Study information

Title	Comparative Effectiveness of Elranatamab (PF 06863135) in Clinical Study C1071003 Versus Standard of Care (SOC) in a Real-World (RW) External Control Arm of Patients with Triple-Class Refractory (TCR) Multiple Myeloma (MM) from TherapyMonitor MM Germany
Protocol number	C1071035
Protocol version identifier	3.0
Date	19 March 2024
Research question and objectives	***Study Objectives*** 1. To compare clinical effectiveness among TCR MM patients treated with elranatamab in Study C1071003 with a comparable RW cohort of TCR patients receiving SOC from the TherapyMonitor MM Germany dataset, including: a. Overall survival (OS) b. Progression-free survival (PFS) c. Time to next treatment (TTNT) d. Time to discontinuation (TTD) 2. To describe treatment patterns of the SOC external control arm

Author(s)	PPD Pfizer Inc. *New York, USA* PPD Pfizer Inc. *New York, USA* PPD Pfizer Pharma GmbH *Berlin, Germany* PPD Pfizer Pharma GmbH *Berlin, Germany* PPD Pfizer Pharma GmbH *Berlin, Germany* PPD Pfizer Pharma GmbH *Berlin, Germany* **Prof. Dr. med** PPD Bucher Evidence Consulting *Botmingen, Switzerland* PPD Cytel Inc. *Berlin, Germany* PPD Cytel Inc. *Wismar, Germany*

1. TABLE OF CONTENTS

1. LIST OF ABBREVIATIONS

Abbreviation	Definition
ADC	Antibody-drug conjugate
AE	Adverse event
ALT	Alanine aminotransferase
AMNOG	Arzneimittelmarkt-Neuordnungsgesetz (Pharmaceuticals Market Reorganisation Act)
ANC	Absolute neutrophil count
AST	Aspartate aminotransferase
ATE	Average treatment effect
ATT	Average treatment effect on the treated
BCMA	B-cell maturation antigen
BMI	Body mass index
CAR	Chimeric antigen receptor
CAR	Censoring at random
CATE	Conditional average treatment effect
CCI	Charlson Comorbidity Index
CDN	CancerDataNet
CI	Confidence interval
CTCAE	Common Terminology Criteria for Adverse Events
ECHO	Echocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic case report form
EHR	Electronic health record
EMA	European Medicines Agency
EMD	Extramedullary disease
FCS	Fully conditional specification
FDA	Food and Drug Administration
FISH	Fluorescence in situ hybridization
FLC	Free light chain
G-BA	Gemeinsamer Bundesausschuss
GBS	Guillain-Barré-Syndrome
GP	General practitioner
GVHD	Graft versus host disease
HBV	Hepatitis B virus
HCP	Healthcare professional
HCV	Hepatitis C virus
HIV	Human immunodeficiency virus
ICD	Informed consent document
IEC	Independent ethics committee
Ig	Immunoglobulin
IMiD	Immunomodulatory drug
IMWG	International Myeloma Working Group
IPTW	Inverse probability of treatment weighting
IQWiG	Institute for Quality and Efficiency in Health Care

IRB	Institutional review board
ISPOR	The Professional Society for Health Economics and Outcomes Research
ISS	International Staging System
Kg	Kilograms
KM	Kaplan Meier
LDH	Lactate dehydrogenase
LOT	Line of therapy
LVEF	Left ventricular ejection fraction
M	Meters
MAR	Missing at random
Max	Maximum
MICE	Multiple imputations by chained equations
Min	Minimum
MM	Multiple myeloma
MNAR	Missing not at random
MRD	Minimal residual disease
MUGA	Multigated acquisition
NICE	National Institute for Health and Care Excellence
NUH	Non-university hospital
OBP	Office-based hemato-oncologists
ORR	Objective response rate
OS	Overall survival
PFS	Progression-free survival
PI	Proteasome inhibitor
POEMS	syndrome characterized by polyneuropathy, organomegaly, endocrinopathy, monoclonal protein, and skin changes
PS	Propensity score
PSM	Propensity score matching
QBA	Quantitative bias analysis
QTcF	Corrected QT (Fridericia method)
RCT	Randomized controlled trial
RDI	Relative dose intensity
RW	Real-world
RWD	Real-world data
RWE	Real-world evidence
RRMM	Relapsed/refractory multiple myeloma
SARS-CoV2	Severe acute respiratory syndrome coronavirus 2
SCA	Synthetic control arm
SCT	Stem cell transplant
SD	Standard deviation
SLiM-CRAB	Criteria involving the following: 60% or more clonal plasma cells (S), light chains (Li), and MRI (M); also elevated calcium levels (C), renal failure (R), anemia (A), and bone lesions (B)
SMD	Standardized mean difference
SOC	Standard of care

SPEP	Serum protein electrophoresis
TCE	Triple-class exposed
TCR	Triple-class refractory
TNXO	TriNetX Oncology GmbH
TT-MM	TherapyMonitor Multiple Myeloma
TTD	Time-to-discontinuation
TTE	Time-to-event
TTNT	Time-to-next treatment
UH	University hospital
ULN	Upper limit of normal
UPEP	Urine protein electrophoresis
β2M	Beta-2-microglobulin

2. RESPONSIBLE PARTIES

Principal Investigator(s) of the Protocol

Name, degree(s)	Job Title	Affiliation	Address
Marco DiBonaventura, PhD	Senior Director, Value & Evidence Team Lead, Hematology & Biosimilars	Pfizer Inc	66 Hudson Boulevard East, New York, NY 10001, USA
Patrick Hlavacek, MPH	Director, Value & Evidence, Hematology	Pfizer Inc	66 Hudson Boulevard East, New York, NY 10001, USA
Stephanie Kauffmann, MSc	Manager, Health Technology Assessment	Pfizer Pharma GmbH	Linkstraße 10, 10785 Berlin, Germany
Niclas Kürschner, MSc	Senior Manager, Health Technology Assessment	Pfizer Pharma GmbH	Linkstraße 10, 10785 Berlin, Germany
Leonie Berger, MSc	Manager, Outcomes Research	Pfizer Pharma GmbH	Linkstraße 10, 10785 Berlin, Germany
Regina Grabow-Schlediger, Dr. med., MBA	Senior Medical Affairs Scientist	Pfizer Pharma GmbH	Linkstraße 10, 10785 Berlin, Germany
Heiner C. Bucher, Prof. Dr. med	Clinical Epidemiologist	Bucher Evidence Consulting	CH4103 Botmingen, Switzerland
Sabrina Müller, MSc	Senior Director & Research Principal, Real-World Evidence	Cytel Inc	Alter Holzhafen 19, 23966 Wismar; Germany
Evi Zhuleku, MSc	Senior Research Consultant, Real-World Evidence	Cytel Inc	Potsdamer Str. 58, 10785 Berlin, Germany

3. AMENDMENTS AND UPDATES

Version Identifier	Date	Amendments Type (substantial or administrative)	Protocol Section(s) Changed	Summary of Amendment(s)	Reason
1.0	01 September 2023	-			
2.0	27 November 2023	substantial	7.4.1	Patients Eligibility Criteria: Exclusion of patients with <3 prior lines of therapy	The approval for Elranatamab is expected to most likely include patients who received at least 3 prior lines of therapy. To meet the anticipated label population, patients with <3 prior lines of therapy will be excluded from both the intervention and the RW control arms.
2.0	27 November 2023	substantial	7.4.3.1	Confounder List: Exclusion of CCI [redacted]	In order to incorporate available serum albumin information, the study will make use of computed R-ISS (Revised International Staging System). CCI [redacted]

					CCI [redacted]
2.0	27 November 2023	substantial	7.5, 7.7	Multiple imputation will serve as the base case, whereas complete case (main TCR cohort) will serve as a sensitivity approach	Complete case analysis does not incorporate all available information and MI is an accepted method by the G-BA to handle missing values.
2.0	27 November 2023	substantial	7.7	IPTW-ATT effect will be determined as an additional sensitivity	The estimand of interest is the ATE, however, in ATT can be better used to standardize to a clinical trial population and is more lax with regards to propensity score overlap than ATE. Since the clinical trial population is heavily pretreated, more than expected in the RW population, some differences between the clinical trial population and

					the RW populatin are to be expected, which could lead to limited PS overlap. Therefore, ATT will be added as a sensitivity analysis.
2.0	27 November 2023	substantial	7.2, 7.5, 7.7	Exclusion of Cohort A+B analyses	Currently, Cohort B is not relevant for the German healthcare system.
2.0	27 November 2023	substantial	7.4, 7.5, 7.7	Exclusion of TCE cohort analyses	TCR was defined as an inclusion criterium for the clinical trial population. To find patients in the RW population that match the clinical trial population as close as possible, TCR will be applied to the RW cohort. Extending the RW cohort to TCE will most probably increase bias that is expected from differences in relevant counfounding variables.

2.0	27 November 2023	substantial	7.7	Subgroup analyses are conditional on the interpretability of the overall main results and may be omitted	If the populations will be too different to produce meaningful results, it will not make sense to investigate subgroups. Only if the overlap between populations is acceptable, leading to interpretability of results on main level, effects in subgroups will be analyzed.
2.0	27 November 2023	administrative	Annex 2	Footnote for factor age added in CCI derivation table	Clarification that age factor is to be omitted in adjustment approaches whereby CCI and age are included as covariated.
3.0	19 March 2024	substantial	4, 5, 6, 7, 7.1, 7.2.2.1.-7.2.2.2., 7.3, 7.4.1-7.4.5., 7.5-7.7, 7.7.5.3., 7.7.5.4., 7.7.5.5., 7.8-7.10, 8.4, 11, ANNEX 2	Flatiron database added as additional data source for the indirect comparison	In addition to the TherapyMonitor Multiple Myeloma German registry, the clinical trial data should be compared to the US-based data from Flatiron.
3.0	19 March 2024	substantial	5, 7.3	Information on extendend follow-up of	Information on clinical trial with longer

				clinical trial added	follow-up added to the clinical trial description and study duration.
3.0	19 March 2024	substantial	7.7	Fine stratification and Overlap Weights will be determined as additional sensitivity analyses	Additional sensitivity analyses added which are efficient and accurate even in the presence of extreme tails or limited overlap of the PS distributions.
3.0	19 March 2024	administrative	4	Date of Final study report changed to 31 July 2024.	Due to substantial changes in the protocol (additional data source and additional statistical analyses) the Final study report will be available later.
3.0	19 March 2024	administrative	7.9	Added information on Quality Control.	Added clarification on how the Quality Control is conducted.

4. MILESTONES

Milestone	Planned date
Final version of study protocol	01 September 2023
Initiation of data analysis	01 September 2023
Completion of data analysis	03 November 2023
Final study report	31 July 2024

5. RATIONALE AND BACKGROUND

Multiple myeloma (MM) is a rare hematologic malignancy characterized by an excess of monoclonal bone marrow plasma cells. With an estimated 3,500 and 2,800 annual cases for men and women, the incidence of MM in Germany corresponds to 5.4 and 3.5 cases per 100,000, respectively [1]. Historically, treatment options for patients with MM were limited to a combination of alkylating agents, such as melphalan, and steroids, such as prednisone [2] [3]. However, the development of new therapies has led to profound changes in the treatment landscape and has increased overall survival (OS) in patients with MM [4], especially with the approval of therapies designed to target specific pathogenic pathways, including proteasome inhibitors (PIs), immunomodulatory agents (IMiDs), and anti-CD38 monoclonal antibodies (anti-CD38). The most important drugs and their approved combinations in Germany have been summarized in Table 1 [5, 6]. Based on the S3 consensus guideline, which summarizes the current standards for diagnosis and treatment of MM in Germany, it is noted that an optimal therapeutic regimen cannot be outlined due to lack of comparative studies [5, 6].

Table 1: Overview of approved agents and combinations for adult patients with MM, according to drug product information (adapted from [6])

Agent	Approved Combination	Indication
Immunomodulatory Drugs (IMiDs)		
Thalomid® (thalidomide| thal, T)	Melphalan/prednisone (MPT)	Untreated MM ≥ 65 years; when HD chemotherapy is not an option
	Bortezomib/dexamethasone (VTD)	Untreated MM, induction therapy; patients eligible for HD with ASCT
	Daratumumab/bortezomib/dexamethasone (Dara-VTd)	Patients with NDMM; patients eligible for ASCT
Revlimid® (lenalidomide| len, R)	Monotherapy	Maintenance therapy; NDMM after ASCT
	Dexamethasone (Rd)	Untreated MM; Non-transplant patients after at least one previous treatment
	Bortezomib/dexamethasone (VRd)	Untreated MM; non-transplant patients
	Melphalan/prednisone (RMP)	Untreated MM; non-transplant patients
	Carfilzomib/dexamethasone (KRd)	After at least 1 previous treatment
	Ixazomib/dexamethasone (Ixa-Rd)	After at least 1 previous treatment
	Daratumumab/dexamethasone (Dara-Rd)	Patients with NDMM, not eligible for ASCT; after at least 1 previous treatment
	Elotuzumab/dexamethasone(Elo-Rd)	After at least 1 previous treatment
Pomalyst® (pomalidomide| pom, P),	Bortezomib/dexamethasone (PVd)	After at least 1 previous treatment, including lenalidomide
	Dexamethasone (Pd)	RRMM, after at least 2 previous treatments, including lenalidomide and bortezomib and progression during the previous treatment
	Daratumumab/dexamethasone (Dara-Pd)	Patients with MM, with 1 previous treatment with a PI and lenalidomide, who were refractory to lenalidomide; or patients who have already received at least 2 previous treatments involving lenalidomide and a PI, and have demonstrated disease progression during or after the previous treatment
	Isatuximab/dexamethasone (Isa-Pd)	Patients with RRMM, who have received at least 2 previous treatments, including lenalidomide and a PI, and have demonstrated disease progression during the previous treatment

	Elotuzumab/dexamethasone (Elo-Pd)	Patients with R/R MM, who have received at least 2 previous treatments, including lenalidomide and a PI, and have demonstrated disease progression during the previous treatment
Proteasome inhibitors (PIs)		
Velcade® (bortezomib| btz, bor, V)	Monotherapy	Progressive MM after at least 1 previous treatment and ASCT or not eligible for ASCT
	Pegylated liposomal doxorubicin	Progressive MM after at least 1 previous treatment and ASCT or not eligible for ASCT
	Dexamethasone (Vd)	Progressive MM after at least 1 previous treatment and ASCT or not eligible for ASCT
	Melphalan/prednisone (VMP)	Untreated MM; patients not eligible for HD and ASCT
	Dexamethasone (VD)	Untreated MM; induction therapy; patients eligible for HD with ASCT
	Thalidomide/dexamethasone(VTD)	Untreated MM; induction therapy; patients eligible for HD with ASCT
	Cyclophosphamide/dexamethasone (VCD)	Induction therapy; NDMM
	Daratumumab/melphalan/prednisone (Dara-VMP)	Patients with NDMM; not eligible for ASCT
	Daratumumab/thalidomide/dexamethasone (Dara-VTd)	Patients with NDMM; not eligible for ASCT
	Daratumumab/dexamethasone (Dara-Vd)	After at least 1 previous treatment
	Panobinostat/dexamethasone (PAN-Vd)	Patients with RRMM, who have received at least 2 previous treatments, including bortezomib and an IMiD
Kyprolis® (carfilzomib| cfz, car, K)	Daratumumab/dexametha-sone (Dara-Kd; KdD)	After at least 1 previous treatment
	Lenalidomide/dexamethasone (KRd)	After at least 1 previous treatment
	Dexamethasone (Kd)	After at least 1 previous treatment
	Isatuximab/dexamethasone (Isa-Kd)	After at least 1 previous treatment
Ninlaro® (ixazomib| I)	Lenalidomide/dexamethasone (Ixa-Rd)	After at least 1 previous treatment
Antibodies		
Darzalex® (daratumumab| dara)	Lenalidomide/dexamethasone (Dara-Rd)	Patients with NDMM, who are not eligible for ASCT; after at least 1 previous treatment
	Bortezomib/melphalan/dexamethasone (Dara-VMP)	Patients with NDMM; not eligible for ASCT
	Bortezomib/thalidomide/dexamethasone (Dara-VTd)	Patients with NDMM; not eligible for ASCT
	Bortezomib/dexamethasone (Dara-Vd)	After at least 1 previous treatment
	Monotherapy	Patients with RRMM, who have already been treated with a PI and an IMiD and demonstrated disease progression during the previous treatment
Darzalex Faspro® (daratumumab and hyaluronidase-fihj)	Pomalidomide/dexamethasone (Dara-Pd)	Patient with MM, who have already received 1 previous PI and lenalidomide and were refractory to lenalidomide, or who have already received at least 2 previous treatments containing lenalidomide and a PI and have demonstrated disease progression during or after the previous treatment
Sarclisa® (isatuximab-ifrc)	Pomalidomide/dexamethasone (Isa-Pd)	Patients with RRMM who have already received at least 2 previous treatments, including lenalidomide and a PI, and had disease progression during the previous treatment
	Carfilzomib/dexamethasone (Isa-Kd)	After at least 1 previous treatment
Empliciti® (elotuzumab, Elo)	Lenalidomide/dexamethasone(Elo-Rd)	After at least 1 previous treatment
	Pomalidomide/dexamethasone (Elo-Pd)	Patients with RRMM who have received at least 2 previous treatments, including lenalidomide and a

		PI, and had disease progression during the previous treatment
Histone deacetylase (HDAC) inhibitor		
Farydak® (panobinostat, PAN)	Bortezomib/dexamethasone (PAN-Vd)	Patients with RRMM who have received at least 2 previous treatments, including bortezomib and 1 IMiD
Antibody-drug conjugate (ADC)		
Blenrep® (belantamab mafodotin)	Monotherapy	Patients with MM, at least 4 previous treatments and disease refractory to at least 1 PI, 1 IMiD and 1 monoclonal anti-CD38 antibody, and demonstrated disease progression during the previous treatment
XPO1 (exportin 1) inhibitor		
Nexpovio® (selinexor)	Dexamethasone	Patients with at least 4 previous treatments, refractory to at least 2 PIs, 2 IMiDs, 1 monoclonal anti-CD38 antibody, and demonstrated disease progression during the previous treatment
CAR-T		
Abecma® (idecabtagen vicleucel)		Patients with RRMM, who have received at least 3 previous treatments, including 1 IMiD, 1 PI, 1 anti-CD38 antibody, and demonstrated disease progression during the previous treatment
Cytostatic agents		
Bendamustine	Prednisone	As primary therapy for MM (Durie-Salmon stage II with progression or stage III), patients > 65 years and not eligible for ASCT, with clinical neuropathy at diagnosis (excludes treatment with thalidomide or bortezomib)
Cyclophosphamide	Prednisone	Remission induction for plasmacytoma (also in combination with prednisone)
	Bortezomib/dexamethasone (VCD)	Induction therapy, NDMM
Doxorubicin		Advanced MM
Pegylated liposomal doxorubicin	Bortezomib	Progressive MM in pats. after at least 1 previous treatment and who have already undergone bone marrow transplantation or are not eligible for it
Melphalan	Prednisone/prednisolone or other anti-myeloma therapeutic agents or as high-dose monotherapy for conditioning before ASCT	MM (plasmacytoma)

Abbreviations: *ASCT, autologous stem cell transplant; HD, high-dose chemotherapy; MM, multiple myeloma; NDMM, newly diagnosed multiple myeloma; RRMM, relapsed/refractory multiple myeloma, XPO1, exportin 1*

Despite these treatment advances, a major proportion of patients with MM will fail all three classes of the current standard of care (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 [7]. Therefore, as TCR patients progress and are exposed to an increasing number of therapies, the duration of response decreases along with overall survival until the disease is ultimately fatal [8].

Pfizer is conducting the C1071003 study (MagnetisMM-3), an open-label, multicenter, non-randomized, single-arm, phase 2 study to evaluate the efficacy and safety of elranatamab (PF-06863135) in patients with MM refractory to at least one PI, one IMiD, and one anti-CD38 (i.e., TCR MM). Elranatamab is an injectable bispecific antibody designed to bind to

B-cell maturation antigen (BCMA) overexpressed on the surface of MM cells as well as to the CD3 receptor on the surface of cytotoxic T cells, bridging them together to activate an immune response [9]. The phase 2 clinical trial (NCT04649359; C1071003) commenced (first patient enrolled) in February 2021, completing enrollment in November 2021. Among evaluable patients in Cohort A (n = 123) of patients without prior BCMA-directed therapy, the results showed an objective response rate (ORR) of 61.0% (95% CI, 51.8 -69.6%), including a stringent complete response (sCR) rate of 15.4%, a complete response (CR) or 21.1%, and partial response (PR) of 4.9%). At 15-months follow-up, progression-free survival (PFS) and overall survival (OS) rates were 50.9% and 56.7% respectively [10, 11]. Extended follow-up of approximately 21 months demonstrated sustained clinical efficacy. The probability of maintaining a response at 18 months was 68.8%, and the median PFS and OS were 17.2 months (95% CI, 9.8 -not estimable) and 21.9 months (95% CI, 13.4 -not estimable), respectively. 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 [12]. To help contextualize the results of Study C1071003, there is a need to generate comparative evidence on clinical outcomes of TCR patients who receive SOC. This has been previously accomplished by creating an external control arm (ECA) 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 and refractory MM (RRMM). The MAMMOTH study identified 275 patients from 14 academic institutions in the US with a diagnosis of active MM who were refractory to an anti-CD38, daratumumab or isatuximab, administered as monotherapy or as part of a combination. Median OS from the time when patients were refractory to an anti-CD38 therapy was 8.6 months (95% CI: 7.5- 9.9). The ORR to the first regimen after patients were refractory to an anti-CD38 was 31%, with median progression-free survival (PFS) and OS of 3.4 and 9.3 months, respectively [13]. KarMMa-RW was a global non-interventional, retrospective study that assessed treatment patterns and outcomes in RW RRMM patients treated with therapies available during the study period and whose characteristics were similar to the KarMMa study, a phase 2 trial of ide-cel in heavily pretreated RRMM patients who were triple-class exposed (TCE) and refractory to the last regimen [14]. Approximately 43% of the KarMMA-RW population were TCR. Given a median follow-up of 13.3 months in KarMMa (n = 129) and 10.2 months among the RW RRMM patients (n = 190), ORR, very good PR or better (≥VGPR), PFS, and OS were significantly improved in KarMMa versus the matched-adjusted eligible RRMM subset (ORR, 76.4% vs 32.2%; ≥VGPR, 57.9% vs 13.7%; PFS 11.6 vs 3.5 months; OS 20.2 vs 14.7 months). Moreover, LocoMMotion (NCT04035226) is a prospective, non-interventional, and multinational study of 248 patients with TCE RRMM, receiving RW standard of care [15]. The study recruited patients who had received at least three prior LOTs or were double refractory to at least one PI and IMiD. Among these RW patients, ORR was 29.8% (95% CI 24.2 – 36.0%), PFS was 4.6 months (95% CI 3.9 – 5.6), and OS was 12.4 months (95% CI 10.3 – not reached).

ECAs have been increasingly used for both regulatory approvals and reimbursement purposes. Most external control arms used in regulatory submissions between 1999 and 2014 supported hematological oncology products [16]. A recent review of new drug applications

submitted to the Federal Drug Agency (FDA) for regulatory approval between 2017 and 2019 confirmed that a significant precedent exists for acceptance of RWE in support of oncology drug approvals, especially for rare or orphan indications. RWE submitted to FDA includes data from electronic health records (EHR), claims, post-marketing safety reports, retrospective medical record reviews, and expanded access studies. However, small sample sizes, data quality, and methodological issues were among the concerns cited by FDA reviewers [17]. Similarly, in a recent review of cancer drug submission approved by the European Medicines Agency (EMA) between 2016 and 2021, ECAs were included in 17% of approved submissions, with 63% of submitted ECAs regarded as supportive evidence for clinical effectiveness [18]. The remaining one-third of ECA submissions were not accepted as supportive evidence because of limitations on patient population heterogeneity, missing outcome assessment with respect to RWD, and unsuitable statistical analysis. Furthermore, a recently published review of the National Institute for Health and Care Excellence (NICE) appraisals found that from 2000 to 2016, 22 technologies (12 in oncology indications) were appraised by NICE based on non-RCT data; 27% of those used observational data to establish comparative effectiveness [19]. When assessing technologies based on non-RCT observational data, NICE committees considered several additional factors, including unmet clinical needs, small patient populations, and large treatment effects. In all 22 technology appraisals, concerns were raised regarding the immaturity of data and the uncertainty associated with the lack of a direct comparator. In Germany, the Gemeinsamer Bundesausschuss (G-BA) has not typically accepted ECA comparisons in the past due to their association with residual bias accompanied by comparing clinically controlled and uncontrolled RW settings. Based on the Institute for Quality and Efficiency in Health Care (IQWiG) guidelines, conclusions from benefit assessments are preferably inferred from results of direct comparative studies. In an internal review of 12 G-BA appraisals for indicated oncological agents using single-arm pivotal trials and comparator efficacy data, the submitted RWE was not considered in the assessment of additional benefit due to several key limitations, including differences in population characteristics, confounder identification, methods of comparisons (i.e., naïve comparisons rendered unsuitable), and transparency in reporting of characteristics and methods (data not published). In a recent study assessing the relevance of RWE in the benefit assessment process in Germany, RWE was incorporated to describe 222 out of 456 different target populations (228 assessments) as a source of epidemiological data [20]. Overall, the quality of the RWD and appropriate method of comparison are crucial when evaluating the contribution of presented RWE.

CCI

[redacted]

To contextualize the clinical profile of elranatamab, retrospective cohort studies (Studies C1071024; NCT05565391, and C1071031; NCT05932290), were conducted to compare the efficacy outcomes observed in the participants of Study C1071003 (with at least 9 months and 15 months of follow-up, respectively) and RW patients selected from 2 US-based oncology electronic health record (EHR) databases, Flatiron Health and COTA. The results of Study C1071024 showed that among TCR MM patients, those treated with elranatamab had significantly higher ORR than those treated with SOC. Thus, in the IPTW analysis comparing participants from Study C1071003 with RW patients from COTA or Flatiron Health, the estimated relative risk (RR) was 2.22 (95% CI: 1.69-2.90, p<.0001) and 1.79 (95% CI: 1.01-3.15, p=.0447), respectively. Differences in the time to response (TTR) between those treated with elranatamab and SOC were not consistently observed, however, DOR was significantly improved for elranatamab compared to SOC (IPTW hazard ratio [HR] 0.11, 95% CI: 0.06 - 0.22, p<.0001 and 0.21, 95% CI: 0.10 - 0.45, p<.0001 when comparing participants from Study C1071003 with RW patients from COTA or Flatiron Health, respectively). Study C1071031 found that patients with TCR MM treated with elranatamab had significantly longer PFS than those treated with SOC (IPTW HR=0.37, 95% CI: 0.22-0.64, p=.0003, when comparing with RW patients from COTA, and IPTW RMST difference 1.55, 2.48, 3.58, 4.74, 7.07 months more for elranatamab, respectively, all p<.05, when comparing with RW patients from Flatiron Health). Similarly, OS was also consistently longer for elranatamab compared with RW patients from COTA (IPTW HR=0.46, 95% CI: 0.27-0.77, p=.0032). When using the cohort with patients from Flatiron Health, in the IPTW analysis, OS was improved with elranatamab, but the results did not reach statistical significance at 9, 12, 15, 18, and 24 months (IPTW RMST difference 0.02, 0.31, 0.72, 1.14, and 2.34 months of additional survival for elranatamab, respectively, all p>.05). Finally, although underpowered, the exploratory analysis comparing PROs among participants treated with elranatamab and patients treated with SOC showed that treatment with elranatamab was generally comparable with no additional decline in patients' quality of life, functional scales, or disease burden through the first 6 months following treatment; indeed, in some cases, there was evidence for an improvement in overall quality of life and overall perceived disease state change.

This study aims to contextualize the outcomes of Study C1071003 by comparing a priori specified clinical effectiveness of patients treated with elranatamab using patient-level data from study C1071003 vs. RW external control of patients with TCE MM treated with SOC therapies from the TM-MM Germany dataset and from the Flatiron Health database. The study is an extension of studies C1071024 and C1071031 with more recent data and an alternative set of therapies included in the ECA, following the G-BA's definition of the appropriate comparator therapy, and with an alternative data base focusing on German patients. To further reduce the potential for bias, appropriate comparative effectiveness methods and statistical techniques will be utilized [23-25].

6. RESEARCH QUESTION AND OBJECTIVES

This study aims to assess the comparative effectiveness of elranatamab investigated in the open-label, multicenter, non-randomized single-arm Phase 2 Study C1071003. versus SOC treatment in TCR MM patients using an ECA of patients from the RW TM-MM Germany dataset and from the Flatiron Health database. Notably, SOC will be defined as patient individual therapy accordingly based on the approved list of comparator therapy in Germany (outlined further in Section 7.4.2, Table 6).

Study Objectives:

1. To compare clinical effectiveness among triple-class refractory (TCR) MM patients treated with elranatamab in Study C1071003 with a comparable RW cohort of TCR patients receiving SOC from the TherapyMonitor MM Germany dataset and from the Flatiron Health database, including:
    a. Overall survival (OS)
    b. Progression-free survival (PFS)
    c. Time to next treatment (TTNT)
    d. Time to discontinuation (TTD)
2. To describe treatment patterns of the SOC external control arm

7. RESEARCH METHODS

7.1. Study design

THIS IS A COMPARATIVE OBSERVATION STUDY USING PROSPECTIVELY AND RETROSPECTIVELY COLLECTED DATA TO EMULATE A TARGET TRIAL. THE STUDY WILL MAKE USE OF PATIENT-LEVEL, PROSPECTIVELY COLLECTED DATA FROM THE PHASE II STUDY C1071003 COMPRISING THE ELRANATAMAB ARM AND RETROSPECTIVE ANONYMIZED DATA FROM MEDICAL RECORDS COLLECTED AS PART OF THE TM-MM PROJECT IN GERMANY BY TRINETX ONCOLOGY (TNXO, FORMERLY ONCOLOGYINFORMATIONSERVICE (OIS) AND FROM ELECTRONIC MEDICAL RECORDS COLLECTED AS PART OF THE FLATIRON HEALTH DATABASE COMPRISING THE ECA. RELATING STUDY PROTOCOLS ARE IN

ANNEX 1. LIST OF STAND-ALONE DOCUMENTS

Target trial emulation principles will be employed to mitigate the risk of bias when performing the comparison [26, 27]. The following stepwise approach will be employed:

1. **Eligibility criteria:** To maximize comparability, the eligibility criteria for the participants from Study C1071003 will be applied to patients from the RWD source (see Section 7.7.2 for more details).
2. **Treatment strategies:** Initiation of elranatamab monotherapy from Study C1071003 versus initiation of SOC in the RW TM-MM dataset and in the Flatiron Health database, defined as patient individual therapy accordingly based on the approved list of comparator therapy, as outlined in Table 6, Section 7.4.2).
3. **Assignment procedure:** Individuals were assigned to the respective treatment groups at the time of treatment initiation in the respective data sources such that both the patient and healthcare provider was aware of the assigned treatment. For this study, assignment will be done when target trial eligibility criteria are met conditional on availability of pre-specified confounders listed in Section 7.4.3.1.
4. **Outcomes:** PFS (main), OS (main), TTD (supportive), and TTNT (supportive) will be comparatively assessed between Study C1071003 and the ECA (see Section 7.4.4)
5. **Follow-up:** Participants will be followed from baseline to the earliest of death, loss to or end of follow-up
6. **Causal contrast of interest:** Observational analog of the intention to treat (ITT) effect.

To compare outcomes between participants of Study C1071003 and the RW external control populations, population adjustment and sensitivity methods will be utilized to account for confounding and selection bias (see Section 7.7, Analysis Plan). Moreover, quantitative bias analysis (QBA) will be applied to address residual bias as well as mismeasured/unmeasured confounding and evaluate the robustness of results in the presence of potential threats to internal validity (see Section 7.7.6).

7.2. Setting

7.2.1. Study C1071003

Study C1071003 is an open-label, multi-center, non-randomized Phase 2 study of elranatamab (PF06863135) monotherapy among participants with TCR MM, which was initiated in February 2021. To determine the effects of prior BCMA-directed therapy on the response to elranatamab monotherapy, Study C1071003 enrolled two independent and parallel cohorts, one with patients who are naïve to BCMA-directed therapies (Cohort A; 123 patients) and the other with patients previously exposed to BCMA-directed therapy (Cohort B; 64 patients). As the results of Study C1071003 showed differences in patient characteristics and outcomes among Cohort A and Cohort B, and due to few patients with prior exposure to BCMA-directed therapy in the RW data source planned for this study, the focus of the comparisons will be on Cohort A in Study C1071003 in order to approximate similar cohorts.

7.2.2. Real-world data source

7.2.2.1. Overview

THE TREATMENT EFFECT OF STUDY C1071003 WILL BE CONTEXTUALIZED VIA A SERIES OF INDIVIDUAL COMPARISONS UTILIZING SOC THERAPEUTIC APPROACHES COMPRISING AN ECA, CONSTRUCTED USING VARIOUS INTERNATIONAL AND NATIONAL REAL-WORLD DATA SOURCES. THE RW DATA SOURCES AND/OR STUDIES HAVE BEEN SELECTED ACCORDING TO DATA AVAILABILITY AND AS FIT-FOR-PURPOSE FOR FULFILLING STUDY OBJECTIVES (TABLE 2) [21, 22, 28]. IN THIS STUDY, THE SAMPLE OF PATIENTS FOR THE CONSTRUCTION OF THE EXTERNAL CONTROL ARM WILL BE EXTRACTED FROM THE GERMAN TM-MM DATASET AND FROM THE FLATIRON HEALTH DATABASE. CCI [redacted]

CCI [redacted]

Table 2: Overview of RWD sources for planned ECA comparisons with Study C1071003

RWD source	Description	Key Endpoints
CCI [redacted]		
Flatiron EHR Data	Retrospective RW study, community & academic sites	ORR, TTR, DOR, PFS, OS
COTA HER Data	Retrospective RW study, academic sites, community centers and hospital systems	ORR, TTR, DOR, PFS, OS
TM-MM Germany	Retrospective RW study	PFS, OS, TTNT, TTD

7.2.2.2. TherapyMonitor Multiple Myeloma (TM-MM) Germany and Flatiron Health database

TM-MM Germany is a representative nationwide data collection of secondary data of patients with MM (ICD-10: C90.0x). The primary objective of TM-MM is to describe demographic features, clinical characteristics, and treatment courses in the clinical reality of patients with MM from the censor date back to the individual initial diagnosis of MM. The dataset consists of retrospective anonymized longitudinal data of therapy courses of MM patients in a representative sample of treatment centers, including university hospitals (UH), non-university hospitals (NUH), office-based hemato-oncologists (OBP) in Germany, capturing >10% of the annual prevalence of MM. Demographic data, data on diagnosis, classification and prognosis, as well as on therapy measures are collected retrospectively over the entire course of therapy. To date, TM-MM in Germany captures data from 01 January 2016 and 31 December 2022. The project is an ongoing effort, with continuous updates of data documentation.

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 US active 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.

Across the clinics in the Flatiron Health Network, data become available in near real time after each clinical encounter and contribute to Flatiron continuously aggregated centralized data set. Its total MM population is estimated at approximately 12,000 patients.

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 [Referenz: Flatiron Health Inc. Analytic guide for Flatiron Health Data. Triple-class refractory multiple myeloma. Version 1. 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 patient's journey.

Due to its timeliness and representativeness, the data provide a good basis for the analysis of RRMM patients in the overall prevalence of patients with MM. Data from TM-MM has been used previously as a source of epidemiological data and standard therapy determination in dossiers for the early benefit assessment in the AMNOG (Arzneimittelmarkt-Neuordnungsgesetz, i.e., Pharmaceuticals Market Reorganisation Act) procedure assessing the marketing of pharmaceutical products in Germany (Modul3A, daratumumab, carfilzomib, isatuximab). Data from TM-MM and from Flatiron Health have been previously used as an external control arm for CARTITUDE-1 and for MajesTEC-1, investigating a similar target population [30, 31, 32].

To set up TM-MM, an electronic case report form (eCRF) was developed together with clinical experts in MM. For data collection, the retrospective documentation of the eCRF is performed online in secuTrial™ (iAS GmbH, Berlin) by the health care professional (HCP) at the treatment center for the entire therapy course of the individual anonymized patients.

Flatiron Health is an oncology-focused health technology company that generates RWD from two EHR-derived primary sources: (i) OncoEMR®, a proprietary oncology-specific EHR used by community oncologists throughout the US, and (ii) EHR data integrations with academic research centers that enable bidirectional transmission of RWD [31].

To capture the complete patient journey in the TM-MM, data for all lines of therapy in MM are documented within the TM-MM dataset in addition to demographics and MM-specific characteristics at initial MM diagnosis and baseline, defined as the start of each line of therapy. New patients or patients with new lines of therapy are documented on a quarterly basis, whereas follow-up of patients is updated every six months retrospectively. A line of therapy is defined as all therapy measures that are carried out after a therapy decision has been made before a progression/relapse occurs and can include systemic therapy for tumor reduction (induction therapy), systemic therapy for stem cell mobilization, high-dose chemotherapy, stem cell transplant (SCT), consolidation or maintenance therapy, or supportive therapy.

An overview of the data contents and the subsequent timepoint of assessment as captured by TM-MM are outlined in Table 3.

Table 3. Summary of data parameters in the existing TM-MM Germany dataset

Category	Variables	Timepoint of Assessment
Patient Characteristics	Age	Initial diagnosis and the beginning of each line of therapy
	Sex	
	Condition (ECOG, individual assessment of patient fitness by physician)	
	Height, weight, body surface area	
	Symptomatology	
	Concomitant diseases requiring treatment	
Multiple myeloma-specific characteristics	Date of initial diagnosis	Initial diagnosis and the beginning of each line of therapy
	Stage (ISS, Durie & Salmon)	
	MM-type	
	M-protein type	
	Bone lesions	
	(SLiM)-CRAB Criteria	
	Presence of EMD (available from Q4 2017)	
	Cytogenetic examination of bone marrow with fluorescence in situ hybridization (FISH) – to be grouped per risk categories	

Category	Variables	Timepoint of Assessment
	Laboratory values: serum M-protein, serum albumin, serum beta-2-microglobulin, serum lactate dehydrogenase	
Therapy line (all lines observable for the patient)	Therapeutic measures including systemic therapy for tumor reduction (induction), stem cell mobilization, high dose chemotherapy, SCT, consolidation therapy, maintenance therapy, surgery, supportive therapy, CAR-T therapy	Each line of therapy
	Duration of remission	
Progression/Relapse	Definition of progression/relapse by variables(M protein levels, new or enlargement of known osteolysis, pathological fractures, development of hypercalcemia etc.) , multiple responses possible. In RWE analyses, progression-free survival will be defined using time to next treatment as the proxy for progression.	Each line of therapy
Therapy details for each substance	Dosage	Each line of therapy
	Number of administrations per cycle	
	Interval (cycle length)	
	Start and end date of each agent	
	Break in administration	
Modification of the therapy of a substance and/or regimen	Dosage per agent	Each line of therapy
	Number of administrations per cycle	
	Interval (cycle length)	
	Therapy breaks	
	Intensification due to addition of a substance	
	De-escalation due to removal of a substance	
	Switch to a different therapy	
	Discontinuation of therapy	
Institutional characteristics	Therapy initiated/performed by health care sector	Each line of therapy
State at end of treatment	Treatment state (ongoing, paused, death and date thereof)	End of each documentation/update

Abbreviations: *CAR, chimeric antigen receptor; ECOG, Eastern Cooperative Oncology Group; EMD, Extramedullary disease; FISH, fluorescence in-situ hybridization; ISS, International Staging System; MM, multiple myeloma; RWE, real-world evidence; SliM-CRAB, criteria involving the following: 60% or more clonal plasma cells (S), light chains (Li), and MRI (M); also elevated calcium levels I, renal failure I, anemia (A), and bone lesions (B); TCE, triple-class exposed*

7.3. Study period

In this study, the TCR eligibility date will be defined as the start of the earliest LOT in which patients have become refractory to at least one IMiD, one PI, and one anti-CD38 agent and have subsequently initiated the treatments of interest as part of the next LOT (index date = initiation of therapy [time zero]). Given that the first anti-CD38 antibody (daratumumab) was granted conditional marketing authorization across the EU by the EMA starting on 20 May 2016, only patients with an index date, as required per selection criteria occurring between 20 May 2016 and 31 December 2021 (inclusion period) will be selected from the TM-MM dataset for the external control in order to allow for a minimum window of 6-months in which time-to-event (TTE) outcomes can be evaluated until the end of the study period on 31 December 2022 (end of TM-MM data availability). Given that the first anti-CD38 antibody (daratumumab) was approved by the FDA in the US on 16 November 2015, only patients with an index date, as required per selection criteria occurring between 16 November 2015 and 31 January 2023 (inclusion period) will be selected from the Flatiron Health database for the external control in order to allow for a minimum window of 6-months in which time-to-event (TTE) outcomes can be evaluated until the end of the study period on 31 July 2023 (end of Flatiron Health data availability). This criterion will be applied to ensure that subjects have the potential for at least some meaningful follow-up to capture the study outcomes.

The study period will be comprised of the baseline period (time preceding or on the index date) and the observational period (time following the index date, Figure 1). The baseline period (screening period) for the participants of Study C1071003 is 28 days before the initiation of elranatamab. In the TM-MM dataset, baseline characteristics are documented at the time of treatment initiation and the beginning of each line of therapy. As such, the baseline for the ECA will be defined as the start of the index LOT (i.e., index date), as documented by physicians in TM-MM and Flatiron Health. Patients will be followed in the observational period, respectively, from the index until the earliest of death, loss to follow-up, or end of the study period. Notably, time from first MM diagnosis or first-line therapy initiation to triple class refractoriness and eligibility for the comparisons of interest is comparable in both the elranatamab and ECA, thereby in this respect, the study design reduces the risk of selection bias and immortal time bias.

The scheme in Figure 1 below summarizes the study periods for Study C1071003 and the external control arm, respectively.

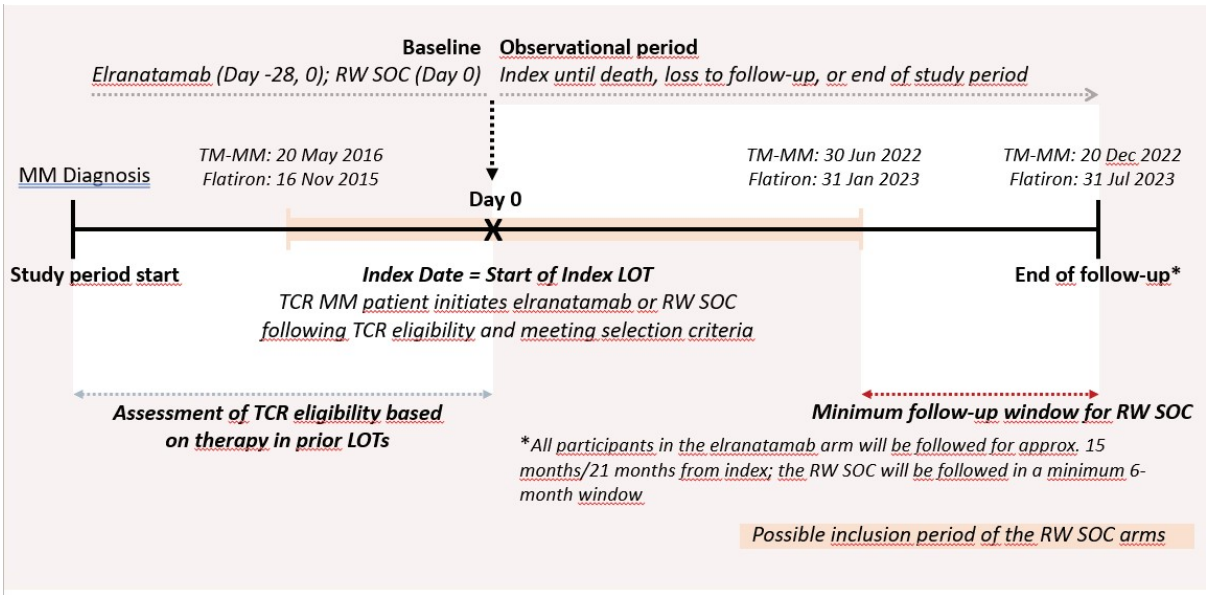

Figure 1. Baseline and Observational Periods of the target trial (Elranatamab in Study C1071003 versus RW SOC in TM-MM and Flatiron Health)

Abbreviations: *LOT, line of therapy; MM, multiple myeloma; OS, overall survival; PFS, progression-free survival; RW, real-world; SOC, standard of care; TCR, triple-class refractory; TTD, time-to-discontinuation; TTNT, time-to-next treatment*

7.4. Target trial emulation

7.4.1. Eligibility Criteria

The selection criteria for TCR patients from Study C1071003 are summarized in detail in Table 4 (inclusion criteria) and Table 5 (exclusion criteria). Patients of the ECA will be selected to maximize comparability with the population of Study C1071003.

However, applying inclusion and exclusion criteria from the clinical trial to the RW databases will require adjustments due to data availability and differences in clinical assessments.

In the Flatiron Health database, refractory status will be identified based on IMWG-derived progression events by comparing subsequent lab values of the same specimen type to its baseline/nadir value using IMWG criteria (i.e., based on information from the Enhanced_MM_progression_v2 data table for Flatiron Health). The indicator of progression uses a baseline/nadir that is identified based on eligible labs i.e., SPEP, UPEP, or FLC occurring ≤ 90 days before or any time after initial MM diagnosis. The closest measurement to the MM diagnosis date occurring on or before first-line therapy initiation after MM diagnosis is prioritized. If there is no valid measurement before initiation of the first line after MM diagnosis, the earliest measurement after the initiation of the first line is used. Only new lab measurements occurring ≥14 days after the initial baseline/nadir measurement are eligible for assessment of progression. The 14-day window is applied to ensure that the same event

(e.g., progression) is not captured multiple times and to ensure a new event is not captured before the effect of the initial therapy can be observed.

Once a baseline/nadir value has been established for a patient, only lab specimens of the same type will be used to evaluate progression in that patient throughout their entire patient journey. If a new lab value under consideration is less than the existing baseline/nadir value, the baseline/nadir value is updated to reflect the new lab value (e.g., if the baseline/nadir M spike is 0.8 g/dL and the new lab value M spike is 0.5 g/dL, the baseline/nadir will be updated to 0.5 g/dL). If the new lab value under consideration is greater or equal to the baseline/nadir value, an evaluation of progression is performed. At least 30 days needs to have elapsed since the previous progression event, in order for a new lab specimen to be eligible for evaluation of progression.

Disease progression is identified based on changes from baseline/nadir value in SPEP, UPEP, or FLC lab values as per IMWG criteria. There are, however, some limitations in directly applying the IMWG criteria to the Flatiron data (or other EHR data) due to data availability and differences in assessment in a clinical trial versus in an RW setting. For example, 24-hour UPEP tests are not widely used in RW settings. There is also limited availability of data for bone marrow biopsies to assess plasma cell percentage and limited availability of radiology data for the assessment of plasmacytomas and lytic bone lesions.

Notably, refractoriness to single agents is not available in the TM-MM dataset for defining the ECA. However, at the start of each LOT (line of therapy) excluding the first LOT, physicians indicate whether the patient is initiating therapy based on the status "refractory," "relapse," or "unknown" relative to their prior line. Hereby, TCR MM patients eligible for selection into the ECA are defined as MM patients who were exposed to at least one IMiD, one PI, and one anti-CD38 therapy in a line of therapy marked as "refractory". The IMiD, PI, and anti-CD38 treatments could be used as part of distinct lines of therapy or within the same line and as part of any therapeutic measure (induction, maintenance, consolidation, etc.). Furthermore, in TM-MM, physicians are asked to document all therapeutic measures in one

line of therapy until the patient progresses, marking the start of documentation of the next line of therapy. As the date of progression is not available directly from the TM-MM data, a proxy of start of the next treatment line after exposure to the above therapies will be used. 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 in Study C1071003. Patients will be considered refractory to all treatments received in a refractory line of therapy that was discontinued due to the progression (i.e., start of the next line of treatment).

As defined in Section 7.3, the TCR eligibility date is marked by the start of the earliest LOT in which patients have become refractory to at least one IMiD, one PI, and one anti-CD38 agent and have subsequently initiated the next LOT. The entire patient history dating back to the initial MM diagnosis is available and will be used to assess TCR eligibility. The start of the earliest LOT where the patient meets the TCR definition will therefore define the index date and qualify as the index LOT.
Detailed mapping to the MagnetisMM-3 inclusion and exclusion criteria can be found in Table 4 and Table 5, respectively. For each criteria, the overlap of the definition used in Study C1071003 and the closest available definition in TM-MM was assessed as “exact” or “partial.” Given the availability of an overlapping definition in TM-MM, all overlapping (partial or exact) criteria were then allocated in two sets: (1) a core set of indispensable criteria that will serve as the basis of the ECA selection and (2) sensitivity set (Table 4 and Table 5). The sensitivity set consists of a subset of partially overlapping criteria whereby the RW definition is comparatively more strict than the definition used in Study C1071003 (ie. potentially applying to a larger subset of patients in the absence of detailed clinical/biochemical values for assessment).

The final core inclusion and exclusion criteria to define the ECA based on data availability in TM-MM and maximized comparability to the selection of Study C1071003 are summarized below. In the sensitivity approach, the main analyses will be repeated utilizing the extended set of criteria (core and sensitivity) to select the ECA. Respectively, patients with cardiovascular diseases (any of the following; coronary heart disease, cardiac insufficiency, and cardiac arrhythmia), renal impairment, depression and other psychiatric diseases, or hepatic impairment at baseline will additionally be excluded as a result.

The approval for Elranatamab is expected to most likely include patients who received at least 3 prior lines of therapy. To meet the anticipated label population, patients with <3 prior lines of therapy will be excluded from both the intervention and the RW control arms.

Inclusion Criteria

- Male or female patients with MM (HCP documented per IMWG criteria)
- Exposure to at least one IMiD, one PI, and one anti-CD38 antibody (independent of sequence of receipt)

- Initiation of a new LOT with one of the patient individualized SOC regimens outlined in Table 6, Section 7.4.2 after triple-class exposure (index date: start of LOT) between 20 May 2016 and 30 June 2022
- Relapsed/refractory to last anti-MM LOT
- Age ≥18 years at index
- ECOG Performance ≤2[1]

Exclusion Criteria

- Smoldering MM documented at initial diagnosis (which is not converted to symptomatic MM at 1L therapy start)
- Plasma cell leukemia documented at initial diagnosis
- If female and documented as "pregnant, breastfeeding"
- Any concomitant malignancy documented at index
- Stem cell transplant within 12 weeks prior to index
- Presence of peripheral neuropathy assessed at index
- Any systemic infection or SARS-CoV2 infection at index
- Any surgery within 14 days of index
- End of therapy with an agent in a line of therapy given as part of a clinical trial is within 30 days of index or index line of therapy is part of a clinical trial
- Patients with <3 prior lines of therapy
- For complete case analysis: patients with missing data of any pre-specified baseline confounders as defined in Section 7.4.3.1[2]
- Prior BCMA-directed therapy

Patient numbers will be reported at each selection step.

[1] Patients with unknown ECOG status will be excluded; based on results of the descriptive study for TM-MM and previous studies with Flatiron Health, missing ECOG was infrequent.

[2] For the complete case analysis (sensitivity), patients with missing data for the pre-specified confounders will be excluded. The base approach will allow for imputation of missing confounders, described further in Section 7.7.3

Table 4: Mapping of inclusion criteria for Study C1071003 onto the external control control cohort (TM-MM)

Study C1071003 inclusion criteria	Definition overlap (TM-MM/Flatiron Health)	ECA (TM-MM / Flatiron Health) inclusion criteria & considerations	Core criteria set	Sensitivity criteria set
Male or female participants age ≥18 years	Exact	Male or female participants age ≥18 years	X	
Participants who are willing and able to comply with all scheduled visits, treatment plan, laboratory tests, lifestyle considerations, and other study procedures.	NA	Not available/not applicable		
Prior diagnosis of MM as defined according to IMWG criteria (Rajkumar et al, 2014).	Exact	MM diagnosis (*TM-MM:* ICD-10 C90.0x, HCP documented per IMWG criteria / *Flatiron Health*: ICD-9 203.0x or ICD-10 C90.0x)	X	
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)	Partial/Exact	*Only partial for TM-MM: assumption that all patients fulfill this criterion and are subsequently receiving treatment for their disease. While serum-M protein values (a) are available, the method of measurement and criteria is not specified.*	X	
Refractory[3] to at least one IMiD	Partial/Exact	Refractory to at least one ImiD (*only partial for TM-MM*: based on relapse/refractory status at start of a subsequent LOT)	X	
Refractory (refer to footnote 1) to at least one PI	Partial/Exact	Refractory to at least one PI	X	

[3] Refractory is defined as having disease progression while on therapy or within 60 days of last dose in any line, regardless of response.

Study C1071003 inclusion criteria	Definition overlap (TM-MM/Flatiron Health)	ECA (TM-MM / Flatiron Health) inclusion criteria & considerations	Core criteria set	Sensitivity criteria set
		(*only partial for TM-MM*: based on relapse/refractory status at start of a subsequent LOT)		
Refractory (refer to footnote 1) to at least one anti-CD38 antibody	Partial/Exact	Refractory to at least one anti-CD38 antibody (*only partial for TM-MM*: based on relapse/refractory status at start of a subsequent LOT)	X	
Relapsed or refractory to last anti-MM regimen	Exact	Relapsed or refractory to last anti-MM LOT	X	
ECOG performance status ≤2	Exact	ECOG performance status ≤2	X	
LVEF ≥40% as determined by a MUGA scan or ECHO	NA	Not available		
Adequate hepatic function is 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	Partial/Exact	Criterion will not be applied *(only partial for TM-MM: Closest match refers to patients marked with hepatic impairment at index*		X
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 local institutional standard method).	Partial/Exact	Criterion will not be applied *(only partial for TM-MM: Closest match refers to patients marked with renal impairment at index*		X
Adequate bone marrow function is characterized by the following: d. 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);	NA/Exact	*TM-MM*: Not available / *Flatiron-Health*: Although data available in Flatiron Health, we will not include in inclusion criteria to align with approaches in previous RWE studies. The presence of		

Study C1071003 inclusion criteria	Definition overlap (TM-MM/Flatiron Health)	ECA (TM-MM / Flatiron Health) inclusion criteria & considerations	Core criteria set	Sensitivity criteria set
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).		criteria will still be measured to describe patients.		
Resolved acute effects of any prior therapy to baseline severity or CTCAE Grade ≤1.	NA	Not available		
Capable of giving signed informed consent, which includes compliance with the requirements and restrictions listed in the ICD and in the MagnetisMM-3 protocol.	NA	Not available/not applicable		

Abbreviations:
ALT, alanine transaminase; ANC, absolute neutrophil count; AST, aspartate transaminase; CTCAE, Common Terminology Criteria for Adverse Events; ECHO, echocardiogram; ECOG, Eastern Cooperative Oncology Group; FLC, free light chain; HCP, healthcare professional; ICD, informed consent document; ImiD, immunomodulator; IMWG, International Myeloma Working Group; LOT, line of therapy; LVEF, left ventricular ejection fraction; MM, multiple myeloma; MUGA, multigated acquisition; NA: not applicable; PI, proteasome inhibitor; SPEP, serum protein electrophoresis; ULN, upper limit of normal; UPEP, urine protein electrophoresis

Table 5: Mapping of exclusion criteria between MagnetisMM-3 and the Trial Cohort

MagnetisMM-3 exclusion criteria	Definition overlap (TM-MM/Flatiron Health)	ECA (TM-MM / Flatiron Health) exclusion criteria & considerations	Core criteria set	Sensitivity criteria set
Female participants who are pregnant or breastfeeding	Exact/NA	TM-MM: If female and documented as “pregnant, breastfeeding”. Captured as one variable	X	

MagnetisMM-3 exclusion criteria	Definition overlap (TM-MM/Flatiron Health)	ECA (TM-MM / Flatiron Health) exclusion criteria & considerations	Core criteria set	Sensitivity criteria set
Smoldering MM	Partial/NA	*TM-MM*: Smoldering MM documented at initial diagnosis. Any smoldering MM which was not converted to active multiple myeloma at start of first line therapy will be excluded / *Flatiron Health*: No specific International Classification Disease (ICD)-9/10 codes that allow identification of smoldering MM. Patients are included in the data if they have evidence of active MM. Since smoldering MM can be a precursor to active MM, patients may have had SMM before progressing to active MM.	X	
Active plasma cell leukemia	Partial/Exact	Plasma cell leukemia documented initial diagnosis and any concomitant malignancy at index will be excluded	X	
Amyloidosis	NA/Exact	*TM-MM*: Not available	X	
POEMS syndrome	NA	Not available		
Stem cell transplant within 12 weeks prior to enrollment or active GVHD	Partial	Stem cell transplant within 12 weeks prior to index *TM-MM: GVHD not available /* *Flatiron Health: GVHD could be identified based on diagnosis codes but there are no specific codes for "active"*	X	

MagnetisMM-3 exclusion criteria	Definition overlap (TM-MM/Flatiron Health)	ECA (TM-MM / Flatiron Health) exclusion criteria & considerations	Core criteria set	Sensitivity criteria set
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); 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).	Partial	Criterion will not be applied *TM-MM: Closest match is the "presence of coronary heart disease, cardiac insufficiency, and cardiac arrythmia" at the start of each therapy line; clinical significance and history of these diseases cannot be ascertained /* *Flatiron Health:* identified based on diagnosis codes (except for QT syndrome)		X
Ongoing Grade ≥2 peripheral sensory or motor neuropathy History of GBS or GBS variants, or history of any Grade ≥3 peripheral motor polyneuropathy	Partial	*TM-MM*: Presence of peripheral neuropathy assessed at index, GBS and grade of polyneuropathy are not available / *Flatiron Health*: ongoing Grade ≥2 peripheral sensory or motor neuropathy cannot be identified in the data as there is no information about Grade. Treatments used to manage neuropathic pain will be used as a proxy to identify presence of more severe neuropathy cases to describe patients. GBS could be identified based on diagnosis codes. There is no information about grade in the data that allows identifying Grade ≥3 peripheral motor polyneuropathy. Treatments used to manage	X	

MagnetisMM-3 exclusion criteria	Definition overlap (TM-MM/Flatiron Health)	ECA (TM-MM / Flatiron Health) exclusion criteria & considerations	Core criteria set	Sensitivity criteria set
		neuropathic pain could be used as a proxy to identify more severe neuropathy cases.		
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	TM-MM: Any systemic infection or SARS-CoV2 infection at index.Active HBV, HCV, HIV, bacterial, fungal, or viral infections specifically are not available. / *Flatiron Health:* could be identified 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.	X	
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.	Partial	Any concomitant malignancy at index *TM-MM:* Timing of other malignancies and specifically adequately treated basal cell or squamous cell skin cancer or carcinoma in situ are not available / *Flatiron Health*: identified based on diagnosis codes or via abstracted data, but there are no specific codes for “active”.	X	

MagnetisMM-3 exclusion criteria	Definition overlap (TM-MM/Flatiron Health)	ECA (TM-MM / Flatiron Health) exclusion criteria & considerations	Core criteria set	Sensitivity criteria set
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.	Partial/NA	TM-MM: Any surgery within 14 days of index. Major surgery or 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 cannot be assessed; the closest equivalent to psychiatric conditions is "depression and other psychiatric diseases" documented at index (exclusion in a combined sensitivity analysis)	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).	Exact	End of therapy with an agent in a line of therapy given as part of a clinical trial is within 30 days of index or the index line is part of a clinical trial	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.	NA	Not available		
Known or suspected hypersensitivity to the study intervention or any of its excipients	NA	Not available		
Live attenuated vaccine must not be administered within 4 weeks of the first dose of study intervention.	NA	Not available		

Abbreviations:
HBV, hepatitis B virus; HCV, hepatitis C virus; HIV, human immunodeficiency virus; GBS, Guillain-Barré syndrome; GVHD, graft versus host disease; MM, multiple myeloma; NA: not applicable; POEMS, syndrome characterized by polyneuropathy, organomegaly, endocrinopathy, monoclonal protein, and skin changes; QTcF, corrected QT interval by Fredericia; SARS-CoV2, severe acute respiratory syndrome coronavirus 2

7.4.2. Treatment strategies

In each comparison, patients will be classified into one of two treatment groups according to the therapy received after TCR eligibility, respectively as those treated with elranatamab 76 mg (Study C1071003) and those treated with a SOC regimen (no pre-specified dose will be defined; TM-MM and Flatiron Health). In the RW setting, no single SOC currently exists for TCR MM patients, and combinations of treatments are frequently used in lieu of monotherapy [5]. In this protocol, the term "SOC" refers to patient individual therapy accordingly based on the approved list of comparator therapy in Germany, outlined in Table 6.

Table 6: Approved standard of care regimens

RW standard of care therapy
Bortezomib monotherapy
Bortezomib + pegyliertes liposomales Doxorubicin
Bortezomib + Dexamethasone
Carfilzomib + Lenalidomide + Dexamethasone
Carfilzomib + Dexamethasone
Daratumumab + Lenalidomide + Dexamethasone
Daratumumab + Bortezomib + Dexamethasone
Daratumumab Monotherapy or
Daratumumab + Pomalidomide + Dexamethasone
Elotuzumab + Lenalidomide + Dexamethasone
Elotuzumab + Pomalidomide + Dexamethasone
Isatuximab + Pomalidomide + Dexamethasone
Ixazomib + Lenalidomide + Dexamethason
Lenalidomide + Dexamethasone
Panobinostat + Bortezomib + Dexamethasone
Pomalidomide + Bortezomib + Dexamethasone
Pomalidomide + Dexamethasone
Cyclophosphamide (in combination with other antineoplastic drugs)
Melphalan (as monotherapy or in combination with prednisone or prednisolone)
Doxorubicin (as monotherapy or in combination with other antineoplastic drugs)
Vincristine (in combination with other antineoplastic drugs)
Dexamethasone (in combination with other antineoplastic drugs)
Prednisolone (in combination with other antineoplastic drugs)
Prednisone (in combination with other antineoplastic drugs)
Best supportive care

7.4.3. Assignment procedure

Individuals will be assigned to the treatment groups based on the date of eligibility of either treatment conditional on the adjustment of the pre-specified list of confounders at baseline, outlined in detail in Section 7.4.3.1.

7.4.3.1. Potential confounders

Ensuring that baseline factors related to treatment assignment and outcomes are captured is of utmost importance for the validity of target trial emulation. The availability of these variables in selected RWD sources determines the feasibility of the external control arm for the target trial. Baseline covariates will be selected to compare patients from elranatamab and the ECA and to perform further statistical adjustments to control for the baseline confounding (e.g., via inverse probability of treatment weighting [IPTW]).

Across both cohorts, baseline covariates to be considered will include those related to patient demographic characteristics, disease characteristics, comorbidity profile, laboratory measurements, and MM treatment patterns (LOTs), as available for both treatment arms. For the external control arm, each baseline variable will be taken on or before the index date; if before, the most recent measurement will be used (physician documentation of baseline characteristics is documented based on the timepoint of treatment initiation). For Study C1071003, baseline characteristics will be considered on the index date or in the 28 days before the index, as applicable. All variables listed below will be considered to describe the study cohorts and to adjust for the baseline confounding. Since a small sample size limits the number of covariates that can be reasonably accounted for in the statistical adjustment (eg, IPTW), a systematic literature review of 57 studies (22 univariate; 35 multivariate analyses to predict outcomes) was previously conducted to identify the variables most strongly and consistently correlated with outcomes in RWD studies and RCT conducted among RRMM patients. Moreover, interviews with medical experts were conducted to identify and discuss all relevant confounders. Collectively, the results of these studies as well as the feedback from medical experts 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. Laboratory values such as hemoglobulin (not available in TM-MM) and albumin levels exhibited limited evidence as prognostic variables.

CCI [redacted] If applicable, multiple imputations will be used to address the presence of missing values for covariates in RWD in a sensitivity analysis prior to patient selection of the main TCR synthetic control arm (SCA) cohorts as outlined in Section 7.4 (see Section 7.7.3 for further details on multiple imputation).

CCI [redacted]

CCI

[redacted]

CCI, *CONTINUOUS MEASURE*. FOR THE ECA, CCI WILL BE CALCULATED BASED ON THE ALGORITHM OUTLINED IN

- ANNEX 2. ADDITIONAL INFORMATION Table 10 employs available comorbidity documentation at the index date.
- Prior SCT, *categorical measure (yes; no)*
- Aspartate aminotransferase (microkat/L) within 90 days before or on the index date, if feasible
- Alanine aminotransferase (microkat/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
- Hemoglobin (g/L) within 90 days before or on the index date, if feasible
- Calcium in serum or plasma (mmol/L) reported within 90 days before or on the index date

Further information of the operational definition of all confounders as defined for the ECA can be found in the ANNEX 2A, Table 9.

7.4.3.2. Other variables

TABLE 9 IN

ANNEX 2. ADDITIONAL INFORMATION provides the operational definition of additional variables that will be used to describe patient characteristics (in addition covariates in population adjustment methods in the section above) and treatment patterns of the external control cohorts, as per availability in the RW datasets. In addition to the description of patient characteristics (extended list of baseline characteristics as defined in Table 9), the external control arms will further be characterized with respect to treatment patterns. Any prior treatment use on the agent and class level will be described, in addition to treatment use/regimens in the line prior to the index LOT (pre-index), the index LOT, and the line following the index LOT (post-index LOT).

7.4.4. Outcomes

Consistency between outcome definitions across data sources is another key consideration to minimize bias. Comparative effectiveness endpoints PFS and OS will be evaluated based on availability within the RW dataset. Definitions of the outcomes will be aligned, where possible, with Study C1071003 (Table 7. Clinical outcome definitions based on Study C1071003 and TM-MM Germany).

Table 7. Clinical outcome definitions based on Study C1071003 and TM-MM Germany

Endpoint	Study C1071003	ECA (TM-MM Germany / Flatiron Health)	Classification
PFS	Time from the date of the first dose until confirmed PD per IMWG criteria or death due to any cause, whichever occurs first. Patients are censored at loss to or end of follow-up. *Sensitivity:* Time from date of the first dose until initiation to either the date of progression (using the proxy: initiation of the next treatment line) or death due to any cause. Patients are censored at loss to or end of follow-up.	Time from initiation of the index LOT (index date) to either the date of progression (proxy in TM-MM dataset: initiation of the next treatment line) or death due to any cause. Patients are censored at loss to follow-up or end of the study period.	Main

OS	Time from the date of the first dose until death due to any cause. Patients are censored at loss to or end of follow-up.	Time from initiation of the index LOT until death due to any cause (18). Patients are censored at loss to follow-up or end of the study period.	Main
TTNT	Time from date of first dose to initiation of the next treatment line. Patients are censored at death, loss to or end of follow-up.	Time from initiation of the index LOT (index date) to start of the next treatment line. Patients are censored at loss to follow-up, death, or end of the study period.	Supportive
TTD	Time from the date of first dose to discontinuation. Patients are censored at death, loss to or end of follow-up.	Time from initiation of the index LOT (index date) to discontinuation (end) of the LOT. Patients are censored at loss to follow-up, death, or end of the study period.	Supportive

Abbreviations: *IMWG, International Myeloma Working Group; LOT, line of therapy; PD, progressive disease*

7.4.5. Follow-up

As per the study period specifications outlined in Section 7.3, patients will be followed from baseline to the earliest of death, loss-to or end of follow-up (including end of the study period on 31 December 2022 for the TM-MM dataset, and on 31 July 2023 for the Flatiron Health database).

7.4.6. Causal contrast

This study will estimate the observational analog of the intention-to-treat effect, more specifically, the effect of initiating the prespecified treatment regimens described above in Section 7.4.2.

7.5. Data source integration

Upon constructing the analytical file for the comparison between Study C1071003 and TM-MM / Flatiron Health, a common data model for Study C1071003 and the RW external control arms will be created with standardized data elements. 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. All data transformation will be documented for transparency.

Combined patient-level study datasets will be created from the entirety of the Study C1071003 arm plus the RW external control arms from TM-MM and Flatiron Health based on the eligibility criteria.

Six combined datasets will be created. For the purpose of all main, sensitivity, subgroup, and supportive analyses outlined in Section 7.7, Table 8:

- Dataset I: Study C1071003 Cohort A arm plus the ECA of TCR patients from TM-MM, naïve to BCMA-directed therapy. All relevant baseline confounders listed in

Section 7.4.3.1 will be imputed in case of missingness (see Section 7.7.3 for further details).

- Dataset II: Study C1071003 Cohort A arm plus the ECA of TCR patients from Flatiron Health, naïve to BCMA-directed therapy. All relevant baseline confounders listed in Section 7.4.3.1 will be imputed in case of missingness (see Section 7.7.3 for further details).

For the purpose of additional analyses, (described further in Section 7.7.5.7):

- Dataset III: Study C1071003 Cohort A arm plus the ECA of TCR patients from TM-MM, naïve to BCMA-directed therapy, with additional exclusion of the following partially matched criteria: cardiovascular disease, renal impairment, hepatic impairment, or depression and other psychiatric diseases at baseline (sensitivity)

- Dataset IV: Study C1071003 Cohort A arm plus the ECA of TCR patients from Flatiron Health, naïve to BCMA-directed therapy, with additional exclusion of the following partially matched criteria: cardiovascular disease, renal impairment, hepatic impairment, or depression and other psychiatric diseases at baseline (sensitivity)

- Dataset V: Study C1071003 Cohort A arm plus the ECA of TCR patients from TM-MM, naïve to BCMA-directed therapy, selected conditional on complete cases of all relevant baseline confounders listed in Section 7.4.3.1

- Dataset VI: Study C1071003 Cohort A arm plus the ECA of TCR patients from Flatiron Health, naïve to BCMA-directed therapy, selected conditional on complete cases of all relevant baseline confounders listed in Section 7.4.3.1

7.6. Study size

The study sample was determined in the previously mentioned descriptive study using TM-MM (protocol in
ANNEX 1. LIST OF STAND-ALONE DOCUMENTS). The descriptive study identified 858 patients eligible matching the selection criteria of Study C1071003 adapted to available RWD parameters as part of sample I (individual patients) and 1,316 observations as part of sample II (individual LOTs). Notably, the study sample is conditional on the final selection criteria and the inclusion period, which will be extended for one additionally quarter to June 2022 as a result of data update (previously: inclusion period end on March 2022). All patients who meet the inclusion/exclusion criteria of the external control arm as defined in Section 7.4.1. will be included in the analyses. For Flatiron Health, the sample size is limited by the duration of the observation window. All patients who meet the inclusion/exclusion criteria defined in Section 7.4.1 will be included in the analysis. All these patients were included in the study. Study C1071003 consists of 187 patients overall (123 Cohort A, 64 Cohort B). No formal sample size estimations have been performed for this observational study.

7.7. Analysis plan

TABLE 8 PROVIDES A SUMMARY OVERVIEW OF THE PLANNED STATISTICAL ANALYSES AND OUTPUTS FOR COMBINED STUDY COHORTS FOR CLINICAL OUTCOMES. IN ADDITION TO OUTCOMES SPECIFIED IN TABLE 8, AVAILABLE PATIENT CHARACTERISTICS (SEE

ANNEX 2. ADDITIONAL INFORMATION Table 9) including demographics, clinical/disease characteristics, laboratory tests/values, and treatment patterns will be analyzed for the ECA.

Table 8: Overview of planned statistical analyses for clinical outcomes

Endpoint	Method	Estimate	Level (Dataset)
Overall survival (OS)	Kaplan-Meier (KM)	KM curve with patients at risk and censored, median & quartiles (95% CI)	Main (I-VI)
	IPTW, Cox proportional hazards regression	HR (95% CI), p-value	Main (I-VI)
	PSM, Cox proportional hazards regression		Sensitivity (I-II)
	Naïve comparison		Sensitivity (I-II)
	Multiple regression		Sensitivity (I-II)
	Doubly robust comparison		Sensitivity (I-II)
	Fine Stratification Weights		Sensitivity (I-II)
	Overlap Weights		Sensitivity (I-II)
	Quantitative Bias Analysis (QBA)	E-values	Sensitivity (I-II)
Progression-free survival (PFS)	Kaplan-Meier (KM)	KM curve with patients at risk and censored, median & quartiles (95% CI)	Main (I-VI)
	IPTW, Cox proportional hazards regression	HR (95% CI), p-value	Main (I-VI)
	PSM, Cox proportional hazards regression		Sensitivity (I-II)
	Naïve comparison		Sensitivity (I-II)
	Multiple regression		Sensitivity (I-II)
	Doubly robust comparison		Sensitivity (I-II)
	Fine Stratification Weights		Sensitivity (I-II)
	Overlap Weights		Sensitivity (I-II)
	Quantitative Bias Analysis (QBA)	E-values	Sensitivity (I-II)
Time to next treatment (TTNT)	Kaplan-Meier (KM)	KM curve with patients at risk and censored, median & quartiles (95% CI)	Supportive (I-VI)
	IPTW, Cox proportional hazards regression	IPTW, Cox proportional hazards regression	Supportive (I-VI)
Time to discontinuation (TTD)	Kaplan-Meier (KM)	KM curve with patients at risk and censored, median & quartiles (95% CI)	Supportive (I-VI)

Endpoint	Method	Estimate	Level (Dataset)
	IPTW, Cox proportional hazards regression	IPTW, Cox proportional hazards regression	Supportive (I-VI)

Abbreviations: *CI, confidence intervals; HR, hazard ratio; IPTW, inverse probability of treatment weighting; PSM, propensity score matching*

In this study, measures taken to reduce potential bias include a pre-specified and detailed study protocol. Additionally, the propensity score (PS) balancing steps will be conducted independently and before availability/knowledge of the study outcomes (see section 7.7.2 for further details).

7.7.1. Descriptive statistics

Tabular summaries of baseline patient characteristics will be presented by treatment arm, including all characteristics specified in Section 7.4.3.1. Categorical variables will be summarized by showing the number and percentage (n, %), and continuous variables will be summarized using mean (standard deviation), median, and range (minimum and maximum) values. Standardized mean differences (SMD) will be calculated for means (continuous variables) and prevalence (dichotomous variables) of each covariate and used to assess imbalances in population characteristics between Study C1071003 patients and patients in the external control arm. We will present counts and standardized differences for both the unweighted and weighted samples (via IPTW, Fine stratification weighting, or overlap weighting) or unmatched and matched samples (via sensitivity propensity score matching [PSM] methodology). The SMD is preferred over p-values because of its robustness to sample size [33]. A standardized mean difference with a cutoff of ≥20% will be used to indicate an imbalance in the covariate between the treatment arms.

7.7.2. Propensity scores and inverse probability of treatment weighting (IPTW)

IPTW is a well-established method for confounding adjustment in non-randomized studies, which aims to eliminate the effect of confounding by observed baseline patient characteristics, improve covariate balance, and thereby obtain unbiased estimates of treatment effects. Estimating the propensity score (PS), on which IPTW relies, 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 an approach often used in non-randomized studies to create a pseudo-population in which the covariates are independent of the treatment assignment, thereby permitting an unbiased estimate of the average treatment effect (ATE), provided that fundamental assumptions of causal inference are satisfied.

PS will be estimated using logistic regression models, whereby the dependent variable is a binary indicator of the treatment arm (trial versus external control). Covariates in the logistic regression will include any relevant patient characteristic or condition that is unbalanced between the treatment arms (see section 7.4.3.1 for list of potential covariates) and will be selected according to prior literature, clinical experience and SMD (≥20%) estimated for

each covariate between treatment arms before and after weighting. Diagnostics of the estimated PS will be applied by examining the distribution of the PS in each treatment group to identify the degree of overlap and region of common support, including an assessment of positivity, overlap and balance [34]. The final specification of the PS model will be selected according to the best balance of covariates and fit.

The estimated PSs are used to generate IPTWs. For the estimation of ATE, the weights will be calculated as:

$$w_{ATE} = \frac{T}{PS} + \frac{(1-T)}{(1-PS)}$$

where *T* is treatment status (external control=0; treated=1). For ATE, the patients in the treatment arm receive a weight of 1/PS, and patients in the comparator arm are weighted by 1/(1 - PS) [35].

To reduce variability due to instability in estimation resulting from patients with large weights, stabilized weights will be calculated by multiplying the IPT-weights by the proportion of individuals in the treatment group (i.e., marginal probability of receiving the actual treatment received). Unlike unstabilized weights (T/PS, (1-T)/(1-PS)), stabilized weights are less likely to show extreme values. Extreme weight values are undesirable as they make the analysis dependent on very few individuals with very extreme probabilities. If extreme values are observed in the stabilized weights, the distribution will be truncated at the 5th and 95th percentiles (i.e., extreme values will be replaced rather than dropped from the analysis, alternatively 1st/99th dependent on model diagnostics).

As a sensitivity approach, IPTW will be used to estimate ATT (average treatment effects on the treated). For estimation of ATT, the weight will be calculated as:

$$w_{ATT} = T + \frac{PS(1-T)}{(1-PS)}$$

For ATT, patients within the treatment arm receive a weight of 1, such that the treated sample is used as the reference population against which the treated and control populations are standardized.

Once the final weighted sample is established, we will proceed with the main analyses. We will fit regression models based on the weighted samples. Standard errors and CI for the IPTW estimator will be obtained via a robust sandwich-type estimator of variance [47]. Variance estimates may alternatively be produced via bootstrapping (if feasible), which is a recommended error estimation procedure when fitting a Cox proportional hazards model in a sample weighted using IPTW [36].

7.7.3. Multiple imputations

Missing data can reduce statistical power and introduce bias in the effect estimates. Any missing data will be summarized as frequency (%) missing for each variable within each treatment arm and in total. Depending on the amount and reasons for missing data, a formal analysis to account for missing data may be conducted. When correctly implemented, multiple imputation procedures produce unbiased estimates and standard errors . If appropriate, multiple imputation by chained equations will be performed.

The complete case analysis will serve as a sensitivity approach for the TCR cohort (main). However, multiple imputation will be performed as the base case for address missing values of important baseline covariates under the assumption of MAR (missing at random). It will assume that there might be systematic differences between the missing and observed values of prognostic characteristics, but these can be entirely explained by other observed prognostic characteristics, the treatment variable, and the observed overall survival. If appropriate, multiple imputations by chained equations (MICE) will be performed using the fully conditional specification (FCS) method []. This approach imputes multivariate missing data on a variable-by-variable basis by specifying an imputation model for each variable. The technique is iterative and proceeds via Gibbs sampling if the initial joint distribution defined by the specified conditional distributions exists []. All comparative effectiveness main analyses will be repeated based on the multiply imputed data.

Within the multiple imputation model, we will include relevant baseline covariates to be used in estimating the propensity score. If applicable, auxiliary variables will be considered based on availability within the data. The imputation model will be run on the total sample of all included patients. Specific details may be further fine-tuned for computational efficiency and proportion convergence, such as the number of imputations, exclusion of a subset of predictors, and sample sizes for imputation, based on the data. Variables will be modeled according to their type; linear regression will be used for quantitative variables, multinomial logistic regression for non-ordered categorical variables, ordinal logistic regression for ordered variables, and logistic regression for binary variables. For any important prognostic variables with substantial levels of missing data (>20%), effect estimates may be calculated under assumptions of MNAR (missing not at random) to identify tipping points using delta adjustment [39, 40]. To reduce complexity and improve the interpretability of results, a maximum of 2-3 of the most important prognostic variables with large amounts of missing observation will be tested (to be determined based on the data).

As with other iterative procedures, diagnostics are essential to assess and evaluate the resulting imputation models obtained via MICE and to determine whether convergence has been achieved []. Visual assessment will be performed using worm plots, strip plots, and density plots to examine the observed and imputed data and compare their distributions. Convergence will be evaluated by plotting the mean and variance of each imputation run across iterations to confirm that there are no apparent trends. If convergence has not been achieved, the number of iterations will be increased until means and variances stabilize.

For each imputed dataset, the comparative analysis of PFS and OS will be conducted, and the individual estimates will be combined using Rubin’s rules [19, 44].

7.7.4. Comparative effectiveness

We will examine comparative effectiveness within the full analysis sets. The target estimand of interest for the primary comparative effectiveness analysis is the ATE, calculated via IPTW.

Each time-to-event outcome (main: OS, PFS; supportive: TTNT, TTD) will be compared between treatment arms using hazard ratios (HR) estimated from Cox proportional hazard models. We will also use the Kaplan-Meier (KM) method to visually examine survival probabilities over the follow-up period. We will report estimates and corresponding 95% CIs.

The proportional hazards assumption will be assessed through visual inspection of log-cumulative hazard plots. When the proportional hazards assumption is met, the log-cumulative hazard plot should show a constant HR over time. If moderate violations of the proportional hazards assumption are observed, HR will be interpreted as a weighted average of the HR over the follow-up period. 95% CIs will be obtained via bootstrapping (if feasible) [45]. If severe violations are observed, a restricted mean survival time model will be applied instead of the Cox proportional hazards regression model [46]. Standard errors and CI for the IPTW estimator will be obtained via a robust sandwich-type estimator of variance [47].

Because median OS likely will not be reached in Study C1071003, IPT-weighted survival probabilities will be described using KM estimator and compared between two treatment groups using the log-rank test [48, 49] at 12 and 15 months of follow-up.

7.7.5. Sensitivity analyses

To evaluate the robustness of results from the primary outcomes (main: OS, PFS), sensitivity analyses will be conducted.

7.7.5.1. Naïve comparison

A naïve comparison of each endpoint (PFS and OS) between treatment arms will be conducted, with estimating hazard ratios (95% CI) from Cox proportional hazards models unweighted and unadjusted for any potential covariates.

7.7.5.2. Multivariable Cox Regression

A comparison of each endpoint will be conducted via a multivariable Cox Regression model, unweighted but adjusted using all baseline covariates of interest (see section 7.4.3.1 for list of potential covariates).

7.7.5.3. Doubly robust comparison

A conditional average treatment effect (CATE) and its associated hazard ratio will be estimated using a semi-parametric approach described by Yadlowsky et al. [50, 51]. This estimator will provide a doubly robust comparison of PFS and OS between treatment arms. Standard errors and 95% confidence intervals are obtained for this estimator using the non-parametric bootstrap. For the TM-MM dataset, the method is implemented in R in the precmed package, using the atefitsurv function.

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 [52]. 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 [53].

Diagnostics will be applied to assess whether the doubly robust model appears to be well-specified [54].

7.7.5.4. Fine Stratification Weights

Fine stratification weighting does not use the PS directly to calculate weights; instead, propensity scores are used to create fine stratums, with the objective of balancing the observed variables between treatment (patients from Study C1071003 receiving elranatamab) and control units within each stratum. Stratums will be created based on the PS distribution of the treatment arm. This approach ensures assignment of all exposed individuals to stratums and minimises loss of information [55, 56]. Following stratification, targeting the ATE, weights for both exposed and reference patients in all stratums with at least one exposed patient and one reference patient are subsequently calculated based on the total number of patients within each stratum. Targeting ATT, weights for the treatment arm are set to 1 and patients in the control arm are reweighted based on the number of exposed patients residing within their stratum, so that patients in the control arm contribute proportionally to the relative number of total patients within a stratum.

7.7.5.5. Overlap Weights

This approach aims to make the distribution of covariates in the treated and reference group similar to each other and similar to the distribution in a subset of the overall study sample where patients are eligible to receive either the treatment of interest or the reference treatment, targeting the ATE in a subset of the overall population with some clinical equipoise [55]. Patients are weighted based on the predicted probability of receiving the opposite treatment., emphasizing the target population with the most overlap in observed characteristics between treatments, by continuously down-weighting the units in the tails of the PS distribution and up-weighting patients who have a substantial probability of receiving either treatment (specifically, patients with propensity scores of 0.5 make the largest contribution to the effect estimate and patients with propensity scores close to 0 and 1 make the smallest contribution) [57]. Extreme weights are impossible as weights are bound between 0 and 1 by design and, therefore, no truncation is necessary. Further, this weighting method yields exact covariate balance between treated and reference groups by construction.

7.7.5.6. Propensity score matching

To estimate the ATT effect, each patient from Study C1071003 will be matched to patients from an observational cohort using greedy nearest neighbor matching on 0.2 standarddeviation of the logit of the PS [585]. Previous studies showed that this caliper allowed the elimination of about 99% of the bias due to the measured confounders. Patients

with no match will be excluded from the analysis. Patients excluded from the analysis due to no match will be described separately in terms of their baseline characteristics, outcomes (unadjusted), and index treatment regimen. PS will be estimated as described above in Section 7.7.2. Descriptive statistics will be reported as described in Section 7.7.1. PFS and OS 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 [46].

7.7.5.7. Additional analyses

Additional analyses will make use of Datasets III-VI. IPTWs will be estimated as described in Section 7.7.2. Descriptive statistics will be reported as described in Section 7.7.1. PFS and OS will be compared between treatment arms using IPT-weighted Cox proportional hazards models. The proportional hazards assumption will be checked, and in case of deviations, a restricted mean survival time model (at 12 months) will be applied instead of the Cox proportional hazards regression model [46].

7.7.6. Quantitative Bias Analysis

Quantitative bias analysis is a general term for any method that quantitatively estimates the direction, magnitude, and uncertainty associated with systematic errors that influence measures of association [59]. When performing analysis on non-randomized data, there is always the possibility of unmeasured confounding. Prior to analysis, a short list of variables of suspected unmeasured confounders should be discussed with Pfizer; these will be variables that were observed in the trial but not available in the RWD sources. The E-value will be computed for each of the confounders identified in the short list for OS and PFS.

The E-value is the minimum strength of association that an unmeasured confounder would need to have with both the exposure and the outcome, conditional on the measured covariates, to fully explain away the observed treatment effect [60]. A large E-value implies considerable unmeasured confounding is required to nullify an effect estimate, whereas a small E-value indicates little unmeasured confounding is required to nullify an effect estimate.

7.7.7. Non-adherence due to discontinuation and dose modifications

Notably, the main comparative approach will utilize the ITT population as defined in Section 7.4.5. Under an alternative estimand, the treatment effect can be estimated with respect to a treatment strategy where patients in the active arm do not discontinue the initiated treatment. To do this, patients can be censored at the time of discontinuation, however under the assumption of censored at random (CAR), the censoring would only cause a reduction in the patients at risk at the time of the censoring event [61]. When the censoring is assumed to not be at random conditional on measured variables, and thereby dependent on the effects of the treatment initiated, a tipping point method can be used to assess the impact of the censoring on the observed treatment effect (PFS, OS). The analysis takes each probability estimated for the KM curve to the power of δ, a predefined factor [61]. After patients drop out of the study,

the δ is interpreted as an increased hazard, under the assumption that individuals who drop out have worse outcomes than the average in the group. The tipping point analysis thereby finds the smallest value of δ which worsens the KM cure of the treatment group so that the difference between the two treatment arms (elranatamab versus control) is no longer significant. In case the tipping point is not clinically reasonable, results are thereby deemed robust to the CAR assumption.

To account for dose modifications among patients in Study C1071003, given appropriate sample sizes, outcomes may be assessed after stratification of the Elranatamab arm based on relative dose intensity (RDI; low vs. high, with cut-off at 50%, 60%, 70%) . The stratification is conditional on the interpretability of the main results, subgroup analyses may be omitted.

7.7.8. Subgroup analysis

Subgroup analyses will be performed for each patient-relevant endpoint (OS, PFS, TTD, TTNT), conditional on a minimum 10 events per subgroup (main approach: IPTW-ATE) following execution of the main results. Conditional on the interpretability of the main results, subgroup analyses may be omitted.

For each endpoint and subgroup, the test for effect modification will be conducted to assess the interaction between treatment and the subgroup of interest. The test for effect modification will be conducted using an IPT-weighted Cox proportional hazards regression model for each subgroup of interest. If the test is significant at 5%, separate results for each subgroup will be presented after re-estimating IPT weights within each subgroup (using the pre-specified covariates, see section 7.4.3.1 for list of potential covariates) to account for potential differences in covariate distribution and treatment assignment patterns. The following subgroups and respective levels will be considered:

Mandatory Subgroups

- Age Group 1: <65 vs. ≥65 years
- Age Group 2: <75 vs. ≥75 years
- Gender: Male vs. Female
- Disease stage (ISS): 1-2 vs. 3

Other Pre-specified Subgroups

- Baseline cytogenetics: high risk vs. standard risk
- Prior stem cell transplant: yes vs. no
- Number of prior lines (≤5 vs >5)
- Type of myeloma (IgG vs. non-IgG vs. light chain only)
- Refractory to last therapy (yes vs. no)
- ECOG (0 vs 1-2)
- Penta-refractory (exposed) (yes vs. no)
- Baseline EMD status (yes vs. no)

Optional Subgroups

- BCMA exposure (yes vs. no)

- Type of prior BCMA therapy (ADC vs. CAR-T)

7.8. Data management

This study involves the use of RWD that exist in an anonymized structured format and contain no patient personal information. The TM-MM project is scientifically accompanied by Prof. PPD (German Myeloma Group) and Dr. PPD (University Hospital Leipzig). Anonymized patient-level data will be provided by TNXO to Cytel for analysis. Furthermore, the names of the study sites participating in TM-MM will be blinded to both Pfizer and Cytel. The Flatiron databases are compliant with both the spirit and the letter of the Health Insurance Portability and Accountability Act of 1996 (HIPAA). The databases meet the criteria for a limited-use dataset and contain none of the data elements prohibited by HIPAA for limited-use datasets. Anonymized patient-level data will be provided to IQVIA for analysis. Furthermore, the names of the study sites participating in Flatiron Health will be blinded to both Pfizer and IQVIA. Pfizer will not receive patient-level data during this study.

This study will further use structured databases from Study C1071003, an open-label, multi-center, non-randomized Phase 2 study of elranatamab (PF-06863135) monotherapy.

7.9. 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. Statistical programming code and summary output will be reviewed by the study team, including a biostatistician, for accuracy and completeness. Final deliverables will be reviewed and verified by a second, independent analyst. All quality checks will be documented.

For the secondary data collected by the TM-MM Germany project, the completeness and plausibility checks have been carried out in three stages during the generation of the dataset include:

a. Online check during data entry (incomplete and non-plausible data entries trigger an error message)

b. Central individual review after completion of data entry by the TNXO team of clinical monitors (in-complete or non-plausible data entry triggers a query process).

c. Central review of data sets by the TNXO team of data analysts to find and exclude any duplicated patients

After the duplicate records are excluded, only complete and plausible records are included in the database. Once the TM-MM dataset is provided by TNXO to Cytel and Pfizer approves data access for Study C1071003, the study data management will adhere to pre-defined process guidelines, which mainly consist of data validation based on computer-assisted checking of variables/values. In the RW dataset, patients with any implausible/counter-intuitive data will be excluded from the sample during the selection of the study population

or for the analysis of specific outcomes, as applicable. The study team will maintain adequate and accurate records to enable the conduct of the study to be fully documented.

7.10. Limitations of the research methods

As the basis for the data collection of the TM-MM dataset is via a retrospective medical chart review and the basis for the data collection of the Flatiron Health dataset is via electronic health records, this study will be subject to certain limitations inherent in retrospective reviews of medical records, including:

a) TM-MM dataset: Patients selected for study inclusion will represent a "convenience" sample in that the records will be obtained from physicians and study sites that are willing to participate in the study. However, the data is collected from a representative sample of sites with respect to both regions and healthcare structures across Germany and therefore is expected to be representative of the wider German population. Moreover, Germany presents relatively uniform healthcare regulations across the nation.

b) Unlike clinical trial settings with specific definitions of study outcomes and scheduled assessments, the assessment of refractoriness and progression to therapies in RW clinical practice settings may not be made consistently across patients and across physicians. Specifically, in RW observational studies, especially those performed retrospectively, it is not possible to implement consistent monitoring and application of homogenous evaluation criteria (e.g., IMWG) that are possible in clinical trial designs. As such, in RW patients, depending on how often clinical assessments are made, the date of disease progression is more likely to be diagnosed later than it would be diagnosed if patients had scheduled assessments similar to clinical study settings [62, 63]. This may result in longer PFS and bias the comparative effectiveness estimates in favor of the RW arms. Also, 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 and assessed for treatment response and/or progression at more frequent intervals (Haut and Pronovost, 2011). Of note, for the TM-MM dataset, refractoriness and progression are reliant on proxies (i.e., refractoriness as exposure; progression as start of the next line of therapy. As progression for the TM-MM dataset will be based on a proxy of initiation of next treatment line and not accurate (clinical/biochemical) dates of progression, this may result in a longer PFS and bias the comparative effectiveness estimates in favor of the RW arm. To account for these differences, PFS will additionally be compared using the proxy definition in both arms for TM-MM.

c) As the data collection is linked to chart reviews, a degree of missing data is expected as only data reported in patient charts can be collected. Data that can be collected is limited to the available information in patient charts accessible to the center that is reporting on behalf of the patient, and as such, the availability of information in records may vary by physician practice and may reflect differences in practice

patterns, recording practices, and medical norms. Furthermore, as TM-MM is an ongoing data collection initiative, the eCRF has been modified at different timepoints to include further variables. As such, some variables were not available for collection by patients documented at earlier time points (i.e., response variables, EMD).

d) Due to data availability, in the TM-MM dataset, data on response was only added to the eCRF and subsequently captured as of Q2 2022 onwards at the end of each line of therapy. As such, it will not be assessable in this study.

To address and/or reduce the impact of potential bias and improve exchangeability between the trial and the external control arms [24, 64, 65], 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.

8. PROTECTION OF HUMAN SUBJECTS

8.1. Patient information

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

8.2. 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.

8.3. Patient Withdrawal

Not applicable.

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

IRB review is not required for this study as it uses de-identified and anonymized secondary data sources available within patient medical records. TM-MM was set up in Germany as health services research, collecting retrospective data that is already available at the surveyed centers within the framework of the patient files. In Flatiron Health, 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”.

8.5. 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 Good Pharmacoepidemiology Practices issued by the International Society for Pharmacoepidemiology, the Good Outcomes Research Practices issued by ISPOR (formerly known as the International Society for Pharmacoeconomics and Outcomes Research) and the International Ethical Guidelines for Epidemiological Research issued by the Council for International Organizations of Medical Sciences.

9. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

This study is retrospective and involves data that exist as structured data by the time of study start or a combination of existing structured data and unstructured data. In these data sources, individual patient data are not retrieved or validated, and it is not possible to link (i.e., identify a potential association between) a particular product and medical event for any individual. Thus, the minimum criteria for reporting an AE (i.e., identifiable patient, identifiable reporter, a suspect product, and event) cannot be met.

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

10. 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 party responsible 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.

11. REFERENCES

1. RKI, *Multiples Myelom. Krebs in Deutschland für 2017/2018*. 2021, Zentrum für Krebsregisterdaten und Gesellschaft der epidemiologischen Krebsregister in Deutschland e. V.: Berlin. p. 138.
2. Cherry, B.M., et al., *Evolving therapeutic paradigms for multiple myeloma: back to the future.* Leuk Lymphoma, 2013. **54**(3): p. 451-63.
3. Jagannath, S., et al., *Real-world treatment patterns and associated progression-free survival in relapsed/refractory multiple myeloma among US community oncology practices.* Expert Rev Hematol, 2016. **9**(7): p. 707-17.
4. Kumar, S.K., et al., *Continued improvement in survival in multiple myeloma: changes in early mortality and outcomes in older patients.* Leukemia, 2014. **28**(5): p. 1122-8.
5. *Leitlinienreport S3-Leitlinie Diagnostik, Therapie und Nachsorge für Patienten mit monoklonaler Gammopathie unklarer Signifikanz (MGUS) oder Multiplem Myelom*. AWMF Online.
6. Piechotta, V., et al., *Patients With Multiple Myeloma or Monoclonal Gammopathy of Undetermined Significance.* Dtsch Arztebl Int, 2022. **119**(14): p. 253-260.
7. Nooka, A.K., et al., *Treatment options for relapsed and refractory multiple myeloma.* Blood, 2015. **125**(20): p. 3085-99.
8. Yong, K., et al., *Multiple myeloma: patient outcomes in real-world practice.* Br J Haematol, 2016. **175**(2): p. 252-264.
9. Pfizer, *Pfizer Initiates Pivotal Phase 2 MagnetisMM-3 Trial of BCMA-CD3 Bispecific Antibody Elranatamab (PF-06863135) in Multiple Myeloma* 2021: Internet.
10. Pfizer, *Pfizer Presents Updated Favorable Elranatamab Data from Pivotal Phase 2 MagnetisMM-3 Trial* 2022.
11. Mohty, M. *Extended Follow-up Data from MagnetisMM-3 With Elranatamab in R/R Myeloma*. in *European Hematology Association Congress*. 2023. Frankfurt, Germany.
12. Pfizer, *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*. 2022.
13. Gandhi, U.H., et al., *Outcomes of patients with multiple myeloma refractory to CD38-targeted monoclonal antibody therapy.* Leukemia, 2019. **33**(9): p. 2266-2275.
14. Jagannath, S., et al., *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, 2020. **38**(15_suppl): p. 8525-8525.
15. Mateos, M.-V., et al., *LocoMMotion: a prospective, non-interventional, multinational study of real-life current standards of care in patients with relapsed and/or refractory multiple myeloma.* Leukemia, 2022. **36**(5): p. 1371-1376.

16. Hatswell, A.J., et al., *Regulatory approval of pharmaceuticals without a randomised controlled study: analysis of EMA and FDA approvals 1999-2014.* BMJ Open, 2016. **6**(6): p. e011666.
17. Feinberg, B.A., et al., *Use of Real-World Evidence to Support FDA Approval of Oncology Drugs.* Value Health, 2020. **23**(10): p. 1358-1365.
18. Wang, X., et al., *Current perspectives for external control arms in oncology clinical trials: Analysis of EMA approvals 2016–2021.* Journal of Cancer Policy, 2023. **35**: p. 100403.
19. Anderson, M., et al., *A review of NICE appraisals of pharmaceuticals 2000-2016 found variation in establishing comparative clinical effectiveness.* J Clin Epidemiol, 2019. **105**: p. 50-59.
20. Borchert, K., et al., *PNS204 THE RELEVANCE OF REAL WORLD EVIDENCE STUDIES IN THE GERMAN BENEFIT ASSESSMENT (AMNOG) PROCESS.* Value in Health, 2019. **22**: p. S795.
21. (IQWiG), I.f.Q.u.W.i.G., *A19-43 Konzepte zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung von Arzneimitteln nach §35a SGB V* 2019.
22. (EMA), E.M.A. and *ICH E10 Choice of control group in clinical trials* 2001: Internet.
23. Carroll, N., *Application of Propensity Score Models in Observational Studies*. 2018: SAS Glob Forum.
24. Ghadessi, M., et al., *A roadmap to using historical controls in clinical trials - by Drug Information Association Adaptive Design Scientific Working Group (DIA-ADSWG).* Orphanet J Rare Dis, 2020. **15**(1): p. 69.
25. Chesnaye, N.C., et al., *An introduction to inverse probability of treatment weighting in observational research.* Clin Kidney J, 2022. **15**(1): p. 14-20.
26. Zhang, Y., et al., *Comparative Effectiveness Research Based on Observational Data to Emulate a Target Trial (CERBOT).* 2019.
27. Hernán, M.A. and J.M. Robins, *Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available.* Am J Epidemiol, 2016. **183**(8): p. 758-64.
28. FDA, *Framework for FDA's Real-World Evidence Program.* . 2018.
29. Palumbo, A., et al., *Revised International Staging System for Multiple Myeloma: A Report From International Myeloma Working Group.* J Clin Oncol, 2015. **33**(26): p. 2863-9.
30. Merz, M., et al., *Adjusted Comparison of Outcomes between Patients from CARTITUDE-1 versus Multiple Myeloma Patients with Prior Exposure to PI, Imid and Anti-CD-38 from a German Registry.* Cancers (Basel), 2021. **13**(23).
31. Xinran, Ma.,et al., *Comparison of Population Characteristics in Real-World Clinical Oncology Databases in the US: Flatiron Health, SEER, and NPCR.* medRxiv 2020.03.16.20037143#
32. Krishnan A, Nooka AK, Chari A, Garfall AL, Martin TG, Nair S, Lin X, Qi K, Londhe A, Pei L, Ammann E, Kobos R, Smit J, Parekh T, Marshall A, Slavcev M, Usmani SZ. Teclistamab versus real-world physician's choice of therapy in triple-m class exposed relapsed/refractory multiple myeloma. J Comp Eff Res. 2023 Jun;12(6):e220186.

33. Austin, P.C., *Using the Standardized Difference to Compare the Prevalence of a Binary Variable Between Two Groups in Observational Research.* Communications in Statistics - Simulation and Computation, 2009. **38**(6): p. 1228-1234.
34. Petersen, M.L., et al., *Diagnosing and responding to violations in the positivity assumption.* Stat Methods Med Res, 2012. **21**(1): p. 31-54.
35. Benedetto, U., et al., *Statistical primer: propensity score matching and its alternatives.* Eur J Cardiothorac Surg, 2018. **53**(6): p. 1112-1117.
36. Austin, P.C., *Variance estimation when using inverse probability of treatment weighting (IPTW) with survival analysis.* Stat Med, 2016. **35**(30): p. 5642-5655.
37. White, I.R., P. Royston, and A.M. Wood, *Multiple imputation using chained equations: Issues and guidance for practice.* Statistics in Medicine, 2011. **30**(4): p. 377-399.
38. Sterne, J.A.C., et al., *Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls.* BMJ, 2009. **338**: p. b2393.
39. van Buuren, S. and K. Groothuis-Oudshoorn, *mice: Multivariate Imputation by Chained Equations in R.* Journal of Statistical Software, 2011. **45**(3): p. 1 - 67.
40 van Buuren, S., *Multiple imputation of discrete and continuous data by fully conditional specification.* Stat Methods Med Res, 2007. **16**(3): p. 219-42.
41. Leacy, F.P., et al., *Analyses of Sensitivity to the Missing-at-Random Assumption Using Multiple Imputation With Delta Adjustment: Application to a Tuberculosis/HIV Prevalence Survey With Incomplete HIV-Status Data.* American Journal of Epidemiology, 2017. **185**(4): p. 304-315.
42. Cro, S., et al., *Sensitivity analysis for clinical trials with missing continuous outcome data using controlled multiple imputation: A practical guide.* Statistics in Medicine, 2020. **39**(21): p. 2815-2842.
43. van Buuren, S., *Flexible Imputation of Missing Data*. 2012, New York: Chapman and Hall/CRC.
44. Marshall, A., et al., *Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines.* BMC Med Res Methodol, 2009. **9**: p. 57.
45. Stensrud, M.J. and M.A. Hernán, *Why Test for Proportional Hazards?* Jama, 2020. **323**(14): p. 1401-1402.
46. Zhao, H. and A.A. Tsiatis, *Efficient estimation of the distribution of quality-adjusted survival time.* Biometrics, 1999. **55**(4): p. 1101-7.
47. Lin, D.Y. and L.J. Wei, *The Robust Inference for the Cox Proportional Hazards Model.* Journal of the American Statistical Association, 1989. **84**(408): p. 1074-1078.
48. Xie, J. and C. Liu, *Adjusted Kaplan-Meier estimator and log-rank test with inverse probability of treatment weighting for survival data.* Stat Med, 2005. **24**(20): p. 3089-110.
49. Statement, S.H.C.W. Internet.
50. Yadlowsky, S., et al., *Estimation and Validation of Ratio-based Conditional Average Treatment Effects Using Observational Data.* Journal of the American Statistical Association, 2021. **116**(533): p. 335-352.

51. *CRAN - Package precmed* Internet.
52. Funk, M.J., et al., *Doubly robust estimation of causal effects.* Am J Epidemiol, 2011. **173**(7): p. 761-7.
53. Bang, H. and J.M. Robins, *Doubly robust estimation in missing data and causal inference models.* Biometrics, 2005. **61**(4): p. 962-73.
54. Austin, P.C. and E.A. Stuart, *Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies.* Stat Med, 2015. **34**(28): p. 3661-79.
55. Desai R J, Franklin J M. *Alternative approaches for confounding adjustment in observational studies using weighting based on the propensity score: a primer for practitioners.* BMJ 2019; 367 :l5657 doi:10.1136/bmj.l5657.
56. Desai RJ, Rothman KJ, Bateman BT, Hernandez-Diaz S, Huybrechts KF. *A Propensity-score-based Fine Stratification Approach for Confounding Adjustment When Exposure Is Infrequent*. Epidemiology. 2017 Mar;28(2):249-257. doi: 10.1097/EDE.0000000000000595. PMID: 27922533; PMCID: PMC5497217.
57. Fan Li, Laine E Thomas, Fan Li. *Addressing Extreme Propensity Scores via the Overlap Weights*, American Journal of Epidemiology, Volume 188, Issue 1, January 2019, Pages 250–257, https://doi.org/10.1093/aje/kwy201.
58. Jagannath, S., et al., *KarMMa-RW: comparison of idecabtagene vicleucel with real-world outcomes in relapsed and refractory multiple myeloma.* Blood Cancer J, 2021. **11**(6): p. 116.
59. Lash, T.L., et al., *Quantitative Bias Analysis in Regulatory Settings.* American Journal of Public Health, 2016. **106**(7): p. 1227-1230.
60. VanderWeele, T.J. and P. Ding, *Sensitivity Analysis in Observational Research: Introducing the E-Value.* Annals of Internal Medicine, 2017. **167**(4): p. 268-274.
61. Lipkovich, I., B. Ratitch, and M. O'Kelly, *Sensitivity to censored-at-random assumption in the analysis of time-to-event endpoints.* Pharm Stat, 2016. **15**(3): p. 216-29.
62. Chakraborty R, Liu HD, Rybicki L, Tomer J, Khouri J, Dean RM, et al. *Progression with clinical features is associated with worse subsequent survival in multiple myeloma.* Am J Hematol [Internet]. 2019 Apr 1 [cited 2023 Mar 20];94(4):439–45. Available from: https://pubmed.ncbi.nlm.nih.gov/30663805/.
63. Goldman-Mazur S, Visram A, Kapoor P, Dispenzieri A, Lacy MQ, Gertz MA, et al. *Outcomes after biochemical or clinical progression in patients with multiple myeloma.* Blood Adv [Internet]. 2023 Mar 28 [cited 2023 Mar 20];7(6):909–17. Available from: https://ashpublications.org/bloodadvances/article/7/6/909/484885/Outcomesafter-biochemical-or-clinical-progression. 64. Hatswell, A., et al., *Summarising salient information on historical controls: A structured assessment of validity and comparability across studies.* Clin Trials, 2020. **17**(6): p. 607-616.
65. Pocock, S.J., *The combination of randomized and historical controls in clinical trials.* J Chronic Dis, 1976. **29**(3): p. 175-88.

12. LIST OF TABLES

13. LIST OF FIGURES

ANNEX 1. LIST OF STAND-ALONE DOCUMENTS

Document Date	Document Name
24 February 2023	20230224_Pfizer_CDN RWD Study_Protocol_final_v2.0_clean
25 March 2023	20230324 Pfizer_CDN RWD Study_Results Report_v1.0
14 April 2022	C1071003_SAP_v7.0_14Apr2022
29 July 2022	C1071003_PA9_clean_29July2022

ANNEX 2. ADDITIONAL INFORMATION

A – OPERATIONAL DEFINTIONS IN TM-MM

Table 9. Operational definitions of descriptive variables for TM-MM

Variable	Operational definition
Demographics	
Age	Continuous; age at index (in years) and categorical as per <65 years, 65-74 years, or ≥75 years at index
Sex	Categorical; female, male
Height	Continuous; height in meters (m) at index
Weight	Continuous; weight in kilograms (kg) at index
Body mass index (BMI)	Continuous; calculated from height and weight as kg/m^2
Practice type	Categorical; institution type in which induction therapy of index LOT was conducted, including university/academic hospital, non-academic hospital, private hospital (clinic MVZ), primary care/specialist practice, outpatient specialist center (practice MVZ), or unknown
Comorbidity profile	
Charlson Comorbidity Index (CCI) score	CCI score reported as a continuous variable as well as a categorical variable for the following score categories: 0, 1, 2, 3, 4, ≥5 assessed based on documented comorbidities at index *See ANNEX 2B CCI derivation using TM-MM data parameters (Table 10)*
Cardiovascular diseases	Categorical; Yes/No determined based on the presence of any of the following at index: coronary heart disease, cardiac insufficiency, and cardiac arrhythmia
Hypertension	Categorical; Yes/No assessed at index
Diseases of the lung (chronic obstructive pulmonary disease)	Categorical; Yes/No assessed at index
Renal impairment	Categorical; Yes/No assessed at index
Hepatic impairment	Categorical; Yes/No assessed at index
Diabetes	Categorical; Yes/No assessed at index
Depression or other psychiatric conditions	Categorical; Yes/No assessed at index
Clinical/disease characteristics	
Time since diagnosis	Continuous; defined as time from the date of initial MM diagnosis and index in months
Symptomatology	Categorical; Yes/No assessed at index for each symptom: Pain in the central area of the vertebral column, Pain in the lower area of the vertebral column, Bone pains, Fractures, Performance loss, Fatigue, Weakness, Bacterial infections
ISS	Categorical; I, II, III, Unknown/not assessed at index.
R-ISS	Categorical; I, II, III, Unknown/not assessed at index. *If missing, R-ISS will be calculated as follows [29, 30]:* *If ISS = I and LDH is not above normal range (low)/missing and no cytogenetic abnormalities were detected/not tested, R-ISS = I* *If ISS = III and LDH is above normal range (high) or cytogenetic abnormalities are detected, then R-ISS = III* *If ISS = III and LDH is missing and cytogenetic abnormalities were not detected/not tested, then R-ISS = III* *Else, R-ISS = II*

Variable	Operational definition
	Note: Variable available towards end of the study period and is not well populated, ISS preferred
Derived ISS stage	Categorical; derived ISS stage assessed at index according to the following: a) Stage I: Patients with serum β2M < 3.5 mg/L; serum albumin ≥3.5 g/dL b) Stage II: Patients with serum β2M < 3.5 mg/L; serum albumin < 3.5 g/dL; OR serum β2M 3.5 to 5.5 mg/L irrespective of serum albumin c) Stage III: Patients with serum β2M > 5.5 mg/L
M-protein type	Categorical; Immunoglobulin (Ig) G, non-IgG (including IgA, IgD, IgE, IgG, IgM), light chains only, and Unknown/not assessed.
ECOG	Categorical; assessed at index as 0, 1, 2, 3, 4, unknown/not assessed.
SLiM-CRAB Criteria	Categorical; Yes/No for each criteria assessed at index including: a) Hypercalcemia (C), > 2,75 mmol/l (> 10,5 mg/dl) or > 0,25 mmol/l higher than the upper norm b) Creatinine level (R), > 2,0 mg/dl (> 173 mmol/l) c) Anemia (A), Hemoglobin < 10,0 g/l or > 2,0 g/l below the lower norm d) Bone involvement (B), osteolyses or osteopenia with compression fractures e) Bone marrow infiltration (S): clonal plasma cell content in the bone marrow > 60% (cytological and histological). f) Free light chains (Li): free light chain quotient in serum > 100 (affected/unaffected light chains). g) Focal lesions on MRI (M): > 1 focal lesion > 1 cm on MRI imaging *Note: SLiM variables available from Q4 2017 onwards*
Extramedullary disease (EMD)	Categorical; Yes/No assessed at index *Note: EMD variable available from Q4 2017 onwards*
Refractory (exposure) status	Categorical; triple-, quad-, penta-refractory at index based on the number of individual IMiD, PI, and anti-CD38 treatments patients were previously exposed to. Quad-exposed, is defined as exposure to at least 2 IMiDs, or 2 PIs, and a CD38; Penta-exposed is defined as exposure to lenalidomide, pomalidomide, bortezomib, carfilzomib, and daratumumab.
Cytogenetics (20, 21)	Categorical; high risk, standard risk, unknown/not assessed based on the following: a) High-risk cytogenetics will be identified based on the presence of at least one of the following markers at index documented in categorical yes/no variables for pre-specified mutations[4] or other (open-text response): Del 17p (or Del17), t(4:14), t(14:16) b) Standard-risk cytogenetics will be identified based on the absence of any high-risk-related markers at index, and the presence of at least one of the markers not used to assess high-risk cytogenetics c) Unknown/not assessed category: patients not tested or unknown entries

[4] TM-MM dataset currently captures whether the following mutations are present (yes/no): Del13, Del17, t(4:14), t(p16:q32), t(11:14), t(6:14), t(14;16), t(14:20), amplification 1q21, hypodiploidy, hyperdiploidy, other (free-text)

Variable	Operational definition
Minimal residual disease (MRD) status at time of best response	Categorical; positive, negative or unknown/not assessed *Note: Variable available from Q4 2018 onwards*
Laboratory tests/results	
Serum M-protein assessment	Categorical; Yes/No based on proportion of patients with a recorded Serum M-protein measurement at index.
Serum M-protein value	Continuous; serum M-protein level in g/dL at index among patients with an assessment
20/20/20 Risk Model: Free light chain (FLC) ratio >20	Categorical; Yes/No assessed at index *Note: Variable available from Q4 2020 onwards and was not well populated*
20/20/20 Risk Model: M-Protein level in serum >20 g/L	Categorical; Yes/No assessed at index *Note: Variable available from Q4 2020 onwards and was not well populated*
20/20/20 Risk Model: Plasma cells in bone marrow >20%	Categorical; Yes/No assessed at index *Note: Variable available from Q4 2020 onwards and was not well populated*
Albumin assessment	Categorical; Yes/No based on proportion of patients with a recorded albumin measurement at index.
Albumin level value	Categorical; Yes/No serum albumin level of <3.5g/dL, or unknown/not assessed at index
Beta-2 microglobulin (β2m) in serum assessment	Categorical; Yes/No based on proportion of patients with a serum β2m measurement at index.
β2m in serum value	Categorical; serum β2m levels in mg/L <3.5 mg/L, 3.5-5.5 mg/L, >5.5 mg/L, or unknown/not assessed at index
MM treatment history	
Year of index date	Categorical; 2016, 2017, 2018, 2019, 2020, 2021, 2022
Duration of follow-up from index	Continuous; defined as time from the index until the earliest of death, loss to follow-up, of end of the study period in months
Time to next treatment on the pre-index LOT (last LOT prior to the index LOT)	Continuous; assessed for the last LOT prior to the index and defined as time from start of the pre-index LOT and next treatment (ie. start of next line of therapy = index LOT) in months
Time from initial MM diagnosis to index date	Continuous; defined as time from initial MM diagnosis until the day before the index in months
Total number of LOTs received (all time)	Categorical; 1, 2, 3, 4 etc., based on number of LOTs from initial MM diagnosis to the censor date
Number of LOTs prior to the index date	Categorical; 0, 1, 2, 3, 4, etc.
Number of LOTs after the index date	Categorical: 0, 1, 2, 3, 4, etc., including the index LOT to the censor date
Prior agents & classes	Categorical; Yes/No based on the proportion of patients that have received an agent or class, per agent and class in any line prior to the index LOT and after initial MM diagnosis

Variable	Operational definition
	See ANNEX 2C for list of MM treatments of interest (Table 12)
Prior SCT	Categorical; Yes/No assessed at index based on any SCT (allogeneic or autologous) between the period of the day before the index date and initial MM diagnosis
Treatment patterns[5]	
Any SCT	Categorical; Yes/No assessed as any SCT (allogeneic or autologous) carried out on the index LOT, pre-index LOT, or post-index LOT
Autologous SCT	Categorical; Yes/No assessed as an autologous SCT carried out on the index LOT, pre-index LOT, or post-index LOT
Allogeneic SCT	Categorical; Yes/No assessed as an allogeneic SCT carried out on the index LOT, pre-index LOT, or post-index LOT
Treatments received by LOT (agents & classes)	Categorical; Yes/No per agent and class in the LOT before the index LOT (pre-index LOT), in the index LOT, and the line after the index LOT (post-index LOT) received as per any therapeutic measure[6] *See ANNEX 2C for list of MM treatments of interest (Table 12)*
Systemic therapy for tumor reduction (induction therapy) regimens by LOT	Categorical; proportion of patients receiving systemic therapy for tumor reduction and the most common regimens in the pre-index LOT, the index LOT and post-index LOT (data-driven list[7] of most common regimen combinations) *See ANNEX 2C for list of MM treatments of interest (Table 12)*
Maintenance therapy regimens by LOT	Categorical; proportion of patients receiving maintenance therapy and the most common maintenance therapy regimens in the pre-index LOT, the index LOT and post-index LOT (data-driven list of most common regimen combinations) *See ANNEX 2C for list of MM treatments of interest (Table 12)*
High-dose chemotherapy regimens by LOT	Categorical; proportion of patients receiving a high-dose chemotherapy regimen and the most common high-dose chemotherapy regimens in the pre-index LOT, the index LOT and post-index LOT (data-driven list of most common regimen combinations) *See ANNEX 2C for list of MM treatments of interest (Table 12)*
Therapy for stem-cell mobilization (regimens) by LOT	Categorical; proportion of patients receiving therapy for stem-cell mobilization an the most common regimens for stem-cell mobilization in the pre-index LOT, the index LOT and post-index LOT (data-driven list of most common regimen combinations) *See ANNEX 2C for list of MM treatments of interest (Table 12)*

[5] Treatment patterns for the Trial Cohort will be described in the index LOT only; for the TCE cohort, treatment patterns will be described in the pre-index LOT, index-LOT, and post-index LOT

[6] Therapeutic measures include systemic therapy for tumor reduction, chemotherapy for stem cell mobilization, high-dose chemotherapy, consolidation therapy, maintenance therapy

[7] Data-driven: the most frequent combinations (regimens) as observed in the dataset. Regimens combinations with <5 patients will be grouped as "Other"

Variable	Operational definition
Consolidation therapy regimens after SCT by LOT	Categorical; proportion of patients receiving consolidation therapy and the most common consolidation therapy regimens in the pre-index LOT, the index LOT and post-index LOT (data-driven list of most common regimen combinations) *See ANNEX 2C for list of MM treatments of interest (Table 12)*
Supportive therapy by LOT	Categorical; based on proportion of patients receiving supportive therapy in the pre-index LOT, the index LOT and post-index LOT

Abbreviations: *β2M, beta-2-microglobulin; BMI, Body mass index; CCI, Charlson Comorbidity Index; ECOG, Eastern Cooperative Oncology Group; EMD, extramedullary disease; FLC, free light chain; GP, general practitioner; IMiD, immunomodulator; ISS, International Staging System; LDH, lactate dehydrogenase; LOT, line of therapy; MM, multiple myeloma; MRD, minimal residual disease; MRI, magnetic resonance imaging; PI, proteasome inhibitor; Q, quarter; SCT, stem cell transplant; SLiM-CRAB, criteria involving the following: 60% or more clonal plasma cells (S), light chains (Li), and MRI (M); also elevated calcium levels (C), renal failure (R), anemia (A), and bone lesions (B); TCE, triple-class exposed*

B – CHARLSON COMORBIDITY INDEX (CCI)

Table 11: Comorbidities and their assigned weights for estimation of the CCI in Flatiron Health

CCI comorbidities	Assigned Weight
Myocardial infarction, Congestive heart failure, Peripheral vascular disease, Cerebrovascular disease, Dementia, Chronic pulmonary disease, Rheumatic disease, Peptic ulcer disease, Mild liver disease,	1
Hemiplegia or paraplegia, Renal disease, Diabetes with chronic complications, any malignancy (including leukemia and lymphoma)	2
Moderate or severe liver disease	3
Metastatic solid tumor, AIDS/HIV	6
Note: Some CCI comorbidities will not be present in the RW TCR cohort due to study exclusion criteria Abbreviations: *AIDS: acquired immune deficiency syndrome; HIV: human immunodeficiency virus*	

See Table 12 in ANNEX 2B for the list of applicable diagnoses in the Flatiron Health database.

C – THERAPEUTIC AGENTS AVAILABLE CAPTURED IN TM-MM/FLATIRON HEALTH

Table 12: Agents of interest to be derived from TM-MM/Flatiron Health

Corticosteroids
Dexamethasone
Prednisone / Prednisolone
Chemotherapy
Adriamycin
Bendamustine
Cyclophosphamide
Doxorubicin
Idarubicine
Melphalan
Vincristine
Proteasome Inhibitors (PI)
Bortezomib
Carfilzomib
Ixazomib
Immunomodulatory agents (IMiDs)
Lenalidomide
Pomalidomide
Thalidomide
Monoclonal antibodies
Anti-CD38
Daratumumab
Isatuximab
Elotuzumab
Belantamab mafodotin
Other
Selinexor
Panobinostat
Supportive therapy
Free-text

Note: *BCMA-directed therapy is defined as Belanatamab mafodotin or CAR-T cell therapy with Idecabtagene vicleucel or Ciltacabtagene autoleuce*