Correlation between CD180 Expression and TP53 Deletion in Chronic Lymphocytic Leukaemia using a Computational Model

By

Tanin Tahvili, Lewis Clark, Ritu Dogra and Jessica Monteiro

In collaboration with
Dr Nino Porakishvili, Dr Emanuela Volpi,
and Dr Alexander Bolotov

Abstract:

Chronic lymphocytic leukaemia (CLL) is the most prevalent adult leukaemia in western countries, with its occurrence increasing with age. B-CLL mainly affects elderly population with median age of 65 at diagnosis. This leukaemia is characterised by the clonal expansion accumulation of malignant CD5+ B cells in blood, peripheral lymphoid organs and bone marrow. It is proposed that onset and progression of CLL is driven by unknown (auto) antigens via Toll-like receptors (TLRs) and/or environmental signals through B Cell Receptors (BCRs). CD180, which is a member of the TLR family is expressed by about 60% of CLL samples. Once ligated with monoclonal antibodies against CD180, half of these CD180+ CLL cells responded by enhancing the activation (phosphorylation) of protein kinases that are involved in BCR signalling leading to activation of pro-survival BTK/PI3/AKT pathway. Studies have also demonstrated that pre-ligation of CD180 can rewire the BCR(IgM)-mediated signalling from the pro-survival BTK-PI3K/AKT pathway towards the pro-apoptotic p38MAPK pathway. This may indicate that patients who co-express both IgM and CD180 have a better prognosis.

Certain chromosomal aberrations have shown to be strong prognostic markers for CLL, particularly regarding response to treatment and survival. TP53 deletion is associated with the worst prognosis of the five most common chromosomal aberrations. Here we investigated the association between CD180+ and TP53 deletion, in order to train a computer algorithm based on experimental data collected to create a model which predicts the outcome of TP53 deletion of patients with CLL. Thus, creating a new and rapid method for therapeutic profiling aiding strategies for prognostic measures in patients with CLL.

Introduction:

B Chronic lymphocytic leukaemia is the most prevalent adult leukaemia in western nations, and its occurrence increases with age. B-CLL mainly affects elderly population with median age of 65 at diagnosis (Sun, Wang and Sun, 2017). This leukaemia is characterised by the clonal expansion of build-up of malignant CD5+ B cells in blood, peripheral lymphoid organs and bone marrow (Durrieu *et al.*, 2011). It is proposed that onset and progression of CLL is driven by unknown (auto) antigens via Toll-like receptors (TLRs) and/or environmental signals through B Cell Receptors (BCRs) (Porakishvili *et al.*, 2011). TLRs are a class of proteins present on antigen-presenting cells which recognise structural conserved molecules derived from pathogens (Van Lent *et al.*, 2010). Previously it was shown that the orphan receptor, CD180, which is a member of the TLR family is expressed by about 60% of CLL samples. Once ligated with monoclonal antibodies against CD180, half of these CD180+ CLL cells responded by enhancing the activation (phosphorylation) of protein kinases that are involved in BCR signalling leading to activation of pro-survival BTK/PI3/AKT pathway. The

same study demonstrated that pre-ligation of CD180 can rewire the BCR(IgM)-mediated signalling from the pro-survival BTK-PI3K/AKT pathway towards the pro-apoptotic p38MAPK pathway. This may indicate that patients who co-express both IgM and CD180 have a better prognosis (Porakishvili *et al.*, 2011).

Certain chromosomal aberrations have shown to be strong prognostic markers for CLL, particularly regarding response to treatment and survival. There are five principal prognostic genetic aberrations in CLL which are deletions 17p, 11q, 13q, trisomy 12 and normal karyotype (summarised in table 1) (Hallek, 2017). p17 deletion and TP53 mutations remain the most significant adverse prognostic features (Canta *et al.*, 2016). In this project the first aim was to investigate the association between P53 deletion and expression of CD180 (and other receptors of established prognostic significance). The second aim was to use the above information to train an algorithm using computer science to help improve the prediction of TP53 deletion thus contributing to the prediction of disease outcome for a patient with CLL.

Table 1: Established genetic prognostic markers

Genetic aberrations		
Abnormality	Incidence	Correlation with CLL progression and survival
13q14.3 l(encompassing the DLEU1 gene)	40-60%	The most common chromosomal aberration in CLL. This mutation is associated with better prognosis and longer overall survival. High percentage (>80%) of cells with 13q deletion leads to shorter survival but low percentage (<80%) is correlated with longer survival (Calin *et al.*,2002).
Trisomy 12	10-20%	Associated with early CLL progression. Previous studies have reported that although this gene abnormality is correlated with shorter progression-free survival overall survival is favourable (Seiffert *et al.*,2011).
11q23 deletion	10-20%	Correlated with poor response to treatment, fast disease progression and short overall survival and involves many genes one of which is the ataxia teleagiectasia mutated (ATM) tumour suppressor gene. ATM has a role in repair of the double stranded DNA break, cell cycle checkpoint control, regulation of tumour suppressor p53 protein and protection of the telomere regions (Hallek, 2017).
17p deletion	3-8%	Deletion always encompasses the TP53 locus but often includes majority of the shot arm of chromosome 17. Strongly associated with aggressive clinical course. Patients usually have lack of response to therapy, short response duration and short overall survival Hallek, 2017.
6q deletion	6%	Generally classed as intermediate-risk feature. It is correlated with greater lymphocytosis and rates of Cd38 positivity (Seiffert *et al.*,2011).

Research plan and experimental approach:

As stated in figure 1, all blood samples were collected from the University College London Hospital under a material transfer agreement (NREC: 08/H0714/6) with University of Westminster. The samples were cultured for FISH, and the Peripheral Blood Mononuclear Cells (PBMCs) were isolated for phenotyping experiment. The purpose of FISH technique is to identify the p53 deletion in CLL patients. Using immunophenotyping technique, the expression of receptors with prognostic significance were measured. Figure 2 shows a GANTT chart, demonstrating how the project was scheduled.
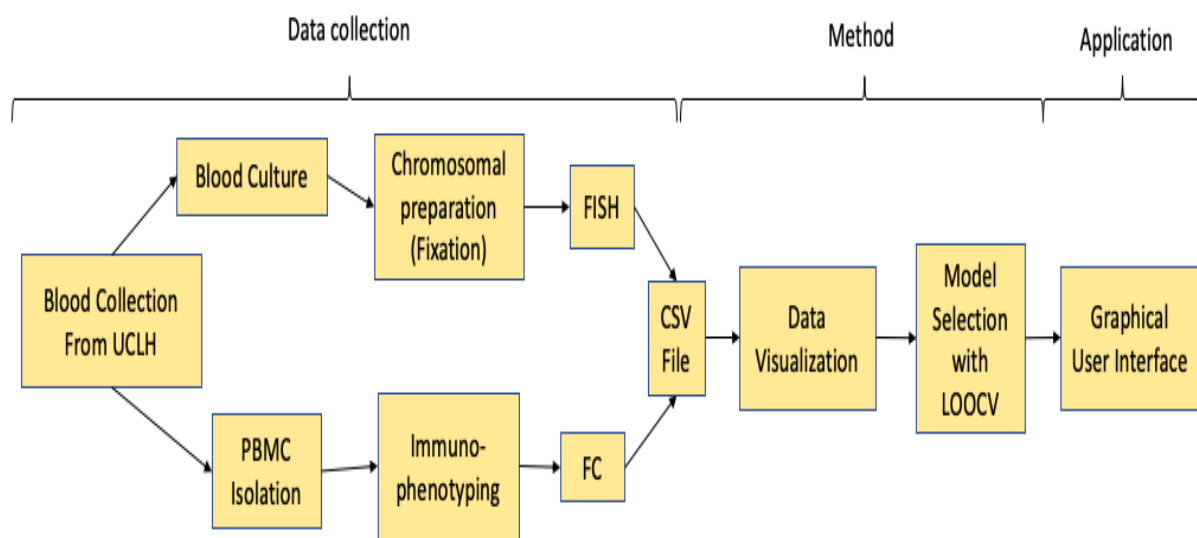

Figure 1: Flowchart demonstrating the sequence of steps of data collection and analysis. Once the sample is collected from UCLH, an aliquot of it is cultured and fixed for FISH analysis. The rest of the sample is processed (their PBMCs are isolated by Histopaque-1077 (Sigma-Aldrich, Dorset, UK)) and prepared for the process of flow cytometric (FC) immunophenotyping. The data produced by FISH and FC were combined into one csv file for statistical analysis in Python. Different train/test splits of the csv data were produced by Leave-One-Out cross validation. The splits were used to train and test the model, and for each one, the feature selection took place. After choosing the best model from CV scores, it was deployed with a graphical user interface.
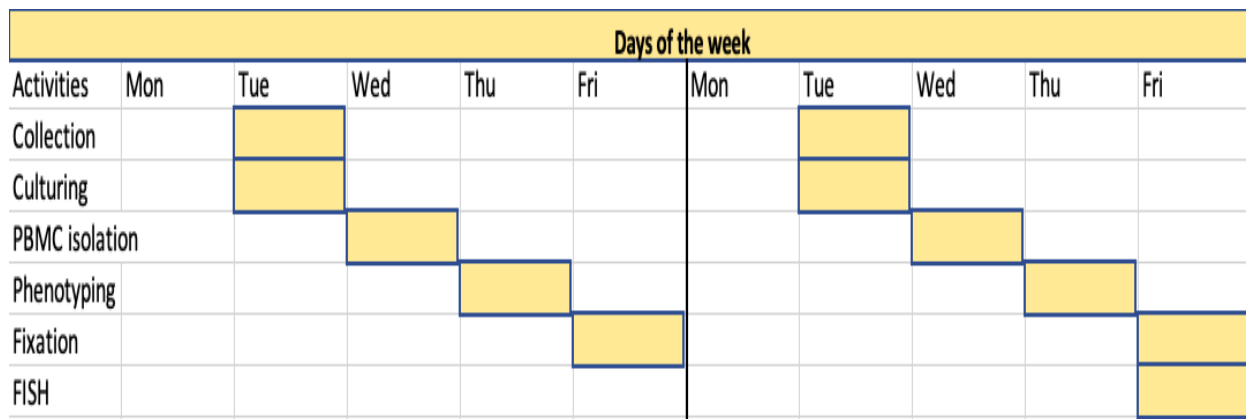
Figure 2: GANTT chart demonstrating the activities for the two-week period. Every Tuesday evening, the samples are collected from the hospital. The blood cultures are prepared on the same day. The cultures are incubated for 72 hours. PBMC takes place on the day after (Wednesdays). PBMCS are then stored in a freezer until the next day for phenotyping. On Friday fixation (chromosome preparation) can take place. The fixed cells can be analysed straight away, or they could be kept at -20 °C for later analysis.

Methods and results:

Patients:

Blood samples were obtained from CLL patients with informed consent following ethical approval from University College London Hospital (UCLH) (08/H0714/6). Once collected from the hospital, for blood culturing purposes an aliquot of each sample was cultured in karyotyping medium (thermo-Fisher) (at 37 °C, in 5% CO_2) for Fixation to be done 72 hours later.

Fixation:

First, Hypotonic solution-0.75: KaryoMAX Potassium Chloride (Thermo-Fisher) was added to the blood cultures and cells were fixed with methanol acetic acid (both Sigma-Aldrich). Next, a drop of the suspension prepared was placed on a microscope slide. Chromosomes were then denatured in 70% Fomamide-9037/2xSS at 72°C. Simultaneously, the probes were denatured at high temperature. Chromosomes are then treated with ice cold ethanol in series (70%, 90% and 100%) followed by air drying. 10 µl of probes are applied to the slides (chromosomes) and incubated at 37 °C for 24 hours. The slides were washed in 2XSSC + Tween20 (at RT), then 0.4SSC+ Tween20 (at 72 °C) and finally followed by 2xSSC + Tween (at RT) (Sigma-Aldrich).

PBMC isolation:

The blood samples were first homogenised using PBS, then the PBMCs were isolated using Histopaque-1077 (all Sigma-Aldrich, UK). These cells were then washed twice with PBS, counted and stored in freezing medium at -80 °C.

Phenotyping:

The previously isolated PBMCs were thawed, counted and distributed in micro-well plates. The antibodies used were anti-IgG1 (used as the negative isotype control), anti-CD180, anti-IgM (BD Bioscience Oxford, UK), anti-CD38, anti-CD79b, andtiCD86, anti-CD40 (Fitzgerald, North Acton, MA) and anti-IgD (Sigma-Aldrich). These antibodies were added to their corresponding previously labelled well. In the next step the cells were incubated with Fluorescein isothiocyanate (FITC) to label the antibodies with fluorescein so they could be viewed by the Flow cytometry (FC). Then mouse serum was added to eliminate the unwanted background produced by extra FITC. Finally, the cells were incubated with conjugated CD19 to solely label the B cells.

Statistical tests:

To analyse significance in association between CD180 expression and p53 deletion, T test should be used. And in future in order to check whether the correlation between the CD180 expression and p53 deletion is independent of other confounding factors such as age, sex and mutational status of CLL, multivariate analysis can be done.

The results obtained demonstrated a positive correlation between the expression of CD180 and P53 deletion. The blood smears of patients who had the deletion also showed presence of micronuclei. The micronuclei are usually a sign of genotoxic events and chromosomal instability. They are commonly observed in cancerous cells and could indicate DNA damage events that increase the chance of degenerative diseases (Geyer and Subramaniyam, 2014). Figure 3 shows the FISH image of one patient who expressed the deletion. Figure 4 illustrates a microscopic image of the peripheral blood smear of the same patient showing the presence of micronuclei.
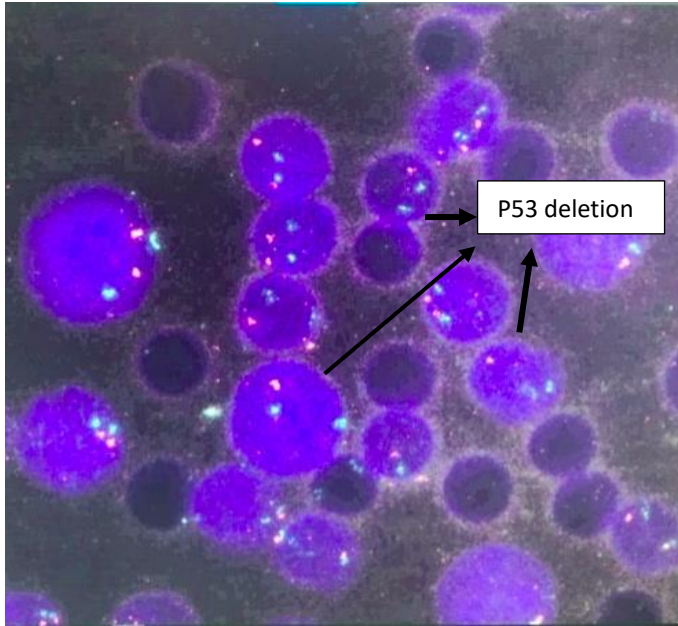
Figure 3: Fluorescence in *situ* hybridisation images for patient 1.

The first probe that emits green colour under the microscope serves as a control and hybridises with each of the chromosomes inside the nuclei of the cells. Each cell has two chromosomes and therefore presence of two green dots indicates that the hybridisation has worked. The second probe binds to the target sequence which in this case is P53 locus on the chromosomes. When the target location is present on chromosomes, the probe will hybridise and fluoresce with a red colour. This means that absence of red dots in the nuclei indicate p53 deletion. The cells that have the deletion are indicated by arrows.

Figure 4: blood smear for the same patient demonstrating micronuclei.

The presence of micronuclei confirms the genetic instability in the patient. Micronuclei are made when a chromosome or a fragment of it is not properly incorporated into one of the daughter nuclei during mitosis (Geyer and Subramaniyam, 2014).

Tools for developing the model:

The computational model was developed in Python, which was ideal for this project because it provided the necessary tools to build the prediction model. The scikit-learn library was used for model selection, feature selection and displaying the test accuracy scores. Numpy and Pandas were used to store the csv data locally and visualise it. Additionally, Tkinter (a Python GUI) was used to create the user interface.

Visualising the patient dataset:

The dataset contained 8 patients with the label column P53 deletion "POSITIVE/NEGATIVE" and four genetic markers: "IgM%, IgD%, CD180% and CD38%" expressed as percentages. In

theory, if any of these genetic markers highly correlate with the outcome (P53 deletion), they could be good predictors for the model.
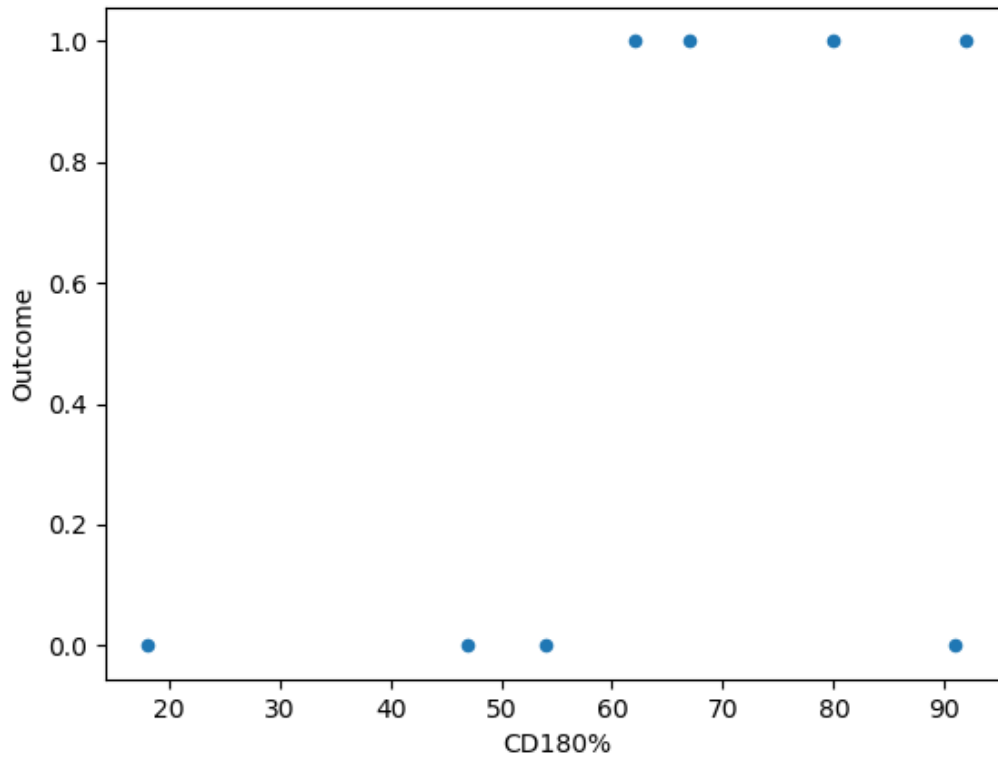

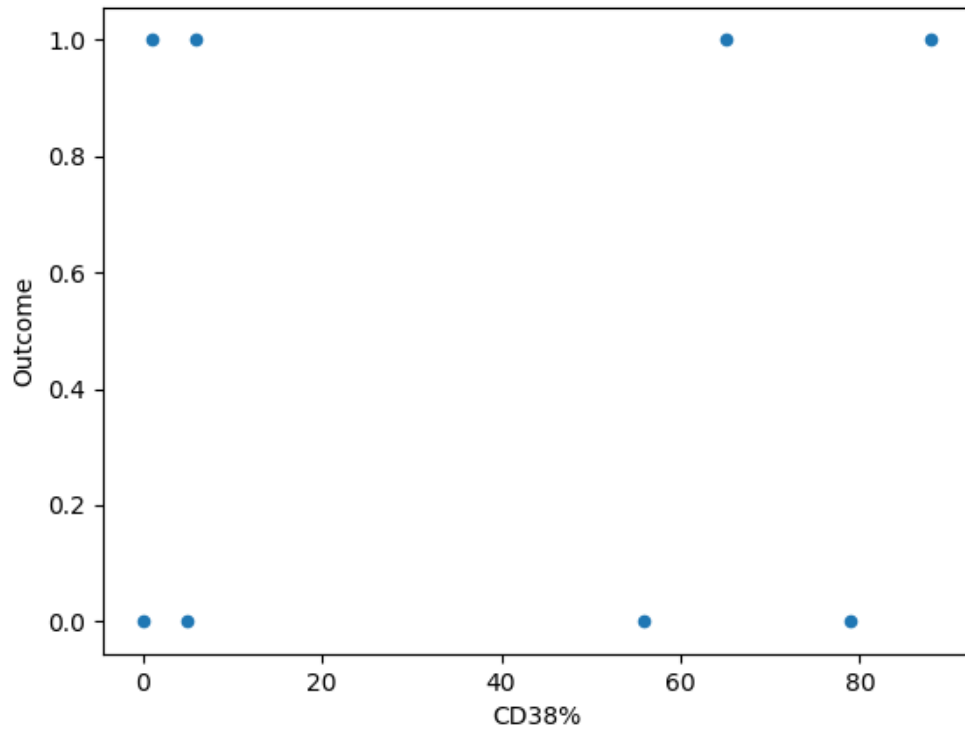Figure 5. **CD180 vs P53 Deletion**. The Pearson Correlation is 0.491390

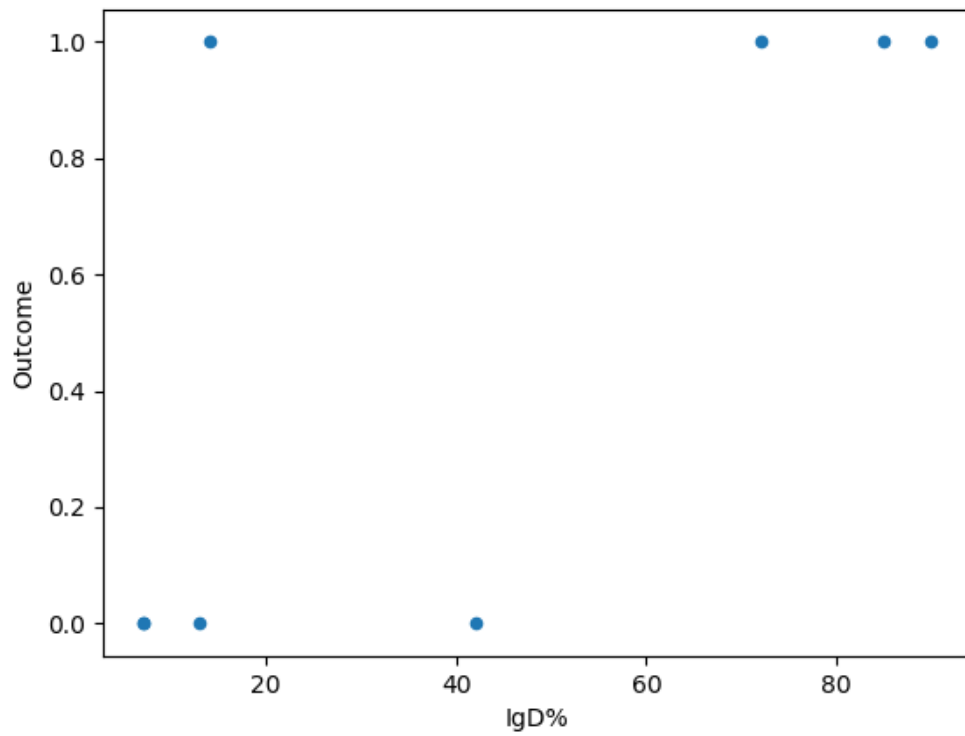Figure 6. **CD38 vs P53 Deletion**. The Pearson Correlation is 0.070159

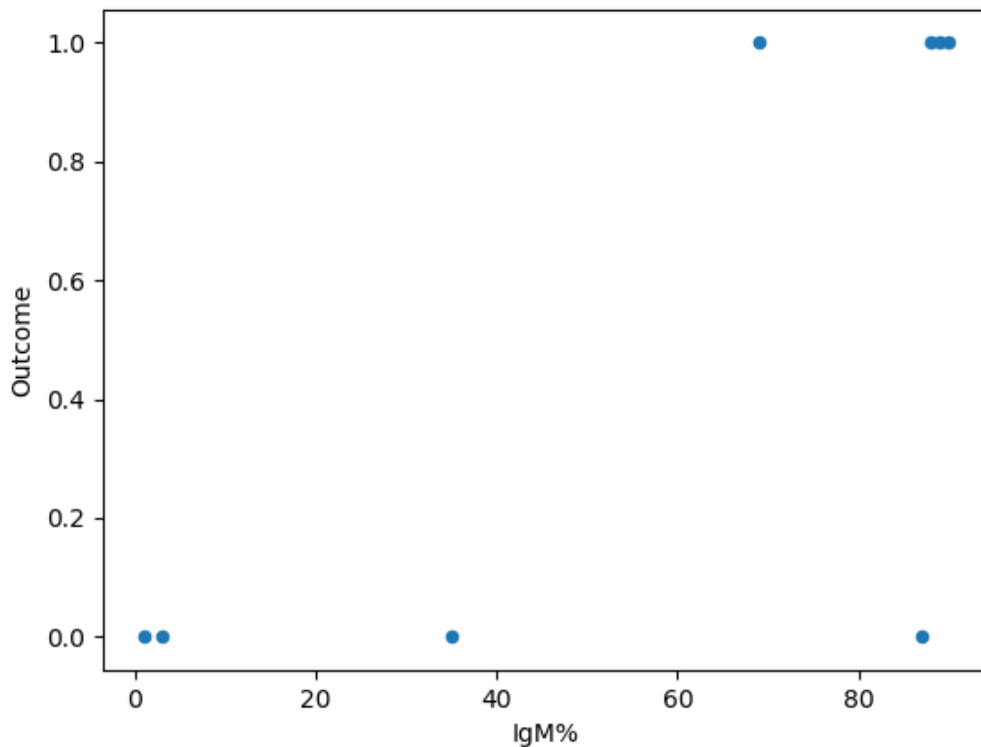Figure 7. **IgD vs P53 Deletion**. The Pearson Correlation is 0.710682

Figure 8. **IgM vs P53 Deletion**. The Pearson Correlation is 0.719465

Model selection:

Model selection is based upon choosing the best model for prediction. The model was grounded on K Nearest Neighbours classifier (KNN) because the algorithm tends to work well with a small dataset with fewer predictors *(Introduction to k-Nearest-Neighbors - Towards Data Science, no date)*. The algorithm must be trained using a dataset with known outputs to build a model. When the trained model accepts new input, it calculates the distance between the input and the training observations, finds the nearest N neighbours, and then votes for a class. The class with the most training observations within the input space will be the predicted outcome.

In order to obtain reliable results, feature selection and cross-validation had to be applied. Feature selection determines the best predictors (in this case, the genetic markers) for prediction. Cross-Validation (CV) is a method used to calculate a reliable testing accuracy from different sets of training and testing observations. Since there was a lack of training samples, Leave-One-Out CV (LOOCV) was used for training and testing the KNN model. In LOOCV, the testing set only contains one testing sample (without the P53 outcome) and the training set is the union of the other folds. It repeats this process until each observation has been included in the testing and training set. After using LOOCV for the KNN model at N neighbours, it would return a mean testing score before incrementing N neighbours.

Feature selection was combined with CV to get the best predictors, but instead of the testing scores, only the columns were returned during the iterations. For each split, the training set was processed by the Recursive Feature Elimination (RFE) algorithm, which ranks the genetic markers according to their importance. The "IgD%" and "IgM%" attributes were the most nominated predictors produced by RFE. However, according to expert opinions, all genetic markers are required for the prediction.
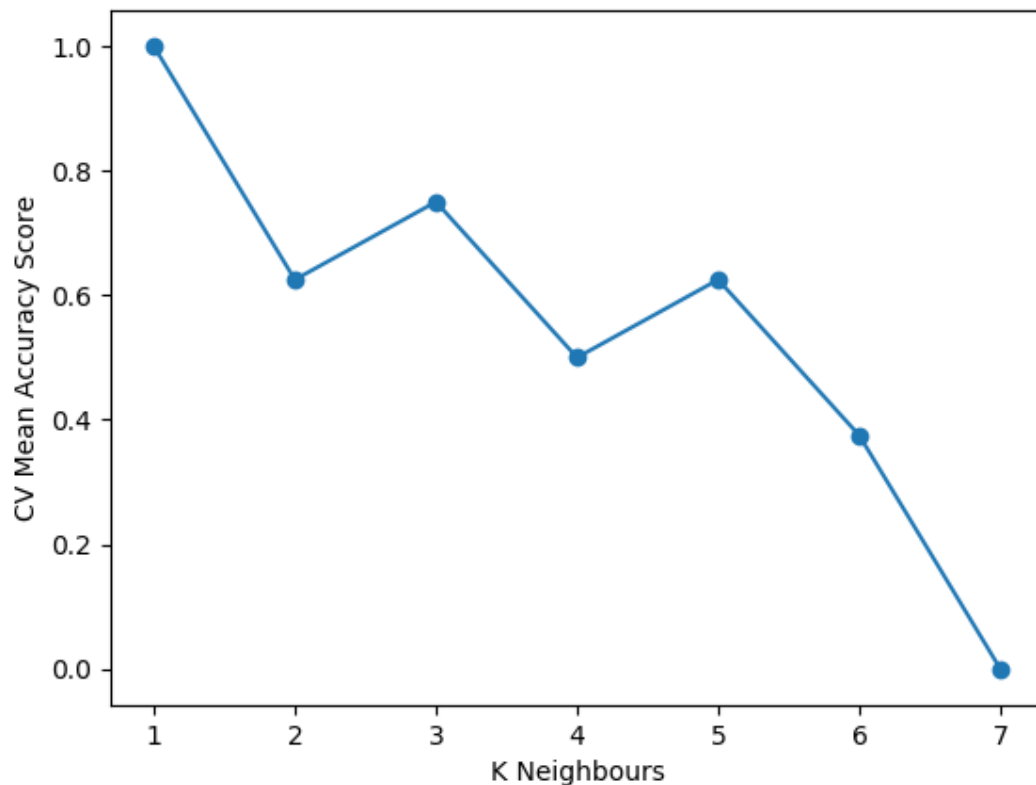

Figure 9. **A graph that visualizes the accuracy score changing as the number of neighbors increase**. K=1, mean score=1.0. K=2, mean score= 0.625. K=3, mean score=0.75. K=4, mean score=0.5. K=5, mean score=0.625. K=6, mean score=0.375. K=7, mean score=0.0.

KNN at N=1 was the best model because it scored the highest mean accuracy score of 1.0. This means that 8 out of 8 testing observations produced by the CV train test splits were predicted correctly when N was equal to 1. Although the scores are perfect, the model could be high in variance because K is very small. If we had more samples, experimentation with more neighbours for the classifier and possibly acquiring better mean scores with lower variance after 7 neighbours could have been possible.
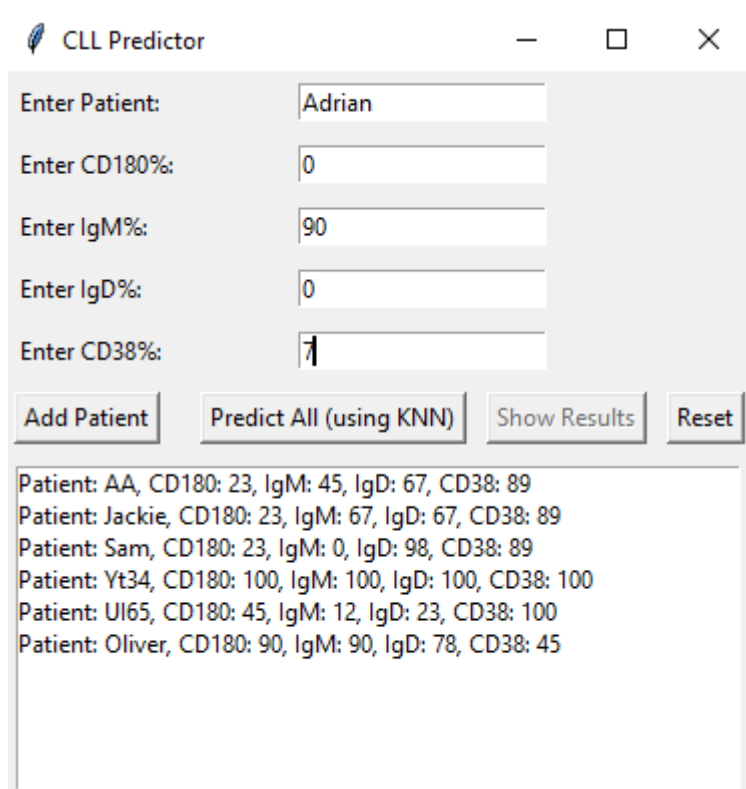
The graphical user interface:

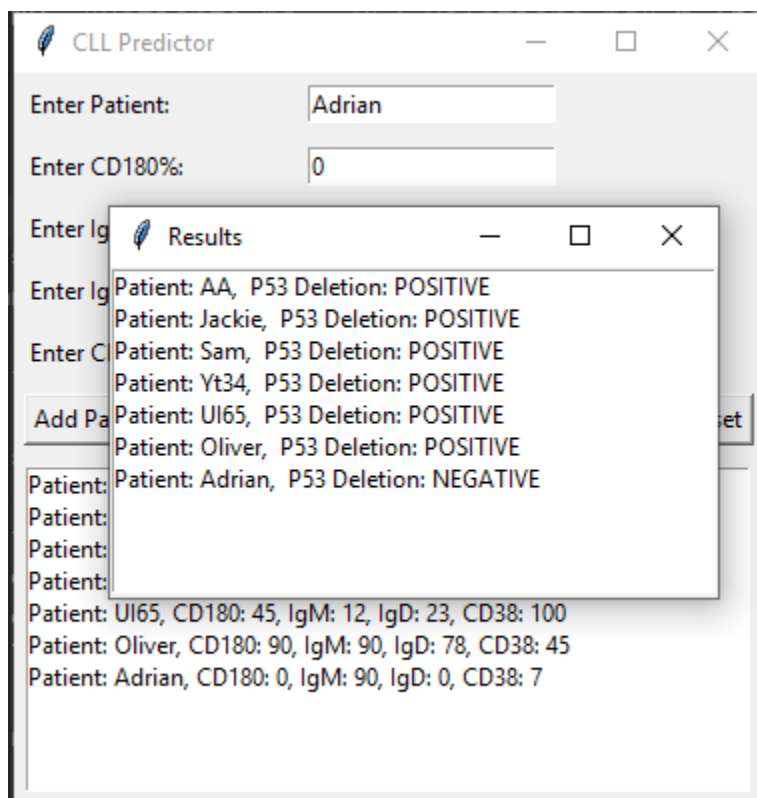Figure 10. **Adding new patients for the prediction.**

Figure 11. **The outcomes for each patient after prediction (P53 Deletion).**

Discussion

Both aims were successfully achieved; finding the association between CD180+ and TP53. Additionally, training the algorithm on the above experimental data collected was successful too. The algorithm was successful in predicting the outcome of TP53 deletion when CD180 and other receptors showed a significant percentage of expression. However, one of the limitations of this project was the sample size (n=8). Although, the model used KNN due to the algorithm working well with a small dataset with fewer predictors, scientific literature for CLL often shows that when a small cohort of patients is used for the sample size results often seem promising/successful. However, when a large cohort is used, these results are negligible due to the nature of CLL; it is very heterogenous for each individual thus difficult to treat. Therefore, to know if this model accurately predicts TP53 deletion and help clinicians in better prognostic measures we need to collect more experimental data based on a large group of patients (n=1000). This would also mean using another model prediction other than KNN, which is better equipped to handle large numbers of data for predictions. This would also mean validation of results using a large cohort of patients and accounting for the differences seen genetically by each patient with CLL.

Experimentally, another limitation is expenses. Each genetic probe is very expensive (~£350) and limits the data collected. If more funding was allowed, other genetics probes could be used with the same computer model. This would generate more predictions for CLL for different genetic aberrations to aid prognostic measures.

The chosen methods in this project was ideal because in order to determine if a sample has TP53 deletion well defined protocols must be followed; blood separation, phenotyping and FISH. This allowed generation of results to then train the algorithm created and assist in checking accurately if the algorithm made "sense/working" if any ambiguous predictions were resulting from the computer model alone. The computer model was effective in predicting the TP53 deletion and was rapid in acquiring results. The speed at which results are given is beneficial to a clinical setting due to large numbers of patients that need information, thus using machine learning is a great method. Conclusively, both methods helped to achieve the results and aims.

Conclusions and Future Work

This project used various scientific methods and machine learning in order to predict the association between CD180+ and TP53 deletions. An algorithm was then trained based on experimental data collected to predict if a patient would have TP53 deletion based on expression of CD180+ and other receptors associated with prognostic significance. Both aims were achieved though the sample size was small. In order to check if this model was accurately working due to the heterogenous nature of CLL, a large cohort of patients must be used whilst consideration of other confounding variables.

Since this project only used TP53 probe due to limited expenses, in future other genetic chromosomal aberrations associated with CLL could be used as probes such as 11q/13q.

This could further help in finding better prognostic outcomes for each patient using this model algorithm. We hope that our algorithm can be used in a clinical setting so that diagnosis/treatment of the disease is made much easier and accessible to all, whilst providing accurate information for those with CLL. Furthermore, perhaps this project can serve as a platform for the prognosis of other diseases and enhance personalised medicine.

Lessons Learned

When working within a multidisciplinary group we found that a new area of research can be formulated due to the expertise and guidance given by academics, and the creativity provided by students. This project was created to provide clinicians and patients better prognostic outcomes when diagnosing/treating CLL, with the use of an algorithm based on experimental data collected. Thus, revolutionising personalised medicine.

Working alongside academics provided an insight into the daily workload and research projects that they are passionate about. It showed us the "world of science" and hard work needed by each project as well as teaching simultaneously. We were able to have an opportunity to understand problems encountered during research such as expenses, time management between students and academics, experimental failure and time required to train the algorithm model numerous times before it worked. Much of the research is trial by error. This requires patience, diligence and communication with those that are well versed in the subject matter.

We learnt new skills mostly in the area of computer science and how the algorithm is created/trained. We all agreed that this was interesting and showed the complexity when designing this project. An opportunity to meet new individuals as well as working with them from various departments within the university was particularly fun. It taught us that creating a project does require numerous people who will give much needed advice/expertise for the success of a project overall. Collaboration is key!

To get this research across to stakeholders, we feel that it is very important to provide the cost expenditure needed for collection of experimental data. The genetic probes are expensive and using other probes to collect more information to gather results, should be accounted for. A GANTT chart is a good idea for providing stakeholders information and a schedule for the timeline of this project. This would also aid in setting a budget for each aspect of the project and prevent over-spending.

To persuade stakeholders that this project is revolutionary and a necessity in healthcare, this model project can be used for a preliminary time period in a hospital clinic. Feedback via clinicians and patients would be collected to understand how effective this model is within a clinical setting. Hopefully, providing the stakeholders a security in knowing that this model is not only working successfully but is effective too. Ultimately providing patients with better care/treatment and faster results. Help would be required by computer experts who train the algorithm, clinicians, patients consent and scientists in the field of CLL to allow for success of this computer model and overall this project. Fundamentally, it is important to acknowledge that health and safety, ethical policies are accounted for and reinforced for everyone involved in this project.

Group Reflection

We jointly agree that the selection process for the team put together was extremely good. All members were repeatedly present and active towards the project. This was most effective during meetings where all members collaborated to resolve a problem via extensive group brain storming, in order to find a solution for the desired outcome of this project.

There were minor barriers to overcome; due to the nature of working within a multidisciplinary group many restraints of the project lied in scheduling meetings. To overcome this, a group chat was created to allow effective communication between all members of the group. This was very successful for those that could not attend meetings at a certain day/time. Therefore, an individual could still receive all the information and give productive input. Due to the essence of this project, a lot of time is necessary to carry out the wide array of experimental protocols for each patient sample for the collection of results. To analyse each sample for end conclusions, adequate time is needed to ensure accuracy and precision of data collected, therefore accumulating valid results. At times, some patients had not produced enough results therefore a repeated analysis of this sample was necessary as well as in the case of complete failed experiments. Therefore, a meticulous amount of time, care and precision was taken when following experimental protocols to ensure no setbacks would occur. Within this project time was a constraint and so in future an accurate GANTT chart should have accounted for the above. Additionally, perhaps more effective communication between members of the group to set a specific day/time each week solely for this project that is scheduled in our official academic calendar.

The biggest enabler for our success was the support of higher education officials involved with this project due to their extensive expertise in both the field of research and producing scientific reports/presentations. Not to mention their guidance when encountering difficulties in the laboratory or with the computer model due to their expertise and extensive experience.

References:

Calin, G., Dumitru, C., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Aldler, H., Rattan, S., Keating, M., Rai, K., Rassenti, L., Kipps, T., Negrini, M., Bullrich, F. and Croce, C. (2002). Nonlinear partial differential equations and applications: Frequent deletions and down- regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences*, 99(24), pp.15524-15529.

Canta, F., Mitrofan, L., Karanis, Y., Mistry, H. and Anger, C. (2016). 17P deletion and TP53 gene mutation (17P/TP53) testing behaviour and treatment patterns for chronic lymphocytic leukemia (CLL) patients in France, Germany, Italy, Spain and UK (EU5). *Annals of Oncology*, 27(suppl_6).

Durrieu, F., Geneviève, F., Arnoulet, C., Brumpt, C., Capiod, J., Degenne, M., Feuillard, J., Garand, R., Kara-Terki, A., Kulhein, E., Maynadié, M., Ochoa-Noguera, M., Plesa, A., Roussel, M., Eghbali, H., Truchan-Graczyk, M., de Carvalho Bittencourt, M., Feugier, P. and Béné, M. (2011). Normal levels of peripheral CD19+CD5+ CLL-like cells: Toward a defined threshold for CLL follow-up-A GEIL-GOELAMS study. *Cytometry Part B: Clinical Cytometry,* 80B(6), pp.346-353.

Geyer, J. and Subramaniyam, S. (2014). Micronuclei and nuclear budding in chronic lymphocytic leukaemia. *British Journal of Haematology*, 167(5), pp.585-585.

Hallek, M. (2017). Chronic lymphocytic leukemia: 2017 update on diagnosis, risk stratification, and treatment. *American Journal of Hematology*, 92(9), pp.946-965.

Porakishvili, N., Memon, A., Vispute, K., Kulikova, N., Clark, E., Rai, K., Nathwani, A., Damle, R., Chiorazzi, N. and Lydyard, P. (2011). CD180 functions in activation, survival and cycling of B chronic lymphocytic leukaemia cells. British Journal of Haematology, 153,(4), pp.486-498.

Seiffert, M., Dietrich, S., Jethwa, A., Glimm, H., Lichter, P. and Zenz, T. (2011). Exploiting biological diversity and genomic aberrations in chronic lymphocytic leukemia. *Leukemia & Lymphoma*, 53(6), pp.1023-1031.

Sun, Y., Wang, S. and Sun, T. (2017). Composite mantle cell lymphoma and chronic lymphocytic leukemia/small lymphocytic lymphoma. *Cytometry Part B: Clinical Cytometry*, 94(1), pp.148-150.

Van Lent, P., Grevers, L., Schelbergen, R., Blom, A., Geurts, J., Sloetjes, A., Vogl, T., Roth, J. and van den Berg, W. (2010). S100A8 causes a shift toward expression of activatory Fcγ receptors on macrophages via toll-like receptor 4 and regulates Fcγ receptor expression in synovium during chronic experimental arthritis. *Arthritis & Rheumatism*, 62(11), pp.3353-3364.

Introduction to k-Nearest-Neighbors - Towards Data Science (no date). Available from https://towardsdatascience.com/introduction-to-k-nearest-neighbors-3b534bb11d26 [Accessed 29 July 2019].

KNN Classification using Scikit-learn (article) - DataCamp (no date). Available from https://www.datacamp.com/community/tutorials/k-nearest-neighbor-classification-scikit-learn [Accessed 29 July 2019].

Towards Data Science (no date). Available from https://towardsdatascience.com/ [Accessed 30 July 2019].