

Mesa R.A., Hudgens S., Floden L., Palmer J., Gupta V., McLornan D., McMullin M.F., Kiladjian J.J., Foltz L., Platzbecker U., Fox M.L., Mead A.J., Ross D.M., Oh S.T., Perkins A., Leahy M.F., Deheshi S., Donahue R., Klencke B., Verstovsek S.

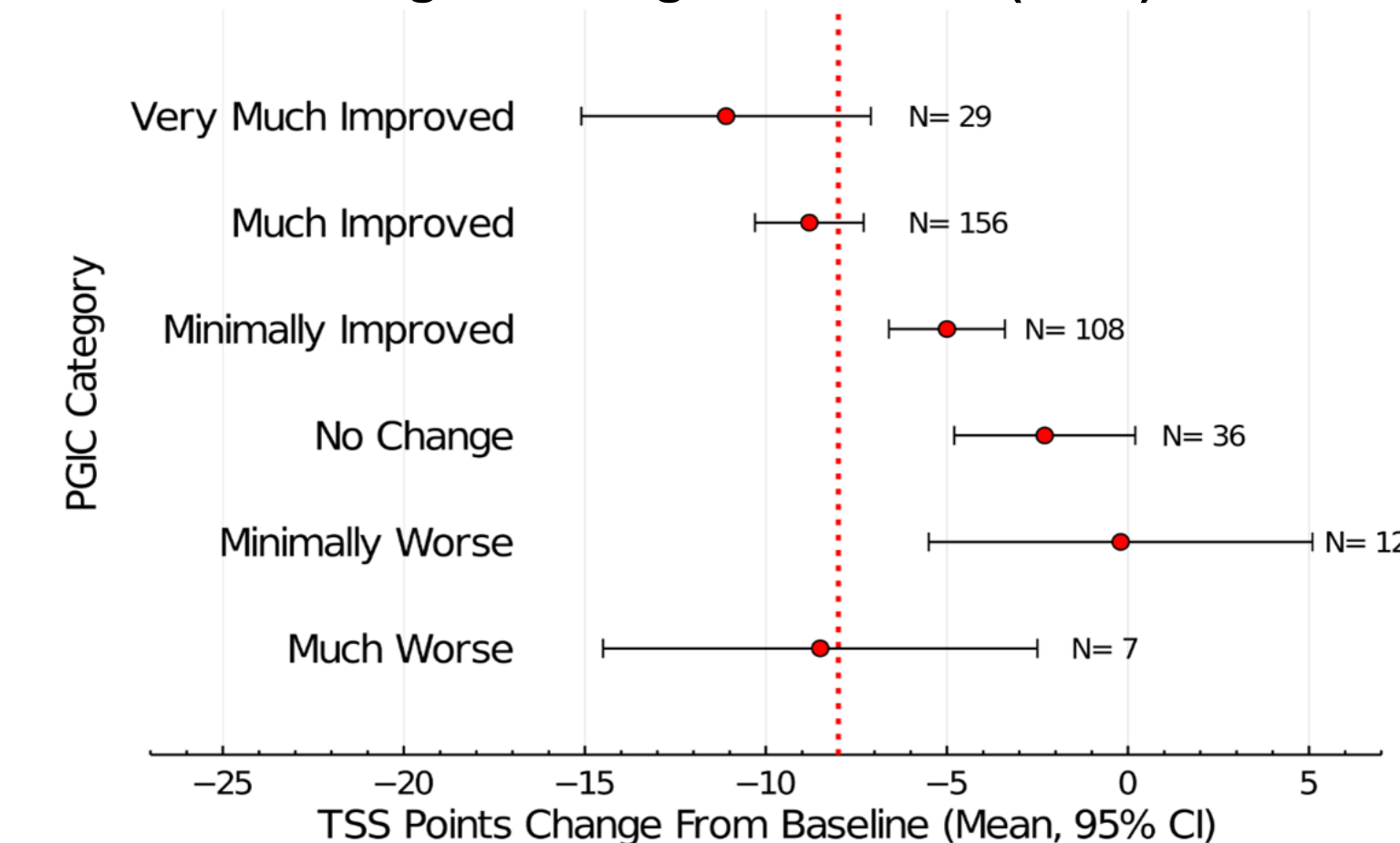
UT Health San Antonio MD Anderson, San Antonio, TX, USA; Clinical Outcomes Solutions, Tucson, AZ, USA; Mayo Clinic, Phoenix, AZ, USA; University Health Network, University of Toronto, Toronto, ON, Canada; Guy's &amp; St Thomas' NHS Foundation Trust, London, UK; Queens University, Belfast City Hospital Trust, Belfast, UK; Hôpital Saint-Louis; Université de Paris, Paris, France; St Paul's Hospital, University of British Columbia, Vancouver, BC, Canada; Leipzig University Hospital, Leipzig, Germany; Hematology Department, Hospital Universitario Vall d'Hebron, Barcelona, Spain; MRC Weatherall Institute of Molecular Medicine, Oxford, UK; Flinders Medical Centre and University, Adelaide, SA, Australia; Washington University School of Medicine, St. Louis, MO, USA; Monash University, Malvern, SA, Australia; University of Western Australia, Perth, WA Australia; Sierra Oncology, Vancouver, BC, Canada; MD Anderson Cancer Center, Houston, TX, USA

## Background and Objective

- Constitutional symptoms cause a significant burden for patients diagnosed with myelofibrosis (MF).
- While patient reported outcomes (PRO) instruments such as the myeloproliferative neoplasm symptom assessment form (MPN-SAF) have been prioritized to interpret efficacy in recent years, the focus has been on domain level rather than individual symptom changes.
- Clinical trials investigating JAK inhibitors for MF have historically measured symptom improvement using a landmark analysis requiring a minimum 50% reduction from baseline in total symptom score (TSS) over the last 28 days of the 24-week treatment period. However, this analysis requires vastly different absolute TSS improvements for patients with very high (symptomatic) or low baseline TSS (asymptomatic) to reach responder status.
- A Phase 3 clinical study, SIMPLIFY-1, randomized 432 intermediate and high risk JAK inhibitor (JAKi) naïve MF patients 1:1 to momelotinib (MMB) or ruxolitinib (RUX). Non-inferiority on the MPN-SAF TSS response rate endpoint at Week 24 (W24) was not met (MMB: 28% vs RUX: 42%); however, improvement in each of the 7 TSS items was similar for MMB vs RUX.
- To understand the discrepancy, we evaluated overall and symptomatic populations and applied item analysis to the MPN-SAF and mixed effect models for repeated measures (MMRM) to the MPN-SAF TSS for SIMPLIFY-1.

## Methods

- Analyses were conducted in the overall population and in a symptomatic subset (selected as patients with baseline TSS  $\geq 10$ ) to remove the distribution of asymptomatic patients.
- The meaningful change threshold (MCT) was determined using Patient Global Impression of Change (PGIC) then applied to a within-subject responder analysis.
- MMRM compared mean change in TSS from BL to W24 using data from all visits.
- The distributions of TSS items were examined at baseline (BL) and shift in scores at W24 (health state shifts) were determined.
- Odds ratios from generalized estimating equations (GEE) for improvement on MMB relative to RUX were calculated for each item of the TSS using multiple predictive imputations for missing data.

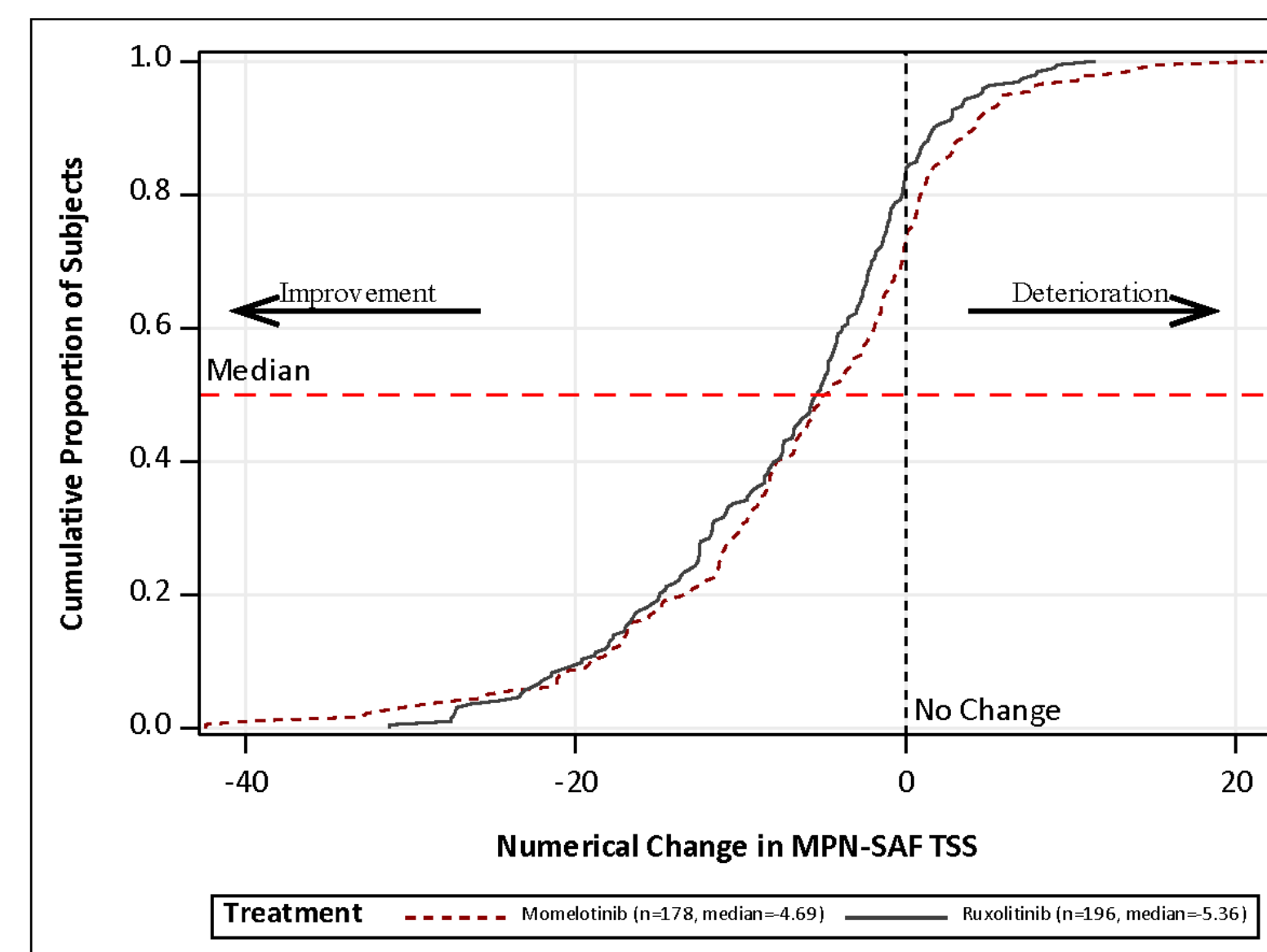
**Figure 1: Anchor-based Method at W24 Demonstrates an 8-Point Meaningful Change Threshold (MCT)**

- An **8-point** change was derived to ensure that patients at the lower bound of the confidence interval (CI) in the No Change category (i.e., symptomatic patients at a value of -7.9, not shown) were not considered to have achieved the MCT (Fig. 1, Overall Population).
- At W24, 39% of MMB and 41% of RUX patients in the overall population, respectively, showed a meaningful improvement of  $\geq 8$  points. In the symptomatic population, 56% and 58% of MMB and RUX patients, respectively, achieved an improvement of  $\geq 8$  points (not shown).

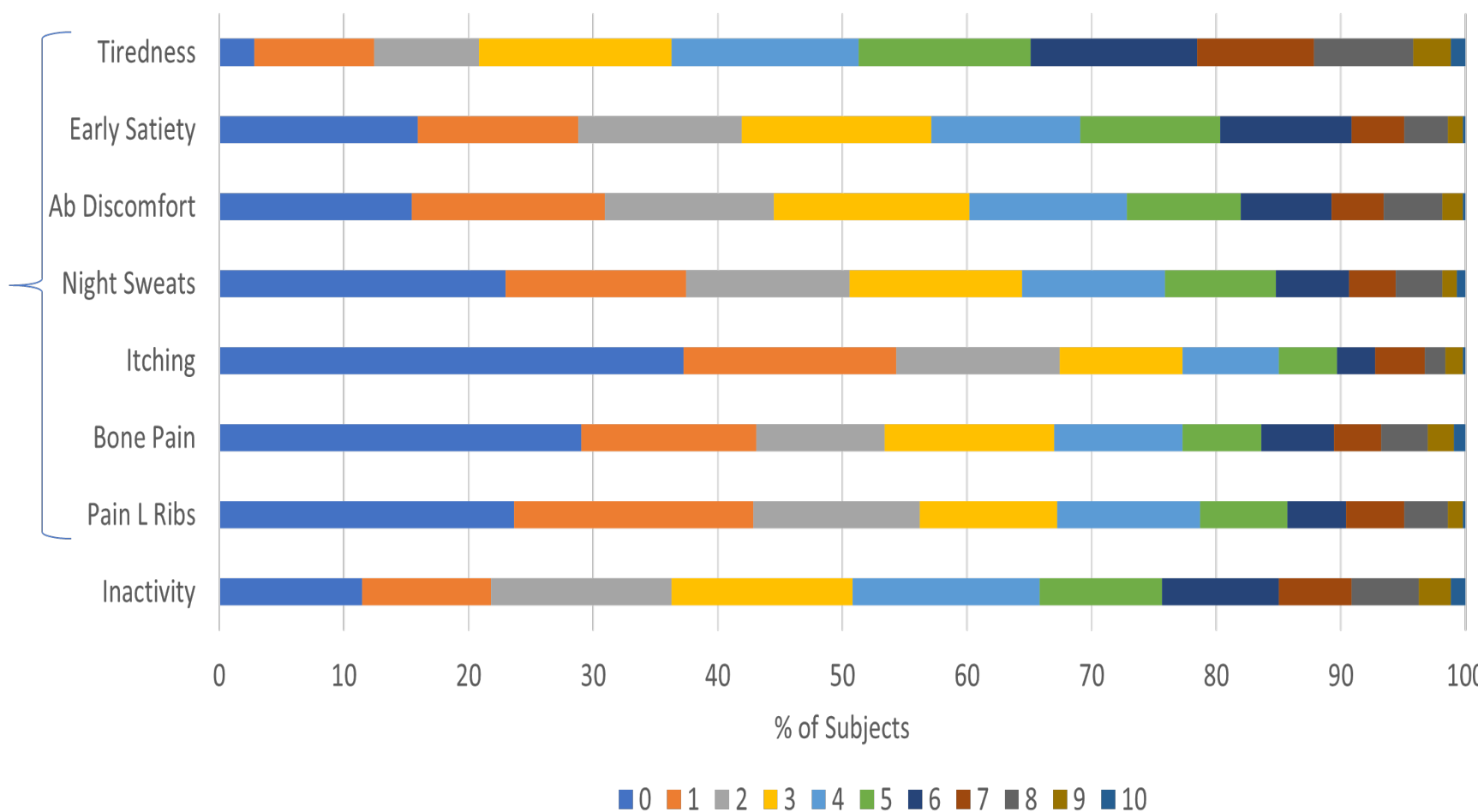
**Table 1. Improvement in Symptom Burden is Near the MCT for Both MMB and RUX**

|                              | Overall Population |               | Symptomatic Population (Baseline TSS $\geq 10$ ) |                |
|------------------------------|--------------------|---------------|--------------------------------------------------|----------------|
|                              | MMB (N=215)        | RUX (N=217)   | MMB (N=145)                                      | RUX (N=147)    |
| <b>Baseline</b>              |                    |               |                                                  |                |
| LS mean (SE)                 | 18.98 (1.726)      | 17.48 (1.738) | 25.13 (1.915)                                    | 23.11 (1.915)  |
| <b>Week 24</b>               |                    |               |                                                  |                |
| Overall Change from Baseline |                    |               |                                                  |                |
| LS mean (SE)                 | -6.35 (1.089)      | -7.87 (1.081) | -8.80 (1.429)                                    | -10.46 (1.408) |
| Difference from RUX          |                    |               |                                                  |                |
| LS mean (SE)                 | 1.52 (0.675)       | -             | 1.67 (0.917)                                     | -              |
| 95% CI                       | 0.196, 2.847       | -             | -0.134, 3.468                                    | -              |

- Table 1 shows the MMRM based TSS change from BL.
- Mean within-subject improvements was 6.35 points (MMB) and 7.87 points (RUX) in the overall population, both near but less than the within-subject MCT of 8-points. In the symptomatic population, the mean was 8.8 points (MMB) and 10.5 points (RUX), both exceeding the MCT.
- Treatment effect of between-treatment group comparison (LS mean) was 1.52 (SE 0.675) on a 70-point scale in the overall population, and 1.67 (0.917) in the symptomatic populations, demonstrating relatively equivalent results.

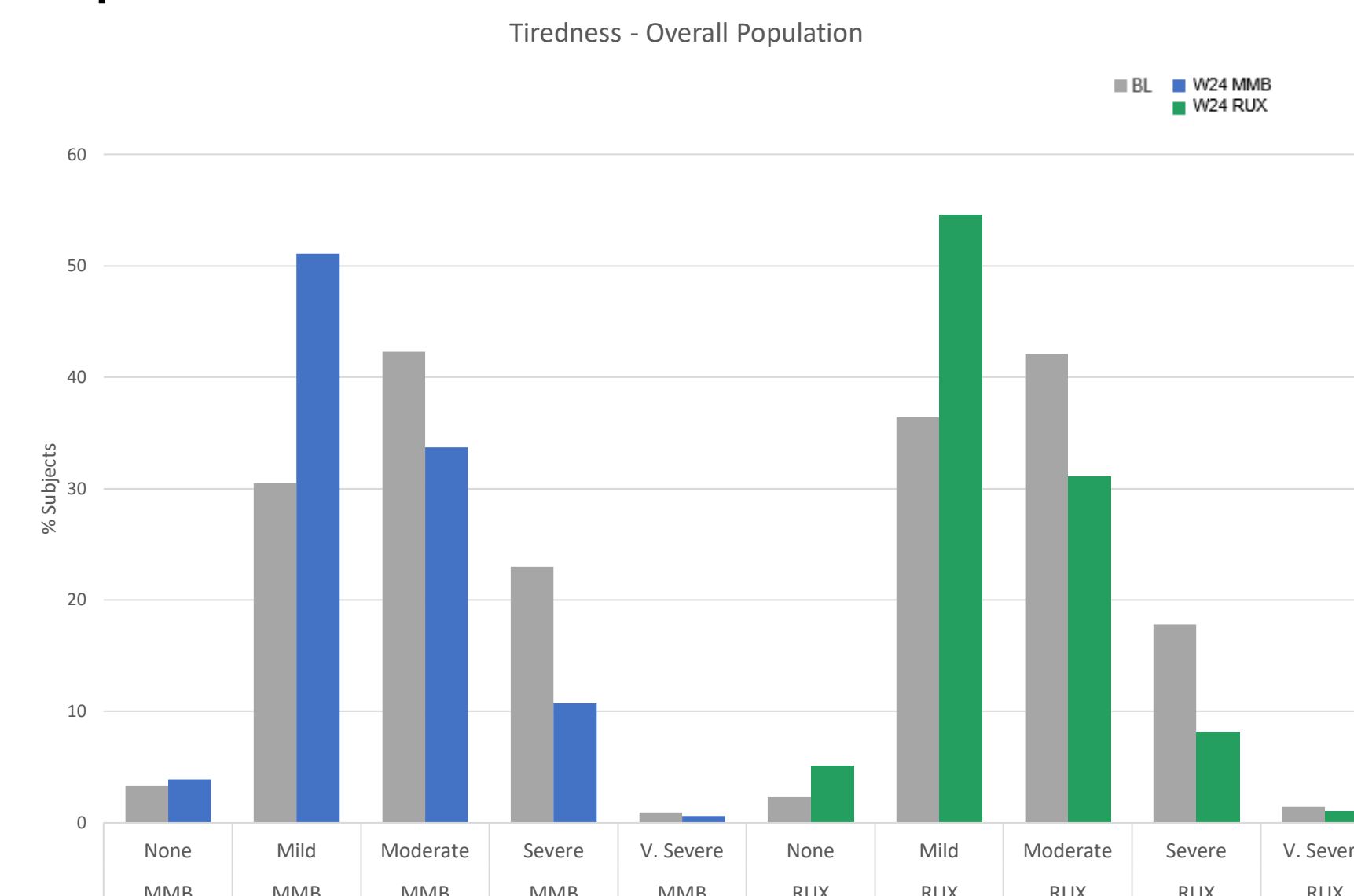
**Figure 2. Patients on MMB and RUX Experienced Similar Meaningful Improvements in Symptom Burden**

- Cumulative distribution function (CDF) curves demonstrated comparable distributions of change from BL for MMB and RUX in both the overall (Fig. 2) and symptomatic population (not shown).

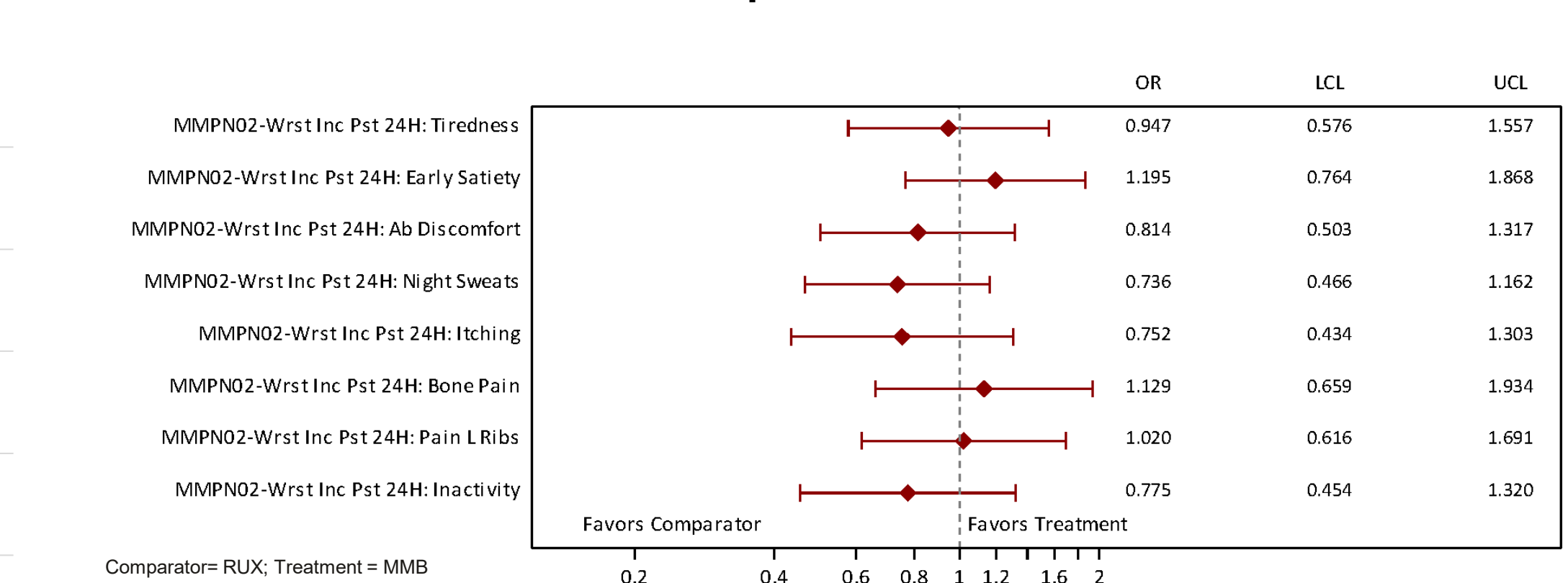
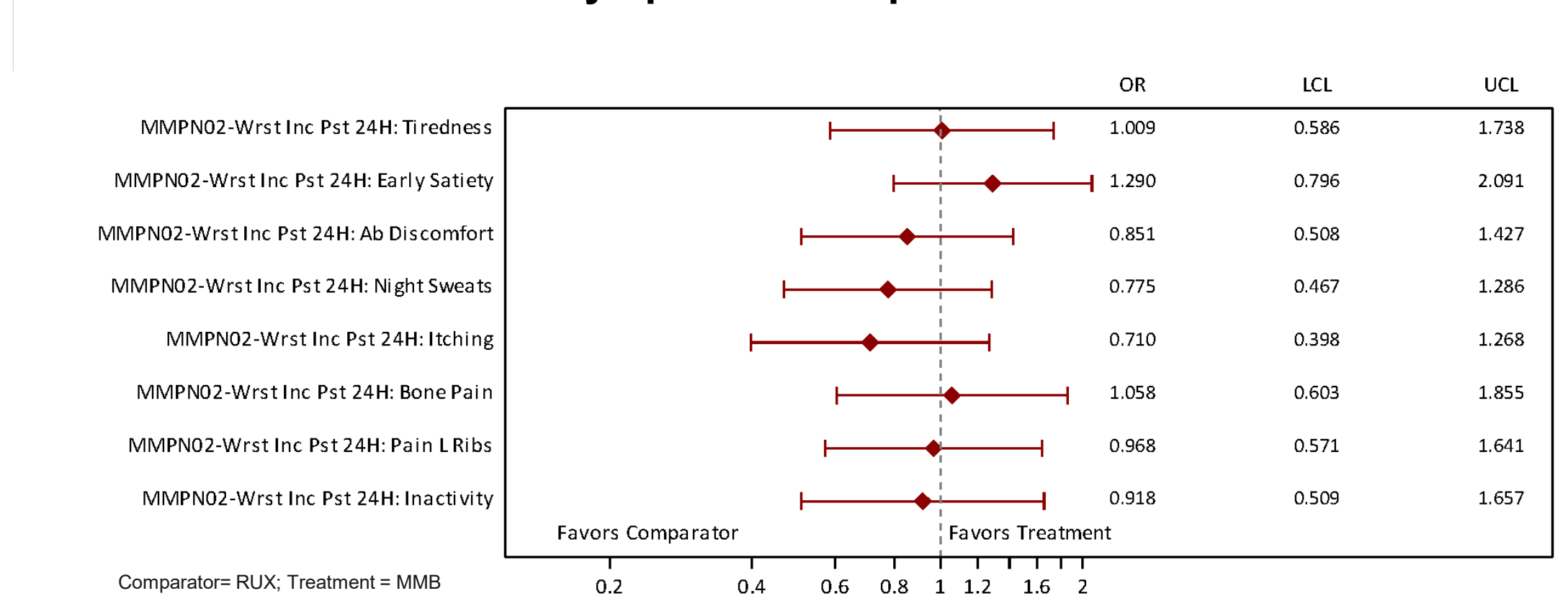
**Figure 3: Heterogeneity of Individual Items of the BL TSS value**

- Figure 3 shows the MPN-SAF TSS item distribution at BL in the overall population using aggregate data from both MMB and RUX treatment arms.
- The proportion of patients at BL with none (0 points) or mild symptoms (1-3 points) ranged from 36% (tiredness) to 77% (itching) (Fig. 3, Overall Population).
- Tiredness was the most prevalent and most severe item reported at BL.

## Results

**Figure 4: Health State Shifts Show Comparable Improvement on Individual Items**

- Baseline TSS in the MMB arm was 18.98 while BL TSS in the RUX arm was 17.48 (Table 1).
- While MMB patients were more likely to report severe or very severe scores on items at BL, the proportion of patients in the severe/very severe categories decreased similarly for both MMB and RUX at W24 (data for all items not shown).
- Correspondingly, the proportion of patients in the none/mild categories increased similarly for both MMB and RUX at W24.
- Figure 4 shows the proportion of patients reporting none/mild tiredness increased in the overall population by 21.2% vs 21.0% in MMB vs RUX.
- The magnitude of improvement in MMB vs RUX was similar for all items in the overall population and symptomatic population.

**Figure 5. Likelihood of Improvement on Individual Items is Similar for MMB and RUX Patients in the Overall Population****Figure 6. Likelihood of Improvement on Individual Items is Similar for MMB and RUX Patients in the Symptomatic Population**

- The odds ratios and confidence intervals that overlap 1.0 for each item in the overall (Fig. 5) and symptomatic population (Fig. 6) demonstrate that likelihood of improvement are similar for MMB and RUX.

## Conclusions

- SIMPLIFY-1 is the only head-to-head Phase 3 study comparing two highly active JAKi in the MF disease setting yet was impacted by low symptom burden at BL because of the failure to restrict enrollment to symptomatic patients.
- SIMPLIFY-1 may have also been impacted by the heterogeneity and imbalance in symptoms across MMB and RUX arms at BL.
- This study evaluated the MPN-SAF TSS using MMRM rather than the SIMPLIFY-1 secondary dichotomous responder endpoint definition (e.g., 50% improvement on the MPN-SAF TSS) and assessed the contribution of each individual item to symptom response.
- The MMRM and individual item results contrast with the % change binary responder endpoint.
- Similar improvements in mean TSS were shown by MMRM. A minimal between-group difference of 1.52 on the 70-point as well as comparable item health state shifts at W24 suggest that MMB provides clinically meaningful and comparable symptom improvements to RUX in JAKi-naïve patients with intermediate and high-risk MF. These analyses require further validation in independent data sets based on the modified MPN-SAF.
- Supplementing the TSS responder rate endpoint with discriminatory methods of TSS analysis quantifies important aspects of treatment response and allows for a more comprehensive understanding of symptom benefit.

\*\*\*Poster previously presented at virtual ISPOR May 17-20, 2021\*\*\*