Supplemental Tables

Table S1. Comparison of participants enrolled at the time of diagnosis versus those enrolled within 12 months of diagnosis.

		Enrolled at Diagnosis (N = 864)	Enrolled within 12 months of Diagnosis (N = 54)
Gender	Female, N (%)	335 (39%)	20 (37%)
	Male, N (%)	529 (61%)	34 (63%)
Age	Mean (SD)	70.0 (12.2)	67.2 (10.2)
Race	Black or African American	47 (5%)	1 (2%)
	White	779 (90%)	50 (93%)
	Other	25 (3%)	1 (2%)
	Unknown/Not Reported	13 (2%)	2 (4%)
Ethnicity	Hispanic or Latino	23 (3%)	1 (2%)
	Not Hispanic or Latino	827 (96%)	51 (94%)
	Unknown/Not Reported	14 (2%)	2 (4%)
Final Study Diagnosis	AML <30% blasts*	14 (2%)	1 (2%)
	ICUS	48 (6%)	1 (2%)
	MDS	226 (26%)	40 (74%)
	MDS/MPN overlap	43 (5%)	2 (4%)
	Other AML	36 (4%)	2 (4%)
	Other Malignancy	90 (10%)	3 (6%)
	Other	407 (47%)	5 (9%)

*AML with <30% blasts without core binding factor or acute promyelocytic leukemia.

Table S2. Disease subclass diagnosis for cases that were classified as Other or with a subclass that did not indicate a myeloid malignancy (N=789). This table is published elsewhere [12] and pertains to a larger sample of the MDS Natural History Study than was included in this paper and has been annotated here to provide additional details of the diagnosis categories.

	N=789	
Diagnosis	**N**	**%**
Normal marrow	159	20.2
Cytopenias NOS	124	15.7
Hypocellular marrow (including marrow failure disorders NOS, aplastic anemia)	118	15.0
Isolated anemia (including vitamin deficiency, GI bleeds, general blood loss)	79	10.0
Reactive marrow	71	9.0
Nondiagnostics with cytopenias	70	8.9
Clonal lymphoid process (including CLL, other NHL, LGL)	57	7.2
Hypercellular marrow	44	5.6
Systemic disease (including autoimmune disorders and renal disease)	33	4.2
Plasma cell neoplasm	33	4.2
Carcinoma	1	0.1

CLL = chronic lymphocytic leukemia; GI = gastrointestinal; LGL = large granular lymphocytes; NHL = Non-Hodgkin's Lymphoma; NOS = Not Otherwise Specified

Table S3. Revised International Prognostic Scoring System (IPSS-R) score distribution for all study participants.

	N = 918	
IPSS-R	**N**	**%**
Very Low	268	29.2
Low	191	20.8
Intermediate	76	8.28
High	42	4.58
Very High	52	5.66
Not Calculated	289	31.5

Score categorization: Very Low: <=1.5; Low: >1.5 - 3;
Intermediate: >3 - 4.5; High: >4.5 - 6; Very High: >6

Table S4. Contingency table of validated local diagnosis and final study diagnosis for patients with local genetics available at diagnosis.

	Final Study Diagnosis							
Validated Local Diagnosis	**AML<30% blasts***	**ICUS**	**MDS**	**MDS/MPN overlap**	**Other**	**Other AML**	**Other Malignancy**	**Total**
AML<30%	0	0	0	0	0	0	0	**0**
ICUS	0	3	1	0	1	0	0	**5**
MDS	0	1	31	5	1	0	0	**38**
MDS/MPN overlap	0	0	1	4	0	0	1	**6**
Other	0	0	5	0	20	0	0	**25**
Other AML	0	0	0	0	0	8	0	**8**
Other Malignancy	0	0	1	0	0	0	5	**6**
Total	**0**	**4**	**39**	**9**	**22**	**8**	**6**	**88**
Agreement rate	-	75%	79%	44%	91%	100%	83%	**81%**
Kappa (95% CI)								**0.73 (0.61-0.85)**

The number of participants assigned to each disease group cross-classified by assignment source is reported in this table, along with the agreement rates and kappa statistic. Agreement rates reflect the proportion of participants in each final study diagnosis category (MDS subtype) that had a matching validated local diagnosis. *AML with <30% blasts without core binding factor or acute promyelocytic leukemia.

Table S5. Contingency table of validated local diagnosis and final study diagnosis for patients without local genetics available at diagnosis.

	Final Study Diagnosis							
Validated Local Diagnosis	**AML<30% blasts***	**ICUS**	**MDS**	**MDS/MPN overlap**	**Other**	**Other AML**	**Other Malignancy**	**Total**
AML<30%	12	0	0	0	0	0	0	12
ICUS	0	28	7	2	11	0	3	51
MDS	2	7	178	5	21	1	3	217
MDS/MPN overlap	0	0	1	24	0	0	2	27
Other	0	9	36	5	358	0	4	412
Other AML	1	0	4	0	0	29	0	34
Other Malignancy	0	1	1	0	0	0	75	77
Total	**15**	**45**	**227**	**36**	**390**	**30**	**87**	**830**
Agreement rate	80%	62%	78%	67%	92%	97%	86%	**85%**
Kappa (95% CI)								**0.78 (0.74-0.81)**

The number of participants assigned to each disease group cross-classified by assignment source is reported in this table, along with the agreement rates and kappa statistic. Agreement rates reflect the proportion of participants in each final study diagnosis category (MDS subtype) that had a matching validated local diagnosis. *AML with <30% blasts without core binding factor or acute promyelocytic leukemia.

Table S6. Contingency table of validated local diagnosis and final study diagnosis by MDS subclass.

	Final Study Diagnosis								
Validated Local Diagnosis	**MDS-EB1**	**MDS-EB2**	**MDS-MLD**	**MDS-RSMLD**	**MDS-RSSLD**	**MDS-SLD**	**MDS-U**	**MDS Isolated del(5q)**	**Total**
MDS-EB1	36	3	1	1	0	0	0	0	**41**
MDS-EB2	1	29	0	1	0	0	0	0	**31**
MDS-MLD	3	0	33	4	1	1	0	0	**42**
MDS-RSMLD	2	1	2	20	3	0	0	0	**28**
MDS-RSSLD	0	0	0	5	10	1	0	0	**16**
MDS-SLD	0	0	9	0	0	3	1	1	**14**
MDS-U	3	1	6	4	0	2	8	0	**24**
MDS Isolated del(5q)	2	1	0	1	0	0	1	8	**13**
Total	**47**	**35**	**51**	**36**	**14**	**7**	**10**	**9**	**209**
Agreement rate	**77%**	**83%**	**65%**	**56%**	**71%**	**43%**	**80%**	**89%**	**70%**
Kappa (95% CI)									**0.65 (0.58-0.72)**

Diagnosis of MDS was based on The World Health Organization (2016) Classification. Total represents the number of participants assigned to MDS for the given assignment source. Agreement rate reflects the proportion of participants in each final study diagnosis category (MDS subtype) that had a matching validated local diagnosis. MDS-SLD = MDS with single lineage dysplasia; MDS-RSSLD = MDS with ring sideroblasts and single lineage dysplasia; MDS-MLD = MDS with multilineage dysplasia; MDS-RSMLD = MDS with ring sideroblasts and multi-lineage dysplasia; MDS-EB1 = MDS with excess blasts 1 (PB: 2-4%, BM: 5-9% blasts); MDS-EB2 = MDS with excess blasts 2 (PB: 5-10%, BM: 10-19% blasts); MDS Isolated del(5q) = MDS with isolated del(5q); MDS-U = MDS, unclassifiable; Other = not classified as MDS.