Received: 29 June 2017 | Revised: 3 July 2017 | Accepted: 6 July 2017

DOI: 10.1002/ajh.24851

RESEARCH ARTICLE

WILEY AJH

A reappraisal of the benefit-risk profile of hydroxyurea in polycythemia vera: A propensity-matched study

Tiziano Barbui[1] | Alessandro Maria Vannucchi[2] | Guido Finazzi[3] | Maria Chiara Finazzi[3] | Arianna Masciulli[1] | Alessandra Carobbio[1] | Arianna Ghirardi[1] | Gianni Tognoni[1]

[1]F.R.O.M. Research Foundation, Papa Giovanni XXIII Hospital, Bergamo, Italy; [2]Department of Experimental and Clinical Medicine, University of Florence, CRIMM, Center of Research and Innovation of Myeloproliferative neoplasms, AOU Careggi, Florence, Italy; [3]Hematology and Bone Marrow Transplant Unit, Papa Giovanni XXIII Hospital, Bergamo, Italy

Correspondence
Tiziano Barbui, F.R.O.M. Research Foundation, Papa Giovanni XXIII Hospital, Piazza O.M.S., 1, 24127 Bergamo (BG), Italy.
Email: tbarbui@asst-pg23.it

Funding information
"Special Program Molecular Clinical Oncology 5x1000" to AGIMM (AIRC-Gruppo Italiano Malattie Mieloproliferative), project number #1005

Abstract

The use of hydroxyurea (HU) as first line therapy in polycythemia vera (PV) has been criticized because no solid demonstration that this drug prevents thrombosis or prolongs survival has been so far produced. Here we present the outcomes of a large cohort of patients with PV included in the European Collaborative Low-dose Aspirin (ECLAP) study. We selected 1,042 patients who, during the follow-up, had received only phlebotomy (PHL) or HU to maintain the hematocrit level < 45%. To assure comparability, we conducted a propensity score matching analysis. The two groups (PHL $n = 342$ and HU $n = 681$) were well balanced for the parameters included in the propensity score (overall balance: $\chi^2 = 2.44$, $P = 0.964$). Over a comparable period of follow-up (PHL = 29.9 vs. HU = 34.7 months), we documented an advantage of HU over PHL consistently significant with respect to the incidence of fatal/non-fatal cardiovascular (CV) events (5.8 vs. 3.0 per 100 person-years in PHL vs. HU group, $P = 0.002$) and myelofibrosis transformation that was only experienced by patients of PHL group. Evolution to acute leukemia was registered in three patients (two in PHL and one in HU group). The excess of mortality and total CV events in the PHL patients was restricted to the high-risk group, and, compared with HU cases, was significant higher in the PHL patients who failed to reach the hematocrit target < 0.45% ($P = 0.000$). In conclusion, this analysis provides reliable and qualified estimates of the therapeutic profile of HU and PHL treatments for future experimental studies and for the management of PV in clinical practice.

1 | INTRODUCTION

Current risk stratification in polycythemia vera (PV) is designed to estimate the likelihood of recurrent thrombosis. Accordingly, PV includes two risk categories: high-risk (age > 60 years or thrombosis history) and low-risk (absence of both risk factors). Hydroxyurea (HU) in association with phlebotomy (PHL), and low-dose aspirin are currently recommended as first line cytoreductive therapy for patients with high-risk disease. In contrast in younger patients without a history of thrombosis, only PHL and aspirin are indicated.[1,2]

It should be noted that these recommendations are based on expert consensus owing that still limited number of randomized clinical trials do exist. In the PVSG 01[3] randomized clinical trial, overall survival was favored by treatment with PHL compared to treatment with radiophosphorus and chlorambucil. In fact, these latter drugs were found to increase the rate of acute leukemia (AL). Conversely, in the PVSG protocol 08, including 51 patients with PV, treatment with HU, regarded as a non-mutagenic myelosuppressive agent, was found to be associated with a lower incidence of thrombosis compared with historical control patients managed with PHL only.[4] Even though the number of patients was low, results of the study also indicated that incidence of leukemia was not apparently increased in comparison with phlebotomized patients.

Since then, no further control or large observational studies were designed to confirm the beneficial role of HU in comparison with phlebotomy. Moreover, the use of HU as first line therapy in PV patients has been criticized because no solid demonstration that HU therapy prevents thrombosis or prolongs survival has been so far produced and there is still a concern that the drug may increase the risk of leukemic transformation, leading to suggest therapeutic PHL as first line therapy, irrespective of patient risk category.[5]

It is also remarkable that after the PVSG studies, only two randomized clinical trials included the reduction of vascular events as primary end-point. These were the ECLAP (European Collaborative Low-dose Aspirin)[6] and the CYTO-PV trial,[7] that demonstrated the efficacy and safety of low dose aspirin and of the HCT control at <45%, respectively, in the reduction of the burden of cardiovascular (CV) mortality and morbidity. Of note, these latter hard end-points were not included as primary end-points in the recent Response trials[8,9] which were designed to assess the superiority of Ruxolitinib in achieving a surrogate end-point, such as haematological response, in HU resistant/intolerant PV patients.

Therefore, many areas of uncertainty still remain in the field that deserve to be addressed for two complementary reasons: to identify the most appropriate settings for the management of PV patients and to better define the risk profile and prognosis for future trials. On the basis of these premises, we carried out an analysis of data of a large prospective cohort of PV patients included in the ECLAP project.[10] where all the outcome events were formally validated over a 2.8 years median period of follow-up.

2 | PATIENTS AND METHODS

2.1 | Patients

This study was approved by the Institutional Review Boards of ECLAP centers. To ensure a representative sample of the spectrum of the disease burden, all patients with new and old diagnoses of PV made according to the criteria established by the PVSG[3] were included with no exclusion criteria with respect to age, therapy, or duration of disease. Treatment strategies had to comply with the recommendation of maintaining the hematocrit value at <0.45 and the platelet count at $<400\times 10^9$/L. Clinical outcomes during the prospective follow-up were recorded at follow-up visits at 12, 24, 36, 48, and 60 months. Out of the 1,638 patients of the whole cohort included in the ECLAP study, 1,042 were selected according to pre-defined criteria for the analyses; that is, patients were included if, during the follow-up, had received only PHL or HU to maintain the HCT level <45% as per ECLAP protocol. Events were diagnosed as previously described.[10]

2.2 | Statistical analysis

Descriptive statistics were used for summarizing the baseline characteristics of the two groups of patients.

To assure comparability between PHL and HU groups, we conducted a propensity score (PS) matching analysis,[11] by forming matched sets of 1 PHL and up to two randomly sampled HU-treated subjects (1:2 matching) who shared a similar values of PS. The PS was estimated by regressing exposure to only PHL conditionally on the baseline covariates (age at enrolment, gender, years from PV diagnosis, prior thrombosis, aspirin use, active smoking and arterial hypertension), with a logistic model. Matching was done using the nearest neighbor method with replacement and with caliper of width equal to 0.2 of the pooled standard deviation of the logit of PS. The standardized difference (STD) was used to quantify differences in proportions between treatment groups: a good balance was assured by standardized difference <0.10.

Incidence rate of each outcome, in the two matched-groups, was calculated considering the observed number of events and the corresponding person-years (PY). The 95% confidence interval (CI) was estimated under the assumption that the number of events followed a Poisson distribution. Rates were stratified according to the treatment group and compared using a conditional Poisson regression. Cumulative incidence curves for death and CV events according to the treatments were evaluated by the Kaplan–Meier approach and the stratified log-rank test was used to compare the curves.

The effect of treatment with PHL and HU on the risk of the selected outcomes was estimated by a Cox proportional-hazard model stratified on the matched pairs, since the matched sample does not consist of independent observations.[12]

For all tested hypotheses, two-tailed P values <0.05 were considered to be significant. Analyses were performed using STATA software, release 13 (Strata Corp LP, College Station, TX).

3 | RESULTS

Baseline characteristics before and after propensity score matching are presented in Table 1. The matched set comprised two groups that were well balanced for the parameters used for the implementation of PS (overall balance test: $\chi^2 = 2.44$, $P = 0.964$). The very good balance between the resulting groups provides a solid background for their assessment in terms of outcomes.

Results of the analysis of main outcome events during the follow-up are presented in Table 2. Over a comparable (yet statistically different in favor of HU: median 30 months and 35 months, respectively, for PHL and HU) period of follow-up, overall mortality was lower in HU patients (rate of 0.3 vs. 0.1 per 100 PY for PHL vs. HU groups, respectively) and the superiority of HU over PHL was consistently significant with respect to the incidence of fatal/non-fatal CV events (rate of 5.8 vs. 3.0 per 100 PY in PHL and HU groups respectively, $P = 0.002$) and of hematologic transformation (rate of 1.1 vs. 0.1 per 100 PY in PHL and HU groups respectively, $P = 0.006$); in particular evolution to AL was documented in three patients (two in PHL group and one in HU-group) while transformation to myelofibrosis was experienced by 8 patients (2.3%) in PHL group and none in the HU group. Five CV deaths (0.7%) occurred in the group treated with HU and 17 (5%) among patients treated with only PHL ($P = 0.000$). By fitting a stratified Cox-proportional hazard model, we estimated that HU reduced by 50% the risk of fatal and non-fatal CV events (HR = 0.50, 95% CI 0.35–0.72) and of 67% the risk of overall death (HR = 0.33, 95% CI 0.20–0.57).

To explain these findings we calculated the proportion of patients who reached the hematocrit, leukocyte and platelet targets in the two groups during the follow-up. We found that the proportion of patients with the target hematocrit < 45% at 12 months was significantly lower in PHL group compared to HU treated patients (31% vs. 52%,

TABLE 1 Baseline characteristics in patients with polycythemia vera treated with phlebotomies (PHL) and hydroxyurea (HU) before and after 1:2 propensity-score (PS) matching

	Before PS-matching			After PS-matching		
	Total cohort ($n = 1,042$)			1:2 random-sample[a] matched cohort ($n = 1,023$)		
	PHL ($n = 342$)	HU ($n = 700$)	STD	PHL ($n = 342$)	HU ($n = 681$)	STD
Age at enrolment ≥ 60, *n(%)*	186 (54.4%)	532 (76.0%)	−0.47	186 (54.4%)	370 (54.3%)	0.00
Male, *n(%)*	238 (69.6%)	374 (53.4%)	0.34	238 (69.6%)	481 (70.6%)	−0.02
Years from diagnosis of PV to enrolment ≥ 5, *n(%)*	102 (29.8%)	260 (37.1%)	−0.16	102 (29.8%)	195 (28.6%)	0.03
Prior thrombosis, *n(%)*	115 (33.6%)	285 (40.7%)	−0.15	115 (33.6%)	222 (32.6%)	0.02
High risk, *n(%)*	221 (64.6%)	588 (84.0%)	−0.46	221 (64.6%)	440 (64.6%)	0.00
Active smoking, *n(%)*	67 (19.6%)	83 (11.9%)	0.21	67 (19.6%)	111 (16.3%)	0.09
Hypertension, *n(%)*	138 (40.4%)	286 (40.9%)	−0.01	138 (40.4%)	261 (38.3%)	0.04
Diabetes mellitus, *n(%)*	25 (7.3%)	52 (7.4%)	0.00	25 (7.3%)	41 (6.0%)	0.05
Aspirin use, *n(%)*	127 (37.1%)	309 (44.1%)	−0.14	127 (37.1%)	262 (38.5%)	−0.03

PHL: phlebotomies; HU: hydroxyurea; PV: polycythemia vera; STD: standardized difference.
[a]1 PHL patient: up to two randomly sampled HU patients in each matched subset (three patients treated with PHL had only one HU matched patient). Matching was done using the nearest neighbor method with replacement and with caliper of width equal to 0.2 of the pooled standard deviation of the logit of PS. STDs <0.1 indicate a good balance between treatment groups.

$P < 0.0001$) and a similar trend was observed along the entire observation period (Figure 1, panel A). Similarly, in the PHL group the proportion of patients who presented $<12 \times 10^9$/L leukocyte count at 12 months was lower than that of HU group (74% vs. 88%, $P < 0.0001$); this difference was maintained up to 36 months (Figure 1, panel B). On the contrary, the number of patients obtaining the target platelet count ($<400 \times 10^9$/L) (Figure 1C) was similar in the two groups.

We then evaluated the main outcome in the two groups of treatments after stratification of the patients by their risk category. Unfortunately, the low number of CV events ($n = 5$), hematologic transformation ($n = 1$), and overall death ($n = 8$) in the low risk patient population prevented any reliable comparison between treatment groups. Conversely, in patients with age over 60 years and/or prior history of thrombosis (high-risk), the rate of fatal and not fatal CV events (8.7 vs. 4.8 per 100 PY in PHL vs. HU group), hematologic transformations (1.5 vs. 0.1 per 100 PY in PHL vs. HU group) and overall mortality (0.5 vs. 0.1 per 100 PY in PHL vs. HU group), were all significantly lower in patients treated with HU in comparison with PHL (Table 3 in Supporting Information).

The overall picture of the comparative profiles of PHL and HU treatments is also well illustrated by the cumulative incidence curves with respect to overall mortality and fatal and non-fatal CV events between the two treatment groups as a whole (Figure 2, panel A,C) and after stratification based on risk category (Figure 2, panel B,D). The excess of mortality and total CV events in PHL patients, restricted to

TABLE 2 Main outcome events during follow-up

	1:2 random-sample[a] matched cohort ($n = 1,023$)		
	PHL ($n = 342$)	HU ($n = 681$)	*P*
Median total follow-up (IQR), months	29.9 (15.1, 41.0)	34.7 (24.1, 45.3)	0.001
Median treatment duration (IQR), months	25.8 (12.7, 37.3)	24.0 (12.0, 36.0)	0.696
Outcome during follow-up			
1. CV events (fatal/non-fatal)	45 (13.2%)	54 (7.9%)	0.006
IR/100 PY (95% CI)	5.76 (4.30–7.72)	3.01 (2.30–3.93)	0.002
2. Hematologic transformation	9 (2.6%)	1 (0.2%)	0.006
IR/100 PY (95% CI)	1.11 (0.58–2.14)	0.05 (0.01–0.38)	0.006
3. Overall death	31 (9.1%)	24 (3.5%)	0.000
IR/100 PY (95% CI)	0.32 (0.23–0.46)	0.11 (0.07–0.16)	0.000

PHL: phlebotomies; HU: hydroxyurea; CV: cardiovascular; PY: person-years.
[a]1 PHL patient: up to two randomly sampled HU patients in each matched subset (three patients treated with PHL had only one HU matched patient). Matching was done using the nearest neighbor method with replacement and with caliper of width equal to 0.2 of the pooled standard deviation of the logit of PS.

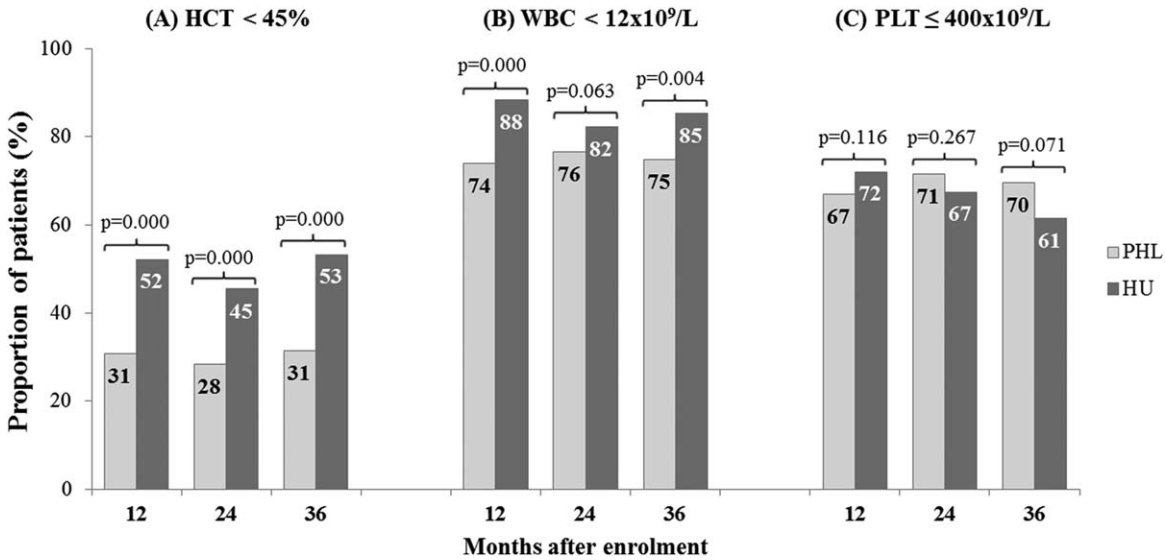

FIGURE 1 Proportion of patients who reached target values of hematocrit (HCT), leukocyte (WBC) and platelets (PLT) counts in the two matched-groups during follow-up. Legend: PHL: Phlebotomies; HU: Hydroxyurea; HCT: Hematocrit; WBC: Leukocyte. (A) Proportion of patients who reached target hematocrit $< 45\%$. (B) Proportion of patients who reached leukocyte counts $< 12 \times 10^9$/L. (C) Proportion of patients who reached platelet target $\leq 400 \times 10^9$/L

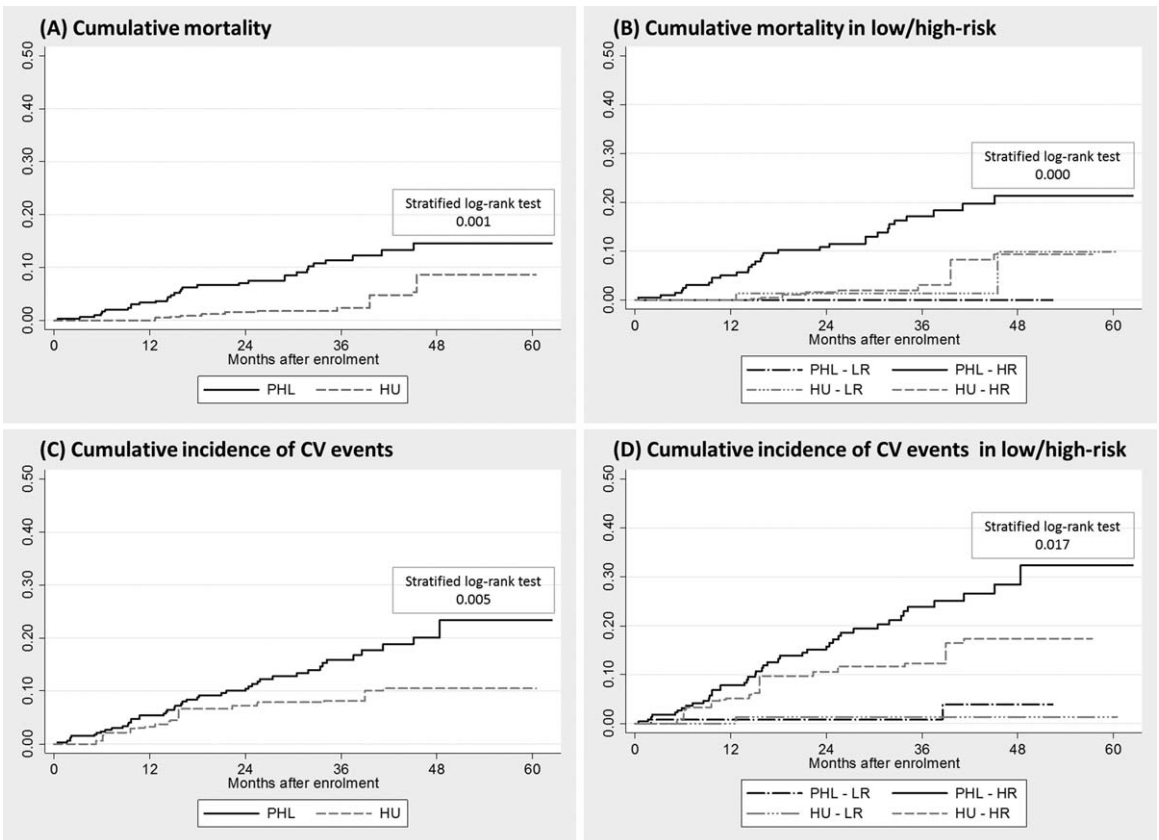

FIGURE 2 Cumulative incidence curves of total mortality and cardiovascular (CV) events among patients treated with PHL/HU at low/high-risk of thrombosis. Legend: PHL: Phlebotomies; HU: Hydroxyurea; LR: Low-risk patients; HR: High-risk patients. (A) Cumulative incidence of total mortality. (B) Cumulative incidence of total mortality among patients at low/high-risk of thrombosis. (C) Cumulative incidence of cardiovascular events. (D) Cumulative incidence of cardiovascular events among patients at low/high-risk of thrombosis

the high-risk group, displays the comprehensive difference in the incidence of these events in PHL and HU treated patients.

4 | DISCUSSION

The results of this study comparing the outcomes of PHL versus HU treatments in a large series of patients included in a prospective clinical trial were obtained by examining two cohorts well-balanced for age at enrolment, gender, years from PV diagnosis, prior history of thrombosis, aspirin use and cardiovascular risk factors. Therefore, the comparison between two treatment groups using the propensity score matching, should be interpreted as a semi-experimental prospective trial approach.

We have shown an efficacy and safety profile of HU strongly suggestive of a positive role in reducing overall mortality, mainly associated with CV events and inferior rate of myelofibrosis evolution, in comparison with PHL treated patients. According to the expectations, the advantages of HU therapy were confined to the population at high-risk, defined by age over 60 years and prior history of thrombosis.

This data confirms the results of the pivotal PVSG 01-trial[3] demonstrating that phlebotomy alone in elderly patients with prior history of thrombosis was the major factor contributing to excess future vascular events. This finding might be interpreted as being associated with a difference in terms of hematologic response in the two groups. In fact, the significant reduction of CV events by maintaining hematocrit $<45\%$ and leukocyte $<12 \times 10^9$/L was demonstrated by the CYTO-PV trial and in a post-hoc analysis of the same study.[7,13] Herein, we report that adequate control of hematocrit and leukocytosis was achieved in a larger proportion of patients treated with HU in comparison with those on PHL only, thus contributing to explain the significantly lower incidence of fatal and non-fatal thrombosis in patients receiving HU. Of note, the benefit of HU was also shown by lower incidence of transformation into myelofibrosis in patients receiving HU. Therefore, we reaffirm the validity of the current guidelines that recommend HU as first line therapy in this high risk category and highlight that in future comparative randomized clinical trials aimed at reducing CV complications in PV, this drug should be the comparative arm.

The low number of events in low-risk patients did not allow a reliable comparison between patients treated by PHL or HU. Regarding this subgroup of PV patients, it should be underscored that current recommended therapy (PHL and aspirin) has not yet been tested in a formal control prospective study. This represents an unmet clinical need as the rate of major arterial and venous thrombosis in younger patients in absence of history of CV events remains exceedingly higher (1.96% per year).[14] than that reported in the general population (0.7% per year).[15–17] It is very likely that the residual risk of vascular complications in patients treated with PHL and aspirin only might be reduced by a more strict control of hematocrit[7] and leukocytosis[18] through cytoreduction therapy. Conceivably, it might be foreseen that alpha-IFN, a non leukemogenic cytoreductive drug shown very active in controlling the disease and reducing the *JAK2* mutated allele burden,[19] might be suitably employed in this category of patients. A phase II randomized clinical trial is ongoing to address whether interferon can accomplish a better and durable control of hematocrit and leukocytosis than phlebotomy plus aspirin, and positively impact on the rate of thrombosis (ClinicalTrials.gov N° NCT03003325).

While the findings of current analysis of a cohort of the original ECLAP study, diagnosed by PVSG criteria and not treated by strictly risk-adapted criteria, are clearly exposed to intrinsic limitations and biases, the rigorous application of a 1:2 propensity-score matching is expected to reduce most of the underlying limitations.[11]

Therefore, we highlight the value of PHL in strictly defined low-risk patients, and confirm the efficacy of HU in high-risk cases, although a residual incidence of vascular events still exists in both groups; moreover, it is reassuring that HU reduced the rate of transformation into myelofibrosis without enhancing the risk of leukemia. While waiting the data of ongoing prospective randomized trials, this analysis provides estimates of the main outcomes observed in a large cohort in the context of a clinical trial and can be considered a reliable and qualified information of HU and PHL treatments for future experimental studies and for the management of PV in clinical practice.

ACKNOWLEDGMENTS

T.B., A.M.V., G.F. were supported by a grant from Associazione Italiana per la Ricerca sul Cancro (AIRC, Milano) "Special Program Molecular Clinical Oncology 5x1000" to AGIMM (AIRC-Gruppo Italiano Malattie Mieloproliferative), project number #1005. A detailed description of the AGIMM project is available at http://www.progettoagimm.it

AUTHORSHIP CONTRIBUTIONS

T.B., A.C., A.G. designed the research; T.B., A.M.V., G.F., M.C.F., A.M. contributed patients; A.C. and A.G. performed statistical analysis; T.B., A.M.V., G.F., A.C., A.G., G.T. participated in data analysis and interpretation, T.B. wrote the article. Everyone read and approved the final draft.

CONFLICT OF INTEREST

None of the authors have any conflict of interest to disclose in regards to the current manuscript.

REFERENCES

[1] Barbui T, Barosi G, Birgegard G, et al. Philadelphia-negative classical myeloproliferative neoplasms: critical concepts and management recommendations from European LeukemiaNet. *JCO*. 2011;29:761–770.

[2] Tefferi A, Barbui T. Polycythemia vera and essential thrombocythemia: 2017 update on diagnosis, risk-stratification, and management. *Am J Hematol*. 2017;92:94–108.

[3] Berk PD, Goldberg JD, Donovan PB, et al. Therapeutic recommendations in polycythemia vera based on Polycythemia Vera Study Group protocols. *Semin Hematol*. 1986;23:132–143.

[4] Kaplan ME, Mack K, Goldberg JD, Donovan PB, Berk PD, Wasserman LR. Long-term management of polycythemia vera with hydroxyurea: a progress report. *Semin Hematol*. 1986;23:167–171.

[5] Spivak JL. Myeloproliferative Neoplasms. *N Engl J Med*. 2017;376:2168–2181.

[6] Landolfi R, Marchioli R, Kutti J, et al. Efficacy and safety of low-dose aspirin in polycythemia vera. *N Engl J Med*. 2004;350:114–124.

[7] Marchioli R, Finazzi G, Specchia G, et al. Cardiovascular events and intensity of treatment in polycythemia vera. *N Engl J Med*. 2013; 368:22–33.

[8] Vannucchi AM, Kiladjian JJ, Griesshammer M, et al. Ruxolitinib versus standard therapy for the treatment of polycythemia vera. *N Engl J Med*. 2015;372:426–435.

[9] Passamonti F, Griesshammer M, Palandri F, et al. Ruxolitinib for the treatment of inadequately controlled polycythaemia vera without splenomegaly (RESPONSE-2): a randomised, open-label, phase 3b study. *Lancet Oncol*. 2017;18:88–99.

[10] Marchioli R, Finazzi G, Landolfi R, et al. Vascular and neoplastic risk in a large cohort of patients with polycythemia vera. *JCO*. 2005;23: 2224–2232.

[11] Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res*. 2011;46:399–424.

[12] Cummings P, McKnight B, Greenland S. Matched cohort methods for injury research. *Epidemiol Rev*. 2003;25:43–50.

[13] Barbui T, Carobbio A, Ghirardi A, Masciulli A, Rambaldi A, Vannucchi AM. No correlation of intensity of phlebotomy regimen with risk of thrombosis in polycythemia vera: evidence from European Collaboration on Low-Dose Aspirin in Polycythemia Vera and Cytoreductive Therapy in Polycythemia Vera clinical trials. *Haematologica*. 2017;102:e219–e221.

[14] Barbui T, Vannucchi AM, Carobbio A, et al. The effect of arterial hypertension on thrombosis in low-risk polycythemia vera. *Am J Hematol*. 2017;92:E5–E6.

[15] Antithrombotic Trialists' (ATT) Collaboration, Baigent C, Blackwell L, et al. Aspirin in the primary and secondary prevention of vascular disease: collaborative meta-analysis of individual participant data from randomised trials. *Lancet*. 2009;373:1849–1860.

[16] Yusuf S, Bosch J, Dagenais G, et al. Cholesterol lowering in intermediate-risk persons without cardiovascular disease. *N Engl J Med*. 2016;374:2021–2031.

[17] Risk and Prevention Study Collaborative Group, Roncaglioni MC, Tombesi M, et al. n-3 fatty acids in patients with multiple cardiovascular risk factors. *N Engl J Med*. 2013;368:1800–1808.

[18] Barbui T, Masciulli A, Marfisi MR, et al. White blood cell counts and thrombosis in polycythemia vera: a subanalysis of the CYTO-PV study. *Blood*. 2015;126:560–561.

[19] Kiladjian JJ, Giraudier S, Cassinat B. Interferon-alpha for the therapy of myeloproliferative neoplasms: targeting the malignant clone. *Leukemia*. 2016;30:776–781.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Barbui T, Vannucchi AM, Finazzi G, et al. A reappraisal of the benefit-risk profile of hydroxyurea in polycythemia vera: A propensity-matched study. *Am J Hematol*. 2017;92:1131–1136. https://doi.org/10.1002/ajh.24851