MAYO CLINIC

ORIGINAL ARTICLE

Bone Marrow Biopsy Operator Experience and Impact on Aspirate, Biopsy, and Ancillary Testing Quality

Lisa M. Marinelli, BS; Hong Fang, MD; Matthew T. Howard, MD; Curtis A. Hanson, MD; Joseph J. Haack, RN; Edward A. Eick, RN; Richard J. Allen, RN; David E. Ruffridge, RN; Colleen M. Byrne, RN; and Rebecca L. King, MD

Abstract

Objective: To assess the relationship between bone marrow (BM) biopsy operator experience and both specimen quality and ancillary testing utilization.

Patients and Methods: We evaluated all referred and in-house (IH) BM biopsy specimens obtained over a contiguous 6-week period from April 3, 2017, to May 19, 2017. The BM specimens were assessed for the length of interpretable marrow, and aspirates were assessed for the presence of spicules. Subgroup comparisons included IH BM obtained by a trained team of nurses within our institution, patients clinically referred (CR) to our institution with outside-obtained BM specimens, and outside pathologist-referred (PR) consultation cases. Ancillary study usage was compared between the first 100 cases of each group.

Results: A total of 1191 BM specimens were analyzed, including 600 IH, 288 CR, and 303 PR cases with biopsies and/or aspirates. The average interpretable biopsy lengths of IH, CR, and PR cases were 16.0 mm, 10.0 mm, and 7.0 mm, respectively ($P<.001$). World Health Organization–recommended length of 15 mm or more was achieved in 61.4%, 26.6%, and 19.1%, respectively ($P<.001$). Of the aspirates analyzed among IH, CR, and PR cases, 93%, 71.3%, and 73.5% contained spicules, respectively ($P<.001$). Use of immunohistochemistry, flow cytometry, karyotype, and fluorescence in situ hybridization was higher in CR and PR cases than in IH cases (all $P<.05$). The IH, CR, and PR cases used on average 1.5, 2.8, and 4.8 immunohistochemistry stains per case ($P<.001$).

Conclusion: Having a dedicated team of BM biopsy operators is likely one factor contributing to improved BM biopsy quality and a reduced need for ancillary testing.

 ■ Mayo Clin Proc Inn Qual Out 2018;2(3):241-247

Adequate bone marrow (BM) biopsy specimens and aspirates are essential to the diagnosis and management of many hematologic diseases. The World Health Organization (WHO) recommends that BM core biopsies include at least 15 mm of evaluable marrow and a 500-cell differential be counted "as close to the particle and as undiluted with blood as possible" in aspirates.[1] In addition, studies indicate that variations in biopsy quality have a negative effect on a pathologist's ability to make a definitive assessment. A considerable correlation has also been observed between length of interpretable BM and rate of a positive diagnosis in cohorts of diffuse large B-cell lymphoma, neuroblastoma, and other lymphomas and metastatic disease in BM.[2-4] Accurate assessment for myeloid malignancy, as well, relies heavily on having adequate aspirates and core biopsies to assess the architecture, cytology of the marrow, and differential count.[1,5,6]

Anecdotally, many factors may affect the quality of BM biopsies and aspirates including patient characteristics (age, body mass index [calculated as the weight in kilograms divided by the height in meters squared], disease state), operator experience, having a dedicated procedural site, use of sedation during the procedure, and needle gauge. However, results

From Mayo Clinic School of Medicine (L.M.M.), Department of Laboratory Medicine and Pathology (H.F.), Division of Hematopathology, Department of Laboratory Medicine and Pathology (M.T.H., C.A.H., R.L.K.), and Division of Hematology (J.J.H., E.A.E., R.J.A., D.E.R., C.M.B.), Mayo Clinic, Rochester, MN.

of published studies vary in terms of which of these factors show true association with sample quality. Operator experience, however, is one factor that has been repeatedly shown to impact BM specimen quality, and, as one recent study has shown, can be addressed with targeted education.[4,7,8]

At our institution, the vast majority of BM specimens are obtained by a trained team of nurses performing BM procedures daily; a small percentage of BM procedures are performed by resident and fellow trainees under the supervision and direction of the nurses. Approximately 75% of our BM procedures are performed under sedation in a dedicated outpatient procedural space, with the remaining performed at the bedside in hospital rooms. Laboratory technologists are present at the bedside to assess BM specimen quality in real time. Our institution averages approximately 4500 in-house (IH) BM procedures per year, divided among 4.5 full-time nurses, giving our operators an optimum level of experience. As a national tertiary referral center with a busy pathology consultation practice, we review BM specimens from a large number of other medical centers. This affords us a unique opportunity to compare specimen quality between BM biopsies and aspirates obtained under a consistent setting within our institution with those obtained at outside institutions, under heterogeneous circumstances. In doing so, we sought not only to assess the quality of our own BM biopsy practice but also to inform the hematology community atlarge about the state of clinical BM specimens today. This type of audit comparing the adequacy of BM biopsy specimens obtained in different clinical settings is lacking in the literature. Second, we compared the use of ancillary testing between case types to determine whether specimen quality (associated with operator experience) might impact the ordering of ancillary testing on BM specimens.

PATIENTS AND METHODS

Specimen Review

We evaluated all IH and referred BM biopsy specimens and aspirates seen in the Division of Hematopathology at Mayo Clinic over a contiguous 6-week period from April 3, 2017, to May 19, 2017. Cases were considered as 1 of 3 types: (1) IH: BM obtained within our institution; (2) clinically referred (CR): patients who were clinically being seen at our institution whose outside-obtained pathology case was sent for review in conjunction with their visit at Mayo Clinic; and (3) pathologist referred (PR): pathology consultations in which a patient's BM specimen was reviewed at our institution at the request of an outside physician (usually a pathologist).

Core biopsies were assessed for the length of interpretable BM on the hematoxylin and eosin–stained slide, excluding areas with crush or aspiration artifact when it completely obscured the marrow space. If multiple biopsies were obtained during the procedure, the lengths of each were added together. Cortical bone was excluded, but areas of subcortical marrow were included in measurements. We compared the median total length of evaluable marrow, percentage of biopsies with a length above the WHO-recommended minimum adequate length of 15 mm, percentage of biopsies below 5 mm, and aspirate quality among the 3 defined groups.[1] Aspirate smears were classified on the basis of an assessment of Wright-Giemsa (or similar)–stained slides as having spicules, marrow elements without spicules, or peripheral blood elements only. Because of logistic limitations, samples were not blinded to the reviewer with regard to origin (IH, PR, CR).

BM Aspirate and Biopsy Procedure

In-house BM procedures are performed primarily while the patient is under sedation and using standard technique with 15-gauge aspirate and 8-gauge Argon T-LOK manual core biopsy needles (Argon Medical Devices) (for further details on needle selection, see Supplemental Table 1, available online at http://mcpiqojournal.org/). After aspiration, the technologist evaluates a small sample on a slide for the presence of spicules, instructing the nurse to redirect up to 3 times to obtain an adequate sample. Once adequacy is verified, the technologist immediately prepares both unit preparation and direct smear slides, ensuring even distribution of units and correct length and thickness and decanting excess fluid. The technologist also provides real-time feedback in evaluation of biopsy core adequacy, redirecting the nurse when the core is less than 1 cm or contains predominantly fat

Mayo Clin Proc Inn Qual Out ■ September 2018;2(3):241-247 ■ https://doi.org/10.1016/j.mayocpiqo.2018.06.007
www.mcpiqojournal.org

TABLE 1. Comparison of Bone Marrow Biopsy and Aspirate Quality Between IH, CR, and PR Cohorts

Characteristic	IH	CR	PR	*P* value			
				Overall	IH vs CR	IH vs PR	CR vs PR
Biopsy							
Total No. of biopsies	596	274	298				
Length (mm), median	16.0	10.0	7.0	<.001	<.001	<.001	.006
Length of ≥15 mm (%)	61.4	26.6	19.1	<.001	<.001	<.001	.03
Length of ≤5 mm (%)	4.5	28.8	50.9	<.001	<.001	<.001	.04
Aspirate							
Total No. of aspirates	600	279	294	<.001	<.001	<.001	.42
With spicules (%)	93.0	71.3	73.5				
With marrow elements but no spicules (%)	6.5	28.0	24.8				
With blood only (%)	0.5	0.7	1.7				

CR = clinically referred; IH = in-house; PR = pathologist referred.

or cartilage (specimen floating in B5/formalin vial). The biopsy core is used to make touch preparations for cases with suspected insufficiency of spicules in aspirate slides.

Ancillary Study Analysis

The first 100 cases of each defined group (IH, CR, PR) (chosen consecutively on the basis of date of signout) were used as a subgroup for analysis of background characteristics and ancillary study usage. Background variables assessed included patient age and indication/diagnosis for BM analysis. Ancillary testing analysis included the number of immunohistochemical (IHC) stains obtained and whether flow cytometry, IHC, karyotype, fluorescence in situ hybridization, next-generation sequencing, and/or other molecular testing was performed.

Statistical Analyses

Statistical analyses were performed using the SAS biostatistical software JMP Pro 13.0.0 (SAS Institute). Categorical variables were compared using χ^2 analysis. Continuous variables were compared using Student *t* test or median tests. A *P* value of less than .05 was considered statistically significant.

RESULTS

BM Biopsy and Aspirate Quality

A total of 1191 BM biopsy specimens were analyzed, including 600 IH, 288 CR, and 303 PR cases with biopsies and/or aspirates. These cases included a total of 1168 biopsies and 1173 aspirates. The CR and PR cases were referred from 177 and 179 different institutions, respectively, resulting in an average of only 1 to 2 cases coming from any single institution.

The median interpretable biopsy lengths of IH, CR, and PR cases were 16.0 mm, 10.0 mm, and 7.0 mm, respectively (Table 1). The IH median biopsy lengths were significantly greater than those of CR and PR cases ($P<.001$), and CR biopsy lengths were significantly greater than those of PR cases ($P=.006$). The WHO-recommended length of 15 mm or more was achieved in 61.4%, 26.6%, and 19.1%, respectively ($P<.001$). A significantly greater proportion of PR cases (50.9%) were 5 mm or less in length, in comparison with CR (28.8%) and IH cases (4.5%) (all $P<.05$).

Of the aspirates analyzed among IH, CR, and PR cases, 93.0%, 71.3%, and 73.5% contained spicules, respectively. The IH aspirates had a significantly greater proportion with spicules than those of CR and PR ($P<.001$), but the difference between CR and PR aspirates was not statistically significant ($P=.42$).

Patient Characteristics

Among the first 100 patients of each group, the median ages were 64.0, 70.5, and 70.0 years, respectively ($P<.001$ [IH vs CR and PR]). Of the indications for BM analysis identified in Table 2, the most common among IH cases, plasma cell disorders (40%), occurred about twice as frequently as in the CR (23%) and PR (17%) cases. The PR cases tended to include a higher frequency of cytopenias/rule

TABLE 2. Background Characteristics of Subgroups Used for Ancillary Study Analysis

Characteristic	IH (N=100)	CR (N=100)	PR (N=100)	*P* value			
				Overall	IH vs CR	IH vs PR	CR vs PR
Age (y), median	64.0	70.5	70.0	<.001	<.001	<.001	.34
Indication for BM (%)				.04	.08	.02	.60
Cytopenia, r/o MDS	13	16	24				
Lymphoma or staging	23	21	26				
MDS	7	7	10				
MPN	3	11	4				
Acute leukemia	6	5	4				
Metastasis	3	2	1				
Cytosis, r/o MPN	3	6	5				
Plasma cell	40	23	17				
MDS/MPN	2	8	8				
Other	0	1	1				

CR = clinically referred; IH = in-house; MDS = myelodysplastic syndrome; MPN = myeloproliferative neoplasm; PR = pathologist referred; r/o = rule out.

out myelodysplastic syndrome (24%) in comparison with the CR (16%) and IH (13%) cases, and the CR cases tended to include a higher frequency of myeloproliferative neoplasm cases. Overall, only differences between IH and PR cases showed statistical significance.

Ancillary Study Analysis

Among the 100 cases of each group assessed for ancillary testing, no significant difference in quality metrics was observed between the subgroup and the respective study group (see Supplemental Table 2, available online at http://mcpiqojournal.org/). The percentage of cases using IHC, flow cytometry, and/or karyotype was higher in CR and PR cases than in IH cases (all $P<.05$) (Table 3). The CR cases also used fluorescent in situ hybridization more frequently than IH cases ($P=.002$). Use of next-generation sequencing and other molecular tests was not statistically different between groups. In comparison of CR and PR cases, CR cases more frequently used karyotype ($P=.008$), but ancillary testing utilization was otherwise similar. The average number of IHC stains for IH, CR, and PR cases was 1.5, 2.8, and 4.8, respectively, with PR cases using statistically more stains than both IH and CR ($P\leq.01$ for both). The number of IHC stains correlated with biopsy length, with a Spearman correlation coefficient of −0.12 ($P=.04$). The average

TABLE 3. Comparison of Ancillary Testing Between Subgroups[a]

Ancillary test	IH (N=100)	CR[b] (N=100)	PR[b] (N=100)	*P* value			
				Overall	IH vs CR	IH vs PR	CR vs PR
No. of IHC stains, mean	1.5	2.8	4.8	<.001	.003	<.001	.008
With ancillary testing (%)							
Flow cytometry	73	83	90	.007	.09	.002	.15
IHC	48	69	63	.008	.003	.03	.37
Cytogenetics/karyotype	37	72	54	<.001	<.001	.02	.008
FISH	23	44	31	.006	.002	.20	.06
NGS	2	2	3	.86	>.99	.65	.65
Other molecular	15	19	17	.75	.45	.70	.71

[a]CR = clinically referred; IH = in-house; IHC = immunohistochemistry; FISH = fluorescence in situ hybridization; NGS = next-generation sequencing; PR = pathologist referred.
[b]The CR and PR ancillary studies by both outside and Mayo Clinic pathologists included.

Mayo Clin Proc Inn Qual Out ■ September 2018;2(3):241-247 ■ https://doi.org/10.1016/j.mayocpiqo.2018.06.007
www.mcpiqojournal.org

number of ancillary tests used for adequate biopsies by WHO standards (≥15 mm) was 2.2, in comparison with 2.7 for inadequate biopsies (P=.008).

DISCUSSION

This study sought to assess the differences in specimen quality between BM biopsies performed IH and those performed at various outside institutions and evaluated as part of a BM pathology consultation practice. We hypothesize that having a trained team of professionals performing biopsies, alongside laboratory technologists assessing specimen quality in real time, leads to better sample quality among our IH practice. Results of our analysis support this assertion and indicate that IH biopsies achieved the WHO-suggested standard of 15 mm or more over twice as often than in both PR and CR cases and had superior aspirate quality as well. This suggests that the overarching hematology and hematopathology communities need to focus on this issue and set expectations on achieving this critical standard.

Interestingly, the median biopsy length of CR cases, at 10.0 mm, was significantly greater than that of PR cases, at 7.0 mm (P=.006), and 50.5% of PR cases had a biopsy length of 5 mm or less in comparison with 28.8% of CR cases (P=.04). These findings raise the question of whether the inferior specimen quality of PR cases contributed to diagnostic difficulty in these cases, leading to the need for pathologist consultation at our institution.

Ancillary test usage was higher for CR and PR cases than for IH cases. In addition, PR cases used, on average, 2 more IHC stains than CR cases. We know from experience that when morphology is suboptimal, IHC may be needed to properly exclude pathology, such as an atypical lymphoid infiltrate or blast population that may be obscured by the poor quality. Within a cohort of lymphomas, for example, a preprocessing biopsy length of 17 to 20 mm rendered a positive diagnosis rate of 68.9%, whereas a preprocessing biopsy length of less than 16 mm rendered a 40.3% positive diagnosis rate.[9] Moreover, without adequate hematoxylin and eosin morphology, IHC may be necessary simply to identify the presence of each normal cell line and assess relative proportions.

Our results support previous studies that point toward operator experience as a key factor in BM specimen quality. One study found that among operators performing a total of 6 biopsies over a 1-year study period, the length of interpretable marrow was on average 5 mm, whereas specimens from operators performing 51 to 75 biopsies in the year averaged 9 mm.[4] Audits of specimens obtained by biopsy operators with varied levels of training at other institutions have reported an average length of evaluable marrow of 7.4 mm and 13.1 mm, in comparison with 16.0 mm obtained by the trained nurse team within our institution.[4,7] In addition, the percentage of particulate aspirates observed at an outside institution from patients with multiple myeloma, 75%, was similar to the 71.3% and 73.5% of CR and PR aspirates but less than the 93.0% of aspirates obtained on IH cases in our study.[10] A report from the University of Wisconsin substantiates the findings of this study, reporting that the individual operator was the most important factor affecting biopsy length, with the presence of sedation and indication for biopsy having no considerable effect. Encouragingly, education and peer comparisons were used to effectively improve specimen quality.[8]

When the patient characteristics among the 3 study groups were compared, some interesting differences did emerge. The average age of IH patients was significantly lower than that of CR patients (64.0 vs 70.5 years). Although a previous study described a negative association between patient age and interpretable biopsy length (r=−0.16; P<.001), this correlation coefficient was small, and the 7-year age difference is unlikely to account for the 6- to 9 -mm difference in biopsy length between groups in our study.[7] The IH patients included twice the number of plasma cell disorders and fewer evaluations for myeloproliferative neoplasms and cytopenias than the CR and PR cases, suggesting some inherent differences in these 3 patient populations. Rudzki et al[7] stratified BM quality by indication, and observed the poorest quality of biopsies among patients with multiple myeloma and the greatest quality among those with chronic myeloproliferative disorders. Although we did not specifically evaluate specimen quality by diagnosis in our cohort, their results would suggest that BM biopsy indication is not the sole cause for increased biopsy length in our IH cases.

Notwithstanding the direct impact on patient care, suboptimal BM quality has implications for health care costs as one factor that may contribute to increased ancillary testing, or increased utilization of pathology consultation services.[11] As such, future research into the health care implications of having a trained team of BM biopsy operators is warranted, specifically as it pertains to the additional cost of increased testing, nonmonetary cost surrounding patient care, and the cost of having a trained operator team.

Although these findings are intriguing, a retrospective study such as this one has several limitations. Although we identified a correlation between ancillary testing utilization and BM quality, these results do not prove that suboptimal specimen quality caused increased ancillary testing. For example, increased diagnostic complexity in PR patients may be an additional factor contributing to additional ancillary testing in this cohort.

In addition, the PR cohort is subject to referral bias, as inferior specimen quality would be expected to contribute to the need for pathology consultation. Comparison of specimen quality between IH and CR cases, however, mitigates this bias, because CR cases were not selectively referred on the basis of histologic challenges. Future research is needed to eliminate potential confounding factors affecting ancillary testing utilization.

We have not formally studied operator experience at the outside institutions whose BM specimens we have reviewed. We acknowledge that several factors in addition to operator experience may contribute to the observed poorer specimen quality in CR and PR cases. This limits our ability to definitively identify operator experience and training as being the sole explanation for lower specimen quality in these groups. Anecdotally, however, we know that most outside-obtained BM specimens are not collected by a specially trained team performing these procedures daily. Although patient feedback was not collected formally as part of this study, we have also received very positive patient feedback on their IH BM procedure experience as compared with those at outside institutions.

As stated above, because of logistic limitations, it was not possible for the reviewer to be blinded to the type of case, creating the potential for observer bias.

Finally, this study cannot account for differences in equipment used for specimen processing and staining that were used at outside institutions. Although quality in preanalytical specimen processing is certainly critical to producing adequate diagnostic material, we could not evaluate this directly in our study.

CONCLUSION

These data suggest that having a dedicated, trained team of BM biopsy operators along with real-time adequacy assessment is a major factor contributing to superior BM biopsy and aspirate quality at our institution. In addition, a shorter biopsy length correlates with the need for referral to our institution, likely due to diagnostic difficulty caused by poorer specimen quality. We hypothesize that suboptimal specimen quality may be one of several factors that contribute to increased ancillary testing. Having an adequate sample is an essential part of BM diagnostics, and increased attention should be paid to ensuring quality samples via proper education and training. Future research into health care outcomes and feasibility of trained BM biopsy operators is warranted.

SUPPLEMENTAL ONLINE MATERIAL

Supplemental material can be found online at http://mcpiqojournal.org/. Supplemental material attached to journal articles has not been edited, and the authors take responsibility for the accuracy of all data.

Abbreviations and Acronyms: **BM** = bone marrow; **CR** = clinically referred; **IH** = in-house; **IHC** = immunohistochemical; **PR** = pathologist referred; **WHO** = World Health Organization

Potential Competing Interests: The authors report no competing interests.

Publication dates: Received for publication May 23, 2018; revisions received June 24, 2018; accepted for publication June 27, 2018.

Correspondence: Address to Rebecca L. King, MD, Division of Hematopathology, Department of Laboratory Medicine and Pathology, Mayo Clinic, 200 First St SW, Rochester, MN 55905 (king.rebecca1@mayo.edu).

REFERENCES

1. Swerdlow SH, Campo E, Harris NL, et al. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*. 4th ed., Vol 2. Lyon, France: World Health Organization; 2017.

2. Campbell JK, Matthews JP, Seymour JF, Wolf MM, Juneja SK; Australasian Leukaemia Lymphoma Group. Optimum trephine length in the assessment of bone marrow involvement in patients with diffuse large cell lymphoma. *Ann Oncol*. 2003; 14(2):273-276.
3. Reid MM, Roald B. Adequacy of bone marrow trephine biopsy specimens in children. *J Clin Pathol*. 1996;49(3):226-229.
4. Bishop PW, McNally K, Harris M. Audit of bone marrow trephines. *J Clin Pathol*. 1992;45(12):1105-1108.
5. Arber DA, Borowitz MJ, Cessna M, et al. Initial diagnostic workup of acute leukemia: guideline from the College of American Pathologists and the American Society of Hematology. *Arch Pathol Lab Med*. 2017;141(10):1342-1393.
6. Foucar K, Reichard K, David C. *Bone Marrow Pathology*. 3rd ed. Chicago, IL: Americal Society for Clinical Pathology; 2010.
7. Rudzki Z, Partyla T, Okon K, Stachura J. Adequacy of trephine bone marrow biopsies: the doctor and the patient make a difference. *Pol J Pathol*. 2005;56(4):187-195.
8. Yang RK, Nazeef M, Patel SS, et al. Improving bone marrow biopsy quality through peer discussion and data comparisons: a single institution experience [published online ahead of print March 25, 2018]. *Int J Lab Hematol*. https://doi.org/10.1111/ijlh.12804.
9. Goyal S, Singh UR, Rusia U. Comparative evaluation of bone marrow aspirate with trephine biopsy in hematological disorders and determination of optimum trephine length in lymphoma infiltration. *Mediterr J Hematol Infect Dis*. 2014;6(1):e2014002.
10. Charles KS, Winfield DA, Angel C, Goepel J. Audit of bone marrow aspirates and trephine biopsies in multiple myeloma—a single centre study. *Clin Lab Haematol*. 2004;26(6):403-406.
11. Robinson A. Rationale for cost-effective laboratory medicine. *Clin Microbiol Rev*. 1994;7(2):185-199.