Fitness for purpose of data sources relevant for real-world data (RWD) studies on CAR-T cell therapy

FWC/EMA/2020/46/TDA/L5.07 – SC01

Objective 3

Protocol / Study Number	EMA/39169/2024
Protocol Date	05 August 2025
Version Number	1.0
Amendment	Not applicable
Therapeutic Area	CAR-T cell therapy
	Certara, Real World Evidence & Modelling Solutions Artak Khachatryan - Responsible Senior Investigator Nadia Quignot - Overall Project Coordinator Stephanie Read - Lead Epidemiologist Giancarlo Pesce - Epidemiologist Kazue Kikuchi - Project Manager **EMA** Carla Jonker (Carla.Jonker@ema.europa.eu) - Project Lead Alexandra Pacurariu (Alexandra.Pacurariu@ema.europa.eu) - Pharmacoepidemiologist

Contents

List of tables

List of figures

Study Glossary

Acronym/Short Name	Meaning
/L	per litre
24h	24 hours
A&E	Accident & Emergency
AE	adverse event
AIDS	acquired immune deficiency syndrome
AL amyloidosis	amyloid light-chain amyloidosis
ALD	affection de longue durée
ALT	alanine aminotransferase
ANC	absolute neutrophil count
anti-HBc	hepatitis B core antibodies
anti-HBs	hepatitis B surface antibody
anti-HCV	hepatitis C antibody
ASCO	American Society of Clinical Oncology
AST	aspartate aminotransferase
B-ALL	B-cell acute lymphoblastic leukaemia
BCL	B-cell lymphoma
BCL-2	B-cell lymphoma 2
BCMA	B-cell maturation antigen
CABG	coronary artery bypass graft
CAR-T	chimeric antigen receptor T-cell
CBR	Clinical Benefit Rate
CCI	Charlson Comorbidity Index
CepiDc	Centre d'épidémiologie sur les causes médicales de décès
CNIL	National Commission for data protection and freedom of information
CNS	central nervous system
COPD	chronic obstructive pulmonary disease
CR	cancer registration
CRS	cytokine release syndrome
CSF	colony-stimulating factor
DARS	Data Access Request Service
DESCAR-T	Dispositif d'Enregistrement et Suivi des patients traités par CAR-T cells
DLBCL	diffuse large B-cell lymphoma
DMSO	dimethyl sulfoxide
DOR	duration of response
DPd	daratumumab, pomalidomide and dexamethasone
ECOG	Eastern Cooperative Oncology Group

Acronym/Short Name	Meaning
EMA	European Medicines Agency
ENCePP	European Network of Centres for Pharmacoepidemiology & Pharmacovigilance
EU	European Union
FEV1	forced expiratory volume in 1 second
g/dL	gram per decilitre
G-CSF	granulocyte colony-stimulating factor
GDPR	General Data Protection Regulation
GPP	Good Pharmacoepidemiology Practices
GVHD	graft-versus-host disease
HBV-DNA	hepatitis B virus deoxyribonucleic acid
HCV	hepatitis C virus
HCV-RNA	hepatitis C virus ribonucleic acid.
HDAC	histone deacetylase inhibitor
HES	Hospital Episode Statistics
HIV	human immunodeficiency virus
HR	hazard ratio
HRA	Health Research Authority
ICD-9	International Classification of Diseases, 9th Revision
ICD-10	International Classification of Diseases, 10th Revision
ICD-O	International Classification of Disease of Oncology
ICF	informed consent form
IMiD	immunomodulatory imide drugs
IMWG	International Myeloma Working Group
ISPE	International Society for Pharmacoepidemiology
LOT	lines of therapy
LVEF	left ventricular ejection fraction
LYSARC	Lymphoma Academic Research Organisation
m^2	square metre
mAb	monoclonal antibody
MCL	mantle cell lymphoma
MDRD	Modified Diet in Renal Disease
mg	milligramme
mg/24h	milligramme per 24 hours
mg/dL	milligramme per decilitre
mL/min	millilitre per minute
MM	multiple myeloma

Acronym/Short Name	Meaning
mmol/L	millimoles per litre
M-protein	monoclonal paraprotein
MRD	minimal residual disease
MUGA	multiple-gated acquisition
N	number
N/A	not applicable
NCRAS	National Cancer Registration and Analysis Service
NHS	National Health Service
NIS	non-interventional studies
NYHA	New York Heart Association
OPCS-4	Office of Population Censuses and Surveys Classification of Surgical Operations and Procedures, Version 4
ORR	overall response rate
OS	overall survival
PD	progressive disease
PFS	progression-free-survival
PI	proteasome inhibitor
POEMs syndrome	polyneuropathy, organomegaly, endocrinopathy, monoclonal protein, and skin changes
PMSI	Programme de médicalisation des systèmes d'information
PR	partial response
PVd	pomalidomide, bortezomib and dexamethasone
QC	quality control
RCT	randomized controlled trial
REC	Research Ethics Committees
RRMM	relapsed/refractory multiple myeloma
RTDS	Radiotherapy Data Set
RW	real-world
RWD	real-world data
SACT	Systemic Anti-Cancer Therapy
SNDS	Système National des Données de Santé
TTNTD	time to next treatment or death
UK	United Kingdom
ULN	upper limit of normal
US	United States
VGPR	very good partial response
β2M	beta-2 microglobulin

Acronym/Short Name	Meaning
β-hCG	β human chorionic gonadotropin

1 Abstract / Study Synopsis

Title	Fitness for purpose of data sources relevant for real-world data (RWD) studies on CAR-T cell therapy
Rationale for Study	Chimeric antigen receptor T-cell (CAR-T) cell therapies have transformed the therapeutic landscape for patients with hematologic malignancies.[1-3] Tisagenlecleucel and axicabtagene ciloleucel were the first CAR-T cell therapies to receive marketing authorisations from the European Medicines Agency (EMA) (August 2018), followed by brexucabtagene autoleucel (December 2020), idecabtageene vicleucel (August 2021), lisocabtagene maraleucel (April 2022), and ciltacabtagene autoleucel (cilta-cel) (May 2022).[4] Generating RWD is important for monitoring the safety and establishing the real-world (RW) effectiveness of CAR-T cell therapies. The EMA has commissioned a project with the aim of identifying and evaluating data sources in Europe and United States (US) that are fit-for-use for conducting regulatory-relevant non-interventional studies (NIS) in CAR-T cell therapies. To support this assessment, the following three theoretical scenarios have been formulated to serve as test cases: 1. To estimate the incidence rates of five safety outcomes (neurotoxicity, cytokine release syndrome (CRS), neutropenia, infections, second primary malignancies) among patients treated with CAR-T cell therapies. 2. To develop an external control arm for CARTITUDE-1, a single-arm phase 1b/2 trial (NCT03548207)[5] to assess the effectiveness of cilta-cel among adult patients with relapsed/refractory multiple myeloma (RRMM). 3. To emulate the CARTITUDE-4 (NCT04181827)[6] clinical trial to compare the effectiveness of cilta-cel to standard of care among adult patients with lenalidomide-refractory multiple myeloma (MM) who had received 1-3 prior lines of therapy (LOT) (CARTITUDE 4 clinical trial emulation). The scenarios are selected to permit a full fitness-for-use evaluation, including evaluating presence of inclusion and exclusion criteria, outcomes and covariates to understand the feasibility of conducting such effectiveness/safety studies in the future. The study does not aim to perform actual analyses of treatment effectiveness or safety. For each scenario, a maximum of three relevant data sources were selected based upon high-level data quality and relevance assessment and based on accessibility and likely willingness to collaborate. This protocol outlines the planned analyses to assess the fitness-for-use of National Cancer Registration and Analysis Service (NCRAS) to address theoretical scenarios 1 and 2, Flatiron Health to address

	theoretical scenarios 1, 2 and 3 and Dispositif d'Enregistrement et Suivi des patients traités par CAR-T cells (DESCAR-T) to address theoretical scenario 1. The fitness-for-use of the Système National des Données de Santé (SNDS) to address each theoretical scenario will be assessed without direct access to the data.
Objectives	The objective of the study is to evaluate the fitness-for-use of NCRAS, Flatiron Health and DESCAR-T data sources to address the following regulatory-relevant theoretical scenarios: 1. To estimate the incidence rates of five safety outcomes (neurotoxicity, CRS, neutropenia, infections, second primary malignancies) among patients treated with CAR-T cell therapy 2. To develop an external control arm for CARTITUDE-1, a single-arm phase 1b/2 trial (NCT03548207)[5] to assess the effectiveness of cilta-cel among adult patients with RRMM 3. To emulate the CARTITUDE-4 (NCT04181827)[6] clinical trial to compare the effectiveness of cilta-cel to standard of care among adult patients with lenalidomide-refractory MM who had received 1-3 prior LOT (CARTITUDE-4 clinical trial emulation).
Study Design	This is a fitness-for-use assessment to test the feasibility of implementing the three theoretical scenarios in NCRAS, Flatiron Health and DESCAR-T data sources. In addition, a hybrid fitness-for-use assessment of the SNDS will be undertaken without direct access to the data.
Source data	**NCRAS and linked datasets** General Data Protection Regulation (GDPR) compliant and pseudonymized data for this study will be sourced from the NCRAS and linked data sources. NCRAS holds nationwide data for patients diagnosed with a malignancy in England from 2008 onwards. Its datasets consist of the Patients Table, Tumour Table and Treatment Table, capturing demographic details of patients diagnosed with cancer, characteristics of the cancer and treatments administered. NCRAS can be linked to other national datasets which include the Systemic Anti-Cancer Therapy (SACT) dataset, Radiotherapy Data Set (RTDS) and the Hospital Episodes Statistics (HES) datasets. As such, NCRAS and its linked data sources represent strong candidate data sources for conducting NIS studies into CAR-T cell therapy. **Flatiron Health**

	De-identified data for this study will be sourced from Flatiron Health. Flatiron Health is a US electronic healthcare records data source. It captures longitudinal, patient-level, RWD of patients receiving care at academic and community cancer centres across the US. Data are collected from over 280 oncology practices at 800+ unique sites of care in the US. For the purposes of theoretical scenarios 2 and 3, the existing MM cohort registry including over 10,000 patients will be utilized. MM-specific data have been abstracted from structured and unstructured fields to provide detailed information about patient demographics, clinical characteristics, treatment and progression data. **DESCAR-T** DESCAR-T is a multicentre registry that has collected data for patients eligible for CAR-T cell therapy from 27 French centres administering CAR-T cell therapy since 01 July 2018. Demographic and clinical data, including treatments and responses are collected from medical records in a standardized electronic case report. Authorization was obtained from the National Commission for data protection and freedom of information (CNIL) to collect patient social security numbers to enable long term follow-up of patient vital status. **SNDS** The SNDS contains comprehensive clinical data, including pharmacy data, with a coverage of 99% of the population in France. It captures demographic data (age, sex, vital status) and diagnoses from the Affection Longue Duree (ALD), for long-term disease database, coded according to the International Classification of Diseases, 10th Revision (ICD-10). These data are linked with the Programme de médicalisation des systèmes d'information (PMSI) database, a national hospital discharge database which captures inpatient data such as diagnoses, medical procedures, imaging, external consultations, external acts performed, as well as some medicines and devices. Data are also linked to Centre d'épidémiologie sur les causes médicales de décès (CepiDc), the database on causes of death. An extract of the SNDS will not be requested for the purpose of this fitness-for-use assessment. Instead, the extent to which the SNDS can be used to address each theoretical scenario will be assessed based upon Certara's prior experience, including publications.
Study Population	For each theoretical scenario, a series of inclusion and exclusion criteria will be specified and the extent to which these eligibility criteria may be implemented in each data source will be evaluated as part of the fitness-for-use assessment. The inclusion and exclusion criteria have been formulated to align with the CARTITUDE-1 and CARTITUDE-4 trials, where applicable. A brief illustrative

	overview of the eligibility criteria for each theoretical scenario is provided below. Full eligibility criteria are provided in Sections 5.2.2 Study Glossary and 5.2.3 Context and Rationale for Study Exclusion Criteria. **Theoretical Scenario 1** The study population will include patients who received at least one dose of a CAR-T cell therapy during the indexing period (23 August 2018 to 31 December 2024) and who had 12 months data source history. Patients will be excluded if they had evidence of any of the safety outcomes of interest or were enrolled in a clinical trial in the 12 months prior to index date. **Theoretical Scenario 2** The study population will include adults with MM, with an Eastern Cooperative Oncology Group (ECOG) performance status score of 0 or 1, who received 3 or more previous LOT or were double refractory to a proteasome inhibitor (PI) and an immunomodulatory drug (IMiD), and had received a PI, IMiD, and an anti-CD38 antibody. Patients who meet the above criteria and initiate a standard of care MM treatment during the indexing period (01 January 2018 to 31 December 2023) will be included. Patients will be excluded if they received CAR-T cell therapy or an agent targeting B-cell maturation antigen (BCMA) prior to index date. **Theoretical Scenario 3** The study population will include adults with MM, with an ECOG performance status score of 0 or 1, who were lenalidomide-refractory and who had received one to three LOT including a PI or an IMiD and who initiated treatment with cilta-cel (treatment arm) or pomalidomide, bortezomib and dexamethasone (PVd) or daratumumab, pomalidomide and dexamethasone (DPd) (comparison arm) during the indexing period (28 February 2022 to 31 December 2023). Patients will be excluded if they received CAR-T cell therapy or an agent targeting BCMA prior to index date.
Study Period	See Section 5.1 Study Design for description of study period and study design figures for each theoretical scenario.
Key Study Variables	Details of variables captured and evaluated for the purpose of assessing the fitness-for-use for addressing each theoretical scenario are provided in Section 5.3 Variables. Examples of exposures, outcomes and covariates captured are provided below Exposures • New use of CAR-T cell therapy *(Theoretical scenario 1)* • New use of standard of care MM *(Theoretical scenario 2)*

	• New use of cilta-cel *(Theoretical scenario 3)* • New use of PVd or DPd *(Theoretical scenario 3)* Outcomes • Diagnosis of CRS, neurotoxicity, infections, neutropenia or second primary malignancies *(Theoretical scenario 1)* • Overall response rate (ORR) *(Theoretical scenarios 2 and 3)* • Minimal residual disease (MRD) negative rate *(Theoretical scenarios 2 and 3)* • Progression-free-survival (PFS) *(Theoretical scenarios 2 and 3)* • Overall survival (OS) (*Theoretical scenarios 2 and 3*) Covariates (necessary for characterising study population as part of addressing each theoretical scenario) • Age at index date • Sex • Cancer type (indication) • ECOG performance status at index date • Receipt of stem-cell transplant prior to index date • Number and type of pharmacotherapy or radiotherapy prior to index date • Refractory status (double, triple, quadruple, other)
Data Analyses	Aligning with the approach recommended in the European Medicine Agency (EMA) "*Data Quality Framework for European Union (EU) medicines: Application to Real-World data*", a series of research question-specific metrics for each dimension of data quality (extensiveness, coherence, reliability, relevance and timeliness) will be evaluated as part of the fitness-for-use assessment of each data source for each theoretical scenario. Please note, metrics relating to foundational determinants (i.e., research-question independent) of data quality of up to 30 data sources for CAR-T cell therapy research are being assessed in a parallel study.[7] Findings from both studies will be reported in a global project report. Some illustrative metrics are provided below. A full list of metrics assessed for each theoretical scenario are provided in Section 5.4 Data Analysis. Extensiveness (coverage and completeness) metrics • Percentage (%) of records for which CAR-T cell administration date is unavailable • Number (%) of patients with incomplete data on key covariate data Reliability (accuracy) metrics

	• Number (%) of patients who received CAR-T cell therapy without a known indication for the given therapy Coherence (relational) metrics • Number of duplicate CAR-T cell therapy records (e.g., records of multiple administrations on single date) • Number (%) of patients with observable data (e.g., treatment data) after date of death Timeliness metrics • Date of last observed data Relevance • Availability of all key data elements (i.e., availability of exposures, outcomes and covariates)

2 Amendments and Updates

Version	Date	Section Changed	Description of Change	Rationale
1.1	05/08/25	Throughout	Addition of SNDS as data source (hybrid)	To expand the fitness-for-use assessment to additional data sources with which EMA has less experience

Abbreviation: EMA = European Medicines Agency; SNDS = Systeme National des Données de Santé.

3 Rationale and Background

Chimeric antigen receptor T-cell (CAR-T) cell therapies have transformed the therapeutic landscape for patients with hematologic malignancies, including B-cell acute lymphoblastic leukaemia (B-ALL), diffuse large B-cell lymphoma (DLBCL), mantle cell lymphoma (MCL), follicular lymphoma, primary mediastinal B-cell lymphoma (BCL), and, more recently relapsed/refractory multiple myeloma (RRMM).[1-3] Multiple CAR-T products, such as axicabtagene ciloleucel, tisagenlecleucel, brexucabtagene autoleucel, lisocabtagene maraleucel, and ciltacabtagene autoleucel (cilta-cel)—have demonstrated significant efficacy in pivotal clinical trials, leading to their regulatory approvals across the United States (US), European Union (EU), and other regions.[8,9] Tisagenlecleucel and axicabtagene ciloleucel were the first CAR-T cell therapies to receive marketing authorisations from the European Medicines Agency (EMA) (August 2018), followed by brexucabtagene autoleucel (December 2020), idecabtageene vicleucel (August 2021), lisocabtagene maraleucel (April 2022), and cilta-cel (May 2022).[4]

Despite encouraging trial results, CAR-T cell therapies present substantial safety challenges, with adverse events (AEs) such as cytokine release syndrome (CRS), neurotoxicity, infections, neutropenia, and the potential for second primary malignancies.[10-14] Understanding the real-world (RW) incidence and impact of these AEs outside controlled trial environments is essential to support risk management strategies, patient care, and resource allocation.

Randomized controlled trials (RCTs), though essential, typically enrol highly selected patient populations and may not fully represent patients encountered in routine clinical practice. Consequently, generalizability of trial outcomes, particularly regarding the effectiveness and safety of CAR-T cell therapies in broader, more diverse patient populations remains uncertain. Furthermore, evolving treatment standards for multiple myeloma (MM) necessitate timely RW evaluations to contextualize the clinical benefit of new therapies against contemporary standard-of-care treatments.

The increasing availability of real-world data (RWD) offers important opportunities to inform the RW effectiveness and safety of approved CAR-T cell therapies, as well as to provide context to findings from single arm trials.

The EMA has commissioned a project with the aim of identifying and evaluating data sources in Europe and US that are fit-for-use for conducting non-interventional studies (NIS) for CAR-T cell therapies. To support this assessment, the following three theoretical scenarios have been formulated to serve as methodological test cases:

1. To estimate the incidence rates of five safety outcomes (neurotoxicity, CRS, neutropenia, infections, second primary malignancies) among patients treated with CAR-T cell therapy
2. To develop an external control arm for CARTITUDE-1, a single-arm phase 1b/2 trial (NCT03548207)[5] to assess the effectiveness of cilta-cel among patients with RRMM
3. To emulate the CARTITUDE-4 (NCT04181827)[6] clinical trial to compare the effectiveness of cilta-cel to standard care among adult patients with lenalidomide-refractory MM who had received 1-3 prior lines of therapy (LOT) (CARTITUDE-4 clinical trial emulation

The scenarios were selected to permit a full fitness-for-use evaluation, which is not possible in the absence of a concrete research question, including evaluating completeness and accuracy of inclusion and exclusion criteria, outcomes and covariates. The study does not aim to perform actual analyses of treatment effectiveness or safety or to generate results that should be used for decision making, except decisions on feasibility and relevance of the data.

The detailed fitness-for-use assessment will aim to assess the specific data for the question at hand in terms of completeness, reliability, extensiveness, timeliness and coherence, aligning with the EMA's Data Quality Framework.[15,16] Specifically, the required data elements to address the research question (theoretical scenario) will be operationalized to create the final dataset (i.e., specific data cut) for the respective theoretical scenario. An assessment of the final dataset will then be conducted by examining the extensiveness, coherence, reliability and timeliness of the required design elements for each theoretical scenario. Please note, a more general quality assurance assessment is being performed as part of the first phase of this project. This will involve the inspection of documentation for several data sources to understand data reliability, timeliness, coherence, along with a high-level assessment of data relevance.

For each scenario, a maximum of three relevant data sources were selected based upon high-level data quality and relevance assessment and based on accessibility and likely willingness to collaborate. This protocol outlines the planned analyses to assess the fitness-for-use of the National Cancer Registration and Analysis Service (NCRAS) dataset to address theoretical scenarios 1 and 2 of Flatiron Health to address theoretical scenarios 1, 2 and 3 and of the Dispositif d'Enregistrement et Suivi des patients traités par CAR-T cells (DESCAR-T) to address theoretical scenario 1. In addition, the fitness-for-use of the French Système National des Données de Santé (SNDS) to address each theoretical scenario will be undertaken using a hybrid approach. This assessment will not involve direct access to SNDS data but will instead rely upon Certara's expertise and prior publications using the SNDS.

4 Research Objectives

The objective of the study is to evaluate the fitness-for-use of NCRAS, Flatiron Health and DESCAR-T to address the following regulatory-relevant theoretical scenarios:

1. To estimate the incidence rates of five safety outcomes (neurotoxicity, CRS, neutropenia, infections, second primary malignancies) among patients treated with CAR-T cell therapy
2. To design an externally controlled study for CARTITUDE-1, a single-arm phase 1b/2 trial (NCT03548207)[5] to assess the effectiveness of cilta-cel among patients with RRMM
3. To emulate the CARTITUDE-4 (NCT04181827)[6] clinical trial to compare the effectiveness of cilta-cel to standard care among adult patients with lenalidomide-refractory MM who had received 1-3 prior LOT (CARTITUDE-4 clinical trial emulation)

5 Research Methods

5.1 Study Design

This is a fitness-for-use assessment to test the feasibility of implementing the three theoretical scenarios in NCRAS, Flatiron Health and DESCAR-T data sources. In addition, a hybrid fitness-for-use assessment of the SNDS will additionally be undertaken without direct access to the data.

A retrospective, observational cohort design will be used as a template for each theoretical scenario, to simulate how such studies would be implemented using the available data. Upon creation of the study cohorts, data quality metrics will then be assessed on this bespoke study specific cohort and the fitness-for-use assessment will be undertaken (outlined in Section 5.4 Data Analysis).

Theoretical Scenario 1: Descriptive cohort analysis to estimate the incidence rates of five predefined safety outcomes of interest among all patients treated with CAR-T cell therapies (regardless of indication). In the theoretical scenario, patients would be followed longitudinally from index date (defined in Section 5.2.1 Context and Rationale for Definition of Index Date) until the earliest of: safety outcome, death, clinical trial enrolment, or end of study period (31 December 2024).

Theoretical Scenario 2: To create an external control arm of patients with RRMM for comparison with the CARTITUDE-1 trial. The CARTITUDE-1 trial was a single arm, phase 1b/2 clinical trial to assess the safety and effectiveness of cilta-cel among adult patients with RRMM. In the Theoretical scenario, patients would be followed longitudinally from index date until the earliest of effectiveness outcome, death, clinical trial enrolment or end of study period (31 December 2024).

Theoretical Scenario 3: Retrospective RW comparative effectiveness analysis to emulate the CARTITUDE-4 trial. The CARTITUDE-4 trial was a phase 3 open-label randomized trial to compare cilta-cel with standard of care therapies among patients with lenalidomide-refractory MM and treated with 1–3 prior LOTs. In the theoretical scenario, patients would be followed longitudinally from index date until the earliest of effectiveness outcome, death, clinical trial enrolment or end of study period (31 December 2024).

5.1.1 Study Design Diagram(s)

Figure 1, Figure 2 and Figure 3 present the study design diagrams for theoretical scenarios 1, 2 and 3, respectively.

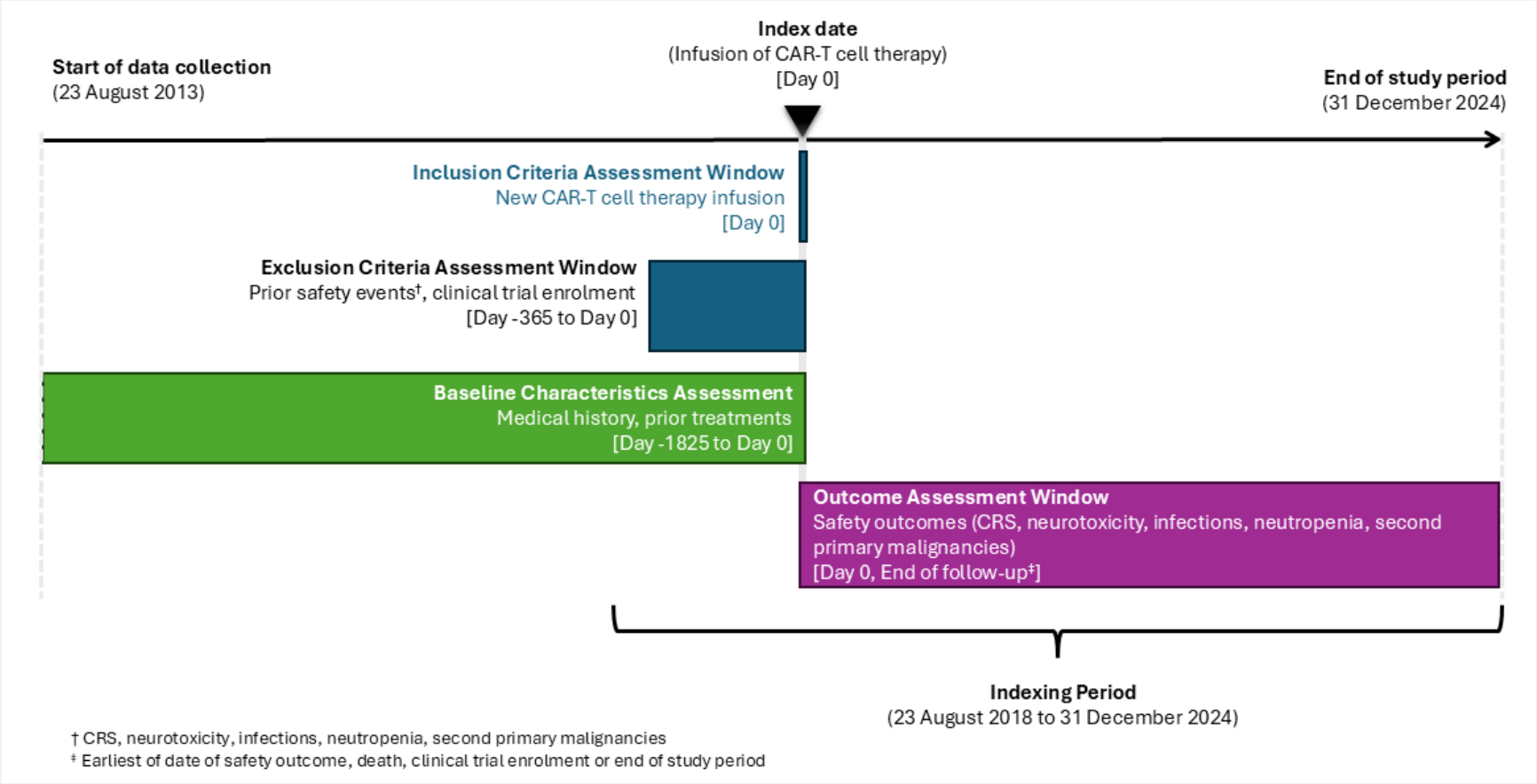

Figure 1. Study Design Diagram for Theoretical Scenario 1

Abbreviations: CAR-T cell = chimeric antigen receptor T-cell; CRS: cytokine release syndrome.

See full inclusion and exclusion criteria in Section 5.2.2 Context and Rationale for Study Inclusion Criteria and 5.2.3 Context and Rationale for Study Exclusion Criteria, respectively.

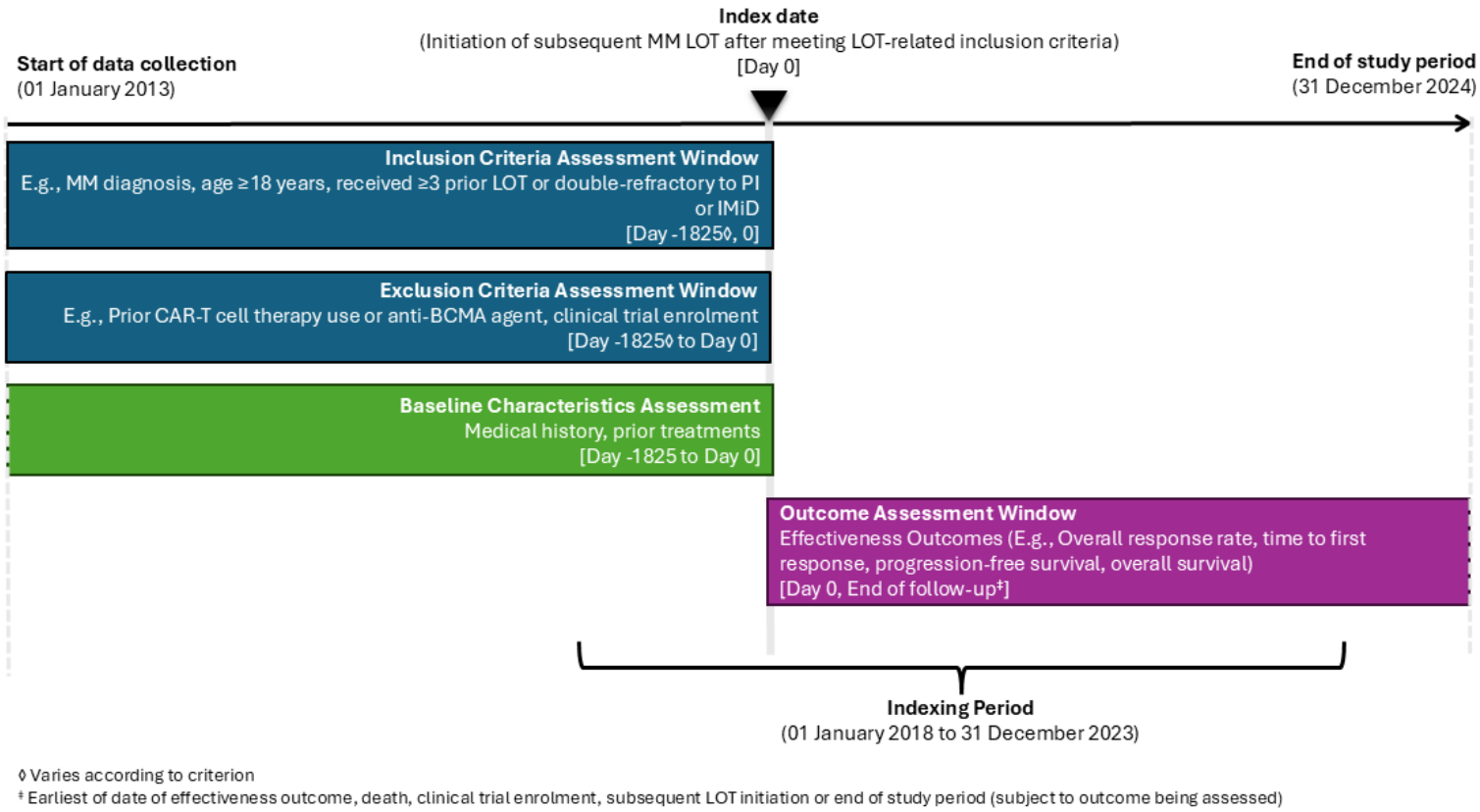

Figure 2. Study Design Diagram for Theoretical Scenario 2

Abbreviations: BCMA = B-cell maturation antigen; CAR-T cell = chimeric antigen receptor T-cell; IMiD = immunomodulatory imide drugs; LOT = lines of therapy; MM = multiple myeloma; PI = proteasome inhibitor.

See full inclusion and exclusion criteria in Section 5.2.2 Context and Rationale for Study Inclusion Criteria and 5.2.3 Context and Rationale for Study Exclusion Criteria, respectively.

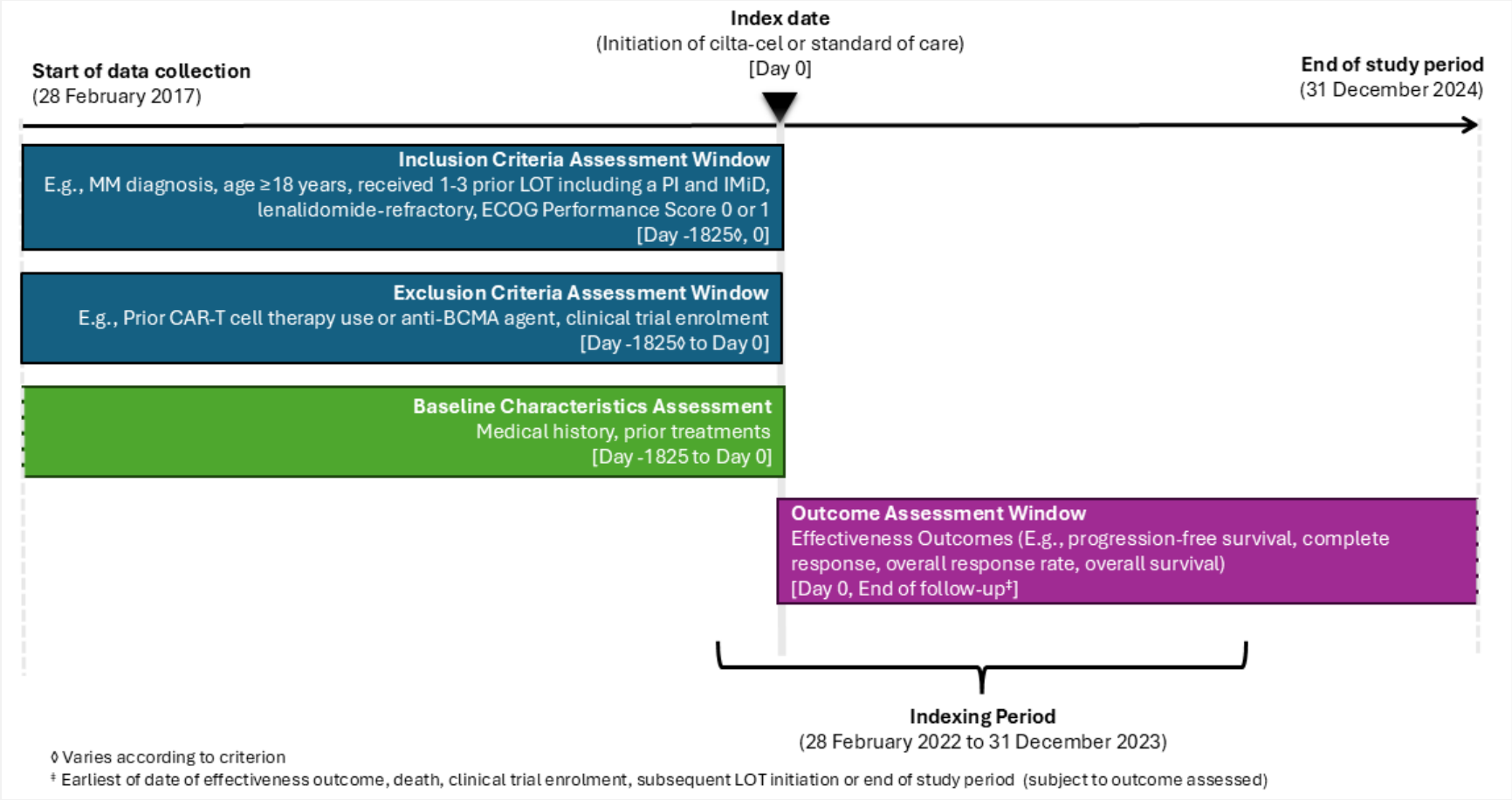

Figure 3. Study Design Diagram for Theoretical Scenario 3

Abbreviations: BCMA = B-cell maturation antigen; CAR-T cell = chimeric antigen receptor T-cell; ECOG = Eastern Cooperative Oncology Group; IMiD = immunomodulatory imide drugs; LOT = lines of therapy; MM = multiple myeloma; PI = proteasome inhibitor.

See full inclusion and exclusion criteria in Section 5.2.2 Context and Rationale for Study Inclusion Criteria and 5.2.3 Context and Rationale for Study Exclusion Criteria, respectively.

5.2 Setting

5.2.1 Context and Rationale for Definition of Index Date

The definition of index date (time zero) is tailored to each of the theoretical scenarios, based on differences in therapeutic context, trial design emulation needs, and expected treatment timelines. These definitions are summarized in Table 1.

Table 1. Operational Definitions of Index Date (Day 0)

Theoretical Scenario	Index Date	Indexing Period (Patient Identification Period)
1: Safety evaluation	Date of CAR-T cell therapy infusion during indexing period	23 August 2018* and 31 December 2024
2: CARTITUDE-1 external control arm [5]	Date of initiation of subsequent standard of care therapy after meeting inclusion criteria relating to prior LOTs (≥3 prior MM LOT OR are double refractory to an IMiD and PI AND received as part of previous therapy a PI, an IMiD, and an anti-CD38 antibody for the treatment of RRMM)	01 January 2018 to 31 December 2023
3. CARTITUDE-4 clinical trial emulation	During indexing period: • Date of cilta-cel infusion (treatment arm) • Date of initiation of PVd or DPd (comparison arm)	28 February 2022** to 31 December 2023

Abbreviations: CAR-T cell = chimeric antigen receptor T-cell; DPd = daratumumab, pomalidomide and dexamethasone; IMiD immunomodulatory imide drugs, LOT = lines of therapy; MM = multiple myeloma; PI = proteasome inhibitor; PVd = pomalidomide, bortezomib and dexamethasone; RRMM = relapsed/refractory multiple myeloma.
* Aligns with date of first EMA CAR-T marketing authorisation[17]
** Aligns with data of US Food and Drugs Administration approval of cilta-cel[18]

5.2.2 Context and Rationale for Study Inclusion Criteria

For the purposes of forming cohorts appropriate to each theoretical scenario that would be subsequently used in fitness-for-use assessments, their respective inclusion criteria are listed in Table 2. To facilitate target trial emulation, inclusion criteria for theoretical scenario 2 and theoretical scenario 3 have been formulated to align with the inclusion criteria of the CARTITUDE-1 and CARTITUDE-4 clinical trial, respectively.

Note, no operational definitions for the inclusion criteria are provided for the SNDS since SNDS will not be accessed for the fitness-for-purpose assessment.

Table 2. Inclusion Criteria for Theoretical Scenarios 1, 2 and 3 and their Potential Operationalization in NCRAS and Flatiron Health

Theoretical Scenario	Trial Inclusion Criteria [Other Criteria]	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Codes / Algorithms	Assessment Window
1: Safety Evaluation	1. N/A [New use of CAR-T cell therapy]	Patients who received at least one dose of any CAR-T cell therapy	Patients who received at least one dose of any CAR-T cell therapy	Patients who received at least one dose of any CAR-T cell therapy	N/A	Indexing period
	2. N/A [With a minimum of 12 months data source history]	Any patient records (i.e., treatments, hospitalisations)	Any patient records (i.e., treatments, diagnosis, hospitalisations)	Any patient records (i.e., treatments, diagnosis, hospitalisations)	N/A	[-365, 0]
2: CARTITUDE-1 External Control Arm [5]	1. Adult (≥18 years)	Aged ≥18 years	Aged ≥18 years	Theoretical scenario not assessed in this data source	N/A	[0,0]
	2. Documented diagnosis of MM according to IMWG diagnostic criteria	Diagnostic record for MM in NCRAS	Diagnostic record for MM in Flatiron Health. Diagnoses of MM are manually reviewed by Flatiron Health data abstractors	Theoretical scenario not assessed in this data source	Appendix 1. Code Lists	[-1825, 0]
	3. Measurable disease at Screening as defined by any of the following: • Serum M-protein level more than or equal to ≥1.0g/dL or urine M-protein level ≥200mg/24h; or • b) Light chain MM without measurable disease in the serum or the urine: Serum immunoglobulin free light chain 10mg/dL and abnormal serum immunoglobulin kappa lambda free light chain ratio	NCRAS and linked datasets do not collect laboratory data therefore criterion **not operational** in this data source	Following laboratory values abstracted from patient health record: • Blood M-protein ≥1.0g/dL OR • Urine M-protein ≥200mg/24h Possibility to identify measurable disease according to light chain MM unknown	Theoretical scenario not assessed in this data source	N/A	[-90, 0]
	4. Received at least 3 prior MM treatment LOT or are double refractory to an IMiD and PI (refractory MM as defined by IMWG consensus criteria). Note: induction with or without hematopoietic stem cell transplant and with or without	Evidence of ≥3 prior MM LOT, defined according to published algorithm OR Double-refractory to an IMiD and PI. Refractoriness is	Evidence of ≥3 prior MM LOT, defined according to published algorithm OR Double-refractory to an IMiD and PI. Refractoriness is	Theoretical scenario not assessed in this data source	Appendix 2. LOT definitions Appendix 3. MM Treatment Groupings	[-1825, 0]

Theoretical Scenario	Trial Inclusion Criteria [Other Criteria]	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Codes / Algorithms	Assessment Window
	maintenance therapy is considered a single LOT	defined as initiating the next LOT ≤60 days after ending the preceding LOT.[19]	defined as discontinuation of drug of interest within 60 days and starting a different drug in the next line or starting a new drug within 60 days after end of previous treatment and by IMWG[20]			
	5. Received as part of previous therapy a PI, an IMiD, and an anti-CD38 antibody	Receipt of an IMiD, a PI, and an anti-CD38 antibody as part of previous LOTs	Receipt of an IMiD, a PI, and an anti-CD38 antibody as part of previous LOTs	Theoretical scenario not assessed in this data source	Appendix 2. LOT definitions Appendix 3. MM Treatment Groupings	[-1825, 0]
	6. Documented disease progression within 12 months of starting the most recent antimyeloma therapy. Defined as evidence of PD based on investigator's determination of response by the IMWG criteria on or within 12 months of their last LOT. Confirmation may be from either central or local testing. Also, participants with documented evidence of PD within the previous 6 months and who are refractory or non-responsive to their most recent LOT afterwards are eligible	Record in SACT dataset indicating treatment discontinuation of prior LOT due to progression within 12 months of LOT initiation OR Record in SACT indicating treatment discontinuation in 6 months prior to index date AND Evidence of being refractory (defined as above) to their most recent LOT	Evidence of progression abstracted from clinician notes within 12 months of initiation of most recent LOT OR Evidence of progression within 6 months of most recent LOT AND Evidence of being refractory (defined as above) to patient's most recent LOT	Theoretical scenario not assessed in this data source	Appendix 1. Code Lists Appendix 2. LOT definitions	[-1825, 0] or as otherwise stated

Theoretical Scenario	Trial Inclusion Criteria [Other Criteria]	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Codes / Algorithms	Assessment Window
	7. ECOG Performance Status grade of 0 or 1	Recorded ECOG Performance Status grade of 0 or 1 derived from SACT data source	Recorded ECOG Performance Status grade of 0 or 1		N/A	[-90,0]
	8. Clinical laboratory values meeting the following criteria during the Screening Phase • Haemoglobin ≥8.0g/dL (≥5mmol/L) (without prior red blood cell transfusion within 7 days before the laboratory test; recombinant human erythropoietin use is permitted) • Platelets ≥50 x 10^9/L (must be without transfusion support in the 7 days prior to the laboratory test) • ANC ≥0.75×10^9/L (prior growth factor support is permitted but must be without support in the 7 days prior to the laboratory test) • AST and ALT ≤3.0× ULN • Creatinine clearance ≥40mL/min/1.73m^2 based upon MDRD formula calculation or a 24-hour urine collection. • Total bilirubin ≤2.0×ULN; except in subjects with congenital bilirubinaemia, such as Gilbert syndrome (in which case direct bilirubin ≤1.5×ULN is required) • Corrected serum calcium ≤12.5mg/dL (≤3.1mmol/L) or free ionized calcium ≤6.5mg/dL (≤1.6mmol/L)	NCRAS and linked datasets do not collect laboratory data therefore criterion ***not operational*** in this data source	Following laboratory values abstracted from clinician notes: • Haemoglobin ≥8g/dL • Platelet count ≥50×10^9/L in participants in whom ≥50% of bone marrow nucleated cells are plasma cells • ANC ≥0.75×10^9/L • AST and ALT ≤3×ULN • Creatinine clearance ≥40mL/min per 1.73m^2 • Total bilirubin ≤2.0×ULN • Serum calcium ≤12.5mg/dL (≤3.1mmol/L) or free ionized calcium ≤6.5mg/dL (≤1.6mmol/L)	Theoretical scenario not assessed in this data source	N/A	[-90, 0]
	9. Women of childbearing potential must have a negative pregnancy test at screening and prior to the first dose of cyclophosphamide and fludarabine using a highly sensitive serum pregnancy test (β human chorionic gonadotropin [β-hCG]).	**Uncertain feasibility** in NCRAS. To be explored whether active pregnancies may be identified in data source.	Record abstracted from clinician notes indicating patient pregnancy		N/A	[-304, 0]

Theoretical Scenario	Trial Inclusion Criteria [Other Criteria]	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Codes / Algorithms	Assessment Window
	10. When a woman is of childbearing potential subject must agree to practice a highly effective method of contraception and agree to remain on a highly effective method of contraception from the time of signing the GVHD until at least 100 days after receiving a JNJ-68284528 infusion. Women and men must agree not to donate eggs (ova, oocytes) or sperm, respectively, during the study and for 100 days after the last dose of study treatment	Criterion **not relevant** for RW database study	Criterion **not relevant** for RW database study	Theoretical scenario not assessed in this data source	-	-
	11. Subject must sign an GVHD indicating that he or she understands the purpose of and procedures required for the study and is willing to participate in the study. Consent is to be obtained prior to the initiation of any study-related tests or procedures that are not part of standard-of-care for the subject's disease.	Criterion **not relevant** for RW database study	Criterion **not relevant** for RW database study	Theoretical scenario not assessed in this data source	-	-
	12. Willing and able to adhere to the prohibitions and restrictions specified in this protocol.	Criterion **not relevant** for RW database study	Criterion **not relevant** for RW database study	Theoretical scenario not assessed in this data source	-	-
	13. N/A [Initiation of new RRMM LOT following satisfying the following criteria: • Received at least 3 prior MM treatment LOT or are double refractory to an IMiD and PI • Received as part of previous therapy a PI, an IMiD, and an anti-CD38 antibody]	After meeting criteria relating to prior LOT, evidence of initiation of a subsequent RRMM LOT	After meeting criteria relating to prior LOT, evidence of initiation of a subsequent RRMM LOT	Theoretical scenario not assessed in this data source	Appendix 2. LOT definitions Appendix 3. MM Treatment Groupings	Indexing window
3: CARTITUDE-	1. Adult (≥18 years)	Theoretical scenario not assessed in this database	Aged ≥18 years	Theoretical scenario not assessed in this data source	N/A	[0,0]

Theoretical Scenario	Trial Inclusion Criteria [Other Criteria]	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Codes / Algorithms	Assessment Window
4 Trial Emulation [6]	2. MM diagnosis according to the IMWG diagnostic criteria	Theoretical scenario not assessed in this data source	Diagnostic record in Flatiron Health. Diagnoses of MM are manually reviewed by Flatiron Health data abstractors	Theoretical scenario not assessed in this data source	Appendix 1. Code Lists	[-1825, 0]
	3. Measurable disease at screening as defined by any of the following: • Serum M-protein level ≥0.5g/dL or urine M-protein level ≥200mg/24h; or • Light chain MM without measurable M-protein in the serum or the urine: Serum free light chain ≥10mg/dL and abnormal serum free light chain ratio	Theoretical scenario not assessed in this data source	Following laboratory values abstracted from clinician notes: • Blood M-protein ≥1.0g/dL OR • Urine M-protein ≥200mg/24h Possibility to identify measurable disease according to light chain MM unknown	Theoretical scenario not assessed in this data source	N/A	[-90, 0]
	4. Received 1 to 3 prior LOT including a PI and an IMiD	Theoretical scenario not assessed in this data source	Record indicating ≤3 LOT, including an IMiD and PI	Theoretical scenario not assessed in this data source	Appendix 2. LOT definitions	[-1825, 0]
	5. Documented evidence of PD by IMWG criteria based on investigator's determination on or within 6 months of their last regimen	Theoretical scenario not assessed in this data source	Evidence of progression abstracted from clinician notes within 6 months after most recent LOT prior to index treatment.	Theoretical scenario not assessed in this data source	Appendix 2. LOT definitions	[-1825, 0]
	6. Subjects with only 1 prior LOT must have progressed within 36 months of a stem cell transplant, or if not transplanted, then within 42 months of starting initial therapy	Theoretical scenario not assessed in this data source	Among patients with 1 prior LOT: Evidence of progression abstracted from clinician notes in patient health record within 36 months of stem cell transplant, defined as a record of	Theoretical scenario not assessed in this data source	Appendix 2. LOT definitions	[-1825, 0]

Theoretical Scenario	Trial Inclusion Criteria [Other Criteria]	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Codes / Algorithms	Assessment Window
			stem cell transplant abstracted from patient health record OR Evidence of progression abstracted from patient health record within 42 months of starting LOT			
	7. Refractory to lenalidomide per IMWG consensus guidelines (failure to achieve minimal response or progression on or within 60 days of completing lenalidomide therapy). Progression on or within 60 days of the last dose of lenalidomide given as maintenance will meet this criterion.	Theoretical scenario not assessed in this data source	Evidence of being lenalidomide refractory defined as having a record indicating a change in treatment within 60 days of last LOT with lenalidomide as a component AND Without lenalidomide as a component of the immediate next LOT.	Theoretical scenario not assessed in this data source	Appendix 2. LOT definitions	[-1825, 0]
	8. ECOG Performance Status score of 0 or 1	Theoretical scenario not assessed in this data source	Recorded ECOG Performance Status grade of 0 or 1	Theoretical scenario not assessed in this data source	N/A	[-90,0]
	9. Clinical laboratory values meeting the following criteria during the Screening Phase (retesting is allowed but the below criteria must be met in the latest test prior to randomization): • Haemoglobin ≥8g/dL (without prior red blood cell transfusion within 7 days before the laboratory test; recombinant human erythropoietin use is permitted); • ANC ≥1×10^9/L (without recombinant human G-CSF within 7 days and without pegylated G-CSF within 14 days of the laboratory test);	Theoretical scenario not assessed in this data source	Following laboratory values abstracted from clinician notes in patient health record: • Haemoglobin ≥8g/dL • ANC ≥1×10^9 /L • Platelet count ≥75×10^9/L in participants in whom <50% of bone marrow nucleated cells are plasma cells;	Theoretical scenario not assessed in this data source	N/A	[-90, 0]

Theoretical Scenario	Trial Inclusion Criteria [Other Criteria]	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Codes / Algorithms	Assessment Window
	• Platelet count ≥75×10^9/L (without prior platelet transfusion within 7 days before the laboratory test) in participants in whom <50% of bone marrow nucleated cells are plasma cells; platelet count ≥50*10^9/L (without prior platelet transfusion within 7 days before the laboratory test) in participants in whom ≥50% of bone marrow nucleated cells are plasma cells; • Lymphocyte count ≥0.3×10^9/L; • AST ≤3×ULN; • ALT ≤3×ULN; • Total bilirubin ≤2.0×ULN; except in participants with congenital bilirubinaemia, such as Gilbert syndrome (in which case direct bilirubin ≤1.5×ULN is required); • Estimated glomerular filtration rate ≥40mL/min per 1.73m^2 (to be calculated using the MDRD formula)		• Platelet count ≥50×10^9/L in participants in whom ≥50% of bone marrow nucleated cells are plasma cells • Lymphocyte count ≥0.3×10^9/L • Aspartate aminotransferase ≤3×ULN • ALT ≤3×ULN • Total bilirubin ≤2.0×ULN • Estimated glomerular filtration rate ≥40mL/min per 1.73m^2			
	10. Women of childbearing potential must have 2 negative pregnancy tests prior to starting PVd or DPd. The first, 10-14 days prior to the start of PVd or DPd and prior to randomization. The second pregnancy test will need to be done within 24 hours prior to the start of PVd or DPd. The investigator must verify that the results of these tests are negative prior to starting PVd or DPd.	Theoretical scenario not assessed in this data source	Record abstracted from clinician notes indicating pregnancy	Theoretical scenario not assessed in this data source	N/A	[-304, 0]
	11. When a woman is of childbearing potential, the subject must commit either to abstaining continuously from heterosexual intercourse or agree to use 2 methods of reliable birth control simultaneously and agree to remain on both methods from the time of signing the ICF until at least 3 months after	Theoretical scenario not assessed in this data source	Criterion **not relevant** for RW database study	Theoretical scenario not assessed in this data source	-	-

Theoretical Scenario	Trial Inclusion Criteria [Other Criteria]	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Codes / Algorithms	Assessment Window
	receiving the last dose of daratumumab or bortezomib or 28 days after the last dose of pomalidomide, whichever is later (Arm A) or at least 1 year after receiving a JNJ-68284528 infusion or at least 3 months after receiving the last dose of daratumumab or bortezomib or 28 days after the last dose of pomalidomide, whichever is later (Arm B).					
	12. A man must commit either to abstaining continuously from heterosexual sexual intercourse or a man who is sexually active with a woman of childbearing potential or a pregnant woman must agree to: • Use a barrier method of contraception from the time of signing the ICF until at least 3 months after receiving the last dose of daratumumab or bortezomib or 28 days after the last dose of pomalidomide, whichever is later (Arm A) or at least 1 year after receiving a JNJ-68284528 infusion or at least 3 months after receiving the last dose of daratumumab or bortezomib or 28 days after the last dose of pomalidomide, whichever is later (Arm B), even if they have undergone a successful vasectomy; • Should agree to practice contraception according to and for the time frame specified in the local pomalidomide pregnancy prevention program	Theoretical scenario not assessed in this data source	Criterion **not relevant** for RW database study	Theoretical scenario not assessed in this data source	-	-
	13. Women and men must agree not to donate eggs (ova, oocytes) or sperm, respectively, during the study and for at least 3 months after receiving the last dose of daratumumab or bortezomib, or 28 days after the last dose of pomalidomide, whichever is later (Arm	Theoretical scenario not assessed in this data source	Criterion **not relevant** for RW database study	Theoretical scenario not assessed in this data source	-	-

Theoretical Scenario	Trial Inclusion Criteria [Other Criteria]	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Codes / Algorithms	Assessment Window
	A) or at least 1 year after receiving a JNJ-68284528 infusion or at least 3 months after receiving the last dose of daratumumab or bortezomib or 28 days after the last dose of pomalidomide, whichever is later (Arm B)					
	14. Must sign an ICF (or their legally acceptable representative must sign) indicating that he or she understands the purpose of, and procedures required for, the study and is willing to participate in the study. Subjects must be willing and able to adhere to the prohibitions and restrictions specified in this protocol, as referenced in the ICF	Theoretical scenario not assessed in this data source	Criterion **not relevant** for RW database study		-	-
	15. Willing and able to adhere to the lifestyle restrictions specified in this protocol	Theoretical scenario not assessed in this data source	Criterion **not relevant** for RW database study	Theoretical scenario not assessed in this data source	-	-

Abbreviations: /L = per litre; ALT = alanine aminotransferase; ANC = absolute neutrophil count; AST = aspartate aminotransferase; CAR-T cell = chimeric antigen receptor T-cell; DESCAR-T = Dispositif d'Enregistrement et Suivi des patients traités par CAR-T cells; DPd = daratumumab, pomalidomide and dexamethasone; ECOG = Eastern Cooperative Oncology Group; G-CSF = granulocyte colony-stimulating factor; g/dL = gram per decilitre; ICF = informed consent form; IMiD = immunomodulatory imide drugs; IMWG = International Myeloma Working Group; LOT = lines of therapy; m^2 = square metre; MDRD = Modified Diet in Renal Disease; mg/24h = milligramme per 24 hours; mg/dL = milligramme per decilitre; mL/min = millilitre per minute; MM = multiple myeloma; mmol/L = millimoles per litre; M-protein = monoclonal paraprotein; N/A = not applicable; NCRAS = National Cancer Registration and Analysis Service; PD = progressive disease; PI = proteasome inhibitor; PVd = pomalidomide, bortezomib and dexamethasone; RRMM = relapsed/refractory multiple myeloma; RW = real-world; SACT = Systemic Anti-Cancer Therapy; ULN = upper limit of normal; β-hCG = β human chorionic gonadotropin.

.

5.2.3 Context and Rationale for Study Exclusion Criteria

The exclusion criteria that will be applied to form the cohorts in which the fitness-for-use assessment will be performed are listed in Table 3. To facilitate target trial emulation, exclusion criteria for theoretical scenario 2 and 3 have been formulated to align with the exclusion criteria of the CARTITUDE-1 and CARTITUDE-4 clinical trial, respectively. Note, no operational definitions for the exclusion criteria are provided for the SNDS since SNDS will not be accessed for the fitness-for-purpose assessment.

Table 3. Exclusion Criteria for Theoretical Scenario 1, 2 and 3 and their Potential Operationalization in NCRAS and Flatiron Health

Theoretical Scenario	Trial Exclusion Criteria [Other criteria]	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Codes / Algorithms	Assessment Window
1: Safety evaluation	1. N/A [Prior history of safety outcomes of interest]	Hospital record in HES with relevant diagnostic code or procedure code for CRS, neurotoxicity, infections, neutropenia or second primary malignancies	Any diagnostic record or laboratory test result indicative of CRS, neurotoxicity, infections, neutropenia or second primary malignancies	Any diagnostic record or laboratory test result indicative of CRS, neurotoxicity, infections, neutropenia or secondary primary malignancies (See Table 5)	Appendix 1. Code Lists	[-365, 0]
	2. N/A [Enrolment in clinical trial at index date]	Participation in a clinical trial, defined as a diagnostic record for trial involvement or a record in SACT indicating patient participating in clinical trial.	Treatment records indicating "clinical study drug" in patient health record.	Record in patient health record indicating participation in clinical trial	ICD-10: Z00.6	[-365, 0]
2: CARTITUDE-1 External Control Arm [5]	1. N/A [Enrolment in clinical trial at index date]	Participation in a clinical trial, defined as a diagnostic record for trial involvement or a record in SACT indicating patient participating in clinical trial.	Treatment records indicating "clinical study drug" in patient health record.	Theoretical scenario not assessed in this data source	ICD-10: Z00.6	[-31,0]
	2. Prior treatment with CAR-T cell therapy directed at any target	Treatment record in SACT indicating CAR-T cell therapy administration	Treatment record indicating CAR-T cell therapy administration	Theoretical scenario not assessed in this data source	Appendix 3. MM Treatment Groupings	[-1825, 0]

Theoretical Scenario	Trial Exclusion Criteria [Other criteria]	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Codes / Algorithms	Assessment Window
	3. Received any therapy that is targeted to BCMA	Treatment record in SACT indicating use of anti-BMCA agent	Treatment record indicating use of anti-BMCA agent	Theoretical scenario not assessed in this data source	Appendix 3. MM Treatment Groupings	[-1825, 0]
	4. Diagnosed or treated for invasive malignancy other than MM, except: • Malignancy treated with curative intent and with no known active disease present for ≥3 years before enrolment; or Adequately treated non-melanoma skin cancer without evidence of disease.	Diagnostic record in HES for a cancer other than MM or non-melanoma skin cancer	**Uncertain feasibility** in Flatiron Health. To explore availability of records indicative of other active malignancies.	Theoretical scenario not assessed in this data source	Appendix 3. MM Treatment Groupings	[-1095, 0]
	5. Prior antitumor therapy as follows, prior to apheresis: • Targeted therapy, epigenetic therapy, or treatment with an investigational drug or used an invasive investigational medical device within 14 days or at least 5 half lives, whichever is less. • mAb treatment for MM within 21 days. • Cytotoxic therapy within 14 days. • PI therapy within 14 days. • IMiD therapy within 7 days. • Radiotherapy within 14 days. However, if the radiation portal covered ≥5% of the bone marrow reserve, the subject is eligible irrespective of the end date of radiotherapy	Treatment record in SACT or RTDS indicating use of: Investigational drug [identified in SACT free text treatments] MAb Cytotoxic therapy PI IMiD Radiotherapy	Treatment record indicating use of: Investigational drug [labelled as “clinical study drug”] MAb Cytotoxic therapy PI IMiD Radiotherapy	Theoretical scenario not assessed in this data source	Appendix 3. MM Treatment Groupings	[-44 ,0] [-55, 0] [-55, 0] [-30, 0] [-37, 0] [-37, 0] Assumes apheresis lasts 30 days.
	6. Toxicity from previous anticancer therapy must resolve to baseline levels or to Grade 1 or less except for alopecia or peripheral neuropathy	**Uncertain feasibility** in NCRAS. To explore whether data are available relating to toxicity from previous anticancer therapy	Record in clinical notes indicating ongoing toxicity	Theoretical scenario not assessed in this data source	-	[-30, 0]

Theoretical Scenario	Trial Exclusion Criteria [Other criteria]	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Codes / Algorithms	Assessment Window
	7. Have following cardiac conditions: • NYHA stage III or IV congestive heart failure • Myocardial infarction or CABG ≤6 months prior to enrolment • History of clinically significant ventricular arrhythmia or unexplained syncope, not believed to be vasovagal in nature or due to dehydration • History of severe non-ischemic cardiomyopathy • Impaired cardiac function (LVEF <45%) as assessed by echocardiogram or MUGA scan (performed ≤8 weeks of apheresis)	Hospital record in HES with relevant diagnostic or procedure code for myocardial infarction or CABG, respectively OR Hospital record in HES with relevant diagnostic code for congestive heart failure, ventricular arrythmia, non-ischemic cardiomyopathy or impaired cardiac function	Diagnostic record for myocardial infarction or abstracted record from clinician notes indicating patient underwent CABG OR Diagnosis record for congestive heart failure, ventricular arrythmia, non-ischaemic cardiomyopathy or impaired cardiac function	Theoretical scenario not assessed in this data source	Appendix 1. Code Lists	[-213, 0] [-1825, 0]†
	8. Received a cumulative dose of corticosteroids equivalent to ≥70mg of prednisone within the 7 days prior to apheresis	**Uncertain feasibility** in SACT dataset. To explore whether records of corticosteroid use are captured in SACT	Treatment records indicating use of corticosteroid with dose equivalent to ≥70mg of prednisone	Theoretical scenario not assessed in this data source	Appendix 4. Corticosteroids and Equivalent Doses	[-37, 0] Assumes apheresis lasts 30 days.
	9. Received either of the following: • An allogenic stem cell transplant within 6 months before apheresis. Participants who received an allogeneic transplant must be off all immunosuppressive medications for 6 weeks without signs of GVHD • An autologous stem cell transplant ≤12 weeks before apheresis	Hospital record in HES with relevant procedure record for allogenic stem cell transplant within 7 months of index date OR Hospital record in HES with relevant procedure record for autologous stem cell transplant within 4 months of index date	Record abstracted from clinical notes for allogenic stem cell transplant within 7 months of index date OR Record abstracted from clinical notes for autologous stem cell transplant within 4 months of index date	Theoretical scenario not assessed in this data source	Appendix 1. Code Lists	[-213, 0] [122, 0] Assumes apheresis lasts 1 month (30 days)

Theoretical Scenario	Trial Exclusion Criteria [Other criteria]	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Codes / Algorithms	Assessment Window
	10. Known active, or prior history of CNS involvement or exhibits clinical signs of meningeal involvement of MM	**Uncertain feasibility** in NCRAS. To explore availability of records indicative of CNS involvement or meningeal involvement in MM	Record in clinical notes indicating CNS involvement or meningeal involvement in MM	Theoretical scenario not assessed in this data source	N/A	[-1825, 0]
	11. Stroke or seizure within 6 months of signing ICF	Hospital record in HES with relevant diagnostic code for stroke or seizure	Diagnostic record for stroke or seizure	Theoretical scenario not assessed in this data source	Appendix 1. Code Lists	[-183, 0]
	12. Plasma cell leukaemia at the time of screening (>2.0 x 10^9/L plasma cells by standard differential), Waldenström's macroglobulinemia, POEMs syndrome, or primary AL amyloidosis	Diagnostic record in NCRAS or HES for plasma cell leukaemia, Waldenström's macroglobulinemia, POEMs syndrome, primary AL amyloidosis	Diagnostic record for plasma cell leukaemia, Waldenström's macroglobulinemia, POEMs syndrome, primary AL amyloidosis	Theoretical scenario not assessed in this data source	Appendix 1. Code Lists	[-90, 0]
	13. Seropositive for human HIV	Hospital record in HES with relevant diagnostic code for HIV	Diagnostic record for HIV	Theoretical scenario not assessed in this data source	Appendix 1. Code Lists	[-1825, 0]
	14. Vaccinated with live, attenuated vaccine within 4 weeks prior to apheresis.	**Uncertain feasibility** in NCRAS. To explore availability of records indicative of vaccination.	Treatment record for live attenuated vaccine	Theoretical scenario not assessed in this data source	-	[-58, 0] Assumes apheresis lasts 1 month (30 days)
	15. Hepatitis B infection as defined according to the ASCO guidelines. In the event the infection status is unclear, quantitative levels are necessary to determine the infection status	Hospital record in HES with relevant diagnostic code for hepatitis B	Diagnostic record for hepatitis B	Theoretical scenario not assessed in this data source	Appendix 1. Code Lists	[-1825, 0]
	16. Hepatitis C or known to have a history of hepatitis C. If positive, further testing of quantitative levels to rule out positivity is required	Hospital record in HES with relevant diagnostic code for hepatitis C	Diagnostic record for hepatitis C	Theoretical scenario not assessed in this data source	Appendix 1. Code Lists	[-1825, 0]

Theoretical Scenario	Trial Exclusion Criteria [Other criteria]	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Codes / Algorithms	Assessment Window
	17. Supplemental oxygen use to maintain adequate oxygenation	Hospital record in HES with relevant diagnosis for dependence on supplemental oxygen or procedure code for supplemental oxygen	Record in clinician of patient records indicating requirement for supplemental oxygen	Theoretical scenario not assessed in this data source	Appendix 1. Code Lists	[-365, 0]
	18. Known life threatening allergies, hypersensitivity, or intolerance to JNJ-68284528 or its excipients, including DMSO	**Uncertain feasibility** in NCRAS. To explore availability of records indicative of allergies, hypersensitivity or intolerance	**Uncertain feasibility** in Flatiron Health. To explore availability of records indicative of allergies, hypersensitivity or intolerance	Theoretical scenario not assessed in this data source	-	-
	19. Serious underlying medical condition, such as: • Evidence of serious active viral, bacterial, or uncontrolled systemic fungal infection • Active autoimmune disease or a history of autoimmune disease within 3 years • Overt clinical evidence of dementia or altered mental status	Hospital record in HES with relevant diagnosis for: Severe infection Autoimmune disease Dementia	Diagnostic record for: Severe infection Autoimmune disease Dementia	Theoretical scenario not assessed in this data source	Appendix 1. Code Lists	[-30, 0] [-1095, 0] [-1825, 0]
	20. Any issue that would impair the ability of the subject to receive or tolerate the planned treatment at the investigational site, to understand informed consent or any condition for which, in the opinion of the investigator, participation would not be in the best interest of the subject (e.g., compromise the well-being) or that could prevent, limit, or confound the protocol-specified assessments.	Criterion **not relevant** for RW database study	Criterion **not relevant** for RW database study	Theoretical scenario not assessed in this data source	-	-
	21. Pregnant or breast-feeding or planning to become pregnant while enrolled in this study or within 100 days after receiving study treatment.	**Uncertain feasibility** in NCRAS. To be explored whether active pregnancies may be	Record abstracted from clinician notes indicating pregnancy	Theoretical scenario not assessed in this data source	Appendix 1. Code Lists	[-304, 0]

Theoretical Scenario	Trial Exclusion Criteria [Other criteria]	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Codes / Algorithms	Assessment Window
		identified in data source.				
	22. Plans to father a child while enrolled in this study or within 100 days after receiving study treatment	Criterion **not relevant** for RW database study	Criterion **not relevant** for RW database study	Theoretical scenario not assessed in this data source	-	-
	23. Major surgery within 2 weeks prior to apheresis, or has surgery planned during the study or within 2 weeks after study treatment administration	Criterion **not relevant** for RW database study	Criterion **not relevant** for RW database study	Theoretical scenario not assessed in this data source	-	-
Scenario 3: CARTITUDE-4 Trial Emulation [6]	1. Treatment with CAR-T cell therapy directed at any target	Theoretical scenario not assessed in this data source	Treatment record indicating CAR-T cell therapy administration	Theoretical scenario not assessed in this data source	Appendix 3. MM Treatment Groupings	[-1825, 0]
	2. Any previous therapy that is targeted to BCMA	Theoretical scenario not assessed in this data source	Treatment record indicating use of anti-BMCA agent	Theoretical scenario not assessed in this data source	Appendix 3. MM Treatment Groupings	[-1825, 0]
	3. Ongoing toxicity from previous anticancer therapy that has not resolved to baseline levels or to Grade 1 or less; except for alopecia	Theoretical scenario not assessed in this data source	Record in clinical notes indicating ongoing toxicity	Theoretical scenario not assessed in this data source	-	[-30, 0]
	4. Participants with Grade 1 peripheral neuropathy with pain or Grade 2 or higher peripheral neuropathy will not be permitted to receive PVd as standard therapy or bridging therapy; however, participants may receive DPd as standard therapy or bridging therapy	Theoretical scenario not assessed in this data source	Diagnostic record for peripheral neuropathy	Theoretical scenario not assessed in this data source	-	[-365, 0]
	5. Received a cumulative dose of corticosteroids equivalent to ≥70mg of prednisone within the 7 days prior to randomization	Theoretical scenario not assessed in this data source	Treatment records indicating use of corticosteroid with dose equivalent to ≥70mg of prednisone	Theoretical scenario not assessed in this data source	Appendix 4. Corticosteroids and Equivalent Doses	[-37, 0] Assumes apheresis lasts 30 days.
	6. Was vaccinated with live attenuated vaccines within 4 weeks prior to randomization	Theoretical scenario not assessed in this data source	Treatment record for live attenuated vaccine	Theoretical scenario not assessed in this data source	-	[-58, 0] Assumes randomizatio n occurs 1 month (30

Theoretical Scenario	Trial Exclusion Criteria [Other criteria]	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Codes / Algorithms	Assessment Window
						days) prior to index
	7. Targeted therapy, epigenetic therapy, or treatment with an investigational drug or used an invasive investigational medical device within 14 days or at least 5 half lives, whichever is less. • mAb) treatment for MM within 21 days. • Cytotoxic therapy within 14 days. • PI therapy within 14 days. • IMiD therapy within 7 days. • Radiotherapy within 14 days. However, if the radiation portal covered ≥5% of the bone marrow reserve, the subject is eligible irrespective of the end date of radiotherapy	Theoretical scenario not assessed in this data source	Treatment record indicating use of: Targeted therapy / epigenetic therapy MAb Cytotoxic therapy PI IMiD Radiotherapy	Theoretical scenario not assessed in this data source		 [-44 ,0] [-55, 0] [-55, 0] [-30, 0] [-37, 0] [-37, 0] Assumes apheresis lasts 30 days.
	8. Active malignancies (i.e., progressing or requiring treatment change in the last 24 months) other than the disease being treated under study. The only allowed exceptions are: • non-muscle invasive bladder cancer treated within the last 24 months that is considered completely cured. • Skin cancer (non-melanoma or melanoma) treated within the last 24 months that is considered completely cured. • Non-invasive cervical cancer treated within the last 24 months that is considered completely cured. • Localized prostate cancer (N0M0) • Breast cancer: adequately treated lobular carcinoma in situ or ductal carcinoma in situ, or history of	Theoretical scenario not assessed in this data source	**Uncertain feasibility** in Flatiron Health. To explore availability of records indicative of other active malignancies.	Theoretical scenario not assessed in this data source		

Theoretical Scenario	Trial Exclusion Criteria [Other criteria]	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Codes / Algorithms	Assessment Window
	localized breast cancer and receiving antihormonal agents and considered to have a very low risk of recurrence. • Malignancy that is considered cured with minimal risk of recurrence.					
	9. Plasma cell leukaemia at the time of screening, Waldenström's macroglobulinemia, POEMs syndrome or primary AL amyloidosis.	Theoretical scenario not assessed in this data source	Diagnostic record for plasma cell leukaemia, Waldenström's macroglobulinemia, POEMs syndrome, primary AL amyloidosis	Theoretical scenario not assessed in this data source	Appendix 1. Code Lists	[-90, 0]
	10. Contraindications or life-threatening allergies, hypersensitivity, or intolerance to JNJ-68284528 or its excipients, including DMSO (refer to Investigator's Brochure), or to fludarabine, cyclophosphamide, tocilizumab, pomalidomide, dexamethasone	Theoretical scenario not assessed in this data source	**Uncertain feasibility** in Flatiron Health. Feasibility to be explored.	Theoretical scenario not assessed in this data source	-	[-1825, 0]
	11. Pregnant or breast-feeding or planning to become pregnant while enrolled in this study or within 3 months of receiving the last dose of daratumumab or bortezomib, or within 28 days after the last dose of pomalidomide, whichever is later (Arm A) or at least within 1 year after receiving JNJ-68284528 infusion or at least within 3 months after receiving the last dose of daratumumab or bortezomib or within 28 days after the last dose of pomalidomide, whichever is later (Arm B)	Theoretical scenario not assessed in this data source	Record abstracted from clinician notes indicating pregnancy	Theoretical scenario not assessed in this data source	-	[-304, 0]
	12. Plans to father a child while enrolled in this study or within 3 months of receiving the last dose daratumumab or bortezomib, or within 28 days after the last dose of pomalidomide, whichever is later (Arm A) or at least	Theoretical scenario not assessed in this data source	Criterion **not relevant** for RW database study	Theoretical scenario not assessed in this data source	-	-

Theoretical Scenario	Trial Exclusion Criteria [Other criteria]	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Codes / Algorithms	Assessment Window
	within 1 year after receiving JNJ-68284528 infusion or at least within 3 months after receiving the last dose of daratumumab or bortezomib or within 28 days after the last dose of pomalidomide, whichever is later (Arm B).					
	13. Stroke or seizure within 6 months of signing ICF	Theoretical scenario not assessed in this data source	Diagnostic record for stroke or seizure	Theoretical scenario not assessed in this data source	-	[-183, 0]
	14. Received either of the following: • An allogenic stem cell transplant within 6 months before apheresis. Subjects who received an allogeneic transplant must have stopped all immunosuppressive medications for 6 weeks without signs of GVHD. Subjects with active GVHD are excluded. • An autologous stem cell transplantation ≤12 weeks before apheresis	Theoretical scenario not assessed in this data source	Record abstracted from clinical notes for allogenic stem cell transplant within 7 months of index date OR Record abstracted from clinical notes for autologous stem cell transplant within 4 months of index date	Theoretical scenario not assessed in this data source	Appendix 1. Code Lists	[-213, 0] [122, 0] Assumes apheresis lasts 1 month (30 days)
	15. Known active, or prior history of CNS involvement or exhibits clinical signs of meningeal involvement of MM	Theoretical scenario not assessed in this data source	Record in clinical notes indicating CNS involvement or meningeal involvement in MM	Theoretical scenario not assessed in this data source	-	[-1825, 0]
	16. Subject with COPD with a FEV1 <50% of predicted normal will not be able to receive DPd as standard therapy or bridging therapy; however, subject may receive PVd as standard therapy or bridging therapy.	Theoretical scenario not assessed in this data source	Diagnostic record for COPD	Theoretical scenario not assessed in this data source	-	[-1825, 0]
	17. Any of the following: • Seropositive HIV • Hepatitis B infection: In the event the infection status is unclear, quantitative viral levels are necessary to determine the infection status.	Theoretical scenario not assessed in this data source	Diagnostic record for HIV, hepatitis B or hepatitis C	Theoretical scenario not assessed in this data source	-	[-1825, 0]

Theoretical Scenario	Trial Exclusion Criteria [Other criteria]	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Codes / Algorithms	Assessment Window
	Subjects who are anti-HBs positive and without history of vaccination and subjects with positive anti-HBc and positive anti-HBs should have HBV-DNA quantification test • Hepatitis C infection (defined as anti-HCV antibody positive or HCV-RNA positive) or known to have a history of hepatitis C. For subjects with known history of HCV infection, confirmation of sustained virologic response is required for study eligibility, defined as ≥24 weeks after completion of antiviral therapy					
	18. Serious underlying medical or psychiatric condition or disease, that is likely to interfere with study procedures or results, or that in the opinion of the investigator would constitute a hazard for participating in this study, such as: • Requirement of supplemental oxygen to maintain oxygen saturation; • Evidence of serious active viral or bacterial infection, requiring systemic antimicrobial therapy, or uncontrolled systemic fungal infection; • Active autoimmune disease or a history of autoimmune disease within 2 years; • Clinical evidence of dementia or altered mental status; • Clinically significant cardiac disease, such as: ○ NYHA Class III or IV congestive heart failure ○ % Myocardial infarction or coronary-artery-bypass graft ≥6 months prior to enrolment	Theoretical scenario not assessed in this data source	Record in clinician of patient records indicating requirement for supplemental oxygen Diagnostic record for: Severe infection Autoimmune disease Dementia Diagnostic record for myocardial infarction or abstracted record from clinician notes indicating patient underwent CABG OR Diagnosis record for congestive heart failure, ventricular arrythmia, non-ischaemic	Theoretical scenario not assessed in this data source	-	[-365, 0] [-30, 0] [-1095, 0] [-1825, 0] [-213, 0] [-1825, 0]

Theoretical Scenario	Trial Exclusion Criteria [Other criteria]	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Codes / Algorithms	Assessment Window
	○ History of clinically significant ventricular arrhythmia or unexplained syncope, not believed to be vasovagal in nature or due to dehydration; ○ History of severe non-ischemic cardiomyopathy; ○ Impaired cardiac function (LVEF<45%) as assessed by echocardiogram or MUGA scan performed ≤8 weeks before randomization.		cardiomyopathy or impaired cardiac function			
	19. Major surgery within 2 weeks before randomization, or has not fully recovered from an earlier surgery, or has major surgery planned during the time the subject is expected to participate in the study.	Theoretical scenario not assessed in this data source	Criterion **not relevant** for RW database study		-	-
	20. Any issue that would impair the ability of the subject to receive or tolerate the planned treatment at the investigational site, to understand informed consent or any condition for which, in the opinion of the investigator, participation would not be in the best interest of the subject (e.g., compromise the well-being) or that could prevent, limit, or confound the protocol-specified assessments.	Theoretical scenario not assessed in this data source	Criterion **not relevant** for RW database study		-	-

Abbreviations: /L = per litre; AL amyloidosis = amyloid light-chain amyloidosis; anti-HBc = hepatitis B core antibodies; anti-HBs = hepatitis B surface antibody; anti-HCV = hepatitis C antibody; ASCO = American Society of Clinical Oncology; BCMA = B-cell maturation antigen; CABG = coronary artery bypass graft; CAR-T cell = chimeric antigen receptor T-cell; CNS = central nervous system; COPD = chronic obstructive pulmonary disease; CRS = cytokine release syndrome; DESCAR-T = Dispositif d'Enregistrement et Suivi des patients traités par CAR-T cells; DMSO = dimethyl sulfoxide; DPd = daratumumab, pomalidomide and dexamethasone; FEV1 = forced expiratory volume in 1 second; GVHD = graft-versus-host disease; HBV-DNA = hepatitis B virus deoxyribonucleic acid; HCV = hepatitis C virus; HCV-RNA = hepatitis C virus ribonucleic acid; HES = Hospital Episode Statistics; HIV = human immunodeficiency virus; ICD-10 = International Classification of Diseases, 10th Revision; ICF = informed consent form; IMiD = immunomodulatory imide drugs; LVEF = left ventricular ejection fraction; mAb = monoclonal antibody; mg = milligramme; MM = multiple myeloma; MUGA = multiple-gated acquisition; N/A = not applicable; NCRAS = National Cancer Registration and Analysis Service; NYHA = New York Heart Association; PI = proteasome inhibitor; POEMs syndrome = polyneuropathy, organomegaly,

endocrinopathy, monoclonal protein, and skin changes; PVd = pomalidomide, bortezomib and dexamethasone; RTDS = Radiotherapy Data Set; RW = real-world; SACT = Systemic Anti-Cancer Therapy.

5.3 Variables

In defining the exposures, outcomes and covariates that would be required for each theoretical scenario, a high-level assessment of NCRAS, Flatiron Health and DESCAR-T, the data source relevance (the extent to which a dataset presents the data elements useful to answer a given research question)[15], a key dimension of data quality, has been made.

5.3.1 Context and Rationale for Exposures of Interest

The exposures of interest are defined based on the therapeutic interventions received by patients and are aligned with the theoretical scenarios (Table 4).

Table 4. Operational Definitions of Exposure for each Theoretical Scenario

Theoretical Scenario	Exposure Group	Definition
1	Exposure: CAR-T cell therapy	Treatment record indicating new use of CAR-T cell therapy (Appendix 3. MM Treatment Groupings) during indexing period, regardless of treatment indication
2	Comparator: Standard of care	Treatment record indicating initiation of subsequent RRMM LOT having met inclusion criteria relating to prior LOTs (≥3 prior MM LOT OR are double refractory to an IMiD and PI AND received as part of previous therapy a PI, an IMiD, and an anti-CD38 antibody for the treatment of RRMM) during indexing period. RRMM treatments defined as a regimen containing any of the treatments listed in Appendix 3. MM Treatment Groupings
3	Exposure: Cilta-cel	New use of cilta-cel during indexing period
	Comparator: Standard of care	Treatment record indicating initiation of standard of care during indexing period, defined as either of the following treatment regimens: • PVd OR • DPd

Abbreviations: CAR-T cell = chimeric antigen receptor T-cell; DPd = daratumumab, pomalidomide and dexamethasone; IMiD = immunomodulatory imide drugs; LOT = lines of therapy; MM = multiple myeloma; PI = proteasome inhibitor; PVd = pomalidomide, bortezomib and dexamethasone; RRMM = relapsed/refractory multiple myeloma.

5.3.2 Context and Rationale for Outcome(s) of Interest

The fitness-for-use of each data source to identify each of the outcomes relevant to theoretical scenarios 1, 2 and 3 and presented in Table 5 will be assessed.

Note, no operational definitions for the outcomes are provided for the SNDS since SNDS will not be accessed for the fitness-for-purpose assessment.

Table 5. Outcomes Assessed in Each Theoretical Scenario and their Potential Operationalization in NCRAS and Flatiron Health

Theoretical Scenario	Trial Outcome [Other Outcome]	Trial Definition, if applicable	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Assessment Window	Follow-up Definition	Codes / Algorithms
1: Safety Evaluation	CRS	N/A	Diagnostic record in HES for CRS or treatment record indicating use of tocilizumab (first line treatment for CRS[21])	Diagnostic record for CRS or treatment record indicating use of tocilizumab	Diagnostic record for CRS or treatment record indicating use of tocilizumab	[0,14]	Follow up until the earliest of date of: • Outcome • Death • End of study period • Enrolment in clinical trial • Assessment window end	Appendix 1. Code Lists
	Neurotoxicity	N/A	Diagnostic record in HES for neurotoxicity	Diagnostic record for neurotoxicity	Diagnostic record for neurotoxicity	[0, 14]		
	Infection	N/A	Diagnostic record in HES for infection	Diagnostic record for infection	Diagnostic record for infection	[0, 365]		
	Neutropenia	N/A	Diagnostic record in HES for neutropenia	Diagnostic record for neutropenia or an ANC $<1.5*10^9$/L	Diagnostic record for neutropenia or an ANC $<1.5*10^9$/L	[0, 365]		
	Second Primary Malignancy	N/A	Diagnostic record in HES or NCRAS for second primary malignancy	Diagnostic record for second primary malignancy	Diagnostic record for second primary malignancy	[0, ∞]		
2: CARTITUDE-1 External Control Arm [5]	ORR	Achievement of PR or better according to IMWG criteria	**Uncertain feasibility** in NCRAS. To be explored	Evidence of PR or better in clinician notes	Theoretical scenario not assessed in this data source	[0, ∞]	Follow up until the earliest of date of: • Outcome • Death • End of study period • Enrolment in clinical trial • Initiation of subsequent LOT • Assessment window end	-
	Very Good or Partial Response or better response rate	Percentage of patients with VGPR or complete response according to IMWG criteria	**Uncertain feasibility** in NCRAS. To be explored	Evidence of a complete response or VGPR in clinician notes	Theoretical scenario not assessed in this data source			
	MRD negative rate	Percentage of patients with negative MRD by bone marrow aspirate	**Uncertain feasibility** in NCRAS. To be explored	Evidence of negative MRD in clinician notes	Theoretical scenario not assessed in this data source			

Theoretical Scenario	Trial Outcome [Other Outcome]	Trial Definition, if applicable	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Assessment Window	Follow-up Definition	Codes / Algorithms
	CBR	Percentage of patients with minimal response or better	**Uncertain feasibility** in NCRAS. To be explored	Evidence of best response or better in clinician notes	Theoretical scenario not assessed in this data source			
	Time to first response	Time between date of the initial infusion of JNJ-68284528 and the first efficacy evaluation that the participant met all criteria for PR or better	**Uncertain feasibility** in NCRAS. To be explored	Time between index date and RW response, defined as record for complete or partial response in clinician notes.	Theoretical scenario not assessed in this data source			
	DOR	Time from partial or better response to first evidence of PD or death	**Uncertain feasibility** in NCRAS. To be explored	Time between clinician-confirmed complete or partial response until date of progression.	Theoretical scenario not assessed in this data source	[Date of response record, ∞]	Evidence of response record until the earliest of date of: • Outcome • Death • End of study period • Enrolment in clinical trial • Initiation of subsequent LOT • Assessment window end	-
	PFS	Time from initial infusion to first documented PD	Two proxies for PFS will be applied: • Time from index date to	• Evidence of clinician note indicating progression	Theoretical scenario not assessed in this data source	[0, ∞]	Index date until the earliest of date of: • Evidence of progression	-

Theoretical Scenario	Trial Outcome [Other Outcome]	Trial Definition, if applicable	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Assessment Window	Follow-up Definition	Codes / Algorithms
			record in SACT indicating treatment discontinuation due to progression. • TTNTD, defined as time from index date to evidence of new treatment for MM or death	• TTNTD, defined as time from index date to evidence of new treatment for MM or death			• Death • End of study period • Enrolment in clinical trial • Initiation of subsequent LOT	
	OS	Time from index date to death from any cause	Time from index date to death from any cause	Time from index date to death (Flatiron Health-derived mortality variable version 2.0)	Theoretical scenario not assessed in this data source	[0, ∞]	until the earliest of date of: • Death • End of study period • Enrolment in clinical trial	-
3. CARTITUDE-4 Trial Emulation 6	PFS	Time from the date of randomization to the date of first documented disease progression as defined in the IMWG criteria, or death due to any cause, whichever occurs first.	Theoretical scenario not assessed in this data source	Evidence of clinician note with OR Evidence of new treatment for MM	Theoretical scenario not assessed in this data source	[0, ∞]	until the earliest of date of: • Evidence of progression • Death • End of study period • Enrolment in clinical trial • Initiation of subsequent LOT	-

Theoretical Scenario	Trial Outcome [Other Outcome]	Trial Definition, if applicable	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Assessment Window	Follow-up Definition	Codes / Algorithms
	Complete Response or stringent complete response	Complete Response or stringent complete response rate is defined as percentage of participants who achieve a Complete Response or stringent complete response according to the IMWG criteria.	Theoretical scenario not assessed in this data source	Evidence of a complete response or stringent complete response in clinician notes	Theoretical scenario not assessed in this data source			
	MRD negative rate	Percentage of patients with negative MRD by bone marrow aspirate	Theoretical scenario not assessed in this data source	Evidence of negative MRD during follow-up	Theoretical scenario not assessed in this data source			
	ORR	Achievement of PR or better according to IMWG criteria	Theoretical scenario not assessed in this data source	RW overall response defined as clinician-confirmed patient response, based upon radiological assessment	Theoretical scenario not assessed in this data source			
	OS	Time from index date to death from any cause	Theoretical scenario not assessed in this data source	Time from index date to death (Flatiron Health-derived mortality variable version 2.0)	Theoretical scenario not assessed in this data source	[0, ∞]	Index date until the earliest of date of: • Death • End of study period Enrolment in clinical trial	-

Theoretical Scenario	Trial Outcome [Other Outcome]	Trial Definition, if applicable	Operational Definition in NCRAS Cohort	Operational Definition in Flatiron Health Cohort	Operational Definition in DESCAR-T Cohort	Assessment Window	Follow-up Definition	Codes / Algorithms
	PFS on next LOT	Time interval between the date of randomization and date of event, which is defined as PD as assessed by investigator that starts after the next line of subsequent therapy, or death from any cause, whichever occurs first.	Theoretical scenario not assessed in this data source	Time from index date until the earliest of progression, death or initiation of new LOT	Theoretical scenario not assessed in this data source	[0, ∞]	Index date until the earliest of date of: • Event (death, progression or initiation of new LOT) • End of study period Enrolment in clinical trial	-

Abbreviations: /L = per litre; ANC = Absolute neutrophil count; CRS = cytokine release syndrome; DESCAR-T = Dispositif d'Enregistrement et Suivi des patients traités par CAR-T cells; CBR = Clinical Benefit Rate; DOR = duration of response; HES = Hospital Episode Statistics; IMWG = International Myeloma Working Group; LOT = lines of therapy; MM = multiple myeloma; MRD = minimal residual disease; N/A = not applicable; NCRAS = National Cancer Registration and Analysis Service; ORR = overall response rate; OS = overall survival; PD = progressive disease; PFS = progression-free-survival; PR = partial response; RW = real-world; SACT = Systemic Anti-Cancer Therapy; TTNTD = time to next treatment or death; VGPR = very good partial response.

5.3.3 Context and Rationale for Covariates

The fitness-for-use of the data sources for capturing key covariates relevant to theoretical scenarios 1, 2 and 3 presented in Table 6 will be assessed. These covariates are applicable to all theoretical scenarios unless otherwise stated.

Table 6. Covariates Required to Address Theoretical Scenarios 1, 2 and 3

Variable	Definition	Theoretical Scenario Applicable To
Demographics:		
Age at index date	Age in years at index date	1, 2, 3
Age at diagnosis of indication	Age in years at first diagnosis of haematological malignancy / indication	1, 2, 3
Sex	Patient sex (male, female)	1, 2, 3
Ethnicity	Patient ethnicity (white, Asian, black, mixed or multiple, other ethnic group)	1, 2, 3
Clinical Characteristics:		
Cancer/disease type	Diagnosis (indication) for CAR-T cell therapy	1
Cancer stage at index date	Cancer stage at index date (I, II, III, IV)	1, 2, 3
ECOG performance status	ECOG performance status at index date (0, 1, 2, 3, 4)	1, 2, 3
CCI	CCI at index date identified using diagnostic codes in Appendix 5. Charlson Comorbidity Score and its Components.	1, 2, 3
Refractory status	Refractoriness based upon number of IMiD, PIs and anti-CD38 therapies that patient had evidence of being refractory prior to index date (double, triple, quadruple, other)	1, 2, 3
Treatment-Related Variables		
Prior stem cell transplant	Evidence of prior stem cell transplant during baseline. Defined as a hospitalization record with evidence of patient undergoing procedure (OPCS-4 codes: W34.2, W34.3, W34.4, W34.5, W34.6, W34.8, W34.9, W35.8 W99.1, W99.8, W99.9, X33.5 X33.6, X33.8, X33.4) (ICD-10 codes Z94.6, T86.0)	1, 2, 3
Prior LOT	Number of LOT prior to index date (definition of LOT provided in Appendix 2. LOT definitions)	1, 2, 3
Time to progression on last regimen prior to the index treatment	Time in months between date of initiation of last LOT and index date	1, 2, 3
Prior IMiD treatment	Receipt of lenalidomide, pomalidomide or thalidomide prior to index date	1, 2, 3
Prior anti-CD38 antibody	Receipt of daratumumab or isatuximab prior to index date	1, 2, 3
Prior PI	Receipt of bortezomib, carfilzomib, ixazomib prior to index date	1, 2, 3
Prior lymphodepletion	Receipt of lymphodepletion therapy (cyclophosphamide, fludarabine, bendamustine, busilvex, alemtuzumab, etoposide)[22] in 7 days prior to index date	1, 2, 3
Prior radiotherapy	Receipt of any radiotherapy prior to index date	1, 2, 3
Concomitant medications	Receipt of other concomitant systemic anti-cancer therapies (e.g. 7 days prior to CAR-T infusion, and after CAR-T infusion until disease progression post-CAR-T) prior to index date	1, 2, 3
MM specific characteristics		
Cytogenetic profile	Risk profile based upon presence of cytogenic abnormalities (t(4;14), t(14;16), t(14;20), del(17p), del(1p), 1q21 prior to index date (high, standard, none)	2, 3

Variable	Definition	Theoretical Scenario Applicable To
International Staging System	Staging of MM according to serum β2M and serum albumin levels at index date (I, II, III)	2, 3
MM type	Type of MM (light chain, immunoglobulin G, other)	2, 3

Abbreviations: β2M = beta-2 microglobulin; CAR-T cell = chimeric antigen receptor T-cell; CCI = Charlson Comorbidity Index; ECOG = Eastern Cooperative Oncology Group; ICD-10 = International Classification of Diseases, 10th Revision; IMiD = immunomodulatory imide drugs; LOT = lines of therapy; MM = multiple myeloma; OPCS-4 = Office of Population Censuses and Surveys Classification of Surgical Operations and Procedures, Version 4; PI = proteasome inhibitor.

5.4 Data Analysis

5.4.1 Context and Rationale for Analysis Plan

Aligning with the approach recommended in the EMAs "Data Quality Framework for EU medicines: Application to Real-World data"[16], a series of metrics for four dimensions of data quality (extensiveness, coherence, reliability and timeliness) will be evaluated as part of the fitness-for-use assessment of NCRAS and Flatiron Health for each theoretical scenario.

5.4.1.1 Theoretical Scenario 1

To assess the fitness-for-use of the data sources for conducting theoretical scenario 1 (safety outcome analysis), the metrics presented in Table 7 will be assessed.

Descriptive statistics for key study elements presented in Table 8 will be derived for the study population for theoretical scenario 1. These descriptives will be subsequently compared with estimates from published literature (e.g., number of recipients of CAR-T cell therapy, number of observed safety events, age distribution) to further evaluate data source representativeness and coverage.

Where possible, the metrics specified in Tables 7 will be described qualitatively based upon Certara's prior experience. For Table 8, only data available from prior publications will be reported.

Table 7. Metrics to be Captured as Part of Detailed Fitness-for-Use Assessment for Theoretical Scenario 1 (adapted from Table 7 of the *Data Quality Framework for European Union (EU) medicines: Application to Real-World data*[16])

Design elements	Operational definition	Data elements for valid capture	Criticality of the quality of the element*	Data Quality Dimension (subdimension)	Metric
Study population	Inclusion criteria: Received a CAR-T cell therapy	Treatment information (CAR-T)	High	Extensiveness (Completeness)	Number and percentage of patients with CAR-T cell therapy available in the data source from the target population (country level).
				Reliability (precision)	Number and percentage of records for which CAR-T cell therapy administration date is present with day/month/year
				Extensiveness (completeness)	Percentage of records for which only CAR-T cell therapy administration date is available
				Reliability (accuracy)	Number of patients who received CAR-T cell therapy without a record for a known indication (haematological malignancy)
				Reliability (accuracy)	Number and percentage of CAR-T cell therapy administrations that occur prior to record for a known indication for the given therapy (first diagnostic record for haematological malignancy after date of first administration of CAR-T cell therapy)
				Reliability (accuracy)	Reliability of start and end dates of CAR-T and bridging therapy (free text, based on insights from data partner or publicly available information)
				Reliability (accuracy)	Number and percentage of CAR-T cell therapy administrations that occur during plausible date range according to date of death (administered CAR-T cell before date of death)
				Reliability (accuracy)	Number and percentage of CAR-T cell therapy administrations that occur during plausible date range according to dates of approval for use
				Coherence (uniqueness)	Number and percentage of patients with multiple CAR-T cell therapy administration records
	Inclusion criteria: 12 months data source history	Patient records	High	Extensiveness	Number and percentage of patients with data source history of at least 12 months
	Exclusion criteria: Participation in clinical trial	Clinical trial indicator variable in SACT or record in HES	Medium	Coherence (relational)	Number and percentage of patients with both an indicator for clinical trial involvement in SACT or diagnostic record for trial participation in HES *(Applicable to NCRAS only)*

Design elements	Operational definition	Data elements for valid capture	Criticality of the quality of the element*	Data Quality Dimension (subdimension)	Metric
	Exclusion criteria: Prior safety outcomes of interest (defined in Section 5.3.2 Context and Rationale for Outcome(s) of Interest)	Diagnostic and treatment records	High		See Key endpoints
	Cohort size	At least 96 patients (Section 9 Study Size)	Medium	Extensiveness	Number of patients meeting inclusion/exclusion criteria in relation to required cohort size
Treatment/exposure	Receipt of CAR-T cell therapy	Treatment information (CAR-T)	High	See above	See above
Comparator group	N/A	N/A	N/A	N/A	N/A
Key endpoints	Diagnostic records of safety outcomes of interest	Diagnostic and treatment records	High	Relevance	Number and percentage of patients in study population who developed each safety outcome at any time point
				Extensiveness (Completeness)	Completeness of each safety outcome (free text, based on insights from data partner or publicly available information)
				Reliability	Date of event captured (Number and percentage)
Other covariates	Type of CAR-T	Treatment information (CAR-T)	Medium	Extensiveness (completeness)	Number and percentage of patients in study population with evidence of CAR-T cell therapy use and with information about CAR-T cell type
	Type of cancer (indication)	Physician diagnosis (ICD-10 code or equivalent)	Medium	Extensiveness (completeness)	Number and percentage of patients in study population with ≥1 record indicating presence of haematological malignancy which is a known indication for CAR-T cell therapy
				Reliability (precision)	Number and percentage of patients in study population with multiple different recorded haematological malignancies
	Age at index date	Date of birth (year)	Medium	Extensiveness (completeness) /	Number and percentage of patients in study population with age data (e.g., complete date of birth)
				Reliability (precision)	Birth date is present with day/month/year available
				Reliability (accuracy)	Number and percentage of patients in study population with plausible age at CAR-T cell administration (i.e., ≤95 years)

Design elements	Operational definition	Data elements for valid capture	Criticality of the quality of the element*	Data Quality Dimension (subdimension)	Metric
	Sex	Sex variable	Low	Extensiveness (completeness)	Number and percentage of patients in study population with complete sex data
	Ethnicity	Ethnicity variable	Low	Extensiveness (completeness)	Number and percentage of patients in study population with complete ethnicity data
	Cancer stage at index date	Cancer variable staging	Medium	Extensiveness (completeness)	Number and percentage of patients with complete cancer stage data at index date
				Extensiveness (completeness)	Number and percentage of patients with date of cancer staging
	ECOG performance status	ECOG performance status variable	Medium	Extensiveness (completeness)	Number and percentage of patients in study population with complete ECOG data at index date
				Extensiveness (completeness)	Number and percentage of patients with date of ECOG performance assessment
				Reliability (accuracy)	Number and percentage of patients with plausible ECOG performance assessment score (i.e. ≥0 or ≤4)
	CCI	Diagnostic records	Medium	Extensiveness (completeness)	Completeness of CCI data (free text, based on insights from data partner or publicly available information)
	Refractory status	Treatment information and/or refractory status variable	Medium	Extensiveness (completeness)	Completeness of refractory data (free text, based on insights from data partner or publicly available information)
				Reliability (accuracy)	Number and percentage of patients meeting criteria for refractory but with subsequent record for therapy patient considered refractory to
	Prior stem cell transplant	Transplant data	Medium	Extensiveness (completeness)	Completeness of transplant data (free text, based on insights from data partner or publicly available information)
	Prior LOT	Treatment information and/or LOT variable	Medium	Extensiveness (completeness)	Completeness of LOT variable (free text, based on insights from data partner or publicly available information)
	Time to progression on last regimen prior to the index treatment	Progression data and/or treatment information	Medium	Extensiveness (completeness)	Completeness of progression data (free text, based on insights from data partner or publicly available information)
				Extensiveness (completeness)	Number and percentage of patients in study population with any record of progression prior to index treatment
	Prior treatments (IMiD, anti-CD38, PI, lymphodepletion,	Treatment information	Medium	Extensiveness (Completeness)	Presence of treatment records indicating use of each treatment (to explore whether each treatment is capturable in data source)

Design elements	Operational definition	Data elements for valid capture	Criticality of the quality of the element*	Data Quality Dimension (subdimension)	Metric
	concomitant medications)				
	Prior radiotherapy	Treatment information	Medium	Extensiveness (completeness)	Completeness of radiotherapy data (free text, based on insights from data partner or publicly available information)
Follow-up time	Time from index date to date of safety event of interest, enrolment in clinical trial, death or end of study period	Diagnostic records, treatment records, death date	High	Coherence (relational)	Number and percentage of patients with observable data (e.g., treatment data) after death

* The criticality of the quality of the element has been pre-specified and will inform the impact of the data quality issue to facilitate the qualification of the data source.
Abbreviations: CAR-T cell = chimeric antigen receptor T-cell; CCI = Charlson Comorbidity Index; ECOG = Eastern Cooperative Oncology Group; HES = Hospital Episode Statistics; ICD-10 = International Classification of Diseases, 10th Revision; IMiD = immunomodulatory imide drugs; LOT = lines of therapy; N/A = not applicable; NCRAS = National Cancer Registration and Analysis Service; PI = proteasome inhibitor; SACT = Systemic Anti-Cancer Therapy.

Table 8. Description of Key Study Elements in Data Sources of Study Population for Theoretical Scenario 1

Study Element	Estimated Statistics
Inclusion Criteria	
Number of CAR-T cell therapy recipients during indexing period	N (%)
Patients with a minimum of 12 months data source history	N (%)
Exclusion Criteria	
Any prior safety outcome of interest	N (% of prior exclusion criterion)
CRS	N (% of all patients with any prior safety outcome)
Neurotoxicity	N (% of all patients with any prior safety outcome)
Infection	N (% of all patients with any prior safety outcome)
Neutropenia	N (% of all patients with any prior safety outcome)
Second Primary Malignancy	N (% of all patients with any prior safety outcome)
Enrolled in clinical trial prior to index date	N (% of prior exclusion criterion)
Study Population	
Final study population	N (% of patients meeting inclusion criterion)
Exposures (among study population)	
CAR-T cell therapy infusion total recipients	N (%)
Tisagenlecleucel	N (%)
Brexucabtagene autoleucel	N (%)
Axicabtagene ciloleucel	N (%)
Cilta-cel	N (%)
Lisocabtagene maraleucel	N (%)
Idecabtagene vicleucel	N (%)
Year of receipt of CAR-T-cell	N (%) (categorical)
Outcomes (during follow up)	
CRS diagnosis	N (%)
Time to CRS diagnosis	Mean, standard deviation, median, interquartile range, range
Neurotoxicity diagnosis	N (%)
Time to neurotoxicity diagnosis	Mean, standard deviation, median, interquartile range, range
Infection diagnosis	N (%)
Time to infection diagnosis	Mean, standard deviation, median, interquartile range, range
Neutropenia diagnosis	N (%)
Time to neutropenia diagnosis	Mean, standard deviation, median, interquartile range, range
Second Primary Malignancy	N (%)
Time to second primary malignancy diagnosis	Mean, standard deviation, median, interquartile range, range
Follow-up time (from index date until date enrolment in clinical trial, death or end of study period)	Mean, standard deviation, median, interquartile range, range
Deaths	N (%) at 1-, 6-, 12- and 24-months and at end of follow-up
Time to death	Median, interquartile range, range
Clinical trial enrolment during follow up	N (%)
Time to clinical trial enrolment	Mean, standard deviation, median, interquartile range, range
Covariates	
Type of cancer (indication)	N (%) (categorical)
Age at index date	Mean, standard deviation, median, interquartile range, range
Age at diagnosis of indication	Mean, standard deviation, median, interquartile range, range
Sex	N (%) (categorical)

Study Element	Estimated Statistics
Ethnicity	N (%) (categorical)
Cancer stage at index date	N (%) (categorical)
ECOG performance status at index date	N (%) (categorical)
CCI at index date	Mean, standard deviation, median, interquartile range, range
Prior stem cell transplant	N (%)
Prior IMiD treatment	N (%)
Prior anti-CD38 antibody	N (%)
Prior PI	N (%)
Prior lymphodepletion	N (%)
Prior LOT	N (%)
Prior radiotherapy	N (%)
Concomitant medications	N (%) (categorical)

Abbreviations: CAR-T cell = chimeric antigen receptor T-cell; CCI = Charlson Comorbidity Index; CRS = cytokine release syndrome; ECOG = Eastern Cooperative Oncology Group; IMiD = immunomodulatory imide drugs; LOT = lines of therapy; N = number; PI = proteasome inhibitor

5.4.1.2 Theoretical Scenario 2

To assess the fitness-for-use of both data sources for conducting theoretical scenario 2 (external control arm – CARTITUDE-1), the metrics presented in Table 9 will be evaluated.

Descriptive statistics for key study elements presented in Table 10 will be derived for the study population for the theoretical scenario 2. These descriptives will be subsequently compared with estimates from published literature including CARTITUDE-1 (e.g., number of observed safety events) to evaluate data source coherence.

Where possible, the metrics specified in Tables 9 will be described qualitatively based upon Certara's prior experience. For Table 10, only data available from prior publications will be reported.

Table 9. Metrics to be Captured as Part of Detailed Fitness-for-Use Assessment for Theoretical Scenario 2 (adapted from Table 7 of the *Data Quality Framework for European Union (EU) medicines: Application to Real-World data* [16])

Design elements	Operational definition	Data elements for valid capture	Criticality of the quality of the element	Data Quality Dimension (subdimension, if applicable)	Metric
Study Population: *Inclusion Criteria*	Age ≥18 years at index date	Age	Medium	Extensiveness (completeness)	Number and percentage of patients in study population with complete age data (e.g., complete date of birth)
				Reliability (precision)	Birth date is present with day/month/year available
				Reliability (accuracy)	Number and percentage of patients in study population with plausible age at index date (i.e., ≤95 years)
	Have documented diagnosis of MM	Diagnostic record for MM	High	Extensiveness (completeness)	Number and percentage of MM patients whose diagnostic record for MM was via death certificate *(applicable to NCRAS only)*
	Measurable disease in 3 months prior to index date: Blood M-protein ≥1.0g/dL OR Urine M-protein ≥200mg/24h	Laboratory values for blood M-protein or urine M-protein	Medium	Extensiveness (completeness)	Completeness of laboratory data (free text, based on insights from data partner or publicly available information)
				Extensiveness (completeness)	Number and percentage of patients in study population with blood M-protein value prior to index date
				Extensiveness (completeness)	Number and percentage of patients in study population with serum M-protein value prior to index date
				Reliability (accuracy)	Number and percentage of patients in study population with plausible blood M-protein value prior to index date
				Reliability (accuracy)	Number and percentage of patients in study population with plausible serum M-protein value prior to index date
	Prior receipt of ≥3 LOTs or double refractory to an IMiD and PI	Treatment records	High	Extensiveness (completeness)	Completeness of LOT variable (free text, based on insights from data partner or publicly available information)
				Extensiveness (completeness)	Completeness of refractory data (free text, based on insights from data partner or publicly available information)
				Extensiveness (completeness)	Presence of treatment records indicating use of each treatment (to explore whether each treatment is capturable in data source)
				Reliability (accuracy)	Number and percentage of patients meeting criteria for refractory but with subsequent record for therapy
	Prior receipt of an IMiD, PI and an anti-CD38 antibody	Treatment records	High	Extensiveness (completeness)	Presence of treatment records indicating use of anti-CD38 treatment (to explore whether this treatment is capturable in data source)

Design elements	Operational definition	Data elements for valid capture	Criticality of the quality of the element	Data Quality Dimension (subdimension, if applicable)	Metric
	Documented disease progression	Progression data	Medium	Extensiveness (completeness)	Completeness of progression data (free text, based on insights from data partner or publicly available information)
				Extensiveness (completeness)	Number and percentage of patients with prior LOTs and with a recorded reason for treatment discontinuation in SACT *(applicable to NCRAS only)*
				Extensiveness (completeness	Number and percentage of patients without clinician notes indicating progression prior to index date
	Recorded ECOG Performance Status grade of 0 or 1 at index date	ECOG variable, including date of assessment	Medium	Extensiveness (completeness)	Number and percentage of patients in study population with missing ECOG data at index date
				Extensiveness (completeness)	Number and percentage of patients with date of ECOG performance assessment
				Reliability (accuracy)	Number and percentage of patients with plausible ECOG performance assessment score (i.e. ≥0 or ≤4)
	Laboratory values within specified ranges	Laboratory values	Low	Extensiveness (completeness)	Number and percentage of patients with complete data for each laboratory test
			Low	Reliability (accuracy)	Number and percentage of patients in study population with plausible values of each laboratory test of interest
	No evidence of pregnancy	Diagnostic or procedure records	Low	Extensiveness (completeness)	Presence of any records relating to active pregnancies during study period (to explore whether active pregnancies are capturable in data source)
	Initiation of a new line of EMA-approved therapy for RRMM	Treatment records	High	N/A	No applicable data quality metrics
Study Population: Exclusion Criteria	Participation in a clinical trial at index date	Clinical trial indicator variable	Medium	Coherence (relational)	Number and percentage of patients with both an indicator for clinical trial involvement in SACT or diagnostic record for trial participation in HES (*applicable to NCRAS only)*
				Extensiveness (Completeness)	Presence of any records relating to clinical trial enrolment during study period (to explore whether this is capturable in data source)
	Prior treatment with CAR-T cell therapy during baseline	Treatment records	High	Extensiveness (Completeness)	Number and percentage of patients with record for CAR-T cell therapy use at any time in the country available in the data source.
	Previous receipt of therapies targeting BCMA during baseline	Treatment records	Medium	Extensiveness (Completeness)	Presence of any records indicating anti-BCMA use during study period (to explore whether these treatments are capturable in data source)

Design elements	Operational definition	Data elements for valid capture	Criticality of the quality of the element	Data Quality Dimension (subdimension, if applicable)	Metric
	Diagnoses or treatment for invasive malignancy other than MM	Diagnostics records	Medium	Extensiveness (Completeness)	Completeness of diagnoses or treatment for other invasive malignancy (free text, based on data partner opinion and likelihood of completeness based on source of data)
	Prior diagnosis of cardiac condition	Diagnostic and procedure records	Medium	Extensiveness (Completeness)	Completeness of information regarding prior cardiac conditions (free text, based on data partner opinion and likelihood of completeness based on source of data)
	Receipt of investigational drug*, mAb, cytotoxic therapy, PI, IMiD or radiotherapy	Treatment records	Medium	Extensiveness (Completeness)	Presence of any records indicating use of each of the listed therapies during study period (to explore whether these treatments are capturable in data source)
	Toxicity from previous anti-cancer therapy	Diagnostic records	Medium	Extensiveness (Completeness)Extensiveness (Completeness)	Completeness of information regarding any toxicity from previous anti-cancer therapy (free text, based on data partner opinion and likelihood of completeness based on source of data)
	Receipt of corticosteroids	Treatment records	Medium	Extensiveness (Completeness)	Presence of any records indicating use of each of listed corticosteroid during study period (to explore whether these treatments are capturable in data source)
	Receipt of allogenic or autologous stem cell transplant	Treatment / procedure records	Medium	Extensiveness (Completeness)	Completeness of transplant data (free text, based on insights from data partner or publicly available information)
	CNS or meningeal involvement in MM	Clinical data	Medium	Extensiveness (Completeness)	Completeness of CNS or meningeal involvement diagnoses (free text, based on data partner opinion and likelihood of completeness based on source of data)
	Stroke	Diagnostic records	Medium	Extensiveness (Completeness)	Completeness of stroke diagnoses (free text, based on data partner opinion and likelihood of completeness based on source of data)
	Plasma cell leukaemia, Waldenström's macroglobulinemia, POEMs syndrome, primary AL amyloidosis	Diagnostic and procedure records	Medium	Extensiveness (Completeness)	Completeness of these diagnoses (free text, based on data partner opinion and likelihood of completeness based on source of data)
	HIV positivity	Diagnostic records	Medium	Extensiveness (Completeness)	Completeness of HIV diagnoses (free text, based on data partner opinion and likelihood of completeness based on source of data)
	Vaccinated with live, attenuated vaccine during pre-specified period prior to index date	Treatment records	Medium	Extensiveness (Completeness)	Completeness of vaccination records (free text, based on data partner opinion and likelihood of completeness based on source of data)

Design elements	Operational definition	Data elements for valid capture	Criticality of the quality of the element	Data Quality Dimension (subdimension, if applicable)	Metric
	Hepatitis B infection	Diagnostic records	Medium	Extensiveness (Completeness)	Completeness of Hepatitis B diagnoses (free text, based on data partner opinion and likelihood of completeness based on source of data)
	Hepatitis C infection	Diagnostic records	Medium	Extensiveness (Completeness)	Completeness of Hepatitis C diagnoses (free text, based on data partner opinion and likelihood of completeness based on source of data)
	Supplemental oxygen use	Treatment / procedure records	Medium	Extensiveness (Completeness)	Presence of any treatment or procedure records indicating use of supplemental oxygen during study period (to explore whether this is capturable in data source)
	Known life threatening allergies, hypersensitivity or intolerance to cilta-cel	Diagnostic records	Medium	Extensiveness (Completeness)	Completeness of information regarding any life-threatening allergies, hypersensitivity or intolerance to cilta-cel (free text, based on data partner opinion and likelihood of completeness based on source of data)
	Serious underlying medical condition	Diagnostic records	Medium	N/A	No applicable data quality metrics
	Pregnancy	Diagnostic records	Medium	N/A	See metric relating to capture of pregnancies above
Study Population	Cohort size	At least 54 patients (Section 9 Study Size)	High	Extensiveness	Number of patients meeting all eligibility criteria in relation to required cohort size
Treatment/exposure	N/A (comparator cohort only)	N/A	N/A	N/A	N/A
Comparator group	Received standard-of-care (excluding CAR-T)	Treatment records	High	N/A (See cohort size row)	N/A (See cohort size row)
Key outcomes	ORR, VGPR, CBR, MRD negative rate	Response data	High	Extensiveness (completeness)	Completeness of response data (free text, based on insights from data partner or publicly available information)
				Extensiveness (completeness)	Number and percentage of patients with any data on treatment response during follow-up
				Extensiveness (completeness)	Number and percentage of patients with any data on treatment response for index treatment
	OS	Death status; Date of death	High	Extensiveness (completeness)	Completeness of death data (free text, based on insights from data partner or publicly available information)
				Extensiveness (completeness)	Number and percentage of patients with complete date of death

Design elements	Operational definition	Data elements for valid capture	Criticality of the quality of the element	Data Quality Dimension (subdimension, if applicable)	Metric
				Reliability (accuracy)	Number and percentage of patients with any record (e.g., treatment or diagnostic records) after date of death
	PFS	Progression-related data	High	Extensiveness (completeness)	Completeness of progression data (free text, based on insights from data partner or publicly available information)
				Reliability (accuracy)	Number and percentage of patients in study population with record in SACT indicating progression as reason for treatment discontinuation during follow-up *(applicable to NCRAS only)*
				Extensiveness (completeness)	Number and percentage of patients in study population with any record of progression during follow-up
				Extensiveness (completeness)	Number and percentage of patients in study population with any record of progression prior to subsequent LOT
Other covariates	Age at diagnosis of MM	Date of birth (year)	Low	Extensiveness (completeness)	Number and percentage of patients with complete date of birth date
				Reliability (accuracy)	Number and percentage of patients in study population with plausible age at diagnosis of MM (i.e., ≤95 years)
	Sex	Sex variable	Low	Extensiveness (completeness)	Number and percentage of patients in study population with incomplete sex data
	Ethnicity	Ethnicity variable	Low	Extensiveness (completeness)	Number and percentage of patients in study population with complete ethnicity data
	Cancer stage at index date	Cancer variable staging	Medium	Extensiveness (completeness)	Number and percentage of patients with complete cancer stage data at index date
				Extensiveness (completeness)	Number and percentage of patients with date of cancer staging
	ECOG performance status	ECOG performance status variable	Medium	N/A	See inclusion criteria
	CCI	Diagnostic records	Medium	Reliability (accuracy)	Number and percentage of patients with plausible ECOG performance assessment score (i.e. ≥0 or ≤4)
	Refractory status	Treatment information and/or refractory status variable	Medium	N/A	See inclusion criteria
	Prior stem cell transplant	Transplant data	Medium	N/A	See inclusion criteria

Design elements	Operational definition	Data elements for valid capture	Criticality of the quality of the element	Data Quality Dimension (subdimension, if applicable)	Metric
	Prior LOT	Treatment information and/or LOT variable	Medium	N/A	See inclusion criteria
	Time to progression on last regimen prior to the index treatment	Progression data and/or treatment information	Medium	Extensiveness (completeness)	Completeness of progression data (free text, based on insights from data partner or publicly available information)
				Extensiveness (completeness)	Number and percentage of patients in study population with any record of progression prior to index treatment
	Prior treatments (IMiD, anti-CD38, PI, lymphodepletion, concomitant medications)	Treatment information	Medium	N/A	See inclusion criteria
	Prior radiotherapy	Treatment information	Medium	Extensiveness (completeness)	Completeness of radiotherapy data (free text, based on insights from data partner or publicly available information)
	MM international staging	Cancer variable staging	Medium	Extensiveness (completeness)	Number and percentage of patients with complete cancer stage data at index date
				Extensiveness (completeness)	Number and percentage of patients with date of cancer staging
	Cytogenetic profile	Record of cytogenic abnormalities	Low	Extensiveness (completeness)	Number and percentage of patients in study population with record relating to cytogenic abnormalities
	MM type	MM type	Low	Extensiveness (completeness)	Number and percentage of patients in study population with record relating to MM type
Follow-up time	From index date until date of effectiveness outcome of interest, subsequent LOT, clinical trial enrolment, death or end of study	Data on clinical trial enrolments and deaths	Medium	Coherence (relational)	Number and percentage of patients with observable data (e.g., treatment data) after death

Abbreviations: AL amyloidosis = amyloid light-chain amyloidosis; BCMA = B-cell maturation antigen; CAR-T cell = chimeric antigen receptor T-cell; CBR = Clinical Benefit Rate; CCI = Charlson Comorbidity Index; CNS = central nervous system; ECOG = Eastern Cooperative Oncology Group; EMA = European Medicines Agency; g/dL = gram per decilitre; HES = Hospital Episode Statistics; HIV = human immunodeficiency virus; IMiD = immunomodulatory imide drugs; LOT = lines of therapy; MM = multiple myeloma; mAb = monoclonal antibody; M-protein = monoclonal paraprotein; MRD = minimal residual disease; N/A = not applicable; NCRAS = National Cancer Registration and Analysis Service; ORR = overall response rate; OS = overall survival; PFS = progression-free-survival; PI = proteasome inhibitor; POEMs = polyneuropathy, organomegaly, endocrinopathy, monoclonal protein, and skin changes; RRMM = relapsed/refractory multiple myeloma; SACT = Systemic Anti-Cancer Therapy; VGPR = very good partial response; mg/24h = milligramme per 24 hours.

Table 10. Description of Key Study Elements in Data Sources of Study Population for Theoretical Scenario 2

Study Element	Estimated Statistics
Inclusion Criteria (applied sequentially, if operational)	
Documented diagnosis of MM	N (%)
Measurable Disease	N (% from prior criterion)
Receipt of ≥3 LOT OR double refractory to PI or IMiD	N (% from prior criterion)
Receipt of PI, IMiD and an anti-CD38 antibody in prior LOT	N (% from prior criterion)
Receipt of subsequent LOT	N (% from prior criterion)
Adult at index date	N (% from prior criterion)
Documented PD	N (% from prior criterion)
ECOG Status 0 or 1	N (% from prior criterion)
Haemoglobin ≥8g/dL	N (% from prior criterion)
Platelet count ≥50×10^9/L in participants in whom ≥50% of bone marrow nucleated cells are plasma cells	N (% from prior criterion)
ANC ≥0.75×10^9/L	N (% from prior criterion)
AST and ALT ≤3×ULN	N (% from prior criterion)
Creatinine clearance ≥40mL/min per 1.73m^2	N (% from prior criterion)
Total bilirubin ≤2.0×ULN	N (% from prior criterion)
Serum calcium ≤12.5mg/dL (≤3.1mmol/L) or free ionized calcium ≤6.5mg/dL (≤1.6mmol/L)	N (% from prior criterion)
No evidence of pregnancy	N (% from prior criterion)
Exclusion Criteria (applied sequentially, if operational)	
Patients meeting inclusion criteria 1 enrolled in clinical trial prior to index date	N (%)
Any prior treatment with CAR-T cell therapy	N (% from prior criterion)
Any prior receipt of anti-BCMA	N (% from prior criterion)
Diagnoses or treated for invasive malignancy other than MM	N (% from prior criterion)
Receipt of investigational drug*, mAb, cytotoxic therapy, PI, IMiD or radiotherapy during pre-specified periods prior to index date	N (% from prior criterion)
Toxicity from previous anti-cancer therapy	N (% from prior criterion)
Pre-specified cardiac condition	N (% from prior criterion)
Receipt of corticosteroids during pre-specified period prior to index date	N (% from prior criterion)
Receipt of allogenic or autologous stem cell transplant during pre-specified period prior to index date	N (% from prior criterion)
CNS or meningeal involvement in MM	N (% from prior criterion)
Stroke during pre-specified period prior to index date	N (% from prior criterion)
Plasma cell leukaemia, Waldenström's macroglobulinemia, POEMs syndrome, primary AL amyloidosis prior to index date	N (% from prior criterion)
HIV positivity	N (% from prior criterion)
Vaccinated with live, attenuated vaccine during pre-specified period prior to index date	N (% from prior criterion)
Hepatitis B infection	N (% from prior criterion)
Hepatitis C infection	N (% from prior criterion)
Supplemental oxygen use	N (% from prior criterion)
Known life threatening allergies, hypersensitivity or intolerance to cilta-cel	N (% from prior criterion)
Serious underlying medical condition	N (% from prior criterion)
Pregnancy	N (% from prior criterion)
Study Population	
Final study population	N (% of first inclusion criterion)
Exposures (among study population)	
Standard of care regimen type	N (%) (categorical)
Year of initiation of regimen	N (%) (categorical)

Study Element	Estimated Statistics
Outcomes during follow-up (among study population, if operational in data source)	
ORR: Patients with PR or better	N (%) at 1-, 6-, 12- and 24-months and at end of follow-up
VGPR: Patients with very good or partial response	N (%) at 1-, 6-, 12- and 24-months and at end of follow-up
MRD negative rate: Patients with negative MRD	N (%) at 1-, 6-, 12- and 24-months and at end of follow-up
CBR: Patients with minimal response or better	N (%) at 1-, 6-, 12- and 24-months and at end of follow-up
Time to first PR or better	Mean, standard deviation, median, interquartile range, range
DOR	Mean, standard deviation, median, interquartile range, range
Progression-free survival: Patients with record of progression	N (%) at 1-, 6-, 12- and 24-months and at end of follow-up
Progression-free survival: Time to record of progression	Mean, standard deviation, median, interquartile range, range
Progression-free survival: Patients with record of subsequent LOT	N (%) at 1-, 6-, 12- and 24-months and at end of follow-up
Progression-free survival: Time to subsequent LOT	Mean, standard deviation, median, interquartile range, range
OS: Deaths	N (%) at 1-, 6-, 12- and 24-months and at end of follow-up
Time to death	Median, interquartile range, range
Follow-up time (from index date until date to enrolment in clinical trial, subsequent LOT, death or end of study period)	Mean, standard deviation, median, interquartile range, range
Clinical trial enrolment during follow up	N (%)
Time to clinical trial enrolment	Mean, standard deviation, median, interquartile range, range
Covariates	
Age at index date	Mean, standard deviation, median, interquartile range, range
Age at diagnosis of indication	Mean, standard deviation, median, interquartile range, range
Sex	N (%) (categorical)
Ethnicity	N (%) (categorical)
MM stage at index date	N (%) (categorical)
MM type	N (%) (categorical)
ECOG performance status at index date	N (%) (categorical)
CCI at index date	Mean, standard deviation, median, interquartile range, range
Prior stem cell transplant	N (%)

Study Element	Estimated Statistics
Prior IMiD treatment	N (%)
Prior anti-CD38 antibody	N (%)
Prior PI	N (%)
Prior lymphodepletion	N (%)
Prior LOT	N (%)
Prior radiotherapy	N (%)
Concomitant medications	N (%) (categorical)
Cytogenic profile	N (%) (categorical)

Abbreviations: /L = per litre; AL amyloidosis = amyloid light-chain amyloidosis; ALT = alanine aminotransferase; ANC = absolute neutrophil count; AST = aspartate aminotransferase; BCMA = B-cell maturation antigen; CAR-T cell = chimeric antigen receptor T-cell; CBR = Clinical Benefit Rate; CCI = Charlson Comorbidity Index; CNS = central nervous system; DOR = duration of response; ECOG = Eastern Cooperative Oncology Group; HIV = human immunodeficiency virus; IMiD = immunomodulatory imide drugs; LOT = lines of therapy; m2 = square metre; mg/dL = milligramme per decilitre; mAb = monoclonal antibody; mL/min = millilitre per minute; MM = multiple myeloma; mmol/L = millimoles per litre; MRD = minimal residual disease; N = number; ORR = overall response rate; OS = overall survival; PD = progressive disease; PI = proteasome inhibitor; POEMs = polyneuropathy, organomegaly, endocrinopathy, monoclonal protein, and skin changes; PR = partial response: ULN = upper limit of normal; VGPR = very good partial response.

5.4.1.3 Theoretical Scenario 3

To assess the fitness-for-use of each data source for conducting theoretical scenario 3 (Target trial emulation – CARTITUDE-4), the metrics presented in Table 11 will be evaluated.

Descriptive statistics for key study elements presented in Table 12 will be derived for the study population for the theoretical scenario 2. These descriptives will be subsequently compared with estimates from published literature, including CARTITUDE-4 trial (e.g., age and sex distributions compared with CARTITUDE-4) to further evaluate data source representativeness and coverage.

Where possible, the metrics specified in Tables 11 will be described qualitatively based upon Certara's prior experience. For Tables 12, only data available from prior publications will be reported.

Table 11. Metrics to be Captured as Part of Detailed Fitness-for-Use Assessment for Theoretical Scenario 3 (adapted from Table 7 of the *Data Quality Framework for European Union (EU) medicines: Application to Real-World data* [16])

Design elements	Operational definition	Data elements for valid capture	Criticality of the quality of the element	Data Quality Dimension (subdimension)	Metric
Study Population: Inclusion Criteria	Age ≥18 years at index date	Age	Medium	Extensiveness (completeness)	Number and percentage of patients in study population with age data (e.g., incomplete date of birth)
				Reliability (accuracy)	Number and percentage of patients in study population with plausible age at index date (i.e., ≤95 years)
				Reliability (precision)	Birth date is present with day/month/year available
	Have documented diagnosis of MM	Diagnostic record for MM	High	Extensiveness (completeness)	Number and percentage of patients in study population without indication prior to index date
				Coherence	Misclassification of MM (free text, based on data partner opinion on likelihood of misclassification of MM)
	Measurable disease in 3 months prior to index date: Blood M-protein ≥1.0g/dL OR Urine M-protein ≥200mg/24h	Laboratory values for blood M-protein or urine M-protein	Medium	Extensiveness (Completeness)	Completeness of laboratory data (free text, based on insights from data partner or publicly available information)
				Extensiveness (completeness)	Number and percentage of patients in study population with value for each laboratory test prior to index date
				Reliability (accuracy)	Number and percentage of patients in study population with plausible values for each laboratory test prior to index date
	Prior receipt of 1 to 3 LOTs including a IMiD and PI	Treatment records	High	Extensiveness (completeness)	Completeness of LOT variable (free text, based on insights from data partner or publicly available information)
				Extensiveness (completeness)	Completeness of refractory data (free text, based on insights from data partner or publicly available information)
				Extensiveness (Completeness)	Presence of treatment records indicating use of each treatment (to explore whether each treatment is capturable in data source)
	Documented disease progression	Progression data	Medium	Extensiveness (completeness)	Completeness of progression data (free text, based on insights from data partner or publicly available information)
				Extensiveness (Completeness)	Presence of records for progression during study period (to assess whether progression is capturable in data source)
	Lenalidomide-refractory	Treatment data	High	Extensiveness (Completeness)	Presence of records indicating lenalidomide-refractory during study period (to assess whether this is capturable in data source)

Design elements	Operational definition	Data elements for valid capture	Criticality of the quality of the element	Data Quality Dimension (subdimension)	Metric
				Reliability (accuracy)	Number and percentage of patients meeting the criteria for lenalidomide-refractory but with subsequent lenalidomide treatment
	Recorded ECOG Performance Status grade of 0 or 1 at index date	ECOG variable, including date of assessment	Medium	Extensiveness (completeness)	Number and percentage of patients in study population with complete ECOG data at index date
				Extensiveness (completeness)	Number and percentage of patients with date of ECOG performance assessment
				Reliability (accuracy)	Number and percentage of patients with plausible ECOG performance assessment score (i.e. ≥0 or ≤4)
	Laboratory values within specified ranges	Laboratory values	Low	Extensiveness (completeness)	Number and percentage of patients with complete data for each laboratory test
			Low	Reliability (accuracy)	Number and percentage of patients in study population with plausible values of each laboratory test of interest
Study Population: Exclusion Criteria	Participation in a clinical trial at index date	Clinical trial indicator variable	Medium	Extensiveness (Completeness)	Presence of any records relating to clinical trial enrolment during study period (to explore whether this is capturable in data source)
	Prior treatment with CAR-T cell therapy during baseline	Treatment records	High	Extensiveness (Completeness)	Number and percentage of patients with record for CAR-T cell therapy use at any time in the country available in the data source.
	Previous receipt of therapies targeting BCMA during baseline	Treatment records	Medium	Extensiveness (Completeness)	Presence of any records indicating anti-BCMA use during study period (to explore whether these treatments are capturable in data source)
	Toxicity from previous anti-cancer therapy	Diagnostic records	Medium	Extensiveness (Completeness)	Completeness of information regarding any toxicity from previous anti-cancer therapy (free text, based on data partner opinion and likelihood of completeness based on source of data)
	Peripheral neuropathy	Diagnostic records	Medium	Extensiveness (Completeness)	Completeness of peripheral neuropathy diagnoses (free text, based on data partner opinion and likelihood of completeness based on source of data)
	Receipt of corticosteroids	Treatment records	Medium	Extensiveness (Completeness)	Presence of any records indicating use of each listed corticosteroid during study period (to explore whether these treatments are capturable in data source)
	Vaccinated with live, attenuated vaccine during pre-specified period prior to index date	Treatment records	Medium	Extensiveness (Completeness)	Completeness of vaccination records (free text, based on data partner opinion and likelihood of completeness based on source of data)
	Receipt of investigational drug*, mAb, cytotoxic therapy, PI, IMiD or radiotherapy	Treatment records	Medium	Extensiveness (Completeness)	Presence of any records indicating use of each of the listed therapies during study period (to explore whether these treatments are capturable in data source)

Design elements	Operational definition	Data elements for valid capture	Criticality of the quality of the element	Data Quality Dimension (subdimension)	Metric
	Diagnoses or treatment for invasive malignancy other than MM	Diagnostics records	Medium	Extensiveness (Completeness)	Completeness of diagnoses or treatment for other invasive malignancy (free text, based on data partner opinion and likelihood of completeness based on source of data)
	Plasma cell leukaemia, Waldenström's macroglobulinemia, POEMs syndrome, primary AL amyloidosis	Diagnostic and procedure records	Medium	Extensiveness (Completeness)	Completeness of these diagnoses (free text, based on data partner opinion and likelihood of completeness based on source of data)
	Contraindications or life-threatening allergies, hypersensitivity to index treatments	Diagnostic records	Medium	Extensiveness (Completeness)	Completeness of information regarding any contraindications, life-threatening allergies or hypersensitivity (free text, based on data partner opinion and likelihood of completeness based on source of data)
	Pregnancy	Diagnostic records	Medium	N/A	See inclusion criteria
	Stroke	Diagnostic records	Medium	Extensiveness (Completeness)	Completeness of stroke diagnoses (free text, based on data partner opinion and likelihood of completeness based on source of data)
	Receipt of allogenic or autologous stem cell transplant	Treatment / procedure records	Medium	Extensiveness (Completeness)	Completeness of transplant data (free text, based on insights from data partner or publicly available information)
	CNS or meningeal involvement in MM	Clinical data	Medium	Extensiveness (Completeness)	Completeness of CNS or meningeal involvement diagnoses (free text, based on data partner opinion and likelihood of completeness based on source of data)
	COPD diagnosis	Diagnostic records	Medium	Extensiveness (Completeness)	Completeness of COPD diagnoses (free text, based on data partner opinion and likelihood of completeness based on source of data)
	HIV positivity, Hepatitis B or Hepatitis C	Diagnostic records	Medium	Extensiveness (Completeness)	Completeness of HIV, Hepatitis B or C diagnoses (free text, based on data partner opinion and likelihood of completeness based on source of data)
	Serious underlying conditions	Diagnostic records	Medium	N/A	No data quality metrics
Study Population	Cohort size	At least 109 patients (Section 9 Study Size)	High	Extensiveness	Number of patients meeting all eligibility criteria
Treatment/exposure	Cilta-cel initiation	Treatment records	High	Extensiveness (Completeness)	Number and percentage of patients with record for CAR-T cell therapy use at any time in the country available in the data source.

Design elements	Operational definition	Data elements for valid capture	Criticality of the quality of the element	Data Quality Dimension (subdimension)	Metric
				Extensiveness (Completeness)	Number and percentage of records for which CAR-T cell therapy as day/month/year available
Comparator group	Standard-of-care initiation (excluding CAR-T)	Treatment records	High	Extensiveness (completeness)	Completeness of standard-of-care data (free text, based on data partner opinion and likelihood of completeness based on source of data)
				Reliability	Reliability of standard-of-care start date (free text, based on insights from data partner or publicly available information)
Key outcomes	PFS	Progression-related data	High	Extensiveness (completeness)	Completeness of progression data (free text, based on data partner opinion and likelihood of completeness based on source of data)
				Extensiveness (completeness)	Number and percentage of patients in study population with any record of progression during follow-up
				Extensiveness (completeness)	Number and percentage of patients in study population with any record of progression prior to subsequent LOT
	Complete response rate, MRD response rate, ORR	Response data	High	Extensiveness (completeness	Completeness of response data (free text, based on data partner opinion and likelihood of completeness based on source of data)
				Extensiveness (completeness)	Number and percentage of patients with any data on treatment response during follow-up
				Extensiveness (completeness)	Number and percentage of patients with any data on treatment response for index treatment
	OS	Death status; Date of death	High	Extensiveness (completeness)	Completeness of death data (free text, based on data partner opinion and likelihood of completeness based on source of data)
				Extensiveness (completeness)	Number and percentage of patients with complete date of death
				Reliability (accuracy)	Number and percentage of patients with any record (e.g., treatment or diagnostic records) after date of death
Other covariates	Age at diagnosis of MM	Date of birth (year)	Low	Extensiveness (completeness)	Number and percentage of patients with complete date of birth data
				Reliability (accuracy)	Number and percentage of patients in study population with plausible age at diagnosis of MM (i.e., ≤95 years)
	Sex	Sex variable	Low	Extensiveness (completeness)	Number and percentage of patients in study population with complete sex data
	Ethnicity	Ethnicity variable	Low	Extensiveness (completeness)	Number and percentage of patients in study population with complete ethnicity data

Design elements	Operational definition	Data elements for valid capture	Criticality of the quality of the element	Data Quality Dimension (subdimension)	Metric
	Cancer stage at index date	Cancer variable staging	Medium	Extensiveness (completeness)	Number and percentage of patients with complete cancer stage data at index date
				Extensiveness (completeness)	Number and percentage of patients with date of cancer staging
	ECOG performance status	ECOG performance status variable	Medium	N/A	See inclusion criteria
	CCI	Diagnostic records	Medium	Reliability (accuracy)	Number and percentage of patients with plausible ECOG performance assessment score (i.e. ≥0 or ≤4)
	Refractory status	Treatment information and/or refractory status variable	Medium	N/A	See inclusion criteria
	Prior stem cell transplant	Transplant data	Medium	N/A	See inclusion criteria
	Prior LOT	Treatment information and/or LOT variable	Medium	N/A	See inclusion criteria
	Time to progression on last regimen prior to the index treatment	Progression data and/or treatment information	Medium	Extensiveness (completeness)	Completeness of progression data (free text, based on insights from data partner or publicly available information)
				Extensiveness (completeness)	Number and percentage of patients in study population with any record of progression prior to index treatment
	Prior treatments (IMiD, anti-CD38, PI, lymphodepletion, concomitant medications)	Treatment information	Medium	N/A	See inclusion criteria
	Prior radiotherapy	Treatment information	Medium	Extensiveness (completeness)	Completeness of radiotherapy data (free text, based on insights from data partner or publicly available information)
	MM international staging	Cancer variable staging	Medium	Extensiveness (completeness)	Number and percentage of patients with complete cancer stage data at index date
				Extensiveness (completeness)	Number and percentage of patients with date of cancer staging
	Cytogenetic profile	Record of cytogenic abnormalities	Low	Extensiveness (completeness)	Number and percentage of patients in study population with record relating to cytogenic abnormalities

Design elements	Operational definition	Data elements for valid capture	Criticality of the quality of the element	Data Quality Dimension (subdimension)	Metric
	MM type	MM type	Low	Extensiveness (completeness)	Number and percentage of patients in study population with record relating to MM type
Follow-up time	From index date until date of effectiveness outcome of interest, subsequent LOT, clinical trial enrolment, death or end of study	Data on clinical trial enrolments and deaths	Medium	Coherence (relational)	Number and percentage of patients with observable data (e.g., treatment data) after death

Abbreviations: AL amyloidosis = amyloid light-chain amyloidosis; BCMA = B-cell maturation antigen; CAR-T cell = chimeric antigen receptor T-cell; CCI = Charlson Comorbidity Index; CNS = central nervous system; COPD = chronic obstructive pulmonary disease; ECOG = Eastern Cooperative Oncology Group; g/dL = gram per decilitre; HIV = human immunodeficiency virus; IMiD = immunomodulatory imide drugs; LOT = lines of therapy; mg/24h = milligramme per 24 hours; MM = multiple myeloma; mAb = monoclonal antibody; M-protein = monoclonal paraprotein; MRD = minimal residual disease; N/A = not applicable; ORR = overall response rate; OS = overall survival; PFS = progression-free-survival; PI = proteasome inhibitor; POEMs = polyneuropathy, organomegaly, endocrinopathy, monoclonal protein, and skin changes.

Table 12. Description of Key Study Elements in Data Sources of Study Population for Theoretical Scenario 3

Study Element	Estimated Statistics
Inclusion Criteria (applied sequentially, if operational)	
Documented diagnosis of MM	N (%)
Measurable Disease	N (% from prior criterion)
Receipt of 1 to 3 LOT including PI or IMiD	N (% from prior criterion)
Documented PD	N (% from prior criterion)
Early disease progression among patients with 1 prior LOT	N (% from prior criterion)
Lenalidomide-refractory	N (% from prior criterion)
ECOG Status 0 or 1	N (% from prior criterion)
Haemoglobin ≥8g/dL	N (% from prior criterion)
ANC $\geq 1\times10^9$ /L	N (% from prior criterion)
Platelet count $\geq 75\times10^9$/L in participants in whom <50% of bone marrow nucleated cells are plasma cells;	N (% from prior criterion)
Platelet count $\geq 50\times10^9$/L in participants in whom ≥50% of bone marrow nucleated cells are plasma cells	N (% from prior criterion)
Lymphocyte count $\geq 0.3\times10^9$/L	N (% from prior criterion)
Aspartate aminotransferase ≤3×ULN	N (% from prior criterion)
ALT ≤3×ULN	N (% from prior criterion)
Total bilirubin ≤2.0×ULN	N (% from prior criterion)
Estimated glomerular filtration rate ≥40mL/min per 1.73m^2	N (% from prior criterion)
No evidence of pregnancy	N (% from prior criterion)
Adult at index date	N (% from prior criterion)
Patients meeting inclusion criteria	N (% of first inclusion criterion)
Exclusion Criteria (applied sequentially, if operational)	
Patients meeting inclusion criteria enrolled in clinical trial prior to index date	N (%)
Any prior treatment with CAR-T cell therapy	N (% from prior criterion)
Any prior receipt of anti-BCMA	N (% from prior criterion)
Toxicity from previous anti-cancer therapy	N (% from prior criterion)
Peripheral neuropathy	N (% from prior criterion)
Receipt of corticosteroids during pre-specified period prior to index date	N (% from prior criterion)
Vaccinated with live, attenuated vaccine during pre-specified period prior to index date	N (% from prior criterion)
Receipt of investigational drug*, mAb, cytotoxic therapy, PI, IMiD or radiotherapy during pre-specified periods prior to index date	N (% from prior criterion)
Diagnoses or treated for invasive malignancy other than MM	N (% from prior criterion)
Plasma cell leukaemia, Waldenström's macroglobulinemia, POEMs syndrome, primary AL amyloidosis prior to index date	N (% from prior criterion)
Contraindications or life-threatening allergies, hypersensitivity to index treatments	N (% from prior criterion)
Pregnancy	N (% from prior criterion)
Stroke during pre-specified period prior to index date	N (% from prior criterion)
Receipt of allogenic or autologous stem cell transplant during pre-specified period prior to index date	N (% from prior criterion)
CNS or meningeal involvement in MM	N (% from prior criterion)
COPD	N (% from prior criterion)
HIV positivity, Hepatitis B or Hepatitis C	N (% from prior criterion)
Serious underlying medical condition	N (% from prior criterion)
Study Population	
Final study population	N (% of first inclusion criterion)
Exposures (among study population)	
Receipt of standard of care regimen type (DPd vs. PVd)	N (%)

Study Element	Estimated Statistics
Receipt of cilta-cel	N (%)
Year of initiation of regimen (by arm)	N (%) (categorical)
Outcomes during follow-up (among study population, if operational in data source)	
Progression-free survival: Patients with record of progression	N (%) at 1-, 6-, 12- and 24-months and at end of follow-up
Progression-free survival: Time to record of progression	Mean, standard deviation, median, interquartile range, range
Progression-free survival: Patients with record of subsequent LOT	N (%) at 1-, 6-, 12- and 24-months and at end of follow-up
Progression-free survival: Time to subsequent LOT	Mean, standard deviation, median, interquartile range, range
Complete response: Patients with complete response or stringent complete response	N (%) at 1-, 6-, 12- and 24-months and at end of follow-up
MRD negative rate: Patients with negative MRD	N (%) at 1-, 6-, 12- and 24-months and at end of follow-up
ORR: Patients with PR or better	N (%) at 1-, 6-, 12- and 24-months and at end of follow-up
OS: Deaths	N (%) at 1-, 6-, 12- and 24-months and at end of follow-up
Time to death	Median, interquartile range, range
PFS on next LOT	Mean, standard deviation, median, interquartile range, range
Follow-up time (from index date until of enrolment in clinical trial, subsequent LOT, death or end of study period)	Mean, standard deviation, median, interquartile range, range
Clinical trial enrolment during follow-up	N (%)
Time to clinical trial enrolment	Mean, standard deviation, median, interquartile range, range
Covariates	
Age at index date	Mean, standard deviation, median, interquartile range, range
Age at diagnosis of indication	Mean, standard deviation, median, interquartile range, range
Sex	N (%) (categorical)
Ethnicity	N (%) (categorical)
MM stage at index date	N (%) (categorical)
MM type	N (%) (categorical)
ECOG performance status at index date	N (%) (categorical)
CCI at index date	Mean, standard deviation, median, interquartile range, range
Prior stem cell transplant	N (%)
Prior IMiD treatment	N (%)
Prior anti-CD38 antibody	N (%)

Study Element	Estimated Statistics
Prior PI	N (%)
Prior lymphodepletion	N (%)
Prior LOT	N (%)
Prior radiotherapy	N (%)
Concomitant medications	N (%) (categorical)
Cytogenic profile	N (%) (categorical)

Abbreviations: /L = per litre; ALT = alanine aminotransferase; ANC = absolute neutrophil count; BCMA = B-cell maturation antigen; CAR-T cell = chimeric antigen receptor T-cell; CCI = Charlson Comorbidity Index; CNS = central nervous system; COPD = chronic obstructive pulmonary disease; DPd = daratumumab, pomalidomide and dexamethasone; ECOG = Eastern Cooperative Oncology Group; g/dL = gram per decilitre; HIV = human immunodeficiency virus; IMiD = immunomodulatory imide drugs; LOT = lines of therapy; m^2 = square metre; mAb = monoclonal antibody; mL/min = millilitre per minute; MM = multiple myeloma; MRD = minimal residual disease; N = number; ORR = overall response rate; OR = overall survival; PD = progressive disease; PFS = progression-free-survival; PI = proteasome inhibitor; POEMs = polyneuropathy, organomegaly, endocrinopathy, monoclonal protein, and skin changes; PVd = pomalidomide, bortezomib and dexamethasone; ULN = upper limit of normal.

6 Data Sources

6.1 NCRAS and Linked Data Sources

General Data Protection Regulation (GDPR) compliant and pseudonymized data for the study will be sourced from the NCRAS and linked data sources, including the Hospital Episode Statistics (HES) data source. NCRAS is responsible for cancer registration (CR) and reporting for England. Access to its data is granted by the National Health Service (NHS) Digital through a data application process. The CR datasets from NCRAS consist of the patient dataset (with demographic details), tumour dataset (on tumour characteristics, including diagnoses codes, morphology, and stages), and treatment dataset (on treatments that were offered around the time of cancer diagnoses). To provide a complete overview of patients care through the NHS, NCRAS links to other national datasets which include the Systemic Anti-Cancer Therapy (SACT) dataset, Radiotherapy Data Set (RTDS) and the HES datasets. These datasets are managed independently but are linked by NCRAS before being released to research organizations through NHS Digital.

The SACT dataset contains clinical information on patients, both adults and children, receiving systemic antineoplastic therapies (including chemotherapies) in NHS England. SACT was initiated in 2012 to compile all systemic antineoplastic treatment data on all cancer patients received in acute inpatient and outpatient settings (including day cases) and the community. It also contains data on systemic therapies received within clinical trials. This study will use SACT data from 01/01/2013 because data collected in the prior year was part of a pilot scheme and potentially incomplete.

The RTDS was initiated in 2009 and collates data from all health service providers of radiotherapy services in England, included those delivered in the private sector under NHS payment.

The HES datasets are comprehensive data of the general care received by patients within NHS England secondary care sector: inpatient, outpatient, critical care, and historical Accident & Emergency (A&E) care. Its primary purpose is to collect data to record resource utilization by health service providers and to facilitate reimbursement as well as monitor healthcare provision. These datasets have been made available as secondary data sources for healthcare research by NHS Digital (data release body of NHS England). NCRAS has an agreement in place with NHS Digital for a data linkage on cancer patients only and the release of this data for research which will benefit cancer patients in general. HES datasets will provide information on comorbidities of hospitalized patients, as well as additional treatment data that can be identified using procedural codes, namely radiotherapies to supplement RTDS (if required), certain systemic antineoplastic agents to supplement SACT (if required) and surgeries.

These datasets are linked at the patient level by NCRAS using a pseudonymized NHS identifier. The linkage enables longitudinal tracking of diagnoses, treatments, outcomes, and healthcare utilization across care settings. The use of these data will potentially allow for a comprehensive RW assessment of the safety and effectiveness of CAR-T cell therapies within the English healthcare system.

The required datasets and their respective date ranges of events of interest and other recorded data items for this study are summarised in Table 13.

Table 13. NCRAS Datasets and their Data Range to be Used for this Fitness-for-Use Assessment

Dataset	Date Range	Required for Selection of Study Population for Either Theoretical Scenario	For Assessment of Study Objectives
CR (Patient dataset)	01/01/2013 to 31/12/2022 (or the latest available)	Yes, to enable the identification of adults (*theoretical scenario 2*)	Yes, to enable the fitness-for-use assessment of key covariates *(theoretical scenario 1 and 2)*
CR (Tumour dataset)	01/01/2013 to 31/12/2022 (or the latest available)	Yes, to enable the identification of MM patients *(theoretical scenario 2)*	Yes, to enable the fitness-for-use assessment of key covariates (*theoretical scenario 1 and 2)*
CR (Treatment dataset)	01/01/2013 to 31/12/2022 (or the latest available)	Yes, to identify index treatments *(theoretical scenario 1 and 2)*	Yes, to enable the fitness-for-use assessment of key covariates *(theoretical scenario 1 and 2)*
SACT	01/01/2013 to 31/12/2024 (or latest available)	Yes, to identify CAR-T cell administrations *(theoretical scenario 1), clinical trial involvement (theoretical scenario 1 and 2)* and RRMM treatment regimens *(theoretical scenario 1 and 2)*	Yes, to enable the fitness-for-use assessment of key outcomes *(theoretical scenario 2)* and covariates *(theoretical scenario 1 and 2)*
RTDS	01/01/2013 31/12/2024 (or latest available)	Yes, to identify prior receipt of radiotherapy *(theoretical scenario 2)*	Yes, to enable the fitness-for-use assessment of key covariates *(theoretical scenario 1 and 2)*
HES admitted care	01/01/2013 to 30/09/2024 (or latest available at the highest achievable completeness)	Yes, to identify prior comorbidities *(theoretical scenario 2)*	Yes, to capture safety outcomes *(theoretical scenario 1)*
HES outpatient	01/01/2013 to 30/09/2024 (or latest available at the highest achievable completeness)	Yes, to identify prior comorbidities *(theoretical scenario 2)*	Yes, to capture safety outcomes *(theoretical scenario 1)*
HES A & E	01/01/2013 to 01/03/2020 (or latest available at the highest achievable completeness)	Yes, to identify prior comorbidities *(theoretical scenario 2)*	Yes

Abbreviations; Accident & Emergency = A&E; CAR-T cell = chimeric antigen receptor T-cell; CR = cancer registration; HES = Hospital Episode Statistics; MM = multiple myeloma; NCRAS = National Cancer Registration and Analysis Service; RRMM = relapsed/refractory multiple myeloma; RTDS = Radiotherapy Data Set; SACT = Systemic Anti-Cancer Therapy.

6.2 Flatiron Health

Flatiron Health is a US electronic healthcare records data source. It captures longitudinal, patient-level, RWD of patients receiving care at academic and community cancer centres across the US. Data are collected from over 280 oncology practices at 800+ unique sites of care from all 50 states in the US and includes data for over 2 million patients from 2011.[20,23] This data source is updated monthly and includes patients from both community and academic care settings who meet the following criteria:

- At least one International Classification of Diseases, 9th Revision (ICD-9) or International Classification of Diseases, 10th Revision (ICD-10) cancer diagnosis code
- At least one cancer clinic encounter with documented clinical activity (e.g., vital signs, treatments)

Trained abstractors manually extract clinically relevant details from a variety of EHR documents, including clinician notes, pathology reports, and radiology results. Mortality data are compiled using a composite

variable utilising information from structured and unstructured EHR information, commercial datasets and the Social Security Death Index.

Flatiron currently maintains 19 disease-specific data sources, including MM. Disease-specific data have been abstracted and curated from structured and unstructured fields to provide detailed information about patient demographics, clinical characteristics, treatment and progression data.[24,25] Unstructured data are processed using both technology-enabled abstraction and artificial intelligence-based extraction methods, including natural language processing, machine learning, and large language models. For the purposes of theoretical scenarios 2 and 3, the existing MM cohort registry including over 10,000 patients will be utilized. Only variables that have already been abstracted will be included in the fitness-for-use evaluation (i.e., unabstracted data from clinical notes will not be derived for the purposes of this study)

6.3 DESCAR-T

DESCAR-T is a multicentre prospective and retrospective registry that has collected data for patients eligible for CAR-T cell therapy from 27 French centres administering CAR-T cell therapy since 01 July 2018 (NCT04328298).[26] The registry is sponsored by the Lymphoma Academic Research Organisation (LYSARC). Demographic and clinical data, including treatments and responses are collected from medical records in a standardized electronic case report. Patients are included in DESCAR-T retrospectively from 01 July 2018 if they received CAR-T cell therapy in a qualified centre prior to the centre's registration to DESCAR-T. Patients are included in DESCAR-T prospectively if they are deemed eligible for treatment with CAR-T for a haematological malignancy according to the contributing centre's multidisciplinary tumour board. Patients are followed from date of confirmed eligibility for CAR-T cell therapy up to 15 years after infusion.

Authorization was obtained from the National Commission for data protection and freedom of information (CNIL) to collect patient social security numbers to enable long term follow-up of patient vital status. DESCAR-T includes a limited number of patients who were not recipients of CAR-T cell therapy and therefore it's utility for theoretical scenarios 2 and 3 are limited.

6.4 SNDS

The SNDS contains comprehensive clinical data, including pharmacy data, with a coverage of 99% of the population in France. It holds pseudonymised individual data, including claims data for >65 million insures. It also includes demographic data (age, sex, vital status) and diagnoses from the affection de longue durée (ALD), for long-term disease database, coded according to the ICD-10. These data are linked with the same pseudonymised identifiers to the Programme de médicalisation des systèmes d'information (PMSI) database. The PMSI dataset on hospitals and other health facilities data includes inpatient data medical information, related diagnoses, medical procedures, imaging, external consultations, external acts performed, as well as some medicines and devices. Data are also linked to Centre d'épidémiologie sur les causes médicales de décès (CepiDc), the database on causes of death.

An extract of the SNDS will not be requested for the purpose of this fitness-for-use assessment. Instead, the extent to which the SNDS can be used to address each theoretical scenario will be assessed based upon Certara's prior experience, including publications.

7 Data Management

Data management procedures will focus on the secure handling, transformation, and quality control (QC) of data obtained from the selected RWD sources for the purpose of assessing their fitness-for-use in NIS on CAR-T cell therapies. The data used in this study will be retrospective, secondary-use patient-level data extracted from structured data sources and registries. All data are pseudonymized at the source prior to access by the study team and will be handled in accordance with applicable data protection legislation, including the GDPR in the EU and other national-level privacy requirements. Data extracts will include only the variables necessary for conducting the feasibility assessments across the three predefined scenarios. These variables will include demographic information, diagnostic codes, procedure codes, treatment records, and mortality data, where applicable. All data transformations, derivations (e.g., line-of-therapy algorithms), and variable mappings will be documented in a data specifications document and feasibility metadata log, in line with best practices for reproducibility and transparency. Data will be stored in secure, access-controlled environments with access provided to only relevant study team members, with all transformation and analysis steps implemented through reproducible and version-controlled scripts. The quality of the data will be assessed across the key dimensions of completeness, plausibility, and consistency, and findings will be incorporated into the fitness-for-use matrices associated with each scenario. No patient-level data will be transferred outside of the secure analysis environments. Only aggregate, de-identified results will be used for reporting the results of the feasibility assessment. All data management activities will be conducted in accordance with Good Pharmacoepidemiology Practices (GPP), the European Network of Centres for Pharmacoepidemiology & Pharmacovigilance (ENCePP) standards, and the EMA's 2023 Data Quality Framework.[16]

8 QC

QC procedures will be implemented to ensure the accuracy, consistency, and reliability of the data preparation, derivation of variables and analyses processes used for the feasibility assessment. As the purpose of this protocol is to assess whether selected RWD sources are suitable for regulatory-relevant studies, all QC measures will focus on ensuring that feasibility conclusions are based on valid and reproducible evidence. The feasibility assessment will involve the systematic evaluation of key data elements across the three scenarios, including population identification, variable availability, coding accuracy, and follow-up completeness. To support this, a multi-level QC process will be established to verify the accuracy of dataset construction, cohort definitions, variable mappings, and derived outcome measures.

All programming scripts used to prepare and analyse the data will be version-controlled and developed using reproducible methods. Independent validation of a sample of code and derived variables will be performed by a second analyst, where feasible, to detect discrepancies or logic inconsistencies. Data specifications, transformations, and intermediate outputs will be reviewed and documented in a centralized project logbook or metadata registry. QC will also include cross-checks of data completeness, missingness patterns, and internal consistency (e.g., logical date ordering between treatment initiation and outcomes). Where inconsistencies or limitations are identified, these will be recorded and flagged in

the scenario-specific feasibility matrices. This transparent approach allows for scenario-level judgments to be made with an understanding of the underlying data quality constraints.

Finally, QC activities will be conducted in alignment with best practices in pharmacoepidemiology, including guidelines from ENCePP and the International Society for Pharmacoepidemiology (ISPE) GPP. Any substantial issues encountered during feasibility evaluation will be documented in the final project report, with recommendations on data source suitability and areas requiring further investigation or refinement.

9 Study Size

The primary aim is to assess whether the data sources under review include a sufficient number of patients meeting the eligibility criteria for each of the three predefined scenarios, and whether the available sample size would be adequate to support future NIS addressing similar objectives.

For **Scenario 1** (safety assessment), the feasibility assessment will evaluate the number of patients with documented CAR-T cell therapy administration across all eligible hematologic malignancy indications. The ability to estimate incidence rates of predefined AEs will depend on both the number of exposed patients and the available follow-up time. While no formal minimum threshold is imposed, the assessment will consider whether sample size and event counts are likely sufficient to yield stable estimates in stratified analyses (e.g., by product, indication, or age group). A sample size of 96 would be sufficient to assess the rate of AEs of 50% with a confidence level of 95% and precision width of ±10%, while 384 patients treated with CAR-T would be required for a precision width of ±5% (Table 14).[27]

Table 14. Sample Size Needed to Assess AE Rates based on AE Frequency and Precision Width (theoretical scenario 1).

Frequency of AE	Precision Width: ±5%	Precision Width: ±10%
10%, 90%	138	35
20%, 80%	246	61
30%, 70%	323	81
40%, 60%	369	92
50%	**384**	**96**

Abbreviation; AE = adverse event.

For **Scenario 2** (external comparator for CARTITUDE-1), the feasibility of conducting meaningful comparative effectiveness analyses will depend on whether a comparator cohort of sufficient size can be identified, with complete outcome data.

To evaluate whether the available sample size in the external comparator arm would be sufficient to detect a meaningful difference in overall survival (OS) when compared with the CARTITUDE-1 trial population, we estimated the minimum number of patients required to detect a range of hazard ratios (HR =0.67, 0.5, and 0.25) using a two-sided log-rank test with 80% power and a 5% significance level (α =0.05). Under the assumption of 1:1 allocation with the CARTITUDE-1 and that 50% of patients in the external control arm will experience the event (next treatment or death) during the follow-up period, censoring rate of 30% per year and average length of follow-up of 2 years, the corresponding minimum sample sizes were calculated to be approximately 169, 54, and 12 patients, respectively, for detecting differences in survival outcomes

of HR =0.67, 0.5, and 0.25.[28] For reference, in the CARTITUDE-4 trial, the cilta-cel treatment arm had 4-fold reduction in the risk of disease progression compared to standard-of-care control arm, with a HR for progression-free survival of 0.26 (95% CI: 0.18-0.38) (Table 15).[9]

Table 15. Sample Size Requirements for Detecting Survival Differences in Theoretical Scenario 2.

HR	Events Required In The Control Arm	Sample Size in Control Arm Assuming Event Rate =50%	Sample Size in Control Arm Assuming Event Rate =30%
0.67	82	169	234
0.5	25	54	73
0.25	6	12	15

Abbreviation; HR = hazard ratio.

For **Scenario 3** (CARTITUDE-4 trial emulation), the study will assess whether an adequate number of patients with lenalidomide-refractory RRMM and 1–3 prior LOT can be identified in both the cilta-cel and comparison groups. Sample size sufficiency will be evaluated with respect to expected event rates for OS, and TTNTD.

Assuming 1:1 allocation between cilta-cel and standard-of-care groups, using a two-sided log-rank test with 80% power and a 5% significance level (α =0.05), an event rate of 50% in the control group, 30% censoring rate and two years of average length of follow-up, the estimated minimum number of patients required to detect a difference in survival outcomes of HR =1.5, 2.0, and 4.0 is of 339, 109, and 25 patients in the two groups combined (Table 16) .[28]

Table 16. Sample Size Requirements for Detecting Survival Differences in Theoretical Scenario 3.

HR	Events Required (both arms)	Event Rate =50% (both arms)	Event Rate =30% (both arms)
1.5	191	339	468
2.0	66	109	146
4.0	17	25	31

Abbreviation; HR = hazard ratio.

10 Limitations of the Methods

This protocol outlines a structured approach to evaluate whether NCRAS and Flatiron Health are fit-for-purpose for conducting NIS on CAR-T cell therapies. While the methodological frameworks for each of the three scenarios are based on established principles of pharmacoepidemiology and comparative effectiveness research, several limitations inherent to this approach must be acknowledged.

This fitness-for-use assessment is limited to three hypothetical use cases (safety surveillance, external control for CARTITUDE-1, target trial emulation of CARTITUDE-4). It does not evaluate the applicability of NCRAS and Flatiron Health for a broader range of CAR-T research questions or future scenarios.

Most CAR-T cell therapies have been introduced in the past few years, limiting the availability of longer-term follow-up data. Also, CAR-T cell therapies are provided at a limited number of specialist centres, leading to small cohorts and possible centre-related biases (e.g. patient selection, supportive care). While

this feasibility study aims to be systematic, it is not fully exhaustive. Moreover, the assessment may require future updates as CAR-T cell therapy use expands, and datasets and coding practices evolve.

These limitations will be systematically documented and evaluated during the feasibility assessment and reported transparently in the final project report. Where applicable, mitigation strategies or alternative analytic approaches will be proposed.

The fitness-for-use of the SNDS will be evaluated without direct access to SNDS data. Instead, a hybrid assessment will be conducted based on Certara's prior experience with this data source, including previously published research. As such, a comprehensive evaluation will not be performed.

11 Protection of Human Subjects

This study will not involve the collection of new data, interaction with patients, or any form of intervention. All data used in the assessment will be fully pseudonymized or de-identified at the source, prior to access by the research team.

11.1 NCRAS and Linked Data Sources

No direct identifiers (e.g., names, social security numbers) will be available at any stage of data processing or analysis. All data handling will take place within secure, access-controlled environments, in full compliance with the EU GDPR. Robust technical and organisational measures such as encryption, strict access controls, and regular security assessments will be implemented to protect data from unauthorised access, disclosure, alteration, or destruction, as required by NHS data security frameworks and GDPR.

In accordance with GPP and ENCePP Code of Conduct, the study team will not attempt to re-identify individuals or link records in a manner that would compromise subject anonymity. No patient-level outputs will be disseminated.

As the study relies exclusively on secondary use of existing data under appropriate legal and ethical authorizations, no additional consent will be sought from data subjects. A privacy notice is available on the Certara website: Privacy Notice for National Health Service (NHS) England Data Access Request Service (DARS) | Certara. This study will be listed in the privacy notice, so the patients can be informed and exercise their data protection rights.

Ethical approval is not required as research is limited to secondary use of information previously collected during normal care (without an intention to use it for research at the time of collection) This study will be classified as a non-Research Ethics Committees (REC) study but Health Research Authority (HRA) approval will be required as it involves access to NHS patient data.

11.2 Flatiron Health

All analyses will be conducted by Flatiron Health and only aggregate level data will be provided.

Data are deidentified in accordance with the HIPAA privacy rule and may include patient demographics, tumour type, diagnosis date, cancer stage, treatment, and other characteristics. Deidentified data are subject to obligations to prevent reidentification and protect patient confidentiality. For example, patient

cohorts of 5 or fewer patients are described as less than or equal to 5 (≤5) patients, and patients aged 85 years and older may have an adjusted birth year in the dataset or data reported as not available.

12 Reporting of AEs

This protocol does not involve the administration of any medical treatments, investigational products, or direct interaction with patients. As a result, no active surveillance, reporting, or monitoring of AE will be conducted during the study. While AEs of special interest are included as part of the hypothetical study scenarios, they will be assessed only in terms of data availability and operational definability. No actual safety analyses or signal detection activities will be performed. Any safety-related findings observed during the feasibility assessment—such as data inconsistencies or unexpected frequencies of coded events will be documented and reported in aggregate form, strictly for the purpose of evaluating data quality. These findings will not be interpreted as individual case safety reports and will not be subject to expedited reporting obligations under pharmacovigilance regulations.

13 Study Reporting

The study results will be communicated in the form of a final global study report and may be disseminated as an abstract and/or manuscript. The global study report will include the findings from this study as well as the parallel study. Accordingly, the general quality and research-question specific data quality will be presented for these data sources.

Individual authors will be expected to contribute intellectually to the conceptualization, design, analysis, and writing of the manuscript. All authors will have the right to remove themselves from the publication should they disagree with the findings or the reporting of the findings. Potential conflicts of interest will be disclosed by all authors.

The publication policy for this project will rigorously follow the “Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication”, which is available at www.ICMJE.org.

All authors should be able to respond affirmatively to the following statements about their participation in the research:

1. Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work
2. Drafting the work or revising it critically for important intellectual content
3. Final approval of the version to be published
4. Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

Acquisition of funding, collection of data, or general supervision of the research group does not justify authorship. Each author should have participated sufficiently in the work to take public responsibility for appropriate portions of the content.

14 References

1. Besliu C, Tanase AD, Rotaru I, et al. The Evolving Landscape in Multiple Myeloma: From Risk Stratification to T Cell-Directed Advanced Therapies. *Cancers (Basel)* 2025; **17**(3).
2. Gagelmann N, Merz M. Fast and furious: Changing gears on the road to cure with chimeric antigen receptor T cells in multiple myeloma. *Semin Hematol* 2024; **61**(5): 306-13.
3. Peery MR, Hill H, Sharps A, Zaver A, Moore DC. B-Cell Maturation Antigen-Directed Immunotherapies for the Treatment of Relapsed/Refractory Multiple Myeloma: A Review of the Literature and Implications for Clinical Practice. *Ann Pharmacother* 2025; **59**(5): 463-72.
4. Jommi C, Bramanti S, Pani M, Ghirardini A, Santoro A. CAR T-Cell Therapies in Italy: Patient Access Barriers and Recommendations for Health System Solutions. *Front Pharmacol* 2022; **13**: 915342.
5. Janssen. A Study of JNJ-68284528, a Chimeric Antigen Receptor T Cell (CAR-T) Therapy Directed Against B-Cell Maturation Antigen (BCMA) in Participants With Relapsed or Refractory Multiple Myeloma (CARTITUDE-1). 2018. https://clinicaltrials.gov/study/NCT03548207 (accessed 18-Apr-2025.
6. Janssen. A Study Comparing JNJ-68284528, a CAR-T Therapy Directed Against B-cell Maturation Antigen (BCMA), Versus Pomalidomide, Bortezomib and Dexamethasone (PVd) or Daratumumab, Pomalidomide and Dexamethasone (DPd) in Participants With Relapsed and Lenalidomide-Refractory Multiple Myeloma (CARTITUDE-4). 2020. https://www.clinicaltrials.gov/study/NCT04181827 (accessed 18-Apr-2025.
7. Certara. Fitness for purpose of data sources relevant for real-world data (RWD) studies on CAR-T cell therapy - Objective 2 Study Protocol, 2025.
8. Berdeja JG, Madduri D, Usmani SZ, et al. Ciltacabtagene autoleucel, a B-cell maturation antigen-directed chimeric antigen receptor T-cell therapy in patients with relapsed or refractory multiple myeloma (CARTITUDE-1): a phase 1b/2 open-label study. *Lancet* 2021; **398**(10297): 314-24.
9. San-Miguel J, Dhakal B, Yong K, et al. Cilta-cel or Standard Care in Lenalidomide-Refractory Multiple Myeloma. *N Engl J Med* 2023; **389**(4): 335-47.
10. Tix T, Subklewe M, von Bergwelt-Baildon M, Rejeski K. Survivorship in Chimeric Antigen Receptor T-Cell Therapy Recipients: Infections, Secondary Malignancies, and Non-Relapse Mortality. *Oncol Res Treat* 2025; **48**(4): 212-9.
11. Xu H, Guan C, Xu P, et al. Clinical efficacy and safety of combined anti-BCMA and anti-CD19 CAR-T cell therapy for relapsed/refractory multiple myeloma: a systematic review and meta-analysis. *Front Oncol* 2024; **14**: 1355643.
12. Yamshon S, Gribbin C, Alhomoud M, et al. Safety and Toxicity Profiles of CAR T Cell Therapy in Non-Hodgkin Lymphoma: A Systematic Review and Meta-Analysis. *Clin Lymphoma Myeloma Leuk* 2024; **24**(6): e235-e56 e2.
13. Zhu F, Wei G, Liu Y, et al. Incidence and Risk Factors Associated with Infection after Chimeric Antigen Receptor T Cell Therapy for Relapsed/Refractory B-cell Malignancies. *Cell Transplant* 2021; **30**: 9636897211025503.
14. Moore DC, Oxencis CJ, Shank BR. New and emerging pharmacotherapies for the management of multiple myeloma. *Am J Health Syst Pharm* 2022; **79**(14): 1137-45.
15. European Medicines Agency. Data Quality Framework for EU medicines regulation. 2023. https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/data-quality-framework-eu-medicines-regulation_en.pdf20/08/2024).
16. EMA. Data Quality Framework for EU medicines regulation: application to Real-World Data (EMA/503781/2024). 2024. https://www.ema.europa.eu/en/documents/other/draft-data-quality-framework-eu-medicines-regulation-application-real-world-data_en.pdf (accessed 22-Apr-2025.
17. European Medicines Agency. First two CAR-T cell medicines recommended for approval in the European Union. 2018.

18. Administation FaD. FDA approves ciltacabtagene autoleucel for relapsed or refractory multiple myeloma. 2022.
19. Elsada A, Zalin-Miller A, Knott C, Caravotas L. A registry study of relapsed or refractory multiple myeloma pre-exposed to three or more prior therapies including a proteasome inhibitor, an immunomodulatory agent and CD38-targeted monoclonal antibody therapy in England. *EJHaem* 2021; **2**(3): 493-7.
20. Martin T, Krishnan A, Yong K, et al. Comparative effectiveness of ciltacabtagene autoleucel in CARTITUDE-1 versus physician's choice of therapy in the Flatiron Health multiple myeloma cohort registry for the treatment of patients with relapsed or refractory multiple myeloma. *EJHaem* 2022; **3**(1): 97-108.
21. EMA. RoActemra, tocilizumab. 2025. https://www.ema.europa.eu/en/medicines/human/EPAR/roactemra (accessed 6-May-2025.
22. Lickefett B, Chu L, Ortiz-Maldonado V, et al. Lymphodepletion - an essential but undervalued part of the chimeric antigen receptor T-cell therapy cycle. *Front Immunol* 2023; **14**: 1303935.
23. Ma X, Long L, Moon S, Adamson BJS, Baxi SS. Comparison of Population Characteristics in Real-World Clinical Oncology Databases in the US: Flatiron Health, SEER, and NPCR. *medRxiv* 2023: 2020.03.16.20037143.
24. Dhakal B, Einsele H, Schecter JM, et al. Real-world treatment patterns and outcomes in relapsed/refractory multiple myeloma (1-3 prior lines): Flatiron database. *Blood Adv* 2024; **8**(19): 5062-71.
25. Castellanos EH, Wittmershaus BK, Chandwani S. Raising the Bar for Real-World Data in Oncology: Approaches to Quality Across Multiple Dimensions. *JCO Clinical Cancer Informatics* 2024; (8): e2300046.
26. Broussais F, Bay JO, Boissel N, et al. [DESCAR-T, a nationwide registry for patient treated by CAR-T Cells in France]. *Bull Cancer* 2021; **108**(10s): S143-s54.
27. Fleiss JL, Levin, B., Paik, M.C. Statistical Methods for Rates and Proportions; 2003.
28. Kohn MA, Senyak, J. Sample Size Calculators [website]. UCSF CTSI. 2025. https://sample-size.net/sample-size-survival-analysis/ (accessed 22-Apr-2025.
29. Ayuketang FA, Jäger, U. Management of Cytokine Release Syndrome (CRS) and HLH. . In: Kroger N, Gribben J, Chabannon C, Yakoub-Agha I, Einsele H, eds. The EBMT/EHA CAR-T Cell Handbook. Cham (CH); 2022.
30. Rees JH. Management of Immune Effector Cell-Associated Neurotoxicity Syndrome (ICANS) In: Kroger N, Gribben J, Chabannon C, Yakoub-Agha I, Einsele H, eds. The EBMT/EHA CAR-T Cell Handbook. Cham (CH); 2022.
31. Subklewe M, Benjamin R. Management of Myelotoxicity (Aplasia) and Infectious Complications. In: Kroger N, Gribben J, Chabannon C, Yakoub-Agha I, Einsele H, eds. The EBMT/EHA CAR-T Cell Handbook. Cham (CH); 2022.
32. Rajkumar SV, Harousseau JL, Durie B, et al. Consensus recommendations for the uniform reporting of clinical trials: report of the International Myeloma Workshop Consensus Panel 1. *Blood* 2011; **117**(18): 4691-5.

15 Appendices

15.1 Appendix 1. Code Lists

Codes for identifying diagnoses & procedures

Condition / Procedures	Codes
MM	ICD-O morphology code: 9732/3 ICD-10 code: C90.0 ICD-9 203.0x
Allogeneic stem cell transplantation	OPCS-4 codes: W34.2-W34.6, W99.x, X33.5 X33.6
Autoimmune Disease	See: Tunnicliffe, L and Warren-Gash, C (2022). *Clinical code list- autoimmune disease ICD10 codes.* London School of Hygiene & Tropical Medicine, London, UK. https://doi.org/10.17037/DATA.00002853.
Autologous Stem-cell transplantation	OPCS-4: X33.4, W34.1
CABG	OPCS-4: K40-K44
Cancer	ICD-10: C00.x–C26.x, C30.x–C34.x, C37.x–C41.x, C43.x, C45.x–C58.x, C60.x–C76.x, C81.x–C85.x, C88.x, C90.x–C97.x, C77.x–C80.x *Excluding C70.0, C44.0*
Congestive heart failure	ICD-10: I50.0
COPD	J40-J44
Dementia	F00.x–F03.x, F05.1, G30.x, G31.1
Hepatitis B	B16.x, B17.0, B18.0, B18.1
Hepatitis C	B17.1, B18.2
HIV	B20.x–B22.x, B24.x
Impaired cardiac function	ICD-10: R93.1, R94.3
Infections	A00-B99, J15-J18
Myocardial infarction	ICD-10: I20.x, I22.x I23.x
Non-ischaemic cardiomyopathy	ICD-10: I42.x
Plasma cell leukaemia	ICD-10: C90.1
POEMs syndrome	None
Pregnancy	To be confirmed
Primary AL amyloidosis	ICD-10: E85.81
Stroke	ICD-10: I60.x-I64.x
Seizure	ICD-10: G40-G41, R56
Supplemental oxygen	ICD-10: Z99.1 OPCS-4: E85.1, X52.8, X58.1
Ventricular arrythmia	ICD-10: I47.0
Waldenström's macroglobulinemia	ICD-10: C88.9
Outcomes	
CRS[29]	T88.7 (unspecified adverse effect of drug) R65.2 (severe sepsis) R50.9 (fever, unspecified)
Neurotoxicity[30]	G93.4 (encephalopathy, unspecified) G93.1 (anoxic brain damage) G40 (epilepsy and recurrent seizures) R41.0 (disorientation)

	R41.8 (other cognitive disturbances)
Infections[2,31]	As above
Neutropenia[2,31]	D70.1 (agranulocytosis) D70.9 (neutropenia) OPCS-4: X29.1 (administration of CSF)
Second primary malignancy[2]	C00-C97 (exclusion of primary haematological malignance recurrence: C90.x)

Abbreviations: AL amyloidosis = amyloid light-chain amyloidosis; CABG = coronary artery bypass graft; COPD = chronic obstructive pulmonary disease; CRS = cytokine release syndrome; CSF = colony-stimulating factor; HIV = human immunodeficiency virus; ICD-9 = International Classification of Diseases, 9th Revision; ICD-10 = International Classification of Diseases, 10th Revision; ICD-O = International Classification of Disease of Oncology; MM = multiple myeloma; OPCS-4 = Office of Population Censuses and Surveys Classification of Surgical Operations and Procedures, Version 4; POEMs syndrome = polyneuropathy, organomegaly, endocrinopathy, monoclonal protein, and skin changes; UK = United Kingdom.

15.2 Appendix 2. LOT definitions

NCRAS

The definition of the LOT in NCRAS will be based on the algorithm published by Elsada and colleagues.[19] Regimens will be extracted from SACT if associated with an ICD-10 diagnosis code C90 and dated between up to one month prior to the first cohort-relevant diagnosis as documented in the NCRAS dataset and the end of follow-up. Supportive therapy and trial regimens will be excluded. Regimens will be restricted to those that included at least one of the drugs in the Appendix 3. MM Treatment Groupings. Melphalan and cyclophosphamide monotherapy regimens (administered without steroids) will be excluded as they are typically used as conditioning regimens prior to transplant and therefore should not be considered for the derivation of LOTs. For the purposes of deriving LOTs, induction therapy, stem cell transplant (autologous or allogeneic) and consolidation/maintenance therapy will be counted as one line.

Change in a LOT is defined when:

- All drugs change between consecutive regimens; or,
- The composition of drug classes change between consecutive regimens; or,
- The composition of drugs change between consecutive regimens (but the drug classes remain unchanged); or,
- Any change in the composition of drugs between consecutive regimens where the treatment-free interval was >60 days. The treatment-free interval is the period between the end of a current, and start of a subsequent, regimen.

Exceptions to the LOT rules above were as follows:

- The addition or removal of steroids between consecutive regimens will not be considered when applying the change in treatment algorithm.
- Reductions from 'bortezomib + daratumumab + dexamethasone' to 'daratumumab' with or without steroids will be handled as movement to maintenance therapy in instances where the gap between regimens is ≤60 days (gaps larger than this will be considered unlikely to be attributable to a move to maintenance therapy). This shift to maintenance therapy will not be considered as an initiation of a new LOT.
- LOTs for patients who receive a stem cell transplant will be derived in the same manner except the treatment-free interval for any change in the composition of drugs (e.g., those used for induction therapy) will be extended to >100 days for determining a change from first-line to second-line therapy. From second line the interval will be reverted to >60 days. This 100-day gap between first- and second-line treatment allows for induction therapy, stem cell transplant and consolidation/maintenance therapy to be counted as one line.

Regimen duration will span the earliest known cycle or administration start date and last known cycle or administration date for the regimen. Where regimens contain a single cycle or administration start date (regimen duration 0 days), the duration will be imputed as the 'usual' cycle duration for that regimen.

In agreement with International Myeloma Working Group (IMWG) criteria, patients will be classified as refractory to an immunomodulatory imide drugs (IMiD) (e.g., thalidomide, lenalidomide, pomalidomide) or a proteasome inhibitor (PI) (e.g., bortezomib, carfilzomib) if a new LOT that excludes the respective IMiD or PI is initiated within 60 days of completing the last line containing that agent.[32] Double-refractory patients are those patients who are refractory to at least one IMiD and at least one PI.

Flatiron Health

The definition of the LOT in Flatiron Health will be based on their in-house algorithm. Consequently, the LOT will be defined as the first eligible drug episode plus other eligible drugs given within 28 days, allowing for up to 3 days of overlap between abstracted agents and other drugs without rolling them into a single line. Systemic treatment, as evidenced by an order or administration of an antineoplastic or steroid (IV dexamethasone, oral dexamethasone, and oral prednisone) recorded in the EHR, are included in LOT. Full documentation of this algorithm (and whether it was validated) will be obtained from Flatiron Health during the study.

15.3 Appendix 3. MM Treatment Groupings

MM Treatment Name	Class
Bendamustine (Belrapzo®/Bendeka®/Treanda® / Levact® / Ribomustin®/ Vivimusta®)	Alkylating agent (Cytotoxic)
Cisplatin	Alkylating agent (Cytotoxic)
Cyclophosphamide	Alkylating agent (Cytotoxic)
Melphalan	Alkylating agent (Cytotoxic)
Melphalan flufenamide	Alkylating agent (Cytotoxic)
Doxorubicin	Anthracycline (Cytotoxic)
Idarubicin (Zavedos®/Idamycin®, Idaru®/ Ondarubin® / Zavedose®)	Anthracycline (Cytotoxic)
Liposomal doxorubicin (Caelyx®/Myocet®)	Anthracycline (Cytotoxic)
Vincristine (Leurocristine®)	Vinca alkaloid (Cytotoxic)
Belantamab mafodotin (Blenrep®)	Anti-BCMA agent
Venetoclax (Venclexta®/Venclyxto®)	BCL2 inhibitor
Daratumumab (Darzalex®)	CD38-directed mAb
Isatuximab (Sarclisa®)	CD38-directed mAb
Elotuzumab (Empliciti®)	mAb
Panobinostat (Farydak®)	HDAC
Lenalidomide (RevlIMiD®)	IMiD
Pomalidomide (Pomalyst®/Imnovid®)	IMiD
Thalidomide (Contergan®/Distaval®/Kevadon®/ Neurosedyn®/Distaval®/ Pantosediv®/ Sedoval K-17®/ Sedoval K17®/ Softenon®/ Talimol®)	IMiD
Selinexor (Xpovio®/Nexpovio®)	Nuclear export inhibitor
Bortezomib (Velcade®)	PI
Carfilzomib (Kyprolis®)	PI
Ixazomib (Ninlaro®)	PI
Etoposide (Vepesid®)	Podophyllotoxin Derivative (Cytotoxic)
Vincristine/Leurocristine (Oncovin®)	Vinca Alkaloid (Cytotoxic)
Tisagenlecleucel (Kymriah®)	CAR-T cell therapy
Brexucabtagene autoleucel (Tecartus®)	CAR-T cell therapy
Axicabtagene ciloleucel (Yescarta®)	CAR-T cell therapy
Cilta-cel (Carvykti®)	CAR-T cell therapy
Lisocabtagene maraleucel (Breyanzi®)	CAR-T cell therapy
Idecabtagene vicleucel (Abecma®)	CAR-T cell therapy

Abbreviations: BCMA = B-cell maturation antigen; BCL-2 = B-cell lymphoma 2; CAR-T cell = chimeric antigen receptor T-cell; HDAC = histone deacetylase inhibitor; IMiD = immunomodulatory imide drugs; mAb = monoclonal antibody; MM = multiple myeloma; PI = proteasome inhibitor.

15.4 Appendix 4. Corticosteroids and Equivalent Doses

Corticosteroid	Equivalent Dose (≥70mg prednisone)
Betamethasone	≥10.5mg
Cortisone	≥350mg
Dexamethasone	≥10.5mg
Hydrocortisone	≥280mg
Methylprednisolone	≥56mg
Prednisolone	≥70mg
Prednisone	≥70mg
Triamcinolone	≥56mg

Abbreviation: mg = milligramme.

15.5 Appendix 5. Charlson Comorbidity Score and its Components

Variable	Scale	Points	ICD-10 code
Diabetes mellitus	Uncomplicated	1	E10.0, E10.1, E10.6, E10.8, E10.9, E11.0, E11.1, E11.6, E11.8, E11.9, E12.0, E12.1, E12.6, E12.8, E12.9, E13.0, E13.1, E13.6, E13.8, E13.9, E14.0, E14.1, E14.6, E14.8, E14.9
	With complications	2	E10.2–E10.5, E10.7, E11.2–E11.5, E11.7, E12.2–E12.5, E12.7, E13.2–E13.5, E13.7, E14.2–E14.5, E14.7
Liver disease	Mild	1	B18.x, K70.0–K70.3, K70.9, K71.3–K71.5, K71.7, K73.x, K74.x, K76.0, K76.2–K76.4, K76.8, K76.9, Z94.4
	Moderate to severe	3	I85.0, I85.9, I86.4, I98.2, K70.4, K71.1, K72.1, K72.9, K76.5, K76.6, K76.7
Cancer	Any malignancy	2	C00.x–C26.x, C30.x–C34.x, C37.x–C41.x, C43.x, C45.x–C58.x, C60.x–C76.x, C81.x–C85.x, C88.x, C90.x–C97.x
	Metastatic	6	C77.x–C80.x
AIDS /HIV	Yes	6	B20.x–B22.x, B24.x
Renal Disease	Yes	2	I12.0, I13.1, N03.2–N03.7, N05.2–N05.7, N18.x, N19.x, N25.0, Z49.0–Z49.2, Z94.0, Z99.2
Congestive Heart Failure	Yes	1	I09.9, I11.0, I13.0, I13.2, I25.5, I42.0, I42.5–I42.9, I43.x, I50.x, P29.0
Myocardial infarction	Yes	1	I21.x, I22.x, I25.2
Chronic pulmonary disease	Yes	1	I27.8, I27.9, J40.x–J47.x, J60.x–J67.x, J68.4, J70.1, J70.3
Peripheral vascular disease	Yes	1	I70.x, I71.x, I73.1, I73.8, I73.9, I77.1, I79.0, I79.2, K55.1, K55.8, K55.9, Z95.8, Z95.9
Transient Ischemic attack OR cerebrovascular accident	Yes	1	G45.x, G46.x, H34.0, I60.x–I69.x
Dementia	Yes	1	F00.x–F03.x, F05.1, G30.x, G31.1
Hemiplegia or paraplegia	Yes	2	G04.1, G11.4, G80.1, G80.2, G81.x, G82.x, G83.0–G83.4, G83.9

Variable	Scale	Points	ICD-10 code
Rheumatic disease	Yes	1	M05.x, M06.x, M31.5, M32.x–M34.x, M35.1, M35.3, M36.0
Peptic ulcer disease	Yes	1	K25.x–K28.x

Abbreviations: AIDS = acquired immune deficiency syndrome; HIV = human immunodeficiency virus; ICD-10 = International Classification of Diseases, 10th Revision.