THE DEVELOPMENT OF PERFORMANCE ASSESSMENT TOOLS TO IMPROVE THE QUALITY OF DIAGNOSIS AND MONITORING OF HAEMOGLOBINOPATHY DISORDERS

B J DE LA SALLE

PhD 2018

The Development of Performance Assessment Tools to Improve the Quality of Diagnosis and Monitoring of Haemoglobinopathy Disorders

BARBARA JOAN DE LA SALLE

A thesis submitted in partial fulfilment of the requirements of the Manchester Metropolitan University for the degree of Doctor of Philosophy

School of Health Sciences,
the Manchester Metropolitan University
in collaboration with the West Hertfordshire Hospitals NHS Trust operating
UK NEQAS Haematology and Transfusion

2018

For Joan Elinor Harrington

Abstract

The haemoglobinopathies are complex inherited disorders of haemoglobin synthesis. Although carriers are generally asymptomatic, homozygous or compound heterozygous inheritance of globin gene mutations may result in severe conditions, identified as a major burden of disease worldwide. The main means of control of the conditions is by carrier screening using specialist laboratory testing, allowing individuals an informed reproductive choice. UK NEQAS Haematology provides a comprehensive external quality assessment service for the haemoglobinopathies and is in a unique position to provide an overview of laboratory performance. The increased emphasis on laboratory quality and a shift in the profile of the participating laboratories from mainly UK-based to more than half non-British Isles in the past decade are drivers for continued improvement in performance assessment methods. The project looked at aspects of performance of both UK NEQAS and the participants. A review of historical quantitation data for haemoglobin (Hb) A_2, the standard biomarker for beta thalassaemia carrier identification, showed there remains a need for standardisation between methods of analysis. The survey material distributed for testing was examined to exclude commutability and stability as confounding factors in performance assessment. Changes to Hb A_2 performance scoring have been suggested including the identification of one-off analytical EQA errors that could represent misdiagnosis of a beta thalassaemia carrier. In newborn screening, a difference was seen in the sensitivity of different analysers at the low levels of adult haemoglobin (Hb A) seen in a newborn infant with homozygous beta thalassaemia. A significant difference in performance in result interpretation was seen between English NHS laboratories and those outside the British Isles. The scheme should extend performance management in interpretation to all laboratories including the non-British Isles group, to ensure that the scheme remains credible to participants and fulfils its remit to improve performance wherever patients are tested. This will however require the scheme to develop the EQA data capture methods better to reflect current laboratory practice.

Acknowledgements

My thanks go to my supervisory team: Professor Bill Gilmore, Dr Barbara Wild, Professor Keith Hyde and Dr Nessar Ahmed who took over as Director of Studies in the final stages of the work. His support has been an enormous source of encouragement.

I also give my wholehearted thanks to my colleagues at UK NEQAS Haematology for keeping the show on the road and giving me the space to work, especially during the final year. I thank Isabella De-Rosa for protecting Fridays in my diary for five years, Gulala Karim for her assistance in preparing dried blood spot cards and keeping them safe and Anne Mahon for her instruction in the preparation and packing of specimens.

I am grateful to Paul McTaggart, Vasileos Rapanakis and Gilbert Sakyi for their patient assistance and support in the extraction of data no matter how many times I asked for help. I also thank Paul McTaggart for his instruction on the use of the STATA package and for leaving me his books.

My gratitude is owed to Dr E J Parker-Williams for the support that only he will remember, to Dr Mitchell Lewis for suggesting I did the work in the first place, to Dr Megan Rowley for her encouragement and to Professor Caroline Doré for statistical advice.

The work was undertaken in collaboration with and on the premises of the West Hertfordshire Hospitals NHS Trust, host to UK NEQAS Haematology and Transfusion and with the support of UK NEQAS Haematology participants, who returned extra results, replied to queries and generally are the best people in the world to work for.

In my personal life, I owes thanks to my family and my children for putting up with whatever I do and to my late father for raising me to believe a girl could do anything.

Finally I remember my lifelong friend Pam Smith for her interest and her affection despite the devastating illness she was facing. I miss her every day.

Contents

List of Figures

List of Tables

List of Equations

Abbreviations

ALTM	All laboratory trimmed mean
ANOVA	Analysis of variance
ANR	Hb A not recorded
APS	Acceptable performance specifications
BP	British Pharmacopoeia
BS	British Standard
BV	Biological Variation
° C	Degrees Celsius
χ^2	Chi squared
CAP	College of American Pathologists
CLIA	Clinical Laboratory Improvement Amendments
CLSI	Clinical & Laboratory Standards Institute
CO_2	Carbon dioxide
CRM	Certified reference material
CPD	Citrate-phospate-dextrose
CV	Coefficient of variation
CZE	Capillary zone electrophoresis
D%	Percent bias
DBS	Dried blood spot
DEAE	Diethylaminoethyl
DI	Deviation Index
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
ENERCA	European Network for Rare and Congenital Anaemias
EQA	External Quality Assessment
Est. CV	Estimated coefficient of variation
Est. SD	Estimated standard deviation
FBC	Full blood count

FDA	Food and Drugs Administration
FOQ	Family Origin Questionnaire
GCV	Geometric coefficient of variation
H	Hydrogen
Hb	Haemoglobin
Hb A	Haemoglobin A or Adult Haemoglobin
Hb A_2	Haemoglobin $A_{2,}$ a minor adult haemoglobin fraction
HbA1c	Haemoglobin A1c or glycated haemoglobin
Hb F	Haemoglobin F or Fetal Haemoglobin
Hb S	Haemoglobin S or Sickle Haemoglobin
HbsAg	Hepatitis B surface antigen
Hct	Haematocrit
HCV	Hepatitis C virus
HIV	Human immunodeficiency virus
HPFH	Hereditary Persistence of Fetal Haemoglobin
HLA	Human leucocyte antigen
HPLC	High performance liquid chromatography
HSCT	Haemopoietic stem cell transplantation
HSD	Historic standard deviation
ICSH	International Council for Standardization in Haematology
IEF	Isoelectric focusing
IFCC	International Federation of Clinical Chemistry and Laboratory Medicine
IQC	Internal quality control
IQR	Interquartile range
ISO	International Standards Organisation
K	Potassium
kDa	KiloDalton
KW	Kruskal-Wallis
L	Litre

LCRA	Locus control region A
LCRB	Locus control region B
LOD	Limit of detection
μg	Microgramme
μL	Microlitre
MCADD	Medium-chain acyl-coenzyme A dehydrogenase deficiency
MCH	Mean cell haemoglobin
MetHb	Methaemoglobin
mg	Milligramme
mL	Millilitre
mm	Millimetre
mRNA	Messenger ribonucleic acid
MTM	Method trimmed mean
NHSBT	National Health Service Blood and Transplant
NHS STP	National Health Service Sickle and Thalassaemia Screening Programme
nm	Nanometre
NS	Non-specified
O_2	Oxygen
p50	Partial pressure
PBS	Phosphate buffered saline
pg	Picogramme
PKU	Phenylketonuria
PRN	Participant Reference Number
PT	Proficiency testing
PUP	Persistent unsatisfactory performance
RBC	Red blood cell
RCPAQAP	Royal College of Pathologists in Australasia Quality Assurance Programmes
RoI	Republic of Ireland

rpm	Revolutions per minute
RSD	Robust standard deviation
SD	Standard deviation
SMTM	Submethod trimmed mean
TE	Total error
TIF	Thalassaemia International Federation
TMS	Tandem mass spectrometry
TQM	Total quality management
Tx	Transposition error
UI	Unable to identify
UK NEQAS	UK National External Quality Assessment Scheme
UP	Unsatisfactory performance
WHO	World Health Organisation
w/v	Weight by volume

Chapter 1: Introduction and Overview

1.1 Introduction

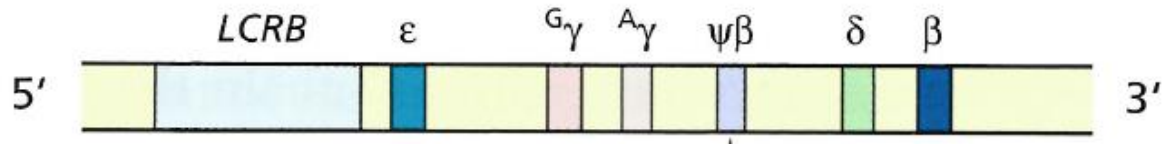
This project examines aspects of the performance assessment tools used by the UK National External Quality Assessment Scheme for Haematology for the external quality assessment (EQA) of the laboratory diagnosis of the haemoglobinopathies. The purpose is to improve the feedback given to participating laboratories and indirectly the quality of the service they provide.

1.2 Haemoglobin structure and function

Haemoglobin (Hb) is a complex, globular, protein tetramer (molecular weight 64-64.5 kDa) comprising two pairs of unlike globin chains, each with a haem group attached. In post-natal life, all normal human haemoglobins contain two alpha and two non-alpha globin chains. Alpha globin chain synthesis is directed by two alpha globin genes (α1 and α2) in the alpha globin gene cluster located on chromosome 16. Synthesis of the non-alpha globin chains is directed by the beta globin gene cluster located on chromosome 11. The beta globin gene cluster contains one beta, one delta and two gamma globin genes (Gγ and Aγ) that code for alternative gamma globin chains with either glycine or alanine respectively at position 136. The arrangement of the globin gene clusters is shown in Figure 1.1. Both globin gene clusters contain pseudogenes (segments of DNA that resemble genes in structure but do not result in the production of a protein) and genes that are only expressed early in intrauterine life.

Haemoglobin carries oxygen (O_2) from the lungs to the tissues, binding one O_2 molecule reversibly to each of the four haem groups. It also transports approximately 10% of the carbon dioxide (CO_2) produced in the body back to the lungs.

Chromosome 11: beta globin gene cluster



Chromosome 16: alpha globin gene cluster

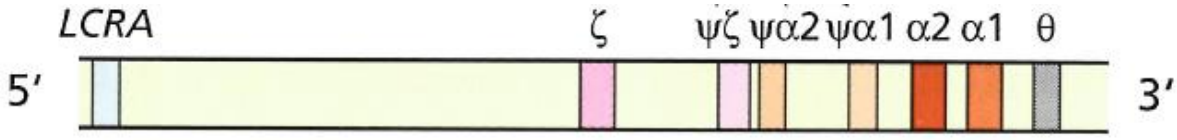


Figure 1.1. Diagram of the alpha and beta globin gene clusters.

The beta globin gene cluster (chromosome 11) shows the locus control region B (LCRB), the ε, Gγ, Aγ, δ and β globin genes with the pseudogene $\psi\beta$. The ε globin gene is expressed early in intrauterine life. The alpha globin gene cluster (chromosome 16) shows the locus control region A (LCRA), the ζ, $\alpha 2$ and $\alpha 1$ globin genes and the pseudogenes $\psi\zeta$, $\psi\alpha 2$ and $\psi\alpha 1$. The ζ globin gene is expressed in early intrauterine life and the function of the θ gene in the alpha globin gene cluster is unknown.

Adapted from: Variant Haemoglobins: A Guide to Identification (Bain, 2010)

The quaternary structure of haemoglobin, shown in Figure 1.2 for Hb A, facilitates the rapid uptake of O_2 (oxygenation) by Hb in the lungs and release (deoxygenation) in the tissues. The oxygen dissociation curve of Hb has a characteristic sigmoid shape, reflecting the alteration of the affinity of the molecule for O_2 according to the partial pressure (p50) of O_2 in the surrounding tissues.

As soon as the first O_2 molecule binds to deoxygenated Hb, the quaternary structure of the molecule undergoes a conformational change that facilitates the binding of the subsequent three O_2 molecules more rapidly than the first (Baldwin, 1980; Perutz, 1980). In the tissues, the reverse happens as the p50 in the surrounding tissues falls, allowing the rapid release of O_2 where it is needed. The rapid deoxygenation of the molecule is supported by stabilisation of the deoxygenated form by the reversible binding of CO_2 to the N-terminal of the alpha globin chains (Kilmartin, 1969). The affinity of Hb for O_2 is also sensitive to changes in pH (Bohr, 1904), with the deoxygenated form of Hb stabilised by hydrogen ions (H^+) in the tissues (Perutz, 1978).

The ability of Hb to bind O_2 reversibly depends upon the iron in the haem group remaining in the reduced or ferrous form, as the oxidised or ferric form combines irreversibly with O_2 forming non-functional methaemoglobin (MetHb). The iron is protected against oxidation by the location of the haem group in a hydrophobic pocket formed by the tertiary folding of the globin chain, which excludes the water necessary for oxidation. The exterior of the Hb molecule is hydrophilic, however, making the molecule soluble.

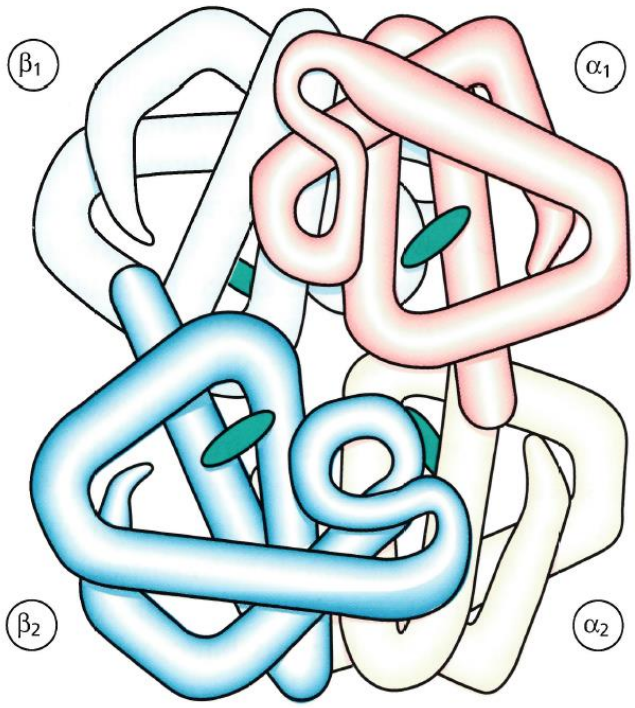


Figure 1.2 Diagram of the quaternary structure of adult haemoglobin (Hb A)

The Hb A molecule comprises two α and two β globin chains arranged in two $\alpha\beta$ dimers. The haem groups are shown in green in the hydrophobic interior of the molecule.

Adapted from: Variant Haemoglobins: A Guide to Identification (Bain, 2010)

1.3 Haemoglobin patterns in infancy and adulthood

Several Hb types are synthesised in the normal individual and the Hb pattern varies at different stages of development. In the first few weeks of gestation embryonic zeta (ζ) and epsilon (ε) globin chains, as well as alpha and gamma chains, are synthesised by primitive erythroblasts in the yolk sac, with the formation of the transient, early intrauterine Hbs Gower 1 ($\zeta_2\varepsilon_2$), Gower 2 ($\alpha_2\varepsilon_2$) (Huehns, 1961) and Portland 1 ($\zeta_2\gamma_2$) (Hecht, 1968). The embryonic Hbs have a particularly high oxygen affinity (Fantoni, 1981). The liver and spleen become the most significant sites of haemopoiesis from six weeks until the sixth to seventh month of fetal life, when the bone marrow takes over, although haemopoiesis continues in the liver and spleen until about two weeks after birth (Kelemen, 1981) as shown in Figure 1.3.

In the normal adult, the major Hb present is haemoglobin A (Hb A), with the remainder of the Hb made up of the minor adult Hb fraction, Hb A_2 (Hb A_2) and traces of fetal Hb (Hb F). First described by Linus Pauling (Pauling, 1949), Hb A comprises two alpha and two beta globin chains ($\alpha_2\beta_2$), of 141 and 146 amino acids respectively. The synthesis of Hb A starts early in the second trimester of pregnancy and forms approximately 20% of the total Hb at birth; thereafter, the proportion of Hb A increases rapidly, reaching adult levels of more than 95% of total Hb within 1 to 2 years of birth, as shown in Figure 1.4.

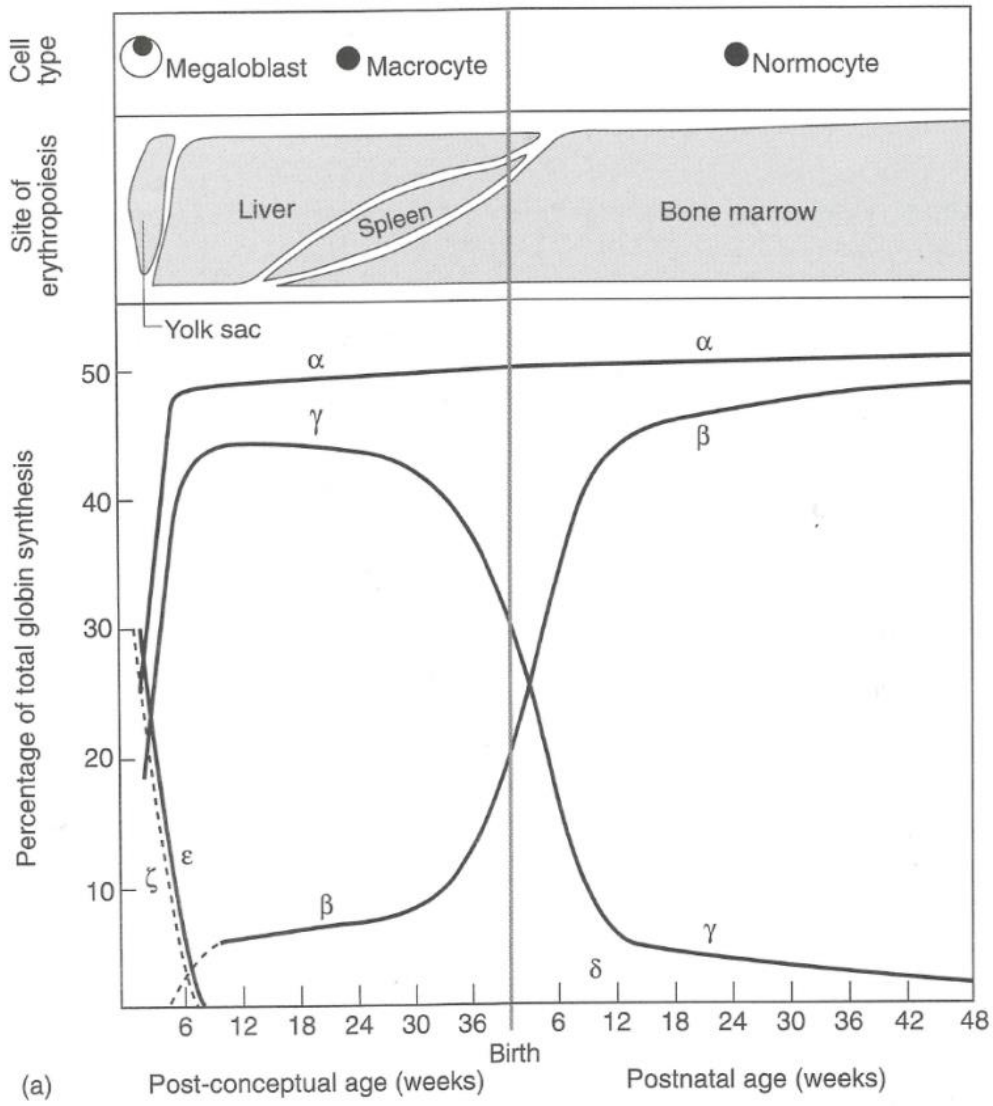


Figure 1.3 Diagram of the changes in sites of erythropoiesis during embryonic life and infancy showing the changes in the proportions of different globin chain synthesised and the erythroid cell type

Adapted from: The Thalassaemia Syndromes (Weatherall and Clegg, 2001)

Fetal Hb comprises two alpha and two gamma chains ($\alpha_2\gamma_2$) and has a higher affinity for O_2 than Hb A, facilitating the transfer of O_2 across the placenta from the maternal to the fetal circulatory systems. At birth, Hb F concentration ranges from 60 to 95% at full term gestation, falling to near adult levels by one year of age (Colombo, 1976). The change from Hb F to Hb A production occurs as a result of switching from the gamma to the beta globin genes during fetal life; in the adult, Hb F forms less than 1% of the total Hb present.

Haemoglobin A_2 (Hb A_2) has two alpha and two delta globin chains ($\alpha_2\delta_2$) (Kunkel and Wallenius, 1955). The mean Hb A_2 level at birth is between zero and 1% but increases during the first year of life to form approximately 2.0 to 3.5% of the total Hb (Steinberg and Adams, 1991; Bain, 2010; Weatherall and Clegg, 2001; Van Delft, 2009). Haemoglobin A_2 is functionally similar to Hb A but its physiological significance, if any, is unclear. The low level of Hb A_2 in the red blood cell is due to decreased transcription of the delta globin gene in comparison with the beta globin gene, as a result of differences in the promoter region (Steinberg and Adams, 1991). The low transcription rate and the lack of apparent physiological significance of Hb A_2 may indicate that the delta globin gene is destined to become another pseudogene. The delta globin chain differs from the beta chain by just ten residues between positions 22 and 115.

Haemoglobin A_2 is found in apes and New World monkeys, as well as Man; the delta globin gene is found in Old World monkeys, although not expressed. Other animals are known to have more than one normal Hb type in adulthood, e.g. horses have two, present in equal amounts, and salmon have two with different physiological attributes, which may reflect the salmon’s occupation of both fresh and saltwater environments (Lehmann and Huntsman, 1974).

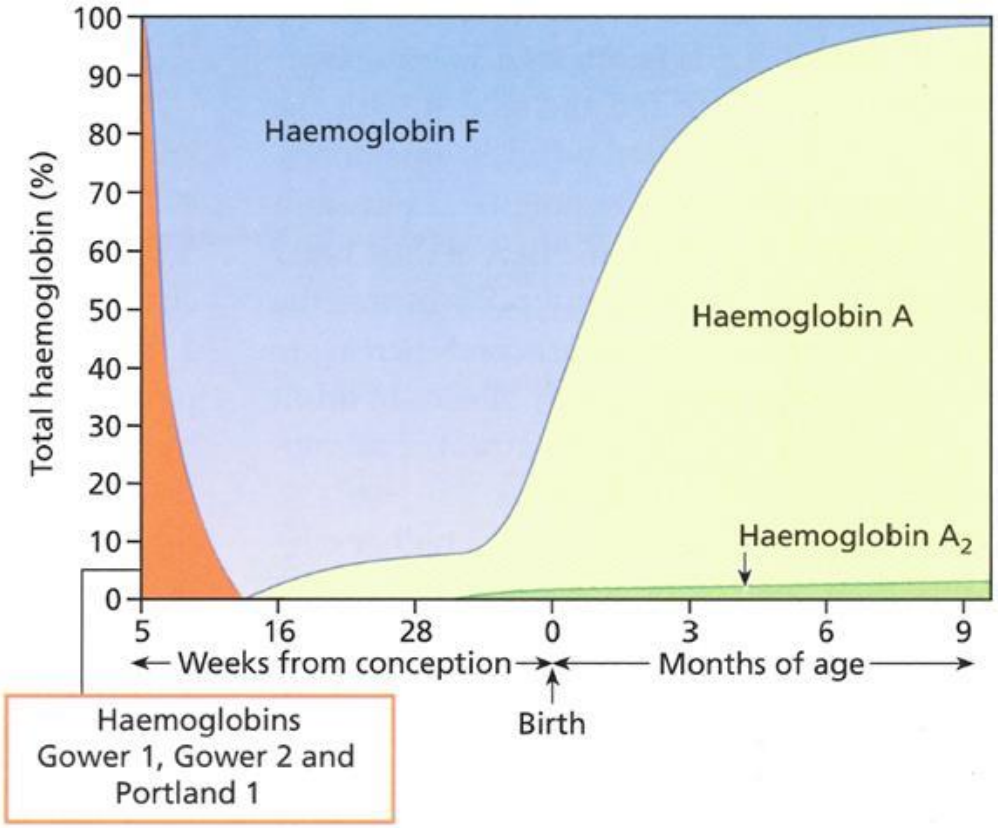


Figure 1.4 Change in the proportions of different haemoglobin types in embryonic life and infancy

Adapted from: Haemoglobinopathy Diagnosis (Bain, 2006)

1.4 The molecular basis of the haemoglobinopathies

In 1910, J B Herrick observed the presence of 'peculiar elongated and sickle shaped' red cells in a medical student of Caribbean origin suffering from anaemia (Herrick, 1910). More cases of sickle cells were reported in African-Americans following Herrick's initial observations (Washburn, 1911; Cook, 1915; Mason, 1922) and studies in subsequent decades showed that there were two types of sickle cells: those present in the untreated blood of individuals with severe anaemia and those that could only be induced by the *in vitro* treatment of blood from asymptomatic relatives of the patients with the anaemia. The asymptomatic condition was called sickle cell trait and it was suggested that the anaemia and the trait were the homozygous and heterozygous state of the same condition (Beet, 1947; Neel 1947; Neel, 1949). Linus Pauling (Pauling, 1949) discovered that the blood of sickle cell carriers contained two Hbs of different electrophoretic mobility, Hb A and Hb S, whereas blood from individuals with sickle cell anaemia contained Hb S but no Hb A.

Mutations in the globin genes may arise as point mutations of the genetic code or from the deletion of a part of the gene. The known Hb variants and thalassaemia mutations are listed on the HbVar database, on the Globin Gene server (Hardison, 1998; Hardison, 2002; Patrinos, 2004; Giardine, 2007).

Globin gene mutations can be grouped into the following main categories:

1.4.1 **Mutations in the coding part of the gene that result in structurally abnormal Hb variants**. More than one thousand structural Hb variants have been described. The clinical significance of the mutation depends upon the nature of the amino-acid

substitution and its location. The majority have no impact; however, substitutions in the haem pocket, substitutions affecting the secondary or tertiary folding of the chain, substitutions affecting the quaternary interaction of the globin chains and substitutions affecting the molecule's solubility may have serious consequences on the molecule's physiological function.

1.4.2 **Mutations in the non-coding part of the globin gene causing defective gene expression** and failure to synthesise the globin chain in normal quantities, resulting in a group of conditions known collectively as the thalassaemia disorders. More than four hundred mutations have been described, most affecting the beta globin and alpha globin genes. Mutations of the gamma or delta globin genes have also been described but are of less clinical significance.

1.4.3 **Mutations that result in the production of a hybrid globin gene,** e.g. a hybrid delta- beta globin gene, as a result of crossing over of the genes. The resulting gene may be defective, with reduced expression and / or a structurally abnormal output.

1.4.4 **Mutations that cause a failure of the normal switch from Hb F to Hb A** in the neonatal period, causing a diverse group of conditions characterised by an increased Hb F synthesis in the adult, known collectively as Hereditary Persistence of Fetal Haemoglobin (HPFH).

Although a small number of dominant mutations exist, globin gene mutations are generally recessive and asymptomatic in the heterozygote. The disease state occurs in the homozygous or compound heterozygous condition and then only if the mutation is of clinical significance.

1.5. Clinically significant structural haemoglobin variants

The majority of structural Hb variants have no impact on the molecule's physicochemical characteristics, i.e. its solubility, stability or oxygen affinity (Wild and Bain, 2004). The most frequently encountered, clinically significant, structural Hb variants are Hb S, Hb C, Hb D^{Punjab}, Hb E and Hb O^{Arab}, all of which affect the non-alpha globin chain, and Hb Lepore, which arises from a cross-over between the beta and delta globin genes. Haemoglobin S in the homozygous state or in combination with Hb C, Hb D^{Punjab} or Hb O^{Arab}, and to a lesser extent Hb E or Hb Lepore, results in the sickling disorders, characterised by severe anaemia and varying pathophysiology depending on the exact genotype and other ameliorating factors. A severe sickling disorder (Hb S-beta thalassaemia) also occurs when Hb S is co-inherited with a beta thalassaemia mutation, which reduces or abrogates the production of any structurally normal beta globin.

Some structural mutations reduce the output from the gene and are described as a 'thalassaemic haemoglobinopathy'. The most common example is Hb E, which results from the substitution of glutamic acid for lysine at position twenty-six of the beta globin chain. As well as being structurally abnormal, the mutation provides an alternative splicing site, resulting in reduced output from the mutated globin gene. Another is Hb Constant Spring, an alpha globin mutation of the termination codon of the alpha-2 globin genes, producing an elongated and unstable alpha globin chain (Clegg, 1971).

1.6. The thalassaemia disorders

The thalassaemias are a group of disorders characterised by a reduction in the output of the affected globin gene because of a deletional mutation or a mutation

affecting gene transcription or the stability of the messenger RNA (mRNA). The severity of the thalassaemia mutation depends on the globin gene affected, the extent to which the gene output is reduced and, in the case of the alpha thalassaemia mutations, the number of genes affected.

Thalassaemia conditions may be described in terms of the clinical phenotype, a practice that arose initially because the molecular basis of the disease was unknown. The carrier or heterozygous state for thalassaemia may be described as thalassaemia minor, often qualified by which globin gene is affected, a largely asymptomatic condition only detectable by laboratory testing. The term thalassaemia minor is no longer widely used but would equate to a thalassemia carrier individual. Thalassaemia intermedia describes a condition associated with a range of genotypes that are symptomatic but not transfusion- dependent, except perhaps in conditions of physiological stress such as pregnancy. Thalassaemia major describes a severe, transfusion dependent condition, usually arising from homozygosity or compound heterozygosity for severe beta thalassaemia mutations. Each of these descriptions is broad and covers a spectrum of clinical conditions (Oliveri and Weatherall, 2001; Weatherall and Clegg, 2001).

1.7. Globin gene mutations and the burden of disease

The impact of a globin gene mutation on the burden of human disease depends on its clinical significance and the frequency of the mutation in a population. Clinically significant mutations that are restricted to members of the same family or an isolated geographical location may be diagnostically and therapeutically challenging but do not pose a significant public health risk.

Despite the severity of the disorders of Hb structure or synthesis in the homozygous or compound heterozygous disease states, they provide protection against *Plasmodium falciparum* malaria infection in carriers, hence the frequency of the most clinically significant mutations is highest in people of differing ethnic backgrounds originating from areas of the world where malaria either is or was endemic (Lehmann and Huntsman, 1974; Clegg and Weatherall, 1999; Williams, 2002; Williams, 2005a; Williams, 2005b; Piel, 2010; Weatherall, 2018).

Within Europe, the Hb disorders have an overall frequency of less than one per two thousand individuals, which classifies them as 'rare' diseases according to the European Union regulations (Regulation (EC) 141/2000); however, this classification reflects the low incidence of Hb disorders in the indigenous Northern European population. The distribution differs by region within Europe and is changing as a consequence of migration both within and from outside the continent (Modell 2008, Gulbis 2010; Piel, 2016). Rund (2005) estimated that 1.67% of the world's population overall carries a gene for thalassaemia. The prevalence of thalassaemia is highly variable by region with the highest carrier frequency seen in Cyprus (fourteen to eighteen percent); in India the frequency overall is approximately three percent but may be up to forty percent in some locations (Bain, 2006). The Thalassaemia International Federation (TIF) estimates that there may be 200,000 transfusion dependent thalassaemia sufferers worldwide with up to 20,000 in Europe (Capellini, 2008; Galanello, 2003; Galanello, 2010). Twenty-three thousand new beta thalassaemia major infants are estimated to be born each year (Weatherall, 2010).

Migration from 'thalassaemic' regions means that non-endemic countries, such as the UK, now recognise thalassaemia as a public health problem (Zeuner, 1999:

Davies, 2000; Weatherall, 2010; Piel, 2016). Henderson (Henderson 2009) noted that sixty-eight different thalassaemia mutations had been reported in the UK and that there are more beta and alpha thalassaemia mutations in the UK than in sixty other countries with published data in this area, reflecting the racially heterogeneous UK population. Some beta thalassaemia mutations are found in the native UK population but these are infrequent (Hall 1992).

1.8. Population screening for the haemoglobin disorders

Population screening and definitive diagnosis for the Hb disorders are different (Streetly and Holland, 2009). Screening requires high throughput, reliable and cost effective procedures to identify a defined range of disorders. Screening is not designed to detect every disorder and only provides a presumptive diagnosis to enable appropriate intervention.

Population screening for the Hb disorders may be undertaken at several opportune stages in life (Giordano, 2009; Streetly and Holland, 2009): at birth, during mandatory schooling, pre-marriage (pre-conception) and in early pregnancy, as well as upon request. Screening may cause unnecessary anxiety and should only be implemented where there is an identified benefit to the health of the individual or the population as a whole through the implementation of an effective medical or social intervention (Wald and Leck, 2000).

In the UK, the Standing Medical Advisory Committee advised in 1993 that universal newborn and antenatal screening was provided in those areas where more than 15% of the local population was of an ethnic background at risk of a severe haemoglobinopathy, with selective screening on the basis of ethnicity in other areas. It was recognised that the provision of screening was patchy and the

concept of a linked newborn and antenatal screening programme, offering universal screening for all newborn infants at birth and all pregnant women in early pregnancy, was proposed with the establishment of the National Health Service Sickle and Thalassaemia Screening Programme (NHS STP) (Streetly, 2000). Universal newborn screening for clinically significant haemoglobins using dried blood spots is undertaken at thirteen specialist laboratories in England (Streetly, 2009), with similar services in the other UK home nations. Although the aim of newborn screening is the detection of infants with sickling disorders, the NHS STP has proposed that 1.5% or less of Hb A should be considered indicative of possible beta thalassaemia major and referred for follow up (NHS STP Newborn Laboratory Handbook, 2017). Antenatal screening is done at district general hospital level, differing in approach according to whether the area is classified as low or high prevalence for haemoglobin disorders. In low prevalence areas, screening is offered on the basis of the full blood count (FBC) and a family origin questionnaire and further testing is only undertaken if either of these indicates a risk. In high prevalence areas, all women are offered laboratory testing. The NHS STP aim is that all women receive effective, high quality, antenatal screening for the haemoglobinopathies, regardless of the expertise of their local hospital. The NHS STP publishes standards and protocols and offers advice and training for laboratory staff and other healthcare professionals. The NHS STP Antenatal Laboratory Handbook (2017) includes an algorithm for antenatal screening, with a Hb A_2 concentration of 3.5% or above set as the action point for the diagnosis of a beta thalassaemia carrier individual. This cut-off was selected on the basis of advice of a panel of expert advisors to the NHS STP and conforms to the author's personal laboratory experience that the majority of 'normal' individuals will have a

Hb A_2 value less than 3.4% and the majority of beta thalassaemia carriers a value greater than 3.7% with very few results in the intermediate zone. This does not mean that a result in the borderline zone is of no significance (Old, 2012).

The provision of screening programmes in other countries is variable. Some Southern European countries have offered antenatal screening for some time (Giordano, 2009) and pre-marital screening has been mandatory in Cyprus since 1973 (Angastiniotis, 1981). Turkish Cyprus introduced pre-marital screening by law in 1980 (Bozkurt, 2007) and Iran in 1991 (Najmabadi, 2006). When first introduced in Iran, the purpose of screening was to avoid marriage between two affected individuals; this was amended to include the option of termination of pregnancy for beta thalassemia major in 1997. In non-endemic countries, a national antenatal screening programme is only offered in the UK, Greece and The Netherlands (Giordano, 2009, Martinez 2014) at the time of writing, although is available on request or as part of a local programme elsewhere (Cousens, 2010). Newborn screening is universal only in the UK, Greece, The Netherlands and the USA amongst non-endemic countries at the time of writing (Streetly 2009, Giordano 2009, Hoppe 2009, Martinez 2014). In France it is offered on a targeted basis (Bardakdjian-Michau, 2009; Thuret, 2010).

The NHS STP requires laboratories to participate in an accredited EQA scheme and to demonstrate satisfactory performance against the criteria set by the EQA provider and the professions (NHS STP Antenatal Laboratory Handbook, 2017). The NHS STP has also established minimum standards for EQA schemes: these include a requirement for a transparent and robust scoring system for participant performance assessment (NHS STP Antenatal Laboratory Handbook, 2017).

1.9. Laboratory diagnosis of the haemoglobin disorders

Carriers of Hb disorders are usually asymptomatic and can only be identified by laboratory testing. Although disease sufferers have clinical symptoms, laboratory testing is required to confirm the precise nature of the disease. The range of disorders seen in an ethnically diverse population requires expert screening and referral laboratory diagnostic services and is the subject of a number of published guidelines (Streetly and Holland, 2009; ICSH, 1978; Ryan, 2010; Traeger-Synodios, 2015).

The laboratory identification of carriers of inherited Hb disorders is presumptive and these are one of the few categories of genetic disorders that can be diagnosed largely on the basis of clinical background and protein chemistry, rather than molecular techniques. Depending on the condition and the reasons for diagnosis, the following information may be used to make a definitive diagnosis (Clarke and Higgins, 2000; Weatherall and Clegg, 2001; Old, 2003; Bain, 2006; Bain 2010; Old 2012; Traeger-Synodios 2015):

- The patient's ethnic background and clinical presentation.
- General laboratory tests, in particular the FBC (especially the Hb concentration, red blood count and mean cell haemoglobin), reticulocyte count, peripheral blood film morphology and iron status.
- Haemoglobin pattern analysis, using tests based on the altered electrical charge of the Hb molecule.
- Quantification of normal and abnormal Hb fractions, which may be undertaken by the same methods or as part of the Hb pattern analysis.

- Demonstration of red cell inclusion bodies.
- Demonstration of the distribution of Hb F within the red cells.
- Functional investigations of the Hb molecule, e.g. demonstration of altered solubility, stability or O_2 affinity.
- Relative rates of globin chain synthesis (mainly historical, rarely used at the time of writing).
- Molecular techniques, including both mass spectrometry and mutational analysis of DNA.

The range and complexity of the laboratory methods used require specialised equipment and a skilled workforce. It is not cost effective to offer the full range of investigations in all testing centres, requiring the establishment of laboratory networks to ensure the provision of high quality services. The establishment of expert networks for the diagnosis and treatment of the rare and congenital anaemias is strongly supported by the European Network for Rare and Congenital Anaemias (ENERCA) (Vives-Corrons, 2014). The referral of patients to expert secondary or tertiary testing centres, sometimes across national borders (European Directive 2011/24/EU on patients' rights in cross border healthcare), requires harmonised laboratory procedures and a clear demonstration of competency by the service provider.

1.10. Laboratory aspects of Haemoglobin A_2 measurement

Shortly after its discovery, Hb A_2 was reported to be raised in the parents of children suffering from Cooley's anaemia, later known as beta thalassaemia major (Kunkel, 1957) and the demonstration of a raised Hb A_2 is the cornerstone of the laboratory identification of the beta thalassaemia carrier (ICSH, 1978: Globin Gene

Disorder Working Party of the British Committee for Standardisation in Haematology General Haematology Task Force, 1994; Bain, 1988; Ryan, 2010). The increased Hb A_2 level in beta thalassaemia carriers is due to an absolute increase in the synthesis of delta globin chains rather than a relative increase in Hb A_2 due to the reduction in beta globin synthesis (Olivieri and Weatherall, 2001).

The observed increase in Hb A_2 in the beta thalassaemia carrier is variable, depending upon the severity of the beta thalassaemia mutation (Huisman, 1997) but is rarely outside the range of 3.5 to 7.0% with the pathological range for most common beta thalassaemia mutations between 4.0 to 6.0% (Stephens, 2012). The normal reference interval is rarely outside 2.0 to 3.5%. Mild (beta plus) mutations may have Hb A_2 levels close to or overlapping the normal range but are clinically significant as they have been shown to result in severe, transfusion-dependent conditions in the homozygous or compound heterozygous state (Angastiniotis, 2013; Old, 2012). The small difference in Hb A_2 between the unaffected individual and the beta thalassaemia carrier demands that the method of measurement of Hb A_2 is both accurate and precise.

The methods currently available and recommended for Hb A_2 quantification depend on the separation of Hb A_2 from Hb A on the basis of charge (Ryan, 2010; Stephens, 2012). The most commonly available automated methods for Hb A_2 quantitation are high performance liquid chromatography (HPLC) (Wild and Stephens, 1997; Van Delft 2009) and automated capillary zone electrophoresis (CZE) (Cotton, 1999; Gulbis, 2003). Earlier manual methods (ICSH, 1978) using anion exchange microcolumns (Efremov 1977) and electrophoresis on cellulose acetate followed by elution (Marengo-Rowe, 1965) remain the recommended

reference methods but are not widely used in screening or diagnostic laboratories because they are labour intensive and require considerable technical expertise.

Electrospray tandem mass spectrometry (TMS) is a method that has been demonstrated to be suitable for definitive identification of haemoglobin variants (Wild 2001; Wild, 2004) and, more recently, Hb A_2 quantification (Daniel 2007). It has been shown to be effective for use in newborn screening (Moat 2014).

An increased Hb A_2 may also be associated with hyperthyroidism (Kendall, 1981), human immunodeficiency virus (HIV) infection treated with anti-retroviral therapy (Routy, 1993) and megaloblastic anaemia (Mehta, 1983). Haemoglobin A_2 measurement is misleading in the presence of Hb S when measured by some HPLC methods due to the co-elution of glycated Hb S fractions with Hb A_2 (Suh, 1996; Head, 2004). Haemoglobin A_2 measurement is also unreliable in the presence of Hb variants that co-elute with Hb A_2 on HPLC, such as Hb E, some Hb D variants and Hb Lepore. Haemoglobin A_2 may be reduced in certain haemoglobinopathies, e.g. delta-beta thalassaemia, some HPFH conditions, alpha thalassaemia, Hb H disease and delta thalassaemia. Both structural and thalassaemic mutations of the delta globin gene have been identified (Bain, 2006); these are of no clinical significance but their coinheritance with a beta thalassaemia mutation may result in a 'split' Hb A_2 peak and the apparent reduction of a pathological Hb A_2 concentration to the normal reference range, possibly confounding diagnosis. Severe iron deficiency accompanied by severe anaemia is known to reduce the Hb A_2 value (Galanello, 1981), although the significance of this has been debated (Passarello, 2012). Severe iron deficiency may depress the Hb A_2 of an individual with a mild beta thalassaemia and a

borderline Hb A_2 into the normal range, resulting in misdiagnosis (Stephens, 2012).

An early study in the UK examined the ability of laboratories to diagnose beta thalassaemia using Hb A_2 quantitation and found that half of the laboratories incorrectly reported a specimen from a beta thalassaemia carrier as normal, reflecting the lack of expertise and a lack of awareness of factors affecting laboratory performance (White and Lewis, 1973).

Evaluation of Hb A_2 data from EQA surveys in the UK NEQAS Haematology scheme between 2006 and 2008 (Batterbee 2010) demonstrated a bias between the most popular HPLC systems in use, although the study was not peer-reviewed. Paleari (2007) undertook a pilot EQA exercise for Hb A_2 measurement using three fresh blood specimens, anti-coagulated with ethylenediaminetetraacetic acid (EDTA) and with normal, borderline and raised Hb A_2 levels, sent to forty-eight Italian laboratories using HPLC machines from the same manufacturer (Bio-Rad Laboratories). This study reported an overall interlaboratory coefficient of variation (CV%) of between 6 to 8% for the three pools, with up to 32% of laboratories reporting results outside the allowable total error (TE) set by the authors. The study did not examine the bias between manufacturers but the authors noted that results varied according to which model of Bio-Rad instrument was used. The numbers in each individual model group were small and the results may not be statistically valid. All laboratories that took part in this exercise correctly differentiated the normal from the beta thalassaemia carrier.

In 2011, UK NEQAS Haematology data showed similar range of geometric CV% (GCV%) to that reported by Paleari in the results of ten EQA samples (all methods Hb A_2 values between 1.9% and 3.6% and GCV values of 6.0 to 11.2%) distributed to more than 270 laboratories (UK NEQAS Haematology Annual Report, 2011). Note that the GCV% is used by UK NEQAS Haematology because the data are log transformed before analysis.

Paleari also gathered information on reference intervals used by the laboratories in the study and reported that there was an overlap between the upper reference interval limit for normal patients (3.0 to 3.5%) and the lower reference interval for beta thalassaemia carrier patients (3.1 to 3.7%) (Paleari, 2007). A variation in normal reference intervals used, even by users of the same instrument type, was also reported in the evaluation of UK NEQAS Haematology data (Batterbee 2010).

In a later study, Paleari evaluated five HPLC analysers (Bio-Rad Classic running the beta thalassaemia short programme, Bio-Rad Variant II using the dual kit programme, Menarini Diagnostics HA8160, Tosoh G7 and Tosoh G8) and three CZE systems (Sebia Capillarys 2, Beckman Coulter P/ACE MDQ with Analis CEofix™-Hb A_2 kit and Proteome PA 800 with Analis CEofix™-Hb A_2 kit) (Paleari 2012). The instruments were based in two expert thalassaemia centres and a total of eighty EDTA samples were analysed, from healthy individuals (n=40), beta thalassaemia carriers (n=29) and individuals with low Hb A_2 (n=11). There was good agreement between the HPLC and CZE methods, and good correlation between the different HPLC methods: however, a bias was observed between the HPLC methods. This study reported better reproducibility than seen previously with CV% values between 0.5% and 4.4% for the HPLC analysers; however, this is expected since a single analyser of each type was used. The CV% would be

expected to be better on a single instrument than on a number of instruments in different locations (De la Salle, 2012a).

There are few published acceptable performance specifications (APS) for Hb A_2. Mosca has published analytical goals for the measurement, giving different values for total error depending on whether it is derived from biological variation (TE = 4.5%), expert opinion (TE >9.0%) or clinical need (TE = 7.0%) (Mosca, 2009; Mosca, 2013) and has suggested a limit of ±0.25% of total Hb A_2 as an APS for a borderline Hb A_2 value of 3.6%. The rationale stated for this was to avoid a Hb A_2 value of 3.6% being misclassified as normal, i.e. less than 3.35%, or a beta thalassaemia carrier, i.e. greater than 3.85%. Throughout this thesis, a single APS of ±0.25% of total Hb A_2 is used as a guide to acceptable EQA performance; which conforms to the clinical need TE of 7.0% at the critical borderline Hb A_2 value of 3.6%. The clinical need TE was chosen to conform to the aspiration of determining APS according to clinical outcome (Sandberg, 2015). Where dealing with actual UK NEQAS data reported to one decimal place (as is the norm for Hb A_2 measurement), this limit was rounded up to 0.3%. This was a pragmatic decision based on the author's experience of the state-of-the-art of performance in Hb A_2 EQA that predicted rounding down to 0.2%, would prove too strict a specification for inter-laboratory assessment. Stephens advised that in the critical borderline Hb A_2 area of 3.0 and 4.0%, a standard deviation (SD) of 0.05 is achievable for Hb A_2, with duplicate estimations within 0.2% of the numerical Hb A_2 result (Stephens, 2012); however, this specification related to within rather than between-analyser performance.

1.14.4 To examine the current 1.11. External quality assessment as part of total quality management

The concept of using statistical analysis for quality control as part of the production process was developed in the decades between the first and second world wars (Shewhart, 1931; Pearson, 1935) with the introduction of mass production methods in industry. The objective was to interact with the production process to ensure a consistent, high quality output of mass-produced, manufactured goods from factory settings with a relatively unskilled workforce. Quality management became integral to the high throughput workplace, where the workload precludes individual attention for each production item. Similar requirements exist in the provision of diagnostic laboratory services, where quality is monitored and maintained through total quality management (TQM) of the testing process rather than the attention of a skilled analyst to each individual result. The importance of TQM in the laboratory is enshrined in the international standard ISO15189 (ISO 15189:2012). The quality of the laboratory's output is demonstrated by the use of internal quality control (IQC) to monitor batch-to-batch or day-to-day quality and EQA to monitor interlaboratory and intermethod quality, amongst other key performance indicators.

EQA, also known as proficiency testing (PT), was introduced in the UK in the late 1960s following studies on interlaboratory performance (Lewis and Burgess, 1969; Whitehead, 1973). EQA schemes vary in design but all entail the distribution of assay material of known but undisclosed content to the participating laboratories, who test the material as a standard test item and report their results back to the EQA provider. The EQA provider analyses the results against agreed performance

criteria, providing an assessment of the quality of the laboratory's output against that of its peers.

EQA is available to some degree in all developed healthcare systems and the World Health Organisation (WHO) provides a manual to support the establishment of national EQA services (Urassa, 2016). ISO 15189:2012 requires the medical laboratory to undertake interlaboratory comparison for all diagnostic procedures in its repertoire and to take part in EQA where this is available. Selection of the EQA scheme is the responsibility of the participating laboratory (James, 2014) but the scheme should be accredited against the international standard ISO 17043 (ISO 17043:2010) where possible. Depending on national and local regulations, EQA participation may be voluntary, a prerequisite for laboratory accreditation as in the UK, or mandatory for licensing, e.g. as in France and the USA. In the USA, the requirement for EQA (or PT) is stipulated in the Clinical Laboratory Improvement Amendments (CLIA) regulations of 1988.

The international standard ISO 17043:2010 (ISO 17043, 2010) requires an element of performance evaluation in the EQA scheme design and EQA providers have devised a number of performance assessment systems (Sandberg, 2015; Jones, 2017b). For participants to derive the greatest benefit from participation and for the EQA provider to manage the data returned effectively, a statistically derived performance score against a target value or model answer is the most commonly used performance assessment mechanism. The ISO standard ISO 13528:2015, provides guidance on the correct use of statistical methods in EQA as does the informative Annex B of ISO 17043 (ISO 13528:2015; ISO 17043:2010)

The EQA target value for a quantitative test may be determined in a number of ways, including the use of known values (determined from manufacture or dilution), certified reference values (determined from a definitive measurement method), reference values (determined by analysis of the assay material against a reference material or standard), consensus values from expert participants (for example reference laboratories) and consensus values from all participants (ISO 17043:2010, Thompson, 2006). The use of the consensus value from all participants is the method associated with the greatest degree of measurement uncertainty; however, in the absence of a certified reference material (CRM) or method, it may be the only option. Performance assessment in qualitative or interpretive schemes may be against an expected result or model answer, expert opinion or a reference method test result. For the scheme to be credible with participants, the EQA provider must demonstrate the quality of the survey material through the uncertainty of measurement of the target and the degree of consensus of the participants' results with the target.

External quality assessment is uniquely placed to gather data on the 'state of the art' in laboratory methodology and may be used to identify technical problems with *in-vitro* diagnostic devices, although the conditions for this are rigorous in terms of the commutability of survey material (BS EN 14136:2004). International standards applicable to EQA require assay material to be as close to patients' samples as possible, traceable to a CRM, where available, and commutable in the event that the EQA scheme is to be used for inter-method comparison (ISO 17043:2010; BS EN 14136:2004).

1.12. The standardisation Haemoglobin A_2 measurement

Metrological traceability is the means by which laboratory results can be made comparable even where different calibration materials or methods of measurement are used, through the use of a CRM or measurement methods/procedures to which the values of other standards, calibrators or controls can be traced through an unbroken chain of calibrations. A CRM is a substance used for calibration, method development or traceability that is certified as meeting the requirements of the measure in terms of homogeneity, stability and characterisation as specified in the relevant ISO standards. CRMs are listed in the catalogues published by the Joint Research Centre of the European Union, the European Reference Material co-operation and the National Institute for Standards and Technology in the United States of America for materials they produce. The Joint Committee on Traceability in Laboratory Medicine provides a searchable database of higher order reference materials.

The only Hb A_2 reference material available at the time of writing is the first WHO Reference Material (89/666) developed in 1989, available from the National Institute for Biological Standards and Controls in the UK. This reference material has an assigned Hb A_2 value of 5.3% and was validated for use with the manual Hb A_2 methods in use at the time of its development. The International Federation of Clinical Chemistry (IFCC) established a working party to develop a commutable standard for Hb A_2 in 2004. Initial work has been reported on a suitable candidate material (Paleari, 2010) but at the time of writing this requires further extensive validation before it is available for use.

1.13. UK NEQAS Haematology external quality assessment provision for the haemoglobinopathies

UK NEQAS Haematology offers a comprehensive programme for the EQA of laboratory performance in the haemoglobinopathies. It is one of the major schemes available worldwide with numbers equivalent to those registered in the College of American Pathologists (CAP) PT programme and the Royal College of Pathologists in Australasia Quality Assessment Programme (RCPAQAP). It is the most comprehensive scheme available in Europe, in terms of the range of investigations covered and the number of specimens distributed annually (Vives-Corrons, 2014). Three individual programmes are available:

1.13.1 The Newborn Sickle Screening programme, which monitors the performance of laboratories screening for clinically significant haemoglobinopathies using dried blood spot (DBS) samples. At the time of writing, twelve distributions of three specimens each are made annually. The numbers registered are relatively small (approximately 30 at December 2015 with six outside the UK) (UK NEQAS Haematology Annual Report, 2015).

1.13.2 The Abnormal Haemoglobins programme, which monitors the performance of laboratories undertaking the screening and diagnostic investigation of adult patients in the haemoglobinopathies using non-molecular techniques. There are several registration options in the scheme: emergency sickle solubility testing (three whole blood specimens distributed six times per year), full participation for Hb fraction identification, quantitation and the interpretation of the results (six distributions of three

whole blood specimens annually) and specimens that simulate capillary blood samples from a newborn infant for laboratories (approximately thirty in the UK) that undertake clinician-directed newborn testing for haemoglobinopathies (six distributions of two specimens annually). At December 2015, there were 317 laboratories registered for full participation, 170 of which were outside the UK (UK NEQAS Haematology Annual Report, 2015).

1.13.3 The DNA Diagnostics for Haemoglobinopathies programme, which monitors the performance of laboratories undertaking alpha and beta globin mutational analysis for the haemoglobinopathies. This is an internationally operated programme with the majority of participating laboratories outside the UK. At December 2015 there were forty-eight laboratories registered, thirty-eight outside the UK (UK NEQAS Haematology Annual Report 2015).

An example report for the Abnormal Haemoglobins scheme is shown in Appendix 1

UK NEQAS Haematology uses a harmonised method for the calculation of performance scores across all its quantitative assay schemes. The system was developed in 1996 for the evaluation of performance in the automated counting schemes (FBC, Automated Differential Count and Reticulocyte Count) and, although ratified by the UK NEQAS Steering Committee and the oversight body, the National Quality Assurance Advisory Panel, in a discussion paper (England 1995), it was not published in a peer-reviewed journal. The system is based on the principle of scoring long-term, retrospective performance (UK NEQAS

Haematology Participants' Manual, 2017), taking between four to six months for persistent unsatisfactory performance in Hb A_2 quantitation to become evident. The length of time taken to detect a laboratory with unsatisfactory performance has been noted as a source of concern by the NHS STP since the unsatisfactory laboratory could screen a large number of antenatal patients in the intervening period. One-off errors that could constitute misdiagnosis if they occurred with a patient's sample are not monitored by the current scoring system.

The NHS STP does not have a formal screening programme for the detection of infants with beta thalassaemia major but has suggested a concentration of less than 1.5 % Hb A may be an indicator of the condition (NHS STP Newborn Laboratory Handbook, 2017). There is no EQA data on how sensitive the screening methods are at this Hb A concentration.

The majority of errors in laboratory medicine occur outside the analytical phase (Plebani 2014) and the performance assessment of the interpretation of results is an increasingly important feature of EQA performance. The provision of an unambiguous interpretation of haemoglobinopathy results is essential to ensure the correct follow up actions are taken. Interpretation is not scored in the Abnormal Haemoglobins scheme and may be a challenge given the large international participant base. For any scoring method in EQA to be effective and credible with participants, the source of errors must be understood and steps taken to eliminate confounding factors that result from the scheme design.

1.14. Project aims

1.14.1 To review the current state of the art Hb A_2 measurement using historical UK NEQAS data, to assess the bias between instrument types and the impact that this might have on the laboratory diagnosis of beta thalassaemia carriers.

1.14.2 To explore the suitability of the currently used EQA survey material for the assessment of laboratory performance in the quantitation of Hb A_2 in comparison to fresh blood.

1.14.3 To examine the sensitivity of current methods for newborn haemoglobinopathy testing for the quantitation of Hb A in dried blood spot specimens, as a means to detect infants with beta thalassaemia major.

performance scoring methods for Hb A_2 quantitation for their effectiveness in detecting laboratories with unsatisfactory performance and how these methods might be improved to allow the earlier identification of problems.

1.14.5 To explore national and international barriers to the extension of performance assessment to Hb fraction identification and case interpretation in the Abnormal Haemoglobins scheme and how these may be overcome

1.15 Declaration

Two sources of data have been used in this thesis: historical EQA data from UK NEQAS exercises, which has been mined and further analysed for performance trends and the modelling of amendments to performance assessment methods, and experimental data gathered as a result of practical exercises.

Raw historical EQA data (participants' results and target values) that had been gathered and calculated as part of routine UK NEQAS exercises were extracted from the UK NEQAS database; statistics calculated in this way are identified as those produced using the bespoke EQA software supplied by KPMD Ltd (Sheffield, UK). All further analysis and data modelling was undertaken by the author. None of the data analyses presented form part of the current data reported in UK NEQAS Haematology exercises.

An initial analysis of performance trends for Hb A_2 data gathered between 2006 and 2008 was commissioned by the National Sickle and Thalassaemia Screening Programme and undertaken by Hannah Batterbee (Batterbee, 2010): this data is referred to for comparison purposes and some of the same raw data is included in the further analysis where it falls into the time periods described in this document.

.

Chapter 2: Review of Historical UK NEQAS Haematology Haemoglobin A_2 Data for Trends in Performance

2.1. Aim

To review the current state of the art in laboratory performance in the measurement of Hb A_2 using historical UK NEQAS external quality assessment data.

2.2. Introduction

Haemoglobin A_2 (Hb A_2) was identified in 1955 as a minor, normal Hb fraction in adults, found at a concentration of approximately 3% (Kunkel and Wallenius, 1955). The percentage of Hb A_2 was later shown to be increased in carriers of beta thalassaemia (Kunkel, 1957) and the laboratory demonstration of an increased Hb A_2, together with characteristic changes in red cell indices, is the standard biomarker used in population screening for adult beta thalassaemia carriers.

Suitable protein chemistry methods for the quantitation of Hb A_2 as a percentage of total Hb are based the charge difference between Hb A_2, a tetramer of two alpha and two delta globin chains ($\alpha_2\delta_2$), and the other haemoglobin fractions found in the normal adult, i.e. Hb A ($\alpha_2\beta_2$) and Hb F ($\alpha_2\gamma 2$). A number of manual methods have been used historically, based on the separation of Hb A_2 from other Hb fractions by electrophoresis or chromatography. The modern medical laboratory requires population screening methods that are cost effective and high throughput, as well as of clinical utility, and the most widely used automated and semi-automated methods for Hb A_2 measurement in use at the time of writing are analysers utilising high performance liquid chromatography (HPLC) or capillary zone electrophoresis (CZE). The use of DNA diagnostic methods is not a cost-effective means of population screening for beta thalassaemia at the time of

writing, even in well-developed health economies, and is used only when a definitive diagnosis is required.

The difference in Hb A_2% between the normal and the beta thalassaemia heterozygote individual is small (less than 0.5% in some cases) and variable, dependent upon the nature of the beta globin gene mutation inherited (Weatherall and Clegg, 2001; Old, 2012). A wide range of beta thalassaemia mutations is seen in the UK, reflecting the diverse ethnic backgrounds of the population, and is greater than may be seen in the regions of the world from which people may have migrated (Henderson, 2009). The NHS Sickle and Thalassaemia Screening Programme's (NHS STP) laboratory screening algorithm includes a cut-off Hb A_2 value of 3.5% or above, in conjunction with a MCH of less than 27 pg, as indicative beta thalassaemia carrier status in an antenatal woman, recommending the screening of the fetus' biological father to assess whether the fetus is at risk of a major haemoglobinopathy. The small difference between the normal and beta thalassaemia carrier state and the use of a fixed cut-off requires methods of measurement that are both accurate and precise.

The UK NEQAS Haematology scheme for Abnormal Haemoglobins was established in 1975 (unpublished UK NEQAS Haematology Steering Committee papers and annual reports). At the time of writing, the scheme makes six distributions annually, each comprising three adult, whole blood specimens for the identification of variant haemoglobin fractions and the quantitation of Hb A_2, Hb S and Hb F, as appropriate to the case. The general pattern of specimens distributed results in one to two specimens for Hb A_2 quantitation per survey; on occasion none or all three may require Hb A_2 measurement (UK NEQAS Haematology Annual Report, 2015).

The UK NEQAS Haematology Abnormal Haemoglobins scheme provides a unique resource of historical data that can be mined to understand better the differences between different methods and to monitor trends in performance over time. The strength of the historical data is that it has been gathered from numerous models of the most commonly used analysers, based in routine diagnostic laboratories and testing aliquots of the same, homogeneous pools of survey material. This gives a good representation of inter-laboratory and inter-method variability in performance over time, assuming that the survey material is commutable. The weakness of using historical data is that it is by its nature retrospective and will always be out-of-date; however, as long as this is borne in mind, historical data provides an important tool for understanding the principles of the potential sources of error in laboratory testing. The use of historical EQA data may be confounded by changes in the method of survey material preparation; however, this is not the case for the preparation of UK NEQAS Haematology Abnormal Haemoglobins material, the preparation of which is described below.

2.3. Preparation of survey material

UK NEQAS Haematology Abnormal Haemoglobins survey material is prepared from whole blood in citrate phosphate dextrose (CPD) anticoagulant, supplied by the National Health Service Blood and Transplant (NHSBT) service. The use of donor blood for quality assessment purposes is approved by NHSBT and donors agree to this use as part of their general consent to blood donation. Each pool of survey material is prepared from a minimum of one unit of whole blood (approximately 450 mL) or concentrated red cells made up to volume with ABO group compatible fresh frozen plasma. Blood is screened by NHSBT as negative for anti-HIV antibodies, anti-hepatitis C virus (anti-HCV) antibodies, HbsAg and

syphilis and is seventy-two hours old upon receipt by UK NEQAS. The blood is not stabilised for use as EQA material but broad spectrum antibiotics are added to maintain sterility (benzyl penicillin BP reconstituted with sterile 0.9% sodium chloride infusion BP, at a final concentration of 5 μg penicillin per mL of blood; gentamicin sulphate injectable solution at a final concentration of 50 μg gentamicin per mL of blood). Specimens are dispensed into 2 mL, plain glass tubes using a bespoke mixing and dispensing system. On occasion, ABO-compatible units of blood may be pooled to increase the volume of the survey material available or to create a specimen with an 'intermediate' Hb A_2 value at or around 3.5% by mixing blood from a beta thalassaemia carrier and a normal donor. The finished EQA specimens are seven to ten days old at distribution, which is by first-class post in the UK and door-to-door courier service to participating laboratories outside the UK. Participants are allowed ten days from the date of despatch to analyse the samples and return their results for analysis.

2.4 Manual methods of Haemoglobin A_2 measurement

2.4.1 Densitometry

Haemoglobin A_2 is separated from Hb A by electrophoresis at alkaline pH on paper, starch gel or cellulose acetate. It is possible to quantitate the resulting peaks by scanning the stained electrophoretic strip and integrating the peak area of the Hb fractions present; however, this has been shown to be inaccurate for the measurement of Hb A_2 and the identification of beta thalassaemia carriers, the major technical challenge being the large difference between the density of the Hb A and Hb A_2 bands and the consequent variance in the Hb A_2 measurement (Schneider, 1974; Schmidt, 1975; Ueda, 1975; Sreedharanunni, 2015).

2.4.2 Cellulose acetate electrophoresis followed by elution

Haemoglobin A and Hb A_2 in a haemolysate are separated on 25x120mm cellulose acetate electrophoresis strips in a Tris-EDTA-borate buffer system at pH 8.9 (Marengo-Rowe, 1965; Old, 2012). The resulting Hb A and Hb A_2 bands are cut out and the haemoglobins eluted in 20 mL and 4.0 mL respectively of the Tris-EDTA-borate buffer (or similar volumes in the same proportion). The Hb A_2 concentration is calculated from the optical density of the eluates at 413 nm, using the formula shown in Equation 2.1.

$$\text{Hb A}_2\ \% = \frac{\text{A413nm Hb A}_2}{(\text{A413nm Hb A}_2\ \text{x}\ 5) + \text{A413 nm Hb A}}\ \text{x}\ 100$$

Equation 2.1 Calculation of Hb A_2 concentration

This method is time-consuming and requires a high level of technical skill; however, it is reproducible and sufficiently accurate for the laboratory diagnosis of

beta thalassaemia carriers (ICSH 1978; Weatherall and Clegg, 2001; Stephens 2012).

2.4.3 Microcolumn chromatography

Haemoglobin A and Hb A_2 in a haemolysate are separated by microcolumn chromatography using a mobile buffer phase and a stationary diethylaminoethyl (DEAE) cellulose anion-exchange resin. In the original method of Efremov (Efremov 1974) haemolysate is applied to a DEAE cellulose column equilibrated with Tris buffer at pH 8.5, at which all the Hb fractions bind to the column; the Hb A_2 fraction is eluted with a fixed volume of 'developer' buffer at pH 8.3 and the eluted Hb A_2 fraction collected into a container of suitable, accurate and known volume (e.g. a 5 mL volumetric flask). The remaining Hb is eluted using a larger volume of Tris buffer at neutral pH and collected into a second container of suitable, known volume (e.g. a 25 mL volumetric flask). The eluates are made up to volume with deionised water, their absorbance read spectrophotometrically at 413 nm and used to calculate the Hb A_2 concentration as described for the electrophoresis and elution method. This method and variations thereof, including commercially-produced, single-use columns, have been shown to perform well for Hb A_2 quantitation (Brosious 1978; Galanello 1977) but again the method demands a high level of technical skill and is labour intensive.

2.5 Automated and semi-automated methods of Haemoglobin A_2 measurement

2.5.1 High performance liquid chromatography

High performance liquid chromatography has emerged as the predominant method for Hb A_2 measurement in recent decades, as the method allows the simultaneous separation and quantitation of the normal Hb fractions and common, clinically significant abnormal Hb variants. The method separates Hb variants according to their different mobilities in a mobile, phosphate buffer phase and a stationary, cation exchange resin phase held in a temperature controlled cartridge. After initial binding of all Hb fractions to the column, the analyser delivers a pre-programmed buffer gradient of increasing pH and ionic strength, eluting the Hb fractions on the basis of their charge. The separated fractions pass through a flow cell where they are putatively identified by their relative retention time on the column and quantitated on the basis of their absorbance.

Automation of HPLC with sample loaders, computer-driven analysis and interfacing with the laboratory information system has made the method relatively high throughput and ‘walk-away', incorporated into commercially produced analysers. Although this method demands less technical skill than the earlier manual methods, the set-up, quality management and diagnostic interpretation of the chromatographic output requires expertise. Haemoglobin A_2 quantitation by HPLC is sufficiently accurate to distinguish most beta thalassaemia carriers for the purposes of screening (Tan, 1993: Wild 1997; Van Delft 2009).

2.5.2 Capillary zone electrophoresis

Capillary zone electrophoresis separates Hb fractions in narrow (25 to 100 mm diameter), buffer-filled capillary tubes on the basis of their charge. The capillary walls are coated with silanol molecules, which dissociate to give the surface a positive charge, binding negatively charged molecules present in the mobile phase. The application of a potential across the capillary causes the cations in the mobile phase to migrate towards the cathode, producing an electro-osmotic flow of the bulk solution through the capillary. Haemoglobin fractions are separated according to the degree of negative charge on the molecule and the strength of their interaction with the positively charged capillary wall. As with HPLC, the separated fractions are quantitated from their absorbance as they pass through a flow cell and putatively identified from their retention time in the capillary.

Commercially produced CZE analysers have a number of capillaries operating in parallel and the method is a viable alternative to HPLC for Hb A_2 quantitation (Mario, 1997; Cotton, 1999; van Delft, 2009; Sangkitporn, 2011).

2.5.3 Electrospray tandem mass spectrometry

Electrospray tandem mass spectrometry (TMS) is a method that has been demonstrated to be suitable for definitive identification of haemoglobin variants (Wild, 2001) and, more recently, Hb A_2 quantification (Daniel, 2007), separating ionised globin peptides on the basis of their charge to weight ratio.

2.6. Materials and methods

2.6.1 Statistical analysis

Statistical analysis was undertaken using either STATA 14 (Stata Corp, Texas, USA) or Microsoft Excel 97 2003 (Microsoft Corp, WA, USA) with the downloaded analysis Toolpak, unless the UK NEQAS all laboratory trimmed mean (ALTM) or method trimmed mean (MTM) target value is quoted, which had been calculated as described in section 2.6.3.

The Hb A_2 data returned by UK NEQAS participants cannot be assumed to be normally distributed since the data sets may include spurious or outlier results (e.g. as a result of specimen transposition or data entry errors, as well as analytical problems) and these will skew the distribution. The outliers are removed by trimming in the calculation of the standard all lalaboratory trimmed mean (ALTM) and method trimmed mean (MTM) target values; however, for some instrument groups, the numbers of participants registered are too few (fewer than ten) for the use of standard trimming. For these reasons, robust, non-parametric statistical methods were used to calculate the non-ALTM and non-MTM descriptive statistics used in this study, i.e. the median, interquartile range (IQR), estimated standard deviation (est. SD) and estimated coefficient of variation (est. CV) (Tukey 1977).

The est. SD was calculated using the interquartile range (IQR) as shown in Equation 2.2.

$$Estimated\ SD = \frac{IQR}{1.349}$$

Equation 2.2 Calculation of the estimated standard deviation

The est. CV% was calculated by dividing the est. SD by the median and expressing the result as a percentage.

The pattern of bias can highlight performance problems in the results returned by individual participants in successive EQA distributions and performance trends in different methods. The percent bias (D%) of the EQA result from the target value was calculated as shown in Equation 2.3, where x_i is the test result and x_{pt} is the target or assigned value.

$$\text{Bias (D\%)} = \frac{x_i - x_{pt}}{x_{pt}} \times 100$$

Equation 2.3 Calculation of percent bias

The Kruskal-Wallis equality of rank test, a non-parametric analysis of variance (ANOVA) test was used to assess the statistical significance of differences between the data from different analyser groups. The Kruskal-Wallis test is a one-way analysis of variance test suitable for use where the data is not normally distributed (Altman 1992, Bland 2015).

2.6.2 Box and whisker plots

Box and whisker plots illustrate the distribution of data and whether the data is skewed. Box and whisker plots show a rectangular ‘box’ equivalent to the IQR with the median marked. The ‘whiskers’ show the range of 1.5 times the upper or lower quartile rolled back to the closest data point, representing the distribution of the data excluding outliers, which are shown as individual data points beyond the whiskers as illustrated in Figure 2.1.

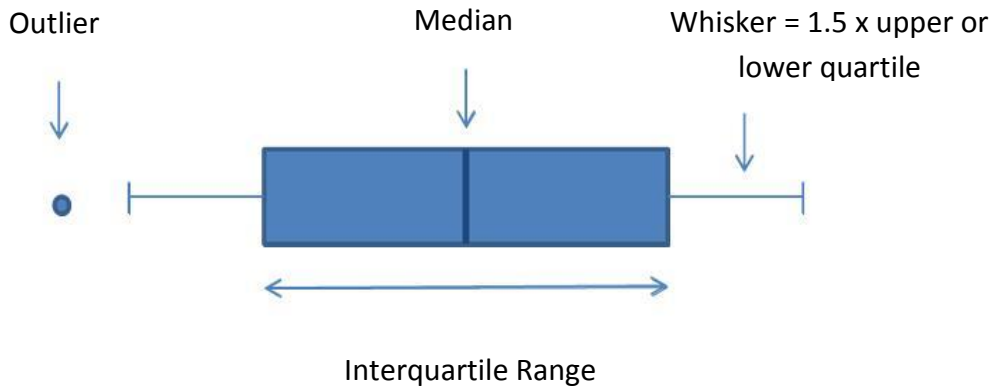


Figure 2.1. Diagram of the 'box and whisker' plot layout

2.6.3 Calculation of the All Laboratory, Method and Sub-method Trimmed Mean values

UK NEQAS target values are generated from a consensus of participants' data using bespoke EQA software (KPMD Ltd, Sheffield, UK). Data is trimmed symmetrically to remove outliers that might otherwise affect the mean value: 10% of data points (5% from the top and 5% from the bottom of the distribution) are removed and data is log-transformed prior to the calculation of the consensus target mean. A robust estimate of the standard deviation (SD) is calculated (Healy 1979). The All Laboratory Trimmed Mean (ALTM) includes data from all participants returning data for the distribution. The Method Laboratory Trimmed Mean (MTM) is calculated for each Hb A_2 method principle, provided there are more than twenty participants registered for the method. At the time of writing, the MTM is calculated for HPLC and CZE method groups.

The appropriate MTM is used in UK NEQAS Abnormal Haemoglobin surveys as the target value (x_{pt} or mean for proficiency testing) for the performance

assessment of Hb A_2 estimation for HPLC and CZE; other method principles with smaller numbers are assessed against the ALTM for statistical robustness. The use of the MTM where there are sufficient numbers registered is largely historical, from the period when manual methods were in use. The separation of HPLC and CZE methods for performance assessment is not evidence-based. The Sub-Method Trimmed Mean (SMTM), e.g. for different HPLC methods, is calculated for information where there are sufficient instruments registered (usually more than twenty) as it provides state-of-the-art performance information by instrument model for manufacturers and participants.

2.6.4 Analysis of changes in methodology (2001 to 2015)

The number of methods used for for Hb A_2 quantitation in December of each year was analysed by method principle from 2001 to 2015. Instruments or methods registered from which results were not returned were excluded since they may no longer have been in use, as were laboratories using methods that are not recommended for Hb A_2 quantitation (Ryan, 2010; Stephens, 2012).

2.6.5 Analysis of trends in Haemoglobin A_2 performance (2001 to 2016)

The est. CV% by HPLC was calculated for each survey material pool distributed between 2001 and 2016 and the change in est. CV% during the period examined, grouping the results according to the ALTM Hb A_2 (2.0% or less, 2.1 to 2.5%, 2.6 to 3.2% and 3.6% or greater; note that there were no specimens with ALTM values between 3.3% and 3.5%).

2.6.6 Evaluation of Haemoglobin A_2 performance data (2006 to 2008)

UK NEQAS Haematology undertook a review of historical Hb A_2 data from the UK NEQAS Abnormal Haemoglobins scheme between 2006 and 2008. This work was

commissioned by the NHS STP and undertaken in collaboration with the Sheffield Teaching Hospitals NHS Trust. This work was not published in a peer-reviewed journal but was made publicly available as a UK NEQAS Haematology report (Batterbee 2010). The 2006 to 2008 study examined the results of twenty-seven specimens issued for the assessment of Hb A_2%, analysed by the most commonly used Hb A_2 quantitation methods in use at the time. The conclusions were that there was an observed difference in the performance of different analysers, with a 'significant negative bias' or a 'significant positive bias' noted with particular analyser models. However, the bias was only documented graphically and did not include statistical analysis.

To evaluate the performance of the different analytical methods further, the bias of the SMTM against the ALTM was calculated in this study for the same twenty-seven specimens distributed between 2006 and 2008 and compared to an 'acceptable EQA performance range' of the ALTM ±0.25% of Hb A_2 for each specimen, established using published performance limits (Mosca, 2013).

2.6.7 Evaluation of Haemoglobin A_2 performance data (2013 to 2016)

UK NEQAS Hb A_2 data for thirty-four survey material pools (distributions 1304AH to 1603AH) distributed between August 2013 and June 2016 was examined for inter-method variation, comparing the SMTM for each instrument method with the trimmed ALTM. For each specimen, only methods with more than ten results were included, resulting in the exclusion of all manual methods but the inclusion of newer HPLC and CZE analysers in comparison to the 2006 to 2008 study. The results from seven HPLC methods (Bio-Rad D10 using the dual kit programme, Bio-Rad Variant II using the beta thalassaemia short programme, Bio-Rad Variant II using the dual kit programme, Menarini HA8160, Primus Ultra, Tosoh G7 and

Tosoh G8) and three CZE methods (Sebia Capillarys, Sebia Capillarys 2 and Sebia Minicap) were included. For two methods (the Tosoh G7 and the Sebia Minicap), there were fewer than ten results returned for some specimens (one specimen for the Tosoh G7 and ten for the Sebia Minicap) and these specimens were excluded from the evaluation for these analysers where individual analyser groups were used. For some parts of the evaluation, all CZE instruments manufactured by Sebia were grouped together to give a more statistically robust data set, as the performance of the three analyser models was sufficiently similar (see later analysis). Miscellaneous HPLC or CZE instrument types or those where no analyser model was quoted were excluded from instrument group or sub-method analysis but were included in the method principle (HPLC or CZE) and all methods analyses.

The variance between the results returned for each of the thirty-four specimens distributed was examined to assess whether the observed differences in the means returned were significant statistically, using the non-parametric Kruskal-Wallis test. The null hypothesis for the analysis was that there was no difference between the different analyser groups, i.e. that any differences observed were the result of sampling.

2.6.8 'Borderline Hb A_2 specimens (2006 to 2011)

Three 'borderline' Hb A_2 specimens with Hb A_2 values of 3.6 to 3.7% (0604AH4, Hb A_2 3.7%; 0902AH1, Hb A_2 3.7% and 1105AH3, Hb A_2 3.6%, distributed in the five years between 2006 to 2011) were examined to assess the possible impact of the use of different methods of Hb A_2 measurement on the correct diagnosis of beta thalassaemia carrier status in patients with Hb A_2 results at or around the 3.5% cut-off value.

2.6.9 Reported Haemoglobin A_2 reference intervals

The reference intervals reported for the 1604AH distribution from June 2016 were collated and the most commonly reported minimum and maximum reference values identified for instrument groups with more than ten analysers.

2.6.10 European Network for Congenital and Rare Anaemias exercise (2010)

In a separate practical exercise in 2010, two specimens, from a beta thalassaemia carrier and a normal donor, were prepared according to standard UK NEQAS methods for Hb A_2 survey material pools and distributed to 243 laboratories in the UK and Europe as part of an EQA study in collaboration with ENERCA. The results were returned within the standard ten days for usual UK NEQAS exercise.

2.7 Results

2.7.1 Change in methodology (2001 to 2015)

The number of laboratories returning Hb A_2 results in December each year rose from 238 in 2001 to 324 in 2015, an increase of 36%. There was a significant change in the methodology used during the period, with a decline in the use of manual methods (microcolumn chromatography and electrophoresis and elution) from 60% of participants in 2001 (144/238) to just 3% in 2015 (10/324) and a corresponding increase in automated methods from 39% (94/238) of participants in 2001 to 97% (314/324) in 2015. The predominant automated method used during the period was HPLC, rising from 34% (94/238) in 2001 to a peak of 90% (253/280) in 2009. From 2009, the proportion of results returned from HPLC analysers declined to 79% (256/324) in 2015, although the absolute numbers of results returned from HPLC analysers increased. During the same period, the results returned from automated CZE instruments increased from nil to 18% (58/324) (see Figure 2.2).

In December 2015 there were 328 HPLC and CZE methods registered for Hb A_2 quantitation, from 31 different countries. The largest numbers of instruments were in the UK (160), Belgium (36), Turkey (24), Portugal (18), Italy (18) and Israel (12) (see Table 2.1). Twenty-two countries had five or fewer instruments registered. Note that the number of instruments registered is greater than the number results returned. Since not all participants return results for all distributions, the exact numbers and range of analysers reporting data for any individual distribution described in this study may vary slightly from the numbers registered.

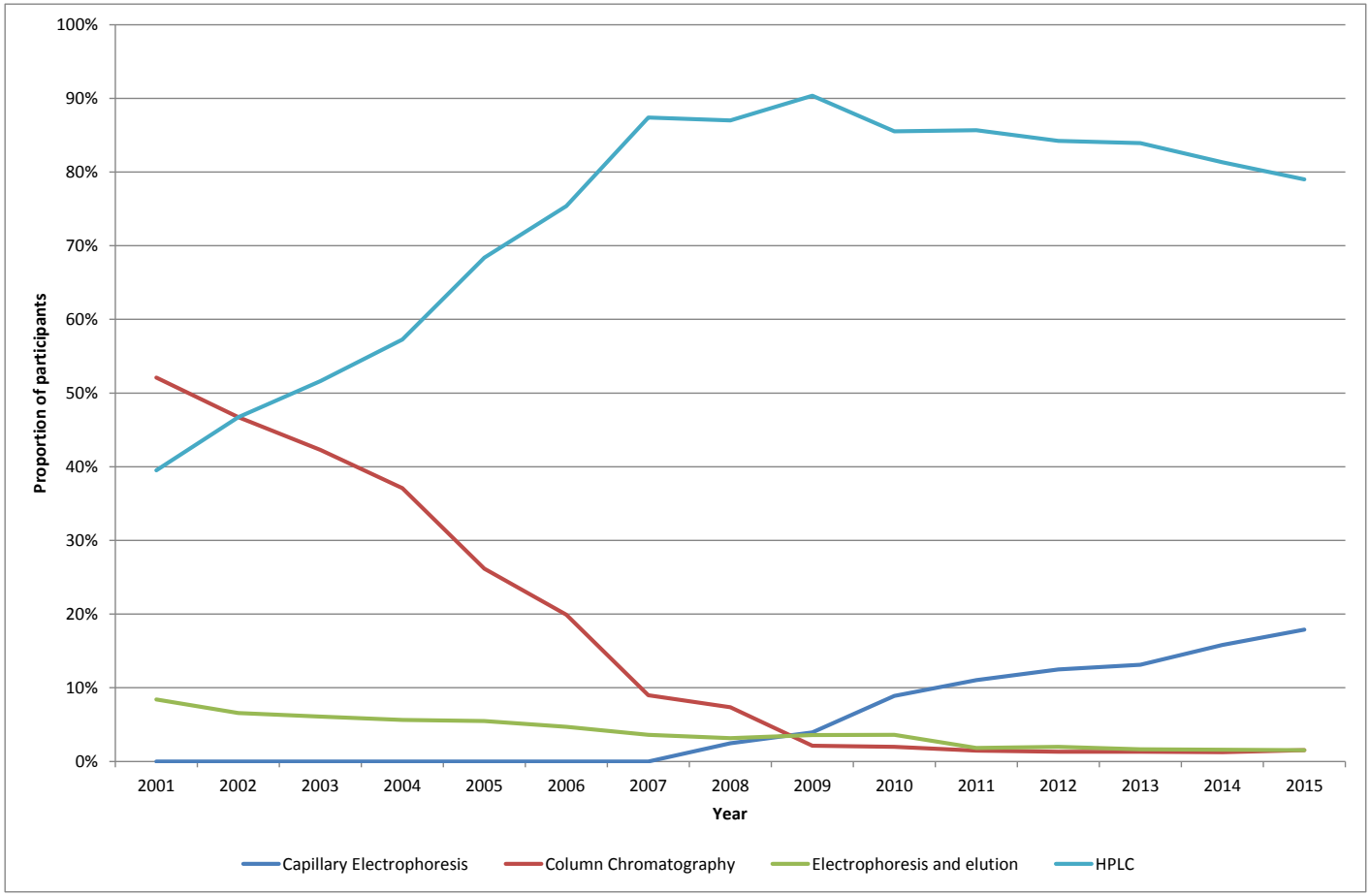


Figure 2.2. Change in method principles used for Haemoglobin A_2 measurement in UK NEQAS Haematology from 2001 to 2015, data showing the methods stated for results returned in December of each year.

Country	Bio-Rad				Menarini HA		Primus	TOSOH		HPLC Other	Sebia Capillarys			CZE Other	TOTAL
	D10	CL	V2 B	V2 DK	8160	8180T		G7	G8		CE	CE2	Mini		
UK	15	2	47	18	16	5	4	6	32	3	6	4	2	0	**160**
Belgium	0	1	0	0	3	0	0	3	3	1	11	7	6	1	**36**
Turkey	1	0	0	0	0	3	15	0	2	2	0	0	0	1	**24**
Italy	2	0	2	6	1	0	0	0	4	0	1	2	0	0	**18**
Portugal	3	2	3	0	6	0	0	1	1	0	1	1	0	0	**18**
Israel	3	0	2	0	0	0	0	1	0	0	4	2	0	0	**12**
RoI	1	0	2	0	0	0	1	1	1	0	1	1	0	0	**8**
Kuwait	1	0	2	0	0	0	0	0	0	3	0	0	0	0	**6**
Norway	0	0	3	0	0	0	0	0	1	0	1	1	0	0	**6**
Switzerland	1	0	1	2	0	0	0	0	0	1	0	0	0	0	**5**
France	1	0	0	0	0	0	0	1	0	0	0	1	1	0	**4**
Austria	0	0	1	0	2	0	0	0	0	0	0	0	0	0	**3**
Germany	0	0	0	0	0	0	0	0	0	1	0	2	0	0	**3**

Table 2.1. Location of HPLC or CZE instruments registered in UK NEQAS Haematology for Haemoglobin A_2 quantitation in December 2015

Continued on next page

Country	Bio-Rad				Menarini HA		Primus	TOSOH		HPLC Other	Sebia Capillarys			CZE Other	TOTAL
	D10	CL	V2 B	V2 DK	8160	8180T		G7	G8		CE	CE2	Mini		
Sweden	0	0	2	0	0	0	0	0	0	1	0	0	0	0	**3**
UAE	0	0	1	0	0	0	1	0	0	1	0	0	0	0	**3**
Greece	0	0	0	1	0	0	0	1	0	0	0	0	0	0	**2**
Luxembourg	0	0	0	1	0	0	0	0	0	0	0	0	1	0	**2**
Oman	0	0	2	0	0	0	0	0	0	0	0	0	0	0	**2**
Australia	0	0	1	0	0	0	0	0	0	0	0	0	0	0	**1**
Bulgaria	0	0	0	0	0	0	0	0	0	0	0	0	1	0	**1**
Cyprus	0	0	0	0	0	0	0	0	0	1	0	0	0	0	**1**
Czech Republic	0	0	0	0	0	0	0	0	0	0	0	0	1	0	**1**

Table 2.1. *continued*

Continued on next page

Country	Bio-Rad				HA		Primus	TOSOH		HPLC Other	Sebia Capillarys			CZE Other	TOTAL
	D10	CL	V2 B	V2 DK	8160	8180T		G7	G8		CE	CE2	Mini		
Denmark	0	0	0	0	0	0	0	0	1	0	0	0	0	0	**1**
Finland	0	0	0	0	0	0	0	0	0	1	0	0	0	0	**1**
India	0	1	0	0	0	0	0	0	0	0	0	0	0	0	**1**
Kenya	0	0	0	0	0	0	0	0	0	0	0	0	1	0	**1**
Malta	0	1	0	0	0	0	0	0	0	0	0	0	0	0	**1**
Netherlands	0	0	0	0	0	0	0	0	0	0	0	1	0	0	**1**
NL Antilles	0	0	1	0	0	0	0	0	0	0	0	0	0	0	**1**
South Africa	0	0	0	0	0	0	0	0	0	0	1	0	0	0	**1**
Spain	0	0	0	0	0	0	0	1	0	0	0	0	0	0	**1**
TOTAL	**28**	**7**	**70**	**28**	**28**	**8**	**21**	**15**	**45**	**15**	**26**	**22**	**13**	**2**	**328**

Table 2.1. *continued*

Continued on next page

Legend to Table 2.1:

Bio-Rad D10:	Bio-Rad D10 HPLC analyser
Bio-Rad CL:	Bio-Rad Classic HPLC analyser
Bio-Rad V2 B:	Bio-Rad Variant II HPLC analyser with the beta thalassaemia short programme
Bio-Rad V2 DK:	Bio-Rad Variant II HPLC analyser with the Dual Kit programme
HA8160/8180T:	Menarini HA8160 and 8180T HPLC analysers
Primus:	Primus Ultra 2 HPLC analyser
Tosoh G7/G8:	Tosoh G7 and G8 HPLC analysers
HPLC Other:	Miscellaneous HPLC instruments or method not stated
Sebia Capillarys:	CE, CE 2 and Minicap CZE analysers
CZE Other:	Miscellaneous CZE instruments or method not stated

2.7.2 Range of specimens distributed (2001 to 2016)

Between 2001 and 2016 a total of 144 specimens were distributed with Hb A_2 values ranging from 1.5% to 5.0%. 14/144 (9.7%) were in the low/normal Hb A_2 range of 1.5 to 2.0%, 114/144 (79.2%) in the mid-normal range of 2.1 to 3.2%, 3/144 (2.1%) in the borderline range of 3.3 to 3.8% and 13/144 (9.0%) in the beta thalassaemia range of greater than 3.9%. Figure 2.3 shows the distribution of the Hb A_2 values distributed by year.

2.7.3 Analysis of estimated CV% trends in Haemoglobin A_2 performance (2001 to 2016)

The coefficient of variation (CV%) is a measure of the variation in a data set relative to the mean and can be used to observe the trends in performance over time. The est. CV% values for Hb A_2 by HPLC were analysed for trends in performance and are shown in Figure 2.4.

For specimens of all values of Hb A_2 there has been a downward trend in est. CV% in the time period reviewed from an average of 14.25% in 2001 to 8.87% in 2016. A similar trend is seen at all Hb A_2 levels, although there is an inverse relationship between the est. CV% and the measurand value.

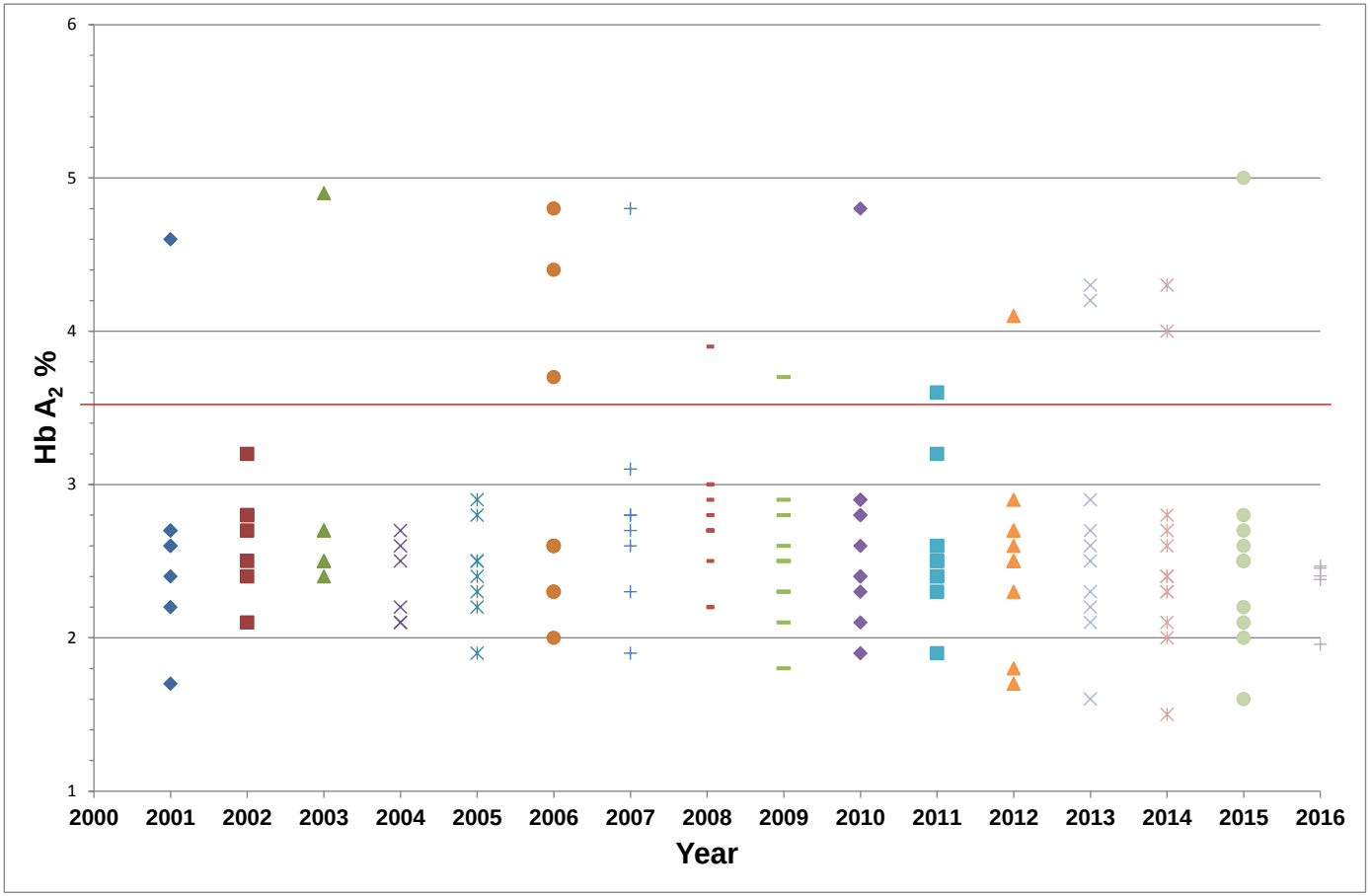


Figure 2.3. The range of Haemoglobin A_2 % values distributed by UK NEQAS Haematology between 2001 and 2016. The Haemoglobin A_2 cut-off of 3.5% for the identification of beta thalassaemia carriers is shown by the solid line.

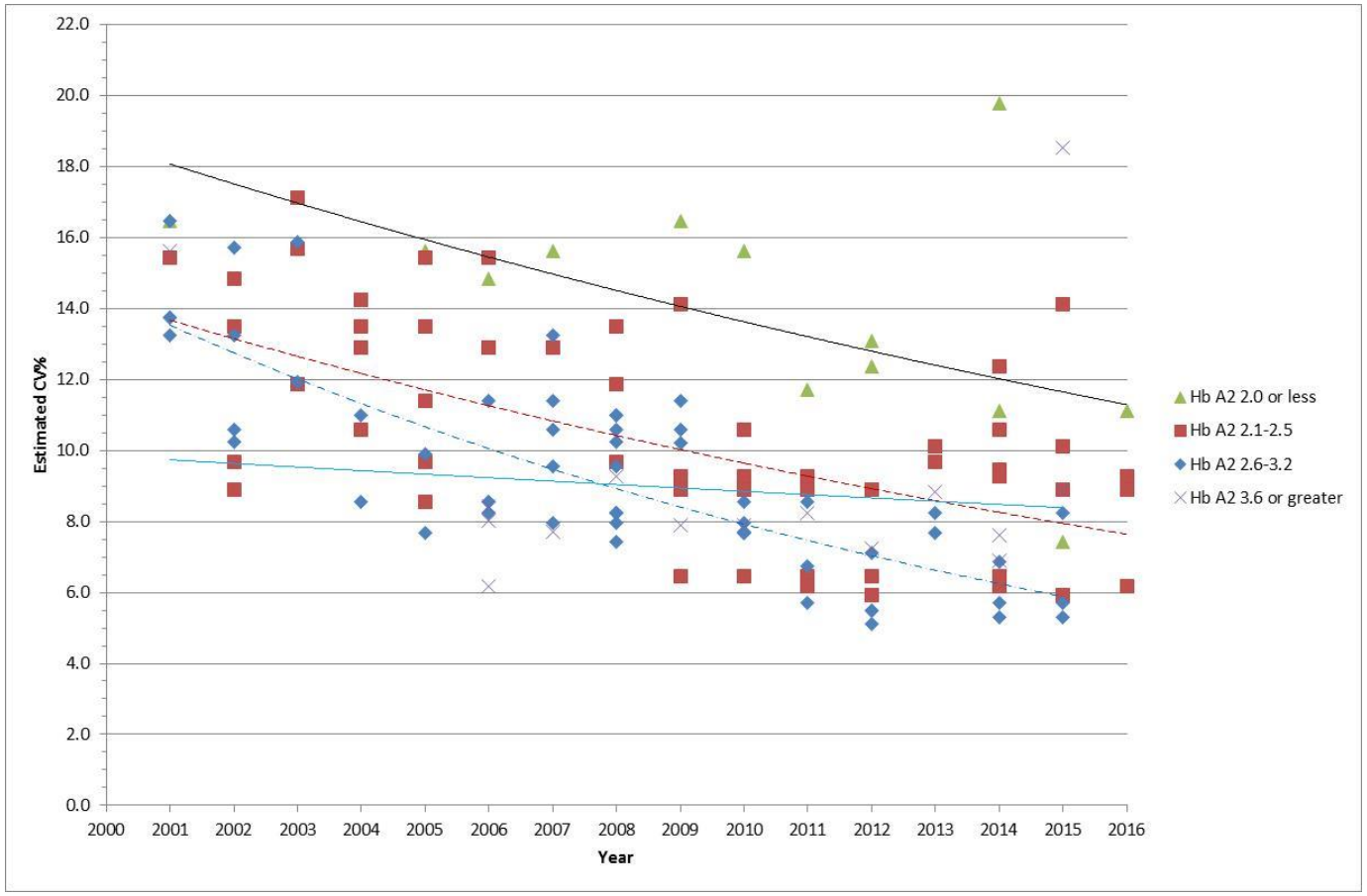


Figure 2.4. Trends in CV% values for HPLC Haemoglobin A_2 method performance in UK NEQAS Haematology from 2001 to 2016, by Hb A_2 value

2.7.4 Evaluation of Haemoglobin A_2 performance data (2006 to 2008)

The six most commonly used Hb A_2 quantitation methods in use by the participants during the period 2006 to 2008 were manual ion exchange microcolumn chromatography, three HPLC systems from Bio-Rad Laboratories (the Variant Classic with the beta thalassaemia short programme, the Variant II with the beta thalassaemia short programme and the Variant II with the dual kit programme), the Menarini Diagnostics HA8160 HPLC analyser and the Tosoh Biosciences G7 HPLC analyser. Figure 2.5 shows a box and whisker plot for the biases observed for each of the six sub-method (analyser) groups in the 2006 to 2008 study, in comparison to the 'acceptable EQA performance range' of ± 0.25% of Hb A_2.

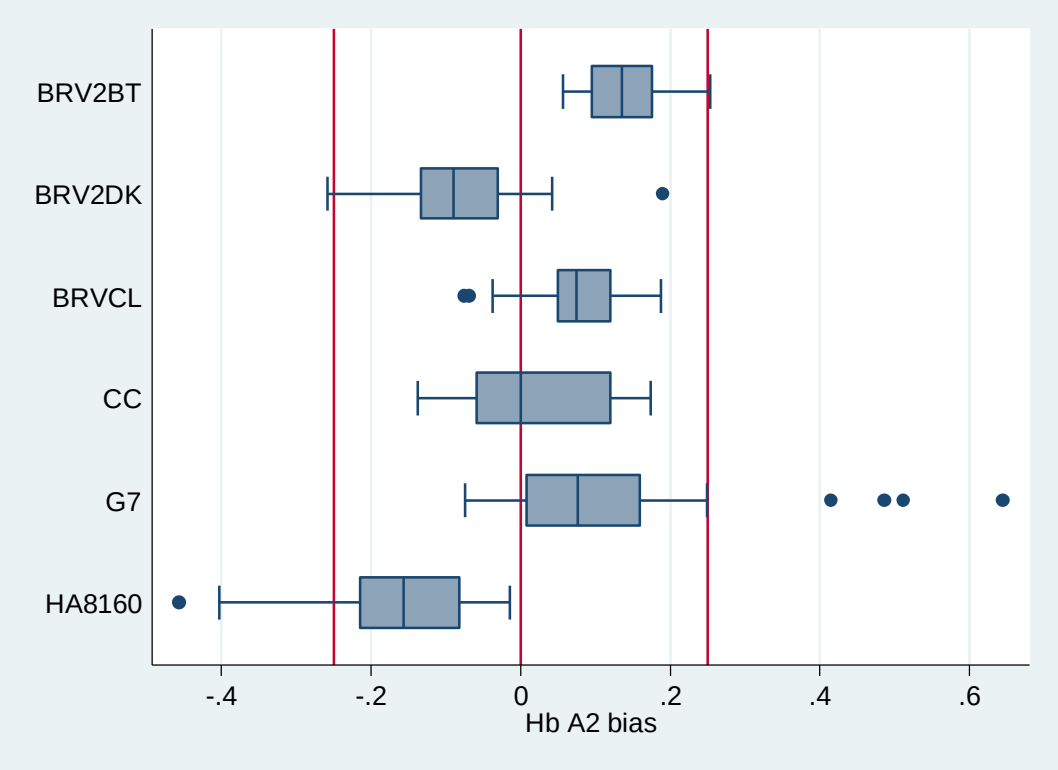


Figure 2.5. Distribution of the bias of sub-method trimmed mean values compared to the all method trimmed mean for Hb A_2 in UK NEQAS surveys from 2006 to 2008. The solid lines show ±0.25% performance limits.

Legend to Figure 2.5:

BRV2BT: Bio-Rad Variant II HPLC analyser with the beta thalassaemia short programme

BRV2DK: Bio-Rad Variant II HPLC analyser with the dual kit programme

BRVCL: Bio-Rad Variant Classic HPLC analyser

CC: Microcolumn chromatography

G7: Tosoh G7 HPLC analyser

HA8160: Menarini HA8160 HPLC analyser

The box and whisker plot shows a variation in the results of the different methods, although the median of each sub-method group and the distribution of the data for five out of six of the sub-method groups, excluding outliers, were within the ±0.25% performance range. For Bio-Rad HPLC analysers, the Bio-Rad Variant II analyser running the dual kit programme gave lower Hb A_2 SMTM values than the other two Bio-Rad methods. The Tosoh G7 sub-method group returned four of the twenty-seven sub-method mean results as outliers, all with a positive bias. The distribution of the means for the HA8160 sub-method group was negatively skewed in comparison to the acceptable EQA performance range, with the bottom whisker and one outlier mean result extending beyond the lower limit of –0.25% of Hb A_2.

2.7.5 Evaluation of Haemoglobin A_2 performance data (2013 to 2016)

During the time period studied (August 2013 to June 2016), thirty-four specimen pools with ALTM Hb A_2 values ranging from 1.47 to 5.02% were distributed. The data used for the study is shown in Table 2.2.

In total, 329 SMTMs were evaluated from ten analyser groups. The SMTM was only calculated where there were more than ten results by analyser group. The bias of the SMTM values against the ALTM by method group is shown in the box and whisker plot in Figure 2.6, using an ‘acceptable EQA performance range’ of ALTM ± 0.25% of Hb A_2.

Distribution	1304		1305	1306		1401		1402		1403		1404		1405		1406		
Specimen	**AH1**	**AH2**	**AH3**	**AH1**	**AH2**	**AH1**	**AH3**	**AH1**	**AH3**	**AH1**	**AH2**	**AH2**	**AH3**	**AH1**	**AH2**	**AH1**	**AH2**	**AH3**
ALTM Hb A_2 %	4.28	2.86	2.15	2.71	2.28	2.81	2.40	2.57	3.99	2.35	1.97	2.66	2.35	4.31	1.47	2.31	2.31	2.13
HPLC MTM Hb A_2 %	4.26	2.84	2.15	2.71	2.25	2.80	2.39	2.57	3.97	2.34	1.97	2.65	2.36	4.30	1.48	2.32	2.33	2.13
CE MTM Hb A_2 %	4.37	2.90	2.17	2.71	2.41	2.88	2.44	2.61	4.06	2.40	1.97	2.67	2.35	4.27	1.42	2.26	2.24	2.11
Analysers	**Number of results returned by specimen pool**																	
BRD10	29	29	29	27	27	29	29	27	27	29	29	27	27	29	29	29	29	29
BRV2 DK	20	20	19	20	20	19	19	20	20	21	20	22	21	24	24	24	24	24
BRV2 BTHAL	58	58	58	62	62	63	63	68	68	67	66	66	65	64	64	65	65	65
HA8160	41	41	42	42	42	40	40	43	43	41	42	39	39	37	36	34	34	34
PRIMUS	11	11	13	12	13	13	13	14	14	17	17	18	18	21	21	22	22	22
TOSOH G7	35	36	31	32	32	29	29	28	28	26	26	24	24	22	22	19	19	19
TOSOH G8	23	23	25	25	25	26	26	30	30	31	31	36	36	38	38	39	39	39
CE	22	22	24	23	23	24	24	24	24	24	24	23	23	25	25	24	24	24
CE 2	10	10	12	12	12	14	14	15	15	13	13	15	16	16	16	16	16	16
CE MINI	~	~	~	10	10	11	11	12	12	12	12	11	11	~	~	~	~	~
Total	**249**	**250**	**253**	**265**	**266**	**268**	**268**	**281**	**281**	**281**	**280**	**281**	**280**	**276**	**275**	**272**	**272**	**272**
	Sub-method trimmed mean Hb A_2 % (SMTM) by specimen pool																	
BRD10	4.40	2.93	2.13	2.77	2.25	2.84	2.33	2.55	3.97	2.39	1.88	2.78	2.48	4.36	1.52	2.43	2.41	2.17
BRV2 DK	4.25	2.76	1.97	2.63	2.14	2.73	2.24	2.47	3.97	2.30	1.84	2.60	2.31	4.37	1.39	2.30	2.28	2.05
BRV2 BTHAL	4.10	2.84	2.25	2.68	2.29	2.81	2.51	2.59	3.86	2.37	2.13	2.65	2.39	4.16	1.68	2.38	2.40	2.28
HA8160	4.03	2.82	2.09	2.71	2.24	2.68	2.22	2.53	3.77	2.38	1.79	2.66	2.38	4.08	1.44	2.34	2.30	2.06
PRIMUS	4.09	2.71	2.07	2.56	2.16	2.70	2.34	2.44	3.90	2.22	1.90	2.45	2.15	4.20	1.39	2.10	2.16	2.02
TOSOH G7	4.78	2.98	2.19	2.89	2.33	3.02	2.57	2.65	4.43	2.36	2.11	2.82	2.37	4.93	1.41	2.38	2.43	2.19
TOSOH G8	4.67	2.90	2.09	2.79	2.19	2.82	2.36	2.59	4.33	2.24	1.85	2.59	2.25	4.63	1.25	2.22	2.25	1.98
CE	4.33	2.88	2.16	2.70	2.41	2.87	2.44	2.59	4.05	2.42	1.96	2.68	2.35	4.26	1.41	2.25	2.23	2.11
CE 2	4.39	2.92	2.17	2.67	2.39	2.87	2.45	2.57	3.98	2.35	1.94	2.65	2.34	4.26	1.41	2.23	2.21	2.05
CE MINI	~	~	~	2.75	2.40	2.89	2.47	2.69	4.15	2.41	2.01	2.70	2.38	~	~	~	~	~

Table 2.2 Analysers and trimmed mean values for specimen pools 1304AH to 1603AH

(continued on next page)

Distribution	1501	1502			1503	1504		1505		1506		1601		1602		1603
Specimen	**AH1**	**AH1**	**AH2**	**AH3**	**AH1**	**AH1**	**AH2**	**AH2**	**AH3**	**AH2**	**AH3**	**AH1**	**AH2**	**AH1**	**AH3**	**AH2**
ALTM Hb A_2 %	2.71	5.02	2.08	1.58	2.59	2.78	2.24	2.48	1.98	2.45	2.52	2.45	2.39	2.37	1.94	2.45
HPLC MTM Hb A_2 %	2.74	5.04	2.10	1.60	2.59	2.78	2.24	2.48	2.00	2.46	2.52	2.45	2.40	2.38	1.96	2.47
CE MTM Hb A_2 %	2.63	4.90	1.97	1.52	2.58	2.77	2.24	2.48	1.94	2.40	2.51	2.43	2.34	2.31	1.90	2.39
Analysers	**Number in each distribution**															
BRD10	28	28	28	28	27	26	26	27	27	27	27	26	26	27	26	28
BRV2 DK	25	27	27	27	28	28	28	28	28	27	27	28	28	25	25	27
BRV2 BTHAL	64	68	68	68	68	66	66	69	69	70	68	70	70	73	73	74
HA8160	34	28	28	28	28	26	26	28	28	27	27	29	29	30	28	29
PRIMUS	20	17	17	17	17	19	19	19	19	20	20	23	23	26	26	25
TOSOH G7	22	18	18	18	16	14	14	15	14	15	15	12	12	10	10	~
TOSOH G8	40	44	44	44	44	44	44	43	43	44	44	45	45	46	46	47
CE	23	21	21	21	23	25	25	25	25	24	23	25	25	24	24	23
CE 2	19	18	18	18	19	18	18	20	19	21	21	29	29	31	30	27
CE MINI	~	10	10	10	~	10	10	10	10	11	11	12	12	12	12	13
Total	**275**	**279**	**279**	**279**	**270**	**276**	**276**	**284**	**282**	**286**	**283**	**299**	**299**	**304**	**300**	**293**
	Sub-method trimmed mean Hb A_2 % (SMTM)															
BRD10	2.87	5.31	2.30	1.66	2.67	2.89	2.30	2.59	2.05	2.62	2.69	2.64	2.50	2.56	2.07	2.75
BRV2 DK	2.73	5.03	1.99	1.37	2.56	2.75	2.12	2.44	1.85	2.38	2.46	2.40	2.30	2.31	1.78	2.75
BRV2 BTHAL	2.73	4.84	2.22	1.80	2.63	2.81	2.30	2.59	2.05	2.62	2.69	2.64	2.50	2.56	2.07	2.75
HA8160	2.73	4.67	2.14	1.46	2.58	2.70	2.11	2.49	1.91	2.44	2.54	2.43	2.33	2.35	1.79	2.51
PRIMUS	2.55	4.80	1.77	1.48	2.43	2.64	2.18	2.42	1.91	2.34	2.35	2.29	2.30	2.18	1.92	2.26
TOSOH G7	2.86	6.01	2.02	1.72	2.66	2.99	2.31	2.71	2.08	2.61	2.64	2.51	2.52	2.37	2.14	~
TOSOH G8	2.69	5.53	1.95	1.39	2.51	2.74	2.08	2.42	1.81	2.30	2.36	2.31	2.26	2.19	1.77	2.29
CE	2.62	4.88	1.97	1.52	2.58	2.77	2.23	2.47	1.92	2.41	2.51	2.42	2.34	2.30	1.89	2.40
CE 2	2.62	4.91	1.95	1.51	2.56	2.78	2.24	2.48	1.94	2.38	2.49	2.41	2.33	2.31	1.90	2.38
CE MINI	~	4.94	1.99	1.56	~	2.74	2.26	2.52	1.96	2.41	2.53	2.49	2.34	2.33	1.95	2.42

Table 2.2 Analysers and trimmed mean values for specimen pools 1304AH to 1603AH *(continued from previous page*)

Legend to Table 2.2

BRD10:	Bio-Rad D10 HPLC analyser
BRV2 DK:	Bio-Rad Variant II HPLC analyser running the Dual Kit programme
BRV2 BTHAL:	Bio-Rad Variant II HPLC analyser running the Beta Thalassaemia Short programme
HA8160:	Menarini HA8160 HPLC analyser
PRIMUS:	Primus Ultra 2 HPLC analyser
TOSOH G7:	Tosoh G7 HPLC analyser
TOSOH G8:	Tosoh G8 HPLC analyser
CE:	Sebia Capillarys CZE analyser
CE2:	Sebia Capillarys 2 CZE analyser
CE MINI:	Sebia Minicap CZE analyser
ALTM:	All laboratory trimmed mean
MTM:	Method trimmed mean
SMTM:	Sub-method trimmed mean

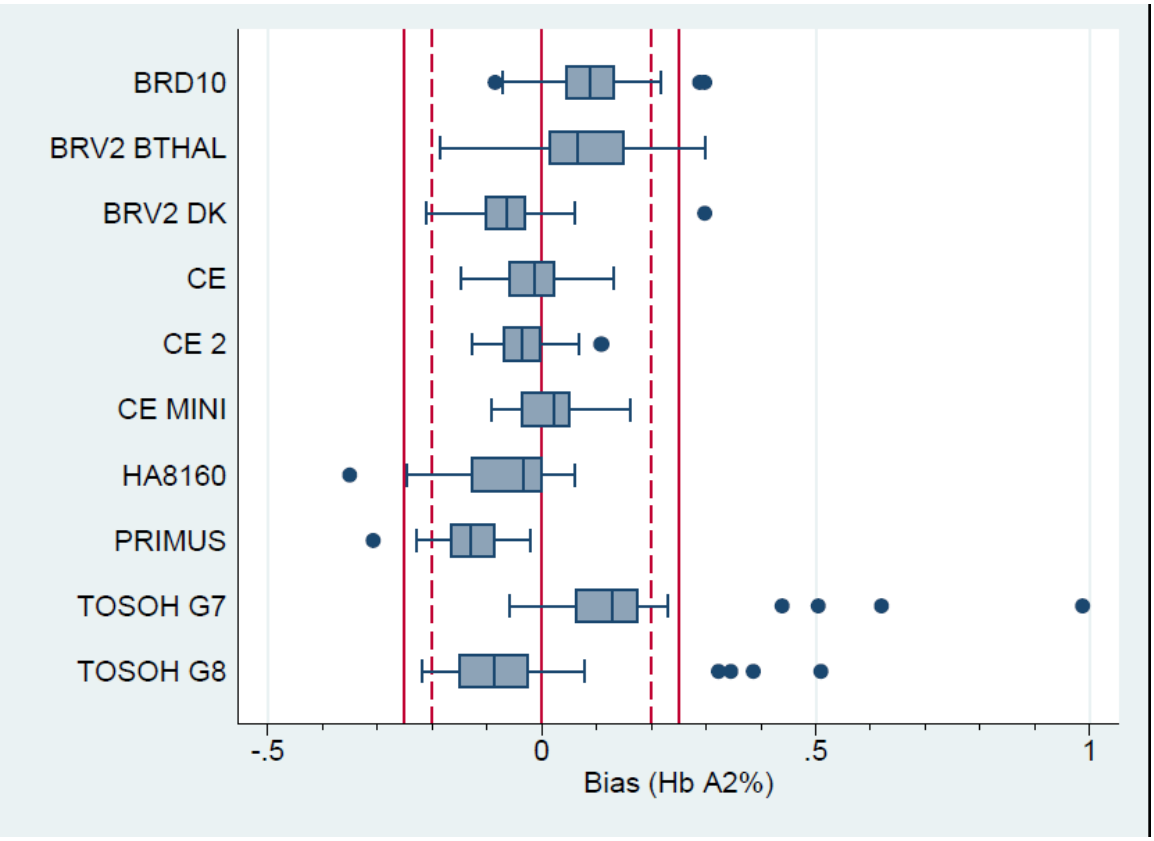


Figure 2.6. Distribution of the bias of sub-method trimmed mean values compared to the all method trimmed mean for Hb A_2 in UK NEQAS surveys from 2013 to 2016. The broken lines show the ±0.2% performance limits and the solid lines ±0.25% performance limits.

Legend to Figure 2.6:

BRD10:	Bio-Rad D10 HPLC analyser
BRV2 DK:	Bio-Rad Variant II HPLC analyser running the dual kit programme
BRV2 BTHAL:	Bio-Rad Variant II HPLC analyser running the beta thalassaemia short programme
HA8160:	Menarini HA8160 HPLC analyser
PRIMUS:	Primus Ultra 2 HPLC analyser
TOSOH G7:	Tosoh G7 HPLC analyser
TOSOH G8:	Tosoh G8 HPLC analyser
CE:	Sebia Capillarys CZE analyser
CE2:	Sebia Capillarys 2 CZE analyser
CE MINI:	Sebia Minicap CZE analyser

For all ten sub-methods, the median and IQRs of the SMTMs were within the ± 0.25% acceptable performance limits. There were a small number of outliers: most notably both the Tosoh G7 and G8 analysers produced four SMTMs that were high outliers. Further examination of the data showed that for both sub-methods, the outliers were specimen pools with raised Hb A_2 values (1304AH1, Hb A_2 ALTM = 4.3%; 1402AH1, Hb A_2 ALTM = 4.0%; 1405AH3, Hb A_2 ALTM = 4.3% and 1502AH1, ALTM Hb A_2 = 5.0%), suggesting a lack of linearity with high Hb A_2 values for the instruments. The HA8160 analysers had a single low outlier, with specimen pool 1502AH1. The Primus sub-method group had a single outlier SMTM (specimen pool 1502AH2, ALTM Hb A_2 = 2.1%) as did the Bio-Rad analysers, all with specimen pool 1603AH2, ALTM Hb A_2 = 2.5%.

The evaluation of the bias of the median results returned by method group takes no account of the range of the results within each sub-method. The potential impact of the variation in the results from one instrument to another of the same model is illustrated in Figures 2.7 and 2.8, where the results for the same survey material pools tested on different analysers of the same type showed a range of up to 2.0% of Hb A_2 for some instrument groups. This would be sufficient to alter the clinical outcome at or around the 3.0 to 4.0% Hb A_2 range. Figures 2.7 and 2.8, which were samples sent out in the same distribution suggest that for some instruments the bias against the ALTM may differ depending upon the Hb A_2 %. For example, there was a far greater proportion of the results from the Tosoh G7 and G8 groups outwith the acceptable performance limits for specimen 1502AH1 (ALTM Hb A_2 = 5.1%, Figure 2.7) than for specimen 1502AH2 (ALTM Hb A_2 = 2.1%, Figure 2.8).

Linear regression analysis of the SMTM (i.e. the analyser group values) against the MTM values (i.e. the 'expected' target results against which performance assessment is undertaken) was performed for the seven HPLC analyser groups and three Sebia CZE analysers (see Figures 2.9, 2.10 for HPLC methods and 2.11 for CZE methods). The linear regression for the HPLC methods confirms that there is lack of alignment between particular instruments and that this varies according to measurand concentration in some cases, e.g. for the Bio-Rad Variant II with the beta thalassaemia short programme (linear regression $y = 0.87x + 0.383$), the slope is lower but the intercept higher than for the other Bio-Rad instruments predicting that the results for raised Hb A_2 values will be lower and results at low Hb A_2 values higher with that model (Figure 2.9). For the Tosoh HPLC analysers (Figure 2.10), the slope is higher but the intercept is lower than for other HPLC analysers, predicting that raised Hb A_2 values will be higher and low Hb A_2 values will be lower than for other HPLC analysers. The results for the three Sebia CZE analysers are very closely aligned (Figure 2.11). The CZE analysers all show linear regression equations with a slope close to 1.0 with a very small intercept, indicating a close relationship between the observed SMTM and the target MTM and very comparable results from the three instrument models. For this reason, the results from the three Sebia CZE analysers (the Capillarys, Capillarys 2 and the Mini cap) have been consolidated for many of the remaining analyses in this and later chapters.

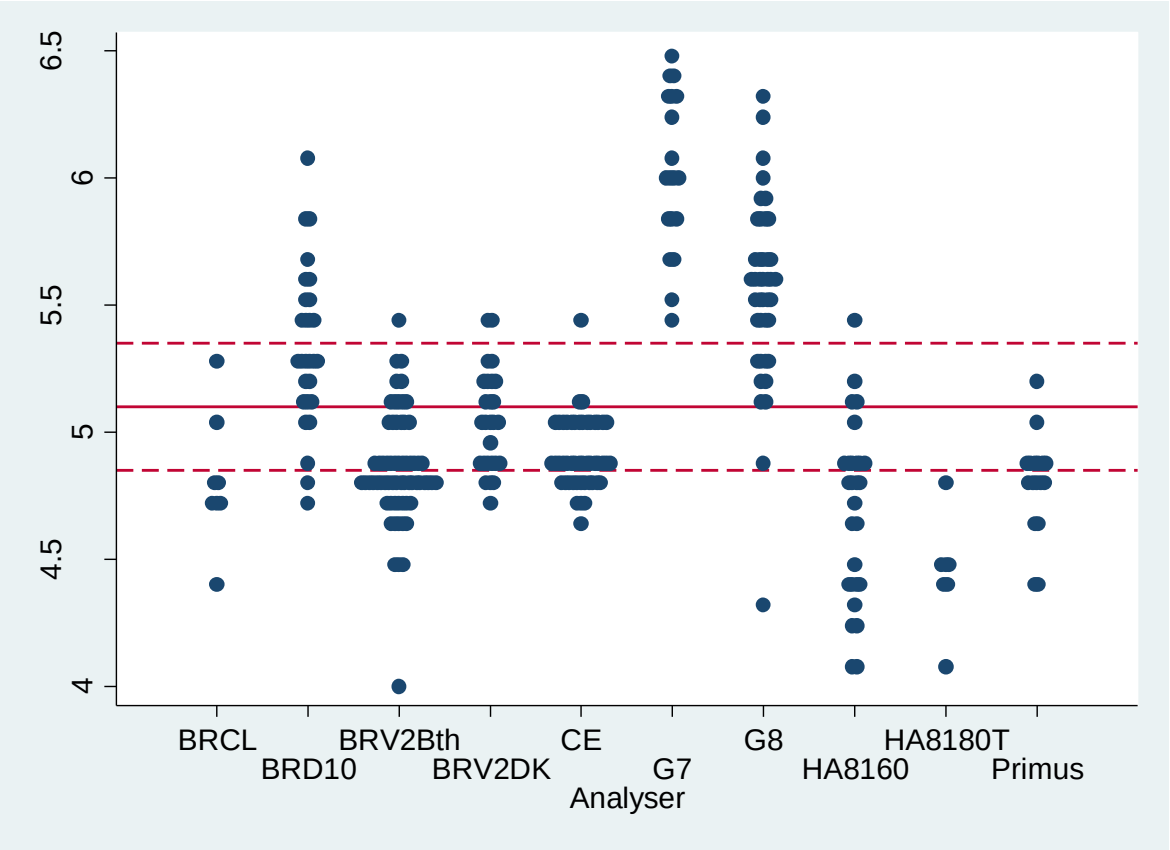


Figure 2.7 Distribution of Haemoglobin A_2 results by instrument group for UK NEQAS specimen 1502AH1 (All methods trimmed mean Hb A_2 value = 5.1%).

Legend to Figure 2.7:

BRCL:	Bio-Rad Classic HPLC analyser
BRD10:	Bio-Rad D10 HPLC analyser
BRV2Bth:	Bio-Rad Variant II HPLC analyser running the beta thalassaemia short programme
BRV2DK:	Bio-Rad Variant II HPLC analyser running the dual kit programme
CE:	Sebia CZE analysers
G7:	Tosoh G7 HPLC analyser
G8:	Tosoh G8 HPLC analyser
HA8160:	Menarini HA8160 HPLC analyser
HA8180T:	Menarini HA8180T HPLC analyser
PRIMUS:	Primus Ultra 2 HPLC analyser

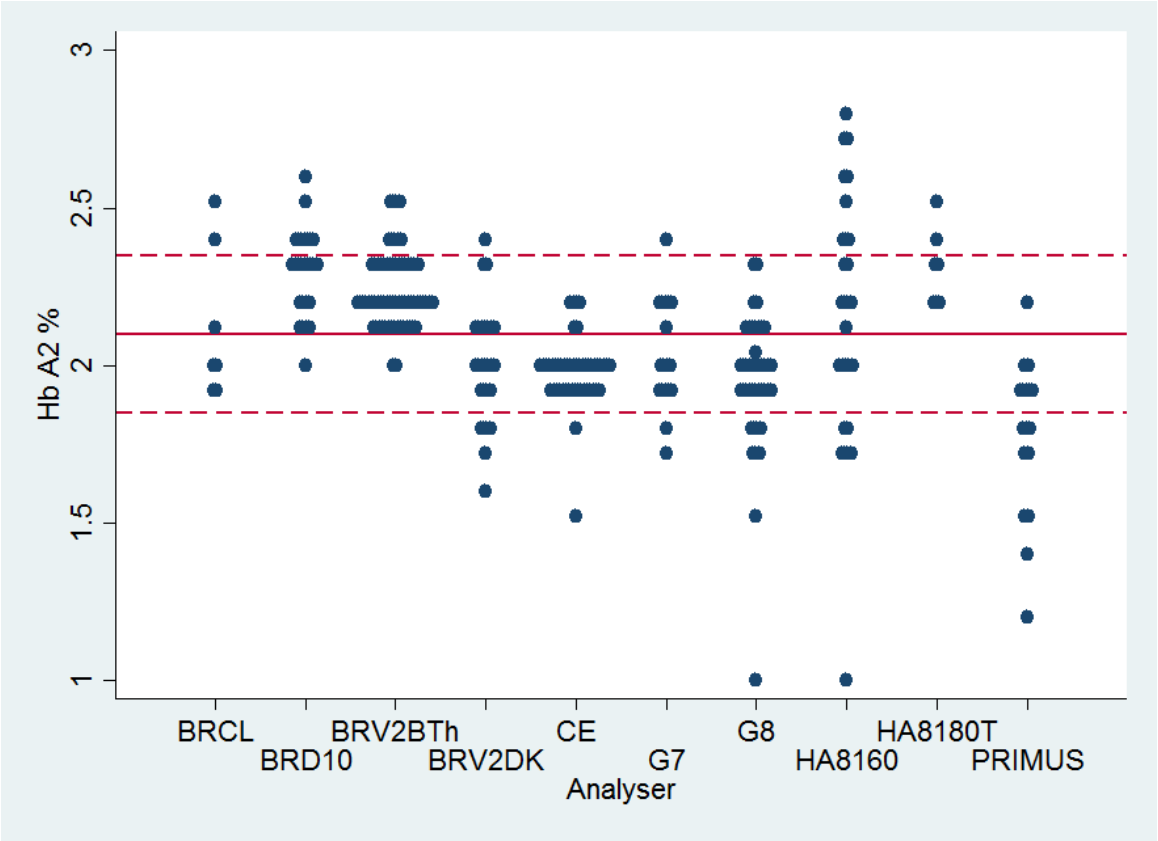


Figure 2.8 Distribution of Haemoglobin A_2 results by instrument group for UK NEQAS specimen 1502AH2 (All methods trimmed mean Hb A_2 value = 2.1%).

Legend to Figure 2.8:

BRCL:	Bio-Rad Classic HPLC analyser
BRD10:	Bio-Rad D10 HPLC analyser
BRV2BTh:	Bio-Rad Variant II HPLC analyser running the beta thalassaemia short programme
BRV2DK:	Bio-Rad Variant II HPLC analyser running the dual kit programme
CE:	Sebia CZE analysers
G7:	Tosoh G7 HPLC analyser
G8:	Tosoh G8 HPLC analyser
HA8160:	Menarini HA8160 HPLC analyser
HA8180T:	Menarini HA8180T HPLC analyser
PRIMUS:	Primus Ultra 2 HPLC analyser

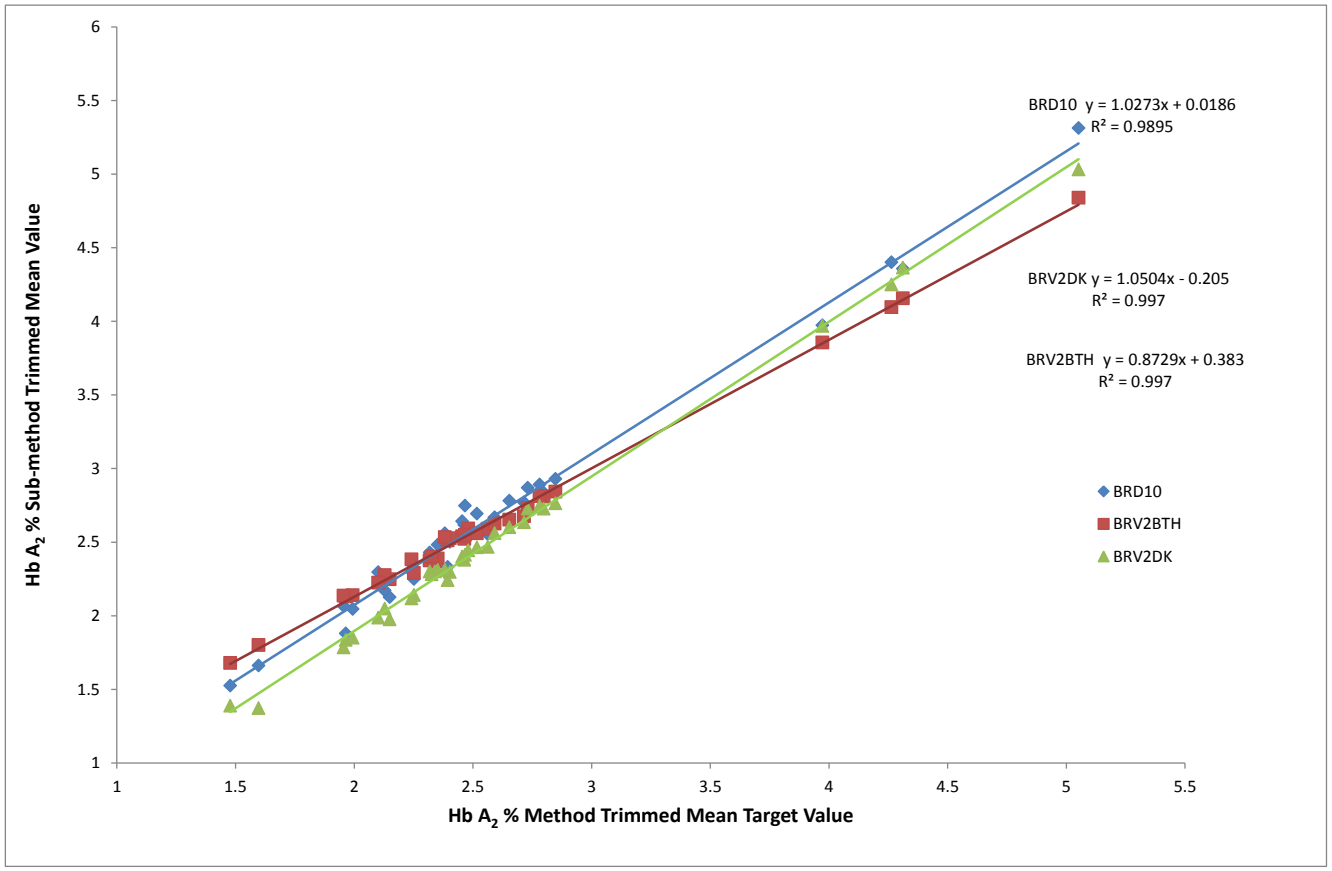


Figure 2.9 Linear regression of the SMTM values for Bio-Rad HPLC analysers against the HPLC MTM Haemoglobin A_2% target value for specimens 1304AH to 1603AH

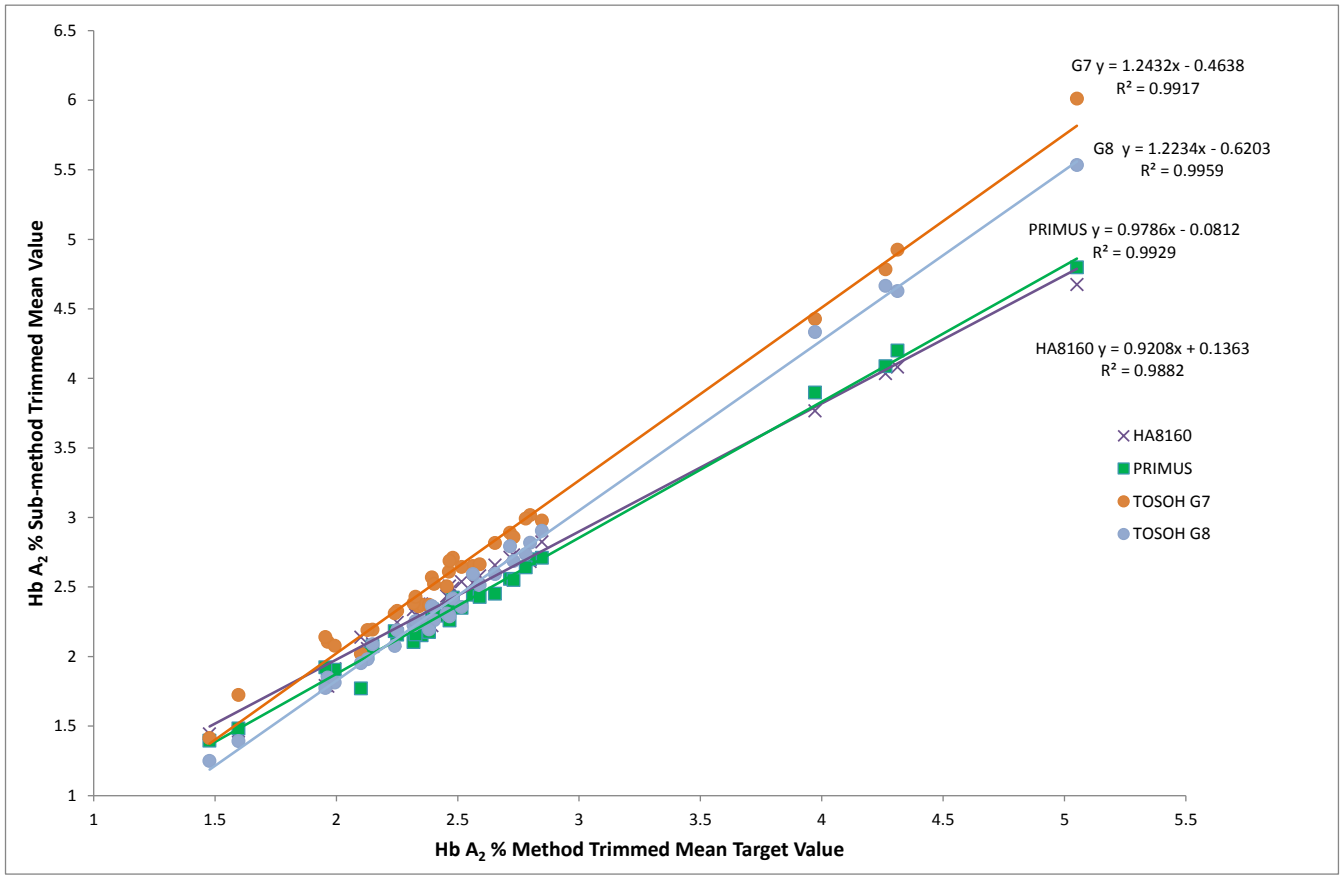


Figure 2.10 Linear regression of the SMTM values for non-Bio-Rad HPLC analysers against the HPLC MTM Haemoglobin A_2% target value for specimens 1304AH to 1603AH

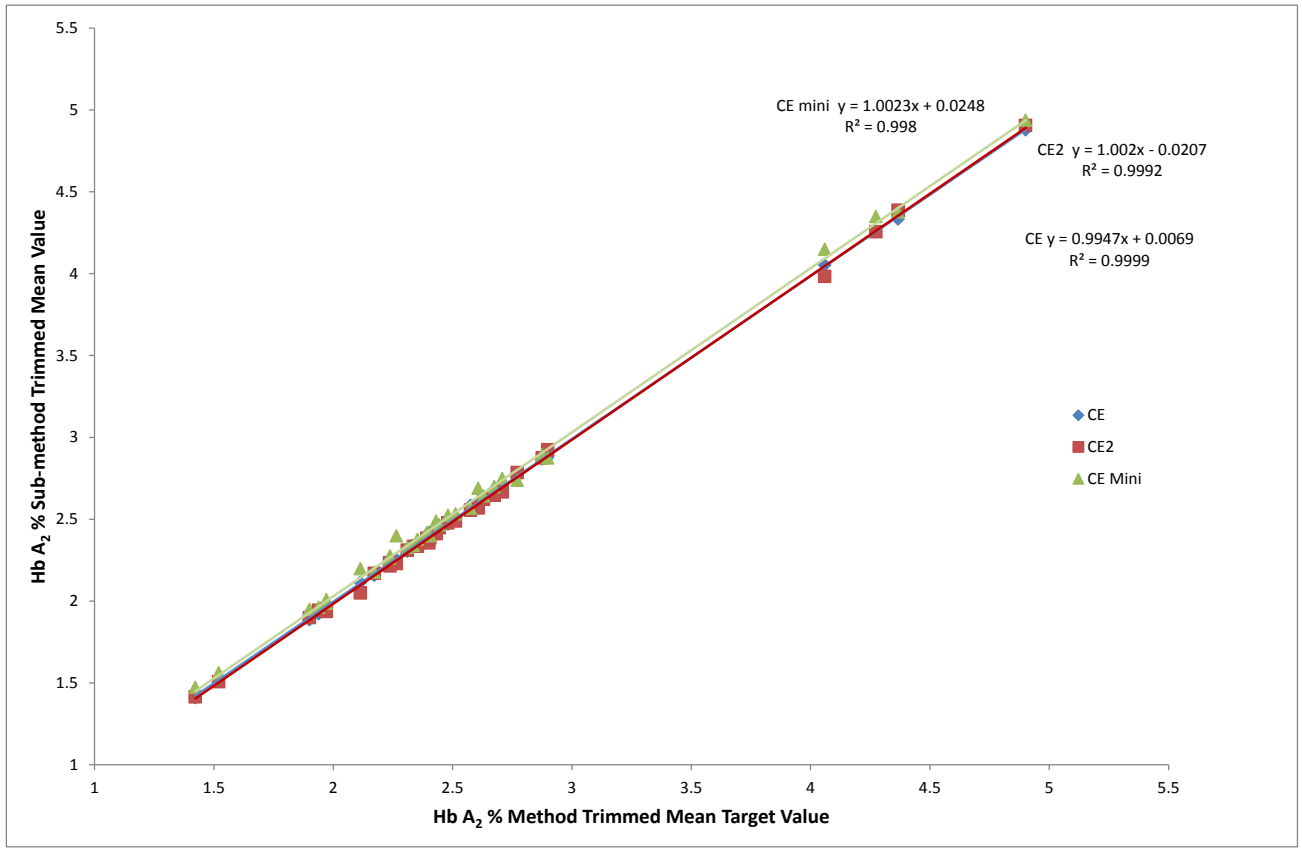


Figure 2.11 Linear regression of the SMTM values for Sebia capillary electrophoresis analysers against the capillary electrophoresis MTM Haemoglobin A_2% target value for specimens 1304AH to 1603AH

The Kruskal-Wallis test was used to test whether the differences between analyser groups was due to sampling alone. Table 2.3 shows the results of the Kruskal-Wallis (KW) test for all thirty-four specimen pools distributed between 2013 (1304AH) to 2016 (1603AH). The table gives the number of analysers in each group, the non-parametric statistics (median and IQR), the Chi-squared (χ^2) values, with the degrees of freedom. For every specimen pool, the KW test gave a probability (p value) of 0.0001 that the null hypothesis, i.e. that the differences observed were due to sampling only, would be true in 1/1000 instances. This indicates that the difference between the different analyser groups is statistically significant. The KW test does not specify which of the groups is/are different.

Specimen	Results by instrument group								χ^2 with ties with 7 degrees of freedom	Probability (*p*)
		BRD10	BRV2 Beta	BRV2 DK	CZE	HA8160	Primus	G7	G8	
1304AH1	N	29	58	20	41	41	11	35	23	133.81	0.0001
	Median	4.4	4.1	4.2	4.4	4.1	4.1	4.8	4.8		
	IQR	0.2	0.2	0.25	0.1	0.4	0.3	0.4	0.4		
1304AH2	N	29	58	20	41	41	11	36	23	38.19	0.0001
	Median	3.0	2.85	2.8	2.9	2.8	2.7	3.0	2.9		
	IQR	0.3	0.1	0.2	0.0	0.3	0.0	0.35	0.2		
1305AH3	N	29	58	19	45	42	13	31	25	38.51	0.0001
	Median	2.2	2.25	2.0	2.2	2.1	2.1	2.3	2.1		
	IQR	0.3	0.1	0.3	0.1	0.4	0.1	0.4	0.2		
1306AH1	N	27	62	20	45	42	12	32	25	45.88	0.0001
	Median	2.8	2.7	2.6	2.7	2.7	2.55	2.9	2.8		
	IQR	0.3	0.1	0.25	0.2	0.3	0.1	0.3	0.3		
1306AH2	N	27	62	20	45	42	13	32	25	51.45	0.0001
	Median	2.3	2.3	2.1	2.4	2.3	2.2	2.4	2.2		
	IQR	0.3	0.0	0.3	0.2	0.2	0.1	0.3	0.2		
1401AH1	N	29	63	19	49	40	13	29	26	56.42	0.0001
	Median	2.8	2.8	2.7	2.9	2.7	2.7	3.1	2.8		
	IQR	0.3	0.1	0.2	0.1	0.3	0.0	0.3	0.2		
1401AH3	N	29	63	19	49	40	13	29	26	86.13	0.0001
	Median	2.3	2.5	2.2	2.5	2.3	2.4	2.5	2.35		
	IQR	0.3	0.2	0.3	0.1	0.3	0.1	0.2	0.3		

Table 2.3 Kruskal-Wallis equality of rank test results for specimens 1304AH1 to 1603AH2 (*continued on next page*)

Specimen	Results by instrument group									χ^2 with ties with 7 degrees of freedom	Probability (*p*)
		BRD10	BRV2 Beta	BRV2 DK	CZE	HA8160	Primus	G7	G8		
1402AH1	N	27	68	20	51	43	14	28	30	33.42	0.0001
	Median	2.6	2.6	2.5	2.6	2.5	2.5	2.7	2.55		
	IQR	0.1	0.15	0.2	0.2	0.3	0.1	0.3	0.2		
1402AH3	N	27	68	20	51	43	14	28	30	133.26	0.0001
	Median	4.0	3.9	4.0	4.0	3.7	3.9	4.5	4.3		
	IQR	0.4	0.1	0.25	0.2	0.3	0.2	0.4	0.3		
1403AH1	N	29	67	21	49	41	17	26	31	35.99	0.0001
	Median	2.4	2.4	2.3	2.4	2.4	2.2	2.35	2.3		
	IQR	0.2	0.1	0.2	0.1	0.3	0.1	0.3	0.2		
1403AH2	N	29	66	20	49	42	17	26	31	107.58	0.0001
	Median	1.9	2.15	1.8	2.0	1.9	1.9	2.1	1.9		
	IQR	0.2	0.1	0.3	0.1	0.3	0.1	0.3	0.1		
1404AH2	N	27	66	22	49	39	18	24	36	60.45	0.0001
	Median	2.8	2.65	2.6	2.7	2.7	2.45	2.8	2.6		
	IQR	0.2	0.1	0.2	0.1	0.4	0.1	0.25	0.2		
1404AH3	N	27	65	21	50	39	18	24	36	60.56	0.0001
	Median	2.5	2.4	2.4	2.35	2.4	2.2	2.3	2.3		
	IQR	0.1	0.1	0.2	0.1	0.4	0.1	0.4	0.3		
1405AH1	N	29	64	24	49	37	21	22	38	131.53	0.0001
	Median	4.4	4.1	4.4	4.3	4.1	4.2	5.0	4.6		
	IQR	0.3	0.1	0.3	0.1	0.4	0.2	0.5	0.3		

Table 2.3 (continued)

Specimen	Results by instrument group								χ^2 with ties with 7 degrees of freedom	Probability (*p*)
		BRD10	BRV2 Beta	BRV2 DK	CZE	HA8160	Primus	G7	G8	
1405AH2	N	29	64	24	49	36	21	22	38	138.69	0.0001
	Median	1.5	1.7	1.4	1.4	1.5	1.4	1.4	1.2		
	IQR	0.1	0.1	0.15	0.1	0.3	0.0	0.1	0.1		
1406AH1	N	29	65	24	49	34	22	19	39	75.04	0.0001
	Median	2.4	2.4	2.35	2.2	2.4	2.1	2.4	2.2		
	IQR	0.2	0.1	0.2	0.1	0.3	0.2	0.3	0.3		
1406AH2	N	29	65	24	49	34	22	19	39	71.92	0.0001
	Median	2.4	2.4	2.3	2.2	2.35	2.2	2.4	2.2		
	IQR	0.2	0.2	0.2	0.1	0.3	0.1	0.3	0.3		
1406AH3	N	29	65	24	49	34	22	19	39	91.64	0.0001
	Median	2.2	2.3	2.1	2.1	2.05	2.0	2.2	2.0		
	IQR	0.1	0.1	0.3	0.0	0.3	0.1	0.2	0.2		
1501AH1	N	28	64	25	49	34	20	22	40	69.53	0.0001
	Median	2.9	2.7	2.7	2.6	2.8	2.55	2.9	2.7		
	IQR	0.15	0.1	0.2	0.1	0.4	0.1	0.3	0.26		
1502AH1	N	28	68	27	49	28	17	18	44	159.46	0.0001
	Median	5.3	4.8	5.0	4.9	4.8	4.8	6.0	5.6		
	IQR	0.4	0.25	0.3	0.2	0.5	0,.1	0.5	0.4		
1502AH2	N	28	68	27	49	28	17	18	44	127.29	0.0001
	Median	2.3	2.2	2.0	2.0	2.2	1.8	2.0	2.0		
	IQR	0.2	0.15	0.2	0.1	0.5	0.2	0.3	0.2		

Table 2.3 (continued)

Specimen	Results by instrument group									χ^2 with ties with 7 degrees of freedom	Probability (*p*)
		BRD10	BRV2 Beta	BRV2 DK	CZE	HA8160	Primus	G7	G8		
1502AH3	N	28	68	27	49	28	17	18	44	136.02	0.0001
	Median	1.65	2.2	1.4	1.5	1.55	1.5	1.8	1.4		
	IQR	0.1	0.15	0.3	0.1	0.45	0.2	0.2	0.2		
1503AH1	N	27	68	28	51	28	17	16	44	55.31	0.0001
	Median	2.7	2.6	2.6	2.6	2.6	2.4	2.65	2.55		
	IQR	0.2	0.1	0.15	0.1	0.25	0.1	0.15	0.3		
1504AH1	N	26	66	28	53	26	19	14	44	50.51	0.0001
	Median	2.9	2.8	2.8	2.8	2.8	2.6	2.95	2.7		
	IQR	0.2	0.1	0.25	0.1	0.4	0.1	0.4	0.35		
1504AH2	N	26	66	28	53	26	19	14	44	95.40	0.0001
	Median	2.3	2.4	2.1	2.2	2.2	2.2	2.3	2.1		
	IQR	0.2	0.1	0.2	0.1	0.4	0.1	0.4	0.3		
1505AH2	N	27	69	28	55	28	19	15	43	69.90	0.0001
	Median	2.6	2.6	2.4	2.5	2.5	2.4	2.7	2.5		
	IQR	0.2	0.2	0.1	0.1	0.3	0.2	0.3	0.2		
1505AH3	N	27	69	28	54	28	19	14	43	117.03	0.0001
	Median	2.1	2.2	1.85	1.9	1.9	1.9	2.1	1.9		
	IQR	0.3	0.1	0.1	0.1	0.3	0.0	0.4	0.2		
1506AH2	N	27	70	27	56	27	20	15	44	69.19	0.0001
	Median	2.6	2.5	2.4	2.4	2.5	2.4	2.6	2.3		
	IQR	0.2	0.1	0.2	0.0	0.3	0.1	0.3	0.35		

Table 2.3 (continued)

Specimen	Results by instrument group								χ^2 with ties with 7 degrees of freedom	Probability (*p*)
		BRD10	BRV2 Beta	BRV2 DK	CZE	HA8160	Primus	G7	G8	
1506AH3	N	27	68	27	55	27	20	15	44	84.99	0.0001
	Median	2.7	2.6	2.5	2.5	2.6	2.35	2.6	2.4		
	IQR	0.2	0.1	0.3	0.0	0.3	0.1	0.3	0.35		
1601AH1	N	26	70	28	66	29	23	12	45	95.36	0.0001
	Median	2.7	2.5	2.45	2.4	2.4	2.3	2.5	2.3		
	IQR	0.3	0.1	0.2	0.1	0.4	0.2	0.3	0.2		
1601AH2	N	26	70	28	66	29	23	12	45	108.20	0.0001
	Median	2.5	2.5	2.3	2.3	2.4	2.3	2.5	2.3		
	IQR	0.2	0.1	0.2	0.1	0.3	0.2	0.4	0.2		
1602AH1	N	27	73	25	67	30	26	10	46	145.97	0.0001
	Median	2.6	2.5	2.3	2.3	2.35	2.1	2.3	2.2		
	IQR	0.1	0.1	0.2	0.1	0.3	0.2	0.2	0.2		
1602AH3	N	26	73	25	66	28	26	10	46	144.34	0.0001
	Median	2.1	2.1	1.8	1.9	1.8	1.9	2.2	1.8		
	IQR	0.2	0.1	0.2	0.0	0.35	0.1	0.5	0.2		
1603AH2	N	28	74	27	63	29	25	10	46	143.59	0.0001
	Median	2.8	2.6	2.4	2.4	2.5	2.3	2.6	2.3		
	IQR	0.3	0.1	0.2	0.0	0.3	0.1	0.5	0.2		

Table 2.3 (continued)

Legend to Table 2.3

BRD10: Bio-Rad D10 with the dual kit programme

BRV2 Beta: Bio-Rad Variant 2 with the beta thalassaemia short programme

BRV2 DK: Bio-Rad Variant 2 with the dual kit programme

CZE: Sebia CZE analysers

HA8160: Menarini HA8160

PRIMUS: Primus Ultra 2

G7: Tosoh G7

G8: Tosoh G8

2.7.6 Borderline Haemoglobin A_2 specimens

The NHS STP's laboratory screening algorithm sets a Hb A_2 value of 3.5% or above with a MCH of less than 27 pg for the identification of beta thalassaemia carriers in antenatal screening. Although it is recognised that haemoglobinopathy screening will not detect 100% of cases, the misdiagnosis of an antenatal woman with a borderline Hb A_2 result, e.g. as a result of her being a 'silent' beta thalassaemia carrier, will result in her partner not being invited for screening and may result in the unexpected birth of an infant with a severe haemoglobinopathy if co-inherited with another beta thalassaemia mutation or clinically significant Hb variant (Weatherall and Clegg, 2001). If there is overestimation, unnecessary additional investigation or partner testing may be undertaken, increasing anxiety in the patient and costs to the laboratory service. Even if the results of all screening methods were aligned, the normal dispersion of results around a statistical mean value of 3.5% will result in 50% of the results falling below the cut-off and 50% above; the concern is that the risk of misdiagnosis may be increased depending on the analysis method used.

The distributions of the data returned for each of three specimens with an ALTM Hb A_2 of 3.6 to 3.7% are shown in the box-and-whisker plots in Figures 2.12 (0604AH4, Hb A_2 3.7%), 2.13 (0902AH1, Hb A_2 3.7%) and 2.14 (1105AH3, Hb A_2 3.6%) for the six main HPLC instrument groups, with twelve to fifty-two instruments in each group. Each plot shows the 'acceptable performance limits' of the expected result ±0.25% of Hb A_2. The distribution of results varies by instrument type in median result, IQR and bias against the MTM value for the specimen. Examination of the actual data returned demonstrated that the proportion of results that would have been identified as 'normal' varied by

instrument. Using a cut-off Hb A_2 value of 3.5%, 16/28 (57%) participants using the Menarini HA8160 would have reported 0604AH4 as normal on the basis of the Hb A_2 result alone, for the other specimens the proportion was 31/43 (72%) for 0902AH1 and 36/52 (69%) for 1105AH3. By comparison, for users of the Bio-Rad Variant II analyser with the beta thalassaemia short programme, the proportions were 6/37 (16%) (0604AH4), 4/46 (9%) (0902AH1) and 28/51 (55%) (1105AH3) and the Tosoh G7 group 2/21 (10%) (0604AH4), 2/29 (7%) (0902AH1) and 2/32 (6%) (1105AH3). The results for all the analyser groups are shown in Table 2.4.

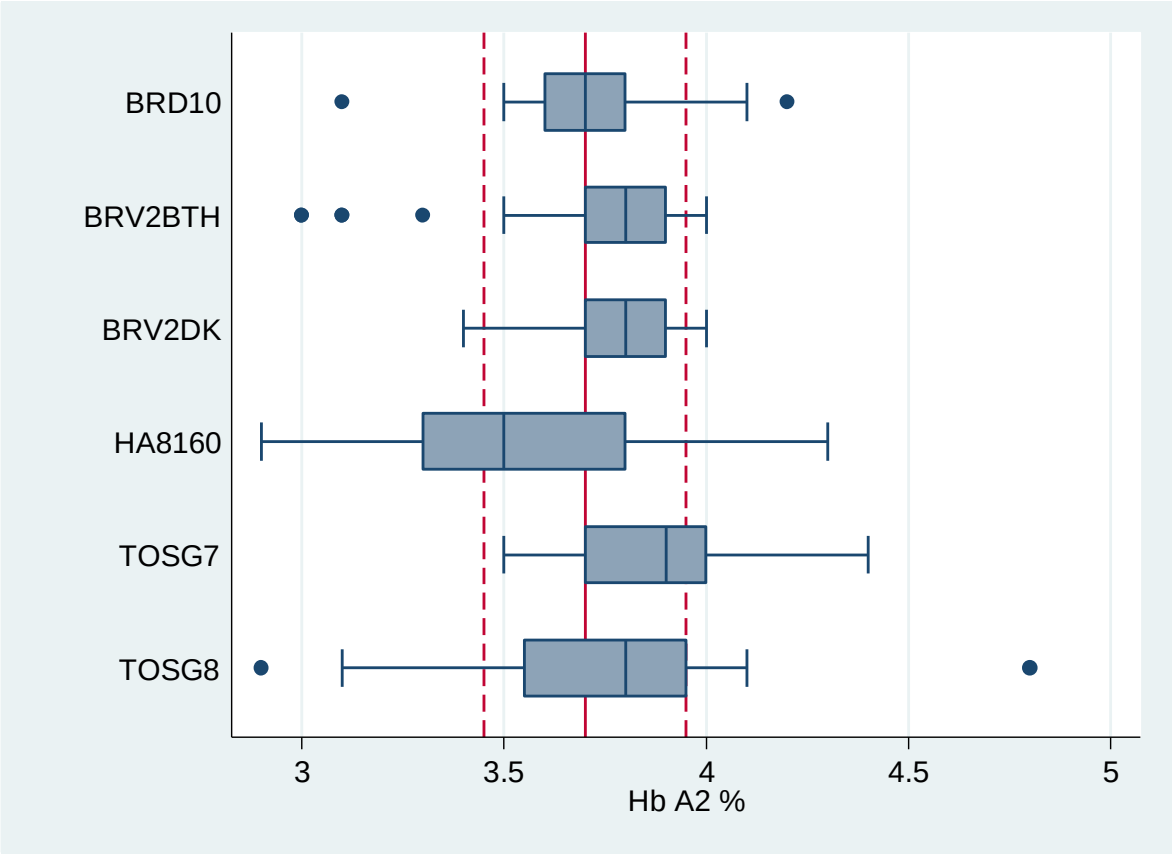


Figure 2.12 Distribution of Haemoglobin A_2 results by analyser type for specimen 0604AH4, MTM Hb A_2 3.7%.
The broken lines show the ±0.25% performance limits.

Legend to Figure 2.12

BRD10: Bio-Rad D10 with the dual kit programme

BRV2BTH: Bio-Rad Variant II with the beta thalassaemia short programme

BRV2DK: Bio-Rad Variant II with the dual kit programme

HA8160: Menarini HA8160

PRIMUS: Primus

TOSG7: Tosoh G7

TOSG8: Tosoh G8

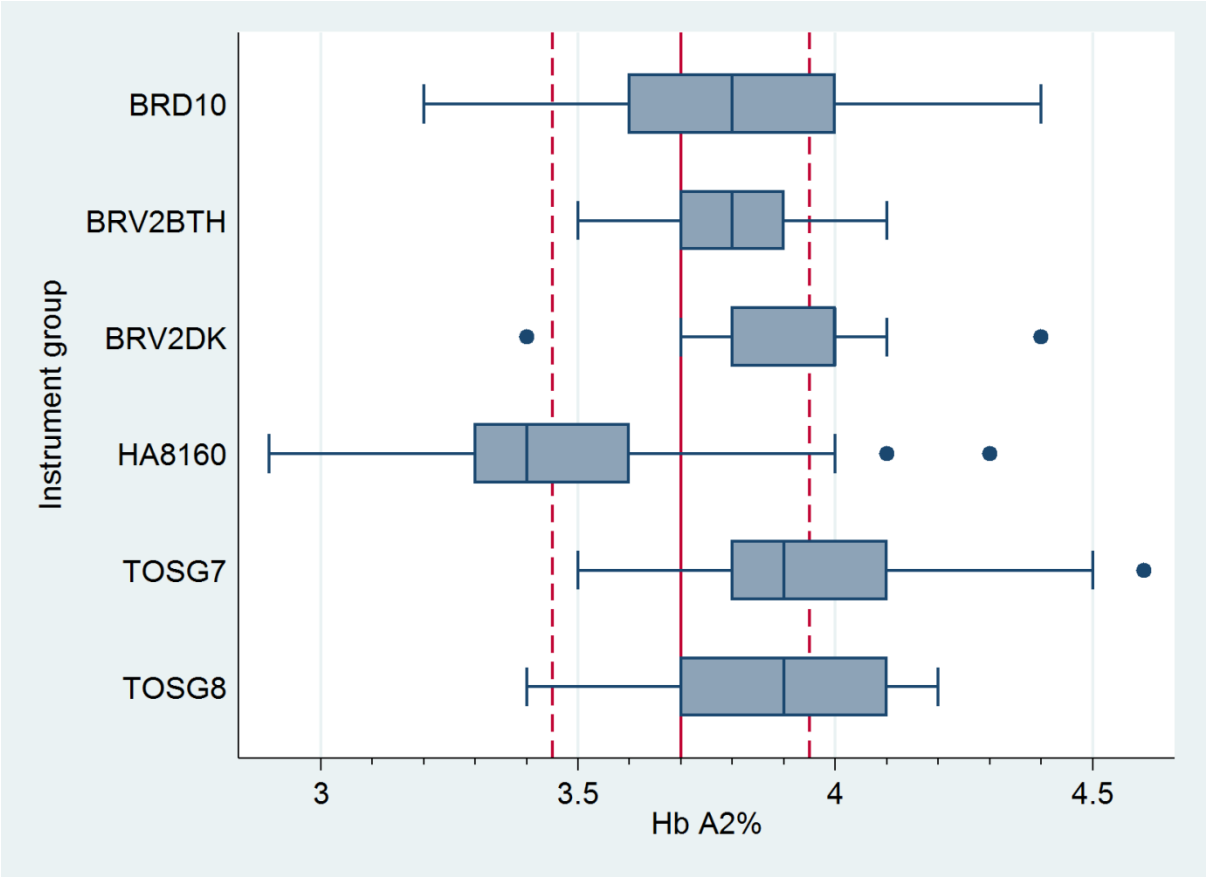


Figure2.13 Distribution of Haemoglobin A_2 results by analyser type for specimen 0902AH1, MTM Hb A_2 3.7%.
The broken lines show the ±0.25% performance limits.

Legend to Figure 2.13

BRD10: Bio-Rad D10 with the dual kit programme

BRV2BTH: Bio-Rad Variant II with the beta thalassaemia short programme

BRV2DK: Bio-Rad Variant II with the dual kit programme

HA8160: Menarini HA8160

PRIMUS: Primus Utra 2

TOSG7: Tosoh G7

TOSG8: Tosoh G8

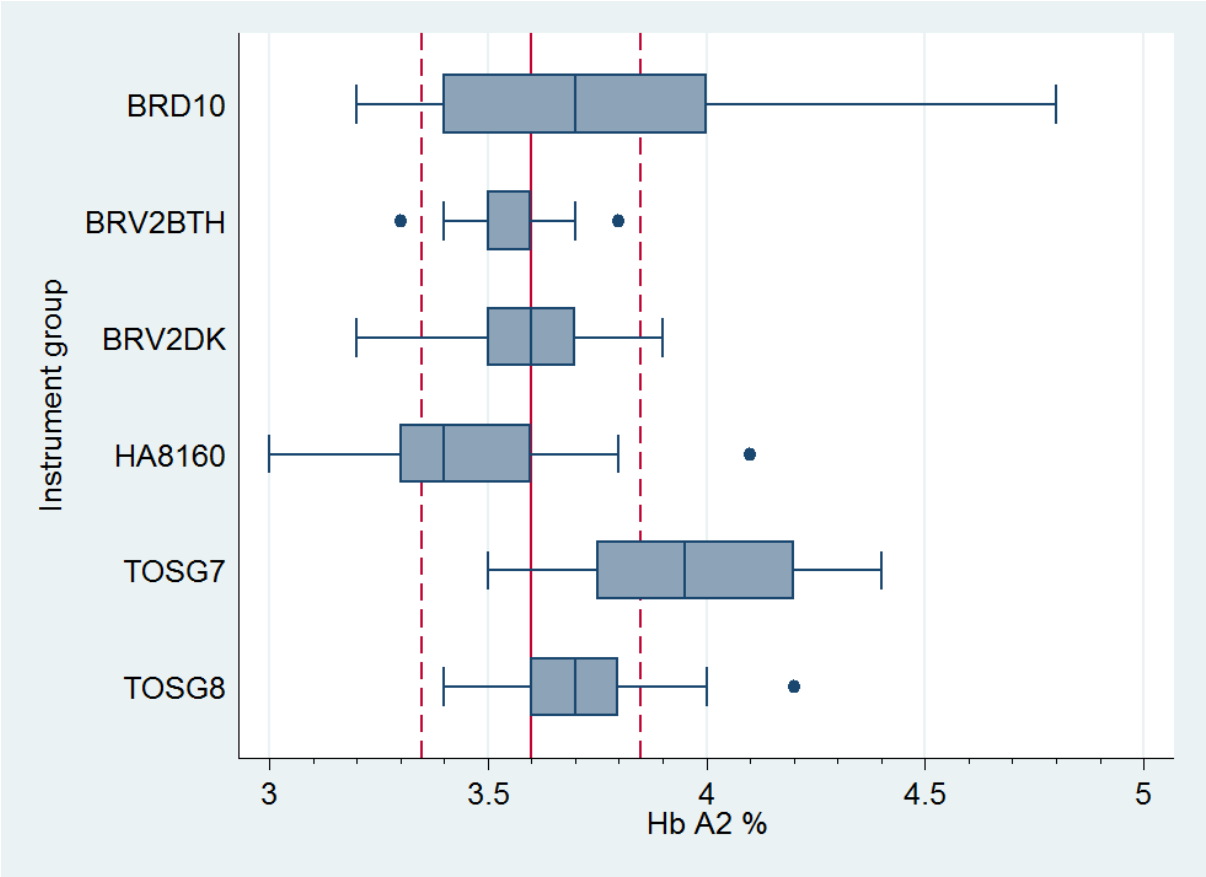


Figure 2.14 Distribution of Haemoglobin A_2 results by analyser type for specimen 1105AH3, MTM Hb A_2 3.6%.
The broken lines show the ±0.25% performance limits.

Legend to Figure 2.14

BRD10: Bio-Rad D10 with the dual kit programme

BRV2BTH: Bio-Rad Variant II with the beta thalassaemia short programme

BRV2DK: Bio-Rad Variant II with the dual kit programme

HA8160: Menarini HA8160

PRIMUS: Primus

TOSG7: Tosoh G7

TOSG8: Tosoh G8

Specimen	MTM Hb A_2 %	Method	No. labs (total)	"Normal" Hb A_2 ≤3.5%	
				N	%
0604AH4	3.7	BRD10	13	2	15%
		BRV2BTH	37	6	16%
		BRV2DK	12	1	8%
		HA8160	28	16	57%
		TOSG7	21	2	10%
		TOSG8	12	3	25%
0902AH1	3.7	BRD10	19	4	21%
		BRV2BTH	46	4	9%
		BRV2DK	17	1	6%
		HA8160	43	31	72%
		TOSG7	29	2	7%
		TOSG8	13	3	23%
1105AH3	3.6	BRD10	27	12	44%
		BRV2BTH	51	28	55%
		BRV2DK	19	6	32%
		HA8160	52	36	69%
		TOSG7	32	2	6%
		TOSG8	17	2	12%

Table 2.4. The numbers of results returned as normal by HPLC for three borderline Haemoglobin A_2 specimens.

Legend to Table 2.4

BRD10: Bio-Rad D10 with the dual kit programme

BRV2BT: Bio-Rad Variant II with the beta thalassaemia short programme

BRV2DK: Bio-Rad Variant II with the dual kit programme

HA8160: Menarini HA8160

TOSG7: Tosoh G7

TOSG8: Tosoh G8

2.7.7 Haemoglobin A_2 reference intervals in use

In the absence of a fixed action point or cut-off for a measurand, laboratories and manufacturers may set or recommend a reference interval peculiar to the instrument or method, acknowledging the difference between methodologies. This approach may cause confusion amongst service users, especially where patients are tested by different hospitals or diagnostic testing providers. The Pathology Harmony project attempted to standardise reference intervals for common measurands in clinical chemistry (Berg and Lane, 2011), although the Haematology Harmony Project only managed to agree nomenclature and units of measurement for the extended FBC (De la Salle, 2012b).

UK NEQAS Haematology requests the reference interval for Hb A_2 used by participants for a normal adult male at each distribution of the Abnormal Haemoglobins scheme as this can be of use when assessing the interpretation of results returned.

Figure 2.15 shows the overall range of reference values reported as in use by participants at the 1603AH distribution in June 2016. The minimum reported reference value by any participant was zero per cent Hb A_2 and the maximum was 4.0%. The mode (most commonly occurring) values for the lower and upper limits of the reference intervals quoted are shown in Table 2.5, the range overall is from 1.5% to 3.5% although there is some variation by instrument type, especially at the lower limit where the mode was 2.0% and 2.2% for some instruments. Only CZE instrument users did not have a mode value of 3.5% for the upper reference interval limit, with a value of 3.2%.

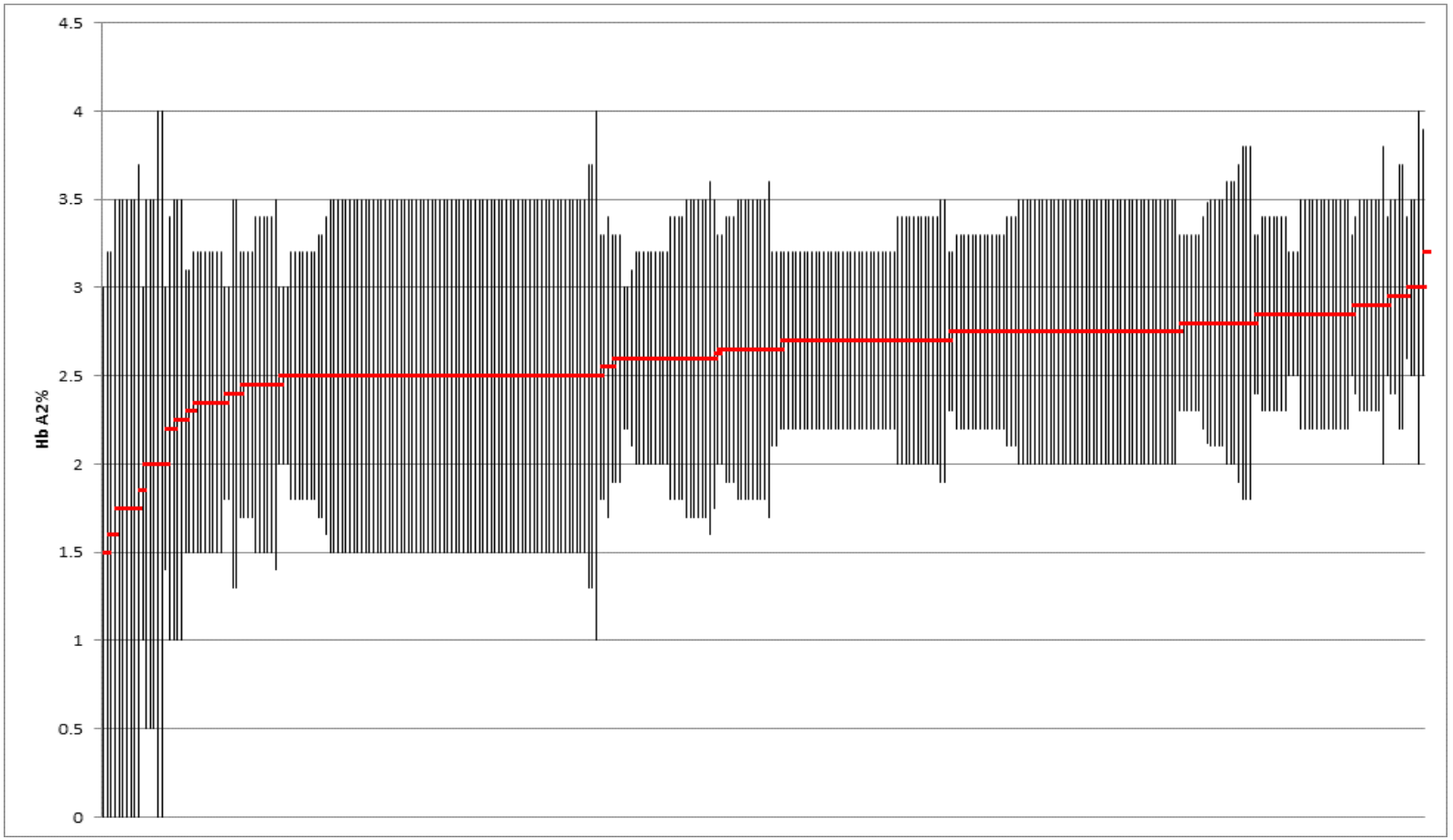


Figure 2.15. Haemoglobin A_2 reference intervals for a normal adult male reported as in use by participants in June 2016

METHOD	Hb A_2% reference interval	
	Min	Max
ALL	1.5	3.5
BRD10	2.2	3.5
BRV2 BTH	2.0	3.5
BRV2 DK	1.7	3.5
HA8160	1.5	3.5
HA8180	1.5	3.5
PRIMUS	1.5	3.5
TOSG7	2.0	3.5
TOSG8	2.0	3.5
CE	1.5	3.5
CE MINI	2.2	3.2
CE2	2.2	3.2

Table 2.5. The most commonly used minimum and maximum values of Haemoglobin A_2 % reference intervals by instrument type, as reported to UK NEQAS Haematology in June 2016

Legend to Table 2.5:

BRD10: Bio-Rad D10 HPLC analyser

BRV2 BTH: Bio-Rad Variant II HPLC analyser running the beta thalassaemia s

BRV2 DK: Bio-Rad Variant II HPLC analyser running the dual kit programme

HA8160: Menarini HA8160 HPLC analyser

HA8180: Menarini HA8180T HPLC analyser

PRIMUS: Primus Ultra 2 HPLC analyser

TOSG7: Tosoh G7 HPLC analyser

TOSG8: Tosoh G8 HPLC analyser

CE: Sebia Capillarys CZE analyser

CE2: Sebia Capillarys 2 CZE analyser

CE MINI: Sebia Minicap CZE analyser

2.7.8 European Network for Congenital and Rare Anaemias exercise (2010)

The median Hb A_2 results were 4.7% and 2.3% for the beta thalassaemia carrier and the normal donor specimen respectively. The results showed 95% confidence limits of 3.96 to 5.44% for the beta thalassaemia specimen and 2.00 to 2.60% for the normal specimen; two laboratories reported results below 3.5% for the beta thalassaemia specimen and no laboratory reported a raised Hb A_2 result for the normal specimen. The observed bias compared to the median for the major instrument types ranged from –8.4% to +17.4% for the beta thalassaemia specimen with 4/7 instrument types showing a bias of more than 5%, with the same pattern of bias by instrument type as shown in the historical UK NEQAS data. The bias shown by individual laboratories was greater, ranging from –30% to +36% across both specimens, reinforcing the need for standardisation of Hb A_2 measurement. The study was part of a presentation at the Thalassaemia International Federation meeting in 2012, the proceedings of which were published (see Publications).

2.8 Discussion

In the UK, population screening for beta thalassaemia carriers is offered in early pregnancy to allow couples an informed reproductive choice. A total of 14,073[1] antenatal women in England had a positive screening result for a significant haemoglobinopathy or thalassaemia in 2015/16, of which 3,755 were identified as beta thalassaemia carriers (NHS STP Data Report 2015/16). Although the newborn screening programme is not designed to detect babies with beta thalassaemia major, there were thirty infants detected with Hb F only at birth in the UK during 2015/16, which are likely to be cases of beta thalassaemia major or intermedia (NHS STP Data Report 2015/16). Although these numbers are relatively small, beta thalassaemia is a disease requiring lifelong transfusion support with life-changing and life-limiting complications (Galanello and Origa, 2010). Beta thalassemia mutations are also significant in the compound heterozygous state with other structural haemoglobin variants, e.g. Hb S and Hb E.

Molecular techniques are recommended for the definitive diagnosis of beta thalassaemia and to confirm an at-risk pregnancy: however, the large number of patients screened initially requires a continued reliance on protein chemistry methods. On a global perspective, the majority of births affected by thalassaemia disorders occur in low-income countries where diagnosis depends upon non-molecular techniques (Piel, 2016). The necessity for simple screening techniques is illustrated by the development of a non-quantitative, one tube osmotic fragility test for thalassaemia screening, utilising the increased resistance to lysis of

[1] These figures are likely to be a slight underestimate because of gaps in data reporting by some laboratories (NHS Sickle and Thalassaemia Programme Data Report 2015/16)

thalassaemic red cells in a buffered 0.36% salt solution, as a cost effective method for population screening in resource limited countries (Chow, 2005; Fucharoen and Winichagoon, 2012).

UK NEQAS historical Hb A_2 data has shown a near total adoption of automated Hb A_2 analysis methods since 2001. Improvement was seen in the performance of Hb A_2 measurement as measured by a reduction in est. CV% in UK NEQAS results in the predominant automated method (HPLC) in the period. The average HPLC method group est. CV% fell from 14.25% in 2001 to 8.87% in 2016, a value comparable to the CV% of between 6.0 to 9.6% reported in a comparison of three HPLC models from the same manufacturer (Paleari 2007). Since only HPLC methods were included in the analysis, the improvement is not the result of a transfer from manual to automated techniques, although the removal of older, less reliable HPLC systems may have had an influence. Other reasons for the improvement in overall performance may include a greater awareness of quality management and sources of error amongst both participants and instrument manufacturers as a result of the educational and oversight role of organisations such as the NHS STP and the European Network for Rare and Congenital Anaemias (ENERCA) project. These initiatives have set standards for screening services, published educational information (the NHS STP Antenatal Laboratory Handbook, 2017; Vives-Corrons, 2014) and offered training. The importance of quality management and accreditation has been promoted by the introduction of the international accreditation standard ISO 15189 for medical laboratories (ISO 15189:2012).

Healthcare science professionals must understand inter-method variation to ensure equity of diagnostic outcome for patients. The study has shown significant

differences when the same specimen is analysed by different Hb A_2 measurement methods, sufficient to produce a different clinical outcome in the borderline Hb A_2 region. The lack of an international standard to which Hb A_2 results are traceable has been recognised as a barrier to quality improvement since it is not possible to say which method gives the 'correct' result. At the time of writing, the only reference material available is the WHO International Reference Reagent for Hb A_2 (WHO 89/666), which was issued in 1989 prior to the widespread introduction of automated methods of Hb A_2 measurement. The IFCC has undertaken preliminary work on the development of a new standard (Paleari, 2010).

Despite the recommendation of a cut-off value for Hb A_2 of 3.5% for the identification of adult beta thalassaemia carriers, a number of laboratories still quote reference intervals with differing upper limits. This may reflect a varied approach to establishment of reference intervals (Barth, 2009; Sinclair, 2014), including a failure to update reference intervals as methods change. Of the twenty laboratories in the study who quoted an upper Hb A_2 reference interval limit of greater than 3.5%, ten were located in the UK or Republic of Ireland. Just three were within the oversight of the NHS STP in England, suggesting that most laboratories within the NHS STP's remit were using an upper reference interval that conformed to (or was less than) the 3.5% cut-off Hb A_2 value for screening for beta thalassaemia carriers.

Historical EQA data is a substantial resource that can be mined for trends in performance and variability between methods.

Interlaboratory comparison of performance through EQA can offer support to oversight bodies and national screening programme initiatives, monitoring trends in performance of both methods and individual laboratories. However, variability in the EQA scheme design may confound the interpretation of trends in performance (Miller 2011).

The first challenge in EQA is the nature and quality of the survey material distributed. The ideal is a material that is commutable across all instruments, stable in transit and for the period of the EQA distribution, homogeneous, traceable to an international standard and in a format that is comparable to patients' material (Miller, 2006; De la Salle, 2017). The EQA provider should also provide cases that test the participating laboratories at the limits of clinical decision making, reporting the results with background educational content where needed to enhance the educational aspects of the scheme. To provide a long-term, retrospective assessment of performance, the scheme must adopt a sensitive, fair and statistically robust performance scoring method that supports the needs of participating laboratories and other stakeholders (Coucke, 2017). Subsequent chapters will examine how well aspects of the services provided by UK NEQAS Haematology for the performance assessment of Hb A_2 measurement meet these ideals in terms of the nature of the survey material provided and the Hb A_2 scoring methods.

Chapter 3: Survey material for the performance assessment of Haemoglobin A_2 measurement

3.1 Aim

To assess the impact of the anticoagulant used and the age of survey material on the performance assessment of Hb A_2 measurement.

3.2 Introduction

The specimen of choice for Hb A_2 measurement is whole blood anti-coagulated with the potassium or sodium salts of ethylenediaminetetraacetic acid (EDTA), although any anticoagulant can be used (Bain 2006, Stephens 2012, Mosca 2009). Shokrani (Shokrani 2000) used heparin or EDTA-anticoagulated samples in a study of Hb A_2 measurement by HPLC and Maharini (Maharini 2014) used citrate phosphate dextrose (CPD) anticoagulated donor blood to assess frequency of Hb variant and thalassaemia carriers amongst blood donors. The use of EDTA for Hb A_2 measurement is largely one of convenience since it is the anticoagulant used for FBC samples in the haematology laboratory (Banfi 2007) and its use for Hb A_2 measurement allows utilisation of the same specimen, reducing the number of samples taken and facilitating reflex testing when a beta thalassaemia carrier is suspected from the FBC results. EDTA is also the anticoagulant recommended for Hb A_{1c} measurement by HPLC by the IFCC (Jeppsson 2002); therefore, the use of the same specimen type for Hb A_2 quantitation simplifies the use of analysers with dual Hb A_{1c} and haemoglobinopathy functionality.

EDTA anticoagulated blood is not a practical solution for haematology EQA in general (Urassa 2016, De la Salle 2017). Patients' samples in EDTA are available in relatively small specimen volumes and pooling of specimens to provide sufficient blood for an EQA exercise would require a large number of specimens from patients with the same clinical condition and compatible ABO blood groups.

Although EDTA blood is the best anticoagulant for the preservation of blood cell parameters and morphology for diagnostic purposes, CPD (as used for donor blood for transfusion) is a better option for stabilisation of red cells for periods of several weeks (Beutler 1983). For an EQA scheme with a substantial number of participants in a range of geographical locations, the practical option for the preparation of whole blood survey material is CPD blood from volunteer donors (De la Salle 2017, Lewis 1995). With more than three hundred participants registered for Hb A_2 measurement, UK NEQAS Haematology requires up to 500 mL of blood per survey material pool in order to provide each laboratory with 1.0 to 1.5 mL of blood for testing with sufficient additional specimens for in-house homogeneity and stability testing and repeat samples. Non-clinical issue blood is collected through NHSBT from anonymised, consented donors and screened for known blood-borne infectious agents prior to issue. A single donor unit is of sufficient volume for one survey material pool. UK NEQAS Haematology uses donations either from donors with no known haemoglobin variant or thalassaemia or from a selected panel of donors known to be beta thalassaemia or Hb variant carriers.

The NHS STP advises that the results of antenatal haemoglobinopathy screening should be available at least as an interim report within three days of receipt of the sample by the laboratory (NHS STP Antenatal Laboratory Handbook, 2017). This turnaround is dictated by the clinical need for the timely review of results and arrangement of testing of the baby's father, if indicated. Although analysis should be undertaken as soon as possible after collection, EDTA-anticoagulated blood specimens for Hb A_2 measurement have been reported to be stable for up to 21 days post venesection if stored at 2 to 8 °C (Tietz 1990). The blood used for the

preparation of UK NEQAS Hb A_2 survey material is seventy-two hours old when issued by NHSBT. The procedure for bottling, packing and shipping of the EQA survey material means that specimens are fourteen days post venesection when distributed as EQA material and twenty-four days old on survey closing day.

A major challenge in EQA is to provide survey specimens that behave in a similar fashion to patients' clinical material and are commutable across all testing platforms and methods (Miller, 2003; Ceriotti, 2014). The purpose of the study described in this chapter was to investigate whether the anticoagulant used and the age of the sample when analysed have an impact on the Hb A_2 results obtained, which might affect the performance of participants in the EQA surveys compared to their testing of clinical material.

The study described in this chapter was undertaken by the author at the commencement of the evaluation of Hb A_2 performance and was approximately ten years old at the time of writing. It was derived only from manual and HPLC methods in use by participating laboratories at the time. This neither compromises the comparison of the UK NEQAS survey material with fresh EDTA blood nor the demonstration of the stability of survey material over time, since it would be a reasonable experimental design to undertake the work with a single testing method. The only question that might arise is the commutability of the survey material with new methods as they come to market, which is a constant challenge in any EQA programme and is discussed later in the chapter.

3.3 Materials and methods

3.3.1 Survey material

A consented, volunteer male donor was venesected under medical supervision in a Haematology clinic for a total of 400 mL of blood: the first 200 mL of the donation was taken into sterile di-potassium ethylenediaminetetraacetic acid (K_2EDTA) solution in 0.9% w/v sodium chloride intravenous infusion BP solution at a final concentration of 1.5 to 2.0 mg EDTA per mL of blood; the second 200 mL of the donation was taken into CPD in a standard blood donor pack with 50% of the CPD anticoagulant removed. The blood was not tested for blood-borne infectious agents as the donor had been tested prior to venesection for HIV and hepatitis B, thus removing any delay to the distribution of the exercise.

3.3.2. Phase One: EDTA specimen preparation and distribution

This phase of the exercise was designed to assess performance of Hb A_2 measurement with a survey material as similar to a patient's EDTA sample as possible. The EDTA-anticoagulated blood was dispensed without further treatment in 1 mL aliquots into 2 mL volume, plain glass specimen vials (no anticoagulant), labelled with a unique survey identifier and despatched by first class post within twenty-four hours of venesection to 160 UK laboratories registered for Hb A_2 quantitation.

3.3.3 Phase Two: CPD specimen preparation and distribution

This phase of the exercise was designed to replicate the usual survey material for Hb A_2 measurement with blood taken from the same donor and at the same time as in phase one but prepared by the same procedures and in the same timeframe as a standard UK NEQAS Abnormal Haemoglobins sample. The CPD anticoagulated blood was stored for seventy-two hours at 2 to 8 °C immediately post-venesection, so that it was the same age when preparation started as the whole blood received from NHSBT for routine survey material pools. At seventy-two hours, the blood was thoroughly mixed by inversion and transferred to a sterile, 500 mL Duran® bottle. Antibiotics were added (0.1 mL per 500 mL of whole blood of 200 mg/mL benzyl penicillin sodium BP reconstituted in 0.9% w/v sterile sodium chloride intravenous infusion BP solution; 0.625 mL per 500 mL of whole blood of 40 mg/mL Genticin injectable solution) (Reardon, 1991) and the blood was mixed for ten minutes on a roller mixer before being dispensed in 1 mL aliquots into 2 mL volume, plain glass (no anticoagulant) specimen vials. The specimens were labelled with a unique exercise identifier and returned to storage at 2 to 8 °C for a further ten days, to simulate the usual time taken for the preparation and packing of a UK NEQAS survey distribution, after which the specimens were dispatched by first class post to the same 160 participants as in phase one.

3.3.4 Survey material analysis schedule

Participants were instructed to test both phase one (EDTA-anticoagulated) and phase two (CPD-anticoagulated) specimens by their standard Hb A_2 method and to record both the date of receipt and testing. To ensure that results were returned in a timely fashion, a closing date of ten days post-distribution was given for each

phase of the exercise. This is the same 'survey open' period used for UK NEQAS Abnormal Haemoglobins exercises. Allowing for twenty-four to forty-eight hours in the post and twenty-four hours for receipt and processing of the samples in the participant's laboratory, it was anticipated that the majority of phase one specimens would be three days and the majority of phase two specimens sixteen days post venesection when tested. Drawing from previous experience of the pattern of testing in regular UK NEQAS exercises, it was anticipated that 25 to 50% of samples would be tested later than the predicted date, ranging up to the closing date of the exercise.

3.3.5 Statistical analysis

Statistical analysis was undertaken using STATA 14 statistical software or Microsoft Excel. Because some of the instrument groups contained fewer than twenty results, non-parametric methods (median, est.SD and est. CV) were used for the calculation of descriptive statistics. Where there were five results or fewer for an instrument, the results were included in the ‘all methods’ group.

One-sample t-tests were calculated as shown in Equation 3.1 where $\overline{x}$ is the mean of the differences between the phase one and phase two results, μ_0 is the hypothesised mean, s is the standard deviation and n is the number of observations. The value for μ_0 (the hypothesised mean) was set at zero, i.e. the null hypothesis is that there is no difference between the phase one (EDTA) and phase two (CPD) specimens.

$$t = \frac{\overline{x} - \mu_0}{s/\sqrt{n}}$$

Equation 3.1 Calculation of one-sample t-tests

3.4 Results

3.4.1 Numbers of results returned by analyser / method

Results were received for each phase of the experiment from 144/160 laboratories (90%) and for both phases from 138/160 laboratories (86%). The Hb A_2 methods used by participants are shown in Table 3.1. Results were returned from manual, microcolumn chromatography and HPLC methods, which were representative of the methods in use at the time.

3.4.2 Analysis of Haemoglobin A_2 results by anticoagulant and specimen age

The summary statistics for phases one and two are shown in Table 3.2. Median and est. SD results were calculated because some of the method groups contained fewer than ten results. Individual method group statistics were not calculated for instruments where there were five or fewer results available.

Two results returned in phase one were excluded from the data analysis as the samples were tested at fifteen days post-venesection, i.e. after the exercise closing date, giving 142 results included in the data analysis for phase one. One result was excluded from the phase two data analysis as the sample was tested at thirty-one days post-venesection, giving 143 results in phase two.

Method	Phase 1	Phase 2
Column Chromatography	60	61
Bio-Rad Variant Classic HPLC	13	14
Bio-Rad Variant II / Dual programme HPLC	9	8
Bio-Rad Variant II / Beta Thalassaemia Short Programme HPLC	32	32
Bio-Rad D10 HPLC	2	2
Menarini HA8160 HPLC	12	11
Primus HPLC	5	5
TOSOH G7 HPLC	11	11
ALL METHODS	**144**	**144**

Table 3.1. Numbers of results returned by analyser or method type for Haemoglobin A_2 measurement in phases one and two

Phase 1	Method	N	Hb A_2 Results					
			Median	IQR	Est SD	Est CV%	Min	Max
	CC	59	2.7	0.5	0.3706	13.73	1.9	3.3
	BRD10	2	~	~	~	~	2.3	2.7
	BRV2Bth	31	2.6	0.2	0.1483	5.70	2	3
	BRV2DK	9	2.4	0.3	0.2224	9.27	2.1	2.5
	BRVClass	13	2.6	0.1	0.0741	2.85	2.2	2.7
	HA8160	12	2.25	0.25	0.1853	8.24	2.1	2.7
	Primus	5	~	~	~	~	2.2	2.4
	Tosoh G7	11	2.7	0.4	0.2965	10.98	2.1	3.2
	ALL METHODS	**142**	**2.6**	**0.4**	**0.2965**	**11.40**	**1.9**	**3.3**
Phase 2	**Method**	**N**	Hb A_2 Results					
			Median	IQR	Est SD	Est CV%	Min	Max
	CC	61	2.7	0.3	0.2224	8.24	2.2	3.3
	BRD10	2	~	~	~	~	2.7	2.9
	BRV2Bth	32	2.6	0.2	0.1483	5.70	2.2	3.1
	BRV2DK	8	2.3	0.15	0.1112	4.83	2.2	2.6
	BRVClass	13	2.5	0.1	0.0741	2.97	2.4	2.7
	HA8160	11	2.2	0.4	0.2965	13.48	2	2.8
	Primus	5	~	~	~	~	2.1	2.3
	Tosoh G7	11	2.7	0.5	0.3706	13.73	2.2	3
	ALL METHODS	**143**	**2.6**	**0.4**	**0.2965**	**11.40**	**2.0**	**3.3**

Table 3.2. Summary statistics for Haemoglobin A_2 results returned in phases one and two

Legend to Table 3.2:

CC:	Microcolumn chromatography
BRD10:	Bio-Rad D10 HPLC analyser
BRVBTh:	Bio-Rad Variant II HPLC analyser, using the beta thalassaemia short programme
BRV2DK:	Bio-Rad Variant II HPLC analyser, using the dual programme kit
BRVClass:	Bio-Rad Variant Classic HPLC analyser
HA8160:	Menarini HA8160 HPLC analyser
Primus385:	Primus HPLC analyser
TOSG7:	Tosoh G7 HPLC analyser
IQR:	Interquartile range
Est. SD:	Estimated standard deviation
Est. CV%:	Estimated coefficient of variation

The all methods summary statistics for both phases were identical, with a median Hb A_2 result of 2.6% and an est. CV% of 11.4%. The range of Hb A_2 results was 1.9 to 3.3% (phase one) and 2.0% to 3.3% (phase two). The median Hb A_2 values for the phase one specimens by instrument group ranged from 2.25% (Menarini HA8160 group) to 2.7% (microcolumn chromatography and Tosoh G7 groups). For the phase two specimens the median Hb A_2 values ranged from 2.2% (Menarini HA8160 group) to to 2.7% (microcolumn chromatography and Tosoh G7 groups). The CV% for most of the individual method groups was less than the overall value of 11.4%; for phase one only the microcolumn chromatography and the Tosoh G7 groups both had an est. CV% greater than 10%, for phase two only the Menarini HA8160 and the Tosoh G7 method groups had an est. CV% greater than 10%.

3.4.3 Effect of specimen age on Haemoglobin A_2 results

In phase one the specimens were between two and ten days old when tested, with two other results returned on specimens at fifteen days post venesection. The mode for the age of the specimens on testing was three days (37/144 specimens, 26%). A total of 68/144 (47%) were three days old or less, 105/144 (73%) were seven days old or less and 142/144 (98.6%) were ten days old or less. Figure 3.1 shows the results returned by specimen age and by instrument group in phase one. The results of the two EDTA specimens tested fifteen days post venesection were in consensus with the other results but were excluded from the data analysis as outliers as the specimens had been tested after the closing date for the exercise. The outlier results are included Figure 3.1 for information.

In phase two, specimens were between fourteen and twenty-three days post venesection when tested, with one other result returned on a specimen tested on day thirty-one. The mode for the age of the specimens at testing was sixteen days

(39/144 specimens, 27.1%). A total of 72/144 (50.0%) were sixteen days old or less, 125/144 (86.8%) were twenty-one days old or less and 143/144 (99.3%) were twenty-three days old or less. Figure 3.2 shows the results returned by specimen age and by instrument or method in phase two. The specimen tested thirty-one days post venesection was excluded from data analysis as an outlier as the specimen had been tested after the exercise closing date. The outlier result is shown in Figure 3.2 for information.

Figure 3.3 shows the box and whisker plots by day of testing for both phase one and phase two specimens. The mean Hb A_2 result for all specimens was 2.58%, with a ± 0.25% acceptable performance range of 2.33 to 2.83%. The median result returned for each day that specimens were reported fell within this range except for specimens tested on day nine in phase one (EDTA specimens), which returned a median Hb A_2 result of 2.3%. The box and whisker plots do not suggest a trend in the results with specimen age when specimens are tested within the survey closing date. Overall there were more results returned that were less than the mean ± 0.25% Hb A_2 range (68/284, 24%) than greater than the mean ± 0.25% Hb A_2 range (46/284, 16%) but this distribution was observed across the full range of testing dates. The proportion of results falling outside the mean ± 0.25% performance range of 2.33% to 2.83% by specimen age at testing for both phases is shown in Figure 3.4.

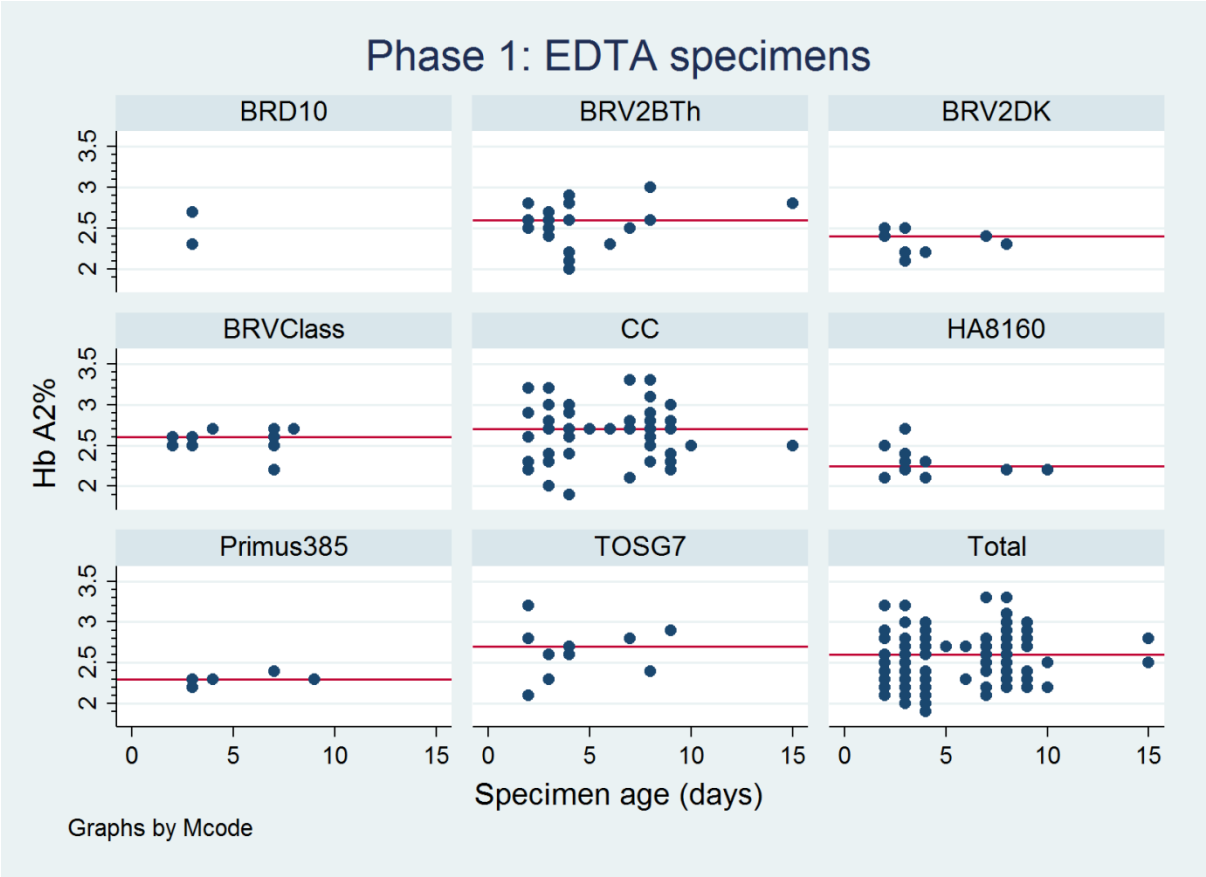


Figure 3.1. Distributional plot showing the Haemoglobin A_2 result returned by specimen age and method for EDTA anticoagulated specimens distributed in phase 1. The reference line shows the median result for each method group.

Legend to Figure 3.1:

CC:	Ion exchange microcolumn chromatography
BRD10:	Bio-Rad D10 HPLC analyser
BRVBTh:	Bio-Rad Variant II HPLC analyser, using the beta thalassaemia short programme
BRV2DK:	Bio-Rad Variant II HPLC analyser, using the dual kit programme
BRVClass:	Bio-Rad Variant Classic HPLC analyser
HA8160:	Menarini HA8160 HPLC analyser
Primus385:	Primus 385 HPLC analyser
TOSG7:	Tosoh G7 HPLC analyser
Total:	All methods results

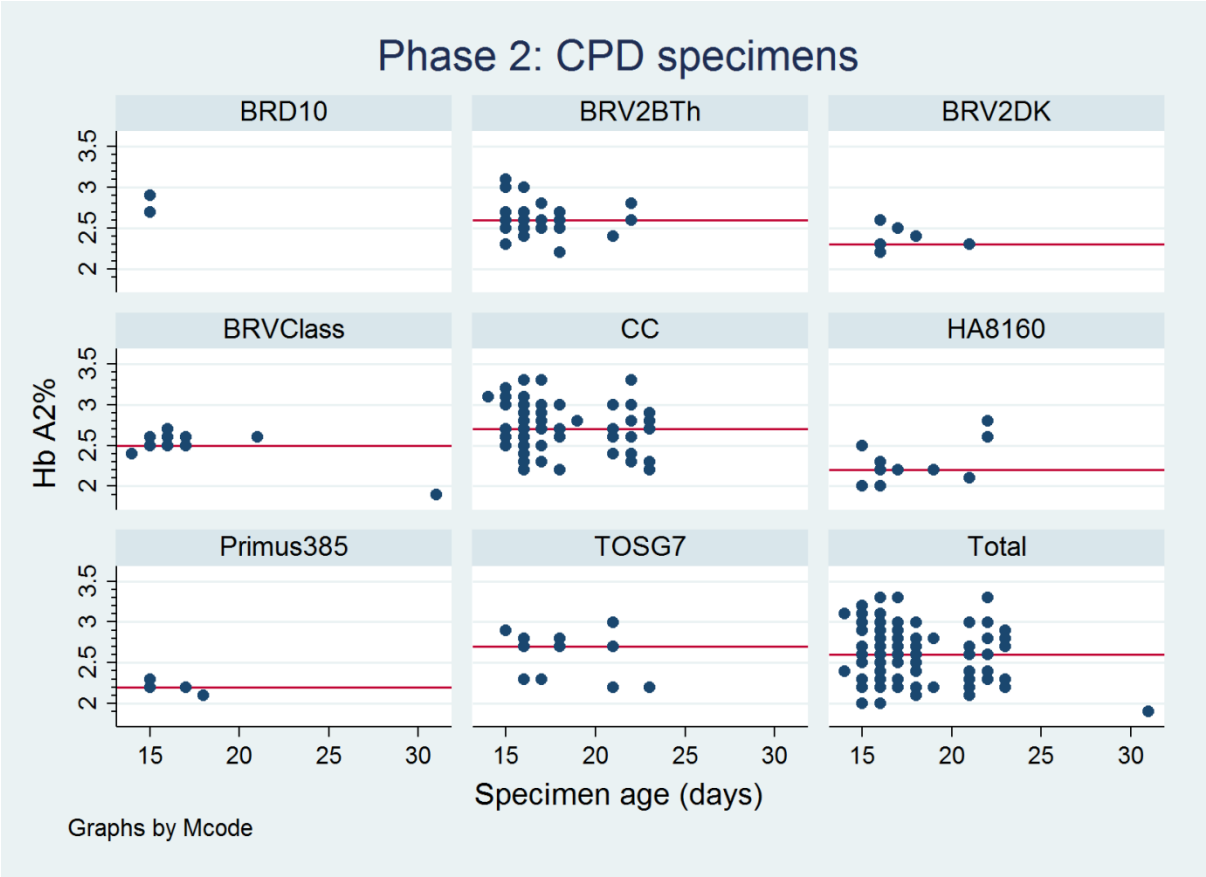


Figure 3.2. Distributional plot showing the Haemoglobin A_2 result returned by specimen age and method for CPD anticoagulated specimens distributed in phase 2. The reference line shows the median result for each method group.

Legend to Figure 3.2:

CC:	Ion exchange microcolumn chromatography
BRD10:	Bio-Rad D10 HPLC analyser
BRVBTh:	Bio-Rad Variant II HPLC analyser, using the beta thalassaemia short programme
BRV2DK:	Bio-Rad Variant II HPLC analyser, using the dual kit programme
BRVClass:	Bio-Rad Variant Classic HPLC analyser
HA8160:	Menarini HA8160 HPLC analyser
Primus385:	Primus 385 HPLC analyser
TOSG7:	Tosoh G7 HPLC analyser
Total:	All methods results

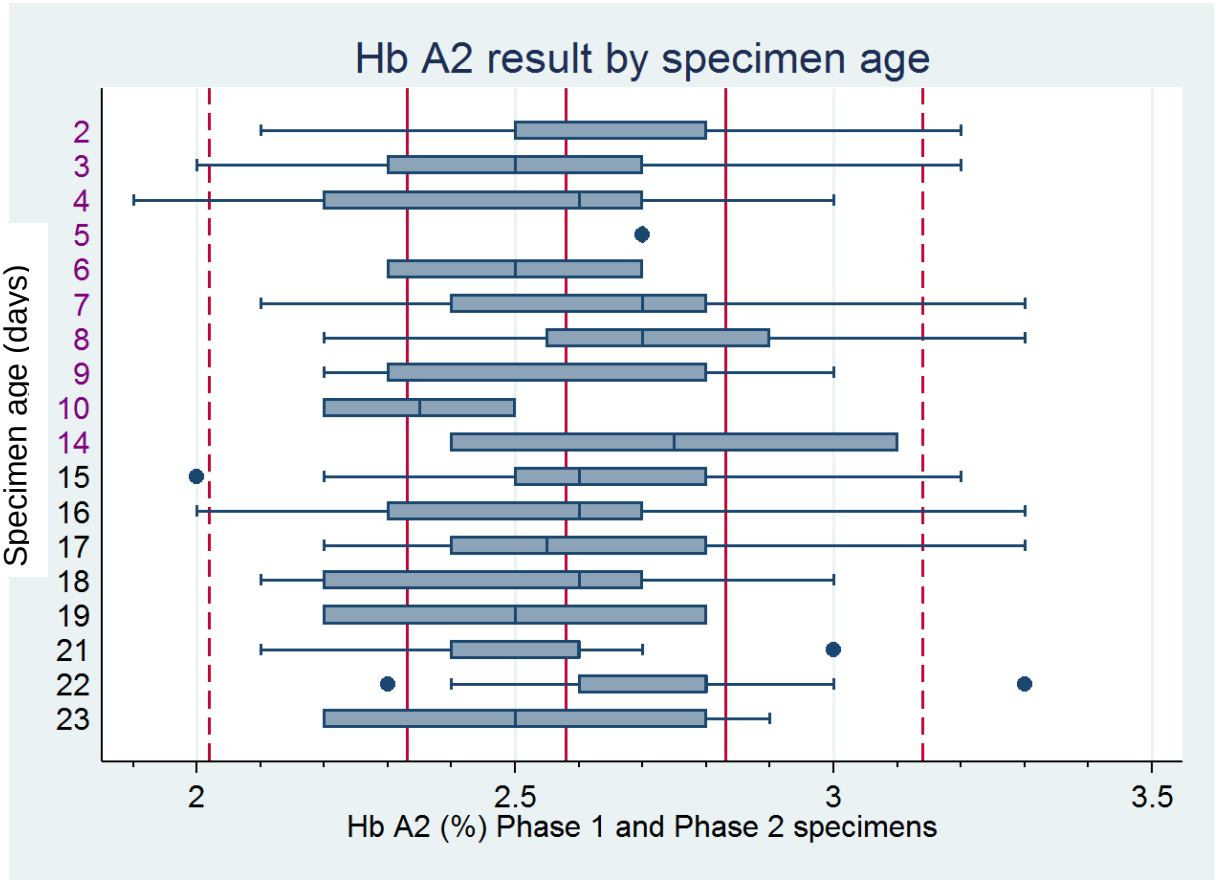


Figure 3.3. Box and whisker plots of Haemoglobin A_2 result returned by specimen age in days. Specimen ages shown in purple are phase 1 (EDTA) specimens (days two to fourteen), those in black are phase 2 (CPD) specimens (days fourteen to twenty-three). The mean result for all specimens is 2.58% with a ± 0.25% performance range of 2.33% to 2.83%, all shown by solid reference lines. The 95% confidence limits (mean ± 2SD) are shown by broken reference lines.

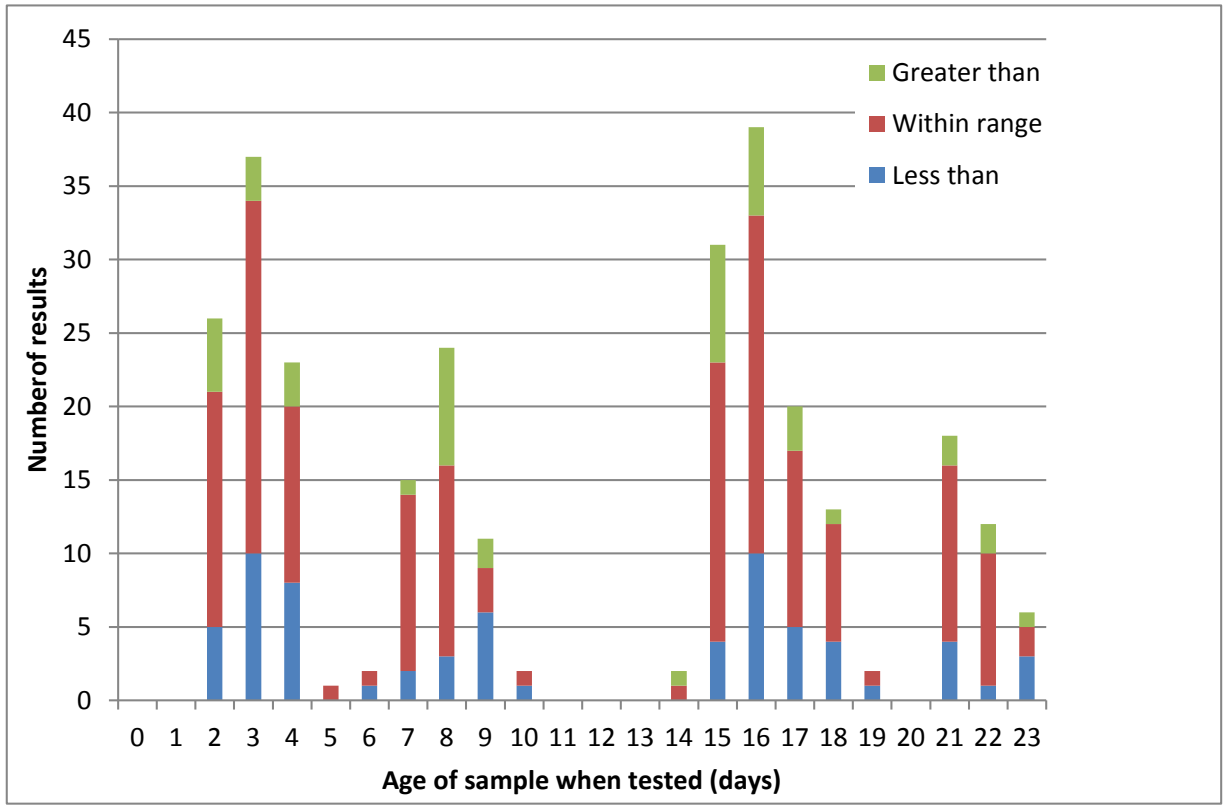


Figure 3.4. Proportion of results returned by specimen age at testing that were less than, within or greater than the mean ± 0.25% performance range of 2.33% to 2.83%. Phase one (EDTA) specimens are those tested at fourteen days or less, phase two (CPD) specimens at fifteen days or more

3.4.4 Effect of anticoagulant on Haemoglobin A_2 performance

T-tests were calculated where results were returned for both phases by the same participants to examine whether the difference between the results returned for phase one (EDTA anticoagulated) and phase two (CPD anticoagulated) specimens was significantly different from zero. Participants who had returned results that were identified as outliers because of the age of the specimens (two participants who had tested the phase one specimen at fifteen days and one participant who had tested the phase two specimen at thirty-one days post venesection) were excluded from the analysis, giving 136 sets of data. T-tests were calculated for the all methods group, by method principle and by sub-method (analyser type) where there were more than five results in the sub-method group.

The results of the t-tests are shown in Table 3.3. Using a probability value of $p \leq 0.05$ as significant, there was no significant difference between the results returned for phase one and phase two for all methods results, by method principle or by the major HPLC sub-methods.

Method	EDTA mean HbA_2	CPD mean Hb A_2	Mean Difference	SD	N	DF	T value	t-critical	P(T≤t)
CC	2.6614	2.7000	-0.0386	0.3288	57	56	-0.8862	2.0032	0.38
HPLC (All)	2.4975	2.5089	-0.0114	0.2088	79	78	-0.4850	1.9904	0.63
BRVClass	2.5750	2.5500	0.0250	0.1357	12	11	0.6383	2.2010	0.54
BRV2BTh	2.5567	2.6133	-0.0567	0.2373	30	29	-1.3077	2.0452	0.20
BRV2DK	2.3250	2.3625	-0.0375	0.1302	8	7	-0.8143	2.3646	0.44
HA8160	2.3091	2.2818	0.0273	0.1348	11	10	0.6708	2.2282	0.52
TOSG7	2.6545	2.6000	0.0545	0.2622	11	10	0.6901	2.2282	0.51
All Methods	2.5662	2.5890	-0.0228	0.2650	136	135	-1.0032	1.9777	0.32

Table 3.3. One sample t-test to test the difference between the Phase One (EDTA) and Phase Two (CPD) specimens

Legend to Table 3.3:

CC: Ion exchange microcolumn chromatography
BRVClass: Bio-Rad Variant Classic HPLC analyser
BRVBTh: Bio-Rad Variant II HPLC analyser, using the beta thalassaemia short programme
BRV2DK: Bio-Rad Variant II HPLC analyser, using the dual kit programme
HA8160: Menarini HA8160 HPLC analyser
TOSG7: Tosoh G7 HPLC analyser
All Methods: All methods results
SD: Standard deviation of the mean difference between EDTA and CPD results
N: Number of results
DF: Degrees of freedom (N-1)
T: Calculated t-test results
t-critical: limit of t-test result for the given the degrees of freedom and the chosen confidence limit (95%)

3.5 Discussion

In this study blood from the same donor was distributed to 160 UK NEQAS participating laboratories as a fresh, EDTA-anticoagulated specimen and after collection into CPD and processing according to the standard protocols for the preparation of survey material for the UK NEQAS Abnormal Haemoglobins scheme.

Although the ideal EQA material is one that is as similar to patients' clinical material as possible, there has to be a compromise in EQA between comparability with clinical material and the logistics of distributing an EQA service to a number of participants in diverse locations. The options for EQA survey material for Hb A_2 measurement for a large, international EQA scheme include whole blood in CPD, as used by UK NEQAS; 'stabilised' blood, as used by the College of American Pathologists (CAP) PT Programme; and lyophilised haemolysate, as used by the Royal College of Pathologists in Australia Quality Assessment Programme (RCPAQAP) at the time of writing. Lyophilised haemolysate was used for the preparation of the WHO Hb A_2 International Reference Reagent and more recently, Pornprasert (2016) and Paleari (2010) described the use of lyophilisation for the production of materials for the control of haemoglobin variant typing and Hb A_2 quantitation. Lyophilised materials offer many advantages for EQA in terms of batch manufacture, long term storage and the stability of material in transit; however, lyophilised materials require reconstitution before use, which may introduce an additional source of error not found with clinical specimens. Lyophilised haemolysate is not suitable for sickle solubility testing, which is still widely used in Hb variant identification and is part of the UK NEQAS Abnormal

Haemoglobins scheme design, as the sickle solubility test requires intact red blood cells.

Oxidation of Hb to methaemoglobin (MetHb) causes deterioration in the quality of the separation of Hb fractions with additional minor peaks and a poor baseline that confound the accurate quantitation of Hb A_2 (Stephens 2012). MetHb is formed physiologically at a rate of two to three percent of total Hb per day but this is reduced, primarily by NADH-linked MetHb reductase in the red cell (Mansouri, 1993; Çimen, 2008). Storage of red cells leads to the exhaustion of the metabolic pathways and a subsequent increase in MetHb, which can be slowed by exclusion of atmospheric oxygen as far as possible and storage of the sample at 2 to 8 °C (Tietz, 1990). Clinical samples in EDTA are stable for up to twenty-one days for Hb A_2 if kept under these conditions (Tietz 1990) and this timeframe is approximately the same as a UK NEQAS Hb A_2 'survey open' period. Louahabi (2006) stored specimens for up to four weeks for the evaluation of the Sebia Capillarys CZE analyser and reported the samples stable for Hb A_2 measurement during that time.

UK NEQAS Haematology survey packages are shipped at ambient temperature to most destinations. The international profile of the scheme means that 45% of participants overall are outside the UK at the time of writing and this proportion is even higher for participants registered for full participation in the Abnormal Haemoglobins programme, with 168/328 (51.2%) of laboratories registered for Hb A_2 measurement located outside the UK at the end of 2015. Of the non-UK participants registered at December 2015, fifty-two (15.8% of the total registrations) are outside Europe, with most non-European participants being located in Turkey, Israel and the Middle East. Delivery outside the British Isles is by courier to maintain the integrity of the specimens and the average transit time

by courier is four days. It is a requirement for accreditation to ISO 17043:2010 (ISO 17043:2010) for the EQA provider to demonstrate that the survey specimens are fit for purpose wherever they are tested.

UK NEQAS Haematology undertakes stability studies routinely on each batch of Hb A_2 survey material produced, testing the material on preparation, despatch and closing day with specimens stored at 2 to 8 °C. This exercise gave the opportunity to examine UK NEQAS survey material during the course of a simulated Abnormal Haemoglobins survey. Both the EDTA specimens distributed in phase one and the CPD specimens distributed in phase two showed no deterioration during the exercise as indicated by the Hb A_2 results returned.

Deterioration of specimens for Hb A_2 measurement is accelerated at high ambient temperatures and daytime temperatures of 50 °C are commonly encountered in some of the locations to which UK NEQAS services are sent. The quality of the results returned from all participants is a matter of concern not just because of the potentially adverse impact on the individual laboratory's performance but also to ensure the reliability of the statistics against which the performance of all participants is monitored. UK NEQAS Haematology uses temperature trackers to monitor the temperatures to which packages are exposed in international transit to selected participants and uses this information to make adjustments to the shipping conditions for future shipments where necessary. To avoid an increase in shipping costs due to the use of unnecessary additional cold-chain packaging, stability studies were undertaken to ascertain the stability of the Hb A_2 survey material at different temperatures in an experiment that was not part of this study but was undertaken under the direction of the author to demonstrate ISO 17043 compliance for accreditation purposes. Fifteen specimens from each of ten

batches of survey material distributed over the course of ten months for Hb A_2 measurement were taken on despatch day. Five were kept at ambient temperature in an air-conditioned laboratory (air temperature ranging from 20 to 25 °C), five were kept at 30 °C and five at 37 °C for up to four days to simulate the average length of time specimens spend in transit. One specimen from each temperature batch was analysed each day in triplicate for Hb A_2 using a Bio-Rad Variant II analyser with the beta thalassaemia short programme. Overall specimens kept for four days at ambient temperature showed a negative bias of 1.2% (range from -12.0% to zero) against the control specimen kept at 2 to 8 °C (note that this is the percent bias and not the change in absolute Hb A_2 result). At 30 °C a negative bias of 0.04% was seen at four days (range -0.2% to 0.11%) and a negative bias of 4.3% (range -14.0% to 0.31%) for specimens kept at 37 °C. These figures indicate that the UK NEQAS survey material is very stable at temperatures of up to 30 °C for four days in transit and that short exposure to temperatures up to 37 °C should not cause the specimens to deteriorate to such an extent that participants incur an unsatisfactory performance score.

The second question asked in this study concerned the impact of the anticoagulant used by UK NEQAS Haematology on the commutability of the survey material. Commutability is the ability of the EQA survey material to give the same result when tested across all platforms. It is essential that both the EQA provider and the participating laboratory understand the limits of commutability when evaluating EQA performance (James, 2014; ISO 17043:2010; BS EN 14136:2004). Where there is a single cut-off Hb A_2 value for the identification of a beta thalassaemia carrier, as in the antenatal screening algorithm of the NHS STP, the ideal is that the EQA performance of all instruments can be assessed in

the same group to reflect ‘real-life’ practice. On the basis of the findings from this exercise, there is no statistical significance in the difference of the results obtained from material prepared according to standard UK NEQAS procedures (distributed in phase two) and fresh EDTA material from the same donor (distributed in phase one) supporting the conclusion that the UK NEQAS Haematology Abnormal Haemoglobins material is commutable for the measurement of $A_{2.}$ Commutability of the survey material makes the current practice of scoring performance against the MTM questionable, since the only justification for grouping instruments for performance assessment is a lack of commutability. The separate performance assessment of HPLC and CZE methods is for historical reasons dating from the use of manual methods and this should be re-evaluated as it is not evidence-based.

Although a representative range of HPLC methods together with microcolumn chromatography were included in this exercise, the changes in methodology of recent years mean that a number of current analysers and methods, including CZE and TMS, were not included in the study. The evaluation of the suitability of EQA material for new techniques and platforms is a constant requirement in all areas of EQA provision. A full study of this nature, with the venesection of a volunteer donor is not a practical means of undertaking this validation on a regular basis; however, the study has demonstrated that in principle the UK NEQAS Hb A_2 survey material performs in a comparable fashion to fresh EDTA blood. Verification of this assumption is possible using comparison studies as new instruments are introduced and the regular data mining of participants’ results on a survey by survey basis.

Chapter 4: The quantitation of Hb A in dried blood spot specimens as a means for screening newborn infants for beta thalassaemia major

4.1. Aim

To assess the measurement of Hb A% as a means to screen newborn infants for beta thalassaemia using dried blood spot cards.

4.2 Introduction

The benefit of early intervention on the course of certain inherited disorders was recognised in the 1950s when the implementation of a modified diet in very young children with phenylketonuria (PKU), an inborn error of amino acid metabolism was found to be beneficial in their management (Horst 1954). Newborn screening for PKU was introduced approximately ten years later and there are now up to fifty disorders that are screened for in the neonatal period (Wilcken, 2008; Sahai, 2009; Mak, 2013). At the time of writing, newborn screening is offered for nine inherited conditions in England using a dried blood spot (DBS) sample taken five to eight days after birth. The conditions are sickle cell disease, cystic fibrosis, congenital hypothyroidism and the inherited metabolic diseases PKU, medium-chain acyl-CoA dehydrogenase deficiency (MCADD), maple syrup urine disease, isovaleric acidaemia, glutaric aciduria type 1 and pyridoxine unresponsive homocystinuria (NHS Newborn Blood Spot Screening Programme Standards, 2017). Of the conditions screened for, sickle cell disease is the most frequently encountered, estimated to affect one in two thousand babies in the UK. Early diagnosis of sickle cell disease and the implementation of penicillin prophylaxis have been shown to be beneficial in reducing mortality in sickling disorders (Rogers, 1978a; Rogers, 1978b; Gaston, 1986).

The main Hb present in the fetus is fetal haemoglobin, Hb F, which is replaced by adult Hb (Hb A), starting in late fetal life with a switch from the production of the

gamma globin chains of Hb F to the production of the beta globin chains of Hb A. The inheritance of a beta thalassaemia mutation results in reduced or absent globin chain synthesis; if both beta globin genes are affected as in beta thalassaemia major, Hb A is absent or severely reduced. The synthesis of gamma globin chains is not affected by the presence of a beta thalassaemia mutation, and the clinical symptoms of beta thalassaemia only become evident as the production of Hb F declines in the neonatal period. Thus an infant with beta thalassaemia major undiagnosed at birth typically presents in the first one to two years of life, with symptoms of severe anaemia, haemolysis and ineffective erythropoiesis (Bain, 2006; Weatherall and Clegg, 2001). An early knowledge of the inheritance of beta thalassaemia major permits diagnostic testing to determine the mutation(s) present and ensures that the infant is referred promptly for appropriate clinical care (Hoppe, 2009). Early identification of the unexpected birth of a beta thalassaemia major infant will also facilitate the prompt investigation of any possible failure in an associated antenatal screening programme where these are linked services.

Universal newborn screening for sickle haemoglobin and other clinically significant haemoglobinopathies using DBS specimens was implemented in 2006 in England (Streetly, 2009b), the same year as it was mandated in all states in America (Benson, 2010). The NHS STP for newborn sickle screening in England requires the identification of infants with sickle cell anaemia (HbSS or homozygous for Hb S) and other sickling conditions resulting from co-inheritance of Hb S with HbC, beta thalassaemia, Hb D^{Punjab}, Hb E, Hb O^{Arab}, Hereditary Persistence of Fetal Haemoglobin (HPFH) or with any other structural Hb variant (NHS STP Newborn Laboratory Handbook, 2017). Over seven million newborn infants were screened

in England in the ten years to 2016, of which 3,600 were found to be positive for significant haemoglobinopathies (NHS STP Data Report 2015/16: trends and performance analysis, 2017). Although it is not one of the aims of the programme, newborn screening will detect infants with an absence of Hb A or an Hb A% that is low for gestational age, which may be indicative of beta thalassaemia major or thalassaemia intermedia. In the first ten years of newborn screening in England, 268 ‘Hb F only’ infants were identified.

Mantikou (Mantikou, 2009) published evidence that beta thalassaemia carriers could be detected at birth using a cut-off value for Hb A of less than 15%, demonstrating a quantitative relationship between Hb A% at birth and reduced output from the beta globin gene(s). The proportion of Hb A at birth is highly variable, even in full-term infants, and an effective screening programme requires a cut-off that reliably identifies the majority of affected infants but with relatively few false positives. A major retrospective analysis of newborn blood spot screening data from 2,288,008 infants born in England in the period 2005 to 2012 has been undertaken to identify the relationship between Hb A% at birth and beta thalassaemia major. The data demonstrated that the adoption of a 1.5% cut-off for Hb A in infants of more than thirty-two weeks’ gestation has a good positive predictive value, specificity and sensitivity, better than that demonstrated for newborn screening for MCADD deficiency (Streetly, 2013). Of the 2,288,008 infants reviewed, 170 were identified as having beta thalassaemia major or beta thalassaemia intermedia using a cut-off of 1.5% Hb A. 119 / 170 were identified prospectively from their Hb A% at birth and 51 / 170 were identified retrospectively in a look-back exercise. Seven of the prospectively identified infants were lost to follow up and fifteen were subsequently found to be false positives (eleven were

premature infants of less than thirty-two weeks' gestation, one a beta thalassaemia carrier and one a compound heterozygote for beta thalassaemia and HPFH). Of the remainder, ninety were confirmed as beta thalassaemia major and seven as beta thalassaemia intermedia. Although there were no 'missed' cases reported from the screening centres involved in the study, one infant would have been a false negative (Hb A = 1.9%) but for the fact that it was followed up as a result of parental results. The NHS STP Newborn Laboratory Handbook for Newborn Screening Laboratories recommends a cut-off of 1.5% Hb A at birth for the identification of infants with beta thalassaemia major.

Newborn screening for haemoglobinopathies requires the initial or primary screening of the DBS and the confirmation of any positive result by a secondary method. Amongst the laboratories registered in the UK NEQAS Haematology Newborn Sickle Screening scheme in December 2016, the primary screening methods were HPLC (the Bio-Rad VARIANT™ newborn screening (Vnbs) analyser, running the sickle cell short programme), CZE (the Sebia Capillarys Neonat Haemoglobin FAST™ CZE system) and electrospray TMS. Confirmatory methods used were from the same range of methods plus isoelectric focusing (IEF). The statements of intended use for both Vnbs and the Neonat Fast instruments are that the instruments are intended for the qualitative detection of the presence of haemoglobins F, A, S, C, D and E (Bio-Rad Vnbs) in neonatal blood and the separation of the normal Hb F and Hb A in human newborns and for the detection of the variant Hbs S, C, E, D and Barts (Sebia Neonat Hb Fast) (Substantial equivalence determination decision summaries K051072 and K091283, US Food and Drug Administration agency; Bio-Rad VARIANT™ nbs Sickle Cell Programme instruction manual; Sebia Capillarys Neonat Hb FAST™ kit

insert Ref 2006). No limit of detection (LOD) has been determined for the Sebia Capillarys Neonat Hb FAST™ but an equivalent performance of 1.1% for Hb S, 2.3% for Hb C and 1.1% for Hb E in comparison with the reference Bio-Rad Vnbs HPLC system is given in a sensitivity study. By comparison, a LOD of 1% is given for the Bio-Rad Vnbs analyser for the detection of Hb S, D, C and E, although no LOD is given for Hb A (Substantial equivalence determination decision summary K051072 and K091283, US Food and Drug Administration agency).

The purpose of the experiments described in this chapter was to review the performance of the main methods of newborn haemoglobinopathy screening (Bio-Rad Vnbs and Sebia Capillarys Neonat Hb kit) for the quantitation of low concentrations of Hb A, as recommended for the detection of infants with beta thalassaemia major by the NHS STP. Although both instruments provide a quantitative result for all Hb fractions detected, the limitations of this quantitation must be understood since the systems are not marketed as quantitative devices.

4.3 Materials and methods

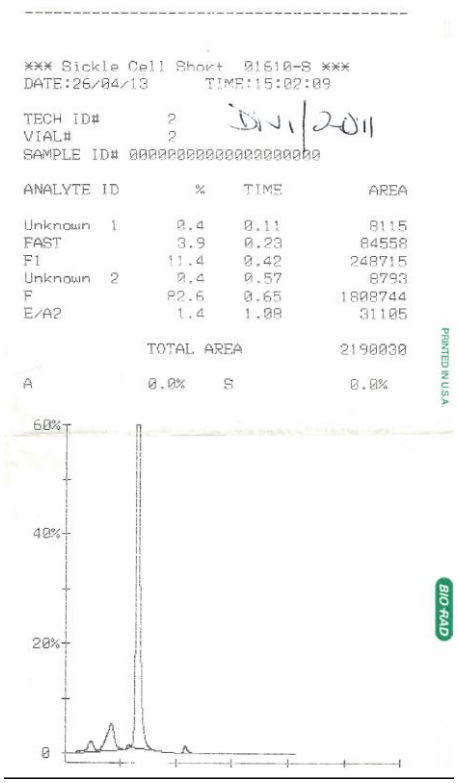
4.3.1 Dried blood spot card preparation

Twelve mL of EDTA-anticoagulated blood was collected with informed consent from a female, adult patient with homozygous delta-beta thalassaemia, who was managed clinically as thalassaemia intermedia. Blood collected for the purpose of quality assurance and instrument or method validation is exempt from the Human Tissue Act regulations (Human Tissue Act, 2004). The patient's haemoglobinopathy screen showed Hb F (82.6%) and Hb A_2 (1.4%) but no Hb A and no abnormal Hb variant present when tested using a Bio-Rad Variant Classic HPLC analyser with the sickle cell short programme (see Figure 4.1a).

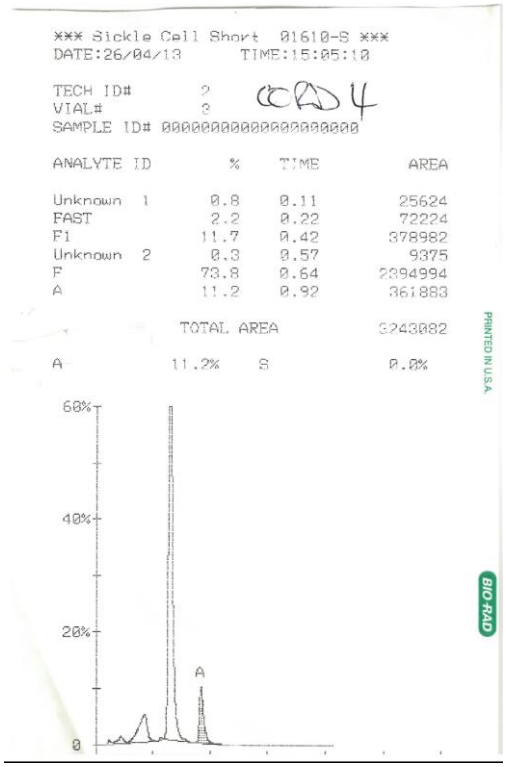
One unit of waste umbilical cord blood was obtained from NHSBT. The unit was unsuitable for therapeutic use and the mother had consented for its use for quality assurance and/or validation purposes. The haemoglobinopathy screen of the umbilical cord blood showed Hb F (73.8%) and Hb A (11.2%), no Hb A_2 and no abnormal Hb variant when tested using a Bio-Rad Variant Classic HPLC analyser with the sickle cell short programme (see Figure 4.1b).

The adult and cord red blood cells (RBC) were washed separately three times in 0.9% w/v phosphate buffered saline (PBS), using a ratio of approximately 10:1 PBS: RBC each time in 10 mL plastic, conical centrifuge tubes, centrifuging at 3000 rpm in a Heraeus Megafuge centrifuge at 20 °C at each wash. The supernatant saline was removed between washes using a water operated Venturi pump, taking care not to disturb the RBC layer. After three washes, the RBCs were resuspended in PBS to give a cell suspension of approximately 50%.

The haematocrit (Hct) of each cell suspension was measured using a Sysmex XE2100 FBC analyser and adjusted to a value in the range 0.45 to 0.55, to simulate a typical Hct for a newborn infant. The actual final Hct of both the adult and cord RBC suspensions was 0.49 L/L. The cell suspensions were adjusted to the same Hct to allow the calculation of the proportions of adult and cord cells required to make a range of Hb A% without having to take account of the Hct of the cell suspensions.



a) Adult sample



b) Cord sample

Figure 4.1. HPLC chromatograms of the adult and cord blood samples used for preparation of the red cell suspensions.

(Note the poor quality of these scans is due to the thermal paper output of the instrument).

The washed, resuspended adult and cord RBCs were mixed to make eight red cell suspensions with a range of Hb A% concentrations calculated to be between 0.5% and 10%. The actual Hb A% of each RBC suspension was measured ten times using a Bio-Rad Vnbs HPLC analyser with the sickle cell short programme, preparing a separate dilution (5 μL cell suspension to 1 mL deionised water) for each, and the mean Hb A% of the ten measurements calculated.

Forty dried blood spot (DBS) cards (Perkin-Elmer Health Sciences, SC, USA) were prepared from each of the eight mixtures with a single spot of 25 μl of RBC suspension per card, applied using a 100 μl air displacement pipette. The cards were checked to ensure that the RBC suspension had penetrated through the filter paper to the back of the card and left to air-dry for thirty minutes. Once dried, a test card was taken from each batch and tested using a Bio-Rad Vnbs HPLC analyser, as a reference for the comparison of the quality of the baseline and the peak separation for the DBS during the different stages of the study.

Table 4.1 shows the volumes of adult and cord blood cells used to prepare the mixtures and the Hb A% of both the RBC suspensions and the DBS cards immediately after preparation.

Once dried, the cards were labelled with the survey material batch number (one to eight) and packed into individual glassine envelopes, supplied with the cards. Each batch of cards was wrapped tightly as a bundle in aluminium foil and then plastic wrapping film and stored at -40 °C for a total of twelve months until all parts of the experiment had been completed. UK NEQAS Haematology experience of storing DBS for use as EQA material has demonstrated that DBS are stable for this period, providing they are kept at -40 °C and tightly wrapped to exclude

moisture (unpublished pre-acceptance testing evidence from ten years of the UK NEQAS Haematology Newborn Sickle Screening EQA programme operation).

When the DBS were removed from storage, the bundle of cards was opened and the required number removed. The remaining cards were rewrapped promptly and returned to storage at -40 °C. Once removed from storage, the cards were laid in a single layer on a clean bench and allowed to equilibrate at ambient temperature (20 to 24 °C in an air-conditioned laboratory) before being packaged for distribution or tested in-house.

Card batch number	Calculated Hb A%	Volume of adult cells (μL)	Volume of cord cells (μL)	Mean Hb A% of the mixed RBC suspension	Hb A% DBS immediately after preparation
1	0.5	1145	55	0.6	0.3
2	1.0	1090	110	1.0	0.7
3	1.5	1040	160	1.5	1.2
4	2.0	990	215	2.0	1.6
5	3.0	880	320	3.1	2.7
6	5.0	665	540	5.0	4.5
7	7.0	450	750	7.0	6.6
8	10.0	100	1100	10.9	10.6

Table 4.1. Preparation of dried blood spot cards with a range of Hb A%. The adult cells were a suspension of RBC from a patient with beta thalassaemia intermedia (Hb F and Hb A_2 only); the cord cells were a suspension of RBC from a unit of umbilical cord blood (Hb F and Hb A). Both adult and cord cell suspensions had a haematocrit of 0.49 L/L.

4.3.2 Bio-Rad VARIANT™ nbs HPLC analyser specimen preparation and analysis

A 3 mm blood spot was punched from each DBS using a hand-held DBS card punch into a labelled 1.5 mL clear microtube and 250 μL deionised water added. The tube was capped and left to stand for ten minutes at room temperature and then agitated gently to encourage elution of the dried blood from the filter paper and to mix the eluate. Three to five drops of eluate were transferred using a pastette to a well in a ninety-six well microtitre plate, which was loaded onto the Bio-Rad Vnbs analyser for analysis together with blank (deionised water), FAES and FADC retention time markers (Bio-Rad Laboratories, CA). The plates were held at 9 ± 2 ^{0}C in the analysis chamber. The samples were injected automatically in sequence into the HPLC cartridge on the analyser and the Hb fractions present in the eluate separated by cation exchange chromatography using a pre-programmed buffer gradient. The separated fractions were detected by their absorbance as they passed through a flow cell in a spectrophotometer monitored at 415 and 690 nm and identified by the retention time window into which they fell. The percentage of each fraction as a part of the whole was calculated automatically by the integration of the area under the absorbance peak.

Samples of liquid blood or RBC suspension were prepared by diluting 5 μL of mixed, whole blood or RBC suspension in 1 mL of deionised water in a 1.5 mL clear microtube and mixed by inversion before being analysed in the same way as the eluate from a DBS.

4.3.3 Distribution of dried blood spot cards for the assessment of Haemoglobin A% quantitation

The DBS cards were distributed for testing on two occasions to fourteen UK Newborn Screening Laboratories registered in the UK NEQAS Newborn Sickle Screening scheme. Both exercises (identified as 1301NHEX and 1401NHEX) took place within a year of preparation of the DBS cards. The timing was selected for logistical purposes. A set of cards was tested in the UK NEQAS Haematology laboratory at the same time, giving a total of fifteen sets of results for each distribution.

At each distribution, fifteen cards were withdrawn from storage from each of the eight batches. One card was tested within twenty-four hours of removal from storage using the Bio-Rad Vnbs HPLC analyser in the UK NEQAS Haematology laboratory and the quality of the chromatogram and the quantitation of the Hb peaks reviewed against the original results of the card tested prior to storage. The cards were packed into fifteen sets with one card from each batch in each set and a set of eight cards was dispatched by first class post to each of fourteen UK newborn screening laboratories. The laboratories were asked to test the cards by their routine primary method for newborn sickle upon receipt in the order one to eight, i.e. from the lowest to the highest Hb A% and to return their results using a proforma result sheet. An example instruction sheet and proforma results sheet are shown in Appendix 2.

4.3.4 Precision of low Haemoglobin A% measurements

Precision is the agreement between replicate analyses of the same specimen and is an essential component of the uncertainty of a measurement. It is not the same as accuracy. Two cards from each batch were withdrawn from storage. The DBS from each card was cut out in its entirety using a clean pair of scissors and left to stand for ten minutes in 3 mL of deionised water in a plastic test tube to elute the Hb from the DBS. The tube was covered with Parafilm™ and mixed by gentle agitation. The resulting eluate was tested using a Bio-Rad Vnbs analyser and its concentration adjusted with deionised water to give the correct concentration for accurate analysis. The eluate from each card was used to make ten replicate analyses of Hb A% using the Bio-Rad Vnbs analyser at UK NEQAS Haematology, i.e. twenty results in total. Two cards from each batch were sent to two volunteer newborn screening laboratories using the Bio-Rad Vnbs analyser (identified as sites X and Y), where the experiment was repeated. The results were analysed using one-way analysis of variance (ANOVA) to examine the variance in the results between different DBS cards and between sites for the same Hb A%.

4.3.5 Carryover of Haemoglobin A

Carryover is the degree of contamination of a specimen by the one immediately preceding it (Broughton 1974) and is of particular relevance when samples of very low analyte concentration may be analysed in a batch of samples with 'normal' analyte concentrations.

One card from each batch was withdrawn from storage and an eluate was prepared from each for analysis on the Bio-Rad vNBS analyser. An umbilical cord blood with a normal Hb variant pattern (Hb F + Hb A) and a Hb A concentration of approximately 25% was selected and 25 μL of cord blood diluted in 5 mL of

deionised water to represent an eluate from a DBS card from an infant with a 'normal' Hb A%. Carryover studies were run with each of the low Hb A% DBS by running three replicate analyses of the normal eluate (designated H1 to H3) followed by three replicate analyses of the low Hb A% DBS (designated L1 to L3 for each DBS batch). The carryover was calculated using the formula shown in Equation 4.1 (Briggs, 2014).

$$Carryover\ \% = \frac{(L1 - L3)}{(H3 - L3)}\ x\ 100$$

Equation 4.1. Calculation of carry over

4.3.6 Statistical analysis

Statistical analyses were undertaken using STATA 14 (Stata Corp, Texas, USA) or Microsoft Excel 97 2003 (Microsoft Corp, WA, USA) with the downloaded analysis Toolpak.

4.4 Results

4.4.1 Accuracy of low Haemoglobin A% measurements from different testing sites and methods

Dried blood spot (DBS) cards prepared from RBC suspensions with Hb A concentrations ranging from 0.6% to 10.9% were distributed to fourteen UK Newborn Screening Laboratories on two occasions for testing by their primary newborn sickle screening method. The DBS cards were also tested by UK NEQAS Haematology. On both occasions, results were returned from two Sebia Capillarys Neonat Hb Fast CZE analysers, one Bio-Rad Variant II HPLC analyser using the beta thalassaemia short programme and twelve Bio-Rad Variant Newborn Sickle screening (Vnbs) HPLC analysers, which included one located at UK NEQAS Haematology. The one Bio-Rad Variant II analyser used was part of the secondary screening methodology used by the one UK Newborn Screening Laboratory that used TMS as its primary method of newborn sickle screening. This laboratory operates a screening protocol that is designed to detect only sickle cell disease.

The results returned on both duplicate studies are shown in Table 4.2 and the descriptive statistics (mean, SD, CV and range) in Table 4.3 for all methods and separately for the largest instrument group (Vnbs analysers). The results are shown graphically against the 'expected' values as determined from the original RBC suspensions from which the DBS were prepared in Figures 4.2 and 4.3, with the results from each methodology analysed separately.

Distribution	1301NHEX results (Hb A%)								1401NHEX results (Hb A%)							
DBS card batch	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
RBC Suspension Hb A%	0.6	1.0	1.5	2.0	3.1	5.0	7.0	10.9	0.6	1.0	1.5	2.0	3.1	5.0	7.0	10.9
Instrument																
Sebia NeoFast	ND	ND	1.8	ND	2.8	3.7	5.5	7.9	ND	ND	1.6	ND	2.7	3.4	5.4	7.0
Sebia NeoFast	ND	ND	ND	ND	2.6	3.3	4.6	6.9	ND	ND	ND	1.9	2.7	3.6	4.8	6.9
Bio-Rad VII	0.7	1.3	1.9	2.4	3.7	5.9	8.2	12.8	0.7	1.2	2.0	2.4	3.6	5.8	8.0	12.0
Bio-Rad VNBS	ND	1.2	1.8	2.2	3.2	5.0	7.0	10.5	ND	1.3	1.4	2.1	3.1	4.9	6.9	10.6
Bio-Rad VNBS	1	1.3	1.4	2.2	3.2	4.5	6.6	9.8	0.9	1.3	1.5	2.3	2.8	4.7	6.5	9.7
Bio-Rad VNBS	ND	0.5	1.1	1.5	2.4	4.1	6.1	9.6	ND	0.9	1.1	1.5	2.4	4.0	6.1	9.7
Bio-Rad VNBS	ND	1.1	1.5	1.8	2.8	4.6	6.2	9.7	0.9	1.2	1.3	1.7	2.9	4.5	6.0	9.8
Bio-Rad VNBS	ND	1.0	1.4	1.8	2.8	4.7	6.4	9.8	ND	1.1	1.4	1.7	2.7	4.8	6.3	9.8
Bio-Rad VNBS	0.8	1.0	1.7	1.8	2.9	4.9	6.6	10.1	0.8	1.0	1.5	1.8	3.1	4.8	6.7	10.3
Bio-Rad VNBS	ND	0.9	1.1	1.8	2.6	4.6	6.4	10	ND	0.8	1.1	1.8	2.8	4.5	6.4	9.9
Bio-Rad VNBS	ND	0.9	1.4	1.8	2.8	4.6	6.2	9.8	ND	1.0	1.2	1.7	2.8	4.4	6.2	10.1
Bio-Rad VNBS	ND	0.9	1.2	1.8	2.7	4.3	6.3	9.6	ND	0.6	1.3	1.8	2.8	4.3	6.2	9.5
Bio-Rad VNBS	ND	0.7	1.2	1.5	2.7	4.3	6.3	9.8	ND	0.7	1.2	1.5	2.9	4.4	6.2	9.6
Bio-Rad VNBS	ND	0.9	1.3	1.8	2.8	4.6	6.5	9.8	0.8	0.9	1.3	1.8	2.7	4.5	6.4	9.7
Bio-Rad VNBS	ND	1.0	1.3	1.7	2.9	4.6	6.5	10.3	ND	1.1	1.3	1.8	2.8	4.6	6.4	10.1

Table 4.2. The results for the DBS with different Haemoglobin A% concentration by instrument type for experiments 1301NHEX and 1401NHEX. The results for each site are shown with the Hb A% of the RBC suspensions from which they were prepared.

ND = not detected

Distribution	1301NHEX results (Hb A%)								1401NHEX results (Hb A%)							
DBS card batch	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
RBC Suspension Hb A%	0.6	1.0	1.5	2.0	3.1	5.0	7.0	10.9	0.6	1.0	1.5	2.0	3.1	5.0	7.0	10.9
ALL METHODS Hb A%																
Number	3	13	14	13	15	15	15	15	5	13	14	14	15	15	15	15
Mean	~	0.98	1.44	1.85	2.86	4.51	6.36	9.76	~	1.01	1.37	1.84	2.85	4.48	6.30	9.65
SD	~	0.22	0.27	0.26	0.31	0.58	0.75	1.25	~	0.22	0.23	0.26	0.27	0.56	0.69	1.25
CV%	~	22.95	18.65	14.21	10.88	12.94	11.82	12.83	~	21.99	17.02	14.23	9.35	12.52	10.98	12.99
Minimum	ND	0.50	1.10	1.50	2.40	3.30	4.60	6.90	ND	0.60	1.10	1.50	2.40	3.40	4.80	6.90
Maximum	1.00	1.30	1.90	2.40	3.70	5.90	8.20	12.80	0.90	1.30	2.00	2.40	3.60	5.80	8.00	12.00
VNBS ONLY Hb A%																
Number	2	12	12	12	12	12	12	12	4	12	12	12	12	12	12	12
Mean	~	0.95	1.37	1.81	2.82	4.57	6.43	9.90	~	0.99	1.30	1.79	2.82	4.53	6.36	9.90
SD	~	0.21	0.22	0.22	0.22	0.25	0.24	0.28	~	0.22	0.13	0.22	0.19	0.25	0.25	0.32
CV%	~	22.22	16.01	11.90	7.99	5.47	3.76	2.79	~	22.53	10.37	12.47	6.57	5.51	3.99	3.22
Minimum	ND	0.50	1.10	1.50	2.40	4.10	6.10	9.60	ND	0.60	1.10	1.50	2.40	4.00	6.00	9.50
Maximum	1.00	1.30	1.80	2.20	3.20	5.00	7.00	10.50	0.90	1.30	1.50	2.30	3.10	4.90	6.90	10.60

Table 4.3. Descriptive statistics for the Haemoglobin A% results for DBS tested by different sites, for all methods used and for the Bio-Rad vNBS HPLC instruments only. Cards were distributed from each batch to all 15 laboratories included in the study; the number of results returned shown excludes results returned as ND or less than (<).

ND = not detected

Most of the analysers failed to detect and quantitate the Hb A present in the batch of cards prepared from the 0.6% Hb A RBC suspension, although all except for the Sebia Neonat Hb Fast analysers quantitated the Hb A in the cards prepared from the 1.0% Hb A RBC suspension. For both exercises 1301NHEX and 1401NHEX, the all methods mean Hb A% value obtained from the DBS specimens showed a negative bias against the Hb A% result of the RBC suspension, with the exception of the results for the lowest Hb A% batch of cards in exercise 1401NHEX. This bias ranged from -2.3% to -10.5% for all instruments in exercise 1301NHEX and +0.8% to -11.5% in exercise 1401NHEX.

The CV% for all instruments ranged from approximately 22% for DBS card batch two (RBC suspension Hb A = 1.0%) to approximately 10-13% for all instruments for DBS card batches five to eight (RBC suspension Hb A 3.1% to 10.9%). The CV% values were lower overall for the Vnbs group, where the CV% was approximately 22% for DBS cards prepared from a RBC suspension with a Hb A concentration of 1% but fell to approximately 3% at a Hb A concentration of 10.9%. The results are illustrative rather than statistically representative for both the Bio-Rad VII instrument and the two Sebia Neonat Hb Fast instruments because of the low numbers of each.

In both exercises, the Bio-Rad VII HPLC analyser showed an apparent over-estimation of the Hb A% in the DBS compared to the RBC suspension. The converse was observed for the Sebia Capillarys Neonat Hb Fast CZE instruments, where the linear regression of the DBS on the RBC suspension Hb A% results reflected an apparent under-estimation of the Hb A% in the DBS specimens (linear regression of y= 0.6067x + 0.738 and y=0.571x + 0.831 for 1301NHEX and 1401NHEX respectively). The linear regression of the results for the Bio-Rad Vnbs

instruments (y=0.91x + 0.02 and y= 0.90x + 0.03) reflected a small observed negative bias of the DBS results compared to the RBC suspension.

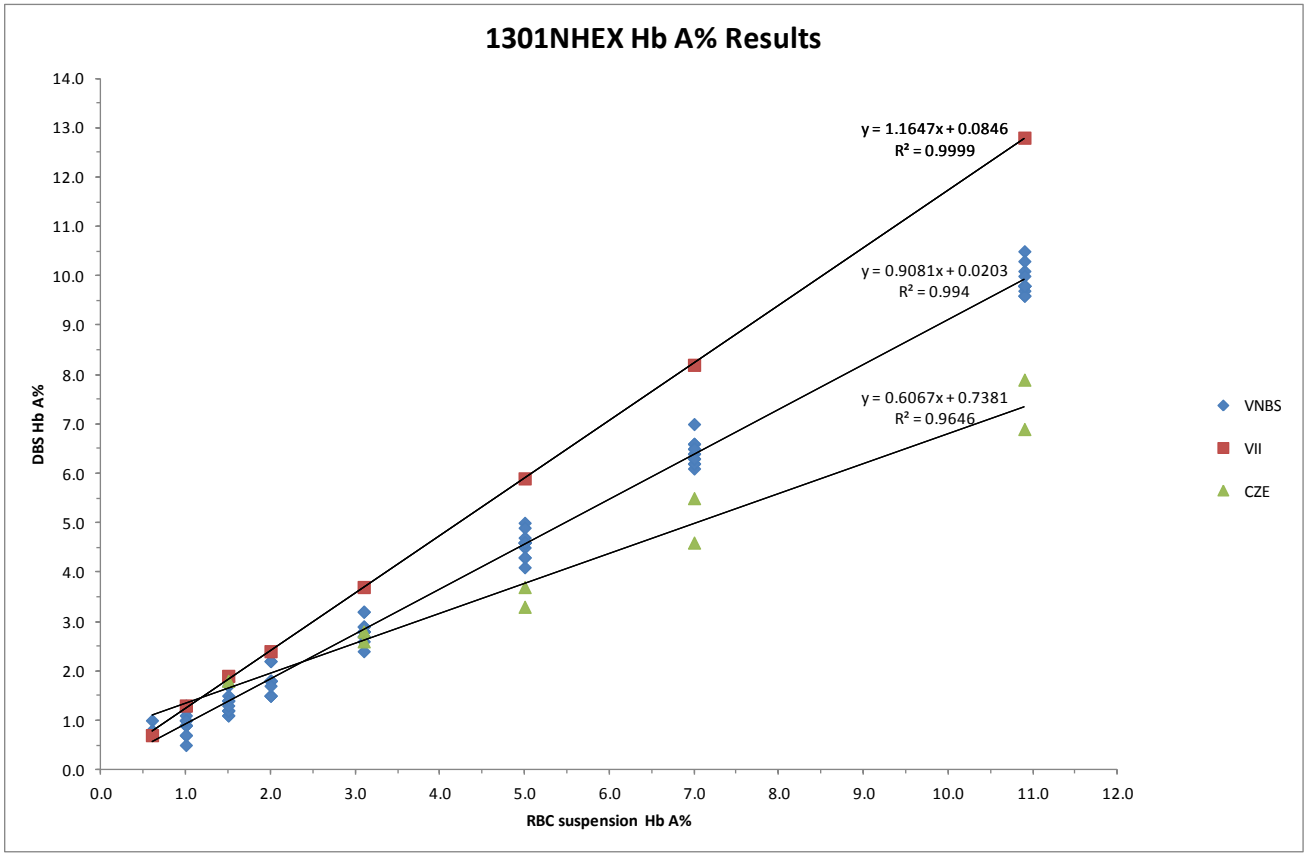


Figure 4.2. Linear regression of the 1301NHEX Haemoglobin A% results

Legend to Figure 4.2: VNBS = Bio-Rad Variant nbs HPLC, VII = Bio-Rad Variant II HPLC, CZE = Sebia Neonat Hb Fast capillary electrophoresis.

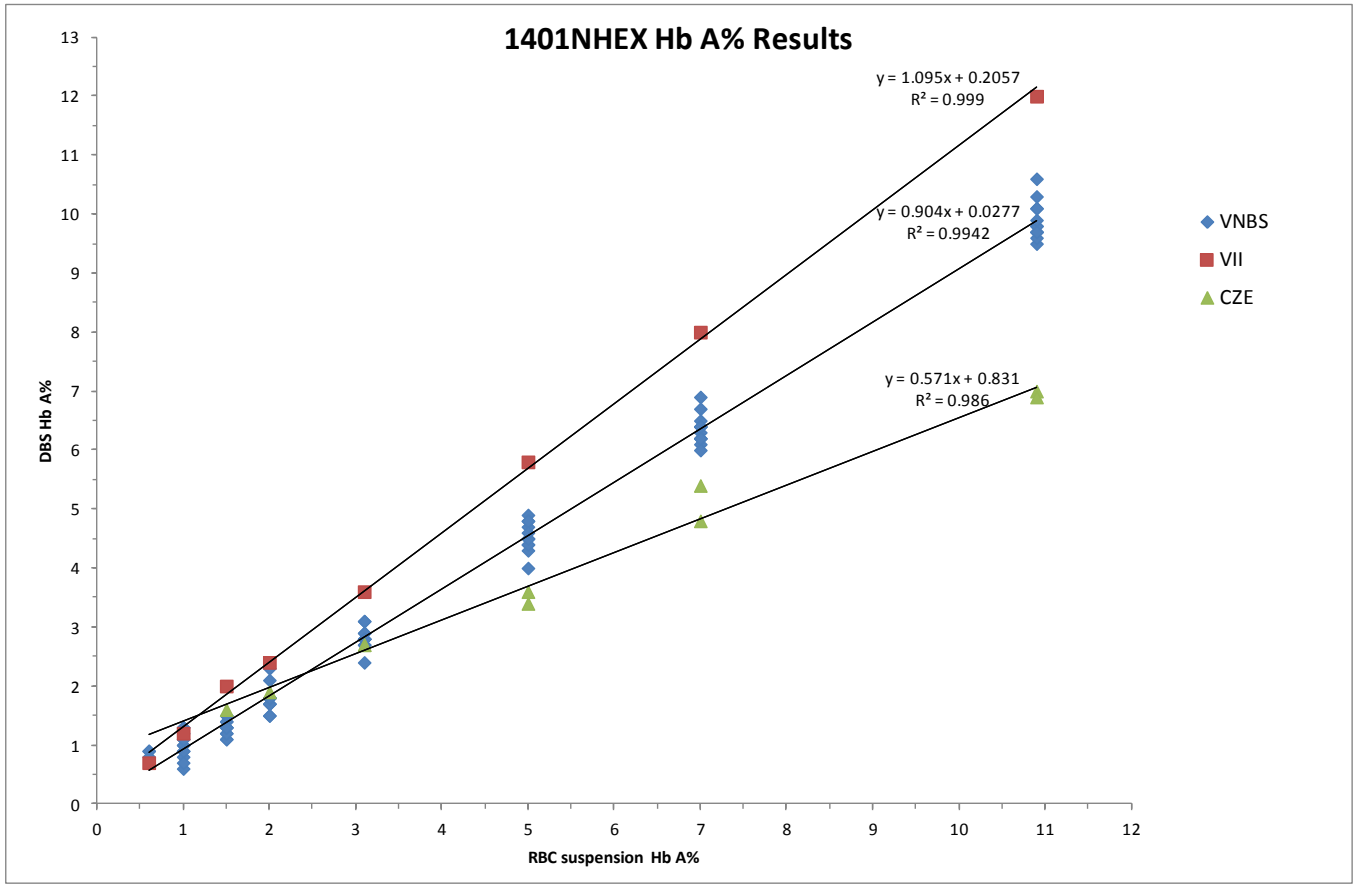


Figure 4.3. Linear regression of the 1401NHEX Haemoglobin A% results.

Legend to Figure 4.3: VNBS = Bio-Rad Variant nbs HPLC, VII = Bio-Rad Variant II HPLC, CZE = Sebia Neonat Hb Fast capillary electrophoresis.

4.4.2 Precision

Ten replicate analyses were made from each of two DBS cards prepared from RBC suspensions with Hb A concentrations from 1.0% to 10.9% (DBS card batches two to eight). The analyses were repeated at three sites (identified as W, X and Y), each using a Bio-Rad Vnbs HPLC analyser. Each site received two cards from each batch, identified as A and B for site W, C and D for site X and E and F for site Y. The results, which are shown in Table 4.4, were analysed by one-way ANOVA firstly for variation between the DBS cards by batch and secondly between sites. There was no significant difference in the Hb A% results from the three sites (W, X and Y) at the $p < 0.05$ level for each Hb A concentration tested ($F(2, 57) = 0.14$ to 2.81, $p = 0.069$ to 0.868). However, there was a statistical difference at the $p < 0.05$ level by individual DBS card for batches two, four, seven and eight, prepared from RBC suspensions with Hb A concentrations of 1.0%, 2.0%, 7.0% and 10.9% respectively. No statistical significance was seen between the results for card batches three, five and six (Hb A concentrations of 1.5%, 3.1% and 5.0% respectively). The descriptive statistics for each card showed little variation of clinical significance between DBS cards, regardless of the apparent statistical significance. The CV% for the DBS are low, ranging from 4.12% to 5.47% for five of six cards tested from batch three (prepared from a RBC suspension of Hb A = 2.0%) to less than 1.0% for all cards tested from batch eight (Hb A% = 10.9%), showing very little variation between the replicate analyses prepared from the same card. These CV% values are similar to those seen in UK NEQAS Haematology Full Blood Count surveys for fully automated cell counters, where the reproducibility is regarded as very good. When the within batch variation

is low, very small differences in the between batch variation may become statistically significant.

4.4.3 Carryover

The experiment to determine the effect of carryover of Hb A was conducted with a specimen prepared to simulate a DBS eluate with a Hb A% of approximately 25% (equivalent to a ‘normal’ infant) analysed immediately before eluates from each of the eight low Hb A% specimens. The results of the carryover experiment are shown in Table 4.5. The carryover from a preceding specimen with a Hb A of approximately 25% ranged from -0.42 to 0.93%. Carryover was detected in three of the eight cases, prepared from RBC suspensions of 1.0%, 3.1% and 7.0% Hb A; the carryover therefore appeared not to be related to the concentration of the Hb A% in the low Hb A specimen. Taking the maximum carryover as 1% of the preceding specimen, a very high specimen (say 25 to 40% Hb A) preceding a specimen from an infant with beta thalassaemia major with an Hb A of 1.4 or 1.5% could be sufficient to push the beta thalassaemia major specimen over the 1.5% Hb A limit and this should be considered when reviewing the results. On the other hand, the variation in results due to carryover from a specimens with a Hb A% of 25% was only 0.1 to 0.2% Hb A, which is the equivalent to the usual reproducibility seen with the instrument in the precision experiment.

Site	Hb A% results by DBS batch and card number													
	2A	**2B**	**3A**	**3B**	**4A**	**4B**	**5A**	**5B**	**6A**	**6B**	**7A**	**7B**	**8A**	**8B**
W	ND	1.3	1.8	1.6	2.3	2.2	3.2	3.1	5	5.1	7.1	6.8	10.7	10.4
	0.7	1.2	1.8	1.7	2.3	2.1	3.2	3.2	5.2	4.9	7	6.8	10.7	10.4
	0.8	1.2	1.8	1.7	2.3	2	3.2	3.1	5	4.9	6.9	6.9	10.7	10.4
	0.8	1.2	1.1	1.7	2.3	2.2	3.2	3.1	5	5.1	7.1	6.8	10.7	10.4
	0.8	1.2	1.8	1.6	2.3	2.1	3.2	3	5.1	4.9	7	6.9	10.5	10.3
	0.8	1.2	1.8	1.8	2.3	2.1	3.1	3.1	5.2	5	7.1	7	10.7	10.4
	ND	1.3	1.9	1.7	2.3	2.1	3.3	3.2	5.1	5	7	6.9	10.5	10.4
	0.7	1.2	1.8	1.6	2.2	2.1	3.1	3	5	5.1	7	6.8	10.6	10.4
	0.8	1.2	1.8	1.8	2.3	2.1	3.2	3.2	5.2	5	7	6.8	10.6	10.2
	ND	1.2	1.8	1.7	2.3	2.1	3.1	3.1	5.2	4.9	7	6.9	10.5	10.4
Mean	0.77	1.22	1.74	1.69	2.29	2.11	3.18	3.11	5.1	4.99	7.02	6.86	10.62	10.37
SD	0.0488	0.0422	0.2271	0.0738	0.0316	0.0568	0.0632	0.0738	0.0943	0.0876	0.0632	0.0699	0.0919	0.0675
CV%	6.33	3.46	13.05	4.37	1.38	2.69	1.99	2.37	1.85	1.75	0.90	1.02	0.87	0.65
	2C	**2D**	**3C**	**3D**	**4C**	**4D**	**5C**	**5D**	**6C**	**6D**	**7C**	**7D**	**8C**	**8D**
X	0.9	1.2	1.6	1.8	2.2	2.2	3.1	3.1	5.1	5	7.1	6.8	10.5	10.7
	0.9	1.1	1.6	1.8	2.2	2.2	3.2	3.1	5.1	5.2	7	6.8	10.5	10.6
	1	1.1	1.6	1.9	2	2	3	3.2	5	5	6.9	6.9	10.5	10.7
	1.1	1.1	1.7	1.8	2.2	2.2	3.1	3.2	4.9	5	7.1	6.8	10.5	10.5
	0.9	1.2	1.6	1.8	2.2	2.1	3.2	3.2	5	4.9	7.1	6.9	10.6	10.5
	0.9	1.2	1.8	1.8	2.2	2.1	3.1	3.2	5	5.2	7.1	7	10.5	10.7
	0.9	1.3	1.8	1.9	2.3	2.3	3.2	3.3	4.9	5.2	7	6.9	10.5	10.5
	0.8	1.1	1.6	1.8	2.3	2.1	3	3.2	5.1	5.1	6.9	6.8	10.5	10.6
	0.9	1.2	1.8	1.6	2.3	2.3	3.2	3.2	5	5	6.9	6.8	10.3	10.5
	1	1.2	1.7	1.8	2.3	2.1	3.1	3.1	4.9	5.2	7	6.9	10.5	10.6
Mean	0.93	1.17	1.68	1.8	2.22	2.16	3.12	3.18	5	5.08	7.01	6.86	10.49	10.59
SD	0.0823	0.0675	0.0919	0.0816	0.0919	0.0966	0.0789	0.0632	0.0816	0.1135	0.0876	0.0699	0.0738	0.0876
CV%	8.85	5.77	5.47	4.54	4.14	4.47	2.53	1.99	1.63	2.23	1.25	1.02	0.70	0.83

Table 4.4. Replicate Haemoglobin A% results returned by sites W, X and Y for DBS card batches 2 to 8 *(continued)*

Site	Hb A% results by DBS batch and card number													
	2E	**2F**	**3E**	**3F**	**4E**	**4F**	**5E**	**5F**	**6E**	**6F**	**7E**	**7F**	**8E**	**8F**
Y	1.3	1.2	1.6	1.8	2.3	2.2	3.2	3.2	5.1	5	7.1	7	10.4	10.6
	1.1	1	1.8	1.6	2.3	2	3.2	3.2	5.2	5.2	7	7.1	10.5	10.1
	1.1	1	1.8	1.8	2.3	2	3.2	3.2	5.1	5	6.9	7	10.5	10.5
	1.2	1	1.8	1.8	2.1	2.2	3.1	3.2	5.1	5	7	6.9	10.3	10.6
	1.1	1.1	1.6	1.8	2.2	2.1	3.2	3.2	4.9	5.1	6.9	7.1	10.3	10.6
	1.1	1	1.8	1.8	2.2	2.2	3	3.1	5.1	5	7	7	10.4	10.6
	1.3	1.3	1.7	1.9	2.3	2	3.2	3.3	5	5.1	6.9	7.1	10.4	10.6
	1.2	1.2	1.6	1.8	2.2	2.1	3.1	3.1	5.1	5	6.9	7	10.7	10.7
	1.2	1.2	1.8	1.8	2.3	2.1	3.2	3.2	5	5.2	7	6.9	10.3	10.3
	1.2	1.2	1.7	1.8	2.3	2.1	3.1	3.1	5	5.2	6.9	6.9	10.4	10.5
Mean	1.18	1.12	1.72	1.79	2.25	2.1	3.15	3.18	5.06	5.08	6.96	7	10.42	10.51
SD	0.0789	0.1135	0.0919	0.0738	0.0707	0.0816	0.0707	0.0632	0.0843	0.0919	0.0699	0.0816	0.1229	0.1792
CV%	6.68	10.14	5.34	4.12	3.14	3.89	2.24	1.99	1.67	1.81	1.00	1.17	1.18	1.71

Table 4.4 (continued)**.**

Legend to Table 4.4: ND = not detected, SD = standard deviation, CV% = coefficient of variation.

Carryover from specimen with Haemoglobin A% ≈ 25%								
DBS batch	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
H1	26.5	26.4	26.4	26.5	26.4	26.3	26.5	26.3
H2	26.2	26.5	26.4	26.5	26	26.3	26.2	26.3
H3	26.4	26.3	26.3	26	26.5	26.5	26.3	26.6
L1	0	1.2	1.6	1.9	3.1	5.1	6.9	10.3
L2	0	1.1	1.4	2.1	3	5	6.8	10.4
L3	0	1	1.6	2	3	4.9	6.9	10.3
L1-L3	0	0.2	0	-0.1	0.1	0.2	0	0
H3-L3	26.4	25.3	24.7	24	23.5	21.6	19.4	16.3
% carryover	**0.00**	**0.79**	**0.00**	**-0.42**	**0.43**	**0.93**	**0.00**	**0.00**

Table 4.5. Carryover calculated from a preceding specimen with a Haemoglobin A concentration of approximately 25%.

4.5 Discussion

The NHS STP has suggested a Hb A of 1.5% or less as indicative of probable beta thalassaemia major in newborn infants with a gestational age of greater than thirty-two weeks, supported by evidence from a major prospective and retrospective study of infants tested at four screening centres in England (Streetly 2013). The experiments described in this chapter attempted to evaluate the performance of instruments used routinely in newborn haemoglobinopathy screening in the UK prepared from RBC suspensions with Hb A concentrations from 0.6% to 10.9%.

Two instruments designed for newborn DBS screening for haemoglobinopathies were used: the Bio-Rad Variant™ Newborn Screening (Vnbs) HPLC analyser with the sickle cell short programme and the Sebia Capillarys Neonat Hb Fast™ system. In addition, one HPLC analyser designed for adult haemoglobinopathy screening, the Bio-Rad Variant II analyser with the beta thalassaemia short programme, was used by one participating laboratory as this was the method used for their secondary investigation of positive screens detected by their primary screening method.

In summary, the results for all instruments demonstrated good linearity, the Bio-Rad VII apparently over-estimated the Hb A% in the DBS specimens compared to the RBC suspensions from which they were prepared, whereas the Sebia Neonat Hb Fast system under-estimated Hb A% in the DBS specimens compared to the cell suspensions. The best correlation between the DBS and the RBC suspension Hb A% results was seen with the Bio-Rad Vnbs instruments. Most instruments failed to detect or quantitate Hb A at a concentration below 1% and the Sebia

Neonat Fast instruments only reliably detected and quantitated Hb A% at concentrations of above 2.0%. Carryover from a preceding DBS simulating a sample from a ‘normal’ infant with an Hb A of 25% was not significant using the Bio-Rad Vnbs instrument and could be accounted for by the uncertainty of the measurement. A limitation of the experiment was that it did not test the carryover from a preceding DBS with much higher Hb A%, e.g. 40%, as might be found with a sample from a very post mature infant or a sample contaminated with a substantial amount of adult blood (or even of adult blood taken in error). The carryover and reproducibility experiments were also only performed using Bio-Rad Vnbs instruments, because of the limited number of low Hb A% DBS cards available, which was a weakness and would warrant repetition with Sebia Neonat Hb Fast instruments.

Verification of a manufacturer's claims for an instrument is required before it is used in the laboratory for diagnostic testing. There are a number of guidelines and recommendations available that are applicable to this process (e.g. Bourner 2014, Briggs 2014, Stephens 2015, CLSI EP-17-A2, CLSI EP-6A, CLSI EP 15-A2, CLSI EP5-A2); although these are mostly relevant to the evaluation of automated cell counting instruments they contain principles relevant to any quantitative procedure.

Both instruments are described as qualitative rather than quantitative by their manufacturers and no LOD is specified for Hb A for either. The detection and quantitation of Hb A by the Bio-Rad Vnbs analyser in all the batches of DBS cards except the batch prepared from a red cell suspension of 0.6% Hb A is comparable to the stated LOD of 1% for Hb fractions S, C, D, E. The detection of Hb A by the Sebia Neonat Hb Fast system was not reliable for DBS cards prepared from red

cell suspensions with Hb A of 2% or less, suggesting that the system was not as sensitive in detecting and quantitating Hb A at very low levels as the Vnbs. A limitation of the study is that only two Neonat Hb Fast instruments were available to take part and hence this observation would benefit from additional work. The Neonat Hb Fast system has been shown to be comparable to the Bio-Rad Vnbs HPLC system when testing both liquid umbilical cord blood samples (Wolff 2012) and capillary DBS samples (Murray 2011). However, the evaluation of the Sebia Neonat Hb Fast system using liquid umbilical cord blood samples noted three discrepancies between the results from the CZE and HPLC systems, one of which was a sample with Hb A of 1.4% by HPLC that was not detected by CZE (Wolff 2012). This is in agreement with the results from the Sebia Neonat Hb Fast systems tested in this study. When the quantitation of different Hb fractions in patients' DBS samples using the Sebia Neonat Hb Fast system was compared with the Bio-Rad Vnbs analyser, the quantitation of Hb A% was good (correlation of $y = 0.75x + 4.54$, $R^2=0.87$) (Murray 2011) but showed a similar under-estimation of Hb A% by the Sebia Neonat Hb Fast as seen in the experiment described in this chapter, where the correlation was $y = 0.61x + 0.74$, $R^2=0.96$, and $y = 0.57x + 0.83$, $R^2=0.97$. The design of the studies was different in that the evaluation with patients' DBS (Murray 2011) examined the correlation between the Sebia Neonat Hb Fast and the Bio-Rad Vnbs using the same DBS samples, whereas the experiment in this chapter compared the results of the DBS samples tested by the Sebia Neonat Hb Fast with the 'expected results' of the red cell suspensions from which the DBS had been prepared. The 'expected' results of the red cell suspensions had been assayed using a Bio-Rad Vnbs. The apparent difference in the results obtained using the Sebia Neonat Hb Fast system is most likely to be

due to a difference in the method and the algorithms used by the two instruments for calculation of the peak areas (Murray 2011).

The difference in the performance between the Bio-Rad and the Sebia newborn screening analysers could lead to additional infants being given a screen positive result for beta thalassaemia major using a cut-off of 1.5% if they were tested using the Sebia Neonat Hb Fast instrument. This is 'failing safe' but might require an adapted strategy for follow-up, depending on the screening instrument used. This could entail a simple confirmation of screen positives by a second biochemical method or by DNA analysis. Screening is not intended to be 100% effective but the level of technical error should be confined to the statistical uncertainty around the cut-off; the risk of an increased number of false positives if reported without consideration of the confounding factors could cause unnecessary anxiety for parents. In the major retrospective and prospective study to examine the appropriateness of the 1.5% cut-off, Streetly noted that a limitation of the work was that it was not a full country survey and more work is in hand to address this (Streetly 2013). There was no information given on the methodology used for Hb A% in the study.

The instruments used in newborn screening are designed for qualitative screening only, although a percentage quantitation is given. It will therefore rest with the laboratory to validate the instrument for this purpose. The challenge for the laboratories will be to obtain sufficient samples at or around the critical cut-off Hb A concentration: there were just 268 'Hb F only' infants born in the UK in the first ten years of the implementation of universal newborn screening (NHS Sickle Cell and Thalassaemia Screening Data Report 2015/16: trends and performance analysis, 2017). This low number will make it difficult for laboratories to rely on

patients' specimens for validation and verification of analysers. It may be that UK NEQAS Haematology could provide support for laboratories by providing sets of cards at a range of Hb A%, prepared in a similar fashion to those used in this study and kept in storage. The cards would need retesting and renewal at intervals to ensure the quality was adequate.

The inclusion of simulated beta thalassaemia DBS cards in the repertoire of cases sent out as part of the EQA scheme would allow better performance assessment of the laboratories; however, the presence of Hb A_2 in the adult blood used in the studies described is a confounding factor and if used as EQA material the source and purpose of the case(s) would have to be made clear in the instructions given to participants. At present, UK NEQAS Haematology does not request the percentage of any Hb fraction present in the cases distributed in the Newborn Sickle Screening scheme and this would require modification to accommodate additional low Hb A% cases to performance assess the detection of infants with beta thalassaemia major.

Chapter 5: The fitness for purpose of UK NEQAS Haematology scoring methods for the assessment of Haemoglobin A_2 performance

5.1 Aim

To examine the effectiveness of current UK NEQAS Haematology performance scoring methods for the detection of out-of-consensus Hb A_2 performance.

5.2 Introduction

External quality assessment (EQA) is a means by which the laboratory demonstrates on a regular basis that its results are comparable to those produced by other testing sites and that they conform to the standards expected for patient care (Libeer, 2001; Plebani, 2005; Miller, 2011). The use of a performance score and acceptable performance specifications ensures the objective evaluation and reporting of large numbers of participants' results by the EQA provider. Measures of performance assessment are required for an accredited EQA scheme (ISO 17043).

Key procedures in the assessment of EQA results are the assignment of a target value and the allowed limits of acceptable performance (Fraser, 1999). The target value may be assigned using a higher-order CRM or reference method, although this requires the EQA material to be commutable and traceable to a CRM or method (Thienpont, 1995; Uldall, 1996; Thienpont, 2003; Vesper, 2009). For most measurands, there is no reference material or method available and the target value in these circumstances may be determined from the formulation of the material (e.g. the known 'weighed-in' quantity of a measurand), the results of a panel of expert laboratories or the consensus mean or median of the results from the participating testing sites after the removal of outlying results (ISO 13528; Jones, 2017b). It has been estimated that 90% of EQA providers use the consensus of participants' results as the target value (Wong, 2005). There are

several alternatives in the choice of consensus target value, depending on the commutability of the survey material, the range of methods in use and the numbers of participants registered. Where the survey material is commutable, it is possible to use the ‘all laboratory consensus mean’, which makes no differentiation between the instruments or methods used for analysis and demonstrates the comparability between different analysers as well as individual laboratories. However, the use of a consensus mean of data from more than one instrument or method favours the method with the greatest number of participants and may make their performance appear superior in any ‘state-of-the-art’ comparison and the mean of the peer group means may be used to give an equal weight to all instrument types in the group (Van Houcke, 2012). The median may also be used as a robust estimator of the consensus target value and the est. SD is calculated from the IQR divided by 1.349 (Tukey, 1977). Since the median is a robust estimator, it is not necessary to trim the data to remove outliers.

Where there is a demonstrated lack of commutability of the survey material the target value for EQA performance assessment is generally derived from the trimmed peer group mean, the peer group being described by the method or analyser type. A minimum of twenty participants is the standard number to form a peer group in UK NEQAS Haematology for statistical robustness, although it has beem suggested that peer groups of as few as six participants may be used (Coucke, 2012). UK NEQAS survey material used for the Abnormal Haemoglobins scheme is commutable for the methods in use at the time of writing, as demonstrated in the study described in Chapter Three.

There are a number of different approaches that have been devised to measure participants’ performance in EQA. These may be grouped into methods based on:

- Passing or failing against a published criterion of acceptable performance,
- The use of a ‘look-up’ principle where a participation score is allocated on the basis of how well the result conforms to the target or clinical outcome.
- The calculation of a score based on a continuous scale such as the z-score (Ley and Ezer, 1974; Coucke, 2012).

How these are applied to the assessment of participants’ performance differs according to the EQA scheme design. In many cases, especially where the score is used for ensuring the laboratory is licensed to practise, the score is reported for a single distribution only and there is no calculation of a cumulative score over a period of time. An example of this approach in haemoglobinopathy EQA is the system used by the College of American Pathologists where the result submitted is reported in terms of the distance from the target value and whether it is within upper and lower acceptable performance limits (CAP, 2017).

The international standards ISO 13528 and ISO 17043 recommend the use of the z-score or similar derivatives as a measure of how far the participant’s result is from the centre of the distribution or target value, as a multiple of the dispersion of the data or the scale (ISO 13528, ISO 17043). A basic z-score describes how many standard deviations from the target value the result is; a score of equal to or greater than 3.0 is regarded as the action signal for unsatisfactory performance and a score of between 2.0 and 3.0 as a warning indicator (ISO 13528). Given a normal distribution of results from a group of competent laboratories, approximately 0.3% would be expected to have a z-score above 3.0 and approximately 5% of participants a z-score of between 2.0 and 3.0 (Coucke, 2017). The use of a z-score and statistically determined performance limits focuses on the technical aspects of EQA performance and facilitates a state-of-the-art comparison of performance

between laboratories and methods. It is important to understand what the performance limits mean in terms of patient safety, especially around the cut-off points for clinical decision-making such as screening for beta thalassaemia carriers. The use of performance limits determined from published biological variation data has been described by a number of authors (Panteghini, 2015, Sandberg, 2015) and analytical goals for Hb A_2 have suggested 7% as a minimum goal for the TE as a compromise between biological variation, expert opinion and clinical need (Mosca, 2013). This TE represents a variation in absolute Hb A_2 measurement of approximately 0.25% for a 'true' Hb A_2 value of 3.6%. Other authors have recommended that duplicate measurements of Hb A_2 should be within 0.2% of Hb A_2 of the final result (Stephens, 2015).

The UK NEQAS Haematology performance assessment and scoring system was initially introduced in 1996 for the assessment of automated cell counting schemes (England, 1995) and later adapted for other schemes, including the haemoglobinopathies. UK NEQAS Haematology uses a consensus trimmed mean as target value: data is made robust to the influence of outliers by 10% trimming (5% of results from the top and bottom), followed by log transformation to give an approximately symmetrical distribution before calculation of the consensus trimmed mean value. Depending upon the scheme design, the consensus mean may be the all laboratory trimmed mean (ALTM), the method trimmed mean (MTM) or individual instrument group (SMTM). A number of methods for outlier removal have been described (Algorithm A of ISO 13528; Grubbs, 1969; Dixon, 1950; Tukey, 1977; Rousseeuw, 1993; Wilrich, 2007) although there is little difference in the methods used where the participant group is greater than twenty (Coucke, 2012).

UK NEQAS Haematology uses a z-score, which it terms the deviation index (DI) to describe how far the individual participant's result is from the target value. The DI relates the difference between the participant's result and the target to a robust estimator of the dispersion of the data, the historic standard deviation (HSD) (Tukey, 1977; Healy, 1979), which is calculated from the results of the six most recent distributions, excluding the current one. The use of a robust estimator such as the HSD is to smooth the variation in the SD from one distribution to another and its consequent impact on the DI.

The UK NEQAS Haematology analytical performance score is a cumulative, retrospective indicator of performance calculated from the sum of the DI values for the most recent six specimens containing the analyte after removal of the arithmetic sign (the 'absolute' DI value). The statistical performance limits are adjusted by including a multiplier in the score that weights the DI values so that a score of equal to or greater than 100 is classified as persistent unsatisfactory performance in any UK NEQAS Haematology scheme (UK NEQAS Participants' Manual, 2017). The multiplier varies by analyte and can be used to adjust the sensitivity of the scoring procedure. To avoid unduly penalising a participant for a single 'blunder', e.g. the transposition of results at data entry that could result in high DI values not related to analytical performance, any DI value greater than 3.5 is truncated to 3.5 before calculation of the analytical performance score. The untruncated DI is still reported in the participant's report. The UK NEQAS Haematology performance tariff automatically refers any score of equal to or greater than 100 for review; scores between 80 and 100 are reported to the participating laboratory as borderline and scores below 80 as satisfactory. Long-term, retrospective scoring of participants' performance is an established feature of

EQA assessment but it requires frequent survey distributions to make the assessment effective within a clinically relevant timeframe. This is a challenge in haemoglobinopathy EQA and frequency may only be sufficient in large schemes. Within Europe, the frequency of distribution of EQA in the haemoglobinopathies is generally relatively infrequent (Vives-Corrons, 2014) making cumulative performance scoring impractical. In these situations, schemes are more likely to assess performance specimen by specimen, without a formal monitoring of performance trends.

UK NEQAS Haematology uses the Hb A_2 method trimmed mean or MTM as the target value for performance assessment, largely for historical reasons. The two main method groups in use for performance assessment at the time of writing were CZE and HPLC, for which performance is scored by method group. Although not used for performance assessment by UK NEQAS, the SMTM, calculated by individual instrument type, is reported for information for analysers with more than twenty instruments registered. Instrument manufacturers and laboratory managers express a wish to see the SMTM value as they feel that this gives a like-for-like comparison on the technical performance of their analyser type (UK NEQAS Haematology unpublished Steering Committee papers, 2015).

The NHS STP's algorithm for the identification of a beta thalassaemia carrier makes no allowance for the analytical method used; therefore, the use of the ALTM may be a more appropriate target against which to assess performance than the MTM. The purpose of the study described is to review UK NEQAS Haematology's participants' data to investigate the effectiveness of the current scoring system for the assessment of performance. Of particular concern is the extent to which an individual, 'one-off' EQA error, which could potentially contribute to an incorrect

diagnosis had it occurred with a patient's specimen, might go undetected by the performance scoring process and whether adjustments to the scoring system could be made to improve the detection of such errors. The study will review the relationship between the performance score and individual DI results, the appropriateness of the target value used for performance assessment, what a DI value means in terms of absolute Hb A_2 measurement and the impact of adjusting the multiplier on the numbers of participants identified as unsatisfactory performers.

5.3 Methods

5.3.1 Data analysis

UK NEQAS Haematology participants' data was analysed using a bespoke external quality assessment (EQA) data management system provided by KPMD Ltd (Sheffield, UK) for the calculation of the ALTM, the MTM and the SMTM values, the DI and the analytical performance scores as part of the regular UK NEQAS Haematology data analysis and reporting process (UK NEQAS Haematology Participants' Manual, 2017).

In brief, trimmed participants' data (10% of results in total, i.e. 5% top and bottom) is log-transformed. A robust standard deviation (RSD) - the historical SD (HSD), calculated for the data set from the most recent six specimens (not including the current distribution), is used to smooth the effect of the variation in the dispersion of data from one distribution to another on the DI (Healy, 1979).

The deviation index (DI) is calculated using log-transformed data as shown in Equation 5.1. At the time of writing, the MTM is used as the target value (x_{pt}).

$$DI = \frac{x_i - x_{pt}}{SD_{pt}}$$

Equation 5.1. Calculation of the Deviation Index (DI)

Where x_i is the participant's result
x_{pt} is the trimmed mean (ALTM, MTM or SMTM, as appropriate)
SD_{pt} is a robust SD (the HSD)

The DI is is used to calculate the UK NEQAS analytical performance score: the DI values for the most recent six results returned by a participant are summed after

removal of the arithmetical sign and truncation of DI values to a maximum of 3.5. The sum of the six values is multiplied by nine for Hb A_2 at the time of writing, to weight the score so that an analytical performance score equal to or greater than 100 is deemed indicative of persistent unsatisfactory performance. The multiplier value of nine was arrived at empirically to place approximately 5% of participants in the persistent unsatisfactory performance category at any one survey. This is deemed a manageable number and allows the scheme to focus support for adverse performance on the laboratories with the greatest problems. Adjustment of the multiplier allows the sensitivity of the scoring to be altered.

Additional statistical analysis for the study was undertaken using STATA 14 (Stata Corp, Texas, USA) and Microsoft Excel 2010 (Microsoft Corporation, California, USA).

5.3.2 Data used for the study, distributions 1304AH to 1603AH

The results of thirty-four survey material pools performance assessed for Hb A_2 in eighteen surveys 1304AH to 1603AH were used in the study. The majority of the results (more than 95%) had been entered by the participants directly into the UK NEQAS Haematology Results service and had not been corrected for data entry errors. The remainder were entered by UK NEQAS Haematology staff using double-data entry. The raw data returned was edited before analysis in the study to remove blank returns, partial data returns and results from manual methods. All data was anonymised before use.

5.3.3 The distribution of Deviation Index results by analytical performance score

The results were sorted by survey material pool into nine categories according to their absolute DI value (equal to or greater than three, equal to or greater than two but less than three, or less than two) and the laboratory's analytical performance score for Hb A_2 (equal to or greater than 100, equal to or greater than 80 but less than 100, or less than 80).

The Hb A_2 results for each specimen in each of the nine categories were screened to identify ‘blunders’ and potential misdiagnoses. ‘Blunders’ in EQA are erroneous results that arise as a consequence of reporting data for the EQA scheme in a format that is different from that used by the laboratory for patients’ results, e.g. the entry of a result in a different order from usual, the conversion of results to a different unit of measurement for submission to the scheme or a typographical data entry error for results that would normally be transmitted automatically via a computer interface. For the purpose of this study, a ‘blunder’ was defined as a result so out-of-consensus that it was unlikely to be a Hb A_2 measurement, identified by scrutiny of each out-of-consensus result in the context of the expected result for the specimen and the others in the survey in which it was distributed. Based on these criteria, any Hb A_2 result less than 0.5% or greater than 10% was reviewed as a potential blunder, such as the transposition of Hb A_2 and Hb F or Hb S results or major clerical errors.

Potential misdiagnoses were defined as results that could lead to a clinically incorrect interpretation had they been reported on a patient’s sample, using the NHS STP’s cut-off value for Hb A_2 of greater than 3.5% as indicative of a beta

thalassaemia carrier. The ALTM was used as the ‘correct’ value for this comparison.

5.3.4 The use of different target values for performance assessment

Four different target values were calculated for each survey material pool. Three were available from the UK NEQAS Haematology database (the ALTM, MTM and SMTM). The fourth, called the ‘Adjusted ALTM’, was calculated for each survey material pool as the mean of the sub-method means for the instrument groups; calculated to eliminate the potential impact of different numbers of instrument groups on the calculation of the ALTM.

The difference between each result in the study and each of the four means was calculated, using the ALTM and adjusted ALTM for the survey material pool and the MTM and SMTM appropriate for the method principle (HPLC or CZE) and the analyser. The mean and SD of the differences against each target were calculated by analyser group for all results in the study.

5.3.5 Calculation of performance scores using different multiplication (weighting) factors

All results and instrument groups for distributions 1304AH to 1603AH were rescored against the usual MTM target values using multipliers of ten, eleven and twelve and the numbers of persistent unsatisfactory performance scores counted by the multiplier value. The impact of different multipliers on the performance scores of participants with out-of-consensus results was assessed.

5.4 Results

5.4.1 Data used for the study, distributions 1304AH to 1603AH

Each survey material pool contained Hb A, Hb A_2 and Hb F with ALTM Hb A_2 results from 1.47% to 5.03%. The numbers of results for each pool ranged from 289 to 332. From 11,607 initial returns, a total of 10,463 Hb A_2 results were used in the study. The most commonly represented HPLC analyser was the Bio-Rad Variant II running either the beta thalassaemia short programme or the dual kit programme. The 'miscellaneous HPLC' group included Agilent 1100 and 1200, Menarini HA8140, Kontron, Shimadzu LC10 and LC20, Varian Prostar, Waters and Zivak HB100 analysers and laboratories that had not provided the HPLC analyser model. The majority of CZE analysers were Sebia Capillarys; a small number of results were returned from a Beckman Coulter P/ACE MDQ analyser and a Helena V8 analyser. For the purposes of this study, all Sebia CZE analysers were analysed as a single group because of the negligible difference between the results produced by the three analysers, as shown in Chapter Two (Figure 2.11). The total numbers of results by analyser group are shown in Table 5.1.

5.4.2 The distribution of Deviation Index results by analytical performance score

The distribution of absolute DI values by score and survey material pool is shown in Table 5.2.

Method group	Sub-method group	No.of results	%
Capillary zone electrophoresis	Sebia	1761	16.8
	Miscellaneous CZE	40	0.4
	Total CZE	**1801**	**17.2**
High Performance Liquid Chromatography	Bio-Rad D10	943	9.0
	Bio-Rad Variant Classic	331	3.2
	Bio-Rad Variant II Beta thal short programme	2245	21.5
	Bio-Rad Variant II Dual kit programme	815	7.8
	Menarini HA8160	1165	11.1
	Trinity Primus 2	629	6.0
	Tosoh G7	712	6.8
	Tosoh G8	1271	12.1
	Miscellaneous HPLC	551	5.3
	Total HPLC	**8662**	**82.8**
TOTAL		**10463**	

Table 5.1. The numbers of results used in the study by method and sub-method group

Specimen	ALTM Hb A_2 (%)	N	DI	Score			Blunders	Potential misdiagnoses
				>100	<100>80	<80		
1304AH1	**4.28**	289	>3	1	2	4	1	4
			>2<3	0	1	7	0	3
			<2	5	11	258	0	0
1304AH2	**2.85**	289	>3	1	1	0	0	1
			>2<3	1	2	1	0	1
			<2	4	11	268	0	0
1305AH3	**2.15**	292	>3	2	1	6	0	3
			>2<3	1	3	6	0	0
			<2	7	8	258	0	0
1306AH1	**2.71**	296	>3	0	2	2	2	1
			>2<3	1	2	2	0	0
			<2	1	11	275	0	0
1306AH2	**2.28**	296	>3	0	2	1	0	0
			>2<3	1	2	4	0	0
			<2	1	11	274	0	0
1401AH1	**2.81**	297	>3	1	0	0	0	0
			>2<3	3	3	2	0	0
			<2	3	7	278	0	0
1401AH3	**2.4**	297	>3	5	2	2	0	0
			>2<3	1	2	3	0	0
			<2	1	6	275	0	0
1402AH1	**2.57**	307	>3	1	1	0	0	0
			>2<3	1	1	4	0	0
			<2	1	9	289	0	0
1402AH3	**3.99**	307	>3	0	1	3	0	3
			>2<3	3	3	7	0	3
			<2	0	7	283	0	6
1403AH1	**2.35**	308	>3	2	1	3	1	0
			>2<3	0	4	15	0	0
			<2	1	9	273	0	0
1403AH2	**1.97**	308	>3	3	5	10	2	0
			>2<3	0	5	19	0	0
			<2	0	4	262	0	0
1404AH2	**2.66**	310	>3	0	2	2	1	0
			>2<3	2	4	9	0	0
			<2	2	7	282	0	0
1404AH3	**2.35**	309	>3	2	3	3	1	1
			>2<3	0	3	8	0	0
			<2	2	7	281	0	0

Table 5.2 Numbers of Haemoglobin A_2 results by absolute Deviation Index (i.e.with arithmetical sign removed) and performance score. *(Continued on next page)*

Specimen	ALTM Hb A_2 (%)	N	DI	Score			Blunders	Potential misdiagnoses
				>100	<100>80	<80		
1405AH1	**4.31**	306	>3	3	1	0	0	2
			>2<3	1	0	15	0	1
			<2	4	22	260	0	0
1405AH2	**1.47**	304	>3	5	5	18	1	0
			>2<3	0	6	28	0	0
			<2	3	12	227	0	0
1406AH1	**2.31**	308	>3	4	1	3	1	0
			>2<3	4	9	5	0	0
			<2	3	9	270	0	0
1406AH2	**2.31**	308	>3	5	3	0	0	1
			>2<3	4	6	5	0	0
			<2	2	10	273	0	0
1406AH3	**2.13**	307	>3	7	2	1	0	1
			>2<3	1	7	2	0	0
			<2	3	10	274	0	0
1501AH1	**2.71**	311	>3	2	1	1	0	0
			>2<3	1	2	4	0	0
			<2	6	18	276	0	0
1502AH1	**5.03**	305	>3	1	2	1	0	2
			>2<3	3	3	13	0	0
			<2	7	19	256	0	0
1502AH2	**2.08**	305	>3	3	6	10	0	2
			>2<3	4	2	12	0	0
			<2	4	16	248	0	0
1502AH3	**1.58**	303	>3	8	9	14	0	1
			>2<3	1	6	26	0	0
			<2	2	9	228	0	0
1503AH1	**2.59**	305	>3	1	2	0	1	0
			>2<3	3	4	8	0	0
			<2	5	15	267	0	0
1504AH1	**2.78**	304	>3	1	2	0	0	1
			>2<3	3	1	0	0	0
			<2	8	9	280	0	0
1504AH2	**2.24**	303	>3	6	4	0	0	1
			>2<3	3	1	5	0	0
			<2	3	7	274	0	0
1505AH2	**2.48**	311	>3	0	0	0	0	0
			>2<3	4	3	3	0	0
			<2	7	10	284	0	0

Table 5.2 (*Continued)*

Specimen	ALTM Hb A_2 (%)	N	DI	Score			Blunders	Potential misdiagnoses
				>100	<100>80	<80		
1505AH3	**1.98**	310	>3	2	2	2	1	1
			>2<3	3	3	9	0	0
			<2	6	8	275	0	0
1506AH2	**2.45**	314	>3	4	0	3	0	2
			>2<3	5	2	5	0	0
			<2	2	5	288	0	0
1506AH3	**2.52**	312	>3	4	0	1	1	0
			>2<3	4	4	4	0	0
			<2	3	3	289	0	0
1601AH1	**2.45**	326	>3	2	0	1	0	1
			>2<3	6	5	5	0	0
			<2	4	5	298	0	0
1601AH2	**2.39**	326	>3	6	1	1	0	1
			>2<3	1	1	0	0	0
			<2	5	8	303	0	0
1602AH1	**2.37**	332	>3	1	2	2	0	0
			>2<3	4	4	7	0	0
			<2	5	11	296	0	0
1602AH3	**1.94**	328	>3	3	4	6	0	2
			>2<3	5	8	27	0	0
			<2	2	4	269	0	0
1603AH2	**2.45**	330	>3	0	2	6	0	3
			>2<3	3	2	9	0	0
			<2	7	7	294	0	0
TOTALS		10463		282	511	9670	13	48
				2.7%	4.9%	92.4%	0.1%	0.5%

Table 5.2 (*Continued*)[2]

[2] Table 5.2 shows the distribution of Haemoglobin A_2 results by Deviation Index (DI) and performance score for surveys 1304AH to 1603AH, with the number of probable blunders and potential misdiagnoses. DI values are shown with the arithmetical sign removed.

A total of 264 / 10,463 (2.4%) results had DI values of greater than 3.0, of which eighty-six had an unsatisfactory performance score of equal to or greater than one hundred. Of the remaining 178 results with DI values greater than 3.0, seventy-two had a borderline performance score (score of eighty to one hundred) and 106 had a satisfactory score of less than eighty. Thirteen results (0.1% of the 10,463 results in the study) were classified as probable 'blunders' and all had DI values greater than 3.0. The causes of the 'blunders' were typographical errors on the part of the participant, e.g. entering 2.5% as 205%, and the transposition of the Hb A_2 result for another parameter, e.g. Hb A%, Hb S% or Hb F%. A total of 470 (4.5%) of the results in the study had DI values of between 2.0 and 3.0 and 9,729 (92.7%) had DI values of less than 2.0. There were no blunders identified with DI values of less than 3.0. These results are summarised in Table 5.3.

Forty-eight results were identified as potential misdiagnoses, i.e. a sample with an ALTM of less than 3.5% reported as greater than 3.5% and *vice versa.* Table 5.4 shows the summary of the distribution by performance score and absolute DI value of the forty-eight results. Thirty-five of the potential misdiagnoses had a DI value greater than 3.0; seven had DI values between 2.0 and 3.0 and six DI values of less than 2.0. Only one of the thirteen potential misdiagnoses with a DI value of less than 3.0 had a score of greater than one hundred and would have triggered an automatic performance review.

Thirty-six of the potential misdiagnoses had DI values greater than 3.0 or performance scores greater than one hundred, so would be easy to screen out for review. Of the remaining twelve results, six had absolute DI values between 2.0 and 3.0; one of these results was a 'false positive', i.e. a Hb A_2 result greater than 3.5% reported for a sample with a normal ALTM Hb A_2 value and five were 'false

negatives', i.e. normal results reported for samples with target ALTM Hb A_2 values greater than 3.5%. Six results had absolute DI values less than 2.0 and performance scores less than eighty (in the red shaded area of Table 5.4), which meant that they would not be identified for further investigation either from a raised or borderline score or DI value. All six were from the same survey material pool (1402AH3, ALTM Hb A_2 value 3.99%) and were 'false negatives', with Hb A_2 results of 3.4 to 3.5%. The DI values ranged from -1.46 to -1.83 and all were HPLC users.

DI	N	Score			Blunders	Potential misdiagnoses
		>100	<100>80	<80		
>3	264 (2.4%)	86	72	106	13 (0.1%)	35
>2<3	470 (4.5%)	77	114	279	0	7
<2	9729 (92.7%)	119	325	9285	0	6
TOTAL	10463	282	511	9670	13	48

Table 5.3. Summary of results by absolute DI values and analytical performance scores with the numbers of blunders and potential misdiagnoses

DI	Score		
	>100	<100>80	<80
>3	5	13	17
>2<3	1	2	4
<2	0	0	6
TOTAL	6	15	27

Table 5.4. Summary of absolute DI values and corresponding analytical performance scores for results that would represent a potential misdiagnosis if they were patients' samples.[3]

[3] The grey shaded area shows the results that might escape review based on a borderline analytical performance score and DI value. The red area results that had both a satisfactory score and DI value. The yellow shaded area shows specimens that would be easily identified for follow up because of a raised score (>100) or DI (>3.0).

Figures 5.1 to 5.6 show the distribution of the DI values by performance score for the results of six survey material pools (1304AH1, 1305AH3, 1402AH3, 1502AH1, 1502AH2 and 1601AH1) relative to the persistent unsatisfactory performance and borderline unsatisfactory performance action points.

In each Figure, results with a score greater than one hundred represent persistent unsatisfactory performance, calculated from the most recent six results. Where the DI for the current survey falls close to zero (or is at least less than 2.0), the performance is satisfactory for the current survey and may indicate a performance issue that has been resolved, as the score 'lags' in its responsiveness to corrective actions. Performance scores above one hundred with DI values greater than +2 or less than -2 for any survey indicate persistent unsatisfactory performance that has not been resolved. Between one and nine laboratories fell into this category at each distribution; since the scores are cumulative, some of these cases may represent the same participant from one survey to the next. In this case, the performance issues will be under review and investigation.

Results with an analytical performance score between eighty and one hundred are classed as 'borderline'. A rising score in this zone, with a DI value greater than 2.0, may represent an adverse performance trend that would benefit from early corrective action. Scores between eighty and one hundred should always be reviewed for trends in performance, if only by the participant laboratory.

Any individual DI value greater than +2 or less than -2 warrants review and values greater than +3 or less than -3 may indicate a serious performance issue (ISO 13528). This action should be undertaken regardless of the performance score.

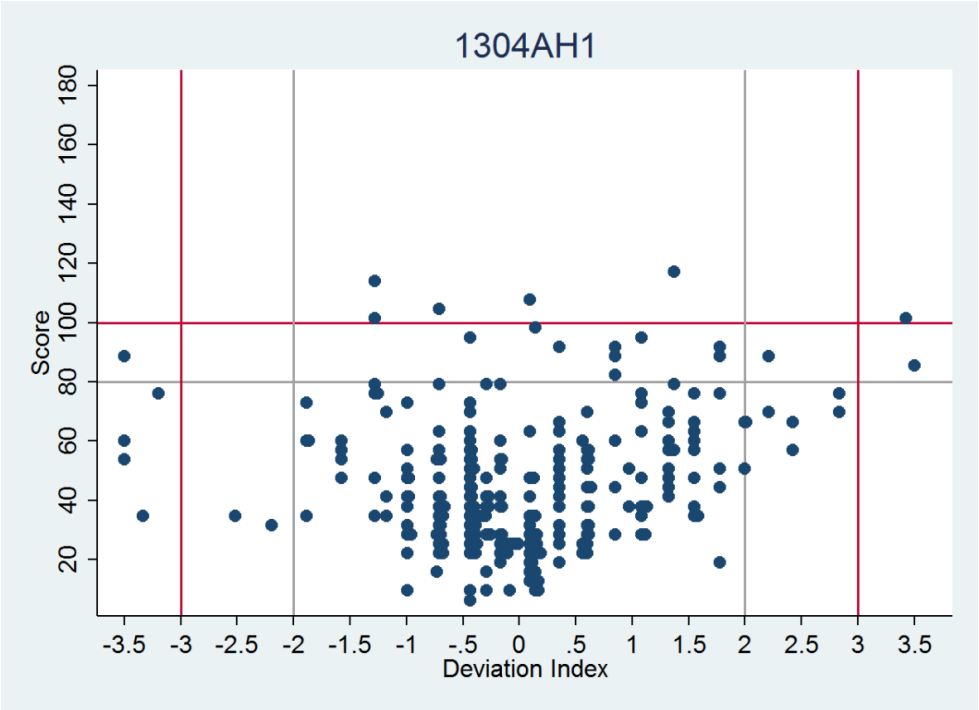


Figure 5.1. Haemoglobin A_2 analytical performance scores by deviation index for specimen 1304AH1 (Hb A_2 ALTM 4.28%). DI values >3.5 truncated to 3.5.

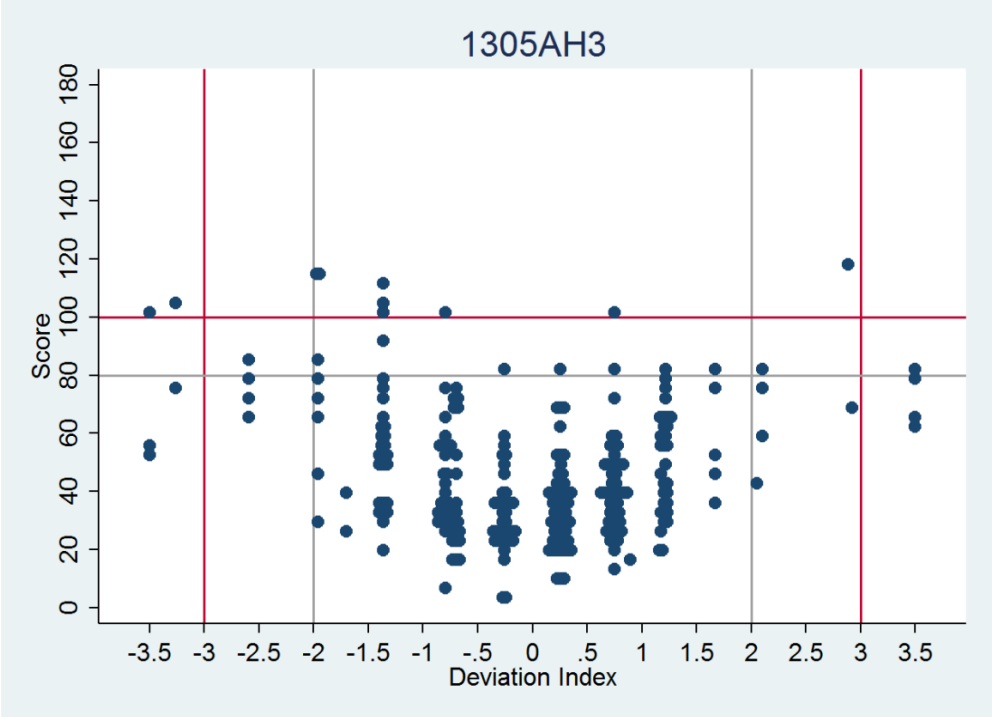


Figure 5.2. Haemoglobin A_2 analytical performance scores by deviation index for specimen 1305AH3 (Hb A_2 ALTM 2.15%). DI values >3.5 truncated to 3.5.

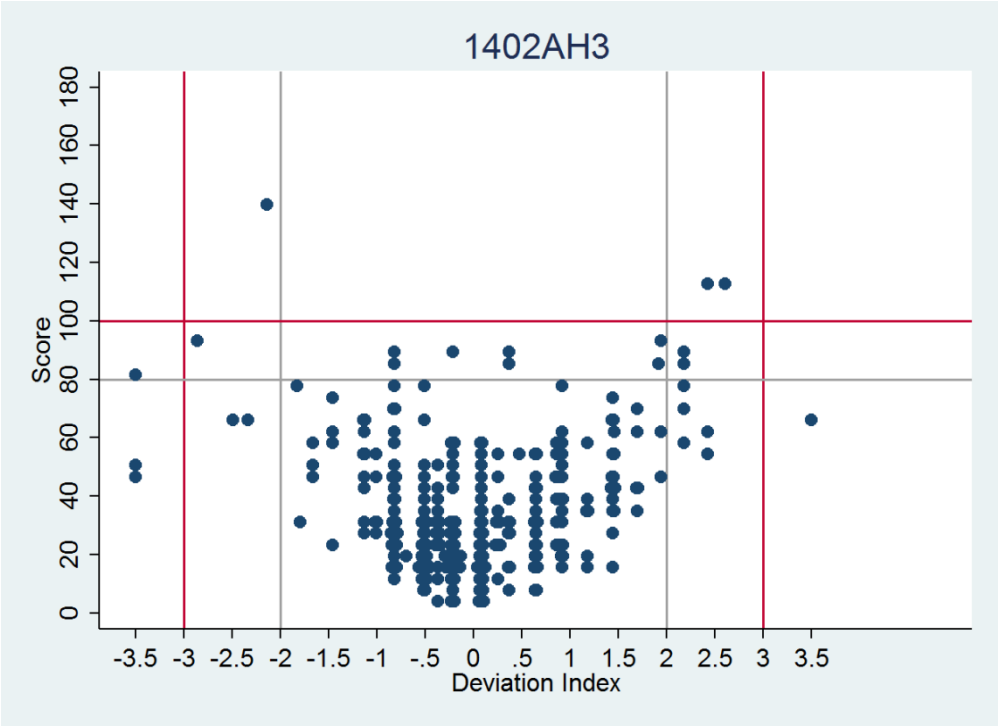


Figure 5.3. Haemoglobin A_2 analytical performance scores by deviation index for specimen 1402AH3 (Hb A_2 ALTM 4.99%). DI values >3.5 truncated to 3.5.

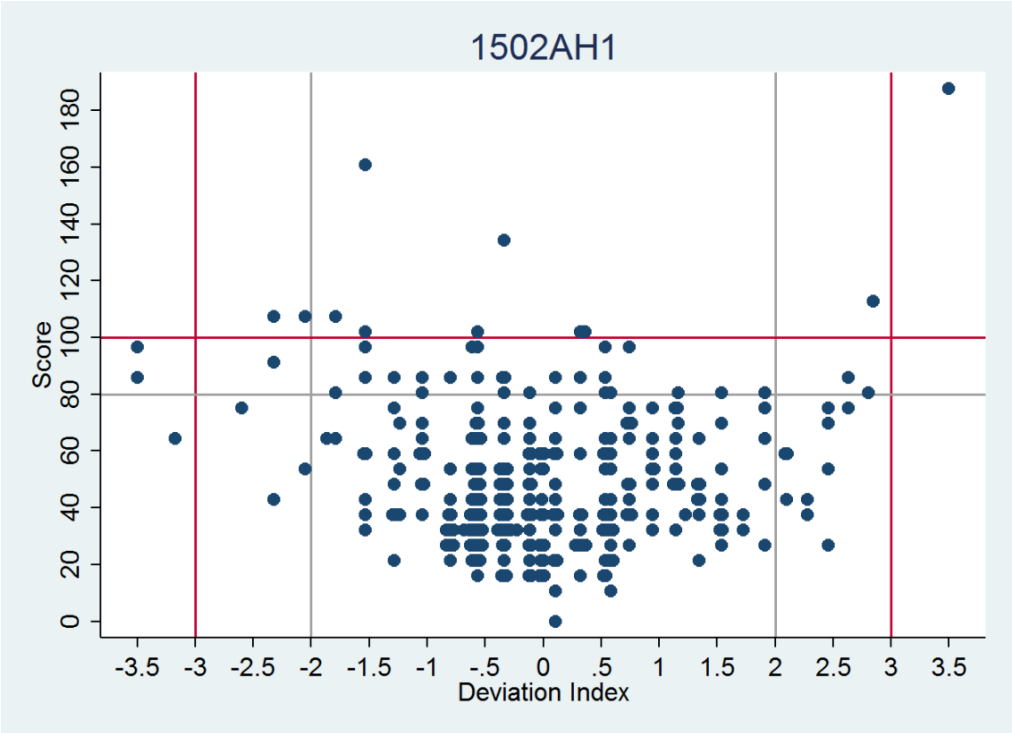


Figure 5.4. Haemoglobin A_2 analytical performance scores by deviation index for specimen 1502AH1 (Hb A_2 ALTM 5.03%). DI values >3.5 truncated to 3.5.

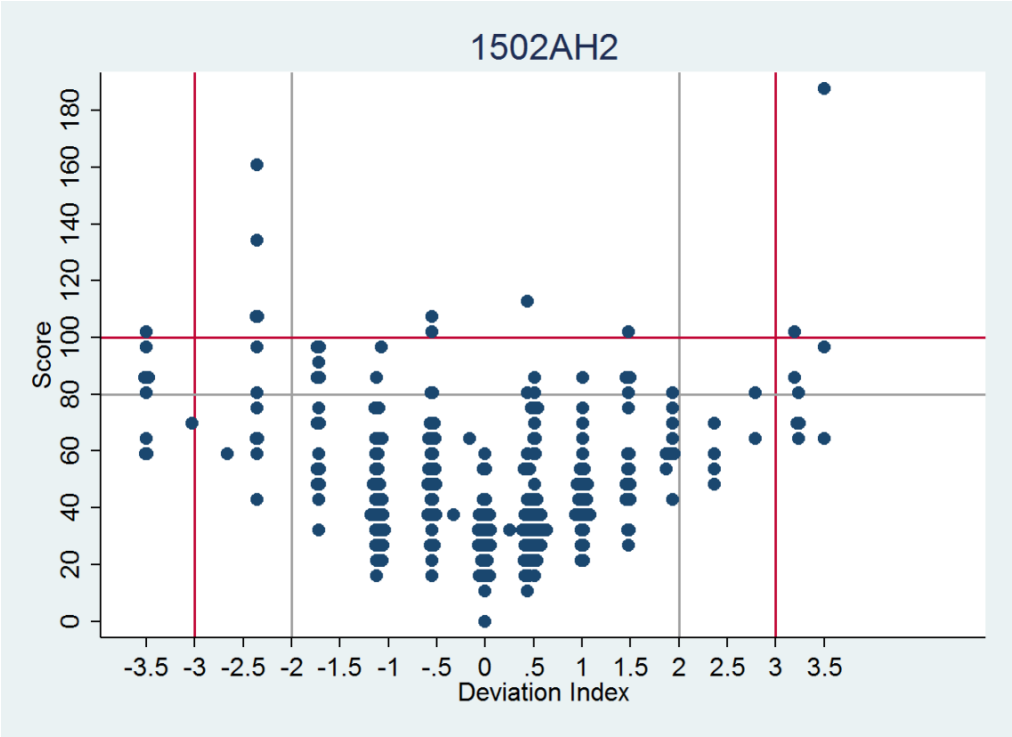


Figure 5.5. Haemoglobin A_2 analytical performance scores by deviation index for specimen 1502AH2 (Hb A_2 ALTM 2.08%). DI values >3.5 truncated to 3.5.

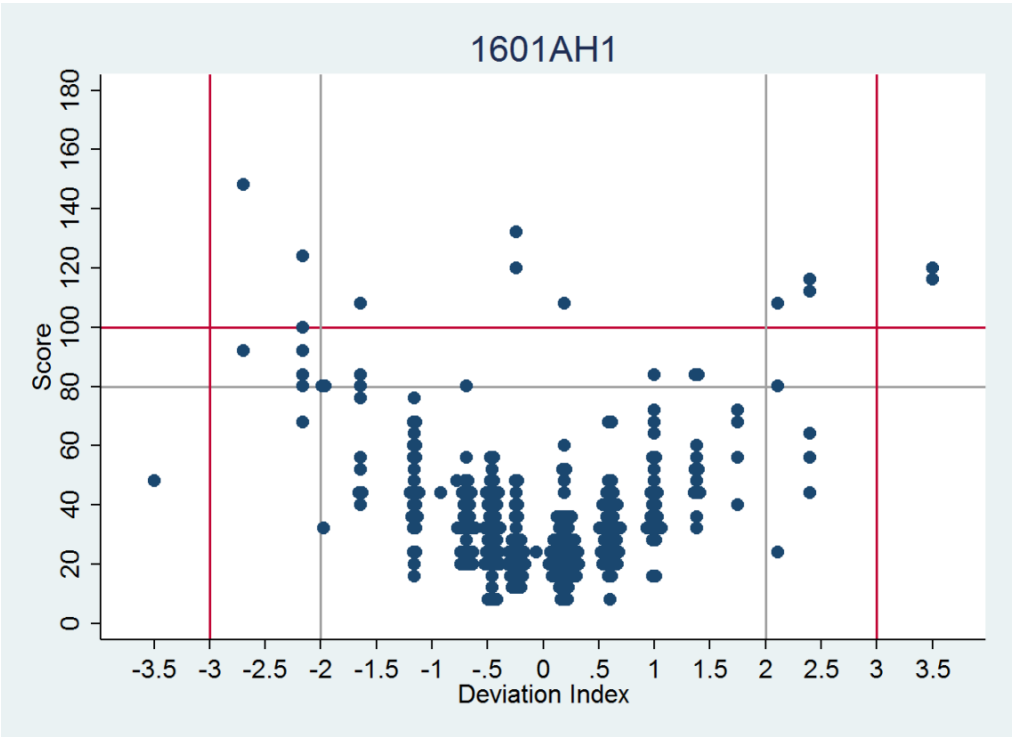


Figure 5.6. Haemoglobin A_2 analytical performance scores by deviation index for specimen 1601AH1 (Hb A_2 ALTM 2.45%). DI values >3.5 truncated to 3.5.

5.4.3 The target value used for performance assessment

The differences between the participants' results and the ALTM, Adjusted ALTM, SMTM and MTM were calculated and the mean of the differences for all data in the study determined by instrument group for each of the four target values. The means of the differences by instrument group are shown in Table 5.5 and Figure 5.7.There was no observable difference in the means of the differences of the participants' results by ALTM, Adjusted ALTM or MTM for any instrument group. There appears to be no advantage in adjusting the target value to correct for the different numbers of each instrument type in the ALTM by using the mean of the sub-method means, suggesting that the ALTM is not significantly skewed by the numbers of different instrument types in the data used in this study. There also is little variation in the mean differences of the ALTM and the MTM consensus target values. The mean differences of the participants' results against the ALTM, adjusted ALTM and MTM reflect the biases observed in the performance of the different analysers.

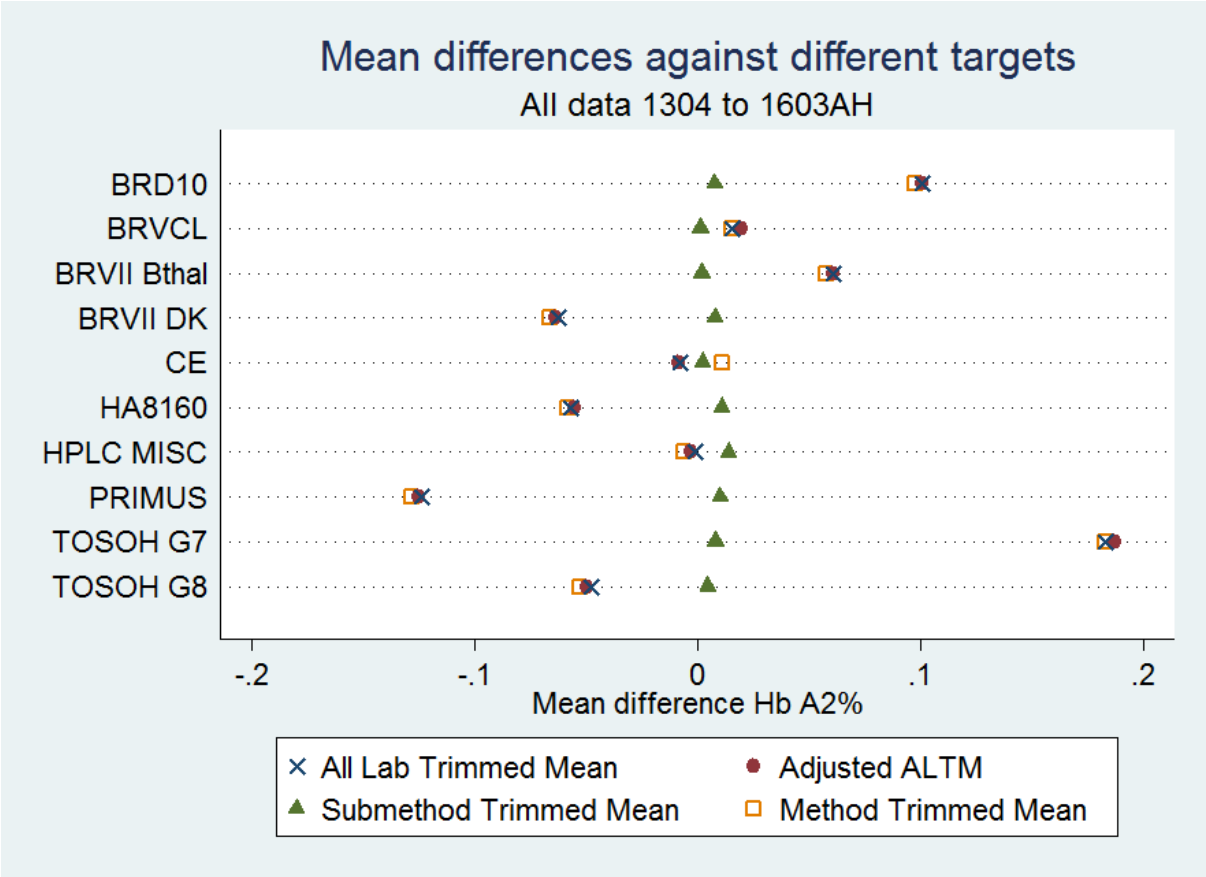
As might be expected, the means of the differences calculated using the SMTM is closer to zero than when using the other target values, by a factor of approximately ten. The only exception to this observation is for the miscellaneous HPLC group. For the CZE group the means of the differences against all four target values are very similar and close to zero; the SMTM and MTM would be expected to be the same for this group since they include the same set of instruments.

Instrument group		Target Value			
		ALTM	Adj ALTM	SMTM	MTM
CZE	N	1801	1801	1801	1801
	Mean difference	-0.0077	-0.0083	0.0023	0.0110
	SD	0.1537	0.1564	0.1445	0.1481
BRD10	N	943	943	943	943
	Mean difference	0.1004	0.1008	0.0076	0.0976
	SD	0.2819	0.2806	0.2656	0.2795
BRVCL	N	331	331	331	331
	Mean difference	0.0152	0.0194	0.0013	0.0152
	SD	0.1685	0.1696	0.1365	0.1671
BRVII Bthal	N	2245	2245	2245	2245
	Mean difference	0.0608	0.0607	0.0018	0.0577
	SD	0.2097	0.2109	0.1802	0.2081
BRVII DK	N	815	815	815	815
	Mean difference	-0.0624	-0.0634	0.0079	-0.0663
	SD	0.2031	0.2019	0.1941	0.2029
HA8160	N	1165	1165	1165	1165
	Mean difference	-0.057	-0.0551	0.0109	-0.0584
	SD	0.3337	0.3353	0.3189	0.3332
HPLC MISC	N	551	551	551	551
	Mean difference	-0.0014	-0.0032	0.0139	-0.0059
	SD	0.3057	0.3058	0.2877	0.3056
PRIMUS	N	629	629	629	629
	Mean difference	-0.1236	-0.1251	0.01	-0.128
	SD	0.2202	0.2194	0.2097	-0.2211
TOSOH G7	N	712	712	712	712
	Mean difference	0.1827	0.1867	0.0081	0.1827
	SD	0.3195	0.3879	0.3416	0.3912
TOSOH G8	N	1271	1271	1271	1271
	Mean difference	-0.0478	-0.0495	0.0045	-0.0524
	SD	0.3017	0.3005	0.2549	0.3027

Legend to Table 5.5:

CZE:	Sebia capillary zone electrophoresis;
BR D10:	Bio-Rad D10 HPLC
BRVCL:	Bio-Rad Variant Classic HPLC
BRVII BThal:	Bio-Rad Variant II, Beta thal short programme HPLC
BRVII DK:	Bio-Rad Variant II, Dual Kit programme, HPLC
HA8160:	Menarini HA8160 HPLC
HPLC MISC:	Miscellaneous HPLC
PRIMUS:	Trinity Primus 2 HPLC
TOSOH G7:	Tosoh G7 HPLC
TOSOH G8:	Tosoh G8 HPLC

Table 5.5 Mean differences of the returned results against different target values, by instrument group.



Legend to Figure 5.7:

CE:	Sebia capillary zone electrophoresis;
BR D10:	Bio-Rad D10 HPLC
BRVCL:	Bio-Rad Variant Classic HPLC
BRVII BThal:	Bio-Rad Variant II, Beta thal short programme HPLC
BRVII DK:	Bio-Rad Variant II, Dual Kit programme, HPLC
HA8160:	Menarini HA8160 HPLC
HPLC MISC:	Miscellaneous HPLC
PRIMUS:	Trinity Primus 2 HPLC
TOSOH G7:	Tosoh G7 HPLC
TOSOH G8:	Tosoh G8 HPLC

Figure 5.7 Distribution of mean differences of results against ALTM, Adjusted ALTM, SMTM and MTM, by instrument group

5.4.4 Calculation of performance scores using different multiplication (weighting) factors

At the time of writing the multiplier used in the calculation of analytical performance score is nine. In this study 2.7% of results (282 / 10,463) had scores greater than one hundred using a multiplier of nine. This multiplier was selected empirically when scoring was introduced for Hb A_2 in 2004 to place approximately 5% of participants in the persistent unsatisfactory performance category (England, 1995). Altering the multiplier up or down adjusts the sensitivity of the score; increasing the multiplier increases the number of participants identified with persistent unsatisfactory performance but this increases the associated performance management workload. The impact of increasing the multiplier on the number of participants identified as persistent unsatisfactory performers was assessed by trial and error using multipliers of ten, eleven and twelve. Table 5.6 shows the average DI value over six specimens that would produce a score of one hundred with different multipliers and the number of participants that would be identified as unsatisfactory performers.

Multiplier	Average DI	No. of PUP scores
9	1.85	282 (2.7%)
10	1.67	457 (4.4%)
11	1.52	708 (6.8%)
12	1.39	1002 (9.6%)

Table 5.6 Average DI values over six specimens required to produce a performance score of one hundred using different multipliers and the number of results in the study identified as persistent unsatisfactory performance (PUP) by multiplier

The DI value varies according to the dispersion of the results for the survey material pool as well as the difference of the individual result from the target mean, even when smoothed by using the HSD. Table 5.7 shows the DI values produced for a range of differences from the target value using data from ten survey material pools with varying Hb A_2 results. Applying an acceptable analytical performance limit of ± 0.25% of Hb A_2 (Mosca 2013) to the figures in Table 5.7 would classify any result that is more than ± 0.3% from the target value as out-of-consensus (note that UK NEQAS Haematology Hb A_2 results are returned to one decimal place only and the value of 0.25% has been rounded up to 0.3%). Using multipliers of nine, ten, eleven and twelve with the average DI from the ten survey material pools suggests that a multiplier of eleven would give a performance score equal to or greater than one hundred for an actual difference from the target value of approximately 0.3% of Hb A_2.

The six potential misdiagnoses identified with scores less than eighty and DI values less than two were reviewed after being re-scored with different multipliers; although their scores increased with the increase in multiplier, all results still had scores less than one hundred with multipliers of ten and eleven and four had scores of less than eighty. One had a score of greater than one hundred and two had scores of greater than eighty with a multiplier of twelve.

R-M Hb A_2%	Deviation index										Avge DI	Score by multiplier			
	Hb A_2 2.31%	Hb A_2 2.48%	Hb A_2 3.99%	Hb A_2 2.85%	Hb A_2 2.15%	Hb A_2 2.57%	Hb A_2 1.94%	Hb A_2 2.31%	Hb A_2 2.52%	Hb A_2 4.31%		9	10	11	12
-0.5	2.8	2.2	1.6	2.1	2.7	2.2	3.2	2.3	2.2	1.0	2.03	109	122	134	146
-0.4	2.2	1.7	1.3	1.7	2.1	1.7	2.4	1.8	1.7	0.8	1.58	85	95	104	114
-0.3	1.7	1.2	1.0	1.2	1.5	1.3	1.8	1.3	1.2	0.6	1.16	63	70	77	84
-0.2	1.2	0.7	0.7	0.8	0.9	0.8	1.1	0.8	0.8	0.3	0.74	40	44	49	53
-0.1	0.7	0.3	0.4	0.4	0.4	0.4	0.5	0.4	0.4	0.1	0.36	20	22	24	26
0	0.2	0.2	0.1	0.1	0.0	0.0	0.1	0.0	0.0	0.1	0.07	4	4	5	5
0.1	0.3	0.6	0.2	0.3	0.6	0.4	0.6	0.4	0.4	0.3	0.37	20	22	25	27
0.2	0.7	1.0	0.5	0.7	1.1	0.8	1.1	0.8	0.7	0.4	0.71	38	43	47	51
0.3	1.1	1.4	0.8	1.0	1.5	1.2	1.6	1.2	1.1	0.7	1.05	57	63	70	76
0.4	-	1.7	1.0	1.3	2.0	1.5	2.0	1.6	1.4	0.9	1.34	72	80	88	96
0.5	-	2.1	1.3	1.7	2.4	1.9	2.5	1.9	1.7	1.1	1.66	90	100	110	120

Table 5.7 Deviation Index values by difference from the target Haemoglobin A_2% for ten survey material pools (Hb A_2 targets of 1.94 to 4.31%) and the analytical performance score produced using different multipliers.

Suggested acceptable performance limits are shown by the shaded areas of the table, using a limit of ±0.3% Hb A_2 from the target. R is the result returned and M the consensus trimmed mean target value. Scores in red indicate persistent unsatisfactory performance.

5.5 Discussion

A number of methods are available for the evaluation of EQA data and there is no ideal method for statistical performance assessment in EQA (Coucke, 2017). There is a recognised need for greater transparency in the setting of analytical performance specifications (APS) and a clear understanding of their limitations, which should be made available to participants and professional, regulatory or oversight bodies (Jones, 2017b). This study examined the current indicators for monitoring participants' performance of Hb A_2 measurement in the UK NEQAS Haematology Abnormal Haemoglobins scheme, how they reflect and report performance to the laboratory and how they might be developed.

The UK NEQAS Haematology analytical performance score is a descriptor of the long-term, retrospective performance of the laboratory for an analyte. Its purpose is to screen the EQA results returned by participants against objective performance specifications and identify laboratories whose performance warrants further investigation and review, either by the scheme or the participant. The scoring system is based on the principles described in international standards (ISO 13528 and ISO 17043); however, it was developed for performance assessment of automated cell counting, where the interpretation of the test results and their application in clinical diagnosis and monitoring are different from the measurement of Hb A_2. Hb A_2 is expressed as a quantitative value but the application of a strict cut-off for the identification of a beta thalassaemia carrier, as in the NHS STP's algorithm, renders its interpretation almost qualitative or outcome based. The use of a strict cut-off Hb A_2 value to screen for beta thalassaemia carriers with the classical raised Hb A_2 phenotype (Hb A_2 greater than 3.8%) is relatively straightforward but the identification of carriers of mild beta

thalassaemia mutations with borderline Hb A_2 values (3.3 to 3.8%) is a greater challenge (Mosca 2008, Giambona 2008; Stephens, 2012).

The APS used by UK NEQAS Haematology for the evaluation of Hb A_2 data are defined as the 'state-of-the-art', i.e. the highest level of analytical performance possible at the time. The scheme is complex in design, however, and includes an interpretation of the results obtained against a clinical scenario. The interpretative element of the scheme is not currently scored but doing so would give an alternative model of APS based on outcome and the application of clinical decision-making (Sandberg 2015).

To assess the state of the art, it is necessary to assign an appropriate target value against which performance is measured. The continued use of the MTM by UK NEQAS on historical grounds must be questioned, as there is no more justification in separating CZE methods from HPLC as separating one HPLC analyser from another. Little difference was shown between the MTM and ALTM for individual analyser models and there is little evidence that the use of the ALTM puts any instrument group at a disadvantage. However, the use of the mean of the sub-method means (described as the 'adjusted ALTM' in the study) should be kept under review as technology and the numbers of participants using different methodologies change. The use of the SMTM as the target value might improve the apparent performance of individual analyser groups; however, its use for performance assessment is counter-intuitive when no allowance for method difference is made in the NHS STP's algorithm and the bias seen between analysers in EQA performance has also been demonstrated with clinical material (Paleari, 2012; Paleari, 2017). There is also the danger that the use of the SMTM will 'normalise' the existence of a bias between different analysers rather than

acting as a driver to improve harmonisation of performance. A barrier to harmonisation is the lack of a reference standard for Hb A_2 other than the now nearly thirty years' old International Reference Material (WHO 89/666).

The Hb A_2 analytical performance score is an indicator of long-term bias and imprecision; the DI on the other hand is an indicator of the TE for an individual result. When these two performance indicators were compared, a number of results with satisfactory DI values were associated with scores indicative of persistent unsatisfactory performance (PUP) and *vice versa.* This serves to illustrate that both comparators need to be taken into account when assessing performance. This is of particular relevance because the score retains a 'memory' of past unsatisfactory performance and remains high in comparison to the DI values in the immediate period following actions to a correct performance problem. A larger proportion of results (2.4%) had DI values greater than 3.0 than predicted from a normal distribution (0.3%) and this reflects the observation that the distribution of EQA data may be leptokurtic, with a taller, more slender peak and heavier tails than a normal distribution, resulting in a larger proportion of participants than expected with z-scores of ±3.0 (Heydorn, 2008). Although DI values are reported to participants in each distribution report, a review of DI values is not currently included in the UK NEQAS Haematology APS for Hb A_2. The scheme should implement a review of DI values as part of the summary of performance made to participants and consider the introduction of a category of a one-off unsatisfactory performance error or UP based on a DI value of more than 3.0.

UK NEQAS Haematology should pilot the use of a different multiplier to increase the sensitivity of the performance score. Increasing the multiplier to eleven would

apply the suggested analytical goals based on biological variation for Hb A_2 (Mosca 2013). This will bring a greater proportion of labs into the zone of borderline or UP and there is a risk that the specificity of the score may be reduced. Using a multiplier of eleven will more than double the number of laboratories with PUP scores according to this study and the use of an automated, rules-based performance assessment system that flags the trends in performance in the survey report may be necessary to manage the workload. This implementation of a change to the multiplier would introduce a step change in the performance scores and require careful shadow scoring and communication to participants, explaining why and how previously satisfactory performance had become borderline or unsatisfactory.

Increasing the multiplier to eleven did not push the six results that represented a potential misdiagnosis into the PUP range, although two moved into the borderline category. All six results were from the same distribution and the survey material pool had the lowest beta thalassaemic Hb A_2 target value in the study (3.99%), although it was above the Hb A_2 range expected for a carrier of a mild beta thalassaemia mutation. The use of the of the 3.5% cut-off value for defining a potential misdiagnosis in this study may have been too blunt as a discriminator, although this is how it is intended for use by the NHS STP. The NHS STP's algorithm includes consideration of the MCH result and all the survey material pools in the study had been issued with a MCH result consistent with a beta thalassaemia carrier, i.e. less than 27 pg. Review of the coded and written interpretive comments provided by the six laboratories in question for this specimen revealed that five of the six had suggested possible beta thalassaemia carrier status for the 'patient'. This raises the idea that the benefit of the scheme

would be enhanced if an element of assessment by outcome or clinical decision making were introduced through the scoring of the interpretive comments as well as the analytical result, i.e. taking an holistic approach to performance assessment. The educational impact of such an approach would benefit from the inclusion of more specimens at or around the cut-off Hb A_2 value of 3.5%.

In summary, the current use of the DI value and performance score detects the majority of PUP participants when the results are assessed in terms of clinical outcome, i.e. whether the laboratory has correctly identified a specimen as from a beta thalassaemia carrier or not. There is scope for increasing the sensitivity of the procedure by amending the multiplier used in scoring, adding a review of the DI values at each distribution and removing the anomaly of using the MTM as a target value. The use of the mean of sub-group means should be kept under review with changes in methods used by participants. Finally, there may be some benefit to participants' understanding of the relationship between the DI and performance score if this could be displayed graphically as in this study, even if on an annual summary basis.

Chapter 6: Errors in fraction identification and interpretation of results

6.1 Aim

To investigate the frequency and types of errors made by different categories of participants when reporting fraction identification and case interpretation in the UK NEQAS Abnormal Haemoglobins external quality assessment scheme and whether these would prove a barrier to scoring of these parameters.

6.2 Introduction

The UK NEQAS Haematology Abnormal Haemoglobins scheme is designed to assess as far as practicable all aspects of the laboratory investigation of the haemoglobinopathies by non-molecular techniques. Haemoglobin fraction identification has been a part of the scheme since it commenced in the 1970s but the interpretation of results against a clinical scenario to replicate the procedure that would be applied to a patient's sample is a more recent introduction.

The identification of variant Hb fractions is presumptive based on the physicochemical features of the Hb molecule. Because of the large number of Hb variants known (Weatherall and Clegg, 2001; Bain, 2006) and their co-migration in protein chemistry-based separation techniques, the identification of a variant Hb using non-molecular methods should be based on the results of two laboratory tests using different principles (NHS STP Antenatal Laboratory Handbook 2017; Ryan, 2010; Traeger-Synodios, 2015). Where laboratories do not have a confirmatory method(s) available to them in-house, they may comply with this requirement by sending patients' samples for testing to a second laboratory. This action is not recommended in external quality assessment (EQA) as it might constitute 'collusion' between participants; the referral laboratory may receive specimens from a number of screening laboratories and therefore will test multiple

identical EQA specimens; an error by the referral laboratory will have an impact on the EQA performance of the screening laboratory, making the scoring of performance and the allocation of adverse penalties difficult.

Result interpretation and recommendation of further actions are essential features of haemoglobinopathy testing and should take into account other laboratory results, e.g. the FBC, iron status and blood film morphology, the patient's ethnic background and the reasons for testing (Clarke, 2000; Bain, 2006; Bain, 2010). Each UK NEQAS Abnormal Haemoglobins specimen is accompanied by a clinical scenario against which participants are required to interpret the results they obtain. Laboratories undertaking antenatal screening in the English NHS receive education, training and advice in this area from the NHS STP and the laboratory diagnosis of the haemoglobinopathies is the subject of national and international guidelines (Ryan, 2010; Traeger-Synodios, 2015). The NHS STP's Laboratory Handbook is freely available as a source of professional guidance to all laboratories.

Some element of performance assessment is required for EQA scheme accreditation against ISO 17043 (ISO 17043:2010) and is an expected part of EQA scheme design (Urassa, 2016). There is no formal performance scoring of the fraction identification and interpretation elements of the UK NEQAS Abnormal Haemoglobins scheme although all laboratories receive an overall educational summary of the incorrect returns in the report of each distribution. An out-of-consensus fraction identification by a UK laboratory is followed up by letter from the scheme; however, in the absence of a performance score, laboratories outside the UK receive no active indication of the need for performance improvement. There are models and precedents for scoring of qualitative and interpretive EQA

results (Kettelhut, 2003; James, 2012) and the performance assessment of the qualitative sickle solubility test option in the scheme is already in place.

The number of non-UK participants registered for full participation in the scheme, i.e. fraction identification, quantitation and interpretation, has increased significantly since the scheme's inception and exceeds the number of UK participants at the time of writing. The extension of performance scoring to fraction identification and interpretation will include all laboratories and there must be sufficient consensus between laboratories for the scoring to be credible, which may be a particular challenge for interpretation because of different oversight standards in place in different countries. The presumptive nature of Hb fraction identification and different levels of service expertise amongst laboratories need to be taken into account for the fair and effective performance assessment of both fraction identification and case interpretation. To make scoring in these areas effective, it is essential to understand the number and possible types of error made by participants, including errors that result from the design of the scheme itself.

6.3 Methods

6.3.1 Data analysis

Data was analysed using Microsoft Excel 2010 with the downloaded Toolpak (Microsoft Corporation, California, USA) and STATA 14 (StataCorp, Texas, USA).

The fraction identification and interpretive comment codes submitted by participants registered in the UK NEQAS Haematology Abnormal Haemoglobins scheme in the six surveys distributed between August 2015 and June 2016 (surveys 1504AH to 1603AH) were used in the study. The raw data was edited before analysis to remove blank and incomplete results, results returned where the specimen quality had been reported as unsatisfactory and those from non-clinical participants, e.g. instrument manufacturers, pharmaceutical companies and quality assurance laboratories. Participants were allocated to one of five laboratory types: English NHS Trusts, non-English NHS Trusts (i.e. NHS Trusts based in Scotland, Wales and Northern Ireland), UK private laboratories, laboratories based in the Republic of Ireland (RoI) and laboratories outside the British Isles (Non-British Isles laboratories). All data was anonymised before use.

6.3.2 Fraction identification assessment

UK NEQAS Haematology identifies the fractions ‘essential’ for diagnosis for each specimen, e.g. Hb A and Hb S for a sickle cell carrier patient, plus other fractions that would be ‘expected’ to be present, e.g. Hb A_2 and Hb F. The essential and expected Hb fractions had been allocated for the specimens in the study in the original survey reports to participants.

The numbers of correct and incorrect reports were determined by laboratory type. Participants' returns were classified as correct if all Hb fractions essential for

diagnosis were reported with no additional fractions other than the expected ones. Incorrect fraction identification results were classified into four categories following scrutiny of the individual returns by the participant including free text comments they had given with their submissions:

1. **Unable to identify (UI)**: reporting a non-specified (NS) fraction or a differential fraction identification (e.g. Hb S or Hb C) stating they were unable to identify further with the method(s) available or that the specimen would be referred to another laboratory for identification. Occasional laboratories reporting their instrument as out of use in a written comment but who had returned a nominal fraction identification (e.g. Hb A or non-specified (NS) fraction) were also included in this category.
2. **Haemoglobin A not recorded (ANR):** reporting a result correct in all respects except for the omission of Hb A. These results are potentially confusing, e.g. reporting a sickle cell carrier specimen (Hb A and Hb S) as 'Hb S' could imply a diagnosis of sickle cell disease. Investigation of the error with participants usually shows that the Hb A fraction was recorded at analysis but overlooked during data entry.
3. **Transposition of specimens or results (Tx):** a pre or post-analytical error with potentially serious consequences. The transposition of EQA specimens can be identified in a distribution where the specimens have different Hb variant patterns or different Hb A_2 or Hb F concentrations. They are identified as a separate category of error because a single transposition will result in two incorrect fraction identification submissions.

4. **All other errors (Other):** this category included all remaining fraction identification errors, e.g.:
    - Failure to report essential fractions (other than failure to report Hb A, which has been counted separately).
    - Reporting fractions in addition to the expected fractions without an adequate explanation, e.g. reporting Hb S and Hb C or Hb C and Hb E without explaining that the laboratory is unable to identify further with the methods available or reporting a NS fraction in addition to the expected fractions.
    - Reporting an incorrect fraction, e.g. reporting Hb D instead of Hb S.
    - A combination of any of these errors or a combination of one or more of these errors with the failure to report Hb A.

6.3.3 Confirmation of fraction identification by more than one method

The methods used for Hb fraction identification were investigated by laboratory type and specimen content for the specimens distributed in surveys 1504AH to 1603AH.

6.3.4 Interpretive comments assessment

Six cases were selected to assess the use of UK NEQAS Haematology interpretive comment codes: 1504AH1, 1504AH2, 1505AH2, 1601AH3, 1603AH1 and 1603AH3, representing a range of the haemoglobinopathy cases distributed in the study period.

Coded comments are used for reporting the interpretation of UK NEQAS Abnormal Haemoglobins specimens to allow the automated analysis of several hundred

participants' results. The coded interpretive comments available fall into two categories: those offering a diagnosis or describing the results obtained, e.g. 'sickle cell carrier', 'delta-beta thalassaemia carrier' or 'raised Hb F', and comments giving recommendations for further action, e.g. 'partner testing indicated'. Essential and acceptable comment codes were allocated to each specimen as part of the study according to the background case details given and the guidance on interpretation and recommendation given in the NHS STP's Antenatal Laboratory Handbook. Essential comments were those considered necessary for unambiguous reporting; acceptable comments were those that did not confuse or contradict the diagnosis. The list of coded comments available is shown in Appendix 3.

The interpretive comments returned by the two largest groups of participants (the English NHS Trusts group and the Non-British Isles laboratories) were compared. Each group was of sufficient size to provide statistically meaningful results and the English NHS Trusts group is subject to active oversight by the NHS STP, in contrast to the Non-British Isles group. Participants who had made errors in fraction identification were still included in the interpretive comment analysis. This may seem counterintuitive and is not practice in other UK NEQAS Haematology schemes; however, it served to illustrate some of the causes of error in interpretation for the purpose of the study. There is a mismatch in some cases between incorrect fraction identification and correct interpretation, e.g. a laboratory may fail to report Hb S in fraction identification, but will report the consensus Hb S% and a correct interpretation, consistent with the incomplete fraction identification being a data entry error.

Interpretive comments and recommendations for action were assessed as 'correct', ‘partially correct’ or 'incorrect' depending on their agreement with the essential and acceptable comment codes, using the following criteria:

- **Correct:** all essential comments listed, with or without the acceptable comments.
- **Partially correct:** one or more essential comments omitted but only essential and/or acceptable comments used.
- **Incorrect:** using a comment that was not on the essential or acceptable list, even if the essential/acceptable comments had been reported.

Diagnostic comments and recommendations for action were assessed separately. Where the interpretive comment was considered incorrect, no assessment of the recommendations for action was undertaken.

The interpretive comments and recommendations for action were assessed in two ways: an overall analysis of the comments returned by the two major laboratory types and an analysis of the combination of comments returned by the individual laboratories. Although the total summary of comments returned by laboratory type is indicative of the overall state-of–the-art, the combination of comments used by the individual laboratory is important for an understanding the provision of unambiguous advice to the clinician.

6.4 Results

6.4.1 Laboratories included in the study

The results from 283 to 308 participants per distribution for the six distributions (three specimens per distribution) between August 2015 and June 2016 were included in the study, giving a total of 1799 sets of results from laboratories in the UK and thirty-three other countries. The laboratories included in each survey are shown by country in Table 6.1. After the UK, the largest numbers of participants were located in Belgium, Turkey, Portugal, Italy and Israel.

Table 6.2 and Figure 6.1 show the numbers of participants by laboratory type. The proportion of results returned by non-British Isles laboratories ranged from 140/283 (49.5%) in August 2015 to 161/307 (52.4%) in June 2016.

Country	Distribution					
	1504AH	1505AH	1506AH	1601AH	1602AH	1603AH
Australia	1	1	1	1	1	1
Austria	3	3	3	3	3	3
Belgium	30	30	29	32	29	29
Bulgaria	1	1	1	1	1	1
Cyprus	1	1	1	1	1	1
Czech Republic	1	1	1	1	1	0
Denmark	1	1	1	1	1	1
Finland	1	1	1	1	1	1
France	1	2	2	2	2	2
Germany	2	3	3	3	3	3
Greece	2	2	2	2	2	2
India	1	1	1	1	1	1
Israel	13	13	13	12	12	13
Italy	17	18	18	19	17	17
Jamaica	0	1	1	1	0	1
Kenya	1	1	1	1	0	0
Kuwait	5	5	6	6	6	6
Luxembourg	1	1	1	1	1	1
Malta	1	1	1	1	1	1
Netherlands	0	0	1	1	1	0
Netherlands Antilles	1	1	1	1	1	1
Norway	6	6	5	6	6	6
Poland	1	1	1	1	1	1
Portugal	16	18	19	21	21	22
Republic of Ireland	6	7	7	6	7	5
South Africa	1	1	1	1	1	1
Spain	0	1	1	1	3	4
Sultanate of Oman	1	1	1	1	1	1
Sweden	3	2	3	3	3	3
Switzerland	5	5	5	5	5	5
Turkey	20	21	23	25	28	30
UK	137	142	141	141	144	141
United Arab Emirates	3	3	3	3	2	2
Viet Nam	0	0	0	0	1	1
TOTAL	**283**	**296**	**299**	**306**	**308**	**307**

Table 6.1. Number of participants returning results by country for distributions 1504AH to 1603AH (August 2015 to June 2016)

Laboratory type	Distribution					
	1504AH	**1505AH**	**1506AH**	**1601AH**	**1602AH**	**1603AH**
NHS (English Trusts)	117	121	120	120	123	120
NHS (Non-English Trusts)	17	17	17	17	17	17
UK private sector	3	3	3	3	3	3
Republic of Ireland	6	6	6	5	6	4
Non British Isles	140	149	153	161	159	163
TOTAL	**283**	**296**	**299**	**306**	**308**	**307**

Table 6.2. Numbers of laboratories returning results in each survey, by laboratory type

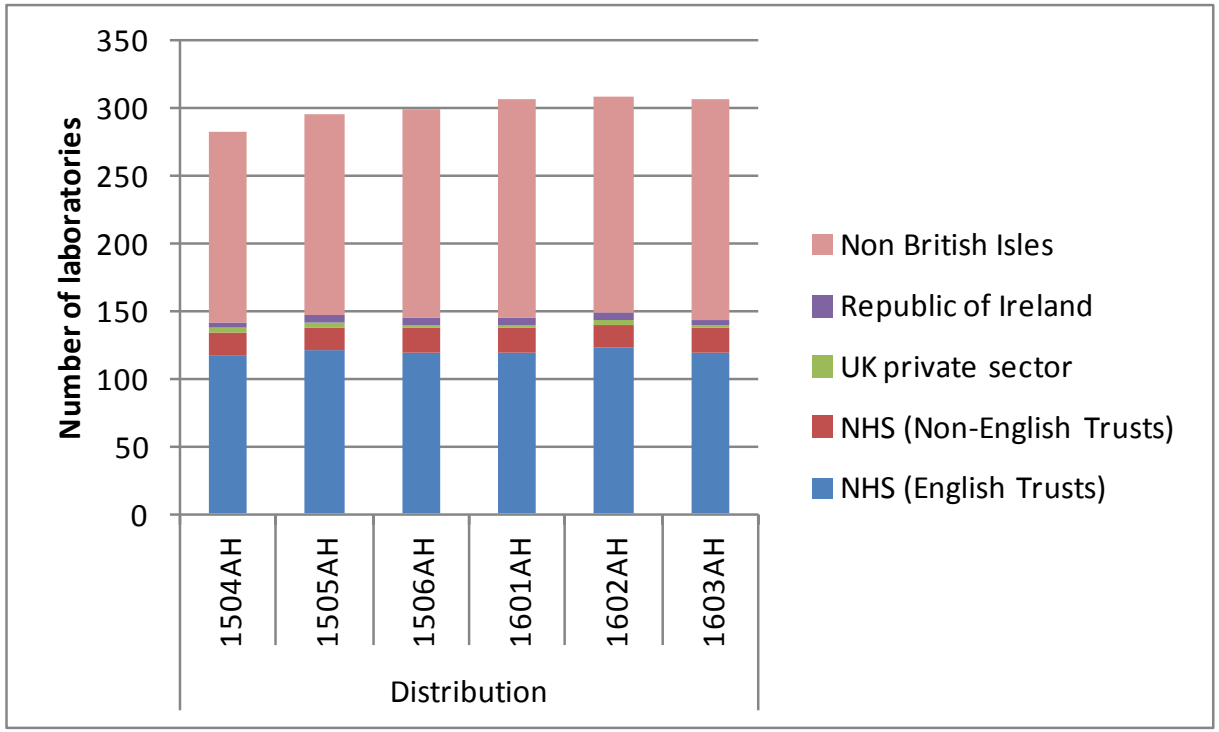


Figure 6.1. The proportion of fraction identification results by laboratory type in the study, distributions 1504AH to 1603AH.

6.4.2 Specimens included in the study

A range of haemoglobinopathy cases were simulated in the eighteen specimens distributed in the study period, including six sickle cell carriers, one Hb C carrier, several alpha thalassaemia carriers (some coexistent with sickle cell carrier status), six with a raised Hb F% (from 5.0% to 21.3%) and two ‘normal’ specimens with no evidence of an abnormal Hb variant or thalassaemia, as shown in Table 6.3, with the essential and expected Hb fractions for each specimen.

Specimen	Case summary	Essential fractions	Expected fractions
1504AH1	29 year old Indian female patient tested antenatally. In conjunction with normal red cell indices, screening showed no evidence of an abnormal Hb variant or thalassaemia carrier. Partner screening was not indicated.	Hb A	Hb A, Hb A_2, Hb F
1504AH2	18 year old Greek female patient screened because of a 'low blood count'. Screening showed a raised Hb F% (11.3%) and a thalassaemic blood count, consistent with a diagnosis of probable delta-beta thalassaemia carrier.	Hb A, Hb F	Hb A, Hb A_2, Hb F
1504AH3	25 year old Black African male patient tested pre-operatively. Haemoglobinopathy screening showed him to be a carrier of Hb S (Hb S = 31.%), blood count indices consistent with a probable coexistent alpha thalassaemia carrier	Hb A, Hb S	Hb A, Hb S, Hb A_2, Hb F
1505AH1	16 year old female patient of Nigerian origin, tested as part of the investigation of anaemia. The Hb concentration low with normal red cell indices. Screening showed her to be Hb S carrier of (Hb S = 32%) with possible coexistent alpha thalassaemia carrier and/or iron deficiency.	Hb A, Hb S	Hb A, Hb S, Hb A_2, Hb F
1505AH2	28 year old Northern European female patient tested antenatally. In conjunction with normal red cell indices, haemoglobinopathy screening showed no evidence of an abnormal Hb variant or thalassaemia carrier; the Hb F% was slightly raised (6.3%), consistent with pregnancy. Partner screening was not indicated.	Hb A, Hb F	Hb A, Hb A_2, Hb F
1505AH3	31 year old Italian female patient tested as part of antenatal screening. Haemoglobinopathy screening showed no abnormal Hb variant and a normal Hb A_2%. The Hb F% (14.6%) was consistent with Hereditary Persistence of Fetal Haemoglobin in conjunction with a normal blood count. Partner testing is indicated.	Hb A, Hb F	Hb A, Hb A_2, Hb F
1506AH1	29 year old male patient of Jamaican origin tested pre-operatively. Screening showed him to be a Hb S carrier. The Hb S% (27.3%) with a microcytic blood count, was consistent with coexistent alpha thalassaemia carrier and/or iron deficiency.	Hb A, Hb S	Hb A, Hb S, Hb A_2, Hb F
1506AH2	26 year old Romanian female patient tested as part of antenatal screening. Haemoglobinopathy screening showed no abnormal Hb variant present and no evidence of thalassaemia carrier in conjunction with a normal blood count. The raised Hb F% (5.0%) was consistent with pregnancy. Partner testing was not indicated.	Hb A, Hb F	Hb A, Hb A_2, Hb F
1506AH3	41 year old African male patient tested as part of the investigation of anaemia. Haemoglobinopathy screening showed no abnormal Hb variant present with a normal Hb A_2 and Hb F%. His blood count was consistent with iron deficiency and/or possible alpha thalassaemia carrier.	Hb A	Hb A, Hb A_2, Hb F

Table 6.3 Cases distributed in surveys 1504AH to 1603AH (August 2015 to June 2016) – *continued on next page*

Specimen	Case summary	Essential fractions	Expected fractions
1601AH1	31 year old Nigerian male whose partner is pregnant and a sickle cell carrier. No evidence of a haemoglobin variant or thalassaemia carrier, with a normal blood count.	Hb A	Hb A, Hb A_2, Hb F
1601AH2	34 year old N. European female patient screened antenatally. Screening showed no abnormal haemoglobin variant or thalassaemia, Hb F% was 10.6% and a normal full blood count. Partner testing was indicated.	Hb A, Hb F	Hb A, Hb A_2, Hb F
1601AH3	27 year old Jamaican female undergoing antenatal haemoglobinopathy screening and found to be a sickle cell carrier (Hb S=38.5%), with normal blood count data. Partner testing was indicated.	Hb A, Hb S	Hb A, Hb S, Hb A_2, Hb F
1602AH1	30 year old Chinese man whose partner is pregnant and an alpha zero thalassaemia carrier. Screening showed no abnormal Hb variant, a normal Hb A_2% and red cell indices consistent with alpha thalassaemia carrier. Accurate genetic counselling and testing of the couple was indicated.	Hb A	Hb A, Hb A_2, Hb F
1602AH2	24 year old African female undergoing antenatal haemoglobinopathy screening as part of antenatal testing and found to be a sickle cell carrier (Hb S = 32.4%). The Hb S% and the low red blood cell indices were consistent with coexistent alpha thalassaemia carrier and/or iron deficiency.	Hb A, Hb S	Hb A, Hb S, Hb A_2, Hb F
1602AH3	26 year old Nigerian female undergoing antenatal haemoglobinopathy screening. There was no evidence of an abnormal Hb variant and the blood count indices were normal; the Hb A_2 was normal but the Hb F was 21.3% consistent with Hereditary Persistence of Fetal Hb. Partner testing was indicated.	Hb A, Hb F	Hb A, Hb A_2, Hb F
1603AH1	29 year old Jamaican female undergoing antenatal haemoglobinopathy screening and found to be a sickle cell carrier. The Hb S was low (24.9%) and the red cell indices reduced, both consistent with coexistent alpha thalassaemia carrier. Partner testing was indicated.	Hb A, Hb S	Hb A, Hb S, Hb A_2, Hb F
1603AH2	18 year old Asian male being investigated for a low haemoglobin concentration. Haemoglobinopathy screening was normal, with a normal Hb A_2 and HbF. Further investigation of possible alpha thalassaemia carrier and/or iron deficiency was indicated.	Hb A	Hb A, Hb A_2, Hb F
1603AH3	29 year old Nigerian female undergoing antenatal haemoglobinopathy screening and found to be a carrier of Hb C, with a blood count at the lower end of the normal reference interval and a normal Hb F%. Partner testing was indicated.	Hb A, Hb C	Hb A, Hb C, Hb A_2, Hb F

Table 6.3 *Continued*

6.4.3 Overall analysis of errors made in fraction identification

Table 6.4 shows the number and type of fraction identification error reported for each specimen by laboratory type.

The use of the 'Non-specified fraction' (NS) category varied by laboratory type. Laboratories within the UK and the RoI used the option thirteen times in 2,619 responses (0.5%) and its use corresponded to 'unable to identify' or 'transposition' errors in fraction identification with the exception of two English NHS laboratories who reported a NS fraction for specimen 1603AH3. 1603AH3 was the Hb C carrier and in both cases the laboratories reported a small NS fraction in addition to the expected fractions. Laboratories outside the UK and the RoI used the NS category seventy-seven times in 2,778 responses (2.8%). Again, it was frequently used for responses in the UI category but there were forty instances of a NS fraction being reported in the 'Other' category compared to two instances among the UK and RoI laboratories, for a similar number of responses overall in each group. There were also more ANR errors among Non-British Isles laboratories for a similar number of responses (116/2,778, 4.2%, of Non-British Isles laboratories compared to 8/2,619, 0.3%, of laboratories in the UK and RoI).

Table 6.5 and Figure 6.2 show the average percentage of the different fraction identification errors for each specimen in the study by laboratory type and essential fractions present. The average percentage by specimen type was used to correct for the varying numbers of specimens with the different Hb fraction patterns in the study.

Laboratory type	Specimen number	Essential fractions	Correct responses	Incorrect responses	Type of FID error			
					UI	ANR	Tx	Other
English NHS	1504AH1	A	117	0	0	0	0	0
	1504AH2	AF	117	0	0	0	0	0
	1504AH3	AS	115	2	2	0	0	0
	1505AH1	AS	118	3	2	0	0	1
	1505AH2	AF	120	1	1	0	0	0
	1505AH3	AF	118	3	1	0	0	2
	1506AH1	AS	117	3	2	0	1	0
	1506AH2	AF	119	1	0	1	0	0
	1506AH3	A	118	2	0	1	1	0
	1601AH1	A	120	0	0	0	0	0
	1601AH2	AF	119	1	0	0	1	0
	1601AH3	AS	118	2	0	0	1	1
	1602AH1	A	121	2	1	1	0	0
	1602AH2	AS	118	5	2	1	2	0
	1602AH3	AF	118	5	1	1	2	1
	1603AH1	AS	116	4	2	0	1	1
	1603AH2	A	119	1	1	0	0	0
	1603AH3	AC	109	11	4	0	1	6
	Totals		**2117**	**46**	**19**	**5**	**10**	**12**
Non-English NHS	1504AH1	A	16	1	0	1	0	0
	1504AH2	AF	16	1	0	1	0	0
	1504AH3	AS	15	2	0	1	0	1
	1505AH1	AS	17	0	0	0	0	0
	1505AH2	AF	17	0	0	0	0	0
	1505AH3	AF	16	1	0	0	0	1
	1506AH1	AS	17	0	0	0	0	0
	1506AH2	AF	16	1	0	0	0	1
	1506AH3	A	16	1	0	0	0	1
	1601AH1	A	17	0	0	0	0	0
	1601AH2	AF	17	0	0	0	0	0
	1601AH3	AS	17	0	0	0	0	0
	1602AH1	A	17	0	0	0	0	0
	1602AH2	AS	16	1	0	0	0	1
	1602AH3	AF	17	0	0	0	0	0
	1603AH1	AS	17	0	0	0	0	0
	1603AH2	A	17	0	0	0	0	0
	1603AH3	AC	15	2	1	0	0	1
	Totals		**296**	**10**	**1**	**3**	**0**	**6**

Table 6.4 (part 1). The number of incorrect fraction identification (FID) results returned for each specimen, by laboratory type, for English and Non-English NHS Trust participants.

Laboratory type	Specimen number	Essential fractions	Correct responses	Incorrect responses	Type of FID error			
					UI	ANR	Tx	Other
UK Private	1504AH1	A	3	0	0	0	0	0
	1504AH2	AF	3	0	0	0	0	0
	1504AH3	AS	3	0	0	0	0	0
	1505AH1	AS	3	0	0	0	0	0
	1505AH2	AF	3	0	0	0	0	0
	1505AH3	AF	3	0	0	0	0	0
	1506AH1	AS	3	0	0	0	0	0
	1506AH2	AF	3	0	0	0	0	0
	1506AH3	A	3	0	0	0	0	0
	1601AH1	A	3	0	0	0	0	0
	1601AH2	AF	3	0	0	0	0	0
	1601AH3	AS	3	0	0	0	0	0
	1602AH1	A	3	0	0	0	0	0
	1602AH2	AS	3	0	0	0	0	0
	1602AH3	AF	3	0	0	0	0	0
	1603AH1	AS	3	0	0	0	0	0
	1603AH2	A	3	0	0	0	0	0
	1603AH3	AC	3	0	0	0	0	0
	Totals		**54**	**0**	**0**	**0**	**0**	**0**
Republic of Ireland	1504AH1	A	5	0	0	0	0	0
	1504AH2	AF	5	0	0	0	0	0
	1504AH3	AS	5	0	0	0	0	0
	1505AH1	AS	6	0	0	0	0	0
	1505AH2	AF	6	0	0	0	0	0
	1505AH3	AF	6	0	0	0	0	0
	1506AH1	AS	6	0	0	0	0	0
	1506AH2	AF	6	0	0	0	0	0
	1506AH3	A	6	0	0	0	0	0
	1601AH1	A	5	0	0	0	0	0
	1601AH2	AF	4	1	0	0	0	1
	1601AH3	AS	5	0	0	0	0	0
	1602AH1	A	6	0	0	0	0	0
	1602AH2	AS	5	1	0	0	1	0
	1602AH3	AF	5	1	0	0	1	0
	1603AH1	AS	4	0	0	0	0	0
	1603AH2	A	4	0	0	0	0	0
	1603AH3	AC	4	0	0	0	0	0
	Totals		**18**	**3**	**0**	**0**	**2**	**1**

Table 6.4 (part 2). The number of incorrect fraction identification (FID) results returned for each specimen, by laboratory type, for UK Private laboratory and Republic of Ireland participants.

Laboratory type	Specimen number	Essential fractions	Correct responses	Incorrect responses	Type of FID error			
					UI	ANR	Tx	Other
Non-British Isles	1504AH1	A	124	17	1	10	0	6
	1504AH2	AF	123	18	3	8	0	7
	1504AH3	AS	124	17	3	8	0	6
	1505AH1	AS	134	15	1	7	1	6
	1505AH2	AF	131	18	1	7	0	10
	1505AH3	AF	131	18	1	7	1	9
	1506AH1	AS	136	17	2	7	0	8
	1506AH2	AF	133	20	1	8	0	11
	1506AH3	A	138	15	1	7	0	7
	1601AH1	A	149	12	1	7	0	4
	1601AH2	AF	146	15	1	7	0	7
	1601AH3	AS	147	14	2	6	0	6
	1602AH1	A	144	15	1	6	1	7
	1602AH2	AS	139	20	4	4	5	7
	1602AH3	AF	140	19	2	5	4	8
	1603AH1	AS	146	17	3	7	0	7
	1603AH2	A	152	11	1	4	2	4
	1603AH3	AC	119	44	16	1	2	25
	Totals		**2456**	**322**	**45**	**116**	**16**	**145**

Table 6.4 (part 3). The number of incorrect fraction identification (FID) results returned for each specimen, by laboratory type, for Non-British Isles participants.[4]

[4] Note to table 6.4: The errors were categorised into unable to identify (UI), Hb A not reported (ANR), transposition of specimens or results (Tx) and other errors (Other)

Essential fractions	Lab Type	Average no. of results	Average % of FID errors			
			UI	ANR	Tx	Other
Hb A	Eng NHS	120	0.3%	0.3%	0.2%	0.0%
	Non-Eng NHS	17	0.0%	1.2%	0.0%	1.2%
	UK Private	3	0.0%	0.0%	0.0%	0.0%
	RoI	5.2	0.0%	0.0%	0.0%	0.0%
	Non British Isles	155.4	0.6%	4.4%	0.4%	3.6%
Hb A + Hb F	Eng NHS	120.3	0.4%	0.3%	0.4%	0.4%
	Non-Eng NHS	17	0.0%	1.0%	0.0%	2.0%
	UK Private	3	0.0%	0.0%	0.0%	0.0%
	RoI	5.7	0.0%	0.0%	2.9%	2.9%
	Non British Isles	152	1.0%	4.6%	0.5%	5.7%
Hb A + Hb S	Eng NHS	120	1.4%	0.1%	0.7%	0.4%
	Non-Eng NHS	17	0.0%	1.0%	0.0%	2.0%
	UK Private	3	0.0%	0.0%	0.0%	0.0%
	RoI	5.3	0.0%	0.0%	3.1%	0.0%
	Non British Isles	154.3	1.6%	4.2%	0.6%	4.3%
Hb A + Hb C	Eng NHS	120	3.3%	0.0%	0.8%	5.0%
	Non-Eng NHS	17	5.9%	0.0%	0.0%	5.9%
	UK Private	3	0.0%	0.0%	0.0%	0.0%
	RoI	4	0.0%	0.0%	0.0%	0.0%
	Non British Isles	163	9.8%	0.6%	1.2%	15.3%

Table 6.5. Average percentage of fraction identification errors by laboratory type and Haemoglobin fraction pattern

Legend to Table 6.5 and Figure 6.2

UI: Unable to identify

ANR: Hb A not reported

Tx: Transposition of specimens or results

Other: All other fraction identification errors

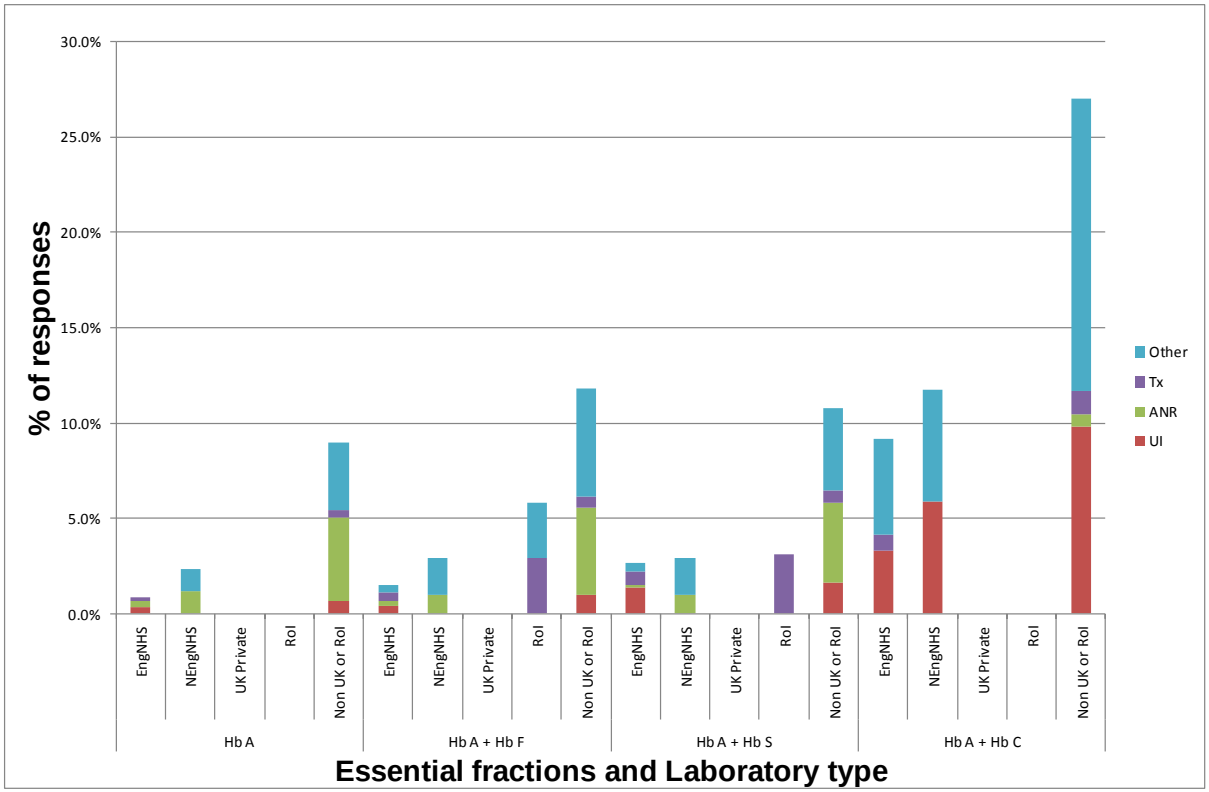


Figure 6.2. The average percentage of fraction identification errors, shown by laboratory type and pattern of essential Haemoglobin fractions.

Figure 6.2 shows:

1. There were no errors made by UK Private laboratories but this group was very small (three). Note also that the number of RoI laboratories was small (between four and six per specimen) and therefore a single error results in a correspondingly large percentage change.
2. For all Hb fraction patterns, Non-British Isles laboratories made the greatest number of errors. In particular, Non-British Isles laboratories made a greater number of errors in the ANR and the Other categories.
3. For the normal (Hb A), raised Hb F (Hb A + Hb F) and sickle carrier (Hb A + Hb S) specimens, the percentage of errors made by laboratories in the UK and RoI was 3% or less in any category.
4. For the Hb C carrier (Hb A + Hb C) specimen, all laboratories made more errors in the UI and Other categories than for the other specimens, perhaps reflecting that this specimen was a greater diagnostic challenge and less familiar to participants as an EQA specimen. (Note that UK NEQAS Haematology distributes a maximum of one Hb C carrier specimen per year compared to between four and six each of the raised Hb F and sickle carrier specimens).

6.4.4 The use of confirmatory methods for Haemoglobin variant identification

Participants are asked to mark the methods used for fraction identification using ‘tick-boxes’ at data entry from the following options: electrophoresis at acid or alkaline pH, HPLC, IEF, CZE, sickle solubility testing and mass spectrometry. The choice of methods is made by the participant depending upon the content of the specimen; there is no restriction on the number of methods that can be used. The

numbers of laboratories using more than one method for fraction identification are shown in Table 6.6 by laboratory type. The numbers of laboratories in the RoI and the UK Private group were small overall: in the RoI group there was little difference by specimen type suggesting that the laboratories tested the samples by all the methods available regardless of content. The UK Private group did not use a confirmatory method for any specimen, suggesting that the three laboratories only had a single method available.

The major difference in the reported use of confirmatory methods was between UK NHS Trusts (both English and Non-English) and the Non-British Isles laboratories:

- For specimens with a Hb variant present, the median percentage of laboratories using more than one method for testing was 69.2% (range 56.7 to 72.5%) for English NHS Trusts and 76.5% (range 70.6 to 82.4%) for Non-English NHS Trusts. By contrast, for Non-British Isles laboratories the median percentage using more than one method was 18.9% (range 16.8 to 19.9%).
- For samples with no Hb variant present and a normal Hb F concentration, the median percentage of laboratories testing by more than one method was 8.3% (range 4.3 to 10.8%) for English NHS Trusts and 35.3% (range 35.3 to 41.2%) for Non-English NHS Trusts, suggesting that a number of Non-English Trust laboratories are testing ‘normal’ specimens by more methods than necessary. The median percentage of Non-British Isles laboratories testing ‘normal’ specimens by more than one method was 10.8% (range 9.9 to 12.7%).
- Where the specimen contained a raised Hb F, the median percentage of laboratories testing by more than one method was slightly higher than for

‘normal’ samples at 17.9% for English NHS Trust laboratories (range 9.1 to 27.0%) and 44.2% (range 35.3 to 64.7%) for Non-English NHS Trusts. For Non-British Isles laboratories the median was 12.2% (range 10.5 to 13.9%).

Specimen	Essential Fractions	Numbers using a confirmatory method by Laboratory type				
		NHS (English Trusts)	NHS (Non-English Trusts)	UK private	RoI	Non British Isles
1504AH1	Hb A	5 4.3%	6 35.3%	0 0.0%	4 66.7%	14 10.0%
1504AH2	Hb A+Hb F	22 18.8%	8 47.1%	0 0.0%	4 66.7%	15 10.7%
1504AH3	Hb A+Hb S	77 65.8%	13 76.5%	0 0.0%	4 66.7%	27 19.3%
1505AH1	Hb A+Hb S	87 71.9%	14 82.4%	0 0.0%	4 66.7%	28 18.9%
1505AH2	Hb A+Hb F	11 9.1%	7 41.2%	0 0.0%	4 66.7%	18 12.2%
1505AH3	Hb A+Hb F	22 18.2%	8 47.1%	0 0.0%	4 66.7%	19 12.8%
1506AH1	Hb A+Hb S	83 69.2%	13 76.5%	0 0.0%	4 66.7%	30 19.9%
1506AH2	Hb A+Hb F	11 9.2%	6 35.3%	0 0.0%	4 66.7%	21 13.9%
1506AH3	Hb A	8 6.7%	6 35.3%	0 0.0%	4 66.7%	19 12.7%
1601AH1	Hb A	10 8.3%	6 35.3%	0 0.0%	4 80.0%	16 9.9%
1601AH2	Hb A+Hb F	21 17.5%	6 35.3%	0 0.0%	4 80.0%	17 10.5%
1601AH3	Hb A+Hb S	78 65.0%	13 76.5%	0 0.0%	4 80.0%	29 17.9%
1602AH1	Hb A	11 8.9%	7 41.2%	0 0.0%	4 66.7%	17 10.8%
1602AH2	Hb A+Hb S	85 69.7%	13 76.5%	0 0.0%	4 66.7%	27 17.1%
1602AH3	Hb A+Hb F	33 27.0%	11 64.7%	0 0.0%	4 66.7%	19 12.1%
1603AH1	Hb A Hb S	87 72.5%	13 76.5%	0 0.0%	3 60.0%	31 19.3%
1603AH2	Hb A	13 10.8%	6 35.3%	0 0.0%	3 60.0%	19 11.8%
1603AH2	Hb A+Hb C	68 56.7%	12 70.6%	0 0.0%	3 60.0%	27 16.8%

Table 6.6 Numbers of laboratories reporting a confirmatory method for fraction identification by laboratory type

6.4.5 Overall analysis of interpretive comments reported

The coded comments returned for each of the six specimens in the study are shown graphically in Figures 6.3 to 6.8 and summarised in Tables 6.7 to 6.12. Each table shows the total number of returns from all participants by comment code, ordered by frequency of use, and the returns for the English NHS Trust and the Non-British Isles laboratories. The five most frequently used comment codes in the English NHS Trust and the Non-British Isles groups are ranked from one (most used) to five. Essential interpretive codes are shown in green and acceptable codes in orange. The number of different coded comments used by the two groups is shown for each specimen together with the average number of codes returned per laboratory.

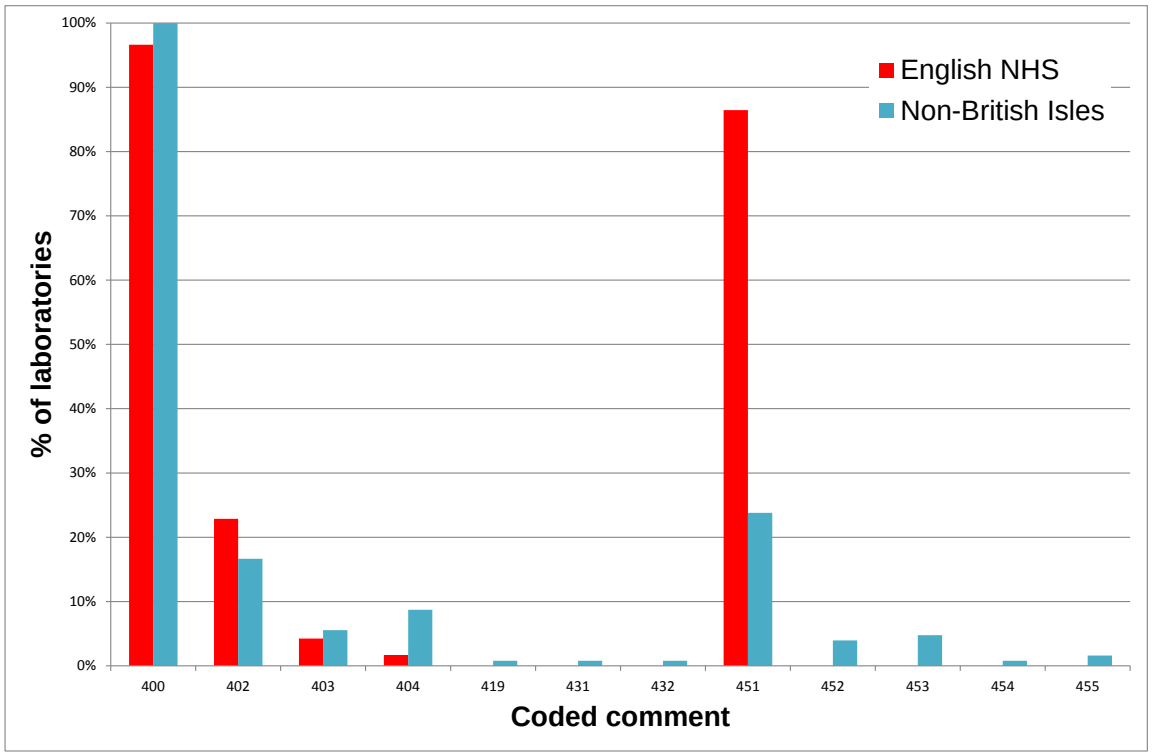


Figure 6.3. Interpretive comment codes returned for specimen 1504AH1 by English NHS Trust laboratories and Non-British Isles laboratories. The case simulated an antenatal patient with no evidence of an abnormal Haemoglobin variant or thalassaemia and red cell indices.

Coded comment	Total all labs	English NHS		Non-British Isles	
		N	Rank	N	Rank
400	262	114	1	126	1
451	146	102	2	30	2
402	53	27	3	21	3
403	15	5	4	7	5
404	14	2	5	11	4
453	6	0		6	
452	6	0		5	
455	3	0		2	
431	2	0		1	
454	2	0		1	
419	1	0		1	
432	1	0		1	
No. codes used	**12**	**5**		**12**	
Avge No. codes/lab	**1.9**	**2.2**		**1.7**	

Table 6.7. Coded comments returned for specimen 1504AH1 by laboratory type.

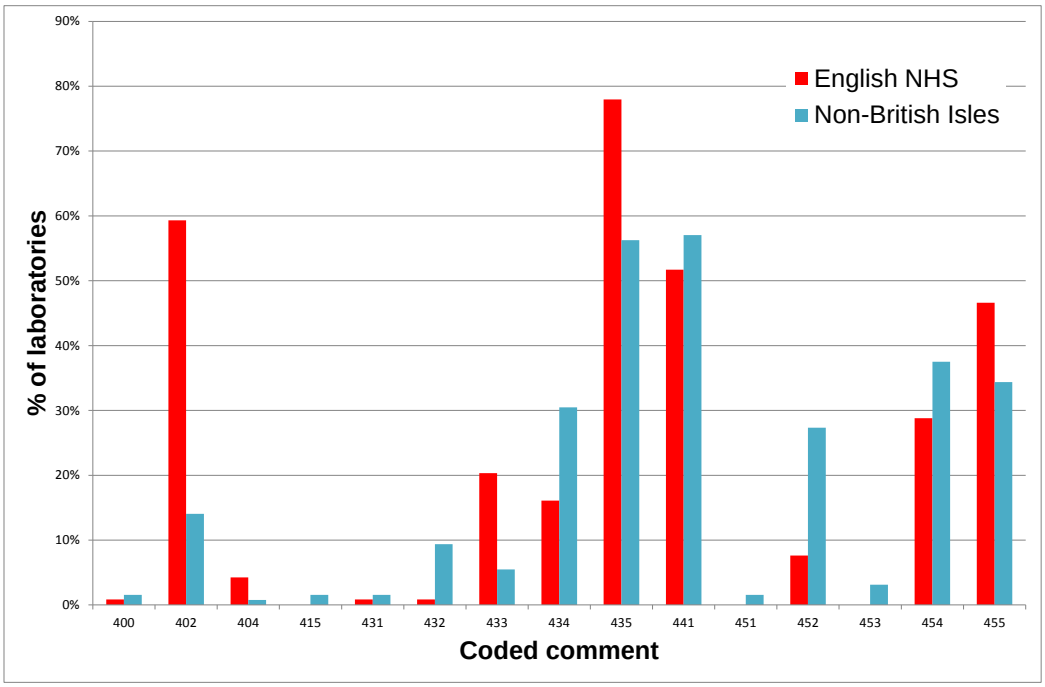


Figure 6.4. Interpretive comment codes returned for specimen 1504AH2 by English NHS Trust laboratories and Non-British Isles laboratories. The case simulated a young female delta-beta thalassaemia carrier (non-antenatal patient).

Coded comment	Total all labs	English NHS		Non-British Isles	
		N	Rank	N	Rank
435	183	92	1	72	2
441	151	61	2	73	1
455	112	55	4	44	4
402	94	70	3	18	
454	92	34	5	48	3
434	65	19		39	5
452	49	9		35	
433	36	24		7	
432	14	1		12	
404	6	5		1	
453	4	0		4	
400	3	1		2	
431	3	1		2	
415	2	0		2	
451	2	0		2	
No. codes used	**15**	**12**		**15**	
Avge No. codes/lab	**3.1**	**3.4**		**2.9**	

Table 6.8. Coded comments returned for specimen 1504AH2 by laboratory type.

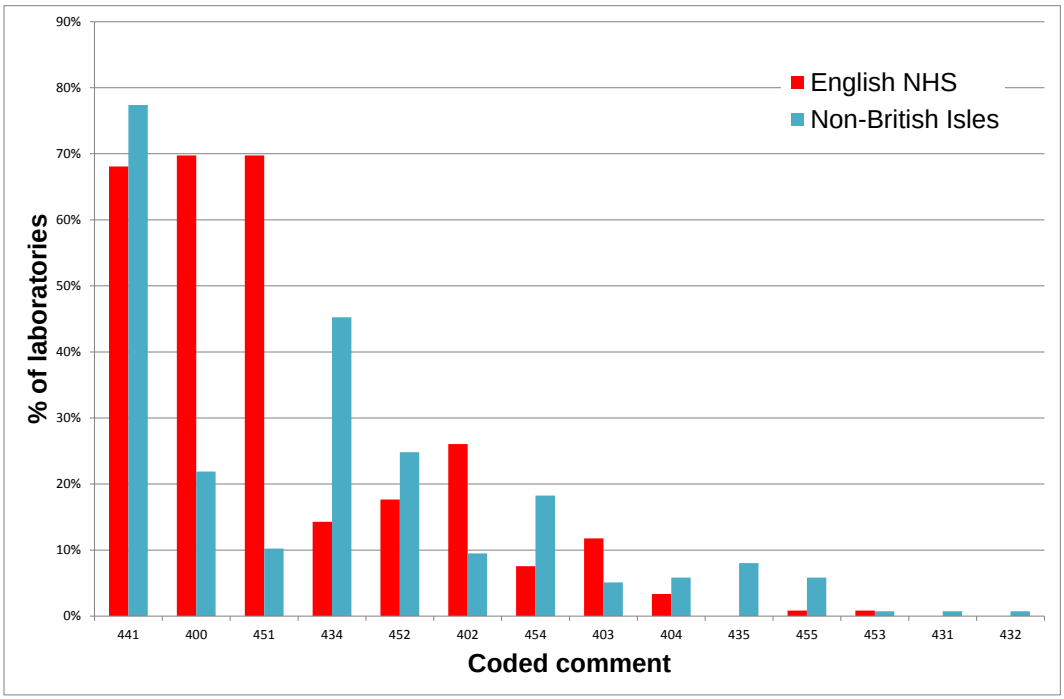


Figure 6.5. Interpretive comment codes returned for specimen 1505AH2 by English NHS Trust laboratories and Non-British Isles laboratories. The case simulated an antenatal patient with normal red cell indices and a slightly raised Haemoglobin F% (6.3%) consistent with pregnancy.

Coded comment	Total all labs	English NHS		Non-British Isles	
		N	Rank	N	Rank
441	212	81	3	106	1
400	127	83	1	30	4
451	99	83	1	14	
434	84	17		62	2
452	57	21	5	34	3
402	44	31	4	13	
454	39	9		25	5
403	20	13		7	
404	12	4		8	
435	12	0		11	
455	12	1		8	
453	2	1		1	
431	1	0		1	
432	1	0		1	
No. codes used	**14**	**11**		**14**	
Avge No. codes/lab	**2.7**	**3.1**		**2.4**	

Table 6.9. Coded comments returned for specimen 1505AH2 by laboratory type.

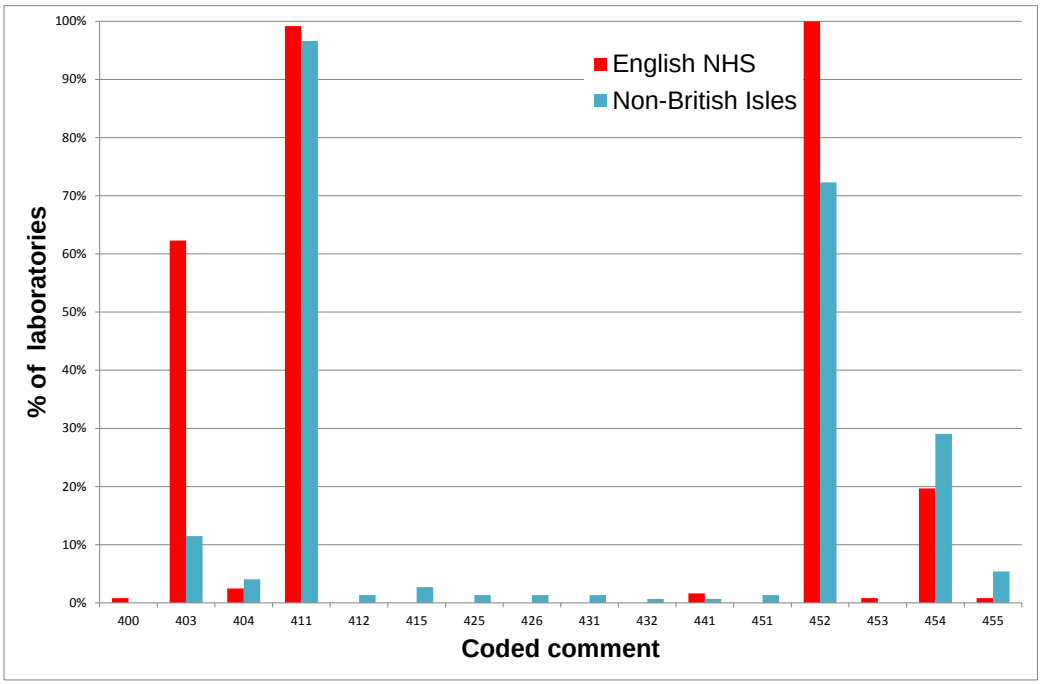


Figure 6.6. Interpretive comment codes returned for specimen 1601AH3 by English NHS Trust laboratories and Non-British Isles laboratories. The case simulated an antenatal sickle cell carrier.

Coded comment	Total all labs	English NHS		Non-British Isles	
		N	Rank	N	Rank
411	294	121	2	143	1
452	250	122	1	107	2
403	105	76	3	17	4
454	67	24	4	43	3
404	9	3	5	6	
455	9	1		8	5
415	4	0		4	
441	3	2		1	
412	2	0		2	
425	2	0		2	
426	2	0		2	
431	2	0		2	
451	2	0		2	
400	1	1		0	
432	1	0		1	
453	1	1		0	
No. codes used	**16**	**9**		**14**	
Avge No. codes/lab	**2.6**	**3**		**2.3**	

Table 6.10. Coded comments returned for specimen 1601AH3 by laboratory type.

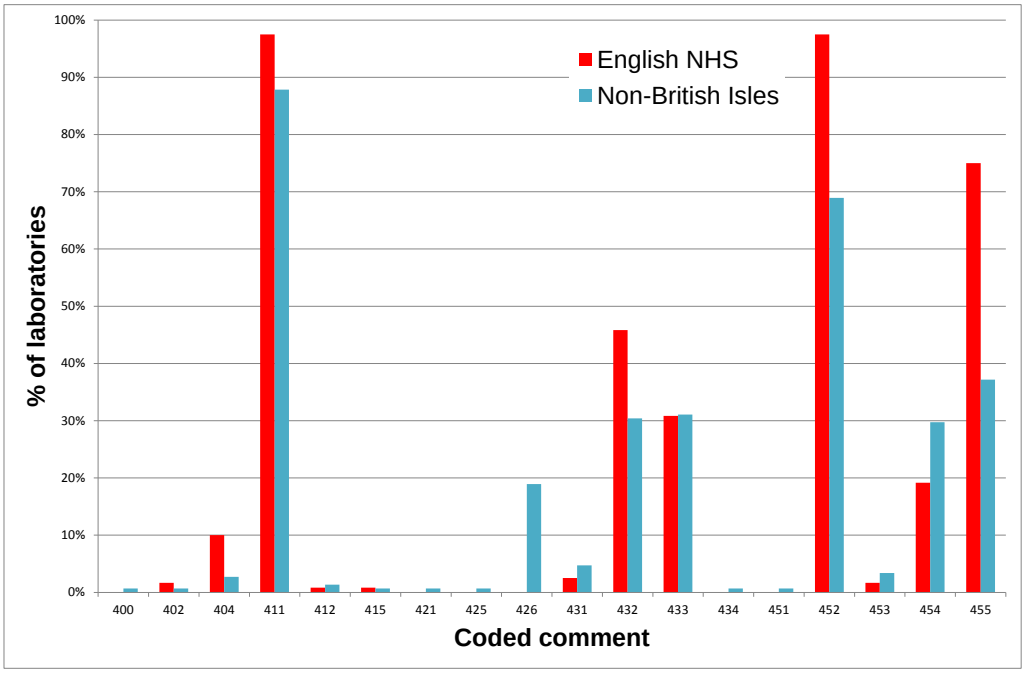


Figure 6.7. Interpretive comment codes returned for specimen 1603AH1 by English NHS Trust laboratories and Non-British Isles laboratories. The case simulated an antenatal sickle cell carrier and probable coexistent alpha thalassaemia trait.

Coded comment	Total all labs	English NHS		Non-British Isles	
		N	Rank	N	Rank
411	273	117	1	130	1
452	219	117	1	102	2
455	145	90	3	55	3
432	111	55	4	45	5
433	92	37	5	46	4
454	69	23		44	
426	28	0		28	
404	19	12		4	
431	11	3		7	
453	7	2		5	
402	3	2		1	
412	3	1		2	
415	2	1		1	
434	2	0		1	
400	1	0		1	
421	1	0		1	
425	1	0		1	
451	1	0		1	
No. codes used	**18**	**12**		**18**	
Avge No. codes/lab	**3.6**	**4**		**3.2**	

Table 6.11. Coded comments returned for specimen 1603AH1 by laboratory type.

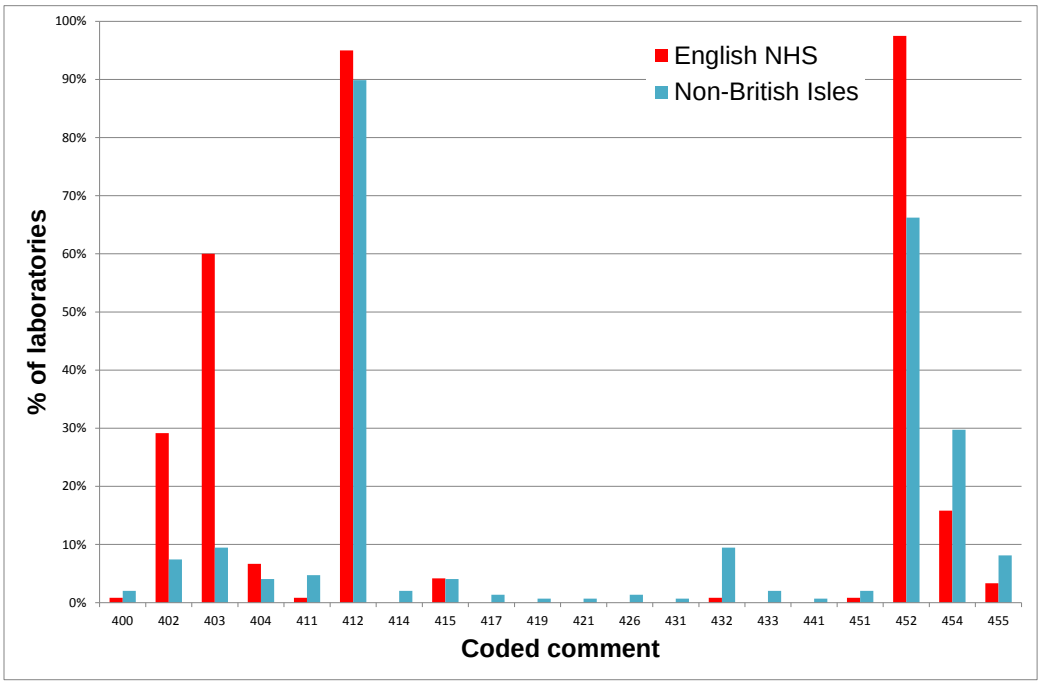


Figure 6.8. Interpretive comment codes for 1603AH3, English NHS Trust and Non-British Isles laboratories. The case was an antenatal patient with Hb C carrier.

Coded comments	Total all labs	English NHS		Non-British Isles	
		N	Rank	N	Rank
412	271	114	2	133	1
452	215	117	1	98	2
403	93	72	3	14	4
454	69	19	5	44	3
402	46	35	4	11	
455	16	4		12	
404	15	8		6	
432	15	1		14	4
415	12	5		6	
411	8	1		7	
400	4	1		3	
451	4	1		3	
414	3	0		3	
433	3	0		3	
417	2	0		2	
426	2	0		2	
419	1	0		1	
421	1	0		1	
431	1	0		1	
441	1	0		1	
No. codes used	**20**	**12**		**20**	
Avge No. codes/lab	**2.9**	**3.3**		**2.5**	

Table 6.12. Coded comments returned for specimen 1603AH3 by laboratory type.

With the exception of specimen 1505AH2, the essential comments were the top-ranked comments by both the English NHS and the Non-British Isles groups and more than 80% of the total comments returned were in consensus with the essential and acceptable comments (see Table 6.13). The percentage agreement varied by laboratory type and specimen: the consensus with the essential and acceptable comments for English NHS Trusts was 100% for the normal specimen (1504AH1), 98.0% to 98.9% for the specimens with Hb variants (1601AH3, 1603AH1 and 1603AH3) and 85.8 and 86.3% for the specimens with raised Hb F (1504AH2 and 1505AH2). The consensus of the Non-British Isles laboratories with the essential and acceptable comments was lower: 92.0% for the normal specimen and 88.8 to 95.3% for specimens with an abnormal Hb variant. The greatest difference was seen in the comments returned by the Non-British Isles laboratories for the raised Hb F specimens: the consensus with the essential and acceptable comments was 80.3% for the delta-beta thalassemia carrier specimen (1504AH2) but only 58.6% for specimen with a slightly raised Hb F (1505AH2).

The consensus agreement with the target values is important as less than 80% consensus may indicate that the specimen will lack credibility for EQA performance assessment and should be reported for educational purposes only. UK NEQAS Haematology experience is that performance assessment of a specimen with a less than 80% consensus will result in an unacceptably high number of objections and complaints from participants.

Specimen	English NHS Trusts			Non-British Isles		
	Total comments	Ess/Acc comments	%	Total comments	Ess/Acc comments	%
1504AH1	250	250	100.0%	212	195	92.0%
1504AH2	372	321	86.3%	361	290	80.3%
1505AH2	345	296	85.8%	321	188	58.6%
1601AH3	351	347	98.9%	340	324	95.3%
1603AH1	460	451	98.0%	475	426	89.7%
1603AH3	378	374	98.9%	365	324	88.8%

Table 6.13. Consensus of coded comments returned with the essential and acceptable comments, by laboratory type.

The number of coded comments reported per laboratory is indicative of the complexity of the case: an average of two comments suggests a simple case with one interpretive comment and one giving recommendations for action, an average of three an intermediate complexity case requiring two or more interpretive comments and four a complex case requiring three or more interpretive comments and / or more than one recommendation for action (e.g. partner testing and referral to a consultant haematologist). The English NHS Trust laboratories reported more codes per specimen in their answers (mean 3.2, range 2.2 to 4.0) than the Non-British Isles group (mean 2.5, range 1.7 to 3.2). In contrast, the Non-British Isles group used a wider range of codes for each specimen (range twelve to twenty) compared to the English NHS Trusts (range five to twelve). Using the Mann-Whitney U test the difference in the numbers of comments/laboratory and the total number of codes used was statistically significant ($p < 0.001$) by laboratory type.

6.4.6 The combination of interpretive comments returned by individual laboratories

The correct combination of comments is essential to report an unambiguous diagnostic interpretation to the clinician. The overall summary of the codes returned by all participants can hide the fact that contradictory codes have been reported in a combination that could prove confusing in the case of a real patient. The numbers of correct, partially correct and incorrect combinations of interpretive comments by laboratory type are summarised in Table 6.14.

As with the overall analysis of the codes returned, there was a difference by both laboratory type and specimen. English NHS laboratories returned fewer incorrect interpretations (70/717, 9.6% of total returns) and recommendations for action (17/647, 2.6% of eligible returns) than the Non-British Isles laboratories (209/835, 25.0% and 120/626, 19.2% respectively). The majority of incorrect interpretations were returned for the two specimens with raised Hb F% (1504AH2 and 1505AH2): 75.7% (53/70) of incorrect interpretations returned by English NHS Trusts and 58.4% (122/209) by Non-British Isles laboratories. The Non-British Isles group also returned a high number of incorrect interpretations for specimens 1603AH1 (sickle cell carrier with probable coexistent alpha thalassaemia trait; 38/209 or 18.2% incorrect) and 1603AH3 (Hb C carrier; 33/209 or 15.8% incorrect).

		Recommendations for action									
		English NHS Trusts					Non-British Isles				
Specimen	**Interpretation**	**N**	**Correct**	**Partial**	**Incorrect**	**N/A**	**N**	**Correct**	**Partial**	**Incorrect**	**N/A**
1504AH1	Correct	116	101	15	0	-	122	30	82	10	-
	Partially correct	2	0	2	0	-	1	0	0	1	-
	Incorrect	0	-	-	-	0	3	-	-	-	3
1504AH2	Correct	75	75	0	0	-	55	55	0	0	-
	Partially correct	8	8	0	0	-	20	20	0	0	-
	Incorrect	35	-	-	-	35	53	-	-	-	53
1505AH2	Correct	50	38	5	7	-	12	1	9	2	-
	Partially correct	51	40	6	5	-	56	9	34	13	-
	Incorrect	18	-	-	-	18	69	-	-	-	69
1601AH3	Correct	118	118	0	0	-	135	102	0	33	-
	Partially correct	0	0	0	0	-	0	0	0	0	-
	Incorrect	4	-	-	-	4	13	-	-	-	13
1603AH1	Correct	55	55	0	0	-	44	39	0	5	-
	Partially correct	60	57	0	3	-	66	46	0	20	-
	Incorrect	5	-	-	-	5	38	-	-	-	38
1603AH3	Correct	112	110	0	2	-	115	79	0	36	-
	Partially correct	0	0	0	0	-	0	0	0	0	-
	Incorrect	8	-	-	-	8	33	-	-	-	33
	TOTAL	**717**	**602**	**28**	**17**	**70**	**835**	**381**	**125**	**120**	**209**

Table 6.14. Correct and incorrect interpretations and recommendations for action returned by English NHS Trusts and Non-British Isles laboratories by specimen distributed. Incorrect interpretations were not assessed for the correctness of the recommendations for action (N/A = Not Applicable)

Looking at each individual specimen, the main observations were:

6.4.6.1 **Specimen 1504AH1** simulated an antenatal patient with no evidence of an abnormal Hb variant or thalassaemia. The specimen was straightforward, with an average of 1.9 coded comments returned per participating laboratory. There were no incorrect interpretations or recommendations returned by English NHS Trusts. The Non-British Isles group of participants returned three incorrect interpretations: a diagnosis of beta thalassaemia carrier (reporting a Hb A_2 result of 3.6%), an unidentified fraction and possible alpha thalassaemia carrier. The incorrect recommendations for action by Non-British Isles laboratories were recommending partner testing in a normal ‘patient’. The major difference in the coded comments returned by English NHS Trusts and Non-British Isles laboratories was the greater use of the coded comment 451 (Partner testing not indicated) by English NHS Trusts.

6.4.6.2 **Specimen 1504AH2** simulated a patient with a ‘low blood count', found to be a carrier of delta-beta thalassaemia. The average number of coded comments used per laboratory was 3.1. The case was a non-antenatal ‘patient’ and there is no NHS STP guidance on the correct recommendations for action; therefore, all recommendations were classified as 'correct'. Given that the ‘patient’ was a young woman, counselling would be advisable but partner testing may be irrelevant unless she is pregnant or planning a pregnancy. Thirty five (35/118, 29.7%) English NHS and 53/128 (41.4%) Non-British Isles laboratories returned an incorrect

interpretation for this case. Eighteen NHS English Trusts and thirty-six Non-British Isles laboratories did not give the essential interpretation of delta-beta thalassaemia carrier. Other incorrect returns by both laboratory types related to the use of coded comments for HPFH or alpha thalassaemia carrier in addition to the essential comments, the use of no evidence of thalassaemia (with or without the coded comment for delta-beta thalassaemia), or the use of coded comment 404 (no evidence of beta thalassaemia) without the use of the coded comment for delta-beta thalassaemia. Occasional Non-British Isles laboratories also used coded comments for the presence of an abnormal Hb variant or beta thalassaemia carrier. None of the laboratories giving the diagnosis of beta thalassaemia carrier reported a Hb A_2 greater than 3.5%. The major difference between the two laboratory types was the greater use of the coded comment 402 (No evidence of sickle haemoglobin) by the English NHS Trusts group.

6.4.6.3 **Specimen 1505AH2** simulated an antenatal patient with normal red cell indices and a slightly raised Hb F (6.3%), consistent with pregnancy. The average number of coded comments used per laboratory was 2.7. Eighteen (18/119, 15.1%) English NHS Trusts reported an incorrect interpretation for the case compared to 69/137 (50.4%) of Non-British Isles laboratories. The incorrect interpretations mainly related to the omission of one or more of the essential diagnostic comments (eighteen English NHS Trusts and fifty-eight Non-British Isles laboratories), the use of coded comments

for HPFH in addition to the essential comments (eleven Non-British Isles laboratories) and occasional laboratories using coded comments for alpha thalassaemia carrier. One Non-British Isles laboratory gave the interpretation of beta thalassaemia trait carrier, although reported a Hb A_2 result of 2.1%. Only two laboratories did not give an interpretation related to a raised Hb F: one reported the consensus Hb F%, the other did not measure Hb F% but noted it as 'normal' in a free text comment. Twelve (12/119, 10.1%) English NHS Trusts gave an incorrect recommendation for action: the recommendations were for partner testing or referral to a consultant haematologist, which is not recommended by the NHS STP unless the Hb F is greater than 10.0%. This case showed the greatest divergence between the two laboratory types in the recommendations given, with a greater use of the coded comment 451 (Partner testing is not indicated) by the English NHS Trust laboratories.

6.4.6.4 **Specimen 1601AH3** simulated an antenatal patient and a sickle cell carrier. The average number of coded comments returned per laboratory was 3.0. Four (4/122, 3.3%) English NHS Trusts reported an incorrect interpretation: all reported the case as a sickle cell carrier but three gave additional out-of-consensus comments of No evidence of a Hb variant or thalassaemia and/or Raised Hb F, reporting Hb F concentrations of 1.0 to 1.5%. The remaining laboratory did not report the case as a sickle cell carrier as they would have referred the specimen to another laboratory for further

testing had this been a patient's sample, i.e. they were 'unable to identify'. Thirteen (13/148, 8.8%) Non-British Isles laboratories returned an incorrect interpretation: five did not report the case as a sickle cell carrier, of which three reported a Hb variant carrier and two homozygous sickle cell disease. The remainder reported a sickle cell carrier but in addition: one reported Hb C carrier (but in fact were unable to identify because the analyser used did not distinguish between Hb S and Hb C), two reported an unidentified or NS fraction but again were unable to identify, two reported beta thalassaemia carrier (reporting Hb A_2 results of 3.6% and 4.1%), one sickle-beta thalassaemia, one raised Hb F (Hb F of 1.1%) and one alpha thalassaemia carrier. These interpretations either alone or in combination with sickle cell carrier could result in confusion if this were a patient's sample. Thirty-three Non-British Isles laboratories gave an incorrect recommendation for action: two reported partner screening as not indicated and the remainder made no recommendation on partner testing. The main difference between the laboratory types in terms of the coded comments used was a greater use of the coded comment 403 (No evidence of thalassaemia) and the near 100% use of the recommendation that partner testing is indicated (coded comment 452) by English NHS Trust laboratories compared to 72% of Non-British Isles laboratories.

6.4.6.5 **Specimen 1603AH1** simulated an antenatal patient and a sickle cell carrier with coexistent alpha thalassaemia trait. The specimen had the highest number of codes used per laboratory (3.6). Five (5/120,

4.2%) English NHS Trusts reported an incorrect interpretation: three incorrectly reported the patient as a beta thalassaemia carrier, either in addition to a sickle cell carrier (two giving Hb A_2 results of 3.8% and 4.0%) or alone (reporting a Hb A_2 of 4.4%), one reported a Hb variant carrier and one Hb C carrier. Both the latter two used an analyser unable to differentiate Hb S and Hb C and were in fact 'unable to identify'. Thirty eight (38/148, 25.7%) Non-British Isles laboratories reported an incorrect interpretation: thirty four reported the case as a sickle-beta thalassaemia carrier or beta thalassaemia carrier, either in addition to or in place of a diagnosis of sickle cell carrier. Two reported a Hb C carrier in addition to sickle cell carrier but were using an analyser unable to differentiate between the two variants, one reported homozygous sickle cell disease (failing to report Hb A in their fraction identification) and one reported no abnormality detected. The use of the coded comments for possible alpha thalassaemia carrier (coded comments 432 and 433) by both groups is of interest. This 'patient' is almost certainly an alpha thalassaemia carrier given the all methods mean Hb S concentration of 24.9%; since she is of Caribbean origin, she would most likely have been a homozygote for alpha plus thalassaemia (Weatherall and Clegg, 2001). Approximately 30% of both laboratory types used coded comment 433 (Possible alpha zero thalassaemia carrier and / or iron deficiency), about the same proportion as used the essential coded comment 432 (Possible alpha plus thalassaemia carrier and / or iron deficiency). For the purposes of this study, code 433 was

considered acceptable as the outcome of both was ‘safe’ and accurate diagnosis of alpha thalassaemia status can only be achieved using molecular techniques; however, this observation suggests that either the coded comments available require clarification or simplification or that additional education of participants may be required. The main difference between the laboratory types with respect to recommendations for action was the advice that partner testing is indicated (coded comment 452), which was given by 98% of English NHS Trusts compared to 79% of Non-British Isles laboratories. The recommendation for testing for iron deficiency (coded comment 455) was used by 75% of English NHS Trusts and 36% of Non-British Isles laboratories.

6.4.6.6 **Specimen 1603AH3** simulated an antenatal patient and a Hb C carrier. The number of coded comments used per laboratory was 3.3. Eight (8/120, 6.7%) English NHS Trusts gave an incorrect interpretation: six did not give the interpretation of Hb C carrier, one gave a normal interpretation, one a diagnosis of sickle cell carrier and four a Hb variant carrier but were in fact probably ‘unable to identify’. One reported a NS fraction in addition to Hb C carrier and one an alpha thalassaemia carrier. All except the laboratory returning a ‘normal’ interpretation advised partner testing. Thirty three (33/148, 22.3%) Non-British Isles laboratories returned an incorrect interpretation making a number of errors: failure to report an interpretation, reporting sickle cell carrier, Hb variant carrier, Hb E or Hb O^{Arab} carrier, alpha thalassaemia carrier, beta thalassaemia

carrier in addition to Hb C carrier or a normal result. The laboratories reporting the specimen as a sickle cell carrier were those that had incorrectly reported Hb S in their fraction identification and two of the three reporting the specimen as ‘normal’ had transposed their specimens or results. Two English NHS laboratories and thirty six Non-British Isles laboratories gave incorrect recommendations for action, failing to indicate partner testing is indicated. The main difference between the comments returned overall by each laboratory type was the greater use of coded comments 402 (No evidence of sickle haemoglobin), 403 (No evidence of thalassaemia) and the recommendation that partner testing is indicated (coded comment 452) by English NHS Trusts.

6.5 Discussion

The interpretation of laboratory results is part of the post-analytical phase of laboratory diagnosis and there is a growing awareness of the need to include pre and post-analytical processes in laboratory quality management, including the quality assurance of interpretation (Plebani, 2009a; Piva, 2009; Plebani, 2010; Sciacovelli, 2011; Hammerling, 2015). Inclusion of interpretation in EQA performance may be as part of an 'end to end' scheme such as the Abnormal Haemoglobins scheme or as a separate interpretive comments scheme (Plebani, 2009b; Vasikaran, 2016), but it is recognised as effective in performance improvement in the diagnostic interpretation of genetic conditions (Hastings, 2008; Hastings, 2010). The provision of a numerical score is highly effective in performance improvement in qualitative EQA (James, 2012). For EQA to be effective, participants must regard the criteria by which their laboratory practice is converted into a numerical performance score as credible, whether for a quantitative, qualitative or interpretive test. Credibility may require demonstration of commutability of survey material or the appropriate grouping of methods in a state-of-the-art quantitative scheme (Sciacovelli, 2010), the comparison of performance to a recognised international guideline or standardised system of nomenclature and/or a scheme that recognises different levels of expertise in the participants. The acceptable performance standards and the design model of the scheme should also be clearly apparent to participants (Jones, 2017b).

This study showed that there are significant differences in the use of interpretive comments by English NHS Trust laboratories in comparison to Non-British Isles laboratories in the numbers of comments used per laboratory and the range of comments reported. The observed differences suggest that the English NHS Trust

laboratories give a more comprehensive interpretation of each case, using more codes per laboratory in each interpretation, e.g. commenting on the Hb variant status and thalassaemia status for each patient, including where one or both of these is 'normal'. The English NHS laboratories use a smaller range of codes overall, suggesting a more accurate and consistent interpretation by this group of laboratories. All laboratory types have access to published guidelines on the laboratory diagnosis of haemoglobinopathies and the superior performance of the English NHS Trust group in interpretation may reflect the proactive oversight of the NHS STP including the guidance provided in the Laboratory Handbook (NHS STP Antenatal Laboratory Handbook, 2017) and face-to-face training days.

For most of the first twenty years of operation of the UK NEQAS Abnormal Haemoglobins scheme, the predominant laboratory type registered was UK NHS-based. In 2001, 215 of 265 (81%) participants were UK NHS laboratories with just 39 (14.7%) non-UK laboratories in the scheme (UK NEQAS Haematology Biennial Report, March 2001). In the study, the proportion of laboratories outside the British Isles registered for Hb fraction identification, quantitation and interpretation had increased to 161 of 307 (52.4%) clinical participants returning results at June 2016. The significant number of laboratories outside the remit of UK oversight bodies requires a review of the scheme to assess how the non-UK laboratories can be incorporated into performance assessment. The non-UK laboratories represent an important source of statistical robustness for the scheme in the face of a shrinking UK participant base; the scheme also has a duty of care to use education to improve the quality of laboratory testing wherever tests are undertaken. This is a particular challenge when dealing with interpretation rather than quantitation, where the scheme is dealing with instruments, reagents and

controls that are manufactured to the same standards on an international basis. Interpretation by contrast relies on the training and expertise of laboratory staff and the professional practice in the country or region in which the laboratory is located. The agreement of both English NHS Trusts and Non-British Isles laboratory types with the top-ranked essential interpretive comments and the greater than 80% consensus of interpretive comments for five of six specimens in the study indicate that there is sufficient agreement between the major groups of participants to make performance assessment credible for both laboratory types.

There are particular differences between the groups however, e.g. the inclusion by Non-British Isles laboratories of the interpretation of Hb S-beta thalassaemia disease and / or beta thalassaemia carrier in sickle cell carrier cases, especially where there are microcytic red cell indices as a result of the commonly-occurring coexistent inheritance of alpha thalassaemia trait. The Hb A_2 measurement by HPLC may be elevated artefactually in the presence of Hb S (Stephens, 2012) but many Non-British Isles participants quote a raised Hb A_2 as evidence for the coexistence of beta thalassaemia. Some of these errors may result from the heuristics of clinical decision-making and improved direct EQA interaction with participants may be a learning tool by which laboratory practice can be modified.

The incidence of the different types of fraction identification errors varied according to the nature of the specimen and geographical location. Detected Tx and UI errors in the study were generally low for all laboratory types, although these errors may be underestimated as a like-for-like Tx will go undetected and an UI error without supporting commentary from the laboratory would be counted as an 'Other' error. The percentage of UI errors was higher for the one Hb C carrier (Hb A + Hb C) specimen distributed and this may reflect unfamiliarity with the

condition, both in the EQA scheme and as a clinical case amongst some participants. The incidence of ANR errors by Non-British Isles laboratories was approximately fourfold higher than by UK laboratories for the frequently distributed specimen types, probably because ANR errors are actively followed up with UK laboratories by UK NEQAS as part of performance management. Interestingly, the incidence of ANR errors for the Hb C carrier case was low for all laboratory types, possibly because a lack of familiarity with the case made all laboratories focus more fully on accurate reporting.

The incidence of 'other' errors was highest amongst Non-British Isles laboratories for all specimens. In some cases, these errors may be clerical in nature. The data capture method for fraction identification in the UK NEQAS Abnormal Haemoglobins scheme was designed when the scheme was first developed. The participant marks the fractions present from a selection of Hb A, Hb S, Hb C, Hb D, Hb E, Hb A_2, Hb F and 'non-specified' (NS) fraction. Where participants are using methods that do not distinguish between Hb C and Hb E, they may select 'Hb C or Hb E', an historic option reflecting the inability to separate these fractions by alkaline electrophoresis, which was the predominant screening method twenty years ago; however, there is no similar allowance for current HPLC methods that do not differentiate between Hb S and Hb C. There is a higher incidence of reporting NS fractions and additional abnormal fractions by non-British Isles laboratories; occasional laboratories regularly mark all the common fractions present, perhaps in the hope that one will be correct, and the use of the NS option may be poorly understood.

The compliance with the recommendation to confirm the identification of an abnormal variant by two methods based on different physicochemical method

principles is apparently poor amongst all laboratory types. Approximately 63% of English NHS Trusts reported using two or more methods for specimens containing Hb S or Hb C compared to approximately 45% of Non-British Isles laboratories. Only five to ten percent reported using more than one method for normal specimens (Hb A, with a normal Hb F%) suggesting that participants are making the effort to mark the methods used on each specimen and not just entering the same options regardless of content. The data entry options for the scheme no longer capture the true workflow in the laboratory since there is no opportunity to enter referral to another laboratory except in written comments, which are difficult to performance assess efficiently from several hundred participants.

An option may be to allow laboratories with only one method available to register as ‘screening’ laboratories and opt out of confirmatory testing for EQA specimens. There is precedent for this in the different levels of registration in the UK NEQAS Haematology Blood Films for Parasitology programme, which allows participants to register for 'screening only' or for 'parasite species identification', depending on their self-declared level of service provision (UK NEQAS Haematology Participants' Manual, 2017). The scheme should consider amendment of the terminology used in data entry to reflect the use of retention time ‘windows’ or ‘zones’ rather than suggesting that a precise fraction identification is possible, capturing data in the ‘what did you see’ followed by ‘how would you report this’ or 'what actions would you take' fashion,. This would allow the removal of the NS fraction option, to be replaced by ‘unable to identify’ or UI with the requirement to give the position and percentage of the peak, an assessment of its significance and what follow up actions are needed. Additional guidance on data entry and tips for successful participation specific to the Abnormal Haemoglobins scheme,

published in the Participants' Manual and linked to the data entry pages or instruction sheets for the scheme could also improve the quality of the data captured.

UK NEQAS Haematology is unable to provide some rarely available but clinically significant Hb variants, e.g. Hb O^{Arab}, which are rare amongst UK blood donors. Even the provision of Hb D and Hb E carriers as EQA specimens is a challenge; none was distributed in the year studied and it is rare for these variants to be distributed more than once every five to ten years, neither did the scheme distribute a beta thalassaemia carrier case during the period due to difficulties in sourcing a donor. Given that good quality management depends upon staff education and training as well as quality control and EQA, the scheme could utilise online case provision to fill these gaps, with background case details with laboratory results from a range of analysers and methods from which the participant can select their own. The cases would not only allow performance assessment in a traditional EQA fashion but would also provide an educational resource for laboratory staff.

A major challenge to the expansion of performance assessment is the additional workload for scheme personnel that it will incur. In this study, there was a total of approximately six fraction identification errors per specimen by UK laboratories in the study, which are currently followed up. There were an additional eighteen per specimen on average by non-British Isles laboratories, i.e. the workload would be quadrupled for fraction identification, without the additional impact of performance assessment of interpretation for all participants. The number of interventions should decrease with time as a result of performance improvement in response to monitoring but the long-term success of the initiative will depend upon the

sustainability of the process. For this reason, it is essential to review the registration options for participants, the data capture mechanism and the introduction of rules-based systems to automate the management of first incidences of unsatisfactory performance. The methods of fraction identification must be linked with interpretive comment assessment to take account of the laboratory's service provision. The scheme must ensure that the data it gathers is as free of ‘blunders’ as possible, so that it provides the best reflection of current laboratory practice.

UK NEQAS has a duty of care to the laboratories and the patients they serve, regardless of where the laboratory is based, and to work to improve laboratory performance through education, especially in the rare anaemias where the relatively small number of participants in an individual country may result in greater reliance on the support provided by international service providers (Vives-Corrons, 2014). The increasing international profile of the scheme means that the scheme must become more proactive in performance assessment in this complex area for all participants. The major barrier is the workload that will be associated with the extension of scoring to the qualitative aspects of the scheme as well as to non-UK laboratories.

Chapter 7: General Discussion

7.1 Discussion

The thalassaemia disorders and other significant haemoglobinopathies are amongst the most common single gene disorders worldwide and predicted to pose an increasing health burden to countries of all income levels (Weatherall, 2010; Piel, 2016). They are regarded as rare in Europe (Regulation (EC) 141/2000) and the majority of haemoglobinopathy and thalassaemia disease patients are born in poorly resourced regions of the world (Weatherall, 2011). According to data from the NHS Sickle and Thalassaemia Screening Programmes there are twenty to thirty babies with possible beta thalassaemia major born in the UK annually (NHS Sickle and Thalassaemia Screening Programme Data Report 2015/16).

Treatment for beta thalassaemia major is palliative, comprising regular blood transfusion and iron chelation, which have significantly improved life expectancy in thalassaemia sufferers, although the regime requires lifelong compliance from the patient and has a major impact on their quality of life (Telfer, 2005; Clarke, 2010). The treatment of transfusion-dependent thalassaemia is also costly: a retrospective study of fifty transfusion-dependent beta thalassaemia major patients estimated the mean total cost of treatment per patient per year in the United States as $128,062 (Sheth, 2017). Recent estimates of the annual cost of treating transfusion-dependent beta thalassaemia major in Turkey, by comparison, were lower, at $14,360; however, the number of transfusion-dependent beta thalassaemia major patients in the country results in annual medical costs of more than 58 million US$ per year (Antmen, 2017). In 1999, the lifetime treatment costs for a patient with beta thalassaemia major in the UK were estimated to be approximately £200,000 (Karnon, 1999). The only potential cure is allogeneic haemopoietic stem cell transplantation (HSCT), of which more than three

thousand have been performed worldwide for beta thalassaemia (Angelucci, 2008; Gaziev, 2008; Angelucci, 2010; Lucarelli, 2012). HSCT requires an HLA-compatible donor but more recent work has suggested the use of genetically corrected autologous HSCT may be a prospect where no HLA-matched donor is available (Acuto, 2014). The majority of beta thalassaemia disease patients are in low and middle-income countries where the cost of palliative support makes HSCT an attractive, although still costly option (Mehta, 2013). A low cost, effective HSCT programme has been described in Pakistan, where costs have been controlled through limiting the highly engineered protective environments associated with the procedure in high income countries without compromising the outcome in the pilot (Faulkner, 2011).

The main means of control for both the thalassaemia disorders and other significant haemoglobinopathies is through screening programmes to detect carriers, with the opportunity this provides for informed reproductive choice (Cao, 2013). In the absence of definitive molecular testing, detection of carriers is presumptive requiring knowledge of the application and interpretation of the appropriate laboratory tests. Detection of beta thalassaemia carriers relies on a small difference between the normal and pathological reference ranges of the minor Hb fraction, Hb A_2. With some mutations, the carrier is described as ‘silent’ because the Hb A_2 concentration is borderline or overlaps the normal reference range (Old, 2012). Consequently, the laboratory detection of beta thalassaemia carriers requires excellent laboratory quality to ensure that testing is both accurate and precise. Quality assurance, comprising internal quality control (IQC) and external quality assessment (EQA), laboratory accreditation and audit, are

objective measures of performance by which the diagnostic testing provider demonstrates the competence of their service.

UK NEQAS Haematology provides one of the most comprehensive programmes of EQA for the laboratory diagnosis of haemoglobinopathies worldwide. Its major comparators in terms of service provision are the College of American Pathologists' (CAP) PT programme and the Royal College of Pathologists in Australasia Quality Assurance Programme (RCPAQAP). UK NEQAS Haematology provides haemoglobinopathy EQA in partnership with other EQA providers in some smaller countries in Europe and the Middle East, who have opted not to develop their own schemes. There are commercial providers, who offer what has been defined in IFCC guidelines as proficiency testing (PT) rather than EQA, to large numbers of participants but these services are more limited in terms of frequency, scope and above all educational content (Maziotta, 2003). A questionnaire undertaken by the author on the provision of EQA in the field of rare anaemias on behalf of the ENERCA project showed a variable provision of haemoglobinopathy EQA in Europe (Vives-Corrons, 2014) in terms of scope and frequency of distribution.

The scope of the UK NEQAS programme had developed considerably in the fifteen years prior to the time of writing, with the addition of emergency sickle screening, newborn screening using both liquid capillary and dried blood spots and molecular testing options, as well as the ‘end-to-end’ testing starting with a clinical scenario and finishing with a suggested diagnostic interpretation. During the same period, there has been a change in the balance of the laboratories registered, from a largely UK base in 2001 (15% of laboratories outside the UK) to just over 50% of laboratories registered for full participation based outside the UK at the time of the

study. The changing profile of the scheme demonstrates the increased significance attached to EQA in laboratory quality management, driven by the implementation of the international accreditation standard ISO 15189:2012, and the importance attached to quality management and EQA participation in professional guidance documents (Ryan, 2010; Vives-Corrons, 2014; Traeger-Synodinos, 2015; Stephens, 2015). A greater number of participating laboratories is a driver in itself for EQA development as it allows more robust statistical analysis and makes the scheme viable where there may be insufficient demand in a single country. However, the increased diversity of participants' geographical location brings challenges, which the scheme must address to ensure it fulfils its duty of care to both the laboratories and the patients they serve, without reduction in the service provided to its core UK participants.

One of the functions of EQA is to provide a 'state-of-the-art' overview of laboratory performance. The field of Hb A_2 measurement has seen the almost complete adoption of automation in the past fifteen years, with high performance liquid chromatography (HPLC) and capillary zone electrophoresis (CZE) replacing earlier manual electrophoretic and chromatographic methods of quantitation. At the time of this study, there were no laboratories registered with UK NEQAS Haematology for Hb A_2 measurement using tandem mass spectrometry (TMS) for Hb A_2 quantitation, although initial enquiries and one request for registration from a non-British Isles laboratory has been received since the study's completion. Improved performance, as measured by a reduction in the coefficient of variation (estimated CV%) from 14.25% to 8.87%, has been seen within the UK NEQAS HPLC users in the period 2001 to 2016. Although the removal of less effective instrument models may have contributed to this, the reduction suggests other factors may also have

had a role, including an increased awareness of the quality of diagnostic testing amongst both instrument users and manufacturers, with the driving force of the NHS Sickle and Thalassaemia Screening Programme and international bodies such as the IFCC and ICSH. Despite the reduction in CV%, the performance of participants still does not meet the published biological variation (BV) standards although at the time of writing the Ricos BV database (See https://www.westgard.com/biodatabase1.htm) only references one BV reference for Hb A_2 (Mosca, 2013), as used for the 'acceptable performance limits' in the evaluation of Hb A_2 data in this study.

This study has shown that despite improvements in performance there remain differences in Hb A_2 measurement between analyser types as well as between users of the same analyser, to the extent that an antenatal patient may receive a different diagnosis depending upon the testing method used by the laboratory. Further work is required on the standardisation of Hb A_2 measurements (Paleari, 2017) as this will provide manufacturers with a product to which they can demonstrate their results are traceable. This study supports the production of new Hb A_2 standards at more than one concentration as the variation between instruments appears to differ according to the Hb A_2 concentration. The only currently available Hb A_2 reference material (the WHO Reference Reagent 89/666) was prepared at two Hb A_2 concentrations (normal and beta thalassaemia range) but only the beta thalassaemia material was supported by the WHO (Hb A_2 concentration 5.3%). This limits its use, although it can be diluted with umbilical cord blood from an infant with no detectable Hb A_2 to check linearity (Stephens, 2015).

The aim of all EQA providers is to use a survey or assay material that is identical to clinical specimens and commutable for all methods, as well as traceable to a reference standard. UK NEQAS Haematology provides a liquid whole blood specimen of human origin for its Abnormal Haemoglobin scheme, smaller in volume than most clinical specimens but otherwise suitable for analysis without reconstitution or further preparation. The study examined whether the difference in anticoagulant and age of the EQA material compared to clinical samples had an impact on Hb A_2 measurement, which would limit the usefulness of the survey material and confound performance assessment. The study provided evidence that UK NEQAS survey material in the Abnormal Haemoglobins is comparable in performance for Hb A_2 measurement to blood taken into ethylenediaminetetraacetic acid (EDTA), the usual anticoagulant for this investigation. Earlier work suggested that specimens for Hb A_2 are stable for up to three weeks if stored at 2 to 8 ºC in EDTA and if taken into other anticoagulants (Tietz, 1990; Shokrani, 2002; Maharini, 2014) but none specifically examined the combined effect of both anticoagulant and storage on Hb A_2 measurement. This part of the study warrants continued work as new Hb A_2 measurement methods come into use. As designed, the EDTA *vs.* CPD anticoagulated sample experiment is not sustainable as a validation method. For future validations it may be possible to undertake parallel testing of a new method or model with the EQA survey material against one with a track record in the EQA scheme, extended over a period of time equivalent to a survey distribution cycle. One consequence of producing a whole blood specimen from a single blood donation for several hundred laboratories is that the specimen volume is limited (approximately 1 mL) and may require analysis in an ‘open’ mode, increasing the opportunity for

specimen transposition errors. The study of fraction identification errors showed a generally low rate of detected transposition errors (from zero to around 1%) for all participants. Transposition is a serious pre-analytical laboratory error; in EQA it is a ‘double-edged’ sword in that it incurs adverse performance penalties but is also an educational opportunity to increase awareness of the impact of pre-analytical errors in screening for genetic disorders. One area that UK NEQAS should consider when detecting a transposition error is to ask the laboratory to confirm that there has been no transposition of the EQA sample and a patient’s sample, which would be a major problem for the scheme in that efforts to improve quality had precipitated an error with a patient’s result.

EQA is often restricted in its geographical reach because of the limitations of the specimens used. A commutable specimen may be desirable but not sufficiently stable or available in sufficient volume for use on an international basis, restricting the effectiveness of EQA in a market where analysers, reagents and calibrators are manufactured and marketed globally (Jones, 2017a). The demonstration that the UK NEQAS Abnormal Haemoglobins survey material is both comparable in performance to clinical samples and stable for sufficient time to allow international comparison of Hb A_2 performance is significant and confirms the scheme’s provision of performance assessment to laboratories internationally. One observation of relevance in the discussion of survey material quality is the greater frequency of reported apparent ‘non-specified’ fractions by Non-British Isles laboratories, seen in the evaluation of fraction identification results in the study. The reporting of NS fractions may indicate a lack of understanding in the purpose of this category and a desire by the participant to ‘cover all options’; however, the scheme should consider whether there is any element of additional deterioration in

the specimen during transit causing additional minor peaks associated with an aging specimen, although deterioration of this sort has not been seen in stability studies undertaken in-house. This could be addressed by asking participants from all geographical locations to return the chromatograms or electropherograms of the EQA specimen, and undertaking a visual assessment of the quality of the separation by location of the participant.

Screening for beta thalassaemia major is not part of the NHS newborn screening programme; however, there is a proposal that a Hb A concentration of less than 1.5% may be indicative of the condition in newborns of greater than thirty-two weeks' gestation (Streetly, 2013). The study assessed the performance of the newborn screening methods available at the time to quantitate Hb A% at this low concentration, using a sample from an adult homozygous delta-beta thalassaemia patient (Hb pattern of Hb F and Hb A_2) spiked with varying amounts of normal umbilical cord blood (Hb F + Hb A). A difference was seen between the HPLC and CZE analysers used for haemoglobinopathy screening in UK newborn screening laboratories. Both methods are designed specifically for newborn screening and marketed as qualitative rather than quantitative. The Bio-Rad vNBS HPLC analyser was capable of measuring Hb A% at levels of 1.0% but the sensitivity of the Sebia Neonat Fast CZE system was not reliable at Hb A concentrations below 2.0%, which might affect quantitation of Hb A at or around 1.5%. Although the number of CZE systems studied was small, the findings were supported by other literature (Wolff, 2012). The first ten years of the NHS Newborn Screening Programme reported the births of 268 ‘Hb F only’ infants in England (approximately twenty to thirty per year), which were possible beta thalassaemia major babies (NHS Sickle and Thalassaemia Screening Programme Data Report,

2015/16). Given that in 2015/16 there were twenty-four ‘thalassaemia affected’ pre-natal diagnosis (PND) fetal results and that pregnancy outcome data shows that 49.7% of pregnancies affected by beta thalassaemia were known to have ended in termination, this suggests that there are still unanticipated beta thalassaemia major births in the UK. There is no immediate medical intervention for beta thalassaemia major in the newborn, as there is in sickle cell disease, but early identification could prepare parents for the impact of a thalassaemic child and enable early referral to a consultant haematologist, prior to the development of clinical symptoms (Hoppe, 2009). The limitations of the technology used in the laboratory must therefore be understood to minimise as far as possible parents being giving an incorrect outcome of newborn screening, either receiving false reassurance or being unduly alarmed, and may assist in the formulation of policy on referral of potentially affected infants for definitive diagnosis.

The provision of samples from a beta thalassaemia major infant as a specimen for newborn screening EQA is a challenge because of the small number of clinical cases and the difficulty of obtaining an umbilical cord sample from an affected infant. The use of a manipulated adult sample without explanation of its provenance may result in different, confused interpretations, e.g. a diagnosis of homozygous Hb E or Hb E-beta thalassaemia as a result of the presence of Hb A_2, limiting the effectiveness of the survey. It may be possible to send out such an adult sample and amend the scoring criteria to reflect that the key indicator was to report a lack of Hb A; alternatively, the scheme could supply samples of differing Hb A%, as an educational exercise to reinforce the participant’s awareness of the performance of their analyser. This fits with the objective of the scheme to improve laboratory performance through education.

The statistical analysis of participants' data is one of the variables that may confound EQA provision (Jones, 2017b; Miller 2011; Kristensen, 2017; Coucke, 2017). Regardless of whether the scheme design is to assess performance against criteria based on regulatory, state-of-the-art, biological variation or clinical outcomes, there is a desire from laboratory managers and instrument manufacturers to see the statistical evaluation of their analyser against its peers, probably for the early identification of adverse performance trends. Within UK NEQAS Haematology, the performance scoring system has been designed to give a long-term and retrospective assessment of performance for quantitative analytes against a consensus mean target of participants' results, trimmed to remove outliers. The system was designed to avoid penalising a participant for a one-off, out of consensus result. The scoring system is familiar and well understood by participants (Participants' questionnaire, 2015, reported in unpublished Steering Committee papers), but there is concern over both the length of time it may take for an unsatisfactorily performing laboratory to be detected for Hb A_2 measurement and the fact that the score will not detect a one-off error that would constitute misdiagnosis if it were a patient's sample. This study found that there were 48/10,463 cases of potential misdiagnosis of beta thalassaemia on the basis of Hb A_2 measurement, forty-two of which were made by participants with satisfactory performance scores, plus an additional thirteen potential misdiagnoses identified as non-analytical errors such as reporting Hb A% or Hb S% instead of Hb A_2% etc. These are relatively small numbers, approximately 0.5% of the total, which might be detected by some other validation step in the case of a patient. However, each of these potentially represented a misdiagnosis of beta thalassaemia, which might result either in unnecessary partner testing or a beta thalassaemia carrier

antenatal patient being missed. From the perspective of the EQA scheme, these potential misdiagnoses are missed learning opportunities that should be used to give feedback on performance to the laboratory. Modifications to the scoring criteria, e.g. by increasing the multiplier to increase the sensitivity of scoring and screening for participants with a deviation index (DI) value of greater than three for any specimen would allow the scheme to utilise these opportunities better. The scheme should also pilot the use of the ALTM as the target value as there is no evidence based reason for continuing to use the MTM, the use of which is based on the differences that may have existed for earlier manual methods. There is no evidence that the use of ALTM benefits or penalises large or small instrument groups but this should be kept under review and consideration given to the use of a ‘mean of means’ in the future.

The inclusion of interpretive EQA is a major area of development for all EQA providers, whether as part of 'end-to-end' EQA or the provision of a separate case interpretation service. The scoring of interpretation has a number of challenges not found in the monitoring of quantitative testing. In particular, the scoring must take into account the scope of the service provided by the laboratory, the methods it uses and its diagnostic expertise to ensure that assessment is objective and effective. This study has highlighted differences in errors of interpretation between English NHS Trust laboratories under the oversight of the NHS STP and Non-British Isles laboratories. English NHS Trust laboratories demonstrated a more comprehensive and accurate use of interpretive comments, supporting the positive impact of an active programme of education and performance oversight. Educational action and feedback by UK NEQAS Haematology may improve the performance of laboratories outside the UK with the introduction of scoring for

interpretation and proactive follow up of errors by all laboratories, regardless of geographical location, as in other specialities (Kettelhut, 2003; Jennings, 2009). There are a number of distinct advantages in implementing active performance management for all participants as it should improve the quality of data gathered, give a better representation of the areas for improvement and add value to participation for laboratories wherever they are based. An initial challenge in the introduction of scoring for interpretation will be how to represent performance if there is a material difference between participant groups; it is possible to assess participants according to geographical location but this may be professionally sensitive where one country or region performs less well than another. However, there was sufficient consensus between the major English NHS and Non-British Isles laboratories to suggest that scoring of interpretation should be piloted and include all participants.

A UK NEQAS Haematology survey of participants' attitudes (Participants' questionnaire 2015, unpublished Steering Committee papers) has shown that participants support the introduction of different levels of registration in the haemoglobinopathies to take account of the repertoire and expertise of the laboratory, for example allowing laboratories with only one analytical method available to register as 'screening' laboratories. This pattern of registration has been used by UK NEQAS Haematology for performance assessment in blood parasite screening and identification by blood film morphology, where again there was a disparity between the scope of services and expertise of different laboratories. An amendment to the scheme design of this nature could also have an educational impact on participants, removing the apparent 'approval' of their identifying a Hb variant as Hb S or Hb C etc. on a single method. In order to

achieve this, UK NEQAS Haematology will need to redesign the workflow for data capture and evaluation to reflect the manner in which the laboratory has generated their results and provide a more holistic approach to EQA performance.

Although the scheme issued a comprehensive range of samples in the study period, the proportions of abnormal cases has deteriorated since 2001 when the annual report showed the inclusion of three beta thalassaemia carriers, one Hb D carrier and five Hb S carriers out of twelve specimens (UK NEQAS Haematology Biennial Scheme Report 2001). The scheme may be able to resolve the provision of common conditions distributed with additional support from NHSBT but there will remain some, e.g. Hb O^{Arab} carrier, Hb E carrier and homozygous or compound heterozygous disease states, that are unlikely to be available in sufficient frequency as a whole blood specimen for performance assessment. In this situation, the gaps could be filled with 'dry' cases, either as an additional, online EQA specimen or on an educational basis. Dry cases will need to be offered as an online programme to allow participants to pick the instrument or method outputs appropriate to their laboratory and to ensure adequate quality of reproduction of chromatograms or other information sources, as well as to manage the data return. This type of programme has proved useful to the laboratory for education, continuing professional development and competency as well as for EQA (Brereton, 2008; Hastings, 2008). Registration by participant reference number (PRN) or laboratory identifier for online cases could be linked to performance with standard 'wet' specimens and an assessment of the overall improvement in scores in the scheme undertaken as a piece of further work.

The demand for high quality EQA for the haemoglobinopathies is likely to grow with the increased requirement for screening and diagnosis of the disorders in

both endemic and traditionally non-endemic countries. The potential development of global EQA will also require EQA providers with expertise in material preparation and data analysis across a wide geographical range. At home, the disorders will remain a diagnostic challenge with an increasingly diverse population and the globin gene mutations encountered (Henderson. 2009). Changes in the provision of diagnostic services with increased 'hub and spoke' configurations and centralisation of specialist diagnostics as described in Lord Carter of Coles' review of Pathology Services in England and a more recent NHS Improvement report (Karakusevic, 2016), may also demand the development of different modes of performance assessment.

The outcomes of this study have shown that the unique resource of the scheme's data can demonstrate areas of concern in laboratory performance. The study has highlighted initial and further areas of development for UK NEQAS Haematology in the tailoring of performance assessment to clinical practice in the laboratory diagnosis of the haemoglobinopathies and the extension of feedback to all laboratories to improve their services to patients.

7.2 Recommendations

This study has covered a number of areas for haemoglobinopathy EQA provided by UK NEQAS Haematology. To ensure that the programme remains fit for purpose for its current participant base, the scheme should consider the following developments:

7.2.1 Review and amendment of the current scoring algorithm for Hb A_2.

The first change should be the amendment of the target value used for Hb A_2 performance assessment from the method trimmed mean (MTM) to the

all laboratory trimmed mean (ALTM). This should be undertaken as soon as possible (i.e. before the end of 2018) subject to the necessary validation of changes to the database and change control. The use of different target values is not evidence-based and may penalise participants using CZE analysers unfairly because of the smaller inter-analyser variation with the CZE group in comparison to the HPLC group. The impact of this change will be evaluated through monitoring the number and analyser type of instruments in the PUP category for Hb A_2, which should be kept to no more than 5% of participants. Once the impact of this change has been assessed, the scheme should pilot increasing the multiplier for Hb A_2 to increase the sensitivity of the performance assessment. Again, the guideline for the proportion of PUP laboratories should be approximately 5%.

7.2.2 The addition of a 'misdiagnosis' category to Hb A_2 performance assessment

In addition to long-term performance assessment, the scoring system should alert participants to errors in EQA that would result in a misdiagnosis if they occurred with a patient's specimen. This is most easily achieved by introducing the flagging and review of any EQA result with a DI value equal to or greater than 3 in addition to the changes to the scoring algorithm in the current year. Not all of these results will need to be notified actively to participants but those that represent a misdiagnosis (e.g. a beta thalassaemia specimen reported as normal and *vice versa*) should be followed up actively.

7.2.3 The introduction of scoring for fraction identification and interpretative comments

Although out-of-consensus fraction identification is notified to UK participants, it is essential that scoring for this and for interpretive comments is introduced within 12 to 18 months. This will entail a redesign of the data entry workflow better to reflect current laboratory practice, with different classes of participant registration, e.g. to allow for screening and diagnostic laboratories. The scoring of interpretation is linked to fraction identification as the laboratory methods used may restrict or dictate the interpretive options available. The major challenge with scoring of interpretation is the establishment of the 'model' or expected answer; although consensus is informative, the most commonly reported item may not be the most clinically significant. The provision of a model answer will require a panel of recognised experts in the field.

7.2.4 The introduction of 'dry' cases for haemoglobinopathy EQA and/or educational purpose

'Dry' cases provide an educational opportunity and performance assessment of scenarios that include rare haemoglobinopathy carriers or disease states. They should be relatively easy to establish in the next phase of the development of the new UK NEQAS Haematology digital platform. They will require a single sample of blood from a patient with the condition of interest, to be tested through the various screening and diagnostic methods available so that the participant can view the results in a familiar output, together with a background clinical scenario and other laboratory results. The results returned may be reported without scoring for

educational or competency assessment or may be scored against the same criteria as the more usual ‘wet’ specimens. The addition of multiple choice questions and additional background information, revealed in a stepwise fashion to the user could allow the assessment to be tailored according to the knowledge, training and experience of individual healthcare scientists. Before commitment to a viable commercial developed software product, this development will require modelling in an open source platform, such as Google drive or Survey Monkey, trialling a minimum of six cases over the space of a year. Once established, the cases could form a library for use by subscribing laboratories or individuals.

7.2.5 The application of active performance assessment to all participants, regardless of geographical location

The extension of active performance assessment to all participants has been identified as necessary to fulfil the duty of care the Scheme has to all participants and to maintain the integrity and quality of the data returned for analysis. Since more than 50% of participating laboratories are outside the UK and RoI at the time of writing, this will significantly increase the volume of performance notifications. English is widely spoken amongst participating laboratories; however, some action may be required to overcome language barriers. This may be as simple as providing a directory of numbered comments translated into the languages most commonly used amongst participants. Additional performance monitoring will increase the Scheme’s workload and will only be manageable with the implementation of an automated, rules-based system of performance assessment. This will require an algorithm of rules to be developed and tested with historical data

before coding by the UK NEQAS Haematology commercial database provider. Initial pilot testing will be with a single scheme and a relatively non-controversial parameter, such as participation performance. Once established for UK participants, the process will be rolled out to non-UK laboratories.

7.2.6 The development of graded Hb A% cases for use in the newborn sickle screening scheme

The extension of the work undertaken on the measurement of Hb A% as part of newborn haemoglobinopathy screening is novel. Clinical, liquid blood specimens from newborn infants with Hb A% less than 1.5% in sufficient volume for EQA are rare and the most reliable source will be simulated material prepared from an adult patient without Hb A (e.g. a patient homozygous for delta-beta thalassaemia) 'spiked' with normal cord blood to give different of Hb A% values. The drawback of such simulated material is that it contains Hb A_2, which may confound the interpretation; in this case, it is best to identify the specimens as suitable for Hb A% only and issued as a separate exercise from other EQA material. The data returned will provide no less information on performance, especially of different analysers and this would be useful work to follow up. Such exercises could be run as part of the regular schedule of exercises, with one annual distribution and surplus dbs cards could be made available to subscribing laboratories on request for use in instrument evaluation.

7.2.7 Evaluation of the recommended changes

The recommended changes are far-reaching and will take a minimum of one to two years to implement fully for a participant base of more than 300

laboratories. The changes will incur costs in terms of software development and staff resources; however, if correctly implemented, they should improve the service to participant laboratories and ensure that the Scheme maintains its competitive position. Evaluation at all stages of implementation is important to ensure that the services are addressing the needs of participating laboratories and other stakeholders, e.g. oversight bodies and commercial manufacturers. Assessment of stakeholder satisfaction and the effectiveness of the changes will take several forms:

1. **Focus group evaluation during planning and pilot stages**

   The key focus groups are the Scheme's expert scientific advisory panel, the strategic Steering Committee and the NQAAP, which is the oversight body made up of representatives of the professional societies. Discussion and feedback from these committees and working or task groups formed from their members will provide expert, up-to-date evaluation of the plans and ensure that professional guideline advice in the field is incorporated into the developments.

2. **Trend analysis of participant performance**

   The Scheme monitors the numbers and types of unsatisfactory and persistent unsatisfactory UK performers as part of the scheme design and the reporting requirements of the professional societies. These numbers will increase initially as a result of the proposed changes and the numbers of UPs and PUPs will increase significantly if the non-UK laboratories are included. However, analysis of the trends in performance should demonstrate an eventual downward trend for

individual laboratories and method groups, with a removal of less effective methodologies. Direct evaluation of individual scenarios, e.g. the improvement in performance with a wet sample following a similar, educational ‘dry’ case scenario, and the evaluation the interpretation of paired cases several months apart will give a measure of the effectiveness of the educationanal initiatives.

3. **Comparison of UK and non-UK performance**

The goal of extending performance assessment to non-UK laboratories is to align their performance better with that of UK laboratories. This will be relatively easy to assess using the numbers of UP and PUP laboratories in both categories, which could be further analysed by type and work profile of the laboratory, compliance with accreditation standards and patterns of practice gathered from questionnaire data distributed with surveys.

4. **Customer satisfaction feedback**

For the Scheme to maintain its competitive position it must respond to participants’ needs. The maintenance and renewal of participation and a continued low number of complaints are measures of satisfaction but other metrics may better reflect the opinions of users, e.g. average satisfaction scores (Morgan, 2006). The use of customer satisfaction surveys can be misleading unless these are well designed, targeted and cover a comprehensive and representative cohort of service users. The Scheme may benefit from taking expert advice on the construction of an effective questionnaire to assess satisfaction with the changes.

References

Acuto, S., Baiamonte, E., Di Stefano, R., Spina, B., Barone, R. and Maggio, A., 2014. Development and recent progresses of gene therapy for β-thalassemia. *Thalassemia Reports*, **4;** DOI**:** https://doi.org/10.4081/thal.2014.2925

Altman, D.,1991. In *Practical Statistics for Medical Research*. Chapman & Hall. London.

Angastiniotis, M. and M. Hadjiminas, 1981. Prevention of thalassaemia in Cyprus. *The Lancet*, **317**: p. 369-371.

Angastiniotis, M., Eleftheriou, A., Galanello, R., Harteveld, C.L., Petrou, M., Traeger-Synodinos, J., Giordano, P., Jauniaux, E., Modell, B. and Serour, G., 2013. In *Prevention of Thalassaemias and Other Haemoglobin Disorders: Volume 1: Principles*. Thalassaemia International Federation, Nicosia, Cyprus.

Angelucci, E., 2010. Hematopoietic stem cell transplantation in thalassemia 2010. *ASH Education Program Book*: p. 456-462.

Angelucci, E. and Baronciani D., 2008. Allogeneic stem cell transplantation for thalassemia major. *Haematologica*, **93:** 1780-1784.

Antmen, A., Angelucci, E., Losi, S., Burrows, N., Bartiromo, C. and Hu, X. H., 2017. Direct Medical Care Cost Associated with β-Thalassemia Care in Turkey. *America Society of Hematology,* **130:** 2094.

Bain, B.J., 1988. Screening of antenatal patients in a multiethnic community for beta thalassaemia trait. *Journal of Clinical Pathology*, **41**: p. 481-5.

Bain, B.J., 2006. In *Haemoglobinopathy Diagnosis.* 2nd edition. Blackwell Publishing, Oxford, UK.

Bain, B J., Wild, B. J., Stephens, A. D. and Phelan, L., 2010. *Variant Haemoglobins: a guide to identification.* Wiley-Blackwell, Chichester, UK.

Baldwin, J., 1980. Structure and cooperativity of haemoglobin. *Trends in Biochemical Sciences,* **5**: p. 224-228.

Banfi, G., G.L. Salvagno, and G. Lippi, 2007.The role of ethylenediamine tetraacetic acid (EDTA) as in vitro anticoagulant for diagnostic purposes. *Clinical Chemical Laboratory Medicine*, **45**: p. 565-576.

Bardakdjian-Michau, J., Bahuau, M., Hurtrel, D., Godart, C., Riou, J., Mathis, M. and Goossens, M., 2009. Neonatal screening for sickle cell disease in France. *Journal of Clinical Pathology*, **62**: p. 31-33.

Barth, J.H., 2009. Reference ranges still need further clarity. *Annals of Clinical Biochemistry*, **46**: p. 1-2.

Batterbee, H., 2010. Evaluation of UK NEQAS (H) Hb A_2 and related performance data. UK NEQAS (H), available on request.

Beet, E.A. 1947.The Genetics of the Sickle Cell Trait in a Bantu Tribe. *Annals of Eugenics*, **1**: p. 279-284.

Benson, J.M. and Therrell B.L., 2010. History and current status of newborn screening for hemoglobinopathies. *Seminars in Perinatology*, **34:** 134-144.

Berg, J. and Lane V., 2011. Pathology Harmony; a pragmatic and scientific approach to unfounded variation in the clinical laboratory. SAGE Publications Sage UK: London, England.

Berge Kristensen, G.B. and Meijer, P., 2017. Interpretation of EQA results and EQA-based trouble shooting*. Biochemia Medica*, **27**: p. 49-62.

Beutler, E. and West, C., 1983. Storage of red cell concentrates in CPD-A2 for 42 and 49 days. *The Journal of Laboratory and Clinical Medicine*, **102**: p. 53-62.

Bland, M., 2015. An introduction to medical statistics. Oxford University Press (UK).

Bio-Rad VARIANT™ nbs Sickle Cell Instruction Manual (Bio-Rad Laboratories, CA, USA).

Bohr, C., 1904. Theoretische Behandlung der quantitativen Verhaltuis bei der Sauerstoffaufnahine der Hamoglobins. *Zbl Physiol*, **17:** p 682.

Bourner, G., De la Salle, B., George, T., Tabe, Y., Baum, H., Culp, N., and Keng, T. B., 2014. ICSH guidelines for the verification and performance of automated cell counters for body fluids. *International Journal of Laboratory Hematology,* **36**: p. 598-612.

Bozkurt, G., 2007.Results from the north cyprus thalassemia prevention program. *Hemoglobin*, **31**: p. 257-264.

Brereton, M.L., De La Salle, B., Burthem, J., Ardern, J., Hickman, L., Seal, L., McTaggart, P., West, M., Swirsky, D., Parker-Williams, J. and Hyde, K., 2008. Review of the UK NEQAS (H) digital morphology pilot scheme for continuing professional development accessed via the internet. *International Journal of Laboratory Hematology*, **30**: p. 365-371.

Briggs, C., Culp, N., Davis, B., d'Onofrio, G., Zini, G. and Machin, S.J., 2014. ICSH guidelines for the evaluation of blood cell analysers including those used for differential leucocyte and reticulocyte counting. *International journal of laboratory hematology,* **36:** p.613-627.

Brosious, E.M., Wright, J.M., Baine, R.M. and Schmidt, R.M., 1978. Microchromatographic methods for hemoglobin A_2 quantitation compared. *Clinical chemistry*, **24:** p.2196-2199.

Broughton, P.M.G., Gowenlock, A.H., McCormack, J.J. and Neill, D.W., 1974. A revised scheme for the evaluation of automatic instruments for use in clinical chemistry*. Annals of Clinical Biochemistry*, **11**: p.207-218.

BS EN 14136:2004 - Use of external quality assessment schemes in the assessment of the performance of in vitro diagnostic examination procedures. British Standards Institute, London, UK.

Cappellini, M. D., Cohen, A., Eleftheriou, A., Piga, A., Porter, J. Taher, A. 2008. In *Guidelines for the Clinical Management of Thalassaemia*, 2nd edition. Thalassaemia International Federation, Nicosia, Cyprus.

Cao, A. and. Kan Y.W, 2013. The prevention of thalassemia. *Cold Spring Harbor perspectives in medicine*, **3**: p. a011775.

Ceriotti, F., 2014. The role of External Quality Assessment Schemes in monitoring and improving the standardization process. *Clinica Chimica Acta,* **432**: p. 77-81.

Chow, J., Phelan, L. and Bain, B.J., 2005. Evaluation of single-tube osmotic fragility as a screening test for thalassemia. *American Journal of Hematology*, **79**: p. 198-201.

Çimen, M.Y. Burak, 2008. Free radical metabolism in human erythrocytes. *Clinica Chimica Acta*. 390: p. 1-11,

Clarke, G.M. and Higgins, T.N., 2000. Laboratory Investigation of Hemoglobinopathies and Thalassemias: Review and Update. *Clinical Chemistry,* **46**: p. 1284-1290.

Clarke, S.A., Skinner, R., Guest, J., Darbyshire, P., Cooper, J., Shah, F., Roberts, I. and Eiser, C., 2010. Health - related quality of life and financial impact of caring for a child with Thalassaemia Major in the UK. *Child: care, health and development*, **36:** pp.118-122.

Clegg, J.B. and Weatherall, D.J., 1999. Thalassemia and Malaria: New Insights into an Old Problem. *Proceedings of the Association of American Physicians*, **111**: p. 278-282.

Clegg, J.B., Weatherall, D.J., and P.F. Milner, 1971. Haemoglobin Constant Spring—A Chain Termination Mutant? *Nature*, **234**: p. 337.

Clinical Laboratory Improvement Amendments of 1988 (CLIA) (https://www.gpo.gov/fdsys/pkg/STATUTE-102/pdf/STATUTE-102-Pg2903.pdf) (accessed December 15, 2017).

Clinical Laboratory Standards Institute (CLSI). User Verification of Performance for Precision and Trueness Approved Guideline, 2nd edn. CLSI document EP15-A2 (ISBN1-56238-574-7). CLSI, 940 West Valley Road, Suite 1400, Wayne, PA, 19087 USA, 2005.

Clinical Laboratory Standards Institute (CLSI). Evaluation of Precision Performance of Quantitative Measurement Methods; Approved Guideline, 2nd edn. CLSI document EP5-A2 (ISBN1-56238- 542-9). CLSI, 940 West Valley Road, Suite 1400, Wayne, PA, 19087 USA, 2004.

Clinical Laboratory Standards Institute (CLSI). Evaluation of Detection Capability for Clinical Laboratory Measurement Procedures; Approved Guideline, 2nd edn. CLSI document EP17-A2 (ISBN1-56238-795-2 [print]; ISBN1-56238-796-0 [electronic]). CLSI, 940 West Valley Road, Suite 1400, Wayne, PA, 19087 USA, 2012.

Clinical Laboratory Standards Institute (CLSI). Evaluation of the Linearity of Quantitative Measurement Procedures: A Statistical Approach; Approved Guideline. CLSI document EP6-A (ISBN1-56238-498- 8). CLSI, 940 West Valley Road, Suite 1400, Wayne, PA, 19087 USA, 2003.

College of American Pathologists (CAP) 2017. Performance report: survey HG-B 2017, Hemoglobinopathy.

Colombo, B., Kim, B., Atencio, R. P., Molina, C. and Terrenato, L., 1976. The Pattern of Fetal Haemoglobin Disappearance after Birth. *British Journal of Haematology*, **32:** 79–88.

Cook, J.E. and Meyer J.,1915. Severe anemia with remarkable elongated and sickle-shaped red blood cells and chronic leg ulcer. *Archives of Internal Medicine*, **XVI**: p. 644-651.

Cotton, F., Lin, C., Fontaine, B., Gulbis, B., Janssens, J. and Vertongen, F., 1999. Evaluation of a capillary electrophoresis method for routine determination of hemoglobins A_2 and F. *Clinical chemistry*, **45:** p.237-243.

Coucke, W., China, B., Delattre, I., Lenga, Y., Van Blerk, M., Van Campenhout, C., Van de Walle, P., Vernelen, K. and Albert, A., 2012. Comparison of different approaches to evaluate External Quality Assessment Data. *Clinica Chimica Acta*, **413**: p.582-586.

Coucke, W. and Rida Soumali, M., 2017. Demystifying EQA statistics and reports. *Biochemia Medica,* **27**: p. 37-48.

Cousens, N.E., Gaff, C.L., Metcalfe, S.A. and Delatycki, M.B., 2010. Carrier screening for beta-thalassaemia: a review of international practice. *European Journal of Human Genetics*, **18**: p. 1077-1083.

Daniel, Y. A., Turner, C., Haynes, R. M., Hunt, B. J. and Dalton N. R., 2007. Quantification of Hemoglobin A_2 by Tandem Mass Spectrometry. *Clinical Chemistry,* **53**: p. 1448-1454.

Davies, S.C., Cronin, E., Gill, M., Greengross, P., Hickman, M. and Normand, C., 2000. Screening for sickle cell disease and thalassaemia: a systematic review with supplementary research. *Health Technology Assessment (Winchester, England)*, **4:** p.i-v.

De la Salle, B., 2017. Survey material choices in haematology EQA: a confounding factor in automated counting performance assessment. *Biochemia Medica*, **27**: p. 63-72.

De la Salle, B.J., McTaggart, P.N., Briggs, C., Harrison, P., Doré, C.J., Longair, I., Machin, S.J. and Hyde, K., 2012a. The accuracy of platelet counting in thrombocytopenic blood samples distributed by the UK National External Quality Assessment Scheme for General Haematology. *American Journal of Clinical Pathology*, **137**: p. 65-74.

De la Salle, B., 2012b. Pathology harmony moves on: progress on implementation in haematology. *British Journal of Haematology*.,**158**: p. 804-805.

Dixon, W.J., 1950. Analysis of extreme values. *The Annals of Mathematical Statistics*, **21**: p. 488-506.

Efremov, C.D., Huisman, T.H.J., Bowman, K., Wrightstone, R.N. and Shroeder, W.A., 1974. Microchromatography of hemoglobins. II. A rapid method for the determination of hemoglobin A_2. *Translational Research,* **83**: p. 657-664.

Efremov, G.D., 1977. An Evaluation of the Methods for Quantitation of Hemoglobin A_2; Results from a Survey of 10,663 Cases. *Hemoglobin*, **1**: p. 845-860.

England, J., 1995. Report to the General Haematology Steering Committee on a scoring system for UK NEQAS General Haematology. *Steering committee paper,* 55.5.

European Directive 2011/24/EU of the European Parliament and the Concil on the application of patients' rignts in cross border healthcare, 2013.

Fantoni, A., M. Farace, and Gambari, R., 1981. Embryonic hemoglobins in man and other mammals. *Blood*, **57**: p. 623-633.

Faulkner, L., Yaqub, N., Khalid, S.K., Zhara, T., Ansari, S., Farzana, T. and Shamsi, T., 2011. Transplantation in low resource countries. *Thalassemia Reports*, **1**: Proceedings of the 12th International Conference on Thalassemia and the Haemoglobinopathies.

Fraser, C.G. and Petersen, P.H., 1999. Analytical performance characteristics should be judged against objective quality specifications. *Clinical Chemistry,* **45: p.** 321-323.

Fucharoen, S. and Winichagoon, P., 2012. New updating into hemoglobinopathies. *International Journal of Laboratory Hematology,* **34**: p. 559-565.

Galanello, R., Melis, M.A., Muroni, P. and Cao, A., 1977. Quantitation of Hb A_2 with DE-52 Microchromatography in Whole Blood as Screening Test for β-Thalassemia Heterozygotes. *Acta Haematologica*, **57**: p. 32-36.

Galanello, R., 2003. Deferiprone in the treatment of transfusion-dependent thalassemia: a review and perspective. *Therapeutics and Clinical Risk Management*, **3**: p795-805.

Galanello, R. and Origa. R., 2010 .Beta-thalassemia. *Orphanet Journal of Rare Diseases,* **5**: p. 11.

Galanello, R., Ruggeri, R., Addis, M., Paglietti, E. and Cao, A., 1981.Hemoglobin A_2 in Iron Deficient βThalassemia Heterozygotes. *Hemoglobin*, **5**: p. 613-618.

Gaston, M.H., Verter, J.I., Woods, G., Pegelow, C., Kelleher, J., Presbury, G., Zarkowsky, H., Vichinsky, E., Iyer, R., Lobel, J.S. and Diamond, S., 1986. Prophylaxis with oral penicillin in children with sickle cell anemia. *New England Journal of Medicine*, **314**: p. 1593-1599.

Gaziev, J., P. Sodani, and Lucarelli, G., 2008. Hematopoietic stem cell transplantation in thalassemia. *Bone marrow transplantation,* **42**: p. S41-S41.

Giambona, A., Passarello, C., Renda, D. and Maggio, A., 2009. The significance of the hemoglobin A_2 value in screening for hemoglobinopathies. *Clinical Biochemistry,* **42**: p. 1786-1796.

Giardine, B., van Baal, S., Kaimakis, P., Riemer, C., Miller, W., Samara, M., Kollia, P., Anagnou, N. P., Chui, D. H.K., Wajcman, H., Hardison, R. C. and Patrinos, G. P., 2007. HbVar database of human hemoglobin variants and thalassemia mutations: 2007 update. *Human Mutation*, **28:** p 206

Giordano, P.C., 2009. Prospective and retrospective primary prevention of Hemoglobinopathies in multiethnic societies. *Clinical Biochemistry*, **42**: p. 1757-1766.

Globin Gene Disorder Working Party of the BCSH General Haematology Task Force. 1994. Guidelines for the fetal diagnosis of globin gene disorders. *Journal of Clinical Pathology*, **47**: p. 199-204.

Grubbs, F.E., 1969. Procedures for detecting outlying observations in samples. *Technometrics,* **11:** p. 1-21.

Gulbis, B., Eleftheriou, A., Angastiniotis, M., Ball, S.,Surrallés, J.,Castella, M., Heimpel, H., Hill, A. and Corrons, J-L., 2010. Epidemiology of rare anaemias in Europe. *Advances in Experimental Medicine and Biology*, **686:** p 375-396

Gulbis, B., Fontaine, B., Vertongen, F. and Cotton, F., 2003. The place of capillary electrophoresis techniques in screening for haemoglobinopathies. *Annals of clinical biochemistry*, **40:** p.659-662

Hall, G.W., R.A. Barnetson and Thein, S.L., 1992. Beta thalassaemia in the indigenous British population. *British Journal of Haematology*, **82**: p. 584-588.

Hammerling, J.A., 2015. A review of medical errors in laboratory diagnostics and where we are today. *Laboratory Medicine*, **43**: p. 41-44.

Hardison, R.C., Chui, D. H. K., Riemer, C. R., Milleg, W., Carver, M. F. H., Molchanova, T. P., Efremov, G. D. and Huisman, T. H. J., 1998. Access to a Syllabus of Human Hemoglobin Variants (1996) Via the World Wide Web. *Hemoglobin*, **22:** p. 113-127.

Hardison, R. C., Chui, D. H.K., Giardine, B., Riemer, C., Patrinos, G. P., Anagnou, N., Miller, W. and Wajcman, H., 2002. HbVar: A relational database of human hemoglobin variants and thalassemia mutations at the globin gene server. *Human Mutation,* **19**: p. 225-233.

Hastings, R.J. and Howell, R.T., 2010. The importance and value of EQA for diagnostic genetic laboratories. *Journal of community genetics*, **1**: p. 11-17.

Hastings, R.J., Maher, E.J., Quellhorst-Pawley, B. and Howell, R.T., 2008. An Internet-based external quality assessment in cytogenetics that audits a laboratory's analytical and interpretative performance. *European Journal of Human Genetics*. **16:** p.1217-1224.

Head, C., Conroy, M., Jarvis, M., Phelan, L. and Bain, B., 2004. Some observations on the measurement of haemoglobin A_2 and S percentages by high performance liquid chromatography in the presence and absence of alpha thalassaemia. *Journal of Clinical Pathology.* **57**: p 276–280.

Healy, M., 1979. Outliers in clinical chemistry quality-control schemes. *Clinical Chemistry.* **25**: p. 675-677.

Hecht, F., Jones, R.T. and Koler, R.D., Newborn infants with Hb Portland 1, an indicator of α-chain deficiency. *Annals of Human Genetics*. 1968. **31**: p. 215-218.

Henderson, S., Timbs, A., McCarthy, J., Gallienne, A., Van, M., Masters, G., May, A., Khalil, M. S., Schuh, A. and Old, J., 2009. Incidence of haemoglobinopathies in various populations — The impact of immigration. *Clinical Biochemistry*. **42**: p. 1745-1756.

Herrick, J.B., 1910. Peculiar elongated and sickle-shaped red blood corpuscles in a case of severe anaemia. *Archives of Internal Medicine.* **6**: p. 517-521.

Heydorn, K., 2008. The distribution of interlaboratory comparison data. *Accreditation and Quality Assurance*. **13**: p. 723-724.

Hoppe, C.C., 2009. Newborn screening for non-sickling hemoglobinopathies. *ASH Education Program Book*. **2009**: p. 19-25.

Horst, B.,Gerrard, J. and Hickmans E.M., 1954. The Influence of Phenylalanine Intake on the Chemistry and Behaviour of a Phenylketonuria Child. *Acta Pædiatrica*. **43**: p. 64-77.

Huehns, E.R., Flynn, F. V., Butler, E. A. and Beaven, G. H., 1961. Two New Hæmoglobin Variants in a Very Young Human Embryo. *Nature.* **189**: p. 496 - 497.

Huisman, T.H.J., Carver, M.F.H. and Baysal, E., 1997. A Syllabus of Thalassemia Mutations, in *The Sickle Cell Anemia Foundation*, Augusta, GA.

Human Tissue Act (2004). Human Tissue Authority. http://www.legislation.gov.uk/ukpga/2004/30/contents (accessed December 12, 2017).

International Committee For Standardization In Haematology, 1978. Recommendations for Selected Methods for Quantitative Estimation of Hb A_2 and for Hb A_2 Reference Preparation. *British Journal of Haematology*. **38**: p. 573-578.

ISO 15189:2012 – Medical laboratories – Requirements for quality and competence. https://www.iso.org/standard/56115.html (accessed December 15 2017).

ISO/IEC 17043:2010 – Conformity assessment – General requirements for proficiency testing. https://www.iso.org/standard/29366.html (accessed December 15, 2017).

ISO 13528:2015 – Statistical methods for use in proficiency testing by interlaboratory comparison. https://www.iso.org/standard/56125.html (accessed December 15 2017).

James, D., Ames, D., Lopez, B., Still, R. and Simpson, W., 2014. External Quality Assessment: best practice. *Journal Clinical Pathology.* **65:** p. 651-655.

James, V.L., 2012. Evaluating participant performance in qualitative PT/EQA schemes. *Accreditation and Quality Assurance*. **17**: p. 375-378.

Jennings, I., Kitchen, D.P., Woods, T.A.L., Kitchen, S., Walker, I.D. and Preston, F.E., 2009. Laboratory performance in the World Federation of Hemophilia EQA programme, 2003–2008. *Haemophilia*. **15**: p. 571-577.

Jeppsson, J.O., Kobold, U., Barr, J., Finke, A., Hoelzel, W., Hoshino, T., Miedema, K., Mosca, A., Mauri, P., Paroni, R. and Thienpont, L., 2002. Approved IFCC reference method for the measurement of HbA1c in human blood. *Clinical Chemistry and Laboratory Medicine.* **40**: p. 78-89.

Jones, G.R., 2017a. The role of EQA in harmonization in laboratory medicine–a global effort. *Biochemia Medica.* **27**: p. 23-29.

Jones G. R., Albarede, S., Kesseler, D., MacKenzie, F., Mammen, J. Pedersen, M., Stavelin, A., Thelen, M. Thomas, A., Twomey P., Ventura, E. and Panteghini, Mauro, 2017b. Analytical performance specifications for external quality assessment – definitions and descriptions. *Clinical Chemistry and Laboratory Medicine (CCLM).* **55:** p. 949 956.

Karakusevic, S., Edwards, N., Lewis, R. and Dayan, M., 2016. The future of pathology services. Nuffield Trust, UK.

Karnon, J., Zeuner, D., Brown, J., Ades, A.E., Wonke, B. and Modell, B., 1999. Lifetime treatment costs of β-thalassaemia major. *Clinical & Laboratory Haematology.* **21**: p. 377-385.

Kelemen, E., 1981. Extraembryonic Haemopoiesis in Man, in *Genetics, Structure and Function of Blood Cells*. Pergamon. p. 45-52.

Kendall, A. and Bastansky, C., 1981. Hemoglobin A_2 in Hyperthroidism. *Hemoglobin*. **5**: p. 571-577.

Kettelhut, M.M., Chiodini, P.L., Edwards, H. and Moody, A., 2003. External quality assessment schemes raise standards: evidence from the UKNEQAS parasitology subschemes. *Journal of Clinical Pathology*. **56**: p. 927-932.

Kilmartin J. V. and Rossi-Benardi L, 1969. Inhibition of CO_2 combination and reduction of the Bohr effect in haemoglobin chemically modified at its α–amino groups. *Nature.* **222:** p. 1243 – 1246.

Kristensen, G.B.B. and Meijer, P., 2017. Interpretation of EQA results and EQA-based trouble shooting. *Biochemia medica*, **2**: p.49-62.

Kunkel, H., Ceppellini, R., Müller-Eberhard, U. and Wolf, J., 1957. Observations on the minor basic hemoglobin component in the blood of normal individuals and patients with thalassemia. *Journal of Clinical Investigation*. **36:** p. 1615.

Kunkel, H.G. and Wallenius, G., 1955. New Hemoglobin in Normal Adult Blood. *Science (Washington).* **122**: p. 288.

Lehmann, H. and Huntsman. R.G., 1974. *Man's haemoglobins, including the haemoglobinopathies and their investigation.* North-Holland Publishing Company, Amsterdam, The Netherlands.

Lewis, S.M., Quality assurance programmes in the United Kingdom. 1995. *Annali - Instituto Superiore di Sanita.* **31**: p. 53-53.

Lewis, S.M. and Burgess, B.J., 1969. Quality control in haematology: Report of interlaboratory trials in britain. *British Medical Journal.* **4**: p. 253-256.

Ley, D.C. and Ezer, S., 1974. The development of an interlaboratory proficiency testing program for the province of Ontario: II. Design of performance graphs. *Clinical Biochemistry*. **7**: p. 239-250.

Libeer, J.-C., 2001. Role of external quality assurance schemes in assessing and improving quality in medical laboratories. *Clinica Chimica Acta*. **309**: p. 173-177.

Louahabi, A., Philippe, M., Lali, S., Wallemacq, P. and Maisin, D., 2006. Evaluation of a new Sebia kit for analysis of hemoglobin fractions and variants on the Capillarys® system. *Clinical Chemical Laboratory Medicine*. **44**: p. 340-345.

Lucarelli, G., Isgrò, A., Sodani, P. and Gaziev, J., 2012. Hematopoietic stem cell transplantation in thalassemia and sickle cell anemia. *Cold Spring Harbor perspectives in medicine.* **2**: p.a011825.

Maharani, E.A., Soedarmono, Y.S. and Nainggolan, I.M., 2014. Frequency of thalassemia carrier and Hb variant and the quality of stored donor blood. *Medical Journal of Indonesia*. **23**: p. 209.

Mak, C.M., Lee, H.C.H., Chan, A.Y.W. and Lam, C.W., 2013. Inborn errors of metabolism and expanded newborn screening: review and update. *Critical Reviews in Clinical* Laboratory *Sciences*. **50**: p. 142-162.

Mansouri, A. and Lurie, A.A., 1993. Methemoglobinemia. *American journal of hematology*, **42:** p.7-12.

Mantikou, E., Arkesteijn, S. G., Beckhoven van, J. M., Kerkhoffs, J-L., Harteveld, C. L. and Giordano, P., 2009. A brief review on newborn screening methods for hemoglobinopathies and preliminary results selecting beta thalassemia carriers at birth by quantitative estimation of the HbA fraction. *Clinical Biochemistry.* **42**: p. 1780-1785.

Marengo-Rowe, A.J., 1965. Rapid electrophoresis and quantitation of haemoglobins on cellulose acetate. *Journal of Clinical Pathology.* **18**: p. 790-792.

Mario, N., Baudin, B., Aussel, C. and Giboudeau, J., 1997. Capillary isoelectric focusing and high-performance cation-exchange chromatography compared for qualitative and quantitative analysis of hemoglobin variants. *Clinical Chemistry*. **43**: p. 2137-2142.

Martinez, P.A., Angastiniotis, M., Eleftheriou, A., Gulbis, B., Pereira, M.D.M.M., Petrova-Benedict, R. and Corrons, J.L.V., 2014. Haemoglobinopathies in Europe: health & migration policy perspectives. *Orphanet Journal of Rare Diseases*. **9**: p. 97.

Mason, V.R., 1922.Sickle cell anemia. *Journal of the American Medical Association*. **79**: p. 1318-1320.

Maziotta. D., Harel, D., Schumann, G. and Libeer, J-C., 2003. Guidelines for the competence of EQAP organizers in medical laboratories. *IFCC/EMD/C-AQ. (http://www.ifcc.org/ifccfiles/docs/EQAP_version_3-2002.pdf, accessed December 2017).*

Mehta, B. and Agarwal, M., 1983. Haemoglobin A_2 in nutritional megaloblastic anaemia. *The Indian journal of medical research*. **77**: p. 478-481.

Mehta, P.A. and Faulkner, L.B. 2013. Hematopoietic cell transplantation for thalassemia: a global perspective BMT tandem meeting 2013. *Biology of Blood and Marrow Transplantation*. **19**: p. S70-S73.

Miller, W.G., 2003. Specimen materials, target values and commutability for external quality assessment (proficiency testing) schemes. *Clinica Chimica Acta.* **327**: p. 25-37.

Miller, W.G., Jones, G. R., Horowitz, G. L. and Weykamp, C., 2011. Proficiency testing/external quality assessment: current challenges and future directions. *Clinical Chemistry*. **57**: p. 1670-80.

Miller, W.G., Myers, G.L and Rej, R., 2006.Why Commutability Matters. *Clinical Chemistry.* **52**: p. 553-554.

Moat, S.J., Rees, D., King, L., Ifederu, A., Harvey, K., Hall, K., Lloyd, G., Morrell, C. and Hillier, S., 2014. Newborn Blood Spot Screening for Sickle Cell Disease by Using Tandem Mass Spectrometry: Implementation of a Protocol to Identify Only the Disease States of Sickle Cell Disease. *Clinical Chemistry*. **60**: p. 373 – 380.

Modell, B. and Darlison, M., 2008.Global epidemiology of haemoglobin disorders and derived service indicators, *Bulletin of the World Health Organization.* **86**: p. 480-487.

Morgan, N. A. and Rego, L. L., 2006. The Value of Different Customer Satisfaction and Loyalty Metrics in Predicting Business Performance. *Marketing Science.* **25:** p 426 – 439.

Mosca, A., Paleari R. and Wild B., 2013. Analytical goals for the determination of Hb A_2. *Clinical Chemistry and Laboratory Medicine*. **51:** p 937-941.

Mosca, A., Paleari, R., Galanello, R., Sollaino, C., Perseu, L., Demartis, F.R., Passarello, C., Giambona, A., Maggio, A. and IFCC Working Group on Standardization of HbA_2, 2008. New analytical tools and epidemiological data for the identification of HbA_2 borderline subjects in the screening for beta-thalassemia. *Bioelectrochemistry*. **73**: p. 137-140.

Mosca, A., Paleari, R., Ivaldi, G., Galanello, R. and Giordano, P.C., 2009. The role of haemoglobin A_2 testing in the diagnosis of thalassaemias and related haemoglobinopathies. *Journal of Clinical Pathology*. **62**: p. 13-17.

Murray, C., Hall, S. and Griffiths, P., 2011. An evaluation of the Sebia capillarys Neonat Haemoglobin FAST™ system for routine newborn screening for sickle cell disease. *International Journal of Laboratory Hematology*. **33**: p. 533-539.

Najmabadi, H., Ghamari, A., Sahebjam, F., Kariminejad, R., Hadavi, V., Khatibi, T., Samavat, A., Mehdipour, E., Modell, B. and Kariminejad, M.H., 2006. Fourteen-year experience of prenatal diagnosis of thalassemia in Iran. *Public Health Genomics.* **9**: p. 93-97.

Neel, J.V., 1947. The clinical detection of the genetic carriers of inherited disease. *Medicine*, **26:** p.115-154.

Neel, J.V., 1949. The detection of the genetic carriers of hereditary disease. *American journal of human genetics*. **1**: p.19-36.

NHS Newborn Blood Spot Screening Programme, 2017. Public Health England publications gateway number 2016573.

NHS Sickle and Thalassaemia Screening Programme: Handbook for antenatal laboratories, 2017. Version 4.2. Public Health England publications gateway number: 2017345.

NHS Sickle and Thalassaemia Screening Programme: Handbook for newborn screening laboratories, 2017. Public Health England publications gateway number: 2016583.

NHS Sickle Cell and Thalassaemia Screening Programme Data Report 2015/16: trends and performance analysis, 2017. Public Health England. PHE publications gateway number: 2016623.

Old, J., 2003. Screening and genetic diagnosis of haemoglobin disorders. *Blood reviews*. **17**: p. 43-53.

Old, J., Harteveld, CL., Traeger-Synodinos, J., Petrou, M., Angastiniotis, M., and Galanello, R., 2012. In *Prevention of thalassaemias and other haemoglobin disorders: volume 2: laboratory protocols.* 2nd edition. Thalassaemia International Federation, Nicosia, Cyprus.

Olivieri, N.F. and Weatherall, D., 2001. Clinical aspects of β-thalassemia. In *Disorders of hemoglobin: genetics, pathophysiology, and clinical management*. Cambridge University Press, Cambridge, UK.

Paleari, R., Caruso, D., Kaiser, P., Arsene, C.G., Schaeffer-Reiss, C., Van Dorsselaer, A., Bissé, E., Ospina, M., De Jesús, V.R., Wild, B. and Mosca, A., 2017. Developing a reference system for the IFCC standardization of HbA_2. *Clinica Chimica Acta*. **467**: p. 21-26.

Paleari, R., Giambona, A., Cannata, M., Leto, F., Maggio, A. and Mosca, A., 2007. External quality assessment of hemoglobin A_2 measurement: data from an Italian pilot study with fresh whole blood samples and commercial HPLC systems. *Clinical Chemical Laboratory Medicine*. **45** p. 88-92.

Paleari, R., Gulbis, B., Cotton, F. and Mosca, A., 2012. Interlaboratory comparison of current high-performance methods for HbA_2. *International Journal of Laboratory Hematology.* **34**: p. 362-368.

Paleari, R., Muñoz, A. and Mosca, A. 2010. Towards the development of a certified reference material for hemoglobin A_2. *Clinical Chemistry and Laboratory Medicine.* **48**: p. 1611-1618.

Panteghini, M. and Sandberg, S., 2015. Defining analytical performance specifications 15 years after the Stockholm conference. *Clinical Chemistry and Laboratory Medicine (CCLM),* **53**: p. 829-832.

Passarello, C., Giambona, A., Cannata, M., Vinciguerra, M., Renda, D. and Maggio, A., 2012. Iron deficiency does not compromise the diagnosis of high HbA_2 β thalassemia trait. *Haematologica*. **97**: p. 472-473.

Patrinos, G.P., B., Riemer, C., Miller, W., Chui, D. H. K., Anagnou, N. P., Wajcman, H. and Hardison, R. C., 2004. Improvements in the HbVar database of human hemoglobin variants and thalassemia mutations for population and sequence variation studies. *Nucleic Acids Research*, **32**: p. D537-D541.

Pauling, L., Itano, H., Singer, S., and Wells, I., 1949. Sickle Cell Anemia, a Molecular Disease. *Science*. **110**: p. 543-548.

Pearson, E.S., 1935: The application of statistical methods to industrial standardisation and quality control. *British standards institution*.

Perutz, M.F., 1978. Hemoglobin Structure and Respiratory Transport. *Scientific American*. **239**: p. 92-125.

Perutz, M.F., 1980. Review Lecture: Stereochemical Mechanism of Oxygen Transport by Haemoglobin. *Proceedings of the Royal Society of London. Series B, Biological Sciences.* **208**: p. 135-162.

Piel, F.B., 2016. The Present and Future Global Burden of the Inherited Disorders of Hemoglobin. *Hematology/Oncology Clinics*. **30**: p. 327-341.

Piel, F.B., Patil, A.P., Howes, R.E., Nyangiri, O.A., Gething, P.W., Williams, T.N., Weatherall, D.J. and Hay, S.I., 2010. Global distribution of the sickle cell gene and geographical confirmation of the malaria hypothesis. *Nature communications*. **1:** p.104.

Piva, E. and Plebani, M. 2009. Interpretative reports and critical values. *Clinica Chimica Acta.* **404**: p. 52-58.

Plebani, M., 2005. External quality assessment programs: Past, present and future. *Jugoslovenska medicinska biohemija.* **24**: p. 201-206.

Plebani, M., 2006. Errors in clinical laboratories or errors in laboratory medicine? *Clinical Chemical Laboratory Medicine*. **44**: p. 750-759.

Plebani, M., 2009a. Exploring the iceberg of errors in laboratory medicine. *Clinica Chimica Acta*. **404**: p. 16-23.

Plebani, M., 2009b. Interpretative commenting: A tool for improving the laboratory–clinical interface. *Clinica Chimica Acta*. **404**: p. 46-51.

Plebani, M., Astion, M. L.,Barth, J. H.,Chen, W.,de, C. A. Ca,Escuer, M. I.,Ivanov, A.,Miller, W. G.,Petinos, P.,Sciacovelli, L.,Shcolnik, W.,Simundic, A. M. and Sumarac, Z., 2014. Harmonization of quality indicators in laboratory medicine. A preliminary consensus. *Clinical Chemistry & Laboratory Medicine.* **52:** p. 951-958.

Plebani, M. and Piva, E. 2010. Medical errors: pre-analytical issue in patient safety. *Journal of Medical Biochemistry*. **29**: p. 310-314.

Pornprasert, S., Tookjai, M., Punyamung, M., Pongpunyayuen, P. and Jaiping, K., 2016. Development of hemoglobin typing control materials for laboratory investigation of thalassemia and hemoglobinopathies. *Clinical Chemistry and Laboratory Medicine (CCLM)*. **54**: p. 81-89.

Reardon, D.M., Mack, D., Warner, B. and Hutchinson, D., 1991. A whole blood control for blood count analysers, and source material for an external quality assessment scheme. *Medical Laboratory Sciences*. **48**: p. 19-26.

Regulation (EC) 141/2000 of the European Parliament and of the Council on orphan medicinal products (2000).

Rogers, D.W., Clarke, J.M., Cupidore, L., Ramlal, A.M., Sparke, B.R. and Serjeant, G.R., 1978a. Early deaths in Jamaican children with sickle cell disease. *British Medical Journal*. **1**: p. 1515.

Rogers, D.W., Vaidya, S. and Serjeant, G.R. 1978b. Early splenomegaly in homozygous sickle-cell disease: an indicator of susceptibility to infection. *The Lancet*. **312**: p. 963-965.

Rousseeuw, P.J. and Croux, C.,1993. Alternatives to the median absolute deviation. *Journal of the American Statistical association*. **88**: p. 1273-1283.

Routy, J.P., Monte, M., Beaulieu, R., Toma, E., St-Pierre, L. and Dumont, M., 1993. Increase of hemoglobin A_2 in human immunodeficiency virus-1-infected patients treated with zidovudine. *American Journal of Hematology.* **43**: p. 86-90.

Rund, D. and Rachmilewitz, E, 2005. β-Thalassemia. *New England Journal of Medicine*. **353**: p. 1135-1146.

Ryan, K., Bain, B. J., Worthington, D., James, J., Plews, D., Mason, A., Roper, D., Rees, D. C., De La Salle, B., Streetly, A. on behalf of the British Committee for Standards in Haematology, 2010. Significant haemoglobinopathies: guidelines for screening and diagnosis. *British Journal of Haematology,* **149**: p. 35-49.

Sahai, I. and Marsden, D., 2009. Newborn Screening. *Critical Reviews in Clinical Laboratory Sciences*. **46**: p. 55-82.

Sandberg, S., Fraser, C.G., Horvath, A.R., Jansen, R., Jones, G., Oosterhuis, W., Petersen, P.H., Schimmel, H., Sikaris, K. and Panteghini, M., 2015. Defining analytical performance specifications: consensus statement from the 1st Strategic Conference of the European Federation of Clinical Chemistry and Laboratory Medicine. *Clinical Chemistry and Laboratory Medicine (CCLM)*. **53**: p. 833-835.

Sangkitporn, S., Sangkitporn, S.K., Tanjatham, S., Suwannakan, B., Rithapirom, S., Yodtup, C., Yowang, A. and Duangruang, S., 2011. Multicenter validation of fully automated capillary electrophoresis method for diagnosis of thalassemias and hemoglobinopathies in Thailand. *Southeast Asian Journal of Tropical Medicine and Public Health*. **42**: p. 1224.

Schmidt, R.M., Rucknagel, D.L. and Necheles, T.F., 1975. Comparison of methodologies for thalassemia screening by Hb A_2 quantitation. *Translational Research.* **86**: p. 873-882.

Schneider, R.G., Hosty, T.S., Tomlin, G. and Atkins, R., 1974. Identification of hemoglobins and hemoglobinopathies by electrophoresis on cellulose acetate plates impregnated with citrate agar. *Clinical Chemistry*. **20**: p. 74-77.

Sciacovelli, L., O'Kane, M., Skaik, Y.A., Caciagli, P., Pellegrini, C., Da Rin, G., Ivanov, A., Ghys, T. and Plebani, M., 2011. Quality Indicators in Laboratory Medicine: from theory to practice. *Clinical Chemistry and Laboratory Medicine*. **49**: p. 835-844.

Sciacovelli, L., Secchiero, S., Zardo, L. and Plebani, M., 2010. The role of the external quality assessment. *Biochemia Medica.* **20**: p. 160-164.

Sebia Capillarys Neonat Hb FAST™ kit insert Ref 2006 (Sebia, France).

Shewhart, W.A., 1931. Economic control of quality of manufactured product: *ASQ Quality Press.*

Sheth, S., Weiss, M., Parisi, M. and Ni, Q., 2017. Clinical and Economic Burden of Transfusion-Dependent Beta Thalassaemia in Adult Patients in the United States. *Oral presentation, ASH Meeting 2017*. https://ash.confex.com/ash/2017/webprogram/Paper101255.html. (Accessed 3 January 2018).

Shokrani, M., Terrell, F., Turner, E.A. and Aguinaga, M.D., 2000. Chromatographic measurements of hemoglobin A_2 in blood samples that contain sickle hemoglobin. *Annals of Clinical & Laboratory Science*. **30**: p. 191-194.

Sinclair, L., Hall, S. and Badrick, T., 2014. A survey of Australian haematology reference intervals. *PathologyJournal of the RCPA*. **46**: p. 538-543.

Sreedharanunni, S., Sharma, P., Murgai, P., Chhabra, S. and Das, R., 2015. Automated alkaline-pH electrophoresis followed by densitometry does not correlate with cation-exchange (CE)-HPLC in quantification of HbA_2 and variant hemoglobins. *Clinical Chemistry and Laboratory Medicine (CCLM).* **53**: p. e301-e303.

Steinberg, M. G. and Adams, J. G. 1991. Hemoglobin A_2: origin, evolution and aftermath. *Blood*. **78**: p. 2165 - 2177.

Stephens, A.D., Angastiniotis, M., Baysal, E., Chan, V., Fucharoen, S., Giordano, P. C., Hoyer, J. D., Mosca, A. and Wild, B. 2012. ICSH recommendations for the measurement of haemoglobin A_2. *International Journal of Laboratory Hematology.* **34:** p. 1-13.

Stephens, A.D., Colah, R., Fucharoen, S., Hoyer, J., Keren, D., McFarlane, A., Perrett, D. and Wild, B. J. 2015. ICSH recommendations for assessing automated high-performance liquid chromatography and capillary electrophoresis equipment for the quantitation of HbA_2. *International Journal of Laboratory Hematology.* **37:** p. 577-582.

Streetly, A., 2000. A national screening policy for sickle cell disease and thalassaemia major for the United Kingdom: Questions are left after two evidence based reports. *British Medical Journal*, **320**: p. 1353-1354.

Streetly, A. and Holland W., 2009a. In *Population screening and public health.* 5th edition. Oxford textbook of public health, Volume 3: the practice of public health, p. 1623-1638.

Streetly, A., Latinovic, R., Hall, K. and Henthorn, J., 2009b. Implementation of universal newborn bloodspot screening for sickle cell disease and other clinically significant haemoglobinopathies in England: screening results for 2005-7. *Journal of Clinical Pathology*. **62**: p. 26-30.

Streetly, A., Latinovic, R., Henthorn, J., Daniel, Y., Dormandy, E., Darbyshire, P., Mantio, D., Fraser, L., Farrar, L., Will, A. and Tetlow, L., 2013. Newborn bloodspot results: predictive value of screen positive test for thalassaemia major. *Journal of Medical Screening*. **20**: p. 183-187.

Suh, D., Krauss, J. and Bures, K., 1996. Influence of hemoglobin S adducts on hemoglobin A_2 quantification by HPLC. *Clinical Chemistry*. **42**: p. 1113-1114.

Tan, G.B., Aw, T.C., Dunstan, R.A. and Lee, S.H., 1993. Evaluation of high performance liquid chromatography for routine estimation of haemoglobins A_2 and F. *Journal of Clinical Pathology*. **46**: p. 852-856.

Telfer, P., Constantinidou, G., Andreou, P., Christou, S., Modell, B. and Angastiniotis, M., 2005. Quality of life in thalassemia. *Annals of the New York Academy of Sciences.* **1054**: p. 273-282.

Thienpont, L., Franzini, C., Kratochvila, J., Middle, J., Ricos, C., Siekmann, L. and Stöckl, D., 1995. Analytical quality specifications for reference methods and operating specifications for networks of reference laboratories. *European journal of clinical chemistry and clinical biochemistry.* **33**: p. 949-957.

Thienpont, L.M., Stöckl, D., Friedecký, B., Kratochvila, J. and Budina, M., 2003. Trueness verification in European external quality assessment schemes: time to care about the quality of the samples. *Scandinavian journal of clinical and laboratory investigation*. **63**: p. 195-202.

Thompson, M., Ellison, S.L. and Wood, R. 2006. The international harmonized protocol for the proficiency testing of analytical chemistry laboratories (IUPAC Technical Report). *Pure and Applied Chemistry.* **78**: p. 145-196.

Thuret, I., Sarles, J., Merono, F., Suzineau, E., Collomb, J., Lena-Russo, D., Levy, N., Bardakdjian, J. and Badens, C., 2010. Neonatal screening for sickle cell disease in France: evaluation of the selective process. *Journal of Clinical Pathology.* **63**: p. 548-551.

Tietz, N., Finley, P. and Pruden, E. 1990. *Clinical Guide to Laboratory Tests.* 2nd edition, WB Saunders Company, Philadelphia.

Traeger-Synodinos, J., Harteveld, C. L., Old, J. M., Petrou, M., Galanello, R.,Giordano, P., Angastioniotis, M., De la Salle, B., Henderson, S. and May, A., 2015. EMQN Best Practice Guidelines for molecular and haematology methods for carrier identification and prenatal diagnosis of the haemoglobinopathies. *European Journal of Human Genetics*. **23:** p. 426-437.

Tukey, J.W., 1977. Exploratory data analysis. Th. MA. Reading. Downloaded from https://pdfs.semanticscholar.org.

Ueda, S., *et al.,* 1975 Routine Hb A estimation by cellulose acetate membrane electrophoresis. *Kawasaki Med. J*,. **1**: p. 113-120.

UK NEQAS Haematology Annual Report to Participants, 2011. Available on request from haem@ukneqas.org.uk

UK NEQAS Haematology Annual Report to Participants, 2015. Available on request from haem@ukneqas.org.uk

UK NEQAS Haematology Participants' Manual, 2017. https://www.ukneqash.org/documents.php (accessed January 2018)

Uldall, A., 1996. External quality assessment schemes for clinical laboratories. *Accreditation and Quality Assurance*. **1**: p. 218-222.

Urassa W., N.M., De la Salle B., Bullock D., Tholen D., Burnett D., Cognat S., Best S., Hurlston M., Kalou M. and Carter J., 2016. In *WHO manual for organizing a* national *external quality assessment programme for health laboratories and other testing sites*. World Health Organisation, Geneva, Switzerland.

US Food and Drugs Administration substantial equivalence decision summary K051072: Bio-Rad VARIANT™ nbs analyser

US Food and Drugs Administration substantial equivalence decision summary K091283:Sebia Neonat Fast™ analyser

Van Delft, P., Lenters, E., Bakker-Verweij, M., De Korte, M., Baylan, U., Harteveld, C. L., and Giordano, P C., 2009. Evaluating five dedicated automatic devices for haemoglobinopathy diagnostics in multi-ethnic populations. *International Journal of Laboratory Hematology.* **31**: p. 484-495.

Van Houcke, S.K., Rustad, P., Stepman, H.C., Kristensen, G.B., Stöckl, D., Røraas, T.H., Sandberg, S. and Thienpont, L.M., 2012. Calcium, magnesium, albumin, and total protein measurement in serum as assessed with 20 fresh-frozen single-donation sera. *Clinical Chemistry*. **58**: p. 1597-1599.

Vasikaran, S., Sikaris, K., Kilpatrick, E., French, J., Badrick, T., Osypiw, J. and Plebani, M., 2016. Assuring the quality of interpretative comments in clinical chemistry. *Clinical Chemistry and Laboratory Medicine (CCLM*). **54**: p. 1901-1911.

Vesper, H.W. and Thienpont, L.M., 2009. Traceability in laboratory medicine. *Clinical Chemistry*. **55**: p. 1067-1075.

Vives-Corrons, J.L., Pereira, M.D.M.M., Romeo-Casabona, C., Nicolás, P., Gulbis, B., Eleftheriou, A., Angastiniotis, M., Martínez, P.A., Bianchi, P., Van Wijk, R. and Heimpel, H., 2014. Recommendations for centres of expertise in rare anaemias. The ENERCA White Book. *Thalassemia Reports*. **4**.

Wald, N., & Leck, I. (Eds.), 2000. In *Antenatal and Neonatal Screening*. Oxford University Press, Oxford, UK.

Washburn, R.E., 1911. Peculiar elongated and sickle-shaped red blood corpuscles in a case of severe anemia. *Virginia Medical Semimonthly*, **15,** p.490-493.

Weatherall. D. J and Clegg, J. B. 2001. In *The Thalassaemia Syndromes* 4th edition. Blackwell Science Ltd. London.

Weatherall, D.J., 2010. The inherited diseases of hemoglobin are an emerging global health burden. Blood*.* **115**: p. 4331-4336.

Weatherall, D.J., 2011. The challenge of haemoglobinopathies in resource-poor countries. *British Journal of Haematology*. **154**: p. 736-744.

Weatherall, D. J., 2018. The Evolving Spectrum of the Epidemiology of Thalassemia. *Hematology/Oncology Clinics of North America*. **32:** p 165-175.

White, J. and Lewis, S., 1973. A report on the interlaboratory quantitation of haemoglobin A_2 and haemoglobin F. *Journal of Clinical Pathology*. **26**: p. 864-867.

Whitehead, T., Browning, D. and Gregory, A., 1973. A comparative survey of the results of analyses of blood serum in clinical chemistry laboratories in the United Kingdom. *Journal of Clinical Pathology*, **26**: p. 435-445.

Wilcken, B. and Wiley, V. 2008. Newborn screening. *Pathology*. **40**: p. 104-115.

Wild, B.J. and Bain, B.J., 2004. Detection and quantitation of normal and variant haemoglobins: an analytical review. *Annals of Clinical Biochemistry.* **41**: p. 355-369.

Wild, B.J., Green, B. N., Cooper, E. K.,Lalloz, M. R., Erten, S., Stephens, A. D. and Layton, D. M., 2001. Rapid identification of hemoglobin variants by electrospray ionization mass spectrometry. *Blood cells, molecules & diseases*. **27:** p. 691-704.

Wild, B.J., Green, B.N. and Stephens, A.D., 2004. The potential of electrospray ionization mass spectrometry for the diagnosis of hemoglobin variants found in newborn screening. *Blood cells, molecules & diseases*. **33**: p. 308-317.

Wild, B.J. and Stephens, A.D., 1997. The use of automated HPLC to detect and quantitate haemoglobins. *Clinical and laboratory haematology*. **19**: p. 171-176.

Williams, T.N., Weatherall, D.J. and Newbold, C.I., 2002. The membrane characteristics of Plasmodium falciparum-infected and-uninfected heterozygous α^0 thalassaemic erythrocytes. *British journal of haematology*. **118:** p. 663-670.

Williams, T.N., Mwangi, T.W., Roberts, D.J., Alexander, N.D., Weatherall, D.J., Wambua, S., Kortok, M., Snow, R.W. and Marsh, K., 2005a. An immune basis for malaria protection by the sickle cell trait. *PLoS medicine*. ***2*****:** p. e128.

Williams, T.N., Wambua, S., Uyoga, S., Macharia, A., Mwacharo, J.K., Newton, C.R.J.C. & Maitland, K., 2005b. Both heterozygous and homozygous α+ thalassemias protect against severe and fatal plasmodium falciparum malaria on the coast of Kenya. *Blood*. **106:** p. 368–371.

Wilrich, P-T., 2007. Robust estimates of the theoretical standard deviation to be used in interlaboratory precision experiments. *Accreditation and Quality Assurance.* **12**: p. 231-240.

Wolff, F., Cotton, F. and Gulbis, B., 2012. Screening for haemoglobinopathies on cord blood: laboratory and clinical experience. *Journal of Medical Screening*. **19**: p. 116-122.

Wong, S.K., 2005. Evaluation of the use of consensus values in proficiency testing programmes. *Accreditation and Quality Assurance.* **10**: p. 409-414.

Zeuner, D., Ades, A., Karnon, J., Brown, J., Dezateux, C. and Anionwu, E.N., 1999. Antenatal and neonatal haemoglobinopathy screening in the UK: review and economic analysis. *Technical Report. Core Research*, Alton.

Appendix 1: Example of a UK NEQAS individual performance report to participants

The report includes pages 1 to 5 of the individual participant's report for distribution 1602AH (April 2016):

Page 1: Sickle solubility testing
Page 2: Cumulative performance score for Hb A_2 and Hb S
Page 3: Hb A_2/Hb F/Hb S data analysis specimen 1602AH1
Page 4: Fraction identification results specimen 1602AH1
Page 5: Interpretive comments results specimen 1602AH1

The full report would include identical pages to p3 to p5 for the remaining specimens in the distribution. They have been omitted from this report to reduce the number of printed pages.

UK NEQAS
Haematology and Transfusion

Abnormal Haemoglobins Scheme	Laboratory:
Distribution: 1606AH Date: 28 Nov 2016	Page 1 of 11
Sickle Cell Screening	

Survey Contents:
Specimen 1606SS1 was from a sickle cell negative donation
Specimen 1606SS2 was from a sickle cell positive donation
Specimen 1606SS3 was from a sickle cell negative donation

Non Participation Penalty: 0

Specimen Quality

	1606SS1	1606SS2	1606SS3
Satisfactory	423	423	423
Unsatisfactory	1	1	1
You reported:	Satisfactory	Satisfactory	Satisfactory

Return Rate
Specimens were distributed to 441 participants.
424 participants returned results.
This represents a 96% return rate.

Sickle Screen (Note: Neg = Negative; Pos = Positive)

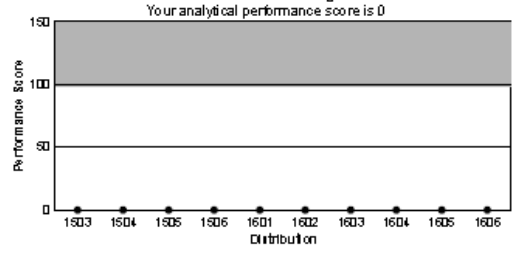
Specimen	1606SS1		1606SS2		1606SS3	
	Neg	Pos	Neg	Pos	Neg	Pos
All Methods	397	3	3	396	397	3
Atlas Medical Sickle Cell Kit	1	0	0	1	1	0
Clin Tech	42	0	1	41	41	1
Coverslip reduction	18	0	0	18	18	0
HD Supplies	21	0	0	21	21	0
Helena Biosciences Europe HbS Solubility	12	0	0	12	12	0
Lorne	2	0	0	2	2	0
Microgen Bioproducts/S-Test/M96-50	79	1	1	79	79	1
Other - please specify	54	0	1	53	54	0
Pacific Haemostasis	1	0	0	1	1	0
Solubility	37	1	0	38	37	1
Streck Sickledex	92	1	0	92	93	0
TCS Biosciences 'Sickle-Check Hb-S'	38	0	0	38	38	0

Your Registered Method	Streck Sickledex

Specimen	1606SS1	1606SS2	1606SS3
Your Results	Negative	Positive	Negative
Expected Results	Negative	Positive	Negative
Sickle Screen score	0	0	0

Sickle Screen Analytical Performance Score	0

Cumulative analytical performance
Sickle Cell Screening
Your analytical performance score is 0



Comments:

Printed at 13:21 on Wednesday, 21 December, 2016 (Final Report)

For information on data analysis and performance assessment see the UK NEQAS Haematology Participants' Manual (www.ukneqash.org)
Scheme Director: Barbara De la Salle
UK NEQAS Haematology, PO Box 14, WATFORD WD18 0FJ, UK
email: haem@ukneqas.org.uk Phone: + 44 (0)1923 217878
Authorised by: Paul McTaggart (Data Manager)

UK NEQAS Haematology and Transfusion	Abnormal Haemoglobins Scheme	Laboratory:
	Distribution: 1606AH Date: 28 Nov 2016	Page 2 of 11
Fraction Identification & Quantification	Performance Score	

Survey Contents:
Specimen 1606AH1 was from a normal adult donation

Specimen 1606AH2 was from a normal adult donation with added umbilical cord blood

Specimen 1606AH3 was from a normal adult donation

Non Participation Penalty: 0

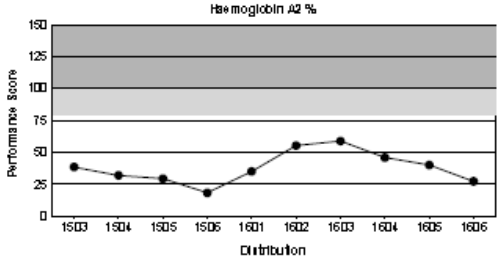
Specimen Quality

	1606AH1	1606AH2	1606AH3
Satisfactory	318	318	319
Unsatisfactory	1	1	0
You reported:	Satisfactory	Satisfactory	Satisfactory

Return Rate
Specimens were distributed to 337 participants.
319 participants returned results.
This represents a 94% return rate.

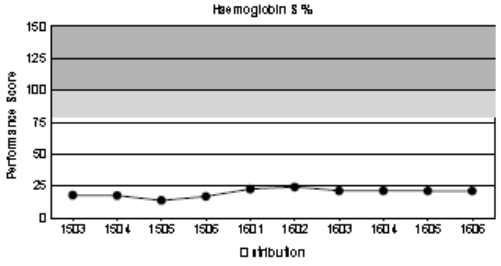
PRN:

Cumulative analytical performance
Your analytical performance score is 27.3
Haemoglobin A2 %



PRN:

Cumulative analytical performance
Your analytical performance score is 21.2
Haemoglobin S %



For information on data analysis and performance assessment see the UK NEQAS Haematology Participants' Manual (www.ukneqash.org)
Scheme Director: Barbara De la Salle
UK NEQAS Haematology, PO Box 14, WATFORD WD18 0FJ, UK
email: haem@ukneqas.org.uk Phone: + 44 (0)1923 217878
Authorised by: Paul McTaggart (Data Manager)

Fraction identification

Fraction	Expected	Essential	Your Results	Reported by all participants
Hb A	Expected	Essential	Present	309
Hb A2	Expected		Present	314
Hb F	Expected		Present	239
Hb S			Absent	2
Hb C			Absent	0
Hb D			Absent	0
Hb E			Absent	0
Hb C or E			Absent	0
Hb Non Specified Fraction			Absent	4

Performance summary for fraction identification

14 laboratories failed to report the fraction identification pattern essential for diagnosis.
Participants are asked to report all fractions present, including the expected ones (HbA, HbA2 & HbF).

Comments:

14 (3 UK) did not provide the fraction identification essential for diagnosis:
3 participants were unable to test the specimen fully:

- 1 (UK) due to analyser failure. Note that another UK laboratory did not test the specimen due to analyser failure but reported 'Hb A' on all 3 specimens in the distribution and was not included in the count of laboratories not providing the fraction identification essential for diagnosis
- 1 no longer performs fraction identification and reported the specimen as 'non-specified fraction' present
- 1 reported the specimen unsatisfactory (leaked in transit) and reported the specimen as 'non-specified fraction' present.

6 did not record Hb A; participants are reminded to record all the fractions present, including the expected ones.
1 (UK) did not submit a fraction identification result for the specimen
4 reported Hb fractions in addition to the expected ones:

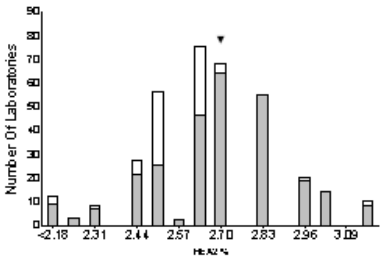
- 2 reported an additional Hb S fraction
- 2 reported an additional non-specified (NS) fraction. 1 (UK) reported interpretive coded comment 419 (Hb variant present of no known clinical significance) and the other reported interpretive coded comment 400 (no evidence of a haemoglobin variant or thalassaemia)

For information on data analysis and performance assessment see the UK NEQAS Haematology Participants' Manual (www.ukneqash.org)
Scheme Director: Barbara De la Salle
UK NEQAS Haematology, PO Box 14, WATFORD WD18 0FJ, UK
email: haem@ukneqas.org.uk Phone: + 44 (0)1923 217878
Authorised by: Paul McTaggart (Data Manager)

Fraction Quantitation

Haemoglobin A2 (%)

	n	Mean	GCV
All Methods	350	2.6	7.40
Capillary Electrophoresis	69	2.5	3.20
Sebia Capillarys	24	2.5	2.55
Sebia Capillarys 2	30	2.6	3.01
Sebia Minicap	14	2.5	4.59
HPLC	273	2.7	7.63
Arkray HA-8180T	17	2.7	6.49
Arkray HA8160	24	2.6	12.08
Bio Rad D10 ; Dual Program Kit	25	2.9	7.23
Bio Rad Variant II; Beta-thal short pro	80	2.7	4.19
Bio Rad Variant II; Dual program Kit	28	2.7	7.62
Primus Ultra 2	20	2.4	6.19
TOSOH G8	49	2.6	7.53

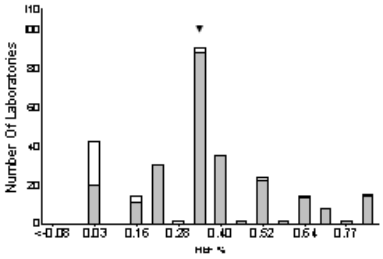


Your registered method is:
HPLC BioRad Variant II;
Beta-thal short program

Your Result :	2.7
DI :	0.13
Uncertainty of Method Mean :	0.08
Perf Score :	27.3
Reported Range (Overall)	
Minimum	0.00
Maximum	206.00
Assessment vs your ref range	
You reported:	Normal
Overall Assessment (%)	
Low	0.6
Normal	99.1
High	
Uncertain	0.3

Haemoglobin F (%)

	n	Mean	GCV
All Methods	235	0.30	57.17
HPLC	225	0.30	54.41
Arkray HA8160	23	0.40	52.99
Bio Rad D10 ; Dual Program Kit	12	0.60	56.15
Bio Rad Variant II; Beta-thal short pro	77	0.30	18.88
Bio Rad Variant II; Dual program Kit	20	0.40	47.42
TOSOH G8	50	0.30	54.92



Your registered method is:
HPLC BioRad Variant II;
Beta-thal short program

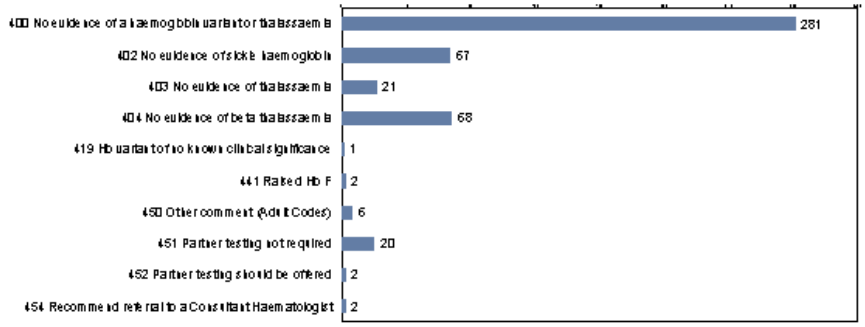
Your Result :	0.3
Reported Range (Overall)	
Minimum	0.00
Maximum	2.30
Assessment vs your ref range	
You reported:	Low
Overall Assessment (%)	
Low	2.2
Normal	96.2
High	1.3
Uncertain	0.3

For information on data analysis and performance assessment see the UK NEQAS Haematology Participants' Manual (www.uknegash.org)
Scheme Director: Barbara De la Salle
UK NEQAS Haematology, PO Box 14, WATFORD WD18 0FJ, UK
email: haem@uknegas.org.uk Phone: + 44 (0)1923 217878
Authorised by: Paul McTaggart (Data Manager)

Interpretation

Sex	Female	RBC	(10^{12}/L)	4.39
Ethnic Origin	Turkish	Hb	(g/L)	136
Age	18 Yrs	MCV	(fL)	90.6
	Sister has been confirmed as a beta thalassaemia carrier	MCH	(pg)	31.0

Analysis of Interpretation Codes reported by Participants



Data Analysis

Top five reported comments (see graph for all reported comments)

Code	Comment	Rank	Number
400	No evidence of a haemoglobin variant or	1	281
404	No evidence of beta thalassaemia	2	68
402	No evidence of sickle haemoglobin	3	67
403	No evidence of thalassaemia	4	21
451	Partner testing not required	5	20

Reported Comments

Your reported comments with the number of participants that reported the same comment

Code	Comment	Rank	Number

Comments:

Specimen 1606AH1 simulated a sample from a 18 years' old woman of Turkish origin, tested as part of family screening, as her sister had been confirmed as a beta thalassaemia carrier. Her blood count was not suggestive of thalassaemia, there was no evidence of a haemoglobin variant present, the all methods trimmed mean Hb A2 result was 2.6% and the Hb F concentration was normal.

281 (95%) of 295 participants who returned an interpretation for this 'patient' used code 400 (no evidence of a haemoglobin variant or thalassaemia), 5 (1.6%) used both code 402 (no evidence of sickle haemoglobin) and 404 (no evidence of beta thalassaemia) and another 7 (2.4%) used code 404 alone. One used code 403 (no evidence of thalassaemia).

One (UK) laboratory reported code 419 (Hb variant present of no known clinical significance) with code 402 (no evidence of sickle haemoglobin) but did not comment on the thalassaemia status.

For information on data analysis and performance assessment see the UK NEQAS Haematology Participants' Manual (www.ukneqash.org)
Scheme Director: Barbara De la Salle
UK NEQAS Haematology, PO Box 14, WATFORD WD18 0FJ, UK
email: haem@ukneqas.org.uk Phone: + 44 (0)1923 217878
Authorised by: Paul McTaggart (Data Manager)

Appendix 2: Instruction sheet and proforma results sheet for experimental exercise 1301NHEX

UK NEQAS United Kingdom National External Quality Assessment Scheme for General Haematology

SURVEY 1301NHEX1: Low Hb A% Exercise, 3 December 2013

CLOSING DATE:

PRN:

1.0 Distribution Package

This package contains survey and comprises a white postal bag containing documentation and a sealed clear plastic bag holding eight newborn screening blood test cards.

In the event of any query, please contact Barbara De la Salle (barbara.delasalle@whht.nhs.uk or 01923 217878).

2.0 Information required for Control of Substances Hazardous to Health (COSHH)

The specimens have been prepared from mixtures of umbilical cord blood and adult blood. Handle the specimens according to the protocol in force in your laboratory for the safe handling of patient samples (Health Notice (86) 25 Annex 2.5). *This information should be reviewed by your COSHH assessor for consideration of any changes necessary to your local work practices*

3.0 Use of packaged material

This material is for use in External Quality Assessment Surveys only.

4.0 Newborn Sickle Cell Screening Survey 13NHEX1

Your laboratory has been notified by email of the exercise. You are provided with eight specimens, -1 to -8, as newborn screening dried blood spot test cards, each containing one drop of blood. There are no patient details with these specimens.

The purpose of this exercise is to assess the measurement of low levels of Hb A% in dbs specimens. You are asked to test each specimen for Hb A%, if it is produced by your instrument, even if you do not report the Hb A% routinely.

Because these are simulated specimens, you may see a minor peak corresponding to Hb A_2, which should be ignored. There are no abnormal haemoglobin fractions in the specimens.

This exercise is non-scoring and no performance assessment of any individual laboratory will be made.

Specimen Handling and Disposal

- Store cards at 2 - 8 °C until tested.
- Test the specimens by your standard protocol for newborn sickle screening.
- Handle and dispose of the specimens as patient material.

5.0 Return of Results

- Specimen quality: indicate the reason for any unsatisfactory comments in the comments box. Please test the specimens, even if the quality does not correspond to that expected of a patient's sample.
- Report the Hb A% to one decimal place on the enclosed results sheet, together with the details of the instrument used.
- Return results sheets by fax (+44 (0)1923 217879) using a header sheet marked for the attention of Barbara De la Salle, by first class post (UK NEQAS (H), PO BOX 14, Watford WD18 0FJ) or scanned and returned by email to barbara.delasalle@whht.nhs.uk.
- We would be grateful if you could also send a copy of the output (e.g. the chromatogram) from your instrument for each specimen.
- A summary report will be sent to your registered laboratory contact by the end of January 2014.

6.0 Date of next survey: Should a follow up survey be undertaken, you will be notified.

Organiser: Professor K Hyde, UK NEQAS (H), PO Box 14, Watford WD18 0FJ, UK
Instruction sheet, exercise 13NHEX1

Appendix 2.1. Sample instruction sheet for exercise 1301NHEX

UK NEQAS NEWBORN SICKLE SCREENING
LOW Hb A % EXERCISE 13NHEX1

PRN	HOSPITAL

DATE RECEIVED	DATE TESTED

INSTRUMENT MAKE AND MODEL

CARD	1	2	3	4	5	6	7	8
QUALITY (please tick)								
Satisfactory								
Unsatisfactory								
RESULTS (to one decimal place)								
Hb A%								

COMMENTS (only if needed)

Please return by December 20, 2013:
By FAX: 01923 217879
By POST: UK NEQAS (H), PO BOX 14, WATFORD, WD18 0FJ
By EMAIL: barbara.delasalle@whht.nhs.uk
We would be grateful for a copy of the output from your analyser (e.g. chromatogram) for each specimen

Appendix 2.2. Example Results sheet for exercise 1301NHEX

Appendix 3: UK NEQAS Haematology coded interpretive comments list for the Abnormal Haemoglobins survey

UK NEQAS Haematology
CODED COMMENTS FOR ABNORMAL HAEMOGLOBINS SURVEY

Code	Comment
400	No evidence of a haemoglobin variant or thalassaemia
402	No evidence of sickle haemoglobin
403	No evidence of thalassaemia
404	No evidence of beta thalassaemia
411	Sickle cell carrier (Hb AS)
412	Hb C carrier (Hb AC)
413	Hb D carrier (Hb AD)
414	Hb E carrier (Hb AE)
415	Hb variant carrier *(state the variant in the comment box, if known)*
416	Hb A_2 variant detected. This is of no clinical importance.
417	Hb O^{Arab} carrier (Hb AO^{Arab})
418	Hb Lepore carrier (Hb A/Lepore)
419	Hb variant carrier of no known clinical significance *(state the variant in the comment box, if known)*
421	Hb SC disease (Hb SC)
422	Hb SD disease (Hb SD)
423	Hb SE disease (Hb SE)
424	Hb SO^{Arab} disease (Hb SO^{Arab})
425	Homozygous sickle cell anaemia (Hb SS)
426	Sickle – beta thalassaemia disease
431	Beta thalassaemia carrier
432	Possible alpha plus thalassaemia carrier and/or iron deficiency
433	Possible alpha zero thalassaemia carrier and/or iron deficiency
434	Carrier of Hereditary Persistence of Fetal Haemoglobin (HPFH)
435	Carrier of delta beta thalassaemia
441	Raised HbF
451	Partner testing not required
452	Partner testing should be offered
453	Partner testing should be offered if he is from a high-risk area *(for alpha zero thalasaemia)*
454	Recommend referral to a Consultant Haematologist
455	Iron status should be checked
450	Other comment (please specify in comment box)

Appendix 4: Publications

De La Salle, B., Mosca, A., Paleari, R., Rapanakis, V. and Hyde, K., 2013. Laboratory diagnosis of the rare anaemias: external quality assessment benefits patient care. *Thalassemia Reports*, *3*(1s), p.31.

http://109.68.152.74/index.php/thal/article/view/thal.2013.s1.e31.