

Guideline-based indicators for adult patients with myelodysplastic syndromes

Kristina Stojkov,^{1,2,*} Tobias Silzle,^{3,*} Georg Stussi,⁴ David Schwappach,^{5,6} Juerg Bernhard,⁷ David Bowen,⁸ Jaroslav Čermák,⁹ Avinash G. Dinmohamed,¹⁰⁻¹² Corien Eeltink,¹² Sabrina Eggmann,¹³ Pierre Fenaux,¹⁴ Ulrich Germing,¹⁵ Manuel Haschke,^{16,17} Eva Hellstrom-Lindberg,¹⁸ Monika Heger,^{1,7} Arjan A. van de Loosdrecht,¹² Jakob Passweg,¹⁹ Michael Pfeilstöcker,²⁰ Uwe Platzbecker,^{21,22} Luca Malcovati,^{23,24} António Medina de Almeida,²⁵ Moshe Mittelman,²⁶ Christine Morgenthaler,^{1,7} David P. Steensma,²⁷ Valeria Santini,^{28,29} Reinhard Stauder,³⁰ Argiris Symeonidis,³¹ Sämi Schär,³² Charlotte Maddox,³² Theo de Witte,³³ Julia Bohlius,⁶ and Nicolas Bonadies,^{1,2} on behalf of the Swiss MDS Study Group and the Swiss Group of Clinical Cancer Research (SAKK)

¹Department of Hematology and Central Hematology Laboratory, Inselspital/Bern University Hospital, and ²Department for BioMedical Research, University of Bern, Bern, Switzerland; ³Department of Medical Oncology and Hematology, Cantonal Hospital St. Gallen, St. Gallen, Switzerland; ⁴Clinic of Hematology, Oncology Institute of Southern Switzerland, Bellinzona, Switzerland; ⁵Swiss Patient Safety Foundation, Zurich, Switzerland; ⁶Institute of Social and Preventive Medicine (ISPM) and ⁷Department of Medical Oncology, Inselspital/Bern University Hospital, University of Bern, Bern, Switzerland; ⁸Department of Hematology, Leeds Teaching Hospitals, Leeds, United Kingdom; ⁹Department of Clinical Hematology, Institute of Hematology and Blood Transfusion, Prague, Czech Republic; ¹⁰Department of Research and Development, Netherlands Comprehensive Cancer Organisation (IKNL), Utrecht, The Netherlands; ¹¹Department of Public Health, Erasmus University Medical Center, Rotterdam, The Netherlands; ¹²Amsterdam University Medical Center, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; ¹³Department of Physiotherapy, Bern University Hospital, Bern, Switzerland; ¹⁴Hôpital St. Louis, Assistance Publique-Hôpitaux de Paris/Paris University, Paris, France; ¹⁵Department of Hematology, Oncology and Clinical Immunology, Heinrich Heine University, Düsseldorf, Germany; ¹⁶Clinical Pharmacology and Toxicology, Department of General Internal Medicine, Bern University Hospital, and ¹⁷Institute of Pharmacology, University of Bern, Bern, Switzerland; ¹⁸Center for Hematology and Regenerative Medicine, Department of Medicine, Karolinska University Hospital, Huddinge, Stockholm, Sweden; ¹⁹Department of Hematology, University Hospital Basel, Basel, Switzerland; ²⁰3rd Medical Department, Hanusch Hospital, Vienna, Austria; ²¹Medical Clinic and ²²Policlinic 1, Hematology and Cellular Therapy, Leipzig University Hospital, Leipzig, Germany; ²³Department of Molecular Medicine, University of Pavia, Pavia, Italy; ²⁴Hematology Oncology, Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS) S. Matteo Hospital, Pavia, Italy; ²⁵Hospital da Luz, Lisboa, Portugal; ²⁶Department of Medicine A, Tel Aviv Sourasky (Ichilov) Medical Center and Sackler Medical Faculty, Tel Aviv University, Tel Aviv-Yafo, Israel; ²⁷Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA; ²⁸MDS Unit, Hematology, Azienda Ospedaliero Universitaria Careggi, University of Florence, Florence, Italy; ²⁹Fondazione Italiana Sindromi Mielodisplastiche (FISiM), Florence, Italy; ³⁰Department of Internal Medicine V (Hematology and Oncology), Innsbruck Medical University, Innsbruck, Austria; ³¹Division of Hematology, Department of Internal Medicine, Medical School, University of Patras, Patras, Greece; ³²SAKK Coordinating Center, Bern, Switzerland; and ³³Department of Tumor Immunology, Nijmegen Center for Molecular Life Sciences, Radboud University Medical Centre, Nijmegen, The Netherlands

Key Points

- GBIs are measurable elements for quality of care and are currently lacking for adult MDS patients.
- We developed a GBI consensus for the domains of diagnosis (n = 14), therapy (n = 8), and provider/infrastructural characteristics (n = 7).

Myelodysplastic syndromes (MDSs) represent a heterogeneous group of hematological stem cell disorders with an increasing burden on health care systems. Evidence-based MDS guidelines and recommendations (G/Rs) are published but do not necessarily translate into better quality of care if adherence is not maintained in daily clinical practice. Guideline-based indicators (GBIs) are measurable elements for the standardized assessment of quality of care and, thus far, have not been developed for adult MDS patients. To this end, we screened relevant G/Rs published between 1999 and 2018 and aggregated all available information as candidate GBIs into a formalized handbook as the basis for the subsequent consensus rating procedure. An international multidisciplinary expert panel group (EPG) of acknowledged MDS experts (n = 17), health professionals (n = 7), and patient advocates (n = 5) was appointed. The EPG feedback rates for the first and second round were 82% (23 of 28) and 96% (26 of 27), respectively. A final set of 29 GBIs for the 3 domains of diagnosis (n = 14), therapy (n = 8), and provider/infrastructural characteristics (n = 7) achieved the predefined agreement score for selection (>70%). We identified shortcomings in standardization of

Submitted 14 May 2020; accepted 29 June 2020; published online 25 August 2020.
DOI 10.1182/bloodadvances.2020002314.

*K.S. and T.S. contributed equally to this work.

Preliminary data presented in abstract form at the 61st annual meeting of the American Society of Hematology, Orlando, FL, 7-10 December 2019 (abstract 4752) and 25th annual meeting of the European Hematology Association, Frankfurt, Germany, virtual edition 11-21 June 2020 (abstract EP808).

Data-sharing requests may be e-mailed to the corresponding author, Nicolas Bonadies, at nicolas.bonadies@insel.ch.

The full-text version of this article contains a data supplement.

© 2020 by The American Society of Hematology

patient-reported outcomes, toxicity, and geriatric assessments that need to be optimized in the future. Our GBIs represent the first comprehensive consensus on measurable elements addressing best practice performance, outcomes, and structural resources. They can be used as a standardized instrument with the goal of assessing, comparing, and fostering good quality of care within clinical development cycles in the daily care of adult MDS patients.

Introduction

Publishing guidelines and recommendations (G/Rs) does not necessarily translate into better quality of care if they are not implemented and adherence is not maintained in daily clinical practice.^{1,2} Adherence to G/Rs is considered a cornerstone for appropriate health service with proven benefit for patients and health care systems.³⁻⁵ Moreover, appropriate care is fundamental to counteract potential negative consequences for patients and society in economically driven health care systems.⁶ To this end, standardized assessment tools for practice performance, outcomes, and structural resources are required to measure, compare, and improve quality in health care systems using clinical development cycles.^{7,8} However, this is not systematically done in daily care, mainly due to a lack of consensus on clinically relevant and measurable elements for quality of care. Hence, quality indicators are being developed for a variety of malignancies in regional and national quality assurance consensus programs such as the Quality Oncology Practice Initiative (QOPI) program initiated by the American Society of Clinical Oncology (ASCO).^{9,10} These programs focus mainly on solid tumors or lymphoma,¹¹⁻¹⁴ but quality indicators addressing hematological neoplasms, such as myelodysplastic syndromes (MDSs), have thus far not been developed.

MDSs represent a heterogeneous group of hematological stem cell neoplasms with a variable risk of transformation into secondary acute myeloid leukemia (AML). MDSs are characterized by cytopenia, dysplasia, and inflammation and occur in all age groups with a median age of onset above 70 years.^{15,16} Due to the aging of the general population, improvements in diagnostics, and increasing survivorship of other cancers, resulting in populations at higher risks for secondary malignancies, MDS is an emerging disease with an increasing burden on health care systems.^{17,18} The heterogeneity of the disease, multimorbidity of the elderly patients, and increasing diagnostic as well as therapeutic options put patients and health care systems at risk for inappropriate use of medical interventions, endangering good quality of care. Several national and international evidence-based guidelines for diagnosis, prognosis, and therapy of adult MDSs have been published to assist and legitimize clinicians in their daily decision-making. Retrospective studies have identified shortcomings in the diagnostic workup and treatment of MDS patients, but the degree of adherence, the reasons for nonadherence and the impact on outcomes remain unknown.¹⁹

Quality indicators are defined as “measurable elements of practice performance for which there is evidence or consensus that they can be used to assess and change the quality of care provided.”^{3(p104)} Quality indicators are usually extracted from published, evidence-based guidelines by a structured selection and consensus process and are referred to as guideline-based indicators (GBIs).^{20,21} GBIs have to be relevant, understandable, measurable, behavior oriented, and attainable (RUMBA criteria) and allow process-based quality of care with impact on patient-centered outcomes to be

operationalized.²² GBIs have not been developed for patients with hematological neoplasms, including MDSs. This lack impedes a standardized and systematic assessment of appropriateness and quality of care using clinical development cycles.⁷ Motivated by the lack of measurable elements for quality of care, we performed a structured consensus process to define a final set of clinically relevant GBIs that can be implemented in the daily care of adult MDS patients.

Methods

We applied a modified Research and Development Corporation (RAND) technique with a 2-step DELPHI rating process (first round, October 2018 to January 2019; second round, March 2019 to June 2019) to find an expert consensus on clinically relevant GBIs (supplemental Figure 1).²³⁻²⁵

Screening of published G/Rs

The core working group (K.S., T.S., J. Bohlius, G.S., D.S., N.B.) identified and screened evidence-based G/Rs and extracted all relevant information (supplemental Table 1).²⁶ With regard to MDS diagnosis and therapy, we retrieved all information for potential GBIs from 11 clinically relevant G/Rs published by national and international MDS working groups between 2011 and 2018 (Table 1 G/Rs from clinical MDS working groups [n = 11]). Because aspects of provider/infrastructural characteristics were not sufficiently addressed in clinical MDS G/Rs, all information for this domain was extracted from an additional 6 guiding documents published by national and international cancer care certification and accreditation boards between 1999 and 2018 (Table 1 G/Rs from cancer care certification/accreditation programs [n = 6]).

Designation of multidisciplinary EPG members

The multidisciplinary expert panel group (EPG) included 28 and 27 members for the first and second rating rounds, respectively, and comprised acknowledged MDS experts (hematologists), health professionals with hemato-oncological expertise in their field (nursing, pharmacology, physiotherapy, psycho-oncology, epidemiology), and MDS patient advocates (Table 2). Experts were identified by the core project working group upon recommendation from chairs of medical or other health professional societies and were selected according to modified European Consortium for Accreditation criteria (supplemental Table 2).²⁷ Each EPG member signed a participation agreement and disclosed any potential conflicts of interest for the rating of candidate GBIs.

Elaboration, aggregation, and structured representation of candidate GBIs

The core working group elaborated, aggregated, and structured all relevant information extracted from G/Rs as candidate GBIs within

Table 1. Guidelines and recommendations screened

Year	Countries	Institution	Links and references
G/Rs from clinical MDS working groups, n = 11			
Annual updates	USA	NCCN	https://www.nccn.org/professionals/physician_gls/default.aspx#site
2018	Germany, Austria, Switzerland	DGHO, OeGHO, SGH-SHH	https://www.onkopedia.com/de/onkopedia/guidelines/myelodysplastische-syndrome-mds/@@guideline/html/index.html
2017	Scandinavia	NMDSG	https://www.nmds.org/index.php/guidelines
2015	France	GFM	http://www.gfmgroup.org/recommandations.php
2017	Europe	EBMT, ELN, BMT-CTN	Allo-HSCT for MDS and CMML: recommendations from an international expert panel ⁵⁰
2014	Europe	ESMO	European Society for Medical Oncology MDS guidelines ⁵¹
2013	Europe	ELN	European LeukemiaNet MDS guidelines ⁵²
2013	The Netherlands/Belgium	HOVON	https://hematologienederland.nl/kwaliteit/richtlijnen/
2013	UK	BCSH	British Committee for Standards in Haematology MDS guidelines ⁵³
2012	Spain	GESMD	https://www.gesmd.es/
2011	Italy	FISIM	https://www.fismonlus.it/en/about-fism/
G/Rs from cancer care certification/ accreditation programs, n = 6			
Annual updates	UK	NICE	https://pathways.nice.org.uk/pathways/blood-and-bone-marrow-cancers#path=view%3A/pathways/blood-and-bone-marrow-cancers/blood-and-bone-marrow-cancers-overview.xml&content=view-index
2018	Germany	DKG	https://www.krebsgesellschaft.de/deutsche-krebsgesellschaft-wtr/german-krebsgesellschaft/zertifizierung/erhebungsboegen/onkologische-zentren.html
2017	USA	ASCO/QOPI	https://practice.asco.org/quality-improvement/quality-programs/qopi-certification-program
2016	USA	ACS/CoC	https://www.facs.org/quality-programs/cancer/coc/standards
2013	Switzerland	SCN	https://www.sgmo.ch/qualitaetsicherung/dokumentation/
1999	Europe	OECI/ACOE	http://www.acoe.be/About-ACOE/Mission-and-Objectives

ACOE, Accreditation Council for Oncology in Europe; ACS/CoC, American College of Surgeons/Commission on Cancers; BCSH, British Committee for Standards in Haematology; BMT-CTN, Blood and Marrow Transplant Clinical Trials Network; CMML, Chronic myelomonocytic leukemia; DGHO, Deutsche Gesellschaft für Hämatologie und Onkologie; DKG, Deutsche Krebs Gesellschaft; EBMT, European Society for Blood and Marrow Transplantation; ELN, European Leukemia Net; ESMO, European Society of Medical Oncology; FISIM, Fondazione Italiana Sindromi Mielodysplastice; GESMD, Grupo Espanol De Sindromes Mielodisplasticos; GFM, Groupe Francophone des Myelodysplasies; HOVON, Hemato Oncology Foundation for Adults in The Netherlands; NCCN, National Comprehensive Cancer Network; NICE, National Institute for Health and Care; NMDSG, Nordic MDS Study Group; OECI, Organization of European Cancer Institutes; OeGHO, Österreichische Gesellschaft für Hämatologie und Onkologie; SCN, Swiss Cancer Network; SGH-SHH, Swiss Society for Hematology; UK, United Kingdom; USA, United States of America.

a formalized handbook (supplemental Handbook). This book represented the basis for the subsequent DELPHI rating procedure and included information on scope, aims, and applied methodologies. Candidate GBIs contained basic information (domain/subdomain, aims, grade of recommendation), definitions (numerator, denominator, relevant time points for assessments, source) and supplementary information (rationale, requirements, and limitations to use) (supplemental Table 3). Depending on the rating performance and feedback from EPG members in the first rating round, candidate GBIs were adapted, merged, or dropped for the second round. All relevant information, including performance, feedback from EPG members, and comments from the core project working group, was reported back to the panelists.

Rating process

Each candidate GBI was rated independently in a DELPHI process by each EPG member using a secured web-based survey.²⁸ Access was granted via a Web link distributed individually to each EPG member. Rating was performed according to the RUMBA criteria²²:

- Relevance, the indicator is relevant and represents key aspects of the quality of MDS patient management;
- Understandability, the indicator is clear, precise, and comprehensively defined;
- Measurability, data collection concerning the indicator in real-life setting is possible and does not require excessive effort;
- Behavioral oriented, the indicator concerns an aspect of MDS patient management that can be targeted by service providers at the personal or institutional level; and
- Attainability, the indicator's goal is realistic and achievable under given conditions.

A 9-point Likert-like scale was used for rating, yielding a maximum score of 45 points per GBI. Each RUMBA criterion could be rated from 1 (do not agree) to 9 (totally agree), where tertiles 1 to 3 represent "no agreement," 4 to 6 "unclear," and 7 to 9 "agreement." Individual ratings were kept anonymous to all others, but all panelists received the aggregated information for each GBI after the first and second rating round. Nonexperts in a given field could

Table 2. Composition of EPG for the 2 DELPHI rating rounds

Expert panel group	First round		Second round	
	Invited	Participated	Invited	Participated
Hematologists	16	14*	17	16
Health professionals	7	6	7	7
Nurses	3	3	3	3
Pharmacologist	1	1	1	1
Physiotherapist	1	1	1	1
Psycho-oncologist	1	1	1	1
Epidemiologist	1	0	1	1
Patient advocates	5	3	3	3
Total	28	23	27	26
Feedback rate, %	82		96	

Represented countries: Hematologists: Austria (2); Czech Republic (1); France (1); Germany (2); United Kingdom (1); Greece (1); Israel (1); Italy (2); The Netherlands (2); Portugal (1); Sweden (1); Switzerland (1); United States (1). Health professionals: The Netherlands (2), Switzerland (5). Patient advocates: Denmark (1); Finland (1); Germany (1); The Netherlands (1); United States (1).

*One with partial completion of the survey.

skip the rating of candidate GBIs and free text boxes were offered for general feedback. Additional subsurveys were included in unclear situations and used to refine the definitions of GBIs.

Selection of GBIs

We used the following metrics for the evaluation of the rating performance of each candidate GBI: (i) median score from all RUMBA criteria; (ii) percentage from the maximum possible score (% max score), representing the percentage of the sum of scores divided by the maximum possible score; and (iii) agreement score defined as the percentage of all numbers of ratings in tertile 7 to 9 ("agreement") divided by the total number of ratings. Scores were calculated for the overall rating performance (all RUMBA criteria from all experts) as well as for each RUMBA item, MDS experts, and non-MDS experts separately. In the first round, candidate GBIs were ranked according to the less stringent criteria (i) and (ii), keeping those with a median score ≥ 7 and percentage from the maximum possible score $\geq 75\%$. In the second round, candidate GBIs were ranked according to the more stringent criteria (iii) (agreement score, tertile 7-9). GBIs from the diagnosis and therapy domains with agreement scores $\geq 70\%$ and GBIs from the provider/infrastructural characteristics domain with agreement scores $\geq 85\%$ were selected for the final set of GBIs. Expert rating disagreement was calculated as the difference of the agreement scores between MDS and non-MDS experts, with a difference $> 30\%$ defined as relevant. To clarify definitions of some GBIs, we used additional subsurveys, where we considered all ratings $< 70\%$ as not sufficiently relevant. We presented and discussed the final set of GBIs at the European LeukemiaNet (ELN) annual meeting of the MDS study group in Amsterdam on 11 October 2019.

Definitions for the use of GBIs

Definitions for the use of GBIs are summarized in supplemental Table 4. We define disease-based risk according to the International Prognostic Scoring System (IPSS) and revised IPSS (IPSS-R); patients with IPSS-R intermediate are allocated to lower risk if high-risk features are absent.²⁹⁻³⁴ Patient-based risk is determined by

the Hematopoietic Cell Transplantation Comorbidity Index (HCT-CI), as it is most frequently used in clinical trials.³⁵

Results

General aspects of the GBI selection process

The core working group extracted, elaborated, and aggregated candidate GBIs from G/Rs in a formalized handbook. GBIs were grouped in 3 main domains and assigned to appropriate subdomains. This process resulted in 61 candidate GBIs for diagnosis ($n = 23$), therapy ($n = 21$), and provider/infrastructural characteristics ($n = 17$), which entered the DELPHI rating procedure. The complete GBI selection process is summarized in Figure 1. The feedback rates from the multidisciplinary EPG members for the first and second DELPHI rounds were 82% (23 of 28) and 96% (26 of 27), respectively (Table 2). Feedback from non-MDS experts was mainly given for the GBIs of provider/infrastructural characteristics as well as GBIs for patient-reported outcomes (PROs) and toxicity assessment (TA) (range, 50% to 90%) (Table 3). Twenty-one candidate GBIs did not meet the less-stringent selection criteria for the first rating round. Three GBIs were rechallenged and 2 merged, leaving 42 candidate GBIs for the second rating round.

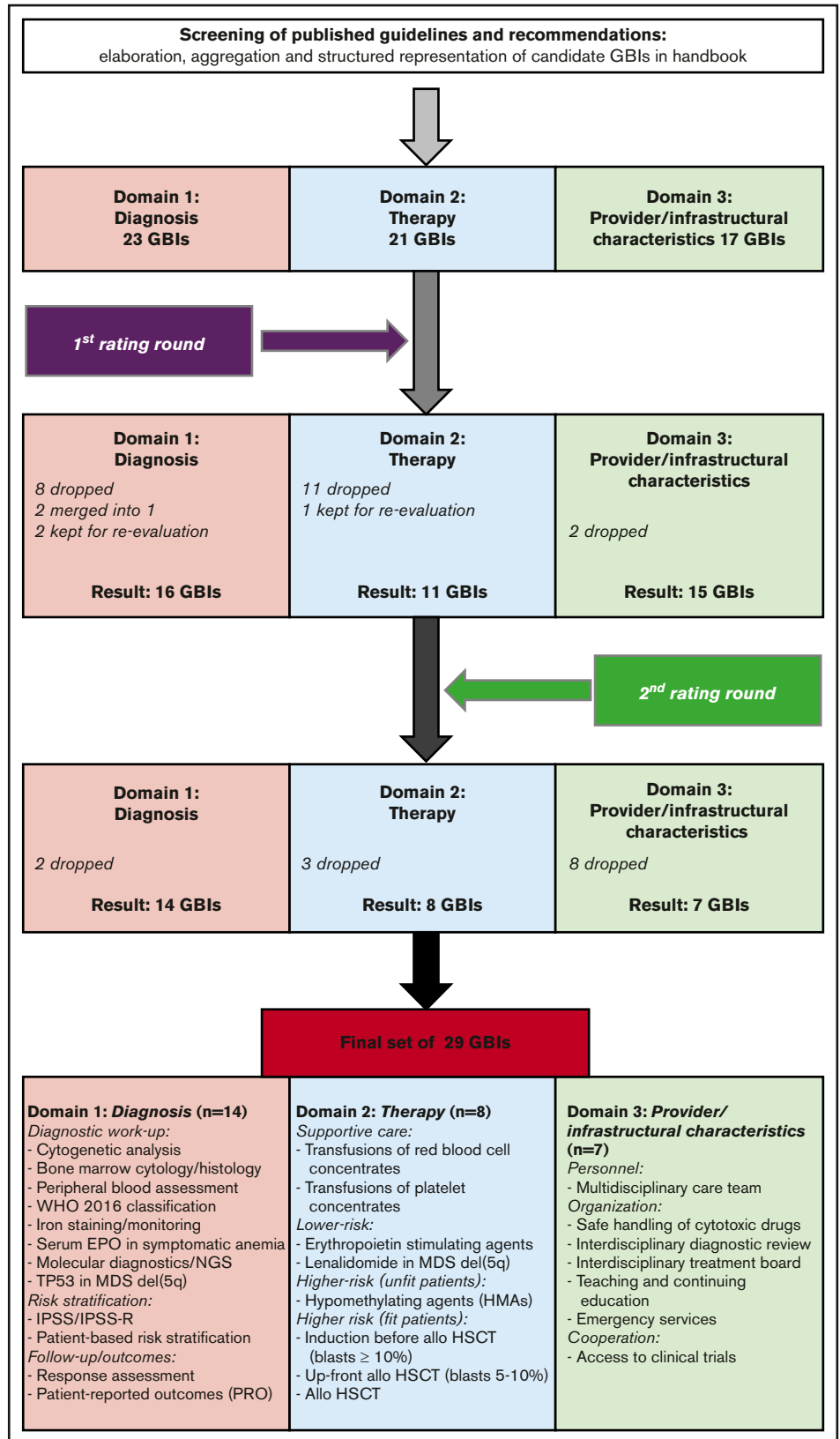
Treatments can be inappropriate either by underuse or overuse. Overtreatment was assessed by the EPG members for each GBI from the therapy domain in the second rating round. To avoid the generation of additional GBIs, overtreatment was integrated into the corresponding GBI with an additional numerator (patients fulfilling the aim of the GBI) and denominator (reference population). The panelists considered only treatment with erythropoietin-stimulating agents (ESAs), lenalidomide (LEN), and hypomethylating agents (HMAs) at relevant risk for overtreatment.

Many GBIs from the provider/infrastructural characteristic domain cover topics generally addressed by certification/accreditation programs. However, some GBIs from this domain also measure performance at the patient level. Additionally, as not all institutions are certified/accredited, a minimal consensus on adequate infrastructure was defined for MDS management.

Diagnosis domain

For the diagnosis domain, 14 of 23 candidate GBIs were finally selected after 2 rating rounds (Table 3). This domain comprises the 5 subdomains basic workup ($n = 6$), supplementary workup ($n = 5$), diagnostic workup related to treatment ($n = 3$), risk stratification ($n = 5$), and follow-up/outcomes ($n = 4$). The agreement scores (percentage of scoring in tertile 7-9) from each GBI in both rating rounds are summarized as a radar plot in Figure 2A for the overall and as a heat map in Figure 2B for each RUMBA item. All GBIs from the subdomain basic workup were selected. Except for the GBI molecular diagnostics/next-generation sequencing (NGS), most GBIs from the subdomain secondary workup had to be dropped after the first rating round. Immunophenotyping (IP) did not reach sufficient agreement and was dropped after the first round, as standardization, and its diagnostic and prognostic implications of IP, seemed to be controversial. In the subdomain risk stratification, the GBIs disease-based risk stratification according to IPSS and disease-based risk stratification according to IPSS-R were merged into 1 GBI after the first rating round. This was based on the panelists' feedback: they considered IPSS still relevant, as most

Figure 1. Disposition and flow diagram of the GBI selection process.



Downloaded from https://ashpublications.org/bloodadvances/article-pdf/4/16/4029/1758394/advancesadv2020002314.pdf by UNIVERSITÄTSSIBLIOTHEK BERN user on 27 August 2020

Table 3. Summary of feedback and performance of candidate GBIs

Domains, subdomains and GBIs	Feedback first round, % (n/n)	Performance first round, % (n/n)	Feedback second round, % (n/n)	Performance second round, % (n/n)	Final rating outcomes
Total	23: 14/9 (100/100)		26: 16/10 (100/100)		
Diagnosis, n = 23					
1.1 Basic diagnostic workup					
1.1.1 WHO 2016 classification	14/4 (100/44)	83 (84/80)	16/4 (100/40)	94 (93/100)	Selected
1.1.2 Peripheral blood count and smear	14/4 (100/44)	86 (86/85)	16/4 (100/40)	89 (89/85)	Selected
1.1.3 Bone marrow cytology	14/4 (100/44)	84 (86/80)	16/4 (100/40)	95 (95/95)	Selected
1.1.4 Iron staining	14/4 (100/44)	78 (76/85)	16/4 (100/40)	96 (96/95)	Selected
1.1.5 Bone marrow histology including immunohistochemistry	13/4 (93/44)	79 (78/80)	16/4 (100/40)	85 (83/95)	Selected
1.1.6 Cytogenetic analysis	13/4 (93/44)	88 (87/70)	16/4 (100/40)	95 (95/95)	Selected
1.2 Supplementary diagnostic workup					
1.2.1 <i>Cytochemistry</i>	13/2 (93/22)	44 (44/50)	—	—	<i>Dropped after first round</i>
1.2.2 <i>Immunophenotyping</i>	13/2 (93/22)	52 (52/60)	—	—	<i>Dropped after first round</i>
1.2.3 <i>Fluorescence in situ hybridization</i>	13/2 (93/22)	62 (65/40)	—	—	<i>Dropped after first round</i>
1.2.4 <i>Microarray comparative genomic hybridization</i>	13/1 (93/11)	41 (39/80)	—	—	<i>Dropped after first round</i>
1.2.5 <i>Molecular diagnostics/NGS</i>	13/1 (93/11)	76 (75/100)	16/3 (100/30)	85 (84/93)	Selected
1.3 Diagnostic workup related to treatment					
1.3.1 Serum erythropoietin in patients with symptomatic anemia	13/3 (93/30)	90 (88/100)	16/4 (100/40)	93 (92/95)	Selected
1.3.2 Monitoring iron status	13/3 (93/30)	92 (91/100)	16/4 (100/40)	87 (85/95)	Selected
1.3.3 <i>7P53</i> mutation status in MDS del(5q)	13/2 (93/22)	85 (85/100)	16/3 (100/30)	86 (88/80)	Selected
1.4 Risk stratification					
1.4.1 Disease-based risk stratification					
1.4.1.1 <i>Disease-based risk stratification according to IPSS</i>	13/2 (93/22)	79 (76/100)	—	—	<i>Merged with 1.4.1.2</i>
1.4.1.2 <i>Disease-based risk stratification according to IPSS/IPSS-R</i>	13/2 (93/22)	93 (92/100)	16/3 (100/30)	98 (100/87)	Selected
1.4.1.3 <i>Disease-based risk stratification according to WPSS</i>	13/2 (93/22)	64 (59/100)	—	—	<i>Dropped after first round</i>
1.4.2 Patient-based risk stratification					
1.4.2.1 Patient-based risk stratification	13/2 (93/22)	74 (73/84)	16/4 (100/40)	72 (70/80)	Selected
1.4.2.2 <i>Assessment for germline predisposition for inherited hematological malignancy syndromes</i>	13/2 (93/22)	56 (55/70)	—	—	<i>Dropped after first round</i>
1.5 Follow-up/outcomes					
1.5.1 PROs*	13/6 (93/67)	62 (59/70)	9/6 (56/60)	73 (64/87)	Selected after reevaluation
1.5.2 Response assessment	13/5 (93/56)	69 (65/76)	16/3 (100/30)	89 (93/73)	Selected
1.5.3 <i>Toxicity assessment</i>	13/7 (93/78)	63 (64/57)	16/8 (100/80)	68 (60/85)	<i>Dropped after second round</i>
1.5.4 <i>Geriatric assessment*</i>	12/5 (86/56)	51 (46/64)	5/4 (31/40)	62 (44/85)	<i>Dropped after reevaluation</i>

GBIs in italics were dropped during the first or second rating rounds. Rating feedback are shown with absolute numbers and fractions for MDS and non-MDS experts separately, where 100% corresponds to the total number of participants within each category. Performance is shown as agreement scores for overall, as well as MDS and non-MDS experts separately.

—, not applicable (as no second rating was performed); RS, ringed sideroblasts; WHO, World Health Organization; WPSS, WHO Classification-based Prognostic Scoring System.

*GBIs that were rechallenged in the second round.

Table 3. (continued)

Domains, subdomains and GBIs	Feedback first round, (%/%)	Performance first round, % (n/n)	Feedback second round, (%/%)	Performance second round, % (n/n)	Final rating outcomes
Therapy, n = 21					
2.1 Supportive care					
2.1.1 Overtreatment of lower-risk MDS patients with asymptomatic cytopenia	13/3 (93/30)	63 (63/67)	16/3 (100/30)	64 (65/60)	Dropped after second round
2.1.2 Transfusions					
2.1.2.1 Transfusions of red blood cell concentrates	13/3 (93/30)	73 (79/47)	16/4 (100/40)	84 (86/75)	Selected
2.1.2.2 Iron chelation	13/3 (93/30)	69 (63/100)	16/3 (100/30)	64 (61/80)	Dropped after second round
2.1.2.3 Transfusions of platelet concentrates	13/4 (93/44)	78 (77/80)	16/4 (100/50)	84 (84/84)	Selected
2.1.3 Antibacterial prophylaxis	13/3 (93/30)	60 (59/81)	—	—	Dropped after first round
2.1.4 Granulocyte colony-stimulating factor for infections	13/2 (93/22)	54 (56/40)	—	—	Dropped after first round
2.2 Lower-risk					
2.2.1 Anemia					
2.2.1.1 ESAs	13/3 (93/30)	89 (88/100)	16/4 (100/40)	91 (94/80)	Selected
2.2.1.2 Granulocyte colony-stimulating factor for ESA refractoriness	13/2 (93/22)	62 (64/30)	—	—	Dropped after first round
2.2.1.3 Overtreatment with growth factors	13/3 (93/30)	67 (67/67)	—	—	Dropped after first round
2.2.1.4 LEN in MDS del(5q)	13/2 (93/22)	91 (90/100)	16/4 (100/30)	89 (91/80)	Selected
2.2.1.5 LEN in non-del(5q) MDS	13/2 (93/22)	51 (49/70)	—	—	Dropped after first round
2.2.1.6 Luspatercept in MDS with RS or mutated SF3B1	13/2 (93/11)	40 (40/40)	—	—	Dropped after first round
2.2.2 Other cytopenias					
2.2.2.1 Immunosuppressive treatment*	13/2 (93/22)	63 (64/50)	8/1 (50/10)	62 (68/20)	Dropped after reevaluation
2.2.2.2 Thrombopoietin-stimulating agents	12/2 (86/22)	60 (59/40)	—	—	Dropped after first round
2.2.2.3 HMAs in lower-risk MDS	13/2 (93/22)	51 (46/73)	—	—	Dropped after first round
2.3 Higher-risk					
2.3.1 Unfit: Noneligible for (or declined) intensive treatment and/or allo-HSCT					
2.3.1.1 HMAs in higher-risk MDS	14/3 (100/30)	92 (97/67)	16/3 (100/30)	83 (85/73)	Selected
2.3.2 Fit: Eligible for intensive treatment and/or allo-HSCT					
2.3.2.1 Induction with chemotherapy or HMA before allo-HSCT	14/2 (100/22)	66 (61/100)	16/3 (100/30)	77 (79/67)	Selected
2.3.2.2 Postinduction chemotherapy for patients without a suitable donor	13/2 (93/22)	58 (54/80)	—	—	Dropped after first round
2.3.2.3 Postinduction chemotherapy for patients with a complicated induction course	13/2 (93/22)	49 (44/80)	—	—	Dropped after first round
2.3.2.4 Upfront allo-HSCT	13/2 (93/22)	68 (62/90)	16/3 (100/30)	79 (84/53)	Selected
2.3.2.5 Allo-HSCT	13/2 (93/22)	79 (75/100)	15/3 (94/30)	89 (93/67)	Selected

GBIs in italics were dropped during the first or second rating rounds. Rating feedback are shown with absolute numbers and fractions for MDS and non-MDS experts separately, where 100% corresponds to the total number of participants within each category. Performance is shown as agreement scores for overall, as well as MDS and non-MDS experts separately.

—, not applicable (as no second rating was performed); RS, ringed sideroblasts; WHO, World Health Organization; WPSS, WHO Classification-based Prognostic Scoring System.

*GBIs that were rechallenge in the second round.

Table 3. (continued)

Domains, subdomains and GBIs	Feedback first round, (%/%)	Performance first round, % (%/%)	Feedback second round, (%/%)	Performance second round, % (%/%)	Final rating outcomes
Provider/infrastructural characteristics, n = 17					
3.1 Certification and accreditation					
3.1.1 Institutions certified in hemato-oncology	11/5 (79/56)	64 (58/100)	—	—	Dropped after first round
3.1.2 Access to a certified/accredited diagnostics laboratory	13/6 (93/67)	78 (76/83)	15/6 (94/60)	76 (69/93)	Dropped after second round
3.2 Personnel					
3.2.1 Board-certified and appropriately trained physicians	13/3 (93/33)	79 (80/80)	15/5 (94/50)	83 (84/80)	Dropped after second round
3.2.2 Nursing expertise	12/8 (86/89)	73 (77/68)	15/8 (94/80)	78 (77/80)	Dropped after second round
3.2.3 Multidisciplinary care team	12/9 (86/100)	70 (72/69)	15/9 (94/90)	89 (92/84)	Selected
3.3 Organization					
3.3.1 Safe handling of cytotoxic drugs	12/7 (86/78)	67 (68/66)	15/7 (94/70)	85 (79/97)	Selected
3.3.2 Isolation ward	13/8 (94/89)	75 (70/85)	15/7 (94/70)	78 (75/86)	Dropped after second round
3.3.3 Interdisciplinary diagnostic review	13/6 (94/67)	79 (86/63)	15/7 (94/70)	87 (84/94)	Selected
3.3.4 Interdisciplinary treatment board	12/5 (86/56)	79 (75/88)	15/7 (94/70)	87 (85/91)	Selected
3.3.5 Multidisciplinary care board	10/8 (71/89)	77 (85/80)	15/7 (94/70)	80 (72/97)	Dropped after second round
3.3.6 Teaching and continuing education	12/8 (86/89)	68 (68/68)	15/7 (94/70)	85 (83/91)	Selected
3.3.7 Emergency services for hemato-oncological diseases	13/8 (93/89)	86 (83/93)	15/7 (94/70)	88 (87/91)	Selected
3.3.8 Standard operating procedures	12/5 (86/56)	75 (75/80)	15/8 (94/80)	79 (75/88)	Dropped after second round
3.4 Cooperation					
3.4.1 Multi-institutional care network	11/7 (79/78)	62 (60/63)	—	—	Dropped after first round
3.4.2 Prospective clinical registries	12/5 (84/56)	66 (69/56)	15/6 (94/60)	84 (84/79)	Dropped after second round
3.4.3 Biobanking	13/4 (93/44)	78 (71/80)	15/5 (94/50)	82 (81/84)	Dropped after second round
3.4.4 Access to clinical trials	12/7 (86/78)	83 (83/80)	15/8 (94/80)	92 (95/88)	Selected

GBIs in italics were dropped during the first or second rating rounds. Rating feedback are shown with absolute numbers and fractions for MDS and non-MDS experts separately, where 100% corresponds to the total number of participants within each category. Performance is shown as agreement scores for overall, as well as MDS and non-MDS experts separately.

—, not applicable (as no second rating was performed); RS, ringed sideroblasts; WHO, World Health Organization; WPSS, WHO Classification-based Prognostic Scoring System.

*GBIs that were rechallenged in the second round.

clinical trials used it for treatment allocation, whereas IPSS-R was found to be better for risk stratification.³⁶ In the subdomain patient-based risk stratification, the GBI assessment for germline predisposition for inherited hematological malignancy syndromes did not reach the agreement threshold in the first round, as it was not considered sufficiently relevant, measurable, and attainable by the EPG members at the time rating was performed. Finally, in the follow-up/outcomes subdomain, all GBIs scored generally as highly relevant but with a low agreement in measurability and behavioral orientation, which affected their overall ranking (Figure 2B). The GBIs monitoring of PROs and geriatric assessment (GA) did not reach the required minimal agreement in the first round, as there were controversies among experts on standardization and practicability. As both are consistently found in G/Rs from national and international MDS working groups, the core project working group decided to keep them for reevaluation in the second DELPHI round. The GBIs GA and TA finally had to be dropped, whereas PROs just reached the required threshold of sufficient agreement after the second rating round. In the expert rating agreement analysis, only GA performed better in non-MDS than MDS experts, all others were comparable (supplemental Figure 2).

Therapy domain

For the therapy domain, 8 of 21 candidate GBIs were finally selected after 2 rating rounds (Table 3). This domain comprises the 3 subdomains supportive care ($n = 6$), lower risk ($n = 9$), and higher risk ($n = 6$), as defined in "Methods." GBI agreement scores are summarized as a radar plot in Figure 3A (overall) and as a heat map in Figure 3B (each RUMBA item). In the subdomain supportive care, only the GBIs transfusion of red blood cell concentrates and transfusions of platelet concentrates reached sufficient agreement. The GBI iron chelation (IC) had to be dropped after the second round. Despite the panelists rating this GBI as relevant and understandable, the behavioral aspect and the attainability seem to be hampered by local guidelines, high treatment costs, fear of adverse effects, and controversial clinical trial data. In the subdomain lower risk and anemia, the GBIs ESAs and LEN in MDS del(5q) reached sufficient agreement, whereas all others, such as granulocyte colony-stimulating factor for ESA refractoriness, LEN in non-del(5q) as well as luspatercept in MDS with ringed sideroblasts or mutated SF3B1, had to be dropped after the first rating round. None of the candidate GBIs in the subdomain other cytopenias reached sufficient agreement. The GBI immunosuppressive treatment (IST) did not reach the required minimal agreement in the first round. Because IST is also consistently found in G/Rs from national and international MDS working groups, it was kept for reevaluation in the second DELPHI round. The performance also remained poor in the second round, as panelists reported insufficient experience and standardization for treatment of the rare lower-risk MDS patients with hypoplastic bone marrow. The GBIs HMAs in higher-risk MDSs for unfit patients as well as induction with chemotherapy or HMAs before allogeneic hematopoietic stem cell transplantation (allo-HSCT), upfront allo-HSCT, and allo-HSCT (after induction chemotherapy) for fit higher-risk MDS patients, respectively, reached sufficient agreement for selection. GBIs defining procedures of postinduction chemotherapy for patients without a suitable donor or complicated induction course were dropped after the first rating round due to insufficient evidence and lack of standardization. In the expert rating agreement analysis, upfront allo-HSCT and IST performed better in MDS

than non-MDS experts, all others were comparable (supplemental Figure 3).

Provider/infrastructural characteristics domain

For the provider/infrastructural characteristics domain, 7 of 17 candidate GBIs were finally selected after 2 rating rounds (Table 3). This domain comprises the 4 subdomains certification and accreditation ($n = 2$), personnel ($n = 3$), organization ($n = 8$), and cooperation ($n = 4$). GBI agreement scores are summarized as a radar plot in Figure 4A (overall) and as a heat map in Figure 4B (each RUMBA item). The higher selection threshold $\geq 85\%$ was justified based on the observation that most candidate GBIs from this domain generally scored $\geq 70\%$ and would have been overrepresented in our final set of relevant GBIs. None of the candidate GBIs from the subdomain certification and accreditation was selected, as they were considered to be a mandatory requirement for clinical practice outside of MDS management. In the subdomain personnel, only the GBI multidisciplinary care team reached the defined agreement threshold. Most GBIs from the subdomain of organization reached the defined agreement, except the GBIs isolation ward, multidisciplinary care board, and standard operating procedures. For the subdomain cooperation, only the GBI access to clinical trials was selected with high agreement. We found a comparable expert rating agreement between MDS and non-MDS experts in all GBIs of this domain (supplemental Figure 4).

Discussion

Here, we provide a first consensus on 29 clinically relevant GBIs for diagnosis ($n = 14$), therapy ($n = 8$), and provider/infrastructural characteristics ($n = 7$) for the care of adult MDS patients. Our GBIs were extracted from published G/Rs and were selected by a structured consensus process involving a multidisciplinary panel with internationally recognized experts.

Surprisingly, some candidate GBIs, generally considered to address relevant aspects of MDS care, did not meet the required threshold for acceptance. These GBIs were uniformly rated as relevant but they scored relatively low with regard to measurability, attainability, and/or behavioral orientation. For those GBIs dropped in the diagnosis domain (ie, IP, TA, and GA), there remains a need for improvement in standardization, practicability as well as validation with future clinical trials. For those dropped in the therapy domain (ie, IC and IST), there are ongoing controversies on efficacy and toxicity as well as insufficient agreement on their adequate use, especially for IST in the rare entity of hypoplastic MDS. In the Swedish population-based MDS registry, for instance, only 0.9% of MDS cases received IST between 2009 and 2017, underlining the low frequency of this MDS subtype and rare use of IST.³⁷ Recently, a core outcome set (COS) for MDS patients has been proposed by the ELN MDS working group (COS-MDS). This set comprises overall survival (OS), performance score (PS), hematological improvement (HI), treatment-related mortality (TRM), safety, and quality of life (PROs).³⁸ Except for TA (safety and TRM), all other outcomes are included in our final set of GBIs. PROs reached the required threshold for acceptance only after rechallenge in the second round. This underlines ongoing controversies and the requirement for further improvements in standardization, practicability, and validation of their clinical utility. The rating performance of the selected GBIs was generally similar between MDS and non-MDS experts (disagreement $\leq 30\%$).

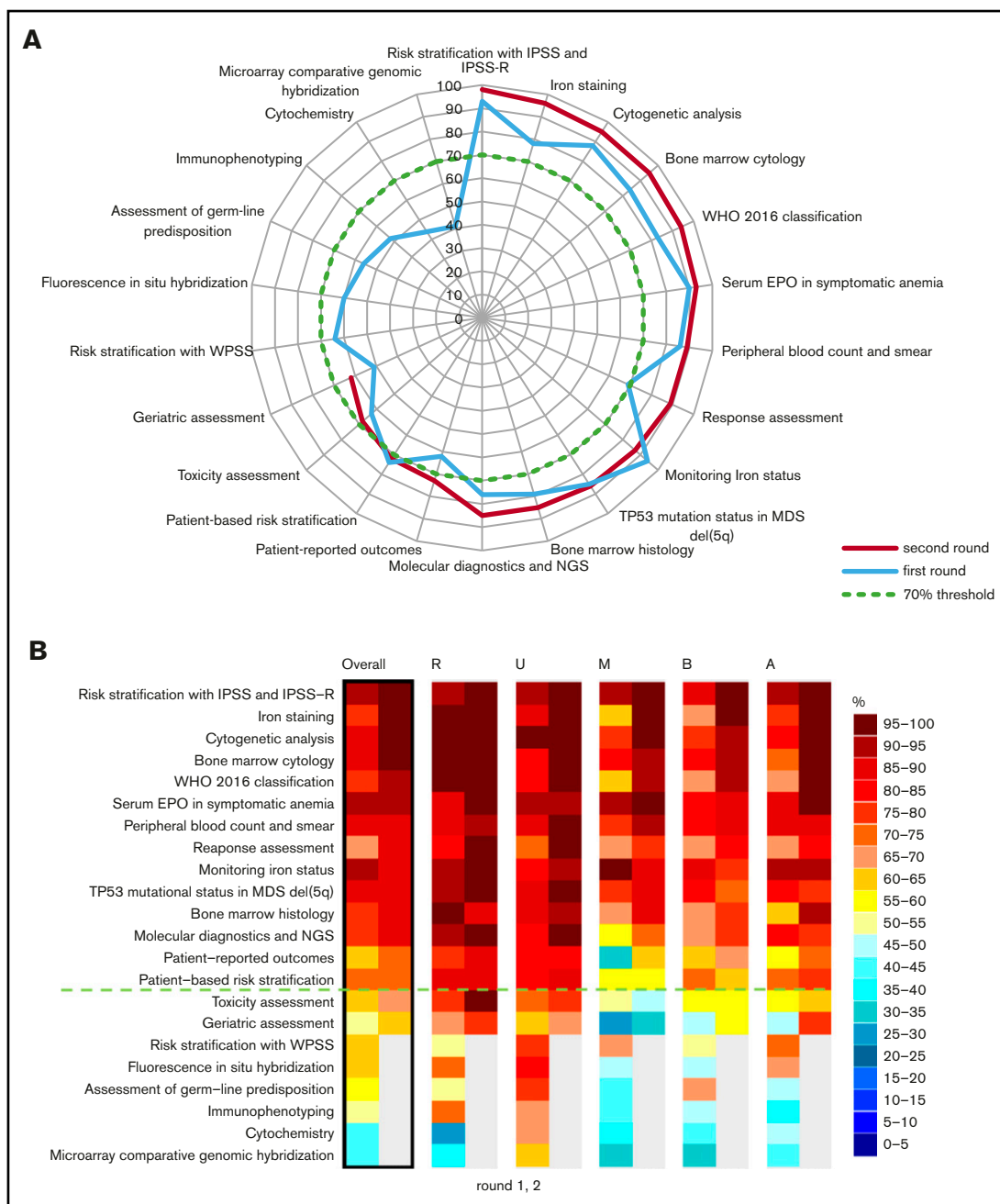


Figure 2. Rating performance of GBIs from the diagnosis domain. (A) Radar plot of overall GBI performance from the diagnosis domain showing agreement scores from the first (blue curves) and second (red curves) DELPHI rating round. GBIs dropped after the first round are shown with single blue curve. (B) Agreement scores for each GBI are shown as a heat map for overall and RUMBA subcategories for the first (left) and second (right) DELPHI rating round. GBIs are ranked according to the performance in the second rating round with the selection threshold highlighted in a green dotted line at 70%. GBIs dropped after the first round are shown in gray for the second round. GBIs selected, 14; merged, 2; dropped, 8.

Only allo-HSCT performed better in MDS experts, suggesting a different perception of its importance between the expert groups.

Currently, there is no generally accepted standard for the development of GBIs. Our strategy may harbor some limitations, but the applied methods are in line with approaches published by many others.^{8,21} The appointment of EPG members and the preselection of candidate GBIs can be potential sources of bias. Our selection of

internationally acknowledged MDS experts as well as multidisciplinary non-MDS experts minimizes regional and discipline-associated biases. The inclusion of non-MDS experts is in line with the current recommendations for the development of GBIs. It stands to reason that their opinion was most prominent for the GBIs from the structural domain and for PROs, whereas for the other domains, most non-MDS expert refrained from rating. Our EPG members were highly committed (feedback rates, 82% and 96%, respectively) and also

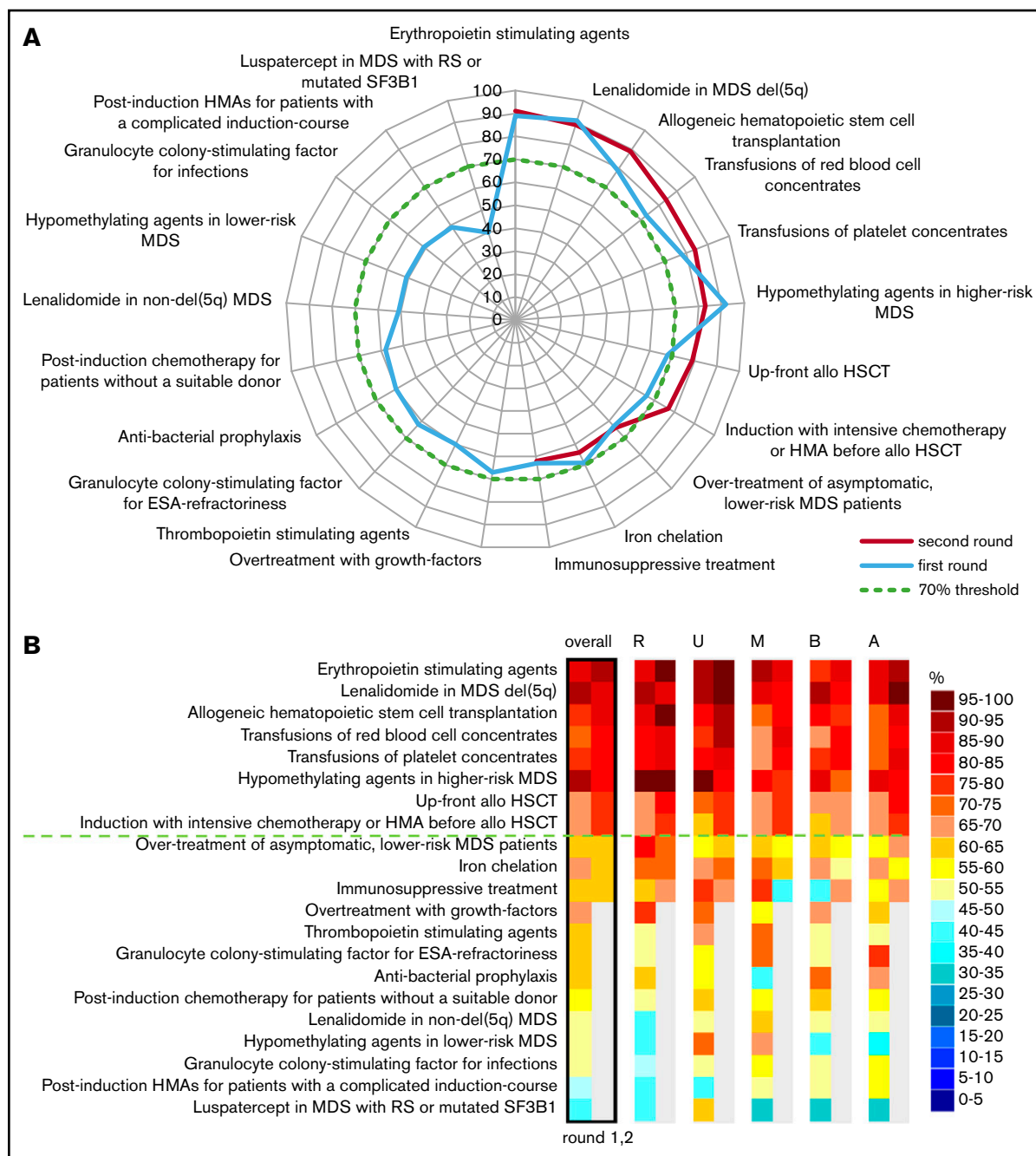


Figure 3. Rating performance of GBIs from the therapy domain. (A) Radar plot of overall GBI performance from the therapy domain showing agreement scores from the first (blue curves) and second (red curves) DELPHI rating round. GBIs dropped after the first round are shown with single blue curve. (B) Agreement scores for each GBI are shown as a heat map for overall and RUMBA subcategories for the first (left) and second (right) DELPHI rating round. GBIs are ranked according to the performance in the second rating round with the selection threshold highlighted in a green dotted line at 70%. GBIs dropped after the first round are shown in gray for the second round. GBIs selected, 8; dropped, 13.

provided general feedback. They did not suggest additional candidate GBIs, indicating that we had covered all relevant aspects of MDS care with our initial preselection. Our GBIs have been developed in the context of high-income countries with mainly European experts. The appropriateness of the selected GBIs for developing countries or other countries, including the United States, not adequately represented by the EPG members, may be debatable. As an example, HMAs may be

more frequently used in lower-risk MDS patients in the United States compared with Europe, which is based on local differences of G/Rs. We also think that assessment of germline predisposition in younger MDS patients as well as treatment with luspatercept in MDS with ringed sideroblasts/*SF3B1* mutations would have performed differently today, as recent G/Rs are increasingly addressing germline predisposition and luspatercept has recently been approved for anemic

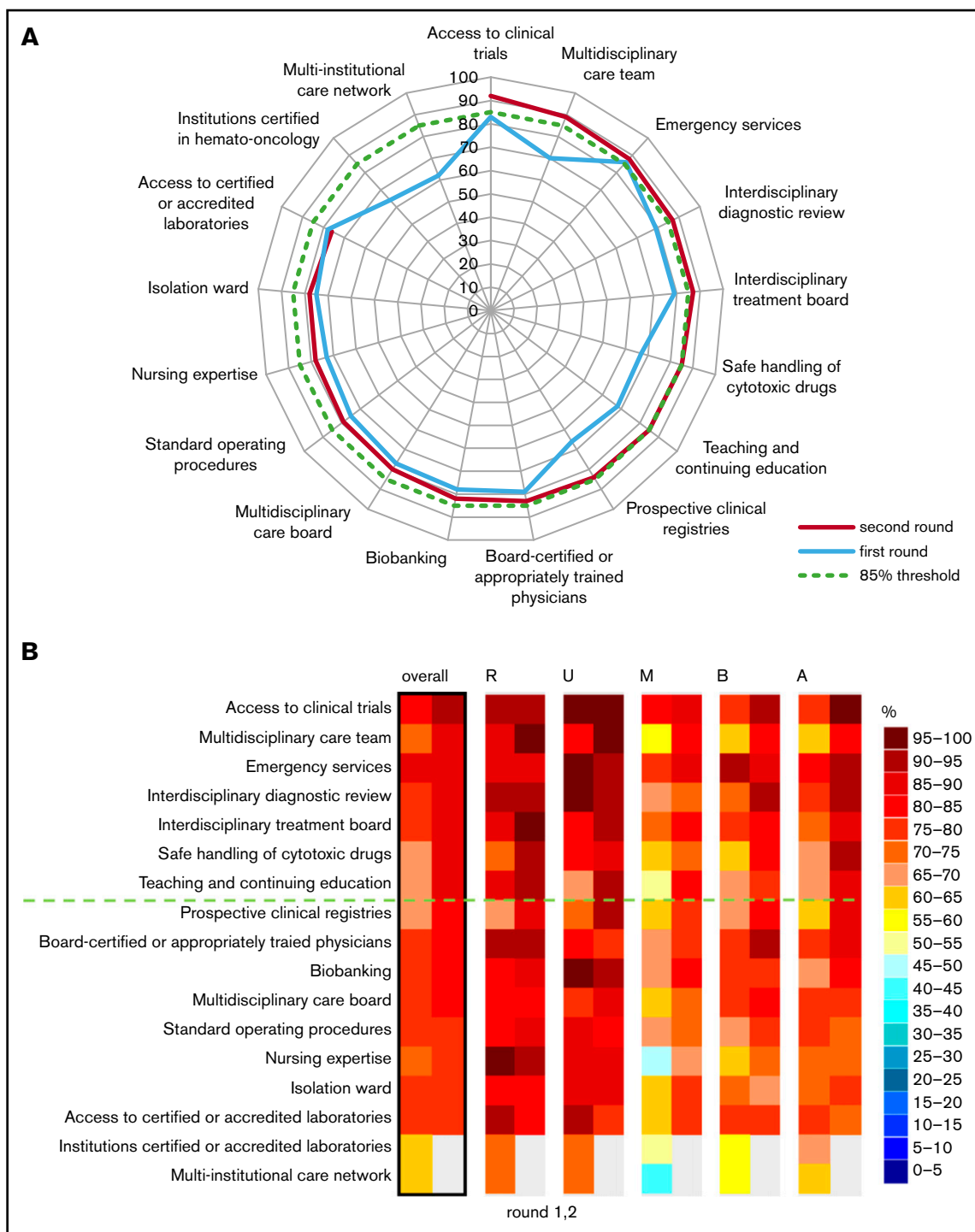


Figure 4. Rating performance of GBIs from the provider/infrastructural characteristics domain. (A) Radar plot of overall GBI performance from the provider/infrastructural characteristics domain showing agreement scores from the first (blue curves) and second (red curves) DELPHI rating round. GBIs dropped after the first round are shown with single blue curve. (B) Agreement scores for each GBI are shown as a heat map for overall and RUMBA subcategories for the first (left) and second (right) DELPHI rating round. GBIs are ranked according to the performance in the second rating round with the selection threshold highlighted in a green dotted line at 85%. GBIs dropped after the first round are shown in gray for the second round. GBIs selected, 7; dropped, 10.

MDS patients with ringed sideroblasts/*SF3B1* mutations.^{39,40} As such, the selection of GBIs remains time-dependent and influenced by regional preferences. Nonetheless, our set represents a first consensus, which can be complemented with additional GBIs that were

currently not included and may be perceived to be relevant in other contexts. The development of GBIs should remain dynamic, follow future revisions of G/Rs, and need to be adapted at regular intervals. Finally and most importantly, GBIs will never replace common sense,

as it is virtually impossible to cover all aspects of irrational behavior, including unnecessary repetition of diagnostic procedures. In any case, our GBIs need future validation for operability and applicability in the daily clinical care of MDS patients and investigations on their impact on patient-centered outcomes. This is planned with the prospective, observational I-CARE for MDS study (SAKK 33/18), which has recently started and will be performed in collaboration with other international MDS study groups.

Assessment of quality in the real-world setting of MDS patient care has thus far not received adequate attention. Available data from retrospective studies indicate common shortcomings in both diagnostic workup and treatment.¹⁹ The analysis of a US Medicare cohort of 4575 MDS patients diagnosed between 2006 and 2011 revealed that only 74% received a bone marrow cytogenetics analysis.⁴¹ Several European population-based registries and MDS cohorts published very similar results.^{36,42-45} The lack of cytogenetics impedes appropriate risk calculation by IPSS or IPSS-R. Our set of GBIs includes all diagnostic methods required for diagnosis and risk stratification and can be used to maintain and improve diagnostic quality.^{16,29,30,34} Unexpectedly, IP did not reach sufficient agreement, as it seems to be increasingly replaced by NGS in cases of unclear diagnoses. Sequencing-based methods are more reproducible and provide additional prognostic information. However, standardization of NGS is ongoing and the clinical benefit remains controversial.^{46,47} Determining the iron status and the endogenous erythropoietin (EPO) level before starting anemia treatment are 2 additional GBIs that were selected. According to reports from the US Medicare cohort and the MDS registry from the Fondazione Italiana Sindromi Mielodisplastiche (FISiM), iron status was measured in 56% and endogenous EPO level in 43% of patients receiving ESAs, respectively.⁴¹ These findings imply that approximately one-half of all lower-risk MDS patients are treated without knowing their chance for a response to ESAs.^{43,48}

Allo-HSCT and HMAs are used in fit and unfit higher-risk MDS patients, respectively. Both treatment modalities are included in our final selection of GBIs. Of note, recent retrospective data from the Italian FISiM registry reported that ~6% of eligible higher-risk MDS patients were transplanted and 36% received HMAs, respectively.⁴³ Reasons for this remarkable underuse need to be explored. A recent prospective and retrospective analysis of MDS patients treated in a tertiary care center in Germany showed that adherence to allo-HSCT G/Rs resulted in improved survival, whereas adherence to other treatments did not.⁴⁹ These data underscore the impact of allo-HSCT G/Rs on survival, whereas the benefit of less-intensive treatments on additional patient-centered outcomes remains difficult to be captured in post hoc studies.

As in many other cancers, the number of MDS patient cases and treatment-associated costs are continuously rising. This phenomenon coincides with increasing regulatory mandates that require a culture of self-examination and quality improvement to maintain value with optimized patient-centered outcomes. Such developments are fostered by the QOPI program initiated by ASCO.⁹ Therefore, our GBIs address an unmet clinical need and provide an instrument for the systematic and standardized assessment of guideline-based adherence in adult MDS patients. This instrument will enable the comparison of quality of care in different health care environments and the identification of potential areas for improvement. Our GBIs represent the first comprehensive consensus of

measurable elements addressing best practice performance, outcomes, and structural resources. They can be used as a standardized instrument with the goal of assessing, comparing, and fostering good quality of care within clinical development cycles in the daily care. This first set of GBIs can be considered as a starting point for more objective evaluation of quality of care in adult MDS patients. It might not be ideal from the individual physician's perspective, however, it is the best compromise that can currently be achieved.

Acknowledgments

The authors thank all expert panel group members for their commitment and participation in the rating procedure. The authors also thank Anne Angelillo-Scherrer and Gabriela Maria Baerlocher, as well as all members of the Swiss MDS Study Group, the ELN MDS Working Group, and the SAKK collaborators for their continuous support of the I-CARE for MDS project. Finally, the authors thank Mariana Ricca and Darren Hart for editing the manuscript and handbook.

This work was supported by Swiss Cancer Research (HSR-4085-11-2016) (N.B.) and the Swiss Group for Clinical Cancer Research (SAKK 33/18) (N.B.).

Authorship

Contribution: K.S. collected and analyzed data and wrote the paper; T.S. analyzed data and wrote the paper; J. Bernhard, D.B., J.Č., A.G.D., C.E., S.E., P.F., U.G., M. Haschke, E.H.-L., M. Heger, A.A.v.d.L., J.P., M.P., U.P., L.M., A.M.d.A., M.M., C. Morgenthaler, D.P.S., V.S., R.S., A.S., and T.d.W. gave GBI ratings and critically revised the paper; C. Maddox and T.d.W. provided essential support to the study and critical revision of the paper; S.S. performed statistical analysis, provided graphs, and critically revised the paper; G.S., D.S., and J. Bohlius designed the study, analyzed data, and wrote the paper; and N.B. initiated and designed the study, analyzed data, and wrote the paper.

Conflict-of-interest disclosure: The authors declare no competing financial interests for the rating and selection of GBIs described in this manuscript. Potentially perceived conflicts of interest according to the definitions and terms of the International Committee of Medical Journal Editors are as follows: P.F. received honoraria and research funding from Celgene Corporation, Jazz Pharmaceuticals, and Astex, and research funding from Aprea. U.G. received honoraria and research funding from Novartis and Celgene, and honoraria from Jazz Pharmaceuticals and Amgen. M. Haschke provided consultancy services to, and held membership on an entity's board of directors or advisory committees for, Sanofi Aventis. A.A.v.d.L. received honoraria and research funding from Celgene; research funding from Roche and Alexion; and honoraria from Novartis and Pfizer. M.P. received honoraria from Janssen-Cilag, and provided consultancy services to, and received honoraria from, Novartis and Celgene. U.P. provided consultancy services to, and received honoraria from, Celgene, AbbVie, and Novartis. A.M.d.A. was on speaker's bureaus for Celgene and Novartis. M.M. received honoraria and research funding from, and was on a speaker's bureau for, Novartis. A.S. held membership on an entity's board of directors or advisory committees for, and received research funding from, Merck Sharp & Dohme, Gilead, Takeda, Janssen, Novartis, and Roche; received research funding from Sanofi and Pfizer; and received honoraria and research funding from, and held membership on an entity's board of directors or

advisory committees for, Celgene. D.P.S. received research funding (to institution, not investigator) from H3 Biosciences; provided consultancy services to Stemline, Pfizer, Onconova, Astex, and Summer Road; held equity ownership in Arrowhead; and received research funding from Aprea. V.S. received honoraria from Menarini, Amgen, Johnson & Johnson, Novartis, Pfizer, and Takeda; and received honoraria and institutional research funding from Celgene Corporation. R.S. received honoraria from, and held membership on an entity's board of directors or advisory committees for, Novartis; received honoraria and research funding from Teva (Ratiopharm); and provided consultancy services to, received honoraria and research funding from, held membership on an entity's board of directors or advisory committees for, and was on an advisory board for Celgene. T.d.W. received research funding from Novartis, Celgene, and Amgen. N.B. receives research funding (to institution, not investigator) from Astellas, Celgene, Novartis, Servier, and Sandoz; and received financial support for travel from Amgen, Celgene,

Janssen, Novartis, Roche, and Sanofi. The remaining authors declare no competing financial interests.

A complete list of the members of the Swiss MDS Study Group and participating centers in the I-CARE for MDS (SAKK 33/18) study can be found at www.MDS-Switzerland.ch. More information on SAKK and the I-CARE for MDS (SAKK 33/18) study can be found at www.sakk.ch/de/krebsarten/leukemias.

ORCID profiles: G.S., 0000-0002-1667-0637; D.S., 0000-0001-8668-3065; S.E., 0000-0002-7379-7343; J.P., 0000-0001-7092-3351; D.P.S., 0000-0001-5130-9284; R.S., 0000-0002-8993-9561; A.S., 0000-0002-3685-3473; N.B., 0000-0001-8761-2066.

Correspondence: Nicolas Bonadies, Department of Hematology and Central Hematology Laboratory, Inselspital, Bern University Hospital, University of Bern, Freiburgstr 18, CH-3010 Bern, Switzerland; e-mail: nicolas.bonadies@insel.ch.

References

1. Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet*. 1993;342(8883):1317-1322.
2. Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA; The Cochrane Effective Practice and Organization of Care Review Group. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. *BMJ*. 1998;317(7156):465-468.
3. Lawrence M, Olesen F. Indicators of quality in health care. *Eur J Gen Pract*. 1997;3(3):103-108.
4. Lugtenberg M, Burgers JS, Westert GP. Effects of evidence-based clinical practice guidelines on quality of care: a systematic review. *Qual Saf Health Care*. 2009;18(5):385-392.
5. Shapiro DW, Lasker RD, Bindman AB, Lee PR. Containing costs while improving quality of care: the role of profiling and practice guidelines. *Annu Rev Public Health*. 1993;14:219-241.
6. Porter ME. What is value in health care? *N Engl J Med*. 2010;363(26):2477-2481.
7. Califf RM, Peterson ED, Gibbons RJ, et al; American Heart Association. Integrating quality into the cycle of therapeutic development. *J Am Coll Cardiol*. 2002;40(11):1895-1901.
8. Shekelle PG. Quality indicators and performance measures: methods for development need more standardization. *J Clin Epidemiol*. 2013;66(12):1338-1339.
9. Chiang AC. Why the quality oncology practice initiative matters: it's not just about cost. *Am Soc Clin Oncol Educ Book*. 2016;35:e102-e107.
10. Kowalski C, Schulte H, Wesselmann S. Reporting program for cancer care quality indicators. *J Oncol Pract*. 2015;11(2):158-160.
11. Wennekes L, Ottavanger PB, Raemaekers JM, et al. Development and measurement of guideline-based indicators for patients with non-Hodgkin's lymphoma. *J Clin Oncol*. 2011;29(11):1436-1444.
12. Wood L, Bjarnason GA, Black PC, et al. Using the Delphi technique to improve clinical outcomes through the development of quality indicators in renal cell carcinoma. *J Oncol Pract*. 2013;9(5):e262-e267.
13. Spencer BA, Steinberg M, Malin J, Adams J, Litwin MS. Quality-of-care indicators for early-stage prostate cancer. *J Clin Oncol*. 2003;21(10):1928-1936.
14. Bilimoria KY, Raval MV, Bentrem DJ, Wayne JD, Balch CM, Ko CY. National assessment of melanoma care using formally developed quality indicators. *J Clin Oncol*. 2009;27(32):5445-5451.
15. Greenberg PL. Myelodysplastic syndromes: dissecting the heterogeneity. *J Clin Oncol*. 2011;29(15):1937-1938.
16. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016;127(20):2391-2405.
17. Cogle CR. Incidence and burden of the myelodysplastic syndromes. *Curr Hematol Malig Rep*. 2015;10(3):272-281.
18. Bonadies N, Feller A, Rovo A, et al; NICER Working Group. Trends of classification, incidence, mortality, and survival of MDS patients in Switzerland between 2001 and 2012. *Cancer Epidemiol*. 2017;46:85-92.
19. Frosch ZA, Abel GA. Assessing quality of care for the myelodysplastic syndromes. *Curr Hematol Malig Rep*. 2016;11(6):402-407.
20. Campbell SM, Braspenning J, Hutchinson A, Marshall M. Research methods used in developing and applying quality indicators in primary care. *Qual Saf Health Care*. 2002;11(4):358-364.
21. Kötter T, Blozik E, Scherer M. Methods for the guideline-based development of quality indicators—a systematic review. *Implement Sci*. 2012;7(1):21.

22. Ketelaars CA. Using the RUMBA requirements in developing criteria. *Nurs Qual Connect*. 1994;4(2):7.
23. Chassin M. How do we decide whether an investigation or procedure is appropriate? In: Hopkins A, ed. *Appropriate Investigation and Treatment in Clinical Practice*. London, United Kingdom: Royal College of Physicians; 1989.
24. Coulter I, Herman P, Ryan G, Hilton L, Hays RD; Members of CERC Team. The challenge of determining appropriate care in the era of patient-centered care and rising health care costs. *J Health Serv Res Policy*. 2019;24(3):201-206.
25. Boukchedid R, Abdoul H, Loustau M, Sibony O, Alberti C. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. *PLoS One*. 2011;6(6):e20476.
26. Shekelle PG, Woolf SH, Eccles M, Grimshaw J. Clinical guidelines: developing guidelines. *BMJ*. 1999;318(7183):593-596.
27. European Consortium for Accreditation. ECA Principles for the Selection of Experts. <http://ecahe.eu/w/images/1/11/Eca-principles-for-the-selection-of-experts.pdf>. Accessed 10 January 2018.
28. Harris PA, Taylor R, Minor BL, et al; REDCap Consortium. The REDCap consortium: building an international community of software platform partners. *J Biomed Inform*. 2019;95:103208.
29. Greenberg P, Cox C, LeBeau MM, et al. International scoring system for evaluating prognosis in myelodysplastic syndromes. *Blood*. 1997;89(6):2079-2088.
30. Greenberg PL, Tuechler H, Schanz J, et al. Revised international prognostic scoring system for myelodysplastic syndromes. *Blood*. 2012;120(12):2454-2465.
31. Malcovati L, Papaemmanuil E, Ambaglio I, et al. Driver somatic mutations identify distinct disease entities within myeloid neoplasms with myelodysplasia. *Blood*. 2014;124(9):1513-1521.
32. Bejar R, Stevenson K, Abdel-Wahab O, et al. Clinical effect of point mutations in myelodysplastic syndromes. *N Engl J Med*. 2011;364(26):2496-2506.
33. Bejar R, Stevenson KE, Caughey B, et al. Somatic mutations predict poor outcome in patients with myelodysplastic syndrome after hematopoietic stem-cell transplantation. *J Clin Oncol*. 2014;32(25):2691-2698.
34. Della Porta MG, Malcovati L, Boveri E, et al. Clinical relevance of bone marrow fibrosis and CD34-positive cell clusters in primary myelodysplastic syndromes. *J Clin Oncol*. 2009;27(5):754-762.
35. Sorror ML, Maris MB, Storb R, et al. Hematopoietic cell transplantation (HCT)-specific comorbidity index: a new tool for risk assessment before allogeneic HCT. *Blood*. 2005;106(8):2912-2919.
36. Moreno Berggren D, Folkvjon Y, Engvall M, et al. Prognostic scoring systems for myelodysplastic syndromes (MDS) in a population-based setting: a report from the Swedish MDS register. *Br J Haematol*. 2018;181(5):614-627.
37. Regionala. Cancercentrum SYD. Kvalitetsregisterrapporter MDS. <https://www.cancercentrum.se/syd/cancerdiagnoser/blod-lymfom-myelom/myelodysplastiskt-syndrom-mds/kvalitetsregister/rapporter/>. Accessed 11 January 2019.
38. Rochau U, Stojkov I, Conrads-Frank A, et al. Development of a core outcome set for myelodysplastic syndromes—a Delphi study from the EUMDS Registry Group. *Br. J. Haematol*. doi:10.1111/bjh.16654
39. University of Chicago Hematopoietic Malignancies Cancer Risk Team. How I diagnose and manage individuals at risk for inherited myeloid malignancies. *Blood*. 2016;128(14):1800-1813.
40. Fenaux P, Platzbecker U, Mufti GJ, et al. Luspatercept in patients with lower-risk myelodysplastic syndromes. *N Engl J Med*. 2020;382(2):140-151.
41. Abel GA, Cronin AM, Odejide OO, Uno H, Stone RM, Steensma DP. Influence of patient and provider characteristics on quality of care for the myelodysplastic syndromes. *Br J Haematol*. 2016;173(5):713-721.
42. Gattermann N, Kündgen A, Kellermann L, Zeffel M, Paessens B, Germing U. The impact of age on the diagnosis and therapy of myelodysplastic syndromes: results from a retrospective multicenter analysis in Germany. *Eur J Haematol*. 2013;91(6):473-482.
43. Santini V, Della Porta M, Enrico B, et al. Evaluation of adherence to treatment recommendations according to Italian and European guidelines in MDS patients enrolled in the Italian FISiM Registry. In: *Proceedings from the 15th International Symposium on Myelodysplastic Syndromes*; 8-11 May 2019; Copenhagen, Denmark.
44. Dinmohamed AG, van Norden Y, Visser O, et al. The use of medical claims to assess incidence, diagnostic procedures and initial treatment of myelodysplastic syndromes and chronic myelomonocytic leukemia in the Netherlands. *Leuk Res*. 2015;39(2):177-182.
45. Dinmohamed AG, Visser O, Posthuma EFM, et al. MDS classification is improving in an era of the WHO 2016 criteria of MDS: a population-based analysis among 9159 MDS patients diagnosed in the Netherlands. *Cancer Epidemiol*. 2017;50(pt A):137-140.
46. Thol F, Platzbecker U. Do next-generation sequencing results drive diagnostic and therapeutic decisions in MDS? *Blood Adv*. 2019;3(21):3449-3453.
47. Sanz GF, Ibañez M, Such E. Do next-generation sequencing results drive diagnostic and therapeutic decisions in MDS? *Blood Adv*. 2019;3(21):3454-3460.
48. Hellström-Lindberg E, Gulbrandsen N, Lindberg G, et al; Scandinavian MDS Group. A validated decision model for treating the anaemia of myelodysplastic syndromes with erythropoietin + granulocyte colony-stimulating factor: significant effects on quality of life. *Br J Haematol*. 2003;120(6):1037-1046.
49. Kasprzak A, Nachtkamp K, Kondakci M, et al. Combined retro- and prospective analysis of adherence to guidelines and patient-tailored therapeutic recommendations in patients with myelodysplastic syndromes (MDS) at a tertiary care centre [abstract]. *Blood*. 2019;134(suppl 1). Abstract 4250.
50. 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.

51. Fenaux P, Haase D, Sanz GF, Santini V, Buske C; ESMO Guidelines Working Group. Myelodysplastic syndromes: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2014;25(suppl 3):iii57-iii69.
52. Malcovati L, Hellström-Lindberg E, Bowen D, et al; European Leukemia Net. Diagnosis and treatment of primary myelodysplastic syndromes in adults: recommendations from the European LeukemiaNet. *Blood*. 2013;122(17):2943-2964.
53. Killick SB, Carter C, Culligan D, et al; British Committee for Standards in Haematology. Guidelines for the diagnosis and management of adult myelodysplastic syndromes. *Br J Haematol*. 2014;164(4):503-525.