
Article:
Cargo, Catherine, Cullen, Matthew, Taylor, Jan et al. (6 more authors) (2019) The use of targeted sequencing and flow cytometry to identify patients with a clinically significant monocytosis. Blood. pp. 1325-1334. ISSN: 1528-0020

https://doi.org/10.1182/blood-2018-08-867333

The use of targeted sequencing and flow cytometry to identify patients with a clinically significant monocytosis

Catherine Cargo[1], Matthew Cullen[1], Jan Taylor[1], Mike Short[1], Paul Glover[1], Suzan Van Hoppe[1], Alex Smith[2], Paul Evans[1], and Simon Crouch[2]

[1]Haematological Malignancy Diagnostic Service, St James's University Hospital, Leeds, United Kingdom; [2]Epidemiology and Cancer Statistics Group, University of York, York, United Kingdom

Short Title:

Molecular diagnosis of CMML

Corresponding Author:
Catherine Cargo, MB
Haematological Malignancy Diagnostic Service
St James's University Hospital
Leeds, United Kingdom

Phone - +44 1132067851
Fax - +44 1132067883

E-mail – catherine.cargo@nhs.net

Text word count – 3979
Abstract word count – 250
Number of figures – 6
Number of tables – 1
Reference count – 24

Presented as a poster presentation at the 58th American Society of Hematology Annual Meeting & Exposition, San Diego, CA, December 3-6, 2016

Key Points

Somatic mutations are detected at high frequency in patients with a monocytosis and are associated with significantly reduced survival

In those without a WHO defined diagnosis, patients with a mutation have both laboratory and clinical features indistinguishable from CMML

Abstract

The diagnosis of chronic myelomonocytic leukaemia (CMML) remains centred on morphology, meaning the distinction from a reactive monocytosis is challenging. Mutational analysis and immunophenotyping have been proposed as potential tools for diagnosis however have not been formally assessed in combination. We aimed to investigate the clinical utility of these technologies by performing targeted sequencing, in parallel to current gold standard techniques, on consecutive samples referred for investigation of monocytosis over a 2-year period (n=283). Results were correlated with the morphological diagnosis and objective outcome measures including overall survival (OS) and longitudinal blood counts.

Somatic mutations were detected in 79% of patients, being invariably identified in those with a confirmed diagnosis (99%) though also in 57% of patients with non-diagnostic BM features. The OS in non-diagnostic mutated patients was indistinguishable from those with CMML (p=0.118) and significantly worse than unmutated patients (p=0.0002). On multivariate analysis age, ASXL1, CBL, DNMT3A, NRAS & RUNX1 mutations retained significance. Furthermore, the presence of a mutation was associated with a progressive fall in haemoglobin/platelet levels and increasing monocyte counts compared with mutation negative patients. Of note, the immunophenotypic features of non-diagnostic mutated patients were comparable to CMML patients and the presence of aberrant CD56 was highly specific for detecting a mutation.

Overall, somatic mutations are detected at high frequency in patients referred with a monocytosis irrespective of diagnosis. In those without a WHO defined diagnosis,

the mutation spectrum, immunophenotypic features and OS are indistinguishable from CMML patients and these patients should be managed as such.

Introduction

Distinguishing a reactive monocytosis from Chronic Myelomonocytic Leukaemia (CMML) is challenging for the haematopathologist. Using current WHO diagnostic criteria, a persistent monocytosis is the hallmark of disease and demonstrating clonality is not a definitive requirement[1]. This leads to a greater risk of mis-diagnoses or mis-classification, particularly in patients with prolonged reactive changes.

More recently alternative techniques, in particular flow cytometry, have provided a potential objective tool to identify patients with disease. Skewing of the distribution of monocyte subsets in the PB (>94% M1 monocytes) has been reported to be both sensitive and specific for detecting CMML[2]. In addition, large studies using targeted sequencing panels have identified recurrent somatic mutations in >90% of patients with CMML[3], providing a further potential tool for diagnosis. The presence of a *TET2* mutation in combination with a *SRSF2* (or *ZRSR2*) mutation has been shown to be highly specific for a myelomonocytic phenotype[4] and these along with *ASXL1* are the most frequently mutated genes within this disease group[3]. Whilst the 2016 WHO diagnostic criteria have stated these mutations can support a diagnosis of CMML, there have been no studies directly assessing the use of this technology in a diagnostic setting. The aim of this study was to determine whether mutational analysis and flow cytometry can provide confirmatory evidence of disease and predict outcome in patients presenting with a monocytosis.

Methods

Patients and Samples

The research was undertaken within HMDS (Haematological Malignancy Diagnostic Service), a fully integrated laboratory that serves a population of ~6 million and is the benchmark for haematopathology services within the UK. All consecutive samples (peripheral blood (PB) or bone marrow (BM)) received between July 2014 and July 2016 from patients ≥18yo for the investigation of monocytosis were included. Patients with a confirmed myeloid diagnosis prior to July 2014 were excluded. The decision to investigate was at the discretion of the referring clinician and the study cohort therefore reflects the variety of samples received in a routine laboratory for the investigation of a monocytosis. An absolute monocyte count was determined on all PB samples when received in HMDS (see Table.1) using flow cytometry (see Flow cytometry methods). Interestingly this was calculated to be $<1x10^9$/L on a proportion of samples (11%), however the vast majority were very close to this threshold and review of local blood count parameters and clinical details confirmed the presence of a PB monocytosis and clinical suspicion of CMML. This highlights the recognised variation in monocyte counts between laboratories and the difficulty when applying arbitrary cut-offs as diagnostic criteria.

A total of 283 patients were referred during this time period (Table.1) of which 121 and 162 had an initial PB and BM sample respectively (Fig.1). A confirmed diagnosis was only made on those cases with an ultimate BM sample (n=207). All samples were processed according to gold standard techniques and dual reported by a team of experienced haematopathologists. Those with a confirmed diagnosis were classified in accordance with the WHO 2008 classification. Those failing to

fulfil the morphological and genetic WHO 2008 criteria, as agreed by 2 haematopathologists were classified as ‘non-diagnostic’.

All samples were taken with full-informed patient consent for investigation of a suspected haematological disorder. This study had local Institutional Review Board approval (REC reference-16/NE/0105) and performed in accordance with the Declaration of Helsinki.

Flow Cytometry

All samples for immunophenotypic analysis were processed within 24 hours. Numerical studies and assessment of monocytic CD56 expression were performed on BM or PB samples following a stain-lyse-wash procedure (FACSLyse, Becton Dickinson, supplementary Table S1 and Fig.S1-2). There was strong correlation between monocyte CD56 expression in the PB and BM (supplementary Fig.S5) enabling analysis to include samples from either source.

PB CD14/CD16 “classical” monocytic subset studies were performed on samples following NH4Cl lysis of erythrocytes using a lyse-stain-wash procedure. A minimum of 10^5 leucocytes were acquired on a single cytometer (FACSCanto II, Becton Dickinson) for all cases. Monocytes were identified using a combination of CD64, CD45, and scatter characteristics and analysed by a single operator (MC) for all analyses (supplementary Fig.S3-4).

DNA Extraction and Targeted Amplicon Sequencing

In parallel to the above analyses, samples were subjected to targeted high throughput sequencing. Referring clinicians and haematopathologists were blinded to the results of this analysis to exclude reporting or treatment bias.

DNA was extracted from fresh blood or BM mononuclear cells using the QIAamp DNA mini kit (QIAGEN, Manchester, UK).

Targeted gene sequencing of 27 genes recurrently mutated in myeloid malignancies was performed on the MiSeq (Illumina, Chesterford, UK). Panel design, validation and variant filtering criteria are included in the supplementary methods and Table S2-3. The mean coverage of identified variants was 1514x (range 52-5605x).

Clinical Follow-up

All follow-up BM assessments were performed as clinically indicated by the referring clinician. These samples were also processed according to gold standard techniques and underwent targeted sequencing in parallel as described above. Any subsequent new diagnoses were recorded.

Survival data was available for all patients and censored on the date of extraction (08/08/2017). Additional clinical information, including serial full blood count data, was collected on a sub-cohort of patients (n=182) either directly from the referring hospital or through the Haematological Malignancy Research Network (HMRN, n=85)[5].

Statistical Analysis

Survival curves were produced using the Kaplan Meier method and simple differences in survival were assessed with the log rank test. The impact of abnormalities on overall survival (OS) and risk of progression were estimated using Cox regression; where variable selection was required to arrive at a multivariable regression, the lasso was used for variable selection and results were reported for the corresponding relaxed lasso model.

Sensitivity, specificity, positive and negative predictive values were calculated using 2x2 contingency tables.

Comparison between flow cytometric parameters in the main cohort was performed using Mann-Whitney U test. Correlation between CD56 expression, M1 monocytes and mutational analysis was performed using both logistic and poisson regression.

The effect of mutations on longitudinal blood counts were assessed using random effects models. Four different models were fitted using a full-factorial interaction between time and mutation status: (i) a random intercept model; (ii) a random intercept and slope model with uncorrelated random effects; (iii) a random intercept and slope model with correlated random effects; and (iv) a random intercept and slope model with correlated random effects, additionally adjusted for age and sex. For each mutation/blood count relationship the best-fitting model was chosen according to a likelihood ratio test. To limit any potential affect from periods of acute illness or intensive treatment, blood count trajectory analysis was restricted to those with <40 measurements over >100 days.

Results

Somatic mutations are detected at high frequency in patients with a monocytosis irrespective of diagnosis

To define the mutation spectrum in patients referred with a monocytosis, targeted sequencing results were analysed for the total cohort and correlated with the final diagnosis in those who underwent BM sampling. Of the total 283 patients, ≥1 mutation was detected in 78% of samples, the spectrum of which is presented in Fig.2A (also see Table S4). Of these patients, 207 underwent BM assessment for a definitive diagnosis. HMDS provides a centralised, integrated haematopathology service and all BMs were reviewed independently by 2 haematopathologists ensuring consistent and high quality BM reporting for this purpose.

In those with a confirmed myeloid malignancy (142/207 cases; 69%) a mutation was almost invariably detected (140/142; 99% of cases). Of the 2 mutation negative cases, one had a complex karyotype including inv3 (involving MECOM), leaving only one case with no demonstrable clonal abnormality. The significant majority of diagnostic cases (80%; 114/142) were classified as CMML. The remaining samples were classified with a spectrum of myeloid malignancies though importantly 11 patients were diagnosed with acute myeloid leukaemia (AML) (n=11) highlighting the importance of a BM assessment in patients referred with a monocytosis.

Somatic mutations were also detected at a high frequency in non-diagnostic samples. At least 1 mutation was detected in 37/65 patients (57%) with indeterminate features. The spectrum of mutations in this group mirrored those detected in the diagnostic group with *TET2, SRSF2* and *ASXL1* being most frequently mutated (Fig.2B). The most notable differences in the non-diagnostic

group were the absence of high risk mutations including *TP53, FLT3* and *NPM1* as well as those associated with specific morphological abnormalities such as *SF3B1* which correlates strongly with the presence of ring sideroblasts[6]. The median and mean number of mutations was higher in those with a confirmed diagnosis (median-3, range 0-8; mean-3) versus those without (median-1, range 0-6; mean-2) (Fig.2C). However, in patients with a confirmed mutation, the number of mutations did not differ significantly between diagnostic and non-diagnostic groups (p=0.62).

The median variant allele fraction (VAF) for all variants was 39% (range 5.2-100%; Supplementary Fig.S6) and there was no difference between the VAF in diagnostic and non-diagnostic cases (p=0.33). In those with an isolated mutation the median VAF was also noted to be high at 38.2% (range 6.3-97.1%) with only 2 variants having a VAF of <10%.

Mutations are therefore found at a very high frequency with a high clonal burden in patients with a monocytosis, and involve a similar spectrum of genes, irrespective of diagnosis.

Overall survival and blood count trajectory correlate strongly with mutation profile

To understand the long-term clinical impact of detecting these mutations, objective outcome measures, including OS and longitudinal blood count analysis, were assessed in the total cohort and correlated with the final diagnosis.

The median survival of all patients from the time of first sampling was 35.2 months (95% CI 25mths-not reached; Fig3A). Survival correlated strongly with the number

of mutations. Those without a mutation had a significantly better OS and even the presence of a single mutation resulted in a significant reduction in survival (p=0.004; Fig3B). On univariate analysis, age was strongly associated with survival. For mutations occurring in >5% of subjects, *ASXL1, CBL, DNMT3A, NRAS & RUNX1* were all strongly associated with survival, as were *EZH2* & *STAG2* amongst the less frequently mutated genes. In order to investigate multivariate significance, all genes mutated in >5% subjects were entered into a lasso survival regression. Taking the 1SE shrinkage parameter, age, *ASXL1, CBL, DNMT3A, NRAS & RUNX1* were selected by the lasso and retained significance in a relaxed lasso regression (Supplementary Table.S5).

In those patients who proceeded to a BM biopsy, survival correlated with the final morphological diagnosis. Those without a confirmed diagnosis had a significantly better OS than those with CMML or another myeloid malignancy. However, this survival benefit was retained only in those without a demonstrable mutation (p=0.0002) with mutated patients having a similar survival to CMML patients (p=0.118; NS) (Fig.4).

Longitudinal blood count data was available for 182 patients though restricted to those with <40 measurements over >100 days (n=133) to exclude periods of acute hospital admissions (due to periods of acute illness/infection) or intensive chemotherapy (median follow-up 465 days; range 119-996days). The presence of a mutation was associated with a significantly lower haemoglobin and platelet count and a higher monocyte count relative to those without a mutation which persisted over time and followed a divergent trajectory (Fig.5). With respect to individual mutations, certain mutations were associated with increasing or declining blood count parameters over time (Supplementary Table.S6). Monocyte counts were

found to increase over time in *TET2, SRSF2, ASXL1, NRAS or RUNX1* mutated subjects relative to non-mutated; similarly, white blood counts increased in *ASXL1, NRAS* & *DNMT3A* mutated subjects; and platelet levels decreased in *ASXL1, CBL* and *RUNX1* mutated subjects relative to non-mutated.

In those without a confirmed diagnosis, follow-up BM biopsies were received on 11 patients. Importantly, of those with a subsequent diagnosis of CMML, all had a confirmed mutation on the original sample. In total 7/37 (19%) non-diagnostic mutated patients had a confirmed diagnosis (6 CMML, 1 MDS). Furthermore, none of the mutation negative cases went on to develop CMML, however 2 patients had confirmed alternative haematological diagnoses - DLBCL and Rosai Dorfman disease.

These findings confirm that the presence of a mutation has a significant impact on outcome with respect to both survival and blood count parameters.

Peripheral Blood mutation profiling is predictive of a bone marrow diagnosis

PB mutational analysis has been shown to correlate strongly with BM analysis in MDS providing a potential alternative to BM sampling. To determine if this is also true in CMML, matched PB and BM samples were analysed. A total of 121 PB samples were received as the initial sample and somatic mutations were detected in 66% (80/121). Forty-five patients (45/121) (37%) had a subsequent BM performed for diagnosis. Sequencing failed on 2 of the matched BM samples. Of the 124 variants detected in the remaining 43 patients, there was high concordance between PB and BM (96%) with only 5 discordant results. Importantly these were low level

variants at the limit of detection for the test or variants detected at areas of poor coverage (*SRSF2/ASXL1*). All 9 mutation negative cases were fully concordant.

The presence of a mutation in the PB was highly predictive of diagnosing a myeloid malignancy in BM with all but 1 case with a demonstrable mutation having a subsequent diagnosis (PPV 0.97, NPV 1.0 (Supplementary Figure.S7)). Of note, none of the mutation negative (n=11) cases had a subsequent confirmed diagnosis.

Immunophenotypic features correlate strongly with the presence of a mutation and a subsequent diagnosis

Flow cytometry has been proposed as a potential diagnostic tool in the investigation of patients with a monocytosis. To determine whether immunophenotyping can predict for the presence of a mutation or a BM diagnosis, flow cytometric analysis was performed alongside sequencing.

Firstly, comparison was made between the immunophenotypic features in the BM of those patients with a confirmed diagnosis of CMML versus non-diagnostic samples. Importantly, non-diagnostic mutated patients had immunophenotypic features indistinguishable from CMML with respect to increased CD64+ monocytes, reduced CD14 expression and aberrant CD56 expression on monocytes (Fig.6). This was most pronounced with CD56 expression (in either PB or BM) which was found almost exclusively in those with a mutation. With respect to individual mutations, aberrant expression of CD56 was strongly associated with TET2 mutations (OR 4.0; 95%CI 2.4-6.8; $p<0.0001$).

Peripheral blood monocyte subsets and CD56 expression are predictive of a somatic mutation

The presence of >94% classical (M1) monocytes has been shown to be highly sensitive and specific for a diagnosis of CMML[2]. PB monocyte subset analysis was not available on every patient in the main cohort and to analyse the relationship between M1 monocytes, CD56 expression and the mutation profile, a separate cohort of 135 patients was investigated. Of these 135 patients, 95 underwent a subsequent BM for definitive diagnosis (CMML=28, MDS=23, MPN=9, non-diagnostic=27, other=8). The presence of aberrant CD56 was again strongly associated with the presence of a mutation (OR 43.9; 95%CI 8.9-793.9; p=0.0003). This was also noted, to a lesser extent, with having >94% M1 monocytes (OR 3.9; 95%CI 1.8-8.7; p=0.0007) (supplementary Table.S7). There was some correlation between the presence of CD56 expression and >94% M1 monocytes (r=0.17, p=0.039) and combining both produced a stronger effect (OR 8.5; 95%CI 3.9-19.5; p<0.00001). Importantly, combining these phenotypic aberrancies did not capture all patients with a mutation. Whilst CD56 was highly specific for the presence of a mutation (98%), sensitivity was only 48%. Similarly, the presence of >94% M1 monocytes had a specificity of 75% for detecting a mutation though a sensitivity of only 56%.

With respect to a confirmed diagnosis, both CD56 expression (OR 4.9; 95%CI 1.9-13; p=0.001) and >94% M1 monocytes (OR 4.2; 95%CI 1.7-11.5; p=0.003) were associated with a final diagnosis of CMML though importantly four patients with CMML had neither of these phenotypic aberrancies.

Discussion

This is the first study to formally examine the use of mutational analysis in the investigation of patients presenting with a monocytosis. This was performed in combination with current gold standard techniques, including recently described flow cytometric analyses, in a large patient cohort. By analysing sequential samples referred to a regional diagnostic laboratory this study has investigated the typical patient population encountered in routine haematology practice. The use of objective outcome measures (longitudinal blood counts and OS) and an unselected patient population have minimised bias and ensured the results are applicable in the 'real-world' setting. Using a targeted sequencing panel of recurrently mutated genes, this study confirms that somatic mutations are not only identified in virtually all patients with a morphological diagnosis of CMML, but also in a significant proportion of patients with a monocytosis and non-diagnostic features. It is possible that the proportion of non-diagnostic samples with detectable mutations was inflated due to referral bias and a high pre-test probability of disease in those undergoing testing, however these patients had a mutation spectrum, immunophenotype and outcome indistinguishable from CMML. The presence of a mutation significantly impacted on survival irrespective of the final diagnosis.

A number of technical limitations of this study should be highlighted. Firstly, as these were routine samples referred for investigation, a corresponding germline sample was not available for analysis. The absence of reference material means that the distinction between germline variants or private single nucleotide polymorphisms (SNPs) and somatic variants is challenging. However, sequencing was limited to well documented driver genes and the landscape of mutations in these genes is well

established. Strict filtering criteria were applied (see Supplementary methods) to ensure only high confidence variants were included. Secondly, the sequencing analysis used amplicon based library preparation which has recognised limitations with respect to PCR errors and false positive results particularly at low VAF. The panel was however internally and externally validated (see supplementary methods) and only reproducible variants were included if detected at low VAF or in areas of low coverage. The results are therefore, to the best of our ability, accurate. In future, sequencing deeper should enable more accurate variant calling at low VAF.

The findings of this study will be key to refining future diagnostic algorithms in the investigation of patients referred with a monocytosis. Mutational analysis has been incorporated into the recent amendment of the WHO diagnostic criteria which now states that the presence of a mutation can support a diagnosis of CMML. Concerns have however been raised regarding the use of mutational analysis in this setting, due to reports of frequent somatic mutations in aging healthy individuals[7–10]. As a result the WHO have stated that the presence of a mutation in either CMML or MDS should not be used alone as proof of disease[1]. Our study has however shown that even in the absence of morphological features, those patients with a mutation had a clinical phenotype and genotype indistinguishable from CMML and a comparably poor outcome. Distinguishing features were also noted between the variants reported in healthy individuals and the mutations detected in our study group. The VAF or clone size of the mutations in our study were significantly higher than in healthy individuals (median 39.2% vs 9-10%), and this was demonstrated across both diagnostic and non-diagnostic samples. This finding has also been described in patients with unexplained cytopenias and several studies have shown that a VAF

>10% and the presence of co-occurring mutations can distinguish clinically significant cytopenias from healthy individuals[11–13]. While the higher VAF is replicated in our patient group, importantly our study has shown that even isolated mutations have a significant impact on survival in patients with a monocytosis. These findings provide strong evidence that in those without diagnostic morphological features the presence of a mutation, irrespective of mutation number, could be disease defining. At the very minimum, it is imperative that these patients are identified and monitored closely.

It has become increasingly feasible to perform mutational analysis in routine clinical practice and this study has demonstrated how modest sized gene panels can provide significant diagnostic and prognostic information. The panel used in the study targeted genes implicated in myeloid malignancies and was incorporated into the routine workload and performed in 'real time' in a cost-effective manner. The genetic profile in CMML is now well established and is noted to be relatively homogeneous involving only a restricted number of genes. Mutation frequencies of >90% of patients have been consistently reported using varying panel sizes ranging from as few as 19 genes[3,14–16]. The mutation profile in our cohort mirrored that reported in the literature and despite the restricted panel, the mutation frequency was high and a significant impact on outcome was demonstrated. The recognised poor prognostic impact of ASXL1 mutations[3,14,17–20], was also replicated across this dataset. Mutational analysis is therefore viable in a routine diagnostic laboratory. It is also likely that a proportion of these patients will have additional mutations in genes not sequenced in this study. To further investigate this would require more extensive sequencing on much larger patient populations.

The potential for PB to be used as a screening tool for monocytosis has also been addressed in this study. This is an attractive option, particularly in a disease commonly presenting in the older patient population. Using flow cytometry, the presence of >94% M1 monocytes in the PB was reported to be both highly sensitive and specific for CMML[2]. Subsequent studies have validated these findings and also confirmed the ability to distinguish CMML from both MDS and MPN cases presenting with a monocytosis[21,22]. These studies however are centred on morphological diagnoses and have not consistently performed mutational analysis. While our study has shown a strong correlation between skewed monocyte subsets and a diagnosis of CMML this did not capture all patients and was neither sensitive nor specific for the presence of a mutation. In contrast aberrant CD56 expression was highly specific for the presence of a mutation (98%), particularly involving *TET2*. CD56 expression has previously been reported to be highly sensitive and specific for a CMML diagnosis (100% and 67% respectively) when combined with other immunophenotypic features including reduced expression of myeloid antigens and ≥20% immature monocytes[23]. Subsequent studies however raised concerns regarding the overexpression of CD56 in reactive conditions[24]. Our data show that CD56 expression at diagnosis is invariably associated with the presence of a somatic mutation though sensitivity was low (48%). Flow cytometry could, therefore, provide a screening tool for the investigation of PB monocytes but, ultimately, mutational analysis will be required to identify patients who require clinical follow-up. Importantly there was high concordance between PB and BM mutational analysis and the presence of a PB mutation was highly predictive of a subsequent BM diagnosis. This suggests that screening of the PB may be a suitable method for

identifying or excluding significant mutations, however this could lead to a rise in inappropriate referrals and a significant burden on laboratory workload. Furthermore, the small proportion of mutated patients in our cohort with other haematological malignancies in the BM, including AML, highlights the importance of a baseline BM assessment to definitively classify the disease. In contrast the negative predictive value of PB screening was 100% suggesting that those without a mutation should not undergo BM assessment. In the first instance, PB screening would be a practical option in those patients unfit for BM assessment or potentially to monitor for treatment response or disease evolution. The latter would require further investigation in a prospective study.

In conclusion, this study has confirmed that mutations are commonly detected in patients referred with a persistent monocytosis. The presence of a mutation impacts significantly on outcome irrespective of diagnosis, and patients with a mutation who fail to meet WHO criteria have CMML disease characteristics. These findings validate the inclusion of somatic mutations in the diagnostic criteria for CMML and at the very minimum those without a confirmed diagnosis require close clinical follow-up. While PB can be confidently used to detect mutations, a baseline BM biopsy is required for definitive disease classification in patients fit for treatment. Immunophenotypic assessment of monocytes may provide a potential screening tool to detect those with a mutation however it will miss a proportion of mutated patients. Ultimately, early identification of patients could provide an opportunity for intervention in this patient group and this requires further investigation.

Acknowledgements

The authors thank all clinicians and laboratory staff who assisted in collecting longitudinal and outcome data.

Dr Alex Smith and Dr Simon Crouch are supported by a Bloodwise programme grant.

Authorship Contributions

CC and SC designed the study. CC, MC, MS, SVH and PE performed the research. CC, MC, JT, PG, AS and SC analysed the data. CC, and SC wrote the paper.

Disclosure of Conflicts of Interest

CC and PE have contributed to advisory boards for Novartis. The remaining authors declare no conflict of interest.

References

1. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016;127(20):2391–2405.
2. Selimoglu-Buet D, Wagner-Ballon O, Saada V, et al. Characteristic repartition of monocyte subsets as a diagnostic signature of chronic myelomonocytic leukemia. *Blood*. 2015;125(23):3618–3626.

3. Elena C, Gallì A, Such E, et al. Integrating clinical features and genetic lesions in the risk assessment of patients with chronic myelomonocytic leukemia. *Blood*. 2016;128(10):1408–1417.
4. Malcovati L, Papaemmanuil E, Ambaglio I, et al. Driver somatic mutations identify distinct disease entities within myeloid neoplasms with myelodysplasia. *Blood*. 2014;124(9):1513–1521.
5. Roman E, Smith A, Appleton S, et al. Myeloid malignancies in the real-world: Occurrence, progression and survival in the UK's population-based Haematological Malignancy Research Network 2004-15. *Cancer Epidemiol.* 2016;42:186–198.
6. Papaemmanuil E, Cazzola M, Boultwood J, et al. Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. *N. Engl. J. Med.* 2011;365(15):1384–1395.
7. Genovese G, Kähler AK, Handsaker RE, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N. Engl. J. Med.* 2014;371(26):2477–2487.
8. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N. Engl. J. Med.* 2014;371(26):2488–2498.
9. McKerrell T, Park N, Moreno T, et al. Leukemia-associated somatic mutations drive distinct patterns of age-related clonal hemopoiesis. *Cell Rep.* 2015;10(8):1239–1245.
10. Xie M, Lu C, Wang J, et al. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat. Med.* 2014;20(12):1472–1478.
11. Cargo CA, Rowbotham N, Evans PA, et al. Targeted sequencing identifies patients with preclinical MDS at high risk of disease progression. *Blood*. 2015;126(21):2362–2365.

12. Kwok B, Hall JM, Witte JS, et al. MDS-associated somatic mutations and clonal hematopoiesis are common in idiopathic cytopenias of undetermined significance. *Blood*. 2015;126(21):2355–2361.
13. Malcovati L, Gallì A, Travaglino E, et al. Clinical significance of somatic mutation in unexplained blood cytopenia. *Blood*. 2017;blood-2017-01-763425.
14. Itzykson R, Kosmider O, Renneville A, et al. Prognostic Score Including Gene Mutations in Chronic Myelomonocytic Leukemia. *J. Clin. Oncol.* 2013;31(19):2428–2436.
15. Mason CC, Khorashad JS, Tantravahi SK, et al. Age-related mutations and chronic myelomonocytic leukemia. *Leukemia*. 2016;30(4):906–913.
16. Kohlmann A, Grossmann V, Klein H, et al. Next-generation sequencing technology reveals a characteristic pattern of molecular mutations in 72.8% of chronic myelomonocytic leukemia by detecting frequent alterations in TET2, CBL, RAS, and RUNX1. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 2010;28(24):3858–3865.
17. Patnaik MM, Padron E, Laborde RR, et al. Mayo Prognostic Model for WHO-Defined Chronic Myelomonocytic Leukemia- ASXL1 and Spliceosome Component Mutations and Outcomes. *Leukemia*. 2013;
18. Patnaik MM, Lasho TL, Vijayvargiya P, et al. Prognostic interaction between ASXL1 and TET2 mutations in chronic myelomonocytic leukemia. *Blood Cancer J.* 2016;6:e385.
19. Patnaik MM, Wassie EA, Padron E, et al. Chronic myelomonocytic leukemia in younger patients: molecular and cytogenetic predictors of survival and treatment outcome. *Blood Cancer J.* 2015;5:e270.

20. Patnaik MM, Itzykson R, Lasho TL, et al. ASXL1 and SETBP1 mutations and their prognostic contribution in chronic myelomonocytic leukemia: a two-center study of 466 patients. *Leukemia*. 2014;28(11):2206–2212.
21. Patnaik MM, Timm MM, Vallapureddy R, et al. Flow cytometry based monocyte subset analysis accurately distinguishes chronic myelomonocytic leukemia from myeloproliferative neoplasms with associated monocytosis. *Blood Cancer J.* 2017;7(7):e584;
22. Talati C, Zhang L, Shaheen G, et al. Monocyte subset analysis accurately distinguishes CMML from MDS and is associated with a favorable MDS prognosis. *Blood*. 2017;129(13):1881–1883.
23. Xu Y, McKenna R, Karandikar N, Pildain A, Kroft S. Flow cytometric analysis of monocytes as a tool for distinguishing chronic myelomonocytic leukemia from reactive monocytosis. *Am. J. Clin. Pathol.* 2005;124(5):799–806.
24. Lacronique-Gazaille C, Chaury M, Le Guyader A, et al. A simple method for detection of major phenotypic abnormalities in myelodysplastic syndromes: expression of CD56 in CMML. *Haematologica*. 2007;92(6):859–860.

Figure Legends

Figure 1. Summary of samples included in study. Flowchart of cases referred for investigation of a monocytosis to HMDS.

*Decision to investigate was at the discretion of the referring clinician

†Focal area of Diffuse Large B-cell Lymphoma noted in bone marrow, likely co-occurring with CMML

Figure 2. Characteristics of mutations detected in patient samples. (A) Spectrum of mutations detected across all patients in study (n=283)

(B) Comparison of mutations detected in those with a diagnostic bone marrow sample vs a non-diagnostic bone marrow sample (diagnostic (n=142) vs non-diagnostic (n=65))

(C) Distribution of no. of mutations according to final diagnostic category. 'Other' denotes those patients with an alternative haematological malignancy.

Figure 3. Overall survival according to mutation number. (A) Overall survival in total cohort from time of initial sample

(B) Overall survival in total cohort by no. of mutations detected at time of initial sample. The p-value represents a log rank test comparing those without a mutation to those with a single mutation.

(C) Overall survival in total cohort by the presence or absence of a mutation

Figure 4. Overall survival according to final diagnosis. (A) Overall survival by diagnosis on bone marrow sample (n=207)

(B) Overall survival by diagnosis with non-diagnostic samples separated by the presence or absence of a mutation. The p-values refer to log-rank tests comparing CMML and non-diagnostic unmutated patients (p=0.0002) and comparing CMML with non-diagnostic mutated patients (p=0.118)

Figure 5. Longitudinal blood count trajectories in relation to mutation status. Plots of all blood count trajectories averaged between mutated (red) and unmutated (black) groups with overlaid linear regression line.
(A) Haemoglobin (g/L) in patients with or without a detectable mutation. (B) Platelet count (log transformed) in patients with or without a detectable mutation. (C) Monocyte count (log transformed) in patients with or without a detectable mutation.

Figure 6. Relationship between immunophenotype and mutations. Box and whisker plots comparing immunophenotypic features of CMML, non-diagnostic mutated (NDM) and non-diagnostic unmutated (NDU) cases. The p-values refer to Mann-Whitney U tests comparing CMML with either non-diagnostic category.
(A) % CD56 expression on monocytes. (B) % CD64+ monocytes of leucocytes
(C) % CD14 expressing monocytes

Table 1. Patient Characteristics				
No. of patients				283
Male:Female				174:109
Median Age (range)				76 (24-96)
Final Diagnosis				
	PB only			76
	CMML			114
	AML			11
	MPN			9
	MDS			4
	Other			4
	Non-diagnostic			65
		CMML	Other haem malignancy	Non-diagnostic
Median Age (Range)		76 (24-91)	76 (42-93)	73 (34-93)
Blood count parameters: Median (Range)				
Haemoglobin (g/L)		105.5 (38-161)	108 (53-174)	122 (84-163)
White cell count ($x10^9$/L)		13.6 (4-104.9)	10.6 (3.9-83.4)	7.9 (4.2-38.2)
Platelets ($x10^9$/L)		90 (1-442)	154 (39-1085)	150 (8-499)
Monocytes ($x10^9$/L)†		2.69 (0.47-23.59)	1.71 (0.23-9.57)	1.29 (0.72-4.08)
Mutation Frequency: no. of patients (%)				
TET2		72 (63%)	9 (32%)	31 (48%)
SRSF2		48 (42%)	9 (32%)	14 (22%)
ASXL1		39 (34%)	13 (46%)	10 (15%)
NRAS		17 (15%)	7 (25%)	5 (8%)
RUNX1		16 (14%)	6 (21%)	4 (6%)
DNMT3A		9 (8%)	5 (18%)	5 (8%)
CBL		18 (16%)	1 (4%)	4 (6%)
KRAS		9 (8%)	3 (11%)	2 (3%)
SETBP1		7 (6%)	2 (7%)	1 (2%)
JAK2		3 (3%)	7 (25%)	1 (2%)
EZH2		8 (7%)	2 (7%)	1 (2%)
SF3B1		6 (5%)	2 (7%)	0 (0%)

Abbreviations: PB, peripheral blood; CMML, chronic myelomonocytic leukaemia; AML, acute myeloid leukaemia; MPN, myeloproliferative neoplasm; MDS, myelodysplastic syndrome

† Monocyte count determined by flow cytometry (see methods)