blood advances

American Society of Hematology
2021 L Street NW, Suite 900,
Washington, DC 20036
Phone: 202-776-0544 | Fax 202-776-0545
bloodadvances@hematology.org

Tumor mutational load is prognostic for progression to therapy among high-count monoclonal B-cell lymphocytosis (HCMBL)

Tracking no: ADV-2023-012242R1

Geffen Kleinstern (Mayo Clinic, United States) Nicholas Boddicker (Mayo Clinic, United States) Daniel O'Brien (Mayo Clinic, United States) Cristine Allmer (Mayo Clinic, United States) Kari Rabe (Mayo Clinic, United States) Aaron Norman (Mayo Clinic, United States) Rosalie Griffin (Mayo Clinic, United States) Huihuang Yan (Mayo Clinic, United States) Tao Ma (Mayo Clinic, United States) Timothy Call (Mayo Clinic, United States) Laura Bruins (Mayo Clinic,) Sochilt Brown (Mayo Clinic Arizona, United States) Cecilia Bonolo de Campos (Mayo Clinic, United States) Curtis Hanson (Mayo Clinic, United States) Jose Leis (Mayo Clinic, United States) Wei Ding (Mayo Clinic, United States) Celine Vachon (Mayo Clinic, United States) Neil Kay (Mayo Clinic, United States) Christopher Oakes (The Ohio State University, United States) Alexander Parker (University of Florida, United States) Danielle Brander (Duke University Health System, United States) J Brice Weinberg (Duke Cancer Institute, United States) Richard Furman (Weill Medical College of Cornell University, United States) Tait Shanafelt (Stanford University School of Medicine, United States) James Cerhan (Mayo Clinic College of Medicine, United States) Sameer Parikh (Mayo Clinic, United States) Esteban Braggio (Mayo Clinic Arizona, United States) Susan Slager (Mayo Clinic, United States)

Abstract:
HCMBL is a precursor condition to chronic lymphocytic leukemia (CLL). We have shown that among individuals with HCMBL the CLL-International Prognostic Index (CLL-IPI) is prognostic for time-to-first therapy (TTFT). Little is known about the prognostic impact of somatically mutated genes among individuals with HCMBL. We sequenced DNA from 371 HCMBL individuals using a targeted sequencing panel of 59 recurrently mutated genes in CLL to identify high-impact mutations. We compared the sequencing results to that of our treatment-naïve CLL cohort(N=855) and employed Cox regression to estimate hazard ratios and 95% confidence intervals (CI) for associations with TTFT. Compared to CLL, the frequencies of any mutated genes were lower in HCMBL (70% versus 52%). At 10-years, 37% of HCMBL individuals with any mutated gene had progressed requiring treatment compared to 10% among HCMBL individuals with no mutations; this led to 5.4-fold shorter TTFT (95%CI:2.6-11.0) among HCMBL with any mutated gene versus none, independent of CLL-IPI. When considering individuals with low-risk of progression according to CLL-IPI, HCMBL individuals with any mutations had 4.3-fold shorter TTFT (95%CI:1.6-11.8) versus those with none. Finally, when considering both CLL-IPI and any mutated gene status, we observed HCMBL individuals who were high-risk for both prognostic factors with worse prognosis compared to low-risk CLL patients (i.e., 5-year progression rate of 32% versus 21%, respectively). Among HCMBL, the frequency of somatically mutated genes at diagnosis is lower than that of CLL. Accounting for both the number of mutated genes and CLL-IPI can identify HCMBL individuals with more aggressive clinical course.

Conflict of interest: COI declared - see note

COI notes: SAP: Research funding has been provided to the institution from Janssen, AstraZeneca, Merck, and Genentech for clinical studies in which Sameer A. Parikh is a principal investigator. Honoraria has been provided to the institution from Pharmacyclics, Merck, AstraZeneca, Janssen, Genentech, Amgen, MingSight Pharmaceuticals, TG Therapeutics, Novalgen Limited, Kite Pharma, and AbbVie for Sameer A. Parikh's participation in consulting activities/advisory board meetings. NEK: Advisory Board for AbbVie, Astra Zeneca, Beigene, Behring, Boehringer Ingelheim Pharmaceuticals, Inc., Dava Oncology, Janssen, Juno Therapeutics, Pharmacyclics. DSMC (Data Safety Monitoring Committee) for Agios Pharm, AstraZeneca, BMS -Celgene, Dren Bio Janssen. Research funding from: AbbVie, Acerta Pharma, Bristol Meyer Squib, Celgene, Genentech, Pharmacyclics, Sunesis, Vincerx.

Preprint server: No;

Author contributions and disclosures: Concept and Study Design was performed by GK, SLS, EB, SAP. Acquisition of Data was performed by SLS, EB, SAP. Analysis of Data was performed by SLS, CA, and KGR. Interpretation of Data was performed by all authors. Manuscript was drafted by GK and SLS. Manuscript was reviewed by all authors.

Non-author contributions and disclosures: No;

Agreement to Share Publication-Related Data and Data Sharing Statement: Deidentified individual participant data are available upon request to corresponding author.

Clinical trial registration information (if any):

Tumor mutational load is prognostic for progression to therapy among high-count monoclonal B-cell lymphocytosis (HCMBL)

Geffen Kleinstern[1,2], Nicholas J. Boddicker[2], Daniel R. O'Brien[2], Cristine Allmer[3], Kari G. Rabe[3], Aaron D. Norman[4], Rosalie Griffin[2], Huihuang Yan[2], Tao Ma[2], Timothy G. Call[5], Laura Bruins[6], Sochilt Brown[6], Cecilia Bonolo de Campos[6], Curtis A. Hanson[7], Jose F. Leis[6], Wei Ding[5], Celine M. Vachon[4], Neil E. Kay[5], Christopher C. Oakes[8,9], Alexander S. Parker[10], Danielle M. Brander[12], J. Brice Weinberg[11,12,13], Richard R. Furman[14], Tait D. Shanafelt[15], James R. Cerhan[4], Sameer A. Parikh[5*], Esteban Braggio[6*], Susan L. Slager[2,5*]

[1]School of Public Health, University of Haifa, Haifa, Israel

[2]Division of Computational Biology, Mayo Clinic, Rochester, MN, USA

[3]Division of Clinical Trials and Biostatistics, Mayo Clinic, Rochester, MN, USA

[4]Division of Epidemiology, Mayo Clinic, Rochester, MN, USA

[5]Division of Hematology, Mayo Clinic, Rochester, MN, USA

[6]Department of Hematology and Oncology, Mayo Clinic, Phoenix, AZ, USA

[7]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA

[8]Division of Hematology, Department of Internal Medicine, The Ohio State University, Columbus OH, USA

[9]The Comprehensive Cancer Center, The James Cancer Hospital and Solove Research Institute, The Ohio State University, Columbus, OH, USA

[10]College of Medicine, University of Florida, Jacksonville, FL, USA

[11]Department of Immunology, Duke University Medical Center, Durham, NC, USA

[12]Department of Medicine, Duke University, Duke Cancer Institute, Durham, NC, USA

[13]Durham Veterans Affairs Medical Center, Durham, NC, USA.

[14]Weill Cornell Medical College/New York Presbyterian Hospital, New York, NY, USA

[15]Department of Medicine, Division of Hematology, Stanford University, Stanford, CA, USA

*These authors jointly supervised the work

Running title: TML is prognostic for high-count MBL

Abstract Count: 250 words

Text Count: 3551 words

Figure/Table count: 7

Reference count: 33

Data Sharing Statement: Deidentified individual participant data are available upon request to corresponding author at slager@mayo.edu.

Correspondence should be addressed to:

Susan Slager, Ph.D.
Professor
Divisions of Hematology and Computational Biology
Mayo Clinic, 200 First Street SW
Rochester, MN 55905
Slager.susan@mayo.edu

Key Points (140 char including space)

- Tumor mutational load is a strong prognostic factor for progression to therapy among individuals with HCMBL, independent of CLL-IPI
- Accounting for both CLL-IPI and TML, we identified HCMBL individuals who have a worse prognosis than low-risk patients with CLL.

Abstract

HCMBL is a precursor condition to chronic lymphocytic leukemia (CLL). We have shown that among individuals with HCMBL the CLL-International Prognostic Index (CLL-IPI) is prognostic for time-to-first therapy (TTFT). Little is known about the prognostic impact of somatically mutated genes among individuals with HCMBL.

We sequenced DNA from 371 HCMBL individuals using a targeted sequencing panel of 59 recurrently mutated genes in CLL to identify high-impact mutations. We compared the sequencing results to that of our treatment-naïve CLL cohort(N=855) and employed Cox regression to estimate hazard ratios and 95% confidence intervals (CI) for associations with TTFT.

Compared to CLL, the frequencies of any mutated genes were lower in HCMBL (70% versus 52%). At 10-years, 37% of HCMBL individuals with any mutated gene had progressed requiring treatment compared to 10% among HCMBL individuals with no mutations; this led to 5.4-fold shorter TTFT (95%CI:2.6-11.0) among HCMBL with any mutated gene versus none, independent of CLL-IPI. When considering individuals with low-risk of progression according to CLL-IPI, HCMBL individuals with any mutations had 4.3-fold shorter TTFT (95%CI:1.6-11.8) versus those with none. Finally, when considering both CLL-IPI and any mutated gene status, we observed HCMBL individuals who were high-risk for both prognostic factors with worse prognosis compared to low-risk CLL patients (i.e., 5-year progression rate of 32% versus 21%, respectively).

Among HCMBL, the frequency of somatically mutated genes at diagnosis is lower than that of CLL. Accounting for both the number of mutated genes and CLL-IPI can identify HCMBL individuals with more aggressive clinical course.

Introduction

Monoclonal B-cell lymphocytosis (MBL) is a precursor state to chronic lymphocytic leukemia (CLL)[1-3] and is characterized by a circulating population of clonal B-cells (<5x10^9/L) in the peripheral blood and in the absence of lymphadenopathy, cytopenias, or organomegaly[4,5]. MBL can be sub-classified based on immunophenotype (the most common being the immunophenotype that is similar to that of CLL) and the size of the MBL clone with high-count MBL (HCMBL) defined as the absolute clonal B-cell count between 0.5 and 4.9×10^9/L[6,7].

On average, progression to CLL requiring therapy among individuals with HCMBL occurs at a rate of ~1-2% per year[8-10]. A number of investigators have evaluated individual biomarkers to predict MBL progression to CLL requiring therapy and overall survival (OS), including mutation status of the immunoglobulin heavy chain variable region (IGHV), high-risk cytogenetic aberrations, high serum β2 microglobulin, CD49d expression, and absolute B-cell count[8-11]. More recently, investigators have integrated various markers into a single prognostic score, called the CLL-International Prognostic Index (CLL-IPI). The CLL-IPI is comprised of 5 individual clinical and leukemic prognostic factors and was found to be highly prognostic for OS and time to first therapy (TTFT) among individuals with CLL[12-14], even early-stage CLL[15]. The CLL-IPI has recently been found to stratify TTFT among individuals with HCMBL[11].

Large sequencing studies of treatment-naïve patients with CLL have identified recurrently mutated genes[16-18], mostly affecting cell cycle, DNA damage response, NOTCH and NF-kB signaling pathways. Investigative teams have also evaluated patients with CLL at later stages of disease, including symptomatic patients with CLL and patients with relapsed CLL, and found, in general, higher gene mutation frequencies at these later disease stages supporting the evidence that these genes impact disease progression[19-22]. Related, researchers have also reported on the clinical impact of individual mutations, individual driver genes, and total number of mutated genes with TTFT and OS among

patients with CLL [16-25]. We have previously shown that the total number of mutated genes (or the tumor mutational load [TML]) is a robust and an independent prognostic marker for progression to therapy among treatment naïve patients with CLL, beyond CLL-IPI[25].

The effect of recurrently mutated CLL driver genes among individuals with HCMBL has been understudied[16,23,25-28]. In the largest study to date of 112 individuals with HCMBL, we reported generally lower gene mutation frequency across the 59 CLL driver genes among individuals with HCMBL compared to that of CLL[25]. We also reported that individuals with 2 or more mutated CLL driver genes had a 4.1-fold increased risk of progression requiring CLL therapy, adjusting for CLL-IPI. These data indicated that the number of genes with high-impact or hotspot mutations may be a prognostic biomarker to identify HCMBL individuals who are at high-risk for progression. Herein, we significantly expanded our cohort with more than triple the number of individuals with HCMBL sequenced (N=371, including the previous 112 HCMBL individuals) and longer follow-up to further evaluate the prognostic impact of these CLL driver genes or their aggregate on disease progression relative to a cohort of 855 patients with CLL, also expanded from our earlier TML study of 445 patients with CLL.

Methods

Study Participants

From the Mayo Clinic CLL Resource[11,25,29], individuals with HCMBL or CLL (who were clinically seen in the Division of Hematology, Mayo Clinic, with available treatment naïve peripheral blood mononuclear cells (PBMC) were selected. Additional patients with CLL with treatment naïve PBMC were also obtained from the CLL clinics at Duke University (N=57), Weill Cornell Medical College (N=18) and from the CLL Research Consortium (N=44)[30]. PBMCs for sequencing studies were collected within a median 0.03 years (range 0-26) from diagnosis. Clinical characteristics were obtained at time of diagnosis or at time of sample collection and included age, sex, Rai stage, serum β2 microglobulin levels, IGHV mutation status, and genetic abnormalities detectable by fluorescence in situ hybridization

(FISH), when available. We also included 54 individuals with HCMBL from the Mayo Clinic MBL Biobank Cohort[31]; these individuals were identified to have HCMBL through MBL screening using 8-color flow cytometry and had the immunophenotype of CLL-like (N=346), atypical MBL (N=6), and non CLL-like MBL (N=19). Abstracted clinical characteristics included age and sex, while del(17p) genetic abnormality was obtained through sequencing.

All individuals provided written informed consent for this research whose protocol was approved by the respective Institutional Review Boards.

Deidentified individual participant data are available upon request to corresponding author.

DNA sequencing

DNA was extracted from PBMCs that had a tumor purity >80% or otherwise from PBMCs enriched for CD5+/CD19+ clonal B-cells. We sequenced the entire coding regions of 59 somatically recurring mutated CLL genes as previously described[25] (**Supplementary Table 1**). In brief, DNA samples were sequenced using Illumina HiSeq 4000 sequencer. The median coverage depth per individual across the 59 recurrently mutated genes was 1,741X with >80% of the individuals having a median coverage depth >1000X per nucleotide, allowing the detection of mutations with variant allelic fraction (VAF) as low as 1%. Somatic mutations were called using MuTect2 in tumor-only mode. After filtering, high-impact mutations (frameshift, nonsense, and splicing variants) and missense mutations in previously identified CLL hotspots (**Supplementary Table 2**) were used for statistical analyses. Raw variants were annotated using GATK Variant Annotator for variant quality, and Biological Reference Repository (BioR) was used for variant annotation. To remove germline polymorphisms, common variants were eliminated based on the minor allele frequencies >0.01% available in the following germline variant databases: 1000 Genomes Project, ExAC and ESP6500 from NHLBI Exome Sequencing Project,

unless present in known CLL/MBL mutation hotspots or in COSMIC. Additionally, we filtered out all somatic variants with <10 supportive reads or <1% VAF.

Statistical analyses

CLL-IPI was calculated based on clinical stage (Rai 0 vs. Rai I-IV), IGHV mutation status (mutated vs. unmutated), *TP53* status (wild type vs. del17p, *TP53* mutations, or both), β2 microglobulin level (>3.5 mg/L), and age (>65 years) as previously described[12]. We classified all individuals with HCMBL as Rai stage 0 when calculating the CLL-IPI score. TML was calculated by counting the number of genes with high-impact and hotspot mutations, but excluding *TP53,* which is used in calculating the CLL-IPI. We then categorized the TML either as binary (none or any mutated gene) or as a categorical variable with 0 mutated genes serving as the reference category. We also considered TML as a continuous variable. TTFT was defined as the time from date of sample to the earliest date of first treatment or date of last follow-up. The event in the TTFT analyses was those individuals who received CLL treatment; otherwise, individuals were censored. Any patients with CLL who were enrolled on early interventional trials were censored on the date of treatment. Deaths in untreated patients were considered a competing risk. OS was defined as time from date of sample to the earliest date of death or last follow-up. The event in OS was those who died, regardless of cause; otherwise, individuals were censored. We used Cox regression models to estimate hazard ratios (HRs) and 95% confidence intervals (CIs) for the effect of individual genes or TML with TTFT and OS. In addition, we evaluated the TML with TTFT and OS stratified by CLL-IPI status (within the HCMBL cohort and in the combined HCMBL and CLL cohort). For the full cohort analyses, we added diagnosis status (CLL or HCMBL) as an additional covariate. OS curves were displayed using the Kaplan-Meier method and the TTFT curves using the cumulative incidence method. To evaluate model discriminative ability, we computed a *c*-statistic and 95% CI for the adjusted Cox regression models. Significant findings were those that had a P<0.05. Analyses were conducted in SAS 9.4 (SAS Institute, Cary, NC).

Results

Baseline characteristics

A total of 371 individuals with HCMBL were sequenced and compared to a cohort of 855 patients with CLL. Among the HCMBL individuals, the median age was 67 years (range 42-95), and 63% were male. Among patients with CLL, the median age was 61 years (range 27-94) and 70% were male (**Table 1, Supplemental Table 3)**. The majority (61%, N=179) of the HCMBL individuals were considered low-risk according to the CLL-IPI; whereas among patients with CLL, 268 (36%) were considered low-risk (**Table 1, Supplemental Table 3)**. A total of 181 individuals (N=76 HCMBL and 105 CLL) were missing CLL-IPI scores and not included in outcome analyses.

Characteristics of somatic variants detected in HCMBL compared to CLL

The type and distribution of high-impact or hotspot mutations among the HCMBL cohort is shown in **Figure 1A**. The most frequently mutated genes were *TP53* (8%), *NOTCH1* (8%), *SF3B1* (6%), and *MYD88* (6%); 49% of the individuals with HCMBL had at least one mutated gene (**Table 2**) and 17 (29%) genes had no evidence of high-impact or hotspot mutations (**Supplementary Table 1**). The majority of the genes were frequently more mutated among HCMBL individuals with IGHV unmutated status compared to HCMBL individuals with IGHV mutated status; exceptions included *TP53, MYD88, CHD2, ITPKB,* and *CHEK2* (**Supplementary Table 1**). When comparing HCMBL to CLL, we observed similar distributions of the variant allele fraction (**Figure 1B**) and the type of mutation (i.e., splice site, frame shift, etc., **Figure 1C**). However, the frequency of individual genes with high-impact or hotspot mutations were higher in CLL compared to HCMBL (**Supplementary Table 1, Figure 2A**). The frequency of any mutated gene (including *TP53*) was 52% compared to 70% among CLL. When looking at the TML, patients with CLL had a higher total number of mutated genes (median=1) compared to individuals with HCMBL (median=0, P<0.001, **Figure 2B**). This higher TML distribution held, even among the 460 CLL Rai 0 patients compared to HCMBL (**Figure 2C,** P<0.001).

Time to first therapy according to the TML

Among individuals with HCMBL, the median follow-up from sample collection was 6 years (range 0-20), and 57 individuals progressed to needing therapy. Overall, the 5- and 10-year cumulative progression to CLL requiring therapy was 9% and 24% respectively (**Figure 3A**). However, this cumulative risk varied according to the TML (**Figure 3B**, P<0.0001). The estimated 5- and 10-year cumulative progression to therapy among those HCMBL individuals with no mutated genes was of 2% and 10%, respectively. In contrast, the estimated 5- and 10-year cumulative progression to therapy among those with 2+ mutated gene was 18% and 37%, respectively. This association between TML and TTFT held in multivariate Cox regression analyses adjusting for CLL-IPI and sex; individuals with any mutated genes had a 5.4-fold increased risk of progression to therapy (95%CI: 2.6-11.0, P<0.0001, **Table 2**). We observed similar effect size for progression to therapy between those with just one mutated gene (HR=5.3, 95%CI 2.5-11.2, P<0.0001) and those with 2+ mutated genes (HR=5.5, 95%CI 2.4-12.4, P<0.0001, **Table 2**). When considering each gene individually, eight genes (*NOTCH1, SF3B1, BRAF, FBXW7, XPO1, EGR2, MED12,* and *ARID1A*) were found to be statistically associated with TTFT univariately and after adjusting for CLL-IPI and sex (**Supplementary Table 4**). Although mutated *TP53* had elevated effect sizes in both univariate and multivariate models, these associations were not statistically significant. However, when we stratified by IGHV mutation status, mutated *TP53* had a 6.35-fold (95%CI 1.71-23.6) shorter time to therapy compared to no *TP53* mutations among HCMBL individuals who were unmutated for IGHV (**Supplementary Table 4**).

Next, we evaluated the association between TML and TTFT stratified by CLL-IPI to see if further separation of risk among individuals with HCMBL could be achieved. We found that TML was prognostic for TTFT within each of the CLL-IPI levels (all continuous TML P<0.05, **Table 2**). Among those HCMBL individuals who were low-risk according to CLL-IPI, the 5-year cumulative incidence to requiring therapy for those with any mutated gene was 9% compared to 3% with no mutated genes, and the 10-year cumulative incidence was 21% compared to 7%, respectively. In Cox regression

analyses, this corresponds to 4.3-fold increased risk of therapy (95%CI 1.6-11.8, P=0.005, **Table 2**). Among those HCMBL individuals who were intermediate CLL-IPI risk, those with any mutation had 3.0-fold increased risk of therapy (95%CI 1.0-9.0, P=0.045, **Table 2**). Among the high- or very high-risk CLL-IPI, nine HCMBL individuals were treated, and all nine of these individuals had a at least one gene with a high-impact or hotspot mutation. None of the HCMBL individuals with no mutated genes were treated, precluding us from evaluating the binary effect of TML (**Table 2**). However, the continuous TML was associated with TTFT (HR=1.5, 95%CI 1.1-2.1, P=0.019).

More aggressive disease among HCMBL compared to low-risk CLL

Next, we evaluated whether there were situations in which individuals with HCMBL had a rate of progression to therapy similar or worse compared to that of low-risk patients with CLL when considering both TML and CLL-IPI status. To answer this, we stratified the HCMBL individuals by CLL-IPI (CLL-IPI low-risk vs CLL-IPI intermediate/high/very-high risk) and TML status (TML=0 vs TML>0) and compared these HCMBL individuals to low-risk patients with CLL (defined as those with CLL-IPI low-risk and TML=0). We found that HCMBL individuals with TML>0 and CLL-IPI intermediate/high/very high-risk had a 5-year rate of progression to therapy of 32%, higher than the 22% among low-risk patients with CLL (i.e., CLL-IPI low-risk, and TML=0, HR=1.93, 95%CI 1.2-3.1, P=0.007, **Figure 4**).

Combined Analysis of HCMBL and CLL cohorts to evaluate TML and CLL-IPI with TTFT and OS

Finally, we combined the CLL and HCMBL cohorts to further evaluate the interaction of TML and CLL-IPI with TTFT and OS, while adjusting for MBL/CLL status. In the combined cohort the median follow-up from sample collection was 7.3 years (range 0-22.5), and 472 individuals progressed to needing therapy. At 5-years, the rate of progression requiring CLL therapy was 40%; but according to the TML, a clear and significant separation of risk of TTFT was observed (P<0.0001, **Figure 5A, Supplementary Table 5**). TML was also strongly associated with TTFT within each of the CLL-IPI levels (all continuous TML P<0.0001, **Figures 5B-D, Supplementary Table 5**). In particular, individuals with low-risk of

progression according to CLL-IPI and with 2+ mutated genes had a 38% cumulative progression to therapy compared to 15% for those with no mutated genes at 5 years. Even among those with a high- or very-high CLL-IPI, a clear separation based on the TML was observed within one year, e.g., estimated 1-year cumulative progression to therapy was 55% among those with 2+ mutated genes compared to 18% among those with no mutated genes (**Figure 4D**), resulting in a 3.4-fold shorter TTFT (95%CI 2.2-5.4, **Supplementary Table 5**).

In the combined cohort, 351 individuals died. The 5-year OS rate was 89% (**Supplementary Figure 1A**). When evaluating OS according to TML, we observed a significant difference in survival (P<0.0001, **Supplementary Figure 1B**); the 5-year OS rate was 92% for those with no mutated genes and 82% for those with 2+ mutated genes. After adjusting for CLL-IPI, sex, and diagnosis (MBL/CLL), the association of TML with OS held in Cox regression analyses; individuals with 2+ mutated genes had 1.4-fold higher mortality compared to those without mutated genes (95%CI 1.1-1.9, P=0.017, **Supplementary Table 6**). TML did not provide additional information for OS within low- and intermediate- CLL-IPI levels. However, in the high or very-high risk CLL-IPI, individuals with any mutated gene had 2.0-fold higher mortality compared to those without a mutated gene (95% CI 1.3-3.2, **Supplementary Table 6**).

Discussion

In the largest study to date of 371 HCMBL individuals, we analyzed the tumor mutations of 59 genes recurrently mutated in CLL and found that a higher total number of mutated genes (or the TML) was associated with greater risk of progression to therapy, independent of CLL-IPI, extending our prior study of 112 HCMBL individuals[25]. Moreover, we found that accounting for both CLL-IPI and TML, we identified HCMBL individuals with worse prognosis compared to low-risk patients with CLL.

HCMBL is a fairly common condition with a prevalence of ~0.2-1% in the general population[31,32] or 2% among individuals with a strong family history of CLL[2]. Typically, individuals with HCMBL are clinically identified due to incidental findings with lymphocytosis or through assessments from another clinical condition. Once identified, individuals with HCMBL are often evaluated annually with a complete blood count (CBC) and physical examination. The understanding is that these individuals with HCMBL have a small risk of progression to CLL requiring therapy of about 10-20% in 10 years (or ~1-2% per year)[8-10]. We recently showed that we could improve this estimate of progression to therapy using the CLL-IPI[11]. Specifically, HCMBL individuals with low-risk according to CLL-IPI had a 15% cumulative risk of progression requiring therapy at 10-years compared to 68% risk among those with intermediate CLL-IPI and 82% progression risk among those with high-risk CLL-IPI. Herein, we demonstrated we can further improve progression estimates, particularly among those individuals in the CLL-IPI low-risk group, by also considering somatic mutations in CLL driver genes.

In our prior work of 112 HCMBL individuals[24], we found that those with 2+ mutated genes had a 4.1-fold increased risk of progression requiring therapy independent of CLL-IPI. Herein, with the larger sample size and longer follow-up, we reported a significant 5.5-fold increased risk, adjusting for CLL-IPI. Importantly, the TML was able to further stratify risk of progression to therapy within the risk levels of CLL-IPI. Among individuals with low-risk according to CLL-IPI, HCMBL individuals with no mutated genes had a 7% cumulative incidence of progression in 10 years compared to 21% among those with detectable mutations. This refinement of stratification compares favorably to the previously reported 15% based only on CLL-IPI (i.e., ignoring TML). Knowledge of a lower risk (on average) of progression based on both CLL-IPI and TML may allow clinicians to reassure a large segment of individuals with HCMBL about their prognosis and potentially may reduce the need for annual follow-up visits. In addition, for future clinical trials wanting to study early-intervention, our results herein support that a combination of TML and CLL-IPI would be invaluable for risk stratification and the possibility of including individuals with HCMBL.

By definition, individuals with HCMBL have lower clonal B-cell counts than patients with CLL and are at a precursor stage to CLL. Our results, based on somatic sequencing, also supported that individuals with HCMBL are, in general, at an earlier pathobiological stage than that of CLL. We saw this by the lower number of high-impact mutations, lower individual gene frequencies, and lower TML distribution in HCMBL compared with CLL. However, based on both CLL-IPI and TML, we observed HCMBL individuals with a more aggressive clinical course as evidenced by shorter time to therapy compared to patients with CLL with low-risk disease based on CLL-IPI and TML. These results suggest that using TML and CLL-IPI, in addition to clonal B-cell count, may better detect individuals at high-risk.

Finally, in the full cohort of individuals with MBL or CLL, the TML was a significant prognostic factor for OS, independent of CLL-IPI, sex, and MBL/CLL diagnosis. Interestingly, we found that individuals with any mutated gene at diagnosis will have a higher risk of dying. Even among the individuals in the high or very-high risk group according to CLL-IPI, any mutated gene was significantly prognostic for OS.

Strengths of this study included using a large cohort of newly diagnosed individuals with HCMBL compared to a cohort of patients with CLL. All of our study participants were treatment naïve at the time of blood sample collection and had extensive annotated clinical data allowing us to account for the effect of CLL-IPI, a well-established CLL prognostic factor that has also been shown to have clinical utility for stratifying TTFT among individuals with HCMBL[11]. Our sequencing depth was over 1000x in most individuals, allowing the detection of small clones on a comprehensive set of recurrently mutated CLL genes. This study also had several limitations. Among the HCMBL individuals, we were not able to fully evaluate the association of TML with TTFT within the high- or very high-risk CLL-IPI category due to limited number of events. Second, our targeted sequencing panel contained 59 genes, while a recent publication of whole exome sequencing in over 1000 patients with CLL identified 37 additional genes not included here[18]. These new putative CLL driver genes had frequencies <1.5%, and thus we

would not expect them to have a significant impact on the conclusions of our study. Finally, our definition of TML assumed that all somatic mutations have equal weight and thus the same clinical impact on outcomes, regardless of type of high-impact mutation, the direction of effect the gene has on TTFT, or which gene the mutation was located. Alternative approaches could provide different weights for each mutation type or mutated gene. By adding weights, the TML score would be expected to better refine prognostication, similar to what has been observed with inherited common variants and the polygenic risk scores[29,33].

In summary, our study highlighted that the TML can identify individuals with HCMBL who are at high risk of progression to CLL requiring therapy beyond CLL-IPI. These findings are similar to and support our previous findings of the effect of TML among patients with CLL. More importantly, we observed individuals with HCMBL considered to have a low-risk of progression based on the CLL-IPI could be further stratified into lower or higher risk of progression based on the TML. Although HCMBL is considered to be a precursor state to CLL, our study identified HCMBL individuals who have a more aggressive clinical course than low-risk patients with CLL based on both TML and CLL-IPI. Our findings may help future clinical practice guidelines.

Acknowledgments

The authors thank the study participants for their generosity of time in participating in our study. This work was supported by the National Institutes of Health (NIH) grants R01AG58266, R21CA256648, R01CA235026, R01CA258465, and P50CA97274; and by the Zuckerman STEM Leadership Program.

Authorship Contributions

Concept and Study Design was performed by GK, SLS, EB, SAP. Acquisition of Data was performed by SLS, EB, SAP. Analysis of Data was performed by SLS, CA, and KGR. Interpretation of Data was

performed by all authors. Manuscript was drafted by GK and SLS. Manuscript was reviewed by all authors.

Disclosures of Conflicts of Interest

SAP: Research funding has been provided to the institution from Janssen, AstraZeneca, Merck, and Genentech for clinical studies in which Sameer A. Parikh is a principal investigator. Honoraria has been provided to the institution from Pharmacyclics, Merck, AstraZeneca, Janssen, Genentech, Amgen, MingSight Pharmaceuticals, TG Therapeutics, Novalgen Limited, Kite Pharma, and AbbVie for Sameer A. Parikh's participation in consulting activities/advisory board meetings. NEK: Advisory Board for AbbVie, Astra Zeneca, Beigene, Behring, Boehringer Ingelheim Pharmaceuticals, Inc., Dava Oncology, Janssen, Juno Therapeutics, Pharmacyclics. DSMC (Data Safety Monitoring Committee) for Agios Pharm, AstraZeneca, BMS –Celgene, Dren Bio Janssen. Research funding from: AbbVie, Acerta Pharma, Bristol Meyer Squib, Celgene, Genentech, Pharmacyclics, Sunesis, Vincerx.

References

1. Landgren O, Albitar M, Ma W, et al. B-cell clones as early markers for chronic lymphocytic leukemia. *N Engl J Med*. 2009;360(7):659-667.
2. Slager SL, Lanasa MC, Marti GE, et al. Natural history of monoclonal B-cell lymphocytosis among relatives in CLL families. *Blood*. 2021;137(15):2046-2056.
3. Kolijn PM, Hosnijeh FS, Spath F, et al. High-risk subtypes of chronic lymphocytic leukemia are detectable as early as 16 years prior to diagnosis. *Blood*. 2022;139(10):1557-1563.
4. Marti GE, Rawstron AC, Ghia P, et al. Diagnostic criteria for monoclonal B-cell lymphocytosis. *Br J Haematol*. 2005;130(3):325-332.
5. Hallek M, Cheson BD, Catovsky D, et al. iwCLL guidelines for diagnosis, indications for treatment, response assessment, and supportive management of CLL. *Blood*. 2018;131(25):2745-2760.
6. Shanafelt TD, Ghia P, Lanasa MC, Landgren O, Rawstron AC. Monoclonal B-cell lymphocytosis (MBL): biology, natural history and clinical management. *Leukemia*. 2010;24(3):512-520.
7. Rawstron AC, Shanafelt T, Lanasa MC, et al. Different biology and clinical outcome according to the absolute numbers of clonal B-cells in monoclonal B-cell lymphocytosis (MBL). *Cytometry B Clin Cytom*. 2010;78 Suppl 1:S19-23.
8. Rawstron AC, Bennett FL, O'Connor SJ, et al. Monoclonal B-cell lymphocytosis and chronic lymphocytic leukemia. *N Engl J Med*. 2008;359(6):575-583.
9. Shanafelt TD, Kay NE, Rabe KG, et al. Brief report: natural history of individuals with clinically recognized monoclonal B-cell lymphocytosis compared with patients with Rai 0 chronic lymphocytic leukemia. *J Clin Oncol*. 2009;27(24):3959-3963.

10. Rossi D, Sozzi E, Puma A, et al. The prognosis of clinical monoclonal B cell lymphocytosis differs from prognosis of Rai 0 chronic lymphocytic leukaemia and is recapitulated by biological risk factors. *Br J Haematol*. 2009;146(1):64-75.
11. Parikh SA, Rabe KG, Kay NE, et al. The CLL International Prognostic Index predicts outcomes in monoclonal B-cell lymphocytosis and Rai 0 CLL. *Blood*. 2021;138(2):149-159.
12. International CLLIPIwg. An international prognostic index for patients with chronic lymphocytic leukaemia (CLL-IPI): a meta-analysis of individual patient data. *Lancet Oncol*. 2016;17(6):779-790.
13. Gentile M, Shanafelt TD, Rossi D, et al. Validation of the CLL-IPI and comparison with the MDACC prognostic index in newly diagnosed patients. *Blood*. 2016;128(16):2093-2095.
14. da Cunha-Bang C, Christiansen I, Niemann CU. The CLL-IPI applied in a population-based cohort. *Blood*. 2016;128(17):2181-+.
15. Molica S, Shanafelt TD, Giannarelli D, et al. The chronic lymphocytic leukemia international prognostic index predicts time to first treatment in early CLL: Independent validation in a prospective cohort of early stage patients. *Am J Hematol*. 2016;91(11):1090-1095.
16. Puente XS, Bea S, Valdes-Mas R, et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2015;526(7574):519-524.
17. Landau DA, Tausch E, Taylor-Weiner AN, et al. Mutations driving CLL and their evolution in progression and relapse. *Nature*. 2015;526(7574):525-530.
18. Knisbacher BA, Lin ZA, Hahn CK, et al. Molecular map of chronic lymphocytic leukemia and its impact on outcome. *Nature Genetics*. 2022.
19. Rossi D, Fangazio M, Rasi S, et al. Disruption of BIRC3 associates with fludarabine chemorefractoriness in TP53 wild-type chronic lymphocytic leukemia. *Blood*. 2012;119(12):2854-2862.
20. Schnaiter A, Paschka P, Rossi M, et al. NOTCH1, SF3B1, and TP53 mutations in fludarabine-refractory CLL patients treated with alemtuzumab: results from the CLL2H trial of the GCLLSG. *Blood*. 2013;122(7):1266-1270.
21. Robbe P, Ridout KE, Vavoulis DV, et al. Whole-genome sequencing of chronic lymphocytic leukemia identifies subgroups with distinct biological and clinical features. *Nat Genet*. 2022;54(11):1675-1689.
22. Ojha J, Secreto CR, Rabe KG, et al. Identification of recurrent truncated DDX3X mutations in chronic lymphocytic leukaemia. *Br J Haematol*. 2015;169(3):445-448.
23. Winkelmann N, Rose-Zerilli M, Forster J, et al. Low frequency mutations independently predict poor treatment-free survival in early stage chronic lymphocytic leukemia and monoclonal B-cell lymphocytosis. *Haematologica*. 2015;100(6):E237-E239.
24. Stilgenbauer S, Schnaiter A, Paschka P, et al. Gene mutations and treatment outcome in chronic lymphocytic leukemia: results from the CLL8 trial. *Blood*. 2014;123(21):3247-3254.
25. Kleinstern G, O'Brien DR, Li X, et al. Tumor mutational load predicts time to first treatment in chronic lymphocytic leukemia (CLL) and monoclonal B-cell lymphocytosis beyond the CLL international prognostic index. *Am J Hematol*. 2020;95:906-917.
26. Barrio S, Shanafelt TD, Ojha J, et al. Genomic characterization of high-count MBL cases indicates that early detection of driver mutations and subclonal expansion are predictors of adverse clinical outcome. *Leukemia*. 2017;31(1):170-176.
27. Agathangelidis A, Ljungstrom V, Scarfo L, et al. Highly similar genomic landscapes in monoclonal B-cell lymphocytosis and ultra-stable chronic lymphocytic leukemia with low frequency of driver mutations. *Haematologica*. 2018;103(5):865-873.
28. Rasi S, Monti S, Spina V, Foa R, Gaidano G, Rossi D. Analysis of NOTCH1 mutations in monoclonal B-cell lymphocytosis. *Haematologica-the Hematology Journal*. 2012;97(1):153-154.
29. Kleinstern G, Weinberg JB, Parikh SA, et al. Polygenic risk score and risk of monoclonal B-cell lymphocytosis in caucasians and risk of chronic lymphocytic leukemia (CLL) in African Americans. *Leukemia*. 2022;36(1):119-125.

30. Coombs CC, Rassenti LZ, Falchi L, et al. Single nucleotide polymorphisms and inherited risk of chronic lymphocytic leukemia among African Americans. *Blood*. 2012;120(8):1687-1690.
31. Slager SL, Parikh SA, Achenbach SJ, et al. Progression and survival of MBL: a screening study of 10 139 individuals. *Blood*. 2022;140(15):1702-1709.
32. Shim YK, Rachel JM, Ghia P, et al. Monoclonal B-cell lymphocytosis in healthy blood donors: an unexpectedly common finding. *Blood*. 2014;123(9):1319-1326.
33. Kleinstern G, Camp NJ, Goldin LR, et al. Association of polygenic risk score with the risk of chronic lymphocytic leukemia and monoclonal B-cell lymphocytosis. *Blood*. 2018;131(23):2541-2551.

Table 1: Demographic, clinical, and tumor characteristics

		Total Cohort		HCMBL		CLL	
		(N=1,226)		(N=371)		(N=855)	
Clinical characteristic		N	%	N	%	N	%
Gender	female	327	31.90%	114	36.80%	213	29.30%
	male	699	68.10%	196	63.20%	503	70.30%
Age	Median (range)	63	(27-95)	67	(42-95)	61	(27-94)
Race	Caucasian	1089	88.80%	364	98.15%	725	84.80%
	Other	137	11.20%	7	1.20%	130	15.20%
Rai stage risk group	Rai 0	831	68.80%	371	100%	460	55.00%
	Rai I-II	326	27.00%	-	-	326	38.90%
	Rai III-IV	51	4.20%	-	-	51	6.10%
	Missing	18				18	
FISH at diagnosis*	13q deletion	587	57.50%	162	54.60%	425	58.90%
	Trisomy 12	164	16.10%	49	16.40%	115	15.90%
	11q deletion	103	10.10%	15	5.00%	88	12.20%
	17p deletion	55	5.40%	8	2.60%	47	6.50%
	Normal	233	22.70%	85	28.10%	148	20.50%
	Other	15	1.50%	3	1.00%	12	1.70%
	Missing	200		68		132	
TP53	Mutated	124	10.10%	30	8.10%	124	10.10%
	Unmutated	921	89.80%	287	92.60%	634	88.50%
	Missing	0		0		0	
IGHV mutation status	Mutated	611	55.90%	221	75.40%	390	48.70%
	Unmutated	483	44.10%	72	24.60%	411	51.30%
	Missing	132		78		54	
Beta-2 microglobulin	≤3.5 mg/L	867	81.60%	278	89.70%	589	78.30%
	>3.5 mg/L	195	18.40%	32	10.30%	163	21.70%
	Missing	164		61		103	
	Median (range)	2.4 (0.2-121.5)		2.1 (1.0-21.5)		2.4 (0.2-16.2)	
CLL-IPI**	Low risk (0-1)	447	42.78%	179	60.70%	268	35.73%
	Intermediate risk (2-3)	313	29.95%	70	23.70%	243	32.40%
	High/Very High risk (4-10)	285	27.27%	46	15.60%	239	31.87%
	Missing	181		76		105	
B-cell count x 10^9/L	Median (range)	6.7 (0.2-394.5)		2.9 (0.5-4.9)		11.8 (0.2-394.5)	
	Missing	164		48		116	

*Individuals can have more than one; HCMBL= high-count monoclonal B-cell lymphocytosis. CLL=chronic lymphocytic leukemia.

Table 2: Association between Tumor Mutational Load and Time-to-First Treatment among individuals with HCMBL

		Total		Event			
CLL-IPI	Mutated genes	N	col %	N	HR	95% CI	P
Overall* (N=288)	None	147	51.0	9	1.00	reference	
	Any	141	49.0	48	5.37	(2.63,10.98)	<0.0001
	0	147	51.0	9	1.00	reference	
	1	90	31.3	31	5.30	(2.51,11.21)	<0.0001
	2+	51	17.7	17	5.50	(2.44,12.36)	<0.0001
	Cont.	288	100	56	1.59	(1.33,1.89)	<0.0001
	c-statistic	0.726				(0.65-0.80)	
Low Risk CLL-IPI** (N=179)	None	103	57.5	5	1.00	reference	
	Any	76	42.5	18	4.30	(1.57,11.79)	0.005
	0	103	57.5	5	1.00	Reference	
	1	56	31.3	14	4.47	(1.56,12.77)	0.005
	2+	20	11.2	4	3.88	(1.04,14.47)	0.043
	Cont.	179	100	23	1.75	(1.15,2.68)	0.009
	c-statistic	0.660				(0.56-0.76)	
Intermediate risk CLL-IPI** (N=70)	None	27	38.6	4	1.00	reference	
	Any	43	61.4	20	3.04	(1.02,9.01)	0.045
	0	27	38.6	4	1.00	Reference	
	1	23	32.9	11	3.51	(1.11,11.14)	0.033
	2+	20	28.6	9	2.57	(0.77,8.55)	0.123
	Cont.	70	100	24	1.48	(1.04,2.11)	0.028
	c-statistic	0.650				(0.54-0.76)	
High-very high risk CLL-IPI** (N=46)	None	24	52.2	0	1.00	reference	
	Any	22	57.8	9	NA	(NA)	NA
	0	24	52.2	0	1.00	Reference	
	1	11	23.9	5	NA	(NA)	NA
	2+	11	23.9	4	NA	(NA)	NA
	Cont.	46	100	9	1.48	(1.07,2.07)	0.019
	c-statistic	0.803				(0.71,0.90)	

*Adjusted for CLL-IPI and sex

**Adjusted for sex

Figure Legends

Figure 1: Characteristics of somatic mutations between HCMBL and CLL. Heat map representing the distribution of high-impact and hotspot mutations in 59 genes frequently mutated in CLL and HCMBL. Each column represents a single individual from our cohort. The external bar indicates individuals with HCMBL (grey) or CLL (red) diagnosis; the second bar stratifies the heat map by CLL-IPI [low risk (green), intermediate risk (yellow), high/very high risk (red)]; the third bar represents the TML distribution (i.e., the number of mutated genes per individual); then the next five bars represent clinical components: Rai stage (0-grey, I-green, II-yellow, III-red, IV-purple), sex, IGHV, B2M, and del(17p) by FISH. Finally, the internal bars represent the list of mutated genes in our study color-coded by the type of mutation: missense (yellow), in frame (grey), frameshift (red), nonsense (purple), splice (blue), more than one mutation (green) (A). Distribution of the variant allele fraction of high-impact and hotspot mutations between individuals with CLL and HCMBL (B). Distribution of mutation type between individuals with HCMBL and CLL (C). HCMBL= high-count monoclonal B-cell lymphocytosis; CLL= chronic lymphocytic leukemia.

Figure 2: Gene frequencies and TML distribution between HCMBL and CLL. Scatter plot of the frequency of mutated genes found in HCMBL and CLL. Red dotted line represents equal mutated gene frequency (A). Distribution of the TML sco re between individuals with CLL and HCMBL (B). Distribution of mutation type between individuals with HCMBL and patients with CLL with Rai stage=0. HCMBL= high-count monoclonal B-cell lymphocytosis; CLL= chronic lymphocytic leukemia; TML= tumor mutational load, which is the number of genes out of 59 with high-impact mutations (excluding *TP53*).

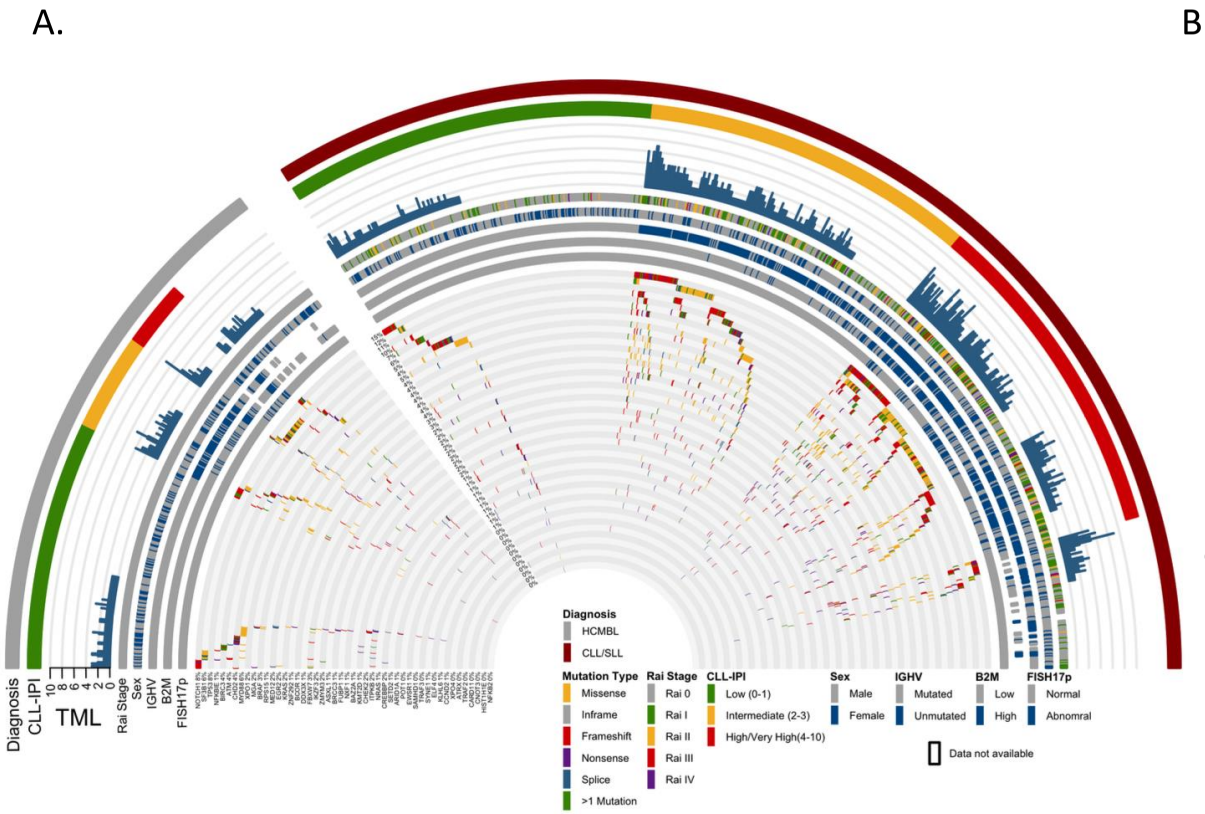
Figure 3: Associations of TML with Time To First Treatment among HCMBL individuals. Cumulative incidence plot of progression to first treatment among individuals with HCMBL (A). Cumulative incidence plot of progression to first treatment among individuals with HCMBL by TML (B). HCMBL= high-count monoclonal B-cell lymphocytosis; TML= tumor mutational load, which is the number of genes out of 59 with high-impact mutations (excluding *TP53*).

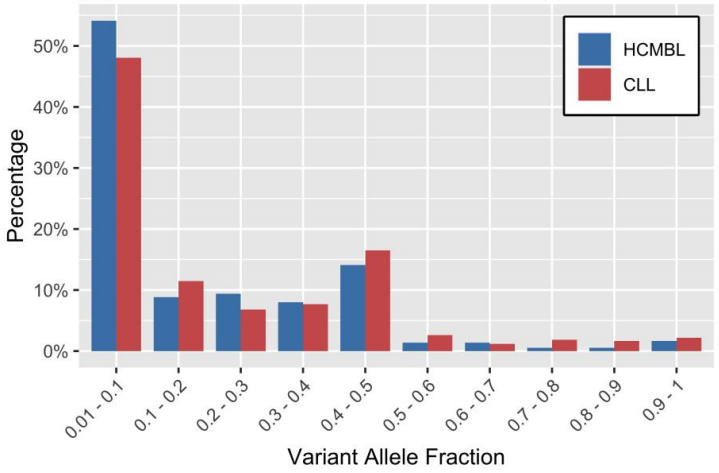
Figure 4: HCMBL and CLL stratified by both TML and CLL-IPI for association with Time to First Treatment. Cumulative incidence plot of progression to first treatment among individuals with HCMBL and CLL stratified by TML (Any versus no mutations) and CLL-IPI (low-risk vs more than low risk). HCMBL= high-count monoclonal B-cell lymphocytosis; TML= tumor mutational load, which is the number of genes out of 59 with high-impact mutations (excluding *TP53*).

Figure 5: Associations of TML with Time to First Treatment among Full Cohort and Stratified by CLL-IPI. Cumulative incidence plot of progression to first treatment among individuals with either HCMBL or CLL by TML (A). Cumulative incidence plot of progression to first treatment among individuals with HCMBL or CLL by TML among low-risk CLL-IPI (B). Cumulative incidence plot of progression to first treatment among individuals with HCMBL or CLL by TML among intermediate-risk CLL-IPI (C). Cumulative incidence plot of progression to first treatment among individuals with HCMBL or CLL by TML among high or very-high risk CLL-IPI (D). HCMBL= high-count monoclonal B-cell lymphocytosis; TML= tumor mutational load, which is the number of genes out of 59 with high-impact mutations (excluding *TP53*).

Figure 1

Figure 1: Characteristics of somatic mutations between HCMBL and CLL





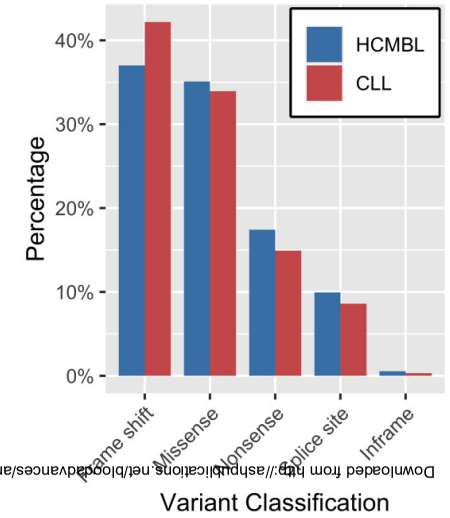


Figure 2

Figure 2: Gene frequencies and TML distribution between HCMBL and CLL

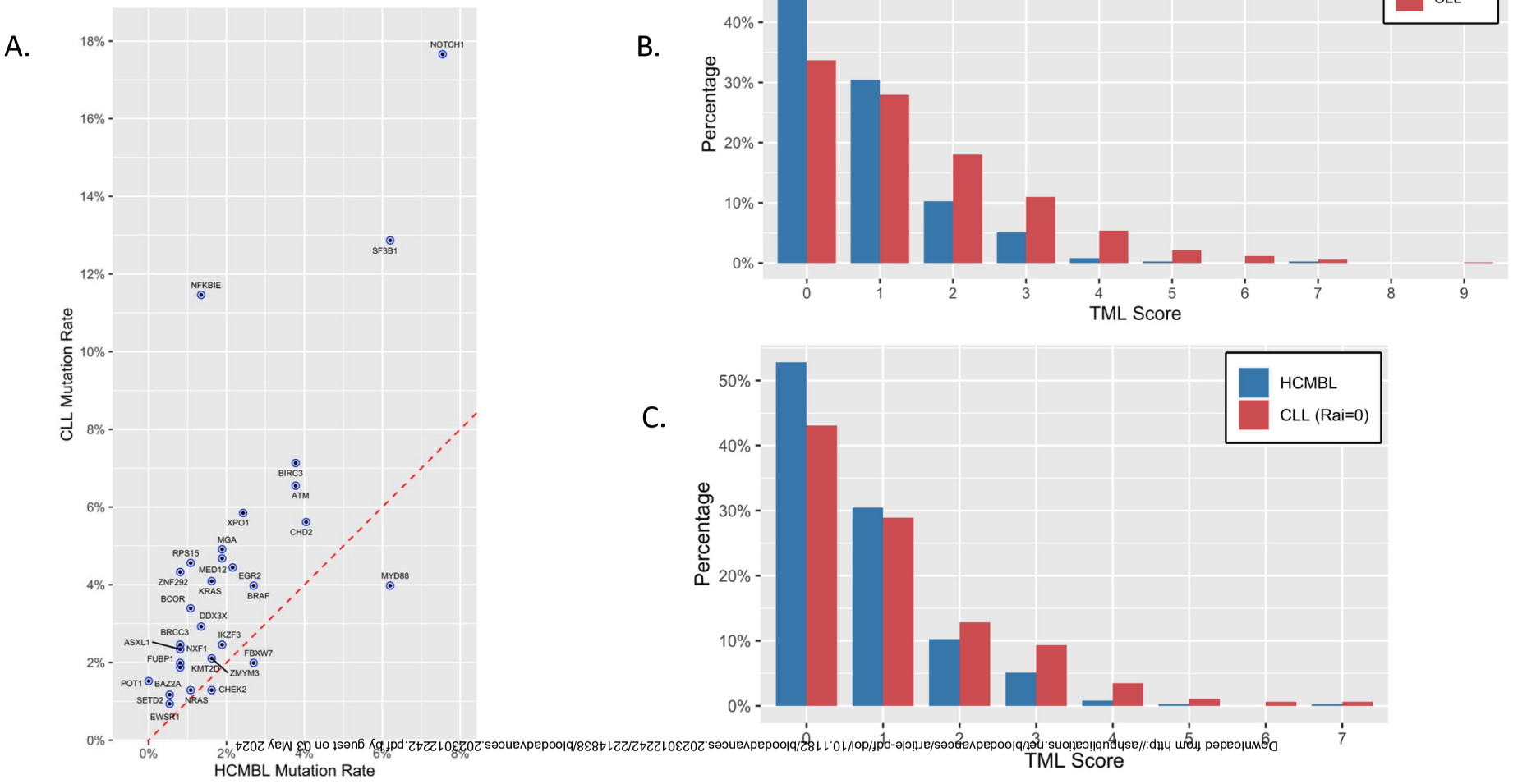
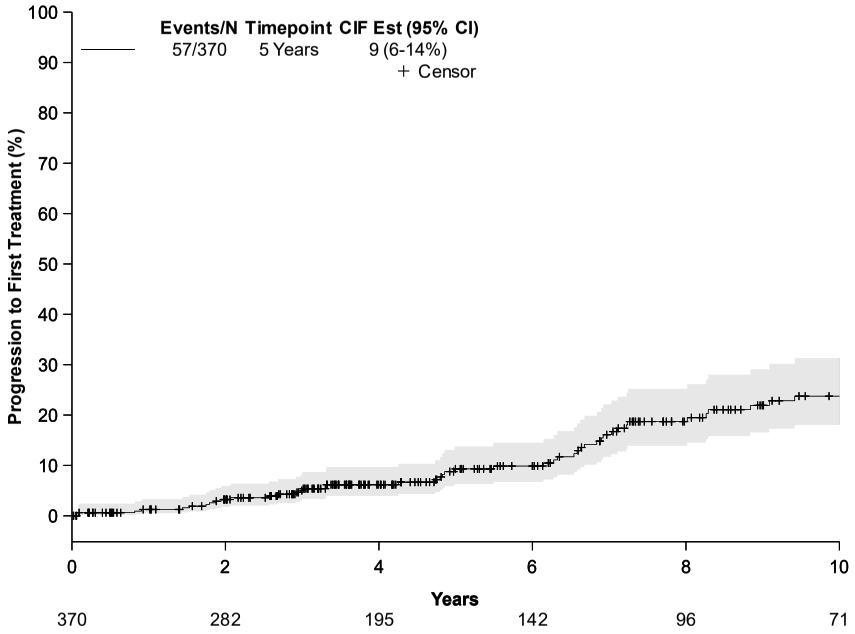


Figure 3

Figure 3: Associations of TML with Time To First Treatment among HCMBL individuals

A.



B.

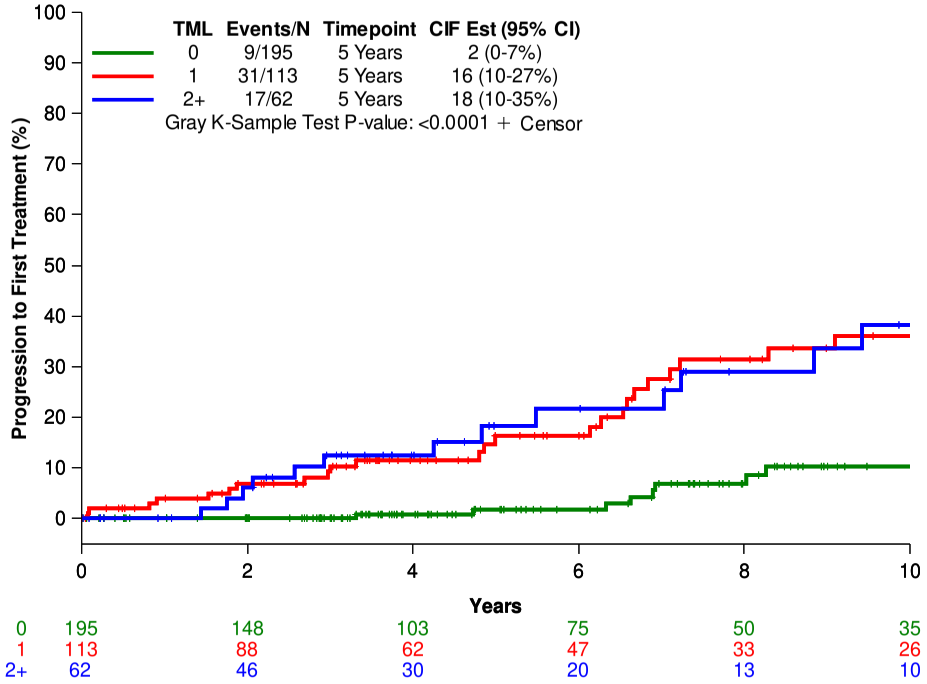


Figure 4

Figure 4: HCMBL and CLL stratified by both TML and CLL-IPI for association with Time to First Treatment

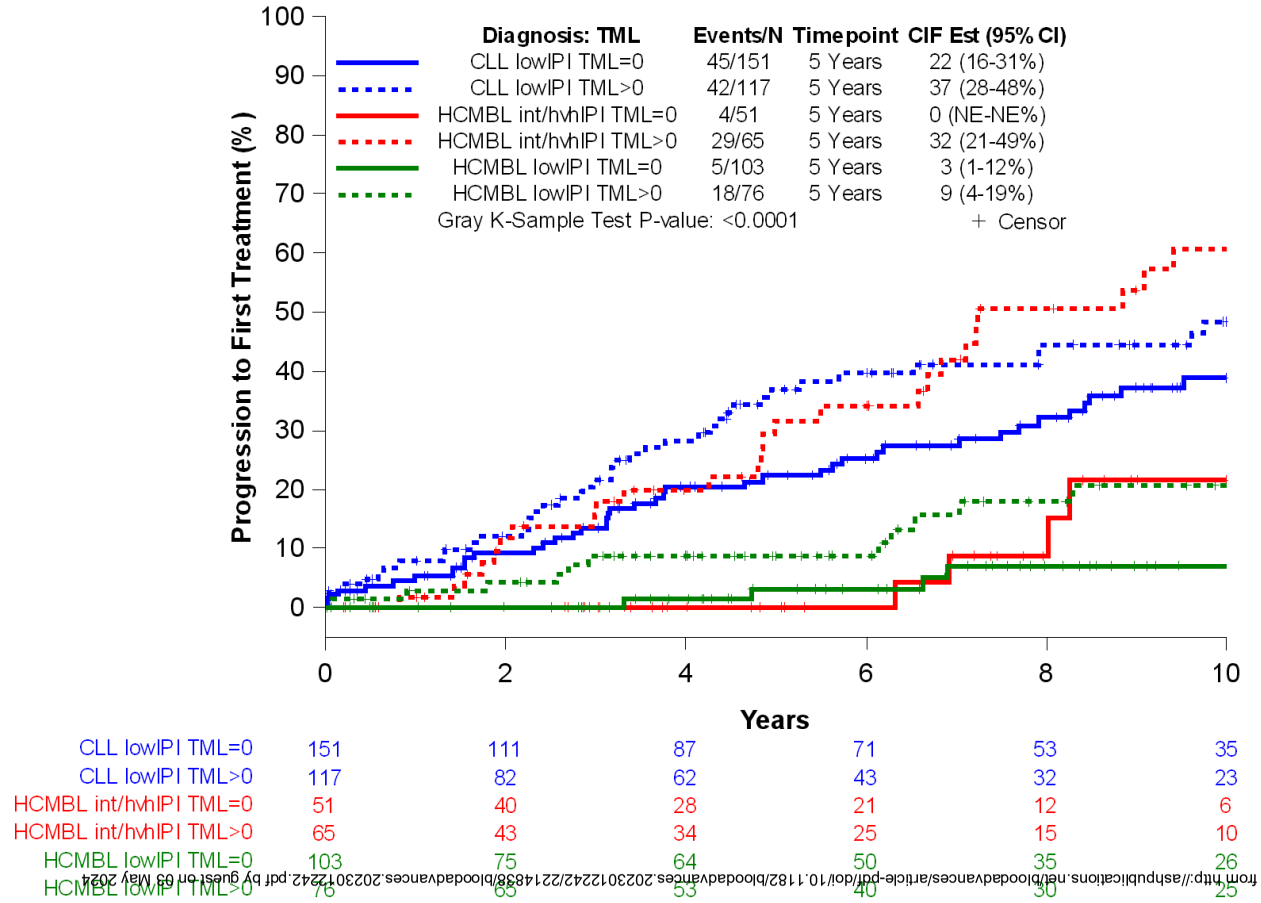
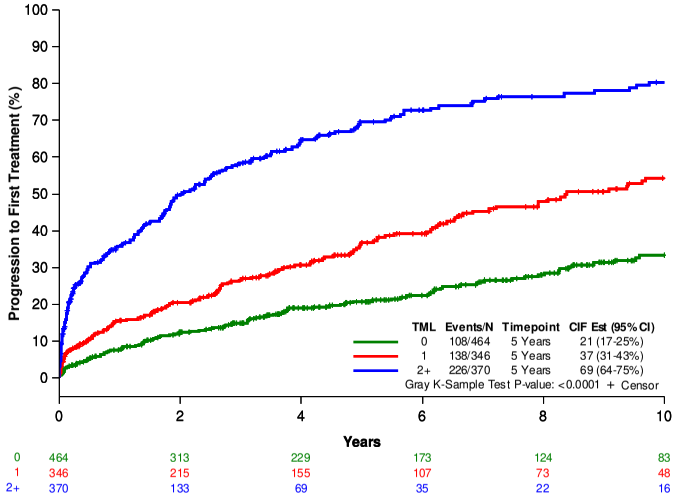


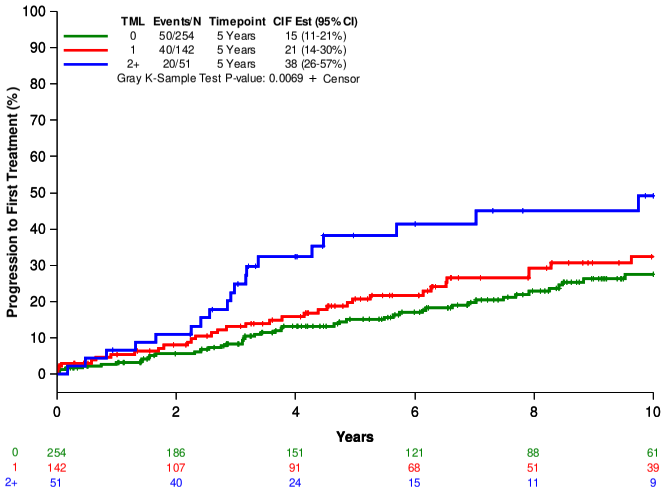
Figure 5

Figure 5: Associations of TML with Time To First Treatment among Full Cohort and Stratified by CLL-IPI

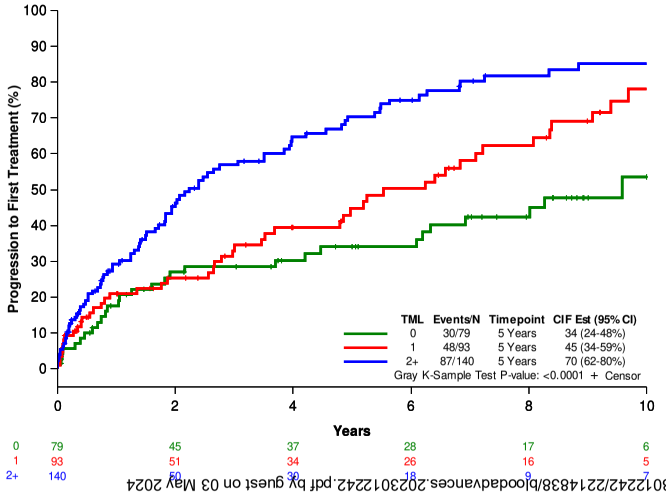
A. Full Cohort



B. Low risk CLL-IPI



C. Intermediate risk CLL-IPI



D. High/Very-high risk CLL-IPI