Why molecular response in polycythemia vera remains an elusive endpoint

Written by The Life Science Feed Editorial Team · Edited by Mara Voss

Polycythemia vera (PV), a chronic myeloproliferative neoplasm, significantly impacts patient outcomes, yet optimal therapy remains unclear. The absence of head-to-head trials comparing newer agents like ropeginterferon alfa-2b-njft against established treatments creates a persistent knowledge gap for clinicians. A recent feasibility assessment for a network meta-analysis, published in J Comp Eff Res, aimed to bridge this gap by indirectly comparing ropeginterferon alfa-2b-njft with peginterferon alfa-2a or ruxolitinib, using hydroxyurea (HU) as a common comparator.1

KEY TAKEAWAYS

- The Pivot: A planned network meta-analysis of polycythemia vera treatments was deemed unfeasible due to profound heterogeneity across existing studies, preventing robust indirect comparisons.
- The Data: Of 40 included studies, only 20 formed connected evidence networks for endpoints, but inconsistencies in patient populations, response definitions, and follow-up durations precluded meaningful synthesis.
- The Action: Clinicians must continue to rely on direct trial evidence and individual patient characteristics, as broad comparative efficacy data for PV treatments remains unavailable from indirect analyses.

Polycythemia vera is a rare, chronic myeloproliferative neoplasm characterized by an overproduction of red blood cells, and often white blood cells and platelets. This uncontrolled proliferation increases the risk of thrombotic events, bleeding, and progression to myelofibrosis or acute myeloid leukemia. Current management strategies focus on controlling hematocrit, alleviating symptoms, and reducing thrombotic risk, typically involving phlebotomy and cytoreductive agents.1

Hydroxyurea (HU) has long served as a first-line cytoreductive therapy, but its use is associated with side effects and a subset of patients develop resistance or intolerance. Newer agents, including interferons and ruxolitinib, have emerged, offering alternatives for patients who do not respond adequately to or tolerate HU. But the direct comparative efficacy of these newer agents against each other, or against HU in specific patient subgroups, remains largely unaddressed by direct head-to-head trials.1

The Attempt to Synthesize Evidence

Walden and colleagues conducted a targeted literature review and feasibility assessment to determine if a network meta-analysis (NMA) could compare ropeginterferon alfa-2b-njft against peginterferon alfa-2a or ruxolitinib, using hydroxyurea as a common comparator.1 The team screened 193 PubMed records and 460 conference abstracts published

between January 2014 and May 2024. They ultimately included 40 records, representing evidence from 11 randomized controlled trials and 10 observational studies.1

The investigators focused on several key endpoints: complete hematologic response (CHR), molecular response (MR), allele burden, event-free survival, and safety. A network meta-analysis requires a connected evidence network, where treatments are linked through common comparators, allowing for indirect comparisons. Of the 40 included studies, only 20 formed connected evidence networks for the endpoints of interest. This initial reduction already highlighted a fragmented evidence base.1

The primary goal was to assess the homogeneity of patient populations, treatment regimens, and endpoint definitions across these studies, which is critical for a robust NMA. A network meta-analysis relies on the assumption that studies are sufficiently similar in design and patient characteristics that any differences in outcomes can be attributed to the treatments themselves, rather than confounding factors. Without this homogeneity, indirect comparisons become unreliable and potentially misleading.1

Why the Numbers Did Not Add Up

The feasibility assessment revealed substantial heterogeneity across the included studies, ultimately precluding a robust network meta-analysis. Patient populations varied significantly, encompassing newly diagnosed patients, high-risk and low-risk cohorts, and patients who were either hydroxyurea-refractory or -intolerant. This variation in baseline characteristics means that the observed treatment effects might not be comparable across studies, as the underlying disease biology and patient responsiveness could differ substantially between these groups.1

Definitions of complete hematologic response also differed, with some requiring the absence of disease-related symptoms in addition to hematologic parameters, while others did not. This inconsistency in outcome definitions makes it difficult to pool data or compare results directly. A CHR definition that includes symptom resolution, for example, sets a higher bar than one focused solely on blood counts, making a direct comparison of CHR rates between studies with different definitions problematic.1

Molecular response thresholds were inconsistent as well. Molecular response, often measured by the reduction in JAK2V617F allele burden, is an increasingly important endpoint in PV, reflecting a deeper disease control beyond hematologic parameters. But without standardized thresholds for what constitutes a molecular response (e.g., a 25% reduction, a 50% reduction, or complete molecular remission), comparing the efficacy of different treatments on this endpoint becomes impossible. This is a recurring issue in molecular testing in other conditions, where consensus on thresholds is still evolving.1

Follow-up durations varied considerably across the studies. Some trials had relatively short follow-up periods, while others extended for several years. This discrepancy is particularly problematic for a chronic disease like PV, where long-term outcomes, such as event-free survival and progression-free survival, are critical. Short follow-up periods may miss late-onset adverse events or long-term benefits, making it difficult to draw definitive conclusions about the comparative long-term efficacy and safety of treatments.1

The definition of standard of care also presented a challenge. In some studies, standard of care comprised almost exclusive use of hydroxyurea, while in others, it involved mixed regimens that could include phlebotomy, aspirin, or other cytoreductive agents. This variability in the comparator arm introduces further confounding, as the baseline treatment

context against which new therapies are evaluated is not consistent. This makes it difficult to isolate the specific effect of the investigational drug.1

The Broader Implications of Heterogeneity

The findings from Walden and colleagues highlight a fundamental challenge in generating robust comparative evidence for rare conditions like polycythemia vera. The lack of standardized clinical trial designs and outcome definitions across studies severely limits the ability to synthesize existing data, even with advanced statistical methods like network meta-analysis. This problem extends beyond PV, affecting many rare diseases where patient populations are small and trial designs often adapt to practical constraints.1

The inability to perform a reliable NMA means that clinicians are left without a clear hierarchy of evidence for choosing between ropeginterferon alfa-2b, peginterferon alfa-2a, ruxolitinib, and hydroxyurea in various patient subgroups. Each drug has its own profile, and individual trial results, while informative, do not provide the head-to-head comparisons that would offer definitive guidance. This forces clinicians to make treatment decisions based on a more fragmented understanding of the evidence, often relying on expert opinion and institutional experience. For a comprehensive overview of haematological conditions, the Oxford Handbook of Clinical Haematology remains a valuable resource.

The study highlights the importance of prospective, harmonized research efforts. Future clinical trials in PV, and indeed in other rare diseases, must adopt standardized patient selection criteria, consistent endpoint definitions, and comparable follow-up durations to enable meaningful comparisons. Without such standardization, the field will continue to struggle with an evidence base that is rich in individual studies but poor in synthesizable, comparative data. This issue is particularly relevant when considering the role of minimal residual disease (MRD) status, which is gaining traction as a prognostic marker but lacks consistent definitions across trials.1

The authors concluded that an NMA for PV treatments was not feasible due to significant clinical and methodological heterogeneity across studies. This included differences in patient characteristics, treatments, outcome definitions, and follow-up times. This is not a failure of the analytical method itself, but rather a reflection of the underlying data quality and consistency. The problem lies not in the inability to perform the statistical calculations, but in the inability to draw valid clinical inferences from such disparate inputs.1

The open-label design of many PV trials is an obvious caveat, introducing potential bias in symptom reporting and investigator assessment. But even beyond blinding, the fundamental differences in what was measured and in whom it was measured proved insurmountable. The trial was not powered to detect differences in molecular response across these varied populations, and that gap matters for a disease where deeper remission is increasingly sought.1

> **CLINICAL IMPLICATIONS**
>
> The inability to conduct a robust network meta-analysis for polycythemia vera treatments leaves clinicians in a familiar bind: an abundance of individual trial data but a scarcity of direct comparative evidence. This means that decisions between ropeginterferon alfa-2b, peginterferon alfa-2a, ruxolitinib, and hydroxyurea will continue to be made on a case-by-case basis, heavily influenced by patient-specific factors, local availability, and individual physician experience rather than a clear, evidence-based hierarchy. The field needs to move beyond simply demonstrating efficacy against placebo or historical controls and embrace trial designs that allow for direct comparisons.

> For patients, this fragmentation of evidence translates into a lack of clear guidance on which therapy offers the best long-term outcomes for their specific disease profile. The inconsistent definitions of molecular response are particularly frustrating; if we cannot agree on what constitutes a meaningful molecular remission, how can we truly compare drugs that aim for it? This is not merely an academic point; it directly impacts patient expectations and the ability to track disease progression or response to therapy with precision.
> The pharmaceutical industry, in collaboration with academic investigators, must prioritize standardization in future PV trials. Harmonizing patient populations, complete hematologic response definitions, and especially molecular response thresholds would significantly enhance the utility of future research. Without this concerted effort, the evidence base will remain a collection of disparate findings, hindering optimal treatment selection and delaying the adoption of truly evidence-based guidelines for this chronic condition.
> The current situation highlights a broader issue in rare disease research: the tension between the practicalities of conducting trials in small populations and the scientific rigor required for robust comparative effectiveness. Until trial designs become more consistent, the promise of molecular response as a guiding principle in PV treatment will remain largely unfulfilled, leaving clinicians to navigate a complex therapeutic landscape (the market, the guidelines, the trial pipeline) with incomplete maps.

EDITORIAL & AI STANDARDS

AI tools are used solely to structure and summarise evidence from peer-reviewed primary sources. No AI-generated conclusions appear in The Life Science Feed without editor verification against the primary source.

REFERENCES

1. 1. Walden P, Hummel N, Kopiec A. Evaluating the feasibility of a network meta-analysis comparing treatment options in polycythemia vera. J Comp Eff Res 2026.

LICENCE & RIGHTS

MEDICAL DISCLAIMER

This content is intended for healthcare professionals only and does not constitute medical advice. Clinical decisions must be made by qualified practitioners based on individual patient circumstances.

FOLLOW THE LIFE SCIENCE FEED

Instagram @lifesciencefeed **LinkedIn** linkedin.com/company/thelifesciencefeed

thelifesciencefeed.com