



Validation of the International Working Group 2023 response criteria for higher risk myelodysplastic syndromes/neoplasms using the International Consortium for Myelodysplastic Syndromes VALIDATE database and a randomized phase III trial

by Jan Philipp Bewersdorf, Tariq Kewan, Luca Lanino, Ondrej Blaha, Yiwen Liu, Jessica M. Stempel, Najla H. Al Ali, Amy E. Dezern, Mikkael A. Sekeres, Geoffrey L. Uy, Samuel Urrutia, Hetty E. Carraway, Pinkal Desai, Elizabeth A. Griffiths, Eytan M. Stein, Andrew M. Brunner, Christine McMahon, Rory M. Shallis, Joshua F. Zeidner, Michael R. Savona, Hayley Hawkins, Namrata Sonia Chandhok, Constantine Logothetis, Aram Bidikian, Ted M. Getz, Gail J. Roboz, Benjamin Rolles, Eunice S. Wang, Amyah C. Harris, Maria L. Amaya, Somedeb Ball, Justin Grenet, Zhuoer Xie, Yazan F Madanat, Yasmin Abaza, Talha Badar, Jaclynn Campos, Torsten Haferlach, Jaroslaw P. Maciejewski, David Sallman, Rahul S Vedula, H. Moses Murdock, Harrison K. Tsai, Evan C. Chen, Lionel Ades, R. Coleman Lindsley, Anoop Enjeti, Kamal Al-Rabi, Khalid Halahleh, Devendra Hiwase, Maria Diez-Campelo, David Valcarcel, Claudia Haferlach, Lisa Pleyer, Ioannis Kotsianidis, Vassiliki Pappa, Valeria Santini, Angela Consagra, Aref Al-Kali, Seishi Ogawa, Yasuhito Nannya, Matteo Giovanni Della Porta, Rami S. Komrokji, Maximilian Stahl and Amer M. Zeidan

Received: May 26, 2026.

Accepted: October 1, 2026.

Citation: Jan Philipp Bewersdorf, Tariq Kewan, Luca Lanino, Ondrej Blaha, Yiwen Liu, Jessica M. Stempel, Najla H. Al Ali, Amy E. Dezern, Mikkael A. Sekeres, Geoffrey L. Uy, Samuel Urrutia, Hetty E. Carraway, Pinkal Desai, Elizabeth A. Griffiths, Eytan M. Stein, Andrew M. Brunner, Christine McMahon, Rory M. Shallis, Joshua F. Zeidner, Michael R. Savona, Hayley Hawkins, Namrata Sonia Chandhok, Constantine Logothetis, Aram Bidikian, Ted M. Getz, Gail J. Roboz, Benjamin Rolles, Eunice S. Wang, Amyah C. Harris, Maria L. Amaya, Somedeb Ball, Justin Grenet, Zhuoer Xie, Yazan F Madanat, Yasmin Abaza, Talha Badar, Jaclynn Campos, Torsten Haferlach, Jaroslaw P. Maciejewski, David Sallman, Rahul S Vedula, H. Moses Murdock, Harrison K. Tsai, Evan C. Chen, Lionel Ades, R. Coleman Lindsley, Anoop Enjeti, Kamal Al-Rabi, Khalid Halahleh, Devendra Hiwase, Maria Diez-Campelo, David Valcarcel, Claudia Haferlach, Lisa Pleyer, Ioannis Kotsianidis, Vassiliki Pappa, Valeria Santini, Angela Consagra, Aref Al-Kali, Seishi Ogawa, Yasuhito Nannya, Matteo Giovanni Della Porta, Rami S. Komrokji, Maximilian Stahl and Amer M. Zeidan. Validation of the International Working Group 2023 response criteria for higher risk myelodysplastic syndromes/neoplasms using the International Consortium for Myelodysplastic Syndromes VALIDATE database and a randomized phase III trial. *Haematologica*. 2026 Oct 8. doi: 10.3324/haematol.2026.301314 [Epub ahead of print]

Publisher's Disclaimer.

E-publishing ahead of print is increasingly important for the rapid dissemination of science.

Haematologica is, therefore, E-publishing PDF files of an early version of manuscripts that have completed a regular peer review and have been accepted for publication.

E-publishing of this PDF file has been approved by the authors.

After having E-published Ahead of Print, manuscripts will then undergo technical and English editing, typesetting, proof correction and be presented for the authors' final approval, the final version of the manuscript will then appear in a regular issue of the journal.

All legal disclaimers that apply to the journal also pertain to this production process.

Validation of the International Working Group 2023 response criteria for higher risk myelodysplastic syndromes/neoplasms using the International Consortium for Myelodysplastic Syndromes VALIDATE database and a randomized phase III trial

Jan Philipp Bewersdorf^{1, #}, Tariq Kewan^{1, #}, Luca Lanino^{1, 2}, Ondrej Blaha³, Yiwen Liu⁴, Jessica M. Stempel¹, Najla H. Al Ali⁵, Amy E. DeZern⁶, Mikkael A. Sekeres⁷, Geoffrey L. Uy⁸, Samuel Urrutia⁸, Hetty E. Carraway⁹, Pinkal Desai¹⁰, Elizabeth A. Griffiths¹¹, Eytan M. Stein¹², Andrew M. Brunner¹³, Christine McMahon¹⁴, Rory M. Shallis⁵, Joshua F. Zeidner¹⁵, Michael R. Savona¹⁶, Hayley Hawkins¹⁶, Namrata Sonia Chandhok¹⁷, Constantine Logothetis¹⁸, Aram Bidikian¹, Ted M. Getz¹, Gail J. Roboz¹⁰, Benjamin Rolles^{1, 17}, Eunice S. Wang¹¹, Amyah C. Harris¹³, Maria L. Amaya¹⁴, Somedeb Ball¹⁶, Justin Grenet¹⁰, Zhuoer Xie⁵, Yazan F. Madanat¹⁸, Yasmin Abaza¹⁹, Talha Badar²⁰, Jaclynn Campos²⁰, Torsten Haferlach²¹, Jaroslaw P. Maciejewski⁹, David Sallman⁵, Rahul S. Vedula¹⁷, H. Moses Murdock¹⁷, Harrison K. Tsai¹⁷, Evan C. Chen¹⁷, Lionel Ades²², R. Coleman Lindsley¹⁷, Anoop Enjeti²³, Kamal Al-Rabi²⁴, Khalid Halahleh²⁴, Devendra Hiwase²⁵, Maria Diez-Campelo²⁶, David Valcarcel²⁷, Claudia Haferlach²¹, Lisa Pleyer²⁸, Ioannis Kotsianidis²⁹, Vassiliki Pappa³⁰, Valeria Santini³¹, Angela Consagra³¹, Aref Al-Kali³², Seishi Ogawa³³, Yasuhito Nannya³⁴, Matteo Giovanni Della Porta^{2, †}, Rami S. Komrokji^{5, †}, Maximilian Stahl^{1, 17, †}, and Amer M. Zeidan^{1, †, *}

1Department of Internal Medicine, Section of Hematology, Yale University, New Haven, CT, USA; 2Humanitas Clinical and Research Center, IRCCS, Rozzano, Italy; 3School of Public Health, Yale Center for Analytical Sciences, Yale University, New Haven, CT, USA; 4: Department of Data Science, Dana Farber Cancer Institute, Boston, MA; 5: Department of Malignant Hematology, H. Lee Moffitt Cancer Center, Tampa, FL, USA; 6Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University, Baltimore, MD, USA; 7Sylvester Comprehensive Cancer Center, University of Miami Miller School of Medicine, Miami, FL, USA; 8Washington University School of Medicine, Saint Louis, MO, USA; 9Department of Hematology and Medical Oncology, Taussig Cancer Institute, Leukemia Program, Cleveland Clinic, Cleveland, OH, USA; 10Weill Cornell Medicine, New York City, NY, USA; 11Roswell Park Comprehensive Cancer Center, Buffalo, NY, USA; 12Department of Medicine, Leukemia Service, Memorial Sloan Kettering Cancer Center, New York City, NY, USA; 13Division of Hematology/Oncology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA; 14University of Colorado, Aurora, CO, USA; 15University of North Carolina, Lineberger Comprehensive Cancer Center, Chapel Hill, NC, USA; 16Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, TN, USA; 17Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA; 18Harold C. Simmons Comprehensive Cancer Center, UT Southwestern Medical Center, Dallas, TX, USA; 19Northwestern University, Chicago, USA; 20Division of Hematology-Oncology, Mayo Clinic, Jacksonville, FL, USA; 21MLL Munich Leukemia Laboratory, Munich, Germany; 22 Hématologie Seniors, Assistance Publique - Hôpitaux de Paris (AP-HP), Université Paris Cité, Paris, France. 23Department of Haematology, Calvary Mater Hospital, Newcastle, New South Wales, Australia; 24King Hussein Cancer Center, Amman, Jordan; 25Department of Haematology, The Royal Adelaide Hospital, Adelaide, Australia; 26Hematology Department, Institute of Biomedical Research of Salamanca, University Hospital of Salamanca, Salamanca, Spain; 27Department of Hematology and Oncology, University Hospital Vall d'Hebron, Barcelona, Spain; 28: 3rd Medical Department with Hematology, Medical Oncology, Hemostaseology, Rheumatology and Infectiology, Oncologic Center, Paracelsus Medical University, Salzburg, Austria; 29Department of Hematology, Democritus University of Thrace Medical School, Alexandroupolis, Greece; 30Hematology Division, Second Department of Internal Medicine, Attikon University Hospital, National and Kapodistrian University of Athens, Athens, Greece; 31MDS Unit, Hematology, AOU Careggi - Department of Experimental and Clinical Medicine, University of Florence, Florence, Italy. 32Mayo Clinic, Rochester, MN, USA, 33DDepartment of Pathology and Tumor Biology, Institute for the Advanced Study of Human Biology (WPI-ASHBi), Kyoto University, Kyoto, Japan; Department of Innovative Medicine, Faculty of Medicine, Kindai University, Osaka, Japan. 34: Institute of Medical Science, the University of Tokyo, Tokyo, Japan

These first authors contributed equally

† These last authors contributed equally

* Corresponding author:

Amer M. Zeidan, MBBS, MHS

Division of Malignant Hematology; Section of Medical Oncology and Hematology

Department of Internal Medicine; Yale School of Medicine

333 Cedar Street

PO Box 208028

New Haven, CT 06520-8028

Email: amer.zeidan@yale.edu

Phone: 203-737-7103

Fax: 203-785-7232

Running title: Validation of IWG 2023 response criteria for HR-MDS

Acknowledgements: JPB is supported by a grant from the Edward P. Evans Foundation and the American Society of Hematology. BR received a Mildred-Scheel scholarship from the German Cancer Aid (“Deutsche Krebshilfe”) (grant no.: 70114570). VS is partially supported by Associazione Italiana per la Ricerca sul Cancro (AIRC) IG-26537-2021. MDP was supported by European Union - Horizon 2020/2023 program and Innovative Health initiative (GenoMed4All #101017549; Synthema #1101095530; Synthia #101172872; Realised #101165912) and AIRC (#29483 and #21267). SO is supported by the Japan Society for the Promotion of Science (JSPS); Grant-in-Aid for Specially Promoted Research, 24H00009.

Author Contributions: JPB, TK, MS, and AMZ conceptualized the project, provided patient data, analyzed the data and wrote the original manuscript. OB provided statistical and scientific input, conducted the primary data analysis of the VALIDATE dataset, and participated in the manuscript writing. YL provided statistical and scientific input and conducted the primary data analysis of the PANTHER dataset. All authors reviewed the manuscript, provided scientific input, and approved the final manuscript.

Funding: VALIDATE is supported by an independent grant from AbbVie. AbbVie had no role in the study design, data analyses, interpretation, or writing of the manuscript.

Competing Interest: The authors declare the following conflict of interests:

DeZern: *Sobi*: Consultancy; *Novartis*: Membership on an entity's Board of Directors or advisory committees; *Geron*: Membership on an entity's Board of Directors or advisory committees; *Caribou*: Membership on an entity's Board of Directors or advisory committees; *Bristol Myers Squibb*: Consultancy; *Appellis*: Consultancy, Membership on an entity's Board of Directors or advisory committees. *Celgene/BMS*, *Agios*, and *Novartis*: Advisory boards and/or consultancy; *Geron*, *Servier*, *Keros*, *Shuattuck labs* and *Celgene/BMS*: Clinical trial committees or DSMB.

Sekeres: *Geron*: Membership on an entity's Board of Directors or advisory committees;

Kurome: Consultancy, Current holder of *stock options* in a privately-held company; *Novartis*: Consultancy, Membership on an entity's Board of Directors or advisory committees; *BMS*: Consultancy, Membership on an entity's Board of Directors or advisory committees. **Uy**: *Jazz*: Other: Advisory Board; *Geron*: consultancy; *Syndax*: consultancy; *Autolus*: consultancy; *Virionexis*: consultancy; **Carraway**: *Agios*: Consultancy, Speakers Bureau; *BMS*: Consultancy, Research Funding, Speakers Bureau; *Jazz Pharmaceuticals*: Consultancy, Other: Travel, Accommodations, Expenses , Speakers Bureau; *Takeda*: Other; *Genentech*: Consultancy; *Daiichi*: Consultancy; *Celgene*: Research Funding; *Syndax*: Other: DSMB; *Astex Pharmaceuticals*: Other; *AbbVie*: Other; *Stemline Therapeutics*: Consultancy, Speakers Bureau; *Novartis*: Consultancy, Other: Travel, Accommodations, Expenses , Speakers Bureau. **Desai**: *Abbvie*: Consultancy, Other: Advisory role; *BMS*: Consultancy, Other: Advisory role, Research Funding; *Janssen Research & Development*: Research Funding; *Servier*: Consultancy, Other: Advisory role; *Janssen Pharmaceuticals*: Current Employment. *Syndax*: Consultancy; *Kura Oncology*: Consultancy. **Griffiths**: *Alexion Pharmaceuticals/AstraZeneca Rare Disease*: Consultancy, Research Funding; *Apellis Pharmaceuticals*: Consultancy, Research Funding; *Blueprint Medicines, Inc*: Research Funding; *Bristol- Myers Squibb*: Consultancy/Research Funding; *Celldex Therapeutics*: Research Funding; *Cogent Biosciences*: Consultancy; *Curio Science*: Honoraria; *Geron Inc.*: Consultancy; *MediCom Worldwide, Inc.*: Honoraria; *MedscapeLIVE*: Honoraria; *NextCure, Inc*: Research Funding; *Novartis*: Consultancy, Research Funding; *Pfizer Inc*: Consultancy; *Refined Oncology*: Consultancy; *Servier Pharmaceuticals*: Consultancy; *Taiho Oncology/Astex Pharmaceuticals*: Consultancy/Research Funding; *Vera and Joseph Dresner Foundation*: Membership on an entity's Board of Directors or advisory committees. *VJHemeOnc*: Honoraria. **Stein**: *Calithera*: Consultancy; *Daiichi*: Consultancy; *Foghorn*: Consultancy; *Servier*: Consultancy; *CTI Biopharma*: Consultancy; *Menarini*: Consultancy; *Genentech*: Consultancy; *Genesis*: Consultancy; *Syndax*: Consultancy; *OnCusp*: Consultancy; *Abbvie*: Consultancy; *Neoleukin*: Consultancy; *Jazz*: Consultancy; *Bristol Myers*

Squibb: Consultancy, Research Funding; *Agios*: Consultancy; *Novartis*: Consultancy; *Gilead*: Consultancy; *Eisai*: Research Funding; *Janssen*: Consultancy; *PinotBio*: Consultancy; *Aptose*: Consultancy; *Syros*: Consultancy; *Astellas*: Consultancy; *Ono Pharma*: Consultancy; *Blueprint*: Consultancy. **Brunner**: *Agios*: Consultancy, Research Funding; *AstraZeneca*: Research Funding; *Celgene/BMS*: Consultancy, Research Funding; *Gilead*: Consultancy; *Taiho*: Consultancy; *Novartis*: Consultancy, Research Funding; *Janssen*: Research Funding; *Keros Therapeutics*: Consultancy; *Takeda*: Consultancy; *Acceleron*: Consultancy; *GSK*: Research Funding. **McMahon**: *Syndax*: Advisory board; *Kura Oncology*: Safety monitoring committee. **Shallis**: *Bristol Myers Squibb*: Consultancy; *Cardinal Health*: Consultancy, Honoraria; *Curio Science*: Consultancy; *Geron*: Consultancy, Honoraria; *Gilead Sciences*: Consultancy, Honoraria; *Kura Oncology*: Consultancy; *Rigel*: Consultancy; *Servier*: Consultancy; *Syndax Pharmaceuticals*: Consultancy; *TScan Therapeutics*: Consultancy. **Zeidner**: *AbbVie*: Consultancy, Research Funding; *Akesobio*: Research funding; *Aptevo*: Research funding; *Arog*: Research Funding; *Ascentage*: Research funding; *Astellas*: Consultancy; *AstraZeneca*: Consultancy, Research Funding; *Auron*: Research Funding; *Crossbow Therapeutics*: Consultancy; *Daiichi Sankyo*: Consultancy, Research Funding; *Faron*: Research Funding; *Foghorn*: Consultancy; *Genentech*: Consultancy; *Genmab*: Consultancy; *Geron*: Consultancy; *Ipsen*: Consultancy; *Jazz*: Consultancy, Research Funding; *Johnson & Johnson*: Consultancy; *Kura Oncology*: Consultancy, Research Funding; *Lin Biosciences*: Research Funding; *Loxo*: Research Funding; *Merck*: Research Funding; *Novartis*: Consultancy, Research Funding; *Relmada Therapeutics*: Consultancy; *Rigel*: Consultancy; *Sellas*: Consultancy, Research Funding; *Servier*: Consultancy; *Shattuck Labs*: Consultancy, Research Funding; *Stemline*: Research Funding; *Sumitomo Pharma*: Consultancy, Research Funding; *Syndax*: Consultancy; *Zentalis*: Research Funding. **Savona**: *Karyopharm Therapeutics Inc.*: Consultancy, Current equity holder in publicly-traded company; *Empath Biosciences*: Current equity holder in publicly-traded company; *Incyte Corporation*: Research Funding; *Sierra Oncology, Inc.*: Consultancy;

Taiho: Research funding and travel costs; *ALX Oncology*: Research Funding; *AbbVie Inc.*: Consultancy; *Boehringer Ingelheim*: Patents & Royalties; *TG Therapeutics, Inc.*: Research Funding; *Takeda Pharmaceutical Company*: Consultancy, Research Funding; *Ryvü Therapeutics*: Consultancy, Current equity holder in publicly-traded company; *Novartis*: Consultancy; *Geron Corporation*: Consultancy; *Forma Therapeutics Inc.*: Consultancy, Consultancy; *Bristol Myers Squibb*: Membership on an entity's Board of Directors or advisory committees; *CTI BioPharma Corp.*: Consultancy. **Roboz**: *Agios*: Consultancy; *Actinium*: Consultancy; *Amgen*: Consultancy; *Astellas*: Consultancy; *AstraZeneca*: Consultancy; *Bristol Myers Squibb*: Consultancy; *Celgene*: Consultancy; *Caribou Biosciences*: Consultancy; *Daiichi Sankyo*: Consultancy; *Ellipses Pharma*: Consultancy; *Geron*: Consultancy; *Glycomimetics*: Consultancy; *Jazz Pharmaceuticals*: Consultancy; *MEI*: Consultancy; *Mesoblast*: Consultancy; *Molecular Partners*: Consultancy; *Morphosys*: Consultancy; *Oncoprecision*: Consultancy; *Oncoverity*: Consultancy; *AbbVie*: Consultancy; *Novartis*: Consultancy; *Pfizer*: Consultancy; *Rigel*: Consultancy; *Roche*: Consultancy; *Syndax*: Consultancy; *Blueprint*: Consultancy; *Jasper Pharmaceuticals*: Consultancy; *Bluebird bio*: Consultancy; *GlaxoSmithKline*: Consultancy; *Janssen*: Consultancy and research support. Funding; *Takeda*: Consultancy. **Wang**: *Gilead*: Consultancy; *GlaxoSmithKline*: Consultancy; *Jazz*: Consultancy; *Kite*: Consultancy, Speakers Bureau; *Novartis*: Consultancy, Speakers Bureau; *Pfizer*: Consultancy, Speakers Bureau; *PharmaEssentia*: Consultancy; *Takeda*: Consultancy; *Dava oncology*: Speakers Bureau; *Kura Oncology*: Speakers Bureau; *BMS*: Consultancy; *Astellas*: Consultancy, Speakers Bureau; *Abbvie*: Consultancy. **Xie**: *Novartis*: Speakers Bureau; *Moffitt Cancer Center*: Current Employment.. **Maciejewski**: *Regeneron*: Consultancy, Honoraria; *Alexion*: Membership on an entity's Board of Directors or advisory committees; *Novartis*: Honoraria, Speakers Bureau; *Omeros*: Consultancy. **Sallman**: *Aprisa*, *Jazz*: Research Funding; *AbbVie*, *Affimed GmbH*, *Gilead*, *Incyte*, *Intellisphere, LLC*, *Molecular Partners AG*, *PGEN Therapeutics, Inc.*, *Takeda*, *Zentalis*; Advisory board for *AvenCell*, *BlueBird Bio*, *BMS*, *Intellia*, *Jasper Therapeutics*, *Kite*,

Magenta Therapeutics, NKARTA, Novartis, Orbita: Consultancy. **Al-Kali**: *Novartis*: Research funding to institution, consultancy; *Astex*: Research funding to institution; *ALX Oncology*: Research funding to institution; *BMS*: Research funding to institution; *Apra*: Research funding to institution; *H3B/Hemavant*: Research funding to institution, *Celgene*: Research funding to institution. **Enjeti**: *Otsuka*: Consultancy, honoraria, research funding to institution; *Servier*: Consultancy; *Jazz*: Consultancy; *RaceOncology*: Consultancy, honoraria; *Amgen*: Consultancy; *AbbVie*: Consultancy. **Madanat**: Honoraria/consulting fees from *BMS, Kura Oncology, Blueprint Medicines, Geron, OncLive and MD Education, VJHemOnc and Medscape Live*. Advisory boards and received honoraria from *Stemline Therapeutics, Blueprint Medicines, Taiho Oncology, SOBI, Rigel Pharmaceuticals, Geron, Cogent Biosciences*. Received travel reimbursement from *MD Education*. **Abaza**: Consultancy, Honoraria: *Astellas, Bristol Myers Squibb, Daiichi Sankyo, Geron, KiTE, Pfizer, Rigel, and Servier*. Research funding to institution: *AbbVie, ALX Oncology, Biomea, Biosight, Curis, and Novartis*. **Badar**: Honoraria from Amgen, Ascentage, Syndax, Geron, Takeda. Research funding from Takeda, Amgen. **Haferlach T**: MLL Munich Leukemia Laboratory: equity ownership **Kotsianidis**: *Abbvie*: Consultancy, Honoraria; *Janssen*: Consultancy, Honoraria, Research Funding; *BMS*: Consultancy, Honoraria; *Novartis*: Consultancy, Research Funding. **Pappa**: Research Support /PI : *Celgene, Janssen, Abvie*; Honoraria/Advisory Board: *Genesis Pharma, Janssen, Roche, Servier, Abvie, Novartis, Gilead, BMS, Astra Zeneca, GSK, Amgen, Sobi, Pfizer*. **Haferlach C**: MLL Munich Leukemia Laboratory: equity ownership. **Stahl**: *Kymera*: Membership on an entity's Board of Directors or advisory committees; *Rigel*: Membership on an entity's Board of Directors or advisory committees; *Curis Oncology*: Other: GME activity ; *Clinical care options*: Other: GME activity ; *GSK*: Membership on an entity's Board of Directors or advisory committees; *Novartis*: Membership on an entity's Board of Directors or advisory committees, Other: GME activity ; *Haymarket Media*: Other: GME activity ; *Boston Consulting*: Consultancy; *Sierra Oncology*: Membership on an entity's Board of Directors or advisory committees; *Dedham group*:

Consultancy. **Della Porta:** *Bristol Myers Squibb*: Honoraria, Membership on an entity's Board of Directors or advisory committees. **Komrokji:** *Geron*: Consultancy; *BMS*: Honoraria, Membership on an entity's Board of Directors or advisory committees, Research Funding; *AbbVie*, *CTI biopharma*, *Jazz*, *Pharma Essentia*, *Servio*: Honoraria, Membership on an entity's Board of Directors or advisory committees, Speakers Bureau; *Rigel*, *Taiho*, *DSI*: Honoraria, Membership on an entity's Board of Directors or advisory committees; *Novartis*: Membership on an entity's Board of Directors or advisory committees. **Zeidan:** *Incyte*: Consultancy, Honoraria; *Agios*: Consultancy, Honoraria; *Boehringer-Ingelheim*: Consultancy, Honoraria; *Schrödinger*: Consultancy, Honoraria; *Celgene/BMS*: Consultancy, Honoraria; *Astellas*: Consultancy, Honoraria; *Syros*: Consultancy, Honoraria; *Jazz*: Consultancy, Honoraria; *Regeneron*: Consultancy, Honoraria; *Novartis*: Consultancy, Honoraria; *Servier*: Consultancy, Honoraria; *Zentalis*: Consultancy, Honoraria; *Tyme*: Consultancy, Honoraria; *Astex*: Research Funding; *Shattuck Labs*: Research Funding; *AbbVie*: Consultancy, Honoraria; *Lox Oncology*: Consultancy, Honoraria; *Pfizer*: Consultancy, Honoraria; *Mendus*: Consultancy, Honoraria; *Foran*: Consultancy, Research Funding; *Otsuka*: Consultancy, Honoraria; *BioCryst*: Consultancy, Honoraria; *ALX Oncology*: Consultancy, Honoraria; *BeyondSpring*: Consultancy, Honoraria; *Chiesi*: Consultancy, Honoraria; *Takeda*: Consultancy, Honoraria; *Kura*: Consultancy, Honoraria; *Ionis*: Consultancy, Honoraria; *Gilead*: Consultancy, Honoraria; *Syndax*: Consultancy, Honoraria; *Amgen*: Consultancy, Honoraria; *Epizyme*: Consultancy, Honoraria; *Janssen*: Consultancy, Honoraria; *Genentech*: Consultancy, Honoraria; *Seattle Genetics*: Consultancy, Honoraria; *Notable*: Consultancy, Honoraria; *Taiho*: Consultancy, Honoraria; *Geron*: Consultancy, Honoraria; *Orum*: Consultancy, Honoraria; *Daiichi Sankyo*: Consultancy, Honoraria, interests.

Data Availability Statement: Deidentified or aggregated study data will be made available upon reasonable request from qualified researchers to the extent permitted by applicable laws and data use agreements signed with study sites. Approval of such requests is at the discretion of the VALIDATE study team and is dependent on the nature of the request, the merit of the research proposed, the availability of the data, and the intended use of the data.

Abstract

Response assessment criteria in patients with higher-risk myelodysplastic syndromes/neoplasms (HR-MDS) have been updated in the 2023 International Working Group (IWG) criteria. However, validation of these criteria has been limited to date.

We analyzed overall survival (OS) by best response according to IWG 2023 criteria in two separate and independent cohorts of patients with HR-MDS who were treated with hypomethylating agents (HMA)-based therapies. The first dataset was the International Consortium for MDS (icMDS) VALIDATE retrospective database, and the second dataset was the cohort of HR-MDS patients treated in the randomized, phase III PANTHER trial of azacitidine + pevonedistat or placebo.

A total of 1,268 patients (VALIDATE: n=1,047; PANTHER: n=221) were evaluated in two separate sets of analyses according to the source dataset. Patients who achieved composite CR (cCR) had improved OS compared to patients not achieving cCR in both datasets and in both univariate and multivariable analyses (VALIDATE: multivariable Hazard Ratio [HR]: 0.65; 95% CI: 0.55-0.76; p<0.001; PANTHER: multivariable HR: 0.69; 95% CI: 0.48-0.99; p=0.046). These findings were also supported by time-dependent and sensitivity analyses. Further, peripheral blood CR (PB-CR) was also associated with improved OS. Hence, the IWG 2023 response criteria can reliably predict meaningful clinical benefit in HR-MDS patients and should replace the IWG 2006 criteria in both clinical practice and clinical trials.

Introduction:

Myelodysplastic syndromes/neoplasms (MDS) are diverse myeloid neoplasms that primarily affect older adults and have a variable prognosis based on cytogenetic, molecular, and hematologic parameters.¹⁻³ Although allogeneic stem cell transplant (allo-SCT) is a potentially curative therapeutic modality for patients with higher-risk MDS (HR-MDS), the majority of older patients are ineligible for this intensive treatment and disease relapse is still a frequent complication.⁴⁻⁶ Outside of allo-SCT, the standard of care for HR-MDS remains monotherapy with the hypomethylating agents (HMA) azacitidine (AZA) and decitabine (DEC), which can prolong overall survival (OS) and delay leukemic progression, but generally does not result in a cure.^{7,8}

Response assessment in MDS in both clinical trials and routine practice has been historically based on the 2006 International Working Group (IWG) response criteria.⁹ While complete remission (CR) has remained the gold standard for treatment response in MDS for decades, focusing solely on CR rate may lead to underestimation of meaningful clinical benefits of a therapy, especially among patients treated with HMA. This is exemplified by the substantial OS benefit of AZA in patients with HR-MDS in randomized clinical trials despite CR rates of generally less than 20-30%.^{7,10} Additional clinical benefits have been documented with AZA that are not captured by the CR rate alone.¹¹

Several limitations related to the variable timing of response assessment across clinical trials, the continuous administration of myelosuppressive therapy, and the arbitrary definition of response thresholds have been recognized leading to the updated IWG response criteria by an international expert panel in 2023.¹²⁻¹⁴ Among the major

changes in the IWG 2023 response criteria was the introduction of several “near-CR” endpoints such as CR with limited hematologic recovery (CR-L) and CR with partial hematologic recovery (CRh) as provisional response categories, as well as the removal of marrow CR (mCR) and stable disease (SD) as response categories.¹⁴ However, attempts at validation of these novel response criteria have been limited to small datasets lacking the necessary statistical power.^{15,16}

Another open question in the field of response assessment in MDS remains whether an invasive bone marrow evaluation for blast count enumeration is necessary for response assessment or if response assessment based on peripheral blood parameters alone is sufficient as recently proposed in an analysis from the Austrian Myeloid Registry.¹⁷ This is especially relevant as patients outside of clinical trials may not undergo bone marrow assessments for response monitoring.

To address these questions and validate the IWG 2023 MDS response criteria, we used patients with HR-MDS who received HMA-based therapies in the frontline setting from two large, international, independent datasets: the International Consortium for MDS (icMDS) VALIDATE retrospective database and the cohort of HR-MDS patients treated in the randomized, phase III PANTHER trial, which compared azacitidine (AZA) + pevonedistat or placebo in patients with HR-MDS, chronic myelomonocytic leukemia (CMML), and oligoblastic acute myeloid leukemia (AML).¹⁸

Methods:

Data sources:

VALIDATE is a retrospective dataset of the icMDS resulting from a collaboration of 31 centers and cooperative groups in 9 countries (**Supplemental Table 1**) and contains patients with MDS and AML who received frontline HMA-based therapies. Patients were entered by participating sites with central validation of baseline risk stratification, molecular disease characteristics, and central response assessments by two investigators independently.

As a second independent validation cohort, we used data for patients with HR-MDS who received treatment with AZA +/- pevonedistat in the randomized, phase III PANTHER trial.¹⁸

Patient selection:

For the VALIDATE cohort, we included adults (≥ 18 years of age) with HR-MDS who received frontline treatment with HMA either as monotherapy or as part of a combination therapy. HR-MDS was defined as either International Prognostic Scoring System (IPSS) score of ≥ 1.5 or an IPSS-revised (IPSS-R) of ≥ 4 at time of diagnosis.^{3,19}

For the PANTHER trial cohort, we included only patients with HR-MDS and excluded patients with CMML or oligoblastic AML. HR-MDS was defined as IPSS-R very high-, high-, or intermediate-risk disease with intermediate-risk disease patients requiring $\geq 5\%$

bone marrow myeloblasts for enrollment.¹⁸ **Supplemental Figure 1** shows a CONSORT diagram of patient selection from VALIDATE and PANTHER trial for our analysis.

Response assessments:

Responses were assessed centrally by two investigators independently and classified according to IWG 2006 and IWG 2023 as well as by peripheral blood – CR (PB-CR) criteria.^{9,14,17}

To evaluate the effect of immortal time bias, we performed time-dependent and modified 6-month landmark analyses. We also conducted sensitivity analyses limited to patients treated with HMA monotherapy to confirm that IWG 2023 response criteria are valid in the context of HMA monotherapy.

Composite CR (cCR) according to IWG 2006 criteria was defined as CR or marrow CR (mCR). cCR per the IWG 2023 criteria is a combination of CR, CR-L, CRh, and CR equivalent (CRequ). CR-L was defined as a composite of CR with bilineage count recovery (CRbi) and CR with unilineage count recovery (CRuni). PB-CR was defined as achieving hemoglobin ≥ 11 g/dL, platelets $\geq 100 \times 10^9$ /L, absolute neutrophil count $\geq 1.0 \times 10^9$ /L, and absence of peripheral blood blasts.

Statistical analysis:

Each dataset was analyzed separately. Baseline comparisons were conducted using chi-square test of proportions with continuity correction in VALIDATE and Fisher's exact

test in PANTHER. Continuous variables were compared using the Kruskal-Wallis rank sum test with three groups or Wilcoxon rank sum test with two groups.

All survival and time-to-event analyses were conducted using Kaplan-Meier estimator, proportional hazards Cox models, and marginal method for estimating adjusted survival curves. A modified 6-month landmark cohort was defined including patients alive at 6 months after treatment initiation (VALIDATE) or randomization (PANTHER). OS was calculated from the landmark time. Patients were categorized according to their overall best response during treatment or follow-up rather than whether the best response had been achieved by the landmark time, and survival was compared using Kaplan–Meier methods. OS was also estimated using the Mantel–Byar method with IWG 2023 response and allo-SCT as time-dependent variables and visualized in the Simon–Makuch plot. Cox models were constructed to adjust for baseline characteristics, initial therapy assignment as well as allo-SCT treatment status and IWG 2023 response as time-dependent variables. Additional details on the statistical analysis plan are provided in the **Supplemental Methods**.

The VALIDATE protocol was exempted by Institutional Review Board (IRB) of the coordinating Center (Yale University) due to minimal risk and need for informed consent from patients was waived. The protocol was exempted, acknowledged, or approved by the other centers according to their local policies. VALIDATE is supported by an independent grant from AbbVie. A de-identified dataset of The PANTHER trial was kindly shared by the trial sponsor, Takeda, with Dana Farber Cancer Institute (DFCI) where the analysis of the dataset occurred. Abbvie and Takeda had no role in the study design, data analyses, interpretation, or writing of the manuscript.

Results:

Patient baseline and treatment characteristics in VALIDATE and PANTHER:

We included a total of 1,047 patients with HR-MDS treated with HMA-based therapy in the VALIDATE database. The median age at diagnosis was 69 years (IQR: 62-75) with 64.4% of patients being male and 84.4% being treated at a study site in the United States. The VALIDATE cohort was enriched for high-risk features including complex karyotype (n=442 patients [42.3%]), *TP53* mutations (n=268 patients [33.0%]) and higher risk disease by both IPSS-R (high: n=357 [34.3%] or very high: n=430 [41.3%]) and IPSS-M (moderate high n=148 patients [14.1%], high n=336 [32.0%], and very high n=451 patients [43.0%]).

Most patients received treatment with AZA monotherapy (n=637 patients [60.8%]) followed by HMA-based combinations (n=234 [22.4%]) including HMA in combination with venetoclax (n = 98 [9.4%]) and decitabine monotherapy (n=176 [16.8%]). The median time from diagnosis to HMA initiation was 33 days (IQR: 18-66 days) and median time on treatment was 6 cycles (IQR: 4-10 cycles). 385 patients (36.8%) proceeded with allo-SCT at a median of 5.5 months from the start of HMA therapy. Additional patient, disease, and treatment characteristics for the entire cohort and by type of HMA treatment are reported in **Table 1**.

In the PANTHER trial dataset, a total of 221 patients with HR-MDS had sufficient data available to adjudicate IWG 2023 responses and were included in this analysis (n=113 treated with AZA monotherapy; n=108 treated with AZA + pevonedistat). Median age was 73 years (IQR: 67- 77 years) and 59% of patients were male. The IPSS-R at

baseline was intermediate, high, and very high in 25%, 39%, and 36% of patients respectively. Mutations in *TP53* were present in 23% of patients. Of this cohort, only 5.9% of patients received a subsequent allo-SCT. A full list of baseline patient, disease, and treatment characteristics for the PANTHER cohort is presented in **Table 2**.

Treatment response assessment according to IWG 2006, IWG 2023, and PB-CR criteria in VALIDATE and PANTHER:

Among the response evaluable VALIDATE cohort, rates of cCR per IWG 2006 and IWG 2023 criteria were 49.5% and 36.2%, respectively. Best responses per IWG 2023 criteria included CR (15.3%), CRequ (1.3%), CRbi (9.2%), CRuni (9.8%), CRh (0.6%), PR (1%), hematologic improvement (HI; 14.2%), no response (28%), and PD (19.0%).

In the PANTHER trial, rates of cCR per IWG 2006 and IWG 2023 criteria were 59.8% and 49.8%, respectively. Best responses per IWG 2023 criteria included CR (23.1%), CRbi (10.0%), CRuni (14.5%), CRh (1.8%), PR (0.5%), no response (43.9%), and PD (6.3%).

Most patients with mCR per IWG 2006 criteria were reclassified as either CRbi or CRuni especially in the context of concurrent HI, while patients with mCR but without HI were most frequently reclassified as “No Response”. Although the definition of CR per IWG 2023 criteria included changes from the IWG 2006 criteria, most patients remained in the CR category when classified per IWG 2023 criteria (VALIDATE cohort: 90.0%; PANTHER cohort: 79.4%). **Supplemental Tables 2 and 3** show the bivariate distribution of response categories according to IWG 2006 and IWG 2023 criteria for the

VALIDATE and PANTHER cohort, respectively. Alluvial plots illustrating the changes in response categories between IWG 2006 and IWG 2023 for VALIDATE and PANTHER cohorts are shown in **Supplemental Figures 2 and 3**, respectively.

Survival outcomes based on best response according to IWG 2023 criteria in VALIDATE and PANTHER:

The VALIDATE cohort had a median follow up time of 16.7 months. The median OS among all response evaluable patients was 19.3 months (95% CI: 17.6 - 21.3 months) with 1-year and 2-year OS rates of 68.8% (95% CI: 66.0 - 71.7%) and 41.5% (95% CI: 38.4 - 44.7%), respectively. In the PANTHER trial, the median OS among all response evaluable patients was 20.2 (95% CI: 17.7 - 22.5) months with 1-year and 2-year OS rates of 72.8% (95% CI: 67.1 - 79.1%) and 37.6% (95% CI: 31.4 - 44.9%).

Unadjusted associations between the IWG 2023 response criteria and OS are provided in **Table 3**. Patients who achieved CR (n=160 patients; 15.3%) had a median OS of 31.3 months (95% CI: 25.5 - 41.8 months). Compared to patients with CR, median OS was similar for patients with CRequ (HR: 1.77; 95% CI: 0.97 - 3.22; p=0.064), CRbi (HR: 1.14; 95% CI: 0.83 - 1.58; p=0.425), CRuni (HR: 1.32; 95% CI: 0.96 - 1.80; p=0.086) and CRh (HR: 0.79; 95% CI: 0.20 - 3.21; p=0.742). Conversely, patients with PR (HR: 2.53; 95% CI: 1.28 - 5.03; p=0.008), HI (HR: 1.68; 95% CI: 1.27 - 2.20; p<0.001), No Response (HR: 1.45; 95% CI: 1.14 - 1.86; p=0.003), PD (HR: 2.15; 95% CI: 1.67 - 2.78; p<0.001), or who received <2 cycles (HR: 3.73; 95% CI: 2.30 - 6.04; p<0.001) had significantly worse OS compared to patients with CR. OS stratified by IWG 2023

response category after adjusting for patients' age, sex, WBC at diagnosis, type of HMA treatment (AZA monotherapy, DEC monotherapy, or HMA combinations), IPSS-M category (Very Low/Low/Moderate Low, Moderate High/High, or Very High) and allo-SCT status is shown in **Figure 1**.

Similar level of association between response per IWG 2023 criteria and OS persisted when follow-up was censored at the time of transplant and for patients who did not proceed to allo-SCT (**Supplemental Table 4**). Conversely, among patients who proceeded to allo-SCT, achieving CR pre-transplant did not impact the risk of all-cause mortality when compared to patients who proceeded to allo-SCT and achieved No Response (HR: 1.11; 95% CI: 0.72 - 1.70; p=0.649) or PD (HR: 1.20; 95% CI: 0.73 – 1.99; p=0.476) as their best response to HMA-based frontline treatment.

Achievement of cCR per IWG 2023 is associated with improved OS in VALIDATE and PANTHER:

We next analyzed whether achieving cCR per IWG 2023 criteria is associated with improved OS. When treating response as time-dependent variable, patients in both VALIDATE and PANTHER who achieved cCR as their best response had improved OS compared to patients who did not achieve cCR (VALIDATE: median survival time (MST) for cCR 23.8 months [95% CI: 21.6 - 27.5 months] vs. 16.9 months [95% CI: 15.7 - 18.2 months]; p<0.001; PANTHER: 21.5 months [95% CI: 18.9 – 25.3 months] vs. 18.3 months [95% CI: 14.8 – 22.2 months]; p=0.041; **Figure 2A and B**).

In a modified 6-month landmark analysis we again found that achieving cCR per IWG 2023 was associated with improved OS compared to no cCR (VALIDATE: MST for cCR 20.4 months [95% CI: 17.0 - 23.9 months] vs. 14.6 months [95% CI: 12.5 - 16.1 months]; $p<0.001$; PANTHER: 17.8 months [95% CI: 14.8 months - not reached] vs. 13.9 months [95% CI: 9.3 - 16.9 months]; $p=0.032$; **Supplemental Figure 4A and B**).

Since HMA monotherapy remains the standard of care for patients with HR-MDS and a subset of patients in both cohorts received treatment with HMA-based combination therapies, we also evaluated the impact of cCR on OS among patients treated with HMA monotherapy to minimize the risk of confounding. Supporting the robust validity of cCR as a predictor of OS, patients who achieved cCR in both study cohorts had improved OS compared to patients who did not achieve cCR (VALIDATE: MST for cCR 23.1 [95% CI: 20.8 - 27.3 months] vs. 17.4 months [95% CI: 15.7 - 19.3 months]; $p<0.001$; PANTHER: 20.8 months [95% CI: 18.5 - 27.6 months] vs. 14.8 months [95% CI 12.2 - 17.7 months]; $p<0.001$; **Figure 2C and D**).

OS benefit is driven by patients achieving CR and – to a lesser extent – CR-L

After having demonstrated that cCR is associated with improved OS, we next sought to elucidate whether this OS benefit is limited to a certain response subcategory within the cCR compound category. In the VALIDATE cohort, achieving both CR and CR-L were associated with improved OS compared to patients who did not achieve either (MST for CR 26.8 months [95% CI: 23.8 - 31.3 months] vs. 22.1 months [95% CI: 19.3 - 25.5 months] for CR-L vs. 16.9 months [95% CI: 15.7 - 18.2 months] for no-CR/CR-L;

$p < 0.001$; **Figure 3A**). There was no statistically significant difference for OS between patients with CR vs. CR-L ($p = 0.11$).

While patients included in the PANTHER dataset who achieved a CR had statistically different OS compared to patients without a CR/CR-L (MST for CR 24.3 months [95% CI: 20.6 months - not reached] vs. no-CR/CR-L 15.1 months [95% CI: 13.1 - 20.1 months]; $p = 0.002$), there was only a numeric, but not statistically significant OS benefit for patients who achieved a CR-L compared to patients without a CR/CR-L (MST for CR-L 21.3 months [95% CI: 15.6 - 27.6 months] vs. no-CR/CR-L 15.1 months [95% CI: 13.1 - 20.1 months]; $p = 0.11$; **Figure 3B**). OS for patients with CR compared to CR-L was similar supporting the inclusion of both categories in the cCR composite endpoint ($p = 0.11$).

In sensitivity analyses, using a modified 6-month landmark approach and restricting the analysis to patients treated with HMA monotherapy we found similar results (**Supplemental Figure 6**).

PB-CR is associated with improved OS in HR-MDS:

We next analyzed whether PB-CR is associated with OS in HR-MDS. In the VALIDATE cohort, MST for patients with and without PB-CR was 25.8 months (95% CI: 22.1 - 30.1 months) and 17.7 months (95% CI: 16.7 - 19.4 months), respectively, with a statistically significant OS benefit seen among patients who achieved PB-CR compared with no PB-CR (HR: 0.65; 95% CI: 0.53 - 0.80; $p < 0.001$; **Figure 4A**). In the PANTHER cohort, we found a similar association between achieving PB-CR and improved OS compared with

patients who did not achieve PB-CR (MST 34.2 months [95% CI: 23.8 months - not reached] vs. 15.7 months [95% CI: 14.8 - 20.7 months]; $p<0.001$; **Figure 4B**). The association between achieving PB-CR and improved OS was maintained in sensitivity analyses using a modified 6-month landmark approach and restricting the analysis to patients treated with HMA monotherapy in both datasets (**Supplemental Figure 6**).

Achieving cCR per IWG 2023 is associated with improved OS in multivariable analysis:

Lastly, we fitted a multivariable model adjusted for patient age, sex, race, ethnicity, treatment type, IPSS-M risk category, baseline WBC, and allo-SCT treatment status to evaluate the impact of cCR as defined by IWG 2023 response criteria on OS. In the VALIDATE dataset, achieving cCR (vs. no cCR) as best response was associated with improved OS (HR: 0.65; 95% CI: 0.55 - 0.76; $p<0.001$; **Figure 5A**). Other variables associated with favorable OS in this model included receipt of allo-SCT (HR: 0.38; 95% CI: 0.31 - 0.46; $p<0.001$) and IPSS-M risk category (very-low, low, or moderate-low risk vs. very high risk; HR: 0.41; 95% CI: 0.31 - 0.53; $p<0.001$ and moderate-high or high risk vs. very high risk; HR: 0.52; 95% CI: 0.44 - 0.61; $p<0.001$). Conversely, decitabine monotherapy (vs. AZA monotherapy; HR: 1.26; 95% 1.03 - 1.53; $p=0.024$) was associated with inferior OS.

Due to the small number of allo-SCT recipients ($n=12$; 5.7%) in PANTHER, we conducted two multivariable models that adjusted for patient age, sex, race, ethnicity, IPSS-R risk category, and baseline WBC with and without adjustment for allo-SCT. In the model without allo-SCT, achieving a cCR was independently associated with

improved OS (HR: 0.69; 95% CI: 0.48 - 0.99; p=0.046). Among the other variables included in the model only IPSS-R risk category was associated with OS (**Figure 5B**). A similar association between cCR and improved OS was found in a model that included allo-SCT (HR: 0.69; 95% CI: 0.48 - 0.99; p=0.042; **Supplemental Figure 7**).

Discussion:

Response assessment for patients with MDS has undergone various updates over the past two decades with the most recent iteration being the IWG 2023 response criteria for patients with HR-MDS.^{9,13,20} The IWG 2023 response criteria were developed with the intent to overcome the shortcomings of the IWG 2006 criteria with regards to capturing the clinical benefits of novel investigational drugs and to serve as a more reliable surrogate for longer-term clinical end points such as OS.²⁰ The importance of response criteria as surrogates of clinical benefits including OS is essential for successful drug development as exemplified by inflated ORR in early, single-arm trials in MDS driven by higher rates of marrow CR, but that eventually did not result in an OS benefit in randomized phase III trials.^{18,21-23}

While the IWG 2023 response criteria are being increasingly included in clinical trials, validation of these criteria is limited.¹⁵ In this international collaboration study utilizing two independent datasets from VALIDATE, a large, international database of HR-MDS patients from the icMDS, and PANTHER, a prospective, phase III randomized controlled trial, we demonstrate that the new definition of cCR can reliably predict meaningful clinical benefit.

Furthermore, the newly proposed category of “near-CR” responses including CR-L and CRh also appears to be associated with improved OS. However, the statistical power for some of these categories was limited because of a small sample size in the PANTHER trial. Additionally, and not unsurprisingly, there was a trend towards inferior outcomes of patients with CRuni compared to patients with a CR or CRbi suggesting that trilineage hematologic recovery in addition to bone marrow blast count reduction is an important contributor to OS. However, the specific thresholds and extent of hematologic recovery that correlate best with long-term outcomes remain unclear and should be investigated prospectively.

One limitation of our VALIDATE dataset is the pragmatic (non-standardized timing of response assessments and treatment schedules), high rates of subsequent allo-SCT and *TP53* mutations, and the retrospective nature of the data collection. To mitigate these concerns, we used a well annotated dataset from the randomized PANTHER trial of AZA +/- pevonedistat as an external validation cohort, which importantly had a low rate of subsequent transplantation. Similar to the VALIDATE dataset, we found that achieving cCR per IWG 2023 response criteria as the best response was associated with improved OS further supporting the IWG 2023 response criteria as the new standard to assess meaningful clinical benefits in HR-MDS.

To further address concerns regarding immortal time bias and the effect of HMA combination therapies, we conducted additional analyses including time-dependent, modified 6-month landmark and subgroup analyses restricted to patients treated with HMA monotherapy and in patients with or without subsequent allo-SCT. These analyses

showed generally supportive findings and suggested that IWG 2023 criteria can apply to a wide range of clinical settings in the treatment of HR-MDS.

Furthermore, we evaluated the utility of PB-CR as another predictor of OS that acknowledges the fact that many patients in real-world practice do not undergo a repeated bone marrow biopsy for response assessment and effectiveness of treatment is gauged based on improvement in CBC parameters. In our analysis from both the VALIDATE and PANTHER dataset, we showed that achieving PB-CR was associated with OS benefit vs. not achieving PB-CR, reflecting the importance of achieving trilineage count recovery in HR-MDS. These findings warrant prospective validation in clinical studies.

Based on the subgroup analyses from VALIDATE and the fact that only 5.9% of patients enrolled in PANTHER proceeded to allo-SCT, our findings support that the OS benefit for patients achieving cCR is especially clear for patients who do not proceed with a subsequent allo-SCT. However, the impact on post-transplant outcomes in our analysis was less pronounced. In a prior analysis from our group, we demonstrated that in patients who receive a subsequent allo-SCT the underlying disease risk as defined by IPSS-M is more important than response to HMA-based therapy for long-term outcomes.²⁴ However, confounding by differences in time-to-transplant, transplant characteristics and post-progression therapies cannot be excluded and prospective studies are needed to fully evaluate the impact of IWG 2023 response criteria for the prediction of post-transplant outcomes.

Despite the use of two separate, large, well annotated, retrospective and prospectively collected datasets, some limitations remain. First, we only collected data on frontline

treatment with HMA-based therapies and therefore cannot comment on the performance of IWG 2023 criteria in the second line settings or using non-HMA based approaches where additional studies are needed. Second, VALIDATE was a pragmatic, real-world study. As such, HMA treatment schedules, timing of response assessments, and use of allo-SCT were not standardized, which might have influenced outcomes. However, timing of response assessments was standardized in the PANTHER trial and the conclusions drawn from VALIDATE remained unchanged.

Third, information on comorbidities and social determinants of health were not available in our study and could have influenced outcomes. However, due to the large, global dataset that originated from two independent cohorts, we believe that our study is broadly representative of patients with MDS treated in contemporary practice and results should be generalizable to other settings. Fourth, information on cause of death was not available in our datasets and we therefore cannot ascertain the impact of response to HMA therapy on cause-specific or AML-related deaths.

Fifth, our cohort was enriched for patients with adverse prognostic features including a complex karyotype (42.3%) and *TP53* mutations (33.0%). Especially, the latter constitutes a subset of patients in whom response to initial treatment does not necessarily correlate with improved OS, as *TP53* mutation status may have greater prognostic significance than initial treatment response.²⁵ Lastly, establishing surrogacy between IWG 2023 and OS in a stringent fashion requires, in addition to individual patient-level validation that we performed, a trial-level validation across multiple prospective trials. We hope this will be accomplished using individual patient-level data from the multiple, recently reported negative phase 3 trials in HR-MDS.

In conclusion, our results from both retrospective and prospective data from VALIDATE and PANTHER databases represent one of the largest efforts that validate the IWG 2023 criteria as a reliable predictor of meaningful clinical benefit in HMA-treated patients with HR-MDS. Beyond CR, we also show that cCR, near-CR endpoints including CRbi and CRh, and PB-CR are also associated with meaningful clinical benefits. Our findings support the use of IWG 2023 criteria as the standard method to assess meaningful clinical benefits in patients with HR-MDS in both clinical practice and clinical trials.

References:

1. Zeidan AM, Shallis RM, Wang R, Davidoff A, Ma X. Epidemiology of myelodysplastic syndromes: Why characterizing the beast is a prerequisite to taming it. *Blood Rev.* 2019;34:1-15.
2. Bernard E, Tuechler H, Greenberg PL, et al. Molecular International Prognostic Scoring System for Myelodysplastic Syndromes. *NEJM Evidence.* 2022;1(7):EVIDoA2200008.
3. Greenberg PL, Tuechler H, Schanz J, et al. Revised international prognostic scoring system for myelodysplastic syndromes. *Blood.* 2012;120(12):2454-2465.
4. de Witte T, Bowen D, Robin M, et al. Allogeneic hematopoietic stem cell transplantation for MDS and CMML: recommendations from an international expert panel. *Blood.* 2017;129(13):1753-1762.
5. Heidenreich S, Ziaikos D, de Wreede LC, et al. Allogeneic stem cell transplantation for patients age ≥ 70 years with myelodysplastic syndrome: a retrospective study of the MDS Subcommittee of the Chronic Malignancies Working Party of the EBMT. *Biol Blood Marrow Transplant.* 2017;23(1):44-52.
6. Nakamura R, Saber W, Martens MJ, et al. Biologic assignment trial of reduced-intensity hematopoietic cell transplantation based on donor availability in patients 50-75 years of age with advanced myelodysplastic syndrome. *J Clin Oncol.* 2021;39(30):3328-3339.
7. Fenaux P, Mufti GJ, Hellstrom-Lindberg E, et al. Efficacy of azacitidine compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes: a randomised, open-label, phase III study. *Lancet Oncol.* 2009;10(3):223-232.
8. Zeidan AM, Stahl M, DeVeaux M, et al. Counseling patients with higher-risk MDS regarding survival with azacitidine therapy: are we using realistic estimates? *Blood Cancer J.* 2018;8(6):55.
9. Cheson BD, Greenberg PL, Bennett JM, et al. Clinical application and proposal for modification of the International Working Group (IWG) response criteria in myelodysplasia. *Blood.* 2006;108(2):419-425.
10. Sekeres MA, Othus M, List AF, et al. Randomized phase II study of azacitidine alone or in combination with lenalidomide or with vorinostat in higher-risk myelodysplastic syndromes and chronic myelomonocytic leukemia: North American Intergroup Study SWOG S1117. *J Clin Oncol.* 2017;35(24):2745-2753.
11. Kornblith AB, Herndon JE, Silverman LR, et al. Impact of azacytidine on the quality of life of patients with myelodysplastic syndrome treated in a randomized phase III trial: a Cancer and Leukemia Group B study. *J Clin Oncol.* 2002;20(10):2441-2452.
12. Silverman LR, Fenaux P, Mufti GJ, et al. Continued azacitidine therapy beyond time of first response improves quality of response in patients with higher-risk myelodysplastic syndromes. *Cancer.* 2011;117(12):2697-2702.
13. Platzbecker U, Fenaux P, Adès L, et al. Proposals for revised IWG 2018 hematological response criteria in patients with MDS included in clinical trials. *Blood.* 2019;133(10):1020-1030.
14. Zeidan AM, Platzbecker U, Bewersdorf JP, et al. Consensus proposal for revised International Working Group response criteria for higher risk myelodysplastic syndromes. *Blood.* 2023;141(17):2047-2061.
15. Urrutia S, Bose P, Alvarado Y, et al. Prospective performance of the IWG-2023 criteria and IPSS-M in a phase 2 trial of guadecitabine for higher-risk MDS or CMML. *Blood Neoplasia.* 2024;1(2):100008.
16. Zugasti I, Lopez-Guerra M, Castaño-Díez S, et al. Hypomethylating agents plus venetoclax for high-risk MDS and CMML as bridge therapy to transplant: a GESMD study. *Exp Hematol Oncol.* 2025;14(1):61.
17. Pleyer L, Vaisband M, Drost M, et al. Cox proportional hazards deep neural network identifies peripheral blood complete remission to be at least equivalent to morphologic complete

remission in predicting outcomes of patients treated with azacitidine-A prospective cohort study by the AGMT. *Am J Hematol.* 2023;98(11):1685-1698.

18. Adès L, Girshova L, Doronin VA, et al. Pevonedistat plus azacitidine vs azacitidine alone in higher-risk MDS/chronic myelomonocytic leukemia or low-blast-percentage AML. *Blood Adv.* 2022;6(17):5132-5145.

19. Greenberg P, Cox C, LeBeau MM, et al. International scoring system for evaluating prognosis in myelodysplastic syndromes. *Blood.* 1997;89(6):2079-2088.

20. Zeidan AM, Platzbecker U, Bewersdorf JP, et al. Consensus proposal for revised International Working Group 2023 response criteria for higher-risk myelodysplastic syndromes. *Blood.* 2023;141(17):2047-2061.

21. Garcia JS, Platzbecker U, Odenike O, et al. Efficacy and safety of venetoclax plus azacitidine for patients with treatment-naïve high-risk myelodysplastic syndromes. *Blood.* 2025;145(11):1126-1135.

22. Swords RT, Erba HP, DeAngelo DJ, et al. Pevonedistat (MLN4924), a First-in-Class NEDD8-activating enzyme inhibitor, in patients with acute myeloid leukaemia and myelodysplastic syndromes: a phase 1 study. *Br J Haematol.* 2015;169(4):534-543.

23. Garcia-Manero G, Platzbecker U, Fenaux P, et al. MDS-1497: primary analysis of the randomized, Phase 3 VERONA Study of venetoclax with azacitidine versus placebo with azacitidine in patients with treatment-naïve, intermediate and higher-risk myelodysplastic syndromes. *Clin Lymphoma Myeloma Leuk.* 2025;25(Supplement 1):S644-S645.

24. Rolles B, Bewersdorf JP, Kewan T, et al. Impact of response to hypomethylating agent-based therapy on survival outcomes in the context of baseline clinical-molecular risk and transplant status in patients with higher-risk myelodysplastic syndromes/neoplasms: an analysis from the International Consortium for MDS (icMDS) Validate Database. *Blood.* 2024;144(Supplement 1):664.

25. Welch JS, Petti AA, Miller CA, et al. TP53 and decitabine in acute myeloid leukemia and myelodysplastic syndromes. *N Engl J Med.* 2016;375(21):2023-2036.

Table 1: Baseline patient and disease characteristics of VALIDATE cohort by treatment type

	Total 1,047 (100.00%)	Azacitidine 637 (60.84%)	Decitabine 176 (16.81%)	Combination 234 (22.35%)	
Parameter	n (%) / Median [Interquartile Range]				p-value
Demographics					
- Age at diagnosis (years)	69.0 [62.0 - 75.0]	69.0 [62.0 - 76.0]	66.0 [60.0 - 71.3]	70.0 [63.3 - 74.0]	< 0.001
- Male sex	675 (64.47%)	419 (65.78%)	110 (62.50%)	146 (62.39%)	0.55
- US Site	885 (84.53%)	500 (78.49%)	176 (100.0%)	209 (89.32%)	< 0.001
- Race					< 0.001
• Asian	20 (1.91%)	10 (1.57%)	6 (3.41%)	4 (1.71%)	
• Black/African American	50 (4.78%)	34 (5.34%)	13 (7.39%)	3 (1.28%)	
• White	838 (80.04%)	483 (75.82%)	148 (84.09%)	207 (88.46%)	
• Unknown	139 (13.28%)	110 (17.27%)	9 (5.11%)	20 (8.55%)	
- Ethnicity					< 0.001
• Hispanic/Latino	83 (7.93%)	55 (8.63%)	13 (7.39%)	15 (6.41%)	
• Non-Hispanic/Latino	816 (77.94%)	468 (73.47%)	151 (85.80%)	197 (84.19%)	
• Unknown	148 (14.14%)	114 (17.90%)	12 (6.82%)	22 (9.40%)	
Baseline disease characteristics*					
- White blood cell count	2.70 [1.90 - 4.21]	2.70 [1.90 - 4.30]	2.60 [1.90 - 4.16]	2.90 [1.85 - 4.11]	0.97
- Absolute neutrophil count	1.10 [0.57 - 2.00]	1.10 [0.57 - 2.01]	1.12 [0.69 - 1.91]	1.04 [0.50 - 1.96]	0.71
- Hemoglobin	8.90 [8.00 - 10.30]	8.90 [8.00 - 10.20]	8.70 [7.97 - 10.20]	9.00 [8.00 - 10.40]	0.51
- Platelet count	66.0 [38.00 - 118.0]	64.0 [37.00 - 114.0]	60.5 [33.00 - 106.0]	79.0 [44.00 - 135.8]	0.01
- Bone marrow blast %	8.00 [5.00 - 13.00]	7.50 [5.00 - 13.00]	7.50 [4.00 - 12.00]	10.00 [5.32 - 15.00]	0.001
- RBC Transfusion dependent	464 (44.27%)	283 (44.29%)	79 (44.89%)	102 (43.59%)	0.96
- Platelet Transfusion dependent	230 (21.95%)	128 (20.03%)	42 (23.86%)	60 (25.64%)	0.17
- IPSS-R (category)					0.001
• Very Low	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	
• Low	4 (0.38%)	3 (0.47%)	0 (0.00%)	1 (0.43%)	
• Intermediate	251 (24.09%)	165 (26.03%)	33 (18.75%)	53 (22.84%)	
• High	357 (34.26%)	235 (37.07%)	53 (30.11%)	69 (29.74%)	
• Very High	430 (41.27%)	231 (36.44%)	90 (51.14%)	109 (46.98%)	
- IPSS-M (category)					< 0.001
• Very Low	1 (0.10%)	1 (0.16%)	0 (0.00%)	0 (0.00%)	
• Low	31 (2.96%)	24 (3.77%)	1 (0.57%)	6 (2.56%)	
• Moderate Low	82 (7.83%)	61 (9.58%)	6 (3.41%)	15 (6.41%)	
• Moderate High	148 (14.14%)	104 (16.33%)	18 (10.23%)	26 (11.11%)	
• High	335 (32.00%)	213 (33.44%)	62 (35.23%)	60 (25.64%)	
• Very High	450 (42.98%)	234 (36.73%)	89 (50.57%)	127 (54.27%)	
- Complex karyotype	442 (42.22%)	237 (37.21%)	93 (52.84%)	112 (48.07%)	< 0.001
- TP53 mutations†	268 (33.00%)	111 (23.57%)	57 (41.30%)	100 (49.26%)	< 0.001
Treatment-Related MDS					0.024
- Yes	253 (24.16%)	136 (21.35%)	53 (30.11%)	64 (27.35%)	
- No	757 (72.30%)	482 (75.67%)	118 (67.05%)	157 (67.09%)	
- Unknown	37 (3.53%)	19 (2.98%)	5 (2.84%)	13 (5.56%)	

Treatment characteristics					
- Time from diagnosis to HMA initiation (days)	33 [18 - 66]	32 [17 - 69]	32 [17 - 53]	36 [20 - 65]	0.089
- Number of cycles	6 [4 - 10]	6 [4 - 11]	5 [3 - 7]	5 [3 - 9]	< 0.001
- Bone marrow transplant	385 (36.77%)	212 (33.28%)	73 (41.48%)	100 (42.74%)	0.014
- Time from HMA to SCT (months)	5.55 [4.20 - 7.67]	5.78 [4.24 - 8.00]	5.22 [3.98 - 7.82]	5.11 [3.98 - 6.76]	0.16

*All values and percentages are from and with respect to the available, non-missing set of records.

#Measurement taken prior to HMA initiation.

†Percentage with respect to the number of patients with baseline molecular data available.

◊Combination therapies include: HMA + Venetoclax [n = 98; (9.36%)] and Other HMA Combination. [n = 136; (12.99%)].

Table 2: Baseline patient and disease characteristics of PANTHER cohort

	Total 221 (100.00%)	Azacitidine 113 (51%)	Azacitidine + pevonedistat 108 (49%)	p-value
Parameter	n (%) / Median [Interquartile Range]			
Demographics				
- Age at diagnosis (years) (N=221)	73 [67 - 77]	73 [67 - 78]	73 [68 - 77]	0.70
- Male sex (N=221)	131 (59%)	66 (58%)	65 (60%)	0.89
Baseline disease characteristics*				
- White blood cell count	2.7 (1.8 – 4.1)	2.7 (1.9 – 4.1)	2.65 (1.7 – 3.9)	0.46
- Absolute neutrophil count (N=210)	1.1 (0.5 – 2.1)	1.1 (0.5 – 2.33)	1.0 (0.48 – 2.00)	0.53
- Hemoglobin (N=212)	9.4 (8.5 – 10.4)	9.2 (8.5 – 10.2)	9.4 (8.5 – 10.6)	0.57
- Platelet count (N=194)	67 (39 - 128)	62 (33 - 117)	75 (41 - 129)	0.38
- Bone marrow blast % (N=214)	10 (7 - 13)	10 (7 - 13)	10 (7.0 – 12.7)	0.83
- IPSS-R (category) (N=221)				0.90
• Very Low	-	-	-	
• Low	-	-	-	
• Intermediate	55 (25%)	29 (26%)	26 (24%)	
• High	86 (39%)	45 (40%)	41 (38%)	
• Very High	80 (36%)	39 (35%)	41 (38%)	
- TP53 mutation (N=221)	51 (23%)	25 (22%)	26 (24%)	0.75
Treatment characteristics				
Bone marrow transplant (N=221)	13 (5.9%)	8 (7.1%)	5 (4.6%)	0.57

Table 3: Overall survival by best response for patients included in VALIDATE

Best Response	n (%)	Median Survival Time (95% CI)	HR (95% CI)	p-value
CR	160 (15.28%)	31.34 (25.46; 41.76)	Ref.	
CRequ	14 (1.34%)	18.51 (10.05; NA)	1.765 (0.968; 3.219)	0.064
CRbi	96 (9.17%)	26.81 (21.32; 34.89)	1.141 (0.825; 1.580)	0.43
CRuni	103 (9.84%)	20.80 (16.56; 29.67)	1.317 (0.962; 1.802)	0.086
CRh	6 (0.57%)	24.51 (19.25; NA)	0.790 (0.195; 3.206)	0.74
PR	11 (1.05%)	11.56 (8.87; NA)	2.533 (1.277; 5.026)	0.008
HI	149 (14.23%)	16.89 (13.77; 22.57)	1.675 (1.273; 2.204)	< 0.001
No Response	288 (27.51%)	21.29 (17.77; 23.43)	1.453 (1.136; 1.858)	0.003
PD	199 (19.01%)	13.21 (11.24; 16.07)	2.153 (1.666; 2.782)	< 0.001
<2 cycles	21 (2.01%)	7.92 (5.16; 15.90)	3.726 (2.297; 6.044)	< 0.001
Overall	1047 (100%)	19.25 (17.61; 21.32)		< 0.001

Figure legends:

Figure 1: Overall survival by best response per IWG 2023 criteria in VALIDATE

Figure 1: Overall survival by best response. Figure 1 shows OS from date of HMA initiation by best response category as defined by IWG 2023 criteria. Due to the small number of patients ($n < 20$), CRh, CRequ, or PR as best response categories were excluded.

Figure 2: Association between cCR and OS in VALIDATE and PANTHER

Figure 2: Association between cCR and OS in VALIDATE and PANTHER. Figure 2 shows the association between OS from date of HMA initiation (VALIDATE) or randomization (PANTHER) by cCR status. (A) and (B) show OS for the entire VALIDATE and PANTHER cohorts, respectively. To minimize the impact of immortal time bias, we performed an adjusted analysis that treated both allo-SCT and IWG response as time-dependent variables. At risk tables for these analyses are not shown as the number of patients at risk is not monotonically decreasing and can potentially be misleading, especially at earlier time points. Analyses were restricted to patients treated with HMA monotherapy to evaluate validity of IWG 2023 criteria in this setting (panel [C] for VALIDATE and [D] for PANTHER).

Figure 3: Overall survival by CR/CR-L status

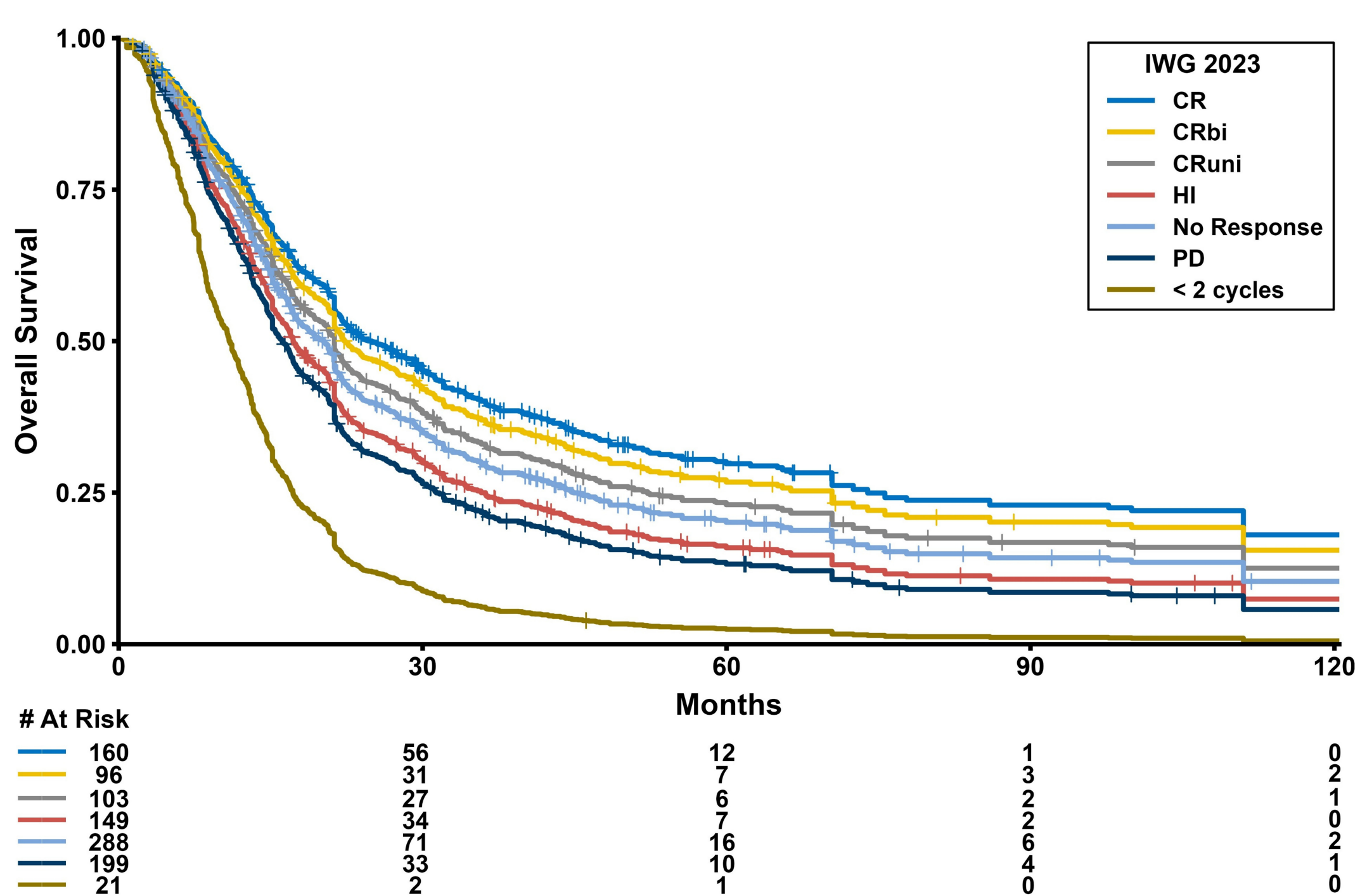
Figure 3: Overall survival by CR/CR-L status. Figure 3 shows overall survival from date of HMA initiation (VALIDATE) or randomization (PANTHER) by best response categorized as CR, CR-L, or non-CR per IWG 2023 criteria. Panel (A) and (B) show results for VALIDATE and PANTHER cohort, respectively.

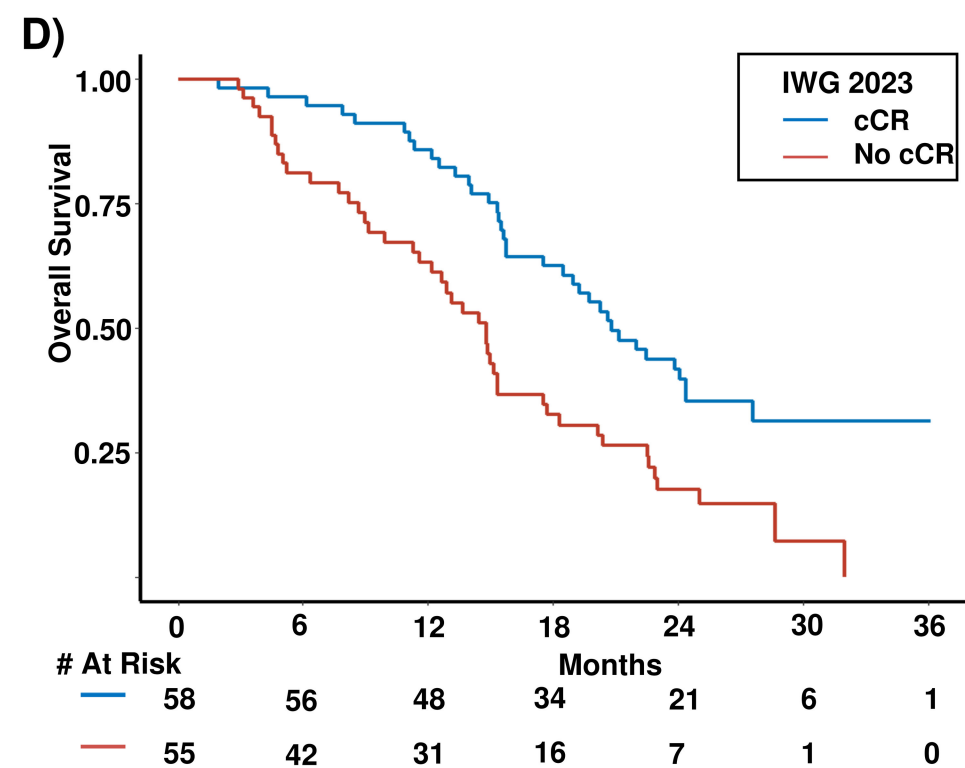
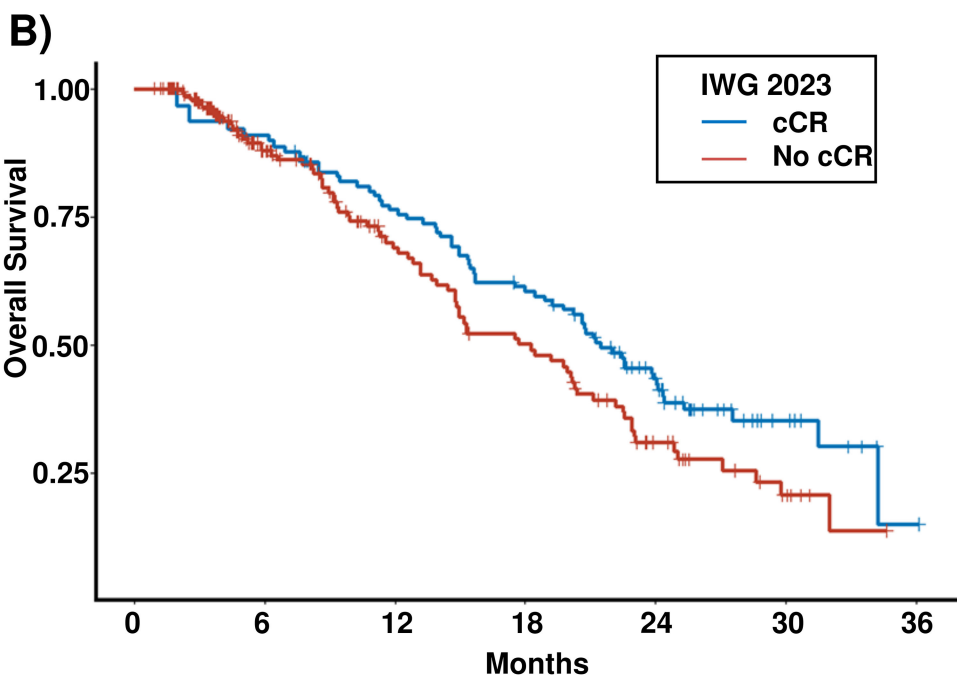
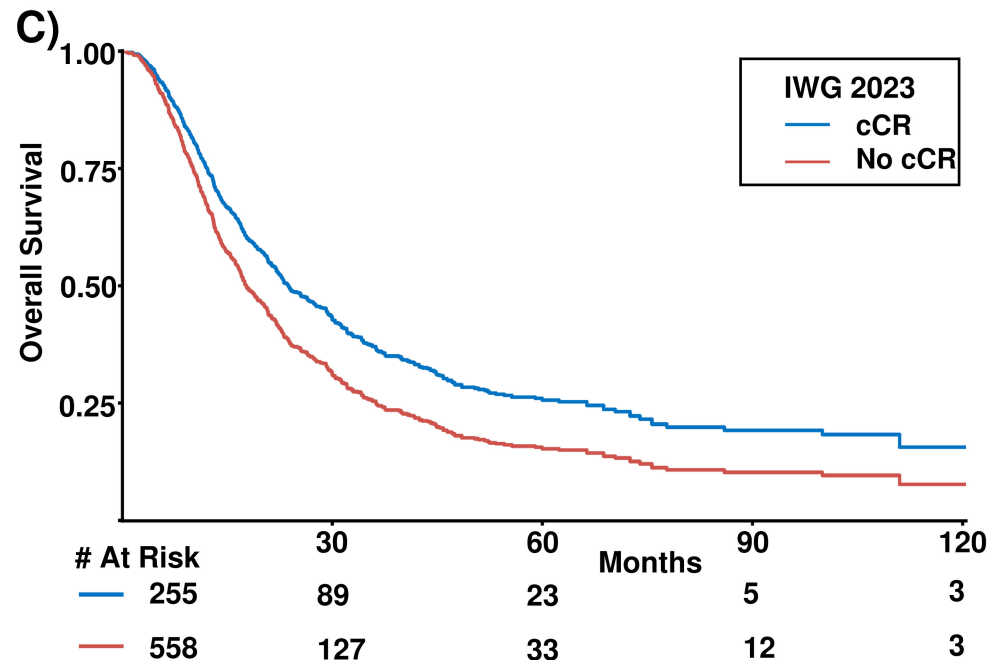
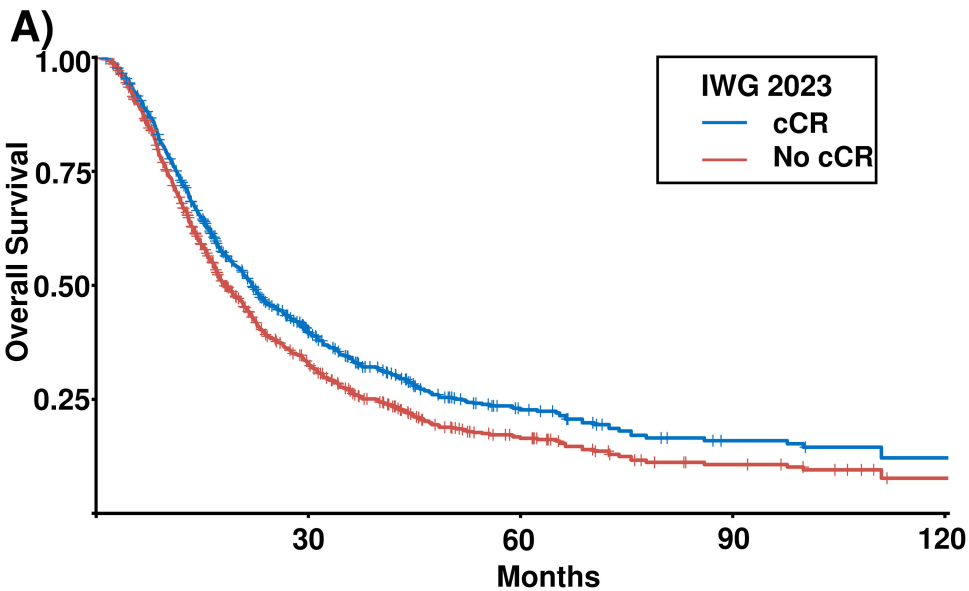
Figure 4: Overall survival by PB-CR status

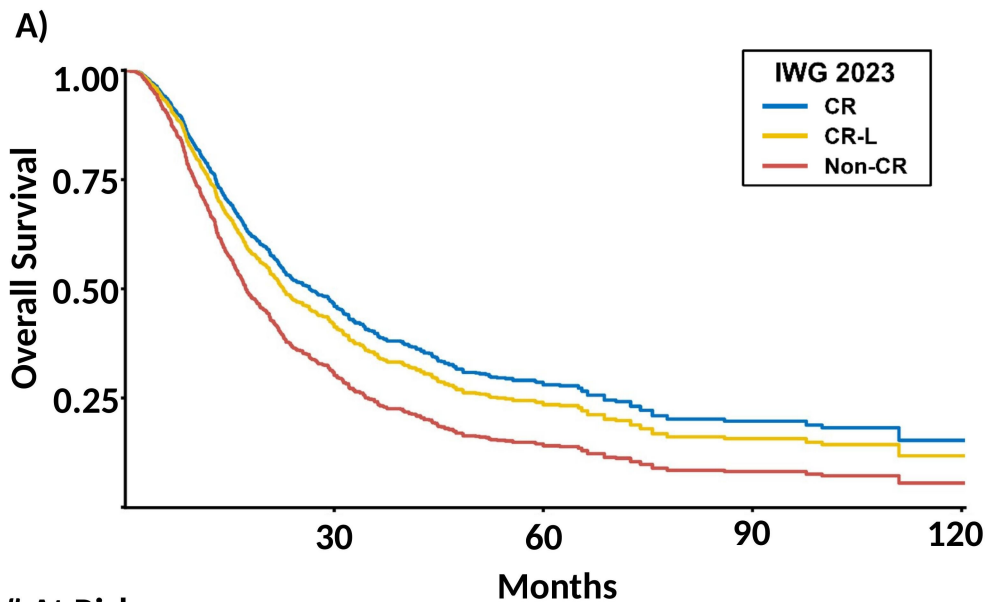
Figure 4: Overall survival by PB-CR status. Figure 4 shows overall survival from date of HMA initiation (VALIDATE) or randomization (PANTHER) by best response categorized as PB-CR vs. no PB-CR per Pleyer et al.¹⁷ Panel (A) and (B) show results for VALIDATE and PANTHER cohorts, respectively.

Figure 5: Multivariable model of predictors OS

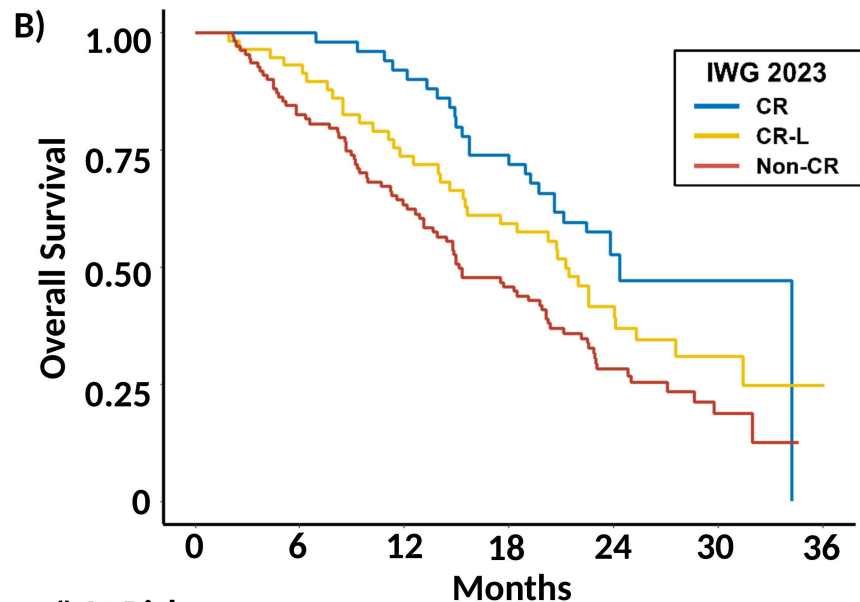
Figure 5: Multivariable model of predictors OS. Figure 5 shows a Cox proportional hazards model of predictors OS for VALIDATE (A) and PANTHER cohorts (B). To reduce the impact of immortal time bias, analyses in the VALIDATE cohort treated both response and allo-SCT as time-dependent variables. Allo-SCT was omitted from the PANTHER model as only a small subset of patients ($n=12$; 5.7%) underwent allo-SCT in this cohort.



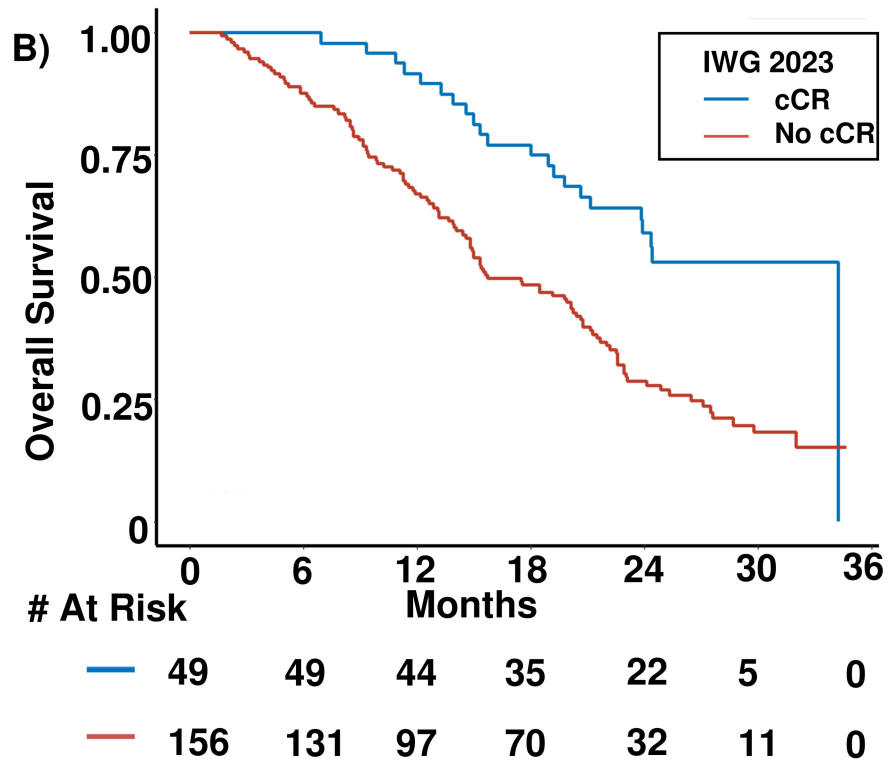
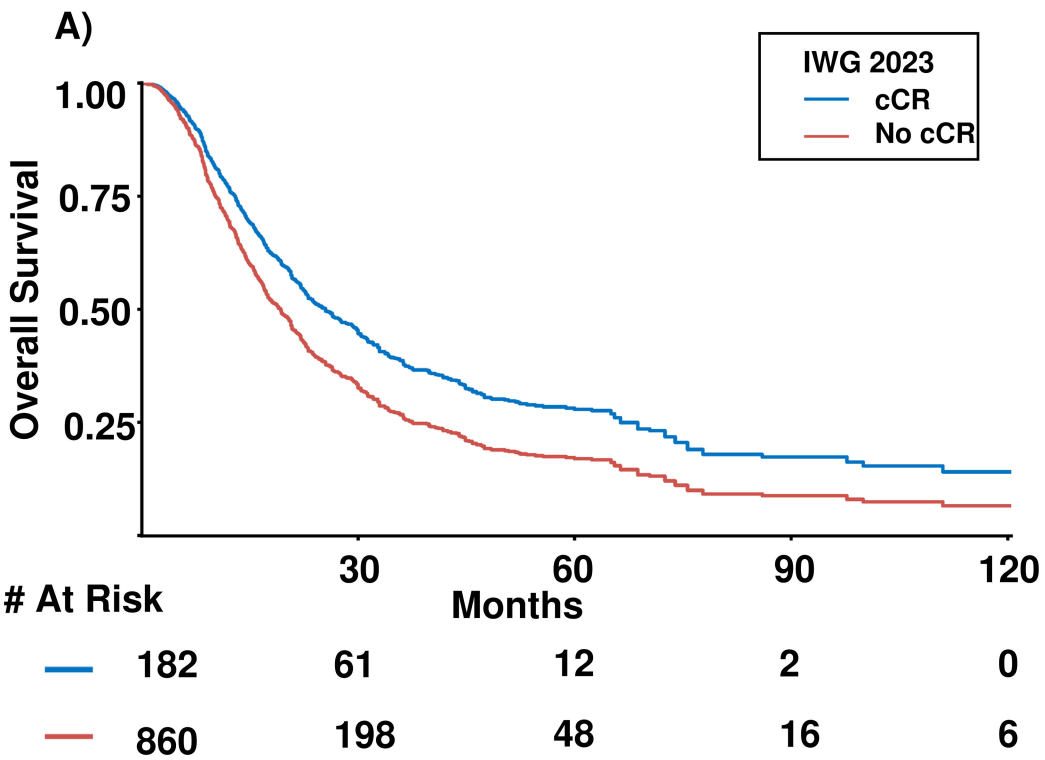


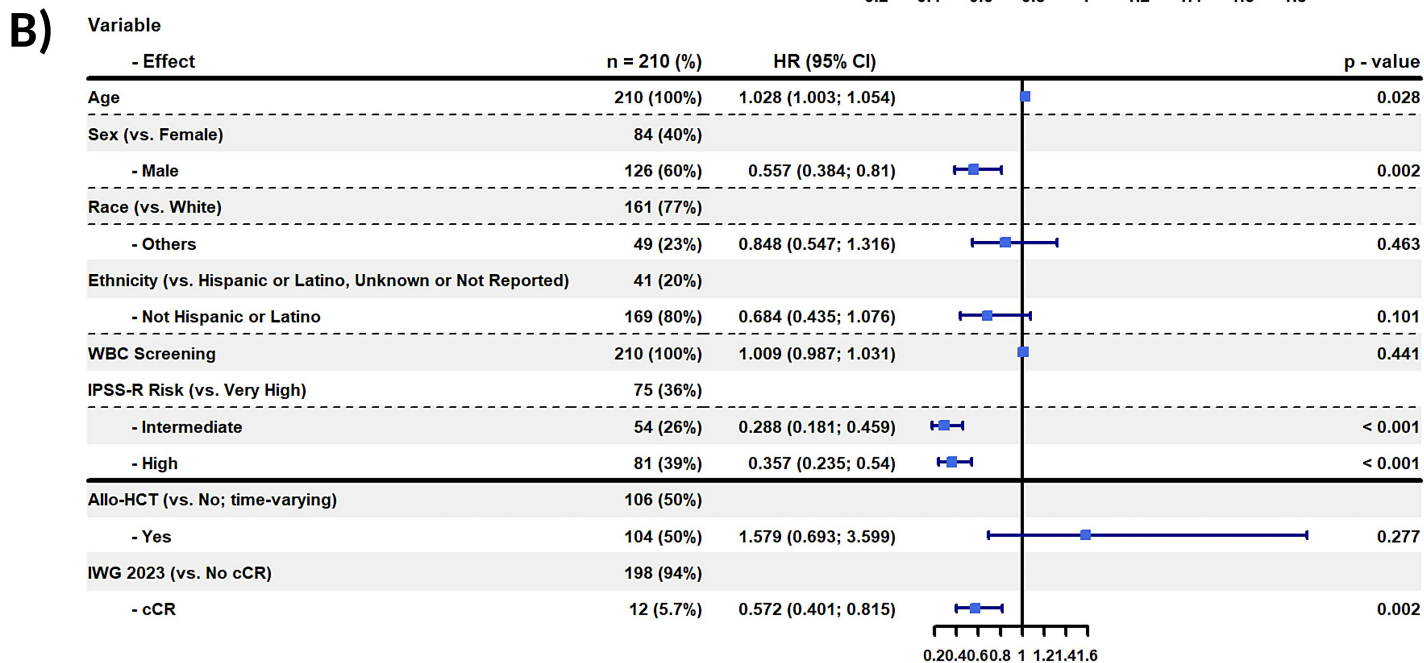
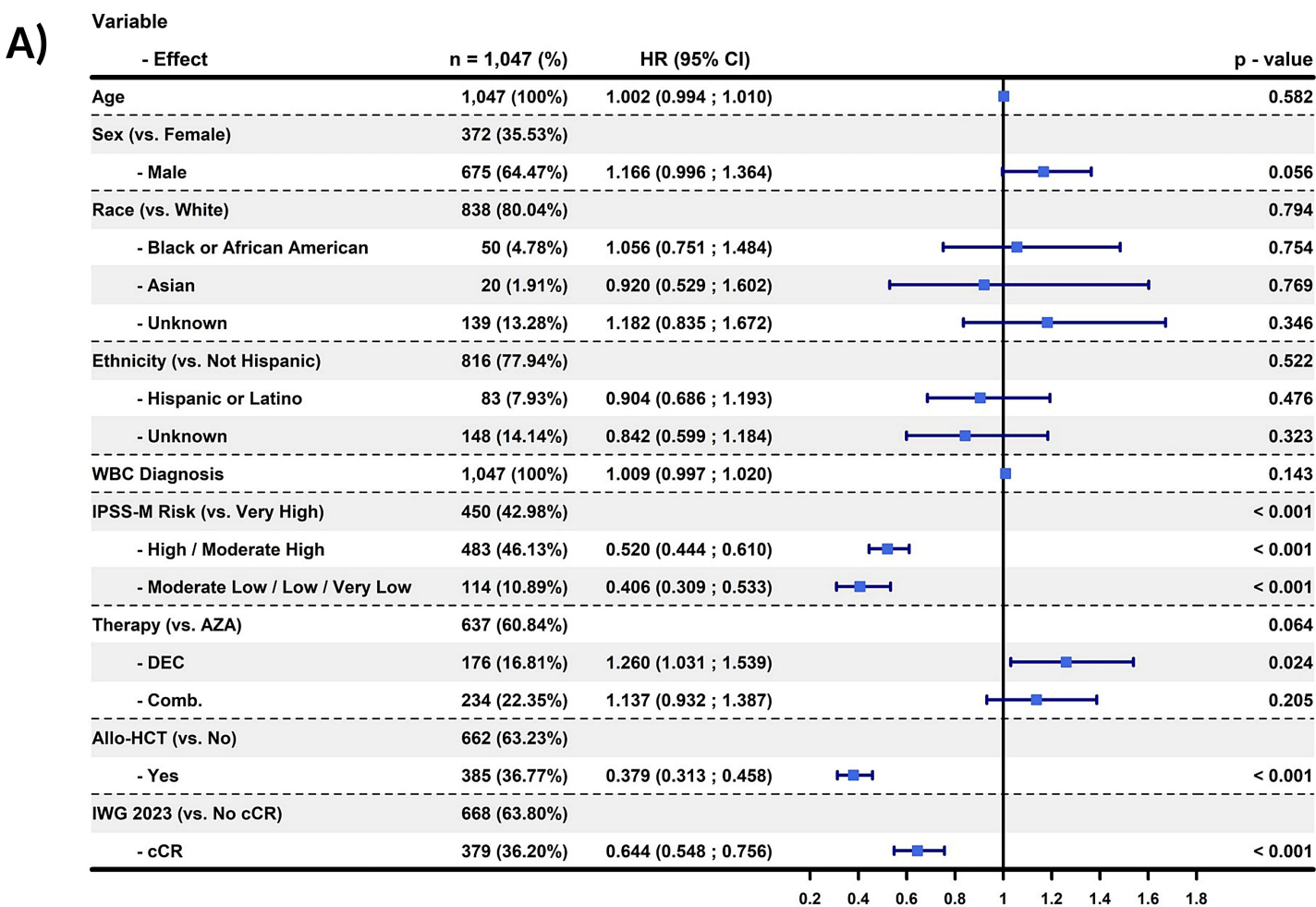


# At Risk				
174	60	13	1	0
205	59	13	5	3
668	141	34	12	3



# At Risk							
51	51	46	46	35	5	0	
58	54	41	41	33	6	1	
112	87	65	65	47	7	0	





Supplement

Validation of the International Working Group 2023 response criteria for higher risk myelodysplastic syndromes/neoplasms using the International Consortium for Myelodysplastic Syndromes VALIDATE database and a randomized phase III trial

Jan Philipp Bewersdorf^{1, #}, Tariq Kewan^{1, #}, Luca Lanino^{1, 2}, Ondrej Blaha³, Yiwen Liu⁴, Jessica M. Stempel¹, Najla H. Al Ali⁵, Amy E. DeZern⁶, Mikkael A. Sekeres⁷, Geoffrey L. Uy⁸, Samuel Urrutia⁸, Hetty E. Carraway⁹, Pinkal Desai¹⁰, Elizabeth A. Griffiths¹¹, Eytan M. Stein¹², Andrew M. Brunner¹³, Christine McMahon¹⁴, Rory M. Shallis⁵, Joshua F. Zeidner¹⁵, Michael R. Savona¹⁶, Hayley Hawkins¹⁶, Namrata Sonia Chandhok¹⁷, Constantine Logothetis¹⁸, Aram Bidikian¹, Ted M. Getz¹, Gail J. Roboz¹⁰, Benjamin Rolles^{1, 17}, Eunice S. Wang¹¹, Amyah C. Harris¹³, Maria L. Amaya¹⁴, Somedeb Ball¹⁶, Justin Grenet¹⁰, Zhuoer Xie⁵, Yazan F. Madanat¹⁸, Yasmin Abaza¹⁹, Talha Badar²⁰, Jaclynn Campos²⁰, Torsten Haferlach²¹, Jaroslaw P. Maciejewski⁹, David Sallman⁵, Rahul S. Vedula¹⁷, H. Moses Murdock¹⁷, Harrison K. Tsai¹⁷, Evan C. Chen¹⁷, Lionel Ades²², R. Coleman Lindsley¹⁷, Anoop Enjeti²³, Kamal Al-Rabi²⁴, Khalid Halahleh²⁴, Devendra Hiwase²⁵, Maria Diez-Campelo²⁶, David Valcarcel²⁷, Claudia Haferlach²¹, Lisa Pleyer²⁸, Ioannis Kotsianidis²⁹, Vassiliki Pappa³⁰, Valeria Santini³¹, Angela Consagra³¹, Aref Al-Kali³², Seishi Ogawa³³, Yasuhito Nannya³⁴, Matteo Giovanni Della Porta^{2, †}, Rami S. Komrokji^{5, †}, Maximilian Stahl^{1, 17, †}, and Amer M. Zeidan^{1, †, *}

1Department of Internal Medicine, Section of Hematology, Yale University, New Haven, CT, USA; 2Humanitas Clinical and Research Center, IRCCS, Rozzano, Italy; 3School of Public Health, Yale Center for Analytical Sciences, Yale University, New Haven, CT, USA; 4: Department of Data Science, Dana Farber Cancer Institute, Boston, MA; 5: Department of Malignant Hematology, H. Lee Moffitt Cancer Center, Tampa, FL, USA; 6Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University, Baltimore, MD, USA; 7Sylvester Comprehensive Cancer Center, University of Miami Miller School of Medicine, Miami, FL, USA; 8Washington University School of Medicine, Saint Louis, MO, USA; 9Department of Hematology and Medical Oncology, Taussig Cancer Institute, Leukemia Program, Cleveland Clinic, Cleveland, OH, USA; 10Weill Cornell Medicine, New York City, NY, USA; 11Roswell Park Comprehensive Cancer Center, Buffalo, NY, USA; 12Department of Medicine, Leukemia Service, Memorial Sloan Kettering Cancer Center, New York City, NY, USA; 13Division of Hematology/Oncology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA; 14University of Colorado, Aurora, CO, USA; 15University of North Carolina, Lineberger Comprehensive Cancer Center, Chapel Hill, NC, USA; 16Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, TN, USA; 17Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA; 18Harold C. Simmons Comprehensive Cancer Center, UT Southwestern Medical Center, Dallas, TX, USA; 19Northwestern University, Chicago, USA; 20Division of Hematology-Oncology, Mayo Clinic, Jacksonville, FL, USA; 21MLL Munich Leukemia Laboratory, Munich, Germany; 22 Hématologie Seniors, Assistance Publique - Hôpitaux de Paris (AP-HP), Université Paris Cité, Paris, France. 23Department of Haematology, Calvary Mater Hospital, Newcastle, New South Wales, Australia; 24King Hussein Cancer Center, Amman, Jordan; 25Department of Haematology, The Royal Adelaide Hospital, Adelaide, Australia; 26Hematology Department, Institute of Biomedical Research of Salamanca, University Hospital of Salamanca, Salamanca, Spain; 27Department of Hematology and Oncology, University Hospital Vall d'Hebron, Barcelona, Spain; 28: 3rd Medical Department with Hematology, Medical Oncology, Hemostaseology, Rheumatology and Infectiology, Oncologic Center, Paracelsus Medical University, Salzburg, Austria; 29Department of Hematology, Democritus University of Thrace Medical School, Alexandroupolis, Greece; 30Hematology Division, Second Department of Internal Medicine, Attikon University Hospital, National and Kapodistrian University of Athens, Athens, Greece; 31MDS Unit, Hematology, AOU Careggi - Department of Experimental and Clinical Medicine, University of Florence, Florence, Italy. 32Mayo Clinic, Rochester, MN, USA. 33Department of Pathology and Tumor Biology, Institute for the Advanced Study of Human Biology (WPI-ASHBi), Kyoto University, Kyoto, Japan; Department of Innovative Medicine, Faculty of Medicine, Kindai University, Osaka, Japan. 34: Institute of Medical Science, the University of Tokyo, Tokyo, Japan

These first authors contributed equally; † These last authors contributed equally

* Corresponding author:

Amer M. Zeidan, MBBS, MHS

Division of Malignant Hematology; Section of Medical Oncology and Hematology

Department of Internal Medicine; Yale School of Medicine

333 Cedar Street

PO Box 208028

New Haven, CT 06520-8028

Email: amer.zeidan@yale.edu

Phone: 203-737-7103

Fax: 203-785-7232

Supplemental Methods:

Methods:

Data sources:

The VALIDATE database is a collaboration of 31 centers and cooperative groups in 9 countries (**Supplemental Table 1**) containing patients with MDS and AML treated with HMA-based therapies in the frontline setting. Patients were entered by the participating study sites with central validation of baseline risk stratification, molecular disease characteristics, and response assessments by two investigators independently. International Prognostic Scoring System (IPSS), IPSS-revised (IPSS-R), and IPSS-molecular (IPSS-M) were calculated centrally.(1-3) Biallelic *TP53* mutation status was classified centrally as defined previously.(4) Next-generation sequencing was performed locally. For patients with missing molecular profiling at time of diagnosis, the IPSS-M score was calculated according to the original published methodology using the publicly available code, which estimates risk under best-case, average-case, and worst-case scenarios for untested variables, thereby permitting assignment of a risk score and corresponding risk category despite absent mutational data.(1)

As an independent validation cohort, we used patients with HR-MDS who received treatment with either AZA monotherapy or AZA + pevonedistat in the randomized, phase III PANTHER trial.(5) As only 5.9% of all patients in the trial received a subsequent allo-SCT, it allowed us to examine the association between IWG 2023 response categories and OS in a population largely treated with non-curative, homogenous medical therapy.

The VALIDATE protocol was exempted by Institutional Review Board (IRB) of the coordinating Center (Yale University) due to minimal risk and need for informed consent from patients was waived. The protocol was exempted, acknowledged, or approved by the other centers according to their local policies. All participating centers signed a data use agreement. VALIDATE is supported by an independent grant from AbbVie. AbbVie had no role in the study design, data analyses, interpretation, or writing of the manuscript. A de-identified dataset of The PANTHER trial was kindly shared by the trial sponsor, Takeda, who also had no role in the study design, data analyses, interpretation, or writing of the manuscript.

Patient selection:

From all patients recorded in the VALIDATE database, we included adults (≥ 18 years of age) with HR-MDS who received frontline treatment with HMA (AZA, DEC, or oral decitabine/cedazuridine) either as monotherapy or as part of a combination therapy. HR-MDS was defined as either IPSS score of ≥ 1.5 or an IPSS-R of ≥ 4 at time of diagnosis.(2, 3) Prior treatment with other MDS-directed therapies other than erythropoiesis stimulating agents was an exclusion criterion. Patients were further excluded if HMA treatment type and start date were unknown or if survival status at last follow-up could not be ascertained.

For the PANTHER trial cohort, we excluded patients with CMML or oligoblastic AML. HR-MDS was defined as IPSS-R very high-, high-, or intermediate-risk disease with intermediate-risk disease patients requiring $\geq 5\%$ bone marrow myeloblasts for enrollment.(5) Prior treatment with HMA was not allowed on the PANTHER trial.(5)

Supplemental Figure 1 shows a CONSORT diagram of patient selection from VALIDATE and PANTHER trial for our analysis.

Response assessments:

Study sites participating in VALIDATE provided variables required for response assessments at baseline and at the time of first and best response. All responses were assessed centrally by two investigators independently and classified according to IWG 2006 and IWG 2023 as well as by peripheral blood – CR (PB-CR) criteria.(6-8) As timing of response assessments in VALIDATE was not standardized due to the retrospective, real-world nature of the analysis, we only included patients who had a bone marrow assessment within 180 days of starting HMA-based therapy unless there was overt evidence of disease progression (e.g., progression based on peripheral blood blast percentage). We used a 180-day window as most responses to HMA-based therapies occur within this timeframe and given that validation of response criteria was the primary objective of this analysis.(9, 10) This was also intended to preserve the validity of response assessments as some patients with initial improvement in their hematologic parameters only underwent a repeat bone marrow at the time of subsequent disease progression and would thus have been incorrectly classified as having progressive disease (PD) as their best response. We further classified patients who received less than 2 cycles of HMA-based therapy as a separate response category to capture early disease progression or toxicity that precluded formal response assessment.

Response assessment in the PANTHER trial occurred on day 22 (± 6 days) of cycles 2, 4, 6, and 9 and every third treatment cycle thereafter. The best response achieved during the entire follow-up was recorded. IWG 2006 response assessment was completed by an independent review committee. IWG 2023 response assessment was conducted by two investigators separately.

To evaluate the effect of immortal time bias, we performed time-dependent and a modified 6-month landmark analyses. As patients in both VALIDATE as well as PANTHER could have received either HMA monotherapy or HMA-based combination treatments, we conducted additional sensitivity analyses limited to patients treated with HMA monotherapy only to confirm that IWG 2023 response criteria are valid in the context of HMA monotherapy, which remains the standard of care for patients with HR-MDS.

Composite CR (cCR) according to IWG 2006 criteria was defined as either CR or marrow CR (mCR). cCR as per the IWG 2023 response criteria is a combination of CR, CR-L, CRh, and CR equivalent (CRequ). CR-L was defined as a composite of CR with bilineage count recovery (CRbi) and CR with unilineage count recovery (CRuni). PB-CR was defined as achieving a hemoglobin of ≥ 11 g/dL, platelet count of $\geq 100 \times 10^9/L$, absolute neutrophil count of $\geq 1.0 \times 10^9/L$, and absence of peripheral blood blasts.

Statistical analysis:

Data for the VALIDATE cohort was collected in a centralized RedCap database and uploaded by accredited research coordinators at each site with validation by two investigators at the coordinating site (Yale School of Medicine).

Baseline patient and treatment characteristics were summarized according to their distribution. Binary and categorical data are reported as counts and percentage in each group with respect to each variable's total for which the data is available. Baseline comparison across the treatment groups was conducted using the chi-square test of proportions with continuity correction in the VALIDATE cohort or using the Fisher's exact test in the PANTHER cohort due to smaller cohort size. Continuous variables were mostly non-normally distributed, reported as median and inter-quartile range (IQR), and compared using the Kruskal-Wallis rank sum test with three groups, or using the Wilcoxon rank sum test with two groups.

All survival and time-to-event analyses were conducted using Kaplan-Meier estimator, proportional hazards, Cox models, and adjusted survival curves. A modified 6-month landmark cohort was also defined including patients alive at 6 months after treatment initiation (VALIDATE) or randomization (PANTHER). OS was calculated from the landmark time. Patients were categorized according to their overall best response during treatment or follow-up rather than whether the best response had been achieved by the landmark time. Accordingly, patients who ultimately achieved the response of interest after the landmark time were classified as responders. Survival was compared between response groups using Kaplan-Meier methods. OS was also estimated using the Mantel-Byar method with IWG 2023 response and allo-SCT as time-dependent variables and visualized in the Simon-Makuch plot. A Cox model was constructed

containing age at diagnosis, patient's sex, white blood cell count at diagnosis, IPSS-M baseline risk category (Very Low/Low/Moderate Low, Moderate High/High, Very High), initial therapy assignment (AZA, DEC, HMA-based combination) as well as allo-SCT treatment status (Yes vs. No) and IWG 2023 response as time-dependent variables for the VALIDATE cohort. Continuous age, sex, WBC at screening, IPSS-R risk category (intermediate and high vs. very high), time-varying IWG 2023 response (cCR vs. no cCR at the time of evaluation) and time-varying allo-SCT were included in the final Cox model for the PANTHER cohort. For VALIDATE, most of the relevant variables were part of the exclusion criteria and therefore all subjects included in the analysis have full data available. Some variables, namely molecular data, have some missing values, but IPSS-M scores were calculated based on established methods accounting for missing data.⁽¹⁾ Similarly, for missingness in the PANTHER data, the exclusion for IWG 2023 criteria is due to missingness in one or more of the IWG 2023 response criteria elements (blasts, blood counts, etc.) and patients with incomplete data were excluded. No imputation was performed for these patients and only complete case analysis was used during modeling. The level of significance was set at $\alpha = 0.05$. All analyses were conducted in R 4.5.2.

Supplemental Tables:

Supplemental Table 1: Study sites participating in VALIDATE

Site Name and PI
Yale University, New Haven, USA Amer Zeidan
Cleveland Clinic, Cleveland, USA Hetty Carraway
Dana Farber Cancer Institute, Boston, USA Maximilian Stahl
Johns Hopkins University, Baltimore, USA Amy Dezern
Massachusetts General Hospital, Boston, USA Andrew Brunner
Memorial Sloan Kettering Cancer Center, New York, USA Eytan Stein
Moffitt Cancer Center, Tampa, USA Rami Komrokji
Roswell Park, Buffalo, USA Elizabeth Griffiths
University of Colorado, Aurora, USA Christine McMahon
University of Miami, Miami, USA Mikkael A. Sekeres
Washington University, Saint Louis, USA Geoffrey Uy
Weill Cornell Medical Center, New York, USA Gail Roboz
University of North Carolina, Chapel Hill, USA Joshua Zeidner
Vanderbilt University, Nashville, USA Michael Savona
University of Texas Southwestern, Dallas, USA Yazan Madanat

Mayo Clinic Florida, Jacksonville, USA Talha Badar
Mayo Minnesota, Rochester, USA Aref Al-Kali
Northwestern University, Chicago, USA Yasmin Abaza
IRCCS Humanitas Research Hospital, Rozzano, Milano, Italy Matteo Della Porta
University of Florence, Florence, Italy Valeria Santini
MLL Munich Leukemia Laboratory, Munich, Germany Torsten Haferlach
Calvary Mater Hospital, Newcastle, New South Wales, Australia and Australasian Leukaemia and Lymphoma Group (ALLG) Anoop Enjeti
King Hussein Cancer Center, Amman, Jordan Kamal Al-Rabi
The Royal Adelaide Hospital, Adelaide, Australia Devendra Hiwase
University Hospital of Salamanca, Salamanca, Spain Maria Diez Campelo
University Hospital Vall d'Hebron, Barcelona, Spain David Valcarcel
Paracelsus Medical University, Salzburg, Austria Lisa Pleyer
Democritus University of Thrace Medical School, Alexandroupolis, Greece Ioannis Kotsianidis
Attikon University Hospital, National and Kapodistrian University of Athens, Athens, Greece Vassiliki Pappa
Kyoto University, Kyoto, Japan Seishi Ogawa
Institute of Medical Science, the University of Tokyo, Tokyo, Japan Yasuhito Nannya

Supplemental Table 2: Bivariate distribution of IWG 2006 and IWG 2023 criteria in VALIDATE

		IWG 2023										
IWG 2006	Best Response	CR	CRequ	CRbi	CRuni	CRh	PR	HI	No Response	PD	<2 Cycles	Total
	CR	144	12	0	0	0	5	12	1	0	0	172
	mCR with HI	12	1	50	43	1	0	29	1	0	0	137
	mCR without HI	5	1	42	59	4	0	6	75	17	0	209
	PR	0	0	0	0	0	3	0	0	0	0	3
	HI	1	0	2	0	0	3	98	1	0	0	105
	SD	0	0	2	1	1	0	0	204	3	0	211
	PD	0	0	0	0	0	0	4	6	179	0	189
	<2 Cycles	0	0	0	0	0	0	0	0	0	21	21
	Total	160	14	96	103	6	11	149	288	199	21	1,047

Supplemental Table 3: Bivariate distribution of IWG 2006 and IWG 2023 criteria in PANTHER

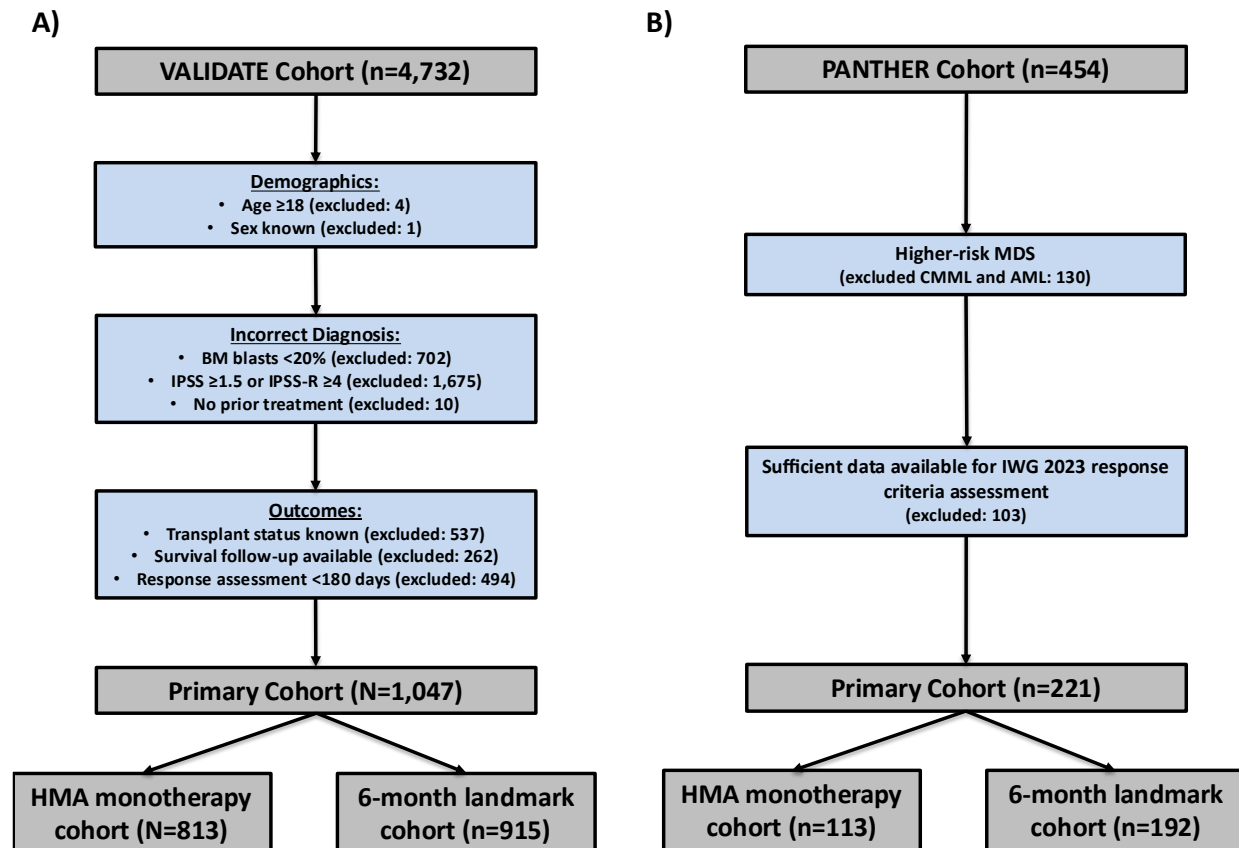
		IWG 2023							
		CR	CRbi	CRuni	CRh	PR	No Response	PD	Total
IWG 2006	Best Response								
	CR	50	7	0	0	1	4	1	63
	mCR	0	8	23	4	0	20	0	55
	HI	1	4	8	0	0	24	0	37
	SD	0	3	1	0	0	49	0	53
	PD	0	0	0	0	0	0	13	13
	Total	51	22	32	4	1	97	14	221

Supplemental Table 4: OS by IWG 2023 response by transplant status in VALIDATE:

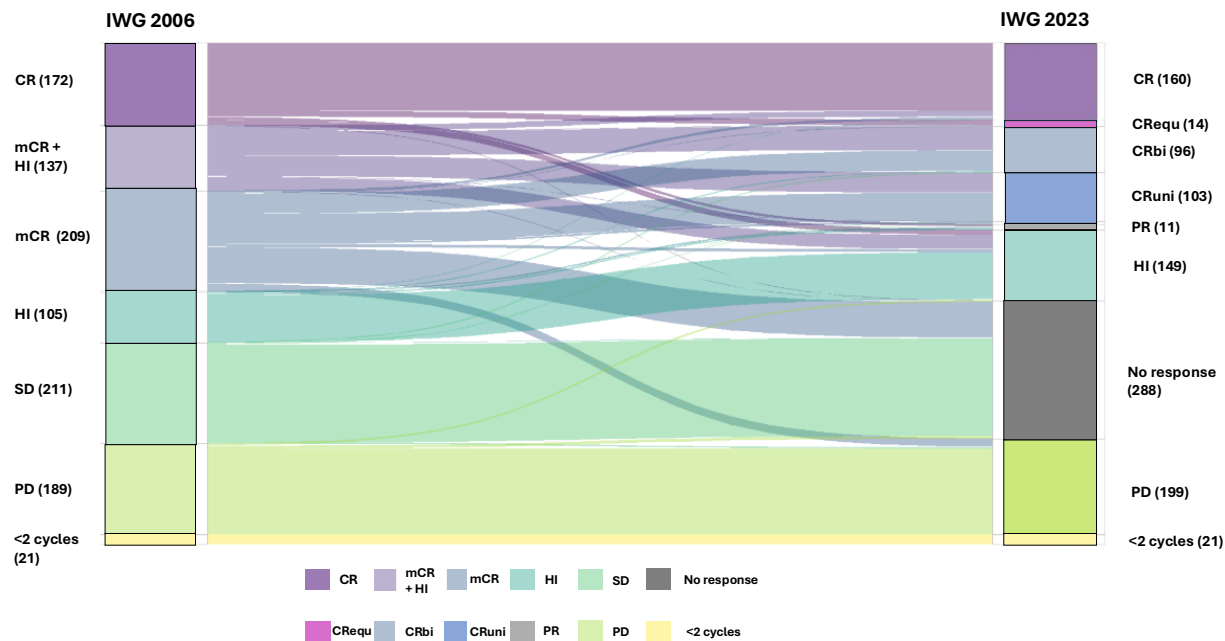
Overall Survival - All Patients (censored at transplant)				
Best Response	n (%)	Median ST (95% CI)	HR (95% CI)	p-value
CR	160 (15.30%)	26.38 (20.67; 35.78)	Ref.	
CRequ	14 (1.34%)	32.72 (22.74; NA)	0.891 (0.385; 2.066)	0.788
CRbi	96 (9.18%)	22.11 (17.22; 32.07)	1.190 (0.778; 1.822)	0.422
CRuni	103 (9.85%)	18.07 (15.11; 26.61)	1.461 (0.998; 2.139)	0.051
CRh	6 (0.57%)	24.51 (NA; NA)	0.703 (0.097; 5.077)	0.727
PR	11 (1.05%)	10.58 (8.77; NA)	3.694 (1.827; 7.471)	< 0.001
HI	148 (14.15%)	16.53 (13.67; 22.01)	1.718 (1.226; 2.405)	0.002
No Response	288 (27.53%)	18.99 (17.08; 22.90)	1.534 (1.133; 2.076)	0.006
PD	199 (19.02%)	11.79 (10.25; 14.72)	2.535 (1.862; 3.452)	< 0.001
<2 cycles	21 (2.01%)	5.85 (3.91; NA)	8.171 (4.361; 15.309)	< 0.001
Overall	1046 (100%)	17.45 (16.13; 19.25)		< 0.001
Overall Survival - Non-transplant Patients				
Best Response	n (%)	Median ST (95% CI)	HR (95% CI)	p-value
CR	84 (12.69%)	25.46 (18.92; 34.40)	Ref.	
CRequ	8 (1.21%)	32.72 (22.74; NA)	0.834 (0.360; 1.933)	0.673
CRbi	44 (6.65%)	21.32 (13.11; 29.37)	1.252 (0.818; 1.917)	0.300
CRuni	64 (9.67%)	15.51 (12.19; 22.01)	1.515 (1.035; 2.219)	0.033
CRh	3 (0.45%)	24.51 (NA; NA)	0.714 (0.099; 5.158)	0.739
PR	9 (1.36%)	10.58 (8.77; NA)	3.146 (1.555; 6.366)	0.001
HI	96 (14.50%)	14.55 (12.45; 19.35)	1.670 (1.192; 2.340)	0.003
No Response	191 (28.85%)	16.99 (13.70; 20.63)	1.552 (1.147; 2.102)	0.004
PD	151 (22.81%)	10.38 (8.64; 13.01)	2.409 (1.768; 3.283)	< 0.001
<2 cycles	12 (1.81%)	4.53 (2.99; NA)	7.993 (4.268; 14.971)	< 0.001
Overall	662 (100%)	15.31 (13.93; 16.92)		< 0.001
Overall Survival - Transplant Patients				
Best Response	n (%)	Median ST (95% CI)	HR (95% CI)	p-value
CR	76 (19.74%)	65.02 (31.11; NA)	Ref.	
CRequ	6 (1.56%)	9.45 (8.64; NA)	6.516 (2.717; 15.628)	< 0.001
CRbi	52 (13.51%)	36.21 (26.42; NA)	1.166 (0.702; 1.937)	0.553
CRuni	39 (10.13%)	48.36 (21.68; NA)	0.984 (0.562; 1.721)	0.954
CRh	3 (0.78%)	19.25 (19.25; NA)	1.047 (0.143; 7.650)	0.964
PR	2 (0.52%)	NA (NA; NA)	0.000 (0.000; Inf)	0.992
HI	53 (13.77%)	27.73 (16.89; NA)	1.549 (0.959; 2.503)	0.074
No Response	97 (25.19%)	41.10 (25.96; NA)	1.106 (0.717; 1.704)	0.649
PD	48 (12.47%)	37.62 (20.44; NA)	1.201 (0.725; 1.988)	0.476
<2 cycles	9 (2.34%)	15.18 (13.27; NA)	2.790 (1.292; 6.027)	0.009
Overall	385 (100%)	37.16 (29.96; 47.05)		< 0.001

Supplemental Figure Legends:

Supplemental Figure 1: Consort diagram of the patient population analyzed

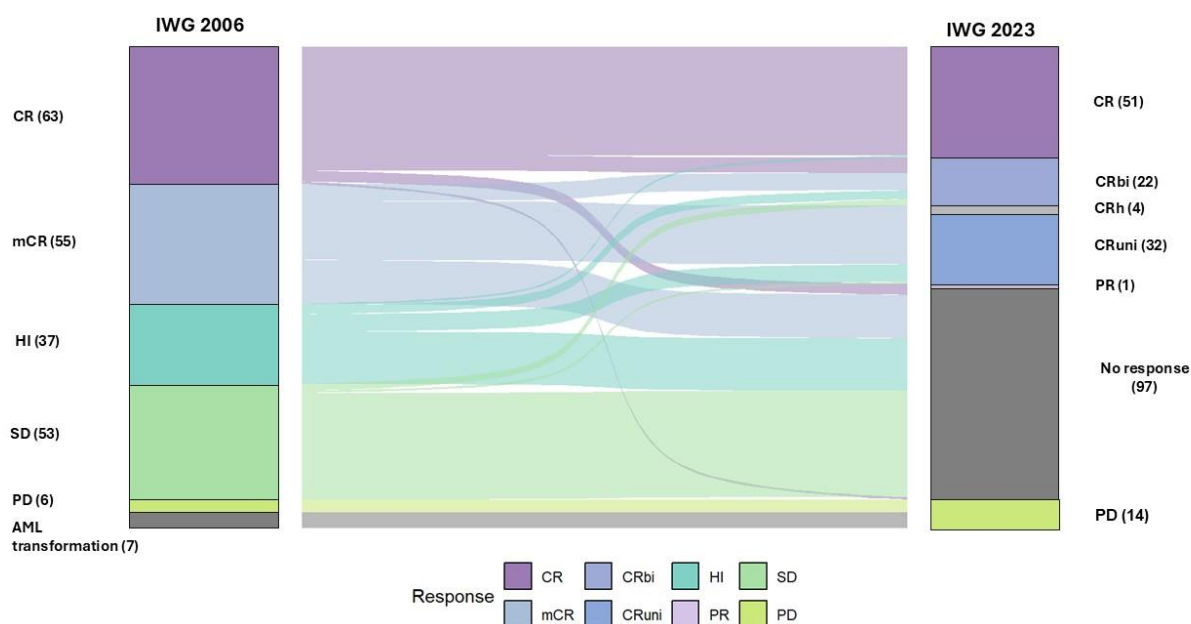


Supplemental Figure 2: Reclassification from IWG 2006 to IWG 2023 response criteria from VALIDATE



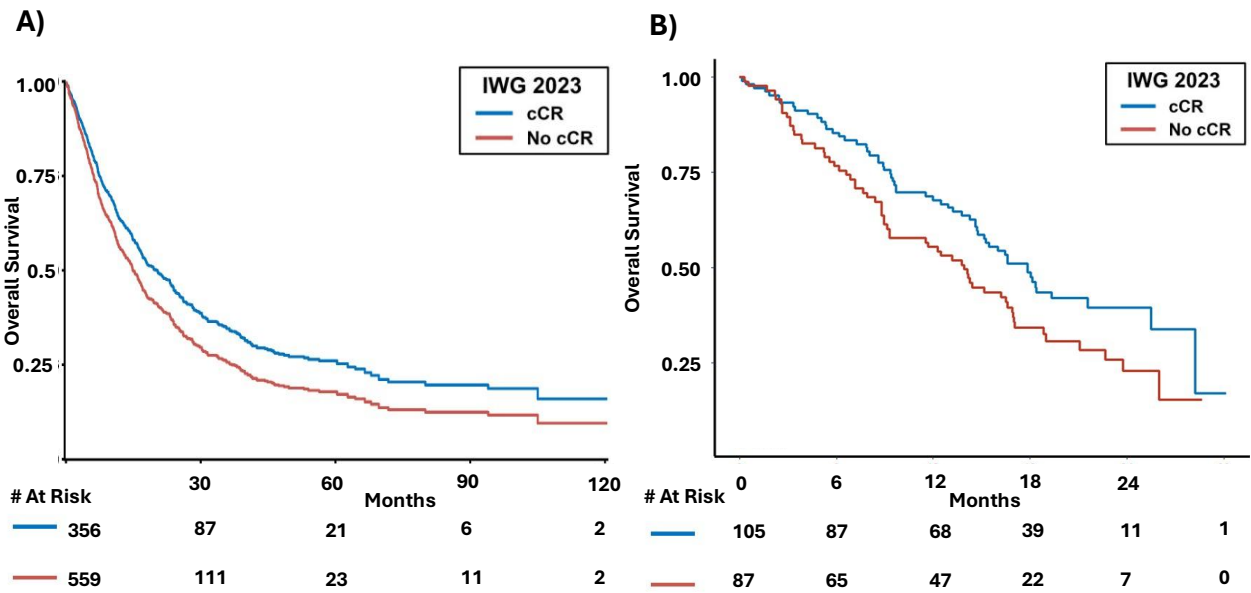
Supplemental Figure 2: Reclassification of responses from IWG 2006 to IWG 2023 response criteria for the VALIDATE dataset is shown as an alluvial plot.

Supplemental Figure 3: Reclassification from IWG 2006 to IWG 2023 response criteria from PANTHER



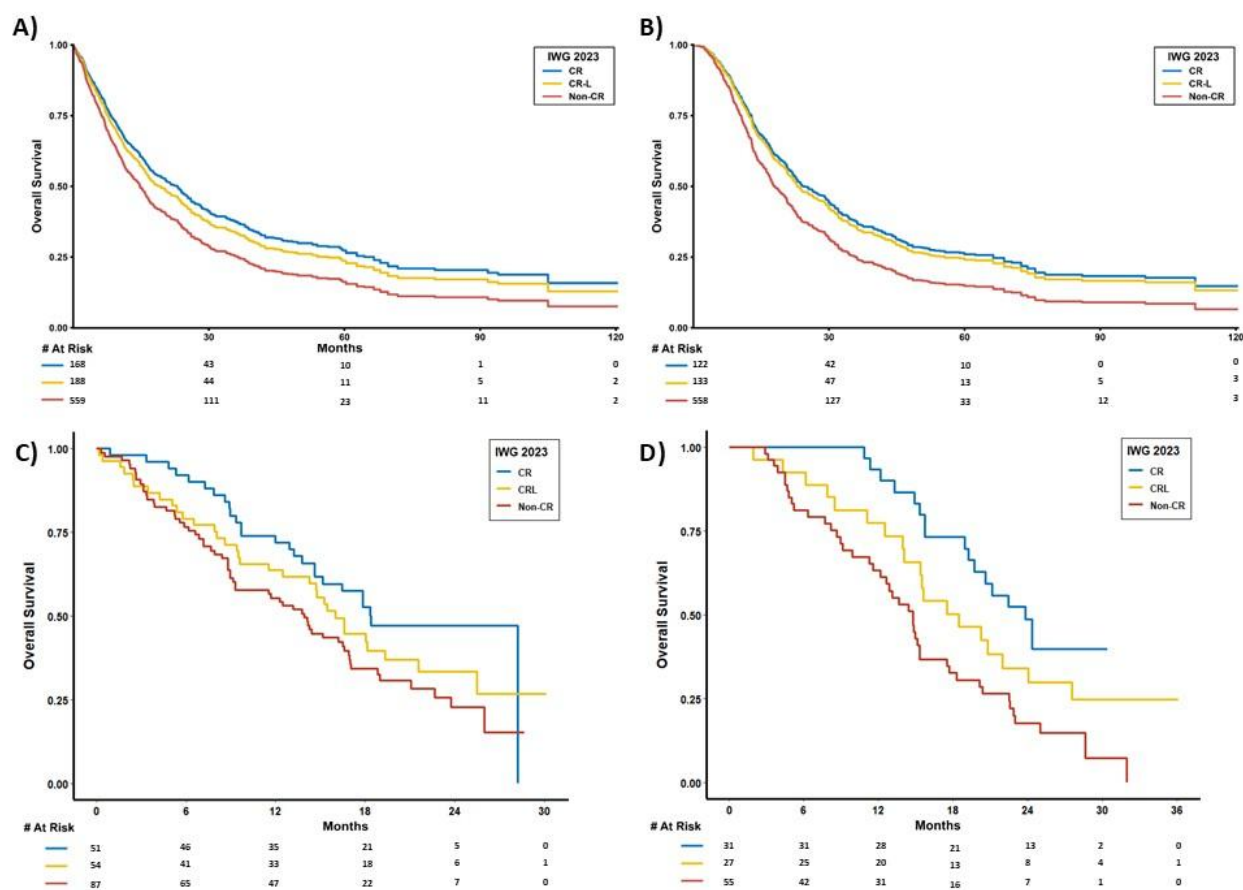
Supplemental Figure 3: Reclassification of responses from IWG 2006 to IWG 2023 response criteria for the PANTHER dataset is shown as an Alluvial plot.

Supplemental Figure 4: Modified landmark analysis of association between cCR and OS in VALIDATE and PANTHER.



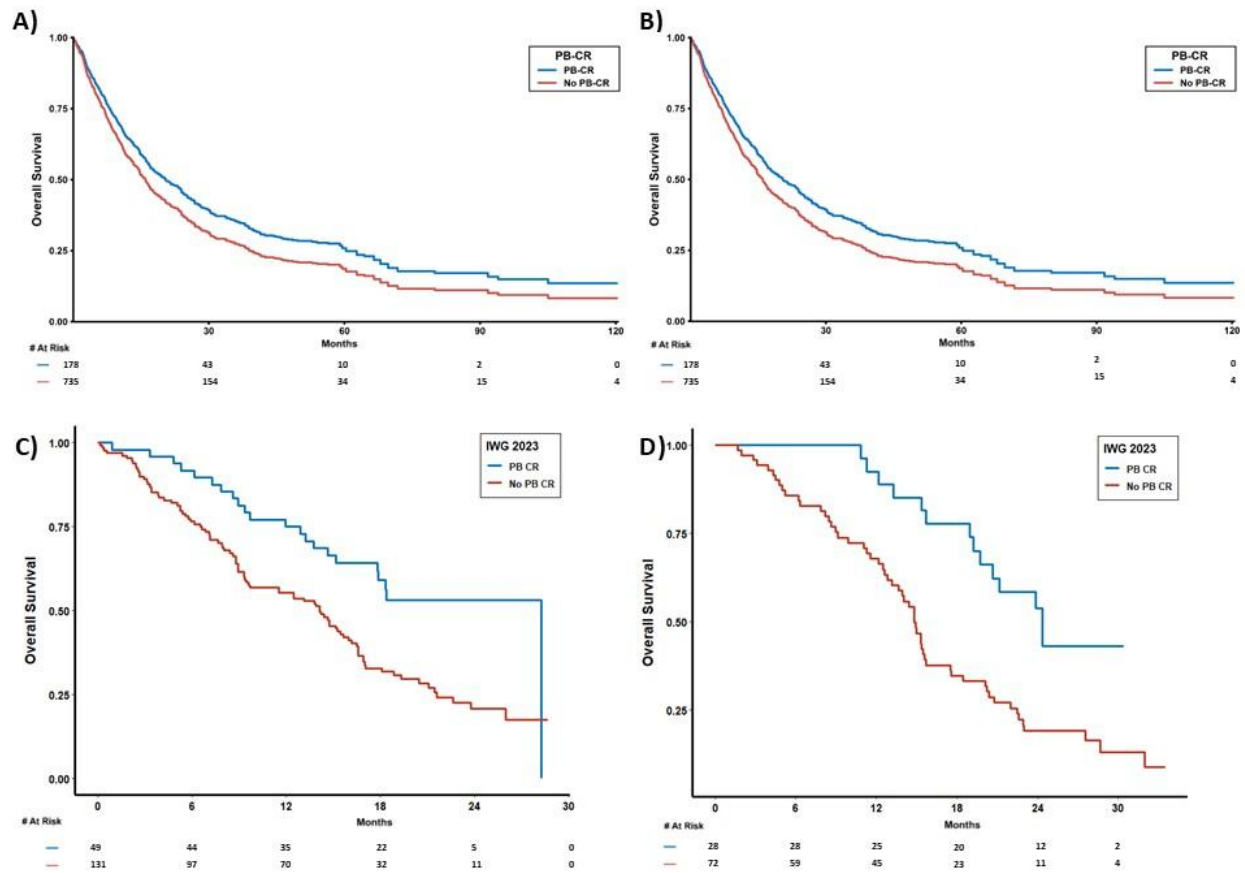
Supplemental Figure 4: Modified landmark analysis of association between cCR and OS in VALIDATE and PANTHER. Supplemental Figure 4 shows the association between OS from date of HMA initiation (VALIDATE; panel A) or randomization (PANTHER; panel B) by cCR status in a modified 6-month landmark analysis.

Supplemental Figure 5: Comparison of CR vs. CR-L no CR/CRL



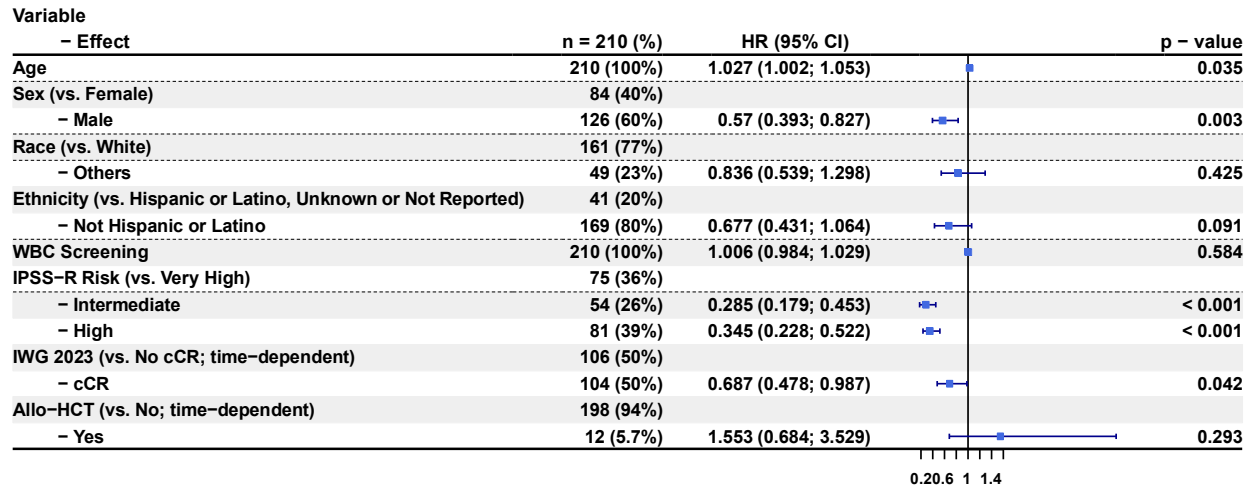
Supplemental Figure 5: Comparison of CR vs. CR-L no CR/CRL. Supplemental Figure 5 shows comparisons of different components of the cCR endpoint (i.e., CR and CR-L) as well as no CR/CR-L for different analysis cohorts: A: VALIDATE all patients with modified 6-month landmark; B: VALIDATE HMA monotherapy only without landmark; C: PANTHER all patients with modified 6-month landmark; D: PANTHER HMA monotherapy only without landmark

Supplemental Figure 6: PB CR vs. NO PB CR



Supplemental Figure 6: Comparison of PB-CR vs. no PB-CR for different analysis cohorts: A: VALIDATE all patients with modified 6-month landmark; B: VALIDATE HMA monotherapy only without landmark; C: PANTHER all patients with modified 6-month landmark; D: PANTHER HMA monotherapy only without landmark

Supplemental Figure 7: Multivariable model of predictors of OS with inclusion of allo-SCT – PANTHER dataset



Supplemental Figure 7: Multivariable model of predictors of OS with inclusion of allo-SCT – PANTHER dataset. Supplemental Figure 7 shows a Cox proportional hazards model of predictors OS for PANTHER cohort with inclusion of all-SCT. To reduce the impact of immortal time bias, analyses were conducted treating both response and allo-SCT as time-dependent variables in the model.

References:

1. Bernard E, Tuechler H, Greenberg PL, et al. Molecular International Prognostic Scoring System for Myelodysplastic Syndromes. *NEJM Evidence*. 2022;1(7):EVIDoa2200008. doi:doi:10.1056/EVIDoa2200008.
2. Greenberg P, Cox C, LeBeau MM, et al. International scoring system for evaluating prognosis in myelodysplastic syndromes. *Blood*. 1997 Mar 15;89(6):2079-88.
3. Greenberg PL, Tuechler H, Schanz J, et al. Revised international prognostic scoring system for myelodysplastic syndromes. *Blood*. 2012 Sep 20;120(12):2454-65.
4. Bernard E, Nannya Y, Hasserjian RP, et al. Implications of TP53 allelic state for genome stability, clinical presentation and outcomes in myelodysplastic syndromes. *Nat Med*. 2020 Oct;26(10):1549-1556.
5. Adès L, Girshova L, Doronin VA, et al. Pevonedistat plus azacitidine vs azacitidine alone in higher-risk MDS/chronic myelomonocytic leukemia or low-blast-percentage AML. *Blood Adv*. 2022 Sep 13;6(17):5132-5145.
6. Cheson BD, Greenberg PL, Bennett JM, et al. Clinical application and proposal for modification of the International Working Group (IWG) response criteria in myelodysplasia. *Blood*. 2006 Jul 15;108(2):419-25.
7. Zeidan AM, Platzbecker U, Bewersdorf JP, et al. Consensus proposal for revised International Working Group response criteria for higher risk myelodysplastic syndromes. *Blood*. 2023. blood.2022018604. doi:10.1182/blood.2022018604.
8. Pleyer L, Vaisband M, Drost M, et al. Cox proportional hazards deep neural network identifies peripheral blood complete remission to be at least equivalent to morphologic complete remission in predicting outcomes of patients treated with azacitidine-A prospective cohort study by the AGMT. *Am J Hematol*. 2023 Nov;98(11):1685-1698.
9. Kantarjian H, Issa JP, Rosenfeld CS, et al. Decitabine improves patient outcomes in myelodysplastic syndromes: results of a phase III randomized study. *Cancer*. 2006 Apr 15;106(8):1794-803.

10. Silverman LR, Fenaux P, Mufti GJ, et al. Continued azacitidine therapy beyond time of first response improves quality of response in patients with higher-risk myelodysplastic syndromes. *Cancer*. 2011 Jun 15;117(12):2697-702.