Automated decision tree to evaluate genetic abnormalities when determining prognostic risk in acute myeloid leukemia

Acute myeloid leukemia (AML) is a heterogeneous disease characterized by multiple genetic mutations and cytogenetic abnormalities.[1-3] Our understanding of AML pathobiology has improved significantly in recent decades, including the ability to predict which patients have a high probability of relapse after chemotherapy. The European LeukemiaNet (ELN) developed guidelines,[4] which were revised in 2017,[5] to guide physicians' decisions as to how to treat and manage AML. However, these guidelines require molecular tests that are not routinely performed at all medical centers,[6,7] limiting their broad applicability, and the complex pattern of co-occurrence among prognostic abnormalities is poorly understood. We created an automated decision tree to stratify patients into prognostic risk categories according to the ELN guidelines. We evaluated 1682 patient records from three datasets, conducting a detailed analysis of the co-occurrence of mutations and karyotypic abnormalities and determining the prognostic impact of each test result on accurately categorizing patients. Our results have implications not only in understanding the real-world distribution of ELN-defined prognostic abnormalities, but

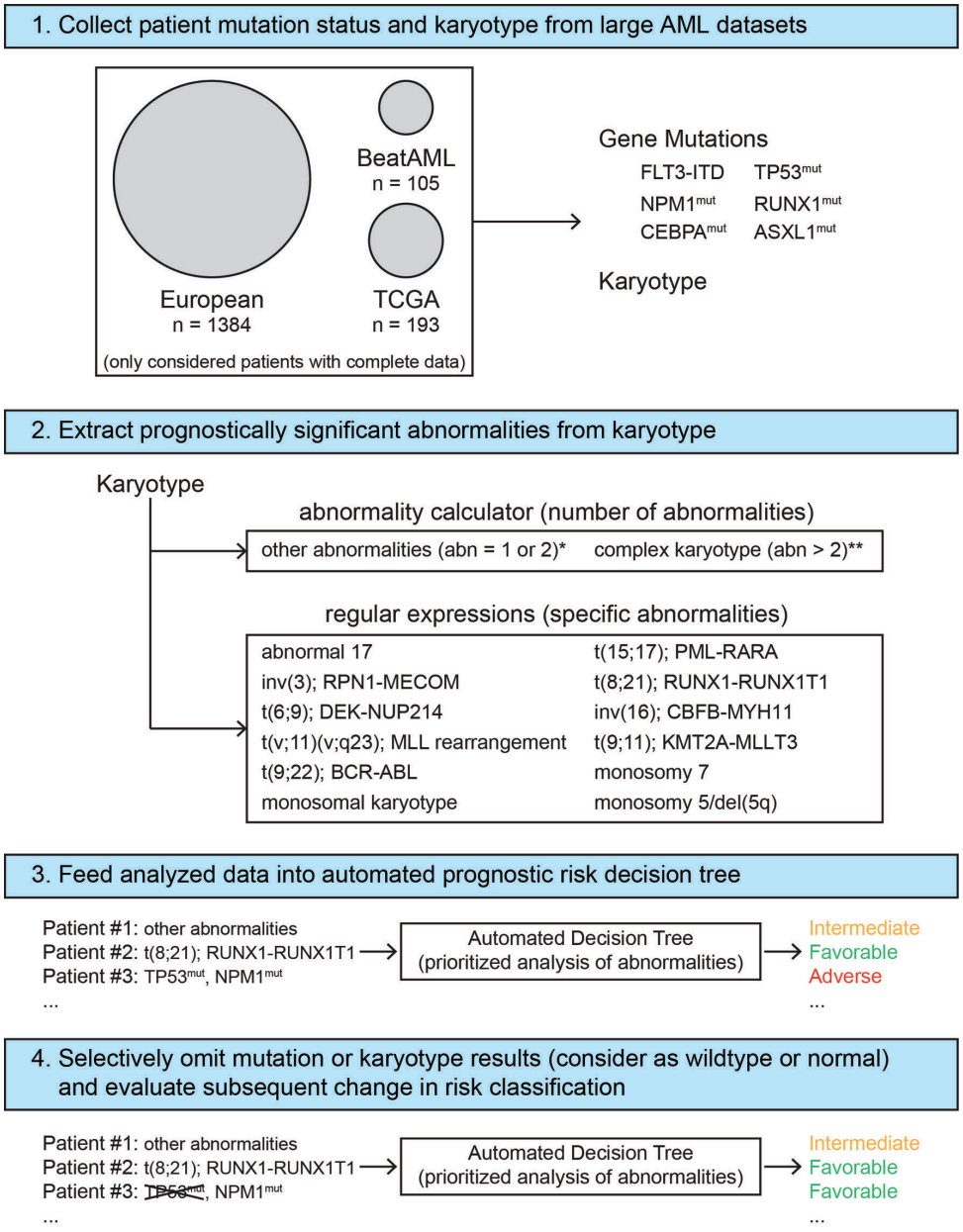


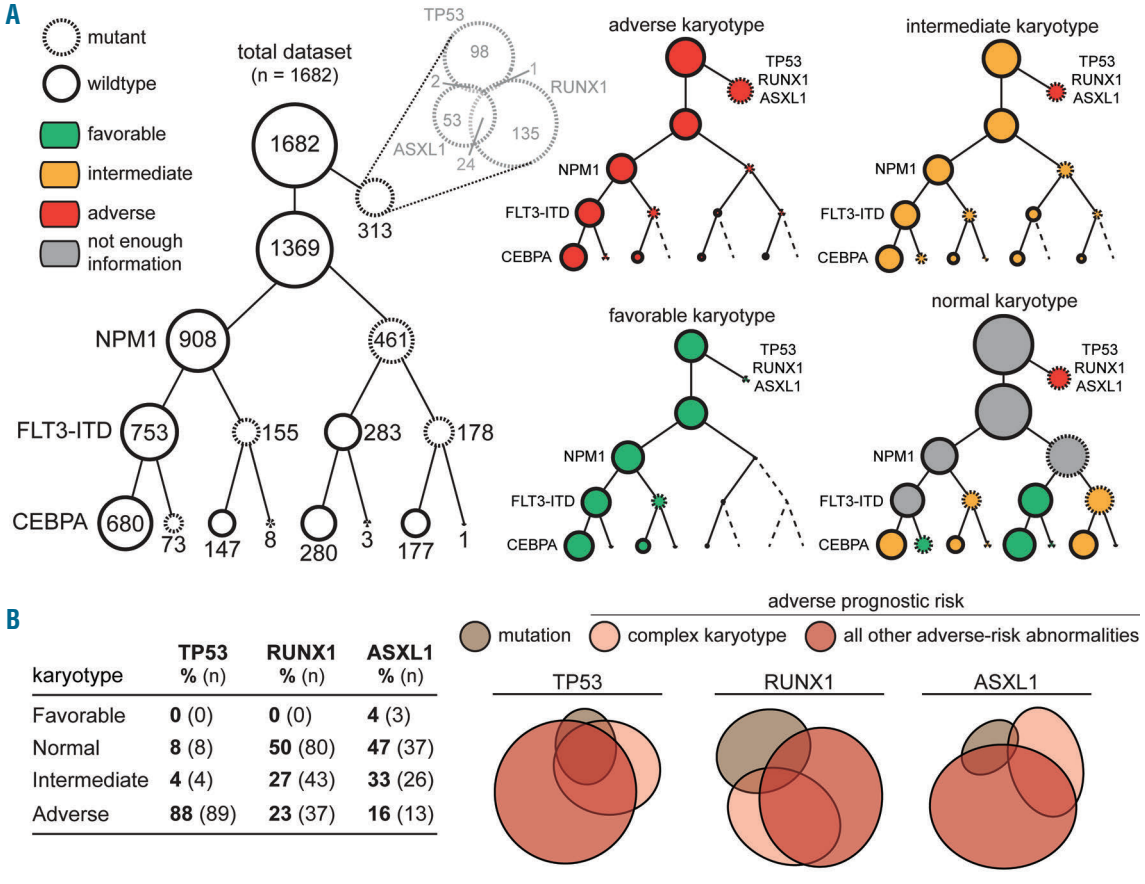
Figure 1. Experimental design to determine the impact of genetic mutation co-occurrence on prognostic risk in acute myeloid leukemia (AML). Patient information from three AML datasets were combined and mutation status and karyotype were extracted. These records were analyzed using an automated script that classified each patient based on their European LeukemiaNet (ELN) prognostic risk score, and the prognostic impact of each genetic test was evaluated. *"Other abnormalities" is defined as the presence of one or two abnormalities. **"Complex karyotype" is defined as the presence of three or more abnormalities.

also in categorizing patients based on incomplete data.

The ELN guidelines stratify patients into three prognostic risk categories (Favorable, Intermediate, and Adverse) by combining the presence of karyotypic abnormalities with genetic mutations (in *FLT3, NPM1, CEBPA, TP53, RUNX1,* and *ASXL1*). To understand the impact of genetic abnormality co-occurrence on prognostic risk, we combined data from three large AML patient datasets (n=1682 patients): the German-Austrian AML Study Group (AMLSG) cohort (n=1384),[3,8] The Cancer Genome Atlas (n=193),[1] and the Beat AML Wave One dataset (n=105) (Figure 1). A complete description of the methods of data analysis, including a comprehensive R Markdown file, is available on GitHub (*https://github.com/WatanabeSmith/AML_ELN_PrognosticRiskClassification*) and in *Online Supplementary Methods*. We filtered out entries without complete karyotype and mutation data, including karyotypes that did not follow standardized nomenclature.[9] The presence or absence of mutations were extracted from these datasets. We extracted prognostically significant cytogenetic abnormalities, along with complex and monosomal karyotypes, using regular expressions and an automated abnormality calculator. The complete dataset is available in *Online Supplementary Table S1.*

To evaluate the pattern of mutational co-occurrence in these patients, we divided the prognostic risk factors by those originating either from cytogenetic testing or gene mutation testing. We subdivided the patient karyotypes into four groups: Favorable, Intermediate, and Adverse (all based on the presence of prognostic cytogenetic abnormalities alone), and Normal (which contain no abnormalities). We sorted patients within each karyotype subdivision by the presence of adverse *TP53/RUNX1/ASXL1* mutations, and the combined presence or absence of *NPM1, FLT3-ITD*, and *CEBPA* mutations (Figure 2A). We confirmed that the co-occurrence patterns from each dataset closely matched one another (*Online Supplementary Figure S1A-C*).

Overall, we observed that *NPM1, FLT3-ITD,* and *CEBPA* mutations are enriched in a Normal karyotype background, especially *NPM1* and *CEBPA*, highlighting their value in providing prognostic information in cytogenetically normal AML. In addition, *TP53* mutations, in contrast to mutations in *RUNX1* and *ASXL1*, more frequently co-occur with complex karyotypes and other adverse-risk markers (Figure 2B). Therefore, within the context of broadly assigning prognostic risk to AML patients, *TP53* mutation status offers little additional prognostic information besides that already provided by



karyotype	TP53 % (n)	RUNX1 % (n)	ASXL1 % (n)
Favorable	**0** (0)	**0** (0)	**4** (3)
Normal	**8** (8)	**50** (80)	**47** (37)
Intermediate	**4** (4)	**27** (43)	**33** (26)
Adverse	**88** (89)	**23** (37)	**16** (13)

Figure 2. Mutational co-occurrence within cytogenetic prognostic risk categories in 1682 acute myeloid leukemia patients. Patients were grouped into four categories according to their karyotype: Adverse, Intermediate, or Favorable (which contain abnormalities that categorize them accordingly, based on the 2017 European LeukemiaNet guidelines); and Normal (which contains no abnormalities). (A) Patients were categorized as having TP53, RUNX1, or ASXL1 mutations, and, for the remaining patients, subdivided by sequential presence or absence of NPM1 mutations, FLT3-ITD, and CEBPA mutations. The area of each circle is proportional to the number of patients within that category. (B) The frequency of TP53, RUNX1, and ASXL1 was tabulated across different karyotype risk categories. Proportional Venn diagrams were drawn showing the overlap between the presence of complex karyotype, an adverse risk gene mutation, and all other adverse cytogenetic abnormalities.

cytogenetic testing. Notably, there are very few patients with adverse/intermediate risk mutations and favorable karyotypic abnormalities, who, according to the ELN conventional care regimens for AML, would normally go forward for transplant.[5]

Next, to categorize the patients into prognostic risk categories, we created an automated decision tree based on the newly published ELN guidelines (Figure 3A). It sequentially processes mutation and cytogenetic data, and returns the corresponding Favorable, Intermediate, and Adverse prognostic risk category. (Cases of acute promyelocytic leukemia were classified as Favorable risk

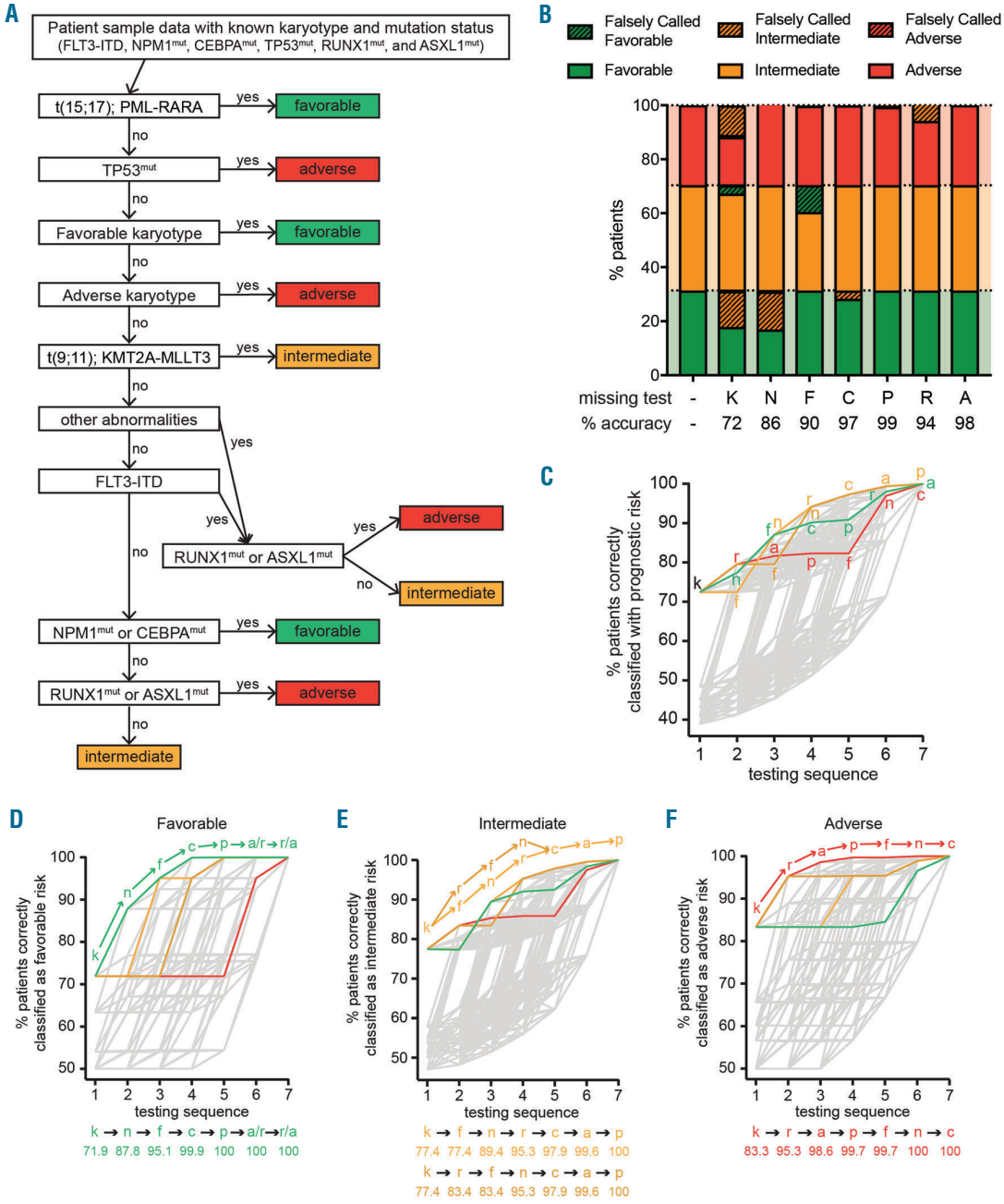


Figure 3. Automated decision tree determines that genetic abnormalities have different prognostic significance individually and sequentially in acute myeloid leukemia patients. (A) Overview of automated decision tree to determine acute myeloid leukemia prognostic risk. After automatically identifying the prognostically significant karyotypic abnormalities, the decision tree returns the assigned prognostic risk category based on the hierarchy of abnormalities described in the European LeukemiaNet (ELN) guidelines (see *Online Supplementary Methods*). (B) Prognostic impact of individual genetic abnormalities, as measured by the omission of data from each of the seven prognostically significant genetic test and the assignment of patients into prognostic risk categories. k: karyotype; f: FLT3-ITD; n: NPM1; c: CEBPA; p: TP53; r: RUNX1; a: ASXL1. (C) Sequential analysis of every possible arrangement of the genetic tests to best determine overall prognostic risk. To mimic an unknown test result, each patient was considered to be wild type for gene mutations and have normal karyotype. For each test in a particular sequence, the prognostic risk category was calculated after successively adding in the test results, and the percentage of patients with a correctly called prognostic risk was recorded. Every possible test sequence (5040 total) was graphed and the optimized sequence to identify patients with favorable (green), intermediate (yellow and dark yellow), and adverse (red) risk were labeled. For all three categories, the optimized sequence began with karyotype, which is labeled in black. (D-F) The sequential analysis for all possible arrangement of tests to best identify patients with (D) favorable, (E) intermediate, and (F) adverse prognostic risk. For these sequences, the percentage of correctly called prognostic risk patients was normalized due to the differences in population size across categories.

based on separate ELN guidelines.[10]) Notably, the decision tree does not consider *FLT3*-ITD allele ratio (*FLT3*-ITD^{high} *vs.* *FLT3*-ITD^{low}) or biallelic CEBPA mutations because those data were not available for many of the patient records (see *Online Supplementary Methods*).

To understand the significance of each test (*NPM1, FLT3-ITD, CEBPA, TP53, RUNX1, ASXL1*, and cytogenetics) in assigning prognostic risk category, we analyzed our dataset using the decision tree while selectively censoring the presence of individual mutations or all karyotypic abnormalities. Censored test results were assumed to be wild-type or cytogenetically normal, simulating an unknown test result. Patients were assigned ELN risk categories based on incomplete data from all possible combinations of genetic test results. These risk categories were compared to the actual risk categories based on complete information to determine the accuracy of categorizing with partial information (Figure 3B). We similarly censored test results in groups that might have more clinical relevance (e.g. all adverse mutations considered at once, simulating results from a mutation panel) (*Online Supplementary Figure S2*).

Karyotype is the most informative single test, as expected,[11] although 72% of patients can still be correctly categorized without cytogenetic information, using only the mutation tests. In most cases, a combination of *NPM1* mutational and *FLT3*-ITD status are the next most impactful tests. Omitting *NPM1* or *FLT3*-ITD results has a similar prognostic impact, but the type of misclassification shifts from incorrectly categorizing favorable-risk patients to incorrectly categorizing intermediate-risk patients, respectively. In addition, *CEBPA* and *RUNX1* mutations each provide significant improvements in classifying favorable- and adverse-risk patients, respectively, while *ASXL1* and *TP53* mutations rarely impact prognostic categorization. Although most genetic mutation tests in these guidelines can be combined in a single mutational panel (greatly decreasing the cost of detecting mutations in additional genes), *NPM1* mutations and *FLT3*-ITDs are identified by PCR amplification and electropherogram, respectively, requiring separate tests and equipment.

Finally, we determined the optimal sequence of the seven tests required for ELN risk classification, in which the next test in the sequence results in the greatest increase in classification accuracy. Each genetic test was evaluated on its ability to classify patients relative to a single risk category (e.g. call patients correctly as Favorable or simply Not Favorable) or in all categories using balanced accuracy (to account for differences in population size) or overall accuracy, respectively. We displayed all 5040 possible sequences for the overall classification and for each risk category, highlighting the optimum sequence (Figure 3C-F and *Online Supplementary Table S2*).

While the test with the greatest impact on prognosis is consistently karyotype, the test with the second biggest impact varies widely depending on risk categories. Favorable-risk patients can be identified with 95% accuracy with the addition of *NPM1* and *FLT3*-ITD, and 99.9% accuracy with the further addition of *CEBPA*, suggesting that significant prognostic impact can be achieved from relatively little information or despite missing information. For intermediate prognostic risk, there were two optimized sequences: one (Intermediate_R) with *RUNX1* as the second test, resulting in more correctly classified patients only after that second test, and another (Intermediate) with *FLT3*-ITD and *NPM1* as the second and third tests, resulting in more correctly classified patients after that third test (since *NPM1* mutations only confer Favorable risk in patients without *FLT3*-ITD). Adverse risk is 95% determined by karyotype and *RUNX1* mutation status, demonstrating the significance and prognostic impact of *RUNX1* over other mutational tests. The optimal overall sequence closely resembles that of intermediate risk and highlights the minimal impact of *ASXL1* and *TP53* on overall ELN risk classification. Because the incidence of *TP53* and *ASXL1* mutations increases with age,[12] we separated our dataset into younger (<60 years old; n=1311) and older (≥60; n=371) patients and evaluated the difference in the prognostic significance of those mutations between both subgroups. We found that the difference in optimized sequences between subgroups was only minimal, and the proportion of patients incorrectly classified as a result of missing *TP53* and *ASXL1* mutation data was only nominally increased in older patients (*Online Supplementary Figure S3* and *Online Supplementary Methods*).

Overall, our automated decision tree has potential applications outside of this investigation. While there are a multitude of clinical or genetic factors outside of the ELN guidelines that have prognostic significance, we focused only on the genetic abnormalities described in the guidelines. The decision tree, however, does allow other researchers to interrogate how additional factors are distributed across the prognostic risk categories.

Moreover, while novel AML treatment strategies will probably include comprehensive mutation testing upon diagnosis, current treatment guidelines remain hampered by financial and technical challenges. This is especially true in low- and middle-income countries (LMIC), where prognostic classification is often made in the presence of incomplete data. A recent Brazilian study assessing the relevance of the 2008 ELN guidelines in LMIC found that 26% of AML patients in their cohort were randomly missing cytogenetic or complete mutational data.[13] This approach could also enable clinicians and healthcare providers in LMIC to evaluate and customize their diagnostic testing depending on the availability of resources and the individual circumstances of each patient. Our analysis enables data-driven decision-making by providing a framework for understanding which genetic features hold the most weight in prognostication given their relative prevalence and co-occurrence in real-world AML datasets. Accordingly, we have made all of our programming scripts and datasets publicly available at *https://github.com/WatanabeSmith/AML_ELN_PrognosticRiskClassification* to encourage further research.

Ultimately, AML prognostic risk involves a constellation of abnormalities, each one carrying more or less weight depending on the abnormalities with which it co-occurs. This analysis provides an important first step in applying this understanding to the newly revised ELN guidelines for AML.

Kevin Watanabe-Smith,[1] Brian J. Druker,[1,2,3] Jeffrey W. Tyner,[1,4] and David K Edwards V[4]

[1]Knight Cancer Institute, Oregon Health & Science University; [2]Howard Hughes Medical Institute; [3]Division of Hematology and Medical Oncology, Oregon Health & Science University and [4]Department of Cell, Developmental & Cancer Biology, Oregon Health and Science University, Portland, OR, USA

Acknowledgments: the authors would like to thank Ted Braun, Jessica Leonard, and Elie Traer for providing important clinical expertise; Marc Loriaux for assisting with Beat AML; and Julia Maxson, Amy Yates, and Brian Foley for helpful guidance and feedback.

Funding: BJD received funding support as an Howard Hughes Medical Institute Investigator. This material is based upon work supported by the Leukemia & Lymphoma Society - SCOR 7005-11 (BJD) and Beat AML (BJD and JWT) - and by the National Science Foundation Graduate Research Fellowship Program (DGE-1448072; DKEV). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Correspondence: edwadavi@ohsu.edu
doi:10.3324/haematol.2018.190926

Information on authorship, contributions, and financial & other disclosures was provided by the authors and is available with the online version of this article at www.haematologica.org.

References

1. Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med. 2013;368(22):2059-2074.
2. Kihara R, Nagata Y, Kiyoi H, et al. Comprehensive analysis of genetic alterations and their prognostic impacts in adult acute myeloid leukemia patients. Leukemia. 2014;28(8):1586-1595.
3. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic Classification and Prognosis in Acute Myeloid Leukemia. N Engl J Med. 2016;374(23):2209-2221.
4. Döhner H, Estey EH, Amadori S, et al. Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. Blood. 2010;115(3):453-474.
5. Döhner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood. 2017;129(4):424-447.
6. Lin TL, Williams T, He J, et al. Rates of complete diagnostic testing for patients with acute myeloid leukemia. Cancer Med. 2015; 4(4):519-522.
7. Pollyea DA, George T, Foucar KM, et al. Molecular genetic testing patterns for patients with newly diagnosed acute myeloid leukemia (AML) enrolled in the CONNECT MDS/AML disease registry. J Clin Oncol. 2017;35(15_suppl):7022-7022.
8. Gerstung M, Papaemmanuil E, Martincorena I, et al. Precision oncology for acute myeloid leukemia using a knowledge bank approach. Nat Genet. 2017;49(3):332-340.
9. McGowan-Jordan J, Simons A, Schmid M. ISCN 2016: An International System for Human Cytogenomic Nomenclature. Basel, New York: Karger, 2016.
10. Sanz MA, Grimwade D, Tallman MS, et al. Management of acute promyelocytic leukemia: recommendations from an expert panel on behalf of the European LeukemiaNet. Blood. 2009;113(9):1875–1891.
11. O'Donnell MR, Abboud CN, Altman J, et al. Acute myeloid leukemia. J Natl Compr Canc Netw. 2011;9(3):280-317.
12. Tsai C-H, Hou HA, Tang JL, et al. Genetic alterations and their clinical implications in older patients with acute myeloid leukemia. Leukemia. 2016;30(7):1485-1492.
13. Benicio MT de L, Ribeiro AFT, Américo AD, et al. Evaluation of the European LeukemiaNet recommendations for predicting outcomes of patients with acute myeloid leukemia treated in low- and middle-income countries (LMIC): A Brazilian experience. Leuk Res. 2017; 60:109-114.