MPN PROGRESSion Registry™ (PROGRESS-MPN): A U.S.-Based Observational Study Tracking Symptoms, Treatments, Clinical Outcomes, and Disease Progression in Myeloproliferative Neoplasms

Protocol Reference Number:
MPNRF-PROG-2025-001

IRB Protocol Number:
20251868

Version Number:
Version 1.1

Version Date:
Nov12, 2025

Sponsor:
MPN Research Foundation
515 N State Street, 14th Floor
Chicago, IL 60654

Principal Investigator:
Raajit Rampal, MD, PhD
Memorial Sloan Kettering Cancer Center, New York, NY

Sub-Investigators:
Douglas Tremblay, MD — Icahn School of Medicine at Mount Sinai, New York, NY
Jan Bewersdorf, MD — Yale School of Medicine, New Haven, CT

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

BMT	bone marrow transplant
CMML	chronic myelomonocytic leukemia
EMR	electronic medical record
ET	essential thrombocythemia
FHIR	Fast Healthcare Interoperability Resources
ICF	informed consent form
IRB	institutional review board
IWG-MRT	International Working Group for Myelofibrosis Research and Treatment
LAR	legally authorized representative
MDS	myelodysplastic syndrome
MDS/MPN-*SF3B1*-T	MDS/MPN with *SF3B1* and thrombocytosis
MF	Myelofibrosis
MPN	myeloproliferative neoplasm
MPN-AP	myeloproliferative neoplasm in accelerated phase
MPN-BP	myeloproliferative neoplasm in blast phase
MPN-U	myeloproliferative neoplasm, unclassifiable
PI	principal investigator
PRO	patient-reported outcomes
pre-PMF	pre-fibrotic primary myelofibrosis
PV	polycythemia vera
QOL	quality of life
RWE	real world evidence
SCT	stem cell transplant
US	United States
WHO	World Health Organization

1. BACKGROUND

Patients diagnosed with myeloproliferative neoplasms (MPNs)—including primary myelofibrosis (MF), essential thrombocythemia (ET), and polycythemia vera (PV)—face progressive bone marrow diseases largely driven by JAK2, CALR, or MPL mutations.[1-4] These conditions lead to overproduction of mature blood elements, marrow fibrosis, and heightened risks of thrombosis, cardiovascular events, and inflammatory complications.[5-7] Although some patients live decades post-diagnosis, quality of life is often significantly reduced, and transformation to acute leukemia (MPN-BP) carries poor outcomes.[8-9]

Despite affecting an estimated 300,000 individuals in the U.S.,[10] much remains unknown about the biological drivers of disease progression, predictive biomarkers, and meaningful surrogate endpoints. Existing MPN registries have provided valuable but limited insights, often constrained to single-center, clinician-reported, or trial-based cohorts.[11] This Registry addresses these gaps by integrating electronic medical records (EMRs), patient-reported outcomes (PROs), claims data, and genomic profiles across a broad patient spectrum—including pre-fibrotic MF (pre-PMF), MPN-AP, MPN-BP or post-MPN AML, and MDS/MPN overlap syndrome (defined as CMML, MDS/MPN-*SF3B1*-T, MDS/MPN with neutrophilia, or MDS/MPN not otherwise specified), including patients originally diagnosed with one of these conditions but who have received one or more SCTs and/or BMTs —supporting comprehensive longitudinal tracking of disease evolution, treatment response, and EMR.

1.1 Study Rationale

Historically, efforts to understand MPN disease progression have been siloed within individual research centers, leaving critical data gaps despite decades of published studies and a few approved therapies. This Registry aims to establish a robust, longitudinal observational research platform to address these unmet needs, benefiting the broader MPN community.

Using a digital health platform, the Registry will collect multimodal data—centered on patient (EMRs) and enriched by self-reported outcomes. The cohort will include patients diagnosed with primary or secondary MF, ET, PV, pre-PMF, MPN-unclassifiable (MPN-U), MPN-AP, and MPN-BP or post-MPN AML, MDS/MPN overlap syndrome as defined above. This will also include patients originally diagnosed with one of these conditions but who have received one or more SCTs and/or BMTs.

By integrating EMR data, PROs, claims data, and linked registries, this study enables comprehensive longitudinal analysis of disease trajectories, treatment patterns, and quality-of-life outcomes. The Registry's real-world evidence (RWE) will help identify early predictors of progression, validate surrogate endpoints, and support future clinical trials and drug development.

Ultimately, this infrastructure will enhance regulatory decision-making, inform treatment guidelines, and improve prognostic modeling—advancing precision care and outcomes for patients with MPNs.

2. PURPOSE

The purposes of the MPN PROGRESSion Registry™ ("Registry") are to:

- **Facilitate high-quality observational research** by collecting and analyzing EMRs, PROs, and linked external datasets — including mortality records, insurance claims, genetic testing results, and affiliated registries — to advance understanding of disease natural history, progression, real-world therapeutic effectiveness, and standards of care.
- **Improve clinical definitions of MPN disease progression** by incorporating updated diagnostic criteria and emerging biomarkers, ensuring refined characterization of pre-fibrotic myelofibrosis (pre-PMF), MPN in accelerated phase (MPN-AP), MPN in blast phase (MPN-BP) or post-MPN AML, and MDS/MPN overlap syndrome as defined above
- **Identify predictive biomarkers and surrogate endpoints** to support drug development, regulatory decision-making, and improved understanding of disease progression.
- **Enhance standards of care and patient outcomes** by leveraging RWE and longitudinal analytics to inform treatment sequencing, monitor therapeutic responses, and guide personalized care strategies.
- **Support participant identification for future research** by enabling review of Registry data to identify and contact individuals eligible for future studies conducted by the MPN Research Foundation or approved collaborators.

3. OVERALL STUDY DESIGN

This is a prospective, non-interventional observational study, designed to evaluate patients with MPNs using both historical and prospective data. Participants will be enrolled over a minimum of three years during the initial run-in phase and followed longitudinally for disease progression, treatment response, clinical outcomes, and survival until withdrawal from the registry or death. Some participants may be followed for over ten years.

The study uses a multimodal data collection strategy, integrating:

- **EMRs:** Structured data extractions, including laboratory results, imaging, clinical notes, and treatment history.
- **PROs:** Longitudinal assessments of symptoms, quality of life, and functional status.
- **Medical Claims and Insurance Data:** Real-world healthcare utilization, treatment adherence, and cost metrics.
- **Mortality Data:** Vital status and cause-specific mortality obtained through national mortality dataset linkages, ensuring accurate survival analyses and minimizing loss to follow-up.
- **Affiliated Registries:** Select collaborations (e.g., CIBMTR, NCDB, SEER) may be used for external validation and cross-registry comparisons, though integration with international registries such as the European LeukemiaNet (ELN) will be considered only after addressing jurisdictional and data governance factors.

In addition to evaluating disease and treatment outcomes, the Registry will systematically monitor enrollment feasibility, participant retention, and heterogeneity across demographic, clinical, and geographic dimensions. These metrics will support reporting obligations to federal registries (e.g., ClinicalTrials.gov) and inform future feasibility assessments for interventional trials.

4. OBJECTIVES

4.1. Primary Objectives

Stage 1: Feasibility Phase
The study will begin with a 12- to 24-month feasibility phase focused on participant retention, data completeness, EMR integration, and engagement strategies. Results from this phase will help optimize study operations, guide participant engagement, and inform potential protocol amendments.

- Evaluate enrollment feasibility and participant characteristics during the first 12–24 months.
- Characterize participant demographics, geographic distribution, and clinical settings.
- Assess inclusion of underrepresented populations to enhance the generalizability and heterogeneity of RWD.
- Assess representativeness of enrolled participants relative to the broader U.S. MPN population, ensuring alignment with clinical and demographic patterns.
- Determine the proportion of participants providing a Minimal Viable Dataset (MVD) at enrollment.
- Monitor consent rates for longitudinal EMR access and secondary data use.
- Identify gaps in baseline data completeness and develop strategies to mitigate missing information.
- Track participant retention at 12, 18, 24, and 36 months, and evaluate completion of at least three consecutive follow-ups.
- Identify early indicators of dropout or disengagement to inform retention strategies.
- Monitor response rates to quarterly PRO surveys (e.g., MPN-SAF TSS, PROMIS Short Form v1.0 – Fatigue 8a) and evaluate follow-up strategies (automated vs. manual).
- Assess the proportion of participants using secondary contacts (caregivers, clinicians) for follow-up.
- Measure success rates of EMR integration across sites using FHIR-based interoperability, compare self-reported and EMR-extracted data, and validate processing scalability.

Stage 2: Surrogate Endpoint Phase
- Establish composite surrogate endpoints integrating genetic markers (e.g., variant allele frequency shifts), clinical indicators (e.g., fibrosis progression, cytopenias), inflammatory markers (e.g., cytokine levels), and PROs (e.g., MPN-SAF TSS, PROMIS trends).
- Ensure surrogate endpoints align with FDA and EMA regulatory standards for clinical trials.
- Incorporate RWD (EMR, claims, PROs) to validate surrogate markers of disease progression.
- Determine time from baseline to first progression event (e.g., ET → PV → MF, blast phase transformation) and assess median time to progression by subtype.
- Measure progression rates per WHO 2022/IWG-MRT criteria using updated risk stratification models.
- Develop surrogate endpoints predictive of disease progression and treatment response to support drug development.
- Compare progression rates in treated vs. untreated cohorts and validate endpoints using external registries (e.g., SEER, NCDB, Mayo Clinic MPN Database, ELN).
- Monitor overall survival (OS) and progression-free survival (PFS), evaluating prognostic significance of surrogate endpoints for future trials.

4.2. Secondary Objectives

- **Characterize Disease Burden and Risk Stratification**
  Develop a composite risk model distinguishing high- and low-risk patients by integrating genetic, inflammatory, and clinical markers. Assess how cytokine levels and inflammatory pathways influence disease trajectory and symptom burden.
- **Evaluate Treatment Response as a Surrogate for Disease Progression**
  Compare progression rates in patients treated with JAK inhibitors, interferons, investigational agents, and other drugs. Determine whether therapy-induced molecular or clinical changes correlate with improved outcomes.
- **Analyze Genetic and Morphological Disease Evolution**
  Monitor available variant allele frequency (VAF) changes in high-risk mutations and evaluate bone marrow fibrosis progression, where data is available, to understand clinical implications.
- **Assess Quality of Life and Symptom Burden at Disease Progression**
  Track longitudinal changes using validated PROs, including MPN-SAF TSS, EORTC QLQ-C30, modified PHQ-9, PGI-S, and PROMIS Short Form v1.0 – Fatigue 8a, to evaluate how symptom burden correlates with disease progression

4.3. Exploratory Objectives

- **Develop Predictive Models for MPN Progression**
  Train and validate artificial intelligence (AI) and machine learning (ML) models to predict time to disease progression, likelihood of transformation to accelerated or blast phase, and overall survival using early clinical and molecular indicators.
- **Validate Findings Across External Data Sources**
  Compare registry data against European MPN registries, Mayo Clinic datasets, and claims-based registries to assess consistency, generalizability, and cross-cohort validation.
- **Assess Healthcare Utilization and Economic Impact**
  Track hospitalization rates, transfusion needs, major interventions, and conduct cost-effectiveness analyses to inform early intervention strategies and support health policy development
- **Assess Early Patterns of Care and Treatment Setting Influence**
  Evaluate how early patterns in healthcare utilization—including differences between academic centers, community hospitals, and rural clinics—impact disease monitoring, treatment choices, referral patterns, and long-term outcomes. Identify disparities in access to specialized care and examine how practice setting influences adherence to evidence-based guidelines and clinical outcomes.

5. PARTICIPATION CRITERIA

5.1 Inclusion Criteria

- Confirmed diagnosis of pre-PMF, primary MF, secondary MF, PV, ET, MPN-U, MPN-AP, MPN-BP or post-MPN AML, and MDS/MPN overlap syndrome as defined above per WHO 2022 criteria, including patients originally diagnosed with one of these conditions but who have received one or more SCTs and/or BMTs
- Willingness and ability to provide written informed consent.

- Willingness and ability to complete patient assessments, independently or with minimal assistance.
- Accessible EMR data from a participating healthcare institution to enable longitudinal data collection.
- U.S. residency, where the Registry has regulatory approvals and data-sharing agreements.
- Reasonable expectation of ongoing medical care and follow-up data availability for longitudinal objectives.

5.2 Exclusion criteria

Participants will be excluded if they meet any of the following:

1. Lack of EMR data access due to technical limitations, non-participating institutions, or missing patient authorization (self-reported or claims-only data is insufficient for inclusion).
2. Missing compulsory data elements: WHO 2022 MPN diagnosis, date of initial diagnosis, or baseline demographics.
3. Lack of capacity to provide informed consent and absence of a legally authorized representative.
4. Significant gaps in clinical history, lab data, treatment records, or follow-up availability that limit longitudinal analysis.
5. Revised diagnosis to a non-MPN hematologic malignancy or other condition outside WHO 2022 criteria (other than post-MPN AML, and MDS/MPN overlap syndrome as defined above); such participants will be removed from future analyses.
6. Enrollment in external studies that restrict medical data access or impose data-sharing limitations incompatible with this Registry.

6. PARTICIPANT RECRUITMENT AND ENROLLMENT

The MPN Research Foundation aims to enroll a representative cohort that reflects the heterogeneity of MPN biology across varied ancestral, socioeconomic, geographic, and clinical backgrounds. Recruitment strategies will include:

- Establishing partnerships with community health centers, primary care providers, and hematology/oncology clinics to broaden outreach.
- Utilizing targeted online advertising through search engines and social media platforms, patient advocacy networks, and Search Engine Optimization (SEO) to enhance visibility.
- Providing standardized educational toolkits to healthcare providers to promote referrals.
- Deploying patient navigators to engage potential participants via phone, email, or online platforms to address barriers.
- Offering multilingual, plain-language materials to improve accessibility and health literacy.
- Monitoring demographic heterogeneity in real time and adjusting strategies to improve representation of underrepresented populations.

6.1. Duration of Observation and Participation

Participants will be followed for a minimum of five years, with the potential for extended participation beyond ten years to support long-term observational analyses.

Standardized data collection will occur at baseline, six months, and every six months thereafter. Collected data will include:

- EMR extractions (clinical history, labs, imaging, treatments)
- Insurance claims and pharmacy records (healthcare utilization, medication use)
- PROs, including MPN-SAF TSS, PROMIS Short Form v1.0 – Fatigue 8a, and quality-of-life assessments

Data integration will prioritize FHIR-based EMR retrieval, supplemented by validated participant surveys and secondary data sources.

Ongoing monitoring and planned assessments include:

- **Year 1:** Baseline validation, data completeness review, and participant engagement check
- **Year 3:** Disease trajectory modeling, retention analysis, and interim outcome assessment
- **Year 5 and beyond:** Predictive analytics on disease progression, long-term outcomes, and composite endpoint evaluation

To promote retention, the Registry will deploy automated reminders, provide quarterly study updates, and offer virtual engagement opportunities (e.g., newsletters, webinars). Participants may withdraw at any time, as outlined in Section 6.2.

6.2. Informed Consent

All participants (or their legally authorized representatives) will provide written informed consent using an IRB-approved form detailing the study's purpose, duration, procedures, risks, benefits, and privacy protections. Informed consent allows the study team to access participant medical records, collect survey and PRO data, and link external datasets (e.g., claims, mortality, genetic testing) for approved research purposes. Participants will have the opportunity to ask questions prior to signing and will receive a copy of the signed consent for their records. Consent can be withdrawn at any time without penalty. Consent materials will be available electronically, with printable copies provided via the MPN Research Foundation website. A secure, encrypted system will capture and store electronic signatures.

6.3. Visit Schedule

The Registry is designed for extended longitudinal follow-up, with no required in-person study visits due to its non-interventional, observational nature. Data will be collected continuously through EMRs, PROs, insurance claims, mortality data, and linkages to affiliated registries.

The visit schedule aligns with structured follow-up intervals to ensure consistent longitudinal tracking of disease progression, treatment response, and post-enrollment clinical markers. Retention and engagement strategies will be employed to maintain data completeness and minimize loss to follow-up.

6.3.1. Data Collection Intervals and Structured Follow-Up Schedule

Participants will be followed for a minimum of five years, with many continuing beyond ten years.

Follow-Up	Data Collected	Collection Method

Timepoint		
Baseline (Enrollment)	Demographics, medical history, WHO 2022 MPN subtype, other hematologic disease information (e.g., post-MPN AML, MDS/MPN overlap syndrome as defined above, history of SCT and/or BMT) baseline labs, molecular testing (if available), treatment history, and PROs (MPN-SAF TSS, EORTC QLQ-C30, modified PHQ-9, PGI-S, PROMIS Short Form v1.0 – Fatigue 8a)	Initial EMR extraction, PROs, insurance claims
Every 3 Months (0-10+ yrs)	Symptom burden updates (MPN-SAF TSS + PGI-S)	PROs
Every 6 Months (0-10+ yrs)	Updates on disease status, treatment changes, adverse events, healthcare utilization, molecular/genetics results (if available), mortality status, progression milestones, and PROs (MPN-SAF TSS, EORTC QLQ-C30, modified PHQ-9, PROMIS Short Form v1.0 – Fatigue 8a)	EMR reauthorization, PROs, claims retrieval, mortality data linkage

6.3.2. Disease Progression and Post-Enrollment Markers
The Registry will monitor disease evolution, therapeutic response, and emerging clinical risks through ongoing data collection from EMRs, laboratory results, imaging reports, mortality linkages, healthcare utilization data, and PROs.

Key post-enrollment markers include:

- **Molecular and Genetic Changes:** Emergence of high-risk mutations and shifts in VAF over time.
- **Bone Marrow and Imaging Biomarkers:** Assessments of fibrosis and splenomegaly progression, based on clinically indicated bone marrow biopsies or imaging.
- **Symptom Burden and Quality of Life:** Trends captured through validated PROs, including MPN-SAF TSS, EORTC QLQ-C30, modified PHQ-9, PGI-S, and PROMIS Fatigue Short Form.
- **Cardiovascular and Thrombotic Events:** Documentation of new complications such as deep vein thrombosis, pulmonary embolism, stroke, or myocardial infarction.
- **Mortality Status:** Vital status and cause of death, obtained through national mortality data linkages.

To maintain complete and up-to-date records, participants will be asked every six months to refresh their EMR data access permissions.

6.3.3. Patient Retention and Engagement Strategies
To support long-term participant retention, the Registry will implement a multi-faceted engagement approach that includes:

- **Automated Communications:**
  Email and SMS reminders will notify participants about upcoming surveys, EMR data refreshes, and study milestones.
- **Aggregate Study Updates:**
  Participants will periodically receive summaries of registry progress, including enrollment trends, geographic representation, heterogeneity metrics, and de-identified PRO findings (MPN-SAF TSS, EORTC QLQ-C30, modified PHQ-9, PGI-S, PROMIS Fatigue Short Form), delivered via secure channels.
- **Educational Materials:**
  Plain-language resources will explain how participant contributions advance MPN research and improve understanding of disease progression.
- **Interactive Opportunities:**
  Invitations to webinars, virtual town halls, or Q&A sessions will allow participants to hear updates directly from investigators.
- **Non-Monetary Incentives:**
  Digital badges, certificates, or eligibility for future research opportunities will help sustain engagement.
- **Patient Support Access:**
  A dedicated patient navigator will assist with survey completion, consent renewals, and EMR data access questions.

As PRO surveys and EMR refreshes occur approximately every six months, communications will reflect cumulative study updates. All outreach will use de-identified data, and participants will have secure access to their own summaries through the registry dashboard.

6.3.4. Data Quality Assurance and Compliance
To ensure the highest quality of data capture across all sources, the Registry will implement the following quality assurance measures:

- **Automated Data Validation Checks**
  Incoming data will be systematically screened for missing values, inconsistencies, and outliers at the point of entry using predefined validation rules aligned with international standards (ICD-10, SNOMED CT, LOINC).
- **Cross-Source Data Verification**
  Clinical data extracted from EMRs will be compared against PROs to confirm consistency, with discrepancies flagged for review.
- **Audit Trails and Data Security**
  All data modifications, including additions, corrections, or deletions, will be tracked through secure, HIPAA-compliant audit trails to ensure full traceability.
- **Regulatory Compliance**
  All procedures will adhere to U.S. FDA and EMA guidelines for the use of RWE in regulatory decision-making, ensuring compliance with relevant national and international standards.
- **Periodic Reviews and Oversight**
  The PI and designated data management personnel will perform regular reviews to assess adherence to protocol-defined quality benchmarks. The Registry's scientific steering

committee will also receive summary reports detailing data completeness, validation outcomes, and unresolved issues.

These layered measures will ensure the Registry produces accurate, reliable, and high-quality datasets suitable for robust observational research and regulatory submission.

6.4. Costs to the Participant and Payment for Participation

Participation in the Registry is voluntary, and participants will not incur direct costs for enrollment, data submission, or study-related activities. To acknowledge the time and effort required, the Registry will offer non-monetary engagement incentives, including entry into periodic randomized incentive distributions, designed to support retention without exerting undue influence or coercion.

6.4.1. No Direct Costs to Participants

Participants will not be charged for completing electronic surveys, EMR data retrieval, or other study-related activities. All Registry tasks, including PRO completion, EMR refreshes, and communications, are provided at no financial burden.

6.4.2. Participant Payment and Engagement Incentives

- **Participation-Based Incentives:**
  Participants who complete biannual PRO surveys and remain actively enrolled will be automatically entered into periodic incentive drawings. Specifically, gift card drawings will be conducted periodically for survey completion and active retention. Winners will be selected at random, and no additional steps are required beyond completing the associated activities.
- **Non-Monetary Engagement Opportunities:**
  Participants will have access to educational materials, aggregate study progress updates, invitations to webinars or virtual events, and optional certificates or badges recognizing their contributions to MPN research.

6.4.3. Potential Indirect Costs

Participants may require internet, smartphone, or computer access to complete study activities. While the Registry will minimize any institutional or medical record retrieval fees, these costs are not typically covered.

6.4.4. Payments from External Researchers for Data Access

The MPN Research Foundation may receive payments from approved external investigators who access de-identified data for secondary research. These funds will support Registry operations, improve data quality, and advance MPN research; they will not affect individual participant incentives.

6.4.5. Ethical Considerations and Regulatory Compliance

All compensation and incentive approaches comply with IRB, FDA, and EMA regulations. Participants will receive IRB-approved disclosures detailing drawing eligibility, prize distribution, data use, and external funding. Compensation is structured to ensure fair acknowledgment of participant contributions without coercion or undue influence.

7. METHODS AND PROCEDURES

Participation in the Registry consists of collection of the participants' information in a research database and the use of this information for observational research studies and for identification and recruitment of potential eligible participants for future research studies.

7.1. Data sources

Data collection for the registry will be done in non-interventional settings. This is not a clinical trial.

1. Registry participants may use a mobile and/or web-based application to add their participant-reported data ("Participant-Reported Data").

2. Registry participants will be asked to provide their written informed consent for their health care providers to disclose relevant EMRs to MPN Research Foundation for inclusion in the Registry and for MPN Research Foundation to access EMRs directly using a medical records technology interface ("EMR Data" and together with PROs and Survey Data, "Registry Data").

3. Registry participants will also provide information in response to surveys periodically administered by MPN Research Foundation ("Survey Data").

4. Additional data may be generated from:
    a. Medical claims and insurance data.
    b. Mortality Data: Vital status and cause-specific mortality outcomes will be captured through linkage to national mortality datasets. This linkage will strengthen the accuracy of survival analyses, enhance longitudinal outcome assessments, and minimize loss to follow-up.
    c. Data from affiliated disease registries.

5. The registry will collect a wide range of data variables, ensuring a holistic understanding of disease progression and treatment effects. These variables include:
    a. **Demographics:** Age, sex, race, ancestral background, socioeconomic indicators, geographic location, history of military service, and insurance status.
    b. **Clinical Data:** WHO-defined MPN subtype, disease stage, duration of disease, history of thrombotic events, splenomegaly, and systemic symptoms.
    c. **Treatment Data:** Medication history, treatment response, adverse events, dose adjustments, and reason for therapy discontinuation.
    d. **Genetic and Molecular Data:** JAK2, CALR, and MPL mutation status, variant allele frequencies, and other emerging genomic biomarkers.
    e. **PROs:** Symptom burden assessment using MPN-SAF TSS, fatigue severity, psychological well-being, and functional status.
    f. **Healthcare Utilization:** Frequency of hospitalizations, emergency department visits, specialist consultations, and treatment adherence.
    g. **Environmental and Lifestyle Factors:** Tobacco use, alcohol consumption, dietary habits, physical activity, and occupational exposures.

6. PROs will be captured through structured survey instruments designed to measure symptom burden, quality of life, and treatment satisfaction. These surveys will include:
    a. **MPN-SAF Total Symptom Score (TSS):** Assesses key MPN symptoms, including fatigue, pruritus, night sweats, bone pain, and early satiety.
    b. **PGI-S:** Captures patient-perceived overall change in health status.
    c. **EORTC QLQ-C30:** Evaluates overall health-related quality of life, including functional status and symptom burden.
    d. **Modified PHQ-9:** Screens for depressive symptoms commonly observed in MPN patients.
    e. **PROMIS Short Form v1.0 – Fatigue 8a:** Assesses fatigue in MPN patients, specifically targeting the experience of fatigue and its impact on daily life.
    f. Data from patient surveys will be integrated with clinical and genomic data to create a robust dataset that allows for multidimensional analyses of disease progression, therapeutic impact, and patient experience.

In the future, participants may be invited to provide optional blood or tissue samples to support additional research on genetic, molecular, or biomarker analyses. Any biospecimen collection will require a separate IRB-approved protocol and additional informed consent. Participation in biospecimen sub-studies will be entirely voluntary and will not affect participation in the main registry.

7.2. Data Storage and Sharing

Data will be stored electronically within the MPN Research Foundation Service Database. Identifiable details, including names, email addresses, dates other than year, photographic images, and free-text notes, will be removed and replaced with a linkage code.

A designated Honest Broker for this study has been established with a third-party through a specific contractual agreement. Serving independently from the study analysis team and outside MPN Research Foundation, the Honest Broker will assign linkage codes, maintain the secure linkage file, and ensure the separation of identifiable data from research datasets. The linkage file will be stored in an encrypted system accessible only to the Honest Broker, with all access documented through audit trails. The Honest Broker will also serve as a resource to the MPN PROGRESSion Registry governance committees, including the Data Access, Use & Publications (DAUP) Committee.

Coded research data will be retained indefinitely by the MPN Research Foundation to support ongoing and future observational research, in accordance with organizational data retention policies and applicable regulatory requirements.

Investigator access to Registry data requires approval by the DAUP Committee, based on scientific quality, validity, and ethical compliance, with all approvals documented in writing. De-identified data may be shared with external researchers for approved secondary studies, subject to appropriate IRB approvals. Investigator access to Registry information for recruitment or eligibility screening will only occur with documented IRB approval.

While participants may receive insights from aggregate analyses, they will not necessarily be informed of all results from every observational study using their data.

7.3. Data Monitoring Plan

The monitoring plan for the Registry will involve routine (i.e., quarterly) monitoring by the PI and/or designees to ensure 1) direct participant identifiers are not contained within the Registry; 2) MPN Research Foundation investigator exports from the Registry are documented; 3) information linking the Registry linkage codes with participant identifiers is stored securely and access to the linkage code key is audited; and 4) any conditions that may negatively impact the confidentiality of information contained within the Registry are reported to the appropriate MPN Research Foundation personnel.

As specified previously, the DAUP Committee must provide approval for an MPN Research Foundation investigator to access the Registry for research studies involving the use of study Data. Access by MPN Research Foundation investigators to the Data for the purpose of identifying potential participants for participation in a research study shall be granted only upon the provision of documentation that an IRB has approved the research study. At the time of annual IRB continuing review, a list of studies conducted using the Registry will be submitted to the IRB. In addition, any unauthorized access to information contained within the Registry or to the information linking the Registry information linkage codes to participant direct identifiers shall be reported to the IRB.

7.4. Research Methods

This study will implement a structured methodology encompassing data acquisition, curation, validation, and governance. Data will be sourced from:

- **EMRs:** Retrospective and prospective clinical data, including laboratory results, imaging studies, physician notes, and treatment history.
- **PROs:** Longitudinal symptom assessments, quality of life metrics, and functional status evaluations.
- **Genomic and Molecular Data:** Mutation profiles, variant allele frequencies, and emerging biomarkers relevant to disease evolution.
- **Administrative Claims Data:** Healthcare utilization patterns, insurance claims, and medication adherence metrics.

To ensure data consistency and quality, all collected data will be standardized using internationally recognized coding systems, such as ICD-10, SNOMED CT, LOINC, and the United States Core Data for Interoperability (USCDI). Rigorous data validation processes will be implemented, including:

- Automated and manual data quality checks.
- Cross-validation between EMRs and PROs.
- Real-time monitoring of data completeness and accuracy.
- Periodic external audits to ensure compliance with best practices.

A governance framework will be established to oversee ethical compliance, data security, and appropriate use of registry data for research and regulatory purposes. Adaptive protocols will be

incorporated, allowing for modifications based on emerging scientific discoveries, technological advancements, and stakeholder feedback.

The study team anticipates that the protocol may be amended over time based on findings from the feasibility phase, input from one or more of the governance committees, stakeholder feedback, or evolving scientific and operational priorities. Potential amendments may include the addition of sub-studies, optional biospecimen collection, adjustments to PRO schedules, or updates to data collection methods. All amendments will be submitted for IRB review and approval prior to implementation

7.5. Withdrawal/Discontinuation

Participants may withdraw consent at any time by submitting a written request (such as a signed form, email, or letter) to the study team. A record of the withdrawal will be maintained for regulatory compliance. Data collected prior to withdrawal will remain in the Registry for approved research; no additional data will be collected unless the participant re-consents.

7.6 Patient-Reported Outcomes Collection

7.6.1 Overview

The Registry will incorporate PROs to assess disease burden, symptom severity, quality of life, mental health, fatigue, and healthcare access. PROs will complement clinical and molecular data collected through EMRs to provide a comprehensive analysis of disease progression and treatment impact from the patient perspective.

PRO collection will provide valuable RWE to:

- Assess the impact of disease and therapy on quality of life.
- Support clinical decision-making and future research.
  Identify early indicators of disease progression and treatment response.

The estimated time to complete the full PRO battery is 8–10 minutes per assessment. Efforts will be made to streamline survey administration and minimize participant burden while ensuring comprehensive data collection.

7.6.2 Selected Patient-Reported Outcomes Measures

Validated instruments have been selected based on their established use in major MPN clinical trials, including COMFORT-I & II, JAKARTA, PERSIST-1 & 2, RESPONSE, and SIMPLIFY trials. These measures align with recommendations for RWE collection in chronic hematologic diseases.

Category	Measure	Purpose	Estimated Completion Time
Symptom Burden	MPN-SAF Total Symptom Score (MPN-SAF TSS)	Assesses key MPN symptoms, including fatigue, pruritus, night sweats, bone pain, and early satiety	5–7 minutes

Category	Measure	Purpose	Estimated Completion Time
Global Change	PGI-S (Patient Global Impression of Severity)	Captures patient-perceived overall change in health status	<1 minute
Quality of Life	EORTC QLQ-C30	Evaluates overall health-related quality of life, including functional status and symptom burden	7–10 minutes
Mental Health	Modified PHQ-9	Screens for depressive symptoms commonly observed in MPN patients. Modified version excludes Question 9, which asks about suicidal ideation.	2–3 minutes
Fatigue	PROMIS Short Form v1.0 – Fatigue 8a	Assesses fatigue in MPN patients, specifically targeting the experience of fatigue and its impact on daily life	5-10 minutes

Special Procedures for modified PHQ-9 Monitoring
The Registry collects patient-reported depressive symptoms using the modified PHQ-9 instrument.. While the registry is a non-interventional, observational study and does not provide clinical care or psychiatric services, the following safeguards are in place to responsibly manage participant well-being:

- All participants will see a message alerting them to the potentially sensitive nature of the questions included in the modified PHQ-9 and reminding them that they may choose to skip any question and may choose to forego completion of the modified PHQ-9 at any time.
- All participants, regardless of modified PHQ-9 score, will immediately see a resource referral pop-up in the online portal upon survey completion. This pop-up provides national mental health and crisis support resources, including the 988 Suicide & Crisis Lifeline, Crisis Text Line, and other relevant support services.
- The informed consent form will clearly communicate these boundaries to participants, including that the registry team does not provide medical or mental health services and that participants are responsible for seeking professional care if they have concerns about their well-being.

7.6.3 Data Collection Schedule
Data collection will follow a structured schedule to balance scientific objectives with participant burden. PROs will be collected on a tiered schedule. MPN-SAF TSS and PGI-S will be collected every 3 months (quarterly) to track changes in symptom burden. EORTC QLQ-C30, modified PHQ-9, and PROMIS Short Form v1.0 – Fatigue 8a will be collected every 6 months (biannually) to assess quality of life, mood, overall health, and fatigue.

Participants will have up to 7 days to complete each survey after it is sent, supported by automated reminders. This approach is designed to maximize data completeness while minimizing burden.

Clinical and EMR data will be updated at the same biannual timepoints and as available through data linkages. Table 2 provides an overview of the assessment schedule, and the recommended completion windows are summarized below.

Timepoint	Recommended Completion Window
Baseline Assessments	7 days
3-Monthy and Quarterly Follow-Ups (MPN-SAF TSS + PGI-S)	7 days
6-Month and Biannual Follow-ups (All PROs + EMR)	7 days

Note: MPN-SAF TSS and PGI-S are assessed every 3 months (quarterly); EORTC QLQ-C30, modified PHQ-9, PROMIS Short Form v1.0 – Fatigue 8a, and EMR data are assessed every 6 months (biannually). All assessments have a 7-day completion window supported by automated reminders.

7.6.4 Data Integration and Analysis

Integration of PROs with clinical and molecular datasets follows established methodologies for multi-source data validation in observational hematology studies[7–9].

Handling of Partial or Incomplete Responses

- If participants complete only a partial assessment, data from completed sections will be retained.
- Participants with consistently incomplete responses will be flagged for additional follow-up.
- A threshold (e.g., ≥50% completion) may be used to determine whether a submission is valid for analysis.
- Participants who submit incomplete assessments will be encouraged to complete the missing sections within the established grace period.
- If the assessment remains incomplete after the grace period, responses will be retained as submitted, and trends will be monitored for persistent gaps in reporting.

Non-Digital Data Entry and Quality Control

- Non-digital submissions (phone-based or paper forms) will be transcribed into the registry's electronic database by designated study personnel.
- A quality control process will be implemented to minimize transcription errors.

Providing PROs to Participants and Clinicians

- Participants may request summary reports of their individual PRO trends, which will be available through the registry portal.
- Where applicable, these reports may also be integrated into EMRs for treating physicians to support shared decision-making.

PROs will be integrated with clinical and molecular data collected through EMRs to create a comprehensive real-world dataset. This will allow for the evaluation of:

- Longitudinal changes in symptom burden and quality of life.
- Associations between trends and disease progression markers.
- Treatment effectiveness and tolerability from the patient perspective.
- Barriers to accessing care and medication adherence.

In the future, an optional online dashboard may be developed to allow participants to view their own reported data alongside general, de-identified study summaries. This dashboard will be intended solely for informational purposes and will not provide medical advice, treatment recommendations, or individualized clinical interpretations. Participation in dashboard access will be entirely optional and will not affect participation in the main registry.

7.7 Study Sites and Coordinating Center

The Principal Investigator, Dr. Raajit Rampal, is based at Memorial Sloan Kettering Cancer Center (MSKCC), which serves as the IRB site of record. All study operations, including participant recruitment, informed consent management, data collection, data storage, and participant communications, are centrally coordinated and administered by the MPN Research Foundation, located in Chicago, Illinois. The MPN Research Foundation functions as the study coordinating center and administrative office. No other clinical sites or healthcare institutions are engaged in participant enrollment, study management, or data operations.

8. RISK/BENEFIT ASSESSMENT

8.1. Potential Risks

Participation in this Registry involves minimal risk. The primary risk is potential loss of confidentiality or unauthorized access to personal information. To minimize this risk, all personally identifiable information (PII) will be separated from clinical and survey data and securely managed by the Honest Broker. Each participant's data will be linked using a unique study identifier, stored separately from the research database.

Although all data will be de-identified and coded, there is a minimal risk of re-identification, particularly when linking with external registries. Only authorized investigators operating under approved data use agreements will have access to de-identified datasets.

Survey responses will be stored in a secure, HIPAA-compliant system, with access restricted to authorized study personnel. As a non-interventional, observational study, there are no foreseeable physical or psychological risks, and no investigational treatments or procedures are involved.

Because the study does not involve clinical interventions or investigational products, no compensation for injury is provided. Participants will not incur any direct costs. In the event of a data breach, institutional policies and applicable laws will be followed, including prompt notification to affected participants and IRB reporting.

8.2. Benefits

As this is a non-interventional, observational study, there are no direct benefits to participants in this registry. However, their contributions will advance understanding of MPNs, disease progression, treatment patterns, and patient-reported experiences. The data collected may inform

clinical decision-making, improve care standards, support therapeutic development, and contribute to public health and regulatory efforts.

As part of the engagement strategy, participants will be eligible for periodic drawings for $50 gift cards, held twice per year for completing PROs and once per year for maintaining active enrollment. Winners will be randomly selected, and no additional steps are required beyond completing the associated activities. These drawings are intended to acknowledge participant time and effort without exerting undue influence or coercion, as described in the ICF.

Participants may also benefit indirectly through access to educational materials, aggregate study updates, and invitations to optional webinars or virtual events related to MPN research.

8.3. Protection of Human Participants

Coded participant data will be retained for at least 10 years after study closure or as required by institutional policy and applicable regulations.

All data collection, storage, and sharing procedures will comply with IRB, HIPAA, FDA, EMA, and national research guidelines. Access to de-identified data for secondary research will require approval from the Registry Steering Committee, DAUP Committee and the IRB, and all requests will be reviewed for ethical use and scientific merit.

This protocol has been reviewed and approved by a central IRB. Ongoing oversight will include annual continuing review and prompt reporting of any unanticipated problems or breaches.

Participation in the Registry is voluntary, and participants may withdraw at any time without penalty. Data collected prior to withdrawal will be retained in coded form and may continue to be used for approved research unless otherwise requested by the participant.

To protect privacy, all electronic data will be stored on encrypted, access-restricted servers. Physical records will be secured in locked facilities. All study staff will be trained in Good Clinical Practice (GCP), human subjects protections, and institutional data privacy procedures.

9. STATISTICAL CONSIDERATIONS

Patient data will come from multi-modal data sources, including academic and non-academic health centers, patient-consented access to medical records, validated survey instruments, and patient self-reports. A durable and sizeable cohort over several years will generate longitudinal data that can be used in studies to allow researchers to analyze data potentially pertinent to disease progression. This cohort will capture a sufficient number of relatively rare events, such as disease progression to blast-phase MPN. At the development stage, the cohort composition and subsequent estimates of the risk of progression are uncertain. The registry aims to enroll approximately 1500 patients within 3 years. Under this sample size assumption, we estimate statistical power to detect an association between various risk factors and progression, as detailed in Table 1

Table 1: Minimum Hazard Ratio (HR) needed for the risk marker with power of 80%, type-1 error 5%, and the assumption of progression of 50% during the follow-up period

Sample Size\Prevalence of marker	10%	20%	30%	50%
1000	1.53	1.37	1.32	1.3
2000	1.35	1.25	1.22	1.2
3000	1.27	1.2	1.17	1.16
5000	1.2	1.15	1.13	1.12

We estimate that the initial run-in part of a study duration will be at least 3 years. After 3 years, a statistical team will review the data and define a statistical analysis plan (SAP) based on the composition and completeness of the data.

Based on experience from similar projects in the past, we estimate an annual drop-out rate of 10-20%. As patients will enroll in this study on an ongoing basis, patients lost to follow up will be replaced over time to maintain adequate statistical power as illustrated in Table 1.

9.1 Statistical Analysis Plan (SAP)
An initial SAP will be developed prior to data collection, outlining intended descriptive and inferential statistical methods, including Kaplan-Meier survival analysis[26], Cox proportional hazards modeling[27] for progression events, and multivariate analyses. Primary endpoints include time to documented MPN progression events and overall survival. Secondary endpoints include PROs, quality-of-life measures (e.g., EORTC QLQ-C30), and healthcare utilization patterns. Overall survival and cause-specific mortality endpoints will be determined using EMR data and validated through linkage with national mortality datasets, ensuring robust and comprehensive survival analyses. The SAP will undergo formal annual reviews and be updated based on interim analysis results, changes in data availability, or emerging scientific questions. Significant amendments will be documented, scientifically justified, and communicated to the IRB. Data will be securely retained for a minimum of 10 years following study completion, in accordance with institutional policies and regulatory requirements.

9.2 Interim Analyses and Data Monitoring
Interim analyses will be conducted annually to evaluate data completeness, quality, and preliminary outcomes. These analyses will be reviewed by the Steering Committee and DAUP Committee, who will advise on data integrity, patient safety, and protocol compliance. Criteria triggering protocol amendments or corrective actions, such as enrollment targets or data quality thresholds, will be generally defined in the SAP. Corrective actions may include targeted data validation audits, protocol amendments, re-training of staff, or adjustments to participant follow-up procedures. Protocol deviations or violations identified during interim analyses will be documented, categorized, and reported to the IRB. Systematic issues leading to protocol deviations will trigger corrective actions.

9.3 Subgroup Analyses
Planned subgroup analyses will examine outcomes based on demographic factors (e.g., age, sex, ancestral background), MPN subtype, treatment modalities, and genetic markers. Subgroup analysis methodology will be explicitly defined in the SAP, including interaction tests in regression models to ensure scientific rigor.

9.4 Handling Multiple Testing

Appropriate statistical methods (e.g., Bonferroni, Holm-Bonferroni, False Discovery Rate[28]) will be employed to adjust for multiple comparisons, ensuring control of Type I error rates. Details of the selected adjustment method will be clearly specified in the SAP.

9.5 Competing Risks and Time-dependent Covariate Analyses
Competing risk analyses, such as Fine and Gray's subdistribution hazard models[29], will be utilized to appropriately account for events that might prevent or alter the probability of the primary outcomes, ensuring unbiased estimates of progression risk. Additionally, time-dependent covariate analyses will evaluate how changes in treatment regimens, symptom severity, or other relevant covariates over time influence disease progression or outcomes.

9.6 Missing Data and Sensitivity Analyses
Patients or their legal representatives lost to follow-up will be contacted via phone or electronically (e.g., email). Missing data will be clearly identified and handled in line with the principles detailed in the SAP. While primary analyses will rely on complete-case data, sensitivity analyses will evaluate the impact of missing data assumptions using multiple imputation methods[30] and other appropriate strategies to ensure robustness of study findings. Additional sensitivity analyses will include assessments of the robustness of results across varying assumptions for distributional assumptions, model specifications, and subgroup effects. Patients not meeting the minimum data requirements will be excluded from analyses.

9.7 Propensity Score Methods
Propensity score matching or adjustment methods[31] will be utilized to minimize confounding and selection bias when analyzing comparative effectiveness among treatment groups or exposure categories. Methodologies and matching criteria will be explicitly described in the SAP.

9.8 Statistical Software and Analysis Environment
Statistical analyses will be conducted using validated statistical software, including SAS (version 9.4 or later), R (version 4.2 or later), or STATA. Analytical scripts and procedures will be documented and version-controlled to facilitate reproducibility and audits by IRB and regulatory authorities.

9.9 Consideration for Potential Data Biases
This study is an observational, non-interventional registry that relies on data collected from multiple sources, including EMRs, PROs, validated survey instruments, and claims data. Such a study design is inherently associated with data biases, including misclassification, recall bias, and missing data, which cannot always be independently verified by investigators.

However, the large sample size, objective disease assessments, and multi-source data validation minimize these limitations. To further enhance data reliability, the Registry will implement a structured quality control framework in collaboration with the data vendor. This framework includes:

- Automated Data Validation Checks – Incoming data will be systematically screened for missing values, internal inconsistencies, and entry errors using predefined validation rules aligned with ICD-10, SNOMED CT, and LOINC terminologies.

- Cross-Source Data Verification – Self-reported patient data will be cross-checked with EMR-derived records, laboratory values, and claims data to ensure consistency and flag discrepancies requiring manual review.
- Longitudinal Data Quality Monitoring – The completeness and accuracy of collected data will be reviewed at predefined study milestones (e.g., 6 months, 12 months, 18 months, annually) to detect unexpected deviations that may indicate systematic errors or reporting inconsistencies.
- Audit Trails and Data Traceability – All modifications to the dataset, including data corrections, updates, and deletions, will be logged through an automated audit trail system to ensure full traceability and compliance with research standards.
- Outlier Detection and Resolution – Advanced statistical methods and machine learning algorithms will be used to flag outlier data points that significantly deviate from expected clinical patterns. Flagged data will undergo secondary review before inclusion in analyses.
- Investigator-Led Quality Assurance Reviews – The PI and designated data management personnel will conduct periodic reviews to assess adherence to protocol-defined data quality standards, document findings, and implement corrective actions.
- Regulatory & Stakeholder Oversight – The data vendor will provide quarterly reports detailing data completeness, validation performance, and unresolved discrepancies. These reports will be reviewed in collaboration with the Registry's scientific steering committee, aligning with regulatory best practices for RWE studies.

By integrating these multi-tiered quality control measures, the Registry ensures that collected data is accurate, complete, and reliable, enabling robust observational research while minimizing bias in study findings.

9.10. Disseminating and Communicating Study Results

Study results will be disseminated through multiple channels, ensuring broad access to findings. Peer-reviewed journal articles will be submitted to peer-reviewed medical and scientific journals. Findings will be presented at relevant hematology and oncology conferences. Lay summaries and newsletters will be distributed to study participants and patient advocacy organizations. Interactive webinars and educational materials will be developed to engage the broader community. Policy briefs will be prepared to inform healthcare decision-makers, regulatory agencies, and industry stakeholders. Real-time access to aggregated, de-identified data insights will be provided to key stakeholders through a secure online platform.

10. APPENDIX

10.1 WHO Criteria for Polycythemia Vera

Major criteria
1. Hemoglobin >16.5 g/dL in men, Hemoglobin >16.0 g/dL in women
or,
Hematocrit >49% in men, Hematocrit >48% in women
or,
increased red cell mass (RCM), more than 25% above mean normal predicted value.
2. BM biopsy showing hypercellularity for age with trilineage growth (panmyelosis) including prominent erythroid, granulocytic, and megakaryocytic proliferation with pleomorphic, mature megakaryocytes (differences in size)
3. Presence of *JAK2*V617F or *JAK2* exon 12 mutation
Minor criterion
Subnormal serum erythropoietin level
Diagnosis of PV requires meeting either all 3 major criteria, or the first 2 major criteria and the minor criterion†

Criterion number 2 (BM biopsy) may not be required in cases with sustained absolute erythrocytosis: hemoglobin levels >18.5 g/dL in men (hematocrit, 55.5%) or >16.5 g/dL in women (hematocrit, 49.5%) if major criterion 3 and the minor criterion are present. However, initial myelofibrosis (present in up to 20% of patients) can only be detected by performing a BM biopsy; this finding may predict a more rapid progression to overt myelofibrosis (post-PV MF).

10.2 WHO Diagnostic Criteria for Essential Thrombocythemia

Major criteria
1. Platelet count ≥450 × 10^9/L
2. BM biopsy showing proliferation mainly of the megakaryocyte lineage with increased numbers of enlarged, mature megakaryocytes with hyperlobulated nuclei. No significant increase or left shift in neutrophil granulopoiesis or erythropoiesis and very rarely minor (grade 1) increase in reticulin fibers
3. Not meeting WHO criteria for *BCR-ABL1*$^+$ CML, PV, PMF, myelodysplastic syndromes, or other myeloid neoplasms
4. Presence of *JAK2*, *CALR*, or *MPL* mutation
Minor criterion
Presence of a clonal marker or absence of evidence for reactive thrombocytosis
Diagnosis of ET requires meeting all 4 major criteria or the first 3 major criteria and the minor criterion

10.3 WHO Diagnostic Criteria for Primary Myelofibrosis

Major criteria
1. Presence of megakaryocytic proliferation and atypia, accompanied by either reticulin and/or collagen fibrosis grades 2 or 3*
2. Not meeting WHO criteria for ET, PV, *BCR-ABL1*$^{+}$ CML, myelodysplastic syndromes, or other myeloid neoplasms
3. Presence of *JAK2*, *CALR*, or *MPL* mutation or in the absence of these mutations, presence of another clonal marker, † or absence of reactive myelofibrosis‡
Minor criteria
Presence of at least 1 of the following, confirmed in 2 consecutive determinations:
a. Anemia not attributed to a comorbid condition
b. Leukocytosis ≥11 × 10^9/L
c. Palpable splenomegaly
d. LDH increased to above upper normal limit of institutional reference range
e. Leukoerythroblastosis
Diagnosis of overt PMF requires meeting all 3 major criteria, and at least 1 minor criterion

†In the absence of any of the 3 major clonal mutations, the search for the most frequent accompanying mutations (eg, *ASXL1*, *EZH2*, *TET2*, *IDH1/IDH2*, *SRSF2*, *SF3B1*) are of help in determining the clonal nature of the disease.
‡BM fibrosis secondary to infection, autoimmune disorder, or other chronic inflammatory conditions, hairy cell leukemia or other lymphoid neoplasm, metastatic malignancy, or toxic (chronic) myelopathies.

10.4. WHO Diagnostic Criteria for pre-PMF

Major criteria
1. Megakaryocytic proliferation and atypia, without reticulin fibrosis >grade 1*, accompanied by increased age-adjusted BM cellularity, granulocytic proliferation, and often decreased erythropoiesis
2. Not meeting the WHO criteria for *BCR-ABL1*$^{+}$ CML, PV, ET, myelodysplastic syndromes, or other myeloid neoplasms
3. Presence of *JAK2*, *CALR*, or *MPL* mutation or in the absence of these mutations, presence of another clonal marker, † or absence of minor reactive BM reticulin fibrosis‡
Minor criteria
Presence of at least 1 of the following, confirmed in 2 consecutive determinations:
a. Anemia not attributed to a comorbid condition
b. Leukocytosis ≥11 × 10^9/L
c. Palpable splenomegaly
d. LDH increased to above upper normal limit of institutional reference range
Diagnosis of prePMF requires meeting all 3 major criteria, and at least 1 minor criterion

†In the absence of any of the 3 major clonal mutations, the search for the most frequent accompanying mutations (e.g., *ASXL1*, *EZH2*, *TET2*, *IDH1/IDH2*, *SRSF2*, *SF3B1*) are of help in determining the clonal nature of the disease.
‡Minor (grade 1) reticulin fibrosis secondary to infection, autoimmune disorder or other chronic inflammatory conditions, hairy cell leukemia or other lymphoid neoplasm, metastatic malignancy, or toxic (chronic) myelopathies

10.5. IWG-MRT Criteria for Secondary Myelofibrosis

Criteria for post-polycythemia vera myelofibrosis	
Required criteria:	
1	Documentation of a previous diagnosis of polycythemia vera as defined by the WHO criteria
2	Bone marrow fibrosis grade 2–3 (on 0–3 scale) or grade 3–4 (on 0–4 scale),[a]
Additional criteria (two are required):	
1	Anemia[b] or sustained loss of requirement of either phlebotomy (in the absence of cytoreductive therapy) or cytoreductive treatment for erythrocytosis
2	A leukoerythroblastic peripheral blood picture
3	Increasing splenomegaly defined as either an increase in palpable splenomegaly of ≥5 cm (distance of the tip of the spleen from the left costal margin) or the appearance of a newly palpable splenomegaly
4	Development of ≥1 of three constitutional symptoms: >10% weight loss in 6 months, night sweats, unexplained fever (>37.5°C)
Criteria for post-essential thrombocythemia myelofibrosis	
Required criteria:	
1	Documentation of a previous diagnosis of essential thrombocythemia as defined by the WHO criteria
2	Bone marrow fibrosis grade 2–3 (on 0–3 scale) or grade 3–4 (on 0–4 scale),[a]
Additional criteria (two are required):	
1	Anemia[b] and a ≥2 mg ml−1 decrease from baseline hemoglobin level
2	A leukoerythroblastic peripheral blood picture
3	Increasing splenomegaly defined as either an increase in palpable splenomegaly of ≥5 cm (distance of the tip of the spleen from the left costal margin) or the appearance of a newly palpable splenomegaly
4	Increased LDH (above reference level)
5	Development of ≥1 of three constitutional symptoms: >10% weight loss in 6 months, night sweats, unexplained fever (>37.5°C)

Abbreviations: IWG-MRT, International Working Group for Myelofibrosis Research and Treatment; LDH, lactate dehydrogenase; post-ET MF, post-essential thrombocythemia myelofibrosis; post-PV MF, post-polycythemia vera myelofibrosis.

[a]Grade 2–3 according to the European classification: 3 diffuse, often coarse fiber network with no evidence of collagenization (negative trichrome stain) or diffuse, coarse fiber network with areas of collagenization (positive trichrome stain). Grade 3–4 according to the standard classification:4 diffuse and dense increase in reticulin with extensive intersections, occasionally with only focal bundles of collagen and/or focal osteosclerosis or diffuse and dense increase in reticulin with extensive intersections with coarse bundles of collagen, often associated with significant osteosclerosis.

[b]Below the reference range for appropriate age, sex, gender and altitude considerations.

10.6: Pre-Screening Questions for Participant Eligibility
The following questions will be presented through the Registry online portal to assess participant eligibility prior to consent. Participants who do not meet the inclusion criteria will be automatically disqualified and will not proceed to the informed consent stage. Participants providing responses to these pre-screening questions will be asked to provide their permission to utilize and retain this basic information by MPN Research Foundation for the sole purpose of understanding the basic characteristics of those interested participants who are eligible versus those who are ineligible. The MPN Research Foundation store data collected as part of this pre-screening process in a secure, HIPAA-compliant system, with access restricted to authorized study personnel.

1. **Are you at least 18 years old?**
   Options:

   - Yes, I am 18 or older
   - No
     → If No → participant is ineligible

2. **Do you currently live in the United States?**
   Options:

   - Yes
   - No
     → If No → participant is ineligible

3. **Have you ever been diagnosed with a myeloproliferative neoplasm (MPN) by a healthcare provider?**
   Options:

   - Yes, confirmed by hematologist/oncologist
   - No, still undergoing testing or unsure
     → If No → participant is ineligible

- **Have you ever had a hematopoietic stem cell transplant (HSCT) — either a bone marrow transplant (BMT), peripheral or umbilical stem cell transplant (SCT) — after an MPN diagnosis?**
  - Yes
  - No
- **When were you diagnosed with your MPN?**
  Options:

  - Enter date (MM/YYYY)
  - If only year known, enter “07/YYYY”
  - If unsure, write “Unsure”
    → No exclusion applied

5. **Which type of MPN have you been diagnosed with? (Select all that apply)**
   Options:

   - Polycythemia Vera (PV)
   - Essential Thrombocythemia (ET)
   - Primary Myelofibrosis (PMF)
   - Post–PV Myelofibrosis (Post-PV MF)
   - Post–ET Myelofibrosis (Post-ET MF)
   - Pre-fibrotic Primary Myelofibrosis (pre-PMF)
   - MPN-Unclassifiable (MPN-U)
   - MPN in Accelerated Phase (MPN-AP)
   - MPN in Blast Phase (MPN-BP) or post-MPN AML
   - Myelodysplatic Syndrome/Myeloproliferative Neoplasm (MDS/MPN) Overlap Syndrome
   - Chronic Myelomonocytic Leukemia (CMML)
   - Other (please specify)
     → No exclusion applied

6. **Has your MPN diagnosis ever been revised to a different condition?**
   Options:

   - No, diagnosis is the same
   - Yes, changed to another condition
     → If Yes → proceed to Question 7

7. **If yes, what is your new diagnosis?**
   Options:

   - Myelodysplastic Syndromes (MDS) not including MDS/MPN overlap syndrome
   - Acute Myeloid Leukemia (AML)
   - Chronic Myeloid Leukemia (CML)
   - Other blood cancer or hematologic malignancy
   - Other non-cancer blood disorder
   - Unsure
   - Other (please specify)
     → If revised to non-MPN malignancy → participant is ineligible

8. **When was your diagnosis revised?**
   Options:

   - Enter date (MM/YYYY)
   - Approximate year or “Unsure”
     → No exclusion applied

Commented [DS1]: Note: we need to specify that sAML is eligible (MPN-BP) to avoid any confusion

Commented [DT2R1]: Agree

9. **Do you currently receive medical care for your MPN and have access to past medical records (e.g., online portal)?**
   Options:

   - Yes
   - No
     → If No → participant is ineligible

10. **Do you have access to your electronic medical records (EMR) from one or more participating healthcare institutions?**
    Options:

    - Yes
    - No
      → If No → participant is ineligible

11. **Has a healthcare provider ever told you that your condition may be reactive (not an MPN)?**
    Options:

    - No, I have an MPN
    - Yes, I have been told it may be reactive
      → If Yes → participant is ineligible

12. **Have you been diagnosed with any of the following blood disorders? (Select all that apply)**
    Options:

- Chronic Myeloid Leukemia (CML)
- Myelodysplastic Syndromes (MDS) not including MDS/MPN overlap syndrome
- None of the above
  → If any except "None of the above" → participant is ineligible

13. **Are you currently enrolled in a clinical trial that does not allow your medical records to be shared with external research studies?**
    Options:

    - No
    - Yes
      → If Yes → participant is ineligible

14. **Are you able to provide informed consent, or do you have a legally authorized representative (LAR) who can provide consent?**
    Options:

    - Yes, I can consent for myself

- Yes, my LAR can consent
- No, I cannot consent and have no LAR
  → If No → participant is ineligible

15. **This registry requires long-term participation, including data sharing, surveys, and follow-up over multiple years. Do you understand and agree to these terms?**
    Options:

    - Yes
    - No
      → If No → participant is ineligible

10.7: Participant Notification Messages
The following participant notifications are used within the Registry portal to guide participants through the screening, consent, enrollment, and withdrawal process. These messages are designed to ensure respectful, clear, and ethically appropriate communication at each decision point. No personally identifiable information is retained for individuals who are screened as ineligible. All content aligns with IRB-approved consent language and complies with applicable legal and regulatory standards.

1. **Disqualified Notification**
   Displayed when a participant does not meet eligibility criteria after pre-screening.

   Thank you for your interest.
   Based on your responses, you are not eligible to participate in the MPN PROGRESSion Registry at this time.

   We appreciate your time and your interest in helping advance MPN research.

2. **Eligible Notification**
   Displayed when a participant meets all eligibility criteria and is cleared to review the informed consent form.

   Thank you for completing the eligibility questions.
   You are eligible to participate in the MPN PROGRESSion Registry.

   Please proceed to the informed consent form (ICF) to learn more about the study and provide consent.

3. **Consent Completion / Enrollment Confirmation Notification**
   Displayed after a participant successfully completes electronic informed consent.

   Thank you for enrolling in the MPN PROGRESSion Registry.
   We appreciate your participation and look forward to working with you.

   Please check your registry dashboard for surveys, updates, and study communications.

4. **Withdrawal or Opt-Out Confirmation Notification**
   Displayed when a participant withdraws or opts out of the registry.

   You have successfully withdrawn from the MPN PROGRESSion Registry.
   Thank you for your contributions to MPN research.

   If you have any questions, please contact the registry team at [insert contact email or phone number].

5. **Welcome / Onboarding Message (Optional)**
   Displayed after enrollment to welcome the participant and guide them to the next steps.

Welcome to the MPN PROGRESSion Registry!
You are now enrolled. Please check your dashboard for upcoming surveys and study updates.

Thank you for helping advance research and improve care for people living with MPNs.

11. REFERENCES

1. Khoury, J.D., Solary, E., Abla, O. et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Myeloid and Histiocytic/Dendritic Neoplasms. *Leukemia*. 2022;36:1703-1719.
2. Tremblay D, Yacoub A, Hoffman R. Overview of Myeloproliferative Neoplasms: History, Pathogenesis, Diagnostic Criteria, and Complications. *Hematol Oncol Clin North Am.* 2021;35(2):159-176.
3. Kralovics R, Passamonti F, Buser AS, et al. A gain-of-function mutation of JAK2 in myeloproliferative disorders. *The New England journal of medicine.* 2005;352(17):1779-1790.
4. Klampfl T, Gisslinger H, Harutyunyan AS, et al. Somatic mutations of calreticulin in myeloproliferative neoplasms. *The New England journal of medicine.* 2013;369(25):2379-2390.
5. Mesa R, Miller CB, Thyne M, et al. Myeloproliferative neoplasms (MPNs) have a significant impact on patients' overall health and productivity: the MPN Landmark survey. *BMC cancer.* 2016;16:167.
6. Tefferi A, Guglielmelli P, Larson DR, et al. Long-term survival and blast transformation in molecularly annotated essential thrombocythemia, polycythemia vera, and myelofibrosis. *Blood.* 2014;124(16):2507-2513; quiz 2615.
7. Moliterno AR, Ginzburg YZ, Hoffman R. Clinical insights into the origins of thrombosis in myeloproliferative neoplasms. *Blood.* 2021;137(9):1145-1153.
8. Ginzburg YZ, Feola M, Zimran E, Varkonyi J, Ganz T, Hoffman R. Dysregulated iron metabolism in polycythemia vera: etiology and consequences. *Leukemia.* 2018;32(10):2105-2116.
9. Hasselbalch HC. Perspectives on chronic inflammation in essential thrombocythemia, polycythemia vera, and myelofibrosis: is chronic inflammation a trigger and driver of clonal evolution and development of accelerated atherosclerosis and second cancer? *Blood.* 2012;119(14):3219-3225.
10. Tefferi A, Mudireddy M, Mannelli F, et al. Blast phase myeloproliferative neoplasm: Mayo-AGIMM study of 410 patients from two separate cohorts. *Leukemia.* 2018;32(5):1200-1210.
11. Shallis RM, Wang R, Davidoff A, Ma X, Podoltsev NA, Zeidan AM. Epidemiology of the classical myeloproliferative neoplasms: The four corners of an expansive and complex map. *Blood Rev.* 2020;42:100706.
12. Titmarsh GJ, Duncombe AS, McMullin MF, et al. How common are myeloproliferative neoplasms? A systematic review and meta-analysis. *American journal of hematology.* 2014;89(6):581-587.
13. Tefferi A, Lasho TL, Jimma T, et al. One thousand patients with primary myelofibrosis: the mayo clinic experience. *Mayo Clinic proceedings.* 2012;87(1):25-33.
14. Masarova L, Alhuraiji A, Bose P, et al. Significance of thrombocytopenia in patients with primary and postessential thrombocythemia/polycythemia vera myelofibrosis. *European journal of haematology.* 2018;100(3):257-263.
15. Baumeister J, Chatain N, Sofias AM, Lammers T, Koschmieder S. Progression of Myeloproliferative Neoplasms (MPN): Diagnostic and Therapeutic Perspectives. *Cells.* 2021;10(12).

16. Barosi G, Mesa RA, Thiele J, et al. Proposed criteria for the diagnosis of post-polycythemia vera and post-essential thrombocythemia myelofibrosis: a consensus statement from the International Working Group for Myelofibrosis Research and Treatment. *Leukemia.* 2008;22(2):437-438.
17. Mesa RA, et al. Myeloproliferative neoplasms have a significant impact on patients' overall health and productivity: the MPN Landmark survey. BMC Cancer. 2016;16:167.
18. Richesson RL, Krischer JP. Data standards in clinical research: gaps, overlaps, challenges, and future directions. J Am Med Inform Assoc. 2007;14(6):687-696.
19. Sherman RE, et al. Real-world evidence: What is it and what can it tell us? N Engl J Med. 2016;375:2293-2297.
20. Verstovsek S, Mesa RA, Gotlib J, et al. A double-blind, placebo-controlled trial of ruxolitinib for myelofibrosis. *N Engl J Med.* 2012;366(9):799-807. doi:10.1056/NEJMoa1110557.
21. Harrison CN, Schaap N, Vannucchi AM, et al. Janus kinase-2 inhibitor fedratinib in patients with myelofibrosis previously treated with ruxolitinib (JAKARTA-2): a single-arm, open-label, non-randomised, phase 2, multicentre study. *Lancet Haematol.* 2017;4(7):e317-e324. doi:10.1016/S2352-3026(17)30088-1.
22. Mascarenhas J, Hoffman R, Talpaz M, et al. Pacritinib vs best available therapy, including ruxolitinib, in patients with myelofibrosis: a randomized clinical trial. *JAMA Oncol.* 2018;4(5):652-659. doi:10.1001/jamaoncol.2017.5818.
23. Vannucchi AM, Kiladjian JJ, Griesshammer M, et al. Ruxolitinib versus standard therapy for the treatment of polycythemia vera. *N Engl J Med.* 2015;372(5):426-435. doi:10.1056/NEJMoa1409002.
24. Mesa RA, Verstovsek S, Gupta V, et al. Overall survival in the SIMPLIFY-1 and SIMPLIFY-2 phase 3 trials of momelotinib in patients with myelofibrosis. *Leukemia.* 2022;36(9):2261-2268. doi:10.1038/s41375-022-01657-2.
25. Masarova L, Bose P, Pemmaraju N, et al. The importance of longitudinal symptom assessment in myeloproliferative neoplasms. *Leuk Lymphoma.* 2018;59(2):372-382. doi:10.1080/10428194.2017.1339873.
26. Kaplan, E.L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282), 457-481.
27. Cox, D.R. (1972). Regression models and life tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2), 187-220.
28. Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289-300.
29. Fine, J.P., & Gray, R.J. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*, 94(446), 496-509.
30. Little, R.J., & Rubin, D.B. (2019). *Statistical Analysis with Missing Data* (3rd ed.). Hoboken, NJ: Wiley.
31. Austin, P.C. (2011). An introduction to propensity score methods for reducing confounding in observational studies. *Multivariate Behavioral Research*, 46(3), 399-424.
32. U.S. Food and Drug Administration (FDA). (2021). *Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products*.
33. Menachemi N, Collum TH. Benefits and drawbacks of electronic health record systems. *Risk Manag Healthc Policy*. 2011;4:47–55.

34. Shah NH, Tenenbaum JD. The coming age of data-driven medicine: translational bioinformatics' next frontier. *JAMA*. 2012;309(14):1417–1418.