HAEMATOLOGICAL AND MOLECULAR CHARACTERISATION OF HIGH HAEMOGLOBIN F AMONG ANAEMIC PATIENTS IN HOSPITAL UNIVERSITI SAINS MALAYSIA

SITI NOR ASSYUHADA BINTI MAT GHANI

UNIVERSITI SAINS MALAYSIA

2020

HAEMATOLOGICAL AND MOLECULAR CHARACTERISATION OF HIGH HAEMOGLOBIN F AMONG ANAEMIC PATIENTS IN HOSPITAL UNIVERSITI SAINS MALAYSIA

by

SITI NOR ASSYUHADA BINTI MAT GHANI

Thesis submitted in the fulfillment of the requirements
for the degree of
Master of Science

DECEMBER 2020

ACKNOWLEDGEMENT

Alhamdullillah. Praise to Allah for giving me strength and His blessing for completing this thesis. I would like to express my special thanks of gratitude to my supervisor, Dr Zefarina Binti Zulkafli and co-supervisor, Dr. Maryam Binti Azlan for all guidance in completing this project. I do appreciate their enthusiasm, support and constant guidance given to me along the way in completing my study. It would not have been possible to write this dissertation without their guidance and persistent helps.

Thank you to Biomedical Department, PPSK, and Haematology Department, PPSP as well as Institute for Research in Molecular Medicine (INFORMM) including the laboratory staffs for providing all the related laboratory equipment as well as the technical guidance throughout this study.

Special thanks to my parents, Mat Ghani bin Dollah and Mek Som binti Mamat for always encourage me in finishing this thesis and finally to all my friends for helping me out with the ideas related to my study.

Thank you.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ACRONYMS, ABBREVIATIONS AND SYMBOLS

β	: Beta
δ	: Delta
γ	: Gamma
°C	: Degree celsius
°C/sec	: Degree celsius per second
°C/min	: Degree celsius per minute
>	: Greater than/ modifier letter right arrowhead
∞	: Infinity
<	: Less than
≤	: Less than or equal to
µg	: Microgram
µL	: Microlitre
µM	: Micromolar
%	: Percent
/	: Solidus
~	: Tilde
X	: Times
A	: Adenine
BQ1	: Lysis buffer
BQ2	: Washing buffer
bp	: Base pair
C	: Cytosine
CE	: Cation exchabge
Cq	: Cycle of quantification

Cd : Codon

dNTPs : Deoxynucleoside triphosphates

HPLC : High performance liquid chromatography

dH_2O : Distilled water

DNA : Deoxyribonucleic acid

EDTA : Ethylenediaminetetraacetic acid

FBC : Full blood count

fL : femtolitre

g : Gram

g : Gravity

G : Guanine

Hb : Haemoglobin

HbA : Adult haemoglobin

HbA_2 : Haemoglobin A2

HbE : Haemoglonin E

HbF : Foetal haemoglobin

HBB : Beta globin

HBD : Delta globin

HBG : Gamma globin

HPFH : Hereditary persistence foetal haemoglobin

HUSM : Hospital Universiti Sains Malaysia

INFORMM : Institute for Research in Molecular Medicine

IVS : Intervening sequence

JEPeM : Jawatankuasa Etika Penyelidikan (Manusia)

k : Kilo

kbp : Kilobase pair

LCR : Locus control region

mg : Miligram

min : Minute

mL : Mililitre

mM : Milimolar

M : Molar

MCV : Mean corpuscular volume

MCH : Mean corpuscular haemoglobin

MCHC : Mean corpuscular haemoglobin concentration

$MgCl_2$: Magnesium chloride

T_m : Melting temperature

n : Number of subjects

ng : Nanogram

ng/mL : Nanogram per microlitre

NTC : Non-template control

pg : petagram

pH : Power of hydrogen

PCR : Polymerase chain reaction

QTLs : Quantitative trait loci

RCF : Relative centrifugal force

RFU : Relative fluorescence unit

SEM : Standard error mean

SEA : Southeast Asia

SNPs : Single nucleotide polymorphisms

T	: Thymine
TBE	: Tris/Borate/EDTA buffer
USM	: Universiti Sains Malaysia
US	: United State of America
V	: Volts

PENCIRIAN HEMATOLOGI DAN MOLEKULAR HEMOGLOBIN F YANG TINGGI DALAM KALANGAN PESAKIT ANEMIA DI HOSPITAL UNIVERSITI SAINS MALAYSIA

ABSTRAK

Anemia adalah satu keadaan yang biasanya dikaitkan dengan pelbagai penyakit. Kebiasaannya, tahap hemoglobin F (HbF) dalam orang dewasa kurang daripada 1.0%. Terdapat beberapa lokus genetik yang mempunyai pengaruh penting terhadap tahap HbF. Tujuan kajian ini adalah untuk menentukan hubungan antara tahap HbF yang tinggi dengan parameter hematologi serta kehadiran *BCL11A* (rs1186868) dan *HMIP-2* (rs9376090) polimorfisme nukleotida tunggal (SNPs) dalam pesakit anemia akibat sebab perolehan. Kajian ini melibatkan 144 pesakit anemia dari Hospital Universiti Sains Malaysia (HUSM) yang mempunyai tahap HbF > 1.0%. Kromatografi cecair berprestasi tinggi (HPLC) digunakan untuk menentukan tahap HbF dan HbA_2. Multipleks ARMS-PCR dan gap-PCR masing-masing telah dilakukan bagi sampel yang mempunyai tahap HbA_2 yang tinggi (> 3.2%) dan tahap HbA_2 yang normal (≤3.2%) untuk mengesan mutasi pada kluster gen β-globin. Diskriminasi alel untuk rs1186868 dan rs9376090 polimorfisme telah dilakukan menggunakan teknik PCR masa nyata untuk sampel yang tidak dapat mengesan mutasi pada kluster gen β-globin. Dalam kajian ini, purata umur pesakit adalah 19.99 ± 1.64 tahun dan didominasi oleh wanita 61.1%. Majoritinya adalah orang Melayu (99.3%). Terdapat korelasi negatif yang sederhana dan signifikan secara statistik antara tahap HbF dengan tahap Hb dan bilangan RBC, (r = -0.348, P < 0.05) (r = -0.377, P < 0.05). Manakala tahap HbF dengan MCV dan MCH menunjukkan korelasi negatif yang lemah tetapi tidak signifikan secara statistik (r = -

0.079, $P > 0.05$) ($r = -0.073$, $P > 0.05$). Daripada multipleks ARMS-PCR, 65 sampel (74.7%) mutasi dikesan dimana 49 adalah heterozigot Cd 26 (75.4%), 10 heterozigot Cd 41/42 (15.4%), 3 kompaun heterozigot Cd 26 and Cd 41/42 (4.6%) and 3 heterozigot IVS 1–1 (4.6%), manakala 22 sampel lagi tidak dapat dikesan. Daripada 57 sampel, hanya satu sampel (1.8%) dijumpai positif Thai (δβ)°-talasemia jenis delesi apabila disaring menggunakan multipleks gap-PCR yang mempunyai empat target delesi; Siriraj J Gγ(Aγδβ)°-talasemia, Thai (δβ)°-talasemia, HPFH-6, and Hb Lepore. Terdapat perbezaan yang signifikan antara purata tahap HbF pesakit yang mempunyai mutasi dan delesi pada kluster gen β-globin dengan yang tidak mempunyai mutasi dan deletion pada kluster gen β-globin ($P < 0.05$). Frekuensi alel kecil (MAF) pada kedua-dua rs1186868 dan rs9376090 adalah sama dengan penduduk Asia Timur (EAS). Tidak terdapat perbezaan yang signifikan pada tahap HbF antara genotip yang mengandungi alel kecil rs9376090 (TC dan CC) jika dibandingkan dengan genotip TT ($P > 0.05$). Sebagai kesimpulan, tahap HbF mempunyai korelasi dengan status anemia subjek. Tahap HbF yang tinggi menunjukkan hubungan dengan pesakit anemia akibat diwarisi dan sebab perolehan. Selain itu, tidak terhadap hubungan antara tahap HbF yang tinggi dengan kehadiran SNPs rs9376090-C dalam kalangan pesakit anemia akibat sebab perolehan. Meskipun memerlukan penyelidikan yang lebih lanjut dalm bidang ini, kajian ini menyediakan data yang boleh digunakan sebagai garis panduan untuk rawatan dan pengurusan yang lebih baik kepada pesakit anemia.

HAEMATOLOGICAL AND MOLECULAR CHARACTERISATION OF HIGH HAEMOGLOBIN F AMONG ANAEMIC PATIENTS IN HOSPITAL UNIVERSITI SAINS MALAYSIA

ABSTRACT

Anaemia is a condition usually associated with variety of diseases. In normal adults, haemoglobin F (HbF) levels are usually less than 1.0%. There are several genetic loci that have significant influence on HbF levels. The aim of this study is to determine the association between elevated HbF level with haematological parameters and the presence of the *BCL11A* (rs1186868) and *HMIP-2* (rs9376090) SNPs in anaemic patients due to acquired causes. This study involved 144 anaemic patients from Hospital Universiti Sains Malaysia (HUSM) with HbF level > 1.0%. High-performance liquid chromatography (HPLC) was used to determine the HbF and HbA_2 levels. Multiplex ARMS-PCR and gap-PCR were performed for those samples with high HbA_2 level (>3.2%) and normal HbA_2 level (≤3.2%) respectively to detect mutation at β-globin gene cluster. Allelic discrimination for rs1186868 and rs9376090 were performed using real-time PCR technique for samples with no mutation detected. In this study, the mean age of patients is 19.99 ± 1.64 years with female 61.1% predominance. Majority were Malays (99.3%). There was a moderate negative correlation and statistically significant between HbF level with Hb level and RBC count, (r = -0.348, P < 0.05) (r = -0.377, P < 0.05) respectively. Meanwhile, the correlation between HbF level with MCV and MCH showed weak negative correlation but not statistically significant, (r = -0.079, P > 0.05) (r = -0.073, P > 0.05) respectively. Following multiplex ARMS-PCR, 65 (74.7%) mutations were detected which comprises of 49 heterozygous Cd 26 (75.4%), 10 heterozygous Cd

41/42 (15.4%), 3 compound heterozygous Cd 26 and Cd 41/42 (4.6%) and 3 heterozygous IVS 1–1 (4.6%), while 22 patients were not detected. Out of 57 samples, only one patient (1.8%) was found positive with Thai (δβ)°-thalassaemia type deletions when subjected to the multiplex gap-PCR consisted of four target deletions; Siriraj J Gγ(Aγδβ)°-thal, Thai (δβ)°-thalassaemia, HPFH-6, and Hb Lepore. There was a significant difference between the mean of HbF level of patients with and without β-globin gene cluster mutation and deletion ($P < 0.05$). The minor allele frequency (MAF) in both rs1186868 and rs9376090 shows similar to the East Asian (EAS) population. There is no significant difference of HbF level between genotypes containing HbF-promoting alleles of rs9376090 (TC and CC) when compared to genotype TT ($P > 0.05$). In conclusion, HbF level correlates with anaemic status of the subjects. Elevated HbF levels showed associations with both inherited and acquired causes. Additionally, there is no relationship between the increased HbF levels with the presence of SNPs rs9376090-C among acquired anaemic patients. Despite the need for further research in this area, this study provides data that can be used as a guideline for better anaemia treatment and management.

CHAPTER 1

INTRODUCTION

1.1 Background of study

Foetal haemoglobin (HbF) is formed by two α- and two γ-globin chains (α2 γ2). HbF is synthesised during foetal development due to high expression of γ-globin genes (Edoh *et al.*, 2006; Stamatoyannopoulos, 2005). The gamma globin (*HBG1* and *HBG2*) genes are usually expressed in the liver, spleen, and bone marrow during foetal life. The normal switch from HbF to HbA synthesis occurs during the first year of life lead to rapidly decreasing of HBF level to reach a concentration of less than 1% in normal children after the age of 1 year and adults (Pace *et al*., 2016; Amato *et al*., 2014; Steinberg *et al*., 2014; Inusa *et al*., 2006). In normal children and adults, the β-globin gene is predominant and γ-globin genes are poorly expressed. According to Koh *et al*. (2017), subjects are considered to have α and β-globin gene cluster mutations when the HbF level is more than 1.0%, HbA_2 level is above 3.2%, and/or show the presence of other abnormal Hb (Koh *et al*., 2017).

The exact time of disappearance of HbF from red blood cells of infants differs and the signal that regulates the switch from HbF to HbA is not known. HbF levels can be evaluated by counting the number of F cells, that is, adult erythrocytes that contain measurable amounts of HbF (Xu *et al.*, 2009).

Elevated HbF level may have its own clinical consequences. Several acquired conditions are associated with increased level of HbF. These include drug-induced illness, micronutrient deficiency, and other medical conditions such as aplastic

anaemia, pernicious anaemia, sideroblastic anaemia, autoimmune diseases, and leukaemia (World Health Organization, 2011). Besides this, a high level of HbF in adults is also linked to inherited factors due to mutations or deletions in the beta (β)-globin gene cluster which cause pathological changes, such as in patients with thalassaemia disorders (Sankaran *et al.*, 2008; Thein *et al.*, 2007). The inheritance of HbF does not always comply with the Mendelian inheritance pattern; hence, the exact underlying cause in the increased level of HbF remains unclear (Sankaran *et al.*, 2008).

Deletions and/or mutation of various sizes and positions within the β-globin gene cluster located on chromosome 11 are responsible for certain clinical disorders, namely hereditary persistence of foetal haemoglobin (HPFH), β-thalassaemia, sickle cell anaemia, HbE disorder, and delta beta (δβ)-thalassaemia (Carrocini *et al.*, 2011). Sickle cell anaemia and β-thalassaemia mutations are characterised by significant elevations of HbF and HbA_2 levels while HPFH and δβ-thalassaemia mutations are characterised by elevations of HbF level but normal HbA_2 level (Galanello, 2013).

Furthermore, several genetic loci have a significant influence on HbF level (Carrocini *et al.*, 2011). According to Sokolova (2019), the *XmnI* polymorphism, *HMIP* locus, and *BCL11A* gene are responsible for 45% of variations in HbF level. It has been shown that the variants in *BCL11A*, *HBS1L-MYB* and *HBB* loci are associated with HbF levels and ameliorate the severity in β-thalassaemia as well as SCA (Gallanello *et al*., 2009; Lettre *et al*., 2008). *BCL11A*, one of the quantitative trait loci (QTLs) that control HbF levels, is located in the 2p16 region of chromosome 2. Fifteen percent of HbF variations are attributed to *BCL11A*

polymorphisms. *BCL11A* gene has been established as a direct regulator of HbF level. Studies shown that the expression of the transcriptional repressor *BCL11A* is regulated by erythroid-specific enhancers that contain 3′ DNase hypersensitive sites (DHS) located +62, + 58, and +55 kb from the transcription initiation site of *BCL11A* (Bauer *et al*., 2013). Bauer *et al*. (2013) reported that common genetic variation at *BCL11A* associated with HbF level lies in noncoding sequences decorated by an erythroid enhancer chromatin signature. In a study of 179 unrelated healthy subjects from a British twin registry, six SNPs in *BCL11A* which are rs243027, rs243081, rs6732518, rs1427407, rs766432 and rs4671393 were reported to be associated with F-cell numbers (Menzel *et al.*, 2007).

Another QTL associated with elevated HbF is *HBS1L–MYB intergenic polymorphism (HMIP)* found on chromosome 6 (6q23). It is located between the *HBS1L* gene (codes for elongation factors and regulates multiple cellular processes) and the *MYB* gene (encodes transcription factors and participates in ontogenesis and erythropoiesis). *HMIP* is unique among the HbF modifier loci because it has marked pleiotropic effects. For example, it affects general hematological parameters including HbF level in healthy individuals (Mtatiro *et al*., 2015). High-risk genotypes of six HbF-associated SNPs, rs9376090, rs7776054, rs9399137, rs9389268, and rs9402685 in the *HMIP* are associated with high HbF level (Lai *et al*., 2016). Previous study in Brazil confirms that common *HMIP-2* variants increase HbF levels in SCA patients (Leonardo *et al*., 2016).

HbF is one of the haematological parameters in evaluating disorders related to blood. HbF parameter is considered as a potential biomarker that indicates the

presence of underlying diseases; however, this requires further investigation (Steinberg *et al.*, 2009). HbF has shown a great importance where it improved the clinical features of an individual with sickle cell disease and β-thalassaemia (Steinberg and Thein, 2018). This nonspecific biomarker is important, denoting a pathophysiological process in patients with silent or non-straight forward disorders which can be associated particularly with anaemic conditions (Steinberg *et al.*, 2009).

1.2 Significance of the study

A number of SNPs have been shown specifically to be beneficial in ameliorating the severity of anaemia due to inherited causes by influencing the levels of adult HbF. To date, the data regarding the role of HbF associated with acquired anaemia among Malaysian population are still lacking. Mapping of QTLs as proposed in this study is one of the tools to identify the susceptible SNPs specifically *BCL11A* and *HMIP-2* and determine its association with elevated HbF in acquired anaemia patients. Thus, the benefit of this study is imperative in considering the role of genetics (SNPs) in *BCL11A* and *HMIP-2* as a potential disease biomarker for HbF in anaemia severity and could provide better approach in management of anemic patients.

1.3 Objective of the study

1.3.1 General objective

To study the role of HbF as a potential disease biomarker among anaemic patients with increased HbF level due to acquired causes.

1.3.2 Specific objectives

1. To study the association between HbF level and anaemic status of the subjects.
2. To determine the increased of HbF levels in anaemic patients due to inherited or acquired causes.
3. To determine the association between elevated HbF and the presence of SNPs rsll86868 dan rs9376090 in anaemic patients due to acquired causes.

1.4 Hypotheses of the study

1. There is a correlation between HbF level and anaemic status of the subjects.
2. There is an association between HbF level due to inherited causes and acquired causes.
3. There is an association between high HbF with SNPs rs1186868 and rs9376090 in anaemic patients due to acquired causes.

CHAPTER 2

LITERATURE REVIEW

2.1 Haemoglobin

Erythrocytes contain a protein known as haemoglobin (Hb) that carries oxygen from the lungs to all parts of the body. The biconcave shape of erythrocytes allows high diffusion rate of oxygen and carbon dioxide across their membrane (Diez-Silva *et al.*, 2010). Hb is a complex protein made up of two pairs of homologous subunits, alpha (α) and beta (β). Each Hb molecule forms a pocket containing a haem group, which has an important role in oxygen-binding capacity (Farashi and Harteveld, 2017). In the arterial circulation, Hb encompasses a high affinity for oxygen and a low affinity for carbon dioxide, organic phosphates, and hydrogen and chloride ions. In the venous circulation, these relative affinities are vice versa (Marengo, 2006). There are seven types of Hb throughout human life. The list of Hb types is shown in Table 2.1.

Each type of Hb comprises different compositions of globin chains. Adult haemoglobin (HbA) is a predominant type of Hb which starts to develop after three months of postnatal life. HbA consists of two alpha chains and two beta chains (α_2 β_2) (Figure 2.1) (Campbell, 2009). HbA is the major component of Hb in normal adults, usually comprising about 97% of the total Hb. The remainder is HbA_2 (α_2 δ_2) and HbF, which usually constitute about 2%–3% and 1% in normal individuals, respectively.

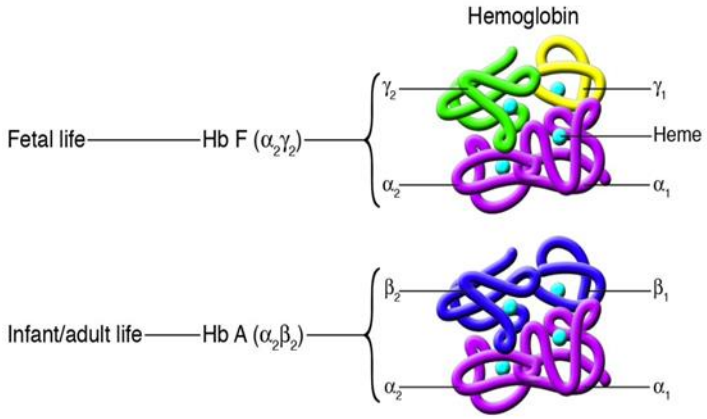

Figure 2.1 Types of haemoglobin. Adapted from Manning *et al.*, 2007.

During the foetal stage, HbF is the predominant type of Hb. Hb switching refers to the developmental process that leads to the silencing of gamma (γ)-globin gene expression and the reciprocal activation of adult beta (β)-globin gene expression (Frenette and Atweh, 2007).

Table 2.1 The types of haemoglobin in humans

Haemoglobin	Subunit/structure	Classification
A	α2β2	Adult
HbA_2	α2δ2	Adult
F	α2γ2	Foetal
Gower I	ζ2ε2	Embryonic
Gower II	α2ε2	Embryonic
Portland I	ζ2γ2	Embryonic
Portland II	ζ2β2	Embryonic

Adapted from Manning *et al.*, 2007.

2.1.1 Foetal Haemoglobin (HbF)

Foetal haemoglobin (HbF) is formed by two α- and two γ-globin chains ($\alpha_2\gamma_2$) (Stamatoyannopoulos, 2005). The oxygen affinity for HbF increase because the absent of 2,3-bisphosphoglycerate (2,3-DPG). 2,3-DPG is a little molecule in the erythrocytes that binds with haemoglobin beta subunits and it decreases the affinity for oxygen and promotes the release of the remaining oxygen molecules bound to the

haemoglobin. The high affinity of HbF for oxygen favors the taking up of oxygen in the placenta. HbF is important in transporting oxygen from maternal to foetal circulation. The HbF oxygen dissociation curve is left-shifted in comparison to HbA. The partial pressure at which HbF is half saturated with oxygen (P50) is 19 mm Hg, compared to 27 mm Hg for HbA. This value indicates that HbF has a high affinity for oxygen, giving HbF the ability to bind oxygen more readily from the maternal circulation. In the foetal systemic circulation, the low oxygen tension allows for proper unloading of oxygen, despite HbF's oxygen affinity. The lower oxygen tension in the foetus is important for development, particularly in angiogenesis. Because foetal blood shows a higher affinity for oxygen than maternal blood, oxygen will diffuse from the pregnant maternal to the foetal circulation within the placenta, allowing for oxygenation of foetal tissues.

In adults, the β-globin gene is predominant; approximately 98% of all Hb consist of HbA ($\alpha_2 \beta_2$). Thus, γ-globin genes are poorly expressed; less than 1% of Hb in adults is made up of HbF (Xu *et al.*, 2009). HbF levels can be evaluated by counting the number of F cells. The HbF and F-cell levels vary considerably in healthy adults, but commonly, there is a good correlation between the two (Menzel *et al.*, 2007). The concentration of HbF depends on several factors. HbF is increased in inherited conditions, such as HPFH, hereditary spherocytosis, sickle cell crisis, and thalassaemia. The level of HbF is also elevated in acquired states, such as pregnancy, aplastic anaemia, thyrotoxicosis, hepatoma, myeloproliferative disorders, and hypoplastic myelodysplastic syndrome. In addition, several genetic loci may also significantly influence HbF level (Carrocini *et al.*, 2011).

Certain mutations within the β-globin gene cluster are associated with increased γ-chain expression (Cao and Galanello, 2010). For example, in β-thalassaemia and related conditions, gamma chain production continues into adulthood for additional γ-globin chains to bind with the excess of α-chains and causes imbalance between α and β-globin chains. Thus, HbF is synthesized to help the erythrocytes survive longer in the circulation. Therefore, induction of HbF expression in erythroid cells is an important therapeutic approach in patients with Hb disease (Fathallah and Atweh, 2006). Patients with HbF levels of ≥20% have a mild phenotype and those with levels of ≥30% are likely to be asymptomatic (Adekile, 2011).

2.2 Anaemia

Anaemia is a global public health problem that affects both developing and developed countries. Anaemia is generally classified based on blood loss, lower erythrocytes production or ineffective erythropoiesis and haemolysis of erythrocytes. More specifically, anaemia is defined as the condition where the normal level of Hb is lower in the body, which reduces the oxygen-carrying capacity (Henry *et al.*, 2004). The World Health Organization (WHO) defines anaemia by a level of Hb less than 13.0, 12.0, and 11.0 g/dL in men, non-pregnant women, and pregnant women, respectively (Table 2.2). Meanwhile, severe anaemia is defined as Hb <7.0 g/dL in children aged 6 to 59 months and pregnant woman while Hb <8.0 g/dL in others. According to WHO, anaemia exists in the following proportions: 3.9% of men, 38.5% of pregnant women, and 17.3% of non-pregnant women. The Western countries have an increased occurrence of anaemia in the elderly population (Gaskell *et al.*, 2008).

Table 2.2 Haemoglobin levels in different age groups (g/dL)

Population	Non-anaemia	Mild	Moderate	Severe
Children 6 - 59 months of age	> 11.0	10.0-10.9	7.0-9.9	< 7.0
Children 5 - 11 years of age	> 11.5	11.0-11.4	8.0-10.9	< 8.0
Children 12 - 14 years of age	> 12.0	11.0-11.9	8.0-10.9	< 8.0
Non-pregnant women (15 years of age and above)	> 12.0	11.0-11.9	8.0-10.9	< 8.0
Pregnant women	> 11.0	10.0-10.9	7.0-9.9	< 7.0
Men (15 years of age and above)	> 13.0	11.0-12.9	8.0-10.9	< 8.0

Adapted from World Health Organization, 2011.

Normal Hb distribution varies in different genders and ethnicities as well as the physiological statuses of an individual. The patient's history and physical examination should be taken into consideration when performing anaemia diagnosis (Conrad, 1990). Several factors affect the concentration of Hb which can be divided into inherited and acquired causes. Examples of inherited disorders are congenital pernicious anaemia, Fanconi anaemia, G6PD deficiency, hereditary spherocytosis, thalassaemia and its variants such as sickle cell anaemia (Mosca *et al.*, 2009; Bain, 2006). Thalassaemia may be treated with blood transfusions as well as other treatments such as iron chelation therapy and folic acid consumption because of the reduced production of healthy erythrocytes and Hb. Sickle cell anaemia is the production of abnormal form of Hb that causes erythrocytes to change from a biconcave to a sickle shape. This abnormal shape of erythrocytes causes them to stick together, resulting in difficulty for them to pass through blood vessels, leading to damage of the body tissues (Jones, 2017).

Meanwhile, acquired conditions of anaemia can be drug-induced anaemia, aplastic anaemia, pernicious anaemia, sideroblastic anaemia, vitamin B12/folate deficiency, autoimmune diseases (a form of anaemia of chronic disease), and leukaemia as well as physiological characteristics such as age, pregnancy status, and sex (Mosca *et al.*, 2009). In some cases, certain drugs can cause the immune system to attack the body's own erythrocytes by producing antibodies. These antibodies attach to erythrocytes and cause haemolysis. Examples of drugs that can cause this type of haemolytic anaemia include cephalosporin, a class of antibiotics (Garratty, 2009). Cephalosporin is an example of drug-dependent antibodies causing production of erythrocyte autoantibodies which affect the immune system. Drugs bind covalently to erythrocyte membrane proteins. Drug antibodies, usually IgG, attach themselves to drug-coated erythrocytes and are subsequently cleared by macrophages (Garratty, 2012). Frequent adverse drug reactions (ADRs) (≥1.0% of patients) associated with cephalosporin therapy include diarrhoea, nausea, skin rashes, electrolyte disturbances, and inflammation at the injection site. Uncommon ADRs (0.1%–1.0% of patients) include headache, dizziness, vomiting, oral and vaginal candidiasis, nephrotoxicity, pseudomembranous colitis, eosinophilia, superinfection, neutropenia, thrombocytopenia, and fever (Shi *et al.*, 2013).

Heavy menstruation, ulcers, injury, or surgery cause blood loss leading to iron-deficiency anaemia. Pregnancy also causes changes in a woman's blood volume which can result in anaemia. Contributions of each of the factors that causes anaemia during pregnancy vary due to geographical location, dietary practice, and season (Stephen *et al.*, 2018). A diet low in iron, folate, or vitamin B12 increases the risk of iron-deficiency anaemia. These nutrients are important in growth and development

(Song *et al.,* 2010). Folate, vitamin B12, and iron have crucial roles in erythropoiesis. Erythroblasts require folate and vitamin B12 for proliferation during their differentiation. Deficiency of folate or vitamin B12 inhibits purine and thymidylate syntheses, impairs DNA synthesis, and causes erythroblast apoptosis, resulting in anaemia from ineffective erythropoiesis (Koury and Ponka, 2004). In addition, anaemia is caused by parasitic infections such as malaria and hookworm or chronic infections like tuberculosis (TB) and human immunodeficiency virus (HIV) (Ononge *et al.,* 2014). People with chronic diseases have the greatest risk of anaemia. Chronic diseases such as kidney disease can affect the body's ability to make erythrocytes. Patients with anaemia of chronic disease have mild to moderate anaemia that tends to correlate in severity with the underlying disease (Smith, 2000).

2.3 Inherited anaemic conditions associated with elevated HbF: Beta (β)-globin gene cluster mutation

The β-globin gene cluster locus is located at chromosome 11, which is composed of epsilon (ε), G-gamma (Gγ), A-gamma (Aγ), delta (δ), and β-globin genes as shown in Figure 2.2. Expression of the β-globin gene cluster locus is controlled by a single locus control region (LCR), the most important regulatory element in the locus, which is located upstream of the globin gene (Frenette and Atweh, 2007). Majority of the β-globin alterations are due to point mutations. To date, numerous mutations affect the β-globin gene cluster with most cases being single nucleotide substitutions, insertions, or short deletions (Thein, 2011). Examples of mutations associated with the β-globin gene cluster are HPFH, δβ-thalassaemia, and β-thalassaemia and its variants.

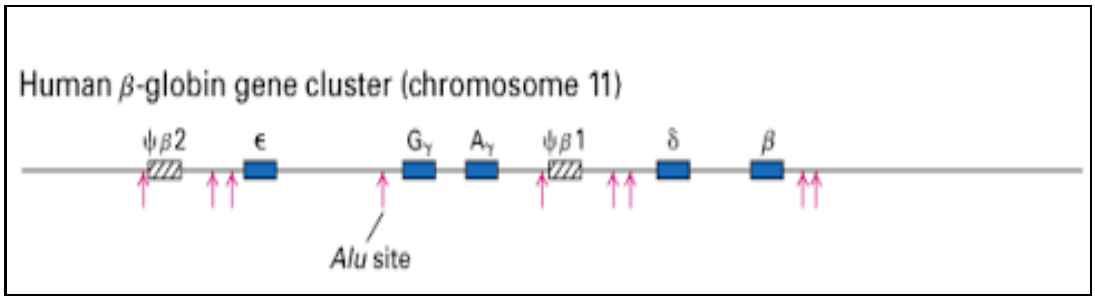

Figure 2.2 The β-globin gene consists of five functional genes (blue boxes) and two pseudogenes (diagonal lines). Adapted from Lodish *et al.,* 2000.

2.3.1 β-globin genes cluster mutation

In a routine clinical setting, HbA_2 level of more than 4.0% indicates β-thalassaemia or severe β-thalassaemia while in mild β-thalassaemia, HbA_2 level is usually more than 3.6% (Ministry of Health, 2009). Fucharoen *et al.* (2002) observed that HPFH-6 and δβ-thalassaemia Thai patients have high HbF (>5%) and normal HbA_2 (2.2%) levels. However, Koh *et al.* (2017), reported that subjects are considered to have α and β-globin gene cluster mutations when the HbF level is more than 1.0%, HbA_2 level is above 3.2%, and/or show the presence of other abnormal Hb. Therefore, in this study, samples with high HbF level (>1.0%) and HbA_2 level (>3.2%) were chosen for detection of β-globin gene cluster mutation while samples with high HbF level (>1.0%) and normal HbA_2 level (≤3.2%) were used for detection of β-globin gene cluster deletion.

β-thalassaemia and its variants are the examples of β-globin gene cluster mutation. β-thalassaemia is a heterogenous group of autosomal recessive disorders that cause a genetically deficient synthesis of β-globin chain in Hb (Thein, 2011). β-thalassaemia phenotype can also arise from structural β-globin chain variants. The common β-globin chain variants are haemoglobin C (HbC), haemoglobin S (HbS or

sickle cell), and haemoglobin E (HbE), while others are rare (Weatherall, 2001). Entries in the database of human Hb variants and thalassaemia mutations (HbVar) show that, at present, at least 228 entries are associated with the deletion mutation. β-thalassaemia is an inherited blood disorder that leads to ineffective erythropoiesis. Genetically, β-thalassaemia can be inherited from parents to a child and can be classified into three groups which are β-thalassaemia minor, β-thalassaemia intermedia, and β-thalassaemia major (Table 2.3).

Table 2.3 The classifications of β-thalassaemia

Classification	Genotype
β-thalassaemia minor	β/β^{+}, β/β°
β-thalassaemia intermedia	β^{+}/β^{+}, β^{+}/β°
β-thalassaemia major	$\beta^{\circ}/\beta^{\circ}$

Adapted from George, 2013. Abbreviations: (β): presence of β-globin chain synthesis; (β^{o}): absence of β-globin chain synthesis; (β^{+}): reduction of β-globin chain synthesis.

Between 1.5% and 7.0% of the world population are estimated to carry one of the genes that cause β-thalassaemia (Roth *et al.*, 2018). More than 800 mutation entries involving β-globin gene are recorded in the database of Hb variants and thalassaemia mutations (HbVar, 2013). Malaysia is a multi-ethnic country consisting of three major ethnic groups (Malays, Chinese, and Indians) and other minor ethnic groups, which include Orang Asli (Department of Statistics Malaysia, 2018). It was reported that 4.5% of the Malaysian population are carriers for β-thalassaemia, where it is commonly found among Malays and Chinese (George *et al.*, 2001), but rare in Indian ethnic (Tan *et al.*, 2006).

Each ethnic group in Malaysia possesses its own specific set of mutations (Tan *et al.*, 2006). According to Tan *et al.* (2004), β-thalassaemia with a structural

Hb variant, HbE is usually found among the Malays by 76.0%. Another study showed that Cd 26 (G–A) or HbE is the most observed β-globin mutation by 84.8% in transfusion-dependent β-thalassaemia Malay patients in Kelantan (Hanafi *et al.*, 2014). This is somehow anticipated since Kelantan geographically shares its border with the Narathiwat Province of Thailand. The prevalence of HbE carriers among Thais is from 10.0% to 60.0% depending on the region and the lowest prevalence rate of thalassaemia and HbE was in the Southern Thailand region (Nuinoon *et al.*, 2014). Furthermore, interracial marriages between the Southern Thais and Kelantanese are common and they have been crisscrossing the border since a long time ago. Indeed, the β-chain mutations among the Kelantanese population are almost equivalent to those in Southern Thailand (Hanafi *et al.*, 2014).

HbE is the most common Hb variant that occurs in Southeast Asia. HbE is a type of structural defect that is caused by a point mutation of the β-globin gene. It is caused by a base substitution at codon 26 of the β-globin gene, GAG–AAG, which results in the substitution of lysine for glutamic acid (Fucharoen and Weatherall, 2012). HbE can be characterised as HbE trait (heterozygous), HbE disease (homozygous), and a variety of compound heterozygous states such as HbE/β-thalassaemia and HbE/sickle cell disease (Bachir and Galacteros, 2004). When β-thalassaemia is co-inherited with HbE (HbE/β°-thalassaemia), these individuals may have severe anaemia and require frequent blood transfusions (Wahed and Dasgupta, 2015). HbE occurs at high frequency in many countries in Asia. HbE is widely distributed in Bangladesh, East India, Southeast Asia, and with high prevalence in the borders of Thailand, Laos, and Cambodia, in which these countries are known as the ‘HbE triangle’. In these areas, its prevalence can reach up to 50%–60%. In

Malaysia, HbE is common among Malays with a carrier rate of 5%. Moreover, HbE has also been identified to be high among the Orang Asli in Peninsular Malaysia (George, 2013). This is supported by a study conducted on Senoi group of Orang Asli, where HbE is the most common mutation with a rate of 18.6% (Koh *et al.*, 2017). Globally, individuals with HbE/β-thalassaemia compose almost 50% of those affected with thalassaemia (Olivieri *et al.*, 2011).

2.3.2 β-globin genes cluster deletion

According to Galanello (2013), deletions affecting the β-globin gene cluster result in the removal of the δ and/or β genes but leave one or both γ genes intact. This leads to the occurrence of β-thalassaemia, HPFH, δβ-thalassaemia, Hb Lepore, and γδβ-thalassaemia. Deletions that remove all or part of the β-globin gene are classified as β-thalassaemia deletions. Meanwhile, deletions that remove the δ- and β-globin genes are termed δβ-thalassaemia, and this is often associated with moderate increases in HbF. This deletion causes the expression of both γ-globin genes at high level to compensate for the absence of β-globin production. The clinical symptoms of δβ-thalassaemia heterozygous have similar presentations to thalassaemia minor with hypochromic microcytic red cells but normal levels of HbA_2. Homozygotes for δβ-thalassaemia or compound heterozygotes with β-thalassaemia are not common but have been reported to have clinical phenotypes ranging from mild anaemia to thalassaemia major (Thein, 2011).

γδβ-Thalassaemia is characterised by a heterogeneous array of large deletions of the β-globin gene complex of sizes ~100 kb or greater, which include all parts of the β-globin gene complex (Sabath, 2017). These deletions prevent the

transcriptional activation of any of the cis-linked β-globin genes at any developmental stage (Bender *et al.*, 2000). Driscoll *et al.* (1989), reported a case of a patient with γδβ-thalassaemia affected by a de novo deletion on the β-globin gene chromosome involving ~30 kb of sequences. This deletion extends from 9.5 to 39 kb of the ε-globin gene. γδβ-thalassaemia is lethal in the homozygous state. However, in the heterozygous state, the disease is transient while manifesting as moderately severe microcytic anaemia in newborns (Thein, 2011).

Meanwhile, HPFH involves large deletions that remove the adult δ and β genes but leave the paired foetal genes (G gamma and A gamma) intact (Collins *et al.*, 1987). Incomplete or failed switching of the HbF to HbA causes a high level of HbF expression in adults (Galanello, 2013). This results in the expression of cis-linked foetal globin genes in adult erythroid cells which leads to higher HbF level in adult life (Bank, 2006). HPFH is commonly seen in the African population (Ngo *et al.*, 2012). These deletions result in elevated HbF level that is significantly higher than that observed in β- or δβ-thalassaemia. HPFH heterozygotes have normal red blood cell indices but higher HbF compared to δβ-thalassaemia patients (Rochette *et al.*, 1994). Individuals with HPFH heterozygotes are asymptomatic (Patel *et al.*, 2015). HPFH homozygotes are clinically normal, and their Hb level may be increased with mildly hypochromic microcytic red cells (Thein, 2011).

Hence, HPFH and δβ-thalassaemia are characterised by the persistent expression of γ-globin genes in adults, and also the presence of the hypochromic and microcytic erythrocytes (Weatherall, 2001). Homozygotes for δβ-thalassaemia as well as compound heterozygotes for δβ-thalassaemia with β-thalassaemia usually

lead to a clinical phenotype of thalassaemia intermedia or major. Although HPFH homozygotes are clinically asymptomatic, compound heterozygotes for HPFH with β-thalassaemia express similar phenotype as in those with thalassaemia intermedia (Panyasai *et al.,* 2004). Individuals with HPFH/β-thalassaemia may have mild anaemia but are clinically asymptomatic (Thein, 2011).

Another type of deletion for β-globin gene cluster is caused by unequal crossover between the linked and partially homologous δ- and β-globin genes, leading to the development of a fusion of δβ-globin gene, also known as the Hb Lepore. Hb Lepore may be easily detected by the electrophoretic method. It constitutes 6% to 15% of the total Hb, with HbA_2 level being normal or discreetly reduced. The haematological profiles from all these variants show a moderate increase of the HbF levels ranging between 2% and 4% while a combination of Hb Lepore and HbA_2 has been observed at the range of 10% to 13% (Guo *et al.*, 2015).

In Asia, approximately 12 types of deletion in the β-globin gene cluster have been defined within the Southeast Asian populations. This includes the β°-thalassaemia 105 bp deletion, the 619 bp deletion, the 3.5 kb deletion, the novel Gγ(Aγδβ)°-thalassaemia, Siriraj J deletion (~118 kb deletion), the Southeast Asian (SEA) deletion (~27 kb deletion), the Filipino deletion (~45 kb deletion), Hb Lepore (~7.4 kb deletion), the Thai (δβ)°-thalassaemia deletion (~12.5 kb deletion), the Chinese Gγ(Aγδβ)°-thalassaemia (~100 kb deletion), and the Asian Indian deletion–inversion Gγ(Aγδβ)°-thalassaemia, as well as HPFH-6 and Vietnamese HPFH (known as HPFH-7 in HbVar) deletions (Tritipsombut *et al.*, 2012). In Asian Indian β-thalassaemia patients, almost 30% of the chromosomes have a 619 bp deletion.

This mutation affects the 3′ end of the large IVS II, the entire exon 3, and a short section of DNA 3′ to the β-globin gene (Thein *et al.*, 2009).

2.4 Acquired anaemic conditions associated with elevated HbF

Several acquired conditions are associated with elevated HbF. These include drug-induced anaemia, aplastic anaemia, pernicious anaemia, sideroblastic anaemia, vitamin B12/folate deficiency, autoimmune diseases, and leukaemia (Mosca *et al.*, 2009). Anaemia in pregnancy is classified into two, which are pathological anaemia and physiological anaemia. Pathological anaemia is further subclassified into haemorrhagic and deficiency anaemia such as vitamin B12 deficiency, iron deficiency, protein deficiency, and folic acid deficiency. Meanwhile, physiological anaemia is due to the joint effect of haemodilution and negative iron balance. Iron deficiency secondary is the main factor of anaemia in pregnant women because of chronic menstruation and inadequate dietary intake (Sabina *et al.*, 2015).

According to the National Organization for Rare Disorders (NORD, 2004), rare disorders such as autoimmune haemolytic anaemia is one of the acquired anaemia. It is characterised by the premature destruction (haemolysis) of erythrocytes. Acquired autoimmune diseases occur when the body's natural defences such as lymphocytes and antibodies destroy its own healthy tissues. Acquired autoimmune haemolytic anaemia is a disorder that occurs in individuals who previously had a normal red blood cell system.

Chemotherapy, radiation, and some drugs used to treat leukaemia may cause aplastic anaemia because some cancer therapies prevent the bone marrow from

synthesising new, healthy blood cells. Leukocyte count is lowered first, followed by platelet count and erythrocyte count. Anaemia due to cancer treatments may be reversible after the treatment is completed (Deborah *et al.*, 2016). Leukaemia itself can also cause anaemia. In leukaemia, leukocytes multiply rapidly, resulting in difficulty for erythrocytes to develop. Other than that, cancer treatments may cause decreased appetite, nausea, and vomiting. This often makes it difficult for patients to eat a nutritious, iron-rich diet, and thus may lead to iron-deficiency anaemia (Weatherspoon, 2016).

2.5 Single nucleotide polymorphisms (SNPs) associated with high foetal haemoglobin (HbF)

Single nucleotide polymorphisms (SNPs) occur when a single nucleotide base pair is substituted after the body makes new cells. SNPs are responsible for the diversity among individuals, genome evolution, interindividual differences in drug response as well as complex and common disease such as diabetes, obesity, hypertension, and psychiatric disorders. SNPs are the most common type of genetic variation among people. The significance of the variants may change the downstream outcomes. In cells, the variant may influence promoter activity (gene expression), messenger RNA (mRNA) stability, and subcellular localisation of mRNAs and/or proteins and hence, may produce the disease (Shastry, 2009). Quantitative trait loci (QTLs) are regions of DNA which are associated with a particular phenotype attributed to polygenic effects such as the product of two or more genes and their environment. QTLs are mapped by identifying the correlation of SNPs with an observed trait (Grisel and Crabbe, 1995).

SNPs can be used as a predictive marker for diseases. SNPs are associated with cancerous or non-cancerous diseases as well as anaemia. Therefore, SNPs offer valuable markers for identifying genes responsible for susceptibility to common diseases (Nassiri *et al.*, 2013). Some genome-wide association studies have reported that there are at least three major loci that play a major role in increased HbF level (Rujito *et al.*, 2016). Previous studies have shown that SNPs in certain QTLs are actively associated with the range of HbF in anaemia as shown in Table 2.4.

Table 2.4 Example of QTLs that are associated with HbF level

Gene	Location	Variant
BCL11A	2p16	rs1186868
		rs1427407
		rs7557939
HBS1L and *MYB* intergenic polymorphisms (*HMIP*)	6q23	rs9389268
		rs9376090
		rs4895441
		rs9402685
		rs7776054
		rs9399137
HBG2-Xmn1	11p15	rs7482144

Adapted from Uda *et al.*, 2018; Lai *et al.*, 2016; Fong *et al.*, 2015; Sheehan, 2013.

The QTL associated with elevated HbF is *BCL11A*, located in the 2p16 region of chromosome 2. The percentage of HbF variations attributed to *BCL11A* QTLs is 15%. *BCL11A* is more frequent in patients with β-thalassaemia intermedia than in β-thalassaemia major (Sokolova, 2019). There is a strong association between genetic variants in the *BCL11A* gene and HbF levels in numerous populations (Sankaran *et al.*, 2008). *BCL11A* gene has been suggested as a direct regulator of HbF level (Basak and Sankaran 2016). A SNP in *BCL11A* was associated with HbF level in Sardinians with β-thalassaemia and in African Americans with sickle cell disease (Uda *et al.*, 2008). Two SNPs in BCL11A genes are associated with increasing HbF

levels in patients with HbE/β-thal patients in Indonesian population (Rujito *et al*., 2016).

Polymorphisms associated with F-cell level in an intergenic region between the genes *HBS1L* and *MYB*, also known as *HMIP*, have been identified (Thein *et al.*, 2007). A previous study showed that there is an association signal of HbF level in the *HBS1L–MYB* intergenic region in a large non-anaemic Sardinian cohort (Menzel *et al.*, 2007). The *HBS1L–MYB* is located on chromosome 6 between the *HBS1L* gene (codes for elongation factors and regulates multiple cellular processes) and the *MYB* gene (encodes transcription factors and participates in ontogenesis and erythropoiesis). This QTL codes for factors that participate in the erythroid maturation pathway. Wahlberg *et al.* (2009), showed a correlation between 6q23 QTL and HbF level in Indian β-thalassaemia patients. It is yet to be determined whether this correlation is secondary to a direct or indirect effect. Craig *et al.* (1996), studied Indian female population with β-thalassaemia and found that those with homozygous 6q23 QTL have a higher HbF concentration (24% compared to 10%). The same correlation was noticed among the healthy population in this study (3%–1%).

The QTL on chromosome 8, 8q region is suggested to influence HbF levels by encoding transcriptional factors, which bind to the *XmnI* site. The SNP rs7482144 at 158 bp 5' of *HBG2* gene on chromosome 11p15.4 is associated with elevated HbF in thalassaemia and sickle cell disease (Gilman and Huisman, 1985). QTLs for HbF level have been defined on chromosomes 6q23, 8q, and Xp22 (Thein *et al.*, 2007). The biological effect of the QTLs on HbF expression includes two plausible

mechanisms. First, the direct effect on *HBG* gene transcription activation or repression which increases or decreases the amount of HbF. Second, the alteration of the kinetics of erythroid maturation and differentiation, mimicking a situation encountered in stress erythropoiesis, resulting in accelerated erythropoiesis with the release of more erythroid progenitors that synthesise predominantly HbF level (Thein *et al.*, 2009).

2.5.1 Methods for single nucleotide polymorphisms (SNPs) genotyping

Single nucleotide polymorphisms (SNPs) are common DNA sequence variations that occur at single bases within a genome. The increase of interest in SNPs is reflected by the furious development of SNPs genotyping methods including array-based hybridisation, allele specific PCR, restriction fragment length polymorphism (RFLP) techniques as well as sequencing. PCR allelic discrimination technologies have broad applications in the detection of SNPs in genetics and genomics. Two different PCR-based allelic discrimination techniques, namely Custom TaqMan SNPs genotyping and high-resolution melting (HRM) assays, have been developed. The allelic specificity of TaqMan assay is provided by two probes, one labelled with FAM dye and the other with VIC dye (Kamau *et al.*, 2012).

Ghomi *et al.* (2014), reported that TaqMan assays have proven to be more sensitive and more reliable than HRM assays. Moreover, the TaqMan SNPs assays have been further improved by developing a rapid and straighforward protocol that includes crude leaf extraction for RNA template preparations. For a project involving a small number of SNPs and a large population, the TaqMan assay is the preferred

technology as it has high throughput and is highly accurate, precise, time efficient, and cost effective (Shen *et al.*, 2009).

One of the most efficient ways to analyse genotypes is through the TaqMan allelic discrimination, which has frequently been used to characterise SNPs. TaqMan SNPs genotyping assays use 5′ nuclease assay chemistry to detect specific SNPs, multinucleotide polymorphism and insertion or deletion alleles (Gaedigk *et al.*, 2015). A relatively small region flanking the target single nucleotide variant is amplified using locus-specific primers and alleles are detected using two TaqMan probes with conjugated minor groove binder (MGB) labelled with VIC dye or FAM dye (Kamau *et al.*, 2012).

2.6 Haematological analysis

Hb level, mean corpuscular volume (MCV), and mean corpuscular haemoglobin (MCH) are the parameters used as a guideline for the first-line screening of anaemia (Alquaiz *et al.*, 2012). A patient who possesses red cell indices are below the normal range is considered to have anaemia as shown in Table 2.5.

Table 2.5 Age and gender related red cell indices for children and adults under normal condition

		Hb (g/dL)	MCV (fL)	MCH (pg)
Male	1-12	120-135	79.0-84.0	27.4-28.7
	12-18	130-160	85.0-87.0	29.1-29.9
	18 and above	130-160	89.0-97.0	30.5-37.8
Female	1-12	120-135	79.0-84.0	27.4-28.7
	12-18	120-160	86.0-88.0	29.4-30.0
	18 and above	120-150	90.0-95.6	30.6-32.6

Adapted from Yip *et al.*, 1984. Abbreviations: Based on the US second National Health and Nutrition Examination Survey (NHANES II) after excluding those with abnormal tests related to iron.