Personalized Medicine Support System for Chronic Myeloid Leukemia Patients

Haneen Reda Banjar

Submitted in fulfillment of the requirements of the degree of Doctor of Philosophy

Supervisor: Professor David Adelson

Supervisor: Professor Deborah White

Supervisor: Dr Alfred Brown

Faculty of Engineering, Computer and Mathematical Sciences

School of Computer Science in collaboration with School of Biological Sciences and

School of Medicine

The University of Adelaide, Australia

Oct 2018

To my parents, my entire family and all the wonderful people who offered their support and encouraged me throughout this dream-come-true journey.

Table of Contents

Abstract

Personalized medicine offers the most effective treatment protocols to the individual Chronic Myeloid Leukemia (CML) patients. Understanding the molecular biology that causes CML assists in providing efficient treatment. After the identification of an activated tyrosine kinase BCR-ABL1 as the causative lesion in CML, the first-generation Tyrosine Kinase inhibitors (TKI) imatinib (Glivec®), were developed to inhibit BCR-ABL1 activity and approved as a treatment for CML. Despite the remarkable increase in the survival rate of CML patients treated with imatinib, some patients discontinued imatinib therapy due to intolerance, resistance or progression. These patients may benefit from the use of second-generation TKIs, such as nilotinib (Tasigna®) and dasatinib (Sprycel®). All three of these TKIs are currently approved for use as frontline treatments. Prognostic scores and molecular-based predictive assays are used to personalize the care of CML patients by allocating risk groups and predicting responses to therapy.

Although prognostic scores remain in use today, they are often inadequate for three main reasons. Firstly, since each prognostic score may generate conflicting prognoses for the risk index and it can be difficult to know how to treat patients with conflicting prognoses. Secondly, since prognostic score systems are developed over time, patients can benefit from newly developed systems and information. Finally, the earlier scores use mostly clinically oriented factors instead of those directly related to genetic or molecular indicators. As the current CML treatment guidelines recommend the use of TKI therapy, a new tool that combines the well-known, molecular-based predictive assays to predict molecular response to TKI has not been considered in previous research. Therefore, the main goal of this research is to improve the ability to manage CML disease in individual CML patients and support CML physicians in TKI therapy treatment selection by correctly allocating patients to risk groups and predicting their molecular response to the selected treatment.

To achieve this objective, the research detailed here focuses on developing a prognostic model and a predictive model for use as a personalized medicine support system. The system will be considered a knowledge-based clinical decision support system that includes two models embedded in a decision tree. The main idea is to classify patients into risk groups using the prognostic model, while the patients identified as part of the high-risk group should be considered for more aggressive imatinib therapy or switched to second-generation TKI with close monitoring. For patients assigned to the low-risk group to imatinib should be predicted using the predictive model. The outcomes should be evaluated by comparing the results of these models with the actual responses to imatinib in patients from a previous medical trial and from patients admitted to hospitals. Validating such a predictive system could greatly assist clinicians in clinical decision-making geared toward individualized medicine.

Our findings suggest that the system provides treatment recommendations that could help improve overall healthcare for CML patients. Study limitations included the impact of diversity on human expertise, changing predictive factors, population and prediction endpoints, the impact of time and patient personal issues. Further intensive research activities based on the development of a new predictive model and the method for selecting predictive factors and validation can be expanded to other health organizations and the development of models to predict responses to other TKIs.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Haneen Banjar

Oct 2018

Acknowledgment

Many thanks to my supervisors: Professor David Adelson, Dr Fred Brown and Professor Deborah White for their valuable reviews and suggestions. I am thankful to Professor Timothy Hughes for his helpful collaboration. Many thanks also to Dr Tamara Leclercq who helped me tremendously in my study years as a graduate student. In addition, I would like to thank Dr Trent Kroger for the challenges he issued regarding my project. Thanks to the anonymous reviewers from Novartis and to the associate editors for their critical comments and suggestions on a previous draft of this work. I would like to thank all past and current supervisors for their encouragement and support. There are also many other academic and administrative staff (past and present) within the University of Adelaide who have facilitated my time as a student and offered advice, guidance and friendship.

My father Reda Banjar taught me the most important lessons in life, determination and optimism, without which my journey would have been less exciting and rewarding. Without the continuous support and encouragement received from my husband Dr Khalid Owaydhah and my mother Jamilah Banjar, I could not have completed my doctoral studies.

List of Tables

List of Figures

List of Acronyms

AIHW	Australian Institute of Health and Welfare
ALL	Acute Lymphocytic Leukemia
AML	Acute Myeloid Leukemia
ANC	Absolute Neutrophil Count
ASH	American Society of Hematology
CART	Classification and Regression Trees
CCR	Complete Cytogenetic Response
CD	Cluster Differentiation
CDSS	Clinical Decision Support System
CLL	Chronic Lymphocytic Leukemia
CM	Combined Method
CML	Chronic Myeloid Leukemia
CMR	Complete Molecular Response
CP	Chronic Phase
CPG	Clinical Practice Guidelines
CRC	Clinical Research Committee
DAS	Dasatinib
DASISION	Dasatinib Versus Imatinib Study in Treatment-Naïve CML
EFS	Event Free Survival
EHR	Electronic Health Records
ELN	European Leukemia Net
ELTS	EUTOS long-term survival scores
ENESTnd	Evaluating Nilotinib Efficacy and Safety in Clinical Trials—Newly Diagnosed Patients
ESMO	European Society of Medical Oncology
EUTOS	EUropean Treatment Outcome Study
FDA	Food and Drug Administration
FN	False Negative

FP	False Positive
G-mean	Geometric mean
IC50	Inhibitory Concentration of 50%
IM	Imatinib
IRIS	International Randomized Study of Interferon Versus Imatinib STI571
IS	International Scale
K-NN	K-Nearest Neighbour
KAMC	King Abdulaziz Medical City
KFSHRC	King Faisal Specialist Hospital and Research Center
MCR	MATLAB Compiler Runtimes
MDT	Meta Decision Trees
MIC	Maximal Information Coefficient
MMR	Major Molecular Response
NCCN	National Comprehensive Cancer Network
NCI	National Cancer Institute
NHMRC	National Statement of Ethical Conduct in Human Research
NIL	Nilotinib
NPV	Negative Predictive Value
OA	OCT-1 Activity
OCT-1	Organic Cation Transporter
OS	Overall Survival
PGx	Pharmacogenomics
Ph	Philadelphia chromosome
PPV	Positive Predictive Value
RAC	Research Advisory Council
REC	Research Ethics Committee
RQ-PCR	Real-time Quantitative Polymerase Chain Reaction
RT-PCR	Real-time Polymerase Chain Reaction
SEER	Surveillance, Epidemiology, and End Results Program

SNP	Single Nucleotide Polymorphisms
SVM	Support Vector Machine
TIDEL	Trial of imatinib with Dose Escalation in chronic myeloid Leukemia
TKI	Tyrosine Kinase Inhibitor
TN	True Negative
TP	True Positive
VPH	Virtual Physiological Human
WCC	White Cell Count

List of Publications

The following publications emerged from this research (as of Oct 2018)

Intelligent Techniques Using Molecular Data Analysis in Leukemia: An Opportunity for Personalized Medicine Support System

Banjar H, Adelson D, Brown F and Chaudhri N. BioMed Research International. 2017

Personalized Medicine Support System for Resolving Conflict in Allocation to Risk Groups and Predicting Patient Molecular Response to Molecular Targeted Therapy

Banjar H, Adelson D, Brown F and Leclercq T. Health Informatics - An International Journal (HIIJ). 2017

Modelling Predictors of Molecular Response to Frontline Imatinib for Patients with Chronic Myeloid Leukemia.

Banjar HR, Ranasinghe D, Brown F, Adelson D, Kroger T, Leclercq T, White D, Hughes T and Chaudhri, C. PloS One. 2017

Consistency Test between Scoring Systems for Predicting Outcomes of Chronic Myeloid Leukemia in a Saudi Population Treated with Imatinib

Banjar HR, Alsobhi E. International Scholarly Research Notices. 2017

Chapter 1: Introduction

Introductory Background

Cancer patients have a strong belief in treatment that will be able to relieve their pain and suffering. However, in many cases, clinicians have uncertain expectations about patient responses. Consequently, treatment failure situations are likely to be recorded, which could be extremely disappointing to patients. In addition, clinicians have rational explanations and scientific justifications, but treatment failure is an undesirable outcome for both sides. Consuming the incorrect drug over a long period of time multiple times per day is harmful. For instance, in childhood Acute Lymphoblastic Leukemia (ALL), an incorrect choice regarding individual treatment could produce serious problems, such as the progression of the disease, relapse and, in some cases, death [1]. The long histories of toxicology and pharmacology have strongly shown the significant research efforts undertaken in studies of drug resistance and toxicity [2]. Many theories and facts have been investigated in order to avoid failed treatments and to study the causes of undesirable side effects [3, 4]. However, many exceptional cases have been recorded for several medical profiles from different populations who would benefit from individualized medicine [2]. This is not new in medical science [2, 5], where drugs are studied in terms of the population, based on some criteria, such as age or sex. Individualized medicine is more precise [6]; the principle is basically to transform the population sample of the medical trials to be involved with their individual parameters in other specific medical trials. This is done to classify them into a very small subpopulation that shares a response to a specific treatment. In these medical trials, studies of biological parameters are considered due to their role in causing differences in responses to drugs. Indeed, some biomedicine researchers [7-10] have discussed the importance of including a combination of biological and clinical data. In some diseases caused by genetic abnormalities, the biological parameters could probably cause diversity in in the patients' responses to drugs with resistance or intolerance.

Medical Background: Chronic Myeloid Leukemia

In 2017, the Australian Institute of Health and Welfare classified leukemia as the sixth-most commonly diagnosed cancer [11]. One cancer that falls under this subcategory of diseases is chronic myeloid leukemia (CML). Patients with CML experience unregulated growth of the predominant myeloid cells occurring in their bone marrow. Those CML patients [12] have a genetic abnormality known as the Philadelphia (Ph) chromosome that has been discovered in their blood cells. The fusion between the tyrosine kinase gene in chromosome 9, known as Abelson (Abl), and the gene at chromosome 22, known as the breakpoint cluster (Bcr), produces the BCR-ABL gene, which encodes for the constitutively active tyrosine kinase Bcr-Abl. Inhibiting Bcr-Abl activity is essential in treating CML patients. As one of the molecular-targeted therapies, tyrosine kinase inhibitors (TKIs) are the most successful therapeutics that have been used over the last decade [13-15].

Three common TKIs approved by the U.S. Food and Drug Administration (FDA) in the treatment of CML include imatinib (Gleevec,® Novartis), nilotinib (Tasigna,® Novartis), and dasatinib (Sprycel,® Bristol-Myers Squibb). These TKIs are used as first-generation imatinib and second-generation nilotinib and dasatinib therapies. The use of imatinib treatment as a firstline therapy is an effective treatment strategy that has shown variable results [16]. Whenever the CML patients are not likely to achieve the targeted response in a specific time frame, it could beneficial to switch them to a second-generation TKI [17]. Many failed cases have been recorded with poor outcomes based on the use of imatinib in frontline therapy [14, 18]. Since then, researchers have investigated a new treatment strategy. This strategy suggested using the second-generation of TKIs in frontline treatment to achieve the optimal treatment for CML [19-21]. Indeed, studies [22, 23] have examined the use of nilotinib and dasatinib as frontline therapies. For example, in the DASISION trial, Kantarjian et al. [23] demonstrated

that the safety profiles of dasatinib in frontline treatment resembled those of imatinib. Therefore, the choice among these TKIs should be based on evidence and reasoning. According to Erba [24], it is necessary to wait and collect valid data to gain a deep understanding of each therapy and provide evidence for choosing the best TKI. Shami and Deininger [18] stated in their conclusion based on current trials that using dasatinib or nilotinib in newly diagnosed CML patients could be reasonable, but due to the limitations of existing data comparing both drugs, it is not preferable to recommend one therapy over the other. In 2010, Wei et al. [20] reviewed a comparison of the three agents as frontline therapies. They advised obtaining better knowledge of the three therapy toxicities and following up with the use of second-generation therapy as frontline treatment over the longer term, as the IRIS trial [16] did with imatinib.

The American Society of Hematology (ASH) has established annual meetings to discuss medical trial outcomes. In 2006, the 48th ASH annual meeting [25] displayed a poster about the International Randomized Study of Interferon Versus imatinib STI571 (IRIS) trial. In this study, it was shown that imatinib is appropriate for CML patients in the chronic phase, as the results indicated that patients who receive imatinib therapy over the long term (5-6 years) achieve MMR in 90% of cases [16]. These results show the efficacy of continuing or receiving imatinib over time for CML patients. In 2010, two significant medical trials, DASISION (Dasatinib Versus Imatinib Study In Treatment–Naïve CML) and ENESTnd (Evaluating Nilotinib Efficacy and Safety in Clinical Trials–Newly Diagnosed Patients), were discussed at the 52nd annual meeting [26]. In DASISION [23], 259 patients were studied regarding responses to 100 mg. DAS once daily versus 260 patients who consumed 400 mg. of IM once daily. The Major Molecular Response (MMR) was lower in the imatinib arm compared to the dasatinib arm. The most important result was the safety profiles of using these drugs, which were similar. The ENESTnd trial focused on the comparison of nilotinib versus imatinib [22]. The samples were derived from newly diagnosed CML patients, with 282 patients given 300

mg. of nilotinib twice daily, 281 patients given 400 mg. of nilotinib once daily or 283 patients administered 400 mg. of imatinib once daily. The median follow-up for these patients was 18 months. The Cytogenetic Response and MMR were lower in the patients treated with imatinib than the patients treated with NIL.

From these trial results, the debate on selecting the optimal TKI to achieve optimal patient response has been established with the recognition of the need for a predictive approach that could accurately predict the response of the patient's upfront treatment. The researchers also recognize the need for a medical trial that could compare each TKI against the others as first-generation treatments. It is clear that collecting the data for CML patients and comparing each TKI against the others is very important. However, the nature of the response to TKI treatment requires a long term to observe the results. Consequently, it is important to find a direction to improve the ability to manage CML disease in individual CML patients by selecting the appropriate TKI.

Response to TKI Therapy

Three types of patent responses can be identified using the criteria employed by the European Leukemia Net (ELN) [27]: the hematologic, cytogenetic and molecular response to treatment (see Appendix 2: ELN Response Definition). Firstly, hematologic response is identified by counts of non-elevated white blood cells, platelets and basophils in the normal range. The second response is cytogenetic response, which is identified by reducing or eliminating the number of cells expressing the Philadelphia chromosome. Finally, for the molecular response, which is the highest degree of remission for CML patients, PCR tests the absence of Bcr-Abl transcripts in the peripheral blood or bone marrow. Measuring the level of *BCR-ABL* mRNA transcript and the Ph-positive cell frequency can accurately monitor the response to treatment [28].

The National Institutes of Health Consensus Group proposed the use of an international scale (IS) to monitor CML patients. The IS standard baseline is taken to represent 100% of definitions from 30 CML patients who enrolled in the (International Randomized Study of Interferon Versus STI571) IRIS study as the median value of *BCR-ABL1* mRNA at the time of diagnosis [29]. Two groups of molecular responses were identified based on this standardized baseline: MMR which refers to a *BCR-ABL1* transcript level of ≤ 0.1%, and complete molecular response (CMR), which refers to an undetectable *BCR-ABL1* transcript level. The optimal response at any time after 12 months is a *BCR-ABL1* transcript level ≤ 0.1[30]. According to the IRIS study [31], achieving MMR is associated with a low risk of disease progression based on the finding that 100% of patients remained free of progression to advanced CML phases at 60 months. In addition, monitoring the level of BCR-ABL mRNA transcript can be a predictor for subsequent loss of response or the development of mutation [32, 33]. These data suggest the relative benefits of predicting MMR in terms of long-term survival.

Recent Personalized Treatment Approaches in CML

Clinicians use patients' characteristics upon diagnosis in a prognostic scoring system to classify CML patients into risk groups. Four scoring systems are historically developed, including Sokal [34], Hasford [35], the European Treatment and Outcome Study (EUTOS) [36], and the EUTOS long-term survival score (ELTS) [37]. However, the Sokal and Hasford scores' ability to identify the newer TKIs' prognosis risk groups is still unclear [38], and the ELTS introduced in 2015 and validated with 2,205 patients is not widely applied in CML treatment decisions.

Many studies [15, 39] have shown that predictive factors could probably assist in predicting patient response. Milojkovic et al. [39] conducted their study to predict the success

or failure of treatment with second-generation tyrosine kinase inhibitors in CML patients by using univariate analyses. They analyzed a cohort of 80 CML patients in the first chronic phase who were treated with dasatinib or nilotinib. Their score system predicted the probability of CML patients achieving a complete cytogenetic response. Their system was based on three factors: cytogenetic response to imatinib, Sokal score and recurrent neutropenia during imatinib treatment. Although this study used simple statistical methods, the system succeeded in classifying three risk categories: good, intermediate and poor risk. In addition, Jabbour et al. [15] also studied the predictive factors for predicting 123 CML patients' responses after imatinib failure. The variables used in this study included age, sex, CML duration, months of performance status, splenomegaly, prior interferon therapy, peripheral blood, bone marrow, best cytogenetic response to imatinib, second-generation therapy, active disease at the start of the second-generation TKIs, clonal evaluation, higher than 90% Philadelphia positivity, and IC50 [42] for nilotinib and dasatinib for in-vitro inhibition of kinase activity of the mutated point in BCR-ABL. Even though they used univariate and multivariate analyses, such as the logistic regression model and the Cox proportional hazard model, to identify prognostic factors associated with cytogenetic response and survival, they succeeded in identifying three risk groups: low, intermediate and high risk. A recent study funded by the Anderson Cancer Centre and National Cancer Institute proposed a model to predict the sustained molecular response for at least two years [40]. However, these current scores and model use mostly clinically oriented factors that are not directly related to the patients' genetic or molecular indicators.

Two of the previously identified predictive factors are highly involved in predicting the molecular response in CML. The first such factor is IC50. In 2005, White et al. [41] studied the inhibitory concentration 50% ($IC50^{imatinib}$) as a predictor of molecular response for CML patients. The results demonstrate that the $IC50^{imatinib}$ is a powerful predictor pre-treatment [42]. The second factor is the activity of Organic Cation Transporter-1 (OA) [43]. There are two

functions for OCT proteins, which are the cellular uptake and excretion of a number of exogenous and endogenous cationic and uncharged substances. OCT-1 protein activity (OA) can be measured by uptake in the presence and absence of a specific OCT-1 inhibitor. It has been found that patients with high OA have better molecular responses than patients with low OA and OA is considered a predictive factor for responses to imatinib, but not for nilotinib or dasatinib [42, 44]. White et al. proposed [43] that in CML patients treated with imatinib, the use of OA pre-therapy was a predictor for the long-term risk of resistance and could be used to individualize dosage strategies. Thus, using OA to estimate the response could lead to better results only for imatinib therapy.

Personalized Medicine Support System

Personalized medicine support systems combine two concepts: personalized medicine and clinical decision support systems. As Muller defined it, "personalized medicine refers to tailoring of medical treatment to the individual characteristics of each patient" [2]. A clinical decision support system refers to any computer program that supports healthcare professionals in their clinical decisions, including any computer system that uses clinical data or knowledge to make a medical treatment decision [45]. A clinical decision support system is a tool used to facilitate personalized healthcare, and is especially useful when dealing with the complexities of genomic information, which could be considered difficult for clinicians to manage manually without the support of computer programs [46]. Similar to pharmacogenomic science, which personalizes treatment selection based on unique patient genetic information [47], personalized medicine support systems personalize treatment selection based on individual patients' clinical, biological and molecular data or knowledge. The personalized medicine support system is a knowledge-based clinical decision support system. Knowledge-based systems typically use expert knowledge in the application domain and definition of the problem [48].

Since the patient data stored in electronic record, it leads to be considered a big data in health. According to Huanng et al. [49], the future of health could be changed with the big data in health. They state that the patient data, such as medical histories and genetic test data is foundation for personalized medicine. The selection of molecular-targeted therapy for cancer patients is a recent issue in personalized medicine. The prediction of response to treatment and risk group identification cannot be addressed manually. Therefore, intelligent techniques could provide insights to these problems. The electronic health records as a source for population can be analyzed by machine learning techniques to solve the selection of molecular-targeted therapy problem.

In the era of personalized medicine, it is vital to have an accurate and personalized medical support system for the selection of molecular-targeted therapy for cancer patients, which is based on patients' molecular, biological, clinical and pathological profiles and their molecular responses [50]. This is crucial to assist clinicians in making accurate decisions about the most effective and least toxic molecular-targeted therapeutic options available.

Currently, clinical decision-making is combined with increasingly sophisticated medical technology, such as equipment and instruments to diagnose many prognostic systems and predictive assays, which makes the decision-making process more complex. For example, medical technology known as Quantitative RT-PCR (Q-PCR) is used to measure the actual percentage of BCR-ABL mRNA transcripts and to show whether a correlation between peripheral blood and the bone marrow results [51]. A significant variability in the results obtained from different laboratories could be due to technical differences in the application of measuring Q-PCR for BCR-ABL. Thus, An et al. suggested minimizing the differences in the Q-PCR to maximize the accuracy of the clinical decision. The complexity of decision-making might result from factors such as increasing treatment options, the multitude of strategies utilized for therapy, heterogeneity in responses, or patients with multiple co-morbidities who

are using multiple medications [52].

Technology may also increase decision complexity due to the development of multiple prognostic score systems for allocation of the risk index to treatments that are equally considered to have efficient scores with different populations [34, 35, 37 , 53]. In addition, the situation may be compounded by medical technology developments, particularly the growth of molecular technology and the introduction of many predictive assays and molecular tests [54]. All these factors may impact decisions in clinical practice. Hunink et al. agreed that medical decisions become complex based on their example of the first edition of The Merck Manual (1899), which includes 192 pages compared with the centennial edition (1999), which includes 2,833 pages [55].

Personalized medicine support systems could offer an effective tool to CML patients who suffer from a history of unsuccessful or failed medication regimens. An opportunity is now available for medical researchers and computing researchers to transform the paradigm toward patient-centered disease management. Human molecular tests and the development of molecular-targeted therapy have provided additional levels of information that need to be processed with the assistance of information technology. The proposed system therefore offers a promising opportunity in the treatment of CML. The biggest advantage of this method is that the selection of medicine is personalized so that an individual can have their own means of treatment, tailored to their needs. This provides hope to CML patients that they can recover from their currently terminal illness and revert back to normal health.

Problem Statement

Why Do Clinicians Need a Personalized Medicine Support System for CML?

There are five central problems that have been observed from a medical perspective:

1. Comparative studies [56- 68] were conducted to establish which prognostic score is most accurate for identifying risk. Since each prognostic score may generate conflicting predictions of the risk index, it is difficult to know how to treat patients with conflicting predictions. As each study recommends using different scores, a tool is needed to resolve conflicts among prognostic score predictions to promote consistent results and increase clinicians' trust in and reliance on these results to make appropriate treatment decisions. The European Leukemia Net (ELN) currently recommends management of CML by achieving at least MMR [30] but previous studies have failed to consider the potential for combined prognostic scores to identify risk categories in CML based on patients' long-term molecular responses.
2. Since prognostic score systems are developed over time, as with the development of the latest ELTS score, patients could benefit from newly developed systems. A valuable score carries information for allocating risk groups that could assist clinicians in CML treatment. The previous research has not considered evolving based on new incoming information.
3. Response information comparing each TKI against the other is not available from the previous medical trials. This challenge could probably be addressed using the available information about frontline therapy from patients' existing profiles to predict the response to TKI.
4. According to a recent economic analysis [69], the cost of treatment with imatinib is approximately **£**57/400 mg., **£**86/600 mg. and **£**114/800 mg., and this dosage is taken every day. Although imatinib is a well-tolerated therapy and leads to a 90% survival rate after five years [2], nilotinib and dasatinib are also clinically proven molecular therapies for CML patients [70]. Approximately 24% of patients develop resistance to imatinib, and approximately 26% of patients discontinue their therapy due to their intolerance for imatinib [21]. Thus, patients should benefit from a tool that recommends using imatinib or

other TKIs before consuming the expensive imatinib and then switching to another treatment after long-term monitoring.

5. Previous studies have not investigated the relationship of molecular response and the predictive assays $IC50^{imatinib}$ and OCT-1 activity (OA) as predictors. There is no study using machine learning in predicting individual responses to TKIs in CML patients or using clinical, biological, molecular-based data and predictive assays in a predictive model.

There are also a number of possible challenges from a computational perspective; we concentrate on four main challenges.

1. This research is a challenge from both a data and a knowledge perspective because in developing medical applications, computer scientists should study the relevant medical background to integrate medical knowledge in the development of the personalized medicine support system. Moreover, the data collection and population identification processes should comprise a particular investigation, and experts in the domain should confirm the exclusion and inclusion criteria. A survey should also be conducted by the computational researchers to generate information relating to medical knowledge about existing prognostic and predictive methods. Defining the predictive factors and reformatting them using domain knowledge is also an important process as the interpretability of the model in clinical prediction rules form will be based on the expert categories and definition.
2. The next challenge originates from selecting the induced models, which should result in several requirements: high performance, interpretable, fewer predictive factors, and the ability to outperform existing methods. The model induction and selection of relevant predictive factors should be implemented with interpretable methods (white-box testing) that are interpretable in the context of expert knowledge and easy to use in practice [71]. Model evaluation, particularly after using predictive factor selection approaches, creates a

challenge around selecting the best models. The models should be selected according to their validation performance. Thus, the computational cost can be very high based on the number of constructed models and the compression of the model's performance. Finally, under this challenge, the generation of clinical prediction rules needs to be represented in such a way that the computer easily understands them in order to prepare rule sets in adapting to new knowledge.

3. This research also generated challenges by studying the level of consistency between existing scoring systems and combining validated prognostic scores to resolve conflicts on predicting risk groups in cancer patients. Before commencing treatment, it is vital to ensure that there is no conflict with decisions based on prognostic scores that have been developed and validated, and that are already in use. In clinical practice, the clinicians rely on their preference to decide which therapy to prescribe based on one prognostic score, or sometimes patients' preferences for cheaper brands [72]. However, consistency is the agreement between the prognostic scores that is a useful output, which clinicians need to receive, but unfortunately, these prognostic scores may generate conflict in their outputs or decisions. The research challenge is to handle this difficulty by studying the level of consistency, and the proposal involves combining methods [73] to reach one output and examine the performance of risk group allocation against patients' actual molecular responses. Several methods should be examined and evaluated to select the best method to combine prognostic scores compared with a single source of prognosis. This problem has been introduced for the first time in the selected medical domain. Therefore, the challenge for this study is to review the current body of work in the medical literature in light of the conflict issues outlined earlier and to propose a solution to resolve conflicts at different combined levels.

4. The fourth challenge is posed by changes in care delivery and the need to keep up with rapid medical and scientific discoveries. More care is required to develop rapid-learning health systems based on a framework to spur delivery in light of recent developments in health information technology, and to access health data and apply evidence in making the right decisions. An important feature in the proposed framework is its ability to capture knowledge from single and multiple sources and to reach agreement on decisions from two sources regarding the same patient.

Contribution to the Discipline

This research addresses the problem of TKI therapy treatment selection by accurately grouping patients into different risk groups and predicting their molecular responses to the selected treatment. The overall thesis aims to research issues that are significant while selecting the most effective treatment from the TKI therapies that are available for CML patients. Most of the empirical research on personalized medicine in leukemia discovered that there is a need to combine different methods to cope with conflicts arising from prior knowledge. The research aims to address this need by framing the problem as a classification problem and come up with a knowledge-based decision support system.

From the discussion of various provided studies on individual models, the research proposed the development of both a predictive model and a prognostic combined model.

1.1. Based on prior medical information, my research has also contributed to knowledge on the TKI treatment method through experimental and automatic comparisons of the study's model behavior and the design of a new structure for a personalized medicine support system.

1.2. My research has contributed to knowledge on TKI treatment by discovering clinical prediction rules from the relationship between the predictive factors and the molecular responses to the selected treatment.

The focus of the research was precise classification, therefore the introduction of a personalized medicine support system structure based on a medical application problem was significant in solving problems related to clinicians' diagnosis decisions. The significant contribution to CML treatment management is that the data included in the research has not been analyzed in any previous studies on personalized medicine because there is no prior research that describes a process used in this study to solve the problem of treating CML. Finally, this study has contributed by helping clinicians with treatment selection problems, therefore expanding the potential applications of TKIs in personalized medicine.

Thesis Objectives

The main objective of this thesis was to improve the ability to manage CML disease in individual CML patients and support CML physicians in TKI therapy treatment selection by correctly allocating patients to risk groups and predicting their molecular responses to the selected treatment. We developed and implemented a knowledge-based system to provide clinicians, patients, and researchers with a platform to allocate risk groups to resolve conflicts on prognostic scores, to predict molecular responses to the standard TKI imatinib, and to extract knowledge to discover the relationship between the predictive factors and achieving a molecular response. The framework has not been considered in previous personalized medicine research; in the existing literature, this process has not been specifically applied to the problem in the medical field being considered in this study. As a result, the current study will add substantial value to the existing literature. Figure 1 shows the hierarchy and linkage of the main aim, the six major objectives, and the 11 sub-objectives of this research. The six journal papers developed to address the main aim of each major objective and sub-objective are also included. A generic framework is presented that includes the following main, major and sub-objectives.

Objective 1 To conduct a systematized literature review on personalized medicine in leukemia and synthesize findings across studies related to intelligent techniques in leukemia.

Sub-objective 1.1 To identify opportunities for further research into personalized medicine support systems in one category of leukemia, namely, CML.

Objective 2 To propose a framework for the selection of treatment in cancer patients based on patient risk profile and molecular response to molecular targeted therapy.

Sub-objective 2.1 To design an algorithm for a predictive model to predict molecular response to molecularly targeted therapy for treatment management by integrating domain knowledge in the learning process.

Sub-objective 2.2 To combine prior prognostic scores and resolve conflict between scoring systems.

Sub-objective 2.3 To automate selection of the predictive factors for predicting the molecular response to molecular targeted therapy and to generate clinical prediction rules.

Objective 3 To study consistency between prognostic score categories used to allocate CML patients to risk groups.

Sub-objective 3.1 To identify risk groups.

Sub-objective 3.2 To identify conflict groups.

Objective 4 To develop a combined prognostic model to resolve conflict between prognostic scores in patients with conflicting predictions.

Sub-objective 4.1 To automatically filter the models and select the highest performance model that learn the best conflict resolution strategy between prognostic scores in patients with conflicting prediction

Sub-objective 4.2 To update entire architecture when adapt a new prognostic model.

Objective 5 To model predictors of molecular response to frontline imatinib for patients with CML.

Sub-objective 5.1 To understand whether rules exist to predict molecular response for CML patients treated with molecular targeted therapy (imatinib) from the clinical, molecular, and cell-count observations collected at diagnosis and categorized based on the available knowledge.

Sub-objective 5.2 To build a predictive model to predict molecular response for molecular targeted therapy (imatinib) in CML patients with better prediction results than those obtained with predictive assays and previous scores.

Objective 6 To discuss the personalized medicine support system's design, implementation, and testing.

Sub-objective 6.1 To implement system aids in identifying the risk group, identifies any conflict in treatment decisions

based on multiple prognostic scores, predicts MMR at 24 months as a long-term endpoint, and supports clinicians as they select patient treatments.

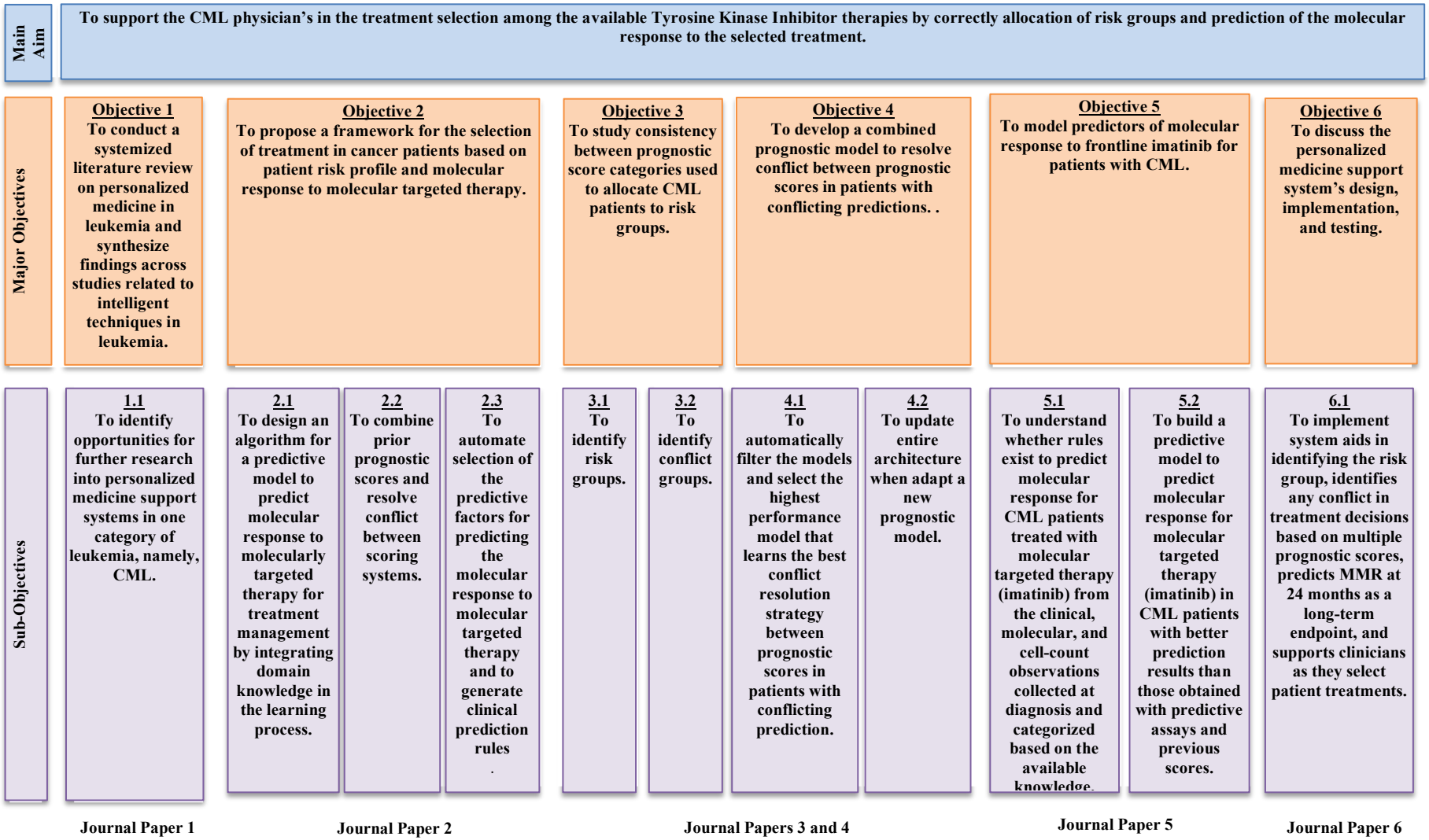


Figure 1 Hierarchy and linkage of the main aim, six major objectives, and eleven sub-objectives of this research. The six journal papers developed to address the main aim and each of the major objectives and sub-objectives, are also depicted

Discussion That Will Be Presented in This Thesis

A personalized medicine support system will be discussed in this thesis. This system is a type of knowledge-based clinical decision support system (see Chapter 3 for details) that can be considerably valuable in the process of creating a personalized medical treatment plan for patients as they reduce costs, improve outcomes and create safer healthcare [46]. This system will increase the effectiveness of the decision by using predictive and prognosis models at diagnosis. Medical decision makers may optimize the patient's journey to avoid possible treatment failures, with assistance from the personalized decision support system.

Using two or more characteristics of patient data and expressing various types of human knowledge can assist healthcare professionals in making appropriate and automated decisions [45 , 74 , 75]. Different prognostic (Chapter 4) and predictive (Chapter 5) factors will be evaluated, as well as their relevance in identifying risk profiles and predicting the response to imatinib in patients who have been diagnosed with CML, and the overall necessity of personalized medical treatments when treating patients with CML. In this thesis, risk stratification methods and the clinical, biological, and molecular-based indicators will be studied to see how these particular factors can be used to help clinicians and other health professionals treat CML patients.

When studying the prognostic scores to identify the risk level and predict the molecular response to TKIs and how they can assist health professionals in making pre-treatment decisions for patients with CML, one must also consider a major work published by the New South Wales Blood and Marrow Transplant Network [76] on acute myeloid leukemia (AML), where the efforts are clear to build six models of care for AML patients: i) presentation and referral, ii) diagnosis and work-up for treatment, iii) treatment, vi) transplant, v) follow-up care, relapse and long-term survivorship, and vi) end-of-life care. To extend the diagnosis and the work-up required for a specific treatment model, our

proposed system (Figure 1, Chapter 3) could be used as it may assist in establishing accurate decisions to select the best treatment for the AML patient.

The National Comprehensive Cancer Network Guideline (NCCN) [29] lists various TKI options and methods that can be utilized to treat cases of CML. These guidelines are backed by various interdisciplinary researchers who have analyzed cancer and blood disorders from a perspective that "treat[s] the entire individual, not just the disease" [77]. In keeping with this information, a change in methodology from automatically using standard TKI treatments (imatinib) for CML to studying patients' unique genetics and prognostics before selecting an appropriate treatment will constitute the contribution of this thesis. The models presented in this thesis (Figure 2, Chapter 6) will primarily either recommend the standard TKI treatment (imatinib) or not recommend it because of its reliability and long track record of precedence [78]. In addition to imatinib's favorable prior results, patients' responses to imatinib are fairly easy to study, as the measurement of *BCR-ABL1* mRNA transcript levels in the peripheral blood of CML patients are monitored by RQ-PCR [79].

Personalized medicine support systems will include different models where each model performs a specific function. The purpose of designing the personalized medicine support system is to provide one system, which integrates different decision support functions that serve within individualized therapy. Furthermore, the personalized medicine support system has a functional integration for any technology, for example, machine learning algorithms (Chapter 4 and 5), feature selections (Chapter 5), combining several classifiers (Chapter 4, Part 2), and conflict analysis (Chapter 4, Part 1) to support treatment-management tasks in general and decision-making in particular. In this system, each function (allocation of risk groups and prediction of molecular response to imatinib) enhances the other functions by increasing the capabilities of the whole personalized medicine support system as each model performs those sub-tasks that it performs best. The expertise from the

medical domain is integrated into the system, which helps to closely define medical concepts in the framework, and it relates to clinicians' preferences. The high quality of a knowledge-based system that integrates knowledge and extracts knowledge is also important to tailor decision outputs for multiple resources. The knowledge-based clinical decision support system (personalized medicine support system) is more difficult to model because integrating the expert domain usually requires computer scientists to have a broad background and knowledge about medical applications, as well as about clinical guidelines and protocols.

The studied information that will be referenced in support of our thesis statement will be of a knowledge-driven and data-driven nature and will include expert background and data sets, such as clinical trials and patient records from hospitals. The future goals and potential methods of validating the information presented in this thesis will be detailed as well.

This research will review intelligent techniques for personalized medicine in leukemia research using molecular data. It then will synthesize findings across studies related to intelligent techniques in leukemia treatment. Specific attention will be paid to particular categories of these studies to help identify opportunities for further research into personalized medicine support systems in one category of leukemia, namely CML. It will address the challenges mentioned earlier by proposing a framework for the selection of treatment in cancer patients based on patient risk profiles and molecular responses to molecular-targeted therapy. This research will not aim to cover all design aspects of intelligent techniques, or all areas of cancer. Rather, it will propose a personalized medicine support system for CML patients. It will apply the theoretical aspects to CML disease and measure its performance in a meaningful way with respect to the field of application. Finally,

this research will compare the selected models with previous methods to provide insight into personalized medicine research questions in CML.

Thesis Outline

This thesis is divided into seven chapters that constitute six journal papers. These publications are classified under the bioinformatics science domain and include two scopes for reading interests to achieve the main objectives of this research: i) the theoretical domain (computational scientists), and ii) the medical applied domain (medical scientists, such as hematologists and clinicians). The theoretical work comprises the framework proposal and the suggested solution to solve the problem from the computational point of view, while the medical application comprises details about CML background from the medical point of view and the experimental work and the results to support the proposed solution.

Chapter 2 comprises journal paper 1 and explores objective 1, which will provide the systemized literature review of intelligent techniques using molecular data analysis in leukemia, as well as identifying the opportunity for a personalized medicine support system. Chapter 3 covers the proposed framework. Chapter 3 comprises journal paper 2 and explores objective 2, which also includes three sub-objectives; it will demonstrate a framework structure for the selection of treatment in cancer patients, based on the patient risk profile and prediction of molecular response to molecular-targeted therapy. Chapter 4 comprises journal papers 3 and 4, explores objectives 3 and 4, and also addresses sub-objectives 3.1, 3.2, 4.1 and 4.2. In addition, this chapter addresses the methods used for allocating CML patients into risk groups and resolves the conflicts resulting from scoring systems. Chapter 5 comprises journal paper 5 and explores objective 5 and sub-objectives 5.1 and 5.2. This chapter emphasizes the modeling predictors and construction of clinical prediction rules that represent the relationship between predictive factors and achieving a major molecular response at 24 months. Finally, Chapter 6 comprises journal paper 6 and discusses objective

6.1. These five chapters are copies of the published journal papers, or unsubmitted work written in manuscript style. Chapter 7 summarizes the principle findings of the thesis, the limitations of the approach and suggestions for further research. Figure 2 outlines the topics of the journal papers included in the thesis.

Appendices are supplied to support the journal papers. Appendix A provides supplementary material for journal paper 4. Appendix B provides supplementary material for journal paper 5.

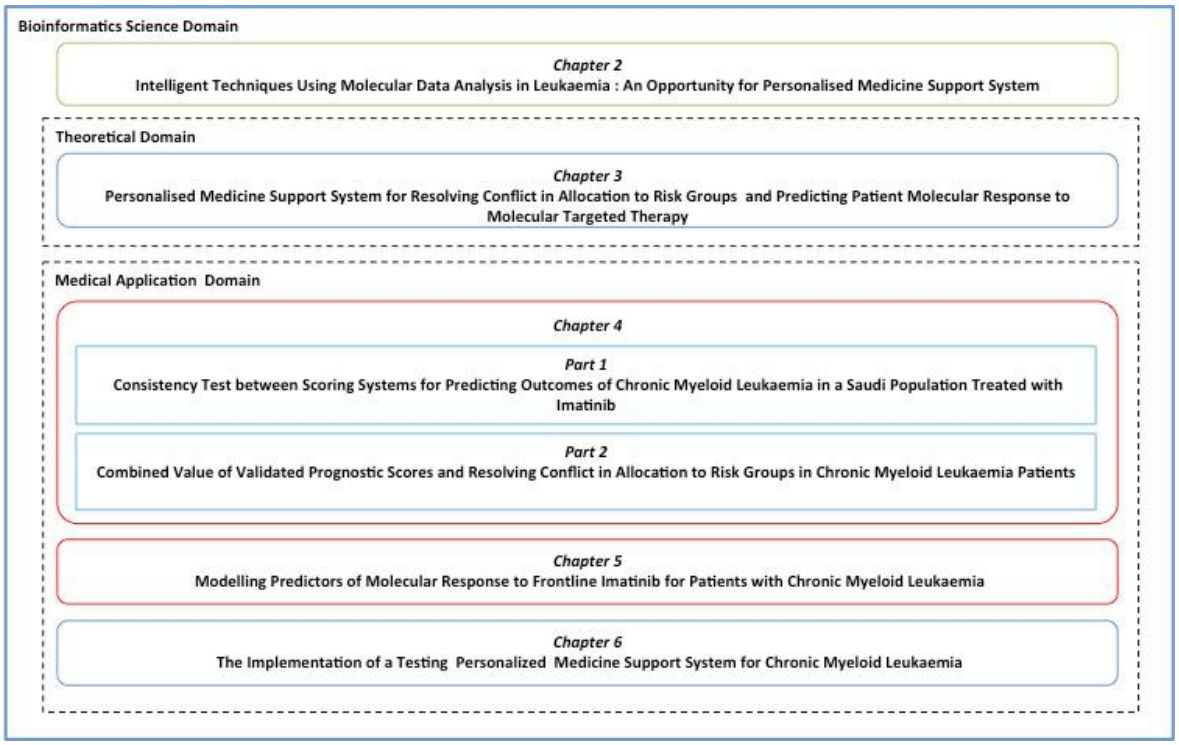


Figure 2 The scopes of the journal papers in the thesis.

References

1. Pedreira CE, Macrini L, Land MG, Costa ES. New Decision Support Tool for Treatment Intensity Choice in Childhood Acute Lymphoblastic Leukemia. IEEE Transaction on Information Technology in Biomedicine. 2009;13(3):284-90. doi: 10.1109/TITB.2008.925965.
2. Müller M. Individualized medicine. Clinical Pharmacology: Current Topics and Case Studies: Springer; 2010.
3. Apperley JF. Part I: Mechanisms of resistance to imatinib in chronic myeloid leukaemia. Oncology. 2007;8:1018-29.
4. Quintás-Cardama A, Kantarjian HM, Cortes JE. Mechanisms of Primary and Secondary Resistance to Imatinib in Chronic Myeloid Leukemia. Cancer Control. 2009;16(2):122-31.
5. Sioud M, Melien. Treatment Options and Individualized Medicine. Target Discovery and Validation Reviews and Protocols. 2007;361:327-40. doi: 10.1385/1-59745-208-4:327.
6. Kohane IS. The twin questions of personalized medicine: who are you and whom do you most resemble? BioMed Central. 2009;1(1):4.1-4.3.
7. Kraft WK, Nelson TJ, Terzic A, Waldman SA. Clinical pharmacology: a paradigm for individualized medicine. Biomarkers in Medicine. 2009;3(6):679. doi: http://dx.doi.org.proxy.library.adelaide.edu.au/10.2217/bmm.09.76.
8. Yang YT, Wiley E, Leppard J. Individualized medicine and pharmacogenomics: ethical, legal and policy challenges. Springer. 2011;9:48-57. doi: 10.1007/s12682-011-0085-7.
9. Ely S. Personalized medicine: individualized care of cancer patients. Translational Research. 2009;154(6):303-8. doi: 10.1016/j.trsl.2009.08.001.
10. Shastry B. Pharmacogenetics and the concept of individualized medicine. The Pharmacogenomics Journal 2006;6:16-21.
11. Cancer in Australia 2017 Canberra: The Australian Institute of Health and Welfare (AIHW); 2017 [2016]. Available from: <http://www.aihw.gov.au/cancer-data>.
12. Deininger MWN, Goldman JM, Melo JV. The molecular biology of chronic myeloid leukemia. Blood. 2000;96(10):3343-56.
13. Guilhot F, Guilhot J. Predicting response in CML. Blood. 2011;117(6):1773-4. doi: 10.1182/blood-2010-11-317123.
14. White DL, Hughes TP. Predicting the response of CML patients to tyrosine kinase inhibitor therapy. Current Hematologic Malignancy Reports. 2011;6(2):88-95.
15. Jabbour E, Kantarjian H, O'Brien S, Shan J, Garcia-Manero G, Wierda W, et al. Predictive factors for outcome and response in patients treated with second-generation tyrosine kinase inhibitors for chronic myeloid leukemia in chronic phase after imatinib failure. Blood. 2010;117:822-1827. doi: 10.1182/blood-2010-07-293977.
16. Deininger M, O'Brien SG, Guilhot F, Goldman JM, Hochhaus A, Hughes TP, et al. International Randomized Study of Interferon Vs STI571 (IRIS) 8-Year Follow up: Sustained Survival and Low Risk for Progression or Events in Patients with Newly Diagnosed Chronic Myeloid Leukemia in Chronic Phase (CML-CP) Treated with Imatinib. Blood (ASH Annual Meeting Abstracts). 2009;114(1126).
17. Marin D, Milojkovic D, Olavarria E, Khorashad JS, Lavallade Hd, Reid AG, et al. European LeukemiaNet criteria for failure or suboptimal response reliably identify patients with CML in early chronic phase treated with imatinib whose eventual outcome is poor. Blood. 2008;112:4437-44. doi: 10.1182/blood-2008-06-162388.
18. Shami P, Deininger M. Evolving treatment strategies for patients newly diagnosed with chronic myeloid leukemia: the role of second-generation BCR–ABL inhibitors as first-line therapy. Leukemia. 2011. doi: 10.1038/leu.2011.217.
19. Agrawal M, Garg RJ, Kantarjian H, Cortes J. Chronic Myeloid Leukemia in the Tyrosine Kinase Inhibitor Era: What Is the “Best” Therapy? Current oncology reports (1523-3790). 2010;12(5):302–13. doi: 10.1007/s11912-010-0116-1.
20. Wei G, Rafiyath S, Liu D. First-line treatment for chronic myeloid leukemia: dasatinib, nilotinib, or imatinib. Journal of Hematology & Oncology. 2010;3(1):47. doi: 10.1186/1756-8722-3-47.

21. Jabbour E, Cortes J, Kantarjian H. Treatment selection after imatinib resistance in chronic myeloid leukemia. Targeted Oncology (1776-2596). 2009;4(1):3-10. doi: 10.1007/s11523-008-0100-y.
22. Hughes TP, Hochhaus A, Saglio G, Kim D-W, Jootar S, Coutre PDl, et al. ENESTnd Update: Continued Superiority of Nilotinib Versus Imatinib In Patients with Newly Diagnosed Chronic Myeloid Leukemia In Chronic Phase (CML-CP). Blood (ASH Annual Meeting Abstracts). 2010;116(Abstract 207).
23. Kantarjian H, Shah NP, Hochhaus A, Cortes J, Shah S, Ayala M, et al. Dasatinib versus imatinib in newly diagnosed chronic-phase chronic myeloid leukemia. The New England Journal of Medicine. 2010;362(24):2260–70.
24. Erba HP. Improving Frontline Treatment for Chronic Myeloid Leukemia: Emerging Evidence for Use of Nilotinib and Dasatinib. Clinical Advances in Hematology & Oncology. 2011;9(10):734-45.
25. Baccarani M, Guilhot F, Larson RA, O'Brien SG, Druker BJ. Outcomes by cytogenetic and molecular response at 12 and 18 months of imatinib in patients with newly diagnosed chronic myeloid leukemia (CML) in chronic phase (CP) in the IRIS trial. The American Society of Hematology 48th Annual Meeting and Exposition; Orlando, FL: Blood (Abstract 2138); 2006. p. Poster Session: Chronic Myeloid Leukemia: Imatinib Therapy.
26. Alexander W, Hughes TP. American Society of Hematology, 52nd Annual Meeting and Exposition: Nilotinib (Tasigna) found superior to imatinib (Gleevec) in chronic-phase chronic myeloid leukemia: ENESTnd update. P and T. 2011;36(2):100-1.
27. Baccarani M. Response Definition, Evaluation and Monitoring. Evolving Concepts in The Management of Chronic Myeloid Leukemia 2006.
28. Clark RE. Facts and uncertainties in monitoring treatment response in chronic myeloid leukaemia. Leukemia Research 2009;33:1151–5. doi: 10.1016/j.leukres.2009.04.001.
29. National Comprehensive Cancer Network Guideline Inc.: National Comprehensive Cancer Network; [cited 2014].
30. Baccarani M, Deininger MW, Rosti G, Hochhaus A, Soverini S, Apperley JF, et al. European LeukemiaNet recommendations for the management of chronic myeloid leukemia: 2013. 2013;122(6):872-84. doi: 10.1182/blood-2013-05-501569.
31. Druker B, Guilhot F, O'Brien S. Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia. N Engl J Med. 2006;355:2408–17.
32. Press R, Galderisi C, Yang R. A half-log increase in BCR-ABL RNA predicts a higher risk of relapse in patients with chronic myeloid leukemia with an imatinib-induced complete cytogenetic response. . Clin Cancer Res. 2007;13:6136–43.
33. Wang L, Knight K, Lucas C. The role of serial BCR-ABL transcriptmonitoring in predicting the emergence of BCR-ABL kinase mutations in imatinib-treated patients with chronic myeloid leukemia. Haematologica. 2006;91:235–9.
34. Sokal J, Cox E, Baccarani M, Tura S. Prognostic discrimination in "good-risk" chronic granulocytic leukemia. Blood 1984;63:789-99.
35. Hasford J, Pfirrmann M, Hehlmann Rd, Allan NC, Baccarani M. A New Prognostic Score for Survival of Patients With Chronic Myeloid Leukemia Treated With Interferon Alfa. Jornal of the National Cancer Institute. 1998;90(11):850-8.
36. Hasford J, Baccarani M, Hoffmann V, Guilhot J, Saussele S, Rosti G, et al. Predicting complete cytogenetic response and subsequent progression-free survival in 2060 patients with CML on imatinib treatment: the EUTOS score. Blood. 2011;118(3):686-92. doi: 10.1182/blood-2010-12-319038. PubMed PMID: WOS:000292967300032.
37. Pfirrmann M, Baccarani M, Saussele S, Guilhot J, Cervantes F, Ossenkoppele G, et al. Prognosis of long-term survival considering disease-specific death in patients with chronic myeloid leukemia. Leukemia. 2016;30(1):48-56. doi: 10.1038/leu.2015.261.
38. Sweet K, Zhang L, Pinilla-Ibarz J. Biomarkers for determining the prognosis in chronic myelogenous leukemia. Journal of Hematology & Oncology. 2013;6(54).
39. Milojkovic D, Nicholson E, Apperley JF, Holyoake TL, Shepherd P, Drummond MW, et al. Early prediction of success or failure of treatment with second-generation tyrosine kinase inhibitors in patients with chronic myeloid leukemia. Haematologica. 2010;95(2):224-31.

40. Sasaki K, Kantarjian H, O'Brien S, Ravandi F, Konopleva M, Borthakur G, et al. Prediction for sustained deep molecular response of BCR-ABL1 levels in patients with chronic myeloid leukemia in chronic phase. Cancer.n/a-n/a. doi: 10.1002/cncr.31187.
41. White D, Saunders V, Lyons AB, al. e. In vitro sensitivity to imatinib-induced inhibition of ABL kinase activity is predictive of molecular response in patients with de novo CML. Blood. 2005;106:2520-6. doi: 10.1182/blood-2005-03-1103.
42. White DL, Hughes TP. Predicting the Response of CML Patients to Tyrosine Kinase Inhibitor Therapy. Current Hematologic Malignancy Reports 2009;4:59–65.
43. White DL, Dang P, Engler J, Frede A, Zrim S, Osborn M, et al. Functional Activity of the OCT-1 Protein Is Predictive of Long-Term Outcome in Patients With Chronic-Phase Chronic Myeloid Leukemia Treated With Imatinib. JOURNAL OF CLINICAL ONCOLOGY. 2010;28(16):2761-7.
44. White D, Saunders V, Grigg A, Arthur C, Filshie R, Leahy MF, et al. Measurement of In Vivo BCR-ABL Kinase Inhibition to Monitor Imatinib-Induced Target Blockade and Predict Response in Chronic Myeloid Leukemia. JOURNAL OF CLINICAL ONCOLOGY. 2007;25(28):4445-51.
45. Musen MA, Shahar Y, Shortliffe EH. Clinical Decision-Support Systems. Biomedical Informatics Computer Applications in Health Care and Biomedicine. 3rd ed. USA: Springer; 2006. p. 698-736.
46. Welch BM, Kawamoto K, Drohan B, Hughes KS. Clinical Decision Support for Personalized Medicine. In: Greenes RA, editor. Clinical decision support : the road to broad adoption: Amsterdam; 2014. p. 383-413.
47. Lyons M. Personalised medicine. Australasian BioTechnology. 2010;20(1):9-11.
48. Fiannaca A, La Rosa M, Urso A, Rizzo R, Gaglio S. A knowledge-based decision support system in bioinformatics: an application to protein complex extraction. BMC Bioinformatics. 2013;14(Suppl 1):S5-S. doi: 10.1186/1471-2105-14-S1-S5. PubMed PMID: PMC3548703.
49. Huang T, Lan L, Fang X, An P, Min J, Wang F. Promises and Challenges of Big Data Computing in Health Sciences. Big Data Research. 2015;2(1):2-11. doi: https://doi.org/10.1016/j.bdr.2015.02.002.
50. Denecke K, Spreckelsen C. Personalized medicine and the need for decision support systems. Stud Health Technol Inform. 2013;186:41-5.
51. An X, Tiwari A, Sun Y, Ding P-R, Chen Z-S. BCR-ABL tyrosine kinase inhibitors in the treatment of Philadelphia chromosome positive chronic myeloid leukemia: A review. Leukemia Research. 2010;34:1255–68. doi: 10.1016/j.leukres.2010.04.016.
52. Rochau U, Sroczynski G, Wolf D, Schmidt S, Conrads-Frank A, Jahn B, et al. Medical decision analysis for fi rst-line therapy of chronic myeloid leukemia. Leukemia & Lymphoma. 2014:1-10. doi: 10.3109/10428194.2013.858149.
53. Hasford J, Baccarani M, Hoffmann V, Guilhot J, Saussele S, Rosti G. Predicting complete cytogenetic response and subsequent progression-free survival in 2060 patients with CML on imatinib treatment: the EUTOS score. Blood. 2011;118(3):686-92. doi: 10.1182/blood-2010-12-319038.
54. Dietel M, Sers C. Personalized medicine and development of targeted therapies: the upcoming challenge for diagnostic molecular pathology. A review. Virchows Archiv. 2006;448:744-55.
55. Hunink M, Glasziou P, J S, Weeks JC, Pliskin JS, Elstein AS, et al. Decision Making in Health and Medicine: Integrating Evidence and Values2001.
56. Kumar SS, Simanti S, Kumar MP, Kumar BN, Abhigyan P, Partha G. A comparative study of Hasford score and Sokal index in prognostication of the novo chronic myeloid leukemia patients and a search for new prognostic markers2013 ; 56(3):[216-20 pp.]. Available from: http://www.ijpmonline.org/text.asp?2013/56/3/216/120369.
57. Tiribelli M, Bonifacio M, Calistri E, Binotto G, Maino E, Marin L, et al. EUTOS Score Identifies Cases with Poor Outcome in Patients with Early Chronic Phase Chronic Myeloid Leukemia, Though Not Predictive for Optimal Response to Imatinib. Blood. 2012;120(21). PubMed PMID: WOS:000314049601287.
58. Pagnano KB, Lorand-Metze I, Miranda ECM, Duarte VO, Delamain MT, Duarte GO, et al. EUTOS Score Is Predictive of Event-Free Survival, but Not for Progression-Free and Overall Survival in Patients with Early Chronic Phase Chronic Myeloid Leukemia Treated with Imatinib: A Single Institution Experience. Blood. 2012;120(21). PubMed PMID: WOS:000314049600358.

59. Breccia M, Finsinger P, Loglisci G, Latagliata R, Mancini M, Salaroli A, et al. The EUTOS score identifies chronic myeloid leukeamia patients with poor prognosis treated with imatinib first or second line. Leukemia Research. 2012;36(9):E209-E10. doi: 10.1016/j.leukres.2012.05.011. PubMed PMID: WOS:000306523600006.
60. Oyekunle AA, Osho PO, Aneke JC, Salawu L, Durosinmi MA. The predictive value of the Sokal and Hasford scoring systems in chronic myeloid leukaemia in the imatinib era. Journal of Hematological Malignancies. 2012;2(2):25-33.
61. Hoffmann VS, Baccarani M, Lindoerfer D, Castagnetti F, Turkina A, Zaritsky A, et al. The EUTOS prognostic score: review and validation in 1288 patients with CML treated frontline with imatinib. Leukemia. 2013;27(10):2016-22. doi: 10.1038/leu.2013.171. PubMed PMID: WOS:000325642600008.
62. Saussele S, Lauseker M, Hoffmann V, Proetel U, Hanfstein B, Baerlocher GM, et al. Prediction of Molecular Response of Chronic Phase CML Patients by the EUTOS Score: Results of the Randomized CML-Study IV. Blood. 2011;118(21):1606-. PubMed PMID: WOS:000299597105490.
63. Dybko J, Jaźwiec B, Haus O, Urbaniak-Kujda D, Kapelko-Słowik K, Wróbel T, et al. The Hasford Score May Predict Molecular Response in Chronic Myeloid Leukemia Patients: A Single Institution Experience. Disease Markers. 2016;2016:7531472. doi: 10.1155/2016/7531472. PubMed PMID: PMC5080519.
64. Ganguly S, Lakshmaiah KC, Jacob LA, Babu S, Dasappa L, Govind Babu KS. Performance of Sokal and Eutos Scores for Predicting Cytogenetic and Molecular Response in Newly Diagnosed Chronic Myeloid Leukemia-Chronic Phase Patients on Imatinib. Indian Journal of Hematology and Blood Transfusion. 2016:1-5. doi: 10.1007/s12288-016-0667-x.
65. Bonifacio M, Binotto G, Calistri E, Maino E, Tiribelli M. EUTOS score predicts early optimal response to imatinib according to the revised 2013 ELN recommendations. Annals of hematology. 2014;93(1):163-4.
66. Höglund M, Sandin F, Hellström K, Björeman M, Björkholm M, Brune M, et al. Tyrosine kinase inhibitor usage, treatment outcome, and prognostic scores in CML: report from the population-based Swedish CML registry2013 2013-08-15 00:00:00. 1284-92 p.
67. Pfirrmann M, Hasford J, Saussele S, Turkina A, Prejzner W, Steegmann JL, et al. The EUTOS Survival Score Is Preferable over the Sokal Score for Prognosis of Long-Term Survival of Patients with Chronic Myeloid Leukemia. Blood. 2015;126(23):595-.
68. Uz B, Buyukasik Y, Atay H, Kelkitli E, Turgut M, Bektas O, et al. EUTOS CML prognostic scoring system predicts ELN-based 'event-free survival' better than Euro/Hasford and Sokal systems in CML patients receiving front-line imatinib mesylate. Hematology. 2013;18(5):247-52. doi: 10.1179/1607845412y.0000000071. PubMed PMID: WOS:000324530700001.
69. Hoyle M, Pavey T, Ciani O, Crathorne L, Jones-Hughes T, Cooper C, et al. asatinib, nilotinib, and standard dose imatinib for the fi rst-line treatment of chronic myeloid leukaemia: systematic reviews and economic analyses. 1st-line TKIs for chronic CML: final NICE report. Exeter, UK: Peninsula Technology Assessment Group (PenTAG), University of Exeter, 2011.
70. Jain N, O'brien S. The frontline treatment of chronic myeloid leukemia in the chronic phase: current clinical decisions and future prospects for treatment. Expert Review of Hematology. 2013;6(5):575-87.
71. Henry SB. An inductive algorithm approach to knowledge acquisition for expert system development. A pilot study. Computers in nursing. 1995;13(5):226-32. Epub 1995/09/01. PubMed PMID: 7585305.
72. Hughes T, White D. Which TKI? An embarrassment of riches for chronic myeloid leukemia patients. 2013.
73. Kuncheva LI. Combining Pattern Classifiers Methods and Algorithms: John Wiley & Sons,; 2004.
74. Ball M, Weaver C, Kiel J. Overview of Clinical Decision Support Systems. In: Third, editor. Healthcare Information Management Systems. New York: Springer-Verlag. p. 463-77.
75. Alonso F, Martínez L, Pérez A, Valente JP. Cooperation between expert knowledge and data mining discovered knowledge: Lessons learned. Expert Systems with Applications. 2012;39:7524–35. doi: 10.1016/j.eswa.2012.01.133.
76. NSW Model of Care for Acute Myeloid Leukaemia. Blood and Marrow Transplant Network [Internet]. 2013. Available from: http://www.aci.health.nsw.gov.au.

77. Chronic Myeloid Leukemia. National Comprehensive Cancer Network, Evidence Blocks [Internet]. 2018. Available from: https://www.nccn.org/professionals/physician_gls/pdf/cml_blocks.pdf.
78. Kalmanti L, Saussele S, Lauseker M, Muller MC, Dietz CT, Heinrich L, et al. Safety and efficacy of imatinib in CML over a period of 10 years: data from the randomized CML-study IV. Leukemia. 2015;29(5):1123-32. Epub 2015/02/14. doi: 10.1038/leu.2015.36. PubMed PMID: 25676422.
79. NCCN Clinical Practice Guidelines in Oncology: Chronic Myelogenous Leukemia: National Comprehensive Cancer Network; 2017 [2017]. Available from: http://www.nccn.org/professionals/physician_gls/pdf/cml.pdf.

Chapter 2: Intelligent Techniques Using Molecular Data Analysis in Leukemia: An Opportunity for Personalized Medicine Support System

Statement of Authorship

Title of Paper	Intelligent Techniques Using Molecular Data Analysis in Leukaemia: An Opportunity for Personalized Medicine Support System
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Banjar, H. , Adelson, D. , Brown, F. and Chaudhri, N., "Intelligent Techniques Using Molecular Data Analysis in Leukaemia: An Opportunity for Personalized Medicine Support System", BioMed Research International, vol. 2017, Article ID 3587309, 21 pages, 2017.

Principal Author

Name of Principal Author (Candidate)	Haneen Reda Banjar		
Contribution to the Paper	HB designed and performed the research, analysed studies, and wrote the manuscript.		
Overall percentage (%)	90%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature	[redacted]	Date	2nd Dec 2017

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate in include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	David Adelson		
Contribution to the Paper	DA revised the manuscript and approved the version of the manuscript to be published.		
Signature	[redacted]	Date	4 Dec 2017

Name of Co-Author	Fred Brown		
Contribution to the Paper	FB revised the manuscript and approved the version of the manuscript to be published.		
Signature	[redacted]	Date	11/12/17

Name of Co-Author	Naeem Chaudhri		
Contribution to the Paper	NC revised the manuscript and approved the version of the manuscript to be published.		
Signature	[redacted]	Date	28/1/18

Hindawi
BioMed Research International
Volume 2017, Article ID 3587309, 21 pages
https://doi.org/10.1155/2017/3587309

Review Article

Intelligent Techniques Using Molecular Data Analysis in Leukaemia: An Opportunity for Personalized Medicine Support System

Haneen Banjar,[1,2] David Adelson,[3] Fred Brown,[1] and Naeem Chaudhri[4]

[1]*School of Computer Science, University of Adelaide, Adelaide, SA, Australia*
[2]*Department of Computer Science, King Abdulaziz University, Jeddah, Saudi Arabia*
[3]*School of Molecular and Biomedical Science, University of Adelaide, Adelaide, SA, Australia*
[4]*Oncology Centre, Section of Hematology, HSCT, King Faisal Specialist Hospital and Research Centre, Riyadh, Saudi Arabia*

Correspondence should be addressed to Haneen Banjar; hrbanjar@kau.edu.sa

Received 5 April 2017; Revised 12 June 2017; Accepted 15 June 2017; Published 25 July 2017

Academic Editor: Junya Kuroda

Copyright © 2017 Haneen Banjar et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The use of intelligent techniques in medicine has brought a ray of hope in terms of treating leukaemia patients. Personalized treatment uses patient's genetic profile to select a mode of treatment. This process makes use of molecular technology and machine learning, to determine the most suitable approach to treating a leukaemia patient. Until now, no reviews have been published from a computational perspective concerning the development of personalized medicine intelligent techniques for leukaemia patients using molecular data analysis. This review studies the published empirical research on personalized medicine in leukaemia and synthesizes findings across studies related to intelligence techniques in leukaemia, with specific attention to particular categories of these studies to help identify opportunities for further research into personalized medicine support systems in chronic myeloid leukaemia. A systematic search was carried out to identify studies using intelligence techniques in leukaemia and to categorize these studies based on leukaemia type and also the task, data source, and purpose of the studies. Most studies used molecular data analysis for personalized medicine, but future advancement for leukaemia patients requires molecular models that use advanced machine-learning methods to automate decision-making in treatment management to deliver supportive medical information to the patient in clinical practice.

1. Introduction

Molecular cytogenetics of hematological malignancies and therapies is under development. Leukaemia is a hematological disorder where two leukaemia patients who may appear identical morphologically may have different molecular profiles and thus the variation in response to the prescribed therapies would be unpredictable [1]. Leukaemia usually begins in the bone marrow, which is the site where all blood cells are formed and produced. When a person has leukaemia, the white blood cells produced are usually numerous, and they are abnormal, which means that they cannot effectively defend the body from diseases, pathogens, or foreign substances. The type of white blood cells affected, either lymphoid or myeloid, can identify the type of leukaemia. Four common types of leukaemia are chronic lymphocytic leukaemia (CLL), chronic myeloid leukaemia (CML), acute lymphocytic leukaemia (ALL), and acute myeloid leukaemia (AML) [2].

The most common modes of treatment for leukaemia involve chemotherapy, radiation therapy, stem cell transplantation, and immunotherapy with interferon [2]. Many patients become disease-free after years of treatment, proceeding to live normal lives. However, these modes of therapy can have disastrous consequences for the victims of leukaemia. For instance, chemotherapy often makes patients lose their hair, it can darken their skin, and it can cause infertility in young patients. Bone marrow transplantation is also very expensive, and it is not often easy to find a matching bone marrow donor, especially if close family members

are not a match. The pharmacogenomics (PGx) aims to replace general modes of treatment, such as chemotherapy and bone marrow transplantation, adopting instead a tailor-made course of medication designed according to a patient's particular genetic makeup [3]. Although multiple targeted therapies are available to use for leukaemia patients [4], it is difficult to select among the available targeted therapies.

Therefore, the use of intelligent techniques in medicine has brought a ray of hope in terms of treating leukaemia patients. Intelligent techniques are able to conduct automatic searches to discover knowledge and learn from data to facilitate the task and achieve the objective. The broad areas frequently defined under intelligence techniques are as follows: knowledge discovery, machine learning, and data mining. These areas use statistics and probability to detect patterns that are difficult to study manually. Intelligent techniques will integrate various molecular technologies and sources of data, information, or knowledge to facilitate the development of personalized medicine and decision-making by physicians.

The personalized decision support system requires personal information or genetic information, such as genetic tests and medical tests, for each patient to integrate, as far as possible, the knowledge gained from genomics research relating to the disease in question [5]. From this definition, it is clear that the need for personalized decision support systems in leukaemia treatment has increased because of the massive amount of available genetic and genomic information. With the use of the personalized medicine support system, leukaemia treatment will no longer be a trial and error game and it will be possible to select which drugs will work at which stage. The outcomes are expected to provide a preventive, next-generation therapy, with better specific interventions for individual patients.

Personalized medicine support systems can use available knowledge resources to deliver just-in-time information to individualize therapy. The existing pharmacogenomics knowledge base (PharmGKB) (available at http://www.pharmgkb.org) is a massive resource that provides researchers with information relevant to genetic variations on drug responses. The second source is to translate PGx knowledge into a rule-based representation where the rules are extracted from the characteristics of PGx knowledge in the US Food and Drug Administration (FDA) drug label database [6]. Another knowledge resource is the Clinical Practice Guidelines (CPG) document, which lists a set of guidelines for cases with specific diagnoses, along with recommended therapeutic action plans [7].

Developing personalized medicine support systems in some medical applications has already made significant progress. First, in cardiovascular diseases, many factors could influence cardiovascular disease, such as genes, environment, and lifestyle (exercise and nutrition). It was important to develop models for prevention, treatment management, or detecting disease to assist clinicians in treating cardiovascular patients. Indeed, the personalized decision support system for cardiovascular patients was constructed using two models. One model was for risk assessment using patients' personal information, and the other was for generating advice to clinicians based on the first model's results [8]. The second application was in type 1 diabetes mellitus. The frequent measurements of glucose levels, monitoring physical activity, and personal information about the genome, such as the genes that could cause obesity and predisposition for diabetes, were used to create a personalized decision support system for treating, preventing, and monitoring patients [9]. Another system for optimizing insulin infusion rates based on nonlinear model predictive control was designed using in silico data from a virtual type 1 diabetes patient, where the system was evaluated using data from a mathematical model of a patient with type 1 diabetes. The results showed the effectiveness of using such methods with diabetes patients [10]. The third application is colon cancer, where selection of the treatment plan was the objective of the personalized medicine support system. The clinical and genomic features were used for early diagnosis in colon cancer using cluster techniques. The platform known as MATCH supported clinicians in making decisions about patients with colon cancer [11]. Adding genetic features improves diagnosis compared to previous research methods. Further methods proposed for colon cancer prediction based on genetic information were studied by Kulkarni et al. [12]. The last example, but not limited to those mentioned in our review, breast cancer classification is an active research application in which the personalized medicine support medicine was the objective in multiple studies [13–31].

One angle of personalized medicine is to identify the correct disease subtype and patient classification. Machine-learning techniques were proven to achieve a high performance classification in identifying patient subtypes by using a support vector machine (SVM) and uncertainty SVM [32], in predicting drug sensitivity in cancer by using Bayesian networks [33, 34], in predicting patient response to therapy by using ensemble methods [13], in predicting risk category using soft computing methods [35], or in recommendations to enhance lifestyle and educate patients about healthy solutions [36]. Thus, intelligent methods should be proposed to create a personalized medicine support system in leukaemia. These methods required information about preknowledge in allocating optimal treatment, responding to each patient's risk category, which should be provided to evaluate the models.

Medical researchers continue to emphasise that their studies are updated with the most effective treatment protocols, which could be used to treat leukaemia patients. The current system for achieving personalized medicine in leukaemia has been established by using the predictive factors to determine upfront treatment. Many groups of researchers have conducted studies by using different techniques to investigate several factors that could affect the drug responses. Studying a single biomarker as a predictive substance could indicate the response pretreatment and predict the risk to the individual [37]. The other approach involves predicting the drug reaction in terms of toxicity or resistance, using an individual's genotype data and clinical data to improve the individual's care [38]. A pharmacogenetics is a field that studies the individual's response to a specific therapy based on the person's genotype information. With the human

molecular tests and the development of molecularly targeted therapy, the system for achieving personalized medicine for leukaemia has additional levels of information that needs to be processed with assistance from information technology.

According to current knowledge, many leukaemia researchers have applied intelligent techniques, but no reviewers have yet undertaken a systematized literature review from a computational perspective concerning the development of personalized medicine intelligent techniques for leukaemia patients using molecular data analysis. This review studies the published empirical research on personalized medicine in leukaemia and synthesizes findings across studies related to intelligence techniques in leukaemia, with specific attention to particular categories of these studies to help identify opportunities for further research into personalized medicine support systems in one category of leukaemia, namely, chronic myeloid leukaemia. A systematic search was carried out to identify studies using intelligence techniques in leukaemia and to categorize these studies based on leukaemia type and also the task, data source, and purpose of the studies. Our review is conducted to support health informatics and biomedical and bioinformatics in order to answer specific technical questions to help develop future research into leukaemia from a technical perspective.

2. Methods

2.1. Search Strategy. Ten databases were searched, including Scopus, PubMed, Web of Science, BIOSIS, Inspec, MEDLINE, Embase, Springer, ACM Digital Library, and IEEE Xplore. The review was restricted to English-language studies published from 2001 to October 2016 because, prior to 2001, molecular targeted therapies and molecular responses for personalized medicine were not approved by the FDA for medical treatments in leukaemia but became more popular around this time [39]. The search was limited to primary research articles. The eligibility of each study was evaluated based on the title and abstract. Only full-text articles were retrieved. The terms searched were leukaemia and leuk*, with different combinations of key words for intelligent systems and techniques (machine learning, data mining, knowledge extraction, and CDS system) and with combined keywords and/or subject headings to identify technical leukaemia articles. Since studies were screened by a single researcher and then reviewed by the coauthors, this work cannot be considered a systematic review, but it could be considered a "systematized review" [40] to demonstrate comprehensive search guidelines and an elaborative quality assessment and synthesis of research evidence.

2.2. Selection of Studies and Data Extraction. The resulting abstracts were evaluated for inclusion. Then, the full text of those identified as meeting the criteria were obtained. Studies were included in the review, if the study

(i) used molecular data from adult leukaemia patients;

(ii) used intelligence techniques to achieve the purpose of the study;

(iii) was implemented as a model for adult leukaemia patients;

(iv) was published in a peer-reviewed journal between 2001 and 2016;

(v) was published in English.

Because the intention was to review the literature to identify whether opportunities currently under clinical development are related to model analysis molecular data for personalized medicine in leukaemia, articles were excluded, if they

(1) published decision-analytic models for economic purposes;

(2) used pediatric leukaemia data;

(3) studied a nonpatient population;

(4) were not written in English;

(5) were doctoral dissertations or pilot studies;

(6) did not include the full text of the study report.

3. Results

3.1. Study Selection. In total, $n = 1{,}929$ citations were retrieved. Excluding duplicates, the search yielded $n = 1{,}747$ articles, and the initial screen of abstracts yielded $n = 745$ articles to undergo a full-text review. Ultimately, 55 studies met the eligibility criteria and underwent data extraction and analysis (Figure 1).

55 studies described 55 unique intelligent techniques (Table 2). The studies were analyzed based on the leukaemia type involved in the study, the task, the data source, and the purpose of the study.

3.2. Based on Leukaemia Type. Of the commonest leukaemia types (Figure 2) involved in previous studies, AML and ALL occupied most of the classification studies because the DNA microarray data can be downloaded online from the Cancer Program Legacy Publication Resources [41].

Some studies [42–44] distinguished ALL origin cell lines from non-ALL leukaemia origin cell lines. A study [45] demonstrated a decision support system that classifies all four types of leukaemia using principal components for feature selection and clustering. A few studies involved chronic leukaemia types to identify the molecular biomarker in CML. For example, Oehler et al. [46] used Bayesian model averaging, while Yeung et al. [47] integrated expert knowledge to predict the functional relationships in gene expression. The capture of disease pathophysiology across patient types was studied by Savvopoulos et al. [48], and temporally and spatially distributed models were built to extract knowledge from CLL patients' blood samples. In CML and CLL studies, some studies were not included in the final selection because most of these studies were prospective studies ending with a description of patient population outcomes, rather than building models or using intelligent techniques. A study of combined prognostic markers using a multivariate model for knowledge extraction is included in this study because it

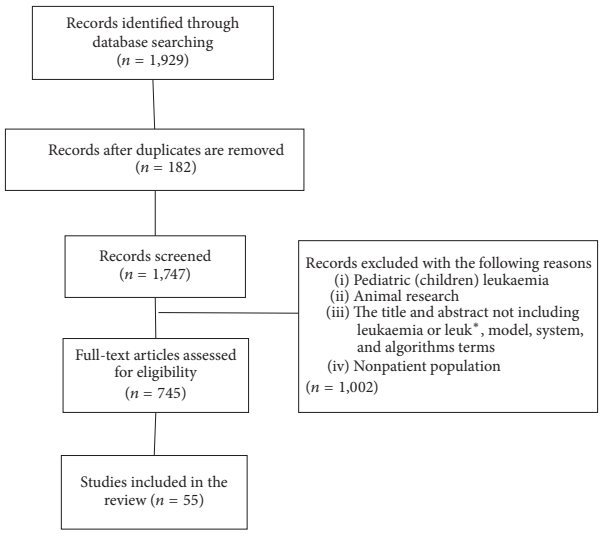


Figure 1: Flow chart showing the article-selection process.

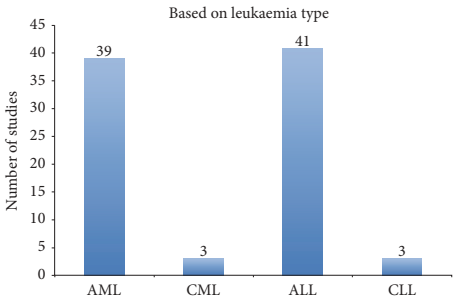


Figure 2: Summary of the frequency of studies based on leukaemia type.

proved that integration of gene expression in a model can predict outcomes in CLL [49].

Emphasis has been placed on CML as a research opportunity because of developments in monitoring CML patients' molecular response to molecular targeted therapy. The Australian Institute of Health and Welfare (AIHW) classified myeloid cancers as the 9th most commonly diagnosed cancer in 2016, with around 3,624 cases in Australia [50]. Chronic myeloid leukaemia is also known as chronic myelogenous leukaemia or chronic granulocytic leukaemia. White blood cells are affected when the bone marrow produces an unusual number of white blood cells. These cells enter the bloodstream and accumulate in organs such as the spleen or liver. If the disease progresses, the bone marrow could produce an excessive number of immature white blood cells. Consequently, the bone marrow cannot make enough red cells, normal white cells, and platelets [51].

In CML, according to White and Hughes [52], the previous studies did not establish criteria for selecting the best molecular targeted therapy for each patient, particularly following the availability of multiple therapies that can represent a perfect application of personalized therapy based on predicting patients' molecular response to molecular target therapy.

3.3. Based on Data Source. Microarray technology is an area of considerable focus for the purpose of cancer diagnosis (Figure 3). One study [45] used exon arrays to classify patients who suffer from different forms of leukaemia at various stages. Another study used human leukaemia tissue to determine different cluster differentiation (CD) markers [53]. Although seven existing studies used bone marrow cell images to classify leukaemia subtype, microarrays are still used to facilitate different types of experiments and clarify the results, making it easy for researchers to propose molecular medicines in contrast to the rate of tumor progression. The other example used gene-expression profiling among adult patients, which proved crucial in treating leukaemia. The researchers stated that, with the help of expression profiling, doctors are able to determine how a patient adapts to treatment methods. In addition, based on these facts, a physician is able to determine the survival probability of the patients in question.

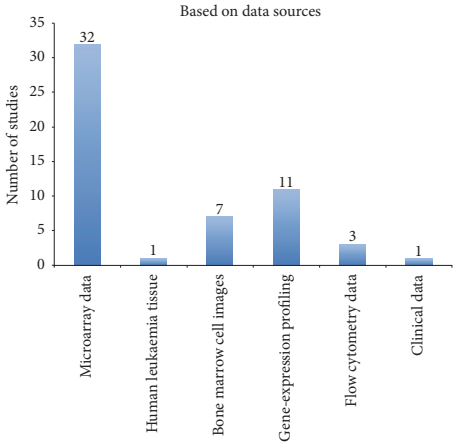


Figure 3: Summary of the frequency of studies based on data sources.

A huge opportunity arises from integrating data sources such as image data, clinical data, lifestyle or family history, SNP, gene-expression profiles, proteomics profiles, and metabolomics profiles. For example, SNPs have been investigated in an attempt to determine the susceptibility rate of patients suffering from leukaemia, which can support cases where patients have been diagnosed with leukaemia. The use of SNPs made it possible for physicians to predict the likely survivability of their patients after treatment, which is useful in determining the most suitable medical interventions.

In terms of patient care and administration, electronic health records (EHR) are often reused in research to answer specific research questions [5, 54–56]. Cases are matched with enquiries based on obtained research criteria for patient inclusion, and a dataset of many matches can then be generated for analysis. The EHR may include sparse data or missing values, as some patients may not seek frequent care. The quality of EHR would likely impact the bias of research findings or modeling performance. Derivation of key variables is also an important aspect when dealing with EHR, as the values may be recorded in different ways in different systems. This arises due to varying definitions between sources. The quality of data and correct values of derived key variables are of concern to researchers, but many algorithms can be investigated during preprocessing procedures to improve the quality of data, which would possibly lead to more reliable results [57].

Yu et al. [58] also divided the source of data and knowledge into three sources: clinical trial, systems biology, and healthcare systems. The meta-analysis and systematic reviews from published clinical trials are the main sources for gathering data and knowledge, although these methods have limitations over time. The difficulty of refining knowledge as new knowledge arises is an issue, as is the length of time required to build a knowledge base using systematic review and meta-analysis. The second source is system biology. The huge amount of data and knowledge collected as panomics for oncology patients come from genomics, transcriptomics, proteomics, and metabolomics data. For example, the Global Alliance for Genomics and Health [59] provides terabytes of genomic and clinical data for researchers. The third source is healthcare systems that provide knowledge in digital formats.

The other important source that has not attracted much interest in leukaemia studies is the data resulting from clinical trials studying healthy populations or epidemiological studies. Future development of clinical decisions can be guided by lessons learned from previous trials. Late-phase clinical trials (phases II, III, and IV) are considered to be massive sources of information that can be used to build personalized models. There is also a rapid increase in the number of electronic medical research databases that provide an opportunity for researchers to reuse medial data to create mathematical models.

The NCI [60] is a US agency that lists ongoing clinical trials that are testing molecular target therapies, including most of the studies conducted by investigators at hospitals and medical centers. The NCI offers full trial descriptions and names of principal investigators, so researchers can contact investigators and collaborate to conduct the proposed research.

The issue with the clinical trial data that it may be biased in several aspects: sampling, referral, selection, method, and clinical spectrum biases. Clinical trials may use sampling methods, sample size, and inclusion and exclusion criteria. Another aspect is referral bias where patients are referred by specialists and the data will represent preselected patients who have high prevalence of disease. Selection bias is clear when the clinical trial data includes groups based on a variety of demographics. In the method aspect, the data may be collected using different measurements, which leads to varying precision and specifications. Finally, the clinical spectrum bias represented in patient records may show other medical problems apart from the disease [61]. For instance, Saussele and Pfirrmann [62] have reported clinical trials in CML. They demonstrated several aspects that may cause challenges to reusing clinical trials. According to Saussele and Pfirrmann, the definition of "remission" varies in clinical trials by major molecular response (MMR) or complete cytogenetic response (CCR). In addition, the clinical trials used different primary endpoints such as 12 months' MMR or 12 months' CCR to judge treatment success. The American Society of Hematology (ASH) has established annual meetings to discuss medical trial outcomes. In 2006, the 48th ASH Annual Meeting [63] displayed a poster about the International Randomized Study of Interferon versus Imatinib STI571 (IRIS) trial. This study showed that Imatinib is appropriate for CML patients in the chronic phase as results indicate that patients receiving long-term (5-6 years) Imatinib therapy achieve MMR in 90% of cases [64]. For CML patients, these results show the efficacy of continuing to receive Imatinib over time. In 2010, two significant medical trials, DASISION (Dasatinib versus Imatinib Study in Treatment-Naïve CML) and ENESTnd (Evaluating Nilotinib Efficacy and Safety in Clinical Trials-Newly Diagnosed Patients), were discussed

at the 52nd Annual Meeting [65]. In DASISION [66], 259 patients were studied for their response to 100 mg Dasatinib once daily versus 260 patients who consumed 400 mg of Imatinib once daily. The MMR was lower for the Imatinib set compared to the Dasatinib set. The most important result was the safety profiles of these drugs, which were similar. The ENESTnd trial focused on the comparison of Nilotinib versus Imatinib [67]. The samples were newly diagnosed CML-CP patients. In the trial, 282 patients were given 300 mg Nilotinib twice daily, 281 patients received 400 mg Nilotinib once daily, and 283 patients received 400 mg Imatinib once daily; the median follow-up for these patients was 18 months. The cytogenetic response (CCyR) and MMR were lower in patients treated with Imatinib than in patients treated with Nilotinib. From these trial results, the debate about selecting the optimal TKI to achieve optimal patient response has been established with recognition of the need for a predictive assay that could predict the patient's response to initial treatment. The researchers also recognize the need for a medical trial to compare each TKI against the others as a first-generation treatment. The introduction of molecular targeted drugs (TKIs) has led to a dramatic improvement in the lifespan of patients affected by this condition. With the three common TKIs, Imatinib, Nilotinib, and Dasatinib currently approved to use as frontline therapy, an important question arises regarding which TKI should be prescribed. White and Hughes [52] stated that, with the lack of clear recommendations about which TKI to select, clinicians may prescribe a particular TKI based on their own preference. It is possible to extract knowledge and modeling systems using medical data and knowledge in leukaemia, but it requires advanced computational methods, such as intelligent systems. Using the available data and knowledge to construct a personalized medicine support system for leukaemia may provide massive amounts of information to use for evaluating therapies and also for potential diagnostic and prognostic markers.

3.4. Based on the Purpose of the Studies. Medical research studies have several purposes, including classification of cancer types or distinguishing healthy cells from unhealthy cells for the purpose of diagnosis, identifying markers to help in the management of treatment, and determining the prognosis of risk. Managing leukaemia patients has gained attention since a successful study by Alvey et al. [68], who developed an expert system by using a tree-structured logical program and produced over 700 clinical diagnostic rules. Even though this study focused on a specific diagnostic aspect of the system, another study [69] provided information for managing patients from registration to diagnosis and through follow-up after treatment. The study integrated history, physical examinations, and laboratory data to develop a decision support system for leukaemia diagnostics. Chae et al. [69] used profiles from patients admitted to Severance Hospital in Seodaemun District, Republic of Korea, and used data from over 490 patients to discover knowledge that helped physicians in decision-making. There is some evidence that interdisciplinary cooperation between biologists, medical scientists, computer scientists, and engineers can be

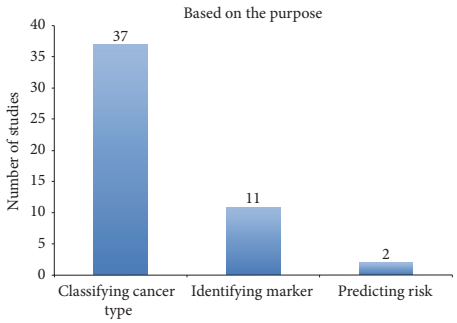


Figure 4: Summary of the frequency of studies based on the purpose of the studies.

productive. However, this research team did not incorporate molecular data in their system.

Among the 55 studies (Figure 4), the common purpose for conducting the studies was classifying the cancer type or subtype, with the aim of diagnosing leukaemia patients. An identifying marker in support of treatment management was observed in 11 studies. The other purpose for conducting the studies was using intelligent techniques and statistical methods for prognosis [70–73]. For example, predicting a relapse prior to transplantation in CML by integrating iterative Bayesian model averaging includes expert knowledge and expected functional connections in expression analyses in order to recognize genes causative of CML evolution [47]. This kind of model provides high-quality results, especially in complex diseases, but has varying levels of classification precision.

A new development is to extract relationships between biomarkers and the outcome in leukaemia patients. Focusing on CML, a predictive factor is a patient characteristic used to predict response to treatment [74]. The predictive factors related to the MMR response include common molecular assays. Other factors depend on peripheral blood counts as well as on the molecular-based and clinical observations of individual patients. In order to select the most effective TKI therapy at the time of diagnosis, various predictive factors in CML have been investigated to distinguish patients at an increased risk of failure with Imatinib, the first-generation TKI [52, 75–77]. Table 1 shows the current predictive assays and score systems, the factors included in the score systems and the methods used, the target prediction, and the published results.

Many studies [77, 129] have shown that predictive factors could probably assist in predicting patient response. Milojkovic et al. [129] conducted their study in order to predict success or failure of treatment with second-generation TKIs in CML patients using univariate analyses. They analyzed a cohort of 80 CML patients in the first chronic phase who were treated with Dasatinib or Nilotinib. Their score system predicted the probability of CML patients achieving a CCR. The system was based on three factors: cytogenetic

Table 1: The current methods used to identify risk in CML.

Previous methods				
Study	Factors	Method	Target prediction	Data and results
Sokal score, Sokal et al. [78]	Age, spleen size (cm), blast (%), and platelets (10^9/L)	Multivariate analysis of survival	Risk groups for chemotherapy	Six European and American sources (n = 813), low 39%, intermediate 38%, and high 23%
Hasford score, Hasford et al. [79]	Age, spleen size (cm), blasts (%), eosinophils (%), basophils (%), and platelets (10^9/L)	Multivariate analysis of survival	Risk groups for interferon alpha alone	14 studies (n = 981), low 40.6%, intermediate 44.7%, and high 14.6%
EUropean Treatment Outcome Study (EUTOS) Score, Hasford et al. [80]	Basophils (%) and spleen size (cm)	Multivariate analysis of response	CCgR at 18 months to Imatinib	Five national study groups (n = 2,060), low 79% and high 21%
EUTOS Long-Term survival (ELTS) score, Hoffmann et al. [81]	Age, spleen size (cm), blast (%), and platelets (10^9/L)	Multivariate analysis of response	Long-term survival	(n = 2,205) low 61%, intermediate 27%, and high 12%

response to Imatinib, Sokal score, and recurrent neutropenia during Imatinib treatment. Although this study used simple statistical methods, the system succeeded in classifying three risk categories: good, intermediate, and poor risk. In addition, Jabbour et al. [77] also studied the factors for predicting 123 CML patients' response after Imatinib failure. The variables used in this study included sex, CML duration, months' performance status, splenomegaly, prior interferon therapy, peripheral blood, bone marrow, best cytogenetic response to Imatinib, second-generation Nilotinib or Dasatinib therapy, active disease at the start of the second course of TKIs, clonal evaluation, higher than 90% Ph positivity, and IC50 for Nilotinib and Dasatinib for in vitro inhibition of kinase activity of the mutated point in *BCR-ABL*. They also used univariate and multivariate analyses, such as the logistic regression model and the Cox proportional hazard model, in order to identify prognostic factors associated with MCyR and survival, and they succeeded in identifying three risk groups: low, intermediate, and high risk.

Previously, two of the predictive factors closely involved in predicting the molecular response in CML were identified. The first such factor is IC50. In 2005, White et al. [130] studied inhibitory concentration 50% ($IC50^{imatinib}$) as a predictor of molecular response for CML patients. The results demonstrate that $IC50^{imatinib}$ is a powerful pretreatment predictor [131]. The second factor is the activity of organic cation transporter 1 (OCT-1). There are two functions for OCT proteins, which are cellular uptake and excretion of a number of exogenous and endogenous cationic and uncharged substances. The OCT-1 protein activity (OA) can be measured by uptake in the presence and absence of a specific OCT-1 inhibitor. It has been found that patients with high OA have a better molecular response than patients with low OA; therefore, OA is considered a predictive factor for response to Imatinib, but not for Nilotinib or Dasatinib [131, 132]. White et al. propose [133] that, in CML patients treated with Imatinib, the use of OA pretherapy was a predictor for long-term resistance risk and could be used to individualize dosage strategies. Thus, involving OA to estimate the response could lead to better results, but only for Imatinib therapy. A recent study has investigated the possible association between molecular response and a number of factors such as Sokal score, age, sex, and Imatinib dose [134]. It was also found that being female is a strong predictor [134]. A recent review of biomarkers that determine prognosis in CML also presented a list of prognostic indicators at diagnosis, such as the three scoring systems, *BCR-ABL1* transcript type, and OA [135]. Another factor is the *BCR-ABL* transcript type; CML patients with the b3a2 *BCR-ABL1* transcript type, compared to those with the b2a2 transcript type, demonstrate greater survival rates, while CML patients with the p190 transcript type are classified as high risk [136, 137].

In practice, clinicians aim to treat individual CML patients with the most beneficial therapy. This can be made possible by using accurate risk assessment methods at diagnosis. When there is any doubt about either the diagnosis or the recommended treatment, a second opinion is often sought before considering any treatment. The need for multiple prognostic scores can occur frequently in a complex problem that has multiple independent experts with varying expertise. When developing prognostic scores have different patient populations, each score can capture different knowledge. There are two general major objectives for combining prognostic scores: first, one prognostic score enhances the decision of another one; and second, it increases the reliability of the final decision. However, integrating multiple prognostic scores could generate conflict in decisions and may not be sufficient to make a final decision.

It is important that clinicians are comfortable with a wide range of prognostic scores that will help to identify risk category because a conflict between scores may be observed in some patients. Consistency is defined as a score that does

Table 2: Review of the studies, data sources, their purpose, and machine-learning algorithms reported from 2001 to 2015.

	Study	Year	Tasks	Data source	Leukaemia types involved in the study	Purpose	Methods
1	Cho [82]	2002	Feature selection and classification	DNA microarray	AML, ALL	Classifying leukaemia types	Pearson's and Spearman's correlation coefficients, Euclidean distance, cosine coefficient, information gain, mutual information and signal-to-noise ratio being used for feature selection
2	Inza et al. [83]	2002	Feature selection and classification	DNA microarray	AML, ALL	Classifying cancer, select genes related to cancer	Feature subset selection, case-based, and nearest neighbor classifier
3	Farag [84]	2003	Feature selection and classification	Blood cells image	AML, ALL	Classifying leukaemia types	A three-layer backpropagation neural network
4	Futschik et al. [85]	2003	Knowledge discovery	Gene expression	AML, ALL	Classifying leukaemia types and select gene expression	Knowledge-based neural networks and evolving fuzzy neural networks and adaptive learning and rule extraction
5	Cho and Won [86]	2003	Feature selection, classification, and ensemble classifiers	DNA microarray	AML, ALL	Classifying leukaemia types and select genes related to cancer	Correlation coefficient, Euclidean distance, cosine coefficient, information gain, mutual information, a feed-forward multilayer perceptron, k-nearest neighbor, self-organizing map, and support vector machine. Majority voting, weighted voting, and Bayesian approach
6	Marx et al. [44]	2003	Feature selection and classification	DNA microarray	AML, ALL	Classifying leukaemia from nonleukaemia	Principal component analysis and clustering
7	Marohnic et al. [87]	2004	Feature selection and classification	DNA microarray	AML, ALL	Classifying leukaemia types	Mutual information and support vector machine
8	McCarthy et al. [88]	2004	Knowledge extraction, classification, feature selection, visualization	Proteomic mass spectroscopy data, and gene expression	Melanoma, leukaemia	Cancer detection, diagnosis, and management	Naïve Bayes, support vector machines, instance-based learning (K-nearest neighbor), logistic regression, and neural networks
9	Rowland [89]	2004	Classification	Gene expression	AML, ALL	Classifying leukaemia types	Genetic Programming
10	Markiewicz et al. [90]	2005	Feature selection and classification	Images of different blast cell	Myelogenous leukaemia	Classifying patients	Support vector machine

Table 2: Continued.

	Study	Year	Tasks	Data source	Leukaemia types involved in the study	Purpose	Methods
11	Tung and Quek [91]	2005	Classification	DNA microarrays	ALL	Classifying leukaemia types	A neural fuzzy system, NN, SVM and the K-nearest neighbor (K-NN) classifier
12	Nguyen et al. [92]	2005	Classification	DNA microarrays	AML, ALL	Classifying leukaemia types	Support vector machine (SVM)
13	Plagianakos et al. [93]	2005	Feature selection and classification	DNA microarrays	AML, ALL	Classifying leukaemia types	artificial neural networks
14	Li and Yang [94]	2005	Feature selection and classification	DNA microarrays	AML, ALL	Classifying leukaemia types	SVM, ridge regression and Rocchio, feature selection in recursive and nonrecursive settings
15	Jinlian et al. [95]	2005	Knowledge extraction	DNA microarray	AML, ALL	Leukaemia gene association structure	Clusters
16	Diaz et al. [96]	2006	Feature selection and classification	DNA microarrays	Acute Promyelo-cytic Leukaemia	Classifying Acute Promyelocytic Leukaemia (APL) from the non-APL leukaemia	Discriminant fuzzy pattern
17	Feng and Lipo [97]	2006	Feature selection and classification	DNA microarrays	AML, ALL	Acute leukaemia types	t-statistics to rank the gene and support vector machines
18	Nguyen and Ohn [98]	2006	Feature selection and classification	DNA microarrays	AML, ALL	Classifying leukaemia types	Dynamic recursive feature elimination and random forest
19	Shulin et al. [99]	2006	Feature selection and classification	DNA microarrays	AML, ALL	Classifying leukaemia types	Independent component analysis and SVM
20	Chen et al. [100]	2007	Feature selection, rule extraction, and classification	DNA microarrays	AML, ALL	Classifying leukaemia types	A multiple kernel support vector machine

Table 2: Continued.

	Study	Year	Tasks	Data source	Leukaemia types involved in the study	Purpose	Methods
21	Ujwal et al. [43]	2007	Feature selection and classification	DNA microarray	ALL	Identifying functional cancer cell line classes, classifying leukaemia from nonleukaemia	p value and clustering
22	Perez et al. [101]	2008	Classification	Gene expression	AML, ALL	Classify leukaemia types	Hybrid fuzzy-SVM
23	Yoo and Gernaey [42]	2008	Feature selection and classification	DNA microarrays data	ALL	Classifying ALL origin cell lines from non-ALL leukaemia origin cell lines	Discriminant partial least squares, principal component and Fisher's linear discriminant analysis, linear discriminant function and SVM, and hierarchical clustering method
24	Avogadri et al. [102]	2009	Knowledge extraction	Gene expression	Myeloid leukaemia	Discovering significant clusters	Stability-based methods
25	Eisele et al. [49]	2009	Knowledge extraction	Gene expression	CLL	Prognostic markers	Multivariate model
26	Chaiboonchoe et al. [103]	2009	Classification	DNA microarrays data	ALL	Identification of differentially expressed genes	Self-organizing maps (neural networks), emergent self-organizing maps (extension of neural networks), the short-time series expression miner (STEM), and fuzzy clustering by local approximation of membership (FLAME)
27	Oehler et al. [46]	2009	Knowledge extraction	Gene expression	CML	Identifying molecular markers	Bayesian model averaging
28	Corchado et al. [45]	2009	Decision support system preprocessing, filtering, classification, and extraction of knowledge	Exon arrays	ALL, AML, CLL, CML	Classifying patients who suffer from different forms of leukaemia at various stages	Principal components, clustering, CART
29	Glez-Peña et al. [104]	2009	Feature selection and classification	DNA microarray	AML	Classifying gene expression	Fuzzy pattern algorithm
30	He and Hui [105]	2009	Classification	DNA microarray	ALL, AML	Classifying leukaemia types	Ant-based clustering (Ant-C) and an ant-based association rule mining (Ant-ARM) algorithms

Table 2: Continued.

	Study	Year	Tasks	Data source	Leukaemia types involved in the study	Purpose	Methods
31	Mukhopadhyay et al. [106]	2009	Feature selection and classification	DNA microarray	ALL, AML	Classifying leukaemia types	GA-based fuzzy clustering, neural network, and support vector machine
32	Torkaman et al. [107]	2009	Classification	Human leukaemia tissue	ALL, AML	Determining different CD markers	Cooperative game
33	Zheng et al. [108]	2009	Feature selection	DNA microarray	ALL	Gene ranking	Knowledge-oriented gene selection
34	Mehdi et al. [109]	2009	Knowledge acquisition	Gene expression	ALL, AML	Pattern clustering	K-nearest neighbor technique
35	Porzelius et al. [110]	2011	Feature selection, classification	Microarray and clinical data	ALL	Risk prediction	Feature selection approach for support vector machines as well as a boosting approach for regression models
36	Chen et al. [111]	2011	Feature selection, data fusion, class prediction, decision rule extraction, associated rule extraction, and subclass discovery	DNA microarray	ALL, AML	Select gene, classify leukaemia types, rule extraction	Multiple kernel SVM
37	Gonzalez et al. [112]	2011	Classification	Bone marrow cells images	ALL, AML	Classifying leukaemia subtypes	Segmentation method to obtain leukaemia cells and extract from them descriptive characteristics (geometrical, texture, statistical) and eigenvalues
38	Tong and Schierz [113]	2011	Feature selection and classification	DNA microarray	ALL, AML	Classifying two-class oligonucleotide microarray data for acute leukaemia	Hybrid genetic algorithm-neural network
39	Chauhan et al. [114]	2012	Classification	Genotype	ALL, AML	Identifying gene-gene interaction	Classification and regression tree
40	Escalante et al. [115]	2012	Feature selection and classification	The morphological properties of bone marrow images	ALL, AML	Classifying leukaemia subtypes	Ensemble particle swarm model selection

Table 2: Continued.

	Study	Year	Tasks	Data source	Leukaemia types involved in the study	Purpose	Methods
41	Yeung et al. [116]	2012	Feature selection and classification	Gene expression	CML	select gene, and predicted functional relationships	Integrating gene expression data with expert knowledge and predicted functional relationships using iterative Bayesian model averaging
42	Manninen et al. [117]	2013	Classification	Flow cytometry data	AML	Prediction method for diagnosis of AML	Sparse logistic regression
43	El-Nasser et al. [118]	2014	Classification	DNA microarrays	ALL, AML	Classifying leukaemia types	Implement enhanced classification (ECA) algorithm, SMIG module, and ranking procedure.
44	Singhal and Singh [119]	2015	Feature selection and classification	Image based analysis of bone marrow samples	ALL	Classifying leukaemia subtypes	Multilayer perceptron (MLP), linear vector quantization (LVQ), k-nearest neighbor (k-NN), and SVM
45	Yao et al. [120]	2015	Feature selection and classification	DNA microarrays	ALL, AML, the mixed-lineage leukaemia (MLL) data	Classifying leukaemia subtypes	Random forests and ranking features
46	Rawat et al. [121]	2015	Computer-aided diagnostic system, feature selection, and classification	Bone marrow cells in microscopic images	ALL	Diagnosis lymphoblast cells from healthy lymphocytes	Support vector machine
47	Kar et al. [122]	2015	Feature selection and classification	DNA microarrays	ALL, AML, the mixed-lineage leukaemia (MLL) data	Classifying leukaemia subtypes	Particle swarm optimization (PSO) method along with adaptive K-nearest neighborhood (KNN)
48	Li et al. [123]	2016	Classification	Gene expression	AML	Identifying feature genes	Support vector machine (SVM) and random forest (RF)
49	Dwivedi et al. [124]	2016	Classification	Microarray gene expression	ALL, AML	Classifying leukaemia subtypes	Artificial neural network (ANN)
50	Krappe et al. [125]	2016	Classification	Image based analysis of bone marrow samples	Leukaemia	Diagnosis of leukaemia and classifying 16 different classes for bone marrow	Knowledge-based hierarchical tree classifier

Table 2: Continued.

	Study	Year	Tasks	Data source	Leukaemia types involved in the study	Purpose	Methods
51	Li et al. [123]	2016	Classification	DNA microarrays	AML, ALL	Classifying leukaemia subtypes	A weighted doubly regularized support vector machine
52	Ocampo-Vega et al. [126]	2016	Feature selection and classification	DNA microarrays	AML, ALL	Classifying leukaemia subtypes	Principal component analysis and logistic regression
53	Rajwa et al. [127]	2016	Classification	Flow cytometry data	AML	Determining progression of the disease	Nonparametric Bayesian framework
54	Ni et al. [128]	2016	Classification	Flow cytometry data	AML	Analyzing minimal residual disease	Support vector machines (SVM)
55	Savvopoulos et al. [48]	2016	Knowledge extraction	CLL cells in peripheral blood	CLL	Capturing disease pathophysiology across patient types	Temporally and spatially distributed model

not contradict other prognostic scores. Consistency among prognostic scores can increase clinicians' trust, as they rely on such results to make appropriate treatment decisions. It is important to study and understand the consistency of scores to help clinicians categorize patients into suitable risk groups and subsequently make better therapeutic decisions.

In light of the aforementioned aspects, it is necessary to conduct a study that can contribute to the CML medical field by solving the previous issues. Using machine-learning techniques and fusion techniques to address these problems could produce promising results. The first proposed solution is to build a personalized medicine support system as a predictive model to combine strong molecular, clinical data, and predictive assays for CML patients that could probably predict an individual molecular response. Moreover, predicting an individual response leads to predicting warning groups for each TKI. From a computer-science perspective, the above issues could be resolved by using a machine-learning algorithm that combines the most effective predictive indicators to predict the outcomes for each TKI, based on existing clinical profiles for individual CML patient characteristics. The main goal of this review is to improve the ability to manage CML disease in individual CML patients. Therefore, CML is an example of a research opportunity to predict the molecular response to TKI treatment. Using intelligent computing techniques could bring about promising results for CML patients.

3.5. Based on the Task. Most of the studies that used machine learning and data mining incorporated two major tasks: feature selection and classification (Figure 5). Although dissimilar feature-selection algorithms may possibly choose dissimilar pertinent genes or diverse numbers of relevant genes or bring about different levels of classification precision [138], feature selection can utilize observations and functions to obtain dimensions for exploring optimal solutions [139]. In addition, selection of a subset in a classifier reduces the computational time and costs of study, thereby increasing classification accuracy.

Many studies [53, 89, 91, 103, 105, 112, 118] used classification algorithms without feature-selection techniques. Since cancer tumors are highly diverse in their genetic patterns and progressions, DNA arrays provide a platform to obtain the best measurements and observations, helping assign objectives to one relevant feature set and hence contributing to precise convergence toward optimal results [140]. However, some researchers [42–44, 82, 86, 87, 94, 96–99, 104, 108, 113, 115, 122, 141, 142] applied two common methods in feature selection: filter and wrapper approaches with classification algorithms. Considering feature-selection techniques is an essential preprocessing method mandated for classification processes [103].

Knowledge extraction or acquisition has been a great challenge for researchers, as they exhibit unusual characteristics in many different genes relative to the number of tumor samples. AML acquires a similar appearance to ALL, which makes it nearly impossible for researchers to distinguish between synonymous patterns. However, Cho et al. [143] proposed an approach to form the optimal linear classifier

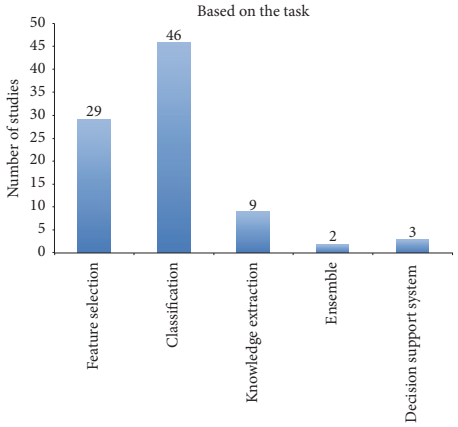


Figure 5: Summary of the frequency of studies based on the task.

by means of gene-expression data. They used discriminant partial least squares and linear discriminant analysis to differentiate between acute leukaemia subtypes. They found that these methods offered a satisfactory level of precision. They concluded that the suggested method builds the optimal classifier made up of a highly accurate, small-size predictor.

Using multiple algorithms for knowledge extraction and classification has not attracted much interest from leukaemia researchers in previous studies [45, 140, 144]. Cho and Won [86] were of the view that conventional machine learning is incapable of delivering accurate information. For this reason, they developed a novel ensemble machine-learning approach for microarray classification. Results indicated that the ensemble machine-learning approaches accuracy of almost 97% in leukaemia classification, which makes it a better alternative to basic machine-learning methods.

Among the 55 studies, three groups of researchers [45, 121, 140] built decision support systems, whose subfunctions included multiple tasks. The first decision support system [45] was built to support leukaemia diagnosis using exon array analysis. The system combined intelligent techniques, such as preprocessing and data-filtering techniques, clustering for classifying patients, and extraction of knowledge techniques. The authors suggested that further study of bone marrow or blood samples may assist in diagnosis of leukaemia stages. The second study [140] was conducted to extract decision rules using a developed SVM. This study comprised carrying out multiple tasks, including data fusion, feature selection, making a prediction model based on gene-expression data, and knowledge extraction. The third study [121] involved developing decision support to identify unhealthy ALL cells using feature-selection techniques with SVM.

From the review of studies based on the task, the need for personalized medicine in CML results in multiple active TKI therapies as molecular targeted therapy available for CML, multiple strategies utilized for frontline CML therapy,

heterogeneity in responses, and multiple prognostic scores and predictive assays.

Therapy takes the form of two major strategies: (i) frontline Imatinib or (ii) frontline second-generation TKIs such as Nilotinib or Dasatinib [4]. Despite the remarkable increase in the survival of CML patients treated with Imatinib, some patients discontinue Imatinib therapy due to intolerance, resistance, or progression. In the IRIS [64] trial, it is demonstrated that variations in molecular response at 12 and 18 months of Imatinib was due to, in about 40% of cases, discontinuing Imatinib because of intolerance or resistance and due to further progression observed in 7% of CML patients.

Hematologic, cytogenetic, and molecular strategies for monitoring patient responses to therapies are used by European LeukaemiaNet [145]. To monitor molecular response, RQ-PCR, which is a sensitive technique, is used to quantify the level of *BCR-ABL1* mRNA transcripts in the peripheral blood of patients. Molecular monitoring is considered to be a standard guide to clinical management in CML [146, 147]. The prediction of long-term molecular response to frontline Imatinib in CML can help clinicians to select the best treatment protocols for CML patients. Patients predicted not to achieve MMR in the long term might be better treated with alternative frontline therapies, such as Nilotinib or Dasatinib. Opportunities to improve individual care for CML patients exist in the appropriate prediction of variation in treatment response to support physicians in treatment decisions.

Prognostic scores are used to personalize CML patient care by predicting responses to therapy. Although the prognostic scores (Sokal, Hasford, EUTOS, and the ELTS scores) remain in use today, they were developed either for identifying risk groups or for predicting cytogenetic response to therapy, but not for molecular response. Although two predictive assays, $IC50^{imatinib}$ and OCT-1 activity (OA), were developed to predict molecular response, according to current knowledge, a model using assays to predict molecular response has not previously been considered. The combination of predictive assays results in greater predictive power than that which each predictor provides alone [148, 149].

4. Conclusion

Modern oncology is experiencing a paradigm shift toward personalized medicine, which aims to direct medical agents toward the tumor site. The field of molecular medicine is also undergoing transformational changes that are bringing a much needed revolution in healthcare. This breakthrough was made possible by technologies in genetic studies that led to the sequencing of the human genome. An analysis of biological samples from whole organisms has now been made possible. In addition, this invention has given a new lease of life to the treatment of cancer. However, the majority of cancer patients have been shown to develop adverse drug reactions due to overreliance on certain medications.

Intelligent techniques may be useful for clinicians in decision-making, warning of specific problems or providing treatment recommendations [150]. In that regard, it would be worthwhile building personalized medicine support system to work as predictive models that integrate molecular-based data to predict cancer susceptibility, including risk assessment, prediction of the probability of developing a type of cancer prior to occurrence of the disease, prediction of recurring cancer, and the prediction of cancer outcomes, such as survivability, life expectancy, and response to therapy or progression. This is highly advantageous since only a quarter of cancer patients respond positively to the drugs prescribed to them. Therefore, it is important to investigate the current development of using molecular information in intelligent models for personalized medicine.

The use of personalized medicine support systems in medicine will bring a ray of hope to the treatment of leukaemia. Other frontiers of personalized medicine research, such as the role of genetics in infectious diseases, proteomics, epigenetics, and metabolomics, were not covered by this review and are out of scope of this research. This review was conducted based on current developments of personalized medicine support systems, and a systematized literature review was carried out on intelligent techniques using molecular data analysis in leukaemia. Both sets of literature led to identifying opportunities for further research for personalized medicine support systems in one category of leukaemia, namely, chronic myeloid leukaemia. We speculate that this paper will assist health informatics and biomedical and bioinformatics in order to answer specific technical questions to help develop future research into leukaemia from a technical perspective.

Disclosure

The funding agreement ensured the authors' independence in publishing the report.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Haneen Banjar designed and performed the research, analyzed data, and wrote the manuscript. All the listed authors contributed substantially to drafts and revisions to the manuscript and approved the current revised version.

Acknowledgments

Financial support for this study was provided in part by a grant from King Abdulaziz University.

References

[1] K. K. Jain, *Basics of Personalized Medicine Textbook of Personalized Medicine*, Springer, New York, NY, USA, 2009.

[2] S. J. Swierzewski, Leukemia Basics: Remedy Health Media; 1999, http://www.healthcommunities.com/leukemia/types.shtml.

[3] G. S. Ginsburg and H. F. Willard, "Genomic and personalized medicine: foundations and applications," *Translational Research*, vol. 154, no. 6, pp. 277–287, 2009.

[4] J. Cortes and H. Kantarjian, "How I treat newly diagnosed chronic phase CML," *Blood*, vol. 120, no. 7, pp. 1390–1397, 2012.

[5] I. Kouris, C. Tsirmpas, S. G. Mougiakakou, D. Iliopoulou, and D. Koutsouris, "E-health towards ecumenical framework for personalized medicine via Decision Support System," in *Proceedings of the 2010 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC'10*, pp. 2881–2885, September 2010.

[6] C. L. Overby, P. Tarczy-Hornoch, I. J. Kalet et al., "Developing a prototype system for integrating pharmacogenomics findings into clinical practice," *Journal of Personalized Medicine*, vol. 2, no. 4, pp. 241–256, 2012.

[7] N. Douali and M.-C. Jaulent, "Genomic and personalized medicine decision support system," in *Proceedings of the 1st International Conference on Complex Systems, ICCS 2012*, November 2012.

[8] S. G. Mougiakakou, I. K. Valavanis, G. Karkalis, S. Marinos, K. A. Grimaldi, and K. S. Nikita, "An integrated web-based platform for the provision of personalized advice in people at high risk for CVD," in *Proceedings of the 9th International Conference on Information Technology and Applications in Biomedicine, ITAB 2009*, November 2009.

[9] D. C. Klonoff, "Personalized medicine for diabetes," *Journal of Diabetes Science and Technology*, vol. 2, no. 3, pp. 335–341, 2008.

[10] K. Zarkogianni, S. G. Mougiakakou, A. Prountzou, A. Vazeou, C. S. Bartsocas, and K. S. Nikita, "An insulin infusion advisory system for type 1 diabetes patients based on non-linear model predictive control methods," in *Proceedings of the Engineering in Medicine and Biology Society, 2007 EMBS 2007 29th Annual International Conference of the IEEE*, pp. 22–26, 2007.

[11] Y. Goletsis, T. P. Exarchos, N. Giannakeas, and D. I. Fotiadis, "Intelligent patient profiling for diagnosis, staging and treatment selection in colon cancer," in *Proceedings of the 8th IEEE International Conference on BioInformatics and BioEngineering, BIBE 2008*, October 2008.

[12] A. Kulkarni, B. S. C. Naveen Kumar, V. Ravi, and U. S. Murthy, "Colon cancer prediction with genetics profiles using evolutionary techniques," *Expert Systems with Applications*, vol. 38, no. 3, pp. 2752–2757, 2011.

[13] H. Moon, H. Ahn, R. L. Kodell, S. Baek, C.-J. Lin, and J. J. Chen, "Ensemble methods for classification of patients for personalized medicine with high-dimensional data," *Artificial Intelligence in Medicine*, vol. 41, no. 3, pp. 197–207, 2007.

[14] V. J. Weigman Jr., *High Level Integration of Genomic Data for Improving Prediction, Prognostication and Classification of Breast Tumors. Proquest Dissertations and Theses*, University of North Carolina, Chapel Hill, NC, USA, 2010.

[15] M. Ture, F. Tokatli, and I. Kurt Omurlu, "The comparisons of prognostic indexes using data mining techniques and Cox regression analysis in the breast cancer data," *Expert Systems with Applications*, vol. 36, no. 4, pp. 8247–8254, 2009.

[16] M. U. Khan, J. P. Choi, H. Shin, and M. Kim, "Predicting breast cancer survivability using fuzzy decision trees for personalized healthcare," in *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS'08*, pp. 5148–5151, August 2008.

[17] S. Lekkas and L. Mikhailov, *Breast Cancer Diagnosis Based on Evolvable Fuzzy Classifiers and Feature Selection. Applications and Innovations in Intelligent Systems XVI*, Springer, London, UK, 2009.

[18] D. Lavanya and K. Usha Rani, "Ensemble Decision Making System for Breast Cancer Data," *International Journal of Computer Applications*, vol. 51, no. 17, pp. 19–23, 2012.

[19] D. Lavanya and KU. Raniv, "Analysis of feature selection with classfication: breast cancer datasets," *Indian Journal of Computer Science and Engineering (IJCSE)*, vol. 2, no. 5, p. 763, 2011.

[20] A. Keleş, A. Keleş, and U. Yavuz, "Expert system based on neuro-fuzzy rules for diagnosis breast cancer," *Expert Systems with Applications*, vol. 38, no. 5, pp. 5719–5726, 2011.

[21] B. Jahn, N. Muhlberger, J. Wurm, and U. Siebert, "Decision-analytic models for breast cancer: Do currently published models meet the requirements of personalized medicine?" in *Proceedings of the 2008 Winter Simulation Conference (WSC)*, pp. 2946-2946, Miami, FL, USA, December 2008.

[22] J. M. Garibaldi, S.-M. Zhou, X.-Y. Wang, R. I. John, and I. O. Ellis, "Incorporation of expert variability into breast cancer treatment recommendation in designing clinical protocol guided fuzzy rule system models," *Journal of Biomedical Informatics*, vol. 45, no. 3, pp. 447–459, 2012.

[23] D. Delen, G. Walker, and A. Kadam, "Predicting breast cancer survivability: a comparison of three data mining methods," *Artificial Intelligence in Medicine*, vol. 34, no. 2, pp. 113–127, 2005.

[24] C. Chen, S.-R. Sun, Y.-P. Gong et al., "Quantum dots-based molecular classification of breast cancer by quantitative spectroanalysis of hormone receptors and HER2," *Biomaterials*, vol. 32, no. 30, pp. 7592–7599, 2011.

[25] M. Arora and D. Tagra, "Neuro-fuzzy expert system for breast cancer diagnosis," in *Proceedings of the 2012 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2012*, pp. 979–985, August 2012.

[26] A. Ali, A. Tufail, U. Khan, and M. Kim, "A survey of prediction models for breast cancer survivability," in *Proceedings of the 2nd International Conference on Interaction Sciences: Information Technology, Culture and Human, ICIS 2009*, pp. 1259–1262, November 2009.

[27] A. Ali, U. Khan, A. Tufail, and M. Kim, "Analyzing potential of SVM based classifiers for intelligent and less invasive breast cancer prognosis," in *Proceedings of the 2nd International Conference on Computer Engineering and Applications, ICCEA 2010*, pp. 313–319, March 2010.

[28] E. A. Rakha, D. Soria, A. R. Green et al., "Nottingham prognostic index plus (NPI+): A modern clinical decision making tool in breast cancer," *British Journal of Cancer*, vol. 110, no. 7, pp. 1688–1697, 2014.

[29] K. Nagasaki and Y. Miki, "Molecular prediction of the therapeutic response to neoadjuvant chemotherapy in breast cancer," *Breast Cancer*, vol. 15, no. 2, pp. 117–120, 2008.

[30] T. Kempowsky-Hamon, C. Valle, M. Lacroix-Triki et al., "Fuzzy logic selection as a new reliable tool to identify molecular grade signatures in breast cancer—the INNODIAG study," *BMC Medical Genomics*, vol. 8, no. 1, article 3, 2015.

[31] Arsene CTC, Lisboa PJG. Chapter 8 - Artificial Neural Networks Used in the Survival Analysis of Breast Cancer Patients: A Node-Negative Study. In: Outcome Prediction in Cancer [Internet]. Amsterdam: Elsevier; [191-239], http://www.sciencedirect.com/science/article/pii/B9780444528551500106.

[32] C. Voichita, *Towards Personalized Medicine Using Systems Biology and Machine Learning*, Wayne State University Dissertations, Detroit, MI, USA, 2013.

[33] D.-C. Kim, X. Wang, C.-R. Yang, and J. X. Gao, "A framework for personalized medicine: prediction of drug sensitivity in cancer by proteomic profiling," *Proteome Science*, vol. 10, no. 1, 2012.

[34] D.-C. Kim, J. Gao, X. Wang, and C.-R. Yang, "A framework for personalized medicine with reverse phase protein array and drug sensitivity," in *Proceedings of the 2011 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2011*, pp. 426–429, November 2011.

[35] P. K. Anooj, "Clinical decision support system: risk level prediction of heart disease using weighted fuzzy rules," *Journal of King Saud University—Computer and Information Sciences*, vol. 24, no. 1, pp. 27–40, 2012.

[36] A. M. Honka, M. J. van Gils, and J. Parkka, "A personalized approach for predicting the effect of aerobic exercise on blood pressure using a fuzzy inference system," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '11)*, pp. 8299–8302, Boston, Mass, USA, August 2011.

[37] H. Michael and M. Crawford, "Individualized medicine by biomarkers," *Clinical Cardiology Alert*, vol. 302, pp. 49–57, 2009.

[38] B. S. Shastry, "Pharmacogenetics and the concept of individualized medicine," *Pharmacogenomics Journal*, vol. 6, no. 1, pp. 16–21, 2006.

[39] GLIVEC Tablets – Summary of Product Characteristics (SPC): electronic Medicines Compendium; 2001, http://www.medicines.org.uk/emc/medicine/15014/SPC/GLIVEC+Tablets/.

[40] M. J. Grant and A. Booth, "A typology of reviews: an analysis of 14 review types and associated methodologies," *Health Information & Libraries Journal*, vol. 26, no. 2, pp. 91–108, 2009.

[41] Cancer Program Legacy Publication Resources, http://www.broadinstitute.org/cgi-bin/cancer/datasets.cgi.

[42] C. K. Yoo and K. V. Gernaey, "Classification and diagnostic output prediction of cancer using gene expression profiling and supervised machine learning algorithms," *Journal of Chemical Engineering of Japan*, vol. 41, no. 9, pp. 898–914, 2008.

[43] M. L. Ujwal, P. Hoffman, and K. A. Marx, "A machine learning approach to pharmacological profiling of the quinone scaffold in the NCI database: A compound class enriched in those effective against melanoma and leukemia cell lines," in *Proceedings of the 7th IEEE International Conference on Bioinformatics and Bioengineering, BIBE*, pp. 456–463, January 2007.

[44] K. A. Marx, P. O'Neil, P. Hoffman, and M. L. Ujwal, "Data Mining the NCI Cancer Cell Line Compound GI50 Values: Identifying Quinone Subtypes Effective Against Melanoma and Leukemia Cell Classes," *Journal of Chemical Information and Computer Sciences*, vol. 43, no. 5, pp. 1652–1667, 2003.

[45] J. M. Corchado, J. F. de Paz, S. Rodríguez, and J. Bajo, "Model of experts for decision support in the diagnosis of leukemia patients," *Artificial Intelligence in Medicine*, vol. 46, no. 3, pp. 179–200, 2009.

[46] V. G. Oehler, Y. Y. Ka, Y. E. Choi, R. E. Bumgarner, A. E. Raftery, and J. P. Radich, "The derivation of diagnostic markers of chronic myeloid leukemia progression from microarray data," *Blood*, vol. 114, no. 15, pp. 3292–3298, 2009.

[47] K. Y. Yeung, T. A. Gooley, A. Zhang, A. E. Raftery, J. P. Radich, and V. G. Oehler, "Predicting relapse prior to transplantation in chronic myeloid leukemia by integrating expert knowledge and expression data," *Bioinformatics*, vol. 28, no. 6, Article ID bts059, pp. 823–830, 2012.

[48] S. Savvopoulos, R. Misener, N. Panoskaltsis, E. N. Pistikopoulos, and A. Mantalaris, "A Personalized Framework for Dynamic Modeling of Disease Trajectories in Chronic Lymphocytic Leukemia," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 11, pp. 2396–2404, 2016.

[49] L. Eisele, R. Prinz, L. Klein-Hitpass et al., "Combined PER2 and CRY1 expression predicts outcome in chronic lymphocytic leukemia," *European Journal of Haematology*, vol. 83, no. 4, pp. 320–327, 2009.

[50] Leukaemia in Australia: The Australian Institute of Health and Welfare, http://www.aihw.gov.au/cancer/leukaemia/.

[51] J. F. Apperley, "Part I: mechanisms of resistance to imatinib in chronic myeloid leukaemia," *The Lancet Oncology*, vol. 8, no. 11, pp. 1018–1029, 2007.

[52] D. L. White and T. P. Hughes, "Predicting the response of CML patients to tyrosine kinase inhibitor therapy," *Current Hematologic Malignancy Reports*, vol. 6, no. 2, pp. 88–95, 2011.

[53] A. Torkaman, N. M. Charkari, and M. Aghaeipour, "A new classification approach based on cooperative game," in *Proceedings of the 2009 14th International CSI Computer Conference, CSICC 2009*, pp. 458–463, October 2009.

[54] B. Welch and K. Kawamoto, "The Need for Clinical Decision Support Integrated with the Electronic Health Record for the Clinical Application of Whole Genome Sequencing Information," *Journal of Personalized Medicine*, vol. 3, no. 4, pp. 306–325, 2013.

[55] J. C. Weiss, S. Natarajan, P. L. Peissig, C. A. McCarty, and D. Page, "Machine learning for personalized medicine: Predicting primary myocardial infarction from electronic health records," *AI Magazine*, vol. 33, no. 4, pp. 33–45, 2012.

[56] CL. Overby, *A Clinical Decision Support Model for Incorporating Pharmacogenomics Knowledge Into Electronic Health Records for Drug Therapy Individualization: A Microcosm of Personalized Medicine*, University of Washington, Seattle, WA, USA, 2011.

[57] B. Hirsch and A. Abernethy, "Structured Decision-Making: Using Personalized Medicine to Improve the Value of Cancer Care," *Journal of Personalized Medicine*, vol. 3, no. 1, pp. 1–13, 2013.

[58] H. Yu, G. Gu, H. Liu, and J. Shen, "A framework for microarray data-based tumor diagnostic system with improving performance incrementally," *Expert Systems with Applications*, vol. 37, no. 9, pp. 6682–6688, 2010.

[59] Global Alliance for Genomics and Health, http://genomicsandhealth.org.

[60] National Institutes of Health, U.S.: Turning Discovery Into Health, https://www.nih.gov.

[61] M. Kwiatkowska, A. S. Atkins, N. T. Ayas, and C. F. Ryan, "Integrating knowledge-driven and data-driven approaches for the derivation of clinical prediction rules," in *Proceedings of the ICMLA 2005: 4th International Conference on Machine Learning and Applications*, pp. 171–176, December 2005.

[62] S. Saussele and M. Pfirrmann, "Clinical trials in chronic myeloid leukemia," *Current Hematologic Malignancy Reports*, vol. 7, no. 2, pp. 109–115, 2012.

[63] M. Baccarani, F. Guilhot, R. A. Larson, S. G. O'Brien, and B. J. Druker, "Outcomes by cytogenetic and molecular response at 12 and 18 months of imatinib in patients with newly diagnosed chronic myeloid leukemia (CML) in chronic phase (CP) in the IRIS tria," in *Proceedings of the The American Society of Hematology 48th Annual Meeting and Exposition*, vol. 2006, Orlando, FL, USA.

[64] M. Deininger, SG. O'Brien, F. Guilhot et al., "International Randomized Study of Interferon Vs STI571 (IRIS) 8-Year

Follow up: Sustained Survival and Low Risk for Progression or Events in Patients with Newly Diagnosed Chronic Myeloid Leukemia in Chronic Phase (CML-CP) Treated with Imatinib," in *Proceedings of the Blood (ASH Annual Meeting Abstracts)*, 2009.

[65] W. Alexander and T. P. Hughes, "American Society of Hematology, 52nd Annual Meeting and Exposition: Nilotinib (Tasigna) found superior to imatinib (Gleevec) in chronic-phase chronic myeloid leukemia: ENESTnd update," *P and T*, vol. 36, no. 2, pp. 100-101, 2011.

[66] H. Kantarjian, N. P. Shah, A. Hochhaus et al., "Dasatinib versus imatinib in newly diagnosed chronic-phase chronic myeloid leukemia," *The New England Journal of Medicine*, vol. 362, no. 24, pp. 2260–2270, 2010.

[67] T. P. Hughes, A. Hochhaus, G. Saglio et al., "ENESTnd Update: Continued Superiority of Nilotinib Versus Imatinib In Patients with Newly Diagnosed Chronic Myeloid Leukemia In Chronic Phase (CML-CP)," in *Proceedings of the Blood (ASH Annual Meeting Abstracts)*, 2010.

[68] P. L. Alvey, N. J. Preston, and M. F. Greaves, "High performance for expert systems: II. A system for leukaemia diagnosis," *Informatics for Health and Social Care*, vol. 12, no. 2, pp. 97–114, 1987.

[69] Y. M. Chae, Q. Park, K. S. Park, and M. Young, "Development of medical decision support system for leukemia management," *Expert Systems with Applications*, vol. 15, no. 3-4, pp. 309–315, 1998.

[70] J. Hasford, M. Pfirrmann, R. Hehlmann et al., "Prognosis and prognostic factors for patients with chronic myeloid leukemia: Nontransplant therapy," *Seminars in Hematology*, vol. 40, no. 1, pp. 4–12, 2003.

[71] N. Mašić, A. Gagro, S. Rabatić et al., "Decision-tree approach to the immunophenotype-based prognosis of the B- cell chronic lymphocytic leukemia," *American Journal of Hematology*, vol. 59, no. 2, pp. 143–148, 1998.

[72] J. Hasford, M. Baccarani, V. Hoffmann et al., "Predicting complete cytogenetic response and subsequent progression-free survival in 2060 patients with CML on imatinib treatment: the EUTOS score," *Blood*, vol. 114, no. 22, pp. 686–692, 2009.

[73] M. Breccia, F. Stagno, A. Gozzini et al., "Hammersmith score application identifies chronic myeloid leukemia patients with poor prognosis before treatment with second-generation tyrosine kinase inhibitors," *American Journal of Hematology*, vol. 116, no. 21, pp. 1395-1396, 2010.

[74] C. N. A. M. Oldenhuis, S. F. Oosting, J. A. Gietema, and E. G. E. de Vries, "Prognostic versus predictive value of biomarkers in oncology," *European Journal of Cancer*, vol. 44, no. 7, pp. 946–953, 2008.

[75] M. Agrawal, R. J. Garg, H. Kantarjian, and J. Cortes, "Chronic myeloid leukemia in the tyrosine kinase inhibitor era: What is the "best" therapy?" *Current Oncology Reports*, vol. 12, no. 5, pp. 302–313, 2010.

[76] F. Guilhot and J. Guilhot, "Predicting response in CML," *Blood*, vol. 117, no. 6, pp. 1773-1774, 2011.

[77] E. Jabbour, H. Kantarjian, S. O'Brien et al., "Predictive factors for outcome and response in patients treated with second-generation tyrosine kinase inhibitors for chronic myeloid leukemia in chronic phase after imatinib failure," *Blood*, vol. 117, no. 6, pp. 1822–1827, 2011.

[78] J. Sokal, E. Cox, M. Baccarani, and S. Tura, "Prognostic discrimination in good-risk chronic granulocytic leukemia," *Blood*, vol. 63, pp. 789–799, 1984.

[79] J. Hasford, M. Pfirrmann, R. Hehlmann et al., "A new prognostic score for survival of patients with chronic myeloid leukemia treated with interferon alfa. Writing Committee for the Collaborative CML Prognostic Factors Project Group," *Journal of the National Cancer Institute*, vol. 90, no. 11, pp. 850–858, 1998.

[80] J. Hasford, M. Baccarani, V. Hoffmann et al., "Predicting complete cytogenetic response and subsequent progression-free survival in 2060 patients with CML on imatinib treatment: the EUTOS score," *Blood*, vol. 118, no. 3, pp. 686–692, 2011.

[81] V. S. Hoffmann, M. Baccarani, D. Lindoerfer et al., "The EUTOS prognostic score: review and validation in 1288 patients with CML treated frontline with imatinib," *Leukemia*, vol. 27, no. 10, pp. 2016–2022, 2013.

[82] S.-B. Cho, "Exploring features and classifiers to classify gene expression profiles of acute leukemia," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 16, no. 7, pp. 831–844, 2002.

[83] I. Inza, B. Sierra, R. Blanco, and P. Larrañaga, "Gene selection by sequential search wrapper approaches in microarray cancer class prediction," *Journal of Intelligent & Fuzzy Systems*, vol. 12, no. 1, pp. 25–33, 2002.

[84] A. Farag, "Computer Based Acute Leukemia Classification," in *Proceedings of the 2003 46th Midwest Symposium on Circuits and Systems*, pp. 701–703, Cairo, Egypt.

[85] M. E. Futschik, A. Reeve, and N. Kasabov, "Evolving connectionist systems for knowledge discovery from gene expression data of cancer tissue," *Artificial Intelligence in Medicine*, vol. 28, no. 2, pp. 165–189, 2003.

[86] S.-B. Cho and H.-H. Won, "Data mining for GENE expression profiles from DNA microarray," *International Journal of Software Engineering and Knowledge Engineering*, vol. 13, no. 6, pp. 593–608, 2003.

[87] V. Marohnic, Z. Debeljak, and N. Bogunovic, "Mutual information based reduction of data mining dimensionality in gene expression analysis. Information Technology Interfaces," in *Proceedings of the 26th International Conference*, pp. 7–10, 2004.

[88] J. F. McCarthy, K. A. Marx, P. E. Hoffman et al., "Applications of machine learning and high-dimensional visualization in cancer detection, diagnosis, and management," *Annals of the New York Academy of Sciences*, vol. 1020, pp. 239–262, 2004.

[89] J. Rowland, "On genetic programming and knowledge discovery in transcriptome data," in *Proceedings of the 2004 Congress on Evolutionary Computation*, pp. 158–165, Portland, OR, USA.

[90] T. Markiewicz, S. Osowski, B. Marianska, and L. Moszczyński, "Automatic recognition of the blood cells of myelogenous leukemia using SVM," in *Proceedings of the International Joint Conference on Neural Networks, IJCNN 2005*, pp. 2496–2501, August 2005.

[91] W. L. Tung and C. Quek, "GenSo-FDSS: a neural-fuzzy decision support system for pediatric ALL cancer subtype identification using gene expression data," *Artificial Intelligence in Medicine*, vol. 33, no. 1, pp. 61–88, 2005.

[92] H. Nguyen, S. Ohn, J. Park, and K. Park, "Combined Kernel Function Approach in SVM for Diagnosis of Cancer," in *Advances in Natural Computation*, vol. 3610 of *Lecture Notes in Computer Science*, pp. 1017–1026, Springer, Berlin, Germany, 2005.

[93] V. P. Plagianakos, D. K. Tasoulis, and M. N. Vrahatis, "Computational intelligence techniques for acute leukemia gene expression data classification," in *Proceedings of the International Joint*

Conference on Neural Networks, IJCNN 2005, pp. 2469–2474, August 2005.

[94] F. Li and Y. Yang, "Analysis of recursive gene selection approaches from microarray data," *Bioinformatics*, vol. 21, no. 19, pp. 3741–3747, 2005.

[95] W. Jinlian, L. Jiangeng, and R. Xiaogang, "Mining leukemia gene association structure with DNA microarray," in *Proceedings of the 2005 International Conference on Neural Networks and Brain Proceedings, ICNNB'05*, pp. 695–701, October 2005.

[96] F. Diaz, F. Fdez-Riverola, D. Glez-Pena, and J. M. Corchado, "Using fuzzy patterns for gene selection and data reduction on microarray data," in *Intelligent Data Engineering and Automated Learning - Ideal*, E. Corchado, H. Yin, V. Botti, and C. Fyfe, Eds., Lecture Notes in Computer Science, pp. 1087-94, 2006.

[97] C. Feng and W. Lipo, "A Novel Approach Searching for Discriminative Gene Sets. Systems, Man and Cybernetics," in *Proceeding of the 2006 SMC '06 IEEE International Conference*, pp. 8–11, 2006.

[98] H. Nguyen and S. Ohn, "DRFE: Dynamic Recursive Feature Elimination for Gene Identification Based on Random Forest," in *Neural Information Processing*, vol. 4234 of *Lecture Notes in Computer Science*, pp. 1–10, Springer, Berlin, Germany, 2006.

[99] W. Shulin, C. Huowang, W. Ji, and Z. Dingxing, "Molecular Diagnosis of Tumor Based on Independent Component Analysis and Support Vector Machines. Computational Intelligence and Security, 2006 International Conference on," in *Proceedings of the Computational Intelligence and Security, 2006 International Conference*, 2006.

[100] Y. Chen, M. Dong, and M. Rege, "Gene expression clustering: A novel graph partitioning approach," in *Proceedings of the 2007 International Joint Conference on Neural Networks, IJCNN 2007*, pp. 1542–1547, August 2007.

[101] M. Perez, D. M. Rubin, L. E. Scott, T. Marwala, and W. Stevens, "A Hybrid Fuzzy-SVM classifier, applied to gene expression profiling for automated leukaemia diagnosis," in *Proceedings of the 2008 IEEE 25th Convention of Electrical and Electronics Engineers in Israel, IEEEI 2008*, pp. 41–45, December 2008.

[102] R. Avogadri, M. Brioschi, F. Ferrazzi, M. Re, A. Beghini, and G. Valentini, "A stability-based algorithm to validate hierarchical clusters of genes," *International Journal of Knowledge Engineering and Soft Data Paradigms*, vol. 1, no. 4, p. 318, 2009.

[103] A. Chaiboonchoe, S. Samarasinghe, and D. Kulasiri, "Using emergent clustering methods to analyse short time series gene expression data from childhood leukemia treated with glucocorticoids," in *Proceedings of the 18th World IMACS Congress and International Congress on Modelling and Simulation: Interfacing Modelling and Simulation with Mathematical and Computational Sciences, MODSIM09*, pp. 741–747, July 2009.

[104] D. Glez-Peña, F. Díaz, F. Fdez-Riverola, J. R. Méndez, and J. M. Corchado, "Fuzzy patterns and GCS networks to clustering gene expression data," *Studies in Fuzziness and Soft Computing*, vol. 242, pp. 103–125, 2009.

[105] Y. He and S. C. Hui, "Exploring ant-based algorithms for gene expression data analysis," *Artificial Intelligence in Medicine*, vol. 47, no. 2, pp. 105–119, 2009.

[106] A. Mukhopadhyay, U. Maulik, and S. Bandyopadhyay, "Unsupervised cancer classification through SVM-boosted multiobjective fuzzy clustering with majority voting ensemble," in *Proceedings of the 2009 IEEE Congress on Evolutionary Computation, CEC 2009*, pp. 255–261, May 2009.

[107] A. Torkaman, N. M. Charkari, M. Aghaeipour, and E. Hajati, "A recommender system for detection of leukemia based on cooperative game," in *Proceedings of the 2009 17th Mediterranean Conference on Control and Automation (MED)*, pp. 1126–1130, Thessaloniki, Greece, June 2009.

[108] Z. Zheng, S. Sharma, N. Agarwal, and L. Huan, "Integrating Knowledge in Search of Biologically Relevant Genes," in *Proceedings of the Data Mining Workshops, 2009 ICDMW '09 IEEE International Conference*, W. Jiangxin and C. Yung, Eds., 2009.

[109] A. M. Mehdi, M. S. Sehgal, A. Zayegh, R. Begg, and A. Manan, "K-means clustering on 3rd order polynomial based normalization of Acute Myeloid Leukemia (AML) and Acute Lymphocyte Leukemia (ALL)," in *Proceedings of the 2009 3rd International Conference on Electrical Engineering, ICEE 2009*, April 2009.

[110] C. Porzelius, M. Johannes, H. Binder, and T. Beissbarth, "Leveraging external knowledge on molecular interactions in classification methods for risk prediction of patients," *Biometrical Journal*, vol. 53, no. 2, pp. 190–201, 2011.

[111] J. Chen, D. Li, B. Zhang, Z. Wang, Y. Wang, and Q. Zheng, "Notice of Retraction Alternol Inhibits Proliferation and Induces Apoptosis in Human Leukemia HL-60 Cells," in *Proceedings of the 2011 5th International Conference on Bioinformatics and Biomedical Engineering*, pp. 10–12, 2011.

[112] J. A. Gonzalez, I. Olmos, L. Altamirano et al., "Leukemia identification from bone marrow cells images using a machine vision and data mining strategy," *Intelligent Data Analysis*, vol. 15, no. 3, pp. 443–462, 2011.

[113] D. L. Tong and A. C. Schierz, "Hybrid genetic algorithm-neural network: Feature extraction for unpreprocessed microarray data," *Artificial Intelligence in Medicine*, vol. 53, no. 1, pp. 47–56, 2011.

[114] P. S. Chauhan, R. Ihsan, A. K. Mishra et al., "High order interactions of xenobiotic metabolizing genes and P53 codon 72 polymorphisms in acute leukemia," *Environmental and Molecular Mutagenesis*, vol. 53, no. 8, pp. 619–630, 2012.

[115] H. J. Escalante, M. Montes-y-Gómez, J. A. González et al., "Acute leukemia classification by ensemble particle swarm model selection," *Artificial Intelligence in Medicine*, vol. 55, no. 3, pp. 163–175, 2012.

[116] K. Y. Yeung, T. A. Gooley, A. Zhang, A. E. Raftery, J. P. Radich, and V. G. Oehler, "Predicting relapse prior to transplantation in chronic myeloid leukemia by integrating expert knowledge and expression data," *Bioinformatics*, vol. 28, no. 6, Article ID bts059, pp. 823–830, 2012.

[117] T. Manninen, H. Huttunen, P. Ruusuvuori, and M. Nykter, "Leukemia Prediction Using Sparse Logistic Regression," *PLoS ONE*, vol. 8, no. 8, Article ID e72932, 2013.

[118] A. A. El-Nasser, M. Shaheen, and H. El-Deeb, "Enhanced leukemia cancer classifier algorithm," in *Proceedings of the 2014 Science and Information Conference, SAI 2014*, pp. 422–429, August 2014.

[119] V. Singhal and P. Singh, "Correlation based feature selection for diagnosis of acute lymphoblastic leukemia," in *Proceedings of the 3rd ACM International Symposium on Women in Computing and Informatics (WCI '15)*, pp. 5–9, Kochi, India, August 2015.

[120] D. Yao, J. Yang, X. Zhan, X. Zhan, and Z. Xie, "A novel random forests-based feature selection method for microarray expression data analysis," *International Journal of Data Mining and Bioinformatics*, vol. 13, no. 1, pp. 84–101, 2015.

[121] J. Rawat, A. Singh, H. S. Bhadauria, and J. Virmani, "Computer Aided Diagnostic System for Detection of Leukemia Using Microscopic Images," in *Proceedings of the 4th International Conference on Eco-friendly Computing and Communication Systems, ICECCS 2015*, pp. 748–756, December 2015.

[122] S. Kar, K. D. Sharma, and M. Maitra, "Gene selection from microarray gene expression data for classification of cancer subgroups employing PSO and adaptive K-nearest neighborhood technique," *Expert Systems with Applications*, vol. 42, no. 1, pp. 612–627, 2015.

[123] J. Li, Y. Wang, Y. Cao, and C. Xu, "Weighted doubly regularized support vector machine and its application to microarray classification with noise," *Neurocomputing*, vol. 173, pp. 595–605, 2016.

[124] A. K. Dwivedi, "Artificial neural network model for effective cancer classification using microarray gene expression data," *Neural Computing and Applications*, pp. 1–10, 2016.

[125] S. Krappe, T. Wittenberg, T. Haferlach, and C. Münzenmayer, "Automated morphological analysis of bone marrow cells in microscopic images for diagnosis of leukemia: Nucleus-plasma separation and cell classification using a hierarchical tree model of hematopoesis," in *Proceedings of the Medical Imaging 2016: Computer-Aided Diagnosis*, March 2016.

[126] R. Ocampo-Vega, G. Sanchez-Ante, M. A. De Luna, R. Vega, L. E. Falcón-Morales, and H. Sossa, "Improving pattern classification of DNA microarray data by using PCA and Logistic Regression," *Intelligent Data Analysis*, vol. 20, no. 1, pp. S53–S67, 2016.

[127] B. Rajwa, P. K. Wallace, E. A. Griffiths, and M. Dundar, "Automated Assessment of Disease Progression in Acute Myeloid Leukemia by Probabilistic Analysis of Flow Cytometry Data," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 5, pp. 1089–1098, 2017.

[128] W. Ni, B. Hu, C. Zheng et al., "Automated analysis of acute myeloid leukemia minimal residual disease using a support vector machine," *Oncotarget*, vol. 7, no. 44, pp. 71915–71921, 2016.

[129] D. Milojkovic, E. Nicholson, J. F. Apperley et al., "Early prediction of success or failure of treatment with second-generation tyrosine kinase inhibitors in patients with chronic myeloid leukemia," *Haematologica*, vol. 95, no. 2, pp. 224–231, 2010.

[130] D. White, V. Saunders, A. B. Lyons et al., "In vitro sensitivity to imatinib-induced inhibition of ABL kinase activity is predictive of molecular response in patients with de novo CML," *Blood*, vol. 106, no. 7, pp. 2520–2526, 2005.

[131] D. L. White and T. P. Hughes, "Predicting the response of CML patients to tyrosine kinase inhibitor therapy," *Current Hematologic Malignancy Reports*, vol. 4, no. 2, pp. 59–65, 2009.

[132] D. L. White, V. A. Saunders, P. Dang et al., "Most CML patients who have a suboptimal response to imatinib have low OCT-1 activity: Higher doses of imatinib may overcome the negative impact of low OCT-1 activity," *Blood*, vol. 110, no. 12, pp. 4064–4072, 2007.

[133] D. L. White, P. Dang, J. Engler et al., "Functional activity of the OCT-1 protein is predictive of long-term outcome in patients with chronic-phase chronic myeloid leukemia treated with imatinib," *Journal of Clinical Oncology*, vol. 28, no. 16, pp. 2761–2767, 2010.

[134] S. Branford, D. T. Yeung, D. M. Ross et al., "Early molecular response and female sex strongly predict stable undetectable BCR-ABL1, the criteria for imatinib discontinuation in patients with CML," *Blood*, vol. 121, no. 19, pp. 3818–3824, 2013.

[135] K. Sweet, L. Zhang, and J. Pinilla-Ibarz, "Biomarkers for determining the prognosis in chronic myelogenous leukemia," *Journal of Hematology and Oncology*, vol. 6, no. 1, article 54, 2013.

[136] W. Prejzner, "Relationship of the BCR gene breakpoint and the type of BCR/ABL transcript to clinical course, prognostic indexes and survival in patients with chronic myeloid leukemia," *Medical Science Monitor*, vol. 8, no. 5, pp. BR193–BR197, 2002.

[137] D. Verma, H. M. Kantarjian, D. Jones et al., "Chronic myeloid leukemia (CML) with $P190^{BCR-ABL}$: analysis of characteristics, outcomes, and prognostic significance," *Blood*, vol. 114, no. 11, pp. 2232–2235, 2009.

[138] J.-Z. Xu and C.-W. Wong, "Hunting for robust gene signature from cancer profiling data: Sources of variability, different interpretations, and recent methodological developments," *Cancer Letters*, vol. 296, no. 1, pp. 9–16, 2010.

[139] Y.-J. Fan and W. A. Chaovalitwongse, "Optimizing feature selection to improve medical diagnosis," *Annals of Operations Research*, vol. 174, no. 1, pp. 169–183, 2010.

[140] Z. Chen, J. Li, L. Wei, W. Xu, and Y. Shi, "Multiple-kernel SVM based multiple-task oriented data mining system for gene expression data analysis," *Expert Systems with Applications*, vol. 38, no. 10, pp. 12151–12159, 2011.

[141] E. J. Yeoh, M. E. Ross, S. A. Shurtleff et al., "Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling," *Cancer Cell*, vol. 1, no. 2, pp. 133–143, 2002.

[142] A. Mukhopadhyay, U. Maulik, and S. Bandyopadhyay, "Refining genetic algorithm based fuzzy clustering through supervised learning for unsupervised cancer classification," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 5483, pp. 191–202, 2009.

[143] J.-H. Cho, D. Lee, H. P. Jin, K. Kim, and I.-B. Lee, "Optimal approach for classification of acute leukemia subtypes based on gene expression data," *Biotechnology Progress*, vol. 18, no. 4, pp. 847–854, 2002.

[144] Z. Chen, J. Li, and L. Wei, "A multiple kernel support vector machine scheme for feature selection and rule extraction from gene expression data of cancer tissue," *Artificial Intelligence in Medicine*, vol. 41, no. 2, pp. 161–175, 2007.

[145] M. Baccarani, G. Saglio, J. Goldman et al., "Evolving concepts in the management of chronic myeloid leukemia: recommendations from an expert panel on behalf of the European LeukemiaNet," *Blood*, vol. 108, no. 6, pp. 1809–1820, 2006.

[146] T. Hughes and S. Branford, "Molecular monitoring of BCR-ABL as a guide to clinical management in chronic myeloid leukaemia," *Blood Reviews*, vol. 20, no. 1, pp. 29–41, 2006.

[147] M. Baccarani, G. Gugliotta, F. Castagnetti, S. Soverini, and G. Rosti, "A Review and an Update of European LeukemiaNet Recommendations for the Management of Chronic Myeloid Leukemia," in *Chronic Myeloid Leukemia*, Hematologic Malignancies, pp. 55–69, Springer International Publishing, Cham, 2016.

[148] S. L. Noble, E. Sherer, R. E. Hannemann, D. Ramkrishna, T. Vik, and A. E. Rundell, "Using adaptive model predictive control to customize maintenance therapy chemotherapeutic dosing for childhood acute lymphoblastic leukemia," *Journal of Theoretical Biology*, vol. 264, no. 3, pp. 990–1002, 2010.

[149] L. M. Fagan and E. H. Shortliffe, *The Future of Computer Applications in Biomedicine. Biomedical Informatics Computer Applications in Health Care and Biomedicine*, Health Informatics, Springer, New York, NY, USA, 3rd edition, 2006.

[150] J. A. Cruz and D. S. Wishart, "Applications of machine learning in cancer prediction and prognosis," *Cancer Informatics*, vol. 2, pp. 59–77, 2006.

Linking Chapter 2 and 3:

We will be linking our papers, “Intelligent Techniques Using Molecular Data Analysis in Leukemia: An Opportunity for Personalized Medicine Support System” and “Personalized Medicine Support System: Resolving Conflict in Allocation to Risk Groups and Predicting Patient Molecular Response to Targeted Therapy.” In the former article, our aim was to study previously published research on personalized medicine intelligent techniques in leukemia. We also aimed to synthesize findings in the research related to leukemia and pay special attention to categories of these studies based on leukemia type and also the task, data source, and purpose of the studies to help identify further opportunities for medical research as it pertains to support systems for leukemia. After conducting the search strategy and applying the selection criteria, the findings were summarized in the Results Section. The emphasis was on CML because a few studies involved chronic leukemia subtypes that use intelligent techniques in the analysis of CML data as well as the recent development in monitoring CML patients’ molecular responses to molecular targeted therapy. Moreover, CML medical research is an active area that has many clinical trial results; one heated debate has concerned selecting the optimal treatment to achieve an optimal patient response. Based upon these results, minor attention was given to risk prediction, while the majority of the study aimed to identify cancer types or markers. Categorizing the selected studies based on the intelligent technique task emphasized the opportunity to develop a system that integrates multiple tasks to solve the medical issue. An intelligent system should be able to support clinicians while making a decision and discovering knowledge from data as well as handling prior knowledge by ensemble techniques. We concluded that the field of molecular medicine is experiencing “transformational changes” that are bringing forth a change in healthcare. We also determined that it would be beneficial to build personal medicine support systems that will bring a ray of hope to the treatment of leukemia.

The second article, “Personalized Medicine Support System: Resolving Conflict in Allocation to Risk Groups and Predicting Patient Molecular Response to Targeted Therapy,” we considered the research opportunities raised from our literature review. Based on categorizing the selected studies based on the leukemia type, data sources, purpose of the studies and intelligent technique task, our paper emphasized on how to propose a system that integrates multiple tasks to solve the medical issue. The intelligent system architecture should be able to support clinicians while making a decision and discovering knowledge from data and combining prior knowledge. In addition, the importance of integrating domain knowledge into predictive and prognostic models for personalized treatment was discussed. We also state within the article that our intelligent system provides support in making complex decisions and can be incorporated into a treatment guide for selecting molecular targeted therapies.

In the Introduction of the second paper, we begin to make our argument that personalized medicine can be used to establish molecular characteristics that are unique to an individual due to his or her specific genetic makeup. We propose that clinicians will eventually be able to offer more tailored diagnoses and treatment protocols for various patients’ individual diseases. After introducing and defining this argument, we then discuss clinical decision support systems (CDSS), which are decision support systems that analyze medical data and help healthcare professionals make effective and timely clinical decisions. After discussing CDSS, we then describe knowledge-based systems. We loosely define these as the wide range of knowledge resources that help clinicians with decision-making. Later in the paper, after defining and describing a few more key working terms and concepts, we discuss the overall framework of the implementation of a personalized medicine support system. This type of system includes an overview of a personalized medicine support system: a prognostic model and a predictive model, a discussion of the prognostic and predictive factors, the use of domain knowledge,

modeling, how prior knowledge was dealt with during research, and how new sources of clinical data were handled. After these sections, the proposed applications of the personalized medicine support systems and conclusions that were made after all of the research and extended study were completed.

Chapter 3:Personalized Medicine Support System: Resolving Conflict in Allocation to Risk Groups and Predicting Patient Molecular Response to Targeted Therapy

Statement of Authorship

Title of Paper	Personalized Medicine Support System: Resolving Conflict In Allocation to Risk Groups and Predicting Patient Molecular Response to Targeted Therapy
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Banjar, H., Adelson, D., Brown, F. and Leclercq, T. Personalized Medicine Support System : Resolving Conflict In Allocation to Risk Groups and Predicting Patient Molecular Response to Targeted Therapy. Health Informatics - An International Journal (HIIJ) Vol.6, No.2, May 2017

Principal Author

Name of Principal Author (Candidate)	Haneen Reda Banjar		
Contribution to the Paper	HB designed and performed the research, formal analysis, investigation, methodology, project administration, resources, software, and validation. She also wrote the original draft.		
Overall percentage (%)	90%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature	[redacted]	Date	2nd Dec 2017

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate in include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	David Adelson		
Contribution to the Paper	DA revised the manuscript and approved the version of the manuscript to be published.		
Signature	[redacted]	Date	4 Dec 2017

Name of Co-Author	Fred Brown		
Contribution to the Paper	FB revised the manuscript and approved the version of the manuscript to be published.		
Signature	[redacted]	Date	11/12/17

Name of Co-Author	Tamara Leclercq		
Contribution to the Paper	TL revised the manuscript and approved the version of the manuscript to be published.		
Signature	[redacted]	Date	3-1-18

PERSONALIZED MEDICINE SUPPORT SYSTEM: RESOLVING CONFLICT IN ALLOCATION TO RISK GROUPS AND PREDICTING PATIENT MOLECULAR RESPONSE TO TARGETED THERAPY

Haneen Banjar[1, 2], David Adelson[3,] Fred Brown[2]and Tamara Leclercq[4]

[1]School of Computer Science, University of Adelaide, Adelaide, Australia.
[2]The Department of Computer Science, King Abdul Aziz University, Jeddah, Saudi Arabia.
[3]School of Molecular and Biomedical Science, University of Adelaide, Adelaide, Australia
[4]Cancer Theme, South Australian Health and Medical Research Institute (SAHMRI), Adelaide. South Australia, Australia

ABSTRACT

Treatment management in cancer patients is largely based on the use of a standardized set of predictive and prognostic factors. The former are used to evaluate specific clinical interventions, and they can be useful for selecting treatments because they directly predict the response to a treatment. The latter are used to evaluate a patient's overall outcomes, and can be used to identify the risks or recurrence of a disease. Current intelligent systems can be a solution for transferring advancements in molecular biology into practice, especially for predicting the molecular response to molecular targeted therapy and the prognosis of risk groups in cancer medicine. This framework primarily focuses on the importance of integrating domain knowledge in predictive and prognostic models for personalized treatment. Our personalized medicine support system provides the needed support in complex decisions and can be incorporated into a treatment guide for selecting molecular targeted therapies.

KEYWORDS

Personalized Medicine Support System, Molecular Targeted Therapy, Predict Molecular Response, Risk Assessment & Cancer Treatment.

1. INTRODUCTION

1.1 Personalized medicine

Personalized medicine could be used to establish molecular characteristics that are unique to an individual, usually because of varying genetic makeup. With this, doctors can offer a more specific diagnosis andeffective treatment protocol for an individual's disease. In clinical decision-making, personalized medicine is used to make decisions that maximize the outcomes and minimize the side-effects of treatment by using available knowledge about an individual [1]. Many studies have focused on personalized care, and its scope of application has increased because it leads to more successful outcomes and minimizes the chances of adverse reactions associated with certain treatment plans. Personalized medicine involves classifying individual

DOI: 10.5121/hiij.2017.6201

patients into sub-populations that pose unique reactions, susceptibilities or general responses to particular treatments to determine the best treatment approaches for every patient [2].

Personalized medicine offers high precision, and it has attracted much interest because it affords advantages such as improved healthcare and reduced need for the development of new medicines [3]. Personalized medicine also helps to minimize unnecessary costs by reducing both the time spent on treatment and the failure rates during clinical trials [4].Economic value increases when caregivers implement personalized medicine in healthcare, because it helps to reduce the unnecessary use of resources.

Traditional personalized medicine contrasts with modern personalized medicine in how suitable therapies are identified for patients. In the former, a diagnosis is predicted based on a patient's family history, social circumstances, environment and lifestyle. In the latter,this prediction is based on a patient's genetic makeup [2]. Medical researchers continue to emphasize that their studies are updated with the most effective treatment protocols used to treat cancer patients. However, different patients show different responses to the same therapy. For example, when subjected to the same therapeutic protocol, the clinical outcomes of two patients may be different in terms of response and survival[5]. This undesirable situation could potentially be avoided by improving predictive assays to predict an individual's response to therapy at diagnosis. Indeed, during recent years, studies have tried to individualize medicine by using predictive factors to determine initial treatments. Many studies have used different techniques to investigate factors that could affect drug responses. Sioud et al. [6]noted that individualized treatment algorithms could depend on integrating the factors that influence drug responses. For example, studying a single biomarker as a predictor could indicate the response before treatment and predict the risk to the individual [7]. Another approach involves predicting a drug reaction in terms of toxicity or resistance by using an individual's genotype data and clinical data to improve the individual's care [8].Inpharmacogenetics, an individual's response to a specific therapy is studied based on his/her genotype information. The European Science Foundation established a roadmap for developing personalized medicine [9] and stated that to build a predictive model, data for individualized treatment options were needed, and response variations in different sub-populations of patients and diseases needed to be stored. Fagan and Shortliffe[10] noted the power of using a computer to (1) analyze the combination of biological and clinical data, which are disparate data sources, and (2) discover the relationships amongst complex database schema that store medical records.

As treatment optimization is an important medical advancement, many computer science researchers have focused on predicting the response to treatment [11]. Indeed, many studies have reported on the development of a prognostic or predictive model that could predict the side-effects of drugs [12-15]; such models use intelligent techniques to address medical issues. This gives researchers a new way of applying the strength of machine-learning techniques to real medical problems. Current intelligent systems could potentially aid in transferring advancements in molecular biology into practice, especially for predicting molecular response to targeted therapy and the prognosis of patients (in different risk groups) in cancer medicine.

1.2 Clinical Decision Support System

A clinical decision support system (CDSS) is a type of decision support system that analyses medical data to assist healthcare providers in making appropriate and automated clinical decisions [16]. Constructing a CDSS requires substantial modeling activity by selecting relevant medical data and a problem-solving strategy to reach appropriate conclusions. For the analysis, collaboration with clinical experts is needed to model relevant application areas. Based on the nature of this collaboration, there are two types of CDSSs: knowledge-based CDSS and non-

knowledge-based CDSS [17]. A key research area in CDSS is the advancement and application of knowledge-based systems because of the usability of expressing various types of human knowledge in an intelligent system. Combining expert knowledge and discovered knowledge in the medical domain maximizes the qualities that they have separately [18]. There are also three classificationsfor CDSS: information management, situational awareness and patient-specific[19]. Information management uses buttonsfor obtaining up to date information; situational awareness uses alerts or dashboard for obtaining information; and patient-specific may be implemented for different purposes such as diagnosis, treatment management or recommendation [20]. Identifying the key needs and functional requirements and determining how to evaluate the system are interrelated issues in designing and implementinga CDSS [21]. Understanding a medical problem and its domain are also primary requirements for implementinga knowledge-based CDSS. Additional requirementsinclude knowing who will deliver the information from the CDSS and how the CDSS can provide support starting from diagnosis and selecting treatment to monitor and follow-up the outcomes.

In the last decade, information technology has been introduced to improve clinical practice [22]. Information systems make data storage and data retrieval easier. This enhances the use of CDSS in clinical practice, and it helps to prevent errors and deliver needed information at the time of request. The development of CDSS has been influenced by the purpose of use and advancements in medical technology. Researches are motivated by the concept of personalized medicine, and they have introduced personalized modeling for developing CDSSs based on the information of individual patients, such as clinical and molecular information [23].

Automatic knowledge acquisition techniques have also attracted much research interest. Knowledge acquisition is the first step in building a knowledge base, and it requires techniques to drive knowledge. Data mining is widely used for knowledge acquisition,and the knowledge discovery process is used as a schema to automatically discover knowledge, which can include the shape of patterns or associations among data features and rules. This process generally includes four steps: problem understanding and data understanding, data pre-processing, model induction, and post-processing [24]. A recent review [25]discussed knowledge discovery in medicine. In this review, the authors identified the primary studies in medical research applications through data mining techniques. They found that the focus has moved to the medical domain more than in the past and that the use of hybrid systems is the future for solving complex problems in medicine. Complex problems in the medical domain began with the success of sequencing the human genome. The morenew biological information was made available, the greater the demand to build intelligent systems that can utilizethis information. Intelligent systems such as CDSSs should be adaptive to any new information made available on an ongoing basis. From adaptive behavior, evolving systems have been introduced to use particularly intelligent systems and online learning algorithms for knowledge acquisition from data and for realizing advanced model structures in data mining tasks and parameters [26].

1.3 Knowledge-based Systems

Researchers have also focused on the integration of knowledge from multiple sources, which is a key issue in many areas such as collaborative knowledge systems, group decisions and distributed expert systems, where knowledge from multiple sources is often contradictory. We need to be comfortable with a wide range of knowledge resources that will help clinicians in decision-making, as a conflict between two sources, rules, or decision-makers is often observed in real cases.

Conflicts in knowledge-based systems can be divided into three types: schema conflicts, data conflicts, and knowledge conflicts [27]. Schema conflicts can arise from the use of different schema definitions such as tables or objects. Data conflicts may ariseowing to incorrect data. A

study compared the ability of several techniques to reduce inconsistency in data [28]. Knowledge conflict may arise when multiple sources of knowledge are integrated. Knowledge conflict is a major issue in distributed artificial intelligence systems [29]. Adrian et al.[30] defined contradiction as an 'inability for all conceived statements or beliefs to be simultaneously true'. A contradiction may also be referred to as an inconsistency. For instance, there may be a contradiction between sentences such that one sentence must be true and the other must be false.

This common situation requires some advanced methodologies to resolve the conflict. One active research direction is the use of the multiple criteria decision analysis (MCDA) approach [31, 32]. This technique is developed to capture the relative preference information of decision-makers who are involved in a conflict. There is a prescriptive approach based on progressive preferential knowledge [33]. This approach resolves the conflicts of rules in a knowledge-based system by using decision analysis techniques, where incorporating a user's preference in judgment about the rules is important. These methods for conflict resolution in a rule-based system always need toincorporate a user's preferences in some specific problem domain that is very difficult to obtain in the medical domain. Experts' opinions, including preferences, are often vague and difficult to estimate using exact numerical values. Noor-E-Alam et al. [34] developed a framework that used qualitative form as a linguistic term to represent experts' opinions. Their approach aimed to use multi-expert multicriteria decision making (ME-MCDM) to reach a single decision by integrating multiple experts' opinions. The linguistic truth value (LTV) can be used to judge alternatives for ME-MCDM. Two suitable algorithms are used to handle conflict aggregation: possibility measure and averaging conflict aggregation. Junming et al. [35] designed a framework using fuzzy case base reasoning for conflict resolution. This approach aimed to find similar cases in the resources and to retrieve the information that resolved the problem. They defined linguistic variables for comparing the retrieved information. Sometimes, it is important to measure the inconsistency; based on the results, the researcher can decide what to do with it. For example, one can measure the conflict and agreement between two knowledge bases [36] or the inconsistency of knowledge bases [37-41].Some studies used machine learning techniques to resolve conflicts. Recent studies on multi-agent systems aimed to resolve conflicts by combining machine learning techniques such as Bayesian network, case-based reasoning and expert systems [42]. This study demonstrated that 'the choice of the specific technique for a given domain depends on the specification of the domain'.

1.4 Medical Data and Knowledge

Clinical trials are usually conducted in a series of phases. Each clinical trial phase aims to answer a specific research question. Phase I is designed to test a new treatment in a small population and to identify side-effects and treatment safety. In Phase II, researchers test the drug on a larger population (several hundred individuals) and evaluate its efficacy and safety. In phase III, the trial is designed to study the efficacy in a population of several hundred to a thousand individuals. Phase IV is designed to monitor the drug in the general population after it is marketed[43]. Late phases of the clinical trials,such as phases II, III and IV, are considered important sources of information that can be used to build mathematical models. There is a rapid increase in the number of electronic medical research databases that provide an opportunity for researchers to reuse medical data to create mathematical models. To access the clinical trial data, the NCI[44]is a US agency that lists ongoing clinical trials of molecular targeted therapies. Investigators within hospitals and medical centers conduct most of the NCI-supported trials. The NCI provides the full trial description and the name of the principal investigator for these studies. Researchers can contact the investigators and collaborate with them.

Clinical trial data may be biasedin several aspects: sampling, referral, selection, method and clinical spectrum. The clinical trial may use sampling methods, sample size and inclusion and exclusion criteria. Another aspect is the referral bias, where a specialist refers patientsand thus the

data represents pre-selected patients who have high prevalence of a disease. The selection bias is clear when the clinical trial data includes a group based on various demographics. Data may be collected using different measurements, leading to different precisions and specifications. Finally, the clinical spectrum bias is represented in the patient's record, which may show other medical problems with the disease [24]. For instance, Saussele and Pfirrmann[45] reported clinical trials of chronic myeloid leukemia (CML). They noted several issues that may challenge the reuse of a clinical trial data. According to Saussele and Pfirrmann, the definition of 'remission'varies in clinical trials depending on the major molecular response (MMR) or complete cytogenetic response (CCR). In addition, clinical trials use different primary endpoints such as 12-month MMR or 12-month CCR to determine the treatment success.

From patient care to patient administration, electronic health records (EHR) are reused in many studies to answer specific research questions [46-49]. Cases were matched with enquires based on obtained research criteria for patient inclusion, and a dataset of many matches can be generated for analysis. The EHR may include sparse data or missing values, as some patients may not seek frequent care. The EHR quality would likely impact the bias in the research finding or the modeling performance. Thederivation of key variables is also an important aspect when dealing with EHR, as the values may be recorded in different ways in different systems. This arises because of varying definitions between sources. The data quality and correct values of derived key variables will concern researchers,and many algorithms can be investigated during pre-processing to improve the data quality,thereby providing reliable results [50].

The reuse of medical data and knowledge for cancer treatment requires advanced computational methods such as intelligent systems. Using multiple sources of data and knowledge for personalized medicine support systemswillcreatelarge amounts of information to deal with for evaluating therapies and potential diagnostic and prognostic markers.

1.5 Predictive vs. Prognostic

In oncology, predictive markers differ from prognostic markers [51, 52]. Predictive markers or factors used to evaluate specific clinical interventions, and they can be useful for selecting a treatment because they can directly predict the response to a treatment.Prognostic markers are used to evaluate a patient's overall outcomes, and can be used to identify the risk or recurrence of a disease. By using advanced molecular technologies, many predictive and prognostic factors have been introduced in cancer medicine. Studying predictive molecular biology in detail is beyond the scope of this study;however, the list of possible predictive factors that have been published previously in separate studies such as clinical, pathological and molecular tests are the primary focus of our study. Based on the available data from medical research and clinical trials, one can create a list of possible predictive factors and prepare it for further investigation and also explore any relationship among the predictive factors.

Several predictive and prognostic tests based on molecular biology have revolutionized cancer medicine. In our study, we propose collecting information about possible predictive factors and prognostic scores that are published in top-ranked medical journals such as Science, Medline and PubMed journals and in other publications indexed under keywords such as 'prediction of molecular response', 'predict outcomes', 'prognostic model or score', 'prognostic indicators' and 'predictive factors'. We propose communicating with several clinical trial investigators who are studying the efficiency of molecular targeted therapies and patient responses by using several predictive and prognostic factors for decision-making. Predictive factors that are regulated for use should be used. According to Mehta[51], predictive markers that are commercialized for decision-making should be reviewed by the FDA and be certified with adequate evidence. Official and ethical approvals on the reuse of medical research data should be obtained before establishing a list of predictive and prognostic factors from clinical trials.

1.6 Molecular Targeted Therapy

Targeted cancer treatments are also called 'molecularly targeted drugs', 'molecularly targeted therapy' and so on [53]. Molecular targeted therapy is a sort of customized therapeutic treatment that aims to treat cancer by acting on unique mutations, overactive kinases, or protein anomalies that drive cancer development. Specifically, the medications used for targeted treatment are intended to act on a specific biochemical pathway vital to the survival of that specific cancer [54]. The FDA has approved numerous targeted cancer treatments to treat specific types of cancer [55]. Others are being examined in clinical trials (human testing), and numerous more are in preclinical testing (animal testing) [56].The National Cancer Institute (NCI) website lists multiple targeted therapies that have been approved for particular types of cancer[57]. For example, two targeted therapies named Bevacizumab (Avastin®)and Everolimus (Afinitor®)have been approved for treating brain cancer. Multiple targeted therapies are also available for leukaemia patients. To personalize targeted therapies, we need to determine the best therapeutic agent for an individual patient at the time of diagnosis. This can be done by studying the efficacy of each targeted therapy and developing a clinical trial. The NCI website provides a list of clinical trials of targeted therapies[57]. These trials are an important resource for shifting information from the laboratory to the implementation of personalized medicine. Developments in molecular technology will play a major role in realizing predictive, preventative and personalized medicine. Here, we also focus on the issue of how to reuse patients' existing profiles, including personal molecular information gathered in clinical trials for resolving conflict in allocation to risk groups and predicting patient molecular response to targeted therapy.

1.7 Research Motivations

To benefit from molecular technology data for treatment and monitoring progress in cancer patients, implementing an intelligent system that integrates medical knowledge, identifies relevant features and suggests drugs or treatments for patientsas well as adapting a new source of data and knowledge to change, remains a future challenge. In this research, we focus on the following main question: 'How can one develop a personalized medicine support system that allocates risk groups and predicts the molecular response tomolecular targeted therapy, where consistency in decisions is paramount?'

Therefore, previous studies have not considered improving knowledge-based CDSS that has a high-quality knowledge base to provide consistency in decisionsmade based on a patient's molecular information for predicting his/her molecular response.The main objective of this framework is to build a personalized medicine support system as a knowledge-based system changes and evolves over time. This intelligent system will provide clinicians, patients and researchers with a platform to extract knowledge and to improve the personalized treatments of patients. Framework functions have not been considered in previous studies on personalized medicine,and they have also not been specifically applied in medical literature to the problem we are considering.Doing so will add substantial value to the work. Ageneric framework is presented that includes the following main sub-functions: (1) consistency test between scoring systems for predicting outcomes in cancer patients treated with molecular targeted therapy, (2) resolving conflicts in validated prognostic scores, (3) developing a predictive model to predict molecular response to molecular targeted therapy, (4) select relevant predictive factors for predicting molecular response to molecular targeted therapy, (5) derive a clinical prediction rule that uses clinical, molecular and cell count observations (these predictive factors were collected atdiagnosis and categorized based on domain knowledge) and (6) adaptive methods that can evolve knowledge-based systems over time.

2. PROPOSED FRAMEWORK

In this section, the system overview ispresented, stages of implementing a personalized medicine support system are discussed and the system structure is described in detail.

2.1 Overview of Personalized Medicine Support System

Patient treatment has become increasingly difficult in recent years as various molecular targeted treatment options have become available. Clinicians aim to treat individual cancer patients with the most appropriate therapy. The ultimate challenge is to decide the most appropriate treatment strategy for an individual patient. To facilitate this, the disease assessment of an individual patient in terms of risk profile and molecular response to initial molecular targeted treatment should be determined. In prognostic models, low-risk patientsare preferentially treated with the least toxic and safest therapeutic options. These patients' molecular responsesare predicted using the predictive model built for the initial molecular targeted treatment. In higher-risk patients, higher-toxicity or combination treatment options may be favoured. High-risk patients or those who fail to respond to initial therapy will benefit from this approach as their response should bepredicted, thereby preventing resistance and/or intolerance to therapy, that may have otherwise ensued. Thus, this system aims to enhance individual disease and treatment management.
A personalized support system is designed for selecting patient therapy. Reused medical research data (clinical trial data) and prior knowledge collected from literature are used with advanced computational techniques to allocate risk groups and predict patients' molecular response to molecular targeted therapy at diagnosis. The personalized medicine support system uses two models as a guide for treatment selection (Figure 1):(1) prognostic model and (2) predictive model.

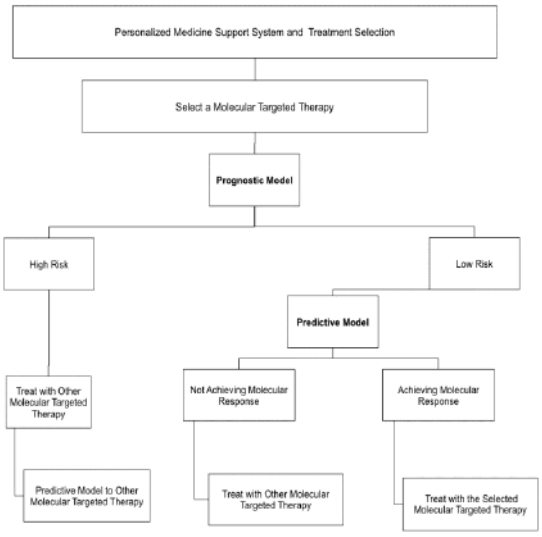


Figure1. Process flow of personalized medicine support system

2.1.1 Prognostic Model

The prognostic model uses four data mining processes: problem understanding and data understanding, data pre-processing, model development and post-processing. In medicine, it is

important to investigate current validated prognostic factors and score systems that are used to stratify cancer patients according to risk profile to ensure appropriate treatment. Historically, prognostication has seen rapid developments, and various scoring systems have been developed to optimize the use of clinical experience in cancer treatment. Our system tests the consistency between these scores while identifying risk categories, and the main procedure in this step is identifying whether conflict groups exist. During data pre-processing, prior knowledge is prepared for the analysis, and necessary prognostic factors are determined to implement previous prognostic scores. The data store contradiction in prior knowledge (validated prognostic scores) is prepared in this process. Feature ranking, conflict analysis and data normalization on a unified scale are performed to help in describing the relation between prognostic scores.

Model development combines validated scoring systems to determine whether the combined model may further improve the prediction compared with a single source of knowledge. In this process, one needs to know what patients' outcomes are as conveyed by previous prognostic scores. Using combined methods to resolve conflict in conflict groups adds another dimension of knowledge. Three levels are used for combining the validated prognostic scores and factors: combination, classification and feature. Prognostic factors are combined in the feature level,and prognostic scores are combined as score values and risk categories. Evaluating the combined methods and selecting the best approach can eliminate contradictions. Highest-accuracy methods will be selected for external evaluation. The final process is to use the selected prognostic models and validation on unseen data. Extending the prognostic model to adopt a new prognostic score or prognostic factor requires an additional process. Figure 2 shows the process flow.

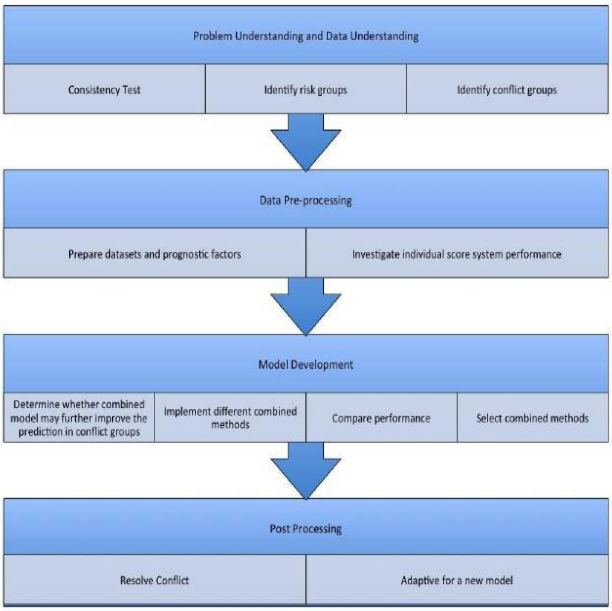


Figure2.Process flow of data mining processes used in prognostic model

2.1.2 Predictive Model

The predictive model also uses four data mining processes: problem understanding and data understanding, data pre-processing, model induction and post-processing. Figure 3 shows the process flow. The first process in knowledge discovery is divided into two stages. In the first

stage, each predictive factor used as an input parameter, studied to understand whether it is relevant to the prediction. When medical research data is reused, it is important to understand the data, such as the comprehensive treatment protocol or specifications for responses.Inclusion and exclusion criteria should also be decided before conducting any analysis to ensure that the material meets clinical requirements for the induction of a model with an accurate population.The second phase is data pre-processing. A statistical method is used to input missing values in predictive factors and prognostic scores from prior knowledge. Removing these patients is not an option in medical research data because most medical research data comprises small populations. For instance, patients who are enrolled in clinical trials mustprovide signed consent to undergo molecular targeted therapy in phase II clinical trials, as the safety profile of the drug is still not confirmed. Furthermore, we use domain knowledge to reformat the predictive factors. This process is performed before model induction by using medical knowledge to split continuous predictive factors into two or more linguistic terms and intervals. Thus, the clinical prediction rules can be interpretedmore easily based on domain concepts.

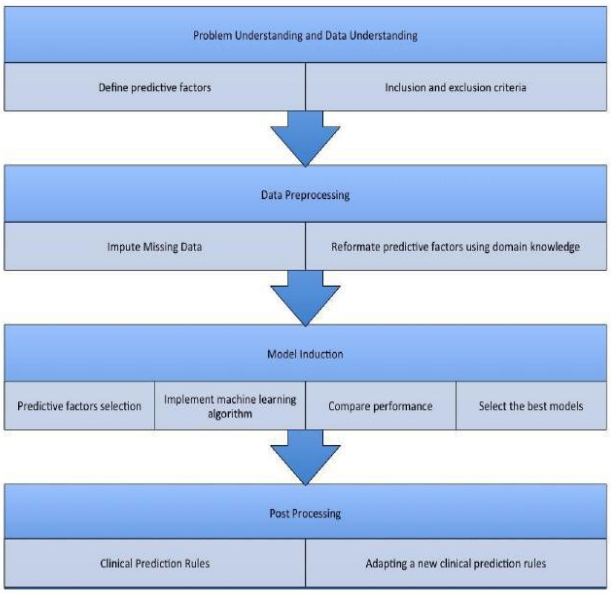


Figure3.Process flow of data mining processes used in predictive model

The third phase is model induction. First, there are three general steps for predictive model induction: training set, learning algorithm and performance evaluation on testing set. The training set represents the patient profiles that store the collected molecular, clinical and blood count information for individuals. We use available factors irrespective of whether we have prior knowledge that a factor's ability to predict patient response has been identified or whether a factor is new in the field. This is because these factors will be used in feature selection techniques [58] to reduce the dimension of the data. Feature selection (predictive factor selection) is important to minimize the cost of the requested tests for a patient. In addition, technically, it can increase the model accuracy because removing unnecessary data may increase the quality of the training set [59]. The second step is the learning algorithm.This is an important step in which the selected machine learning techniques are used for learning [60]. The selected algorithm should outperform previous methods and should be interpretable. Interpretability is an important aspect for medical experts. The selected machine learning algorithm is popular in the medical domain as a primary analytical model for discovering the relation among variables, dividing a dataset into

groups based on shared characteristics (classification) and providing a set of clinical prediction rules that can be applied to new patients. The model performance is evaluated by comparing patients' actual outcomes (molecular response) with those predicted by the model. A confusion matrix can help in measuring the performance based on accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), G-mean(geometric mean) and F-score(weighted harmonic mean of sensitivity and PPV). Two sets are used in the model induction processfor validating a small dataset: training and testing setand cross-validation set[61]. When a dataset is divided into a number of equal folds, one fold is used for validation and the others, for training. The process is repeated to ensure that all folds are used one time for validation. The standard deviation is measured after completing this cross-validation approach.

In post-processing, we compare the performance of the predictive models using validation sets and select the model that achieves the highest validation performance. The predictive model generates clinical prediction rules and discovers unknown knowledge by selecting relevant predictive factors (molecular tests, clinicalbase and cell count observations collected at diagnosis and categorized based on available knowledge) and extracting the relationship between the predictive factors and the molecular outcomes.This would prove that the approach could be applied in practice for treatment management by integrating domain knowledge in the learning process.A machine learning algorithm is used to discover unknown knowledge and perform classification. The predictive models present interesting properties of the data and use historical data to predict the behavior of unseen data.

2.2 Prognostic and Predictive Factors

The factors can be selected based on experts' choices and publications. All predictive factors and prognostic factors have an equal chance of being studied by the selected feature selection techniques to compare the performance of each or group of predictive factors to predict the molecular response, and prognostic factors to allocate risk group. Many factors may influence the correlation between baseline data and end-point molecular response.There are manycategories of factors that maybe clinically based, biologically (molecularly) based, environmentally based, assay results and family history data etc. In a personalized medicine support system, we collect information from literature and create possible predictive and prognostic factors based on available data. We select the predictive and prognostic factors that include the patient's individual information. For each patient$p_i has\ pf_z = \{pf_1, pf2, \dots, pfz\}$, where$pf_z$ is the predictive and prognostic factors, z is the total number of factors, $i = \{1, 2, 3, \dots, n\}$is the number of patients and n is the total number of patients in dataset D. Finally, D contains rows of patient $P_i$and the columns arepf_z. Ideally, the predictive and prognostic values should have been previously studied and validated by using a group of patients with known responses to molecular targeted therapy.The following are possible factors that can be used as inputs for the system:

- Clinical Factors: clinical factors are any specifications or measurements related to the patient, such as spleen size or agethat can be observed or obtained by clinicians. Clinical factors are collected on the day of diagnosis.

- Biological Factors: patients are not biologically identical. Biological factors include blood cell counts or the activity levels of molecular parts. Multiple biological factors may influence the treatment response, and these may not be directly related to treatment outcomes. In practice, pathological tests are considered a primary pre-treatment step for individual patients. Including an individual's information helps in realizing more specific treatment. In addition, currently, most interesting developments are related to real-time quantitative polymerase chain reaction (RQ-PCR) analysis for DNA sequences. To

monitor the molecular response, RQ-PCR, a sensitive technique, is used to quantify the level of mRNA transcripts in the peripheral blood of patients.

- Predictive Assays: assays are used to predict possible outcomes following treatment [62]. Predictive assays can guide a clinician in making a rational choice of therapy at diagnosis. Identifying predictive assays, especially genetics-based ones, in a diverse group of individuals may provide substantial knowledge about the mechanisms of individual differences in responses to molecular targeted therapy.Groups whose assays show similar results can be treated similarly. Studying significant differences by the statistics of different thresholds for identifying groups of response and using a threshold less than a cut-off point (e.g. $p< 0.5$) are commonly used approaches in medical study. Here, we used a predefined threshold from publications and applied it to our dataset.

2.3 Inclusion and Exclusion Criteria

Inclusion criteria identify the base standard for including patient data in the study,and exclusion criteria exclude patient data based on predefined requirements. The reuse of medical research data requires refinement of patients in the secondary analysis. Inclusion and exclusion criteria help researchers in optimizing existing medical data to make it suitable for new research. In clinical trial data, patients adhere to research protocols and some standards. This clinical trial data usually uses information available for reuse in a second analysis if it first satisfies a new study's inclusion and exclusion criteria. Understanding data is important to obtain accurate results and establish effective research. Building personalized medicine, as a predictive model to predict molecular responses to molecular targeted therapy should follow these criteria in the reuse of medical research data:

- In clinical trials of molecular targeted therapy, patients from different trials must follow the same treatment protocol.
- Molecular responses must be monitored according tointernational standards based on a pre-identified treatment guide.
- Input data should be filtered at the diagnosis or pre-treatment stage.
- Output prediction end-point should be identified.
- If there is a disease phase, patients should be in the same phase or stage.

Some clinical trials provide useful information by changing the values of some stored data under the guidance of human experts. For example, a patient group is switched to a second-line treatment during the trial. This group may have responded unsuccessfully to the frontline treatment and will therefore be defined as a negative responder to the selected treatment.

The previous list of inclusion and exclusion criteria vary from study to study based on the specific research question under investigation and how much information can be analyzed from the existing resources by collaborating with the domain expert.

2.4 Reformat using Domain Knowledge

Defining the prognostic and predictive factors and reformatting them using domain knowledge is an important process as the final rules' interpretability will be based on the earlier categories. We reformatted the factor values of text, numeral or mixed-type stored data by using existing knowledge such as standard boundaries of blood counts, domain knowledge of clinical expertise, risk categories and previous medical publications. Although some machine learning techniques can handle continuous predictor values, reformatting the data by categorizing each factor in the dataset into subgroups can help improve the comprehensibility of the final model. For each

predictive factor,we reformat the values of factors into categoryC_d,where$d$ is the number of subcategories that represent values in the range using existing knowledge and confirmation by experts in the field.For example, we categorized the value of $pf = IC50$ intofour categories,$d = 4$: $C_1 \leq 0.5$µMasGroup 1, $C_2 > 0.5\ and \leq 0.7$µMas Group 2, $C_3 > 0.7\ and \leq 0.95$µMasGroup 3 and $C_4 > 0.95$µMas Group 4. Then, we reformatted the index number of the category. If the final model selects a relevant predictive factor, we can use these categories to distinguish the predictive group on test patients.

2.5 Modeling

The framework is designed to deal with prior knowledge and extracting knowledge from a new source of data. Here, prognostic model used the prior knowledge that is collected from existing prognostic score systems while predictive model used medical data to extract knowledge in the form of clinical prediction rules.

2.5.1 Dealing with Prior Knowledge (Prognostic Model)

The development of the prognostic model was divided into two stages: (1) primary analysis and (2) combined model development.Figure 4. shows the schema for the prognostic model. In primary analysis, available prognostic factors, scores and models are surveyed.

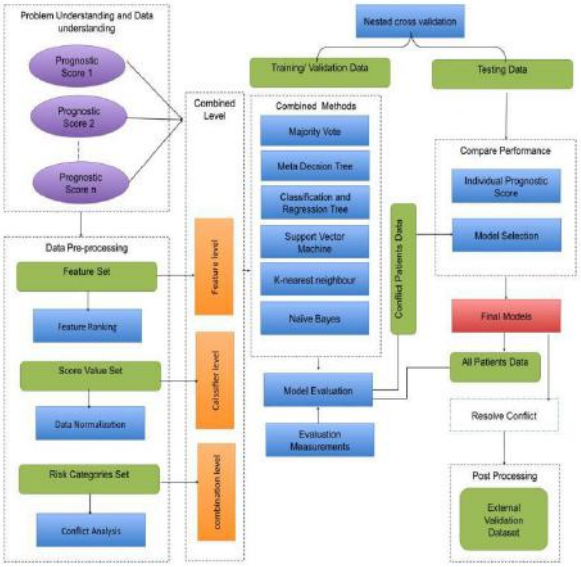


Figure 4.The schema for the prognostic model.

An information matrix is created by adding the validated prognostic method in rows and columns representing the available information from each method. Primary analysis involves investigating the consistency between the prognostic methods by identifying the obtained risk outcomes using this prognostic method. Inconsistency occurs when two different risk categories are applied to the same patient,and it is observed that one prognostic score classifies the patient in one group and the other one contradicts the first classification. The consistency test can be briefly described as follows:

1. Calculate the risk outcomes using prognostic factors included in the prognostic score equation (if available) for all patients in dataset D.

2. Create dataset DPcontaining rows of patient $P_i$and the columns arePS_m, where PS is the prognostic score methodsand m the number of validated prognostic methods.
3. Repeat step (1) for all m.
4. Create an information matrix$IM_1$for risk outcomes from DP to identify the number of risk groups resulting from step (1) calculation.
5. Analyze the risk categories obtained fromIM_1, where R_1 is the minimum and R_2 is the maximum number of categories obtained from all prognostic methods.
6. If R_1 is equal to R_2, go to step (11).
7. CombinedR_2 risk groups to be equal to R_1 groups, where the groups can include all possible combinations of risk groups from R_2.
8. Evaluate the accuracy performance after combining risk groups.
9. Select the combined risk group with the highest accuracy.
10. Combine the risk categories to be equal toR_1in the PS_mthat have more than R_1risk groups.
11. Test consistency in risk outcomes obtained from prognostic methods for each P_i in DP.
12. List in the information matrix $IM_2$the possible combination of risk outcomes, which the number of rows is R_1 raised toPS_m.
13. If $PS_1 = PS_2 = \dots\dots PS_m$results have similar risk categories, P_i is considered in a consistent risk group. However, if at least one PS_m has a different risk outcome for the same patient, P_i is considered in a conflict group.

Moving to the second stage of combining models, the steps briefly described as follows:

1. In step (1), three possible data types could be generated:

    - DP_1: risk categories set as categorical data type resulting from validated risk groups in prognostic model outcomes.
    - DP_2: score values set as a continuous data type resulting from applying the equation used to calculate the score for the prognostic method.
    - DP_3: prognostic factors (feature set) as both data types (categorical or continuous).

Normalize DP_2 and DP_3: Normalization involves mapping the data values into the interval [0-1], where the minimum value is 0 and the maximum value, 1[63]. Normalization is an important process in this study for several reasons. First, the data from prognostic scores is usually not of the same scale. Grouping and comparing patients into logical descriptions using different intervalsin the problem space could make it difficult for the learning algorithm(especially the clustering base algorithm) to learn. Therefore, arranging the data into logical groups based on a unified scale among scores can help in describing the relation between risk groups. Second, for visualization, normalized prognostic score data is easy to represent in the space dimension. Finally, pre-processing real values into scaled values could help in finding a learning function from a space, such as finding the solution using the Euclidean distance between samples.

2. Select the strategies to combinemodels (prognostic methods): model selection or model fusion[64].

    - Model selection:Each model is supposed to be an expert in a specific domain of the feature space, and the selected model decides the output of the ensemble.
    - Model fusion: Each model is supposed to have complete information on the whole feature space, and we apply combiners to all outputs from the systems.

3. The performance of single PS_m would be the same in model selection. Therefore, Select the classifier fusion method[64], and two possible combiner approach can be used:

- Non-trainable combiner (data-independent) such as majority voting.
- Trainable combiner (data-dependent) methods can be used for building combinations. Here, we used common levels to combine models (prognostic scores) [64]:
    - Combination level:Mainly focuses on possible ways of combining the risk outputs of prognostic methods in an ensemble based on the information obtained from a single method.
    - Classifier level:Different machine learning techniques can be employed.
    - Feature level: Use patients' prognostic factors for machine learning.

4. Create a subset $D_c$of conflict group of patients from DP as rows represent patientP_i and columns representPS_mdata. The last column is the actual molecular responseY_i:
$$Y_{i=}\begin{cases}0 & notachivingmolcularresponse\\ 1 & achiveingmolcularresponse\end{cases}$$

5. Divide$D_c$into training and testing datasets, and apply internal validation on training data.

6. Create the combined prognostic model $M_c$specifically for$D_c$to classifyY_i using different combined functions that maximize the accuracy of the classification.

7. Evaluate the combined method by the combined function (classifier) by comparing the actual outcome Y_i withthe method outcomeY_iM_c, and measure the model performance.

8. Generate the evaluation dataset$E$, where the rows are all possible models$M_c$that are built from different combined methods and the columns contain the performance measurements accuracy, sensitivity, specificity, PPV, NPV, G-mean and F-score.

9. Rank the model $M_c$based on selected criteria.

10. Select the best models$M_c$that resolve the conflict.

11. Evaluate the selected model $M_c$onall $P_i$in DP.

12. Compare the performance of$M_c$with the previous methods using testing data.

13. Perform external validation on an unseen dataset.

2.5.2 Dealing with a New Source of Clinical Data (Predictive Model)

The schema for the predictive model is illustrated in Figure 5. Algorithms are often used to select a relevant subset of input features (in our problem, a subset of predictive factors that will deliver a highly predictive model)[59, 65].

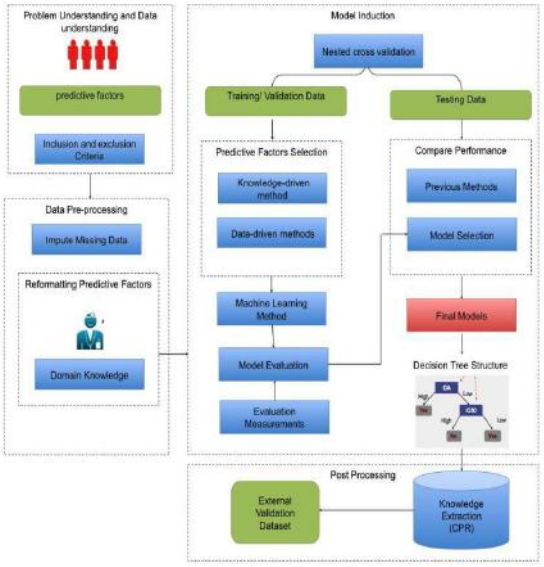


Figure 5. The schema of the predictive model

This is also very important in the context of healthcare costs, where fewer input factors imply fewer diagnostic tests to obtain the relevant predictive factors [66]. We also need to extract the relations between the most related predictive factors and try to understand whether there exist clinical rules for prediction. We divided the feature selection process into two main types:

- A knowledge-driven method for feature selection, such as existing medical literature that has published predictive factors as an informative feature or clinical expert judgment on molecular factors associated with predicting molecular response, known as manual feature selection [67].
- Data-driven methods for feature selection, known as automatic feature selection. We used the wrapper approach [58], where all subsets of the features are evaluated using a given machine learning approach.

Model induction can be briefly described as follows:

1. Apply predictive factorPf_z selection on dataset D to identify which predictive factors are important to be used as an input vector.
2. Create sample $x$from$D$where rows represent patient P_i, and the columns are predictive factorPf_z . The last column is the actual molecular responseY_i:
3. $Y_i = \begin{cases} 0 & notachivingmolcularresponse \\ 1 & achiveingmolcularresponse \end{cases}$
4. Divide x into training and testing datasets, and apply internal validation to the training data.
5. Create the predictive model$M_x$specifically for$x$to predict Y_i using the learning function that maximizes the predictionaccuracy.
6. Evaluate the predictive factor subset by the learning function (classifier) by comparing the actual outcome Y_i withthe predicted outcomes$Y_i M_x$ and measure the modelperformance.
7. Generate the evaluation dataset$E$, where the rows are all possible models$M_c$that are built from different combined methods and the columns contain the performance measurements accuracy, sensitivity, specificity, PPV, NPV, G-mean and F-score.

8. Rank the model based on selected criteria.
9. Select the best models.
10. Compare the performance of M_x with the previous methods using testing data.
11. Perform external validation on an unseen dataset.
12. Visualize the structure of the model.
13. Generate the clinical prediction rules.

Clinicalprediction rules from data: in the clinical decision process, clinical prediction rules play an important role after increasing the acceptance of evidence-based medicine. This is because of the advantages that can be gained by using clinical prediction rules. First, these rules can support and guide clinicians in patient casesinvolving complex decisions. Second, these rules,which represent expert knowledge, can be shared with clinicians in primary care. Finally, these rules can be used to train and guide new experts.The decision tree structure represents extracted production rules [68]. We used the training set of patients from which a decision tree was generated to create a structure. The path from the root of the tree to the leaf node was used to establish the conditions (IF parts), whereas every leaf of a decision tree corresponded to a response (THEN parts). The decision rules R were in the form, IF $Pf_1$is C_1 AND $Pf_2$is C_2, THEN Response is Y, where Pf_1andPf_2 are the predictive factors, C_1 and C_2 are the subcategories that belong to Pf_1andPf_2,respectively, and Y, is the class (molecular response achieved or not achieved).

2.6 Evolving Approaches

In a personalized medicine support system, there are three possible schemas for the evaluation of new knowledge: confirmation, contradiction and contribution [24]. In confirmation, no changes need to be made to the existing knowledge base. In contribution, the new knowledge is used for constructing predictive models and updating clinical prediction rules (predictive modeling steps in this study). In contradiction, the new knowledge conflicts with the old knowledge, and combined methods are used to resolve the conflict in knowledge (prognostic model steps in this study).

The intelligent techniques used in the proposed framework are selected to generalize the predictive and prognosis performance. The predictive factor wrapper approach evaluates all possible combinations of predictive factors. Therefore, new predictive factors will be equally tested in the construction of predictive models. Prognostic modeling evaluates different combination levels. Therefore, a new level of knowledge, either prognostic factors, prognostic scores or prognostic model, can be combined and evaluated. Nested cross validation is a suitable validation method for evaluation, as internal validation can be used to evaluate the new knowledge and inform the decision of whether to enhance the performance, refine the modeling or ignore the new knowledge in case of no further improvement.The following suggestions are given as solutions and will be applied in a personalized medicine support system automatically if the internal validation of the selected model does not outperform a previous method in the unseen dataset (testing data).An evolving system can change in the following ways:

- Data pre-processing phase in predictive model: the expert range of the predictive factors of conflict cases can be changed, and this should be confirmed by domain experts.
- Predictive factor selection approaches: predictive factors or prognostic factors can be updated or added.
- Combined method: different combination levels can be selected.

3. PROPOSED APPLICATIONS

Personalized medicine support systems will provide great benefits in the future as molecular data analysis improves. This type of framework based on individual data will become a basic tool used by healthcare professionalsfor selecting personalized treatments. Clinicians will be able to allocate risk profiles and predict molecular responses to selected molecular targeted therapy, thereby improving treatment selection and avoiding disease progression. The treatment of many types of cancer can be explored using this framework. Table 1 shows some of the cancers that can be treated with multiple molecular targeted therapies as well as the available studies for prognostic and predictive factors.

Table 1 of the cancers that can be treated with multiple molecular targeted therapies as well as the available prognostic and predictive factors to use for the same

Type of Cancer	**FDA-approved molecular targeted therapy [69]**	**Prediction and Prognosis**
Breast cancer	Ado-trastuzumabemtansine (Kadcyla), Everolimus (Afinitor), Lapatinib (Tykerb), Palbociclib (Ibrance), Pertuzumab (Perjeta), Trastuzumab (Herceptin)	[70]
Lung cancer	Afatinib (Gilotrif), Alectinib (Alecensa), Atezolizumab (Tecentriq), Ceritinib (Zykadia),Erlotinib (Tarceva), Gefitinib (Iressa), Necitumumab (Portrazza), Nivolumab (Opdivo), Osimertinib (Tagrisso), Pembrolizumab (Keytruda), Ramucirumab (Cyramza)	[71]
Melanoma	Aldesleukin (Proleukin),Dabrafenib (Tafinlar),Ipilimumab (Yervoy), Nivolumab (Opdivo), Pembrolizumab (Keytruda), Trametinib (Mekinist), Vemurafenib (Zelboraf)	[72]
Chronic myeloid leukemia	Imatinib (Gleevec), Nilotinib (Tasigna), Dasatinib (Sprycel), Ponatinib (Iclusig), Bosutinib (Bosulif)	[73 , 74]
Chronic lymphocytic leukemia	Alemtuzumab (Campath), Idelalisib (Zydelig), Obinutuzumab (Gazyva), Ofatumumab (Arzerra, HuMax-CD20), Rituximab (Rituxan, Mabthera), Venetoclax (Venclexta)	[75]

4. CONCLUSIONS

The framework of current learning methodologies basically includesthe essential concept of dynamic changes in the input structure of models based on domain knowledge and discoveringthe relation of importance predictive factors with patient responses. In our framework, we conduct precise modeling based on prior knowledge with interpretable and consistent meanings,and we focus on developing models with high predictive and prognostic performance.

Owing to space and time limitations, the implementation and some computational approaches have not been included in this paper. The initial future work is to implement a framework for real cancer patient data that stores molecular responses to molecular targeted therapy. Important areas for extending our proposed approach are as follows:(1) optimization techniques should be tested for model selection;(2) we trained models to maximize accuracy, but algorithms could also be developed to maximize the G-mean or F-score performance, especially in the imbalance dataset;(3) the use of a wider range of combined methods such as decision templates and the Dempster-Shafer method[76] should be enabled; and(4) online adaptive models that dynamically react to a new piece of knowledge should be developed.

In this paper, we have described a novel framework for a personalized medicine support system. A personalized medicine support system is a hybrid knowledge-based system that integrates multiple models to provide automated selection tools for individual treatment in one integrated dynamic environment. This system has been proposed to meet the requirements of clinicians and to provide a treatment guide. This system can work on oncology-based data using a data mining process for modeling; this approach can be implemented for treating other types of cancer aswell.

REFERENCES

[1] I. S. Kohane, "The twin questions of personalized medicine: who are you and whom do you most resemble?," BioMed Central, vol. 1, pp. 4.1-4.3, 2009.

[2] M. Verma, "Personalized Medicine and Cancer," Journal of Personalized Medicine, vol. 2, pp. 1-14, 2012.

[3] M. Müller, "Individualized medicine," in Clinical Pharmacology: Current Topics and Case Studies, ed: Springer, 2010.

[4] D. L. Mitchell, "Successful Implementation of Personalized Medicine: The Value, Challenges , and Effect on Patient Care," Master of Science in Health Care Administration, the Faculty of Utica College, ProQuest, 2013.

[5] M. Dietel and C. Sers, "Personalized medicine and development of targeted therapies: the upcoming challenge for diagnostic molecular pathology. A review," Virchows Archiv, vol. 448, pp. 744-755, 2006.

[6] M. Sioud and Ø. Melien, "Treatment Options and Individualized Medicine," Target Discovery and Validation Reviews and Protocols, vol. 361, pp. 327-340, 2007.

[7] H. Michael and M. Crawford, "Individualized Medicine by Biomarkers," JAMA, vol. 302, pp. 49-57, 2009.

[8] B. Shastry, "Pharmacogenetics and the concept of individualized medicine," The Pharmacogenomics Journal vol. 6, pp. 16-21, 2006.

[9] J. Corander, T. Aittokallio, S. Ripatti, and S. Kaski, "The rocky road to personalized medicine: computational and statistical challenges," Personalized Medicine: Future Medicine part of Future Science Group, vol. 9, pp. 109–114, 2012.

[10] L. M. Fagan and E. H. Shortliffe, "The Future of Computer Applications in Biomedicine," in Biomedical Informatics Computer Applications in Health Care and Biomedicine, 3 ed USA: Springer, 2006, pp. 829-847.

[11] R. T. Ng and J. Pei, "Introduction to the special issue on data mining for health informatics," SIGKDD Explor. Newsl., vol. 9, pp. 1-2, 2007.

[12] M. Kurosaki, "Pretreatment prediction of response to peginterferon plus ribavirin therapy in genotype 1 chronic hepatitis C using data mining analysis," Journal of gastroenterology (0944-1174), vol. 46, p. 104, 2011.

[13] A. D. Revell, D. Wang, M. A. Boyd, S. Emery, A. L. Pozniak, F. D. Wolf, et al., "The development of an expert system to predict virological response to HIV therapy as part of an online treatment support tool," AIDS, vol. 25, pp. 1855–1863 2011.

[14] C. E. Pedreira, L. Macrini, M. G. Land, and E. S. Costa, "New Decision Support Tool for Treatment Intensity Choice in Childhood Acute Lymphoblastic Leukemia," IEEE Transaction on Information Technology in Biomedicine, vol. 13, pp. 284-290, 2009.

[15] H. Ying, F. Lin, R. D. MacArthur, J. A. Cohn, D. C. Barth-Jones, H. Ye, et al., "A Fuzzy Discrete Event System Approach to Determining Optimal HIV/AIDS Treatment Regimens," IEEE Transactions on Information Technology in Biomedicine vol. 10, p. 663, 2006.

[16] M. A. Musen, Y. Shahar, and E. H. Shortliffe, "Clinical Decision-Support Systems," in Biomedical Informatics Computer Applications in Health Care and Biomedicine, 3rd ed USA: Springer, 2006, pp. 698-736.

[17] M. Ball, C. Weaver, and J. Kiel, "Overview of Clinical Decision Support Systems," in Healthcare Information Management Systems, Third, Ed., ed New York: Springer-Verlag, pp. 463-477.

[18] F. Alonso, L. Martínez, A. Pérez, and J. P. Valente, "Cooperation between expert knowledge and data mining discovered knowledge: Lessons learned," Expert Systems with Applications, vol. 39, pp. 7524–7535, 2012.

[19] M. A. Musen, B. Middleton, and R. A. Greenes, "Clinical Decision-Support Systems," in Biomedical Informatics: Computer Applications in Health Care and Biomedicine, E. H. Shortliffe and J. J. Cimino, Eds., ed London: Springer London, 2014, pp. 643-674.

[20] L. Aleksovska-Stojkovska and S. Loskovska, "Clinical Decision Support Systems: Medical knowledge acquisition and representation methods," in Electro/Information Technology (EIT), 2010 IEEE International Conference on, 2010, pp. 1-6.

[21] E. S. Berner, "Clinical Decision Support Systems: State of the Art," Agency for Healthcare Research and Quality, 2009.

[22] D. Levinson, "Information, computers, and clinical practice," Jama, vol. 249, pp. 607-9, 1983.

[23] Y. Hu, N. Kasabov, and W. Liang, "Personalized Information Modeling for Personalized Medicine," Springer Handbook of Bio-/Neuroinformatics, pp. 533-553, 2014.

[24] M. Kwiatkowska and A. S. Atkins, "Integrating knowledge-driven and data-driven approaches for the derivation of clinical prediction rules," in Machine Learning and Applications, 2005. Proceedings. Fourth International Conference on, 2005, p. 6 pp.

[25] N. Esfandiari, M. R. Babavalian, A.-M. E. Moghadam, and V. K. Tabar, "Knowledge discovery in medicine: Current issue and future trend," Expert Systems with Applications, vol. 41, pp. 4434-4463, 7// 2014.

[26] N. Kasabov and D. Filev, "Evolving Intelligent Systems: Methods, Learning, & Applications," International Symposium on Evolving Fuzzy Systems, IEEE, 2006.

[27] S. T. C. Wong, "Coping with Conflict in Cooperative Knowledge-Based Systems," IEEE Tranaction on Sytems, man, And Cybernetics-Part Asystems and Humans, vol. 27, 1997.

[28] M. K. Yusof and A. Azlan, "Comparative Study of Techniques in Reducing Inconsistent Data," International Journal of Database Theory and Application, vol. 5, pp. 37-47, 2012.

[29] A. H. Bond and L. Gasser, "An analysis of problems and research in DAI," in Readings in Distributed ArtiJicial Intelligence, ed, 1988.

[30] W. T. Adrian, A. Lig eza, and G. J. Nalepa, "Inconsistency Handling in Collaborative Knowledge Management," in Federated Conference on Computer Science and Information Systems, 2013, pp. 1221-1226.

[31] G. Ke, B. Fu, M. De, and K. Hipel, "A hierarchical multiple criteria model for eliciting relative preferences in conflict situations," Journal of Systems Science and Systems Engineering, vol. 21, pp. 56-76, 2012/03/01 2012.

[32] F. J. J. d. Santos and H. d. A. Camargo, "Decision Support Systems in Multicriteria Groups: an Approach Based on Fuzzy Rules," IEEE, 2010.

[33] B. S. Ahn and S. H. Choi, "Conflict resolution in a knowledge-based system using multiple attribute decision-making," 2009.

[34] M. Noor-E-Alam, T. F. Lipi, M. Ahsan Akhtar Hasin, and A. M. M. S. Ullah, "Algorithms for fuzzy multi expert multi criteria decision making (ME-MCDM)," Knowledge-Based Systems, vol. 24, pp. 367-377, 2011.

[35] H. Junming, S. Chong, L. Shuang, and W. Wanshan, "Fuzzy Case-Based Reasoning for Conflict Resolution in Collaborative Design," in Computing, Communication, Control, and Management, 2008. CCCM '08. ISECS International Colloquium on, 2008, pp. 233-237.

[36] G. Qi, W. Liu, and D. Bell, "Measuring conflict and agreement between two prioritized knowledge bases in possibilistic logic," Fuzzy Sets and Systems, vol. 161, pp. 1906-1925, 7/16/ 2010.

[37] K. Mu, K. Wang, and L. Wen, "Approaches to measuring inconsistency for stratified knowledge bases," International Journal of Approximate Reasoning, vol. 55, pp. 529-556, 1// 2014.

[38] K. Mu, Z. Jin, W. Liu, D. Zowghi, and B. Wei, "Measuring the significance of inconsistency in the Viewpoints framework," Science of Computer Programming, vol. 78, pp. 1572-1599, 9/1/ 2013.

[39] J. Grant and A. Hunter, "Measuring inconsistency in knowledgebases," Journal of Intelligent Information Systems, vol. 27, pp. 159-184, 2006.

[40] A. Hunter and S. Konieczny, "Approaches to Measuring Inconsistent Information," in Inconsistency Tolerance, L. Bertossi, A. Hunter, and T. Schaub, Eds., ed Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 191-236.

[41] N. Kasabov, "Soft Computing Methods for Global, Local and Personalized Modeling and Applications in Bioinformatics," Soft ComputingBased Modeling in Intel. Systems, vol. 196, pp. 1-18, 2009.

[42] K. M. Khalil, M. Abdel-Aziz, T. T. Nazmy, and A.-B. M. Salem, "Intelligent Techniques for Resolving Conflicts of Knowledge in Multi-Agent Decision Support Systems," presented at the Sixth International Conference on Intelligence Computing and Information Systems, Cairo, Egypt, 2014.

[43] . Australian Clinical Trials. Available: https://www.australianclinicaltrials.gov.au/what-clinical-trial/phases-clinical-trials
[44] . National Institutes of Health. https://www.nih.gov.
[45] S. Saussele and M. Pfirrmann, "Clinical Trials in Chronic Myeloid Leukemia," Curr Hematol Malig Rep pp. 109–115, 2012.
[46] B. M. Welch and K. Kawamoto, "The Need for Clinical Decision Support Integrated with the Electronic Health Record for the Clinical Application of Whole Genome Sequencing Information," J. Pers. Med., vol. 3, 2013.
[47] J. C. Weiss , S. Natarajan, P. L. Peissig, C. A. McCarty, and D. Page, "Machine learning for personalized medicine: predicting primary myocardial infarction from electronic health records," AI Magazine, vol. 33, p. 33, 2012.
[48] C. L. Overby, "A Clinical Decision Support Model for Incorporating Pharmacogenomics Knowledge Into Electronic Health Records for Drug Therapy Individualization: A Microcosm of Personalized Medicine," Doctor of Philosophy, Division of Biomedical and Health Informatics, University of Washington, ProQuest, 2011.
[49] I. Kouris, C. Tsirmpas, S. G. Mougiakakou, D. Iliopoulou, and D. Koutsouris, "E-Health towards ecumenical framework for personalized medicine via Decision Support System," in Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE, 2010, pp. 2881-2885.
[50] B. R. Hirsch and A. P. Abernethy, "Structured Decision-Making: Using Personalized Medicine to Improve the Value of Cancer Care," Journal of Personalized Medicine, vol. 3, pp. 1-13, 2013.
[51] S. Mehta, A. Shelling, A. Muthukaruppan, A. Lasham, C. Blenkiron, G. Laking, et al., "Predictive and prognostic molecular markers for cancer medicine," Therapeutic Advances in Medical Oncology, vol. 2, pp. 125-148, 2010.
[52] R. Alsolami, S. J. L. Knight, and A. Schuh, "Clinical application of targeted and genome-wide technologies: can we predict treatment responses in chronic lymphocytic leukemia?," Personalized Medicine, vol. 10, pp. 361-376, 2013/06/01 2013.
[53] A. Lopez-Chavez, A. Thomas, A. Rajan, M. Raffeld, B. Morrow, R. Kelly, et al., "Molecular profiling and targeted therapy for advanced thoracic malignancies: a biomarker-derived, multiarm, multihistology phase II basket trial," J Clin Oncol, vol. 33, pp. 1000-7, Mar 20 2015.
[54] J. M. Llovet and J. Bruix, "Molecular targeted therapies in hepatocellular carcinoma," Hepatology, vol. 48, pp. 1312-27, Oct 2008.
[55] P. P. Piccaluga, G. Martinelli, and M. Baccarani, "Advances in the treatment for haematological malignancies," Expert Opin Pharmacother, vol. 7, pp. 721-32, Apr 2006.
[56] G. M. Grimshaw, A. Szczepura, M. Hulten, F. MacDonald, N. C. Nevin, F. Sutton, et al., "Evaluation of molecular tests for prenatal diagnosis of chromosome abnormalities," Health Technol Assess, vol. 7, pp. 1-146, 2003.
[57] . National Comprehensive Cancer Network Guideline Inc.
[58] A. G. Karegowda, M. A. Jayaram, and A. S. Manjunath, "Feature Subset Selection Problem using Wrapper Approach in Supervised Learning," International Journal of Computer Applications, vol. 1, pp. 13–17, 2010.
[59] P. Perner, "Improving the accuracy of decision tree induction by feature preselection " Applied Artificial Intelligence: An International Journal, vol. 15, 2001.
[60] L. Breiman, J. Friedman, R. Olshen, and C. Stone, Classification and regression trees: Belmont, Calif : Wadsworth International Group, 1984.
[61] J. Lasserre, S. Arnold, M. Vingron, P. Reinke, and C. Hinrichs, Predicting the outcome of renal transplantation vol. 19, 2012.
[62] N. S. Russell and A. C. Begg, "Editorial radiotherapy and oncology 2002: predictive assays for normal tissue damage," Radiotherapy and Oncology, vol. 64, pp. 125-129, 8// 2002.
[63] D. Steinley, "Standardizing Variables in K-means Clustering," in Classification, Clustering, and Data Mining Applications: Proceedings of the Meeting of the International Federation of Classification Societies (IFCS), Illinois Institute of Technology, Chicago, 15–18 July 2004, D. Banks, F. R. McMorris, P. Arabie, and W. Gaul, Eds., ed Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 53-60.
[64] L. I. Kuncheva, Combining Pattern Classifiers Methods and Algorithms: John Wiley & Sons,, 2004.
[65] R. C. Prati, "Combining feature ranking algorithms through rank aggregation," in Neural Networks (IJCNN), The 2012 International Joint Conference on, 2012, pp. 1-8.

[66] M. Qian, I. Nahum-Shani, and S. A. Murphy, "Dynamic Treatment Regimes," in Modern Clinical Trial Analysis, W. Tang and X. Tu, Eds., ed New York: Springer, 2013, pp. 127-148.
[67] T.-H. Cheng, C.-P. Wei, and V. S. Tseng, "Feature Selection for Medical Data Mining: Comparisons of Expert Judgment and Automatic Approaches," in Proceedings of the 19th IEEE Symposium on Computer-Based Medical Systems (CBMS'06), 2006.
[68] J. R. Quinlan, "Generating production rules from decision trees," in Proceeding IJCAI'87 Proceedings of the 10th international joint conference on Artificial intelligence, USA, 1987, pp. 304-307
[69] R. Abramson. (2016). Overview of Targeted Therapies for Cancer. Available: https://www.mycancergenome.org/content/molecular-medicine/overview-of-targeted-therapies-for-cancer/
[70] T. Foukakis and J. Bergh. (2016). Prognostic and predictive factors in early, nonmetastatic breast cancer. Available: https://www.uptodate.com/contents/prognostic-and-predictive-factors-in-early-nonmetastatic-breast-cancer
[71] M. Paesmans, "Prognostic and predictive factors for lung cancer," Breathe, vol. 9, pp. 112-121, 2012.
[72] D. Weinstein, J. Leininger, C. Hamby, and B. Safai, "Diagnostic and Prognostic Biomarkers in Melanoma," The Journal of Clinical and Aesthetic Dermatology, vol. 7, pp. 13-24, 2014.
[73] K. Sweet, L. Zhang, and J. Pinilla-Ibarz, "Biomarkers for determining the prognosis in chronic myelogenous leukemia," Journal of Hematology & Oncology, vol. 6, 2013.
[74] D. L. White and T. P. Hughes, "Predicting the response of CML patients to tyrosine kinase inhibitor therapy," Current Hematologic Malignancy Reports, vol. 6, pp. 88-95, 2011.
[75] D. Mertens and S. Stilgenbauer, "Prognostic and Predictive Factors in Patients With Chronic Lymphocytic Leukemia: Relevant in the Era of Novel Treatment Approaches?," Journal of Clinical Oncology, vol. 32, pp. 869-872, 2014.
[76] A. Altmann, M. Rosen-Zvi, M. Prosperi, E. Aharoni, H. Neuvirth, E. Schülter, et al., "Comparison of Classifier Fusion Methods for Predicting Response to Anti HIV-1 Therapy," PLoS ONE, vol. 3, 2008.

Linking Chapter 3, 4 and 5:

In the previous paper, “Personalized Medicine Support System: Resolving Conflict in Allocation to Risk Groups and Predicting Patient Molecular Response to Targeted Therapy,” we found that the use of prognostic and predictive factors are important prior to treatment. The personalized medicine support system framework was proposed as a knowledge-based clinical decision support system where knowledge is the main component in constructing the system. Our main question was how could one develop a personalized medicine support system that allocates risk groups and predicts the molecular response to molecular targeted therapy, where consistency in decisions is paramount. We also outlined the main processes of our personalized medicine support system. Later, the steps were demonstrated in detail to help computing researchers follow them when constructing such a system. We also concluded that there was not enough time and space to include the implementation and the full amount of computational approaches using real medical data. As a result, we decided that future work on this subject should include the implementation of the framework for actual cancer patient data that will store molecular responses to molecular targeted therapy.

In chapter 3, figure 1 demonstrates the process flow of a personalized medicine support system and connects two models: prognostic model details in chapter 4 and predictive model details in chapter 5. Figure 2 depicts the data mining process for developing a prognostic model of the study in chapter 4, and figure 3 shows the data mining process for developing a predictive model of the study in chapter 5. Finally, figure 4 contains the schema for a prognostic model that will be covered in chapter 4, while figure 5 presents the schema for a predictive model included in chapter 5.

In the upcoming chapters, we evaluate the proposed framework for a personalized medicine support system. Chapter 4 has two parts, which represent the two stages for building a prognostic model using prior knowledge in our proposed framework (Section 2.5.1 in Chapter 3), while Chapter 5 explains how to deal with a new source of data and build a predictive model (Section 2.5.2 in Chapter 3).

In Chapter 4, our first paper entitled "Consistency Test between Scoring Systems for Predicting Outcomes of Chronic Myeloid Leukemia in a Saudi Population" is detailed. We evaluate the common scoring systems that are used to identify risk groups based on patients' molecular responses. In this study, we analyzed 104 CML patients and monitored them for any major molecular responses using RQ-PCR. This cohort was used as an example to present the issue of inconsistency between existing score systems. In this paper's introduction, we provided background information related to CML and how it affects the bone marrow. We also discussed how the prognostic scores being evaluated are used in developing and identifying the best CML treatments for individual patients. In the Materials and Methods Section of the paper, we presented the study population that was evaluated and the scoring systems that were used on CML patients. This section includes a two-step analysis process that consisted of studying the prognostic index using combined groups and a consistency analysis between the risk categories obtained from the scoring systems. Model selection and model fusion were both acceptable for resolving any conflict generated by the score systems' outcome. In the Results and Discussion Sections, we evaluated CML scoring systems based on patients' molecular responses to determine which prognostic scores best apply when a conflict prognosis was generated from prognostic scores. This section also contained multiple data tables and visuals that display factors, equations, risk categories, and other variables used during this particular study. Another table within this section displays data pertaining to patients that were placed in various risk groups depending on their scores. We claim in this section that the study was the first to investigate the conflicts within the

data and also to compare the four validated scoring systems. At the end of the paper, the Hasford score outperformed all other scores in cases where the clinician depended on a model selection approach prior to treatment. However, the conflict remained an issue when the clinician used common score system outcomes. Therefore, in the second paper, we implemented methods to investigate any further enhancement of using the prognostic scores to resolve the conflict in decision.

In our second paper, entitled ”Combined Value of Validated Prognostic Scores and Resolving Conflict in Allocation to Risk Groups in Chronic Myeloid Leukemia Patients,” we aimed to improve the current level of knowledge by adding an additional level of analysis and discovering a conflict group. This group has not yet been studied clinically in other current medical research. The combined method improved clinical knowledge as it increased our confidence that we were making the right decision because various scores are included and combined to reach a final conclusion. We began the paper with medical background information about CML and available prognostic scoring systems that have been used in newly diagnosed CML patients to assess their risk profile. The CML medical literature validates the use of the Sokal, Hasford, EUTOS and ELTS. Therefore, we combined only the common scores. Our study attempts to develop a framework that automatically filter the models and select the highest performance model that learn the best conflict resolution strategy between prognostic scores in patients with conflicting prediction and update entire architecture when adapt a new prognostic model. The study presents several methods in detail to address the conflict at different levels. The method section contains two different datasets used to perform the analysis, while the results section uses tables and visuals to present the results from our experiments. Any disagreement between common score systems was displayed in the figures. The performance for individual scores and also the comparisons between the single score performance and the proposed combined methods in the conflicted group were discussed in detail. The automatic selection for the final models was applied to

external data to examine their performance. The strength of our study is that it was the first time a conflicted group was introduced and also the first time in the CML literature that research attempted to resolve this conflict. Our proposed combined method for prognostic scores is expected to improve the quality of treatment decisions. Our findings suggested that combined methods have a great impact in identifying risk groups before treatment in conflicted CML patients.

Chapter 4 Part 1: Consistency Test between Scoring Systems for Predicting Outcomes of Chronic Myeloid Leukemia in a Saudi Population Treated with Imatinib

Statement of Authorship

Title of Paper	Consistency Test between Scoring Systems for Predicting Outcomes of Chronic Myeloid Leukemia in a Saudi Population Treated with Imatinib
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Haneen R. Banjar and Enaam Alsobhi, "Consistency Test between Scoring Systems for Predicting Outcomes of Chronic Myeloid Leukemia in a Saudi Population Treated with Imatinib," International Scholarly Research Notices, vol. 2017, Article ID 1076493, 6 pages, 2017. doi:10.1155/2017/1076493

Principal Author

Name of Principal Author (Candidate)	Haneen Reda Banjar		
Contribution to the Paper	HB designed and performed the research, analyzed the data, and wrote the manuscript.		
Overall percentage (%)	90%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature	[redacted]	Date	14 Feb 2017

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate in include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Enaam Al Sobahi		
Contribution to the Paper	EA is a principle investigator of the data. EA revised the manuscript and approved the version of the manuscript to be published.		
Signature	[redacted]	Date	18 Feb 2017

Hindawi
International Scholarly Research Notices
Volume 2017, Article ID 1076493, 6 pages
https://doi.org/10.1155/2017/1076493

Research Article

Consistency Test between Scoring Systems for Predicting Outcomes of Chronic Myeloid Leukemia in a Saudi Population Treated with Imatinib

Haneen R. Banjar[1,2] and Enaam Alsobhi[3]

[1]*The Department of Computer Science, King Abdulaziz University, Jeddah, Saudi Arabia*
[2]*School of Computer Science, University of Adelaide, Adelaide, SA, Australia*
[3]*Pathology and Laboratory Medicine, King Abdulaziz Medical City, National Guard Health Affairs, Jeddah, Saudi Arabia*

Correspondence should be addressed to Haneen R. Banjar; hrbanjar@kau.edu.sa and Enaam Alsobhi; ensobhi@hotmail.com

Received 1 September 2016; Revised 27 December 2016; Accepted 15 January 2017; Published 13 February 2017

Academic Editor: Haifa Al-Ali

Copyright © 2017 Haneen R. Banjar and Enaam Alsobhi. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Inconsistency in prognostic scores occurs where two different risk categories are applied to the same chronic myeloid leukemia (CML) patient. This study evaluated common scoring systems for identifying risk groups based on patients' molecular responses to select the best prognostic score when conflict prognoses are obtained from patient profiles. We analyzed 104 patients diagnosed with CML and treated at King Abdulaziz Medical City, Saudi Arabia, who were monitored for major molecular response (achieving a *BCR-ABL1* transcript level equal to or less than 0.1%) by Real-Time Quantitative Polymerase Chain Reaction (RQ-PCR), and their risk profiles were identified using Sokal, Hasford, EUTOS, and ELTS scores based on the patients' clinical and hematological parameters at diagnosis. Our results found that the Hasford score outperformed other scores in identifying risk categories for conflict groups, with an accuracy of 63%.

1. Introduction

The Australian Institute of Health and Welfare (AIHW) classified myeloid cancers as the ninth most commonly diagnosed cancer in 2016, with more than 3,600 cases in Australia [1]. Chronic myeloid leukemia (CML) is also known as chronic myelogenous leukemia or chronic granulocytic leukemia. The bone marrow produces an unusual number of white blood cells. The bone marrow could produce an excessive number of immature white blood cells and lead to progressive disease. Consequently, the bone marrow cannot make enough red cells, normal white cells, and platelets [2].

Prognostic scores in patients with CML are used to stratify CML patients according to risk profile to ensure appropriate treatment. Historically, the science of prognostication has evolved rapidly, and various scoring systems have been developed to optimize the use of clinical experience in CML treatment. These scores were developed using logistic regression with the selection of the patients' clinical and hematological parameters at diagnosis. The common prognostic scores have shown variable correlation with complete cytogenetic response (CCyR) [3–8] and major molecular response (MMR) [9–12]. Although the investigation compared the prognostic value of the validated scoring systems in overall survival (OS), event free survival (EFS) or optimal response in CML patients who receive frontline imatinib, applying the established prognostic scores in a comparative fashion and questioning the value of scoring systems, especially with regard to inconsistency in risk category, has not been considered in previous studies.

The European LeukemiaNet (ELN) current recommendations for the management of CML are basically addressed to the goal of achieving an at least MMR [13]. As newly diagnosed CML patients should be stratified based on the available prognostic scoring systems, we considered the risk

Table 1: Characteristics of 95 patients with CML at diagnosis.

Factor	Median	Range	SD
Age (yrs)	40.21	18–74	15.13
Spleen size (cm, BCM)	8.33	0–25	7.53
Platelet count ($\times 10^9$/L)	510.97	4.42–2876	439.88
Basophils (%)	1.32	0–7	1.10
Eosinophils (%)	0.83	0–0.07	1.24
Peripheral blast (%)	1.50	0–10	1.39

SD = standard deviation; BCM = below costal margin.

groups might be studied based on the MMR outcomes. This is needed to evaluate the clinical impact of the existing prognostic scores by comparison of prognostic risk groups with primary concern on consistency in prognostic scores outcomes. Inconsistency occurs when two different risk categories are applied to the same CML patient; that is, one prognostic score classifies the patient in one group and the other score contradicts the first classification. Consistency in prognostic scores used to estimate the risk group of CML patients before therapy commencement can increase clinician trust in the treatment decision and play important role in modern medicine for CML changing treatment modalities [14, 15]. However, conflict between prognostic scores is observed in some CML patients. Thus, it is important to study consistency between prognostic score categories used to allocate CML patients to risk groups in order to support clinician decision-making.

Our analysis evaluated the different scores outcomes with the long-term molecular response in patients treated with imatinib to determine which was the best prognostic score to apply where a conflict prognosis was generated by prognostic scores.

2. Materials and Methods

2.1. Study Population. Participants in this study were members of the Saudi population diagnosed with CML and treated at King Abdulaziz Medical City, Jeddah [16]. A total of 104 CML patients received 400 mg imatinib as the initial therapy. Patient characteristics are described in Table 1. All of the patients monitored their MMR in time points defined by ELN [13] where MMR is defined as achieving a *BCR-ABL1* transcript level equal to or less than 0.1% at 12 months by RQ-PCR.

2.2. Scoring Systems in CML. Four common prognostic scoring systems are available for CML patients prior to commencing therapy: (1) the Sokal score [17], (2) the Hasford score [14], (3) the European Treatment and Outcome Study (EUTOS) score [15], and (4) the EUTOS long-term survival (ELTS) score [18]. These four scores ascertain the level of risk for CML patients by running multivariable regression analysis. Prognostic scores were calculated using formulas in Table 2, based on the patients' clinical and hematological parameters at diagnosis.

The analysis is conducted in two steps: (1) studying the prognostic index using combined groups and (2) consistency analysis between the risk categories obtained from the scoring systems. First, from Table 2, the EUTOS score is the only score that classifies CML patients into low risk and high risk. The number of categories in comparative prognostic scores in Sokal, Hasford, EUTOS, and ELTS was three, three, two, and three, respectively. Accuracy was measured on prognostic score data by assuming two different combined groups: (1) low and intermediate risk in Sokal, Hasford, and ELTS scores as low risk and (2) intermediate and high risk in Sokal, Hasford, and ELTS scores as high risk.

Secondly, in consistency analysis, the combined category is selected based on the higher-accuracy results from combined groups to study the inconsistency between scoring systems. We are dealing with two models advising on the same patient. Each score may provide an index that conflicts with the other. The patients were classified into a consistency group or an inconsistency group. The consistency group included patients who observed consistent risk categorization from scoring systems, while the inconsistency group included patients who observed inconsistent risk categorization from scoring systems. The possible combination of risk categories for S scoring systems is N (number of the risk categories) raised to S power. The number of patients belongs to each molecular response groups is included to calculate the accuracy and determine which is the most accurate scoring system that can be used in a conflict group.

3. Results and Discussion

This study presents the analysis of each scoring system for distinguishing patients. We evaluated scoring systems in CML for identifying risk categories based on patients' molecular responses to determine which was the best prognostic score to apply where a conflict prognosis was generated by prognostic scores.

Of the 104 CML patients included in this study, the data of 9 patients were removed due to incomplete MMR data, to improve overall data quality. Of the 95 patients with complete data, 33 (34%) did not achieve MMR, while 62 (65%) did achieve MMR. The number of CML patients per prognostic score included in the two different combined methods is shown in Table 3.

It is clearly observed that the combined method of low and intermediate risk in Sokal, Hasford, and ELTS score as low risk achieved higher accuracy than the second combined method of intermediate and high risk in Sokal, Hasford, and ELTS score as high risk. Comparison of the accuracies in Sokal was 62.10% versus 48.42%, Hasford was 67.37% versus 58.94%, and ELTS was 62.10% versus 61.05%. Indeed, the ELN [13] recommended dividing patients into low-risk (including intermediate) and high-risk populations in the management of CML. Basically, there is insufficient evidence to prove intermediate risk patients behave differently from low-risk patients. A study used the combined method of low and intermediate in one risk group to evaluate Sokal and

Table 2: The current methods used to identify risk in CML.

Study	Factors	Scoring systems		Risk categories	Target prediction	Data and results
		Equation	Method			
Sokal score, Sokal et al. [17]	Age, spleen size (cm), blast (%), and platelets (10^9/L)	$\exp(0.0116 \times (\text{age [years]} - 43.4)) + 0.0345 \times (\text{spleen size [cm]} - 7.51) + (0.188 \times ((\text{platelets } [10^9/\text{L}]/700)^2 - 0.563)) + (0.0887 \times (\text{blasts [\%]} - 2.10))$	Multivariate analysis of survival	Low-risk score < 0.8, of patients, intermediate-risk in 0.8–1.2, and high-risk > 1.2	Risk groups for chemotherapy	Six European and American sources ($n = 813$), low 39%, intermediate 38%, and high 23%
Hasford score, Hasford et al. [14]	Age, spleen size (cm), blasts (%), eosinophils (%), basophils (%), and platelets (10^9/L)	$(0.6666 \times \text{age [0 when age} < 50 \text{ years; 1 otherwise}]) + (0.0420 \times \text{spleen size [cm]}) + (0.0584 \times \text{blasts [\%]}) + (0.0413 \times \text{eosinophils [\%]}) + (0.2039 \times \text{basophils [0 when basophils} < 3\%; \text{1 otherwise}]) + (1.0956 \times \text{platelet count [0 when platelets} < 1500 \times 10^9/\text{L; 1 otherwise}]) \times 1000$	Multivariate analysis of survival	Low-risk score ≤ 780, intermediate-risk in 781–1480, and high-risk ≥ 1481	Risk groups for interferon alpha alone	14 studies ($n = 981$), low 40.6%, intermediate 44.7%, and high 14.6%
EUTOS score, Hasford et al. [15]	Basophils (%) and spleen size (cm)	$(7 \times \text{basophil [\%]}) + (4 \times \text{spleen [cm]})$	Multivariate analysis of response	Low-risk score < 87 and high-risk ≥ 87	CCgR at 18 months to imatinib	Five national study groups ($n = 2,060$), low 79% and high 21%
ELTS score, Pfirrmann et al. [18]	Age, spleen size (cm), blast (%), and platelets (10^9/L)	$0.0025 \times (\text{age in completed years})^3 + 0.0615 \times \text{spleen size below costal margin} + 0.1052 \times \text{blasts in peripheral blood} + 0.4104 \times (\text{platelet count}/1000)^{-0.5}$	Multivariate analysis of response	Low-risk score ≤ 1.5680, intermediate-risk in 1.5680–2.2185, and high-risk > 2.2185	Long-term survival	($n = 2,205$), low 61%, intermediate 27%, and high 12%

Table 3: The number of patients in different risk groups as per calculated scores.

	n	Not achieving MMR	Achieving MMR	Accuracy
		Combined groups (1)		
Sokal score risk group				
High	25	11	14	62.10
Low and intermediate	70	22	48	
Hasford score risk group				
High	6	4	2	67.37
Low and intermediate	89	29	60	
EUTOS score risk group				
High	10	4	6	63.15
Low	85	29	56	
ELTS score risk group				
High	17	7	10	62.10
Low and intermediate	78	26	52	
		Combined groups (2)		
Sokal score risk group				
Intermediate and high	62	23	39	48.42
Low	33	10	23	
Hasford score risk group				
Intermediate and high	46	20	26	58.94
Low	49	13	36	
EUTOS score risk group				
High	10	4	6	63.15
Low	85	29	56	
ELTS score risk group				
Intermediate and high	42	19	23	61.05
Low	53	14	39	

EUOS to predict optimal response [12]. Therefore, we used the first combined method in the consistency analysis.

In Table 4, there will be sixteen rows in our analysis ($2^4 = 16$). The consensus group involved 65 (68.42%) patients, and there were 30 (31.58%) patients in the conflict group. To identify the most appropriate prognostic score to use when there is conflict between prognostic scores, we compared the number of patients belonging to each group. Table 4 shows that, in the consensus group, both prognostic scores incorrectly predict CML risk group in 21% (19 patients did not achieve MMR, while all scores classified them in the low-risk group, and 1 achieved MMR, while all scores classified this patient in the high-risk group) of cases. In the conflict group, the Sokal and ELTS scores predicted MMR accurately in 46.67% (14 of 30) of patients, while the EUTOS score predicted MMR accurately in 50% (15 of 30) of patients. The highest accuracy of 63.33% (19 of 30) of patients was obtained by the Hasford score for predicting the risk category. However, the accuracy achieved by the Hasford score in both groups (consensus and conflict groups) was the lowest (58.95%) among the other scores (Sokal's accuracy: 62.11%, EUTOS's accuracy: 63.16%, and ELTS's accuracy: 62.11%).

Although the results show that the Hasford performance in the consensus and conflict groups was not recommended, the Hasford score accuracy percentage (63%) shows that Hasford may be useful in identifying risk group in conflict CML patients. In the conflict group, the Hasford prognostic score identified more low-risk categories for CML patients and few high-risk patients, while the Sokal score identified more high-risk patients and few low-risk patients. Only one study [3] reported conflict in 22 CML patients. This study also supports our finding as they found that a majority of patients corroborated better with the Hasford score [14] than the Sokal and EUTOS scores. Previous studies compared and assessed the Sokal, Hasford, and EUTOS but not ELTS scores in investigating consistency between the scoring systems. Our study is the first to investigate the conflict and compare the four validated scoring systems. Comparison of prognostic scores shows the diversity in scoring, but in future work, we intend to implement advanced methods from computer science to resolve conflict. Thus, a new scoring system combining the power of currently available prognostic scores may further help increase accuracy of identifying risk groups.

Table 4: The consistency/inconsistency of prognostic scores for predicting major molecular response.

Analysis groups	Score prediction				MMR		Total
	Sokal	Hasford	EUTOS	ELTS	Achieving MMR, $n = 62$	Not achieving MMR, $n = 33$	
Conflict group	Low and intermediate	Low and intermediate	Low	High	5	2	30
	Low and intermediate	Low and intermediate	High	Low and intermediate	0	0	
	Low and intermediate	Low and intermediate	High	High	0	1	
	Low and intermediate	High	Low	Low and intermediate	0	0	
	Low and intermediate	High	Low	High	0	0	
	Low and intermediate	High	High	Low and intermediate	0	0	
	Low and intermediate	High	High	High	0	0	
	High	Low and intermediate	Low	Low and intermediate	7	5	
	High	Low and intermediate	Low	High	0	1	
	High	Low and intermediate	High	Low and intermediate	1	1	
	High	Low and intermediate	High	High	4	0	
	High	High	Low	Low and intermediate	1	1	
	High	High	Low	High	0	1	
	High	High	High	Low and intermediate	0	0	
Consensus group	Low and intermediate	Low and intermediate	Low	Low and intermediate	43	19	65
	High	High	High	High	1	2	
Single score in consensus group and conflict group	Low and intermediate				48	22	95
	High				14	11	
		Low and intermediate			36	13	
		High			26	20	
			Low		56	29	
			High		6	4	
				Low and intermediate	52	26	
				High	10	7	
Single score in conflict group	Low and intermediate				5	3	30
	High				13	9	
		Low and intermediate			17	10	
		High			1	2	
			Low		13	10	
			High		5	2	
				Low and intermediate	9	7	
				High	9	5	

Competing Interests

There is no conflict of interests regarding the publication of this paper.

Authors' Contributions

Haneen R. Banjar designed and performed the research, analyzed the data, and wrote the manuscript. Enaam Alsobhi is a principle investigator of the data. Both authors contributed substantially to drafts and revisions to the manuscript. They also approved the current revised version.

References

[1] Leukaemia in Australia: The Australian Institute of Health and Welfare, 2016, http://www.aihw.gov.au/cancer/leukaemia/.

[2] J. F. Apperley, "Part I: mechanisms of resistance to imatinib in chronic myeloid leukaemia," *The Lancet Oncology*, vol. 8, no. 11, pp. 1018–1029, 2007.

[3] S. K. Sinha, S. Sinha, P. K. Mandal, N. K. Bhattacharyya, A. Pandey, and P. Gupta, "A comparative study of Hasford score and Sokal index in prognostication of the novo chronic myeloid leukemia patients and a search for new prognostic markers," *Indian Journal of Pathology and Microbiology*, vol. 56, no. 3, pp. 216–220, 2013.

[4] M. Tiribelli, M. Bonifacio, E. Calistri et al., "EUTOS score identifies cases with poor outcome in patients with early chronic phase chronic myeloid leukemia, though not predictive for optimal response to imatinib," *Blood*, vol. 120, no. 21, article 3778, 2012.

[5] K. B. Pagnano, I. Lorand-Metze, E. C. M. Miranda et al., "EUTOS score is predictive of event-free survival, but not for progression-free and overall survival in patients with early chronic phase chronic myeloid leukemia treated with imatinib: a single institution experience," *Blood*, vol. 120, no. 21, article 1681, 2012.

[6] M. Breccia, P. Finsinger, G. Loglisci et al., "The EUTOS score identifies chronic myeloid leukeamia patients with poor prognosis treated with imatinib first or second line," *Leukemia Research*, vol. 36, no. 9, pp. e209–e210, 2012.

[7] A. A. Oyekunle, P. O. Osho, J. C. Aneke, L. Salawu, and M. A. Durosinmi, "The predictive value of the Sokal and Hasford scoring systems in chronic myeloid leukaemia in the imatinib era," *Journal of Hematological Malignancies*, vol. 2, no. 2, pp. 25–33, 2012.

[8] V. S. Hoffmann, M. Baccarani, D. Lindoerfer et al., "The EUTOS prognostic score: review and validation in 1288 patients with CML treated frontline with imatinib," *Leukemia*, vol. 27, no. 10, pp. 2016–2022, 2013.

[9] S. Saussele, M. Lauseker, V. Hoffmann et al., "Prediction of molecular response of chronic phase CML patients by the EUTOS score: results of the randomized CML-Study IV," *Blood*, vol. 118, no. 21, article 3762, 2011.

[10] J. Dybko, B. Jaźwiec, O. Haus et al., "The hasford score may predict molecular response in chronic myeloid leukemia patients: a single institution experience," *Disease Markers*, vol. 2016, Article ID 7531472, 5 pages, 2016.

[11] S. Ganguly, K. C. Lakshmaiah, L. A. Jacob, S. Babu, L. Dasappa, and K. S. Govind Babu, "Performance of sokal and eutos scores for predicting cytogenetic and molecular response in newly diagnosed chronic myeloid leukemia-chronic phase patients on imatinib," *Indian Journal of Hematology and Blood Transfusion*, pp. 1–5, 2016.

[12] M. Bonifacio, G. Binotto, E. Calistri, E. Maino, and M. Tiribelli, "EUTOS score predicts early optimal response to imatinib according to the revised 2013 ELN recommendations," *Annals of Hematology*, vol. 93, no. 1, pp. 163–164, 2014.

[13] M. Baccarani, M. W. Deininger, G. Rosti et al., "European LeukemiaNet recommendations for the management of chronic myeloid leukemia: 2013," *Blood*, vol. 122, no. 6, pp. 872–884, 2013.

[14] J. Hasford, M. Pfirrmann, R. Hehlmann et al., "A new prognostic score for survival of patients with chronic myeloid leukemia treated with interferon alfa," *Journal of the National Cancer Institute*, vol. 90, no. 11, pp. 850–858, 1998.

[15] J. Hasford, M. Baccarani, V. Hoffmann et al., "Predicting complete cytogenetic response and subsequent progression-free survival in 2060 patients with CML on imatinib treatment: the EUTOS score," *Blood*, vol. 118, no. 3, pp. 686–692, 2011.

[16] E. Alsobhi, M. B. Abrar, M. Abdelaal et al., "Response to imatinib therapy in adult patients with chronic myeloid leukemia in Saudi population: A Single-center Study," *Leukemia & Lymphoma*, vol. 56, no. 4, pp. 882–886, 2015.

[17] J. E. Sokal, E. B. Cox, M. Baccarani et al., "Prognostic discrimination in 'good-risk' chronic granulocytic leukemia," *Blood*, vol. 63, no. 4, pp. 789–799, 1984.

[18] M. Pfirrmann, M. Baccarani, S. Saussele et al., "Prognosis of long-term survival considering disease-specific death in patients with chronic myeloid leukemia," *Leukemia*, vol. 30, no. 1, pp. 48–56, 2016.

Chapter 4 Part 2: Combining Validated Prognostic Scores and Resolving Conflict in Allocation to Risk Groups in Chronic Myeloid Leukaemia Patients

Statement of Authorship

Title of Paper	Combined Value of Validated Prognostic Scores and Resolving Conflict in Allocation to Risk Groups in Chronic Myeloid Leukaemia Patients
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Banjar, H., Hughes, T., Brown, F., Adelson, D., White, D., Alsobhi, E. and Chaudhri. N. Combined Value of Validated Prognostic Scores and Resolving Conflict in Allocation to Risk Groups in Chronic Myeloid Leukaemia Patients.

Principal Author

Name of Principal Author (Candidate)	Haneen Reda Banjar	
Contribution to the Paper	HB designed and performed the research, data curation, formal analysis, investigation, methodology, project administration, resources, software, and validation. HB wrote the original draft.	
Overall percentage (%)	90%	
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.	
Signature	[redacted]	Date 2nd Dec 2017

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate to include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Timothy Hughes	
Contribution to the Paper	TH contributed to the research, validation, data curation, drafts, and revised the manuscript and approved the version of the manuscript to be published.	
Signature	[redacted]	Date 9/1/2018

Name of Co-Author	Fred Brown
Contribution to the Paper	FB revised the manuscript and approved the version of the manuscript to be published.

Signature	[redacted]	Date	11/12/17

Name of Co-Author	David Adelson		
Contribution to the Paper	DA revised the manuscript and approved the version of the manuscript to be published.		
Signature	[redacted]	Date	4 Dec. 2017

Name of Co-Author	Deborah White		
Contribution to the Paper	DW contributed to data curation, and revised the manuscript and approved the version of the manuscript to be published.		
Signature	[redacted]		11/1/2018

Name of Co-Author	Enaam Alsobhi		
Contribution to the Paper	EA contributed to data curation, and revised the manuscript and approved the version of the manuscript to be published.		
Signature	[redacted]	Date	3 Feb 2018

Name of Co-Author	Naeem Chaudhri		
Contribution to the Paper	NC contributed to the validation, data curation and the conception, interpretation of data, drafts, and revised the manuscript and approved the version of the manuscript to be published.		
Signature	[redacted]	Date	28/1/18

Combining Validated Prognostic Scores and Resolving Conflict in the Allocation to Risk Groups of Chronic Myeloid Leukemia Patients

Haneen Banjar[1, 2]*, Timothy Hughes[3, 4,5,6,7], Fred Brown[2¶], David Adelson[8¶], Deborah White[3, 4,9,5,6&], Enaam Alsobhi[10, 11,12&] ,Naeem Chaudhri[13¶]

[1]The Department of Computer Science, King Abdul Aziz University, Jeddah, Saudi Arabia, email.

[2]School of Computer Science, University of Adelaide, Adelaide, South Australia, Australia.

[3]Cancer Theme, South Australian Health and Medical Research Institute (SAHMRI), Adelaide, South Australia, Australia.

[4] University of Adelaide, Discipline of Medicine, Adelaide, South Australia, Australia.

[5] Centre for Cancer Biology, University of South Australia, Adelaide, South Australia, Australia.

[6] Centre for Personalised Cancer Medicine, University of Adelaide, Adelaide, South Australia, Australia.

[7] South Australian Health and Medical Research Institute, Adelaide, South Australia, Australia.

[8] School of Molecular and Biomedical Science, The University of Adelaide, Adelaide, Australia.

[9] University of Adelaide, Discipline of Paediatrics, Adelaide, South Australia, Australia.

[10] Pathology and Laboratory Medicine, King Abdulaziz Medical City-Western Region National Guard Health Affairs, Jeddah, Saudi Arabia.

[11]Princess Noura Oncology Centre, Jeddah, Saudi Arabia.

[12] King Saud Bin Abdulaziz University for Health Sciences, King Abdulaziz Medical City, Jeddah, Saudi Arabia.

[13]Oncology Centre, Section of Hematology/HSCT, king Faisal Specialist Hospital and Research Centre, Riyadh, Saudi Arabia

*Corresponding Author

E-mail: Hrbanjar@kau.edu.sa (HRB).

¶ These authors contributed equally to this work

& These authors also contributed equally to this work

Short Title: Resolving Conflict in Allocation to Risk Groups in CML

Keywords: *Chronic Myeloid Leukemia, Prognostic Score, Conflict Outcomes, Combined Model, Sokal Score, Hasford Score, EUTOS Score, ELTS Score, and CML Treatment Guideline.*

Abstract

Background: Obtaining accurate risk stratification for chronic myeloid leukemia (CML) at the time of diagnosis is essential for identifying patients who will most likely benefit from aggressive therapy. However, inconsistent results have been reported when two different risk categories have been applied to the same CML patient, i.e., when one prognostic score classifies the patient in one risk group and another score contradicts the first classification.

Aim: To overcome this, we have proposed to develop a framework that automatically i) combines the common prognostic scores in CML and filters the models to select the highest performing model for developing the best conflict resolution strategy between prognostic scores in patients with conflicting prognoses and ii) updates the entire architecture when adapting for a new prognostic model.

Method: Australian and Saudi populations were included in this study, who received imatinib 600 mg/800 mg as the initial therapy. A consistency test extracted 60 CML patients whose profiles generated conflicting risk categories. Several combined score models were used and the auto filtered to select the highest performing model in term of the accuracy, G-mean and F-score.

Results: Most combined models outperformed the single prognostic scores with an accuracy performance above 78%-80%. The highest accuracy performance achieved by a single score was the Hasford score at 74%.

Conclusion: We explored a new group of patients, known as "conflict patients," whose profiles generated an outcome from one prognostic score that contradicted the outcomes using other prognostic scores. Current CML treatment guidelines should be updated to include our combined model when prognostic scores generate a conflicting prognosis for the same patient.

Introduction

According to the Australian Institute of Health and Welfare (AIHW), myeloid cancer was the ninth-most commonly diagnosed cancer in 2016 with more than 3,600 cases in Australia [1]. Chronic myeloid leukemia (CML) is a malignant blood cancer that produced an excessive number of immature white blood cells circulating in an affected person's blood. The causative chromosomal translocation in CML gives rise to the *BCR-ABL1* gene, which encodes for the constitutively active tyrosine kinase Bcr-Abl [2]. An accurate risk stratification for CML at the time of diagnosis is essential for identifying high-risk patients and those less likely to experience unfavorable outcomes in achieving a major molecular response (MMR: *BCR-ABL1* transcript≤0.10%) after imatinib (Glivec®) frontline therapy. These patients are more likely to benefit from frontline second-generation tyrosine kinase inhibitors (TKI), such as nilotinib (Tasigna®) or dasatinib (DAS) (Sprycel®).

Newly diagnosed CML patients should be assessed based on the available prognostic scoring systems. In practice, scoring systems are used to stratify CML patients according to their risk profile to ensure appropriate treatment. Historically, four common prognostic scores, namely the Sokal score [3], the Hasford score [4], the European Treatment and Outcome Study (EUTOS) score [5], and the EUTOS long-term survival (ELTS) score [6], have been used to clinically identify high-risk groups. These four scores ascertain the level of risk for CML patients by running a multivariable regression analysis. Clinicians aim to treat individual CML patients with the most appropriate therapy. In low-risk patients, imatinib is often preferred because it has proved to be the least toxic and safest option. In higher risk patients, second-generation TKI or combination approaches are preferred because the higher toxicity and higher risk of organ damage is counterbalanced by greater potency and lower propensity to drug resistance. Thus, using an accurate risk assessment method at diagnosis can assist in effective treatment assignment.

Three central problems that have been observed in previous studies are i) establishing which prognostic score is the most accurate for identifying risk, ii) since each prognostic score may generate conflicting prognoses for the risk index, it is difficult to know how to treat patients with conflicting prognoses, and iii) since prognostic score systems are developed over time, how patients can benefit from newly developed systems and information.

First, scoring systems have shown a variable correlation with complete cytogenetic response [7-12], major molecular response (MMR) [13-16], overall survival [17 , 18], and event free survival (EFS) [19] in CML patients who receive frontline imatinib. Several studies have compared three scores to predict molecular response but have recommended different scores each time (Table 1).

Table 1. Summary of comparative studies in predicting major molecular response (MMR) endpoints.

Study	Year	Country	Source	No. of patients	Median age (range)	Median follow-up months (range)	Study Type	Scores involved in study	Recommended score for predicting molecular response	Months
Dybko et al. (Jaroslaw Dybko et al., 2015)	2015		Single centre	88	51(21–83)	47 months	Comparative study	Sokal, Hasford and EUTOS scores	Hasford	at 18 months
Yamamoto et al. (Yamamoto et al., 2014)	2014	Yokohama	Multicentre	145	54 (15–83)	57 (9–130)	Validation study	Sokal, Hasford and EUTOS scores	None.	at 18 months
Bonifacio et al. (Bonifacio et al., 2014)	2014	Italy	Single centre	314	57 (19–85)	N/A	Validation study	Sokal and EUTOS scores	EUTOS score	at 3 and 6 months
Suttorp et al. (Suttorp et al., 2013)	2013	Germany	Single centre	90	11.6 (1–18)	N/A	Comparative study	Sokal, Hasford and EUTOS scores	Refined risk categorization in EUTOS from 87 to 64	at 3 months
Jabbour et al. (Jabbour et al., 2011)	2012	USA(Texas)	Single centre	465	47 (15–85)	117 (16–130) imatinib 400mg 88 (4–118) imatinib 800mg	Validation study	EUTOS scores	The EUTOS score was not predictive of the outcome.	N/A
Than et al.(Than et al., 2012)	2012	Singapore	Single centre	139	45 (16–88)	N/A	Validation study	EUTOS scores	EUTOS score	at 12 and 18 months
Yahng et al. (Yahng et al., 2012)	2012	Korea	Single centre	255	42 (19–77)	57 (13–102)	Comparative study	Sokal, Hasford and EUTOS scores	Sokal and EUTOS scores	at 18 months
Saussele et al.(Saussele et al., 2011)	2011	German	Single centre	1252	N/A	N/A	Validation study	Sokal, Hasford and EUTOS scores	EUTOS score	at 12 months
Marin D et al. (Marin, Ibrahim, & Goldman, 2011)	2011	UK	Single centre	282	43 (13–80)	68 (16–130)	Comparative study	Sokal and EUTOS scores	Sokal score	N/A

Second, prognostic scores for CML have been developed that optimize the use of clinical experience in CML treatment. However, the science of prognostication has evolved rapidly, with various scoring systems established with independent aims and different patient populations. Unfortunately, given the variety of prognostic scores, conflicting conclusions can be generated, thus making it difficult to reach a final prognosis. Consistency is defined as one prognostic score that does not contradict other prognostic scores. It is important to study and understand consistency between scores and resolve prognostic conflict in order to help clinicians categorize patients into suitable risk groups and subsequently make better therapeutic decisions. Consistency among prognostic scores can increase clinician confidence because healthcare providers depend upon these results for making appropriate treatment decisions. Finally, when a new prognostic score is developed, a prognostic model structure should be flexible enough to be updated to integrate this new information. This way the prognostic model structure can be continuously updated and have its data contents modified.

There have been no previous methods for detecting conflicts between the current validated scoring systems for CML, and clinicians need to specify which prognostic score to use based upon their preferences for each patient. Such an approach does not guarantee an optimal result and does not integrate the information from all existing prognostic scoring systems. To overcome this, we have proposed to develop a framework that automatically i) combines the common prognostic scores in CML and filters the models to select the highest performing model for developing the best conflict resolution strategy between prognostic scores in patients with conflicting prognoses and ii) updates the entire architecture when adapting for a new prognostic model.

Materials and Methods

In this section, we discuss the materials and methods used in the datasets, inclusion and exclusion criteria, data pre-processing, and examine the performance of the individual prognostic scores. Then, we describe the proposed framework, which contains two processes: one that resolves prognostic conflict using combined score models and an adaptive process for new prognostic models. The first process automatically filters the models and selects the highest performance models that develop the best conflict resolution strategy between prognostic scores in patients with conflicting prognoses. The second process updates the entire architecture when adapting to a new prognostic model.

Datasets

The data used in our experiments came from Australian and Saudi populations. The Australian population was sourced from the Therapeutic Intensification in De Novo Leukemia (TIDEL) II [20] clinical trial, and the Saudi population was sourced from a tertiary care hospital, the King Faisal Specialist Hospital and Research Centre (KFSHRC), Riyadh, Saudi Arabia [21] and from King Abdulaziz Medical City (KAMC), Jeddah, Saudi Arabia [22].

Ethics Statement

The data were analysed anonymously. All study participants provided written informed consent prior to participation. The TIDEL II trial is registered at www.ANZCTR.org.au as ACTRN12607000325404 and funded by Novartis Australia. TIDEL II was carried out with the approval of human research ethics committees (RAH Protocol No. 070718c), and ethically approved by the National Statement on Ethical Conduct in Human Research (NHMRC) 2007 and in accordance with the Declaration of Helsinki. The KFSHRC data were ethically approved by the Clinical Research

Committee (CRC) and Research Ethics Committee (REC) 2005, Research Advisory Council (RAC) reference: ORA/1811/26, Proposal No. 2051056. The KAMC data is an ethically approved prospective study, Protocol No.: RCJ0210-134, 2010.

Inclusion and exclusion criteria

Patient outcomes were divided into two broad outcome groups: i) positive group (CML patients able to remain on imatinib and achieve target MMR at 24 months) and ii) negative group (CML patients who did not achieve MMR at 24 months on imatinib or switched therapies and patients found to be intolerant to imatinib, who switched to nilotinib/dasatinib and then achieved MMR at 24 months since MMR was not achieved by administering standard treatment).

Eligible patients from TIDEL II [20], KFSHRC [21], and KAMC [22] who were included in the study were:

- Patients who received imatinib 600 mg/800 mg as the initial therapy.
- Patients considered in the chronic phase with a blast percentage of less than 10%, based upon the standards of the American Cancer Society medical and editorial content team [23].
- Patients monitored for achieving a BCR-ABL1 transcript level ≤ 0.1% at 24 months using RQ-PCR.
- Patients who were switched to a second line treatment because of intolerance or toxicity to imatinib were considered as a negative group in our analysis.

Patients who were excluded were:

- Pregnant patients.
- Patients who used a second line treatment (nilotinib or dasatinib) as a frontline treatment.
- Patients who were in advanced disease phases.

Data pre-processing

A dataset with missing data could affect the performance of a combined model if using that dataset as training sample. Three methods were proposed to treat the missing data issue: deletion, classification, or imputation [24]. Some patient data had been missed while collected from white blood cell counts that were obtained from blood samples at diagnosis. If the option were to delete the missing data, this information would be lost, and the learning procedure could be affected by providing fewer samples [24]. Therefore, we imputed the missing data by using multiple imputations. This method replaced the missing values using linear regression for a given prognostic factor [25].

For individual CML patients, we collected prognostic factors, calculated prognostic score values, obtained the prognostic risk category, collected the actual MMR at 24 months, and defined the *conflictIndex*. First, the prognostic factors were the parameters that were collected at diagnosis before commencing treatment. For the Sokal and ELTS scores, the prognostic factors are age, spleen size, platelet count, and peripheral blood blasts. The Hasford score adds peripheral blood eosinophil and basophil percentages to these, and the EUTOS score is based on the percentage of basophils and spleen size.

Second, prognostic score values were calculated using each patient's clinical and hematological prognostic factors as shown in the prognostic scores' formulas [26]. It has been argued that the EUTOS score predicts the probability of no complete cytogenetic response at 18 months and not overall survival [5]. However, the EUTOS score can be used to identify CML patients with a significantly lower probability of responding to therapy [8] and survival [18, 27] and predict molecular response [13]. After calculating the score values, we used the Min-Max normalization technique [28],

and the resulting values came from the minimum and maximum values that the prognostic score models could generate for this patient population. Each prognostic score kept its cut off values for categorizing CML patients into risk groups and was mapped into its normalized value between 0 and 1. We used the normalized score values for discovering the relationships between the data and minimizing misclassification errors when training machine learning algorithms [28]. Thus, all data was assumed to have the same distribution. We plotted the distribution of pairs of scores to examine the data to see whether a linear or non-linear relationship exists between the scores.

Third, the Sokal, Hasford, and ELTS scores classify patients into three risk categories of low, intermediate, and high. In contrast, the EUTOS score classifies patients into two risk groups of low and high. Here, the risk groups were analyzed using consistency test as suggested in [26]. We compared two risk groups:

- Group 1 (low and intermediate risk) vs. high risk.
- Group 2 low risk vs. (intermediate and high risk).

Then, we selected the group that achieved the highest accuracy for stratifying patients into risk groups.

Fourth, the international scale (IS) [29] divides molecular responses into major molecular response (MMR; *BCR-ABL1* transcript ≤ 0.10%) and complete molecular response (CMR; *BCR-ABL1* transcripts not detectable). The optimal response at any time after 12 months is a *BCR-ABL1* transcript level ≤ 0.1[30]. Therefore, we selected the comparison endpoint for long-term molecular response at 24 months.

Finally, *conflictIndex* represented the consistency test [26] among prognostic scores: 1 indicates conflicting observations between prognostic score categories and 0 indicates consistent observations.

Examination of the performance of the individual prognostic scores

Four common prognostic scores (classifiers) for CML patients were examined using risk categories for predicting molecular response. The performance of the classifiers was measured using a coincidence matrix [31]. The problem was presented as a binary classification where the test results were placed in either a positive group (achieved MMR using imatinib) or a negative group (did not achieve MMR). The results were divided into four conditions: (a) high-risk patients correctly identified for not achieving MMR ("true negative"), (b) low-risk patients achieving MMR but wrongly identified as high-risk patients that would not achieve MMR ("false negative"), (c) low-risk patients correctly identified for achieving MMR ("true positive"), and (d) high-risk patients not achieving MMR but wrongly identified as low-risk patients that would achieve MMR ("false positive"). Accuracy reflects the number of correctly classified patients belonging to a group.

Resolving conflict using a combined model

Two strategies can be used to combine classifiers (prognostic scores): classifier selection and classifier fusion [32]. In classifier selection, each classifier is intended to be an expert for a specific domain of the feature space, and the selected classifier decides the output of the ensemble. This method has been used previously in many comparative studies as is shown in Table 1. Classifier selection is only guaranteed to give the same training accuracy as the best individual classifier, whereas in classifier fusion, each classifier is intended to have complete information on the entire feature space, and we apply combined methods to all outputs from the systems. Therefore, we focused on using the fusion method to build a model that would increase our understanding of the process that should be taken in the conflicting prognoses cases. Methods for classifier fusion can be a non-trainable combiner (data-independent) or a

trainable combiner (data-dependent) [32]. The use of data-independent methods, such as majority voting, is simple compared to the complex methods required for a training process. In the data-dependent methods, we built a combination based on the information obtained from the individual classifier (prognostic score), and we could select a possible approach for combining the outputs of the classifiers. The approaches to building a combined classifier include combining classifiers at the combination, classifier, feature, or data levels [32]. The combination level mainly focuses on the possible ways of combining outputs (risk categories) of classifiers (prognostic scores) in an ensemble based on information obtained from a single classifier. For the classifier level, different machine learning techniques can be employed. For the feature level, different feature subsets result in different performances by training the same classifier. Finally, the data level ensemble created by using different training sets should be generated randomly from the distribution and then sampled with a replacement.

In this study, we aimed to develop a combined model that could resolve conflicts and outperform a single prognostic score. We decided to train multiple methods until this objective was achieved. First, we trained a meta decision tree [33] as an example of the combination level method. The difference between meta decision trees and the ordinary decision trees is that their leaves specify which classifier (in our case, prognostic scores) should be used rather than class label. Second, we implemented three base classifiers, a support vector machine [34], K-nearest neighbor algorithm [32], and naive Bayes classifier [32], as examples for classification levels (more details about each base classifier are discussed in the appendix A). Third, on the feature level, two features could be carried out by an automatic procedure: forward selection or backward elimination [35]. Forward selection started with no features in the model and added each feature using model comparison criterion. We stopped the addition if no feature

improved the model. Conversely, backward elimination started with all candidate features, and each feature was deleted using model comparison criterion. This process was repeated until either no further deletion could improve the model or there were only two candidate features. In our experiments, using a subset of prognostic scores to train the classifiers affected the conflict group and moved the patients to a consistent group. Finally, we did not implement any model for the data level as the performance would be based on mean predictions from the subset training data [36]. Thus, we focused on the combination level and classification level. Table 2 shows the techniques used for combining prognostic scores and dataset descriptions.

Table 2. Combining levels and techniques used for combining prognostic scores using patients in the conflict group.

Machine learning	Dataset description
Meta decision trees. [33]	Risk categories from each prognostic score and the patient outcome based on achieving/not achieving MMR at 24 months. For example, Sokal,=1 Hasford=1, EUTOS=0, ELTS=0, where 1 indicates low risk and 0 indicates high risk and the outcome of 1 indicates a positive group and 0 indicates a negative group.
Support vector machine [34]	Using three score values calculated by the formula assigned to each prognostic score [26] and the patient outcome based on achieving/not achieving MMR at 24 months. For example, Sokal= 0.2 Hasford=1200, EUTOS=0.8, ELTS=0.7, and the outcome of 1 indicates a positive group and 0 indicates a negative group. Sokal score values were categorized as low-risk and intermediate-risk (Sokal score<1.2) and high-risk (Sokal score>1.2), the Hasford scores were categorized as low-risk and intermediate-risk (score≤1480) and high-risk (score≥1481), and the EUTOS scores were categorized as low-risk (score<87) and high-risk (score≥87).
K-nearest neighbor [32]	
Naive Bayes [32]	

Nested cross-validation

Nested cross-validation [37] was used to automate finding the best conflict resolution strategy and to select the final model. This validation method suggested dividing the data into two parts: a training/validation part (~75%) and testing part (~25%). Since the testing part would be saved for evaluating the selected model, the training data itself was used for conducting an extra layer of 10-fold cross-validation to automatically select the highest performing combined model. This method is preferred

when the data are scarce [38]. The combined models were trained on a training set, the validation set was used for selecting the best model, and this model's performance was evaluated with the testing set. Each model was trained to minimize the misclassification rate between the prognostic scores categories for risk groups and actual patient outcome (positive or negative). The overall performance of a model was calculated using the average performance over 10 folds. We compared the training performance with the mean of a 10-fold cross-validation to avoid an over-fitting issue in the models [39]. We used the confidence interval obtained from the mean and standard deviation at a 95% confidence level.

The training data was imbalanced, meaning that the patients in one group were greatly outnumbered by those in the other group [40]. Therefore, for additional accuracy, we used another two measures, a G-mean (geometric mean) and F-score (weighted harmonic mean of sensitivity and positive predictive value), to assess the performance of a model trained on imbalanced data [31]. We reported the accuracy, G-mean, and F-score for the training and cross-validation performance. Finally, the recommended combined model was evaluated on test data and external unseen data.

Adaptation of a new prognostic model

For the first process of developing a combined model to resolve prognostic conflicts, a model architecture was chosen that outperformed individual prognostic score performance. For the second process, the model learned from new data and automatically created and updated the first process (resolve conflict using a combined model). The evaluation methods were applied after the incoming information updated the selected combined model architecture. The full procedure is shown in Figure 2, which is the schema for developing the prognostic combined model, evaluation, final model selection, and addition of incoming information.

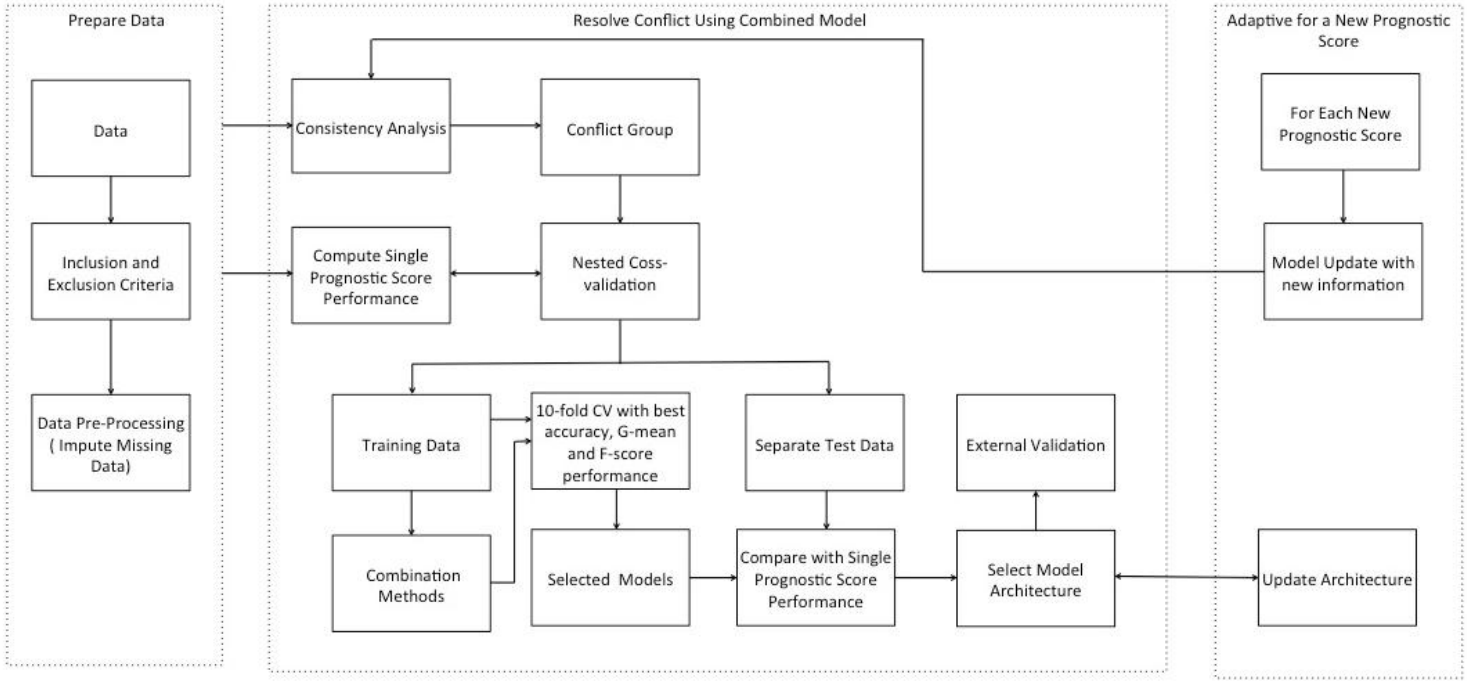


Figure 1. **The schema for developing the prognostic combined model, evaluation, final model selection, and addition of incoming information**. To build the prognostic model, we collected data from the Australian and Saudi populations and applied inclusion and exclusion criteria to include eligible patients for the analysis. In data pre-processing, we imputed missing data and conducted a consistency analysis. We included a conflict group in the analysis and used a nested cross-validation method to build, evaluate, and select the model. In the nested cross-validation, the dataset was divided into two parts: a training/validation part and testing part. The performance of individual prognostic scores was evaluated in both parts. The different combined models were implemented with training data. Selection of the best accuracy, G-mean, and F-score performance using mean values was a result of a 10-fold cross validation. The highest performance model was automatically selected and compared with the individual performances of separate testing data. The choice of the final model architecture would depend on the highest performance. Further, the entire model architecture was updated by repeating the processes from consistency analysis to the final choice of the architecture with the adaption of newly incoming prognostic score information. Finally, the final model was validated externally on unseen data.

Results

In this section, first we report basic information about the number of patients in both prognostic groups. Then we describe the dataset and present figures that show agreement between prognostic scores. Pre-processing is discussed in regards to imputing missing data and conducting consistency analysis. We reveal single prognostic score performance and the combined value of validated prognostic scores for conflict patients using Sokal, Hasford, and EUTOS scores. Then we describe how the recommended model was evaluated on unseen data. Finally, we describe the method for adding new information with the ELTS score and validation on independent datasets.

Insight into the data

This study included data on 486 patients from the Australian and Saudi populations. The patients were compromised of 210 CML patients from TIDEL II [20], 172 KFSHRC [21], and 104 from KAMC [22]. Applying the inclusion and exclusion criteria as shown in Appendix Table A, this report is based on data from 403 CML patients in the chronic phase. Table 3 describes patient characteristics regarding the mean and range for each clinical and hematological parameter category.

Table 3. Patient characteristics for the training, testing, and validation data. The mean and range for each parameter.

Patient Characteristics in Training and Testing Data	n	Mean	Range	No. missing	Imputed values mean
Age, yr.	314	49	17–81	0	0
Spleen size, cm	312	5.30	0-30	2	11.81
Platelet count (x 10^9/L)	309	494.26	4.41–2876	5	424.94
Peripheral blood blasts (% of leukocytes)	306	1.35	0–10	8	3.04
Peripheral blood eosinophils (% of leukocytes)	311	1.11	0–17.81	3	1.93
Peripheral blood basophils (% of leukocytes)	311	1.89	0–38.89	3	3.94
Patient Characteristics in Validation Data	n	Mean	Range	No. missing	Imputed values mean
Age, yr.	172	37	13–80	0	0
Spleen size, cm	123	7.34	0–99	49	11.16
Platelet count (x 10^9/L)	160	414.319	0–1902	12	415.4
Peripheral blood blasts (% of leukocytes)	159	1.91	0–10	13	4.13
Peripheral blood eosinophils (% of leukocytes)	159	3.01	0–75	13	6.81
Peripheral blood basophils (% of leukocytes)	159	3.09	0–21	13	5.68

To prepare the data, we used multiple imputations to impute prognostic factors (continuous data). We used the impute command in SPSS v21 to replace missing values with linear regression estimation values. The parameter to be imputed was to be used as a dependent variable in the regression model. In multiple imputations, the missing value is imputed multiple times using the linear regression method. A continuous parameter may have an imputed value outside the range of the observed values. We restricted the value to fall within a minimum and maximum range. Then, we divided the data into two datasets:

1. The score values set, which was compromised of the actual values calculated from the individual prognostic score formula using age, spleen size, platelets count, peripheral blood blasts, and eosinophil and basophil percentages of leukocytes parameters.
2. The risk categories were set at this stage, and we conducted an analysis to compare two risk groups. Table 4 shows the accuracy for each prognostic score

belonging to group 1and group 2. It was clear that the accuracy achieved by the prognostic score of group 1, with Sokal 60%, Hasford 72%, and ELTS 66%, outperformed the accuracy achieved by group 2, with Sokal 45%, Hasford 59%, and ELTS 60%. We got data that was compromised of data came from mapping the actual scores into risk categories. Therefore, we categorized the Sokal score values into low-risk and intermediate-risk (Sokal score<1.2) and high-risk (Sokal score>1.2), the Hasford scores into low-risk and intermediate-risk (score≤1480) and high-risk (score≥1481), and the EUTOS scores into low-risk (score<87) and high-risk (score≥87).

Table 4. The number of patients in the different risk groups per calculated scores.

1	Combined groups (1)				Combined groups (2)				
	n	Positive outcome	Negative outcome	Accuracy		n	Positive outcome	Negative outcome	Accuracy
Sokal Score Risk Group					Sokal Score Risk Group				
High	126	85	41	60.54%	Intermediate and high	254	179	75	45.65%
Low and Intermediate	277	203	74		Low	149	109	40	
Hasford Score Risk Group					Hasford Score Risk Group				
High	26	12	14	71.96%	Intermediate and high	145	98	47	58.80%
Low and Intermediate	377	276	101		Low	258	190	68	
EUTOS Score Risk Group					EUTOS Score Risk Group				
High	43	24	19	70.22%	High	43	24	19	70.22%
Low	360	264	96		Low	360	264	96	
ELTS Score Risk Group					ELTS Score Risk Group				
High	54	38	16	66%	Intermediate and high	154	100	54	60.04%
Low and Intermediate	349	250	99		Low	249	188	61	

Moving to the consistency test, we defined a new column for both datasets (the score values set and the risk categories set), namely *ConflictIndex*. Each patient was tested for an outcome prognosis with each prognostic score. Then we separated the conflicting data from the consistent data. In Figure 2, we represented the distribution of all 403 patients in three prognostic scores of Sokal, Hasford, and EUTOS respectively, and Figure 3 showed the distribution in four prognostic scores of Sokal, Hasford, EUTOS, and ELTS, respectively. Disagreement in prognostic scores risk categories for patient outcome occurred more frequently in the negative group. We also noticed that about 62% of patients were misclassified by the three prognostic scores and stratified into the low and intermediate risk group when their observed outcomes were in the negative group. Similarly, the four prognostic scores classified 60% of the patients in low and intermediate risk group when their outcomes were in the negative group. In the training set, we observed that, when considering agreement among the three scores, 50 patients belonged to the conflict group, while, when considering agreement among the four scores, 60 patients belonged to the conflict group. In the testing set, nine patients had conflicting prognoses using three or four scores.

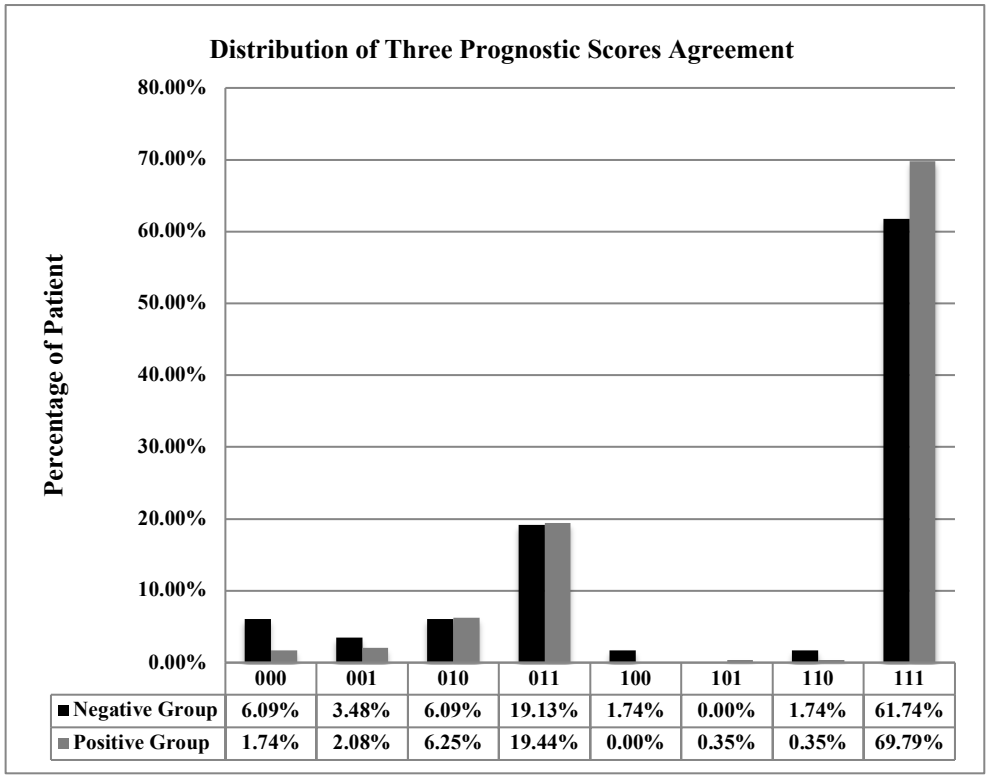
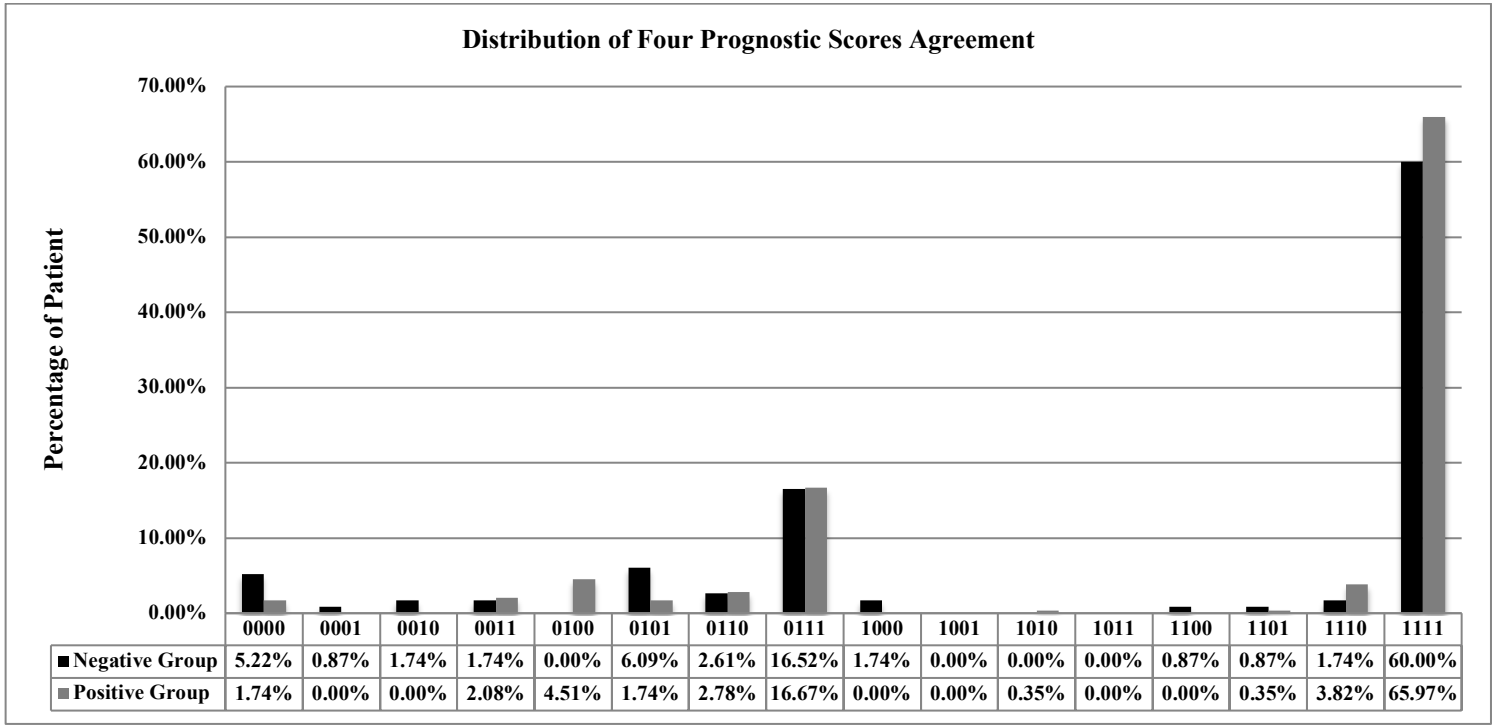


Figure 2. **Patient distribution for the three prognostic scores of Sokal, Hasford, and EUTOS, respectively**. 0 indicates high risk and 1 indicates low or intermediate risk. The three indexes refer to the three scores. For example, 110 represent Sokal and Hasford patient score categories of low or intermediate risk, respectively, while the last index refers to the EUTOS outcome that categorizes the same patient in a high-risk group.



	0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
Negative Group	5.22%	0.87%	1.74%	1.74%	0.00%	6.09%	2.61%	16.52%	1.74%	0.00%	0.00%	0.00%	0.87%	0.87%	1.74%	60.00%
Positive Group	1.74%	0.00%	0.00%	2.08%	4.51%	1.74%	2.78%	16.67%	0.00%	0.00%	0.35%	0.00%	0.00%	0.35%	3.82%	65.97%

Figure 3. **Patient distribution for the four prognostic scores of Sokal, Hasford, EUTOS, and ELTS, respectively.** 0 indicates high risk and 1 indicates low or intermediate risk. The four indexes refer to the four scores. For example, 1110 represents Sokal, Hasford, and EUTOS patient score categories of low or intermediate risk, respectively, while the last index refers to the ELTS outcome that categorizes the same patient in a high-risk group.

Performance of individual prognostic scores

We examined the performance of the individual prognostic scores because later we compared the single prognostic scores with several combined models to select the highest accuracy, G-mean, and F-score performance.

Table 5 shows the performance of each prognostic score in identifying risk categories based on the training data from the CML patients whose profiles demonstrated a conflict between prognostic risk scores.

Table 5. The performance of each prognostic score for conflict patents using three and four prognostic scores.

	Risk Categories				Single Score Performance				
Three Score Performance on Training Set, n=50									
	Risk Categories				Positive outcome, n=36	Negative outcome, n=14	Accuracy	G-mean	F-score
Sokal Score	Low and Intermediate				0	3	0.22	0	0.35
	High				36	11			
Hasford Score		Low and Intermediate			33	10	0.74	0.51	0.83
		High			3	4			
EUTOS Score			Low		26	8	0.64	0.56	0.74
			High		10	6			
Four Score Performance Training Set, n=60									
	Risk Categories				Positive outcome, n=43	Negative outcome, n=17	Accuracy	G-mean	F-score
Sokal Score	Low and Intermediate				7	5	0.32	0.34	0.25
	▪ High				36	12			
Hasford Score		Low and Intermediate			40	12	0.75	0.52	0.84
		High			3	5			

Table 5. Continued.

	Risk Categories				Single Score Performance				
	Risk Categories				Positive outcome, n=43	Negative outcome, n=17	Accuracy	G-mean	F-score
EUTOS Score			Low		33	10	0.67	0.56	0.76
			High		10	7			
ELTS Score				Low and Intermediate	23	10	0.50	0.47	0.60
				High	20	7			
Score Performance on Testing Set, n=9									
	Risk Categories				Positive outcome, n=6	Negative outcome, n=3	Accuracy	G-mean	F-score
Sokal Score	Low and Intermediate				2	0	0.56	0.58	0.50
	High				4	3			
Hasford Score		Low and Intermediate			5	2	0.67	0.53	0.77
		High			1	1			
EUTOS Score			Low		5	3	0.56	0.00	0.71
			High		1	0			
ELTS Score				Low and Intermediate	4	1	0.67	0.67	0.73
				High	2	2			

Table 5. Continued.

Score Performance on External Validation Set, n=65									
	Risk Categories				Positive outcome, n=44	Negative outcome, n=21	Accuracy	G-mean	F-score
Sokal Score	Low and Intermediate				4	1	0.37	0.29	0.16
	High				44	20			
Hasford Score		Low and Intermediate			41	19	0.66	0.30	0.79
		High			3	2			
EUTOS Score			Low		36	15	0.65	0.48	0.76
			High		8	6			
ELTS Score				Low and Intermediate	33	20	0.52	0.19	0.68
				High	11	1			

For conflict patients (n=50), the Hasford scores correctly identified risk groups with an accuracy of 74% and F-score of 83%. Similarly, for conflict patients (n=60), the highest accuracy and F-score were observed in the Hasford score as 75% and 84%, respectively. The EUTOS score achieved the highest G-mean with 56% in both conflict groups (n=50 and n=60) of patients. The EUTOS score G-mean performance indicated its ability to identify both risk groups from imbalanced data (three score dataset: positive group=36 patients vs. negative group=14 patients and four score dataset: positive group=43 patients vs. negative group=17 patients). The results showed that the Hasford score outperformed the other scores among conflict patients.

Resolve conflict using combined methods in conflict patients

We divided our experiments into two stages: the development of the combined model using three prognostic scores and the development of the adaptive combined model with four prognostic scores. We used Matlab 2016b statistics and the Machine Learning Toolbox for classifications [41]. The following four machine learning techniques were used: meta decision tree, support vector machine, K-nearest neighbor and naive Baye. We used C# code to calculate the performance and automate the process for selecting the highest accuracy, G-mean, and F-score models.

Table 6 summarizes the results achieved by combining the methods of the three prognostic scores during the first stage, and Figure 4 depicts the improvement in performance in the training and validation set of the three combination methods compared to a single prognostic score.

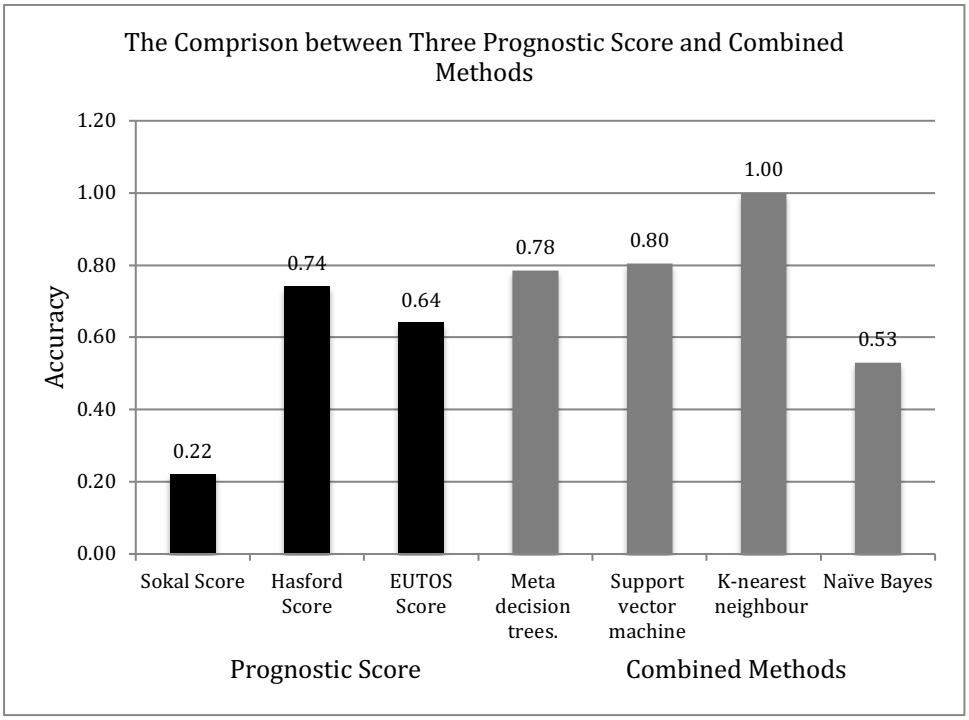


Figure 4. The performance in the training and validation set of the three combination models compared to single prognostic scores.

Most of the combination models improved performance significantly over the single scores. However, only the naive Bayes could resolve the conflict significantly. Overall, G-mean and F-score performances for the meta decision tree and support vector machine were lower than the single scores. However, the G-mean was higher in K-nearest neighbor and naive Bayes (100% and 58%, respectively) models. While for the F-score, only the K-nearest neighbor outperformed the single scores. From the 10-fold cross-validation results, it was clear that the K-nearest neighbor model had an over-fitting issue, but the reminder of the combined methods fit the data well. A comparison between the cross-validation results suggested selecting the meta decision tree and support vector machine based on their accuracy (75% and 76%, respectively) as good combined models to resolve conflicts. Details on how to reuse the structure of the model appear in the appendix A.

The new prognostic score ELTS was used during the second stage of our experiment for adapting newly incoming information. Table 7 summarizes the results

achieved by combining the methods of four prognostic scores, and Figure 5 depicts the improvement in the performance in the training and validation set of the three combination methods compared to single prognostic scores. Similarly, most combination methods improved performance significantly over a single score. However, only the naive Bayes could resolve the conflict significantly. The G-mean in all combined methods (57%, 64%, 100%, and 63% respectively) outperformed the highest G-mean for the single scores, which was achieved by the EUTOS score (56%). From the 10-fold cross-validation results, it was clear that the K-nearest neighbor model also had an over fitting issue, but in the remainder of the combined methods, the data fit will. A comparison between the cross-validation results suggested selecting the support vector machine over the other combined methods to resolve conflicts.

Testing the combination models on a test set (n=9) revealed unexpected results from the combined scores where the testing data results were not improved. Therefore, the generalization capability was quite similar to a single score because the highest accuracy performances (67%) were achieved by the Hasford and ELTS scores while the accuracy performances were (66%) in the three combination models for four scores.

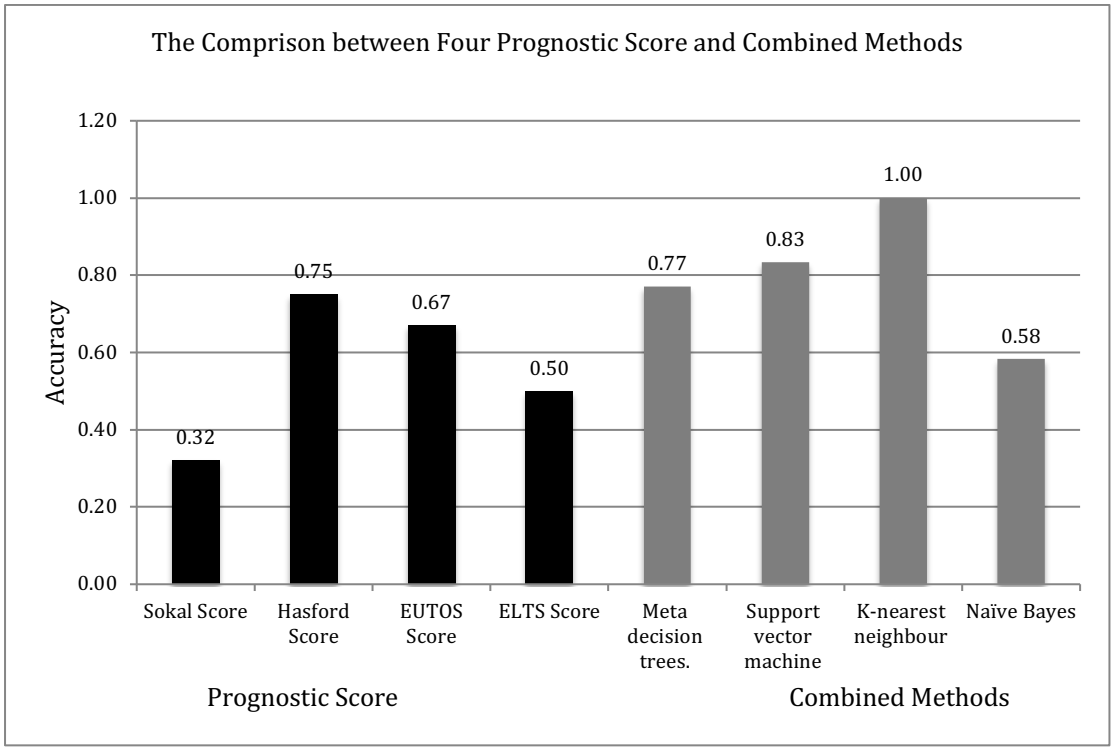


Figure 5. The performance in the training and validation set of the combination methods compared to the single prognostic scores.

Table 6. Results for the combination models in the training and validation set.

			Training Performance			10 Fold Cross-Validation			Testing Performance		
Model ID	Machine learning	Inputs	Accuracy	G-mean	F-score	Accuracy	G-mean	F-score	Accuracy	G-mean	F-score
1	Meta decision trees. [33]	Sokal, Hasford, EUTOS	0.78	0.46	0.35	0.75 ±0.11	0.16±0.35	0.51±0.33	0.55	0.4	0.34
2	Support vector machine [34]	Sokal, Hasford, EUTOS	0.80	0.53	0.44	0.76±0.08	0.21±0.34	0.20±0.32	0.55	0.4	0.34
3	K-nearest neighbor [32]	Sokal, Hasford, EUTOS	1	1	1	0.69±0.18	0.46±0.42	0.40±0.37	0.55	0.47	0.33
4	Naive Bayes [32]	Sokal, Hasford, EUTOS	0.52	0.58	0.47	0.50±0.23	0.38±0.34	0.53±0.25	0.55	0.4	0.34

Table 7. Results for the combination models in the training and validation set after adapting the newly incoming information.

			Training Performance			10 Fold Cross-Validation			Testing Performance		
Model ID	Machine learning	Inputs	Accuracy	G-mean	F-score	Accuracy	G-mean	F-score	Accuracy	G-mean	F-score
5	Meta decision trees. [33]	Sokal, Hasford, EUTOS ELTS	0.77	0.57	0.46	0.70 ±0.23	0.33 ±0.43	0.31 ±0.41	0.66	0.52	0.4
6	Support vector machine [34]	Sokal, Hasford, EUTOS ELTS	0.83	0.64	0.58	0.75 ±0.07	0.07 ±0.30	0.41 ±0.28	0.66	0.52	0.4
7	K-nearest neighbour [32]	Sokal, Hasford, EUTOS ELTS	1	1	1	0.57 ±0.20	0.21 ±0.31	0.20 ±0.27	0.66	0.66	0.57
8	Naive Bayes [32]	Sokal, Hasford, EUTOS ELTS	0.58	0.63	0.52	0.50 ±0.19	0.44±0.32	0.43±0.26	0.55	0.57	0.5

External validation of the final models

The generalization behavior in the testing set led to a further investigation of the combined methods in an external set. We compared our recommended model support vector machine from the four score combination set to resolve the conflict prognoses generated by the prognostic scores on unseen data from the KFSHRC. The 103 patients included 65 patients from the conflict group.

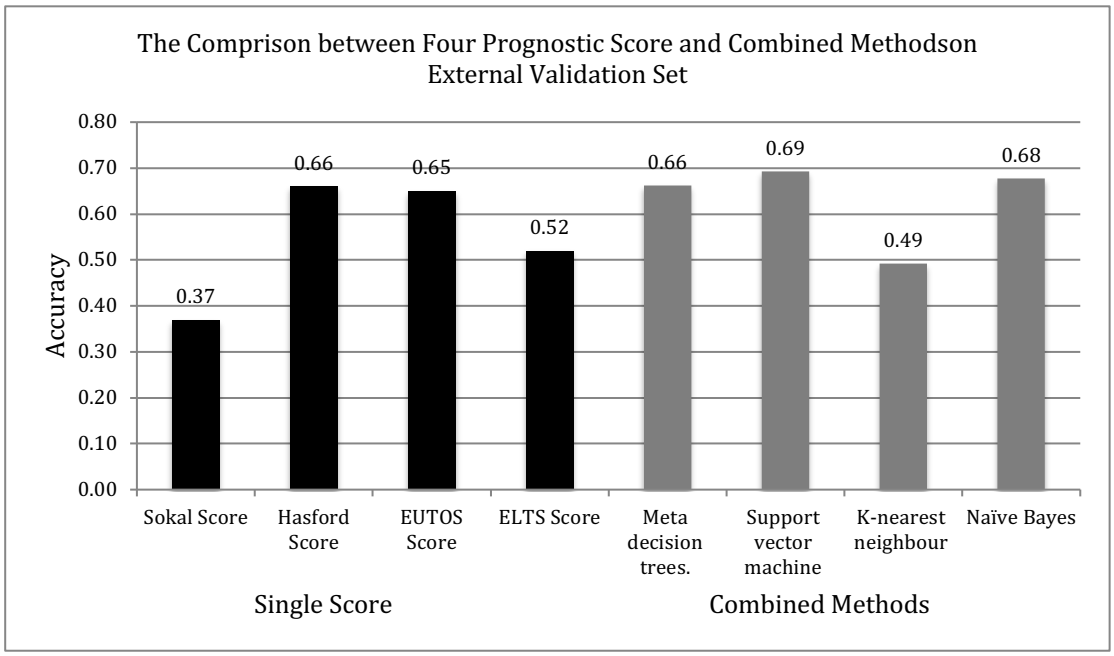


Figure 6. Comparison between combined model performances and single score systems in an external dataset.

Figure 6 shows that the support vector machine model correctly identified the risk group with an accuracy of 69% for the conflict group, while the meta decision tree achieved a performance similar to the Hasford score. Overall, most combined methods outperformed the accuracy performance of a single score, except for the K-nearest neighbor model.

Discussion

When dealing with CML, there is a range of validated prognostic scores established to help clinicians identify risk categories. However, current prognostic risk scores can provide conflicting prognoses in some cases, depending on the patient's profile. Thus, it is

desirable to achieve consistency among prognostic scores. This paper provides evidence that using combined score models could resolve inconsistencies between prognostic scores in conflict patients and outperform individual prognostic scores.

Learning how to deal with patients who have conflicting prognoses was the focus of this study. Previous studies have compared the prognostic value of validated scoring systems for overall survival, event free survival, or optimal response for CML patients who receive frontline imatinib, and then have recommended the best score. However, the problem of identifying risk in the group of patients with conflicting risk scores (conflict patients) has remained unsolved. This study used combination models to resolve conflicts in decision making.

These combined methods performed quite well in the conflict group. The suggested models were the combination scores using meta decision tree and support vector machine. The resulting structure can be used by other health organizations. Comparisons of the test data showed that the combined models had not achieved a high performance but were similar to a single score performance. The external validation data showed that the recommended model outperformed the single scores in the conflict group. Thus, the combined method can identify a risk group among conflict patients.

In clinical practice, there are treatment guidelines that have been published by the National Comprehensive Cancer Network [42], European Society of Medical Oncology [43], and European LeukemiaNet (ELN) [30]. Clinicians must specify their prognostic score based on their preferences for each patient. Clinicians must determine risk groups by using either Sokal [3], Hasford [4], or EUTOS [5] scores to distinguish high-risk patients who may need more aggressive initial imatinib or other TKIs (nilotinib and dasatinib) therapies that are available for frontline treatments. However, these approaches do not guarantee an optimal result and do not involve the information from all existing prognostic score systems.

Among patients who generated inconsistent risk indexes, our recommended combination methods correctly predicted the risk for 60 CML patients based on their molecular risk profiles. Although a comparative study [7] showed a conflict in the risk indexes of 22 patients, the study recommended one score as better than the others based on single prognostic score accuracy. Another study of 255 CML patients demonstrated conflict in 27 patients identified by the EUTOS score as high risk, but not by the Hasford score. Additionally, 17 CML patients were identified as low or intermediate risk by the Sokal score but as high risk by the EUTOS score. Thus, our study analyzed more CML patients and proved that conflict is a common problem.

Therefore, to detect a conflict between current validated scoring systems for CML, our study contributed to updating the current guidelines when the prognostic scores generate conflicting prognoses for the same patient. Our study analyzed the patients by fusing consistency tests and developing combined models to resolve conflicts between prognostic scores in patients with conflicting prognoses. We reached the same findings regarding combining risk groups as the ELN [30], who recommended dividing patients into low-risk (including intermediate) and high-risk populations in the management of CML.

Most of combined models outperformed single score performance in conflict patients. Cross-validation performance in the training sets showed that combining individual prognostic scores by the mean achieves better accuracy with fewer training sets (nine-fold rather than a full training set) than single scores. This leads to the conclusion that combined methods are more robust than single score performances. However, performance accuracy was not always significantly improved.

Strengths and limitations

As no single prognostic score is accurate, prognostic scores are judged by their utility. To our knowledge, no study has applied methods for resolving conflict or combined methods for prognostic scores to improve the quality of treatment decisions. The strength of

our study comes from combined models used in a conflict group of 60 CML patients. The models can be validated on another single or multicenter cohort once data for the prognostic scores become available.

This study contains common algorithm limitations because we focus on four common machine-learning techniques to combine prognostic scores at the classifier level. However, these methods do not always give an optimized solution. For example, we selected *k* in the K-nearest neighbor method based on the lowest misclassification rate as suggested by trial and error. In addition, computation procedures for performance are usually performed on training data and can initially recommend a weaker or stronger technique. However, using different datasets may lead to different conclusions, as these methods are data-driven. The majority vote analysis is the first attempt clinicians make to overcome conflicting observations in prognosis risk indexes when the number of the prognostic score (L/2)+1 can give a majority vote. Otherwise, the use of a weighed majority vote [32] might assist in resolving conflict. However, they are not always accurate in predicting and correctly identifying actual outcomes.

Future work has been grouped under two views: the use evolving intelligent systems and further application of combined models. First, the more that new prognostic scores are made available, the more demand to validate and test their consistency with existing prognostic scores. Evolving intelligent systems have been introduced that use online learning algorithms that acquire knowledge from data, perform advanced model structuring for data mining tasks and parameters, change system features or components over time, and develop knowledge in a system over time [44]. The evolving intelligent system must respond to changes in knowledge and deal with the scope of the problem through automated behavior.

Future studies should ideally conduct randomized controlled clinical trials to determine whether our combined models clinically improve patients risk identification in conflict patients. It is also important to study other medical outcomes for conflict patients,

specifically whether those patients need specific treatment based on their overall survival, event free survival, or type of mutation.

References

1. Cancer in Australia 2017 Canberra: The Australian Institute of Health and Welfare (AIHW); 2017 [2016]. Available from: <http://www.aihw.gov.au/cancer-data>.
2. Saglio G, Ulisciani S, Fava M, al. e. Molecular Monitoring in Patients with Chronic Myelogenous Leukemia. 2008.
3. Sokal J, Cox E, Baccarani M, Tura S. Prognostic discrimination in "good-risk" chronic granulocytic leukemia. Blood 1984;63:789-99.
4. Hasford J, Pfirrmann M, Hehlmann Rd, Allan NC, Baccarani M. A New Prognostic Score for Survival of Patients With Chronic Myeloid Leukemia Treated With Interferon Alfa. Jornal of the National Cancer Institute. 1998;90(11):850-8.
5. Hasford J, Baccarani M, Hoffmann V, Guilhot J, Saussele S, Rosti G. Predicting complete cytogenetic response and subsequent progression-free survival in 2060 patients with CML on imatinib treatment: the EUTOS score. Blood. 2011;118(3):686-92. doi: 10.1182/blood-2010-12-319038.
6. Pfirrmann M, Baccarani M, Saussele S, Guilhot J, Cervantes F, Ossenkoppele G, et al. Prognosis of long-term survival considering disease-specific death in patients with chronic myeloid leukemia. Leukemia. 2016;30(1):48-56. doi: 10.1038/leu.2015.261.
7. Kumar SS, Simanti S, Kumar MP, Kumar BN, Abhigyan P, Partha G. A comparative study of Hasford score and Sokal index in prognostication of the novo chronic myeloid leukemia patients and a search for new prognostic markers2013 ; 56(3):[216-20 pp.]. Available from: http://www.ijpmonline.org/text.asp?2013/56/3/216/120369.
8. Tiribelli M, Bonifacio M, Calistri E, Binotto G, Maino E, Marin L, et al. EUTOS Score Identifies Cases with Poor Outcome in Patients with Early Chronic Phase Chronic Myeloid Leukemia, Though Not Predictive for Optimal Response to Imatinib. Blood. 2012;120(21). PubMed PMID: WOS:000314049601287.
9. Pagnano KB, Lorand-Metze I, Miranda ECM, Duarte VO, Delamain MT, Duarte GO, et al. EUTOS Score Is Predictive of Event-Free Survival, but Not for Progression-Free and Overall Survival in Patients with Early Chronic Phase Chronic Myeloid Leukemia Treated with Imatinib: A Single Institution Experience. Blood. 2012;120(21). PubMed PMID: WOS:000314049600358.
10. Breccia M, Finsinger P, Loglisci G, Latagliata R, Mancini M, Salaroli A, et al. The EUTOS score identifies chronic myeloid leukeamia patients with poor prognosis treated with imatinib first or second line. Leukemia Research. 2012;36(9):E209-E10. doi: 10.1016/j.leukres.2012.05.011. PubMed PMID: WOS:000306523600006.
11. Oyekunle AA, Osho PO, Aneke JC, Salawu L, Durosinmi MA. The predictive value of the Sokal and Hasford scoring systems in chronic myeloid leukaemia in the imatinib era. Journal of Hematological Malignancies. 2012;2(2):25-33.
12. Hoffmann VS, Baccarani M, Lindoerfer D, Castagnetti F, Turkina A, Zaritsky A, et al. The EUTOS prognostic score: review and validation in 1288 patients with CML treated frontline with imatinib. Leukemia. 2013;27(10):2016-22. doi: 10.1038/leu.2013.171. PubMed PMID: WOS:000325642600008.
13. Saussele S, Lauseker M, Hoffmann V, Proetel U, Hanfstein B, Baerlocher GM, et al. Prediction of Molecular Response of Chronic Phase CML Patients by the EUTOS Score: Results of the Randomized CML-Study IV. Blood. 2011;118(21):1606-. PubMed PMID: WOS:000299597105490.
14. Dybko J, Jaźwiec B, Haus O, Urbaniak-Kujda D, Kapelko-Słowik K, Wróbel T, et al. The Hasford Score May Predict Molecular Response in Chronic Myeloid Leukemia Patients: A Single Institution Experience. Disease Markers. 2016;2016:7531472. doi: 10.1155/2016/7531472. PubMed PMID: PMC5080519.
15. Ganguly S, Lakshmaiah KC, Jacob LA, Babu S, Dasappa L, Govind Babu KS. Performance of Sokal and Eutos Scores for Predicting Cytogenetic and Molecular Response in Newly Diagnosed Chronic Myeloid Leukemia-Chronic Phase Patients on Imatinib. Indian Journal of Hematology and Blood Transfusion. 2016:1-5. doi: 10.1007/s12288-016-0667-x.
16. Bonifacio M, Binotto G, Calistri E, Maino E, Tiribelli M. EUTOS score predicts early optimal response to imatinib according to the revised 2013 ELN recommendations. Annals of hematology. 2014;93(1):163-4.
17. Höglund M, Sandin F, Hellström K, Björeman M, Björkholm M, Brune M, et al. Tyrosine kinase inhibitor usage, treatment outcome, and prognostic scores in CML: report from the population-based Swedish CML registry2013 2013-08-15 00:00:00. 1284-92 p.
18. Pfirrmann M, Hasford J, Saussele S, Turkina A, Prejzner W, Steegmann JL, et al. The EUTOS Survival Score Is Preferable over the Sokal Score for Prognosis of Long-Term Survival of Patients with Chronic Myeloid Leukemia. Blood. 2015;126(23):595-.

19. Uz B, Buyukasik Y, Atay H, Kelkitli E, Turgut M, Bektas O, et al. EUTOS CML prognostic scoring system predicts ELN-based 'event-free survival' better than Euro/Hasford and Sokal systems in CML patients receiving front-line imatinib mesylate. Hematology. 2013;18(5):247-52. doi: 10.1179/1607845412y.0000000071. PubMed PMID: WOS:000324530700001.
20. Yeung DT, Osborn MP, White DL, Branford S, Braley J, Herschtal A, et al. TIDEL-II: first-line use of imatinib in CML with early switch to nilotinib for failure to achieve time-dependent molecular targets2015 2015-02-05 00:00:00. 915-23 p.
21. Khalil SH, Abu-Amero KK, Al Mohareb F, Chaudhri NA. Molecular monitoring of response to imatinib (Glivec) in chronic myeloid leukemia patients: experience at a tertiary care hospital in Saudi Arabia. Genetic testing and molecular biomarkers. 2010;14(1):67-74. Epub 2009/12/01. doi: 10.1089/gtmb.2009.0126. PubMed PMID: 19943786.
22. Alsobhi E, Abrar MB, Abdelaal M, Alsaeed A, Absi A, Alzahrani Z, et al. Response to imatinib therapy in adult patients with chronic myeloid leukemia in Saudi population: a single-center study. Leuk Lymphoma. 2014:1-5. Epub 2014/06/24. doi: 10.3109/10428194.2014.935365. PubMed PMID: 24956142.
23. How Is Chronic Myeloid Leukemia Staged? : The American Cancer Society 2017. Available from: https://www.cancer.org/cancer/chronic-myeloid-leukemia/detection-diagnosis-staging/staging.html.
24. Cismondia F, Fialho AS, Vieira SM, Retic SR, Sousab JMC, Finkelstein SN. Missing data in medical databases: Impute, delete or classify? Artificial Intelligence in Medicine. 2013;58(1):63-72. doi: 10.1016/j.artmed.2013.01.003.
25. Starkweather J, Herrington R. Replace Missing Values. Secondary Replace Missing Values 2014. Available from: http://www.unt.edu/rss/class/Jon/SPSS_SC/Module6/SPSS_M6_1.htm.
26. Banjar HR, Alsobhi E. Consistency Test between Scoring Systems for Predicting Outcomes of Chronic Myeloid Leukemia in a Saudi Population Treated with Imatinib. International Scholarly Research Notices. 2017;2017:6. doi: 10.1155/2017/1076493.
27. Than H, Kuan L, Seow CH, Li W, Allen JC, Jr., Chuah C. The EUTOS Score Is Highly Predictive for Clinical Outcome and Survival in Asian Patients with Early Chronic Phase Chronic Myeloid Leukemia Treated with Imatinib. Blood. 2012;120(21). PubMed PMID: WOS:000314049601307.
28. Saranya C, Manikandan G. A Study on Normalization Techniques for Privacy Preserving Data Mining. International Journal of Engineering and Technology (IJET). 2013;5(3).
29. National Comprehensive Cancer Network Guideline Inc.: National Comprehensive Cancer Network; [cited 2014].
30. Baccarani M, Deininger MW, Rosti G, Hochhaus A, Soverini S, Apperley JF, et al. European LeukemiaNet recommendations for the management of chronic myeloid leukemia: 2013. 2013;122(6):872-84. doi: 10.1182/blood-2013-05-501569.
31. Kohl M. Performance Measures in Binary Classification. International Journal of Statistics in Medical Research. 2012;1:79-81.
32. Kuncheva LI. Combining Pattern Classifiers Methods and Algorithms: John Wiley & Sons,; 2004.
33. Todorovski L, Dzeroski S. Combining Multiple Models with Meta Decision Trees. Proceedings of the Fourth European Conference on Principles of Data Mining and Knowledge Discovery; Berlin, Germany: Springer; 2000. p. 54-64.
34. Christianini N, Shawe-Taylor J. An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press; Cambridge, UK2000.
35. Draper NR, Smith H. Applied Regression Analysis Hoboken, NJ: Wiley-Interscience; 1998.
36. Sun Y, Kamel MS, Wong AKC, Wang Y. Cost-sensitive boosting for classification of imbalanced data. Pattern Recognition. 2007:3358–78.
37. Lasserre J, Arnold S, Vingron M, Reinke P, Hinrichs C. Predicting the outcome of renal transplantation. J Am Med Inform Assoc. 2012;19(2):255-62. Epub 2011/08/31. doi: 10.1136/amiajnl-2010-000004. PubMed PMID: 21875867; PubMed Central PMCID: PMCPmc3277611.
38. Banjar H, Ranasinghe D, Brown F, Adelson D, Kroger T, Leclercq T, et al. Modelling Predictors of Molecular Response to Frontline Imatinib for Patients with Chronic Myeloid Leukaemia. PloS one. 2017;12(1):e0168947. doi: 10.1371/journal.pone.0168947. PubMed PMID: PMC5207707.
39. Varma S, Simon R. Bias in error estimation when using cross-validation for model selection. BMC Bioinformatics. 2006;7:91. PubMed PMID: 16504092.
40. Gu Q, Zhu L, Cai Z. Evaluation Measures of the Classification Performance of Imbalanced Data Sets. Computational Intelligence and Intelligent Systems2009. p. 461-71.
41. Choose Classifier Options: MathWorks; 2016 [cited 2016 1 Aug]. Available from: https://au.mathworks.com/help/stats/choose-a-classifier.html - bunt0p6-1.
42. NCCN Clinical Practice Guidelines in Oncology: Chronic Myelogenous Leukemia: National Comprehensive Cancer Network; 2017 [2017]. Available from: http://www.nccn.org/professionals/physician_gls/pdf/cml.pdf.
43. Baccarani M, Pileri S, Steegmann JL, Muller M, Soverini S, Dreyling M. Chronic myeloid leukemia: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up†. Annals of Oncology. 2012;23(suppl_7):vii72-vii7. doi: 10.1093/annonc/mds228.

44. Kasabov N, Filev D. Evolving Intelligent Systems: Methods, Learning, & Applications. International Symposium on Evolving Fuzzy Systems, IEEE. 2006.

Linking Chapter 4 and 5:

Treatment of patients with CML has become increasingly difficult in recent years due to the variety of treatment options available and the challenge of deciding on the most appropriate treatment strategy for an individual patient. The previous papers, "Consistency Test between Scoring Systems for Predicting Outcomes of Chronic Myeloid Leukemia in a Saudi Population," and "Combined Value of Validated Prognostic Scores and Resolving Conflict in Allocation to Risk Groups in Chronic Myeloid Leukemia Patients," were used to determine the risk groups for CML patients; however, considering the available treatment options, there are no published criteria for selecting the appropriate treatment. Therefore, predicting the molecular response and identifying the relationship between predictive factors, such as peripheral blood counts and molecular based and clinical based data, is a step forward in managing CML treatment.

In the following paper, entitled "Modelling Predictors of Molecular Response to Frontline Imatinib for Patients with Chronic Myeloid Leukaemia" we described our CML study in great detail. Within this study, we explain the nature of CML, how it usually develops, and the genes and chromosomes that are typically associated with the disease. Along with a thorough description of CML, the paper examines predictive factors and treatment methods that could be used in addressing the ever-growing problem of how to treat CML due to the large number of treatment options available for clinicians to use. In the Introduction, molecular monitoring is identified as the standard guide to the clinical management of CML. In the related works section of this paper, we define and discuss predictive factors and how they are used to predict treatment responses, including how predictive factors are related to MMR responses, how other factors depend on peripheral blood counts, how to select the most effective TKI therapies at the time of a diagnosis, and more. The Related Works Section of the paper also contains tables, definitions, and

discussions about the two common descriptive assays and the three common prognostic scoring systems used to identify CML patient risk groups. In the materials and methods section, the datasets, predictive factors, missing data, machine learning methods, and other methods used during the investigation are discussed in great detail. Near the end of the paper, in the Results Section, we explain how the number of patients and their various outcomes were reported in the TIDEL II dataset. We also describe our methods of data preparation for analysis and demonstrate the predictive factors associated with MMR while extracting the rules for prediction. The Results Section gives insight into the data collected and manipulated during the study, how any missing values were handled, and the method used to select predictive factors. In the paper, we determined that common prognostic scores yield similar results in sensitivity performance and therefore are good predictors for our designated positive group that was used during the study. We also found that the study's limitations included the fact that prior knowledge may change due to varying expert opinions. Finally, we elaborate on the research and planning methods that were used during the study. The Discussion Section will also shed light on the results of the study and how they can be applied in future clinical trials and professional clinical use.

This paper has investigated deeper into the study of CML, how it can be treated, and how the molecular responses of individuals can affect how their treatment plans are created.

Chapter 5: Modelling Predictors of Molecular Response to Frontline Imatinib for Patients with Chronic Myeloid Leukaemia

Statement of Authorship

Title of Paper	Modelling Predictors of Molecular Response to Frontline Imatinib for Patients with Chronic Myeloid Leukaemia
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Banjar, H., Ranasinghe, D., Brown, F., Adelson, D., Kroger, T., Leclercq, T., White, D., Hughes, T. and Chaudhri, N. (2017). Modelling Predictors of Molecular Response to Frontline Imatinib for Patients with Chronic Myeloid Leukaemia. PLoS ONE, 12(1), e0168947. http://doi.org/10.1371/journal.pone.0168947

Principal Author

Name of Principal Author (Candidate)	Haneen Reda Banjar		
Contribution to the Paper	HB designed and performed the research, data curation, formal analysis, investigation, methodology, project administration, resources, software, and validation. HB wrote the original draft.		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature	[redacted]	Date	14 Feb 2017

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate in include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Damith Ranasinghe		
Contribution to the Paper	DR contributed to the experimental design and the conception, and interpretation of data. DR revised the manuscript and approved the version of the manuscript to be published.		
Signature	[redacted]	Date	

Name of Co-Author	Fred Brown		
Contribution to the Paper	FB contributed to the experimental design and the conception, and interpretation of data. FB revised the manuscript and approved the version of the manuscript to be published.		
Signature	[redacted]	Date	17/2/17

Name of Co-Author	David Adelson		
Contribution to the Paper	DA contributed to the conception and interpretation of data, and revised the manuscript. DA also approved the version of the manuscript to be published.		
Signature	[redacted]	Date	17/2/17

Name of Co-Author	Trent Kroger		
Contribution to the Paper	TK contributed to the experimental design and the conception, interpretation of data and revised the manuscript.		
Signature		Date	

Name of Co-Author	Tamara Leclercq		
Contribution to the Paper	TL contributed to validation, drafts, and revised the manuscript. She also approved the version of the manuscript to be published.		
Signature	[redacted]	Date	15-2-17

Name of Co-Author	Deborah White		
Contribution to the Paper	DW contributed to the research, data curation, drafts, and revised the manuscript. She also approved the version of the manuscript to be published.		
Signature	[redacted]	Date	15/2/2017

Name of Co-Author	Timothy Hughes		
Contribution to the Paper	TH contributed to the research, data curation, resources, validation, interpretation of data, drafts, and revised the manuscript. TH also approved the version of the manuscript to be published.		
Signature	[redacted]	Date	15-2-17

Name of Co-Author	Naeem Chaudhri		
Contribution to the Paper	NC contributed to the validation, data curation and the conception, interpretation of data, drafts, and revised the manuscript. He also approved the version of the manuscript to be published.		
Signature	[redacted]	Date	20th February 2017

PLOS ONE

RESEARCH ARTICLE

Modelling Predictors of Molecular Response to Frontline Imatinib for Patients with Chronic Myeloid Leukaemia

Haneen Banjar[1,2]*, Damith Ranasinghe[1,3☯], Fred Brown[1], David Adelson[4], Trent Kroger[1†], Tamara Leclercq[5,6☯], Deborah White[5,6,7,8,9‡], Timothy Hughes[5,6,8,9,10‡], Naeem Chaudhri[11]

1 School of Computer Science, University of Adelaide, Adelaide, South Australia, Australia, **2** The Department of Computer Science, King AbdulAziz University, Jeddah, Saudi Arabia, **3** Auto-ID Lab, School of Computer Science, University of Adelaide, Adelaide, South Australia, Australia, **4** School of Molecular and Biomedical Science, The University of Adelaide, Adelaide, Australia, **5** Cancer Theme, South Australian Health and Medical Research Institute (SAHMRI), Adelaide. South Australia, Australia, **6** University of Adelaide, Discipline of Medicine, Adelaide, South Australia, Australia, **7** University of Adelaide, Discipline of Paediatrics, Adelaide, South Australia, Australia, **8** Centre for Cancer Biology, University of South Australia, Adelaide, South Australia, Australia, **9** Centre for Personalised Cancer Medicine, University of Adelaide, Adelaide, South Australia, Australia, **10** Haematology Department, SA Pathology, Adelaide, South Australia, Australia, **11** King Faisal Specialist Hospital and Research Centre, Oncology Center, Riyadh, Saudi Arabia

☯ These authors contributed equally to this work.
† Deceased.
‡ These authors also contributed equally to this work.
* Hrbanjar@kau.edu.sa

Check for updates

OPEN ACCESS

Citation: Banjar H, Ranasinghe D, Brown F, Adelson D, Kroger T, Leclercq T, et al. (2017) Modelling Predictors of Molecular Response to Frontline Imatinib for Patients with Chronic Myeloid Leukaemia. PLoS ONE 12(1): e0168947. doi:10.1371/journal.pone.0168947

Editor: Matthaios Speletas, University of Thessaly Faculty of Medicine, GREECE

Received: August 16, 2016

Accepted: December 8, 2016

Published: January 3, 2017

Copyright: © 2017 Banjar et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Due to ethical restrictions, de-identified datasets are available upon request to tim.hughes@sahmri.com for authorization. The Saudi data is restricted to only a few authorised users who are permitted to access the database with Clinical Research Committee (CRC) and Research Ethics Committee (REC) permission. The Saudi dataset is available upon request to chaudhri@kfshrc.edu.sa for authorization.

Abstract

Background

Treatment of patients with chronic myeloid leukaemia (CML) has become increasingly difficult in recent years due to the variety of treatment options available and challenge deciding on the most appropriate treatment strategy for an individual patient. To facilitate the treatment strategy decision, disease assessment should involve molecular response to initial treatment for an individual patient. Patients predicted not to achieve major molecular response (MMR) at 24 months to frontline imatinib may be better treated with alternative frontline therapies, such as nilotinib or dasatinib. The aims of this study were to i) understand the clinical prediction 'rules' for predicting MMR at 24 months for CML patients treated with imatinib using clinical, molecular, and cell count observations (predictive factors collected at diagnosis and categorised based on available knowledge) and ii) develop a predictive model for CML treatment management. This predictive model was developed, based on CML patients undergoing imatinib therapy enrolled in the TIDEL II clinical trial with an experimentally identified achieving MMR group and non-achieving MMR group, by addressing the challenge as a machine learning problem. The recommended model was validated externally using an independent data set from King Faisal Specialist Hospital and Research Centre, Saudi Arabia.

Principle Findings

The common prognostic scores yielded similar sensitivity performance in testing and validation datasets and are therefore good predictors of the positive group. The G-mean and F-

Funding: Financial support received by HB for this study was provided in part by a grant from King Abdulaziz University. The funder had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. TH received funds from Novartis Australia for the TIDEL II trial. Novartis had no role in the decision to publish and preparation of the manuscript.

Competing Interests: HB has no conflict of interest to declare. DLW receives honoraria and research funds from Novartis Pharmaceuticals, and BMS. DLW is also a member of Advisory Boards for Novartis. TPH receives honoraria and research funds from Novartis Pharmaceuticals, BMS and Ariad. TPH is also a member of Advisory Boards for Novartis. NAC receives honoraria from Novartis Pharmaceuticals and BMS, and research funds from Novartis. However, Novartis, BMS and Ariad had no role in the design of the study, collection and analysis of data. This does not alter the authors' adherence to PLOS ONE policies on sharing data and materials.

score values in our models outperformed the common prognostic scores in testing and validation datasets and are therefore good predictors for both the positive and negative groups. Furthermore, a high PPV above 65% indicated that our models are appropriate for making decisions at diagnosis and pre-therapy. Study limitations include that prior knowledge may change based on varying expert opinions; hence, representing the category boundaries of each predictive factor could dramatically change performance of the models.

Introduction

Chronic myeloid leukaemia (CML) is a malignant blood cancer that results in overproduction of myeloid cells in bone marrow, leading to significant increase in the number of immature white cells circulating in an affected person's blood. The causative chromosomal translocation, known as the Philadelphia chromosome (*Ph*), in CML gives rise to the *BCR-ABL1* gene, which encodes for the constitutively active tyrosine kinase Bcr-Abl [1]. The need for personalised medicine in CML refers to multiple active tyrosine kinase inhibitor (TKI) therapies available for CML, multiple strategies utilised for frontline CML therapy, and heterogeneity in responses. The first generation TKI, imatinib (IM) (Glivec®), is a standard strategy used over the past decade [2], but some patients exhibit a poor response to this therapy. These patients may benefit from a second generation TKI, such as nilotinib (NIL) (Tasigna®) or dasatinib (DAS) (Sprycel®). Each of these TKIs are currently approved for use as frontline treatment in CML. Therefore, frontline CML therapy occurs via one of two major strategies: i) frontline IM or ii) frontline second generation TKIs, such as NIL or DAS[3]. Hematologic, cytogenetic, and molecular strategies for monitoring patient responses to therapies are used by European LeukaemiaNet [4]. To monitor molecular response, real-time quantitative polymerase chain reaction (RQ-PCR) is used to quantify the level of *BCR-ABL1* mRNA transcripts in peripheral blood of patients. According to the international scale, the two main molecular responses are major molecular response (MMR; *BCR-ABL1* transcript $\leq$ 0.10%) and complete molecular response (CMR; *BCR-ABL1* transcripts not detectable). Molecular monitoring is considered a standard guide to clinical management in CML [5, 6]. Prediction of the long-term molecular response to frontline IM in CML can support clinicians to select optimum treatment protocols for CML patients. Patients predicted not to achieve MMR at 24 months may be better treated with alternative frontline therapies, such as NIL or DAS.

The application of machine learning in medicine reduces the gap between clinical research and clinical practice. This type of model may be useful for clinicians in decision-making by warning of specific problems or providing treatment recommendations [7]. There is a need to adapt machine learning technology to deal with the high complexity of the medical domain. Coping with the complexity of cancer patient management, we used machine learning to drive hidden information and transfer evidence into practice. Incorporating prior domain knowledge and raw data into machine learning algorithms makes the best use of information from various data sources. Using machine learning algorithms as a 'white box' model results in easier and more interpretable mathematical models that lead to simple and clear decisions. The decision is generated on the basis of expert experience. Thus, it is important to incorporate prior knowledge to classify medical data and identify the relation between different predictive factors.

CML treatment predictive models emulate the decision-making ability of a human expert and provide recommendations for clinicians based on early prediction of patient molecular responses to specific treatment. Thus, predictive models predicting MMR to TKI therapy from

a CML patient's clinical, molecular, and blood count factors at diagnosis have the potential to support clinicians manage CML treatment more effectively. This study aimed to i) understand whether there exist rules for predicting MMR at 24 months for CML patients treated with IM from the clinical, molecular, and cell count observations collected at diagnosis and categorised based on the available knowledge and ii) build a predictive model to predict MMR for IM in CML patients with better prediction results than those obtained with predictive assays and previous scores. CML patients predicted not to achieve MMR at 24 months may be better treated with alternative frontline therapies, including second generation TKIs, such as NIL and DAS.

Related work

A predictive factor is a patient characteristic used to predict treatment response [8]. Predictive factors related to MMR response include common molecular assays. Other factors depend on peripheral blood counts as well as molecular-based and clinical observations of the individual patient. In order to select the most effective TKI therapy at the time of diagnosis, various predictive factors in CML have been investigated to distinguish patients at increased risk of failure with IM, the first generation TKI [9–12]. Table 1 shows the current predictive assays and score systems, factors included in the score systems and methods used, target prediction, and published results.

Two common predictive assays, namely OCT-1 activity (OA)[13, 18, 19] and $IC50^{IM}$[14], have been studied to distinguish CML patients likely to achieve a good molecular response to IM. CML patients with the b3a2 *BCR-ABL1* transcript type, compared to those with the b2a2 transcript, demonstrate greater survival rates, while CML patients with the p190 transcript type are classified as high risk [20, 21].

Three common prognostic scoring systems have been developed to identify CML patient risk groups: the Sokal [15], Hasford [16], and European Treatment and Outcome Study (EUTOS)[17] scores. The Sokal score is derived using age, spleen size, platelet count, and peripheral blood blasts; the Hasford score also uses peripheral blood eosinophil and basophil percentage; and the EUTOS score is based on percentage of basophils and spleen size. These three scores ascertain the level of risk for CML patients by running multivariable regression analysis. However, these scoring systems were developed in the era when chemotherapy was the only therapy available. The EUTOS score was developed to predict cytogenetic response to IM therapy and failed to predict MMR[22]. Although prognostic scores are currently used to

Table 1. The current predictive assays and score systems, the factors included in score systems and the methods used; the target prediction and final results.

Previous methods				
Study	**Factors**	**Method**	**Target prediction**	**Data and Results**
White *et al.* [13]	OA (ng/200,000 cells)	Kaplan Meier Analysis	MMR by 60 months to IM	TIDEL I clinical trial (n = 56), High OA: 89%, and low OA: 55%
White *et al.* [14]	$IC50^{IM}$ (μM)	Kaplan Meier Analysis	MMR by 12 months to IM	TIDEL I clinical trial (n = 116), Low $IC50^{IM}$: 65%, and High $IC50^{IM}$: 39%
Sokal Score, Sokal *et al.* [15]	Age, spleen Size (cm), blast (%), and platelets (10^9/L)	Multivariate analysis of survival	Risk groups to chemotherapy	Six European and American sources (n = 813), Low 39%, intermediate 38%, and high 23%
Hasford Score, Hasford *et al.* [16]	Age, spleen size (cm), blasts (%), eosinophils (%), basophils (%)and platelets (10^9/L)	Multivariate analysis of survival	Risk groups to interferon alpha alone	14 studies (n = 981), Low 40.6%, intermediate 44.7%, and high 14.6%
EUTOS Score, Hasford *et al.* [17]	Basophils (%) and spleen Size (cm)	Multivariate analysis of response	CCgR at 18 months to IM	Five national study group (n = 2060), Low 79%, and high 21%

doi:10.1371/journal.pone.0168947.t001

personalise the care of CML patients by predicting response to therapy, they were developed either for identifying risk groups or for predicting cytogenetic response to therapy, but not for molecular response.

A recent study investigated the possible association between molecular response and a number of factors, such as Sokal score, age, sex, and IM dose [23], and found that female sex is a strong predictor. A recent review of biomarkers that determine prognosis in CML also presented a list of prognostic indicators at diagnosis, such as the three scoring systems, BCR-ABL1 transcript type, and OA[24]. However, to our knowledge, relations between predictive factors to predict molecular response have not previously been considered.

Materials and Methods

Dataset

Patients from the Therapeutic Intensification in De Novo Leukaemia (TIDEL II) [25] clinical trial were eligible for the study. Two sequential cohorts of 105 CML patients (total of 210) received IM 600 mg as the initial therapy. All patients were in chronic phase and monitored for time-dependent molecular response targets. For the long-term prediction, we considered achieving a *BCR-ABL1* transcript level $\leq$ 0.1% at 24 months using RQ-PCR. When CML patients failed to achieve molecular response targets, they were either dose-escalated to 800 mg IM or switched to NIL. Where intolerance or toxicity to IM was observed, patients were switched to NIL.

Patients enrolled in the TIDEL II trial were divided into two broad outcome groups: i) positive outcome (CML patients able to remain on IM and achieve target MMR at 24 months), and ii) negative outcome (CML patients who did not achieve MMR at 24 months on IM therapy). This study used inclusion and exclusion criteria; pregnant patients were excluded. Patients found to be intolerant to IM, who switched to NIL and then achieved MMR at 24 months, were removed from analysis as they could not be assessed as MMR failures to IM because they may have switched for non-biological reasons. Patients who achieved MMR on IM constituted the positive group and patients who did not achieve MMR on IM were considered the negative group, which included i) patients who did not achieve MMR at 24 months on IM; ii) patients, who had a suboptimal response to IM, switched to NIL and went on to achieve MMR at 24 months since MMR was not achieved by administering IM; and iii) patients who received IM followed by NIL and did not achieve MMR at 24 months.

Predictive Factors

We investigated the relation between MMR at 24 months and common predictive factors in the medical literature as mentioned in the Related Work section. Table 2 shows the list of predictive factors including description, factor type, and median with range values. All clinical, molecular, and predictive assays and peripheral blood factors were collected at the time of diagnosis, as follows: i) clinical factors: age, gender, and spleen size measured in centimetres below the costal margin; ii) molecular factors: *BCR-ABL1* transcript level pre-therapy and *BCR-ABL1* transcript type; iii) predictive assays: OA and $IC50^{IM}$; and iv) peripheral blood factors: white cell count (WCC), absolute neutrophil count (ANC), and eosinophil, basophil, monocyte, lymphocyte, platelet, and blast counts.

Impute Missing Data

Missing values are a common problem in clinical trials. Medical data is usually collected for specific purposes (diagnosis, monitoring or treatment) and medical research aims to achieve desired outcomes by designing clinical studies to test and validate specific medical hypotheses.

Table 2. Predictive factor descriptions, factor type and median with range values.

Factors	Description	Type	Median (Range)
Age (years)	Clinical factor recorded at the time of diagnosis	Continuous	49 (17–81)
Gender	Clinical factor recorded at the time of diagnosis	Categorical	
Spleen (cm)	Clinical factor measured by observation at diagnosis	Continuous	3.8 (0–30)
***BCR-ABL1* Transcript Type**	Genetic factor identified by quantitative PCR analysis to *BCR-ABL1*. Transcript b2a2 or b3a2 is distinguished only by the absence of 75 nucleotides. Both b2a2 and b3a2 occur in patients with linked polymorphisms within exon 13 (b2) and intron 13 of the *BCR* gene.	Categorical	
OA (ng/200,000 cells)	The OCT-1 protein activity as a protein function can be measured by uptake in the presence and absence of a specific OCT-1 inhibitor in mRNA.	Continuous	4.7 (0–16.32)
$IC50^{IM}$ (μM)	Biological factor measured as the concentration of IM producing a 50% decrease in the level of p-Crkl.	Continuous	1 (0.2–4.5)
***BCR-ABL1* level pretherapy (at diagnosis)**	Real-time quantitative polymerase chain reaction (RQ-PCR) can measure the level of *BCR-ABL1* transcripts in the peripheral blood of the patient.	Continuous	107.14 (1.96–969)
ANC (10^9 /L)	Biological factor (neutrophil, granulocytes) that can be measured from peripheral blood	Continuous	32.59 (0.5–219.2)
Monocytes (10^9 /L)	Biological factor in white blood cells that can be measured from peripheral blood	Continuous	1.7 (0–13.02)
Lymphocytes (10^9 /L)	Biological factor in white blood cells that can be measured from peripheral blood	Continuous	3.37 (0–13.9)
Basophils (10^9 /L)	Biological factor in white blood cells that can be measured from peripheral blood	Continuous	2.16 (0–38.89)
Eosinophils (10^9 /L)	Biological factor in white blood cells that can be measured from peripheral blood	Continuous	1.11 (0–17.81)
WCC (10^9 /L)	Biological factor and important cells in CML that can be measured from peripheral blood.	Continuous	49.6 (1.1–353.50)
Blasts (10^9 /L)	Biological factor that can be measured from peripheral blood	Continuous	1.27 (0–13.9)
Platelets (10^9 /L)	Biological factor that can be measured from peripheral blood	Continuous	485 (91–1219)
Sokal Score[15]	Risk score developed in 1984	Continuous	1 (0.45–8.08)
Hasford Score[16]	Risk score developed in 1998	Continuous	801 (0–2137.6)
EUTOS Score[17]	Risk score developed in 2011	Continuous	46.94 (0–228.5)

doi:10.1371/journal.pone.0168947.t002

Medical data may involve incomplete data or missing values which may be caused by a lack of information or discontinuation of study. There are three types of missing data: missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR). In the first type, MCAR, the absence of predictive factors value is unrelated to the outcome or other predictive factors, while the second type, MAR, the missing predictive factor is not related to missing values but may be related to observed data. The third type, MNAR, is considered non-ignorable missingness where the predictive factor missing value depends on the factor itself [26]. We aimed to have high quality predictive performance on average for all future cases using the population-based model, in contrast to the construction of patient-specific models, which are influenced by the particular predictive factor of the patient case at hand. Our population-based model is commonly used to perform well on average in all future cases. Therefore, we considered 10% as the cut-off for including a predictive factor in the analysis, since removing patient data has the undesired effect of reducing patient datasets, thus affecting the performance of the learned predictive model [27]. Missing data were treated using imputation [28] and not by removing patients who might carry information for prediction from other molecular data available. In imputation techniques, missing values are replaced with values estimated from suitable statistical methods based on information available in the dataset. For example, the imputation of missing values for blood counts were derived

from known values of blood count range, and the effect of missing data on estimate of variance was beyond the scope of this study. We used linear interpolation [29] values to impute continuous and categorical data. We used the impute command in SPSS to replace missing values with linear interpolation estimation values from the last valid value before the missing value and the first valid value after the missing value. The corresponding correlation values for the original dataset (with missing values) and completed dataset (without missing values) were calculated to indicate whether imputation affected the dataset.

Reformatting Predictive Factors Using Domain Knowledge

We reformatted factor values of data stored as text, numerals, or mixed type using existing knowledge, such as standard boundaries of blood counts, domain knowledge of clinical expertise, and previous medical publications [13, 14, 30]. Although the selected machine learning techniques have the capacity to handle continuous predictor values, reformatting the data by categorising each factor in the TIDEL II dataset into subgroups assisted comprehensibility of the final predictive model [31]. For example, we categorised the value of $IC50^{IM}$ equal to 0.5 $_{\mu M}$ as low $IC50^{IM}$ and reformatted the index number of the category. If the final model selected $IC50^{IM}$ as a relevant predictive factor, we used these categories to distinguish the predictive group on test patients.

Machine Learning Methods

The main goal was to produce a useful predictive model that is understandable for all users. Machine learning, such as SVM and neural networks, are difficult to interpret, while logistic regression and naive Bayes allow easy interpretation of results. However, the main issue with these algorithms are the strong assumptions of conditional independence between predictive factors. They also assess the contribution of each predictive factor to classification, but not the relations. The *k*-nearest neighbours[32] technique is another well-known machine leaning algorithm, but it is sensitive to local structure of data. Classification and regression trees (CART) [33] tend to be easily interpreted by clinicians. This method has the ability to learn relationships between predictive factors and molecular response. The importance of a clear predictive model stems from the need to trust the computation to predict response. In addition, clinicians need to understand model recommendations to explain the reasons for their decision [31].

CART is a binary recursive partitioning process capable of processing continuous and categorical data as predictors or outcomes. The CART mechanism produces the optimal tree after pruning based on a cost function to avoid overfitting in the maximal tree. The steps are provided below for the basic algorithm of a decision tree, previously having been described in [33]:

1. The top-down recursive and divide-and-conquer style is used to construct the tree.
2. The root node is located in the top-most node of the tree.
3. Each node denotes a test on a factor and each branch indicates an outcome of the test, where the leaf nodes represent classes. For selected factors, the data are recursively partitioned. Here, a splitting criterion called the Gini Impurity Measure is used to determine the best split in each node.
4. For given node t, the Gini index calculates the relative frequency of class c at node t as in (1):

$$Gini(t) = 1 - \sum_{c} [p(c|t)]^2 \quad (1)$$

The following scenarios demonstrate the possible indication of using the Gini index

measure. In the worst scenario, patient outcomes in training data using the examined split value of predictive factor are equally distributed between both classes at the node that maximises Gini value to indicate the least interesting information. However, in the best scenario, the minimum Gini value is the most interesting information for ascertaining when all patient outcomes in training data belong to one class using the examined split value of predictive factor [34].

5. The following three conditions are used to stop splitting:

- For the given node, all the tested data belong to the same class.
- No factors remain for splitting.
- No tested data are left for splitting

The starting point for this paper was a predictive model developed from a training set of patient cases. We used CART for re-expressing the decision tree as a clearly expressed set of clinical prediction rules that in the (IF..Then form) to identify relation between predictive factors and patient outcomes. CART has the ability to learn complex and non-linear relationships between factors and the response. The decision tree structure represented the extracted production rules [35]. Classifying a patient using a decision tree is effected by following a path of predictive factors through the tree to one of the leaves (patient response). This path from the root of the tree to a leaf establishes conditions which must be satisfied by any patient classified by that leaf. Thus, each leaf of a decision tree corresponds to a prediction rule. These rules are easier and lead to simple and clearer decisions which are more interpretable by clinicians than 'black box' mathematical models, such as SVM. The decision is generated on the basis of expert experience. The prediction rules were in the form, IF A *is S1* AND B is *S2*. . . THEN Response is X, where A and B are the predictive factors, *S1* and *S2* are the subcategories that belong to A and B, and X is the class (achieving MMR or not achieving MMR).

Predictive Factor Selection

Prior to using the machine learning algorithm feature selection, algorithms were often used to select a relevant subset of input features (in our problem, a subset of predictive factors to deliver a highly predictive model) [36, 37]. This is also very important in the context of healthcare costs where fewer input factors imply fewer diagnostic tests to obtain relevant predictive factors [38]. We also need to extract relations between the most related predictive factors and to understand whether prediction rules exist. We divided the feature selection process into two main types:

i) Knowledge-driven method for feature selection, such as existing medical knowledge regarding the predictive factor as an informative feature or clinical expert judgment on molecular factors associated with predicting MMR, known as manual feature selection [39], including

- Predictive assays: OA, $IC50^{IM}$.
- Molecular predictive factors: OA, $IC50^{IM}$, *BCR-ABL1* transcript level pre-therapy, *BCR-ABL1* transcript type.

ii) Data-driven methods for feature selection, known as automatic feature selection. We used the wrapper approach[40] where all subsets of features are evaluated using a given machine learning approach. The models resulting from the wrapper with each machine learning algorithm *i* were $i = 1,2,\ldots 2^n$, where *n* is the number of predictive factors. CART was run on the training data repeatedly using those subsets where predictive factor selection was only in the root node.

Evaluation Measurements

Performance of predictive models was measured by using a coincidence matrix. The problem was presented as a binary classification where the test outcome was positive (achieved MMR) or negative (did not achieve MMR). Results were divided into four conditions: (a) CML patients correctly identified as not achieving MMR ('True negative' (TN)); (b) patients achieving MMR wrongly identified as not achieving MMR ('False negative' (FN)); (c) patients correctly identified as achieving MMR ('True positive' (TP)); and (d) patients not achieving MMR wrongly identified as achieving MMR ('False positive' (FP)). We reported 'accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) using the following Eqs (2–6):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{3}$$

$$Specificityy = \frac{TN}{TN + FP} \tag{4}$$

$$PPV = \frac{TP}{TP + FP} \tag{5}$$

$$NPV = \frac{TN}{TN + FN} \tag{6}$$

All predictive models were trained to minimise misclassification rate between predicted MMR and actual MMR. In our predictive model, data were imbalanced in the positive and negative patient groups. Two measures, G-mean (geometric mean) and F-score (weighted harmonic mean of sensitivity and PPV), have often been used to assess performance of a predictive model trained on imbalanced data as in (7 and 8):

$$G - mean = \sqrt{sensitivity * Specificity} \tag{7}$$

$$F - score = \frac{2 * Sensitivity * PPV}{Sensitivity + PPV} \tag{8}$$

It is important to measure the balance between sensitivity and specificity, which means the model correctly predicted both response groups (achieving MMR or not achieving MMR). We also reported PPV (probability that a patient achieving MMR was correctly predicted to achieve MMR) because high PPV means that few patients will be unpredicted, which is crucial when making decisions in diagnosis pre-therapy. The PPV was calculated from the study test data population, in which the prevalence was 48%.

Nested Cross-Validation

In traditional cross-validation, 10% of data (1 fold) is used for testing and the remainder for training (9 folds), with training and testing performance repeated 10 times. Standard deviation of performance of the 10 predictive models was estimated, considering they were independent. The confidence interval was obtained from the mean and standard deviation at 95%

confidence level. The summary of test performances calculated on unseen folds was considered as the final performance.

However, when data are scarce, an extra layer of cross-validation should be performed. Since the test set cannot be touched (it is saved to evaluate the final models), new cross-validation was conducted on the training set. This technique is known as nested cross-validation [41]. In this case, we separated the training set into training and validation (~75%). The remaining 25% was used as the testing set. The model was trained on the training set; features were selected on the validation set; and performance was evaluated on the test set. We treated inner cross-validation as part of the model fitting procedure. To avoid overfitting, we compared the difference between training performance and inner cross-validation of the selected model. This procedure may be a good estimator of error for finding the best predictive model and predictive factor selection [42]. Finally, we reported accuracy, sensitivity, specificity, PPV, NPV, G-mean, and F-score of the testing set.

Model Selection

To compare resulting models we used the G-mean and F-score as criteria for model selection. The wrapper approach often needs evolutionary calculations, leading to extensive processing expense. Here, we ranked the resulting models based on G-mean and F-score measurements and selected the highest values.

The last step in our method was to compare the models generated using different feature selection techniques with current predictive assays and score systems (Table 1) based on G-mean and F-score performance to make a final recommendation. The full procedures are shown in (Fig 1), which is the schema of the CML predictive model.

External Validation Dataset

The external dataset was obtained from a tertiary care hospital, King Faisal Specialist Hospital and Research Centre (KFSHRC), Riyadh, Saudi Arabia. There were 172 adult CML patients who used frontline TKI[43]. Only patients using frontline IM with observed MMR at 24 months were selected. We performed pre-process steps to prepare the dataset for validation. We applied inclusion and exclusion criteria, imputed missing values and reformatted predictive factors using domain knowledge. We used evaluation measurements on predicted response by the recommended model versus observed response: accuracy, sensitivity, specificity, PPV, NPV, G-mean, and F score.

Ethics Statement

The data were analysed anonymously. All study participants provided written informed consent prior to participation. The TIDEL II trial is registered at www.ANZCTR.org.au as ACTRN12607000325404 and funded by Novartis Australia. The TIDEL II was carried out with the approval of human research ethics committees (RAH Protocol No 070718c, and ethically approved by the National Statement on Ethical Conduct in Human Research (NHMRC), 2007 and in accordance with the Declaration of Helsinki. The KFSHRC data ethically approved by Clinical Research Committee (CRC) and Research Ethics Committee (REC), 2005, Research Advisory Council (RAC) reference: ORA/1811/26, Proposal No 2051056.

Results

In this section, we first report the number of patients for various outcomes in the full TIDEL II dataset. Then, we describe data preparation for analysis, including imputing missing values

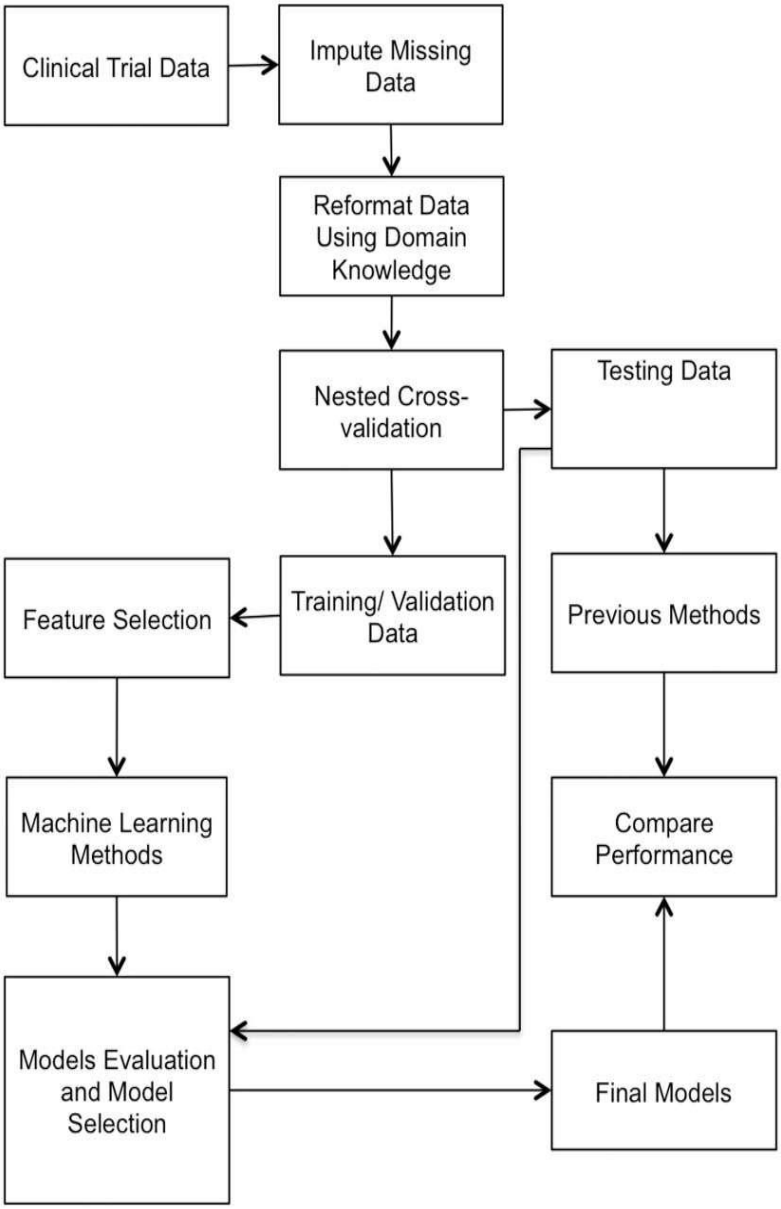


Fig 1. The schema for the CML predictive model, building, evaluation, and final model selection. To build the predictive model, we studied a clinical trial, preparing data for analysis by imputing missing values and reformatting factors using comprehensive standard boundaries to create subcategories for each predictive factor based on domain knowledge. For evaluation and final model selection, the nested design was used to split the dataset into training, validation and testing sets. The model was trained on the training set, features were selected on the validation set, and performance was evaluated on the test set. The final models were compared with previous methods.

doi:10.1371/journal.pone.0168947.g001

and reformatting predictive factors using domain knowledge. We reveal predictive performance of the machine learning technique and the effect of feature selection methods on selecting the final model. We also present evaluation results on unseen data based on a comparison

with previous methods. Finally, we demonstrate the strong predictive factors associated with MMR at 24 months and extract the rules for prediction.

Insight into the Data

In the TIDEL II clinical trial, among the 210 patients and under the inclusion/exclusion criteria, analysis included 173 (82.3%) CML patients. A positive outcome was observed in 102 (48.5%) patients able to remain on IM and achieve the target of MMR at 24 months (positive outcome), compared with 71 (33.8%) patients showing a negative outcome. The remaining 37 (34.7%) patients were excluded from analysis due to pregnancy or intolerance to IM, who were switched to NIL and then achieved MMR at 24 months. We split the TIDEL II cohort into training and testing sets. The training set included 123 (58.5%) patients and was used for inner cross-validation; the remaining 50 (23.8%) patients (testing set) were utilised for comparison between published predictive methods and our final models. The positive group was comprised of 76 training (74%) and 26 testing (26%) patients, while the negative group was compromised of 47 training (66%) and 24 testing (34%) patients. The 71 patients in negative group included i) 15 training and 4 testing of 19 (11%) patients who did not achieve MMR at 24 months on IM; ii) 13 training and 6 testing of 19 (11%) patients who had suboptimal response to IM, were switched to NIL and went on to achieve MMR at 24 months as MMR was not achieved by administering IM; and iii) 19 training and 14 testing of 33 (19%) patients who received IM followed by NIL and did not achieve MMR at 24 months. Additional details about the number of patients in the study and inclusion and exclusion criteria are shown in Fig 2.

Imputation for Missing Values

Correlation coefficient results obtained with original data (with missing values) and complete data (missing values imputed by linear interpolation) are presented in (Tables A-C in S1 Text). We used the complete dataset for analysis. In TIDEL II, missing values were <10% as well as the external validation dataset.

Reformat Predictive Factors Using Domain Knowledge

Knowledge was derived from standard boundaries of blood counts, clinical expertise, and previous medical publications [13, 14, 30]. The categories for each predictive factor used to transform data into categorical data and number of patients are shown in Table 3.

Predictive Factor Selection and Prediction Results

The machine learning algorithm was trained for each feature selection method, once with all features, once with molecular features, and once with all subsets using the wrapper approach. In the wrapper approach, all subsets were trained to include different predictive factors. Here, model selection criteria were accuracy, G-mean, and F-score, which resulted from inner cross-validation performance on the training set. Predictive performance of feature selection methods with the machine learning technique on the training set is provided in Table 4. In the table, training performance and inner cross-validation performance are also presented to show fit of the models. Models A, B, and C have overfitting problems where training accuracy (A: 81%, B: 64%, C: 70%) is much larger than the cross validation mean of accuracy (A: 51%, B: 57%, C: 50%). However, the three models selected based on achieving high cross-validation mean performances (accuracy, G-mean, F-score) achieved training accuracy (D: 78%, E: 77%, F: 73%), compared with the cross-validation mean of accuracy (D: 76%, E: 75%, F: 76%); therefore, these were better models.

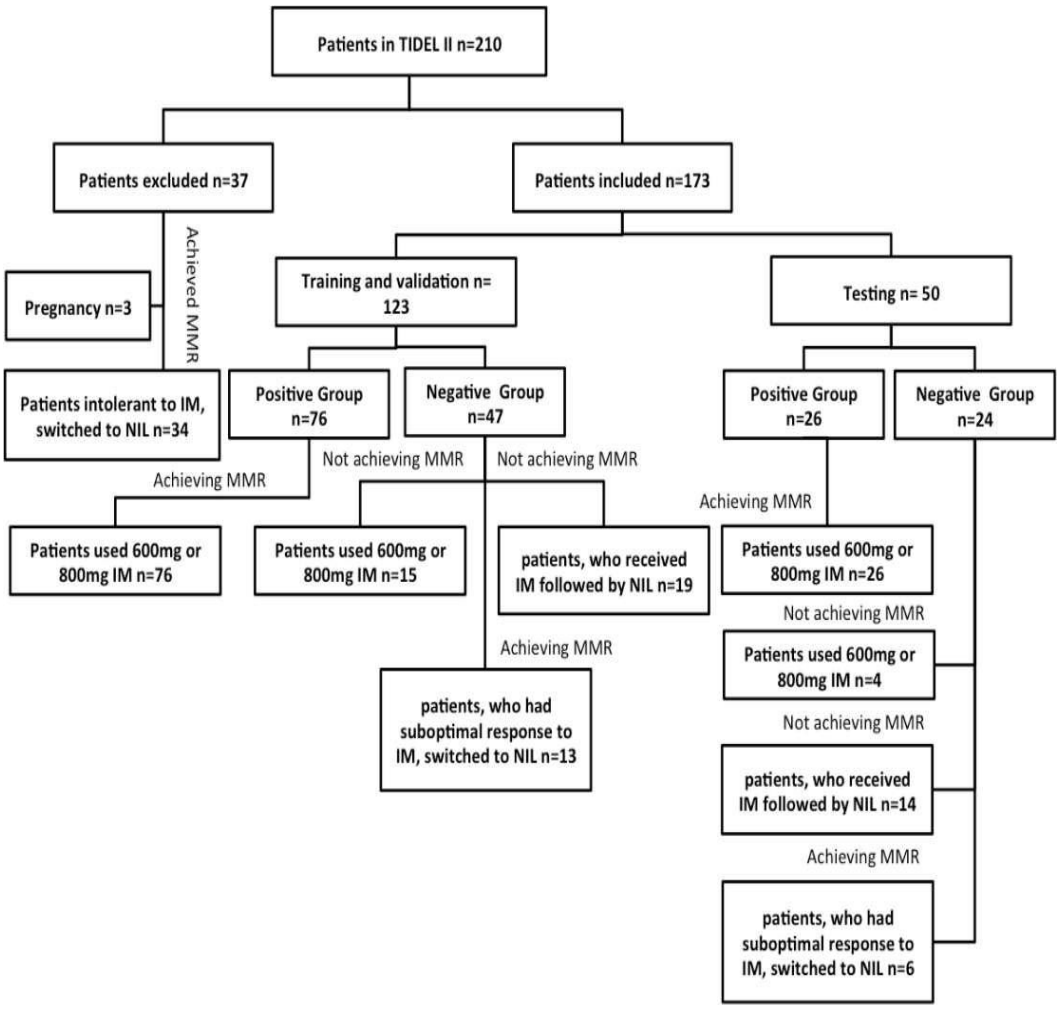


Fig 2. TIDEL II patients in this study. Inclusion and exclusion criteria.

doi:10.1371/journal.pone.0168947.g002

Comparison of Final Models and Previous Methods

We measured performance on the testing set of single predictive assays commonly used to predict MMR (OA and $IC50^{IM}$) and common prognostic risk scores (Sokal, Hasford, EUTOS) (Table 5). OA achieved the highest accuracy (68%) compared with other previous methods ($IC50^{IM}$: 54%, Sokal: 58%, Hasford: 56%, EUTOS: 52%). However, Model D (accuracy: 72%) outperformed OA (accuracy: 68%). Although, the Hasford score accurately predicted the positive group (those that achieved MMR at 24 months) with a sensitivity of 92%, and $IC50^{IM}$ accurately predicted the negative group (those that did not achieve MMR at 24 months) with a specificity of 75%, G-mean ($IC50^{IM}$: 50%, Hasford: 39%) did not exceed that of OA (67%). The three scores achieving high sensitivity (Sokal: 84%, Hasford: 92%, EUTOS: 84%) are therefore good predictors of the positive group. The highest G-mean and F-score values among previous methods were achieved by OA (G-mean: 67%, F-score: 69%). On the other hand, Models D, E, and F had G-mean and F-score values that outperformed OA performance. In addition, our models achieved PPV values better than previous methods. The PPV performances in our

Table 3. The categories for each predictive factor used to transform data into categorical data and number of patients.

Factors	Categories	Patient in TIDEL II		Patient in Saudi Population	
		No. of patient	Patient %	No. of patient	Patient %
Age (years)	Young ≤30	21	12.21%	36	33.03%
	Middle Age>30, ≤60	104	60.47%	67	61.47%
	Older>60	47	27.33%	6	05.50%
Gender	Male	92	40.70%	45	41.28%
	Female	118	59.30%	64	58.72%
Spleen (cm)	Not palpable ≤1	99	57.56%	52	47.71%
	Small >1, ≤10	44	25.58%	34	31.19%
	Large >10	28	16.28%	23	21.10%
***BCR-ABL1* Transcript Type**	b2a2	68	39.53%	None	None
	b3a2	68	39.53%	None	None
	Both	34	19.77%	None	None
	e1a2	2	1.16%	None	None
OA (ng/200,000 cells)	Low ≤4	80	46.51%	None	None
	Standard >4	87	50.58%	None	None
$IC50^{IM}$ (μM)	Group 1 ≤0.5	19	11.05%	None	None
	Group 2 >0.5 ≤0.7	31	18.02%	None	None
	Group 3 >0.7 ≤0.95	31	18.02%	None	None
	Group 4 >0.95	79	45.93%	None	None
***BCR-ABL1* level pretherapy (at diagnosis)**	Low ≤20	8	4.65%	None	None
	Moderate>20, ≤100	96	55.81%	None	None
	High>100	66	38.37%	None	None
ANC (10^9 /L)	Low <1.8	3	1.74%	None	None
	Normal ≥1.8, ≤7.5	35	20.35%	None	None
	High >7.5, ≤ 50	97	56.40%	None	None
	Very High >50	36	20.93%	None	None
Monocytes (10^9 /L)	Low <0.2	16	9.30%	12	11.01%
	Normal ≥0.2, ≤0.8	46	26.74%	23	21.10%
	High >0.8	109	63.37%	74	67.89%
Lymphocytes (10^9 /L)	Low <1	12	6.98%	None	None
	Normal ≥1, ≤3.5	98	56.98%	None	None
	High >3.5	61	35.47%	None	None
Basophils (10^9 /L)	Normal ≤0.1	28	16.28%	30	27.52%
	High >0.1, ≤1	76	44.19%	10	9.17%
	Very High >1	67	38.95%	69	63.30%
Eosinophils (10^9 /L)	Normal ≤0.5	91	52.91%	28	25.69%
	High >0.5	79	45.93%	81	74.31%
WCC (10^9 /L)	Low <10	29	16.86%	None	None
	Normal >10, <100	122	70.93%	None	None
	High >100	20	11.63%	None	None
Blasts (10^9 /L)	Normal >0, ≤5	154	89.53%	99	90.83
	High >5	13	7.56%	10	9.17%
Platelets (10^9 /L)	Low ≥20, <150	3	1.74%	8	7.34%
	Normal >150, ≤400	91	52.91%	53	48.62%
	High >400	75	43.60%	48	44.04%
Sokal Score	Low, intermediate≤1.2	132	79.51%	92	84.40%
	High>1.2	34	20.48%	17	15.60%

(*Continued*)

Table 3. (*Continued*)

Factors	Categories	Patient in TIDEL II		Patient in Saudi Population	
		No. of patient	Patient %	No. of patient	Patient %
Hasford Score	Low, Intermediate<1480	154	93.9%	101	92.66%
	High≥1481	10	6%	8	7.34%
EUTOS Score	Low<87	140	83.33%	92	84.40%
	High≥87	28	16.67%	17	15.60%

Each predictive factor and the number of CML patients included in the study. Haematologist experts and previous publications identified the categories.

doi:10.1371/journal.pone.0168947.t003

models (A: 73%, B: 84%, C: 80%, D: 88%, E: 84%, F: 96%) were higher than PPV in OA (67%). These high PPV values indicated that our models could be trusted in making decisions at diagnosis pre-therapy.

Clinical Prediction Rules and Extraction of Relations between the Predictive Factors and MMR Predictions

In this section, we demonstrate two findings: i) the selected predictive factors by CART algorithm and ii) clinical prediction rules that clearly expressed the (IF..Then) conditional relationship between the predictive factors and the MMR predictions.

Firstly, the CART uses only those factors that help separate response groups, while other factors are not considered. All predictive factors were examined at each node to assess splitter effectiveness. The decision tree generated by all features (Figure A in S1 File) had eight related factors of 15 predictive factors after pruning the maximal tree: age, spleen size, platelets, eosinophil, WCC, monocytes, $IC50^{IM}$, and *BCR-ABL1* transcript type.

Secondly, we extracted clinical prediction rules from selected models. The rule set is attached in (S1 Table) and (Figures B and C in S1 File) represented the tree structure for model E and model F respectively. To express the conditional relationship between predictive factors and MMR predictions we presented examples from the molecular decision trees and recommended model (model D). The first example is model B and model C that used molecular predictive factors, $IC50^{IM}$ divided patients into two groups. The groups with $IC50^{IM}$ >0.5 **μM** and $IC50^{IM}$ <0.95 **μM** were classified as the positive group, while for patients with $IC50^{IM}$ <0.5 **μM** and $IC50^{IM}$ >0.95 **μM**, OA values can help identify MMR group. In model C, adding *BCR-ABL1* transcript level and transcript type identified further relations. In the group with $IC50^{IM}$ >0.5 **μM** and $IC50^{IM}$ <0.95 **μM**, patients who had the b2a2 type may achieve MMR, while the other type need more information about OA and *BCR-ABL1* transcript level to identify MMR group. We also noticed that different MMR groups were identified based on the same *BCR-ABL1* transcript level. Thus, $IC50^{IM}$, OA, and *BCR-ABL1* transcript type affected the role of *BCR-ABL1* transcript level. We also demonstrate these relations in Fig 3.

The second example is model D. Although the CART with wrapper approach recommended three models, model D showed best performance. In Fig 4, the conditional relation between age, spleen size and MMR prediction is simply structured in the tree's right branches. For example, the first clinical prediction rule was IF spleen size belongs to the large size group >10 cm AND age belongs to the young group ≤30 or older group >60 THEN the patient may achieve MMR. On the other hand, IF spleen size belongs to the large size group >10 cm AND age belongs to the middle age group >30 and ≤60 THEN the patient may not achieve MMR. The accuracy of models measures how likely extracted rules are to correctly identify the MMR group (Table 4, training performance).

Table 4. Predictive performance of feature selection methods with the machine-learning technique from nested cross-validation methods, wrapper method used the highest cross validation performances for selecting the final models.

Feature selection approaches	Model name	Features	Training Performance							Cross-validation Performance						
			Accuracy	Sensitivity	Specificity	PPV	NPV	G-mean	F-score	Accuracy	Sensitivity	Specificity	PPV	NPV	G-mean	F-score
All features	**A**	1	0.81	0.83	0.77	0.86	0.72	0.80	0.84	0.51 (0.43,0.59)	0.60 (0.53,0.67)	0.31 (0.17,0.45)	0.63 (0.49,0.77)	0.33 (0.15,0.51)	0.37 (0.20,0.54)	0.61 (0.46,0.76)
Molecular features	**B**	8,9	0.64	0.67	0.54	0.80	0.38	0.60	0.72	0.57 (0.5,0.64)	0.62 (0.55,0.69)	0.42 (0.23,0.61)	0.75 (0.63,0.87)	0.28 (0.15,0.41)	0.46 (0.29,0.63)	0.67 (0.55,0.79)
	C	8,9,10,16	0.70	0.73	0.64	0.81	0.53	0.68	0.76	0.50(0.38, 0.62)	0.60 (0.48,0.72)	0.31 (0.17,0.45)	0.59 (0.43,0.75)	0.36 (0.16,0.56)	0.38 (0.21,0.55)	0.59 (0.43,0.75)
The highest Cross-validation																
Accuracy	**D**	2,3,7,13,15	0.78	0.77	0.81	0.92	0.57	0.79	0.83	**0.76 (0.62,0.78)**	0.78 (0.71,0.85)	0.71 (0.51,0.91)	0.88 (0.78,0.98)	0.57 (0.39,0.75)	0.71 (0.54,0.88)	0.82 (0.65,0.99)
G-mean	**E**	2,3,6,7,8,10,15,16	0.77	0.77	0.75	0.88	0.59	0.76	0.82	0.75 (0.69,0.81)	0.79 (0.72,0.86)	0.75 (0.61,0.89)	0.85 (0.76,0.94)	0.59 (0.42,0.76)	**0.76 (0.67,0.85)**	0.81 (0.69,0.93)
F-score	**F**	3,7,8,15	0.73	0.72	0.82	0.94	0.40	0.77	0.81	0.76 (0.68,0.84)	0.78 (0.71,0.85)	0.71 (0.51,0.91)	0.88 (0.78,0.98)	0.57 (0.39,0.75)	0.71 (0.54,0.88)	**0.83 (0.75,0.91)**

The features indexes are: 1 = all feature, 2 = Age, 3 = Spleen Size, 4 = Platelets, 5 = Basophils, 6 = Eosinophils, 7 = Blast, 8 = OA, 9 = IC50IM, 10 = BCR-ABL1 transcript level pre therapy, 11 = WCC, 12 = ANC, 13 = Monocytes, 14 = Lymphocytes, 15 = Gender, and 16 = BCR-ABL1 Transcript type. For each model the table gives the training and Cross-validation Performances. In cross validation performance, the means obtained from 10-fold cross validation and 95% confidence intervals.

doi:10.1371/journal.pone.0168947.t004

Table 5. The comparison between previous methods and, our predictive models.

		Testing Performance						
		Accuracy	Sensitivity	Specificity	PPV	NPV	G-mean	F-Score
Previous Methods	OA[13]	**0.68**	0.73	0.62	**0.67**	**0.68**	**0.67**	**0.69**
	$IC50^{IM}$[14]	0.54	0.34	**0.75**	0.60	0.51	0.50	0.43
	Sokal score[15]	0.58	0.84	0.29	0.56	0.63	0.49	0.67
	Hasford Score [16]	0.56	**0.92**	0.16	0.54	0.66	0.39	0.68
	EUTOS Score [17]	0.52	0.84	0.16	0.52	0.50	0.37	0.64
Our Model s	Model A	0.60	0.59	0.61	**0.73**	0.45	0.60	0.65
	Model B	0.62	0.59	0.69	**0.84**	0.37	0.64	0.69
	Model C	0.58	0.56	0.61	**0.80**	0.33	0.59	0.65
	Model D	**0.72**	**0.67**	**0.81**	**0.88**	**0.54**	**0.74**	**0.76**
	Model E	0.66	0.62	0.73	**0.84**	0.45	**0.67**	**0.71**
	Model F	0.64	0.59	**0.87**	**0.96**	0.29	**0.72**	**0.73**

The bolded value indicated the comparative values between our methods and the previous methods, Model A = all predictive factors, Model B = OA and $IC50^{IM}$, Model C = OA, $IC50^{IM}$, *BCR-ABL1* Transcript level Pretherapy and *BCR-ABL1* Transcript Type, Model D = CART algorithm with the highest accuracy value, Model E = CART algorithm with the highest G-mean value, and Model F = CART algorithm with the highest F-score value.

doi:10.1371/journal.pone.0168947.t005

External Validation of the Final Model

The Model D was validated in the Saudi cohort. Inclusion and exclusion criteria were applied to the dataset and the (Figure D in S1 File) shows the number of patients in the validation. Table 3 shows the number of patients included in the validation for each predictive factor. We included 109 patients who used frontline 400mg IM, while 63 (36%) patients were excluded due to missing MMR values at 24 months (27 patients) or using frontline NIL or DAS (36 patients). There were 78 (72%) patients in the positive group who achieved MMR at 24 months and 31 (28%) patients in the negative group as follows: i) 24 (77%) patients who did not achieve MMR at 24 months on 400mg IM; ii) 2 (6%) patients who had a suboptimal response to IM, switched to NIL and went on to achieve MMR at 24 months as MMR was not achieved by administering IM; and iii) 5 (17%) patients who received IM followed by NIL or DAS and did not achieve MMR at 24 months.

We compared performance of the recommended model (Model D) and common prognostic risk scores (Sokal, Hasford and EUTOS) (see Table 6). Hasford score achieved the highest accuracy (68%) compared with Sokal and our model (Sokal: 63%, Hasford: 67% and Model D: 50%). In addition, the EUTOS score achieved a slightly lesser G-mean and F-score performance than our model D (the Saudi population's G-mean: 44% vs. the EUTOS's G-mean: 41%; and the Saudi population's F-score: 29% vs. the EUTOS's F-score: 25%); furthermore, compared to the prognostic scores, our model D achieved the highest G-mean (44%) and F-score (29%) overall. Although the common prognostic risk scores achieved better accuracies, highest specificity (35%) was found in our model compared with common prognostic scores (Sokal: 13%, Hasford: 6% and EUTOS: 19%) which confirmed that model D accurately predicted the negative group (those that will not achieve MMR at 24 months) with a specificity of 35%. Validation data from the Saudi population were performed similar to testing data from TIDEL II as the three scores achieving high sensitivity (Sokal: 83%, Hasford: 92%, EUTOS:

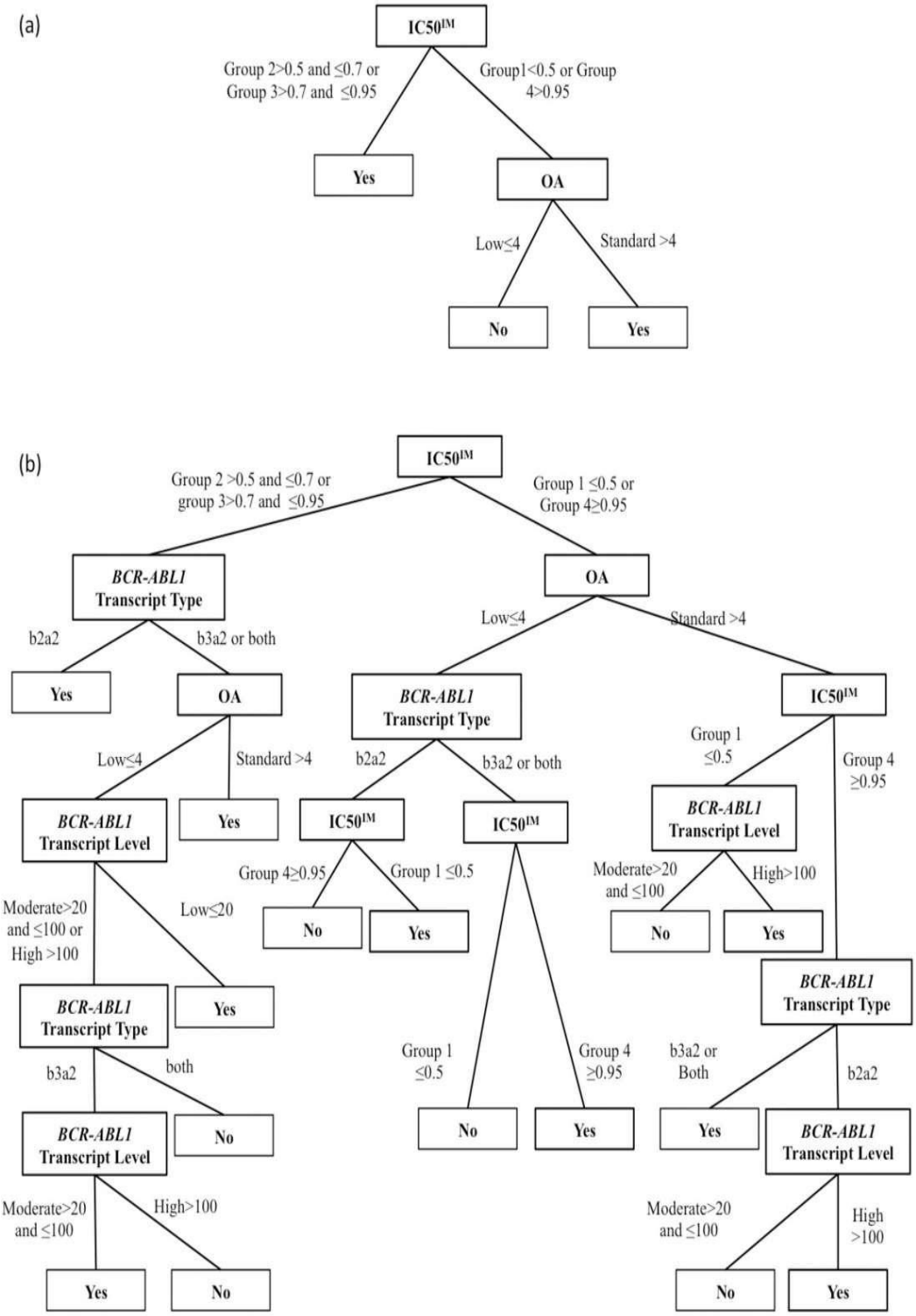


Fig 3. Tree structures for a) Model B and b) Model C. Relations between molecular predictive factors and MMR. The tree represents each predictive factor in nodes. The node has two possible splits: it is connected to either the second predictive factor or the MMR group of patients in the positive or

negative group. This graphical structure illustrates the predictive rules. Predictive rules can be used on unseen data to predict the target. A predictive rule of the form: IF (conditions) THEN (class) is equivalent to a path from the root node to leaf in the decision tree, (Yes: achieve MMR at 24 months) and (No: did not achieve MMR at 24 months).

doi:10.1371/journal.pone.0168947.g003

86%) are therefore good predictors of the positive group. Although PPV value in the EUTOS score (73%) outperformed model D (68%), PPV above 60% indicated that our model could be trusted in making decisions at diagnosis pre-therapy.

Discussion

We employed a widely used and practical machine learning technique to develop a predictive model to support decisions related to treatment strategies for CML. Results indicate that the predictive model presented in this paper should be evaluated for potential clinical use. The early prediction of MMR at 24 months could be effective at reducing failure rate of TKI. Although the study analysed TIDEL II trial data, the methodology can be applied for different clinical trials.

Results of our study suggest that CML patients predicted not to achieve MMR at 24 months owing to IM could then be treated with alternative therapies, including the second generation TKIs NIL and DAS, or with more aggressive IM therapy, such as switching to NIL therapy and close monitoring. By contrast, CML patients predicted to achieve MMR could safely be treated with standard IM therapy with good clinical outcomes expected. The consequence of mistaken classification and subsequent treatment with IM is most likely to be treatment failure and higher risk of mortality.

OA with high G-mean accurately predicts both the positive group (a predictor for patients who will achieve MMR at 24 months) and the negative group (a predictor for patients who will not achieve MMR at 24 months). In addition, the high F-score for OA indicates it is a good predictor for the positive group (sensitivity) and can be trusted in clinical practice. However,

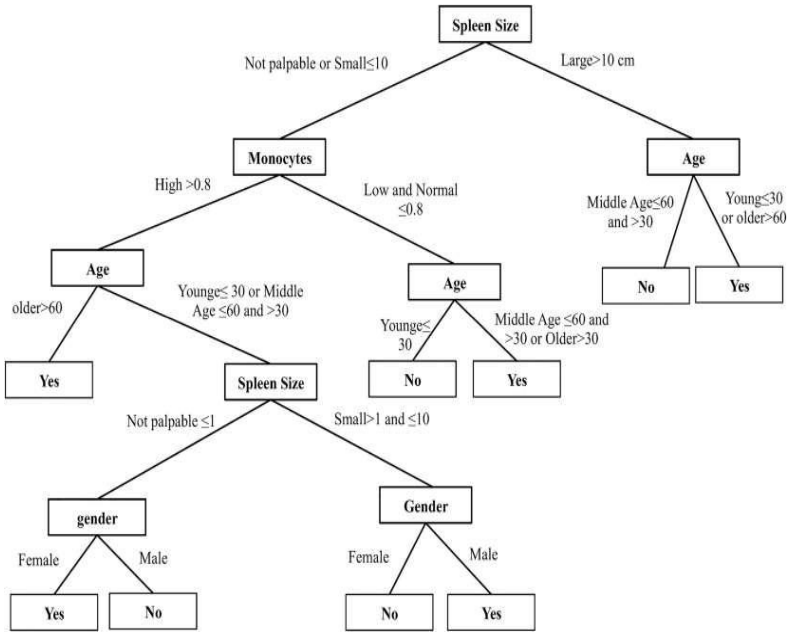


Fig 4. Model D structure. The final model in the tree graph that achieved high accuracy performance.

doi:10.1371/journal.pone.0168947.g004

Table 6. The comparison between previous methods and our recommended model on Saudi population.

	Validation Performance						
	Accuracy	Sensitivity	Specificity	PPV	NPV	Gmean	F-score
Sokal score[15]	0.63	0.83	0.13	0.71	0.24	0.33	0.17
Hasford Score[16]	**0.68**	0.92	0.06	0.71	0.25	0.24	0.1
EUTOS Score[17]	0.67	0.86	0.19	**0.73**	**0.35**	0.41	0.25
Model D	0.50	0.55	**0.35**	0.68	0.24	**0.44**	**0.29**

doi:10.1371/journal.pone.0168947.t006

the machine learning models often fitted the data better than previous methods (CART models fit the data better than OA). An important aspect of the decision tree is being informed by historical data and predicted unseen data so that data fits the model. However, in developing previous methods, researchers started with a model and checked whether data fit the proposed model and predictions were obtained by assuming that data are normally distributed or linearly associated. In conventional statistical methods, as data is collected by researchers to examine specific medical hypotheses or answer specific clinical questions, the approach is about a model fitting the data[44]. Thus, This predictive model was developed by addressing the challenge as a machine learning problem.

The results displayed in Tables 5 and 6 show that the small number of patients' predictive factors included in our study may lead to overfitting and may consequently affect their generalization ability on unseen examples. We observed that the G-mean and F-score performances were reduced from 74% and 76% in the internal validation of the testing data to 44% and 29% in the external validation of the Saudi data. In the present study, it is particularly interesting to notice the difference in performance between the prognostic scores and our model D that can be beneficial for the molecular response prediction, owing to the reduced risk of misclassified patients, who will not achieve MMR at 24 months. In addition, the difference in the prediction accuracy obtained from the nested cross validation of the CART was small (Table 4), which indicates that our models were not overfit. We pruned our trees based on a cost function in order to avoid overfitting in the maximal tree. Also, we have constructed predictive models to predict MMR positive and negative groups based on the pre-therapy predictive factors (i.e. the clinical, molecular, and blood count factors) in order to learn patterns as clinical prediction rules that are associated with a response to IM.

Indeed, the recommended decision tree model D was validated internally on TIDEL II clinical trial data, and externally on the Saudi dataset. The highest G-mean, and F-score values in the testing data (TIDEL II's G-mean: 74% and F-score: 76%) and the external validation data (Saudi population's G-mean: 44% and F-score: 29%) compared with the prognostic scores' G-mean and F-score confirmed that our models are good predictors for positive and negative groups, while the highest sensitivities in the testing data and the external validation data were observed in prognostic scores in comparison with our model D, confirming that prognostic scores were good predictors for a positive group. As the Hasford score achieved the highest sensitivity in TIDEL II and the Saudi population (sensitivity: 92%), and the Sokal and EUTOS scores achieved higher sensitivity than our model (TIDEL II: Sokal and EUTOS scores: 84%; model D: 67%), similarly, the sensitivity of Sokal and EUTOS were higher than our model D in the external validation data (Sokal score: 83%; EUTOS score: 86%; model D score: 55%).

Strengths and Limitations

The main strength of our study is the development of predictive models using domain knowledge in construction of clinical prediction rules which can be applied for validation on

different cohorts. We extracted hidden knowledge in the form of prediction rules. Another strength is interpretability and high prediction performance of the decision tree model developed using age, gender, spleen size, blasts, and monocytes. Also, the number of predictive factors used in the model is another advantage in the context of healthcare costs as fewer input factors imply fewer diagnostic tests to obtain relevant predictive factors [38]. From the wrapper approach, predictive factors determined to be relevant predictors of MMR at 24 months were age, gender, spleen size, eosinophils, blasts, monocytes, OA, *BCR-ABL1* transcript level, and *BCR-ABL1* transcript type. Indeed, considerable evidence supports the significant impact of age[15, 16], gender[23], OA[13], spleen size[15–17], blasts, and eosinophils [15, 16].

This study is the first to investigate relations between molecular predictive factors and MMR. The levels of *BCR-ABL1* transcript pre-therapy and *BCR-ABL1* transcript type as molecular tests play a role in investigating prediction of MMR achievement at 24 months. We found that molecular factors could significantly increase model performances (Models E and F), whereas Model D could be sufficient for prediction of MMR at 24 months. This study does not ignore the significance of all predictive factors involved in the analysis because different criteria (not maximising accuracy, G-mean, and F-score) may result in different predictive factors.

We reported the performance of the three common prognostic scores in the Saudi population (KFSHRC) and TIDEL II as part of the evaluation of model performance, but this appears to be the first study to report these results, helping overcome the lack of research in the area of comparing the performance of existing prognostic scores in both Saudi and Australian populations. Although the Australia group reported their Sokal score result from TIDELL II [25], it is best to obtain the most reliable of the three prognostic performance scores and compare it with the new development model. As our results demonstrate, the previous prognostic scores are good predictors of the positive group, but our models are good predictors for both the positive and negative groups.

The tree structures were significantly influenced as the majority of patients had the same outcome. A larger dataset may increase prediction accuracy. Additionally, varying expert opinions in representing category boundaries of each predictive factor could dramatically change performance of the models.

With the available options for CML treatment, development of a method to predict CML patient response to treatment at diagnosis is critically important. We can conclude that this predictive model assists managing CML treatment by predicting the likelihood that *de novo* CML patients will achieve MMR at 24 months when treated with IM. The decision tree prediction model presented here offers medical practitioners an additional tool to provide patients with improved, individualised treatment plans. For future work, this method may also be extended to other TKIs available for use as frontline CML therapy.

Supporting Information

S1 Text. Supplementary Results.
(DOCX)

S1 File. Decision Tree Structures.
(ZIP)

S1 Table. Clinical Prediction Rules, dataset includes list of clinical prediction rules that constructed from recommended models.
(XLSX)

Acknowledgments

The authors wish to thank the anonymous reviewers from Novartis and the associate editors for their critical comments and suggestions on a previous draft of this paper.

Author Contributions

Conceptualization: HB.

Data curation: HB DW TH NC.

Formal analysis: HB DR FB TK DA.

Investigation: HB DW TH NC.

Methodology: HB DR FB TK.

Project administration: HB.

Resources: HB DW TH NC.

Software: HB.

Validation: HB DR FB DA TL DW TH NC.

Visualization: HB DW TL TH.

Writing – original draft: HB.

Writing – review & editing: HB DR FB TK DA TL DW TH NC.

References

1. Apperley JF. Part I: Mechanisms of resistance to imatinib in chronic myeloid leukaemia. Oncology. 2007; 8:1018–29. doi: 10.1016/S1470-2045(07)70342-X PMID: 17976612
2. White DL, Hughes TP. Predicting the Response of CML Patients to Tyrosine Kinase Inhibitor Therapy. Current Hematologic Malignancy Reports 2009; 4:59–65. doi: 10.1007/s11899-009-0009-2 PMID: 20425416
3. Cortes J, Kantarjian H. How I treat newly diagnosed chronic phase CML. Blood. 2012; 120:1390–7. doi: 10.1182/blood-2012-03-378919 PMID: 22613793
4. Baccarani M, Saglio G, Goldman J, Hochhaus A, Simonsson B, Appelbaum F, et al. Evolving concepts in the management of chronic myeloid leukemia: recommendations from an expert panel on behalf of the European LeukemiaNet. Blood. 2006; 108(6):1809–20. doi: 10.1182/blood-2006-02-005686 PMID: 16709930
5. Hughes T, Branford S. Molecular monitoring of BCR–ABL as a guide to clinical management in chronic myeloid leukaemia. Blood. 2006; 20:29–41.
6. Baccarani M, Deininger MW, Rosti G, Hochhaus A, Soverini S, Apperley JF, et al. European LeukemiaNet recommendations for the management of chronic myeloid leukemia: 2013. 2013; 122(6):872–84. doi: 10.1182/blood-2013-05-501569 PMID: 23803709
7. Cruz JA, Wishart DS. Applications of Machine Learning in Cancer Prediction and Prognosis. Cancer Informatics 2006; 2:59–77.
8. Oldenhuis CNAM, Oosting SF, Gietema JA, de Vries EGE. Prognostic versus predictive value of biomarkers in oncology. European Journal of Cancer. 2008; 44:9 4 6–9 5 3.
9. Agrawal M, Garg RJ, Kantarjian H, Cortes J. Chronic Myeloid Leukemia in the Tyrosine Kinase Inhibitor Era: What Is the "Best" Therapy? Current oncology reports (1523–3790). 2010; 12(5):302–13.
10. White DL, Hughes TP. Predicting the response of CML patients to tyrosine kinase inhibitor therapy. Current Hematologic Malignancy Reports. 2011; 6(2):88–95. doi: 10.1007/s11899-011-0087-9 PMID: 21448598
11. Guilhot F, Guilhot J. Predicting response in CML. Blood. 2011; 117(6):1773–4. doi: 10.1182/blood-2010-11-317123 PMID: 21310930

12. Jabbour E, Kantarjian H, O'Brien S, Shan J, Garcia-Manero G, Wierda W, et al. Predictive factors for outcome and response in patients treated with second-generation tyrosine kinase inhibitors for chronic myeloid leukemia in chronic phase after imatinib failure. Blood. 2010; 117:822–1827.
13. White DL, Saunders VA, Frede A, Dang P, Zrim S, Osborn MP, et al. The Functional Activity of the OCT-1 Protein Is Predictive of Molecular Response and Survival in CP-CML Patients Treated with Imatinib: A 5 Year Update of the TIDEL Trial. Blood [0006–4971]. 2009; 114(22):507–.
14. White D, Saunders VA, Kalebic T, Hughes TP. The IC50 Assay Is Predictive of Molecular Response, and Indicative of Optimal Dose in De-Novo CML Patients. Blood [0006–4971]. 2008; 112(11):405 –.
15. Sokal J, Cox E, Baccarani M, Tura S. Prognostic discrimination in "good-risk" chronic granulocytic leukemia. Blood 1984; 63:789–99. PMID: 6584184
16. Hasford J, Pfirrmann M, Hehlmann Rd, Allan NC, Baccarani M. A New Prognostic Score for Survival of Patients With Chronic Myeloid Leukemia Treated With Interferon Alfa. Jornal of the National Cancer Institute. 1998; 90(11):850–8.
17. Hasford J, Baccarani M, Hoffmann V, Guilhot J, Saussele S, Rosti G. Predicting complete cytogenetic response and subsequent progression-free survival in 2060 patients with CML on imatinib treatment: the EUTOS score. Blood. 2011; 118(3):686–92. doi: 10.1182/blood-2010-12-319038 PMID: 21536864
18. White DL, Saunders VA, Dang P, Engler J, Zannettino ACW, Cambareri AC, et al. OCT-1−mediated influx is a key determinant of the intracellular uptake of imatinib but not nilotinib (AMN107): reduced OCT-1 activity is the cause of low in vitro sensitivity to imatinib. Blood. 2006; 108:697–704. doi: 10.1182/blood-2005-11-4687 PMID: 16597591
19. White DL, Saunders VA, Dang P, Engler J, Venables A, Zrim S. Most CML patients who have a suboptimal response to imatinib have low OCT-1 activity: higher doses of imatinib may overcome the negative impact of low OCT-1 activity. Blood. 2007; 110:4064–72. doi: 10.1182/blood-2007-06-093617 PMID: 17761829
20. Prejzner W. Relationship of the BCR gene breakpoint and the type of BCR/ABL transcript to clinical course, prognostic indexes and survival in patients with chronic myeloid leukemia. Medical science monitor (1234–1010). 2002; 8(5).
21. Verma D, Kantarjian HM, Jones D, Luthra R, Borthakur G, Verstovsek S, et al. Chronic myeloid leukemia (CML) with P190BCR-ABL: analysis of characteristics, outcomes, and prognostic significance. Blood (0006–4971). 2009; 114(11):2232–5.
22. Jabbour E, Cortes J, Nazha A. EUTOS score is not predictive for survival and outcome in patients with early chronic phase chronic myeloid leukemia treated with tyrosine kinase inhibitors: a single institution experience. Blood. 2012; 119:4524–6. doi: 10.1182/blood-2011-10-388967 PMID: 22431574
23. Branford S, Yeung DT, Ross DM, Prime JA, Field CR, Altamura HK, et al. Early molecular response and female sex strongly predict stable undetectable BCR-ABL1the criteria for imatinib discontinuation in patients with CML. Blood. 2013; 121:3818–24. doi: 10.1182/blood-2012-10-462291 PMID: 23515925
24. Sweet K, Zhang L, Pinilla-Ibarz J. Biomarkers for determining the prognosis in chronic myelogenous leukemia. Journal of Hematology & Oncology. 2013; 6(54).
25. Yeung DT, Osborn MP, White DL, Branford S, Braley J, Herschtal A, et al. TIDEL-II: first-line use of imatinib in CML with early switch to nilotinib for failure to achieve time-dependent molecular targets2015 2015-02-05 00:00:00. 915–23 p.
26. Ambler G, Omar RZ, Royston P. A comparison of imputation techniques for handling missing predictor values in a risk model with a binary outcome. Stat Methods Med Res. 2007; 16(3).
27. Acuna E, Rodriguez C. The treatment of missing values and its effect in the classifier accuracy. Classification, Clustering, and Data Mining Applications. Studies in Classification, Data Analysis, and Knowledge Organisation: Springer Berlin Heidelberg; 2004. p. 639–47.
28. Cismondia F, Fialho AS, Vieira SM, Retic SR, Sousab JMC, Finkelstein SN. Missing data in medical databases: Impute, delete or classify? Artificial Intelligence in Medicine. 2013; 58(1):63–72. doi: 10.1016/j.artmed.2013.01.003 PMID: 23428358
29. Starkweather J, Herrington R. Replace Missing Values. Secondary Replace Missing Values 2014. Available from: http://www.unt.edu/rss/class/Jon/SPSS_SC/Module6/SPSS_M6_1.htm.
30. Thamas D. CBE What the White Cell. IMVS Newsletter. 2007.
31. Freitas AA. Comprehensible classification models: a position paper. ACM SIGKDD Explorations. 2013; 15(1):1–10.
32. Mitchell T. Machine learning: McGraw-Hill; 1997.
33. Breiman L, Friedman J, Olshen R, Stone C. Classification and regression trees: Belmont, Calif: Wadsworth International Group; 1984.

34. Timofeev R. Classification and Regression Trees (CART) Theory and Applications: Humboldt University; 2004.

35. Quinlan JR, editor Generating production rules from decision trees. Proceeding IJCAI'87 Proceedings of the 10th international joint conference on Artificial intelligence; 1987; USA: Morgan Kaufmann Publishers Inc.

36. Perner P. Improving the accuracy of decision tree induction by feature preselection Applied Artificial Intelligence: An International Journal. 2001; 15(8).

37. Prati RC, editor Combining feature ranking algorithms through rank aggregation. Neural Networks (IJCNN), The 2012 International Joint Conference on; 2012 10–15 June 2012.

38. Qian M, Nahum-Shani I, Murphy SA. Dynamic Treatment Regimes. In: Tang W, Tu X, editors. Modern Clinical Trial Analysis. New York: Springer; 2013. p. 127–48.

39. Cheng T-H, Wei C-P, Tseng VS, editors. Feature Selection for Medical Data Mining: Comparisons of Expert Judgment and Automatic Approaches. Proceedings of the 19th IEEE Symposium on Computer-Based Medical Systems (CBMS'06); 2006: IEEE Computer Society.

40. Karegowda AG, Jayaram MA, Manjunath AS. Feature Subset Selection Problem using Wrapper Approach in Supervised Learning. International Journal of Computer Applications. 2010; 1(7):13–7.

41. Lasserre J, Arnold S, Vingron M, Reinke P, Hinrichs C. Predicting the outcome of renal transplantation2012 2012-03-01 00:00:00. 255–62 p.

42. Varma S, Simon R. Bias in error estimation when using cross-validation for model selection. BMC Bioinformatics. 2006; 7:91. doi: 10.1186/1471-2105-7-91 PMID: 16504092

43. Khalil SH, Abu-Amero KK, Al Mohareb F, Chaudhri NA. Molecular monitoring of response to imatinib (Glivec) in chronic myeloid leukemia patients: experience at a tertiary care hospital in Saudi Arabia. Genetic testing and molecular biomarkers. 2010; 14(1):67–74. Epub 2009/12/01. doi: 10.1089/gtmb.2009.0126 PMID: 19943786

44. Shouval R, Bondi O, Mishan H, Shimoni A, Unger R, Nagler A. Application of machine learning algorithms for clinical predictive modeling: a data-mining approach in SCT. Bone marrow transplantation. 2014; 49(3):332–7. doi: 10.1038/bmt.2013.146 PMID: 24096823

Linking Chapter 5 and 6:

In the previous article, “Modelling Predictors of Molecular Response to Frontline Imatinib for Patients with Chronic Myeloid Leukaemia,” we employed a widely used and practical machine learning technique to develop a predictive model to support decisions related to treatment strategies for CML. The results indicated that the treatments suggested in the paper could be provided for clinical use in the future. It was suggested that patients who were not expected to reach a stable MMR at 24 months could be considered prospects for alternative treatments. This study contributed a new knowledge-based predictive model to the area of individualized TKI treatments in CML patients. This work could assist clinicians in deciding to use frontline imatinib or frontline second-generation TKIs at diagnosis. The study analyzed the outcomes of CML patients enrolled in the TIDEL II clinical trial. The predictive factors selection approach, based on prior knowledge and a decision tree, was used to predict MMR and to select the relevant factors. The models were evaluated by nested cross-validation in terms of accuracy, G-mean, and F-score. Finally, the decision tree marked a significant advance because its predictive power is more accurate than the existing methods.

The next article, “The Implementation and Testing of a Personalized Medicine Support System for Chronic Myeloid Leukaemia,” discusses the personalized medicine support system’s design, implementation, and testing. The paper begins with the background of CML and treatment management and progresses to the current treatment selection issues at diagnosis. Then, the solution is implemented in the personalized medicine support system. The design demonstrates that the inputs of prognostic and predictive factors are fed into the inference engine. This inference performs a consistency test, presented in Chapter 4, Part 1; it also develops a combined prognostic model, as presented in Chapter 4, Part 2, and develops a predictive model as presented in Chapter 5. In addition, the knowledge-based

relationship is displayed both ways: receiving prior knowledge to develop a combined prognostic model and adding discovered knowledge from the predictive model structure and clinical predication rules. Finally, the inference generates treatment-specific advice and recommends if imatinib is a suitable TKI treatment. The study applies a multistage process in developing personalized medicine support system software. This software is available online, and the guide is presented in figures attached in the paper to display the graphic user interfaces for the tool. The personalized medicine support system is tested, and it has proved its ability in selecting the best TKI treatment, which may improve CML patient health care.

Chapter 6: The Implementation of a Testing Personalized Medicine Support System for Chronic Myeloid Leukaemia

Statement of Authorship

Title of Paper	The Implementation and Testing a Personalized Medicine Support System for Chronic Myeloid Leukaemia.
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Banjar, H. , Adelson, D. , "The Implementation and Testing a Personalized Medicine Support System for Chronic Myeloid Leukaemia".

Principal Author

Name of Principal Author (Candidate)	Haneen Reda Banjar		
Contribution to the Paper	HB designed and performed the research, analysed studies, and wrote the manuscript.		
Overall percentage (%)	90%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature	[redacted]	Date	2nd Dec 2017

The Implementation of Testing Personalized Medicine Support System for Chronic Myeloid Leukaemia

Haneen Banjar[1,2]

[1]School of Computer Science, University of Adelaide, Adelaide, South Australia, Australia.

[2]The Department of Computer Science, King AbdulAziz University, Jeddah, Saudi Arabia.

*Corresponding Author

E-mail: Hrbanjar@kau.edu.sa (HRB).

Abstract

After the second generation of Tyrosine Kinase Inhibitor therapy (TKI) was approved for use in frontline treatments, three TKIs became available for treating chronic myeloid leukaemia (CML) patients. A personalized medicine support system has been developed that can help healthcare organizations determine the right treatments for their CML patients. This study discusses the personalized medicine support system's design, implementation, and testing. This system is able to identify the risk group, resolve any conflicts in common prognostic scores outcomes, and predict CML patients' molecular responses to imatinib at 24 months at the time of their diagnoses. The personalized medicine support system is more effective in selecting the best TKI treatment than the single scoring system, which may improve CML patient healthcare. In the future, these findings in research will be validated on large CML patient cohorts.

Background

Due to the very intricate nature of CML, one standardized treatment cannot be recommended for most patients, as the disease often requires molecular interventions to produce positive results in a patient. This leads to the utilization of personalized medicine by health care professionals to treat cancers and diseases similar to CML. To better support medical professionals who agree with personalized medicine's concept and capabilities, the stratification of risk groups and also the prediction of molecular responses to TKIs can be used to provide information on how new treatments can be selected that will minimize the amount of treatment failures.

For many CML cases, the use of one treatment is not the best route to pursue. For example, three TKIs (imatinib, nilotinib, and dasatinib) have been approved for use in many countries to treat CML patients; imatinib is a highly effective TKI, but a prior study proved

that the number of mutations detected with imatinib was twice as high as the number of mutations identified in patients being treated with other TKIs [1]. It is better if CML patients receive more potent treatments because they appear to be more effective in reducing the number of mutations [2]. Indeed, the development of CML treatment management, rational decisions are required based on molecular predictive tests [3, 4]. Thus, the undesirable option is to use the same treatment in every case.

How Can Personalised Medicine Be Utilized in CML?

Personalized medicine can be defined as the tailoring of medical treatment to the individual characteristics of each patient [5]. With personalized medicine and treatments, doctors and other professionals use the patient's genetic profile to select a mode of treatment. This process uses molecular technology and machine learning to determine the best approaches to treating Leukemia patients, who may appear identical but actually possess differing molecular properties that can cause unpredictability in the response to prescriptions [6].

With personalised medicine, clinicians can offer more specific treatment plans for individuals who are combating CML. Personalised medicine can also be used to make treatment decisions that maximize outcomes and minimize side effects [7]. This is done with use of collected data that can then be studied and sorted to allow statisticians and health professionals to make various inferences about the best possible CML therapy. Personalised medicine involves classifying patients into subpopulations that pose unique reactions or responses to particular treatments to determine the best treatment approaches available for those individuals [8]. As stated by Mitchell [9], one result of this very detailed and meticulous process is that "unnecessary costs" can be minimized by reducing the time spent on treatment and also the failure rate during clinical trials, which will increase the economic value.

To help clinicians and other healthcare professionals discover the best individualized treatments for CML patients, the current mode of personalized treatment approaches can assist to design a more complete picture of what patients need from their care providers such as predictive assays, clinical, pathological, molecular based markers and prognostic scores. For example, prognostic scores: Sokal [10], Hasford [11], the European Treatment and Outcome Study (EUTOS) [12], and the EUTOS long-term survival score (ELTS) [13] help clinicians to stratify patients into population groups according to their risk profiles and therefore ensure that they receive the best treatment [14]. This is just an example of the methods and studies that are conducted in order to find the correct personalised medical treatments for CML patients as an alternative to a standard treatment for leukaemia-like cancers that may not benefit all patients who receive it.

The Problems with the Treatment in CML

The growing number of treatment options makes it difficult to narrow down the correct treatments that should be used for individual patients [2]. The first-generation TKI treatment, imatinib, is effective in that approximately 85% of CML patients respond positively [15]. However, some patients develop resistance to imatinib, which has led to the development of the second-generation TKI treatment that has already been approved for frontline use [16]. When further investigating imatinib, accurate risk stratification can potentially identify patients that can benefit from very aggressive therapies. These patients are important to identify because they must be distinguished from the overall majority who respond favourably to imatinib [17].

According to Hughes [17], bone marrow morphology and cytogenetics as well as a baseline real-time quantitative PCR (RQ-PCR) for BCR-ABL should be included in baseline assessment pre-therapy. This will help identify patients who may benefit from an increase in the imatinib dose, a switch to another TKI, and/or exploring if anallogeneic stem cell transplantation is an option. This is a prime example of how collecting data from an

individual's genetic makeup can help determine if a new intervention should be used for a patient's CML treatment.

A personalized medicine support system has been developed that can help healthcare organizations determine the right treatments for their CML patients [18]. This study discusses the personalized medicine support system's design, implementation, and testing. This system is aimed to identify the risk group, resolve any conflicts in common prognostic scores outcomes, and predict CML patients' molecular responses to imatinib at 24 months at the time of their diagnoses.

Personalized Medicine Support System for CML

The software development life cycle, includes seven multistage processes: i) requirement specification, ii) analysis, iii) design, iv) implementation, v) testing, vi) deployment, and vii) maintenance [19]. In the personalized medicine support system, the software multistage are demonstrated as follow:

Stage 1 Requirement Specification

CML patients need clear TKI treatment recommendations. Prognostic scores were used to stratify the risk groups. However, over time, a new prognostic score was validated in multicentre, single centre, and clinically diverse cohorts, which confirmed the ability to predict risk groups. If there are multiple scores, each prognostic score may generate a conflicting outcome for the same patient profile. In addition, no predictive model uses genetic or molecular indicators to predict CML patients' molecular response to imatinib.

Based on the above issues, our program must satisfy the following requirements:

- It must let the clinician enter the prognostic and predictive factors from which the treatment selection will be made;
- It must use the proposed analysis and final models published in [20-22] to stratify the risk group and predict the patients' molecular response to imatinib.

Stage 2 System Analyses

The output is a treatment-specific recommendation for imatinib or another TKI based on the patients' profile.

A new group of patients (conflict patients) whose profiles generated an outcome from a prognostic score that contradicted the outcomes from other prognostic scores [21] should be distinguished. Our findings [22] suggest that the integration of validated prognostic scores may enable better identification of high-risk patients who should be considered for more aggressive therapies.

Modelling predictors for CML patients' molecular responses to frontline imatinib were discovered in [21]. The published paper developed three predictive models using 15 clinical, biological, and molecular factors: i) the suggested predictive factors that showed the highest accuracy performance ii) the suggested predictive factors that showed the highest G-mean performance, and iii) the suggested predictive factors that showed the highest F-score performance. We let the clinician select the model that fit their patients' available data.

Stage 3 System Design

The system design details were published in [18]. After following the modelling steps to produce the framework, and the methods were published in [21, 22]. The personalized medicine support system includes six main components. Figure 1 illustrates these components and the published model's connections.

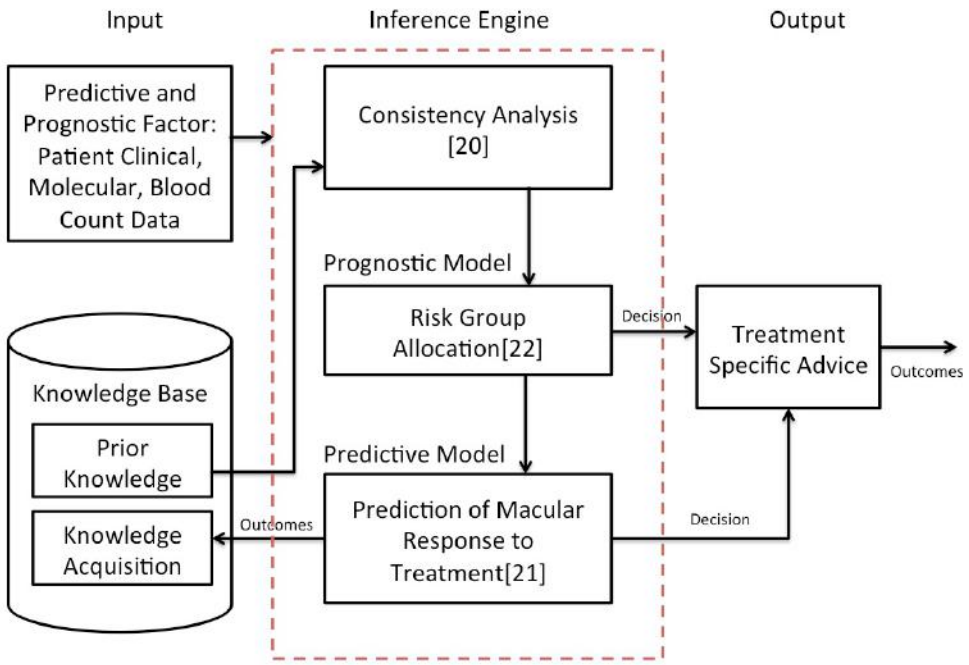


Figure 1 Conceptual framework of the personalized medicine support system

Objective and Scope

This system aims to support clinicians in their decisions when they select their CML patients' frontline TKI treatments. The system predicts the patients' risk class and molecular responses based on evidence of from clinical trials and patients admitted to hospitals. This system can overcome limitation to the current treatment guidelines for CML because it provides specific advice about whether to use frontline imatinib or another available TKI.

This model targets newly diagnosed adult patients with CML. The patients' profiles show their eligibility for any of the three TKI treatments. The patients should have completed molecular, clinical, and peripheral blood count tests at the time of their diagnoses. Other types of leukaemia are not the focus of this work. However, they may be considered in the boarder context of haematology services.

A successful system must:

- Outperform current prediction methods;

- Resolve treatment decision conflicts and stratify a risk group;
- Consider existing and newly discovered knowledge to aid in clinicians' final decisions.

The need for our system

We believe our work is novel and is a step forward in personalized medicine solutions and best practices in the prediction of CML patients' MMR to imatinib. Our study contributes a new knowledge-based predictive model for individualised TKI treatment in CML patients. This work will assist physicians in determining whether imatinib or a second-generation TKI should be used as a frontline treatment for CML patients at the time of their diagnoses. This system will do the following:

- Enable better identification of high-risk patients who should be considered for more aggressive therapies or allografting;
- Increase the effectiveness of CML treatment management. CML patients who will fail to respond to frontline imatinib therapy will benefit from this system because it predicts their response before they experience any resistance or intolerance to their therapies;
- Suggest accurate treatment, preventing the patient from paying unnecessary treatment costs;
- Prevent the prescription of the incorrect medication;
- Make accurate decision at diagnosis to help save patients' lives and time;

What is needed for this system

The system needs predictive and prognostic factors. Then, two options will be generated as treatment-specific recommendation. The system's input and the output are displayed in figure 2.

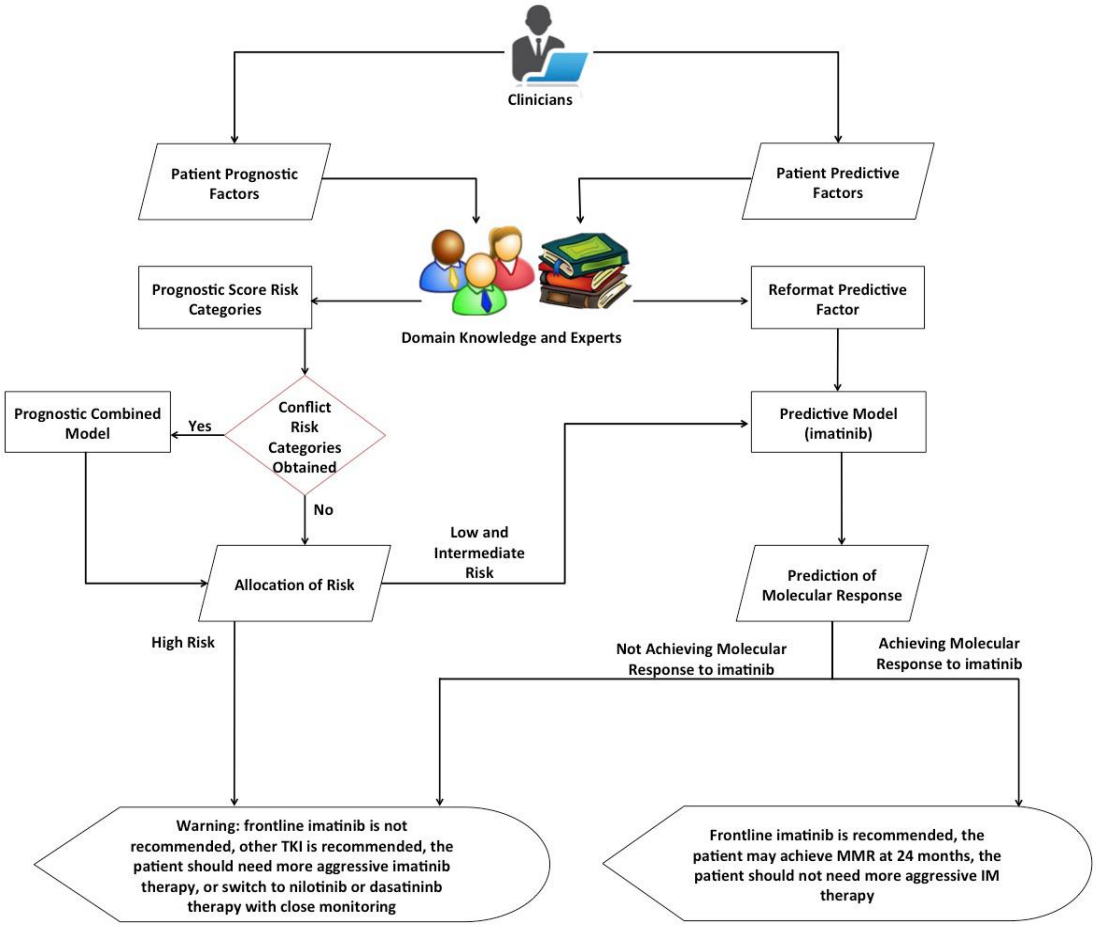


Figure 2 displays the flow chart of the final recommendation procedures from the personalised medicine support system. The clinicians insert the predictive and prognostic factors for the new CML patients, which are all collected prior to treatment. The system automatically categorised the predictive factors using prior knowledge and calculates the risk index using the existing formula published with the prognostic score systems. Then, the system checks the conflict between the generated risk indexes. If the conflict obtained from the common scores on the prognostic combined model outcome is created, the recommendation will be either 'Warning: frontline imatinib is not recommended; other TKIs are recommended. The patient should need a more aggressive imatinib therapy,' or 'Switch to nilotinib or dasatinib therapy with close monitoring.' Otherwise, the clinicians may test the low- and intermediate-risk groups in the second model (the predictive model). The categorised predictive factors will return the index of the selected value into the predictive factors, and the patients' molecular response to imatinib will be predicted. If the patient's outcome does not achieve the MMR at 24 months to imatinib, a warning recommendation will be displayed. However, if the patient's outcome from the predictive model shows a molecular response at 24 months to imatinib, the system will state, 'Frontline imatinib is recommended. The patient may achieve MMR at 24 months. The patient should not need more aggressive IM therapy.'

- i) TKI= Tyrosine Kinase Inhibitor Therapy, MMR=Major Molecular Response and IM=imatinib.

Predictive and prognostic factors

All the patients' clinical, molecular, and predictive assays and peripheral blood factors are collected at the time of their diagnoses. The prognostic model inputs include age (years), spleen size (cm), blast (%), platelets (10^9/L), eosinophils (%), and basophils (%).

While The prediction model inputs include the following: as follows: i) clinical factors: age, gender, and spleen size measured in centimetres below the costal margin; ii) molecular factors: *BCR-ABL1* transcript level pre-therapy and *BCR-ABL1* transcript type; iii) predictive assays: OA and $IC50^{IM}$; and iv) peripheral blood factors: white cell count (WCC), absolute neutrophil count (ANC), and eosinophil, basophil, monocyte, lymphocyte, platelet, and blast counts.

Options from the system

The output is either: i) 'frontline imatinib is recommended. The patient may achieve MMR at 24 months, and the patient should not need more aggressive imatinib therapy' or ii) 'warning: frontline imatinib is not recommended. Other TKIs are recommended, and the patient should need more aggressive imatinib therapy, or switch to nilotinib or dasatinib with close monitoring.'

Stage 4 Implementation

This section discusses two forms (the prognostic and predictive forms) [21, 22] that were written in C# with the Toolboxes from MATLAB R2016a (The MathWorks, Inc.). The software was implemented after studying the users' behaviour. Suitable error and warning messages with red star pointers where data is missing are provided if the form is used incorrectly.

Stage 5 Testing

From our publications [21, 22], we select 5 of 282 patients (173 Australian patients and 109 Saudi patients) to display the output from the models. Figures 3–6 show the print screens and the users' application guide.

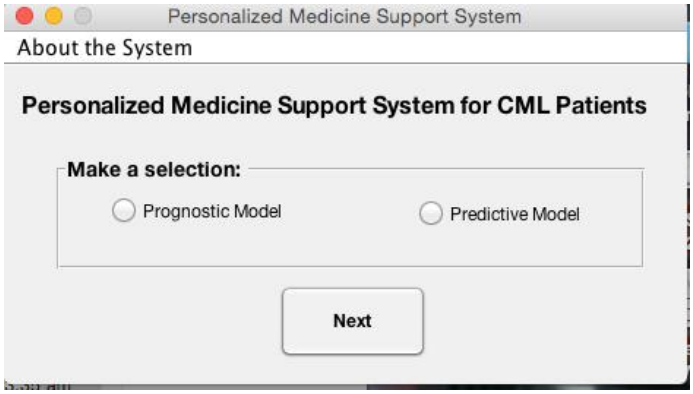


Figure 3 Make a selection of the models form. This form allow user to select either allocation of risk and resolve conflict using the prognostic model, or predict a molecular response using the predictive model.

The available forms will be listed in the menu bar, so the user can navigate between forms from there. If the user would like more information about the system, the 'about the system' link opens a form that displays the information and publications on the system.

From the menu, the user can click on 'Home' to return to the starter form, 'prognostic combined model,' 'predictive model,' and 'about the system.' The prognostic combined model form requests the pre-treatment prognostic values at the time of the patient's diagnosis, and the form calculates the common prognostic scores (figure 4). In the case of a conflict in the decision based on the scores, Sokal [10], Hasford [11], EUTOS [12], and ELTS [13], the combined model displays a warning message and allocates the patient to a risk group using the model published in [22] (see figure 5). The predictive model form requests the pre-treatment predictive factor values at diagnosis based on the model selection (figure 6). The models' predictive factors were published in [21]. The model will predict MMR and provide a suitable recommendation plan.

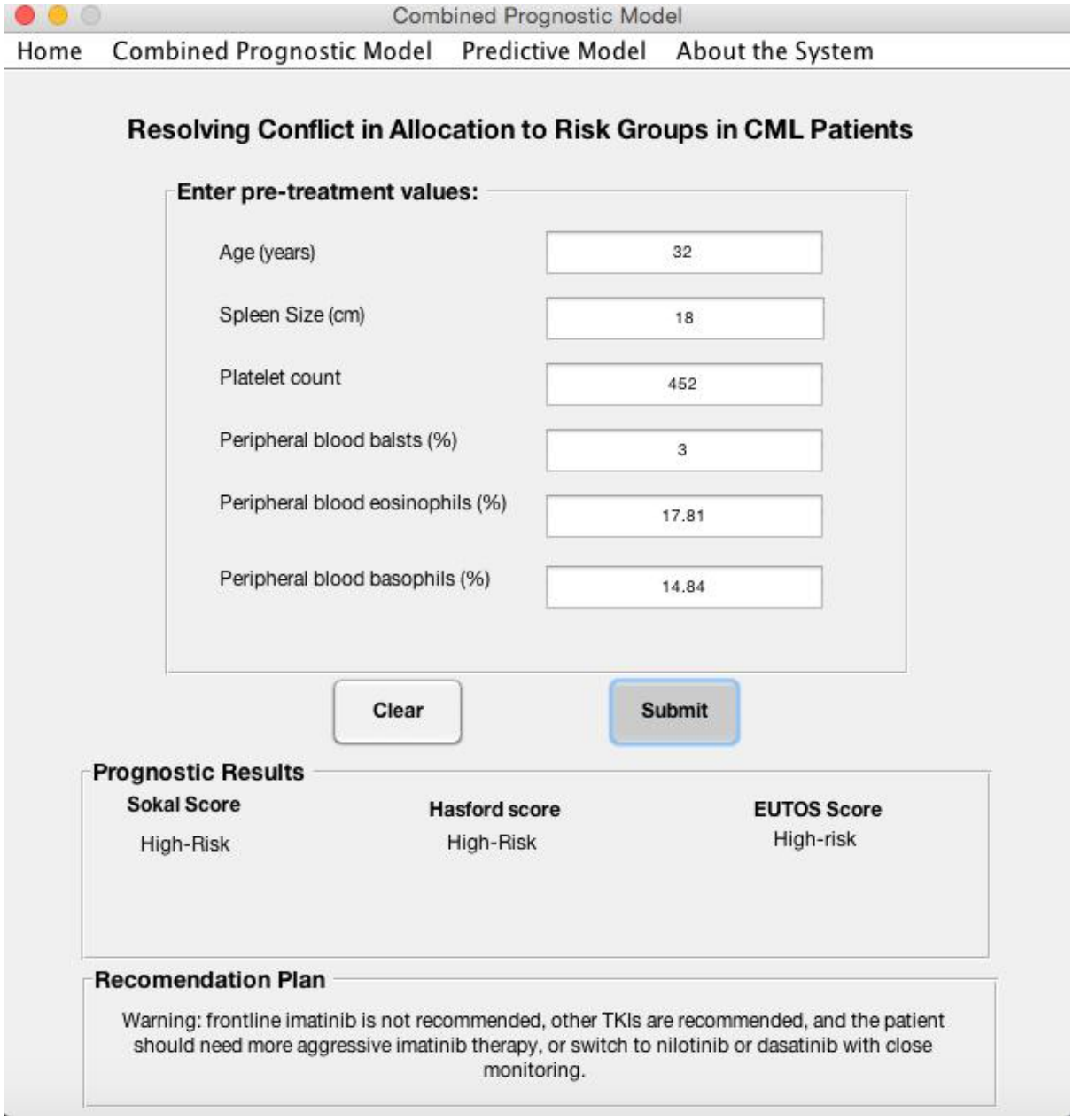


Figure 4 The prognostic combined model shows that the results of the three common scores are in agreement as well as a suitable treatment plans recommendation.

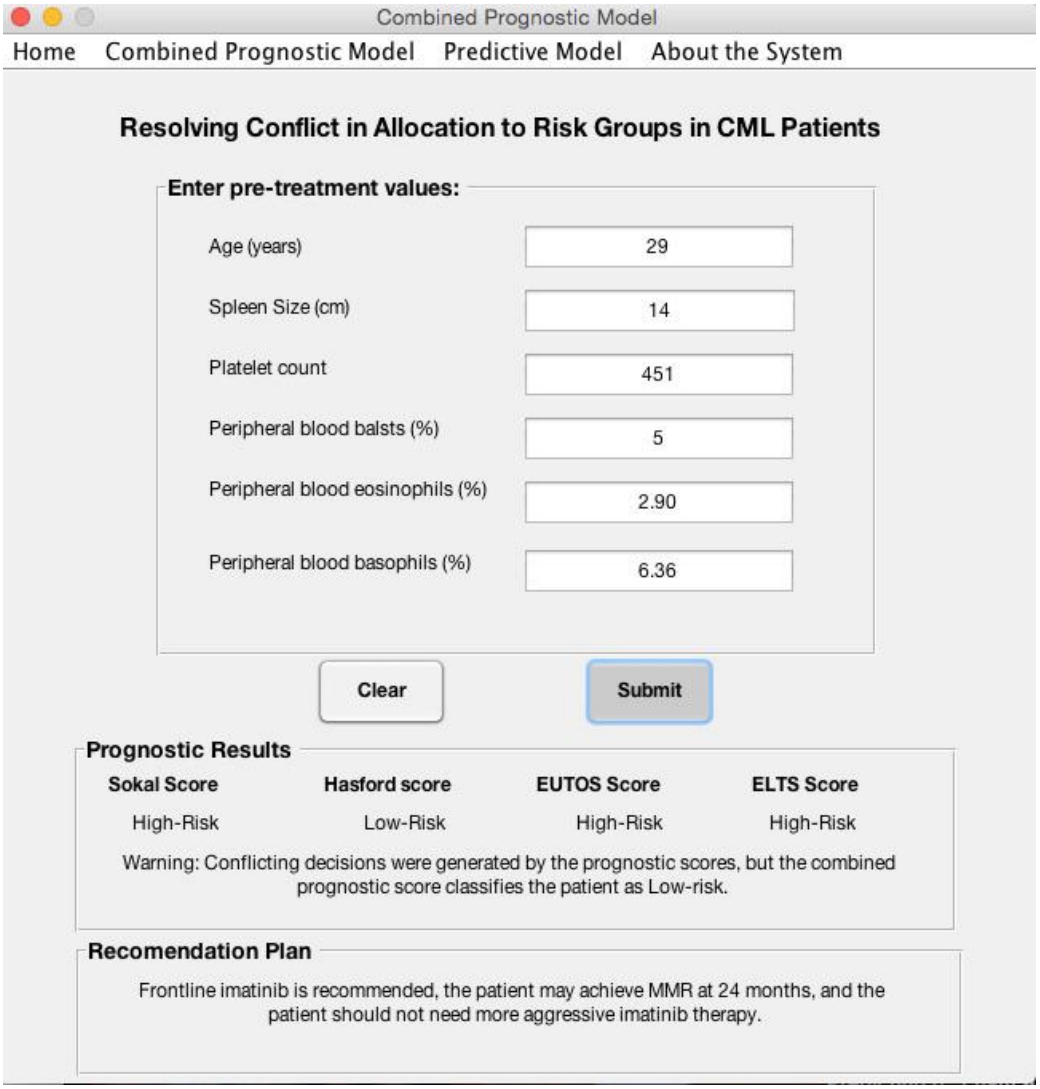


Figure 5 The prognostic combined model shows that the results of the four common scores generate conflicting decisions and provides a suitable treatment plan recommendation.

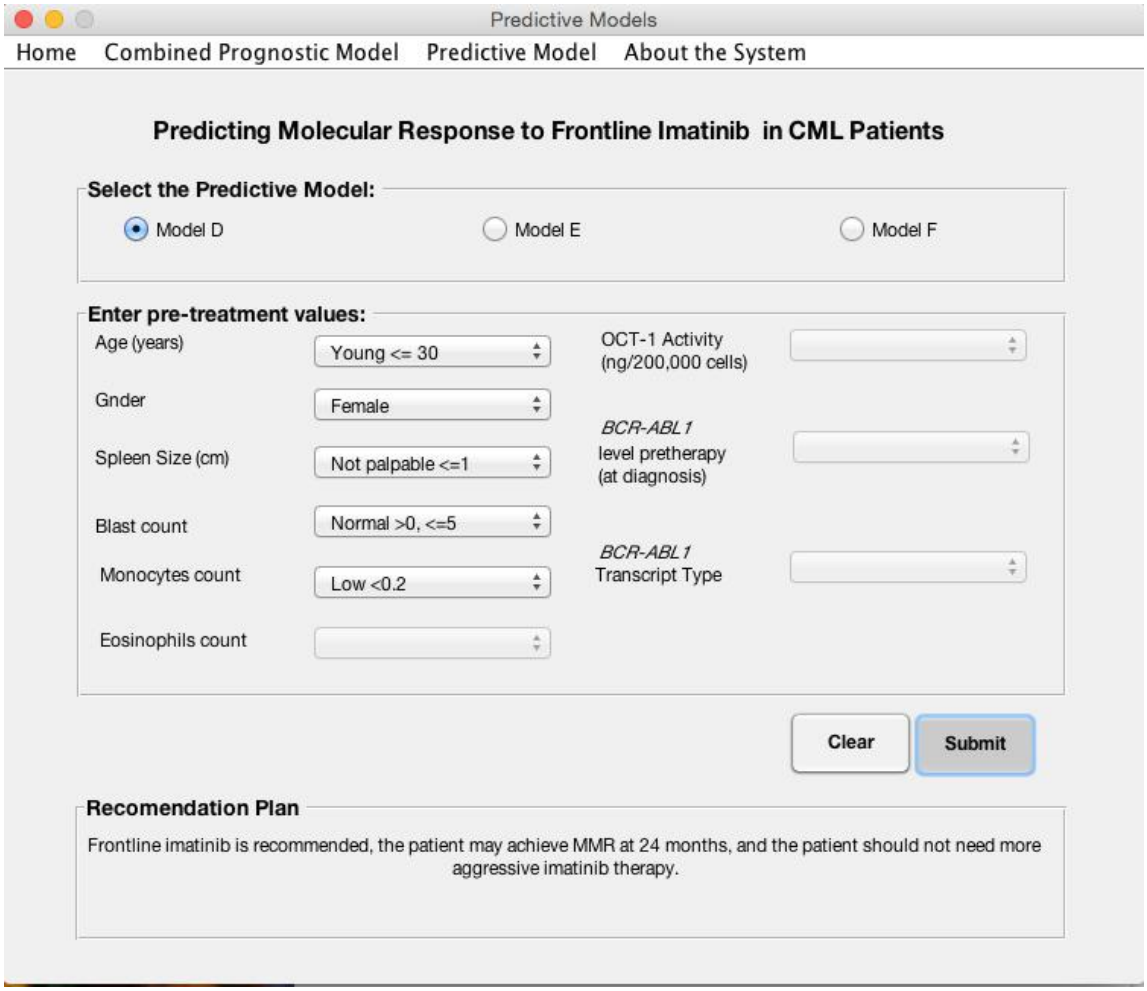


Figure 6 The selection of the predictive model will activate the related predictive factors, and the submit button displays the recommended plan based on the CML patients' predicted major molecular response at 24 months.

Stage 7 Deployment

The deployed application consists of a collection of MATLAB functions and data packaged as C# shared libraries. Any programming environment that supports C functions can compile these functions. The application can be installed on a machine without MATLAB, but the application requires a runtime library called MATLAB Compiler Runtimes (MCR) [23]. The package is available on (https://github.com/Hrbanjar/CML.git) or upon request from the author.

Stage 8 Maintenance

In [18], evolving approaches were suggested. Our software should continue to add and include new models and knowledge in accordance with the evolving environment. The author can fix newly discovered bugs and incorporate changes.

Discussion

This study provides evidence about the success of the design, implementation, and testing of the personalized medicine support system for CML patients. It guides decision makers as they select the best TKI therapy for CML patients. From [21], using the predictive model is sufficient to prescribe imatinib. However, it is important that conflicts generated by the existing score systems be handled, and if clinicians would like to identify TKI risk groups, they should use the combine prognostic model outcome [22]; the high-risk group should also be closely monitored and may need a combination of TKIs.

We recommend validating the use of the personalized medicine support system. The results of an ongoing clinical trial will provide additional information about the impact of our system on clinical decision making.

References

[1] Hochhaus A, Saglio G, Larson RA, Kim DW, Etienne G, et al. (2013) Nilotinib is associated with a reduced incidence of BCR-ABL mutations vs imatinib in patients with newly diagnosed chronic myeloid leukemia in chronic phase. Blood 121: 3703-3708.

[2] Hughes T and White D (2013) Which TKI? An embarrassment of riches for chronic myeloid leukemia patients.

[3] Baccarani M, Deininger MW, Rosti G, Hochhaus A, Soverini S, et al. (2013) European LeukemiaNet recommendations for the management of chronic myeloid leukemia: 2013. 122: 872-884.

[4] White DL and Hughes TP (2011) Predicting the response of CML patients to tyrosine kinase inhibitor therapy. Current Hematologic Malignancy Reports 6: 88-95.

[5] Müller M (2010) Individualized medicine. Clinical Pharmacology: Current Topics and Case Studies. Springer.

[6] Jain KK (2009) Basics of Personalized Medicine Textbook of Personalized Medicine. Springer New York.

[7] Kohane IS (2009) The twin questions of personalized medicine: who are you and whom do you most resemble? BioMed Central 1: 4.1-4.3.

[8] Verma M (2012) Personalized Medicine and Cancer. Journal of Personalized Medicine 2: 1-14.

[9] Mitchell DL (2013) Successful Implementation of Personalized Medicine: The Value, Challenges , and Effect on Patient Care. the Faculty of Utica College. ProQuest.

[10] Sokal J, Cox E, Baccarani M and Tura S (1984) Prognostic discrimination in "good-risk" chronic granulocytic leukemia. Blood 63: 789-799.

[11] Hasford J, Pfirrmann M, Hehlmann Rd, Allan NC and Baccarani M (1998) A New Prognostic Score for Survival of Patients With Chronic Myeloid Leukemia Treated With Interferon Alfa. Jornal of the National Cancer Institute 90: 850-858.

[12] Hasford J, Baccarani M, Hoffmann V, Guilhot J, Saussele S, et al. (2011) Predicting complete cytogenetic response and subsequent progression-free survival in 2060 patients with CML on imatinib treatment: the EUTOS score. Blood 118: 686-692.

[13] Pfirrmann M, Baccarani M, Saussele S, Guilhot J, Cervantes F, et al. (2016) Prognosis of long-term survival considering disease-specific death in patients with chronic myeloid leukemia. Leukemia 30: 48-56.

[14] Yahng S-A, Jang E-J, Choi SY, Oh Y-J, Bang J-H, et al. (2012) Comparison of Sokal, Hasford and EUTOS Scores in Terms of Long-Term Treatment Outcome According to the Risks in Each Prognostic Model: A Single Center Data Analyzed in 255 Early Chronic Phase Chronic Myeloid Leukemia Patients Treated with Frontline Imatinib Mesylate. Blood 120: 2794-2794.

[15] Wei G, Rafiyath S and Liu D (2010) First-line treatment for chronic myeloid leukemia: dasatinib, nilotinib, or imatinib. Journal of Hematology & Oncology 3: 47.

[16] Mauro M (2009) Tailoring Tyrosine Kinase Inhibitor Therapy in Chronic Myeloid Leukemia. Cancer control [1073-2748] 16: 108 -121.

[17] Hughes T and Branford S (2006) Molecular monitoring of BCR–ABL as a guide to clinical management in chronic myeloid leukaemia. Blood 20: 29-41.

[18] Banjar H, Adelson D, Brown F and Leclercq T (2017) Personalized Medicine Support System : Resolving Conflict in Allocation to Risk Groups and Predicting Patient Molecular Response to Targeted TherapyPersonalized Medicine Support System : Resolving Conflict in Allocation to Risk Groups and Predicting Patient Molecular Response to Targeted Therapy. Health Informatics - An International Journal (HIIJ) 6: 21.

[19] Liang YD (2015) Introduction to Java Programming, Comprehensive. Pearson Education.

[20] Banjar HR and Alsobhi E (2017) Consistency Test between Scoring Systems for Predicting Outcomes of Chronic Myeloid Leukemia in a Saudi Population Treated with Imatinib. International Scholarly Research Notices 2017: 6.

[21] Banjar H, Ranasinghe D, Brown F, Adelson D, Kroger T, et al. (2017) Modelling Predictors of Molecular Response to Frontline Imatinib for Patients with Chronic Myeloid Leukaemia. PloS one 12: e0168947.

[22] Banjar H, Hughes T, Brown F, Adelson D, White D, et al. (2017) Combined Value of Validated Prognostic Scores and Resolving Conflict in Allocation to Risk Groups in Chronic Myeloid Leukaemia Patients.

[23] Working with the MCR. In: R. b. Documentation, editor editors. Matlab Compiler. MathWorks.

Chapter 7: Conclusion

Summary

The development of TKI (imatinib, nilotinib, and dasatinib) therapy has improved molecular responses in CML patients. This is significant because achieving major molecular response (MMR) is the highest level of remission in CML patients. This work aimed to support clinicians in choosing the most appropriate treatment from among the following current targeted strategies: i) frontline imatinib or ii) frontline second-generation tyrosine kinase inhibitors (TKIs), that is, nilotinib or dasatinib, which are both used to reach molecular targets in the treatment of CML. This innovative work constitutes a step forward from recently implemented personalized medicine solutions and provides best practices for allocating risk groups and for predicting MMR to TKI in CML patients.

This study addresses issues that are crucial in selecting the most effective treatment from the TKI therapies available for individual CML patients. It also presents a framework for a personalized medicine support system that can be used to correctly allocate risk groups and predict patients' molecular responses to imatinib as a standard TKI treatment. High-risk patients who fail to respond to frontline imatinib will benefit from this system, which predicts molecular response before the patient faces any resistance or intolerance to therapy.

First, a review of the published empirical research on personalized medicine in leukemia was conducted; this was followed by synthesizing the findings of a number of previous studies related to intelligence techniques to detect and treat leukemia. A systematic search was carried out to identify leukemia studies using intelligence techniques and to categorize these studies based on leukemia type as well as on the task, data source, and purpose of the studies; leading to particular attention on chronic myeloid leukemia (CML).

A modeling approach based on data-mining phases was then developed for both prognostic and predictive model designs. A framework was designed for a personalized

medicine support system that included information resulting from prognostic factors, prognostic scores, and predictive factors, such as molecular and clinical data, predictive assays, and blood cell counts.

We found that the domain knowledge as well as the data was essential to understanding the problem. In particular, the processes for collecting medical background details and data were investigated, and domain experts were involved to confirm the exclusion and inclusion criteria. Conducting a systemized review was fundamental to understanding the medical issues associated with leukemia; the review assisted with selecting the leukemia subtype and with defining the predictive factors that were used in this system. During all processes, verification by experts was required for the combined prognostic model and predictive model to construct the final models. Thus, knowledge-based systems were an applicable choice for designing a personalized medicine support system.

This study demonstrates the validity of combined methods for coping with conflicts in prior knowledge. An early determination of risk categories could provide a warning at the diagnosis stage, where certain medications and preventive action could increase the span of a patient's healthy life. A study was carried out on the combined values of previously validated and published prognostic scores, which were then used to identify risk groups. The combined prognostic models could assist in determining the most appropriate therapy for individual patients. For low-risk patients, clinicians prefer therapy that has proved to be the least toxic and safest option. In higher risk patients, they prefer combination approaches or second-generation molecular targeted therapy, because the higher toxicity and higher risk of organ damage are counterbalanced by greater potency and a lower propensity for drug resistance [1]. Chapter 4 introduced new models for facilitating and resolving conflicts, which were obtained by providing scoring systems with accurate risk-assessment methods for use in diagnosis.

In addition, a framework was designed to adapt the system to accommodate newly emerging knowledge. The framework addresses changes in care delivery and the need to keep up with rapid medical scientific discoveries. Considerable care was required to design a rapid-learning health system based on the framework and thereby enable it to take account of recent developments in health information technology, to access health data, and, finally, to apply the evidence effectively and thus make the right decision. Our design was able to automatically refine the entire structure of the final model while adding new information.

Later, we found that the predictive model was required to the model predictors of molecular response to TKI**.** Chapter 5 contributed new predictive models to the area of individualized TKI treatment for CML patients. The outcomes of CML patients who were enrolled in the TIDEL II clinical trial were analyzed. The predictive-factors selection approach based on prior knowledge and using a decision tree, was employed to predict MMR, to select relevant factors, and to generate interpretable clinical prediction rules. Nested cross-validation was used to evaluate the models in terms of accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), G-mean, and F-score. Finally, the recommended decision tree marks a significant advance, because its predictive power was stronger than existing methods for both positive and negative groups. However, the prognostic methods achieved a good performance in predicting the positive group.

The models have been reported to achieve high performance with fewer predictive factors and to outperform existing methods. The interpretability of these models and the selection of relevant predictive factors are based on user comprehension in the context of expert knowledge and on their ease of use in practice. An important part of predictive modeling is to obtain accurate relative predictive factors and predictive assays. How these factors act jointly in the prediction of MMR is also important. The relationships between the

predictive factors and MMR were presented in the form of clinical prediction rules that offer promising opportunities in personalized treatment research.

The development of new models must ensure that the model is appropriately applied, correctly developed, and rigorously evaluated. To increase the reliability of the results, the models trained, validated, and tested on real CML data were representative examples for evaluating the personalized medicine support system. Model performance was internally validated via the use of nested cross-validation methods. Those models showing the most promising performance were selected for independent validation to evaluate them meaningfully with respect to the field of application.

Finally, the implementation of the personalized medicine support system served as a tool for use in assisting clinicians to test the model on their patients' data.

Research Strengths

The benefits of using a personalized medicine support system include reductions in treatment costs, higher treatment success rates, and safe health care [2]. Indeed, Australia's national agency for health and welfare statistics and information estimated that leukemia treatment has a high cost of approximately over $50,000 per patient [3]. However, predicting the most appropriate treatment for patients offers less waste in terms of using ineffective and costly therapies. As well as ensuring financial savings, this approach will reduce undesirable adverse effects for leukemia patients. This is because the treatment plan would be designed according to the molecular profile of each patient and to how that patient would respond to certain medications.

An interesting finding from this study is that predictive and prognostic methods typically outperform the single prognostic factor, prognostic score system, predictive factor, or predictive assays. This suggests that a personalized medicine support system can enhance treatment decisions. This study could be involved in the CML treatment guidelines of the

following organizations: i) National Comprehensive Cancer Network (NCCN) [4], ii) European Society of Medical Oncology (ESMO) [5], and iii) European LeukemiaNet (ELN) [6]. The three guidelines recommended for determining the risk groups using either Sokal [7], Hasford [8], or EUTOS [9] scores and three TKIs (imatinib, nilotinib, and dasatinib) are available for use. However, the framework used in this study would improve the guidelines when the combined prognostic model determines the risk level and handles the conflict in prognostic score outcomes that may be obtained from the same patient profile. The predictive model for primary treatment then predicts the molecular response and suggests the best recommendations for supporting clinicians in diagnosis. It is expected that improvements in CML management and the ability to select the most suitable frontline TKI will make treatment decisions more accurate in the near future.

Research Limitations

- Impact of diversity in human expertise: The models are derived from human knowledge held by domain experts, from literature, and from other knowledge that was computationally discovered. Thus, the final models are limited to human knowledge, which could be incomplete or biased.
- Impact of changing factors, population, and prediction endpoints: Modeling predictors used only those predictive factors available in TIDEL II clinical trials; depending on the absence or presence of predictive factors, clinical prediction rules may change. The models also resulted from the use of a potential population. In addition, clinical prediction rules would be affected if different prediction end-points from clinical trials were used.
- Impact of time: In developing these models, analysis is a time-consuming process if using various predictive factors. The personalized medicine support system framework requires high-performance computation to

provide the results within a short period of time. This is particularly important during the analysis phase, as possible clinical misinterpretation of results during analysis development phases may result in the need to repeat the analysis process.

- Impact of patient personal issues: Molecular testing should be performed for CML patients planning to undergo imatinib treatment and to benefit from the personalized medicine support system. Regardless of the nature of the health insurance system, if the patients are unable to participate in the assays or molecular test pre-therapy for financial reasons, they are also unable to enjoy the benefits of this technology.

Future Work

Molecular techniques now play a pivotal role in monitoring several types of disease and in determining suitable therapeutic strategies. Many researchers and scholars have dedicated time and effort to investigate biomarkers and predictive factors in recommending the best treatment plan for patients. The availability of data sets containing prognostic or predictive factors and the molecular response data for pretreatment with molecular targeted therapy has facilitated the development of a personalized medicine support system for predicting the molecular response to molecular targeted therapy. There are currently over 4,900 diseases and 66,880 tests available in the GENETests medical genetics information resource [10]. A personalized medicine support system for identification of risk groups and for the prediction of molecular response to molecular targeted therapy is a solution for the frequent updates in molecular sources, such as new molecular tests and drugs, as well as the predictive or prognostic factors that are continually being discovered, which continue to expand the available medical knowledge. Research on how to predict the molecular response of individual patients treated with molecular targeted therapy has revealed several techniques that can be used by researchers in various ways across the methodology.

The framework presented needs to be developed further as follows:

- Development of a new predictive model: Model induction and the generation of clinical prediction rules can be developed by using a wide variety of techniques to construct rules derived from the machine-learning research area. One of the most interesting applications of the technological aspects of development in medicine is the use of soft computing techniques, such as the Adaptive Neuro Fuzzy Inference System (ANFIS) in predicting molecular response. ANFIS is powerful when integrating fuzzy logic; the rule base could be generated by a genetic algorithm, or by adapting the neural network to reason according to optimal clinical prediction rules. Clustering techniques could also be used to discover the best form for the membership function of linguistic categories instead of using domain knowledge to reformat the data. Therefore, the use of soft computing techniques offers a range of efficient knowledge-adaption methods.
- The potential relevance of the selected predictive factors depending on the domain knowledge and wrapper approach: The potential relevance of the selected predictive factors may be conducted using different statistical methods to optimize selection.
- Supporting validation: Validation can be expanded by collecting high-quality data from other health organizations.
- Strengthening our model development by encouraging coordination from other leukemia research units.
- Obtaining data to develop models to predict response to other TKIs: Data are available from clinical trials that have compared frontline imatinib with nilotinib [11, 12] or dasatinib with imatinib [13, 14], or from clinical trials that have reported the efficacy and safety of imatinib [15]. The personalized medicine support system can

be enhanced by adding more data from the clinical trials to predict the molecular response to each of the TKIs.

Finally, further intensive research activities based on the medical domain issues can be conducted to determine the effectiveness of using a wide variety of intelligent techniques in the treatment of individual cancer patients.

Conclusion

In this thesis, we addressed the problem of TKI therapy treatment selection by correctly allocating patients to risk groups and predicting their molecular response to the selected treatment. The main contributions of our work include expressing this task as a classification problem and proposing a framework for use in solving it based on a knowledge-based decision support system. A discussion of different studies on individual models is provided; in particular, we propose developments of both the combined prognostic model and the predicative model.

Two additional contributions are notable: an experimental comparison of our models' behavior has been performed automatically, and the design of a new structure for a personalized medicine support system based on the available medical knowledge.

The main focus of our thesis was on the accurate classification itself. A new approach based on a medical application problem was introduced for solving issues related to a clinicians' diagnosis decisions. The foundations of this approach rely on the implementation of a knowledge-based system to provide clinicians, patients, and researchers with a platform to allocate risk groups and resolve conflicts in prognostic scores, to predict molecular response to the standard TKI imatinib, and to extract the relationship between the predictive factors and achieving a molecular response. The system would increase the possibility of personalized CML treatment in CML. During the experiments, our combined prognostic model automatically suggested the best combined method performance of validated

prognostic scores and enabled better identification of high-risk patients who should be considered for more aggressive therapies or allografting. The auto-comparison between the models assisted in adapting the incoming information and in updating the final model selection based on performance. Additionally, the selected predictive models outperformed the current methods in term of G-mean and F-score and are therefore good predictors for patients who respond to treatment and for those who do not respond to treatment. The G-mean balances the sensitivity and the specificity, which is necessary, both clinically and economically, to avoid prescribing imatinib to CML patients who are less likely to achieve MMR at 24 months of imatinib therapy. Using the personalized medicine support system for selecting a TKI rather than current methods would result in more effective CML treatment and improved health care for CML patients.

In the second part of the study, we used a framework that has not been considered in previous personalized medicine research; moreover, the existing literature contains no research describing a process that was specifically applied to the problem in the chronic myeloid leukemia treatment field and that, to our knowledge, had never been addressed before. This contribution can certainly assist clinicians in treatment selection problems, thereby enlarging the potential application field of personalized medicine. Another contribution involves conflict resolution in patient outcomes. Up-to-date consistency testing and intelligent techniques have not been applied to CML treatment guidelines. This contribution enables the correct allocation of risk groups without conflict prognosis outcomes. Finally, the present study findings reveal that because patients' various genetic and molecular make-ups react differently to imatinib, its use is not a catch-all solution for CML patients. Molecular factors are important in this subject area, primarily because polymerase chain reactions can measure genetic transcripts in the blood of CML patients. When monitoring these reactions, clinicians use this data to predict the best long-term plans

for their patients. By utilizing these methods, a CML patient is more likely to receive the treatment that is genetically recommended for his or her illness and thus is most effective.

New opportunities will arise for further applications of a personalized medicine support system in different cancer domains, especially in the diagnosis of cancer patients treated with multiple molecular targeted therapy options. This trend will be accompanied by a higher demand for researchers who can design, adapt, and evaluate this framework and for clinicians to interpret results. Medical decision making in selecting the best treatment option for cancer patients is a complex process in which many predictive and prognostic factors should be taken into account for the prior selection of molecular targeted therapy; predictive and prognosis models for use in diagnosis may increase the effectiveness of the decision. Medical decision makers should optimize the patient's potential for a successful treatment plan and outcome, with assistance from the personalized decision support system.

References

1. Wei G, Rafiyath S, Liu D. First-line treatment for chronic myeloid leukemia: dasatinib, nilotinib, or imatinib. Journal of Hematology & Oncology. 2010;3(1):47. doi: 10.1186/1756-8722-3-47.
2. Balas E, Boren S. Managing clinical knowledge for health care improvement: Patient-Centered Systems; 2000.
3. Goss J. Leukaemia the most expensive cancer per case: Australia's national agency for health and welfare statistics and information 2005 [cited 2014 12 Dec]. 1-7]. Available from: http://www.aihw.gov.au/media-release-detail/?id=6442464580.
4. NCCN Clinical Practice Guidelines in Oncology: Chronic Myelogenous Leukemia: National Comprehensive Cancer Network; 2017 [2017]. Available from: http://www.nccn.org/professionals/physician_gls/pdf/cml.pdf.
5. Baccarani M, Pileri S, Steegmann JL, Muller M, Soverini S, Dreyling M. Chronic myeloid leukemia: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up†. Annals of Oncology. 2012;23(suppl_7):vii72-vii7. doi: 10.1093/annonc/mds228.
6. Baccarani M, Deininger MW, Rosti G, Hochhaus A, Soverini S, Apperley JF, et al. European LeukemiaNet recommendations for the management of chronic myeloid leukemia: 2013. 2013;122(6):872-84. doi: 10.1182/blood-2013-05-501569.
7. Sokal J, Cox E, Baccarani M, Tura S. Prognostic discrimination in "good-risk" chronic granulocytic leukemia. Blood 1984;63:789-99.
8. Hasford J, Pfirrmann M, Hehlmann Rd, Allan NC, Baccarani M. A New Prognostic Score for Survival of Patients With Chronic Myeloid Leukemia Treated With Interferon Alfa. Jornal of the National Cancer Institute. 1998;90(11):850-8.
9. Hasford J, Baccarani M, Hoffmann V, Guilhot J, Saussele S, Rosti G. Predicting complete cytogenetic response and subsequent progression-free survival in 2060 patients with CML on imatinib treatment: the EUTOS score. Blood. 2011;118(3):686-92. doi: 10.1182/blood-2010-12-319038.
10. Genetests medical genetics information resource (database online): University of Washington; [cited 2016]. Available from: https://www.genetests.org.
11. Kantarjian HM, Hochhaus A, Saglio G, De Souza C, Flinn IW, Stenke L, et al. Nilotinib versus imatinib for the treatment of patients with newly diagnosed chronic phase, Philadelphia chromosome-positive, chronic myeloid leukaemia: 24-month minimum follow-up of the phase 3 randomised ENESTnd trial. The Lancet Oncology. 12(9):841-51. doi: 10.1016/S1470-2045(11)70201-7.
12. Martinelli G, Castagnetti F, Poerio A, Breccia M, Palandri F. Molecular responses with nilotinib 800 mg daily as first-line treatment of chronic myeloid leukemia in chronic phase: Results of a Phase II trial of the gimema cml wp. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2009;27:7074.
13. Kantarjian H, Shah NP, Hochhaus A, Cortes J, Shah S, Ayala M, et al. Dasatinib versus imatinib in newly diagnosed chronic-phase chronic myeloid leukemia. The New England Journal of Medicine. 2010;362(24):2260–70.
14. Borthakur G, Kantarjian HM, O'Brien SM, Jones D, Koller C, Nicaise C, et al. Efficacy of dasatinib in patients (pts) with previously untreated chronic myelogenous leukemia (CML) in early chronic phase (CML-CP). Journal of Clinical Oncology. 2009;26(15S).
15. Deininger M, O'Brien SG, Guilhot F, Goldman JM, Hochhaus A, Hughes TP, et al. International Randomized Study of Interferon Vs STI571 (IRIS) 8-Year Follow up: Sustained Survival and Low Risk for Progression or Events in Patients with Newly Diagnosed Chronic Myeloid Leukemia in Chronic Phase (CML-CP) Treated with Imatinib. Blood (ASH Annual Meeting Abstracts). 2009;114(1126).

Appendix

Supplementary Material A: Combining Validated Prognostic Scores and Resolving Conflict in Allocation to Risk Groups in Chronic Myeloid Leukaemia Patients

Haneen Banjar, Timothy Hughes, Fred Brown, David Adelson, Deborah White, Enaam Alsobhi, Naeem Chaudhri

S1. Patient Data Exploratory

To display the distribution of pairs of scores, we represented examples in Figures A-C of the inconsistencies between portions of patients. The scores were normalized [0-1] to display risk boundaries. We found a correlation between the Sokal and Hasford scores of 0.74, while the correlation between the Sokal and EUTOS scores was 0.32 and the correlation between the Sokal and ELTS scores was 0.47. These results show linear correlation between Sokal and Hasford but nonlinear relationship between Sokal and the other scores. Regarding the Hasford scores, the correlation was 0.64 and 0.77 with EUTOS and ELTS respectively. These results showed the linear relationship between Hasford, EUTOS and ELTS scores. The correlation between EUTOS and ELTS was 0.65 that showed the linear relationship.

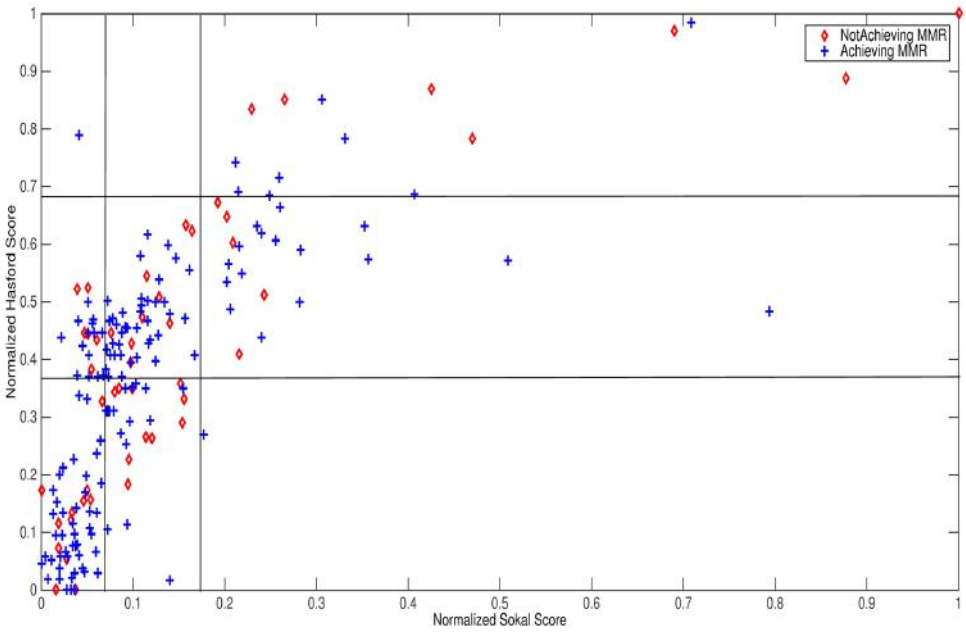


Figure A. Agreement between Sokal and Hasford scores. (a) Distribution of the Sokal and Hasford scores for achieving and not achieving MMR at 24 months. (b) The vertical lines show the boundaries for low,

intermediate and high risk groups for the Hasford score. (c) The horizontal lines demonstrate the boundaries for low, intermediate and high risk groups for the Sokal scores.

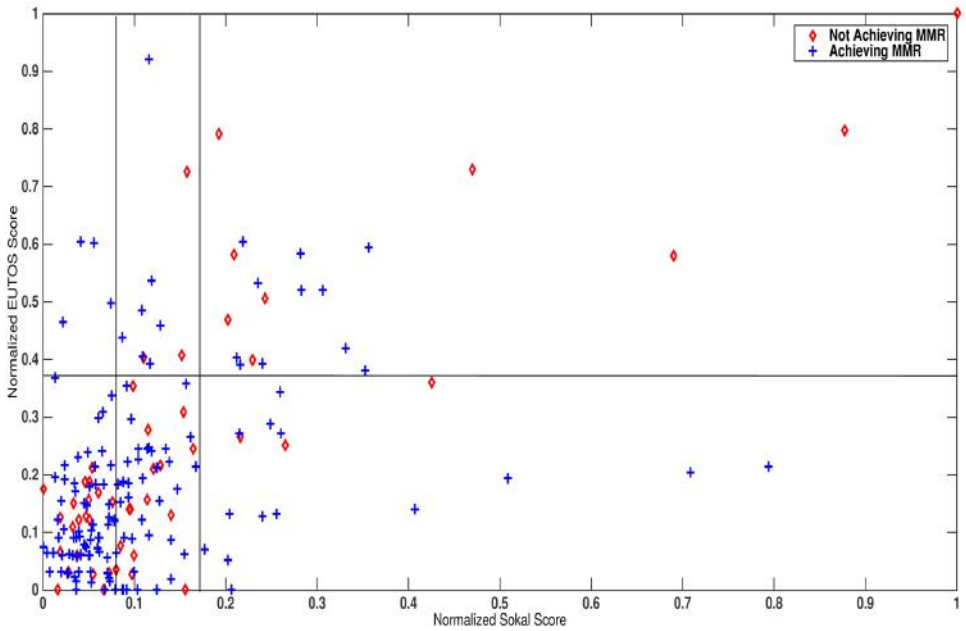


Figure B. Agreement between Sokal and EUTOS scores. (a) Distribution of the Sokal and EUTOS scores for achieving and not achieving MMR at 24 months. (b) The vertical lines show the boundaries for low and high risk groups for the EUTOS scores. (c) The horizontal lines demonstrate the boundaries for low, intermediate and high risk groups for the Sokal scores.

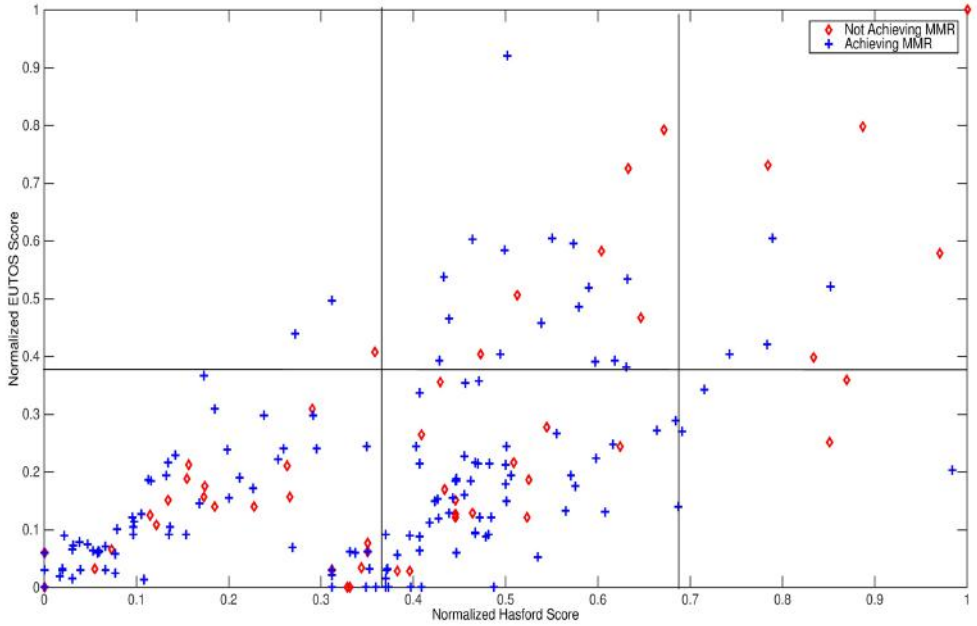


Figure C. Agreement between Hasford and EUTOS scores. (a) Distribution of Hasford and EUTOS scores for achieving and not achieving MMR at 24 months. (b) The vertical lines show the boundaries for low and high risk groups for the EUTOS scores. (c) The horizontal lines demonstrate the boundaries for low, intermediate and high risk groups for the Hasford scores.

S2. Inclusion and exclusion criteria

The study includes CML patients from Australia and Saudi population: 210 patients

from the TIDEL II [1], 172 patients from the KFSHRC [2] and 104 patients from the KAMC [3]. We apply the following criteria as shown in table A. The final dataset compromised 403 patients.

Table A. The criteria and the number of patients for selecting the study population.

Inclusion Criteria	**No. of Patients**
Patients received imatinib 600 mg/ 800 mg as the initial therapy.	**403**
Patients were in chronic phase	**403**
Patients monitored for achieving a BCR-ABL1 transcript level ≤ 0.1% at 24 months using RQ-PCR.	**403**
Patients who were switched to second line treatment because of intolerance or toxicity to imatinib were considered as a negative group in our analysis.	**38**
Exclusion Criteria	
Pregnant patients	**3**
Patient used the second line treatment (nilotinib or dasatinib) as a frontline treatment.	**73**
Patients were in advance phases.	**10**

S3. Base Classifiers

S3.1 Support Vector Machine

Support vector machine finds the best hyper plane that separates the patients of one outcome from those of the other outcome. The problem is to find a hyper plane that maximizes the margin between data points in the binary classification. The margin is the maximal width of the paralleled hyper plane that has no interior data points [4]. We tried linear, quadratic, cubic, fine, medium and coarse Gaussian kernel function.

S3.2 k-Nearest Neighbour

k-NN [5] is an approach that requires pre-defined core elements, such as a set of patients with a known MMR group, distance, or similarity metric to compute the distance between patients and the value of the nearest neighbour *k*. We tried multiple *k* values starting with

odd numbers and selected *k* with the minimum misclassification rate. To predict a test patient, the set of *k* closest training data to this patient will be *k*-NN. Majority voting in training data chooses the output prediction.

S3.3 Naïve Bayes

Naïve Bayes is a family of probabilistic classifiers that use Bayes' theorem [5]. We defined conditional probability $P(X|Y)$ to be the score of the probability that a patient with prognostic scores value $X=(X_1,X_2,.....X_n)$ belongs to group Y_k (Y_1: achieving MMR; and Y_2: not achieving MMR), where *n* is the number of prognostic scores and *k* is the possible outcome. We used Equation (1) to compute the probability:

$$P(Y_k|X)=\frac{f(X|Y_k)*P(Y_k)}{P(X)} \quad (1)$$

S3.4 Machine Learning Techniques

We used C# to implement the algorithm and obtain results using three datasets. The main file Start.m included the starter function to load the datasets and create the objects. Each object has properties (Model, IndexInputNum, Training Data, Testing Data, Validation Data, Training Results, Cross-Validation Results, Testing Results and Validation Results) and functions (Performance and Evaluation).

The main Starter calls the algorithms DTcode, SVMCode, KNNCode and NBCode. The following are the C# codes for each algorithm:

- DTCode.m

We implemented the CART algorithm to plot our tree structure. We attached part of the code for 10-fold cross-validation:

```
for k=1:10
    test = (indices == k); train = ~test;
    obj.Tree = fitctree(...
    xdata(train,obj.IndexInputNum), ...
    group(train,1), ...
    'SplitCriterion', 'gdi', ...
    'MaxNumSplits', 20, ...
    'Surrogate', 'off', ...
    'ClassNames', [0; 1]);
    TestData=obj.Trainingdata(test,:);
    z=size(TestData);
        for i=1:z(1,1)
```

The final structure is presented using

```
    view(obj.Tree)
```

- SVMCode.m

In SVM, we used several kernel functions to find the best hyperplan.

```
Obj.SVM = fitcsvm(...
    obj.Trainingdata (train,obj.IndexInputNum), ...
    group(train,1), ...
    'KernelFunction', 'polynomial', ...
    'PolynomialOrder', 2, ...
    'KernelScale', 'auto', ...
    'BoxConstraint', 1, ...
    'Standardize', true, ...
    'ClassNames', [0; 1]);
```

- KNNCode.m

```
obj.KNN =fitcknn(...
    obj.Trainingdata (train,obj.IndexInputNum), ...
    group(train,1), ...
    'Distance', 'Euclidean', ...
    'Exponent', [], ...
    'NumNeighbors', 1, ...
    'DistanceWeight', 'Equal', ...
    'Standardize', true, ...
    'ClassNames', [0; 1]);
```

- NBCode.m

```
obj.KNN= fitcnb(obj. Trainingdata (train,obj.IndexInputNum),obj. Trainingdata
(train,end));
```

S4. Automate the Selection of the highest performance

All the results fro the 10 fold cross validation were saved in *crossvalidation.csv* file. We read the table that contained the mean of the accuracy from 10 fold cross-validation. Then, repeat the same code for returning the highest G-mean and F-score results from training all combined methods.

S5. How to Use Combined Methods to Predict Risk Groups

The process can be briefly described as follows:

1. Gather the required features to calculate the prognostic scores in the features set.
2. Calculate the score values using the equations of each prognostic score and store the results in the score values set.
3. Categorize the score values based on risk indexes belonging to each prognostic score.
4. If you used three prognostic score systems (Sokal, Hasford and EUTOS) apply one of the combined models as follows:
   4.1. Figure A shows the meta decision tree model using the tree structure. After identifying the risk index from each prognostic score, the path from the first node to the leaf represents which score should be used. For example, a patient was low risk according to the Sokal and Hasford scores but high risk according to the EUTOS score. The right path from the tree classifies that this patient is in the high risk group.

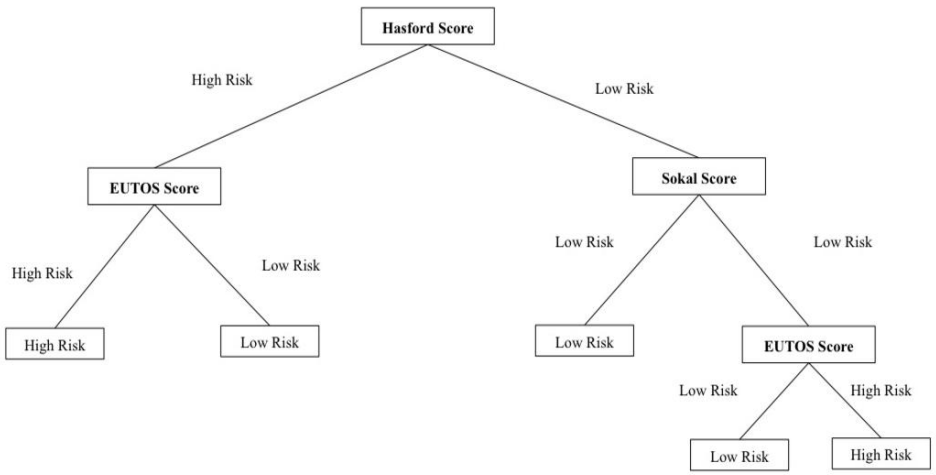


Figure D. The meta decision tree structure constructs relations between the prognostic scores and risk categories. The node represents each score and has two possible splits: It is connected to either the second prognostic score or the risk group in the low or high risk groups.

4.2. The SVM structure:

Kernel Function: polynomial

Kernel Polynomial Order: 2

Scale: 0.3623

Bias: 1.2304

Mu: [0.0723 0.4885 0.1833]

Sigma: [0.1310 0.1619 0.1730]

The property Score Transform= '@(S)sigmoid(S,-3.041237e+00,1.912769e+00)' of the classifier obj.SVM contains the optimal transformation function.

After adapting the new prognostic score ELTS, the SVM model structure was updated to:

4.3. The SVM structure:

Kernel Function: polynomial

Kernel Polynomial Order: 2

Scale: 0.6276

Bias: 1.1235

Mu: [0.0723 0.4885 0.1833 0.2708]

Sigma: [0.1310 0.1619 0.1730 0.1756]

The property Score Transform= '@(S)sigmoid(S,-2.222958e+00,1.004196e+00)' of the classifier obj.SVM contains the optimal transformation function.

References

1. Yeung DT, Osborn MP, White DL, Branford S, Braley J, Herschtal A, et al. TIDEL-II: first-line use of imatinib in CML with early switch to nilotinib for failure to achieve time-dependent molecular targets2015 2015-02-05 00:00:00. 915-23 p.
2. Khalil SH, Abu-Amero KK, Al Mohareb F, Chaudhri NA. Molecular monitoring of response to imatinib (Glivec) in chronic myeloid leukemia patients: experience at a tertiary care hospital in Saudi Arabia. Genetic testing and molecular biomarkers. 2010;14(1):67-74. Epub 2009/12/01. doi: 10.1089/gtmb.2009.0126. PubMed PMID: 19943786.
3. Alsobhi E, Abrar MB, Abdelaal M, Alsaeed A, Absi A, Alzahrani Z, et al. Response to imatinib therapy in adult patients with chronic myeloid leukemia in Saudi population: a single-center study. Leuk Lymphoma. 2014:1-5. Epub 2014/06/24. doi: 10.3109/10428194.2014.935365. PubMed PMID: 24956142.
4. Christianini N, Shawe-Taylor J. An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press; Cambridge, UK2000.
5. Kuncheva LI. Combining Pattern Classifiers Methods and Algorithms: John Wiley & Sons,; 2004.

Supplementary Material B: Modelling Predictors of Molecular Response to Frontline Imatinib for Patients with Chronic Myeloid Leukaemia

S1 Text. Supplementary Results. doi:10.1371/journal.pone.0168947.s001

Supplement Results

Imputation for Missing Values

The identifiers IDs of the patients those imputed values of the factors of patients with missing values is listed in Table A.

Table A The identifiers of the patients those imputed values of the factors of patients with missing values.

Patient ID	Factors with Missing value	Missing value imputed
4	Spleen, Sokal score, Hasford Score	0, 0.63, 690.713
10	$IC50^{IM}$	0.6
12	OA, $IC50^{IM}$	4.02, 0.8
25	$IC50^{IM}$	0.86
26	Eosinophils and Hasford score	0.65, 449,91
43	EUTOS score	90.525
48	Blast, Sokal Score, and Hasord Score	0.5, 0.61, 1006.42
53	$IC50^{IM}$	0.61
64	EUTOS score	30
69	$IC50^{IM}$	0.85
71	*BCR-ABL1* Level pretherapy	48.5
77	$IC50^{IM}$	0.82
79	$IC50^{IM}$	0.87
88	$IC50^{IM}$	0.73
93	*BCR-ABL1* Level pretherapy	125
95	Age, Sokal score, Hasford score	52, 1.27, 1436.31
97	Blast, Sokal score, Hasford score	0, 0.76, 603.52

102	$IC50^{IM}$	2.788
124	Platelets, Blast, Sokal score, Hasford score	399, 0, 0.77, 661.594
126	OA	6.44
151	OA, $IC50^{IM}$	4.97, 0.57
153	OA	4.72
174	Platelets, Blast, Sokal score, Hasford score	311.33, 3.3, 1.22, 1250.59
175	Platelets, Blast, Sokal score, Hasford score	355, 1.66, 0.87, 682.16
187	OA	3.55
206	$IC50^{IM}$, EUTOS score	1.47, 57
207	Basophils, Eosinophils, Hasford score	7.65, 8.905, 1024.75
208	$IC50^{IM}$	1.5
210	EUTOS score	57

The imputation is done by using the linear interpolation for the continuous and categorical data.

Correlation Coefficients for MMR at 24 months with Original Data and Completed Data

The original data included missing values and completed data after imputation of missing values of factors. The correlation coefficient was calculated between each predictive factor and the MMR at 24 months. Table B shows that there were not large differences in correlation coefficients.

Table B The correlation Coefficient in Original data and completed data

Original Data		**Completed Data**	
Predictor	**Correlation p**	**Predictor**	**Correlation p**
Spleen	0.31521833	Spleen	0.31732076
Age	0.19547668	Basophils	0.19958623
Basophils	0.19309652	Age	0.19328342
Monocytes	0.18853493	**WCC**	0.18869802
WCC	0.17740892	**Monocytes**	0.18803225
Eosinophils	0.16655321	Eosinophils	0.18417037
ANC	0.14582059	ANC	0.15514001
Blast	0.110962816	Blast	0.12231851
IC50imatinib	0.10926288	**Hasford Score**	0.11979281
Hasford Score	0.108925685	**Platelets**	0.09735691
OA	0.10400609	**IC50imatinib**	0.096847385
Platelets	0.0939224	**OA**	0.09405362
Gender	0.08926122	Gender	0.08926122
EUTOS Score	0.0820541	EUTOS Score	0.079521775
***BCR-ABL1* level at diagnosis**	0.03623853	**Sokal Score**	0.038800742

Sokal Score	0.03104846	***BCR-ABL1* level at diagnosis**	0.030398324
BCR-ABL1 Transcript Type	0.02592691	*BCR-ABL1* Transcript Type	0.02592691
Lymphocytes	0.008399159	Lymphocytes	0.022077713

Overall Summary of Missing Values in TIDEL II and Saudi Population

Table C Missing values percentage included in TIDEL II and Saudi datasets

Factors	TIDEL II (n= 210)	Saudi Population (n=172)
Age (years)	0	0
Gender	0	4(2.32%)
Spleen (cm)	1(0.48%)	0
BCR-ABL1 Transcript Type	0	Not palpable
OA (ng/200,000 cells)	6(2.86%)	Not palpable
IC50IM (µM)	14(6.67%)	Not palpable
BCR-ABL1 level pretherapy (at diagnosis)	4(1.90%)	Not palpable
ANC (109 /L)	1(0.48%)	Not palpable
Monocytes (109 /L)	1(0.48%)	0
Lymphocytes (109 /L)	1(0.48%)	Not palpable
Basophils (109 /L)	1(0.48%)	Not palpable
Eosinophils (109 /L)	3(1.43%)	Not palpable
WCC (109 /L)	1(0.48%)	Not palpable
Blasts (109 /L)	7(3.33%)	Not palpable
Platelets (109 /L)	4(1.90%)	9(5.23%)
Blasts (% of leukocytes)	0	0
Basophils (% of leukocytes)	3(1.43%)	10(5.81%)

Eosinophils (% of leukocytes)	3(1.43%)	10(5.81%)

Machine Learning Implementation

The actual response is (*ResMMR*: Actual MMR, 1 refers to achieving MMR and 0 refers to not achieving MMR).

First, the wrapper approach selects the prognostic subset (*InputFactorsIndex*: refers to the input index in the data set). For example, if the OA is located in the second column and $IC50^{IM}$ is located in the fifth column of the data set, then the *InputFactorsIndex*=[2 5]. Next, the *classregtree* is a keyword in Matlab to implement (*DTStruct*) CART structure; *fitcknn* is a keyword to build KNN. In addition *NaiveBayes.fit* is reserved for building Naive Bayes. Finally, the *P* refers to the prediction results; 1: Yes, and 0: No. For construction and evaluation of the CART model:

DTStruct = classregtree(TrianData(*InputFactorsIndex)*, *ResMMR*, *'method'*, *'classification'*, *'splitcriterion'*, *'gdi'*, *'categorical'*,1: length(*InputFactorsIndex*));

P= eval(*DTStruct,TestData*(*InputFactorsIndex*));

S1 File. Decision Tree Structures. doi:10.1371/journal.pone.0168947.s002

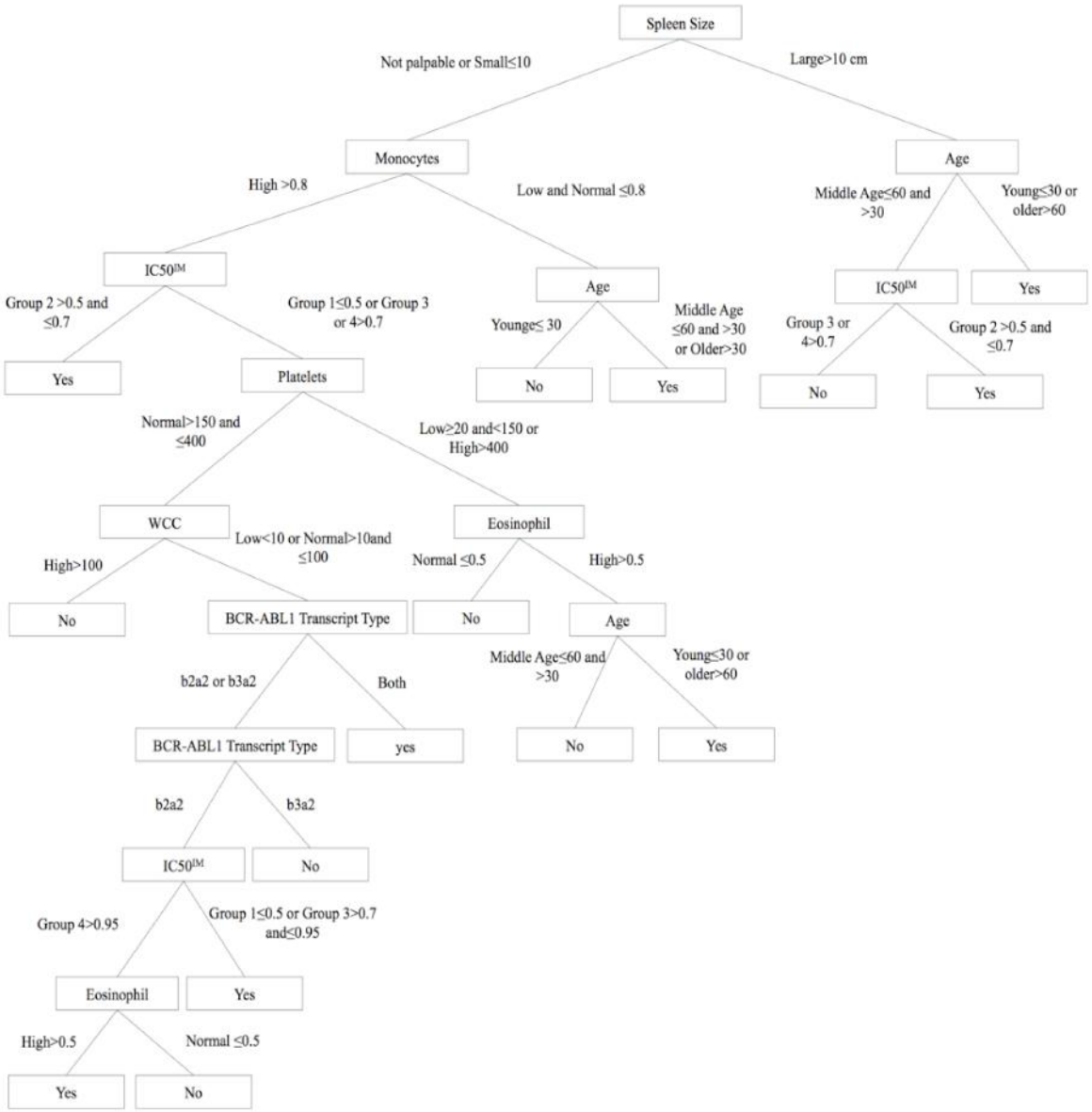


Fig A. Model A Structure. All the predictive factors model in the tree graph.

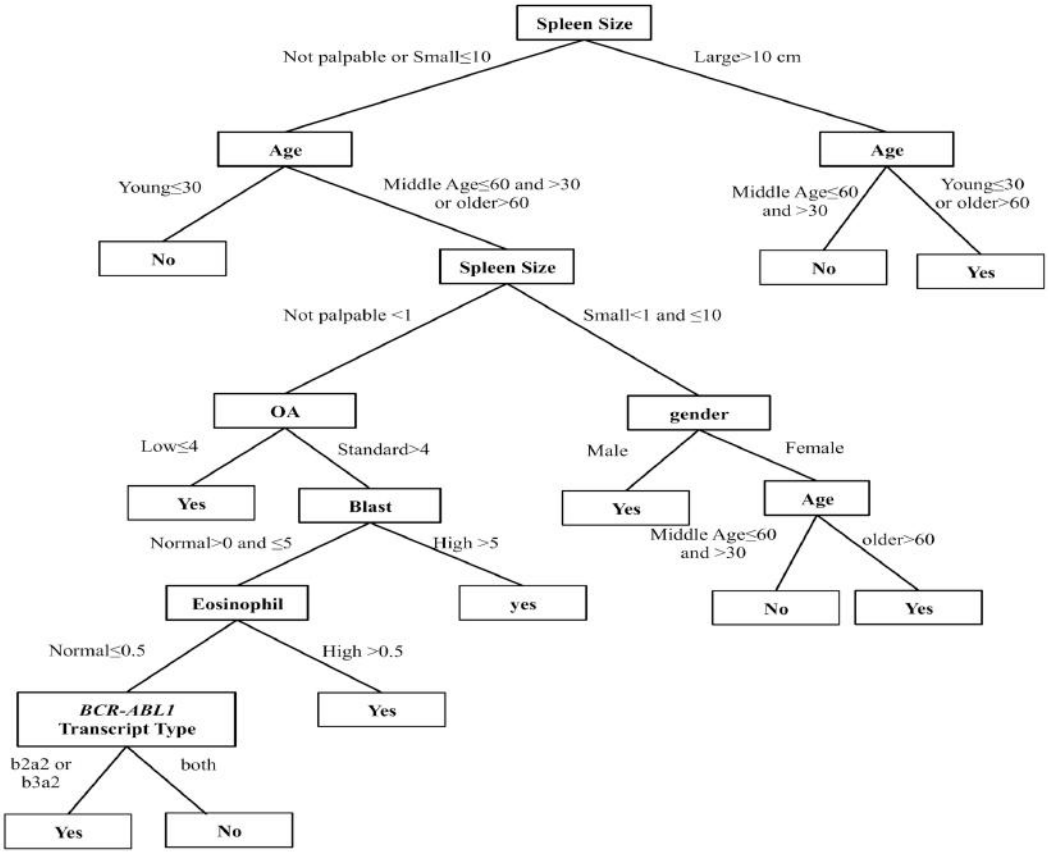


Fig B. Model E structure, the final model in the tree graph that achieved high G-mean performance.

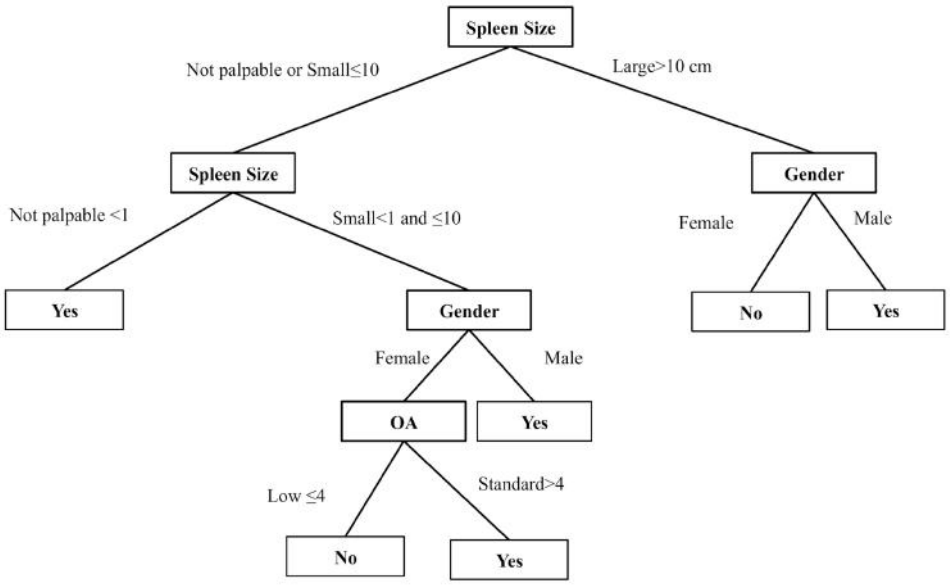


Fig C. 7 Model F structure, the final model in the tree graph that achieved high F-score performance.

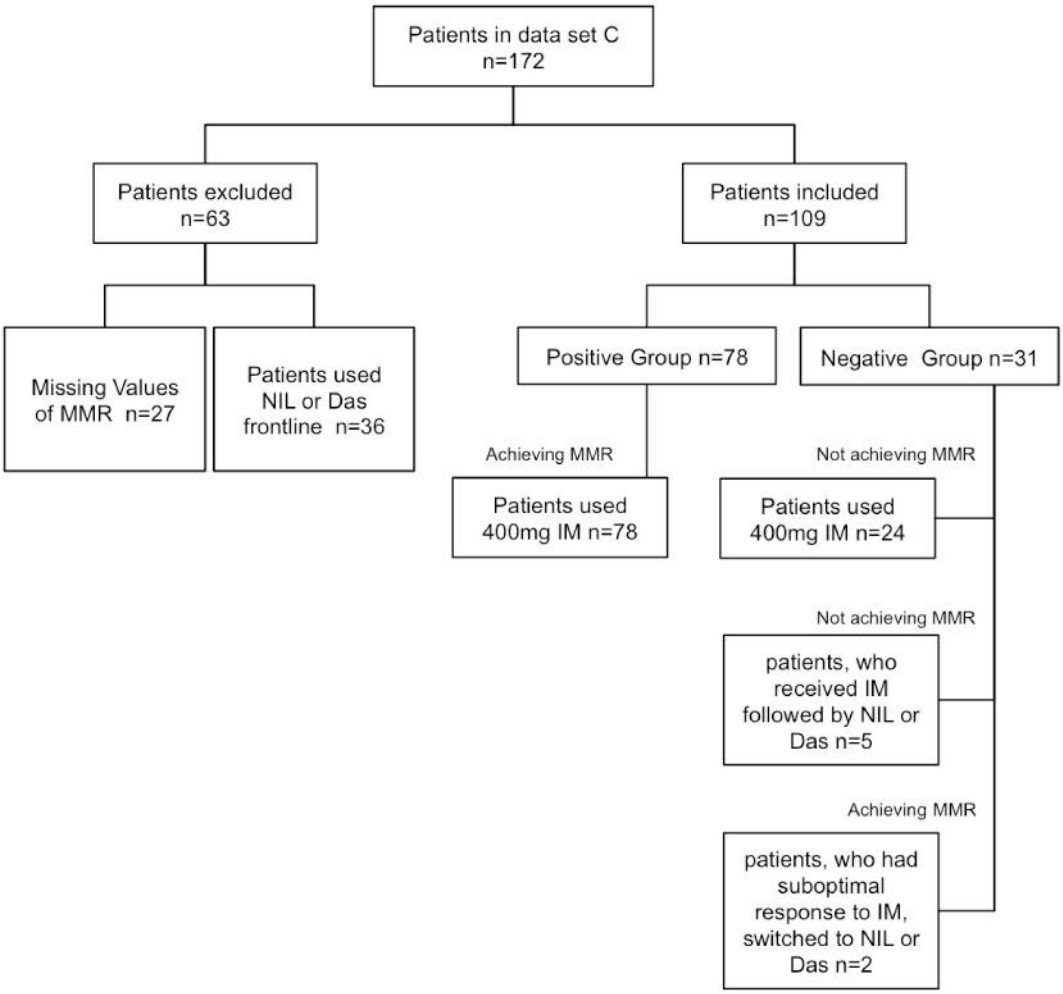


Fig D. Saudi patients used in this study, inclusion and exclusion criteria.

S1 Table. Clinical Prediction Rules, dataset includes list of clinical prediction rules that constructed from recommended models. doi:10.1371/journal.pone.0168947.s003. Fore, example, the following table displays a part of the S1 Table.

	Rule ID	If Age	AND Gender	AND Spleen Size	…	…	Then MMR
Model A	1	Young	Female	Not Palpable			Yes
. . .	2	Young	Female	Small			No
Model *n*	Number of clinical prediction rules	Age Categories	Gender Categories	Spleen Size Categories			MMR Outcome Yes or No

……The End…..