Turun yliopisto
University of Turku

QUANTITATIVE MAGNETIC RESONANCE IMAGING OF THE LIVER AND HEART IN IRON OVERLOAD

Methodology and the role of iron excess in hematological malignancies and allogeneic stem cell transplantation

Johanna Virtanen

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
Sarja - ser. D osa - tom. 1139 | Medica - Odontologica | Turku 2014

University of Turku

Faculty of Medicine

Department of Radiology

Medical Imaging Centre of Southwest Finland, Turku University Hospital

Turku Doctoral Programme of Clinical Sciences

Supervised by

Professor Riitta Parkkola,
Department of Radiology
University of Turku
Turku, Finland

Reviewed by

Docent Nina Lundbom
Department of Radiology
University of Helsinki
Helsinki, Finland

Docent Kaarina Partanen
University of Kuopio
Docrates Cancer Center
Helsinki, Finland

Opponent

Professor Osmo Tervonen
Department of Radiology
University of Oulu
Oulu, Finland

ISBN 978-951-29-5877-1 (PRINT)
ISBN 978-951-29-5878-8 (PDF)
ISSN 0355-9483
Painosalama Oy - Turku, Finland 2014

To my dear family and friends

“Gold is for the mistress – silver for the maid –
Copper for the craftsman cunning at his trade.”
"Good!" said the Baron, sitting in his hall,
"But Iron – Cold Iron – is master of them all."

Cold Iron
Rudyard Kipling

ABSTRACT

Johanna Virtanen

QUANTITATIVE MAGNETIC RESONANCE IMAGING OF THE LIVER AND HEART IN IRON OVERLOAD
Methodology and the role of iron excess in hematological malignancies and allogeneic stem cell transplantation

Department of Radiology, University of Turku, Turku University Hospital, Turku, Finland

Annales Universitatis Turkuensis
Painosalama Oy, Turku, Finland 2014

Systemic iron overload (IO) is considered a principal determinant in the clinical outcome of different forms of IO and in allogeneic hematopoietic stem cell transplantation (alloSCT). However, indirect markers for iron do not provide exact quantification of iron burden, and the evidence of iron-induced adverse effects in hematological diseases has not been established.

Hepatic iron concentration (HIC) has been found to represent systemic IO, which can be quantified safely with magnetic resonance imaging (MRI), based on enhanced transverse relaxation. The iron measurement methods by MRI are evolving.

The aims of this study were to implement and optimise the methodology of non-invasive iron measurement with MRI to assess the degree and the role of IO in the patients. An MRI-based HIC method (M-HIC) and a transverse relaxation rate (R2*) from M-HIC images were validated. Thereafter, a transverse relaxation rate (R2) from spin-echo imaging was calibrated for IO assessment. Two analysis methods, visual grading and rSI, for a rapid IO grading from in-phase and out-of-phase images were introduced. Additionally, clinical iron indicators were evaluated. The degree of hepatic and cardiac iron in our study patients and IO as a prognostic factor in patients undergoing alloSCT were explored.

In vivo and *in vitro* validations indicated that M-HIC and R2* are both accurate in the quantification of liver iron. R2 was a reliable method for HIC quantification and covered a wider HIC range than M-HIC and R2*. The grading of IO was able to be performed rapidly with the visual grading and rSI methods. Transfusion load was more accurate than plasma ferritin in predicting transfusional IO. In patients with hematological disorders, the prevalence of hepatic IO was frequent, opposite to cardiac IO. Patients with myelodysplastic syndrome were found to be the most susceptible to IO. Pre-transplant IO predicted severe infections during the early post-transplant period, in contrast to the reduced risk of graft-versus-host disease. Iron-induced, poor transplantation results are most likely to be mediated by severe infections.

Key words: hematological malignancy, allogeneic stem cell transplantation, hepatic iron concentration, liver iron overload, cardiac iron overload, transverse relaxation, magnetic resonance imaging.

TIIVISTELMÄ

Johanna Virtanen

MAKSAN JA SYDÄMEN MAGNEETTIKUVAUS ELIMISTÖN RAUTAKUORMAN KVANTITOINNISSA
Menetelmät ja rautaylimäärän merkitys veritaudeissa ja kantasolusiirrossa

Diagnostinen radiologia, Turun Yliopisto, Turun Yliopistollinen Keskussairaala, Turku, Suomi

Annales Universitatis Turkuensis
Painosalama Oy, Turku, Finland 2014

Elimistön kokonaisrautakuormaa pidetään kliinisen lopputuloksen tärkeimpänä tekijänä eri raudan kertymäsairauksissa ja kantasolusiirroissa. Epäsuorilla rautakuorman markkereilla ei saada kuitenkaan tarpeeksi luotettavasti mitattua rautaylimäärää eikä rautakuorman haitallisuutta hematologisissa tautitiloissa ole tähän mennessä varmasti osoitettu.

Maksan rautapitoisuus kuvastaa hyvin elimistön rautakuormaa. Se on mahdollista mitata maksan magneettikuvauksella perustuen raudan aiheuttamaan poikittaisen relaksaation tehostumiseen.

Tämän tutkimuksen tarkoituksena oli analysoida ja kehittää kajoamattomia magneettikuvaukseen perustuvia kliinisiä sovelluksia rautakuorman mittaamiseksi, määrittää rautakuorman astetta ja arvioida sen merkitystä. Otimme käyttöön ja kehitimme edelleen maksan rautapitoisuuden magneettikuvausmenetelmää (M-HIC). Validoimme M-HIC menetelmästä lasketun poikittaisen relaksaationopeuden (R2*) ja kalibroimme spin-kaikukuvauksella saadun poikittaisen relaksaationopeuden (R2) maksan rautapitoisuuteen. Kehitimme kaksi uutta analysointitapaa (visual grading ja rSI) rautakuorman nopeaksi analysoimiseksi magneettikuvista. Arvioimme myös kliinisiä raudan markkereita. Määritimme validoiduilla menetelmillä maksan ja sydämen rautakuorman astetta ja arvioimme rautakuorman aiheuttamia haittavaikutuksia hematologisilla kantasolusiirtopotilailla.

Tutkimuksessa totesimme että M-HIC ja R2* ovat luotettavia menetelmiä maksan rautapitoisuuden määrittämisessä. Kalibroimalla R2 pääsimme mittaamaan luotettavasti maksan rautapitoisuutta laajemmalta määritysalueelta. Rautamäärän arviointi onnistui molemmilla kehitetyillä rautakuorman nopeilla arviointitavoilla magneettikuvista. Potilaan saamien punasoluyksiköiden määrä ennusti verensiirroista aiheutunutta rautakuormaa paremmin kuin ferritiini. Veritauteja sairastavilla potilailla maksan rautaylimäärä oli yleistä, kun taas sydämen rautaylimäärä harvinaista. Rautakuorma kertyi herkimmin myelodysplastista syndroomaa sairastaville potilaille. Korkea rautakuorma näytti lisäävän vaikeita infektioita heti kanta-solusiirron jälkeen, mutta vähentävän käänteishyljintää. Näin ollen raudan haittavaikutukset kantasolusiirrossa välittyvät todennäköisimmin lisääntyneen infektioherkkyyden kautta.

Avainsanat: veritaudit, kantasolusiirto, maksan rautapitoisuus, sydämen rautakuorma, maksan rautakuorma, poikittainen relaksaatio, magneettikuvaus.

TABLE OF CONTENTS

ABBREVIATIONS

AL	Acute leukemia
ALL	Acute lymphoblastic leukemia
alloSCT	Allogeneic hematopoietic stem cell transplantation
AML	Acute myeloid leukemia
AUC	Area under curve
B-HIC	Biochemical hepatic iron concentration
CI	Confidence interval
CLL	Chronic lymphoblastic leukemia
CML	Chronic myeloid leukemia
COR	Cumulative odds ratio
CT	Computed tomography
FA	Flip angle
GVHD	Graft-versus-host disease
HDS	Number of high dose sessions
HDU	Number of high dose units
HIC	Hepatic iron concentration
HL	Hodgkin's lymfoma
HSCT	Hematopoietic stem cell transplantation
IO	Iron overload
IQR	Interquartile range
LPI	Labile plasma iron
LIP	Labile iron pool
MAC	Myeloablative conditioning
M-HIC	Magnetic resonance imaging-based hepatic iron concentration, SI-derived method
M-HIC(R2)	Magnetic resonance imaging-based hepatic iron concentration measured with R2
M-HIC(R2*)	Magnetic resonance imaging-based hepatic iron concentration measured with R2*
M-PIC	Magnetic resonance imaging-based phantom iron concentration, SI-derived method
MPN	Myeloproliferative neoplasms
MR	Magnetic resonance
MRI	Magnetic resonance imaging
MUD	Matched unrelated donor
NHL	Non-Hodgkin's lymphoma

NMR	Nuclear magnetic resonance
NMV	Net magnetisation vector
NPV	Negative predictive value
NTBI	Non-transferrin-bound iron
OR	Odds ratio
OS	Overall survival
PPV	Positive predictive value
R2	Transverse relaxation rate (1/s) from spin echo
R2*	Transverse relaxation rate (1/s) from gradient echo
RARS	Refractory anemia with ringed sideroblasts
RBC	Red blood cell
RE	Reticuloendothelial
RF	Radio frequency
RIC	Reduced intensity conditioning
RIT	Reduced intensity of toxicity
ROS	Reactive oxygen species
ROC	Receiver-operator characteristic
rSI	Relative SI
SAA	Severe aplastic anemia
SD	Standard deviation
SI	Signal intensity
SNR	Signal-to-noise ratio
SQUID	Superconducting quantum interference device
T	Tesla
T1	Longitudinal relaxation time (ms) from spin echo
T2	Transverse relaxation time (ms) from spin echo
T2*	Transverse relaxation time (ms) from gradient echo
TE	Time to echo
TR	Time of repetition
TRM	Transplant-related mortality

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following original publications, which are referred to in the text by Roman numerals I-IV. Additional unpublished data analyses are also presented referred to by a symbol ([a]).

I Virtanen, J.M., Komu, M.E., Parkkola, R.K. 2008. Quantitative liver iron measurement by magnetic resonance imaging: in vitro and in vivo assessment of the liver to muscle signal intensity and the R2* methods. Magn Reson Imaging 26 (8): 1175-1182.

II Virtanen, J.M., Pudas, T.K., Ratilainen, J.A., Saunavaara, J.P., Komu, M.E., Parkkola, R.K. 2012. Iron overload: accuracy of in-phase and out-of-phase MRI as a quick method to evaluate liver iron load in haematological malignancies and chronic liver disease. Br J Radiol 85 (1014): e162-167. Epub 2011 Mar 8.

III Virtanen, J.M., Remes, K.J., Itälä-Remes, M.A., Saunavaara, J.P., Komu, M.E., Partanen, A.M., Parkkola, R.K. 2012. The relationship between cardiac and liver iron evaluated by MR imaging in haematological malignancies and chronic liver disease. Blood Cancer Journal 2(1): e49.

IV Virtanen, J.M., Remes, K.J, Itälä-Remes, M.A., Saunavaara, J.P., Vahlberg, T., Sinisalo, M., Parkkola, R.K. 2013. Prognostic impact of pretransplant iron overload measured with magnetic resonance imaging on severe infections in allogeneic stem cell transplantation. Eur J Haematol 91(1): 85-93. Epub 2013 May 31.

1. INTRODUCTION

Iron is an essential metal in the body due to its ability to accept and donate electrons between ferrous (Fe^{2+}) and ferric (Fe^{3+}) forms and its key role in various biochemical reactions in respiratory chain, cell growth and multiplication. However, iron excess can be toxic to the host and lead to tissue damage (Evens et al., 2004; Kohgo et al., 2008; Majhail et al., 2008). In a physiologically balanced state, circulating iron is bound to transferrin, which protects the body from the toxic sequels of iron, as no or minimal free, more unstable iron is available. When the binding capacity of transferrin is exceeded, non-transferrin-bound-iron (NTBI); i.e., free iron appears, and the toxic ferrous form of iron can catalyse the conversion of hydrogen peroxide into highly toxic free radicals, such as hydroxyl. Iron overload (IO) can cause organ dysfunction and damage through the production of these reactive oxygen species (ROS) and oxidation of proteins, peroxidation of membrane lipids and modification of nucleic acid. These toxic iron-induced effects are thought to eventually lead to organ dysfunction and fibrosis (Evens et al., 2004; Majhail et al., 2008).

Systemic IO has been considered a principal determinant of clinical outcome in all forms of iron overload. Although there is a wide variety of hematological malignancies, where red blood cell (RBC) transfusions lead to transfusional IO (Rose et al., 2007; Di Tucci et al., 2008), the studies have mainly been conducted in patients with thalassemia. In these patients, IO has best been characterised and cardiac IO has been the leading cause of mortality (Olivieri et al., 1994; Brittenham and Badman, 2003; Leitch, 2011). Indeed, iron-chelation for reducing IO is a recommended treatment for thalassemia, but there are still controversies about chelation in myelodysplastic syndrome (MDS) and in other hematological diseases (Leitch 2011). Among hematological disorders, iron-induced, adverse effects have been suggested to occur in patients with MDS (Fenaux and Rose, 2009; Cuijpers et al., 2010; Leitch, 2011).

Despite the considerable post-transplant morbidity and mortality in hematopoietic stem cell transplantation (HSCT) recipients, transplantation has been used increasingly as a curative treatment for hematological malignancies. This treatment has variable post-transplant complications. Infections and graft-versus-host disease (GVHD) are the two main causes of transplant-related mortality (TRM). The poor prognosis after HSCT is thought to be associated with IO, although the causality is not established (Evens et al., 2004; Majhail et al., 2008; Kataoka et al., 2009). IO in these patients is a consequence of frequent RBC transfusions, but it is also thought to be related to ineffective erythropoiesis and possibly to other transfusion-related factors, such as, destruction of bone marrow and tumour cells resulting in iron release from these tissues (Evens et al., 2004; Majhail et al., 2008). Indeed, free iron in blood; i.e., labile plasma iron (LPI) has been found to increase tremendously during HSCT (Sahlstedt et al., 2001; Yegin et al., 2011). Because the impact of IO in HSCT is still not established, future prospective works are needed to assess, not the role of ferritin, but of iron burden itself in transplantation-related mortality and co-morbidities (Kamble and Mims, 2006; Majhail et al., 2008; Koreth and Antin, 2010; Malcovati and Cazzola 2011).

These unanswered questions about the degree and the role of IO in hematological malignancies prompt research and development of novel methods for evaluating the causality between iron and its potential adverse effects. Iron load measurement is also required for the evaluation and monitoring of iron-chelation treatment. Thus, different methods for IO quantification have been developed. Many of those have shown to be either non-specific (for example ferritin) or semi-quantitative (for example, histological analysis of bone marrow or liver) and invasive, as liver biopsy.

Magnetic resonance imaging (MRI) has become the choice of IO measurement. It offers a safe, non-invasive method for IO quantification. The modifications for liver and cardiac iron measurement by MRI have been evolving towards an optimal non-invasive method, which would be robust, accurate, have a wide quantitative range and should also be available for the patients (Brittenham and Badman, 2003; Alústiza et al., 2007, Wood, 2007; Wood and Ghugre, 2008, Koreth and Antin, 2010). The measurement of iron concentration by MRI is based on the paramagnetic feature of iron, which shortens the transverse relaxation time in the proportion of iron concentration. Thus, an increase in the iron concentration darkens proportionally MR images (Wood, 2007).

The initiative behind this work rose from the clinical need for a non-invasive hepatic iron concentration (HIC) measurement method in patients with hematological malignancies. Gandon's work with the calibration of MRI method for liver biopsies (Gandon et al., 2004) was performed with a similar MRI scanner that we used in our institution. This facilitated the method transfer into our centre and initiated an idea of validating and introducing this quantitative liver imaging method for a reliable HIC measurement method, magnetic resonance imaging based hepatic iron concentration (M-HIC). After the validation of M-HIC and the transverse relaxation rate (R2*) derived from the M-HIC images, we began to evaluate the use of transverse relaxation rate (R2) method from spin echo imaging, which enables measuring-higher HIC values. The calibration of R2 and consequent measurement with the wider quantitative range was required in the clinical studies in hematological patients with severe IO. Because the use of the different quantitative MRI methods needs special expertise, it would be beneficial if rough detection of IO would be possible from routine abdominal imaging. This would be of importance worldwide, especially in the centres, which treat iron-overloaded population, but are lacking the definite quantitative MRI-based methods. Thus, we evaluated the capability of in-phase and out-of phase-imaging, in addition to clinical parameters, to detect IO. Finally, the measurement of IO from the liver by R2 method in seriously ill patients before allogeneic hematopoietic stem cell transplantation (alloSCT) was conducted in a follow-up study, where the clinical significance of high IO on the post-transplant outcome was assessed.

2. REVIEW OF THE LITERATURE

2.1. IRON METABOLISM

2.1.1. Iron balance

Iron is one of the most important compounds in human metabolism. It is bound to larger molecules, mainly hemoglobin, transferrin, ferritin or hemosiderin, because of the toxicity and precipitation of free iron. The total human body iron content is approximately three to four grams. In physiological conditions, most of it (70%) is incorporated in the hemoglobin of red blood cells; i.e., erythroid precursors and matured RBC. The remainder is stored in the liver, spleen and muscles, in hepatocytes and reticuloendothelial (RE) macrophages as ferritin or hemosiderin. The circulating proportion of iron is approximately 0.05 per cent of the total body iron and bound to transferrin. The daily loss of iron is rather small, only one to two mg in adults (<0.1% of total body iron), and is replaced from dietary sources to maintain iron balance (Andrews, 1999; Kohgo et al., 2008; Majhail et al., 2008; Brissot et al., 2011; Ganz and Nemeth, 2012).

2.1.2. Iron cycle

The dietary iron absorption is only a minor source for plasma iron. It is absorbed from the diet in the intestinal tract, mainly in the duodenum. Both non-heme iron from dietary plants and heme iron from meat are absorbed by enterocytes. In the small bowel wall, iron within enterocytes is released into the circulation *via* a metal transporter, ferroportin in the ferrous form. Excreted ferrous iron is thereafter oxidised to ferric, by hephaestin, a homologue of ceruloplasmin. Then, the circulating ferric iron is bound to serum transferrin.

Most of the iron for the body needs is recycled from senescent RBC. The degradation of RBC is done mainly within the spleen, and to a lesser degree within the liver, in approximately 120 days after production of the cell. It is done by RE macrophages, which scavenge iron from hemoglobin and load it onto transferrin for reuse. Then, transferrin delivers its iron to cells *via* an interaction with transferrin receptor 1 found on the plasma membrane of most cells. Intracellular iron is used in many enzymatic reactions, particularly within the respiratory chain. The transferrin-bound iron is transferred and utilised mainly for bone marrow erythropoiesis or is alternatively stored in hemosiderin and ferritin (Andrews, 1999; Kohgo et al., 2008; Brissot et al., 2011). Normally, the same amount of iron that is lost daily from the skin, the intestine and, in women, through menstruation, is also absorbed. Humans do not have any active mechanism for iron excretion. (Andrews, 1999; Fleming and Bacon, 2005; Kohgo et al., 2008; Majhail et al., 2008; Brissot et al., 2011).

2.1.3. Iron regulation

Systemic iron balance is a semi-closed system and tightly regulated. Plasma iron concentration is normally maintained stable despite fluctuations in diet and iron losses

in order to maintain adequate but non-toxic amounts of iron stores (Majhail et al., 2008; Ganz, 2011; Ganz and Nemeth, 2012). Because of the lack of mechanism to excrete excess iron, iron homeostasis is dependent on the regulatory feedback between the body's iron needs and interstitial iron absorption and distribution (Fleming and Bacon, 2005).

In the last decade, there has been growing evidence of newly identified genes and mechanisms involved in iron homeostasis (Brissot et al., 2011; Ganz and Nemeth, 2012). A new regulatory compound hepcidin was found. Hepcidin is a peptide that is transcribed in the liver and has a central role in iron regulation. Hepcidin controls iron level and distribution with negative feedback by inducing degradation of its receptor, the cellular iron exporter ferroportin. During iron overload, hepcidin transcription and secretion from hepatocytes increase, and inhibit intestinal absorption and iron release into the plasma, whereas its expression is markedly decreased during iron deficiency. It is produced in the liver and secreted by hepatocytes and it acts through three pathways. Hepcidin interacts with ferroportin, and inhibits iron export from: 1) recycling macrophages of spleen and liver, 2) bowel wall enterocytes and 3) liver hepatocytes. Thus, hepcidin lowers plasma iron concentration (Ganz, 2011; Ganz and Nemeth, 2012). This 25-amino acid peptide was first described by Park et al. and Krause et al, independently (Krause et al., 2000; Park et al., 2001). The peptide was named hepcidin due to the site of hepatic synthesis (hep-) and the antibacterial properties of this peptide (-cidin) (Park et al., 2001; Ganz, 2003).

2.1.4. The role of liver

The liver is a central organ in iron metabolism, storage and also in regulation; i.e., determining iron distribution, plasma concentration and release from the storage *via* the feedback system of hepcidin (Fleming and Bacon, 2005; Ganz and Nemeth, 2012).

The liver is the major storage organ of excess iron and liver iron content is regarded as an accurate measure of systemic iron content (Angelucci et al., 2000). Thus, the liver has become the most important target for direct iron content measurement. Excess iron in the liver is stored in hepatocytes and RE Kupffer cells together with macrophages, as ferritin and hemosiderin, in order to keep iron more soluble and in a non-toxic form (Massover, 1993; Andrews, 1999; Kohgo et al., 2008). Ferritin is a large metalloprotein, which can load and unload Fe atoms as a biomineral form into its core. The apoprotein, apoferritin consists of 24 polypeptide subunits and forms a spherical shell of 13 nm in diameter. Iron itself is stored in the centre of ferritin, in an 8 nm cavity, which consists of a ferrihydrite mineral core ($5Fe_2O_3$-$9H_20$) (Massover, 1993; Gossuin et al., 2004). Hemosiderin is a large metalloprotein cluster of spicular crystals with a variable size of less than 1 nm to over 20 nm. It is a partially denatured form of ferritin, more heterogeneous and less soluble than ferritin and appears especially in pathologically iron-overloaded cells (Iancu, 1992; Gossuin et al., 2004).

More recently, the liver has also been found to determine how much iron is absorbed and released from storage sites into plasma (Kohgo et al., 2008; Majhail et al., 2008; Ganz, 2011). Although there are several molecules involved in the body's iron

metabolism: in iron absorption, bone marrow iron uptake, reutilisation of senescent RBC and iron storage (Fleming and Bacon, 2005), only the liver-derived hepcidin has been found to directly regulate body iron level. Hepcidin synthesis in the liver is transcriptionally regulated by three main components. Firstly, the hepcidin levels are regulated by extracellular and intracellular iron concentrations. Secondly, through yet undefined pathways, hepcidin is also homeostatically regulated by the iron requirements of erythroid precursors for hemoglobin synthesis, ertyhropoesis. Thirdly, the hepcidin synthesis is regulated by inflammation. This is related to the host defence mechanism, where iron deprivation is used as a defence mechanism in infections (Fleming and Bacon, 2005; Kohgo et al., 2008; Majhail et al., 2008; Ganz and Nemeth, 2012).

2.2. PATHOPHYSIOLOGY OF IRON OVERLOAD

Both excess and deficiency of iron can cause cellular damage and organ dysfunction. Disorders of iron homeostasis are the most common diseases in humans. Low plasma iron causes anemia, as hemoglobin synthesis is limited. On the other hand, excess iron can lead to cell and tissue damage, which is thought to be mediated by the iron-catalysed generation of ROS and also more recently by impaired immunity (Ganz and Nemeth, 2012).

Circulating iron is bound to transferrin and other binding proteins, and no or very little free plasma iron is normally present. However in IO, when high plasma iron levels exceed transferrin binding capacity, NTBI occurs (Ganz and Nemeth, 2012). The fraction of NTBI, labile plasma iron (LPI), is considered to be most reactive and toxic. Hepatocytes have essentially two pathways for uptake of iron from the circulation: transferrin-bound iron mainly in physiological concentrations and NTBI in excess of iron (Kohgo et al., 2008). NTBI is avidly taken up by hepatocytes causing hepatic IO. Cardiac and endocrine iron uptake is thought to need a more rapid increase in iron concentration than hepatic iron accumulation (Ganz and Nemeth, 2012). Within the tissue, mainly in the liver, iron is accumulated either in ferritin or a labile iron pool (LIP). Ferritin is post-transcriptionally upregulated by the excess of iron and bounds iron in non-toxic form, while LIP is biologically active. Most of LIP consists of ferric (Fe^{3+}) iron-bound complexes with other plasma proteins and organic anions, such as citrate or adenosine diphosphate, and a small but important LIP proportion is reduced to the ferrous (Fe^{2+}) form (Kohgo et al., 2008; Pullarkat, 2010; Ganz and Nemeth, 2012).

2.2.1. Oxidative stress

Because iron can donate and accept electrons between two oxidation states: reduced ferrous (Fe^{2+}) and oxidised ferric (Fe^{3+}) forms, it is able to take part in various biochemical reactions involving oxygen transport and electron transfer, essential to biological organisms. This feature also makes iron highly toxic and may lead to tissue damage by oxidation of proteins, peroxidation of membrane lipids and structural changes to nucleic acids (Pullarkat 2010, Majhail, Lazarus, and Burns 2008). Iron, in

its ferrous form, is considered to be a mediator of ROS and tissue injury. Fe^{2+} ions in aqueous solution oxidise and form Fe^{3+}. Already in 1894, Fenton first discovered that ferrous iron could catalyse the oxidation of tartrate by hydrogen peroxide into toxic free radicals, although the importance of ROS was recognised much later. Nowadays, this is thought to be the principle explanation for the damage that iron generated by ROS in living cells. The reaction of iron-induced free radical formation is known as Fenton's reaction:

$Fe^{2+} + H_2O_2$ ---- > $Fe^{3+} + OH^{\bullet} + OH^-$,

where H_2O_2 is hydrogen peroxide, $OH^{\bullet}$ is hydroxyl radical and OH^- is hydroxide. Especially the hydroxyl radical is capable of causing damage to biomolecules, for example, DNA. Other ROS, such as hydrogen peroxide, are less reactive, but are also thought to cause pro-oxidant state and cell damage (Evens et al., 2004).

2.2.2. Innate immunity

IO associates with increased susceptibility to different infections. An association between excess iron and infections have been found both in RBC transfusion-dependent and in hereditary IO (Andrews, 1999; Evens et al., 2004). For example, *Pseudomonas, Staphylococcus, Mucormycosis, Vibrio Vulnificus, Listeria Monocytogenes* and *Yersinia enterocolitica* have been associated with transfusional IO (Evens et al., 2004). Patients with genetic hemochromatosis have been reported to be susceptible to infections, such as *Vibrio vulnificus, Listeria monocytogenes, Yersinia enterocolitica, Salmonella enteritidis, Klebsiella pneumoniae, Escherichia colo* and *Rhizopus arrhizus (*Andrews, 1999*)*. Opportunistic bacteria and fungi have been found to need free iron for a growth advantage (Evens et al., 2004). In addition, susceptibility to infections due to RBC transfusions, containing iron, has been widely acknowledged clinically (Brand, 2002).

In recent years, the association of IO with, especially, innate immunity; i.e., natural resistance to infections has been recognised. Disturbances in iron homeostasis and an increase of free iron were found to impair innate immunity and host defence (Ganz, 2009; Nairz et al., 2010; Wang and Cherayil, 2009). Innate immunity is a crucial element for the survival to our species. It involves the immunity that is not adaptive or dependent on antibiotics, including, for example, phagocytic cells and tissue fluids. In order to function normally, innate immunity needs circumstances of extremely low levels of free ionic iron (10^{-18} M) in body fluids. This low-iron environment is maintained with iron-binding proteins (transferrin and lactoferrin). Three main sources responsible for free iron and disrupting the iron homeostasis in clinical infections are thought to be: 1) excess of iron due to systemic IO, hepatic disease, cancer of chemotherapy; 2) free heme compounds from lyses of red cell and 3) hypoxia of injured tissue (Bullen et al., 2006). Iron-dependent effects on pathogen virulence and immunity are suggested to explain the susceptibilities for infections, such as salmonella, tuberculosis and malaria (Wang and Cherayil, 2009).

Although the underlying mechanisms or iron-related immunity are still not thoroughly understood, iron deprivation was found to be the key factor in the antimicrobial host

defence. Mammals have evolved strategies for restricting the availability of iron to pathogens. One of these is a regulatory feedback mechanism (Wang and Cherayil, 2009) affecting the hepcidin-ferroportin axis. Hepcidin restricts the iron availability in the plasma *via* ferroportin and keeps iron accumulated in storage sites of RE macrophages, small bowel wall enterocytes and hepatocytes (Ganz and Nemeth, 2012). A potent stimulus for upregulation of hepcidin expression is the cytokine, IL-6. This cytokine is involved in inflammation and thus hepcidin level is raised in many acute and chronic infectious and inflammatory diseases. As a result, plasma iron levels decrease and iron deprivation helps the host to defend against pathogens, but it contributes also to anemia. For example, macrophages are studied as potential components of iron-mediated host defence, as they have been proposed to restrict iron availability from pathogens by exporting iron and manganese out of the phagosome into the cytosol, inhibiting the pathogen growth (Wang and Cherayil, 2009).

2.3. FORMS OF IRON OVERLOAD

Excess iron in the body is a common disorder, because human iron balance is dependent on the regulatory feedback without an active iron excretion mechanism. There is a wide variety of different diseases resulting in IO. Iron accumulation can be either hereditary or secondary (Kohgo et al., 2008). Most of the secondary forms are due to transfusional IO. In addition, liver iron deposition may occur in miscellaneous hepatopathies, such as chronic viral hepatitis, alcohol-induced hepatopathy and non-alcoholic fatty liver disease. The degree of iron loading in these diseases is generally lower than in hereditary hemochromatosis, as the principal cause of the disease is not iron accumulation, but the underlying liver condition ((Bonkovsky et al., 1997; Gentry-Nielsen et al. 2001; Hernando et al., 2014). The IO-related secondary and hereditary disorders are categorised in Table 1.

2.3.1. Secondary IO

Secondary IO occurs in a large number of patients with hematological disorders suffering from bone marrow failure and transfusion-dependent anemia leading to transfusional IO. Accordingly, in patients with genetic transfusion-dependent anemia, especially in thalassemia, IO is frequently found due to ineffective erythropoiesis and repeated RBC transfusions (Kohgo et al., 2008). Accumulation of cardiac iron has been found to be the leading cause of death in thalassemia major (Olivieri et al., 1994). Therefore, IO has been best characterised, studied and treated with iron-chelating therapy in patients with thalassemia, and less is known about the transfusion-dependent hematological diseases or chronic hepatopathies (Olivieri and Brittenham, 1997; Wood, 2008).

2.3.1.1. Transfusional IO

Long-term RBC transfusion is a routine treatment in many types of anemia resulting in rapid iron loading and transfusional IO, known also as transfusion hemosiderosis (Andrews, 1999). Earlier it has been estimated mainly by serum ferritin, and has been found in patients with RBC transfusion-dependent anemia, especially in patients with

bone marrow failure or genetic anemia (Kohgo et al., 2008; Pullarkat, 2010). Among hematological diseases, acute leukemia (AL), MDS and aplastic anemia were suggested to have the highest IO (Pullarkat et al., 2008).

The main cause of a progressive accumulation of body iron in these patients is RBC transfusion treatment, which brings excess iron into the body. Each unit of packed RBC contains approximately 200-250 mg iron (Pullarkat, 2010), which is at least ten times more than normal daily dietary absorption (Kohgo et al., 2008). In Finland, the volume of one unit of RBC is 261 ml and contains 173 mg of iron on average (http://extranet.libris.fi/proweb/Verivalmisteiden_kayton_opas2_2013/).

Iron from RBC transfusions accumulates in RE system: Kupffer cells and macrophages of the liver, spleen and bone marrow, and also through plasma NTBI to hepatocytes and other parenchymal organs (Kohgo et al., 2008; Pullarkat, 2010). Already in 1981, a small, preliminary study reported clinical consequences of transfusional IO in patients with acquired anemia. These patients were found to suffer from liver iron accumulation leading to hepatic fibrosis, and IO was also associated to left ventricular cardiac dysfunction. The pattern of organ involvement was then suggested to be similar that in hereditary hemochromatosis (Schafer et al., 1981). Approximately 10 to 20 RBC transfusions have been suggested to lead to clinically significant iron accumulation (Majhail et al., 2008; Pullarkat, 2010).

2.3.2. Hereditary hemochromatosis

Hereditary hemochromatosis is an autosomal recessive disorder and the most common genetic predisposition in the Western countries. Even every tenth person in the Western population has a mutated allele, thus being heterozygous for this mutation. The prevalence of homozygous individuals is 1-3/1000 (Brissot et al., 2011), which roughly corresponds to an estimated incidence of hereditary hemochromatosis of 1/227 inhabitants in Caucasian populations (Adams, 2006). However, clinically significant IO usually develops only in the patients who are homozygous for the mutation, and due to a variable clinical penetration of the disease, not even all homozygous patients encounter a clinical disease (Andrews, 1999). Clinical symptoms are often non-specific, such as arthralgias, fatigue and diabetes (Adams, 2006), and a severe clinical presentation is rare (Brissot et al., 2011).

HFE gene-related type 1 hemochromatosis is the most common type of all genetic IO disorders. The C282Y mutation in both alleles explains more than 90 percent of the type 1 hemochromatosis and is even more prevalent in certain areas, such as Ireland, Brittany and France. This mutation is of Celtic origin, and has occurred before 4000 BC. Many other mutations have also been recognised in recent years. The different mutations in hemochromatosis result in increased export properties of ferroportin from the duodenum and spleen. This induces increased iron release from the cells and hypersideremia. Owing to the appearance of NTBI in the plasma, iron is avidly taken up by various parenchymal cells, first and foremost, the hepatocytes. The effect is, in most cases of hemochromatosis, mediated by a low hepcidin level (Brissot et al., 2011).

Table 1. Classification of the causes and disorders in different forms of IO. Modified from Kohgo et al., 2008 and Brissot et al., 2011.

Secondary iron overload	
RBC transfusions	Hematological diseases
Ineffective erythropoiesis	Thalassemia, MDS, sideroblastic anemia
Administration of iron *i.v.* or per os	
Dietary iron overload	
Liver dysfunction	
Others	Porphyria
Hereditary hemocromatosis and related disorders	
Hereditary hemochromatosis	Type 1 *-HFE gene mutation (location 6p21.3)*
	Type 2 *-Subtype A: hemojuvelin gene mutation (location 1q21)* *-Subtype B: hepcidin gene mutation (location 19q13)*
	Type 3 *-Transferrin receptor 2 gene mutation (location 7q22)*
	Type 4 *-Ferroportin gene mutation (location 2q32)* *-Hepcidin-deficient phenotype (different mutations)*
Others	*H-ferritin gene mutation*
	Divalent metal transporter 1 gene mutation
Aceruloplasminemia	*Ceruloplasmin gene mutation*
Atransferrinemia	*Transferrin gene mutation*

2.4. HEMATOLOGICAL DISEASES

Hematological malignancies are severe disorders characterised with long-term RBC transfusion need, which leads to transfusional IO. The malignant cells of hematological diseases, such as leukemia and bone marrow infiltrating lymphomas, accumulate in bone marrow leading to encroachment of the bone marrow reserves until bone marrow failure sets in impairing the development of clonal cells. This results in cytopenias and increased risk of infections, bleeding and anemia, which is commonly treated with repeated RBC transfusions. Generally, the treatment of hematological malignancies can be either curative or symptomatic depending on the patient's age and disease type. Acute hematological malignancies have a short course of disease, but in many cases they can be treated curatively. Patients with chronic malignancies, on the other hand, cannot be treated curatively with chemotherapy, and they are often transfusion-dependent for longer periods. Stem cell transplantations are needed in some acute

malignancies as part of curative treatment, and also certain chronic hematological malignancies can be treated curatively with alloSCT (Palva, 2005). The reported five-year survival rates in Europe from 2000 to 2002 have been 37 per cent for all myeloid malignancies (Maynadié et al.2013) and 57 per cent for all lymphoid malignancies (Marcos-Gragera et al., 2011).

The classification of World Health Organization is used for hematologic diseases. According to the lineage, classification into lymphoid and myeloid malignancies can be made. Myeloid malignancies are subcategorised mainly on their degree of maturation and biologic properties into entities such as acute myeloid leukemia (AML), MDS and myeloproliferative neoplasms (MPN), which include chronic myelogenous leukemia (CML) and myelofibrosis (Vardiman, 2010). Lymphoid neoplasms are a heterogeneous group of diseases, such as non-Hodgkin's lymphoma (NHL), Hodgkin's lymphoma (HL), lymphoid leukemias and plasma cell neoplasms, which include multiple myeloma (Marcos-Gragera et al., 2011). The classification is based on a combination of clinical, morphologic, immunophenotypic, genetic and other biologic features. Furthermore, the neoplasms containing more mature cells (for example MPN, MDS, mature B-cell and T/NK cell lymphoma, HL) are separated from neoplasms consisting of precursor cells with impaired maturation, such as AML and lymphoblastic leukemia/lymphoma.(Vardiman, 2010)

2.4.1. Myelodysplastic syndrome

MDS is a myeloid malignancy, which consists of a heterogeneous group of clonal stem cell disorders of varying prognoses depending on the subtype. MDS is characterised by ineffective hematopoiesis, morphological dysplasia, cytopenias and a high risk of progression to AML (Malcovati, 2007). Myelodysplastic changes can be found in all three myeloid lineages (Palva 2005). More than 90 per cent of patients suffer from anemia (Malcovati, 2007) and most of the patients become RBC transfusion-dependent. The excess of blasts and evolution to AML have the worst prognosis and are treated like acute leukemia (Palva, 2005).

The patients with MDS are thought to be more susceptible to IO than other patients with hematological disease and transfusional IO. The patients who would otherwise have more favourable prognosis according to the disease subtype are more likely to receive loads of RBC transfusions, thus having high risk of developing significant IO (Malcovati et al., 2007; Fenaux and Rose, 2009; Cuijpers et al., 2010). Another cause of IO is ineffective, though active dysplastic erythropoiesis, which occurs in MDS and especially in patients with RARS (refractory anemia with ringed sideroblasts). Even mitochondrial iron accumulation has been found within the ringed sideroblasts in these patients, but also in some other subtypes of MDS. This increased erythropoietic activity in itself can lead to IO, because it suppresses inappropriately the hepcidin synthesis. The downregulation of hepcidin increases intestinal iron absorption, RE iron recycling and plasma iron levels. This is suggested to lead to parenchymal iron loading, similarly as has been reported in thalassemia, where iron has been found mostly in hepatocytes and less in RE Kupffer cells. The interactions between bone marrow and hepcidin-producing liver hepatocytes are thought to be mediated *via*

humoral factors, such as growth differentiation factor 15 and twisted gastrulation 1 (Cazzola et al., 2008; Cuijpers et al., 2010).

High IO as estimated by serum ferritin and transfusion need have been associated with reduced survival in these patients and iron-chelation treatment is recommended, although somewhat controversial (Malcovati, et al. 2007). Furthermore, it is not clear whether these imperfect surrogates for iron, ferritin and transfusion need, reflect actual IO or more the disease severity. Thus, direct IO quantification, such as liver iron load measurement with MRI, has been proposed to be used in order to clarify the actual impact of IO on the prognosis in these patients (Malcovati, 2007).

2.4.2. Other hematological diseases

AL consists of clonal malignancies that originate from a primitive progenitor cell, which does not mature normally and remains at the blast stage. These clonal immature blast cells accumulate and overcrowd the bone marrow leading to anemia, granulocytopenia, and thrombocytopenia. AL originates from either myeloid or lymphoid lineage and is divided into AML and acute lymphoblastic leukemia (ALL) (Palva, 2005). AML is not only the most common AL in adults, but also the commonest myeloid malignancy, and one of the most iron-overloaded disorders among all hematological malignancies (Pullarkat et al., 2008; Sant et al., 2010; Maynadié et al. 2013).

Aplastic anemia rises typically from immune-mediated destruction of the hematopoietic progenitor cells and the disturbance of the bone marrow microenvironment, leading to pancytopenia. The progenitor cells of the bone marrow are destroyed and replaced by the fat cells (Palva, 2005; Young et al., 2008). Thus, aplastic anemia is not classified as a malignancy, although it can occur in a very severe form. Especially severe aplastic anemia (SAA) is a life-threatening bone marrow failure disorder, which can be treated with alloSCT (Palva, 2005; Young et al., 2008).

Chronic MPN are clonal hematological disorders derived from the progenitor cell proliferation in one or more of the myeloid lineages, leading to distinct disease entities as myelofibrosis, CML, polycythemia vera or essential thrombocytosis. These diseases are typically diagnosed in elderly population and they are the most common chronic myeloid malignancies (Palva, 2005). In myelofibrosis, the bone marrow fibrosis occurs as a response of megakaryocyte-derived fibrogenic growth factors leading to bone marrow failure.

Lymphomas are solid tumours of the immune system. NHL accounts for about 90 per cent of the lymphomas, and 10 per cent consist of HL (Shankland et al., 2012). Eighty-five to 90 per cent of NHL arise from B lymphocytes and the remainder derive from T lymphocytes or natural killer cells. It usually develops in the lymph nodes, but can occur in almost any tissue (Shankland et al., 2012). HL is a malignancy of lymph nodes. The tumour consists of few neoplastic Reed-Sternberg's cells of the B-cell origin, and reactive mass of more numerous T-cells. AlloSCT with reduced intensity is nowadays increasingly used in the patients who do not respond to standard salvage therapy. Treatment-related mortality of this procedure in expert centres is roughly 20 per cent (Townsend and Linch, 2012).

Myeloma is a malignant disease of B-lymphocytes in which clonal primitive stem cells mature only at the level of plasma cells. These monoclonal tumour cells then excrete monoclonal immunoglobulin. Symptoms include anemia and later on also thrombocytopenia. Typical findings are lytic bone lesions due to activated osteoclasts. AlloSCT is performed quite rarely in highly selected cases.

2.4.3. Cardiac IO

Cardiac iron overload has been found to be the leading cause of death and an important prognostic factor in transfused patients with thalassemia. Thus, these patients have been treated with iron-chelation treatment over 20 to 30 years (Olivieri et al., 1994; Olivieri and Brittenham, 1997; Wood, 2008). In thalassemia, cardiac iron accumulation occurs later in the course of disease compared to liver IO (Wood et al., 2004). Both the rate and the primary mechanism of cardiac iron uptake and release seem to be different from those in the liver. The iron uptake is generally more regulated in the liver. Cardiac iron overloading occurs as transferrin binding capacity of iron saturates, NTBI appears in the circulation and enters *via* voltage-dependent channels into cardiac myocytes (Wood, 2008). However, in hematological malignancies the incidence and importance of cardiac iron excess has not been established. Only a few cases of cardiac IO have been reported in patients with MDS, but these results have been contradictory and the role of cardiac IO in MDS is still widely debated (Jensen et al., 2003; Chacko et al., 2007; Konen et al., 2007; Di Tucci et al., 2008; Wood, 2008; Fenaux and Rose, 2009; Park et al., 2009; Armand et al. 2011). No distinct association between cardiac and hepatic iron content has been found in cross-sectional studies (Anderson et al., 2001; Wood et al., 2004), although potential causal relationship in a longitudinal study between liver and cardiac IO has been suggested in thalassemia (Noetzli et al., 2008). However, chelation might have been a confounding factor in these results, as it may change the magnetic resonance (MR) calibration curve (Wood et al., 2008), and the liver to cardiac iron proportions due to different efficiency (Anderson et al., 2002; Berdoukas et al., 2009), and accordingly influence the results.

2.4.4. HSCT

HSCT, and especially alloSCT, is an increasingly used curative treatment for hematological malignancies and SAA (Evens et al., 2004; Copelan, 2006; Majhail et al., 2010). The most common indications for alloSCT are AL, MDS, lymphomas, myeloma, SAA, myelofibrosis, CML and chronic lymphomas (Copelan, 2006).

Despite the recent advances, there still is a substantial mortality left in alloSCT (Evens et al., 2004; Majhail et al., 2010). TRM has more recently been at the level of 26 per cent, and the overall long term mortality has been 47 per cent between 2003 to 2007, reported in a study of 1146 alloSCT recipients suffering from hematological diseases (Gooley et al., 2010). The most important causes of TRM, in these patients, are severe infections and GVHD. IO has been suggested as a contributor for poor transplantation results (Evens et al., 2004; Majhail et al., 2008; Kataoka et al., 2009).

2.4.4.1. Basic concept of HSCT

HSCT can be either myeloablative or non-myeloablative (Treleaven and Barrett, 2009). Myeloablative conditioning (MAC) regimens include fractionated total-body irradiation combined with a toxic and immunosuppressive conditioning agent, such as cyclophosphamide. A less toxic reduced-intensity conditioning (RIC) regimen has been developed more recently. It consists of immunosuppressive, low-dose cytotoxic regimens with or sometimes without low-dose total body irradiation. This conditioning regimen is better tolerated and it has lower TRM than MAC. Thus, it was first designed for older recipients or for recipients with organ dysfunctions, but is also increasingly used as a standard conditioning (Copelan 2006).

During HSCT, the graft, either autologous or allogeneic, is infused intravenously into the recipient. Thereafter, stem cells circulate to the bone marrow and start homing into the hematopoietic sites, followed by stem cell engraftment and finally reconstitution of hematopoiesis. A complete recovery of the immune system takes several months or even years (Treleaven and Barrett, 2009). In alloSCT, the graft is derived from a suitable sibling donor or a HLA-compatible registry donor (Copelan 2006). The possible adverse post-transplant alloimmune reactions (acute and chronic GVHD) can be both adaptive and innate immunity responses (Treleaven and Barrett, 2009).

2.4.4.2. AlloSCT and IO

Transplantation recipients are suggested to have an increased susceptibility to iron. Excess of iron in alloSCT has been proposed to predict poor outcome, i.e., increased morbidity and mortality. Indeed, the NTBI concentration in blood increases tremendously during transplantation (Dürken et al., 1997, Dürken et al., 2000; Sahlstedt et al., 2001; Evens et al., 2004; Sahlstedt et al., 2009; Yegin et al., 2011) and high post-transplant IO has been measured in these patients by HIC analysis (Rose et al., 2007).

The possible reasons for the rapid increase of NTBI during HSCT have been speculated to be several: increased intestinal iron absorption due to low hepcidin (i.e., in MDS, thalassemia, HFE gene mutations), increased RE macrophage iron due to transfusional IO, underutilisation of plasma iron due to decreased erythropoiesis as a result of cytotoxic conditioning regimen, and finally the release of cellular iron due to the destruction of bone marrow (Pullarkat, 2010). The possible mechanisms of NTBI-mediated adverse effects are pro-oxidative state followed by tissue injury (Dürken et al., 2000; Evens et al., 2004; Yegin et al., 2011) and increased infection risk due to suppression of host immune responses together with the role of iron as a cofactor of pathogen growth (Bullen et al., 2006; de Witte, 2008; Koreth and Antin, 2010).

Although the rise of NTBI in transplantation is well established, the role of IO in post-transplant complications is still being debated (Majhail et al., 2008; Kataoka et al., 2009). Studies done with non-specific or semi-quantitative iron indicators, such as ferritin, suggest that IO leads to increased infection risk (Miceli et al., 2006; Kontoyiannis et al., 2007; Ozyilmaz et al., 2010; Pullarkat, 2010; Tachibana et al., 2011; Ali et al., 2012) and decreased survival (Altès et al., 2002; Armand et al., 2007; Platzbecker et al., 2008, Pullarkat et al., 2008, Kataoka et al., 2009; Alessandrino et al.,

2010; Sucak et al., 2010; Armand et al., 2011) in alloSCT recipients. The poorer survival in the allotransplanted patients with hyperferritinemia has been reported with both myeloablative and non-myeloablative conditioning (Mahindra et al., 2009a, Mahindra et al., 2009b).

Until very recently, there has been no prospective data concerning the effect of actual iron concentrations on patients' outcome after alloSCT. Pre-transplant IO and its impact on post-transplant mortality have been measured directly with MRI in two studies reaching contradictory results and without analysing post-transplant infections (Armand et al. 2012; Wermke et al. 2012). They both used quantitative MRI for hepatic iron concentration assessment. Armand et al. used an R2* measurement, which was calibrated with HIC using a previously presented calibration equation (Wood et al., 2005). Wermke et al. used the method based on liver-to-muscle SI ratios (Gandon et al., 2004), which we validated in this study. The effect of IO on GVHD has been even more controversial in the studies, which used non-direct iron indicators for IO assessment (Platzbecker et al., 2008; Mahindra et al., 2009a; Pullarkat, 2010) and studies on GVHD with a direct IO assessment by MRI have been lacking.

Despite the lacking evidence and controversial results in these studies, it has been proposed that to improve outcome, IO should be reduced before alloSCT with iron-chelation treatment and/or post-transplant with phlebotomy or chelators (de Witte, 2008; Sucak et al., 2010). Because ferritin as an acute-phase reactant also reflects inflammation and infection, the analysis of IO has been recommended to be done by MRI or biopsy before initiating the therapy for IO (Majhail, Lazarus, and Burns 2008). Therefore, further prospective studies with direct, non-invasive measurement of IO by MRI have been strongly suggested to be performed in order to establish the incidence, clinical significance and therapeutic need of IO (Brittenham and Badman, 2003; Koreth and Antin, 2010; Leitch, 2011; Armand et al., 2012).

2.5. THE METHODS FOR BODY IRON CONTENT ASSESSMENT

Many different methods for the IO assessment have been developed. These methods vary, for example, in methodology, target of assessment, accuracy, availability and safety. Liver iron content can be regarded as an accurate measure of whole body iron stores (Angelucci et al., 2000). Therefore, the liver has been the main focus of iron concentration quantification and MR imaging is considered the most promising non-invasive method for iron quantification (Brittenham and Badman, 2003; Wood, 2007).

2.5.1. Liver biopsy

Liver biopsy with biochemical analysis has previously been the golden standard for IO quantification (Bonkovsky et al., 1999). In this method, HIC measurement is done typically by atomic absorption spectrophotometer of a liver specimen. Usually, a simultaneously performed biopsy for histological analysis enables detection of fibrosis and cirrhosis (Alústiza et al., 2007). However, as an invasive procedure, it may not be safe especially in the patients with thrombocytopenia or coagulopathy (Majhail et al., 2008). Due to complication risk and discomfort, it is not suitable in treatment

monitoring with multiple measurements (Brittenham and Badman, 2003; Alústiza et al., 2007). The variability of HIC measurement from needle-biopsy specimens is relatively high. The intra-individual coefficient of variation for HIC in patients with cirrhosis was 25 per cent on average varying from 11 to 44 per cent (Villeneuve et al., 1996). Even in patients without a liver disease, the reported on average variation was 19 per cent (Hernando et al., 2014). Formalin fixation of the liver specimen was shown to reduce HIC significantly (Villeneuve et al., 1996), which might contribute to the reported high variability. Iron load can also be estimated in histological specimens with Perl's Prussian Blue staining, but this method is only semi-quantitative and not all forms of tissue iron are demonstrable (Bancroft and Stevens, 1990).

2.5.2. SQUID

Magnetic susceptibility measurement of HIC with SQUID (superconducting quantum interference device) was developed by Bauman and Harris in 1967. It was the first method that was calibrated and validated to be used as an accurate non-invasive measure of liver iron concentration. The method has excellent correlations with chemically measured HIC from biopsies. However, it is not feasible in clinical use due to complexity, high cost and limited number of equipment worldwide (Brittenham and Badman, 2003; Alústiza et al., 2007; Majhail et al., 2008).

2.5.3. CT

Computed tomography (CT) was first proposed as a potential method for IO assessment three decades ago (Howard et al., 1983). It is a relatively cheap, feasible, non-invasive and robust imaging method in clinical use despite the radiation dose. It has been thought that liver density measured with CT would increase with liver iron content, although other causes for liver density variations, such as fatty liver is regarded as a confounding factor. Dual energy CT has been suggested as a potential technique, and it has been promising in an *ex vivo* study in liver phantoms (Fisher et al., 2011). However, the validation of the CT method in humans has been limited. The sensitivity of the CT method for IO assessment has not been adequate especially in the lower concentration range either in human (Howard et al., 1983; Bonkovsky et al., 1990) or rat studies with or without dual energy (Nielsen et al., 1992). Thus, it has hardly been used in clinical setting. More recently, the capability of CT has been re-evaluated using modern equipment in a retrospective study in 37 patients with transfusional IO (Wood et al., 2011). They calibrated liver attenuation values with a hydroxyapatite phantom used for bone marrow density calculations and correlated the calibrated liver density values with a liver iron concentration, measured with a validated MRI-based R2 method (Wood at al., 2005). The results showed the lacking diagnostic value of the CT method (<76 Hounsfield units) in the liver iron concentration range lower than 143 mikromol/g, but a linear correlation in the higher liver attenuation and iron concentration values. However, the relatively high cut-off value currently prevents the wider clinical use of this method for IO assessment and leaves only a complementary role for CT in this regard.

2.5.4. Serum and plasma iron studies

Ferritin and other serum or plasma laboratory parameters, such as transferrin saturation, are very feasible and non-invasive iron indicators, but they are indirect and not specific for iron. Ferritin is widely used. However, as an acute-phase reactant, ferritin rises also in infection, inflammation, hepatocellular necrosis or neoplasms, chemotherapy and liver injury (Nielsen et al., 2000, Evens et al., 2004; Alústiza et al., 2007; Majhail et al., 2008). Thus, the correlation between ferritin and HIC has been poor in thalassemia, hemochromatosis (Nielsen et al., 2000), sickle-cell anemia (Karam et al., 2008) and in iron-overloaded patients including hematological disorders with inflammation (Othlof et al., 2007). In addition, ferritin is insufficient in detecting cardiac iron (Kolnagou et al., 2006). The published data concerning the correlation of serum markers other than ferritin with HIC is more limited. The transferrin saturation (serum iron divided by iron-binding capacity) is thought to lack specificity as an iron load indicator, as it rises, for example, in chronic hepatopathies (Alústiza et al., 2007).

2.5.5. RBC transfusion indices

The total cumulative number of RBC units received, i.e., transfusion burden, is thought to proportionally add iron into the body and cause transfusional IO, but transfusion burden is not generally used in the clinical setting for IO estimation, and validations of the method have been lacking. Transfusion need, however, has been taken into the prognostic scoring system of MDS (Malcovati et al., 2007), but it is more related to the severity of the disease and anemia than to IO (Malcovati and Cazzola, 2011; Malcovati, 2007). In one study on patients with hematological disorders, who survived alloSCT, the correlation between transfusion burden and HIC was evaluated. The results of the first published correlation were promising as the number of RBC units received seemed to cause a proportional increase in HIC (Rose et al., 2007).

2.5.6. MR imaging

Currently, MRI has become the most widely used tool for non-invasive measurement of total body iron stores. The improvements in MR instrumentation, software and methodology have enabled robust and reliable non-invasive methods for tissue iron quantification. Along with the increasing number of treatment options, this facilitates iron overload assessment in a large number of patients and helps the clinicians to choose the correct treatment for the patients and to monitor the treatment effect.

The measurement of iron concentration by MRI is based on the paramagnetic feature of iron, which can shorten transverse relaxation time in the proportion of iron concentration. Thus, iron will cause signal loss in MR images (Wood, 2007). This phenomenon is used to measure iron concentrations quantitatively with different methods. The quantification of iron concentration can be done by measuring relative signal intensities (Gandon et al., 2004) or the transverse relaxation. Transverse proton relaxation time is measured either from spin echo images (T2) or gradient echo images (T2*). The reciprocals for these relaxation times are proton transverse relaxation rates

from spin (R2) and gradient echo images (R2*), which increase proportionally with iron accumulation (St Pierre et al., 2005; Wood et al., 2005).

2.6. BASIC CONCEPTS OF MR IMAGING

In 1946, Felix Bloch (Bloch, 1946) and Edward Purcell et al. (Purcell et al., 1946) independently discovered the basic concepts of MR imaging. The alignment of nuclear magnetic moments with the external magnetic field, their precession at Larmor frequency, and the resonance phenomenon that enables the transfer of the energy and the change in the direction of nuclear net magnetisation caused by a radio frequency (RF) pulse. Bloembergen et al. continued the work on the basis of NMR absorption, creating the concept of T1 and T2 relaxation times (Bloembergen et al., 1948).

2.6.1. Instrumentation of MR scanner

Three different electromagnetic systems are needed in all clinical MR scanners. The largest one with a strong current is used to generate the main magnetic field (Figure 1a). The other two are coils, which are used either in obtaining three-dimensional information (Figure 1b) or transmitting and receiving the MR signal (Atlas, 1996).

The main magnetic field of a MR scanner is constantly on in modern MR systems. It is produced by a large electric current in loops of superconducting wire. These wires are immersed in liquid helium, where the low temperature enables superconducting state of the material. As a result, very large currents can be used to gain a strong magnetic field and retain it for many years without recharging the equipment. The strength of this magnetic field is 1.5 Tesla (T) in typical clinical MR scanners (Atlas, 1996; Pooley, 2005). It is a relatively strong magnetic field, since 1 T corresponds 20,000 times the average of Earth's magnetic field (MR Glossary, 2009).

Gradient coils are used to encode the location of a voxel in three-dimensional space as illustrated in Figure 1b (Atlas, 1996; Pooley, 2005; Bitar et al., 2006). These gradients create linear differences in a selected region of the magnetic field strength (Bitar et al., 2006). The information about the location of the voxel is obtained in three perpendicular dimensions by 1) the excitation of a certain section of the target tissue and encoding, 2) the phase and 3) frequency of the MR signal in that tissue. For a typical 1.5T cylindrical-bore MR scanner, the longitudinal z-axis is aligned with the main magnetic field. A perpendicular plane is typically a transverse plane forming the x- and y-axes (Atlas, 1996; Pooley, 2005; Bitar et al., 2006).

RF coils are required for two functions, transmitting and receiving the MR signal. Transmitter RF coils are applied for causing a change in the direction of the nuclear spins. As a result, the signal produced by resonating nuclear spins in a tissue can be detected by the receiver RF coil, which can be different in size depending on the target (Atlas, 1996).

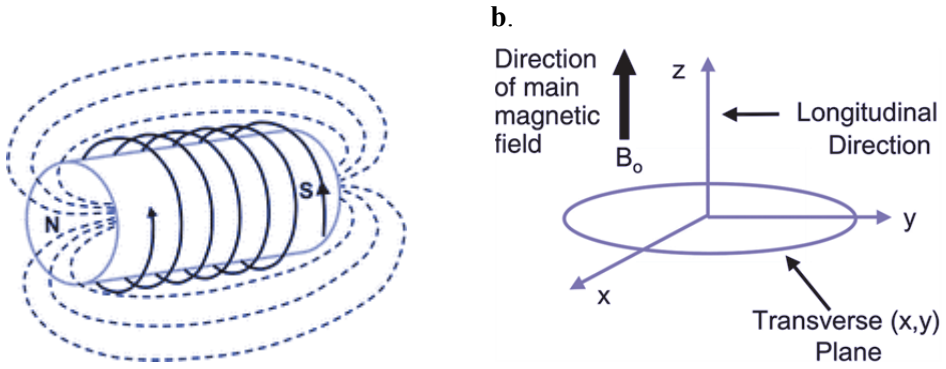


Figure 1. a) Powerful main magnetic field of MR scanners. N=North, S=South. **b)** Three-dimensional coordinate system for MR scanners. The longitudinal magnetisation is typically horizontal and aligned with the patient longitudinal axis, for example, head to feet. Accordingly, the transverse plane corresponds to patient's axial direction. Modified and reprinted with permission from Pooley et al., 2005.

2.6.2. Precession and nuclear magnetic resonance

Within an external magnetic field, the nuclear magnetic dipole moment, often described as the spin, rotates around its axis and produces a miniature magnetisation vector (Figure 2). This nuclear precession and NMR are the key elements in the efficient transfer of energy to the spin (Pooley, 2005). The precessional frequency is characteristic of different nuclei and can be determined by Larmor equation,

$$\omega_0 = B_0 \cdot \gamma, \qquad \text{[Eq. 1.]}$$

which shows that the frequency of precession (ω_0) is directly dependent on external magnetic field strength (B_0) and gyromagnetic ratio (γ), a characteristic constant for each atom. For example, hydrogen proton (^{1}H) processes at 1.5T with a frequency of, ω_0=1.5 T × 42.57 MHz/T = 63,86 MHz. ^{1}H is the most commonly used nucleus in clinical scanners, as it is the most abundant nucleus in water-rich human tissues. Also other odd numbered nuclei, such as ^{13}C, ^{19}F, ^{23}Na and ^{31}P, can be used in MR imaging, because they possess a magnetic moment (Schild, 1990; Atlas, 1996; Pooley, 2005; Bitar et al., 2006).

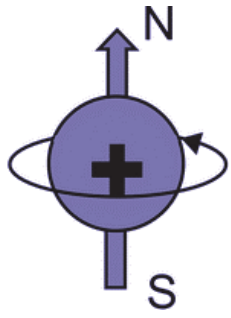


Figure 2. The positively charged hydrogen nucleus induces a small magnetic field. The thin arrow around the nucleus illustrates the spinning motion as the nucleus rotates around its axis. Modified and reprinted with permission from Pooley et al., 2005.

2.6.3. Magnetic nucleus in changing external magnetic field

When an external magnetic field (*B*o) is applied in the longitudinal z-direction, the magnetic moments of nuclei align with it, either in parallel or in an opposite direction. These magnetic moments in opposite directions are at different energy levels. The lower level is preferred resulting in an approximate difference of $0.8 \cdot 10^{-6}$ in the number of protons in opposite directions at 1.5T (Schild, 1990). The cumulative sum of all these magnetic moments of nuclei is the net magnetisation vector (NMV). After applying an RF pulse, nuclei absorb the energy at the Larmor frequency and the NMV starts flipping towards transversal plane. The change in the direction of the NMV is called flip angle (FA). After the RF pulse, the NMV has two components, longitudinal (along the applied field) and transversal (perpendicular to the applied field) magnetisation (Li et al., 2004; Bitar et al., 2006,). This phenomenon is illustrated in Figure 3.

When the RF pulse is turned off, two important processes occur simultaneously: Transversal magnetisation decreases (T2 decay) and longitudinal magnetisation increases (T1 recovery), as nuclei dephase and realign with the main magnetic field, since in nature the lower energy level is preferred. The constants of T1 and T2 are the basis of SI, which is different in different tissue types thus forming the MR contrast (Atlas, 1996; Pooley, 2005; Bitar et al., 2006).

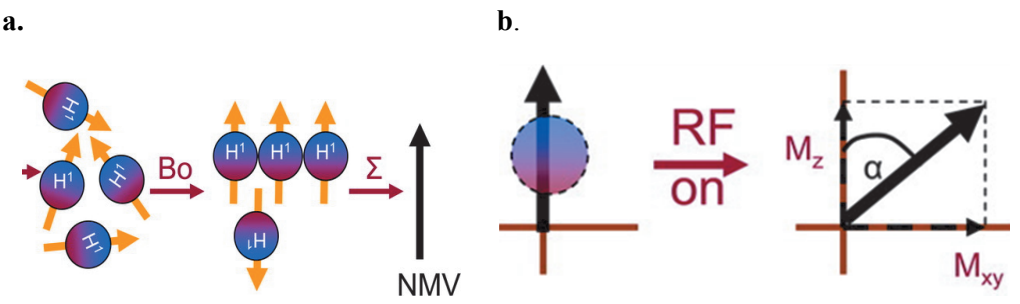


Figure 3. **a)** Initially the magnetic vectors of nuclei point to various angles (left), but these nuclear vectors rapidly align with an external magnetic field (*B0*). The sum of these miniature vectors, a net magnetisation vector (NMV), is used for MR image production. **b)** The RF pulse brings energy to the system when it excites the nuclei and flips the magnetisation towards the transversal plane. The transversal component of NMV induces a current in the receiver coil and is used for signal creation. Mz = longitudinal magnetisation, Mxy = transversal magnetisation, and α = flip angle. Modified and reprinted with permission from Bitar et al., 2006.

2.6.3.1. T1 recovery

The increasing longitudinal magnetisation is a process known as T1 recovery (Bitar et al., 2006). It was initially called *spin-lattice relaxation* because spins interact with their surroundings, i.e., lattice during this return to equilibrium (Atlas, 1996). Currently, it is more often called *longitudinal relaxation or T1 recovery*. T1 relaxation begins when the RF pulse is no longer applied. Then, the nucleus starts releasing its excess energy to the environment and the nuclei start to realign their longitudinal direction (Atlas, 1996; Pooley, 2005; Bitar et al., 2006). T1 relaxation time (ms) is a constant that

defines how much of this relaxation has occurred in each tissue. The T1 time differs between tissues. T1 recovery and an equation fitted to normal human liver, where the typical T1 time is 586 ms (de Bazelaire et al., 2004), is illustrated in Figure 4.

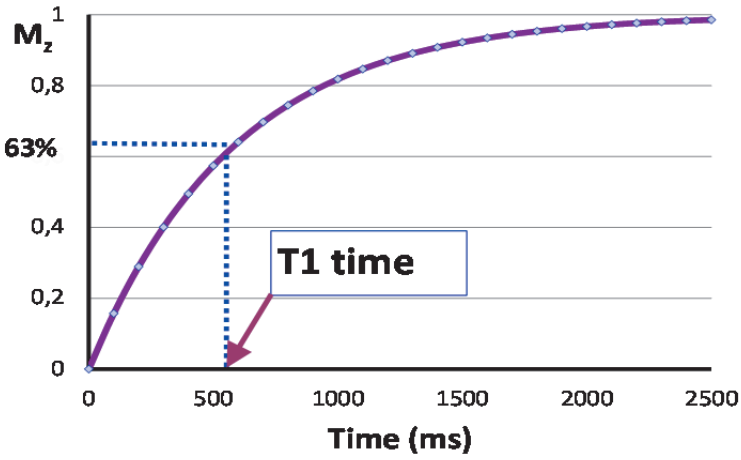


Figure 4. T1 recovery. Longitudinal relaxation occurs after the RF pulse has been turned off and the magnetic vector starts to return to equilibrium, towards longitudinal direction (Mz). 63% of the longitudinal magnetisation has increased at the time that is defined as T1 time (ms). This particular curve and T1 time chosen are characteristic of normal human liver tissue.

2.6.3.2. T2 decay

T2 decay is a simultaneous phenomenon with T1 recovery and essential in contrast formation. It has been defined as *transverse* or *spin-spin relaxation,* namely because spins interact with each other and no energy is transported to the surroundings during this transverse relaxation (Atlas, 1996; Bitar et al., 2006). When a 90º RF pulse is applied, nuclear spins are coherent and in the same phase of the procession, termed *in-phase*. After the RF pulse is turned off, the transverse magnetisation starts to reduce at a certain speed, as nuclei start to lose the phase coherence; i.e., the spins begin to *dephase.* The nuclear transverse net magnetisation is used for the MR signal production. T2 time is a constant, which defines how rapid this dephasing is in certain tissue (Atlas. 1996; Pooley. 2005; Bitar et al., 2006). It is related to the intrinsic field, which is caused by adjacent protons and their spins (Chavhan et al., 2009). Especially large molecules, such as paramagnetic hemoglobin, enhance T2 dephasing and shorten the T2 time (Atlas, 1996). Moreover, ferritin, hemosiderin and other paramagnetic iron deposits have been found to enhance predominantly T2 relaxation over T1 relaxation (Stark, 1991). The reciprocal of T2 is R2 (1/T2), defined as *transverse relaxation rate*, which increases proportionally with iron concentration (Stark 1991; St Pierre et al. 2005; Wood, 2007). T2 decay, with an equation fitted to normal human liver T2 of 46 ms (de Bazelaire et al. 2004) is illustrated in Figure 5.

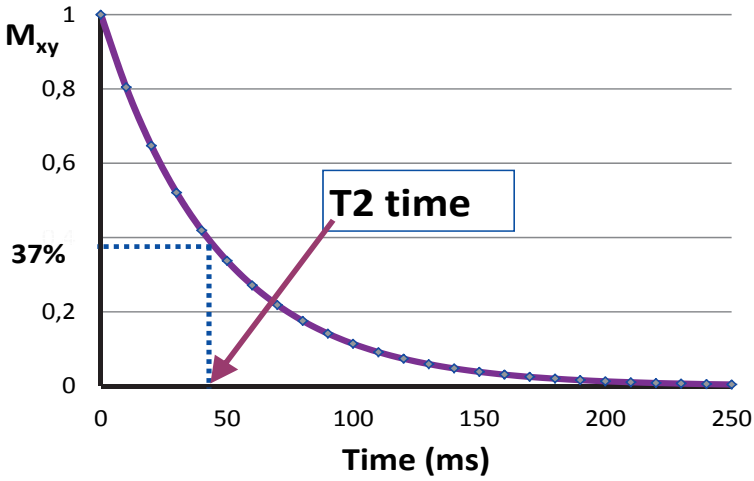


Figure 5. T2 decay. Transverse relaxation occurs after the RF pulse has been turned off. Transversal magnetisation (Mxy) starts to reduce as nuclei lose their coherence phases. 63% of the transversal magnetisation has decreased at the time that is defined as T2 time (ms) and 37% of this magnetisation is left for signal production. This curve and T2 time chosen are characteristic of normal human liver tissue.

2.6.3.3. T2 decay*

In both T2 and T2* relaxation, the dephasing is caused by irreversible spin-spin relaxation within the tissue. In addition, T2* has contributions from magnetic field inhomogeneities arising from the main magnetic field and from differences in magnetic susceptibility, chemical shift effects and even gradients applied for spatial encoding (Pooley, 2005; Chavhan et al., 2009). These inhomogeneities in the external magnetic fields are reversed and eliminated at a spin-echo T2 imaging with a 180º rephasing pulse, which is absent in gradient echo imaging (Pooley, 2005; Bitar et al., 2006). Thus, T2* from gradient echo sequences are more sensitive to magnetic field heterogeneities than T2 form spin-echo sequences (Ernst et al., 1997; Chavhan et al., 2009).

The relationship between T2 and T2* is illustrated with an equation,

$$1/T2^* = 1/T2 + 1/T2', \qquad \text{[Eq. 2.]}$$

where $1/T2' = \gamma \Delta B_{inhom}$ refers to magnetic field inhomogeneity in relation to the gyromagnetic ratio (γ). T2* is characteristically shorter than T2 in a tissue (Chavhan et al., 2009). This is illustrated with corresponding transverse relaxation curves in Figure 6. The T2* decay of transverse magnetisation is faster than T2 decay. Thus, T2*-weighted images typically show less signal than T2-weighted images, as the transverse magnetisation that is left after T2 or T2* relaxation is used for the signal production (Bitar et al., 2006). T2* time can be measured from gradient echo images with different TE, whereas T2 decay is measured using spin-echo imaging (Chavhan et al., 2009).

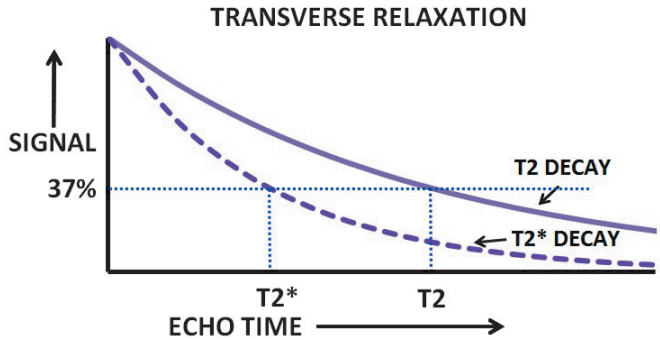


Figure 6. T2* and T2 in similar tissue. Due to the basic features in the formation of gradient echo imaging, T2* is always shorter than corresponding T2 time.

2.6.4. Tissue contrast

T1 and T2 or T2* along with proton density are responsible for the differences of signal intensities among different tissues, thus creating contrast in MR images. The inherent SI of a tissue is assessed with MR images of different parameters. Time to echo (TE), time of repetition (TR) and flip angle (FA) are key parameters in MR imaging and take part in creating the contrast. TE; i.e., *echo time* is the time that is used from the first RF pulse to the time when the echo is created. For example, in spin-echo imaging, after a 90º RF pulse and dephasing, spins are rephased with a 180º RF pulse at TE/2. It takes another TE/2 for the nuclei to rephase at the time of TE, when the maximal echo is created for signal production (Atlas, 1996; Pooley, 2005; Bitar et al., 2006). In short, TE time defines the time that is used for T2 decay. The longer TE is, the more T2-weighted is the sequence and the smaller proportion of transversal magnetisation is left for signal production (Bitar et al., 2006; Westbrook et al., 2011). TR time; i.e., *repetition time* is the time between the first RF excitation pulse and the start of the second excitation pulse. During TR, the longitudinal spin-lattice relaxation recovers and this repetition time defines the amount of T1 recovery that has occurred until the signal is read (Atlas, 1996; Pooley, 2005; Bitar et al., 2006; Westbrook et al., 2011).

TE and TR are used to emphasise a particular type of contrast (Bitar et al.. 2006). At long TE and subsequently long TR, water having a long T2, i.e., slow transverse relaxation, has higher SI than tissues, since less signal decay has occurred in water during TE. This kind of MR imaging is typically called *T2-weighted*, where the tissue contrast is mainly formed by the differences in T2 rather than the differences in T1 times. Consistently, short TE and TR result in a mainly *T1-weighted* image, where the contrast between the tissues is governed mainly by the differences in T1 and to a lesser degree with T2. If both T1 and T2 relaxation effects are minimised with long TR and short TE time, *a proton density* weighting is produced (Pooley, 2005; Bitar et al., 2006). TE and TR can be used also in measuring T1 and T2 times of a certain tissue. Imaging with variable TE times with constant TR and other parameters can be used to assess tissue-specific T2 time by plotting the tissue signal intensities against TE times.

Similarly, T1 time can be measured with different TR times and a constant TE time and other parameters.

2.6.5. Gradient echo versus spin-echo imaging

The gradient echo sequences are faster than spin-echo imaging due to many reasons. The angle of the RF pulse (FA) in gradient echo imaging is smaller than the 90º RF pulse in spin-echo imaging, and the duration of the pulse is shorter. In addition, the transverse relaxation is faster in gradient echo compared to spin-echo imaging and there is no 180º refocusing pulse. The rephasing and dephasing at gradient echo imaging is done with fast gradient pulses (Pooley, 2005). The characteristics of gradient echo imaging and T2* imaging are faster acquisition times, lower signal and susceptibility to artefacts and field inhomogeneity compared to spin-echo imaging.

Most importantly, this susceptibility to local magnetic field inhomogeneities can also be taken advantage of, for example, in the assessment of IO (Chavhan et al., 2009). R2*, as a reciprocal of T2*, has been used as a robust method for IO assessment, as R2* is linearly dependent on tissue iron concentrations (Wood et al., 2005).

2.6.6. In-phase and out-of-phase imaging

Hydrogen nuclei in fat and water molecules precess at Larmor frequencies that differ from each other, due to their different chemical environments. This *chemical shift* can be taken advantage of by a special dual-echo imaging, *in-phase and out-of-phase imaging*. In this gradient echo imaging, the tissue is imaged at two TE times and two different images are produced: at the TE time when hydrogen nuclei of fat and water are spinning *in phase* with each other (approximately TE of 4.2 ms) and *out of phase* with each other (approximately TE of 2.1 ms) at 1.5 T.

The in-phase and out-of-phase imaging was originally developed and used for fat detection. It detects mobile, intracellular fat. If fat is present within a voxel at the same time with water, the signals will cancel each other out on the out-of-phase images. Correspondingly, these signals of water and fat are added up on the in-phase images (Bitar et al., 2006).

In addition to this, in-phase and out-of–phase imaging has been found useful in liver iron assessment due to the dual-echo approach. The two images obtained approximately at 2.1 and 4.2 ms have different T2* effects. The image obtained at longer TE is more susceptible to local inhomogeneities, such as iron depositions. In hepatic iron accumulation, hepatic tissue signal decreases on the image with longer TE due to the decay of transverse magnetisation. The signal drop in the more T2*-weighted images is seen at 4.2ms, which is in-phase at 1.5T (Merkle and Nelson, 2006). Simultaneous steatosis may potentially affect the diagnostic performance. However, one retrospective report suggested that in-phase and out-of-phase imaging can be used for detection of significant hepatic IO (Lim et al., 2010), while prospective validation and correlation to HIC have been lacking.

2.7. QUANTITATIVE MRI FOR IRON MEASUREMENT

MRI has become the dominant and feasible technique in measuring tissue iron content. All different methods used in IO quantification are based on the same empirical observation, ability of iron to enhance transverse relaxation. The optimal use of quantitative IO measurement with different methods needs local expertise (Wood, 2007).

2.7.1. Target organs

The liver is the main focus of quantitative iron measurement, with increasing investigation and many different methods being developed. Non-invasive HIC measurement by MRI enables the evaluation of systemic IO in a wide variety of diseases. Liver biopsies are routinely performed in clinical practice, which has made it possible to calibrate quantitative liver MRI with liver tissue iron concentrations (Brittenham and Badman, 2003; Alústiza et al., 2007; Wood, 2007; Wood and Ghugre, 2008).

Soon after the first HIC calibrations, cardiac iron content was reported of being measured reliably with MRI (Ghugre et al., 2006). Cardiac iron predicts ventricular dysfunction (Anderson et al., 2001) and is a major cause of death in patients with β-thalassemia (Wood, 2008). Because cardiac IO has not been found to have any significant correlation to other non-invasive iron indicators, it has been recommended to be measured directly by MRI (Chirnomas et al., 2008; Leung et al., 2009). Furthermore, cardiac biopsies are generally not recommended for screening as an invasive procedure and for sampling error possibility. Hence, MRI has become the method of choice in assessing cardiac IO. Nevertheless, the frequency and clinical importance of cardiac IO in other disorders, except for thalassemia, is yet not fully established and further studies are required (Wood, 2008).

In addition to liver and heart, iron accumulation can be measured in other target organs with the same basic principle of transverse relaxation enhancement, e.g., in kidney, pancreas and brain, although these measurements are less frequently used (Wood, 2007; Papakonstantinou et al., 2009).

2.7.2. Biophysical relaxation mechanisms of iron

Tissue iron accumulation causes enhancement of transverse relaxation and gradual signal intensity (SI) reduction, which is more pronounced on heavily T2- or T2*-weighted MR images. This SI reduction can be measured directly from a single MR image or calculated with the MR imaging constants. T2 and T2* correlate inversely and their reciprocals (R2 and R2*) correlate directly with tissue iron concentration. Empirical findings indicate that tissue iron deposits in an external magnetic field are magnetised so that the protons influenced by this magnetic heterogeneity disrupt coherence, thus enhancing transverse relaxation (Wood, 2007). This magnetic heterogeneity occurs when paramagnetic iron-containing particles, such as ferritin and hemosiderin, become partially aligned with the applied field, inducing microscopic

field gradients that dephase nearby water protons (Brooks et al., 2001). Although the strong signal drop due to iron particles is empirically well established and qualitatively understood, the exact physical relaxation mechanisms are still somewhat under consideration (Atlas, 1996; Wood, 2007; Gossuin et al., 2004).

The darkening of liver SI in hepatopathy was detected already in the first liver MRI publications (Doyle et al., 1982), although at the time, it was not certain which MRI constant of the target tissue was responsible for the signal drop. SI alterations can be caused by either proton density or by exponential proportions of TR/T1 or TE/T2 to the following equation (Anderson 2011):

$$SI = NH \cdot (1 - e^{-TR/T1}) \cdot e^{-TE/T2}, \qquad \text{[Eq. 3.]}$$

where NH is proton density, the first exponent reflects T1 component and second T2 component in the SI creation. This is the principle equation for MR imaging and quantitative MR measurements. As the MR instrumentation evolved, the specific constants could be quantified more precisely, and namely the transverse T2 relaxation was found to be surprisingly strong compared to much weaker enhancement of longitudinal T1 relaxation. Quite early on, T2 time was found to be the principal cause of the iron-induced MR effect (Stark et al. 1985). Thus the classic dipole-dipole theory by Bloembergen et al. in 1948 could not explain this evidently more versatile transverse relaxation mechanism. Furthermore, the relaxivity of iron-loaded tissue was found to be superior in solid tissues compared to solutions (Stark, 1991). The increased relaxation in tissue was thought to partly be explained by protein clustering (Gossuin et al., 2007).

The diffusion of water in gradients of this inhomogeneous iron-rich magnetic environment was proposed as an explaining mechanism. It is not cancelled by the 180º refocusing pulse, and subsequently can be used to explain T2 relaxation in spin-echo imaging (Stark, 1991; Atlas, 1996). Other theories were also suggested (Atlas, 1996; Gossuin et al., 2004; Ghugre et al., 2005; Gossuin et al., 2007), such as proton or chemical exchange models, which have been used to explain the ferritin relaxation, especially in solutions *ex vivo*, although the individual models *per se* may be incomplete (Ghugre et al., 2005; Ghugre and Wood, 2011). In certain limits, the chemical exchange theory was also found to match the two diffusion theories of outer sphere and mean gradient diffusion. Along with the evolution of these theories, more precise knowledge was obtained. For example, in the multi-echo T2 imaging performed with train of echoes, echo spacing was demonstrated to be one of the factors affecting T2 time (Brooks et al., 2001). In iron-loaded tissue, the water diffusion related dephasing is considered a key factor, which seems to even explain the variation of the relaxivity among different investigators with different modification of T2 sequence (Ghugre et al., 2005). This hypothesis was recently modelled with a simulated virtual liver phantom, which assessed susceptibility based diffusion losses (Ghugre and Wood, 2011). Importantly, the model could also predict the empirically determined linear and curvilinear relationship of R2* and R2 with HIC, respectively (Wood et al., 2005; St. Pierre et al., 2005). Variation in iron particle scale contributed to the curvilinearity of the R2-iron relationship. They postulated that diffusion-mediated susceptibility losses dominate contrast mechanisms for clinical R2-iron

measurement. Furthermore, iron susceptibility, iron distribution, and proton mobility (diffusion) are the most important characteristics in modelling the transverse relaxation in clinical HIC measurement range, and also in other iron-overloaded tissues. Thus, the improved knowledge of the relaxation mechanisms facilitate the understanding of R2/R2*-iron concentration calibration curves and may help in distributing the method to different imaging centres and to different tissues (Ghugre and Wood, 2011).

2.7.3. The development of MRI-based iron measurement

At the beginning of the development of clinical MR scanners in 1983, it was recognised quite early on that the signal drop caused by IO could be correlated to chemical measurement of HIC. The feasibility of T2 measurement in patients was first demonstrated in 1985 (Stark et al., 1985). Thereafter, a relative liver SI measurement of a conventional T2-weighted single spin-echo image was calibrated in HIC of 48 patients. At that time, the variation of the method, which was obtained with developing instrumentation and only one echo sequence, was regarded too wide for clinical use. A CT-based method was also concluded to be inefficient in accuracy (Bonkovsky et al., 1990).

Therefore, gradient echo-based MR imaging was first evaluated as a more rapid technique, potentially more sensitive to low iron concentrations. Ernst et al. reported the first promising results concerning the calibration of two gradient echo sequences with HIC (Ernst et al., 1997). Bonkovsky et al. then optimised further gradient echo imaging for iron quantification (Bonkovsky et al., 1999). They tested the efficiency and sensitivity of six different gradient sequences of different TE, TR and FA. With the most optimal sequence, they were able to produce encouragingly good correlations between relative SI ratios of liver (in relation to background noise) and B-HIC measurement.

More recently, with the use of modern 1.5T equipment, all the imaging methods were developed side by side; the relative SI methods (Gandon et al., 2004), gradient echo imaging with the calculation of a constant of T2* (Anderson et al. 2001; Anderson, 2011) and spin-echo T2 imaging with multiple echoes and calculations of T2 constants (Clark et al., 2003; St Pierre et al., 2005; Wood et al., 2005). The increase in the magnetic field strength enabled better sensitivity to iron, as iron relaxation rates R2 and also R2* were found to be linearly dependent on the main magnetic field strength (Vymazal et al., 1992; Li et al., 2004: Storey et al.2007). The challenge was to optimise a feasible and robust method. Optimising a method would require validation for accuracy but also for inter-scanner and inter-centre transferability to gain a reliable and transferable method to be used in different scanners of variable instrumentation, software and gradients (Brittenham and Badman, 2003).

2.7.4. SI ratio method

SI ratio-based methods in iron load measurement are relatively straightforward to introduce. However, the assessment of the absolute SI measured from the target tissue, especially from gradient echo sequences, would be equivocal, due to inter-scanner and

inter-study differences. Because of that, the SI of a target organ should be divided with a signal of a reference tissue or by background noise. The reference tissue chosen should not be affected by iron accumulation (Alústiza et al., 2007). Several authors focused on the liver-to-paraspinous-muscle ratio, although also liver-to-air, noise, spleen and fat have been used. Gradient echo based sequences are thought to be more sensitive than spin-echo imaging in iron assessment with SI ratio methods (Bonkovsky et al., 1990; Bonkovsky et al., 1999).

Gandon et al. (Gandon et al., 2004) introduced the first widely used and calibrated feasible MR method based on the SI ratio. The calibration was done in 174 biopsy specimens with B-HIC assessment. Liver-to-paraspinous-muscle SI ratios from five different gradient echo based MR images sensitive to hepatic iron were integrated in an algorithm. Thus, actual hepatic iron concentrations (μmol/g dry liver weight) were evaluable non-invasively and reliably (Gandon et al., 2004; Olthof et al., 2007). However, the gradient echo methods are sensitive to iron, and they might saturate, i.e., have too low a signal in the patients with highest IO. One study did raise questions about the accuracy among the patients with the highest IO, i.e., over 11.2 mg/g =201μmol/g (Rose et al., 2006), but the imaging with a measurement modification showed an excellent reproducibility by Alústiza et al. in a study of 112 patients (Alústiza et al., 2004) and also in the validation group of 35 patients with IO less than 349 μmol/g (Gandon et al., 2004). It was still suggested that the calibration of different equipment with phantoms might be recommendable for validating the transferability to other sites (Alústiza et al., 2004).

2.7.5. T2* method

T2* or the reciprocal R2* (=1/T2*) is calculated from gradient echo images, which makes it more susceptible than T2 to iron concentrations. If the signal reduction due to this phenomenon is very rapid, it can lead to virtually total loss in signal intensity due to strong local variations in magnetic field, defined as *susceptibility effects*. For example, iron oxide particles have extremely high magnetic permeability (*susceptibility*), i.e., large magnetisation, which is most efficiently observed in strongly T2*-weighted images (Atlas, 1996; Chavhan et al., 2009).

In the liver, relaxation rate R2* has been validated as a potential method for IO measurement, with linear dependence with tissue iron concentrations. The first T2* assessment, which was *in vivo* liver biopsy-calibrated, showed a promising correlation in 30 patients and the method seemed feasible for both the cardiac and hepatic iron assessment (Anderson et al., 2001). Afterwards, a larger study was conducted and R2* was found to correlate linearly with HIC (Wood et al., 2005). The iron-dependent component of R2* is demonstrated to be linearly dependent also on the magnetic field strength (Storey et al., 2007).

The T2* method is used as the principal method for cardiac iron assessment. R2* and T2* are better suited to iron quantification in tissues where iron concentrations are relatively low, such as in myocardium (Porter, 2005). T2* sequences are also faster than spin-echoes, which additionally favour the T2* over T2 imaging in cardiac iron

quantification (Anderson, 2011). A *postmortem* study showed that R2* and R2 increase proportionally with myocardial iron concentrations, thus being suitable for cardiac iron quantification (Ghugre et al., 2006). Cardiac T2* has been validated widely in iron-overloaded patients with thalassemia (Wood, 2008) with good inter-study (Westwood et al., 2003a; Ghugre et al., 2006), inter-scanner (Westwood et al., 2003b) and inter-centre (Westwood et al., 2005; Tanner et al., 2006) reproducibility.

2.7.6. T2 method

T2 and R2 for iron measurement are used mainly in the liver (Wood, 2007). In the myocardium, T2 has been promising, but the use has been more limited (He et al., 2008; He et al., 2009; Wood, 2008).

Numerous studies have demonstrated the feasibility of liver R2 imaging in patients (Clark and St. Pierre, 2000; Voskaridou et al., 2004; Carneiro et al., 2005; Alexopoulou et al., 2006). R2 did not seem to be affected by underlying microarchitecture of the tissue (Clark et al., 2003). Furthermore, R2 measurements yielded to higher tissue iron concentrations than corresponding gradient echo measurements. T2 and R2 were able to quantify iron concentrations as high as 40 mg/g, *i. e,.* 700 μmol/g liver dry weight (St. Pierre et al., 2005). This wider quantitative concentration range is a clear advantage of R2 imaging compared to gradient echo-based imaging (Wood, 2007).

The studies performed by St. Pierre et al. and Wood et al., were the largest and most tightly controlled for liver R2 measurement (St. Pierre et al., 2005; Wood et al., 2005). They showed that the R2 method is a feasible and accurate imaging method for IO measurement. The method of St. Pierre et al. was the first commercial post-processing method for iron concentration measurement, and also obtained FDA approval (St. Pierre et al., 2005; Wood and Ghugre, 2008). These studies showed that R2 correlated non-linearly with HIC. The non-linear nature of the R2 calibration curve was consistent despite of the differences in technique and instrumentation. This was recently validated by Ghugre et al. using a Monte Carlo model, which took account of iron scale and distribution, intrinsic magnetic susceptibility of iron particles and the type of MRI method. This experiment was able to explain the biophysical elements of the linear R2* and curvilinear R2 calibration curves. Still, the calibration values can be affected by the type of R2 method, so it was suggested that the R2 calibrations gained cannot be straightforwardly used with any other R2 method (Ghugre and Wood, 2011).

3. AIMS OF THE PRESENT STUDY

The objectives of this thesis were to establish a quantitative and accurate MRI-based method for hepatic iron concentration measurement, to develop modifications suitable for daily clinical practice and to explore possible iron-related complications. The specific aims were:

1. To *in vitro* and *in vivo* validate two quantitative MRI-based methods, liver-to-muscle SI and R2*, for hepatic iron concentration measurement, and to calibrate R2 in order to implement a robust, transferable and non-invasive method to measure hepatic iron concentration.

2. To determine the accuracy and inter-reader reliability of in-phase and out-of-phase MR imaging in liver iron overload assessment for a rapid body iron overload estimation. Previously validated quantitative MRI-based hepatic iron concentration measurement was used as a reference standard.

3. To evaluate the number of transfused red blood cell units and ferritin in the estimation of liver iron concentration in patients with hematological disorders, and ferritin in all patients by using quantitative liver MRI as a reference standard.

4. To assess the degree and mutual relationship of cardiac and hepatic iron overload with quantitative MRI.

5. To assess the prognostic impact of pre-transplant iron overload measured with MRI on post-transplant complications: severe infections, graft-versus-host disease and mortality in patients with hematological disorders undergoing allogeneic stem cell transplantation.

4. SUBJECTS AND STUDY DESIGN

The overall study outline with material and methods is presented in Table 2. The MR imaging was conducted between 2004 and 2011 in four 1.5T MR scanners. The study IV patients were followed up until March 2012.

4.1. Patients and healthy subjects

In total 159 subjects, 142 patients and 17 healthy subjects took part in this study (Table 3). 108 patients suffered from hematological diseases, 33 patients suffered from chronic liver diseases and one patient had thalassemia intermedia. Adult patients were included in the study due to a clinical suspicion of either transfusional IO or IO-related chronic liver disease (I-III). Exclusion criteria were age under 18 years (I-IV) and chelation treatment (III). AlloSCT recipients with hematological disease were included in the follow-up study (IV).

Hematological diseases consisted of hematological malignancies (AML, ALL, MDS, myelofibrosis, chronic leukemia, lymphoma, myeloma) and SAA. Chronic liver diseases consisted of hepatopathy (for example, chronic viral hepatitis, alcohol-induced hepatopathy and non-alcoholic fatty liver disease) and hereditary hemochromatosis (type 1, n=4). This single-centre study was conducted according to the Declaration of Helsinki guidelines. Written informed consent was obtained and the study protocol was approved by the local Ethics Committee of the Hospital District of Southwest Finland.

4.2. Design of the studies I-III

The patients for liver and cardiac IO assessment by MRI were screened to join the study at the Department of Medicine, Turku University Hospital, Finland. The MR images were obtained after 4 hours of fasting at the Department of Radiology, Imaging Centre of Southwest Finland, Turku University Hospital. The two different MR studies for the same subject were principally scheduled for the same day with the maximum interval of 30 days in all study patients or subjects. The mean interval of two MR studies conducted was 2 days in M-HIC inter-scanner validation (study I), 1 day in the inter-scanner validation of in-phase and out-of-phase imaging (study II) and 0.25 days between cardiac and liver calibration (study III). The R2 imaging and in-phase and out-of-phase imaging were conducted simultaneously with the reference standard imaging, M-HIC. The liver phantoms and pork muscle mimicking paraspinous muscles were imaged fresh in all three MR scanners. All phantoms were imaged within 8 hours after they were made at the standard temperature of 21° C.

First, the liver MR imaging of the patients and voluntary healthy subjects and liver phantom MR imaging were conducted to validate the gradient echo-based M-HIC method for liver iron concentration measurement at the standard scanner (Gandon et al., 2004). The instrumentation, scanner vendor and type of our standard scanner were matching to the MR scanner used in the calibrations of the method by Gandon et al.

The methods of liver-to-muscle SI-based M-HIC and R2* from M-HIC images were *in vivo* inter-scanner validated with the subgroup of the patients in two different MR scanners. and *in vitro* validated in three different MR scanners for transferability testing. M-HIC was correlated *in vivo* with the R2* and calibrated with biochemical HIC assessment in patients with chronic liver disease.

The method was optimised in order to reach higher HIC values. This was done by calibrating M-HIC to the spin-echo transverse relaxation method R2 in order to introduce a new parameter, M-HIC(R2). This methodological development enabled us to measure the HIC values of over 350 μmol/g liver dry weight. The liver R2-method with a wide quantitative range of HIC was suitable to be used in the cross-sectional study to evaluate the relationship between liver and cardiac iron content by MRI and in the prognostic assessment of excessive IO in alloSCT recipients.

After the validation of M-HIC, it was used as a reference standard to evaluate the alternative tools for more rapid and straightforward estimation of IO principally for screening purposes. These rapid means of IO estimation included commonly used MR sequence of in-phase and out-of-phase imaging in addition to clinical parameters of ferritin and transfusion indices. In-phase and out-of-phase imaging was analysed with two novel methods, which we validated.

4.3. Design of the study IV

Adult patients with hematological disorders, who were to be treated with alloSCT, were recruited to this follow-up study at Turku University Hospital, which is one of the two alloSCT centres in Finland (Figure 7). Before the alloSCT, liver and cardiac IO was assessed with quantitative pre-transplant MRI to get accurate information on the iron burden. After the alloSCT, the post-transplant outcome of the patients was followed up to evaluate the relationship between pre-transplant IO and outcome after alloSCT. The median interval between pre-transplant MRI and alloSCT was 11 days (interquartile range of 7-21 days), the mean was 15±14 days. One patient was excluded because the interval increased to nine months. The disease- and patient-depended pre-transplant variables and differences in transplantation procedure were recorded and statistically evaluated as possible confounding factors. In addition to mortality, GVHD and severe infections were recorded during the early post-transplant period days (0-100) and late post-transplant period (101-365).

4.4. Design of the additional data analysis[a]

Additional correlations of HIC with ferritin data (I-IV) and cardiac R2* (III, IV) were performed in the study patients in order to examine the consistency of our results in a larger number of patients. The latest MRI and ferritin data were recorded. Double studies of ferritin and cardiac R2* were excluded. The ferritin values closest to the corresponding MRI were recorded, and the maximum interval between the MRI and ferritin was set as three months.

As a result, a total of 130 patients were included in the correlation analysis between ferritin and HIC and a total of 117 subjects were included in the correlation between cardiac R2* and HIC, including 93 patients with hematological disorders.

Table 2. Study outline.

Study	I	II	III	IV	Additional data[a]
Subjects (n)	33	54	87 45 calibration subgroup	67	130 ferritin subgroup 117 cardiac MRI subgroup
Phantoms (n)	16				
Liver biopsy (n)	7				
MRI methods	M-HIC[1] M-HIC(R2*) M-PIC Phantom R2*	M-HIC[1] M-HIC(R2*) Visual grading rSI	M-HIC(R2) Cardiac R2* [2]	M-HIC(R2) Cardiac R2* [2]	M-HIC[1] M-HIC(R2*) M-HIC(R2) Cardiac R2* [2]
Other methods	B-HIC		Ferritin RBC units		Ferritin
MR scanners	GE Signa Horizon LX[3] Siemens Magnetom Symphon[4] Siemens Magnetom Avanto I[4]	Siemens Magnetom Symphony[4] Siemens Magnetom Avanto II[4]	Siemens Magnetom Symphony[4] Siemens Magnetom Avanto II[4]	Siemens Magnetom Avanto II[4]	GE Signa Horizon LX[3] Siemens Magnetom Symphony[4] Siemens Magnetom Avanto II[4]
Aim	Quantitative MRI method validation	In-phase and out-of phase method validation	Liver IO to heart IO Liver IO by clinical markers R2 calibration: Supplementary Information	Prognostic impact of IO on post-transplant mortality and morbidity	Liver IO to ferritin Liver IO to cardiac IO

[1] HIC method introduced by Gandon et al., 2004
[2] Cardiac T2* introduced by Anderson et al., 2001
[3] GE Medical Systems, Milwaukee, WI, USA
[4] Siemens Medical Solutions, Erlangen, Germany
[a] Additional data analysis in all patients

Table 3. General characteristics of the study subjects

Study	Study group	n (female/male)	Plasma ferritin µmoll/l median (range)
I	Patients	27 (9/18)	968 (19-15261)
	Hematological diseases	10 (4/6)	2120 (968-15261)
	Chronic liver diseases	17 (5/12)	581 (19-2171)
	Healthy subjects	6 (4/2)	
II	Patients	42 (12/20)	1654 (61-8084)
	Hematological diseases	29 (10/19)	2227 (204-8084)
	Chronic liver diseases	13 (2/11)	948 (61-4873)
	Healthy subjects	12 (9/3)	
III	Patients[1]	75 (32/43)	1568 (17-6680)
	Hematological diseases	59 (27/32)	1799 (17-6680)
	Chronic liver diseases	15 (4/11)	802 (61-3105)
	Healthy subjects	12 (8/4)	
IV	alloSCT recipients	67 (34/33)	1130 (17-8878)

[1]Including a thalassemia intermedia (n=1)

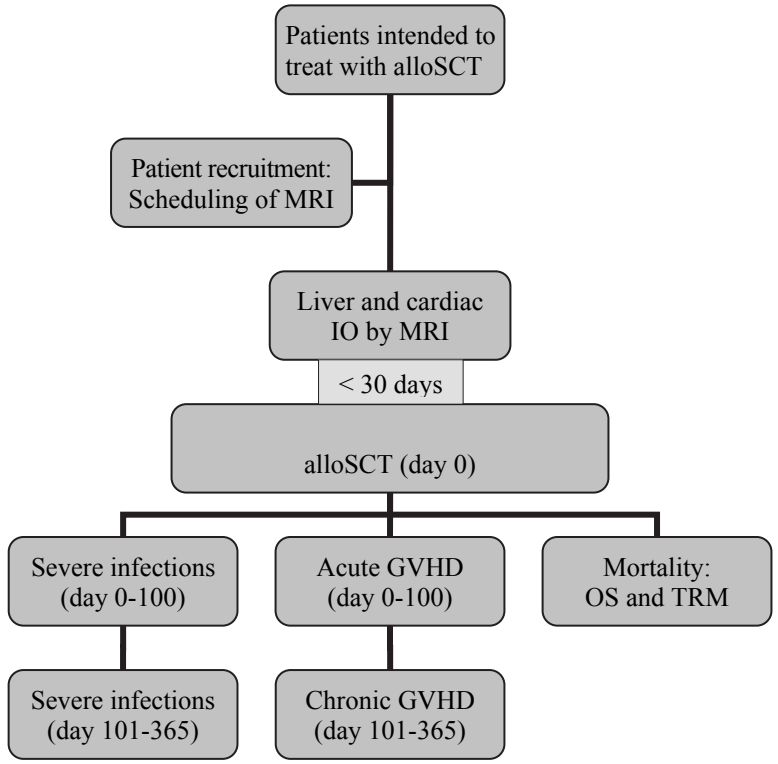


Figure 7. Study design (IV) outlined. The outcome was recorded as severe infections, GVHD and survival; overall survival (OS) and TRM were recorded separately in the follow-up time of surviving patients (median: 13.1 months, IQR: 6.8-23.9)

5. MATERIALS AND METHODS

5.1. Phantoms (I)

The liver phantoms were gel-based solid phantoms that were prepared to simulate the NMR characteristics of normal and iron-overloaded human liver tissue with increasing iron concentrations.

First we prepared the phantom corresponding to normal liver T1 and T2 times. This was performed by adding gradual amounts of $MnCl_2$ and agarose solution in distilled water and performing repeated T1 and T2 measurements to gain the optimal concentrations: 0.5% agarose (QA-Agarose TM, Molecular Biology Grade, Q-BioGene) solution with 100 ml distilled water was doped with 0.17 mM $MnCl_2$ ($MnCl_2$ x $4H_2O$, min. 99%, Sigma-Aldrich). The mixture was then heated in a microwave oven to produce gel. The magnetic features of iron in aqueous solutions are not comparable with iron in solid tissue. Therefore, we performed the phantom in the form of gel. These compounds were chosen because agarose makes the phantom more solid and additionally shortens especially the T2, whereas Mn shortens particularly the T1. The chosen T2 (47 ± 0.4 ms) and T1 (690 ± 15 ms) for the phantoms were comparable to the values measured for normal human liver (Thomsen et al. 1994). T1 and T2 were measured with 1.5T GE Signa Horizon LX MRI scanner by using the spin-echo SE sequence with various TE and TR.

In order to mimic the human liver tissue with IO, increasing amounts of iron were added into the phantoms during the preparation process. Sixteen samples were prepared by dissolving increasing amount of ammonium iron (II) sulphate ($(NH_4)_2[Fe(SO4)_2]$ x $6H_2O$, puriss. p.a. min 99%, Riedel-de Haen, Sigma-Aldrich Laborchemikalien GmbH) in $MnCl_2$ spiked agarose gel base phantoms before the sample heating. The iron concentration of samples varied between 19 μmol/g and 273 μmol/g.

5.2. Biochemical HIC (I)

Seven patients underwent liver biopsy with clinical indications. All biopsies for biochemical HIC (B-HIC) measurement were performed within a 30-day interval from M-HIC imaging at GE Signa scanner. The biopsy patients were suspected of having parenchymal disease and possible iron accumulation of the liver. Biopsies were obtained in a routine clinical manner from the right liver lobe: Two biopsy specimens were obtained with an 18-gauge biopsy needle (Temno, Cardinal Health, Inc., Dominican Republic) and placed in a 10% buffered formalin solution. Both histological analysis and B-HIC were performed.

For the B-HIC measurement, an approximately 1 cm core of the needle biopsy specimen was dried in a microwave oven, ashed and weighed. The average weight of

the specimen in biochemical analysis was 1.34 mg. The iron concentration analysis of the specimens was performed in Åbo Academy by using a mass spectrometer, ICP-MS (PerkinElmer, Canada). Because the liver tissue specimens were relatively small and it was known that tissue iron could dissolve in the formalin solution according to Villeneuve et al. (Villeneuve et al. 1996), the dissolved iron mass needed to be assessed in order to avoid this confounding factor. To achieve this, we measured the baseline iron concentration of the formalin solution by the mass spectrometer. Then, the volume and iron concentration of the formalin solution of each sample were measured, the amount of dissolved iron was calculated and the final B-HIC values were assessed in each sample.

5.3. Clinical parameters of the patients

5.3.1. Ferritin, CRP and ALT (III,[a])

Plasma concentration of ferritin (µg/l) was measured electrochemically (MEIA). CRP (mg/l) was measured with immunoturbidimetric assay and ALT (U/l) according to IFCC recommendation on Roche Modular PPEE analyser (Roche Diagnostics GmbH, Mannheim, Germany).

The median interval between the blood tests (ferritin, CRP and ALT) and MRI was 18 days (95% CI, 18.4-30.0; range 0-126 days), recorded in 68/75 patients (III). Ferritin[a] was additionally recorded in all eligible patients within three months from the liver MRI: in 21/27 (I), 38/42 (II), and in 67/67 patients (IV).

5.3.2. RBC transfusion indices and MDS (III)

The transfusion dates and the number of RBC units transfused were collected in patients with hematological disorders via a transfusion database (Trace Line®) founded in 1996, and patient reviews.

In addition to the impact of transfusion load on IO, we also investigated whether receiving RBC transfusions in a large volume or a definite subgroup of MDS would *per se* have an additional impact on iron accumulation. The transfusion variables for backward linear regression analysis were: (1) Transfusion load defined as total cumulative number of RBC units received, (2) The number of high-dose sessions (HDS) of transfusions, which equalled to the number of transfusion sessions, in which ≥4 RBC units were transfused and (3) The number of high-dose units of RBC received (HDU), which equalled to the number of cumulative RBC units of >2 transfused in a single session. The patients with MDS were included in the analysis as an independent variable because they were suggested of having increased susceptibility to iron due to active but ineffective hematopoiesis (Fenaux and Rose, 2009). Transfusion dependency, defined as duration of regular transfusions (years), was used to characterise the patient RBC requirement and the severity of the anemia at the specific point in time.

5.3.3. Other definitions for patients with hematological disorders (III, IV)

Disease duration was defined as the time interval between the diagnosis from histological analysis and the MRI (III) or alloSCT (IV). Disease severity was defined as standard or advanced according to prior risk score analysis (Gratwohl et al., 2009) (IV). Diagnoses of the patients were categorised as ALL, AML, AA, lymphoma, MDS and myelofibrosis (III) and the alloSCT recipients were categorised into four clinically distinct groups: (1) AL, (2) MDS and myelofibrosis, (3) AA and (4) other hematological malignancy, i.e., chronic leukemia, lymphoma and multiple myeloma (IV).

After alloSCT (IV), patient outcomes were recorded for the time periods covering 0-100 days (early post-transplant) and 101 days-12 months (late post-transplant). The severity of infections was graded from 0 to 5 according to the Common Terminology Criteria for Adverse Events (CTCAE) version 4.0 issued by National Cancer Institute, May 29, 2009 (CTCAE, 2009). Infections were categorised as severe infections (grade 4-5) and non-severe infections (grade 0-3). The modified Seattle Glucksberg criteria were used for grading the severity of acute GVHD in grades 0 to 4 on the basis of liver, bowel or skin involvement (Przepiorka et al. 1995), and the maximum score of acute GVHD of liver, bowel or skin was categorised into two groups: grade 0-1 and 2-4 acute GVHD. Chronic GVHD was categorised into three groups as no GVHD, limited GVHD or extensive GVHD (Ratanatharathorn et al., 2001).

5.4. MR imaging methods

5.4.1. Liver-to-muscle SI and R2* methods (I, II)

5.4.1.1. MR imaging (I, II)

44 gradient echo-based MR imaging examinations of liver-to-muscle SI method were conducted in 33 subjects using the 1.5T GE Signa Horizon LX referred as the standard scanner. A subgroup of patients (n=14) was scanned using the same method at the standard scanner and Siemens Magnetom Symphony for *in vivo* inter-scanner validation. Sixteen phantoms and pork muscle were imaged with the same imaging method for *in vitro* inter-scanner validation at the three MR scanners: the standard scanner, Siemens Magnetom Symphony and Avanto (I). 54 subjects underwent this MRI protocol using the Siemens Magnetom Symphony (n=37) or Avanto scanners (n=17) in study II.

The MR imaging protocol consisted of five sequences introduced earlier (Gandon et al., 2004). In all subjects and phantoms, the sequences were obtained in the transverse plane with body coil. The four T2*-weighted breath-hold gradient echo sequences were scanned with a body coil and the following parameters: TR of 120 ms, TE of 4, 9, 14 and 21 ms and flip angle (FA) of 20°. The T1-weighted gradient echo sequence was scanned with TR of 120 ms, TE of 4 ms and FA of 90°. The matrix was 256x128, the field of view (FOV) 400 mm and slice thickness 10 mm.

In all subjects, the sequences of four slices were positioned at the portal level in transverse plane in order to reach the optimal area of homogenous right liver lobe, which is also the target in routine liver biopsies. In the MR imaging of 16 liver phantoms, three slices were positioned in transverse plane to cover the phantoms. The phantoms of different iron concentrations were piled upon each other, and the phantoms formed an area of 28 x 25 cm in axial plane. The fresh pork muscle was positioned below the phantoms in order to mimic the actual transversal positions of the liver and paraspinous muscles in human subjects (Figure 8).

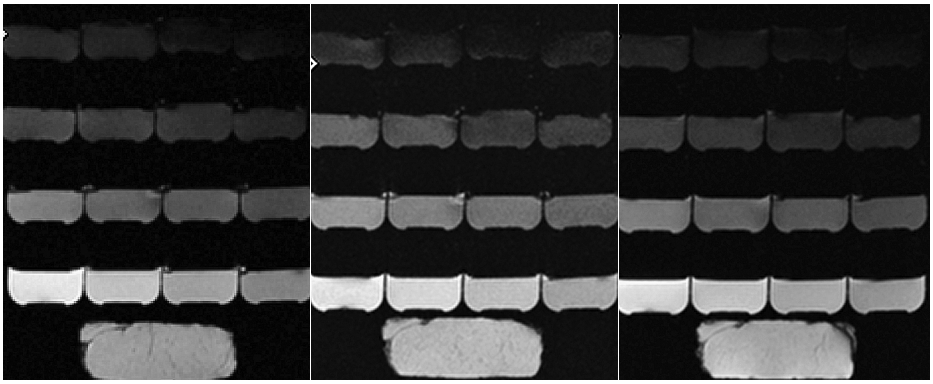


Figure 8. *In vitro* validation of the gradient echo-based methods, M-HIC and R2*. Phantoms with increasing iron concentration imaged at the TE of 14 at the 1.5T standard scanner (left), Siemens Symphony (middle) Avanto (right) scanners. Fresh muscle at the lowest. Reprinted from study I.

5.4.1.2. Data analysis

We measured the mean SI of region of interests (ROI) for the liver phantoms, the pork muscle (ROI of 253 mm² in Siemens and 256 mm² in the standard equipment), right part of the right lobe of the human liver (average ROI of 300 mm^2) and both paraspinous muscles (ROI of 100-300 mm^2). The measurements were optimised for patients with various muscle and liver sizes for a homogenous ROI. Five consistent ROIs were obtained from each of the five sequences in each subject and phantom: three liver and phantom SI measurements and two paraspinous and pork muscle measurements. Liver and muscle measurements were made in the same slice *in vivo,* corresponding to phantom and pork muscle *in vitro* measurements.

5.4.1.2.1. M-HIC and M-PIC

Based on the SI ratios of liver-to-paraspinous-muscle and liver-phantom-to-pork-muscle, the MRI-based iron concentrations for the liver and liver phantoms were calculated. We performed the calculations applying the algorithm by Gandon et. al (http://www.radio.univ-rennes1.fr/Sources/EN/HemoCalc15.html), which is presented in the University of Rennes website (Gandon et al., 2004). As a result of the analysis, we obtained iron concentrations for the human liver, M-HIC (MRI derived hepatic iron concentration, μmol/g dry liver weight) and for liver phantoms, named as M-PIC (MRI derived phantom iron concentration, corresponding μmol/g dry liver weight). The

method measured iron in 5-10 μmol/g increments according to the algorithm. The sequences were designed to have increasing susceptibility to iron. The least susceptible sequences were responsible for the measurement of the highest iron concentrations, but became saturated in the most severe IO due to the low liver signal-to-noise ratio (SNR). Thus, the quantitative range of the M-HIC method at the standard scanner was limited to approximately 5-350 μmol/g. The upper normal limit was 35 μmol/g (Gandon et al. 2004). HIC>35 μmol/g was defined as liver IO.

*5.4.1.2.2. Liver and phantom R2**

We assessed the mean R2* (1/T2*, s^{-1}) for every patient, healthy subject and liver phantom calculated from the MR images, which were originally obtained with liver-to-muscle SI ratio method. The transverse relaxation times (T2*, ms) of the human livers and liver phantoms were calculated using the ROI measurements from the four T2*-weighted sequences described above with TE range of 4-21 ms. The average liver and phantom SI of the three ROIs at different TEs was used for the T2* analysis. The T2* was assessed with monoexponential function (Anderson et al., 2001):

$$SI(TE) = A \cdot e^{-TE/T2^*}, \quad \text{[Eq. 4.]}$$

where SI(TE) is the signal intensity at particular TE, A is the computational SI at TE of 0.

M-HIC(R2*) defined as MRI-based hepatic iron concentration measured with R2* were assessed: The R2* was expressed as a function of iron concentration (M-HIC) in at the standard scanner. The slope of the straight line gives the value of transversal relaxivity, r2* ((s*μmol/g)-1).

This M-HIC imaging sequence with corresponding liver R2*, i.e., M-HIC(R2*) assessment, was used as a reference standard (II, III Supplementary Information). In 54 patients and healthy subjects, 48 subjects were within the quantitative range of the reference standard, which was 5-320 μmol/g after the standardisation of scanner-specific differences according to study I. Three M-HIC(R2*) values of normal controls below 5 μmol/g were set at the lower quantitative limit of 5 μmol/g. The six values ≥ quantitative range were excluded from the analyses of continuous variables. For the semi-quantitative evaluation of visual grading, the values above 320 μmol/g were set at 325μmol/g.

5.4.2. In-phase and out-of-phase imaging (II)

5.4.2.1. MR imaging

Liver in-phase and out-of-phase imaging was conducted simultaneously with the liver reference standard imaging in each subject. In a subgroup of 19 patients, the in-phase and out-of-phase imaging was conducted at both Magnetom Symphony and Avanto scanners with equal parameters for inter-scanner validation of the rSI method.

In-phase and out-of-phase imaging of the liver was obtained during a breath-hold in transverse plane with a body coil using a T1-weighted dual echo spoiled gradient-

recalled echo sequence. The parameters were the following: TR 168 ms, TE 2.38 ms at out-of-phase and TE 4.76 ms at in-phase, FA 70°, slice thickness 10 mm, matrix size 146 x 256, FOV 400 mm and time of acquisition 25 seconds.

5.4.2.2. Data analysis

The IO assessment from in-phase out-of phase images was evaluated using two methods: visual grading and relative SI (rSI). The analysis methods apply to modern clinical 1.5T equipment, where the first echo pair is routinely chosen, consequently in-phase is obtained with the longer TE and out-of-phase with shorter TE. The visual grading method was applied to evaluate the degree of iron overload semi-quantitatively in all the patients. The rSI was applied for evaluating IO alternatively as a continuous variable within the quantitative range of the validated reference standard M-HIC.

5.4.2.2.1. Visual grading

The visual grading analysis was done according to criteria presented in Table 4. The images were analysed separately by two independent radiologists, reader 1 (two years of expertise in abdominal radiology) and reader 2 (three years of expertise in musculoskeletal radiology). They were blinded to patients' clinical history, laboratory findings and iron status. The readers were instructed to emphasise in-phase criteria especially if they suspected fatty liver disease after comparing the image pair.

5.4.2.2.2. rSI

The rSI was calculated as SI difference between the images at different TE in relation to the signal level of the images with the shorter TE. It can be assessed with the following equation:

$$rSI = \frac{SI(out) - SI(in)}{SI(out)} \cdot 100\%, \qquad \text{[Eq. 5.]}$$

where *SI(out)* is signal intensity of out-of-phase and *SI(in)* is signal intensity of in-phase images. This index was applied as a continuous iron indicator to test if there was enhanced transverse relaxation and SI drop found on *SI(in)* with the longer TE compared to *SI(out)*. The SI difference between out-of-phase and in-phase was calculated by subtraction of the whole liver as *SI(out*) minus *SI*(in). The most homogenous slice of the subtracted sequence was chosen for measurements. Three regions of interest with an area of 100-200 mm² were obtained from subtracted images and from out-of-phase images of the axial slice at the right part of the right liver lobe, avoiding vessels and possible artefacts. The rSI difference of three regions of interest and the average rSI were calculated for each subject. If there was no signal in the subtracted images, the SI difference was calculated additionally as *SI(in)* minus *SI(out)* to detect fatty liver, if present.

Table 4. The criteria of visual grading method for IO assessment

Grade	Description	Criteria	
		In-phase images	In-phase and out-of-phase images
0	normal iron concentration	Liver SI is greater than muscle SI	Liver SI of in-phase images and out-of-phase images is about the same
1	minor IO	Liver SI is about as low as muscle SI	Liver SI of in-phase images is slightly lower than out-of-phase images
2	moderate IO	Liver SI is lower than muscle SI, although not as low as the background noise	Liver SI of in-phase images is clearly lower than out-of-phase images
3	severe IO	Liver SI is lower than muscle SI, and about as low as background noise	Liver SI of in-phase images is clearly lower than in out-of-phase images

5.4.3. Liver M-HIC(R2) (III, IV)

Spin-echo-based liver R2 was used for HIC assessment, defined as M-HIC(R2), in all 87 study III and 67 study IV subjects. Before the use of the method, R2 was calibrated to our reference standard in study II subjects (III Supplementary Information).

5.4.3.1. MR imaging

The subjects underwent T2-weighted spin-echo-based MR imaging for the calculation of T2 time. The sequence was positioned at the portal vein level in transverse plane in the same manner it was performed in the imaging of signal intensity method of M-HIC. A T2 multi-echo spin-echo imaging sequence with fat suppression was obtained in each subject at 11 different TE times (TE 7.2-86.4 ms, 7.2 ms interval) with the following parameters: TR 2000 ms, FA180°, three slices with the thickness of 10 mm, matrix size 192 x 256, FOV 400 mm, time of acquisition 4:50 minutes.

5.4.3.2. Data analysis

The images of twelve different echo times of each subject were analysed using constant freehand ROI. For SI measurement the ROI was obtained from the right part of the right liver lobe in segments 6 and 7 avoiding vessels and possible artefacts. The mean size of the region of interest was 15.3 cm²±9.7 (III). T2 was assessed with the monoexponential fitting with the following equation (Wood et al., 2005):

$$SI(TE) = A \cdot e^{-TE/T2} + C, \qquad \text{[Eq. 6.]}$$

where C is an offset constant, which is used to compensate image noise and iron-poor regions within the ROI (Wood and Chugre 2008). A is the computational SI at TE of 0.

5.4.3.3. Calibration of R2 with reference standard (III Supplementary Information)

Forty-eight patients and healthy subjects underwent both R2 and validated M-HIC within the quantitative range of 5-320 μmol/g. One R2 value due to incorrect TE values and two incalculably low R2 values of healthy subjects were excluded, leaving

45 subjects into the correlation. The R2 and M-HIC imaging were obtained within a 0-8-day interval.

The calibration of the R2 and M-HIC was performed with an empirical nonlinear analytical fitting. We used the equation introduced in a single spin-echo R2 method calibration (St. Pierre et al., 2005):

$$y = a + bx^{d} + cx^{2d}, \qquad \text{[Eq. 7.]}$$

where y is R2 (s^{-1}), x is the reference standard, M-HIC (μmol/g), a (s^{-1}), b ((μmol/g)$^{-d}$ s^{-1}), c ((μmol/g)$^{-2d}$ s^{-1}) and d are the fitting constants.

5.4.4. Cardiac R2* (III, IV)

5.4.4.1. MR imaging

The MR imaging of cardiac R2* was performed in study III and study IV subjects at 1.5T Siemens Magnetom Avanto scanner. Two patients (III) and seven alloSCT recipients (IV) declined the cardiac MRI. The imaging was conducted directly after liver R2 at the same scanner.

We used the validated T2* method introduced earlier (Anderson et al., 2001) applying ECG gated breath-hold, spoiled gradient-recalled echo sequence (2D FLASH) at increasing echo times (TE 3, 5, 6.5, 8, 10, 12, 14, 17, 20, 23, 26, 29 ms) with the following parameters: TR 500ms, FA 20º, slice thickness 10 mm, matrix size 128x256, FOV 350 mm, seven phase-encoding steps per cardiac cycle, time of acquisition 15-25 s. 40 subjects were scanned with TE of 5-29 and the others were scanned with TE of 3-29. Black-blood preparation pulse (He et al., 2007; He et al., 2008) was applied to reduce artefacts in 79 subjects (III) and all patients (IV). After 3D imaging of ventricles, a single short-axis mid-ventricular slice was acquired repetitively with different TE values in each patient.

5.4.4.2. Data analysis

A 1.0-1.3 cm² region of interest (ROI) was chosen for the signal intensity (SI) measurement from left ventricular septum (Figure 9) (Ghugre et al., 2006). All images were evaluated and selected to avoid movement artefacts. A monoexponential equation (Anderson et al., 2001; He et al., 2007; He et al., 2008) was applied to the cardiac T2* assessment. Because R2* is directly proportional to the iron concentration (Wood et al., 2005), it was chosen as the principal cardiac iron measurement instead of T2*, which correlates nonlinearly and inversely with iron concentrations in a tissue. Cardiac IO was defined as R2*>50 1/s (corresponding T2*<20 ms), which was used as a cut-off point for normal and pathologically elevated cardiac iron (Anderson et al., 2001).

5.5. Statistical analysis

The statistical analyses were performed with GraphPad InStat ver. 3.06 (I-III), GraphPad Prism versions 2.01 5.00, 5.02 (I-IV) (GraphPad Software, Inc., San Diego,

CA, USA) and SAS System for Windows, version 9.2 (SAS Institute Inc., Cary, NC, USA) (II-IV). The normality test of the data and the pairing effectiveness were tested. The level of significance was set at $p<0.05$.

In study I, least square measures linear regression analysis with Pearson correlations (r) was used. The comparison between the means of groups of *in vitro* study were performed using one-way ANOVA to test if there was a significant difference among the results defined for the three 1.5T MRI scanners. The paired T test with two-tailed p values was used for the method comparison with liver biopsies and in the *in vivo* study to compare the results between the two MRI scanners. The Bland-Altman plot and line of agreement was applied to test the *in vivo* transferability testing (Bland and Altman 2003).

In study II, T tests with two-tailed p values and means±SD (standard deviation) were used to express the ages and rSI results in different scanners. Non-parametric Spearman rank correlation coefficients (r) with two-tailed p values were used to evaluate the correlations. The Kruskal-Wallis test was used between the grades of visual grading method. The results were expressed as medians and IQR (interquartile range). The sensitivity, specificity, PPV (positive predictive value), NPV (negative predictive value) with 95% confidence interval (CI) and accuracy (%) of the methods for detecting liver iron overload were measured. The agreement between the readers was evaluated with weighted kappa (κ) coefficient.

In study III, all correlations were obtained using Pearson correlation (r). Unpaired T test and one-way ANOVA with Tukey-Kramer multiple comparisons post-test were used between the groups. The effect of transfusion variables and MDS on hepatic iron concentration in patients with hematological disorders was analysed with univariate and backward linear regression analysis with exclusion criteria of $p<0.10$. The normality of residuals was tested with Shapiro-Wilks test and collinearity diagnostics were performed.

In study IV, the associations of outcomes between categorical groups of low IO and excessive IO were tested using the $\chi 2$ test. All the variables were analysed with logistic univariate and subsequent multivariate regression analysis. First univariate analysis was used to examine the association between the explanatory variables (liver M-HIC(R2), cardiac R2*, age, disease duration, gender, diagnostic subgroup, disease severity, conditioning regimen, donor type and ABORh blood group incompatibility) with severe infections, acute GVHD and chronic GVHD. The explanatory variables with $p<0.10$ by univariate analysis were further included in the subsequent multivariate logistic regression analysis. Pre-transplant M-HIC(R2), cardiac R2*, age and disease duration were analysed as continuous explanatory variables and the others were analysed as categorical explanatory variables. Infections were analysed in two defined categories. The maximum score of acute GVHD of liver, bowel or skin was analysed in two defined categories. Chronic GVHD was analysed with cumulative logistic regression in three categories: no GVHD, limited GVHD or extensive GVHD. The results of the logistic regression analyses are expressed as odds ratios (OR) or cumulative odds ratios (COR) with 95% confidence intervals (CI). The effect of the explanatory variables on mortality was analysed with Cox's regression model.

A receiver-operator characteristic (ROC) curve was applied to search for the optimal threshold for the rSI difference method (II) and for iron indicators of RBC and ferritin (III) for IO detection and to assess the optimal cut-off value of HIC for predicting severe infections (IV): The sensitivity and specificity of each value and the accuracy of the methods (area under curve (AUC)) were calculated and the closest value from the upper left corner of the ROC curve was selected as the cut-off value.

The subgroups of HIC in the study were analysed with one-way ANOVA. The results were expressed as mean±SD. All the linear and nonlinear dependencies of the scatter plots in the study were analysed as goodness-of-the-fit (r^2 and R^2, respectively). The additional data was analysed with least square measures linear regression analysis with goodness-of-the-fit and Pearson correlation.

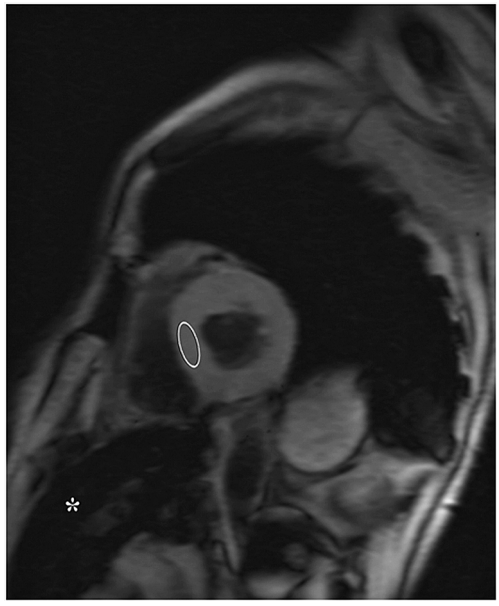


Figure 9. An example of a short axis slice at TE of 8 ms from human heart. The ROI is positioned to left ventricle septum. Reprinted from study III.

6. RESULTS

6.1. Quantitative liver iron measurement methods by MRI (I,III)

6.1.1. In vitro validation of liver-to-muscle SI and R2* (I)

Liver-to-muscle SI (M-PIC) and phantom R2* showed a strong linear dependency with linearly increasing liver phantom iron concentration. The correlations of the fit for M-PIC ($r=0.989$, $p<0.0001$) and for R2* ($r=0.981$, $p<0.0001$) were excellent at the standard scanner. The linear dependency was equally strong for both methods at all three scanners (Figure 10). There were no statistical differences between the mean values of the R2* and M-PIC at the three scanners ($p=0.846$ and $p=0.482$), but the difference between the slopes among the scanners was larger in M-PIC than in R2*. This difference in transferability of the methods between the standard scanner and the two other scanners is illustrated in Figure 11 with the line of equality.

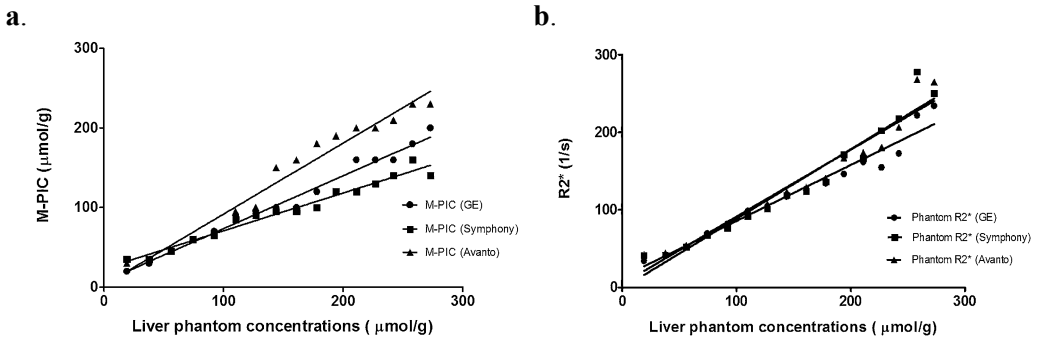


Figure 10. The dependence of **a)** M-PIC and **b)** R2* on the increasing iron concentrations of 16 phantoms MR imaged with different scanners. The GE was defined as the standard scanner. Reprinted from study I.

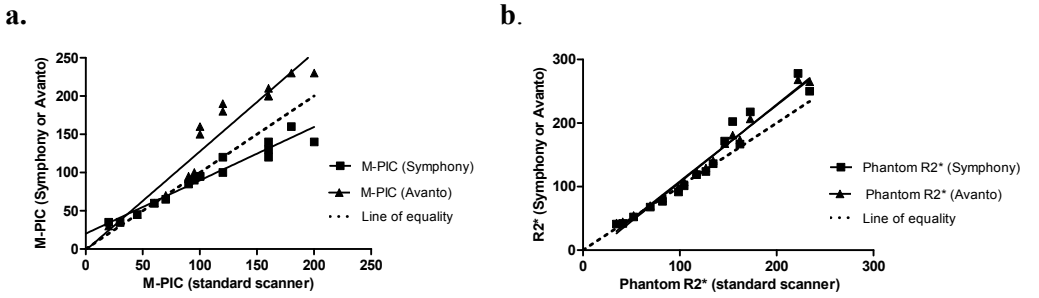


Figure 11. Inter-scanner validation between the standard scanner and Symphony and Avanto scanners of both methods **a)** M-PIC and **b)** R2* in 16 phantoms. The R2* values were closer to the line of equality than the M-PIC values at both scanners.

6.1.2. In vivo validation of liver-to-muscle SI and R2* (I)

In the *in vivo* validation, R2* showed a strong linear dependency with the liver-to-muscle SI method M-HIC at the standard scanner ($r=0.981$, $p<0.0001$). The accuracy of the slope starting point was tested with biopsy-derived B-HIC at the low iron levels of 4-38 µmol/g. All five normal values of M-HIC were true negative, one was true positive IO and one borderline false positive IO (M-HIC: 45 µmol/g). Seven R2* values were incalculably high in iron concentration and excluded due to the low SNR.

For inter-scanner validation, R2* and M-HIC were obtained at the standard and Symphony scanners. M-HIC values were found to be scanner-dependent ($p<0.0001$) and biased in Bland Altman with the mean difference of 32±18 µmol/g (60±56%). In contrast, the mean difference in R2* was 3.5±17 s^{-1} (2.2±9.7%), and no significant difference in R2* was found between the groups imaged at different scanners ($p=0.505$). In order to measure iron concentrations using the different scanners, the two following equations were made, also illustrated in Figure 12: The M-HIC inter-scanner calibration equation

$$\text{M-HIC(standard)}=(\text{M-HIC(Symphony)} -28.3\ \mu\text{mol/g})/1.04 \quad \text{[Eq. 8.]}$$

and the calibration curve between the validated R2* from either one of the scanners and standard scanner M-HIC with the equation

$$\text{M-HIC}(R2^*)=(R2^*- R2^*(0))/ r_2^*, \quad \text{[Eq. 9.]}$$

where $R2^*(0)$ is y-intercept (27±3.5 s^{-1}) and r^*_2 is the relaxivity (1.2±0.04 $(s\ \mu mol/g)^{-1}$). M-HIC(R2*) is M-HIC derived from R2*measurement and thus more transferable among the scanners than the liver-to-muscle SI method.

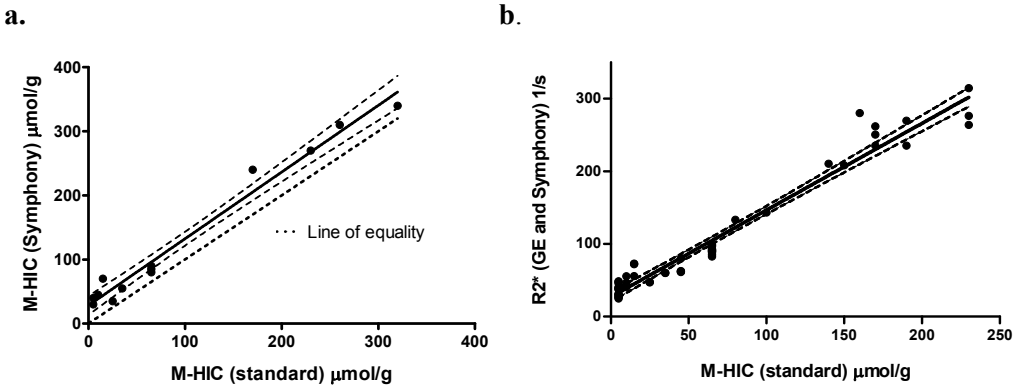


Figure 12. M-HIC inter-scanner and R2* calibration. **a)** Inter-scanner calibration of the liver-to-muscle SI method (M-HIC) illustrates a strong dependency between the two scanners ($r=0.989$, $p<0.0001$) and the systematic bias from the line of equality. **b)** Calibration curve for the assessment of M-HIC(R2*) with a strong dependency between M-HIC and R2* ($r=0.979$, $p<0.0001$). Modified and reprinted from study I.

6.1.3. R2 calibration for the assessment of M-HIC(R2) (III Supplementary Information)

In order to measure HIC in a wider quantitative range, a spin-echo imaging was obtained for transverse relaxation rate R2 evaluation. Then the R2 was calibrated with the reference standard, the validated M-HIC imaging. The empirical nonlinear fitting in a wide quantitative range according to an earlier introduced R2 formula was used (St. Pierre et al. 2005). We found the best fit values for the constants a, b, c, and d as follows: a=31.14 s^{-1}, b=-20.18 s^{-1} $(\mu mol/g)^{-0.2542}$, c=6.772 s^{-1} $(\mu mol/g)^{-0.5084}$ and d=0.2542. Using this curve, R2-based hepatic iron concentration (M-HIC(R2)) with constants A, B, C and D can be calculated with following equation:

$$\mathrm{M-HIC(R2)}=\left(A+\sqrt{B\cdot(\mathrm{R2})-C}\right)^{D}, \qquad \text{[Eq. 10.]}$$

where the values for the constants are: A=1.490, B=0.1477, C=2.378 and D=3.934. This fitting is illustrated in Figure 13.

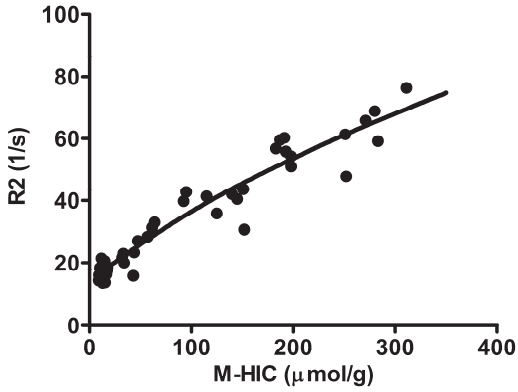


Figure 13. The nonlinear empirical calibration curve between R2 and reference standard M-HIC providing R2-based HIC, M-HIC(R2), with a strong dependency (R^2=0.929 p<0.0001). Reprinted from Supplementary Information.

6.2. In-phase and out-of-phase imaging in IO assessment (II)

The visual grading by reader 1 (r=0.953, p<0.0001), reader 2 (r=0.946, p<0.0001) and rSI method (r=0.772, p<0.0001) correlated significantly with the reference standard HIC (Figure 14).

In the visual grading of in-phase and out-of-phase imaging (Figure 15) the differences between the medians of the four grades were significant with excellent inter-reader agreement (weighted κ=0.962). Both readers detected grades 1-3 liver IO with sensitivity of 85 per cent (95% CI: 74-92%), specificity of 100 per cent (95% CI: 81-100%), PPV of 100 per cent (95% CI: 94-100%), NPV of 64 per cent (95% CI: 44-81%), and accuracy of 88 per cent. Grouped grades of 2 and 3 were accurate in finding

HIC>151μmol/g with a sensitivity and PPV of 100 per cent (95% CI: 91-100%) and a specificity and NPV of 100 per cent (95% CI: 92-100%).

In the rSI method, the correlation to M-HIC is demonstrated (Figure 14 b) with the following linear fit:

$$\text{M - HIC(rSI)} = 3.562 \cdot \text{rSI} + 26.99, \qquad \text{[Eq. 11.]}$$

where M-HIC(rSI) is an estimation of HIC by rSI. The optimal threshold for rSI method to detect iron overload was 10 per cent to reach the best possible sensitivity of 85 per cent (95%CI: 68-95%) and specificity of 100 per cent (95%CI: 84-100%), PPV of 100 per cent (95%CI: 88-100%), NPV of 81 per cent (95% CI: 61-93%) and accuracy of 91 per cent. Three out of seven subjects with detectable liver fat, suffered also from IO (mean M-HIC values of 45.7±45.3 μmol/g) and were false negative results decreasing the sensitivity. In the inter-scanner validation, the mean of the difference was 0.005±0.051 and no significant difference was found between the scanners (p=0.675).

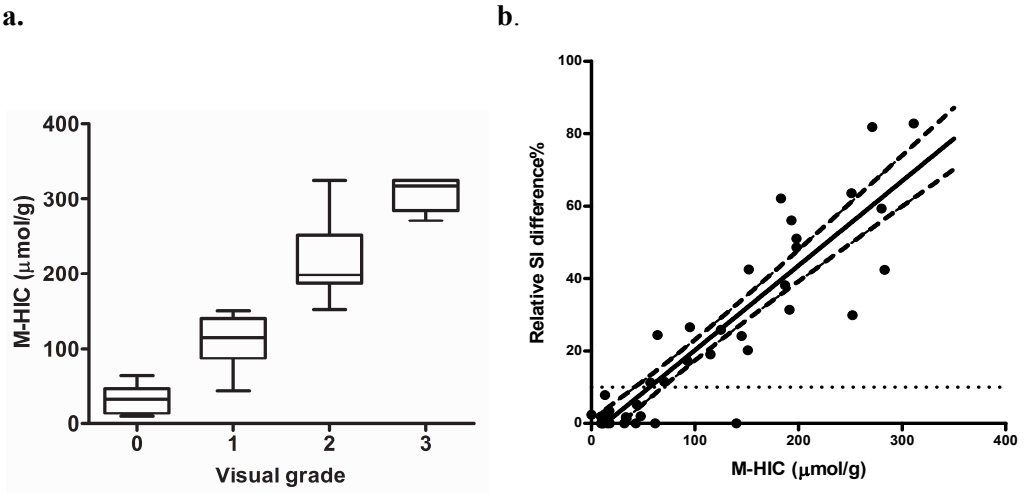


Figure 14. Visual grading and rSI. **a)** Visual grading of liver IO in the four grades. Box plot of the HIC in the four grades: normal iron concentration (grade 0), mild (grade 1), moderate (grade 2) and severe IO (grade 3). The degree of IO in each grade is illustrated with medians (horizontal lines), IQR (boxes) and range (whiskers). There was no overlapping between the grades 1 and 2 at the 151 μmol/g threshold. **b)** Linear dependence between rSI and M-HIC (r^2=0.832, p< 0.0001). Dotted line is set at 10%, which was the detection threshold of rSI method for IO (>36 μmol/g).

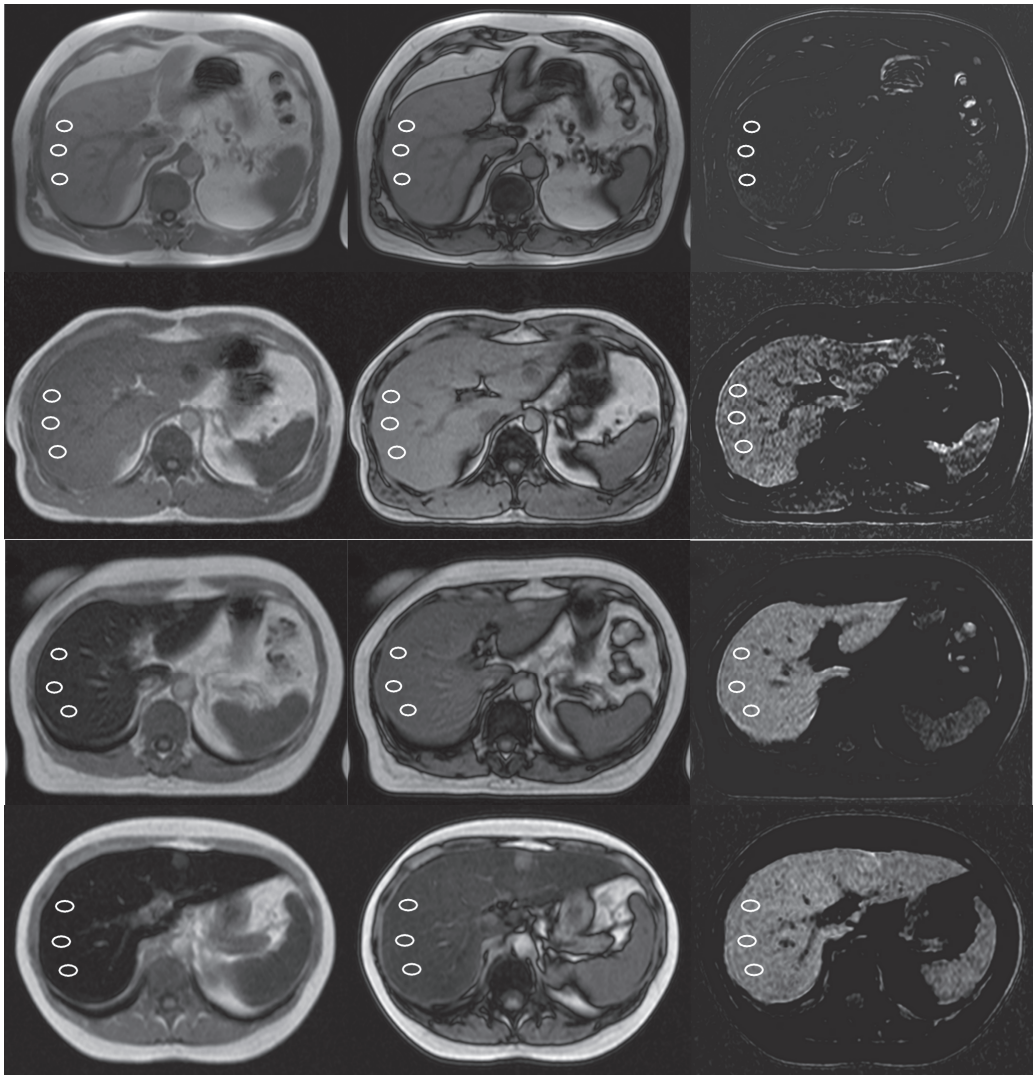


Figure 15. MR images of visual grading of liver IO in the four grades: In-phase (left), out-of-phase MR images (middle) and the subtracted images (right) of typical patients in each of the four grades presented in each row: First row (grade 0), second (grade 1), third (grade 2), last row (grade 3). The images demonstrate the basic principle of the gradual signal drop with increasing IO pronounced in the in-phase images. The ROIs of the liver rSI measurements are also demonstrated. Reprinted from study II.

6.3. RBC and ferritin as iron indicators (III, [a])

We evaluated the RBC transfusion load and plasma ferritin as iron indicators in hematologic patients with transfusional IO (III). The correlation of M-HIC with the number of RBC units was higher ($r=0.840$, $p<0.0001$) than with ferritin ($r=0.676$, $p<0.0001$). In the additional analysis, the increase in the number of subjects did not increase the dependency between ferritin and HIC ($r=0.614$, $p<0.0001$). The

correlations of RBC transfusion load and ferritin in patients with hematological disorders and correlation of ferritin in all subjects are demonstrated in Figure 16.

In patients with hematological disorders, RBC transfusion load was found to be more accurate than ferritin in predicting liver iron overload by the ROC analysis. The respective optimal cut-off values for RBC load and ferritin to predict iron overload were 13 RBC units (AUC=0.940, 100% specificity, 87.5% sensitivity, 100% PPV, 62.5% NPV) and 1019 µg/l (AUC=0.843, 80% specificity, 81% sensitivity, 93% PPV, 55% NPV).

In addition, we evaluated the effect of different transfusion indices (RBC transfusion load, HDS and HDU) and MDS itself on transfusional hepatic iron accumulation in patients with hematological disorders. The total number of RBC transfused ($p<0.001$) and MDS ($p=0.002$) were the only explaining variables left in the final model, HDS and HDU had no effect on IO. This model suggested that MDS is an independent variable, that increases the transfusional liver iron overload with an average of 79 µmol/g compared to non-MDS patients.

The other laboratory parameters (ALT and CRP) did not correlate with HIC ($p=0.192$ and $p=0.700$). In the group of patients with hematologic and chronic liver disease, the respective mean ALT (65 U/l and 66 U/l) was at the same level ($p=0.959$) and the respective means CRP were 7.1 mg/l and 61.5 mg/l ($p=0.004$).

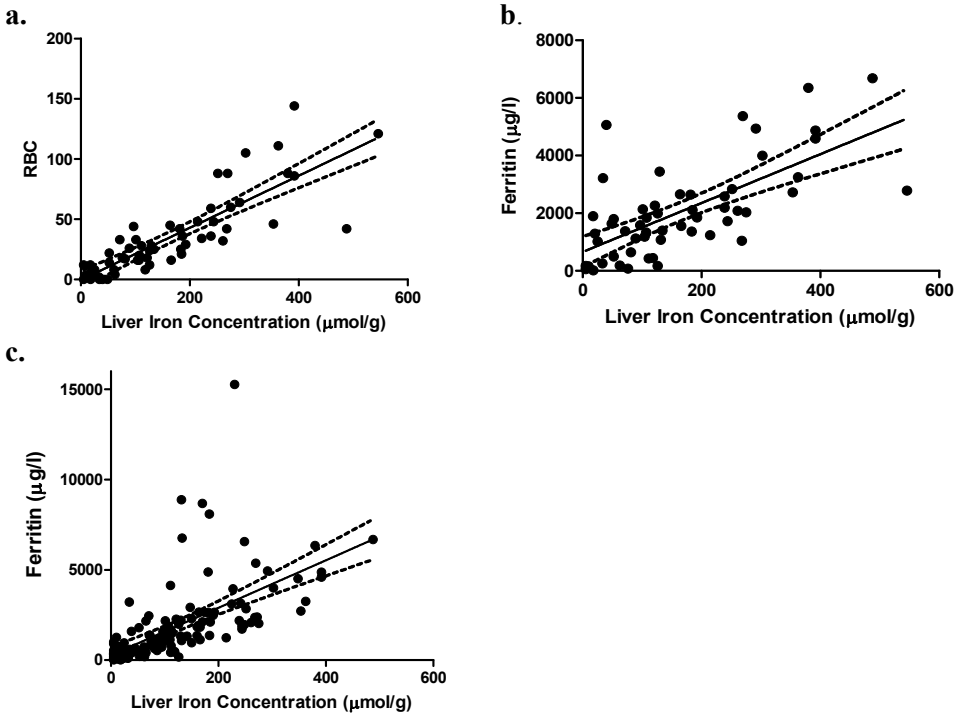


Figure 16. RBC transfusion load and ferritin as iron indicators. In the patients with hematological disorders, the dependency of a) RBC units was stronger ($r^2=0.706$, $p<0.0001$) than the dependency of b) ferritin ($r^2=0.457$, $p<0.0001$). Reprinted from study III. c) In the additional analysis in all study patients, the dependency of ferritin on IO is reduced, but remains highly significant ($r^2=0.377$, $p<0.0001$).

6.4. Cardiac and hepatic IO (III, [a])

6.4.1. The degree of cardiac and hepatic IO

In all study subjects, the prevalence of hepatic IO was higher in patients with hematological disorders (83%, 90/108) than in patients with chronic liver disease (45%, 15/33). The mean HIC was 139±102 µmol/g in the patients with hematological disorders, 47±51 µmol/g in the patients with chronic liver disease, and at the normal level (13±10µmol/g) in the healthy subjects, with significant differences ($p<0.001$) (Figure 17). The range of HIC and cardiac R2* in all study subjects varied between 5 to 488 µmol/g and 14 to 85 s^{-1}, respectively. The highest means HIC (III) among the subgroups of patients with hematological disorders (ALL, AML, SAA, lymphoma, MDS, myelofibrosis) were found in MDS (254 µmol/g), SAA (232 µmol/g) and AML (177 µmol/g) ($p=0.060$).

The prevalence of cardiac IO was low in all study subjects. Only three patients with hematological disorders had cardiac IO (3.2%, 3/93). None of the patients with chronic liver disease or the patients undergoing alloSCT or healthy controls had cardiac IO. The degree of cardiac R2* (III) did not differ between the patients with hematological disorders, patients with chronic liver disease and healthy subjects (36.6±10.2, 36.0±5.4, 34.2±4.9 s^{-1}, respectively, $p=0.714$).

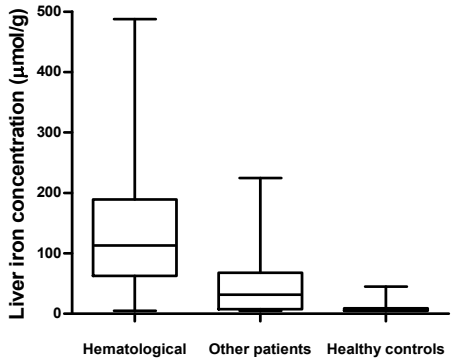


Figure 17. The degree of hepatic IO of all study subjects in the additional analysis. Liver iron concentration was higher in the patients with hematological disorders than in the other patients with chronic liver disease and healthy controls. Medians (horizontal lines), IQR (boxes) and range (whiskers) are shown.

6.4.2. The correlation between cardiac and hepatic IO

In the patients and healthy subjects (III), there was a correlation between cardiac R2* and M-HIC(R2) ($r=0.322$, $p=0.003$). Patients with MDS ($r=0.905$, $p=0.005$) and with myelofibrosis ($r=0.863$, $p=0.027$) showed a correlation between hepatic and cardiac iron overload. In the other diagnostic subgroups of AA, ALL, AML, CLL, CML,

lymphomas, and chronic liver disease, cardiac iron content did not correlate with liver IO (p>0.43).

When adding all patients to the analysis, no significant correlation was found between cardiac R2* and M-HIC(R2) (r=0.235, p=0.011), but in MDS, the correlation remained (r=0.893, p=0.003) (Figure 18).

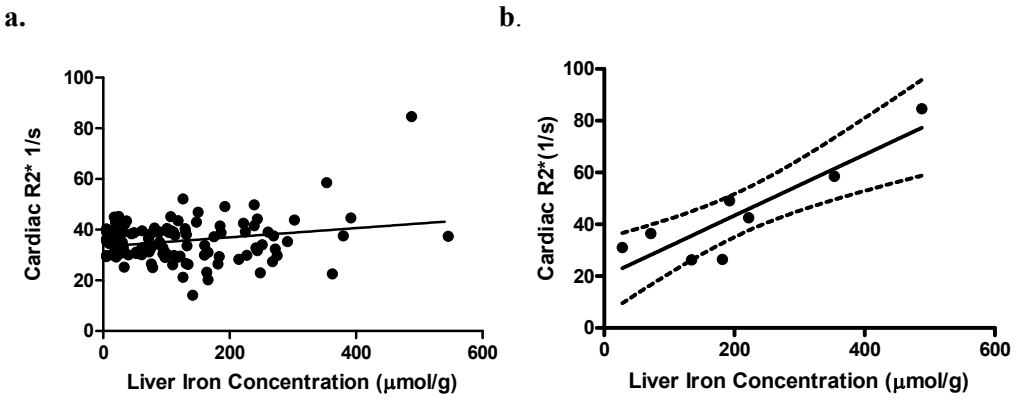


Figure 18. The correlation between cardiac and liver iron content in all study subjects in the additional analysis. **a)** The cardiac IO is rare and not significantly dependent on liver IO in all subjects. **b)** In MDS, the cardiac IO was overrepresented and correlated with liver IO. The two patients with the most clearly elevated cardiac R2* (R2*=85 and 59 s^{-1}) were MDS patients, who suffered from simultaneous severe liver IO (HIC=488 and 353 µmol/g, respectively).

6.5. The prognostic impact of pretransplant IO in alloSCT (IV)

In patients undergoing alloSCT, liver IO was present in 78 per cent and the upper quarter limit of the M-HIC(R2) was 150 µmol/g, which was set as the cut-off value for excessive and non-excessive IO. This cut-off was used in the categorical analysis, performed in addition to the principle multivariate analysis.

Categorically, severe infections occurred more frequently in the patients with excessive IO (35%) compared to patients with non-excessive IO (8%) in the early post-transplant period (p=0.013). In contrast, acute (12%) and chronic GVHD (31%) occurred less frequently in these iron-overloaded patients compared to acute (52%) and chronic GVHD (61%) in the patients with non-excessive IO (p=0.004 and 0.052, respectively). No difference in late period or in all severe infections, OS or TRM was found between the group of patients with or without excessive IO (p=0.201-1.00). Typical alloSCT recipients with excessive and non-excessive IO with corresponding T2 decay for M-HIC(R2) assessment are demonstrated in Figure 19.

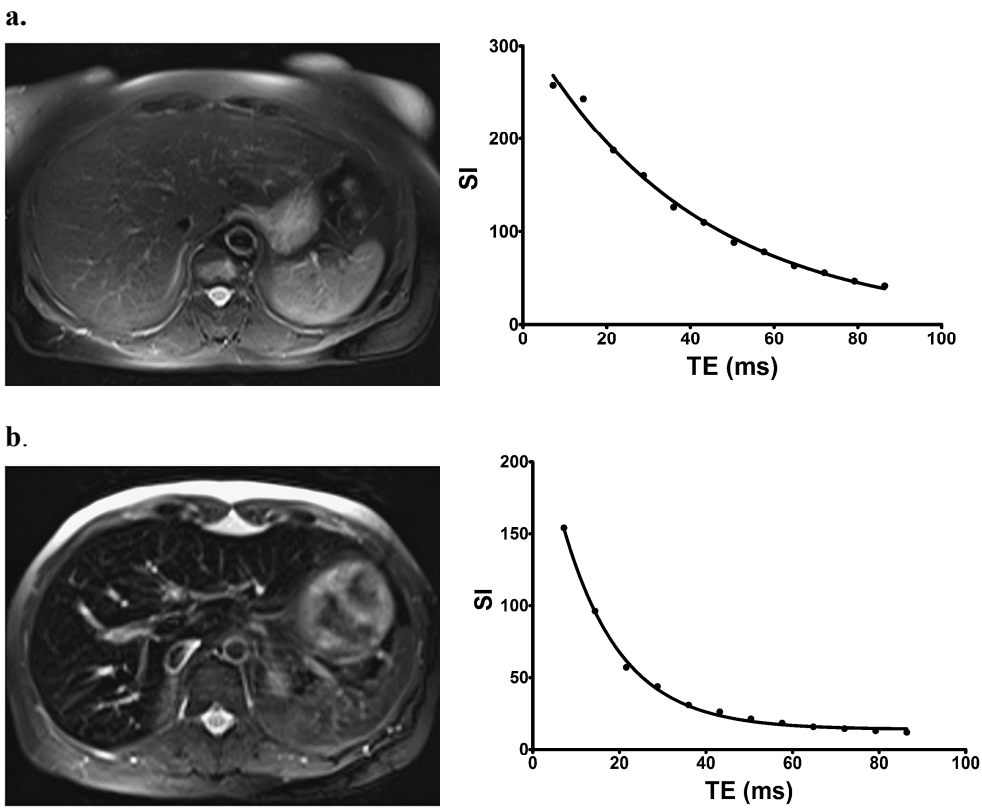


Figure 19. Distinct differences of two representative study patients illustrated with axial liver MR images (TE=64.8ms) and corresponding T2 decay for R2 quantification. **a)** Patient with mild IO (M-HIC(R2)=45 µmol/g). Liver SI of the MR image (left) was higher and T2 decay (right) correspondingly slower. The patient suffered from post-transplant acute grade 3 GVHD of skin and extensive chronic GVHD, together with grade 2 non-severe infection. **b)** Patient with severe IO (M-HIC(R2)=348 µmol/g). Liver SI of the MR image (left) was considerably lower and T2 decay (right) faster due to the iron load. The patient suffered from severe grade 5 post-transplant infection (sepsis), followed by death due to the infection at the day 12 after alloSCT.

6.5.1. Severe post-transplant infections

In the early post-transplant period, pre-transplant M-HIC(R2) was the only variable, which was significantly associated with severe infection (for every 10 µmol/g increase OR: 1.15, 95% CI 1.05-1.26, p=0.003). The M-HIC(R2) was higher in the patients with severe post-transplant infections than in other patients (Figure 20). The odds for severe infections increased 1.15-fold with every 10 µmol/g increase in M-HIC(R2). At the threshold of M-HIC(R2)>125 µmol/g (OR: 6.54, 95% CI 1.49-28.57, p=0.013) and M-HIC(R2)>200 µmol/g (OR: 13.33, 95% CI 2.67-66.67, p=0.002) the odds increased distinctly and reached 14-fold at the threshold of M-HIC(R2)>269 µmol/g (OR: 14.08, 95% CI 1.14-166.67, p=0.040). In the ROC analysis, the most accurate cut-off value of M-HIC(R2) for predicting severe infections was 141.6 µmol/g, which yielded a sensitivity of 70 per cent and a specificity of 79 per cent (Figure 21). Age, gender,

diagnostic subgroup, disease severity, disease duration, conditioning regimen, donor type, ABORh blood group incompatibility and cardiac R2* were not significantly associated with severe infections in the early post-transplant period (p=0.111-0.772).

In the whole one-year follow-up, both M-HIC(R2) and cardiac R2* tended to show an association with severe infections (p=0.084 and p=0.099, respectively), but these variables did not show any association during the late post-transplant period (p=0.878 and p=0.149, respectively) by the initial univariate analysis. In the subsequent multivariate analysis, no statistically significant associations between severe infections and M-HIC(R2) (for every 10 μmol/g increase OR: 1.06, 95% CI 0.99-1.14, p=0.067) or cardiac R2* (for every 10 μmol/g increase OR:0.93; 95% CI 0.85-1.02, p=0.146) were found during the one-year follow-up. Age, gender, diagnostic subgroup, disease severity, disease duration, condition regimen, donor type and ABORh blood group incompatibility were not significantly associated with the occurrence of infections either in the late post-transplant period or in the one-year follow-up (p=0.069-0.929 and p=0.105-0.966, respectively).

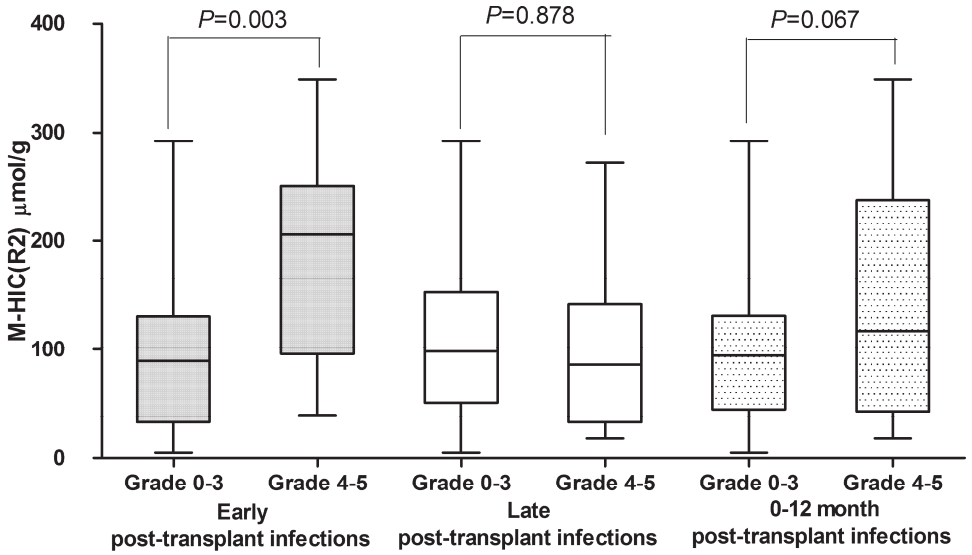


Figure 20. Associations between M-HIC(R2) and post-transplant infections presented with medians (horizontal lines), IQR (boxes) and ranges (whiskers). Significantly higher HIC values occur in patients with severe infections in the early post-transplant period. Reprinted from study IV.

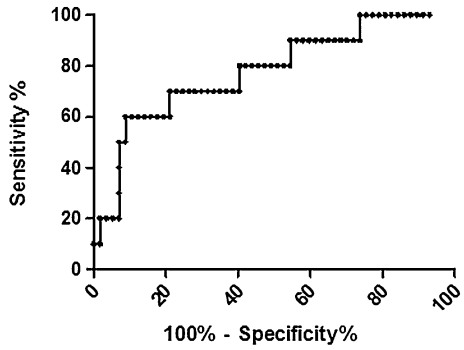


Figure 21. ROC-curve on the accuracy of IO to predict severe infections in the early post-transplant period. The least distance from upper left corner provides the best accuracy (70% sensitivity, 79% specificity, AUC=0.779). Lowering the cut-off to 101 μmol/g increases sensitivity to 80%, and decreases specificity to 60%.

6.5.2. GVHD

M-HIC(R2) was the only variable showing an association with acute GVHD (for every 10 μmol/g increase OR: 0.92, 95% CI 0.85-0.99, p=0.025). This inverse association suggests that every 10 μmol/g increase in M-HIC(R2) reduces the odds for acute GVHD 0.92-fold. M-HIC(R2) was lower in patients with grade 2-4 GVHD than in other patients (Figure 22). Age, gender, diagnostic subgroup, disease severity, disease duration, condition regimen, donor type, blood type difference and cardiac R2* were not associated with acute GVHD (p=0.215-0.950).

In chronic GVHD, the preliminary univariate analysis suggested that M-HIC(R2) was inversely, and sibling donor and male gender were positively associated with chronic GVHD (p=0.011, p=0.034 and p=0.079, respectively). The further multivariate analysis concluded that only M-HIC(R2) showed an independent association (for every 10 μmol/g increase COR: 0.91, 95% CI 0.85-0.98, p=0.015). M-HIC(R2) was accordingly lower in patients with GVHD than in patients without GVHD (Figure 22). Donor type and gender were not associated with chronic GVHD after multivariate analysis (p=0.122 and p=0.101, respectively). Age, diagnostic subgroup, disease severity, disease duration, condition regimen, blood type difference and cardiac R2* were not associated with chronic GVHD (p=0.130-0.841).

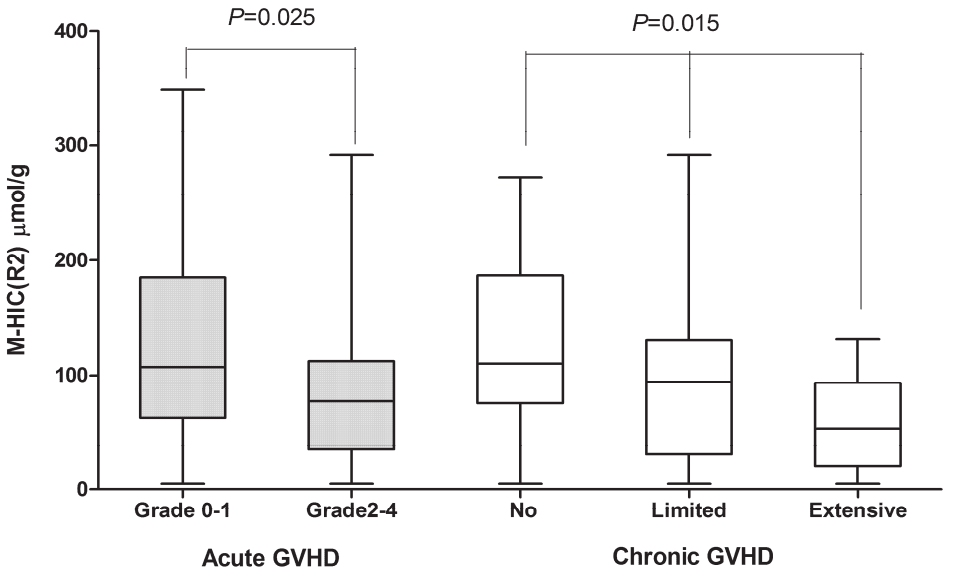


Figure 22. Associations between M-HIC(R2) and acute/chronic GVHD presented with medians (horizontal lines), IQR (boxes) and ranges (whiskers). The lowest HIC values occur in patients with grade 2-4 acute GVHD and extensive chronic GVHD. Reprinted from study IV.

6.5.3. Mortality

There was no association between any of the variables and mortality. During the follow-up, TRM was low (13%) and these patients also suffered from severe infections. The patients who died due to transplantation complications had higher M-HIC(R2) than the other patients (median 133, IQR 52-185 and median 96, IQR 45-148, respectively), but M-HIC(R2) was not associated with TRM, (p=0.166) nor with OS (p=0.436). Age, gender, diagnostic subgroup, disease severity, disease duration, condition regimen, donor type, blood group incompatibility and cardiac R2* were not significantly associated with TRM (p=0.110-0.933) or OS (p=0.153-0.997).

7. DISCUSSION

In this study, MR methods for HIC quantification were validated, optimised and used to investigate IO in the study patients. The quantitative MR imaging methods M-HIC, M-HIC(R2*) and M-HIC(R2)) were found to be reliable iron measures having high correlation with HIC, and performed better than rSI and semiquantitative visual grading. However, all MR imaging-based methods in this study were more reliable in iron concentration measurement than the non-direct iron indicators. In the clinical studies in hematological and allotransplanted patients, a wider quantitative range was found to be essential to cover especially the patients suffering from excessive IO.

7.1. Quantitative liver iron measurement methods by MRI (I,III)

The gradient echo-based methods M-HIC and R2* and the spin-echo-based R2 were strongly dependent on HIC and thus were found to be accurate tools for liver iron concentration measurement.

7.1.1. Liver-to-muscle SI and R2* methods (I)

Our *in vitro* results showed that both the liver-to-muscle SI method M-HIC and R2*, derived from gradient echo imaging sequences and first introduced by Gandon et al., were strongly and linearly dependent with the gel-based liver phantom iron concentrations. We were the first to apply R2* in the analysis of these gradient echo sequences and thus were able to measure *in vivo* iron concentrations derived from R2* rather than liver-to-muscle SI. The linear dependency of R2* on iron concentrations was in agreement with the earlier studies performed with a different sequence type (Storey et al., 2007; Wood et al., 2005). The *in vitro* correlations of both methods at the three scanners remained strong with gel-phased iron-loaded phantoms and are novel results, while the *in vivo* validation of these sequences is in agreement with earlier results (Alústiza et al., 2004). The standard scanner in our centre used to produce the *in vivo* reference M-HIC values was similar to the scanner that was previously calibrated to 174 liver biopsy-derived HIC (Gandon et al., 2004). This similarity in the scanner type facilitated the transfer of HIC quantification in our centre. Although the HIC quantification by MRI is affected by several variables, such as the sequence type, parameters, instrumentation, differences in coil elements, field strength, data reconstruction and analysis, there are less concerns with the differences if the scanners are of similar type and vendor. Although there were only seven biopsies with biochemical HIC analysis in our study, these biopsies confirmed the reliability of the method in the lower quantitative HIC range. Thus, our results along with the previous validations indicate that the methods used in this study are robust and reliable for iron concentration quantification.

In the inter-scanner validation, R2* was found to have a better reproducibility among the scanners than M-HIC. In the study subjects, liver-to-muscle method showed a distinct systematic mean difference of 32 μmol/g (60%) between the scanners, whereas

the mean difference in R2* was relatively small (2.2%). Thus, we corrected the scanner dependency with the inter-scanner equation. To our knowledge, this is the first study with *in vitro* and *in vivo* validations of the inter-scanner differences of the liver-to-muscle SI method. Only one small study has reported favourable intra- and inter-reader agreement of the liver-to-muscle method (Olthof et al., 2009). An obvious explanation for the better transferability of R2* over M-HIC is using another tissue as a reference. Thus, scanner-specific spatial variation is more likely to have an influence in the liver-to-muscle SI with larger spatial difference rather than in liver SI decay-based R2*, where the SI decay is obtained from more similarly located liver ROI. The results on the favourable transferability of R2* is in agreement with the previous studies, where R2* or T2* (1/R2*) has been found to be a transferable method with good inter-scanner (Westwood et al., 2003) and inter-centre (Westwood et al., 2005; Tanner et al., 2006) reproducibility.

7.1.1.1. The influence of spatial variations of B(0)

B(0) field variations are an important source of systematic R2* alterations (Hernando et al., 2012). These intravoxel field differences cause additional enhanced transverse relaxation, SI drop and an increase in R2* values in gradient echo imaging (Hernando et al., 2014), where no 180° refocusing pulse is applied. The field variations make R2* distinctly sequence-dependent, and B(0) differences are typically also dependent on the scanner type, instrumentation, coil and patient. The scanner-dependent B(0) variations most likely contributed to the inter-scanner bias of the M-HIC method. Also the intra-individual values can be influenced, if B(0) differences vary in the region of interest of a target tissue. Intravoxel field differences increase with the voxel size. Because the in-plane resolution of an MR image is typically higher than in the through-slice direction (slice thickness), the field difference in through-slice direction is a principal contributor to the intravoxel inhomogeneity (Hernando 2012). B(0) variations could be tested, for example, by evaluating the phantom SI from different spatial locations in-plane and through-slice directions. However, simultaneous imaging of phantoms and human liver is not preferred, as the air/phantom boundaries would create more inhomogeneity to the field of view. Recently, distinct B(0) field maps have been produced for susceptibility measurement, which correlated directly with R2* (Hernando et al., 2013).

There are two important aspects regarding the B(0) variations in clinical IO measurement. Firstly, the high resolution makes the R2* or SI measurement less sensitive to the field variations and systematic reproducibility errors. Reducing slice thickness should then be favourable, although it would also decrease SNR, which is to be considered, especially in abdominal imaging. Secondly, these variations are most prominent in near the interfaces between tissue and air. Thus subphrenic area is not favourable for ROI position in SI measurements from gradient echo imaging (Hernando et al., 2012; Hernando et al., 2014). If B(0) mapping is not available, the ROI position should be visually chosen as most optimal and homogenous avoiding the areas of high susceptibility effects.

7.1.2. Liver R2 method in comparison to other methods (III Supplementary Information)

With the use of M-HIC(R2) a wider quantitative range was obtained than with the use of the gradient echo methods. However, gradient echo imaging can be faster and sensitive especially to lower iron concentrations (Gandon et al., 2004; St. Pierre et al., 2005; Wood et al., 2005). The empirical, non-linear nature of the calibration curve between R2 and HIC, first found in single echo imaging studies, (St. Pierre et al. 2005) was successfully reproduced in our study. The new calibration was required because the earlier calibrations could not be directly generalised to other type of R2 sequence. We used multi-echo imaging for R2 assessment and thus our R2 values of the calibration curve were lower compared to the studies that used single-echo imaging (St. Pierre et al., 2005; Wood et al., 2005). The R2 level and nonlinearity of the curve agrees with the previous observations. T2 sequence type (single- or multi-echo) and in particular the echo spacing within a train of echoes have been found to lower R2 values compared to single-echo imaging (Brooks, 2001; Wood et al., 2005). When considering the accuracy, the wide quantitative range, and transferability of the methods, the M-HIC(R2) can be suggested to be the most preferable method especially in the most severe IO. There are few publications concerning the transferability of R2 on modern scanners. This is possibly due to the fact that R2 should be an even more transferable method than R2*, which can be explained by the basic MR characteristics of R2 and R2*. They are both dependent on the rate of transverse relaxation, but only R2* has contributions from the local magnetic inhomogeneities, such as changes in the gradients. R2 measures more strictly transverse relaxation of the tissue. However, R2 inter-scanner reproducibility was tested in patients with IO in one study (St. Pierre et al., 2005) and more recently, a multicentre validation of the same method was conducted, in which no significant bias was found, as was to be expected (St. Pierre et al., 2014).

Thus, it can be concluded that R2* and R2 are both transferable and robust methods for iron concentration quantification, R2 having a wider quantitative range and a non-linear calibration curve for HIC compared to a linear dependence of R2*. Transferability testing between different types of scanners is, however, recommended for SI ratio method, M-HIC.

7.2. In-phase and out-of-phase imaging in IO assessment (II)

We assessed the clinically widely used routine in-phase and out-of-phase sequence for a rapid IO assessment. The use of this sequence for IO evaluation requires no special expertise in the imaging or methodology, but the knowledge of the basic principles of the sequence is needed. Two novel methods for the assessment of IO were introduced: a semi-quantitative visual grading method and a relative SI method, rSI.

We found that in-phase and out-of-phase imaging accurately detects clinically significant IO. Both methods of visual grading and rSI showed comparable accuracy and particularly high specificity in the IO detection. The optimal threshold of the rSI method was at 10 per cent for liver iron overload detection. All patients having an rSI

of over 10 per cent or a visually detectable liver signal drop (grade 1-3) were true positive and had hepatic IO. With the use of the 10 per cent threshold, the limit of IO detection for rSI was at 63 μmol/g. The limit of IO detection for visual grading was at the same level, 63.8 μmol/g (upper range of grade 0). Interestingly, our detection limits corresponded to the published observation by Alústiza (Alústiza and Castiella, 2008). In one retrospective study in patients with chronic liver disease by Lim et al, the in-phase and out-of-phase imaging was also found to be of value for the detection of significant hepatic IO, although they used only a semi-quantitative reference standard (Lim et al. 2010). In clinical view, 71 μmol/g liver dry weight has been found to indicate hereditary hemochromatosis, measured from liver biopsies (Kowdley et al., 1997). In thalassemia, the optimal HIC limits for the need of chelation therapy have been studied and the risk of iron-induced complications has been found to be at 125 μmol/g, whereas the threshold of 269 μmol/g has been found to predict cardiac disease and early death (Olivieri and Brittenham, 1997). Thus, the detection limit of 64 μmol/g found in our study seems to be more than sufficient to reveal clinically significant IO.

Both of our methods rSI and visual grading were able to differentiate the different degrees of IO, even if they did not quite reach the level of R2 or R2*, due to either the semi-quantitative nature of visual grading or the accuracy of rSI as a continuous index. However, especially visual grading shows the potential of being a first-hand guideline for radiologists to estimate IO, in addition to more commonly performed liver fat detection. The visual grading was able to differentiate semi-quantitatively minor grade 1 IO from grade 2-3 IO very accurately at the threshold of 151 μmol/g with 100 per cent PPV and NPV. Thus, according to the criteria, a lower liver than muscle SI predicts reliably HIC greater than 151 μmol/g. The visual grading was accurate and comparable between the readers, suggesting robustness of the method. It seems to be relatively easy to introduce and perform.

The two main limitations of iron assessment by in-phase and out-of-phase imaging are liver fat fraction and the fact that currently the use is limited almost solely to 1.5T. In order to evaluate fat and iron reliably, one has to be familiar with the more pronounced susceptibility effect on in-phase and the chemical shift effect due to fat in out-of-phase images (Merkle and Nelson, 2006). Even with that knowledge, the simultaneous presence of fat and iron could lower the sensitivity of the imaging in iron detection. In our study, the rSI was influenced by liver fat content and caused false negative results in these patients, who suffered from simultaneous fatty liver and IO. Consequently, in the presence of clearly elevated liver fat content, the use of rSI is limited and the analysis should alternatively be made with in-phase criteria of visual grading, which was found to be a more robust tool. In addition, significant liver oedema might be a potential pitfall, as it decreases the signal on T1-weighted images (Noone et al., 2000) and might affect the visual grading of the in-phase images. It should be noted that the use of this sequence for iron assessment currently applies to 1.5T scanners, where the longer echo time is routinely and constantly obtained at in-phase among different vendors. In contrast at 3T, the use of the first echo pair has not been constant (Merkle and Nelson, 2006), which is a limitation for a widespread use of this method. If the longer and more iron sensitive echo were obtained at out-of-phase, the SI drop from iron and fat would occur on the same image and could not be separated. Thus, in-phase

and out-of-phase imaging at 3T should be interpreted with caution and requires further validations before reliable use.

7.3. RBC and ferritin as iron indicators (III,[a])

In transfusional iron overload, both transfusion load, i.e., total cumulative number of RBC transfused, and ferritin can be used in a rough clinical estimation of IO, although the transfusion load correlated more closely with HIC than ferritin (III). The additional analysis in a larger number of all study patients did not alter the results, but confirmed the relative limitations of ferritin as an iron indicator.

However, in patients with hematological disorders, we were able to identify the optimal IO detection (HIC≥36 μmol/g) limits of 13 units of RBC and 1019 μg/l of ferritin, with relatively high accuracy and PPV for both indicators. The detection limit of RBC was still more accurate than that of ferritin in predicting IO, as especially the NPV of ferritin was rather limited (55%). Our study results are in good agreement with the correlation results of one prior study in alloSCT recipients (Rose et al., 2007). Infectious or inflammatory conditions were found to weaken the correlation between ferritin and liver iron in a study with a cohort resembling ours (Olthof et al., 2007). However, these conditions were not considered a major factor in our study patients, since the mean CRP collected simultaneously with the ferritin samples was low. Our results suggest that in clinical practice, RBC transfusion load would be more reliable than plasma ferritin for detecting and estimating IO in patients with hematological disease.

In addition, our results suggested that IO was independently increased not only due to transfusions but also due to MDS, whereas large RBC transfusion volumes, patient ages or CRP did not affect IO. The susceptibility of MDS to iron accumulation is to be explained with the pathology of the disease. Apparently, high erythropoietic activity can lead to down-regulation of hepcidin and paradoxically increased iron absorption (Fenaux and Rose, 2009). Especially MDS patients with isolated erythroid dysplasia and low risk of leukemic evolution are thought to be more likely to develop parenchymal iron overload and organ damage (Cazzola et al., 2008). Thus, some precautions should be taken in the estimation of IO by transfusion load in patients with MDS, as transfusions seem to be responsible for only a part of the iron load in this particular disorder (Fenaux and Rose, 2009). In addition, in the patients with chelation or phlebotomy treatment, the transfusion load might overestimate IO, depending on the specific treatment effect. Excessive bleeding might also decrease the IO values. Thus, in those cases of treated transfusional IO and in patients with non-transfusional component of IO, for example in MDS, an indication for ferritin measurement remains as well.

7.4. The degree and relationship of cardiac and hepatic IO (III,[a])

Both the degree and the prevalence of IO were higher in patients with hematological disorder than in patients with chronic liver disease. In the patients with hematological

disorders, 83 per cent suffered from liver IO and three of them had cardiac IO, while only 45 per cent of the patients with chronic liver disease had liver IO and none of them had cardiac IO. Our high liver IO prevalence and degree in hematologic disorders are in good agreement with the few prior studies, in which HIC was measured from post-mortem samples with a biochemical analysis (Altes et al., 2004) or with MRI (Rose et al., 2007). Also in alloSCT recipients, our hepatic IO was at the corresponding level to the results of two other recent MRI studies in these patients (Wermke et al., 2012; Armand et al., 2012). Cardiac iron measurements with MRI have been studied earlier mainly in thalassemia and in some smaller studies in MDS (Wood, 2008).

To our knowledge, this is the first study to investigate the relationship between cardiac and hepatic IO in adult patients with different hematological disorders, even if cardiac IO has been thought to be the main cause of death and focus of treatment in thalassemia for almost two decades (Olivieri et al., 1994; Olivieri et al., 1995). Our novel findings indicate that in patients with hematological disorders other than MDS, the risk of elevated cardiac iron is low. The faint correlation between hepatic and cardiac IO was diminished when all study patients were taken into the analysis. Ultimately, the weak correlation was due to the strong correlation in the subgroup of MDS patients. Thus, the lacking correlation between cardiac and hepatic iron burden in all patients is in agreement with the cross-sectional studies in thalassemia (Anderson et al., 2001; Wood et al., 2004; Leung et al., 2009). In our study, this result most likely also reflects the low occurrence of cardiac IO.

We were able to identify a specific diagnostic subgroup, patients with MDS, which had the highest liver and cardiac IO. The transfused units of RBC, patient age or the high transfusion volumes did not explain the results in MDS. The occurrence of cardiac iron overload in MDS, in our study, is principally in agreement with other studies, although there has been some inconsistency in the results earlier (Wood, 2008; Fenaux and Rose, 2009). Some of the inconsistencies are to be explained with methodological differences, as Jensen et al. found cardiac iron in 9 out of 12 patients with MDS using a different MR method than ours (Jensen et al., 2003). Other studies used a validated T2* method, corresponding to ours and found 3 out of 22 (Di Tucci et al., 2008), 1 out of 11 (Chacko et al., 2007) and 1 out of 10 (Konen et al., 2007) cardiac iron-overloaded patients. In our study, both of the MDS patients with severe liver iron overload (>269 μmol/g) also had cardiac iron overload unlike the other subgroups. This result for MDS agrees with a previously detected threshold of 269 μmol/g for greatly increased risk for cardiac disease and early death in patients with thalassemia (Olivieri and Brittenham, 1997). Earlier, two small studies have also reported on a critical level of liver iron concentration in non-chelated MDS patients, after which cardiac iron overload might occur (Jensen et al., 2003; Di Tucci et al., 2008).

The correlation we found between hepatic and cardiac iron in MDS was clearly stronger than previously reported in chelated patients (Konen et al., 2007; Chacko et al., 2007). The occurrence of cardiac IO and the high degree of IO in MDS found in our study can be explained with two factors, which are thought to occur in certain subgroups of MDS and are similar to thalassemia, namely high number of transfusions and active yet insufficient erythropoiesis. The MDS patients, who survive longer, have

higher number of RBC transfusions for a longer period of time than an average hematological patient resulting in higher transfusional IO. Secondly, especially in the MDS subgroup of RARS, ineffective, though active dysplastic erythropoiesis results in increased intestinal iron absorption. Dysplastic erythropoiesis increases unnecessary iron requirements via hepcidin down-regulation, even if the patient already suffers from severe IO. When iron loading increases, transferrin saturates resulting in increasing amount of circulating NTBI, which can enter cardiac tissue by non-controlled pathways (Cazzola et al., 2008; Wood, 2008; Cuijpers et al., 2010). This is thought to cause cardiac and parenchymal IO in general. Our results suggested that the correlation between liver and cardiac IO might also occur in myelofibrosis in addition to MDS. A possible explanation for this might be related to the common features of these entities. Even the diagnosis of MDS is not always straightforward, especially in the fibrotic or hypoblastic variants, which may be not distinguished from myelofibrosis on the basis of bone marrow smears (Horny et al., 2007). In earlier studies, chelation treatment with variable efficiency in different iron-overloaded tissues might have biased the proportions of the heart and liver IO (Berdoukas et al., 2009; Anderson et al., 2002) and chelation is known to also affect the relaxation rates (Wood et al., 2008). By excluding the chelated patients, we were able to explore the iron accumulation without the possible bias from iron unloading treatment, which may have contributed to our correlations.

7.5. Prognostic impact of pre-transplant IO in alloSCT (IV)

This present study is the first prospective evaluation to show an association between the pre-transplant IO and the post-transplant severe infections after alloSCT. We have demonstrated that IO associates strongly with severe infections during the early post-transplant period. Multivariate analysis showed that no other variable; i.e., cardiac R2*, age, disease duration, gender, diagnostic subgroup, disease severity, conditioning regimen, donor type and ABORh blood group incompatibility, was associated with the severe infections. The association between pre-transplant IO and severe post-transplant infections was significant whether the HIC analysis was performed either as continuous or categorical variable. The severe, grade 4-5 infections included blood stream infections and deaths from infection of various pathogens. The infections were caused by either bacterial, viral or fungal pathogens, typical for SCT recipients (Treleaven and Barrett, 2009), but they more frequently caused a severe disease pattern in the patients with excessive IO. Our findings confirm the results of the earlier retrospective studies concerning blood stream infections, in which similar results were found, although IO was measured with indirect iron indicators, such as ferritin (Tachibana et al., 2011) or semi-quantative measures (Ali et al., 2012; Miceli et al., 2006). In one prospective study of 190 allo-SCT recipients, pre-transplant hyperferritinemia was associated not only with bloodstream infections but also with acute GVHD and death (Pullarkat et al., 2008). In our study, there was an association between pre-transplant IO and severe infections only in the early post-transplant period. Some immunomodulatory factors, other than iron, might have counterbalanced the effect in the late post-transplant period. For example, the anti-inflammatory

treatment of frequently found GVHD in low IO group might have increased the infection risk also in these patients.

In this study, the risk for infections increased proportionally with increasing IO. The odds for severe infections increased from 6.5-fold to 14-fold within the HIC range of 125 μmol/g to 269 μmol/g. This agrees well with earlier results on the complication risk, reported in thalassemia, as this very HIC range has been associated with increasing risk of iron-induced complications (Olivieri and Brittenham, 1997). A gradual increase in the post-transplant complication risk with increasing IO has also been suggested earlier in studies conducted with iron surrogates. Risk for TRM increased in RBC range of 20 to 40 units, ferritin range of 1000 to 3000 μg/l (Alessandrino et al., 2010) and in the groups of patients, in which ferritin was categorised as greater than 1000 μg/l and greater than 2500 μg/l (Platzbecker et al., 2008).

We found a critical cut-off value, HIC >142 μmol/g, to predict severe infections in the early post-transplant period. This value corresponds to plasma ferritin level of 2000 μg/l and 30 RBC units, based on our study III. This agrees with the other earlier study results on the prognostic impact of corresponding levels of iron surrogates, i.e., ferritin or transfusions (Mahindra et al., 2009a; Mahindra et al., 2009b; Ali et al., 2012,). In addition, HIC greater than 150 μmol/g, our limit for excessive IO, has been suggested to predict invasive aspergillosis in patients who died after HSCT, which also agrees with our results (Altes et al., 2004). Pending our study, two other studies where pre-transplant IO was measured directly with liver MRI were published, but the results on post-transplant mortality were contradictive between these studies and infections were not analysed (Wermke et al., 2012; Armand et al., 2012). The observational study of 78 patients (Wermke et al., 2012) reported that HIC of greater than 125 μmol/g predicts post-transplant mortality. Their pre-transplant IO was measured with the same reliable M-HIC method (Gandon et al., 2004), which has been validated and calibrated with R2 by us. When a lower HIC cut-off value of 90 μmol/g was used in the study of 45 allo-SCT recipients with a lower HIC median, no association between pre-transplant IO and mortality was found (Armand et al., 2012). Based on the results of our study and others, we suggest that the clinically critical HIC range for increased infection risk in these patients is at the level of 125 to 150 μmol/g. Thus, it can be even suggested that IO assessment before alloSCT might be preferred to be done with MRI, especially in the selected group of alloSCT recipients, who are suspected to have significant IO and increased infection risk.

High body iron load seems to be a cofactor in severe transplant-related infections in this study, which can be explained by the different adverse effects of iron. In addition to the earlier introduced oxidative stress, iron-related alterations in immunity; i.e., immunomodulatory effects may play a crucial part in this phenomenon (Majhail et al., 2008). Accordingly, iron deprivation has been found to be an essential element of the host defence and immunity. The actual mechanisms are not thorougly revealed, but hepcidin has been found to take part in the iron concentration maintenance and redistribution between extra and intracellular space of macrophages, and is thus considered to be an important link between immunity and iron (Ganz, 2009; Wang and

Cherayil, 2009; Nairz et al., 2010). Natural resistance to infections is suggested to be damaged by high concentration of free iron. IO is thought to have a role in the excessive formation of free ionic iron, which can thus alter normal bactericidal mechanisms and enhance the growth of bacteria and fungi (Bullen et al., 2006).

GVHD is another major cause of TRM after allo-SCT, in addition to severe infections. In our study, liver IO was the only variable that showed an independent association both with acute and chronic GVHD in the multivariate analysis. According to the inverse association, HIC would decrease the occurrence of GVHD during the whole follow-up period, which is a novel finding. The results of our categorical analysis were quite consistent with the multivariate results, as the patients with excessive IO had significantly less acute GVHD and an apparent tendency towards less chronic GVHD. Our finding supports the hypothesis presented earlier that IO might lead to decreased risk of GVHD through its ability to impair host immune responses (Majhail et al., 2008). Furthermore, recent results agreeing with this hypothesis were found in a liver transplantation study, where allograft acceptance was associated with higher ferritin and hepcidin levels together with increased hepatocyte iron content, indicating that iron plays a potential role in the regulation of alloimmune responses (Bohne et al., 2012). Earlier results concerning the association between GHVD and IO have been contradictory. Our results agree with two studies, in which hyperferritinemia was associated with a lower incidence of GVHD (Mahindra et al., 2009a; Wahlin et al., 2011), whereas two other ferritin studies showed an opposite association (Pullarkat et al., 2008; Platzbecker et al., 2008), and two studies did not find any association, when IO was assessed independently by MRI (Armand et al., 2012) or ferritin (Kataoka et al., 2009). When interpreting these results one must remember the double role of ferritin, on one hand as an iron indicator and on the other hand as an acute phase reactant, which might have contributed to these results.

IO associated proportionally with infections and inversely with GVHD, but no distinct association was found between survival and IO in this study. The patients with excessive IO were at an increased infection risk, but the other patients with lower iron load were at increased risk for GVHD. As both severe infections and GVHD can be lethal, the risk of infection-related deaths was probably compensated with the risk of principally GVHD-related deaths. A number of previous studies, which used an indirect iron indicator, have found an association between ferritin and post-transplant survival (Altes et al., 2002; Majhail et al., 2008; Pullarkat et al., 2008; Kataoka et al., 2009; Mahindra et al., 2009a; Mahindra et al. 2009b; Alessandrino et al., 2010; Sucak et al., 2010). The two studies, where pre-transplant IO was measured with MRI, found contradictory results. A recent study of 88 allo-SCT recipients reported significant associations between TRM and IO measured with an M-HIC method corresponding to ours (Wermke et al. 2012). This was contradicted with another recent prospective study of 44 allo-transplanted patients with a lower HIC median and cut-off (Armand et al., 2012). They found no association between pre-transplant IO, measured with MRI, and mortality. When putting together the results of these studies along with the present study, the IO-related poor transplantation results seem to be mediated by infections, not by GVHD.

7.6. Strengths and limitations

The *in vitro* results agreed with the *in vivo* results on the linear dependence of the methods of M-HIC and R2*, and also concerning the transfer validation. Thus, our *in vitro* phantoms were found to be reliable in testing the clinical method, because they were designed to mimic liver both in relaxation values and consistency. We first performed the methodological validation before applying the methods in patients with transfusional IO and addressing the clinical questions about the role of IO. The phantoms were prepared to follow T2 and T1 characteristics of a normal liver. Furthermore, gel-based phantoms were chosen instead of simple liquid samples to optimise the iron relaxivity (r2*), which is weaker in solutions (Stark, 1991). The relaxivity of the liver was slightly higher than in the phantoms. The phantoms were imaged at 21°C, whereas the liver at the body temperature. Ferritin-induced transverse relaxation is faster at lower temperatures (Gossuin et al., 2007) and the relaxivity of liquid phantoms is also thought to increase with diminishing temperature (Clark et al., 2003). Water diffusion, on the other hand, is known to enhance transverse relaxation with increasing temperature, but many other factors, such as iron particle size and distribution, take part in relaxation modulation of the phantoms and liver tissue (Ghugre and Wood, 2011). However, the *in vitro* and *in vivo* studies were conducted separately. Thus, the phantom relaxation did not contribute or bias the liver iron measurement. In addition, the phantoms were imaged after the temperature was stabilised at all the scanners for reliable transferability testing.

The number of biopsy-proven B-HIC for *in vivo* in our centre was limited, and thus the principal SI-based MR method, M-HIC, was based on a calibration made previously (Gandon et al., 2004). Since our M-HIC was conducted at a similar scanner that was used by Gandon et al., the reference standard iron concentration quantification by M-HIC can be regarded to be free from a scanner-type bias. The phantoms were imaged with a single excitation in the field of view, but in different spatial locations in the xy plane. As a limitation, systematic local differences of B(0) field, which can affect the SI, were not measured. However, the use of single excitation for phantom imaging can also be regarded as strength, as then the variations between the excitations at clinical scanners did not confound the results. The ROIs of the phantoms were taken from spatially different locations in xy plane, but within the area routinely tested for homogeneity by the vendors. Also, the possible effect of scanner-dependent magnetic drift of B(0) over time was not evaluated in this study.

Our results on the MR methodology followed consistently the earlier empirical or theoretical models and suggestions for the behaviour of R2 and R2*. Firstly, our calibration curve of gradient echo-based R2* method was directly dependent on iron concentration, and R2 showed a non-linear correlation to HIC. Both of these linear and non-linear calibrations were consistent with the previous validations and models (Ghugre and Wood, 2011). Secondly, R2 yielded higher iron concentrations than the gradient echo-based methods, according to the theoretical basis of R2 and R2*. The wide quantitative range of R2 was crucial in assessing the HIC also in the patients with the highest IO in clinical studies III and IV. Thirdly, the range and number of TE times; i.e., 11 echo times used between 7.2 to 86.4 ms for R2 method, and 12 echo

times for cardiac T2* between 3 to 29 ms, can be regarded sufficient, which was also seen in both previous validations (Wood and Ghugre, 2008). The liver R2* from M-HIC data was measured with a relatively limited number of TE times (4-21 ms), but these could not be increased without changing the validated and calibrated method. This limitation in the shortest TE was reflected also to the somewhat limited quantitative range of this M-HIC(R2*) method. Finally, the model of data fitting for T2 (1/R2) was chosen as mono-exponential nonlinear fit with a constant C, which is considered the most robust way of correcting the undesired SI from noise or low-iron tissue distortion (Wood and Ghugre, 2008). This offset constant C is used to compensate image noise and iron-poor and H_20-rich anatomical areas, i.e., vessels, within the ROI, which seemed to be beneficial to the R2 method. This noise level, which we corrected, is demonstrated in the examples of T2 decay of the two study patients for R2 quantification (Figure 19), where the fitting curve remains above the noise level even at the longest TE. This model of data fitting probably facilitated the use of the empirical calibration curve by St. Pierre et al. between the correlation of HIC and R2. If left uncorrected, this might have led to a systematic underestimation of R2 (Wood and Ghugre, 2008). However in cardiac R2*, the noise floor was not addressed, because a change in the fitting model would have affected R2* values and thus probably also the 50 s^{-1} (20 ms) cut-off value (Anderson et al. 2001) of the upper-range normal.

The variability of R2 values is found to be considerably large especially at the higher part of the HIC range (St. Pierre et al., 2005; Wood et al., 2005), which is also illustrated in our study. The variability has been comparable to the biopsy-based HIC results. In a post-mortem study of thalassemia patient with cirrhosis, the R2 variation within the liver tissue corresponded to HIC variations (Clark et al., 2003). Thus the HIC heterogeneity, most prominent in cirrhotic liver, explains largely the variability of R2. The intra-individual variability of HIC increases with decreasing size of a biopsy sample. This favours larger samples and MR imaging as liver biopsy represents only a diminutive part of the liver volume compared to liver ROI, and by visual assessment the whole liver can be assessed**.** Because of the non-linearity of R2 calibration curve, the R2 change per concentration unit decreases as HIC increases. This, in its part, addresses the R2-based HIC variability. However, there are also other factors, which may partly explain the variability of the R2 results, for example, steatosis, fibrosis, diffusion or inflammatory states. Among those, liver fat fraction can modulate SI and should thus be taken into consideration in IO measurements (Hernando et al., 2014). Fat saturation is recommended if suitable to add in the sequence, as was done in the R2 multi-echo in our study. Also, the use of in-phase TE counterbalances the major fat fraction (Hernando 2012 MRM) and was used in our gradient echo-based methods R2* and M-HIC. However, the effect of fat fraction or cirrhosis on the variability of the measurements was not measured in our study. The analysis of fatty liver by spectroscopy as confounding factor would have been especially beneficial in the in-phase and opposite-phase imaging. Fibrosis or inflammatory conditions were not found to affect HIC measurements by the R2 in a recent multi-centre work in patients with thalassemia (St. Pierre et al., 2014), where fibrosis was found to increase the variability of R2, reflecting the heterogeneous iron distribution. The effect of fibrosis and oedema

on the liver relaxivity seems to be rather diminutive compared to the iron-induced increased relaxivity, and is, for that reason, not easily shown.

In the *in vivo* analysis, the dual studies were not excluded from the correlation plots (I), which might be regarded as statistical limitation. However, these studies were imaged and analysed separately from the first imaging after an intervention of phlebotomy or RBC transfusions, and not regarded as typical repetitive and similar data points.

The number of all patients was satisfactory in the whole study, although some of the different diagnostic subgroups were relatively small. Thus, the conclusions, for example, on MDS are based on a limited number of patients. Thus, further evaluation in a larger number of patients with MDS is required. This study did include prospective and novel data in addition to cross-sectional and validation studies, which are considered more reliable than retrospective studies. However, one must remember that the results of the negative impact of IO on infections in this study are restricted to the most severe, life-threatening infections in the early period. The pathogens causing the infections in study IV were diverse, and thus the results cannot be used to explain the behaviour of an individual pathogen. The patients were also treated with prophylactic antimicrobial agents, which might bias the incidences of the pathogens. Finally, due to the limited number of deaths, the effect of IO on survival could not be established. Accordingly, a larger prospective study or pooled data of multiple studies using direct quantitative measurement of HIC for IO assessment is required to conclude the specific iron-induced effect on mortality.

7.7. Future considerations

The non-invasive IO measurement by MRI is preferred to liver biopsies especially in patients with hematological disorders, as the patients are at the increased risk for thrombocytopenia and bleeding. In the future, the quantitative IO measurement will most likely become a widespread tool in IO treatment monitoring and clinical decision-making.

3T scanners are increasingly used in clinical practice. However, all the validated methods with a wider clinical use are optimised at 1.5T. We are lacking validated methods for IO measurement at 3T, and larger, controlled trials are needed especially in this regard. The higher field strength has its known advantages, such as faster acquisition time, but it is also more sensitive to susceptibility and motion artefacts (Storey et al., 2007; Peng et al., 2013). At 3T, there remains a challenge in iron-induced increased signal decay as R2* increases near linearly with field strength causing faster SI drop and lower SNR in highest HIC range (Storey et al., 2007; Hernando et al., 2012). This is to be addressed in future research in order to maintain the quantitative range at 3T. New promising methods, such as B(0) field independent magnetic susceptibility measurements (Hernando et al., 2013) are evolving as potential IO measure at different field strengths, although they still need optimisation before wider clinical use for tissue iron measurement.

Overall, a wide range of non-invasive tools for IO measurement is required in clinical monitoring of the patient. The quantitative MR imaging methods are to be used principally in specialised centres, while more straightforward visual grading and non-direct clinical methods are more easily introduced.

8. CONCLUSIONS

The conclusions of this study are:

1. Hepatic iron concentration can be measured with three quantitative MR methods: (1) liver-to-muscle SI (M-HIC), (2) transverse relaxation rate R2* or (3) R2. M-HIC and R2* are both accurate and directly proportional to iron concentration, and R2* is more reproducible than M-HIC between scanners. R2 can measure higher iron concentrations than M-HIC or R2*.

2. We introduced two novel methods, visual grading and rSI, which can be used to detect and grade clinically relevant IO from routine abdominal MRI without special expertise. The visual grading is a robust and semi-quantitative iron indicator with excellent inter-reader agreement, while rSI is a continuous measurement but limited by liver fat.

3. Both RBC transfusion load and ferritin can estimate IO, but less reliably than quantitative MRI. In patients with hematological diseases, transfusion load is more accurate than ferritin. Adding all patients into the analysis did not increase the accuracy of ferritin.

4. Both the degree and prevalence of IO are higher in hematological diseases than in chronic liver disease. In hematological patients, hepatic IO is common but cardiac IO is rare and does not correlate with hepatic IO. MDS patients seem to be the most susceptible to cardiac IO, which also correlates with hepatic IO.

5. In alloSCT recipients, pre-transplant IO measured with MRI predicts severe infections during the early post-transplant period. In contrast, IO reduces the risk for GVHD. The impact on mortality needs further assessment. Iron-induced poor transplantation results seem to be mediated by severe infections.

9. ACKNOWLEDGEMENTS

This study was carried out at the Imaging Centre of Southwest Finland, Turku University Hospital and the Department of Radiology, University of Turku during the years 2005-2014. It was financially supported by grants from Turku University Hospital and a grant from TYKS-säätiö.

I wish to express my gratitude to Professor Hannu Aronen, University of Turku, for his encouragement, interest in research and inspiring ambition.

I owe my deepest gratitude to the supervisor of this study, Professor Riitta Parkkola for her continuous support and encouragement, excellent guidance with positive and energetic attitude throughout these years. Her optimism and pragmatism were essential in completing this thesis.

I wish to thank the official reviewers, Docent Nina Lundbom and Docent Kaarina Partanen, for their skilful and thorough work. Their comments resulted in distinct improvements in this thesis.

I was privileged to work with outstanding physicists, Markku Komu and Jani Saunavaara. I want to express many thanks to both of them. We shared the interest in MR methodology. Markku acquainted me closely with methodology, MR physics and experimental methods. Jani further carried out the methodological work characterised by skilfulness and resourcefulness. I appreciate all the conversations, support and help that I received for the challenges throughout this study. This dialogue was delightful and precious for carrying out this work.

I was fortunate to collaborate with skilful hematologists and stem cell transplantation specialists, Docent Maija Itälä-Remes and Professor Kari Remes. I wish to thank both of them for their time and interest in iron-related effects and wide clinical perspective. For meaningful research, the knowledge about the clinical views is essential. A decade ago, Maija first brought up the idea of whether radiologists cold start obtaining iron measurements with MRI at our centre. This prompted our research. I appreciated the initiative and took the challenge. After the studies in this thesis were completed, our research continued with overseas collaboration we were asked to join to conduct a further meta-analysis.

I wish to thank the management at the Imaging Centre of Southwest Finland, at the Department of Radiology. I am grateful for Professor Roberto Blanco, the Chairman at the Imaging Centre of Southwest Finland, Docent Anu Alanen and Juha Sjöval M.D. for the encouragement and opportunity to do research between the clinical work periods and for continuous maintenance of competent equipment and facilities for clinical work and research. Special thanks to Sakari Salo M.D. for the first quantitative study in our centre in the T2 multiecho imaging and the pragmatic support during these years.

I felt privileged to have skilled co-workers. I wish to thank my dear friend from the choir Sirkku Viitanen-Vanamo and Robert Badeau for the linguistic revisions of this manuscript. Sirkku's skilful comments were particularly important in the efficient completing of the final revisions. I am grateful for the radiographers at our centre for the excellent and precise MR imaging. I also wish to thank the staff at the Department of Biostatics for the excellent help and at the Galilaeus Oy for the laboratory facilities and professional *in vitro* phantom preparation, and many thanks to the Department of Chemistry at Åbo Academy for outstanding collaboration in the precise measurement of small liver biopsy samples

I wish to thank Professor Juhani Knuuti, director of the Turku PET Centre, for the encouragement and professional views during the time, in which I have been privileged to attend to the outstanding work in clinical PET imaging. I also wish to thank Professor Pirjo Nuutila and Professor Mika Teräs for kind support and interest in my thesis. During my stay at the Turku PET Centre I have appreciated the excellent company of my co-workers and colleagues in a great working environment. I wish to thank Marko Seppänen, Jukka Kemppainen, Terhi Tuokkola, Virva Saunavaara, Tuula Janatuinen, Sami Kajander, Marita Kailajärvi, Maija Mäki, Kirsi Virtanen, Maria Wendelin-Saarenhovi, and Jussi Pärkkä for practical views, positive and friendly attitude and encouragement towards the completion of my theses.

I am grateful for all my colleagues and friends at A-röntgen over the past years for their support, help and delightful conversations. I wish to thank Teemu Paavilainen, Ia Kohonen, Pirkko Sonninen, Jussi Kankare, Jari Karhu, Juha Ratilainen and Tomi Pudas for listening to my stories and creating a pleasant and friendly atmosphere and for active approach to challenges. Juha and Tomi were excellent readers in the visual grading method, which was an important part of this study. I am deeply grateful for Kirsti Dean, who unfortunately passed away, but from whom I learnt a lot and had the great privilege to work and study with. I was also privileged to get to know her husband, Professor Peter Dean, who gives me a great example in both having a great interest in science and warm attitude towards the family and the fellow-men. Special thanks to Kimmo Mattila, for the outstanding knowledge and energetic interest in the research ideas. I wish to thank the skilled co-workers at the abdominal imaging, Timo Tuunanen and Miia Virtanen for many important discussions and for sharing the clinical workload on daily basis.

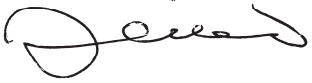
I am deeply grateful for the continuous support and confidence from my friends. Special thanks to my dear friends and colleagues Minna Saarinen, Susanna Hirsimäki, Jutta Paananen and Niina Hekkala starting from the days at the University of Oulu, and to Maarit Hakonen, whom I first met at a choir practise. I really appreciate the dear singing friends at the choir Sekakuoro Kulkuset for giving me the inspiration and enthusiasm also through the most difficult times. Many choir members have encouraged me in these past years. I am grateful for the support from Juha Laakso, our bass chorister and my delightful soprano friends Mari Schildt and Sirpa Ääri-Vähätalo in addition to songwriter Kimmo Stjärnstedt among many others.

I owe my sincere thanks to my dear parents Sirkka and Jussi Kallio for having a great confidence in me from the start and for the understanding upbringing in a steady,

secure and comfortable atmosphere, where equality, education and determination were appreciated but where imaginative play, arts, poetry and music of different styles were easily reached. I am also grateful for the kind help and support from my in-laws Maila and Heikki Virtanen for taking care of the children, as well I am grateful for the support from my brother-in-law Timo Virtanen and his family.

I am most deeply grateful to my whole loving family for the support and for the patience during the writing process. I sincerely want to thank my dear husband Tero Virtanen for his pragmatic and objective views, help and support and also for the patience, persistence and encouragement. The presence of our two children, Sanni and Eero, bring the most precious moments in my life.

Turku, September 2014



Johanna Virtanen

10. REFERENCES

Adams, P.C. 2006. Review article: the modern diagnosis and management of haemochromatosis. Aliment Pharmacol Ther 23 (12):1681-91.

Alessandrino, E.P., Della Porta, M.G., Bacigalupo, A., Malcovati, L., Angelucci, E.M. Van Lint, T., Falda, M., Onida, F., Bernardi, M., Guidi, S., Lucarelli, B., Rambaldi, A., Cerretti, R., Marenco, P., Pioltelli, P., Pascutto, C., Oneto, R., Pirolini, L., Fanin, R., Bosi, A. 2010. Prognostic impact of pre-transplantation transfusion history and secondary iron overload in patients with myelodysplastic syndrome undergoing allogeneic stem cell transplantation: a GITMO study. Haematologica 95 (3):476-84.

Alexopoulou, E., Stripeli, F., Baras, P., Seimenis, I., Kattamis, A., Ladis, V., Efstathopoulos, E., Brountzos, E.N., Kelekis, A.D., Kelekis, N.L. 2006. R2 relaxometry with MRI for the quantification of tissue iron overload in beta-thalassemic patients. J Magn Reson Imaging 23 (2):163-70.

Ali, S., Pimentel, J.D., Munoz, J., Shah, V., McKinnon, R., Divine, G., Janakiraman, N. 2012. Iron overload in allogeneic hematopoietic stem cell transplant recipients. Arch Pathol Lab Med 136 (5):532-8.

Altes, A., Remacha, A.F., Sarda, P., Sancho, F.J., Sureda, A., Martino, R., Briones, J., Brunet, S., Canals, C., Sierra, J. 2004. Frequent severe liver iron overload after stem cell transplantation and its possible association with invasive aspergillosis. Bone Marrow Transplant 34 (6):505-9.

Altès, A., Remacha, A.F., Sureda, A., Martino, R., Briones, J., Canals, C., Brunet, S., Sierra, J., Gimferrer, E. 2002. Iron overload might increase transplant-related mortality in haematopoietic stem cell transplantation. Bone Marrow Transplant 29 (12):987-9.

Alústiza, J.M., Artetxe, J., Castiella, A., Agirre, C., Emparanza, J.I., Otazua, P., García-Bengoechea, M., Barrio, J., Mújica, F., Recondo, J.A. 2004. MR quantification of hepatic iron concentration. Radiology 230 (2):479-84.

Alústiza, J.M., and Castiella, A. 2008. Liver fat and iron at in-phase and opposed-phase MR imaging. Radiology 246 (2):641.

Alústiza, J.M., Castiella, A., De Juan, M.D., Emparanza, J.I., Artetxe, J., Uranga, M. 2007. Iron overload in the liver diagnostic and quantification. Eur J Radiol 61 (3):499-506.

Anderson, L.J. 2011. Assessment of iron overload with T2* magnetic resonance imaging. Prog Cardiovasc Dis 54 (3):287-94.

Anderson, L.J., Holden, S., Davis, B., Prescott, E., Charrier, C.C., Bunce, N.H., Firmin, D.N., Wonke, B., Porter, J., Walker, J.M., Pennell, D.J. 2001. Cardiovascular T2-star (T2*) magnetic resonance for the early diagnosis of myocardial iron overload. Eur Heart J 22 (23):2171-9.

Anderson, L.J., Wonke, B., Prescott, E., Holden, S., Walker, J.M., Pennell, D.J. 2002. Comparison of effects of oral deferiprone and subcutaneous desferrioxamine on myocardial iron concentrations and ventricular function in beta-thalassaemia. Lancet 360 (9332):516-20.

Andrews, N.C. 1999. Disorders of iron metabolism. N Engl J Med 341 (26):1986-95.

Angelucci, E., Brittenham, G.M., McLaren, C.E., Ripalti, M., Baronciani, D., Giardini, C., Galimberti, M., Polchi, P., Lucarelli, G. 2000. Hepatic iron concentration and total body iron stores in thalassemia major. N Engl J Med 343 (5):327-31.

Armand, P.,. Kim, H.T, Cutler, C.S., Ho, V.T., Koreth, J., Alyea, E.P., Soiffer, R.J., Antin, J.H. 2007. Prognostic impact of elevated pretransplantation serum ferritin in patients undergoing myeloablative stem cell transplantation. Blood 109 (10):4586-8.

Armand, P., Kim, H.T., Rhodes, J., Sainvil, M.M., Cutler, C., Ho, V.T., Koreth, J., Alyea, E.P., Hearsey, D., Neufeld, E.J., Fleming, M.D., Steen, H., Anderson, D., Kwong, R.Y., Soiffer, R.J., Antin, J.H. 2011. Iron overload in patients with acute leukemia or MDS undergoing myeloablative stem cell transplantation. Biol Blood Marrow Transplant 17 (6):852-60.

Armand, P., Sainvil, M.M., Kim, H.T., Rhodes, J., Cutler, C., Ho, V.T., Koreth, J.,. Alyea, E.P, Neufeld, E.J., Kwong, R.Y., Soiffer, R.J., Antin, J.H. 2012. Does iron overload really matter in stem cell transplantation? Am J Hematol 87 (6):569-72.

Atlas, S.W. 1996. Magnetic resonance imaging of the brain and spine. 2nd ed. Philadelphia: Raven Press.

Bancroft, J.D. and Alan, S. 1990. Theory and practice of histological techniques. 3rd ed.: Churchill Livingstone.

Berdoukas, V., Chouliaras, G., Moraitis, P., Zannikos, K., Berdoussi, E., Ladis, V. 2009. The efficacy of iron chelator regimes in reducing cardiac and hepatic iron in patients with thalassaemia major: a clinical observational study. J Cardiovasc Magn Reson 11 (1):20.

Bitar, R., Leung, G., Perng, R., Tadros, S.,. Moody, A.R, Sarrazin, J., McGregor, C., Christakis, M., Symons, S., Nelson, A.,. Roberts, T.P. 2006. MR

pulse sequences: what every radiologist wants to know but is afraid to ask. Radiographics 26 (2):513-37.

Bland, J.M., and Altman, D.G. 2003. Applying the right statistics: analyses of measurement studies. Ultrasound Obstet Gynecol 22 (1):85-93.

Bloch, F. 1946. Nuclear Induction. Physical Review 70 (7-8):460-474.

Bloembergen, N., Purcell, E.M.,. Pound. R.V. 1948. Relaxation Effects in Nuclear Magnetic Resonance Absorption. Physical Review 73 (7):679-712.

Bohne, F., Martínez-Llordella, M., Lozano, J.J, Miquel, R., Benítez, C.,. Londoño, M.C., Manzia, T.M., Angelico, R., Swinkels, D.W., Tjalsma, H., López, M., Abraldes, J.G., Bonaccorsi-Riani, E., Jaeckel, E., Taubert, R., Pirenne, J., Rimola, A., Tisone, G., Sánchez-Fueyo, A. 2012. Intra-graft expression of genes involved in iron homeostasis predicts the development of operational tolerance in human liver transplantation. J Clin Invest 122 (1):368-82.

Bonkovsky, H.L., Rubin, R.B., Cable, E.E., Davidoff, A., Rijcken, T.H., Stark, D.D. 1999. Hepatic iron concentration: noninvasive estimation by means of MR imaging techniques. Radiology 212 (1):227-34.

Bonkovsky, H.L.,. Slaker, D.P., Bills, E.B.,. Wolf, D.C. 1990. Usefulness and limitations of laboratory and hepatic imaging studies in iron-storage disease. Gastroenterology 99 (4):1079-91.

Bonkovsky, H.L., Banner, B.F., Rothman, A.L. 1997. Iron and chronic viral hepatitis. Hepatology 25 (3):759-68.

Brand, A. 2002. Immunological aspects of blood transfusions. Transpl Immunol 10 (2-3):183-90.

Brissot, P., Bardou-Jacquet, E., Jouanolle, A.M., Loréal, O. 2011. Iron disorders of genetic origin: a changing world. Trends Mol Med 17 (12):707-13.

Brittenham, G.M., and Badman, D.G. 2003. Noninvasive measurement of iron: report of an NIDDK workshop. Blood 101 (1):15-9.

Brooks, R.A., Moiny, F., Gillis, P. 2001. On T2-shortening by weakly magnetized particles: the chemical exchange model. Magn Reson Med 45 (6):1014-20.

Bullen, J.J., Rogers, H.J., Spalding, P.B., Ward. C.G. 2006. Natural resistance, iron and infection: a challenge for clinical medicine. J Med Microbiol 55 (Pt 3):251-8.

Carneiro, A.A., Fernandes, J.P., de Araujo, D.B., Elias, J., Martinelli, A.L., Covas, D.T., Zago, M.A., Angulo, I.L., St Pierre, T.G., Baffa, O. 2005. Liver iron concentration evaluated by two magnetic methods: magnetic resonance imaging and magnetic susceptometry. Magn Reson Med 54 (1):122-8.

Cazzola, M., Della Porta, M.G., Malcovati, L. 2008. Clinical relevance of anemia and transfusion iron overload in myelodysplastic syndromes. Hematology Am Soc Hematol Educ Program:166-75.

Chacko, J., Pennell, D.J., Tanner, M.A., Hamblin, T.J., Wonke, B., Levy, T., Thomas, P.W., Killick, S.B. 2007. Myocardial iron loading by magnetic resonance imaging T2* in good prognostic myelodysplastic syndrome patients on long-term blood transfusions. Br J Haematol 138 (5):587-93.

Chavhan, G.B., Babyn, P.S., Thomas, B., Shroff, M.M., Haacke, E.M. 2009. Principles, techniques, and applications of T2*-based MR imaging and its special applications. Radiographics 29 (5):1433-49.

Chirnomas, D.S., Geukes-Foppen, M., Barry, K., Braunstein, J., Kalish, L.A., Neufeld, E.J., Powell, A.J. 2008. Practical implications of liver and heart iron load assessment by T2*-MRI in children and adults with transfusion-dependent anemias. Am J Hematol 83 (10):781-3.

Clark, P.R., and St Pierre, T.G. 2000. Quantitative mapping of transverse relaxivity (1/T(2)) in hepatic iron overload: a single spin-echo imaging methodology. Magn Reson Imaging 18 (4):431-8.

Clark, P.R., Chua-Anusorn, W., and St Pierre, T.G. 2003. Proton transverse relaxation rate (R2) images of liver tissue; mapping local tissue iron concentrations with MRI [corrected]. Magn Reson Med 49 (3):572-5.

Copelan, E.A. 2006. Hematopoietic stem-cell transplantation. N Engl J Med 354 (17):1813-26.

CTCAE. 2009. Common Terminology Criteria for Adverse Events (CTCAE) version 4.0. National Cancer Institute 2009. U.S. Department of Health and Human Services http://evs.nci.nih.gov/ftpl/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf.

Cuijpers, M.L., Raymakers, R.A., Mackenzie, M.A, de Witte, T.J., Swinkels, D.W. 2010. Recent advances in the understanding of iron overload in sideroblastic myelodysplastic syndrome. Br J Haematol 149 (3): 322-33.

de Bazelaire, C.M., Duhamel, G.D., Rofsky, N.M., Alsop, D.C. 2004. MR imaging relaxation times of abdominal and pelvic tissues measured in vivo at 3.0 T: preliminary results. Radiology 230 (3):652-9.

de Witte, T. 2008. The role of iron in patients after bone marrow transplantation. Blood Rev 22 Suppl 2:S22-8.

Di Tucci, A.A., Matta, G., Deplano, S., Gabbas, A., Depau, C., Derudas, D., Caocci, G., Agus, A., Angelucci, E. 2008. Myocardial iron overload assessment by T2* magnetic resonance imaging in adult transfusion dependent patients with acquired anemias. Haematologica 93 (9):1385-8.

Doyle, F.H., Pennock, J.M., Banks, L.M., McDonnell, M.J., Bydder, G.M.,. Steiner, R.E, Young, I.R., Clarke, G.J., Pasmore, T., Gilderdale, D.J. 1982. Nuclear magnetic resonance imaging of the liver: initial experience. AJR Am J Roentgenol 138 (2):193-200.

Dürken, M., Herrnring, C., Finckh, B., Nagel, S., Nielsen, P., Fischer, R., Berger, H.M., Moison, R.M., Pichlmeier, U., Kohlschütter, B.,. Zander, A.R, Kohlschütter, A. 2000. Impaired plasma antioxidative defense and increased nontransferrin-bound iron during high-dose chemotherapy and radiochemotherapy preceding bone marrow transplantation. Free Radic Biol Med 28 (6):887-94.

Dürken, M., Nielsen, P., Knobel, S., Finckh, B., Herrnring, C., Dresow, B., Kohlschütter, B., Stockschläder, M., Krüger, W.H., Kohlschütter, A., Zander, A.R. 1997. Nontransferrin-bound iron in serum of patients receiving bone marrow transplants. Free Radic Biol Med 22 (7):1159-63.

Ernst, O., Sergent, G., Bonvarlet, P., Canva-Delcambre, V., Paris, J. C., L'Herminé, C. 1997. Hepatic iron overload: diagnosis and quantification with MR imaging. AJR Am J Roentgenol 168 (5):1205-8.

Evens, A.M., Mehta, J., Gordon, L.I. 2004. Rust and corrosion in hematopoietic stem cell transplantation: the problem of iron and oxidative stress. Bone Marrow Transplant 34 (7):561-71.

Fenaux, P., and Rose, C. 2009. Impact of iron overload in myelodysplastic syndromes. Blood Rev 23 Suppl 1:S15-9.

Fisher M.A., Reiner C.S., Raptis D., Donati O., Goetti R., Clavien P-A., Alkadhi H. 2011. Eur Radiol 21:1727-32.

Fleming, R.E., and Bacon, B.R. 2005. Orchestration of iron homeostasis. N Engl J Med 352 (17):1741-4.

Gandon, Y., Olivié, D., Guyader, D., Aubé, C., Oberti, F., Sebille, V., Deugnier, Y. 2004. Non-invasive assessment of hepatic iron stores by MRI. Lancet 363 (9406):357-62.

Ganz, T. 2003. Hepcidin, a key regulator of iron metabolism and mediator of anemia of inflammation. Blood 102 (3):783-8.

Ganz, T. 2009. Iron in innate immunity: starve the invaders. Curr Opin Immunol 21 (1):63-7.

Ganz, T. 2011. Hepcidin and iron regulation, 10 years later. Blood 117 (17):4425-33.

Ganz, T., and Nemeth, E. 2012. Hepcidin and iron homeostasis. Biochim Biophys Acta 1823 (9):1434-43.

Gentry-Nielsen, M.J., Preheim, L.C., Lyman, K.N., McDonough, K.H., Potter, B.J. 2001. Use of rat models to mimic alterations in iron homeostasis during human alcohol abuse and cirrhosis. Alcohol 23 (2):71-81.

Ghugre, N.R., Coates, T.D., Nelsonm M.D., Wood, J.C. 2005. Mechanisms of tissue-iron relaxivity: nuclear magnetic resonance studies of human liver biopsy specimens. Magn Reson Med 54 (5):1185-93.

Ghugre, N.R., and Wood, J.C. 2011. Relaxivity-iron calibration in hepatic iron overload: probing underlying biophysical mechanisms using a Monte Carlo model. Magn Reson Med 65 (3):837-47.

Ghugre, N.R., Enriquez, C.M., Gonzalez, I., Nelson, M.D. Jr., Coates, T.D., Wood, J.C. 2006. MRI detects myocardial iron in the human heart. Magn Reson Med 56 (3):681-6.

Gooley, T. A., Chien, J.W., Pergam, S.A., Hingorani, S., Sorror, M.L., Boeckh, M., Martin, P.J., Sandmaier, B.M., Marr, K.A., Appelbaum, F.R., Storb, R., McDonald, G.B. 2010. Reduced Mortality after Allogeneic Hematopoietic-Cell Transplantation. New England Journal of Medicine 363 (22):2091-2101.

Gossuin, Y., Gillis, P., Muller, R.N., Hocq, A. 2007. Relaxation by clustered ferritin: a model for ferritin-induced relaxation in vivo. NMR Biomed 20 (8):749-56.

Gossuin, Y., Muller, R.N., Gillis, P. 2004. Relaxation induced by ferritin: a better understanding for an improved MRI iron quantification. NMR Biomed 17 (7):427-32.

Gratwohl, A., Stern, M., Brand, R., Apperley, J., Baldomero, H., de Witte, T., Dini, G., Rocha, V., Passweg, J., Sureda, A., Tichelli, A. Niederwieser, D., European Group for Blood and Marrow Transplantation and the European Leukemia Net. 2009. Risk score for outcome after allogeneic hematopoietic stem cell transplantation: a retrospective analysis. Cancer 115 (20):4715-26.

He, T., Gatehouse, P.D., Kirk, P., Tanner, M.A., Smith, G.C., Keegan, J., Mohiaddin, R.H., Pennell, D.J., Firmin, D.N. 2007. Black-blood T2* technique for myocardial iron measurement in thalassemia. J Magn Reson Imaging no 25 (6):1205-9.

He, T., Gatehouse, P.D., Smith, G.C., Mohiaddin, R.H., Pennell, D.J., Firmin, D,N. 2008. Myocardial T2* measurements in iron-overloaded thalassemia:

An in vivo study to investigate optimal methods of quantification. Magn Reson Med 60 (5):1082-9.

He, T., Kirk, P., Firmin, D.N., Lam, W,M,, Chu, W.C., Au, W.Y., Chan, G.C., Tan, R.S., Ng, I., Biceroglu, S., Aydinok, Y., Fogel,,M.A., Cohen, A.R., Pennell, D.J. 2008. Multi-center transferability of a breath-hold T2 technique for myocardial iron assessment. J Cardiovasc Magn Reson 10 (1):11.

He, T., Smith. G.C, Gatehouse, P.D., Mohiaddin, R.H., Firmin, D.N., Pennell, D.J. 2009. On using T2 to assess extrinsic magnetic field inhomogeneity effects on T2* measurements in myocardial siderosis in thalassemia. Magnetic Resonance in Medicine 61 (3):501-6.

Hernando, D., Cook, R.J., Diamond, C., Reder, S,B., 2013. Magnetic susceptibility as a B0 field strenght independent MRI biomarker of liver iron overload. Magnetic Resonance in Medicine 70:648-56.

Hernando, D., Levin, Y.S., Sirlin, C.B., Reeder, S.B. 2014. Quantification of liver iron with MRI: State of the art and remaining challenges. J. Magn. Reson. Imaging. doi: 10.1002/jmri.24584.

Hernando, D., Vigen, K.K., Shimakawa, A., Reeder, S.B. 2012. R2*mapping in the presence of macroscopic B0 field variations. Magnetic Resonance in Medicine. 68:830-40.

Horny, H-P., Sotlar K., Valent P. 2007. Diagnostic value of histology and immunohistochemistry in myelodysplastic syndromes. Leukemia research 31 (12): 1609-16

Howard J.M., Ghent C.N., Carey L.S., Flanagan P.R, Valberg L.S. 1983. Diagnostic efficacy of hepatic computed tomography in the detection of body iron overload. Gastroenterology 84 (2):209-15.

Iancu, T. C. 1992. Ferritin and hemosiderin in pathological tissues. Electron Microsc Rev 5 (2):209-29.

Jensen, P.D., Jensen, F.T., Christensen, T., Eiskjaer, H., Baandrup, U., Nielsen, J.L. 2003. Evaluation of myocardial iron by magnetic resonance imaging during iron chelation therapy with deferrioxamine: indication of close relation between myocardial iron content and chelatable iron pool. Blood 101 (11):4632-9.

Kamble, R., and Mims, M. 2006. Iron-overload in long-term survivors of hematopoietic transplantation. Bone Marrow Transplant 37 (8):805-6.

Karam, L. B., Disco, D., Jackson, S. M., Lewin, D., McKie, V., Baker, R. D., Baker, S. S., Laver, J. H., Nietert, P. J., Abboud, M. R. 2008. Liver biopsy results in patients with sickle cell disease on chronic transfusions: poor correlation with ferritin levels. Pediatr Blood Cancer 50 (1):62-5.

Kataoka, K., Nannya, Y., Hangaishi, A., Imai, Y., Chiba, S., Takahashi, T., Kurokawa, M. 2009. Influence of pretransplantation serum ferritin on nonrelapse mortality after myeloablative and nonmyeloablative allogeneic hematopoietic stem cell transplantation. Biol Blood Marrow Transplant 15 (2):195-204.

Kohgo, Y., Ikuta, K., Ohtake, T., Torimoto, Y., Kato, J. 2008. Body iron metabolism and pathophysiology of iron overload. Int J Hematol 88 (1):7-15.

Kolnagou, A., Economides, C., Eracleous, E., Kontoghiorghes, G. J. 2006. Low serum ferritin levels are misleading for detecting cardiac iron overload and increase the risk of cardiomyopathy in thalassemia patients. The importance of cardiac iron overload monitoring using magnetic resonance imaging T2 and T2*. Hemoglobin 30 (2):219-27.

Konen, E., Ghoti, H., Goitein, O., Winder, A., Kushnir, T., Eshet, Y., Rachmilewitz, E.. 2007. No evidence for myocardial iron overload in multitransfused patients with myelodysplastic syndrome using cardiac magnetic resonance T2 technique. Am J Hematol 82 (11):1013-6.

Kontoyiannis, D.P., Chamilos, G., Lewis, R.E., Giralt, S., Cortes, J., Raad, I.I., Manning, J.T., Han, X. 2007. Increased bone marrow iron stores is an independent risk factor for invasive aspergillosis in patients with high-risk hematologic malignancies and recipients of allogeneic hematopoietic stem cell transplantation. Cancer 110 (6):1303-6.

Koreth, J., and. Antin, J.H. 2010. Iron overload in hematologic malignancies and outcome of allogeneic hematopoietic stem cell transplantation. Haematologica 95 (3):364-6.

Kowdley, K.V., Trainer, T.D., Saltzman, J.R., Pedrosa, M., Krawitt, E..L, Knox,T.A., Susskind, K., Pratt, D., Bonkovsky, H.L., Grace,.N.D., Kaplan, M.M. 1997. Utility of hepatic iron index in American patients with hereditary hemochromatosis: a multicenter study. Gastroenterology 113 (4):1270-7.

Krause, A., Neitz, S., Mägert, H.J., Schulz, A., Forssmann, W.G., Schulz-Knappe, P., Adermann, K. 2000. LEAP-1, a novel highly disulfide-bonded human peptide, exhibits antimicrobial activity. FEBS Lett 480 (2-3):147-50.

Leitch, H.A. 2011. Controversies surrounding iron chelation therapy for MDS. Blood Rev 25 (1):17-31.

Leung, A.W., Chu, W.C., Lam, W.W., Lee, V., Li, C.K. 2009. Magnetic resonance imaging assessment of cardiac and liver iron load in transfusion dependent patients. Pediatr Blood Cancer 53 (6):1054-9.

Li, T.Q., Aisen, A.M., Hindmarsh, T. 2004. Assessment of hepatic iron content using magnetic resonance imaging. Acta Radiol 45 (2):119-29.

Lim, R.P., Tuvia, K., Hajdu, C.H., Losada, M., Gupta, R., Parikh, T., Babb, J.S., Taouli, B. 2010. Quantification of hepatic iron deposition in patients with liver disease: comparison of chemical shift imaging with single-echo T2*-weighted imaging. AJR Am J Roentgenol 194 (5):1288-95.

Mahindra, A., Bolwell, B., Sobecks, R., Rybicki, L., Pohlman, B., Dean, R., Andresen, S., Sweetenham, J., Kalaycio, M., Copelan, E. 2009a. Elevated pretransplant ferritin is associated with a lower incidence of chronic graft-versus-host disease and inferior survival after myeloablative allogeneic haematopoietic stem cell transplantation. Br J Haematol 146 (3):310-6.

Mahindra, A., Sobecks, R., Rybicki, L., Pohlman, B., Dean, R., Andresen, S., Kalaycio, M., Sweetenham, J., Bolwell, B., Copelan, E. 2009b. Elevated pretransplant serum ferritin is associated with inferior survival following nonmyeloablative allogeneic transplantation. Bone Marrow Transplant 44 (11):767-8.

Majhail, N.S.,. Lazarus, H.M, Burns, L.J. 2008. Iron overload in hematopoietic cell transplantation. Bone Marrow Transplant 41 (12):997-1003.

Majhail, N.S., Lazarus, H.M., Burns, L.J. 2010. A prospective study of iron overload management in allogeneic hematopoietic cell transplantation survivors. Biol Blood Marrow Transplant 16 (6):832-7.

Malcovati, L. 2007. Impact of transfusion dependency and secondary iron overload on the survival of patients with myelodysplastic syndromes. Leuk Res 31 Suppl 3:S2-6.

Malcovati, L., and Cazzola, M. 2011. The relevance of transfusion-dependency in the prognostic assessment of patients with myeloid neoplasms. Am J Hematol 86 (3):241-3.

Malcovati, L., Germing, U., Kuendgen, A., Della Porta, M.G., Pascutto, C., Invernizzi, R., Giagounidis, A., Hildebrandt, B., Bernasconi, P., Knipp, S., Strupp, C., Lazzarino, M., Aul, C., Cazzola, M. 2007. Time-dependent prognostic scoring system for predicting survival and leukemic evolution in myelodysplastic syndromes. J Clin Oncol 25 (23):3503-10.

Marcos-Gragera, R., Allemani, C., Tereanu, C., De Angelis, R., Capocaccia, R., Maynadie, M., Luminari, S., Ferretti, S., Johannesen, T.B., Sankila, R., Karjalainen-Lindsberg, M.L., Simonetti, A., Martos, M.C., Raphaël, M., Giraldo, P., Sant, M., HAEMACARE Working Group. 2011. Survival of European patients diagnosed with lymphoid neoplasms in 2000-2002: results of the HAEMACARE project. Haematologica 96 (5):720-8.

Massover, W.H. 1993. Ultrastructure of ferritin and apoferritin: A review. Micron. 24 (4):389-437.

Maynadié, M., De Angelis, R., Marcos-Gragera, R., Visser, O., Allemani, C., Tereanu, C., Capocaccia, R., Giacomin, A., Lutz , J.M., Martos, C., Sankila, R., Johannesen, T.B., Simonetti, A., Sant, M., HAEMACARE Working Group. 2013. Survival of European patients diagnosed with myeloid malignancies: a HAEMACARE study. Haematologica 98 (2):230-8.

Merkle, E.M., and Nelson, R.C. 2006.Dual gradient-echo in-phase and opposed-phase hepatic MR imaging: a useful tool for evaluating more than fatty infiltration or fatty sparing. Radiographics 26 (5):1409-18.

Miceli, M.H., Dong, L., Grazziutti, M.L., Fassas, A., Thertulien, R., Van Rhee, F., Barlogie, B., Anaissie, E.J. 2006. Iron overload is a major risk factor for severe infection after autologous stem cell transplantation: a study of 367 myeloma patients. Bone Marrow Transplant 37 (9):857-64.

|MR Glossary. 2009. Siemens Medical Solutions©. USA. Inc.http://usa.healthcare.siemens.com/siemens_hwem-hwem_ssxa_websites-context-root/wcm/idc/groups /public/@us/@clinicalspec/documents/download/mdaw/ndi1/~edisp/mr_glossary_of_terms-00305952.pdf

Nairz, M., Schroll, A., Sonnweber, T., Weiss, G. 2010. The struggle for iron - a metal at the host-pathogen interface. Cell Microbiol 12 (12):1691-702.

Nielsen P., Engelhardt R., Fisher R., Heinrich H., Langkowski J., Bucheler E. 1992. Noninvasive liver-iron quantification by computed tomography in iron-overloaded rats. Investigative Radiology 27(4):312-6.

Nielsen, P., Günther, U., Dürken, M., Fischer, R., Düllmann, J. 2000. Serum ferritin iron in iron overload and liver damage: correlation to body iron stores and diagnostic relevance. J Lab Clin Med no. 135 (5):413-8.

Noetzli, L.J., Carson, S.M., Nord, A.S., Coates, T.D., Wood, J.C. 2008. Longitudinal analysis of heart and liver iron in thalassemia major. Blood 112 (7):2973-8.

Noone, T.C., Semelka, R.C., Siegelman, E,S., Balci, N.C., Hussain, S.M., Kim, P.N., Mitchell, D.G. 2000. Budd-Chiari syndrome: spectrum of appearances of acute, subacute, and chronic disease with magnetic resonance imaging. J Magn Reson Imaging 11 (1):44-50.

Olivieri, N.F., and Brittenham, G.M. 1997. Iron-chelating therapy and the treatment of thalassemia. Blood 89 (3):739-61.

Olivieri, N.F.,. Brittenham, G. Matsui, M,D., Berkovitch, M.,. Blendis, L.M, Cameron, R.G., McClelland R.A., Liu, P.P., Templeton, D.M., Koren, G. 1995. Iron-chelation therapy with oral deferipronein patients with thalassemia major. N Engl J Med 332 (14):918-22.

Olivieri, N.F., Nathan, D.G., MacMillan, J.H., Wayne, A.S., Liu, P.P., McGee, A., Martin, M., Koren, G., Cohen, A.R. 1994. Survival in medically treated patients with homozygous beta-thalassemia. N Engl J Med 331 (9):574-8.

Olthof, A.W., Sijens, P.E., Kreeftenberg, H.G., Kappert, P., Irwan, R., van der Jagt, E.J., Oudkerk, M. 2007. Correlation between serum ferritin levels and liver iron concentration determined by MR imaging: impact of hematologic disease and inflammation. Magn Reson Imaging 25 (2):228-31.

Olthof, A.W., Sijens, P.E., Kreeftenberg, H.G., Kappert, P., van der Jagt, E.J., Oudkerk, M. 2009. Non-invasive liver iron concentration measurement by MRI: comparison of two validated protocols. Eur J Radiol 71 (1):116-21.

Ozyilmaz, E., Aydogdu, M., Sucak, G., Aki, S.Z., Ozkurt, Z.N., Yegin, Z.A., Kokturk, N. 2010. Risk factors for fungal pulmonary infections in hematopoietic stem cell transplantation recipients: the role of iron overload. Bone Marrow Transplant 45 (10):1528-33.

Palva, I. 2005. Veritaudit. Edited by Vilpo, J. Helsinki. Medivil Oy.

Papakonstantinou, O., Alexopoulou, E., Economopoulos, N., Benekos, O., Kattamis, A., Kostaridou, S., Ladis, V., Efstathopoulos, E., Gouliamos, A., Kelekis, N.L. 2009. Assessment of iron distribution between liver, spleen, pancreas, bone marrow, and myocardium by means of R2 relaxometry with MRI in patients with beta-thalassemia major. J Magn Reson Imaging 29 (4):853-9.

Park, C.H., Valore, E.V., Waring, A.J, Ganz, T. 2001. Hepcidin, a urinary antimicrobial peptide synthesized in the liver. J Biol Chem 276 (11):7806-10.

Park, J., Ohyashiki, K., Akata, S., Takara, K., Uno, R., Kakizaki, D., Miyazawa, K., Kimura, Y., Tokuuye, K. 2009. Evaluation of cardiac iron overload in transfusion-dependent adult marrow failure patients by magnetic resonance imaging. Leuk Res 33 (6):756-8.

Peng, P.,Huang, Z., Long, L., Zhao, F., Li C., Li, W., He, T. 2013. Liver iron quantification by 3 Tesla MRI: Calibration on a rabbit model. J Magn Reson Imaging 38 (6):1585-90.

Platzbecker, U., Bornhäuser, M., Germing, U., Stumpf, J., Scott, B.L., Kröger, N., Schwerdtfeger, R., Böhm, A., Kobbe, G., Theuser, C., Rabitsch, W., Valent, P., Sorror, M.L., Ehninger, G., Deeg, H.J. 2008. Red blood cell transfusion dependence and outcome after allogeneic peripheral blood stem cell transplantation in patients with de novo myelodysplastic syndrome (MDS). Biol Blood Marrow Transplant 14 (11):1217-25.

Pooley, R.A. 2005. AAPM/RSNA physics tutorial for residents: fundamental physics of MR imaging. Radiographics 25 (4):1087-99.

Porter, J.B. 2005. Liver iron measurement by MRI. Blood 105 (2):437-438.

Przepiorka, D., Weisdorf, D., Martin, P., Klingemann, H. G., Beatty, P., Hows, J., Thomas, E.D. 1995. 1994 Consensus Conference on Acute GVHD Grading. Bone Marrow Transplant 15 (6):825-8.

Pullarkat, V. 2010. Iron overload in patients undergoing hematopoietic stem cell transplantation. Adv Hematol 2010: doi: 10.1155/2010/345756.

Pullarkat, V., Blanchard, S., Tegtmeier, B., Dagis, A., Patane, K., Ito, J., Forman, S.J. 2008. Iron overload adversely affects outcome of allogeneic hematopoietic cell transplantation. Bone Marrow Transplant 42 (12):799-805.

Purcell, E.M., Torrey, H.C., Pound, R.V. 1946. Resonance Absorption by Nuclear Magnetic Moments in a Solid. Physical Review 69 (1-2):37-38.

Ratanatharathorn, V., Ayash, L., Lazarus, H.M., Fu, J., Uberti, J.P. 2001. Chronic graft-versus-host disease: clinical manifestation and therapy. Bone Marrow Transplant 28 (2):121-9.

Rose, C., Erns, O., Hecquet, B., Maboudou, P., Renom, P., Noel, M.P., Yakoub-Agha, I., Bauters, F., Jouet, J.P. 2007. Quantification by magnetic resonance imaging and liver consequences of post-transfusional iron overload alone in long term survivors after allogeneic hematopoietic stem cell transplantation (HSCT). Haematologica 92 (6):850-3.

Rose, C., Vandevenne, P., Bourgeois, E., Cambier, N., Ernst, O. 2006. Liver iron content assessment by routine and simple magnetic resonance imaging procedure in highly transfused patients. Eur J Haematol 77 (2):145-9.

Sahlstedt, L., Ebeling, F., von Bonsdorff, L., Parkkinen, J., Ruutu, T. 2001. Non-transferrin-bound iron during allogeneic stem cell transplantation. Br J Haematol 113 (3):836-8.

Sahlstedt, L., von Bonsdorff, L., Ebeling, F., Parkkinen, J., Juvonen, E., Ruutu, T. 2009. Non-transferrin-bound iron in haematological patients during chemotherapy and conditioning for

autologous stem cell transplantation. Eur J Haematol 83 (5):455-9.

Sant, M., Allemani, C., Tereanu, C., De Angelis, R., Capocaccia, R., Visser, O., Marcos-Gragera, R., Maynadié, M., Simonetti, A., Lutz, J.M., Berrino, F., HAEMACARE Working Group. 2010. Incidence of hematologic malignancies in Europe by morphologic subtype: results of the HAEMACARE project. Blood 116 (19):3724-34.

Schafer, A.I., Cheron, R.G., Dluhy, R., Cooper, B., Gleason, R.E., Soeldner, J.S., Bunn, H.F. 1981. Clinical consequences of acquired transfusional iron overload in adults. N Engl J Med 304 (6):319-24.

Schild H.H. 1990. MRI made easy © Schering AG. Germany. Nationales Druckhaus Berlin.

Shankland, K.R., Armitage, J.O., Hancock, B.W. 2012. Non-Hodgkin lymphoma. Lancet 380 (9844):848-57.

St Pierre, T.G., Clark, P.R., Chua-anusorn, W., Fleming, A.J., Jeffrey, G.P., Olynyk, J.K., Pootrakul, P., Robins, E., Lindeman, R. 2005. Noninvasive measurement and imaging of liver iron concentrations using proton magnetic resonance. Blood 105 (2):855-61.

St Pierre, T.G., El-Beshlawy, A., Elalfy, M., Al Jefri, A., Al Zir, K., Daar, S., Habr, D., Kriemler-Krahn, U., Taher, A. 2014. Multicenter validation of spin-density projection-assisted R2-MRI for noninvasive measurement of liver iron concentration. Magnetic Resonance in Medicine 71:2215-23.

Stark, D.D., Moseley, M.E., Bacon, B.R., Moss, A.A., Goldberg, H.I., Bass, N.M., James, T.L. 1985. Magnetic resonance imaging and spectroscopy of hepatic iron overload. Radiology 154 (1):137-42.

Stark, D.D. 1991. Hepatic iron overload: paramagnetic pathology. Radiology 179 (2):333-5.

Storey, P., Thompson, A.A., Carqueville, C.L., Wood, J.C., de Freitas, R.A., Rigsby, C.K. 2007. R2* imaging of transfusional iron burden at 3T and comparison with 1.5T. J Magn Reson Imaging 25 (3):540-7.

Sucak, G.T., Yegin, Z.A., Ozkurt, Z.N., Aki, S.Z., Yağci, M. 2010. Iron overload: predictor of adverse outcome in hematopoietic stem cell transplantation. Transplant Proc 42 (5):1841-8.

Tachibana, T., Tanaka, M., Takasaki, H., Numata, A., Ito, S., Watanabe, R., Hyo, R., Ohshima, R., Hagihara, M., Sakai, R., Fujisawa, S., Tomita, N., Fujita, H., Maruta, A., Ishigatsubo, Y., Kanamori, H. 2011. Pretransplant serum ferritin is associated with bloodstream infections within 100 days of allogeneic stem cell transplantation for myeloid malignancies. Int J Hematol 93 (3):368-74.

Tanner, M.A., He, T., Westwood, M.A., Firmin, D.N., Pennell, D.J. 2006. Multi-center validation of the transferability of the magnetic resonance T2* technique for the quantification of tissue iron. Haematologica 91 (10):1388-91.

Thomsen, C., Becker, U., Winkler, K., Christoffersen, P., Jensen, M., Henriksen, O. 1994. Quantification of liver fat using magnetic resonance spectroscopy. Magn Reson Imaging 12 (3):487-95.

Townsend, W., and Linch, D. 2012. Hodgkin's lymphoma in adults. Lancet 380 (9844):836-47.

Treleaven, J. and Barrett, A.J. 2009. Hematopoietic stem cell transplantation in clinical practise. Edinburgh. Churchill Livingstone.

Twomey, P.J., and Kroll, M.H. 2008. How to use linear regression and correlation in quantitative method comparison studies. Int J Clin Pract 62 (4):529-38.

Vardiman, J.W. 2010. The World Health Organization (WHO) classification of tumors of the hematopoietic and lymphoid tissues: an overview with emphasis on the myeloid neoplasms. Chem Biol Interact 184 (1-2):16-20.

Villeneuve, J.P., Bilodeau, M., Lepage, R., Côté, J., Lefebvre, M. 1996. Variability in hepatic iron concentration measurement from needle-biopsy specimens. J Hepatol 25 (2):172-7.

Voskaridou, E., Douskou, M., Terpos, E., Papassotiriou, I., Stamoulakatou, A., Ourailidis, A., Loutradi, A., Loukopoulos, D. 2004. Magnetic resonance imaging in the evaluation of iron overload in patients with beta thalassaemia and sickle cell disease. Br J Haematol 126 (5):736-42.

Vymazal, J., Brooks, R.A., Zak, O., McRill, C., Shen, C., Di Chiro, G. 1992. T1 and T2 of ferritin at different field strengths: effect on MRI. Magn Reson Med 27 (2):368-74.

Wahlin, A., Lorenz, F., Fredriksson, M., Remberger, M., Wahlin, B.E., Hägglund, H. 2011. Hyperferritinemia is associated with low incidence of graft versus host disease, high relapse rate, and impaired survival in patients with blood disorders receiving allogeneic hematopoietic stem cell grafts. Med Oncol 28 (2):552-8.

Wang, L., and Cherayil, B.J. 2009. Ironing out the wrinkles in host defense: interactions between iron homeostasis and innate immunity. J Innate Immun 1 (5):455-64.

Wermke, M., Schmidt, A., Middeke, J.M., Sockel, K., von Bonin, M., Schonefeldt,C., Mair, S., Plodeck, V., Laniado, M., Weiss, G., Schetelig, J., Ehninger, G., Theurl, I., Bornhauser, M., Platzbecker, U. 2012. MRI-based liver iron content predicts for non-relapse mortality in MDS and AML patients

undergoing allogeneic stem cell transplantation. Clin Cancer Res 18 (23):6460-68.

Westbrook, C., Roth, C.K., Talbot, J. 2011. MRI in practice. 4th ed. / Catherine Westbrook, Carolyn Kaut Roth and John Talbot. ed. Oxford: Wiley-Blackwell.

Westwood, M., Anderson, L.J., Firmin, D.N., Gatehouse, P. D., Charrier, C.C., Wonke, B., Pennell, D.J. 2003a. A single breath-hold multiecho T2* cardiovascular magnetic resonance technique for diagnosis of myocardial iron overload. J Magn Reson Imaging 18 (1):33-9.

Westwood, M.A., Anderson, L.J., Firmin, D.N., Gatehouse, P.D., Lorenz, C. H., Wonke, B., Pennell, D.J. 2003b. Interscanner reproducibility of cardiovascular magnetic resonance T2* measurements of tissue iron in thalassemia. J Magn Reson Imaging 18 (5):616-20.

Westwood, M.A., Firmin, D.N., Gildo, M., Renzo, G., Stathis, G., Markissia, K., Vasili, B., Pennell, D.J. 2005. Intercentre reproducibility of magnetic resonance T2* measurements of myocardial iron in thalassaemia. Int J Cardiovasc Imaging 21 (5):531-8.

Wood, J.C. 2007. Magnetic resonance imaging measurement of iron overload. Curr Opin Hematol 14 (3):183-90.

Wood, J.C. 2008. Cardiac iron across different transfusion-dependent diseases. Blood Rev 22 Suppl 2:S14-21.

Wood, J.C., Aguilar, M., Otto-Duessel, M., Nick, H., Nelson, M.D., Moats, R. 2008. Influence of iron chelation on R1 and R2 calibration curves in gerbil liver and heart. Magn Reson Med 60 (1):82-9.

Wood, J.C., Enriquez, C., Ghugre, N., Tyzka, J.M., Carson, S., Nelson, M.D., Coates, T.D. 2005. MRI R2 and R2* mapping accurately estimates hepatic iron concentration in transfusion-dependent thalassemia and sickle cell disease patients. Blood 106 (4):1460-5.

Wood, J.C., and Ghugre, N. 2008. Magnetic resonance imaging assessment of excess iron in thalassemia, sickle cell disease and other iron overload diseases. Hemoglobin 32 (1-2):85-96.

Wood J.C., Mo A., Gera A., Koh M., Coates T., Gilsanz V. 2011. Quantitative computed tomography assessment of transfusional iron overload. British Journal of Haematology 153:780-5.

Wood, J.C., Tyszka, J.M., Carson, S., Nelson, M.D., Coates, T.D. 2004. Myocardial iron loading in transfusion-dependent thalassemia and sickle cell disease. Blood 103 (5):1934-6.

Yegin, Z.A., Paşaoğlu, H., Aki, S.Z, Ozkurt, Z.N., Demirtaş, C., Yağci, M., Acar, K., Sucak, G.T. 2011. Pro-oxidative/antioxidative imbalance: a key indicator of adverse outcome in hematopoietic stem cell transplantation. Int J Lab Hematol 33 (4):414-23.

Young, N.S., Scheinberg, P., Calado, R.T. 2008. Aplastic anemia. Curr Opin Hematol 15 (3):162-8.