codes for “active” or “uncontrolled”, but “active” infections will be closely approximated using clinically plausible covariate assessment periods. HBV or HIV any time before index. Other infections within 30 days before index.	Critical and expanded criteria	Partial match HBV, HCV, and HIV at any time before index (comorbidities including these infections are only captured once per patient). No information for SARS-CoV2. Information on infections is not available through lab data.	Critical and expanded criteria

Patient exclusion criteria per C1071003 Study	**Flatiron Spotlight RW alternative approach to be used in the external control arm**	*Implementation*	**COTA RW alternative approach to be used in the external control arm**	*Implementation*
Any other active malignancy within 3 years prior to enrollment, except for adequately treated basal cell or squamous cell skin cancer, or carcinoma in situ.	Perfect or close match. Will be identified based on the presence of ≥1 diagnosis code of any other malignancy, except for adequately treated basal cell or squamous cell skin cancer, or carcinoma in situ within the 3 years prior to or on the index date. "Adequately treated" is not identifiable therefore all basal or squamous cell skin cancer or carcinoma in situ will be considered adequately treated and not reason for exclusion.	Critical and expanded criteria	Perfect or close match. Any secondary malignancy developed up to 3 years before index treatment. Only secondary malignancies diagnosed between initial MM diagnosis and index date are identifiable. COTA will censor patients who develop a secondary cancer after initial MM diagnosis.	Critical and expanded criteria
Other surgical (including major surgery within 14 days prior to enrollment), medical or psychiatric conditions including recent (within the past year) or active suicidal ideation/behavior or laboratory abnormality that may increase the risk of study participation or, in the investigator's judgment, make the participant inappropriate for the study.	Not available	N/A	N/A	N/A
Prior/Concomitant Therapy:				
Previous treatment with an anti-BCMA bispecific antibody.	Partial match Only non-investigational agents are identifiable, therefore to ensure no exposure to anti-BCMA bispecific antibody treatment, all patients treated	(Baseline covariate - investigational agents), expanded criteria	Partial match Only non-investigational agents are identifiable, therefore to ensure no exposure to anti-BCMA bispecific antibody treatment, all patients treated	(Baseline covariate as an investigational agents), expanded criteria

Patient exclusion criteria per C1071003 Study	**Flatiron Spotlight RW alternative approach to be used in the external control arm**	*Implementation*	**COTA RW alternative approach to be used in the external control arm**	*Implementation*
	with investigational agents would need to be excluded		with investigational agents would need to be excluded	
Prior/Concurrent Clinical Study Experience:				
Previous administration with an investigational drug within 30 days (or as determined by the local requirement) or 5 half-lives preceding the first dose of study intervention used in this study (whichever is longer).	Perfect or close match. Previous administration with an investigational drug within 30 days before the index date or on index date.	Critical and expanded criteria	Perfect or close match. Previous administration with an investigational drug within 30 days before the index date or on index date.	Critical and expanded criteria

Other Exclusions:				
Investigator site staff or Pfizer employees directly involved in the conduct of the study, site staff otherwise supervised by the investigator, and their respective family members.	Not applicable	N/A	N/A	N/A
Known or suspected hypersensitivity to the study intervention or any of its excipients.	Not applicable	N/A	N/A	N/A
Live attenuated vaccine must not be administered within 4 weeks of the first dose of study intervention.	Not available.	N/A	N/A	N/A

17. ANNEX 1. LIST OF STAND ALONE DOCUMENTS

Number	Document Reference Number	Date	Title
001	001	11-Mar-2019	Protocol for Flatiron database ("Retrospective study of treatment patterns and clinical outcomes in patients with triple class refractory MM treated in a real-world setting")
002	002	17-May-2022	Protocol for COTA database ("Retrospective study of treatment patterns and clinical outcomes in patients with triple class refractory MM treated in a real-world setting")