Insights from the 2023 Global Leukemia Experience Survey: CLL specific findings, unmet needs and recommendations for action.

Pathways, Immunity Awareness, and Treatment Decisions

September 2024

Nick York, Kathryn Huntley, Jan Rynne, AJ Poots, Sarah Gunn

CLL Advocates Network (CLLAN)

The CLL Advocates Network (CLLAN) was founded in 2014 by representatives from Canada, Czech Republic, France, Ireland, Italy, Portugal, UK and the United States. It is hosted under the umbrella of the Leukaemia Patient Advocates Foundation (LePAF), a patient-led non-profit foundation based in Switzerland acting as a legal platform for self-sustained patient advocacy initiatives. The CLL Advocates Network is governed by a Steering Committee consisting of the 10 members, of whom 6 are patients and 1 is a carer.

CLLAN Mission

- Improve CLL patient outcomes as a global network of CLL patient advocates.

Picker

Picker is a leading international health and social care charity. We carry out research to understand individuals' needs and their experiences of care. We are here to:

- Influence policy and practice so that health and social care systems are always centred around people's needs and preferences.
- Inspire the delivery of the highest quality care, developing tools and services which enable all experiences to be better understood.
- Empower those working in health and social care to improve experiences by effectively measuring, and acting upon, people's feedback

Contents

Figures

Tables

Boxes

Executive Summary

The Chronic Lymphocytic Leukaemia Advocates Network (CLLAN) conducted an analysis of the 2023 Global Leukemia Experience Survey data to examine diagnostic pathways, immunity status, and treatment decision experiences of people with CLL. The survey was a collaboration between Acute Leukemia Advocates Network (ALAN), CML Advocates Network (CMLAN) and CLLAN. It was released online in 13 languages to international respondents aged 18 or more, distributed via the advocacy networks.

We provide the key findings and implications for people with CLL, and the key findings and recommendations for clinicians and advocacy groups:

Key findings and implications for people with CLL

Lack of knowledge of symptoms and delay in visiting healthcare professionals

The proportion of respondents with CLL who were not aware that the symptoms they were experiencing could be due to leukemia and waited to visit a clinician is highly concerning. It is important that people seek medical attention at the earliest opportunity. Improving symptom awareness in the general population and educate primary care clinicians.

Understanding of and engagement with the active monitoring strategy

Respondents do not fully understand the active monitoring strategy and may lack adequate support in managing symptoms. People need to be educated on how to recognise CLL progression and when to seek further medical advice. Clear information, education and support to aid understanding of active monitoring can empower patients and impact on emotional psychological anxiety burden.

Impact of immunity awareness on quality of life

Respondents who did not receive clear information about what steps to take to stay healthy and avoid infections reported worse quality of life. Respondents who received complete information about CLL immunisation / vaccination protocols reported better quality of life.

Key findings and recommendations for clinicians and advocacy groups

Enhance symptom awareness

Develop and implement targeted educational campaigns, with emphasis on primary care empowerment, to reduce clinical delays in referral escalation and diagnosis, and emphasis early symptom recognition and timely consultations to reduce delays in patients presenting.

Improve diagnostic clarity

Many respondents with CLL did not receive clear information on their diagnosis. Once a diagnosis is made, ensure that explanations given are clear and comprehensive, including understanding the nature of the disease, progression, potential compromised immunity and varied prognoses.

Enhance communication and support for active monitoring plans

Prioritise clear, consistent communication about active monitoring plans, ensuring that people fully understand what it entails and how to manage their symptoms and progression.

Focus on person-centred immunity status education

Given the links between immunity awareness and quality of life, clinicians could focus on delivering clear and practical guidance on managing compromised immunity by providing personalised advice, including vaccination / immunisation protocols for CLL.

Strengthen involvement in treatment decisions

Clinicians should actively involve people in treatment decisions, offering them choices and ensuring they understand the implications of each option. This may be achieved by empowering clinicians in communicating in this way with patients and by empowering patients with clear information and learning opportunities to be more able and confident to become involved in their own healthcare discussions and decisions with their doctors.

Background and research aim

The Chronic Lymphocytic Leukemia Advocates Network (CLLAN) commissioned Picker to examine diagnostic pathways, immunity status, and treatment decision experiences of people with CLL using data from the 2023 Global Leukemia Experience Survey. The survey was commissioned by a collaboration of advocacy groups (Acute Leukaemia Advocates Network, ALAN; CML Advocates Network, CMLAN; and CLLAN). It was released online in 13 languages to international respondents aged 18 years or more, distributed via the advocacy networks. The 2023 survey received 2,260 patient responses across all leukemia types, of which 846 responses were from patients with CLL, who were from 30 countries, mainly from UK, USA, and Canada; 54% (454/846) were female; 84% (709/846) were more than 56 years old (Appendix A1: Respondent characteristics).

CLLAN's commitment to patient advocacy underscores the importance of understanding unmet needs and ensuring that all patients have a voice in their care; that their choices are respected and supported. CLLAN initiated this study to explore the experience of people with CLL, a condition characterised by its often-slow progression and complex treatment decisions. Unlike other leukaemia types, patients with CLL are more likely to experience periods where active treatment is deferred in favour of an ‘active monitoring’, ‘active surveillance, or ‘watch and wait’ strategy[1]. Active monitoring closely monitors a patient’s condition without giving treatment until symptoms appear or change and may be used to treat asymptomatic, symptomatic or progressive CLL. While medically justified, this approach may lead to anxiety and uncertainty in patients if they do not understand their prognosis, how to manage their symptoms, how to monitor their CLL progression, or the rationale for delaying treatment. Thus, understanding patient experiences during this phase will provide insight into unmet needs during this period.

People with CLL are vulnerable to infections due to their compromised immune systems. Understanding how people perceive and manage this risk, and how it impacts their quality of life is vital to informing support mechanisms for people living with CLL. People with CLL who are highly aware of their immunity status may experience heightened anxiety and take measures that, while protective, might lead to social isolation or other lifestyle changes that diminish wellbeing. Conversely, people who lack adequate knowledge about their immune health may unknowingly expose themselves to greater risks.

There is a broader concern for how treatment decisions are made and communicated to people with CLL. Given the nature of CLL, it is essential that people are informed and actively involved in their treatment planning. Empowering people through clear communication and shared decision-making is fundamental to improving their overall experience, quality of life and treatment outcomes.

[1] For the purposes of this report, active monitoring will be used interchangeably with the terms ‘active surveillance’ and ‘watch and wait’.

Research questions

To address the research aim, the following research questions (RQs) were established:

RQ1 What is the active monitoring and diagnostic pathway experience for people with CLL?

RQ1.1 How does this contrast with other types of leukaemia?

RQ2. Are people who are more aware of their immunity status, and thus risk of infection, more likely to have lower Quality of Life (QoL) than people who are less aware?

RQ3. Can we classify and predict experiences of treatment decision making?

Analytical methods for this report are outlined in Appendix A2: Methods.

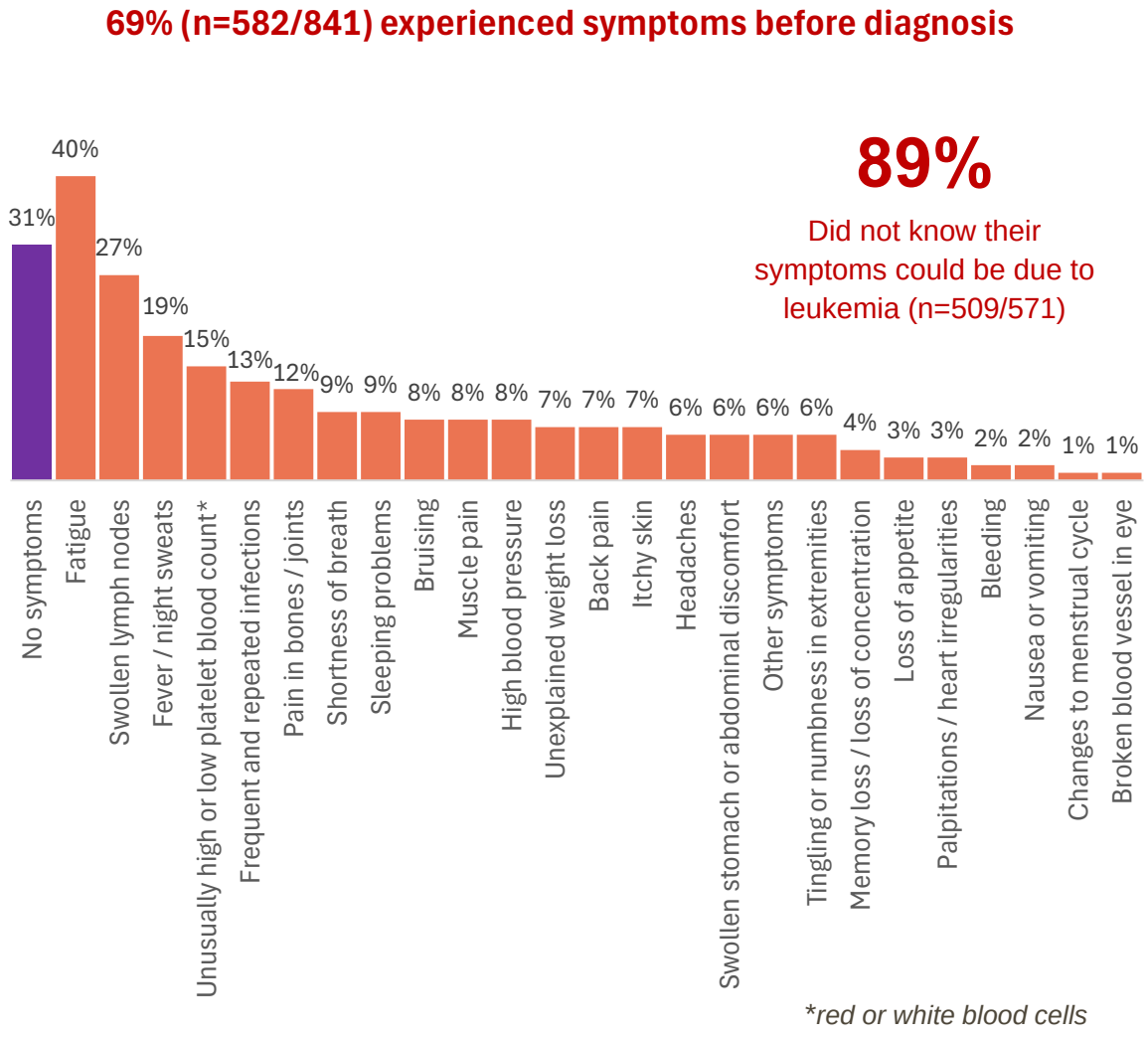
Findings

Top level summaries and visualisations of RQs are provided in this section. Detailed visualisations can be found in Appendix A3: RQ1, Appendix A4: RQ2, Appendix A5: RQ3.

RQ1. What is the active monitoring, and diagnostic pathway experience for people with CLL?

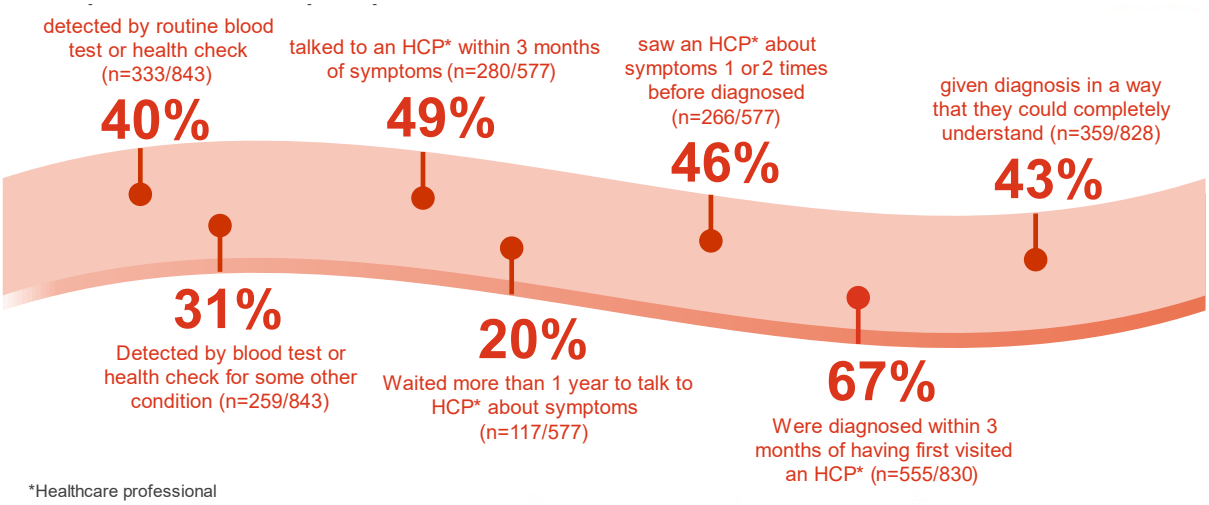
69% (n=582/841) of CLL respondents experienced symptoms before diagnosis. 89% (n=509/571) did not know their symptoms encountered before diagnosis could be due to leukemia. The most commonly experienced symptoms were fatigue (40%, n=340/841), swollen lymph nodes (27%, n=223/841) and fever and night sweats (19%, n=161/841) (Figure 1, Appendix A3: RQ1).

Figure 1. Symptoms encountered before CLL diagnosis



Of CLL respondents, 71% (n=592/843) had their leukemia detected by a blood test or health check. 49% (n=280/577) spoke to a Health Care Professional (HCP) within 3 months of first experiencing symptoms; 20% (n=117/577) waited more than 1 year. Nearly half (46%, n=266/577) stated that they saw a HCP about their symptoms one or two times before they were diagnosed. 67% (n=555/830) were diagnosed within 3 months of first talking to a HCP, and 43% (n=359/828) said they were given their diagnosis in a way that they could completely understand (Figure 2, Appendix A3: RQ1).

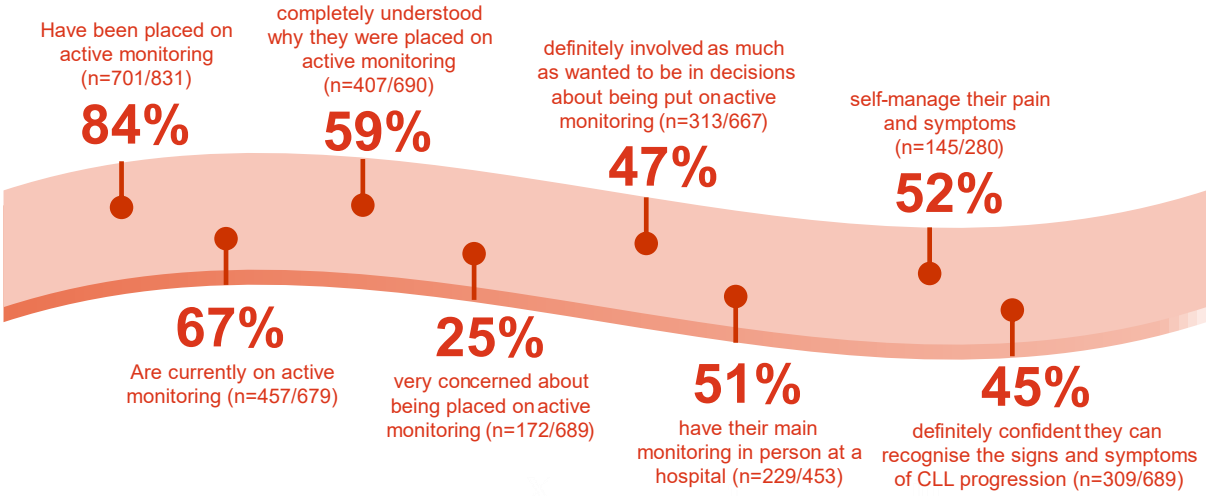
Figure 2. Selected experiences of the diagnostic pathway for people with CLL. HCP = Healthcare professional



The majority (84%, n=701/831) of CLL respondents to the 2023 Global Leukemia Experience Survey have been placed on active monitoring, with 67% (n=457/679) indicating they are currently on active monitoring. Six in ten (59%, n=407/690) completely understood the explanation they received about why they were placed on active monitoring for the first time. A quarter (25%, n=172/689) of respondents with CLL stated they were *very concerned* about being placed on active monitoring. Almost half (47%, n=313/667) said they were *definitely involved* as much as they wanted to be in decisions about being put on active monitoring.

Half (51%, n=229/453) stated that the main way their monitoring visits take place was in person at the hospital. 38% (n=173/453) reported that they do not experience any pain or symptoms. Where pain or symptoms were present, 52% (n=145/280) self-manage, 22% (n=62/280) receive help from their hospital doctor, and 21% (n=59/280) receive help from their GP or family doctor. Near half (45%, n=309/689) respondents with CLL were *definitely confident* in their ability to recognise the signs and symptoms of CLL progression (Figure 3, Appendix A3: RQ1).

Figure 3. Selected experiences of the 'active monitoring' pathway for people with CLL



Information provision throughout patient journey

Only 76% (n=588/770) of CLL patients reported being given or directed to information or support during their treatment journey. As shown in

Table 1, this was most likely to be written information / booklets / leaflets (44%, n=336/770), direction to leukemia charities / associations / organizations (34%, n=263/770) and information on side effects and risks of treatment (31%, n=239/770).

*Table 1. Information provision throughout treatment journey, respondents could select multiple categories. **

Base (n=)	CLL 770
Written information / booklets / leaflets	44%
Leukemia charities / associations / organizations	34%
Side effects and risks of treatment	31%
Patient support groups	29%
Online forums	21%
Clinical trials	19%
Access to a Clinical Nurse Specialist (CNS)	19%
Mental wellbeing (emotional) including referral to psychological support, counselling or psychotherapy	18%
General wellbeing (including spiritual wellbeing)	14%
Patient rights	12%
Financial information	7%
Buddying or befriending	6%
End of life care	2%
I was not given or directed to information or support	**24%**

please note access to a Clinical Nurse Specialist (CNS) was shown only to respondents from the United Kingdom.

RQ1.1 How does this contrast with other types of leukaemia?

Compared to other leukemia types, CLL patients were more likely, than statistically expected, to (**Error! Reference source not found.**Appendix A3: RQ1):

- Not experience symptoms
- Wait more than 2 years to speak to a HCP from the point of first experiencing symptoms
- Visit a HCP more than 4 times before being given a diagnosis
- Wait more than 2 years from first speaking to a HCP to receiving their diagnosis
- Report that their diagnosis was not explained to them in a way they could understand.

RQ2. Are people more aware of their immunity status, and thus risk of infection, more likely to have lower Quality of Life than people who are less aware?

The Hematological Malignancy Patient-Reported Outcome Measure (HM-PRO) [2] is a validated questionnaire. It is a composite measure combining impact of disease and treatment on the quality of life (QoL) of a patient, Part A, and the resulting signs and symptoms, Part B. HM-PRO is the only generic hematological malignancy specific patient-reported outcome measure covering quality of life and symptoms. It has been developed directly from the experience of patients, for patients.

Our analysis focused on HM-PRO Part A which has 24 items and a score range of 0-48. The higher the total score, the greater the effect on a patient's QoL.

Three questions were included to ascertain knowledge of immunity and impact on quality of life (Appendix A4: RQ2):

- Do you know your immunity status? (Q72)
- Were you given clear information about what steps you can take to stay healthy and avoid infections? (Q73)
- Has a health professional spoken to you about the protocols for CLL immunisations / vaccinations, including which you should receive and which you should avoid? (Q74)

No significant relationship was detected between a patient knowing their immunity status (Q72) and the patient's quality of life (HM-PRO).

There was a significant relationship between patient's being given clear information about what steps they can take to stay healthy and avoid infections (Q73) and their quality of life. Respondents who indicated they were not given clear information about what steps they can take to stay healthy and avoid infections, were more likely than statistically expected to report a higher impact on their QoL (HM-PRO Part A score of 29-48).

There was a significant relationship between patient's receiving complete information about the protocols for CLL immunisations / vaccinations, including which they should receive and which they should avoid (Q74), and their quality of life. Respondents who indicated they received complete information about the protocols for CLL immunisations / vaccinations, including which they should receive and which you should avoid, were more likely than statistically expected to report a lower impact on their QoL (Part A score of 0-8). Interestingly, respondents who indicated they received partial information were *less* likely than statistically expected to report a lower impact on their QoL (Part A score of 0-8).

Respondents who do not receive clear information about what steps they can take to stay healthy and avoid infections report a higher impact to their psychological wellbeing, while respondents who received complete information about protocols on immunisations / vaccinations reported little no to impact on their psychological wellbeing.

[2] HM-PRO, https://hmpro.co.uk/ [accessed: 12 September 2024]

RQ3. Can we classify and predict experiences of treatment decision making?

We can predict people's responses to ***Q40 Were you involved as much as you wanted to be in decisions about your treatment?*** with moderate accuracy (68%,Appendix A5: RQ3). Our model has an Area Under Curve (AUC) of 0.76 overall, which is moderately high in performance, but would not be sufficient for clinical decisions.

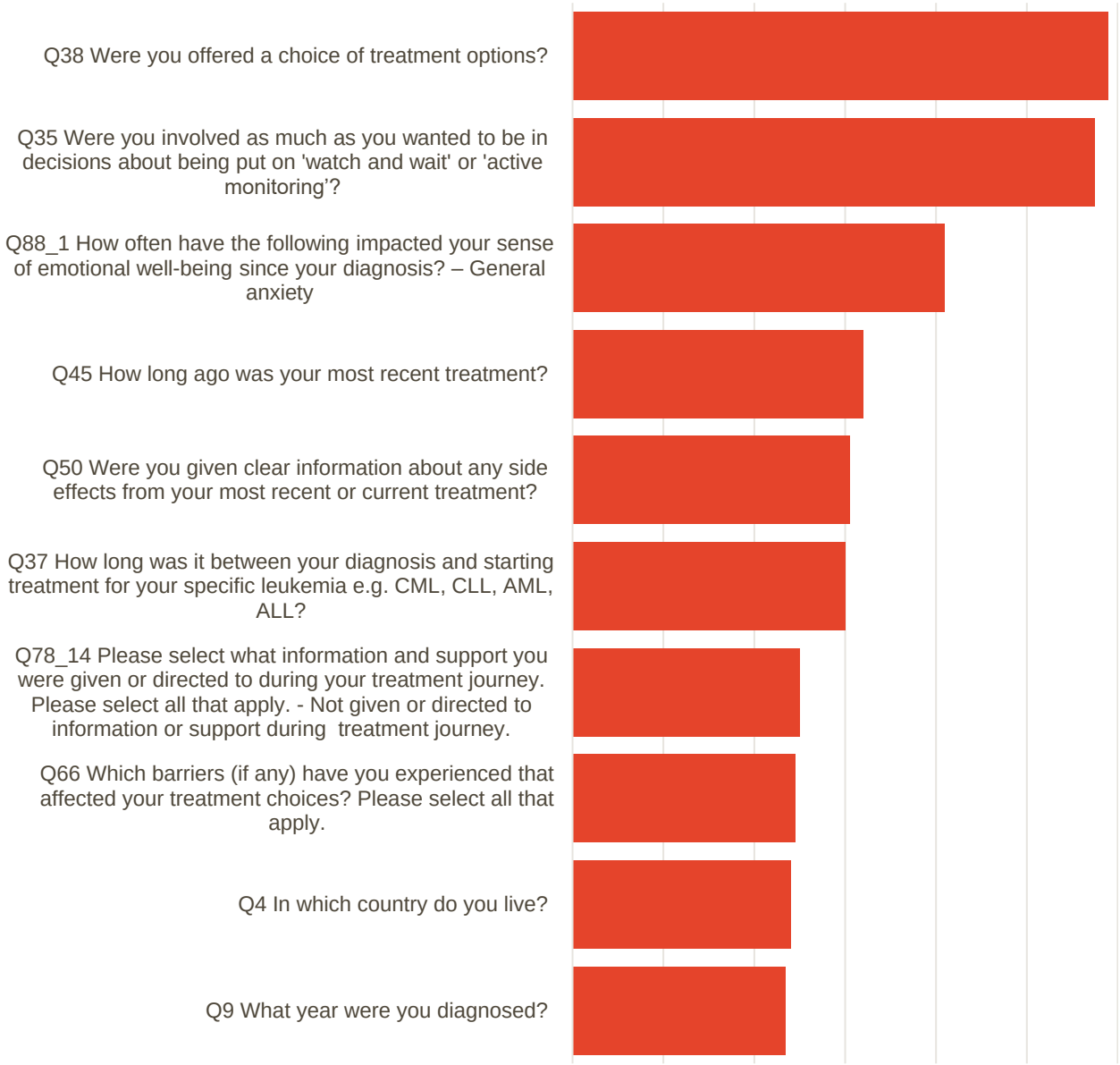
Given the complex nature of the outcome variable and a questionnaire feature data set it is perhaps not surprising we do not achieve a clinical decision sufficiency. This indicates there are additional unobserved effects on our outcome that necessitate further investigation to improve the model.

Nonetheless, our model is sufficient to highlight where gains might be had in improving people's stated feeling about their involvement in treatment decision-making. Examining the top five most important features (Figure 4

Figure 4) in the model's prediction in more detail:

- ***Q38 Were you offered a choice of treatment options?***
  - Not being offered a choice means more likely to select ***No*** for Q40 (i.e. not as involved as much as wanted to be in decisions: "not involved"), being offered a choice more likely to select ***Yes, definitely*** for Q40 (i.e. involved as much as wanted to be: "involved").
- ***Q35 Were you involved as much as you wanted to be in decisions about being put on 'watch and wait' or 'active monitoring'?***
  - Not involved in that decision means more likely to select ***No*** (not involved) for Q40, being involved more likely to select ***Yes, definitely*** (involved) and ***Yes, to some extent*** (involved) for Q40.
- ***Q88_1 General anxiety***
  - Selecting *Never* having general anxiety, more likely to select ***Yes, definitely*** (involved) for Q40 and less likely to select ***Yes, to some extent*** (moderately involved) for Q40. Selecting *Very often* anxious were more likely to select ***No*** (not involved) for Q40**.**
- ***Q45 How long ago was your most recent treatment?***
  - No obvious pattern.
- ***Q50 Were you given clear information about any side effects from your most recent or current treatment?***
  - Not being given clear information means more likely to select ***No*** (not involved) for Q40, being given clear information means more likely to select ***Yes, definitely*** (involved) for Q40.

Figure 4. The top 10 features that were most influential in our model's decisions. This is not just the number of times the variable is used in the models, but how much they improve the predictions. The relative size of the bars is the focus of interpretation with larger bars being more important.



Discussion and Recommendations

Our findings provide insight and understanding of the challenges faced by people with CLL. These results have several important implications for patient care and highlight areas where healthcare practitioners can improve patient outcomes.

Implications for people with CLL

Lack of knowledge of symptoms and delay in visiting HCPs

The proportion of respondents with CLL who were not aware that the symptoms they were experiencing could be due to leukemia is highly concerning. Respondents may be misattributing symptoms to other causes before seeking medical attention. Moreover, 1 in 5 respondents with symptoms delayed seeking medical attention for 1 year or more.

It is important that people seek medical attention at the earliest opportunity, to aid timely diagnosis and access to appropriate care to manage symptoms and disease progression.

Understanding of and engagement with the active monitoring strategy

The active monitoring strategy is a common strategy for CLL, yet the survey results indicate that respondents do not fully understand this strategy and may lack adequate support in managing symptoms. This uncertainty can lead to heightened stress and a sense of helplessness, which may negatively impact overall quality of life.

The low levels of patient confidence in recognising the signs of disease progression suggest that more needs to be done to educate people about what to monitor and when to seek further medical advice. People who are better informed and supported are more likely to be empowered, which can lead to a more proactive approach to managing their condition and thus, better health outcomes.

Impact of immunity awareness on quality of life

Understanding of immunity related information and its implications for health is striking. Respondents who did not receive clear information about what steps they can take to stay healthy and avoid infections were more likely to report worse quality of life. While respondents who received complete information about the protocols for CLL immunisations / vaccinations, including which they should receive and which they should avoid, were more likely to report better quality of life.

This underscores the importance of providing complete information, and ensuring it is clear, actionable, and tailored to the person’s needs. Well-informed people are likely better equipped to manage the psychological burden of CLL and their compromised immunity, and in turn have improved overall well-being.

Recommendations for clinicians and advocacy groups

Enhance symptom awareness

To reduce delays in people with CLL seeking medical attention for their symptoms, clinicians and advocacy groups could develop and implement targeted educational campaigns, with emphasis on primary care empowerment, to reduce clinical delays in referral escalation and diagnosis, and emphasis early symptom recognition and timely consultations to reduce delays in patients presenting.

Improve diagnostic clarity

Findings show that once respondents with CLL sought medical attention, two-thirds are diagnosed within 3 months. However, many respondents with CLL did not receive clear information about their diagnosis. Once a diagnosis is made, clinicians should ensure that the explanation given is clear and comprehensive, including information on understanding the nature of the disease, progression, potential compromised immunity and varied prognoses. There appear to be missed opportunities to direct people with CLL to additional information and support resources, including for mental and general wellbeing, buddying or befriending and financial information.

Enhance communication and support for active monitoring plans

Clinicians and advocacy groups should prioritise clear, consistent communication about the active monitoring strategy, ensuring that people fully understand what it entails and how to manage their symptoms and monitor progression. Provision of ongoing support, including regular check-ins and access to resources for symptom management is paramount, as this could help alleviate anxiety and improve quality of life.

Focus on person-centred immunity status education

Given the links between immunity awareness and quality of life, clinicians and advocacy groups could focus on delivering clear and practical guidance on managing immunity. This may include discussing the importance of vaccinations, providing personalised advice on how to stay healthy, and addressing people’s anxieties related to their immunity status.

Strengthen involvement in treatment decisions

Clinicians should actively involve people in treatment decisions, offering them choices and ensuring they understand the implications of each option. This may be achieved by empowering clinicians in communicating in this way with their patients and by empowering patients with clear information and learning opportunities to be more able and confident to become involved in their own healthcare discussions and decisions with their doctors. By doing so, people may be more likely to feel engaged and satisfied with their care, which can improve adherence to treatment plans and health outcomes.

Conclusion

The insights gained from the 2023 Global Leukemia Experience Survey highlight several areas where healthcare for people with CLL can be improved. By enhancing communication, providing targeted education, and involving people more deeply in their care decisions, clinicians could improve the quality of life, experiences, and clinical outcomes for people with CLL.

Appendix A1: Respondent characteristics

This appendix provides response counts for respondents who selected CLL as their leukemia type in the Global Leukemia Patient Experience Survey 2023, for country, age, and gender.

Country	Number of responses
Australia	12
Belgium	6
Brazil	32
Canada	69
Chile	2
China	29
Czechia (Czech Republic)	24
Denmark	52
Finland	5
France	3
Germany	10
Greece	7
India	1
Ireland	27
Israel	68
Italy	35
Kenya	1
Monaco	2
Netherlands	50
New Zealand	18
Norway	5
Peru	1
Portugal	1
Republic of Korea	2
South Africa	1
Spain	3
Sweden	1
Tanzania	1
United Kingdom	272
United States of America	105
No response provided	1
Total	**846**

Gender	Number of responses
Female	454
Male	357
Prefer not to say	20
Prefer to self-describe (please specify)	15
Total	**846**

Age	Number of responses
26 to 35	3
36 to 45	24
46 to 55	110
56 to 65	270
66 to 75	316
76 and over	123
Total	**846**

Appendix A2: Methods

Using data from the 2023 Global Leukemia Experience Survey regarding the ‘patient’ focussed questions and derived variables thereof.

Ethics

There is no intervention. The survey was online only and anonymous at the point of data collection. There is no personal identifiable data (PID) recorded as part of the survey and data collection is in accordance with data governance principles, as advised by Economit Ltd. data protection consultancy (company registration number: 07613723).

Statistical methods overview

Methods RQ1

We reported the percentage of respondents selecting each response category for each of a selection of questions (agreed in advance with CLLAN members) relating to the active monitoring pathway, the broader pathway, and a question on information provision (Box 1). This approach is known as univariate descriptive statistics. We restricted the analysis to people identifying themselves as having CLL.

Methods RQ1.1

For RQ1.1, we returned to the full data set, and cross-tabulated the leukaemia type with variables in RQ1, excluding the active monitoring questions that were exclusive to CLL (Box 1). We used statistical tests to examine how variables affect each other (called bivariate tests), specifically we use chi-squared tests of associations to examine these associations. We made an adjustment to the analyses to account for multiple tests on the same outcome variable, which makes us less likely to wrongly accept a conclusion of an association if that association does not exist.

Methods RQ2

For CLL respondents only, we cross-tabulated and analysed associations between questions about knowledge of immunity status and receipt of related information that were asked of CLL respondents only, with the HMPRP-A grouped score (

Box 2).

Methods RQ3

We created a model to investigate factors that contribute towards respondents stating that they were involved in treatment decisions as much as they wanted (Box 3). We used an approach that allows the computer to recognise and classify patterns (machine learning) by creating a series of simple decision rules in a method called decision tree modelling. It works by splitting data into different paths to reach a conclusion, in a similar way to a flow chart. We used an advanced version, called Gradient Boosted Decision Trees (GBDT), which improves accuracy by combining multiple decision trees in a way that mitigates errors and fine-tunes predictions. We restricted the model to people with CLL.

Box 1. RQ1 variables.

RQ1 Active monitoring questions

- Q29 Have you ever been placed on a 'watch and wait' or 'active monitoring' monitoring plan?
- Q30 When you were placed on 'watch and wait' or 'active monitoring' for the FIRST time, was it explained to you in a way you could understand?
- Q31 What best describes your current 'watch and wait' or 'active monitoring' situation?
- Q32 What is the main way your monitoring visits take place? A monitoring visit is a regular check up with a member of your medical team to check if your leukemia needs treatment.
- Q33 While on 'watch and wait' or 'active monitoring', who helps you with pain or symptoms related to your diagnosis? Please select all that apply.
- Q34 How did you feel about being put on 'watch and wait' or 'active monitoring'? If you have been on 'watch and wait' or 'active monitoring' more than once, please think back to the first time.
- Q35 Were you involved as much as you wanted to be in decisions about being put on 'watch and wait' or 'active monitoring'?
- Q36 Are you confident you can recognise the signs and symptoms of CLL progression?

RQ1 Information provision question

- Q78 Please select what information and support you were given or directed to during your treatment journey. Please select all that apply.

RQ1 & RQ1.1 Diagnostic pathway questions

- Q10 Was your leukemia detected as a result one of the following? Please select one answer.
- Q11 What symptoms did you encounter before your diagnosis? Please select all that apply.
- Q12 Did you know that the health problems you were experiencing could be symptoms of leukemia?
- Q13 How long was it from the time you first experienced symptoms until you first spoke to a healthcare professional about them? By this we mean any medical professional you saw, this could be a GP / family doctor, hospital doctor, etc.
- Q14 How many times did you see a healthcare professional about your symptoms before you were diagnosed? By this we mean any medical professional you saw, this could be a GP / family doctor, hospital doctor, etc.

- Q15 How long was it from the time you first saw a healthcare professional until you were diagnosed with your condition? By this we mean any medical professional you saw, this could be a GP / family doctor, hospital doctor, etc.
- Q26 Was your diagnosis explained to you in a way you could understand?

RQ1.1 Outcome variable:

- Q01_leukemia_detail - derived variable; categorised as ALL, AML, CLL & CML.

Box 2. RQ2 variables.

RQ2 Immunisation status questions

- Q72 Do you know your immunity status?
- ROUTED Q73 Were you given clear information about what steps you can take to stay healthy and avoid infections?
- Q74 Has a health professional spoken to you about the protocols for CLL immunisations / vaccinations, including which you should receive and which you should avoid?

RQ2 HMPRO A (Quality of life, QoL, measure)

- 24 items and a score range of 0-48:
- Used a derived grouped variable:
    - Score 0 to 8 (low impact on QoL); 9 to 18; 19 to 28; 29 to 48 (high impact).

Box 3. RQ3 variables.

RQ3 outcome variable

- Q40 Were you involved as much as you wanted to be in decisions about your treatment?

RQ3 candidate predictor variables

- All other variables in the data set, excluding HMPRO questions.

Statistical detail

Software and modelling approach

We invoked the analyses using Python 3.11, via Spyder 5.5.1.

We generated cross tabulations of data and ran chi-squared tests of association, with a null hypothesis of no association in each case. We applied a Bonferroni correction to mitigate multiple testing of the same outcome (with an adjusted test threshold $\alpha = 0.05$ / number of tests).

Specifically, for GBDT, with a multi-category outcome and multiple categorical predictors invoked the model via the *LightGBM* package, which is suited to such structures and has computational efficiency.

GBDT modelling is an ensemble technique that creates a combined classification of the data set that can be used to explore the importance of independent variables ("features") in predicting an outcome. The underlying models in the ensemble are decision trees that split the data set recursively into "leaves" that are sub-groups of the data set, with the idea that the final leaves contain cases that are as similar as possible to each other, by minimising a "log loss function".

Whilst the method can deal with complex data, it can provide solutions that are over-fit to the data, to attempt to mitigate this we trained a machine learning model by splitting the data in to testing and training sections, and optimised "hyper-parameters" (parameters about how the method is applied, such as gradient method, maximum number of leaves, feature fraction, minimum tree depth). Hyperparameter optimisation was conducted using the *Hyperopt* package, and a defined training space, to search for the most suitable ("best") parameters. We applied these optimised parameters to train the model (including an early stopping rule, when the function did not improve for 100 iterations), evaluated the model performance using diagnostic plots of Area Under the Curve (AUC), a "confusion matrix" of classification pattern, performance metrics of accuracy and F1-score, and a "gain" importance graph to find the most influential features in the model's predictions. Gain is concerned with how much better the model performance as a result of including a feature.

Missing data approach

Frequency tables report missing responses, which will include people who dropped out of the survey and item non-responders.

Bivariate testing drops cases where the respondents answered, "don't know / cannot remember", and on a question-by-question basis respondents who provide an answer to both variables in the cross tabulation.

Decision tree analyses are robust to missing data and do not need to be adjusted unless missingness is very high, e.g. 40% or more, at which point there would be other concerns around missing data and suitability for analysis.

Diagnostic metrics

These metrics assess the performance of classification models. In practice, the choice of which metric to prioritise depends on the specific context of the problem. For instance, in a spam email detection system, high precision is desirable to ensure that legitimate emails are not classified as spam (minimising false positives). However, in a medical diagnosis scenario, high recall may be preferred to ensure that no real positive cases are missed, even if it means more erroneous prediction of a positive.

Precision

A high precision score indicates that the model's positive predictions are mostly accurate and have fewer false positives (FP).

Formula: Precision = TP / (TP + FP)

Recall (Sensitivity or True Positive Rate)

A high recall score indicates that the model can successfully capture most of the positive instances and has fewer false negatives (FN).

Formula: Recall = TP / (TP + FN)

F1-score

The F1-score is the harmonic mean of precision and recall. It provides a balance between precision and recall and is useful when you want to consider both false positives and false negatives. This ranges from 1 (best) to 0 (worst).

Formula: F1-score = (2 * Precision * Recall) / (Precision + Recall)

Accuracy

An overall accuracy can be calculated as being the number of cases correctly classified.

Formula: Accuracy = (TP +TN) / (TP + FP + TN +FP)

ROC and AUC

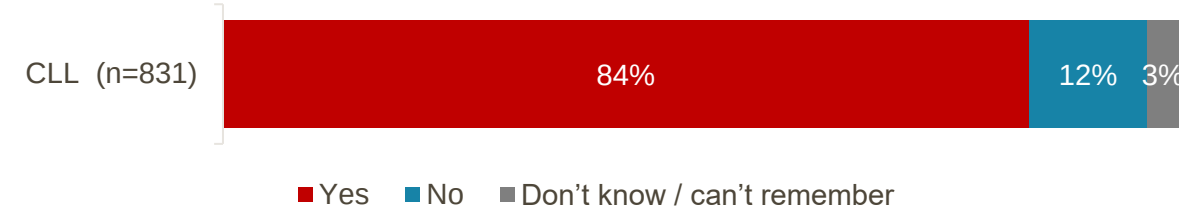
We visualised true positive and false positive rates across different discrimination thresholds as Receiver Operator Characteristic (ROC) curves and determine the area under the curve (AUC). AUC ranges 0 to 1, with 1 indicating a perfect classification, 0.5 as no better than chance, and less than 0.5 as an “inverted” model – incorrectly classifying the data more often than chance.

Formula: AUC = ½ * ([TP / (TP + FN)] + [TN / (TN + FP)])

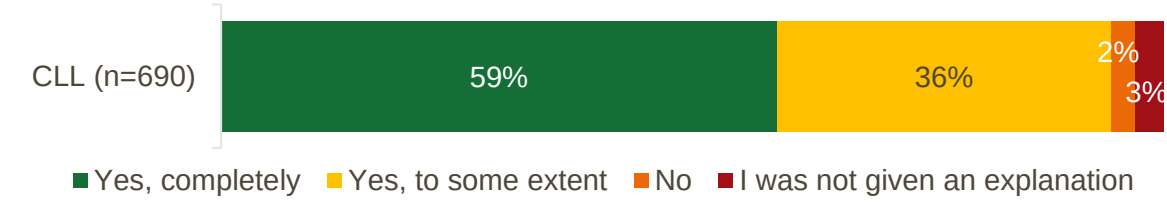
Appendix A3: RQ1

Active monitoring selected questions

Q29 Have you ever been placed on a 'watch and wait' or 'active monitoring' monitoring plan?

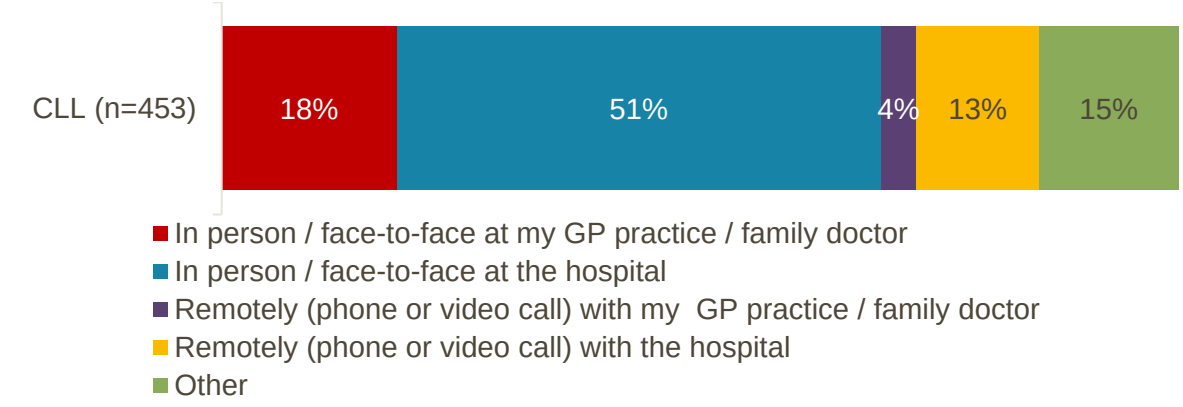


Q30 When you were placed on 'watch and wait' or 'active monitoring' for the FIRST time, was it explained to you in a way you could understand?

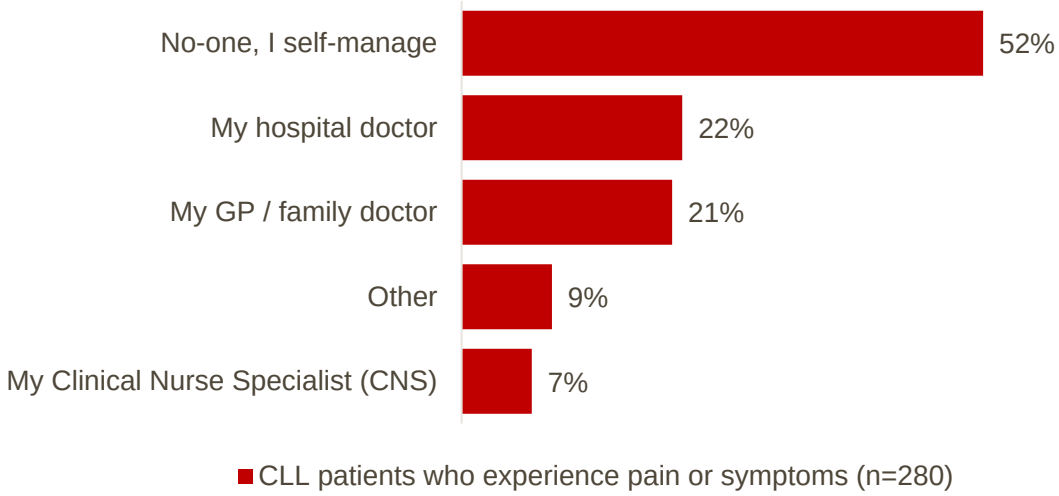


Q31 What best describes your current 'watch and wait' or 'active monitoring' situation?

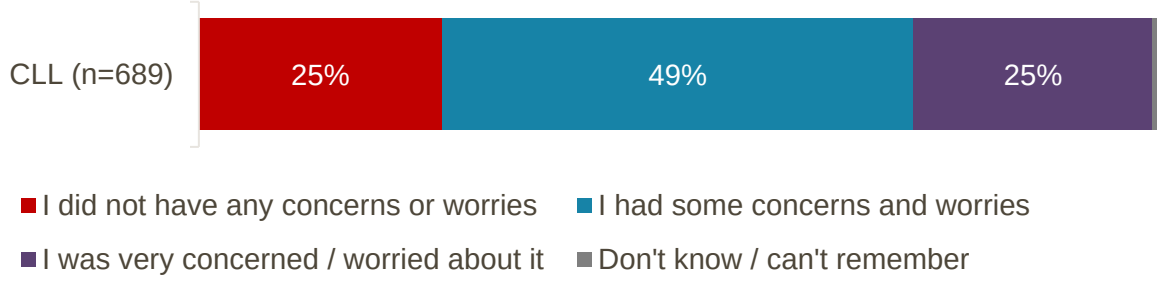
Q32 What is the main way your monitoring visits take place? A monitoring visit is a regular check up with a member of your medical team to check if your leukemia needs treatment.



Q33 While on 'watch and wait' or 'active monitoring', who helps you with pain or symptoms related to your diagnosis? Please select all that apply.



Q34 How did you feel about being put on 'watch and wait' or 'active monitoring'? If you have been on 'watch and wait' or 'active monitoring' more than once, please think back to the first time.



Q35 Were you involved as much as you wanted to be in decisions about being put on 'watch and wait' or 'active monitoring'?

Q36 Are you confident you can recognise the signs and symptoms of CLL progression?

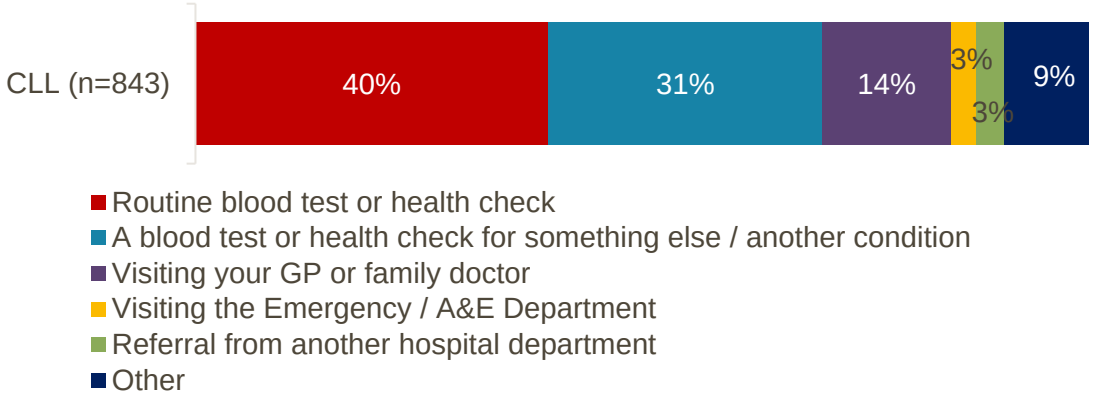
Information provision selected question

Q78 Please select what information and support you were given or directed to during your treatment journey. Please select all that apply.

Base (n=)	CLL 770
Written information / booklets / leaflets	44%
Leukemia charities / associations / organizations	34%
Side effects and risks of treatment	31%
Patient support groups	29%
Online forums	21%
Clinical trials	19%
Access to a Clinical Nurse Specialist (CNS)	19%
Mental wellbeing (emotional) including referral to psychological support, counselling or psychotherapy	18%
General wellbeing (including spiritual wellbeing)	14%
Patient rights	12%
Financial information	7%
Buddying or befriending	6%
End of life care	2%
I was not given or directed to information or support	**24%**

Diagnostic pathway selected questions

Q10 Was your leukemia detected as a result one of the following? Please select one answer.

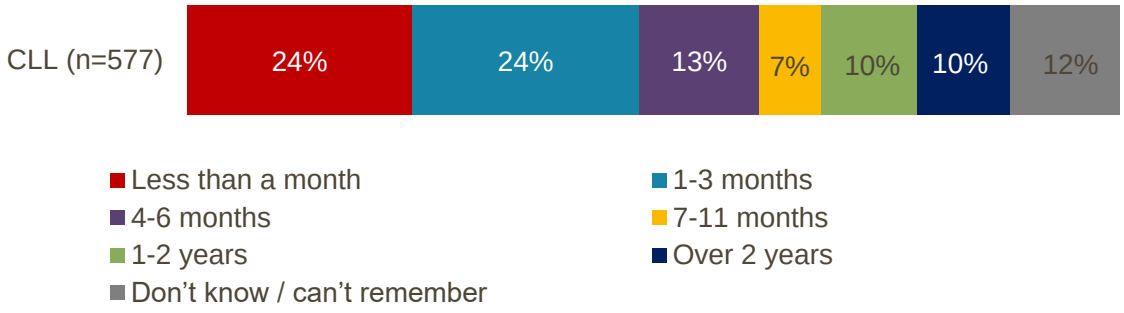


Q11 What symptoms did you encounter before your diagnosis? Please select all that apply.

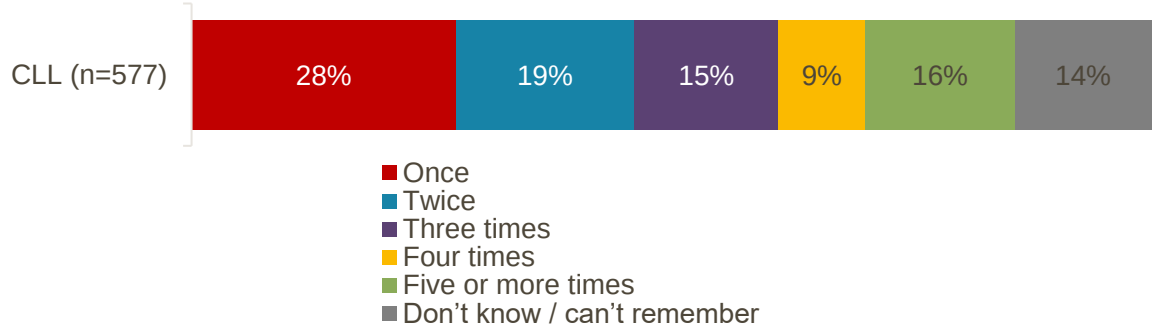
	CLL
Base (n=)	**841**
No symptoms	31%
Fatigue	40%
Swollen lymph nodes	27%
Fever / night sweats	19%
Unusually high or low red / white / platelet blood count	15%
Frequent and repeated infections	13%
Pain in bones / joints	12%
Shortness of breath	9%
Sleeping problems	9%
Bruising	8%
Muscle pain	8%
High blood pressure (Hypertension)	8%
Unexplained weight loss	7%
Back pain	7%
Itchy skin	7%
Headaches	6%
Swollen stomach or abdominal discomfort	6%
Other symptoms	6%
Tingling or numbness in extremities	6%
Memory loss / loss of concentration	4%
Loss of appetite	3%
Palpitations / heart irregularities	3%
Bleeding	2%
Nausea or vomiting	2%
Changes to menstrual cycle	1%
Broken blood vessel in eye (Subconjunctival hemorrhage)	1%

Q12 Did you know that the health problems you were experiencing could be symptoms of leukemia?

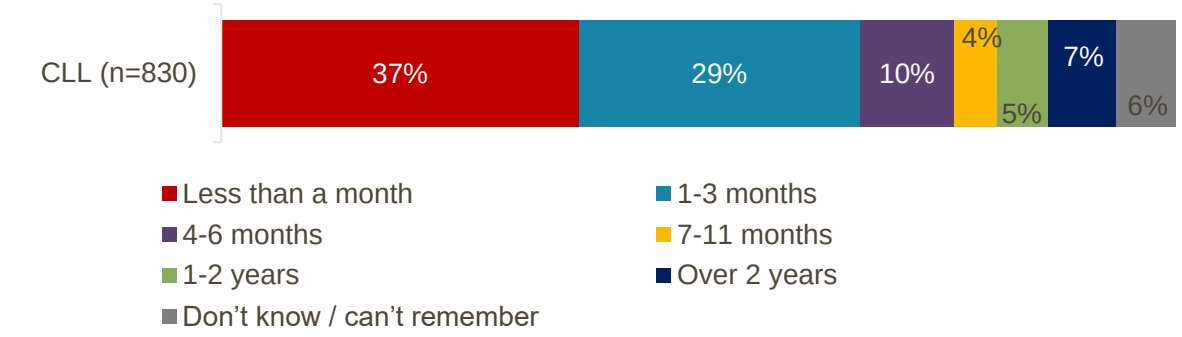
Q13 How long was it from the time you first experienced symptoms until you first spoke to a healthcare professional about them? By this we mean any medical professional you saw, this could be a GP / family doctor, hospital doctor, etc.



Q14 How many times did you see a healthcare professional about your symptoms before you were diagnosed? By this we mean any medical professional you saw, this could be a GP / family doctor, hospital doctor, etc.



Q15 How long was it from the time you first saw a healthcare professional until you were diagnosed with your condition? By this we mean any medical professional you saw, this could be a GP / family doctor, hospital doctor, etc.



Q26 Was your diagnosis explained to you in a way you could understand?

RQ1.1 How does this contrast with other types of leukaemia?

Chi-square results for diagnostic pathway questions against classified leukemia type (Q1_leukemia_detailed), *df* = degrees of freedom, adjusted alpha is the test threshold under the Bonferroni correction for multiple testing, *Adj, Sig*., describes whether the p-value is less than the threshold and thus if a statistically significant association has been detected between the variables.

All indicate an association between the variables and the outcome (leukemia type), except for knowing the health problems could be leukaemia.

	X^2 statistic	df	P-value	Adjusted alpha	Adj. Sig.
Q10_leukemia_detection	208.35	25	7.52E-31	7.14E-03	Significant
Q11_symptoms_binary	183.93	5	7.75E-38	7.14E-03	Significant
Q12_knew_health_problems_could_be_leukemia	14.62	5	1.21E-02	7.14E-03	Not Significant
Q13_from_first_symptoms_how_long_to_speak_to_healthcare_professional	330.14	25	5.12E-55	7.14E-03	Significant
Q14_times_HCP_before_diagnosed	66.32	20	7.18E-07	7.14E-03	Significant
Q15_how_long_to_diagnose	174.39	25	2.36E-24	7.14E-03	Significant
Q26_diagnosis_explained_understandably	29.89	10	8.94E-04	7.14E-03	Significant

Appendix A4: RQ2

Chi-square results for knowledge of immunity status questions against HMPRO_A grouped score.

	X^2 statistic	df	P-value	Adjusted alpha	Adj. Sig.
HMPRO-A total grouped vs					
Q72_know_immunity_status	12.05	6	6.08E-02	1.67E-02	Not Significant
Q73_clear_information_to_avoid_infection	34.55	6	5.28E-06	1.67E-02	Significant
Q74_hcp_spoken_about_CLL_immunisation_protocols	30.37	6	3.35E-05	1.67E-02	Significant

Appendix A5: RQ3

Overall accuracy is 68% with F1-score of 0.646 and AUC of 0.76, indicating moderate performance, but suggestive of unobserved influences. The poorest performance in predicting the no group due to the recall ability being low.

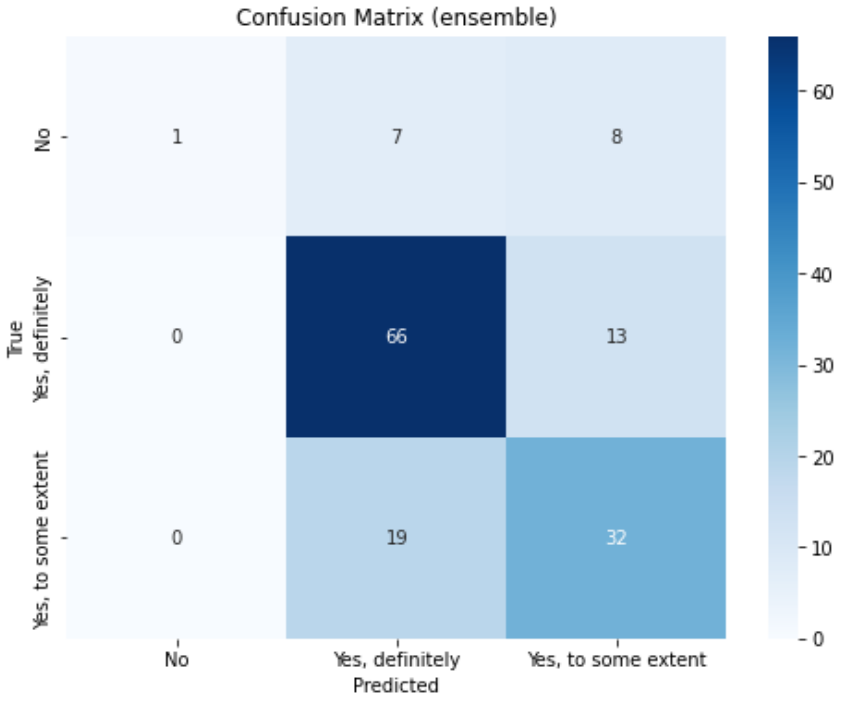
Diagnostic metrics

Ideally all these scores would be 1.

<table>
<tr><td colspan="3">Model
Overall Accuracy: 0.678
Weighted Average Precision: 0.709
Weighted Average Recall: 0.678
Weighted Average F1-score: 0.646</td></tr>
<tr><td>Category: No
Precision: 1
Recall: 0.063
F1-score: 0.118</td><td>Category: Yes, definitely
Precision: 0.717
Recall: 0.835
F1-score: 0.772</td><td>Category: Yes, to some extent
Precision: 0.604
Recall: 0.627
F1-score: 0.615</td></tr>
</table>

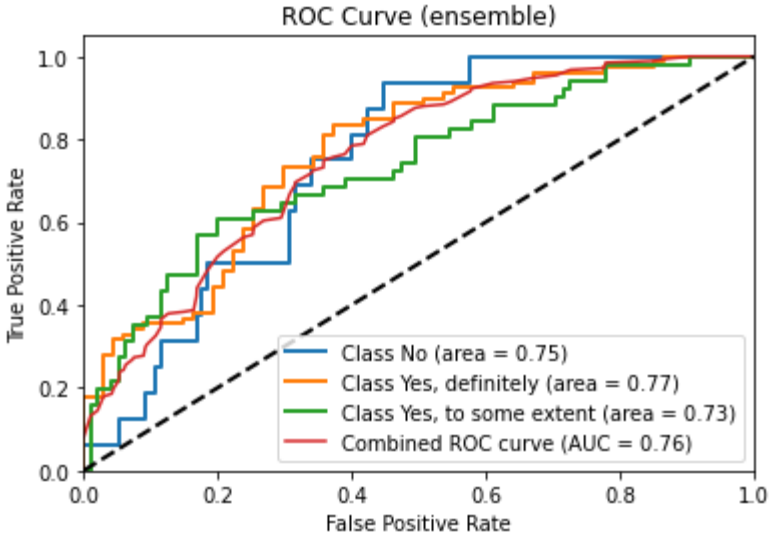
Confusion matrix

Comparing the predicted outcome category with the actual outcome category, using the test data fraction. Ideally everything would lie on the diagonal top left to bottom right.



Area Under Curve

Ideally the curve for the categories of outcome (class) and model overall would be smooth and close to the vertical axis immediately, swiftly rising towards one and staying that way as you move along the horizontal axis. In the ideal case described, the Area Under the Curve (AUC) for the plot would be 1. An AUC of 0.5 indicates a model no better than chance. Our model has an AUC of 0.76 overall.



Analytical References

Gradient Boosted Decision Modelling and LightGBM

Friedman JH. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **29**, 1189–1232. https://doi.org/10.1214/aos/1013203451 (2001).

Ke GL, Meng Q, Finley T, Wang TF, Chen W, Ma WD, Ye QW, Liu TY. LightGBM: a highly efficient gradient boosting decision tree. Advances in Neural Information Processing Systems 30 (Nips 2017) 2017, 30. https://papers.nips.cc/paper_files/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html

LightGBM documentation for python: https://lightgbm.readthedocs.io/en/latest/Python-Intro.html [accessed August 2024].

Sagi O, Rokach L. Ensemble Learning: A Survey. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2018**, *8*, e1249

Bergstra J, Yamins D, Cox DD, “Hyperopt: A python library for optimizing the hyperparameters of machine learning algorithms,” in Proceedings of the 12th Python in science conference, vol.13, p20, Citeseer, 2013 **DOI** 10.25080/Majora-8b375195-003

CLL Advocates Network
Leukemia Patient Advocates Foundation
P.O.B. 453
CH-3000 Bern 7
Switzerland
info@clladvocates.net

Picker Institute Europe
Suite 6, Fountain House,
1200 Parkway Court,
John Smith Drive,
Oxford OX4 2JY

Tel: +44 (0) 1865 208100

info@pickereurope.ac.uk
picker.org

Charity registered in England and Wales: 1081688
Charity registered in Scotland: SC045048
Company limited by guarantee registered in England and Wales: 3908160