Identifying Optimal Anemia Management Practices in Hemodialysis

by

Angelo Karaboyas

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Epidemiological Science)
in the University of Michigan
2019

Doctoral Committee:

Professor Emeritus Hal Morgenstern, Chair
Assistant Professor Nancy Fleischer
Dr. Ronald L. Pisoni, Arbor Research Collaborative for Health
Dr. Bruce M. Robinson, Arbor Research Collaborative for Health
Professor Douglas E. Schaubel

Angelo Karaboyas

akaraboy@umich.edu

ORCID iD: 0000-0001-8846-5360

DEDICATION

For Joseph Nowak, in memoriam. You can call me Doc now, Papa.

ACKNOWLEDGMENTS

This dissertation was written by Angelo Karaboyas, with extensive contributions from my faculty advisor and mentor, Hal Morgenstern. Additional contributions were made by Bruce M. Robinson, Ronald L. Pisoni, Nancy L. Fleischer, Douglas E. Schaubel, Brian A. Bieber, Raymond C Vanholder, Stefan H. Jacobson, Elke Schaeffner, Manish M. Sood, Tadao Akizawa, Masaaki Inaba, Nafeesa N Dhalwani, Glen James, Marvin V Sinsakul, Sandra Waechter, and all members of the Dialysis Outcomes and Practice Patterns Study (DOPPS) analytical, programming, editorial, and project support staff.

Funding to support their dissertation work was obtained from a consortium of organizations. Funding for the research described in Chapter II was supported by Vifor Fresenius Medical Care Renal Pharma. Funding for the research described in Chapter III was supported by AstraZeneca. Funding for the DOPPS Program – data utilized for the entirety of this dissertation research – is supported by Amgen (since 1996, founding sponsor), Kyowa Hakko Kirin (since 1999 for Japan DOPPS), and Baxter Healthcare Corp. Additional support for specific projects and countries is provided by Akebia Therapeutics, AstraZeneca, European Renal Association-European Dialysis & Transplant Association (ERA-EDTA), Fibrogen, Fresenius Medical Care Asia-Pacific Ltd, Fresenius Medical Care Canada Ltd, German Society of Nephrology (DGfN), Italian Society of Nephrology (SIN), Janssen, Japanese Society for Peritoneal Dialysis (JSPD),

Kidney Care UK, MEDICE Arzneimittel Pütter GmbH & Co KG, Otsuka America, Proteon Therapeutics, the Association of German Nephrology Centres, and Vifor Fresenius Medical Care Renal Pharma. Public funding and support is provided for specific DOPPS projects, ancillary studies, or affiliated research projects by National Health & Medical Research Council (NHMRC) in Australia, Belgian Federal Public Service of Public Health in Belgium, Cancer Care Ontario (CCO) through the Ontario Renal Network (ORN) in Canada, French National Institute of Health and Medical Research (INSERM) in France, Thailand Research Foundation (TRF), Chulalongkorn University Matching Fund, King Chulalongkorn Memorial Hospital Matching Fund, and the National Research Council of Thailand (NRCT) in Thailand, National Institute for Health Research (NIHR) via the Comprehensive Clinical Research Network (CCRN), and Kidney Research UK (KRUK) in the United Kingdom, and the Agency for Healthcare Research and Quality (AHRQ) and National Institutes of Health (NIH) in the US.

All grants are made to my employer, Arbor Research Collaborative for Health, and not to Angelo Karaboyas directly. All support is provided without restrictions on publications.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Optimal anemia management strategies for end-stage kidney disease patients treated with hemodialysis are unknown, with controversies over how best to utilize erythropoiesis-stimulating agents (ESA) and intravenous iron to support hemoglobin levels and minimize adverse events. With large randomized trials rare in nephrology, it is thus crucial that research questions are clearly defined, study designs are appropriately selected, and analytic techniques are properly implemented when using observational data. The three aims of this dissertation attempt to address current controversies in anemia management using innovative statistical methods, leveraging data from the Dialysis Outcomes and Practice Patterns Study (DOPPS), an international prospective cohort study of hemodialysis patients.

Aim 1 focused on anemia management during the transition period to hemodialysis. Among patients who initiated hemodialysis with hemoglobin <10 g/dL, the highest (vs. low) doses of ESA and intravenous iron were each associated with elevated mortality. To assess the impact of pre-dialysis anemia treatment, a seemingly counterintuitive design – restricting to patients who achieved target hemoglobin (>=10 g/dL) four months later – was used to limit inclusion of patients whose low hemoglobin at hemodialysis initiation was likely confounded by poor health status. Even in this subset, anemia at hemodialysis initiation was common and associated with elevated mortality. A more proactive approach to anemia management prior to end-stage kidney disease may thus

avoid aggressive correction of hemoglobin levels during the early dialysis period and improve survival.

Aim 2 focused on how hemoglobin response to ESA therapy may be blunted by inflammation. Hemoglobin and ESA doses were compared over the 3 months before and after detection of new inflammation, defined as an acute C-reactive protein increase from <=5 to >10 mg/L. Confounding due to baseline characteristics, whether measured (age, sex, comorbidity history) or unmeasured (genetic or environmental factors), was avoided by this longitudinal self-matched design. Patients experiencing new inflammation had *both* higher ESA doses and lower hemoglobin (vs. pre-inflammation levels), supporting the hypothesis that inflammation increases resistance to ESA treatment. Quicker recognition of new inflammation in hemodialysis patients could help identify the cause of worsening anemia and guide ESA and intravenous iron dosing decisions more proactively.

Aim 3 focused on applying the parametric g-formula, an extension of standardization to longitudinal data, to replicate a randomized trial using observational data. DOPPS data were used to compare iron supplementation strategies, with the goal of mimicking the recently published PIVOTAL randomized trial. Comparing the proactive high-dose vs. reactive low-dose strategy, 1-year mortality risk was 20% greater under the parametric g-formula simulation, but similar in the PIVOTAL trial. Simulated differences for all secondary outcomes were directionally consistent but of lesser magnitude than in the PIVOTAL trial. Success in mimicking the PIVOTAL trial was mixed, and potential explanations for the divergent results include model misspecification and/or differences in the study populations. This example illustrates the potential of the parametric g-

formula to evaluate many variations of complex interventions across different populations, which could prove enormously informative in the age of big data.

This dissertation outlines critical gaps in the literature on anemia management in hemodialysis patients, and describes three studies that utilize innovative designs and complex statistical analyses to address these gaps. These studies attempt to advance both the optimization of anemia management strategies in hemodialysis patients and the use of causal inference principles to guide epidemiologic research using observational data.

CHAPTER I

Introduction

Anemia in Chronic Kidney Disease

For patients with chronic kidney disease (CKD) who progress to end-stage kidney disease (ESKD), the most common treatment in the United States (US) is in-center hemodialysis (HD) for 3-4 hours three times per week. During HD treatment, a filter is used to clean toxins from the blood to maintain electrolyte balance and remove extra fluid from the body to avoid the dangerous complications of kidney failure. These complications include (1) fluid overload, leading to swelling, high blood pressure, and fluid in the lungs, (2) hyperkalemia, leading to life-threatening arrhythmia, (3) cardiovascular disease, leading to stroke or heart attack, (4) mineral and bone disorder, leading to bone fractures and vascular calcification, and (5) anemia, which is discussed in detail below.[1,2]

When blood oxygen levels are low, a signal is sent to the kidneys to make erythropoietin (EPO), stimulating the bone marrow to create more red blood cells by using the body's iron.[3,4] Thus, requisite levels of both EPO and iron are needed to avoid anemia, which is reflected by a low red blood cell count. However, damaged kidneys do not produce enough EPO, hampering the erythropoiesis process.[5,6] Further, ESKD impacts iron levels on two fronts: by impairing dietary iron absorption, and by increasing iron losses

through frequent blood draws, gastrointestinal bleeds, and through the HD process itself, with total negative iron balance of about 3 grams per year.[7–9] Without enough EPO and/or iron to support erythropoiesis, many CKD patients suffer from anemia, which is defined as hemoglobin levels below 13 g/dL (in men) or 12 g/dL (in women).[5] Anemia can result in symptoms such as fatigue, weakness, and dizziness, and also lead to cardiovascular complications.[6]

The first-line treatment for anemia of CKD prior to the late 1980's was frequent blood transfusions,[10] which had side effects including iron overload, infections, and sensitization that could potentially impede transplantation.[3] Recombinant human EPO, an erythropoiesis stimulating agent (ESA), was introduced in 1989, revolutionizing anemia treatment in ESKD.[11] EPO production declines with severity of CKD, and most HD patients now require ESA treatment to promote erythropoiesis; only 2% of patients maintain hemoglobin > 12 g/dL without any ESA therapy over 4 months.[12] ESA's work by utilizing more of the body's iron to create red blood cells, thereby reducing iron levels.[13] Iron deficient HD patients do not respond well to ESA therapy due to the shortage of available iron for erythropoiesis, and thus iron supplementation is often prescribed in combination with an ESA. By 1998, 60% of US HD patients were receiving intravenous (IV) iron, up from <1% in 1992.[14] Iron deficiency can be described as absolute or functional, and is most often quantified by two markers: serum ferritin, a measure of iron stores, and transferrin saturation (TSAT), a measure of available iron. Low levels of both ferritin and TSAT reflect an absolute iron deficiency, while a low TSAT plus high ferritin reflect a functional iron deficiency.[4,9,10,15] In both cases, the iron supply to the bone marrow is inadequate, obstructing the process of erythropoiesis.

Changes in anemia management strategies

Early studies showed that ESA administration was effective in raising hemoglobin levels[11] and avoiding blood transfusions.[16] Uptake of the drug was swift and by 1992, 70% of HD patients in the US were receiving an ESA; by 2002, the proportion was 90%.[14] As a result, transfusions decreased by more than 50% from 1992 to 2005.[14] However, concerns with ESA toxicity began in 1998 with the Normal Hematocrit Cardiac Trial (NHCT), which showed higher mortality in HD patients who were ESA-treated to higher hemoglobin targets (14 vs. 10 g/dL).[17] Additional randomized trial results published in 2006 also failed to demonstrate a benefit of higher hemoglobin targets in non-dialysis CKD: the Cardiovascular Risk Reduction by Early Anemia Treatment with Epoetin Beta (CREATE) trial showed no difference in risk of cardiovascular events or CKD progression in patients treated to a hemoglobin target of 13.0-15.0 vs. 10.5-11.5 g/dL,[18] while the Correction of Hemoglobin and Outcomes in Renal Insufficiency (CHOIR) trial showed an increased risk of death in CKD patients treated to a higher hemoglobin target (13.5 vs. 11.3 g/dL).[19] In 2009, another randomized study in non-dialysis CKD, the Trial to Reduce Cardiovascular Events with Aranesp Therapy (TREAT), showed an increased risk of stroke for patients treated to a hemoglobin target of 13 g/dL vs. placebo with rescue therapy at 9 g/dL.[20]

In the US, there was a strong regulatory response to this growing evidence base pointing to the potential harm of a more complete correction of hemoglobin to levels considered normal in the general population. Starting in 2011, a new “bundled” payment system for dialysis patients was implemented by the Centers for Medicare and Medicaid Services (CMS) so that medications administered intravenously at the dialysis facility –

including ESA and iron – were no longer separately billable.[21] Later in 2011, the Food and Drug Administration (FDA) modified the ESA label, replacing the hemoglobin target of 10-12 g/dL with a recommendation to use the lowest ESA dose needed to avoid transfusions.[22] The Quality Incentive Program (QIP) introduced with the bundle that rewarded dialysis facilities based on performance metrics later switched from penalizing facilities with an excessive fraction of patients with hemoglobin too low (<10 g/dL) to penalizing facilities with too many patients with hemoglobin too high (>12 g/dL).[23]

The potential impact of these regulatory changes – particularly the bundled payment system – in the US garnered great interest in the nephrology community because dialysis facilities were provided a financial incentive to rely less on expensive ESA therapy to support hemoglobin levels.[24,25] The result was an immediate and dramatic decrease in ESA dosing.[26–29] Predictably, mean hemoglobin levels also decreased due to less ESA support,[26–29] though hemoglobin may not be a valid surrogate for clinical outcomes.[24,30,31] Two studies that analyzed the impact of the bundled payment system on clinical outcomes both found that trends in all-cause and cardiovascular-related mortality rates were unchanged, while stroke rates decreased.[26,28] An increase in transfusions rates, as might be expected with lower hemoglobin levels, was also observed by both studies,[26,28] although the bundled payment system coincided with an increase in the number of fields available on each Medicare claim, creating uncertainty in the interpretation of trends in transfusion rates.[32]

A continuum of strategies for treating anemia in HD patients are available: higher doses of ESA with lower doses of iron, or lower doses of ESA with higher doses of iron.[33] IV iron is less expensive than ESA therapy and often results in reduced ESA dose

requirements;[34,35] thus an increase in IV iron dosing was observed following implementation of the bundled payment system in 2011.[36] TSAT levels were largely unchanged, but serum ferritin levels increased dramatically.[29] While the increase in IV iron dosing was transient, the ferritin increase was sustained; this pattern was partially explained by the lower ESA dosing, resulting in patients utilizing less iron for erythropoiesis and thus more iron remaining in stores.[36] Increasing acceptance of higher ferritin and/or TSAT targets at many US centers may have also contributed to the increased prevalence of high ferritin levels. These historically high ferritin levels prompted concern that ESA toxicity was being replaced with iron toxicity, citing theoretical long-term safety issues with IV iron dosing.[37–39]

Anemia treatment options today

While serum ferritin was found to be the best marker of iron stores based on hepatic MRI in a study of dialysis patients excluding those with overt inflammation,[40] others have argued that ferritin has several disadvantages as an index of iron status and is inadequate for guiding iron repletion therapy.[10,41–43] Serum ferritin is elevated when patients are inflamed, leading to strong correlations with C-reactive protein (CRP) and other markers of acute illness.[44–47] Kalantar-Zadeh et al.[48] showed that while high ferritin levels were strongly associated with elevated mortality in a crude analysis, this association was almost eliminated after adjustment for patient characteristics and markers of malnutrition and inflammation. The utility of a single measurement of serum ferritin as a marker of iron stores may be further limited by extreme within-patient variability over time.[49,50] These issues complicate recommended guidelines that advise holding IV iron when ferritin levels reach a certain threshold.[5,51–53]

Randomized trials have consistently shown harmful effects of administering large doses of ESA to reach and maintain higher hemoglobin levels,[54] but controversy remains in identifying optimal strategies for iron supplementation in CKD and ESKD.[8,55–60] The Dialysis Patients' Response to IV Iron with Elevated Ferritin (DRIVE) trial demonstrated that IV iron effectively raised hemoglobin levels in patients with functional iron deficiency (TSAT < 25% and ferritin 500-1200 ng/ml).[61] The DRIVE II study, a 6 week observational extension of the DRIVE trial, showed that patients treated with IV iron required lower ESA doses to maintain hemoglobin levels.[34] Large cohort studies investigating the association between IV iron dosing and adverse events in HD patients have yielded mixed results.[48,62–64] Higher mortality risk with larger doses of IV iron was observed by Bailie et al.[62] (≥300 mg/month) and Kalantar-Zadeh et al.[48] (>400 mg/month). In contrast, no association between IV iron dose and all-cause mortality was observed by Miskulin et al.[63] and Feldman et al.[64] While the threshold for iron toxicity is unclear, Horl[33] notes that there has been no epidemic of iron overload in HD patients with many years of high ferritin levels.

The long-term safety of IV iron was recently assessed in the Proactive IV Iron Therapy in Haemodialysis Patients (PIVOTAL) study, a large, open-label, UK-based randomized controlled trial.[65] In PIVOTAL, the IV iron dose assigned each month depended on the most recent values of ferritin and TSAT. In the proactive high-dose arm, 400 mg IV iron was administered monthly unless upper thresholds of ferritin (>700 ng/mL) or TSAT (>40%) were reached. In the reactive low-dose arm, lower IV iron doses (100 or 200 mg/month) were administered unless lower thresholds of ferritin (<200 ng/mL) or TSAT (<20%) were reached. Comparing the high vs. low dose arms, the authors observed a

hazard ratio (HR) of 0.85 (95% CI: 0.73, 1.00) for the composite primary end point of nonfatal myocardial infarction, nonfatal stroke, hospitalization for heart failure, or death over the 42-month follow-up period, and concluded that a proactive high-dose IV iron treatment regime was superior to a reactive low-dose regime.[65]

Despite potential warning signs of harm due to high ESA doses, some patients still receive very high doses due to a lack of hemoglobin response to lower doses. This hyporesponsiveness to ESA affects about 10% of HD patients[66] and is commonly defined as one of (1) a decrease in hemoglobin level at constant ESA dose; (2) an increase in ESA dose to preserve a similar hemoglobin level; or (3) a failure to raise hemoglobin into target range despite large ESA doses.[67] Inflammation, most often measured in population studies by high CRP, is a strong predictor of mortality.[47] CRP elevations can be transient in many HD patients,[68] but high levels could be sustained in cases of chronic inflammation. CRP level is also associated with morbidity and acute illnesses, which may worsen response to ESA therapy.[69]

Anemia management guidelines

Formal clinical practice guidelines for anemia management in ESKD are released every few years by a variety of organizations: Kidney Disease Improving Global Outcomes (KDIGO) worldwide, with follow-up commentary from Kidney Disease Outcomes Quality Initiative (KDOQI) in the US; European Renal Best Practice (ERBP) in Europe; National Institute for Health and Social Care (NICE) in the UK; Caring for Australasians with Renal Impairment (CARI) in Australia; and Japanese Society for Dialysis Therapy (JSDT) in Japan. Because it is well-established that targeting high hemoglobin levels with large ESA doses can lead to serious adverse events,[17,19,20,54] guidelines relating to

ESA doses and hemoglobin levels are in general agreement, with hemoglobin targets settling in the 10-11 g/dL range.[5,51–53] For patients who start HD with very low hemoglobin levels, often due to inadequate treatment in CKD stages 4 and 5, ESA dosing to avoid transfusions is recommended.[22] In contrast to ESA therapy, iron supplementation strategies remain controversial.[8,55–60]

There is great regional heterogeneity in upper targets for ferritin levels at which discontinuing IV iron is advised. Ferritin levels are recommended to be kept very low in Japan (median 73 ng/mL[70]); in Europe, values are higher,[71,72] with guidelines recommending to discontinue IV iron at an upper ferritin target of 500 ng/mL;[52] in the US, the 2013 KDOQI commentary[51] on the 2012 KDIGO guidelines[5] expressed a comfort with IV iron dosing up to a ferritin level of 800 ng/mL. This threshold may have been regularly exceeded in practice at many facilities, resulting in a historically high median ferritin of 810 ng/mL as of June 2017.[29]

Rationale

The three aims of this dissertation address current controversies in anemia management practices for HD patients using innovative methods in observational data to minimize confounding. The regular schedule of maintenance HD (typically three times per week) provides a great opportunity to explore large databases of longitudinal information on patient condition, treatments, and laboratory measurements. To address these specific aims, data from the Dialysis Outcomes and Practice Patterns Study (DOPPS), a multiphase prospective cohort study of center-based, adult chronic hemodialysis patients in >20 countries ongoing since 1996, will be leveraged.

Aim 1 is to determine whether it is beneficial to manage anemia in advanced CKD (before starting HD). A substantial proportion of patients initiate HD with severe anemia[73] and the most likely causes include (1) lack of any pre-ESKD nephrologist care, (2) lack of adequate anemia treatment despite nephrologist care, (3) lack of responsiveness to anemia treatment, and (4) generally poor health or acute illness. The optimal strategy to treat these patients is unknown, as reflected by the variation in practice during the first few months of HD, especially across global regions.[74] Some nephrologists may choose to administer very high doses of ESA and/or large bolus doses of IV iron to quickly achieve hemoglobin increases into target range, but these rapid hemoglobin increases from ESA therapy may be potentially harmful.[75] Others may instead provide moderate amounts of ESA and/or IV iron, achieving hemoglobin increases more slowly to avoid adverse events that have been linked with higher hemoglobin targets.[17,19,20] While ESA and IV iron doses among long-term HD patients have likely been titrated based on individual patient responsiveness to treatment, doses among new HD patients may be more likely driven by nephrologist practice patterns and regional guidelines, creating an opportunity to leverage these discretionary practices analytically.[5,51–53] Evaluation of these anemia treatment strategies during the early months of HD is needed, specifically whether a more aggressive approach may lead to better short-term surrogate outcomes (i.e., swift increase of hemoglobin) but worse clinical outcomes due to large, and potentially harmful, doses of ESA and/or IV iron. Further, while patients who initiate HD with very low hemoglobin levels have higher mortality rates on HD, it's unknown whether this disadvantage is sustained even among patients whose hemoglobin was corrected into target range within 4 months on HD.

Using a seemingly counterintuitive approach by restricting on patients who were able to achieve hemoglobin ≥10 g/dL after 4 months on HD helps avoid bias by excluding cases in which hemoglobin was low at HD initiation due to lack of responsiveness to anemia treatment or generally poor health, and thus attempts to isolate the reason related to the question of interest – lack of anemia treatment prior to ESKD.

Aim 2 is to find out whether, and to what extent, new inflammation impairs responsiveness to ESA therapy in HD patients. Cross-sectional analyses have consistently shown that patients with higher CRP have greater ESA resistance,[76–78] but prospective studies have been less frequent and yielded mixed results using a variety of analytic approaches.[79–82] Financial incentives and clinical risk mitigation strategies to reduce ESA doses motivate the need to distinguish between HD patients who require higher ESA doses to achieve hemoglobin ≥10 g/dL and those who are over-treated. A better understanding of how ESA dose requirements may change in response to inflammation occurrences could lead to more proactive changes in ESA utilization. A self-matched longitudinal design to quantify the magnitude of within-patient changes in hemoglobin and ESA dose in patients who experienced a rise in CRP, relative to pre-inflammation levels, will help address this research question in a manner that avoids confounding due to fixed patient characteristics, e.g., sex, baseline age, and comorbidity history, as well as other unmeasured confounders such as genetic or environmental factors.

Aim 3 is to replicate key findings from a randomized trial of IV iron supplementation by applying the parametric g-formula to longitudinal data from a large prospective cohort study. The recently published PIVOTAL trial[65] showed that a proactive high-dose (vs.

reactive low-dose) IV iron treatment regime resulted in lower ESA doses and fewer adverse events. While randomized controlled trials are the gold standard for causal inference, they are generally inflexible to different selection criteria and intervention protocols. One practical alternative is to apply the parametric g-formula, an extension of standardization to longitudinal data, to compare complex and dynamic treatment strategies using observational data.[83] Cohort studies of HD patients are particularly suitable for g-formula analyses that depend on rich longitudinal data. While the ultimate goal is to assess a broad spectrum of IV iron treatment strategies using this method, an important first step is to mimic the PIVOTAL trial using real-world data. This proof-of-concept could then be extended to treatment strategies that remain unaddressed by randomized trials, which are generally designed to test only one or two specific hypotheses. If trial results can be replicated in this g-formula simulation, the potential to evaluate many variations of complex intervention strategies across different populations could prove to be enormously informative in the age of big data.

This dissertation outlines critical gaps in the literature on anemia management in HD patients, and stresses the importance of clearly defining the research question and choosing the appropriate study design. The three unique aims utilize novel applications of complex statistical analyses to contribute meaningfully to the literature, and may ultimately help optimize anemia management strategies in HD patients.

CHAPTER II

Low Hemoglobin at Hemodialysis Initiation: An International Study of Anemia Management and Mortality in the Early Dialysis Period

Coauthors who contributed to this work include: Hal Morgenstern, Sandra Waechter, Nancy L. Fleischer, Raymond Vanholder, Stefan H. Jacobson, Manish M. Sood, Douglas E. Schaubel, Masaaki Inaba, Ronald L Pisoni, and Bruce M. Robinson.

INTRODUCTION

Most chronic kidney disease (CKD) patients suffer from anemia due to deficiencies in iron and erythropoietin, often resulting in fatigue, weakness, and an increased risk of cardiovascular complications.[5,84] A substantial proportion of CKD patients reach end-stage kidney disease (ESKD) with very low hemoglobin levels.[73] The most likely causes include: (1) uremic intoxication inhibiting erythropoiesis; (2) lack of pre-ESKD nephrology care; (3) lack of adequate anemia treatment, despite nephrologist care, including insufficient correction of iron deficiency, as illustrated by low levels of serum ferritin or transferrin saturation (TSAT); or (4) lack of responsiveness to anemia treatment, often due to poor general health or acute illness. Regular treatments with intravenous (IV) iron and erythropoiesis-stimulating agents (ESA) are standard for thrice-weekly in-center hemodialysis (HD) patients, but are more difficult to carry out for non-dialysis CKD patients who receive care intermittently.

ESA and IV iron therapy are effective in raising hemoglobin levels and avoiding blood transfusions.[11,16,61,85] However, concerns regarding ESA toxicity emerged following a number of randomized clinical trials (RCT).[17–20] Studies have consistently demonstrated harmful effects of administering large doses of ESA to reach and maintain higher hemoglobin levels,[54] resulting in clinically acceptable hemoglobin targets that are now generally in the range of 10-12 g/dL.[5,51–53,86] The practice of starting anemia therapy at 9.5-10.0 g/dL, based on results from the Time to Reconsider Evidence for Anemia Treatment (TREAT) trial (glomerular filtration rate [GFR] 20-60),[20] may not be directly applicable to the dialysis transition period. Further, two recent studies showed conflicting results regarding whether there may be a benefit[87] or no benefit[88] to starting dialysis therapy with a higher hemoglobin level. Controversy also remains in identifying optimal strategies for iron supplementation in non-dialysis CKD and ESKD.[55,57,58,60,89] In the non-dialysis CKD setting, the efficacy of IV iron to raise or sustain hemoglobin levels has been well-established, but most RCTs do not have sufficient follow-up to evaluate long-term safety.[90] In the HD setting, large cohort studies of IV iron and mortality have yielded mixed results.[48,62–64,91,92] While the recently published Proactive IV Iron Therapy in Haemodialysis Patients (PIVOTAL) RCT[65] demonstrated non-inferiority of a proactive high IV iron dose strategy, results may not be generalizable to all Dialysis Outcomes and Practice Patterns (DOPPS) countries – especially not the US, where average iron doses and ferritin levels are greater than even the high IV iron dose PIVOTAL arm.[29]

ESA and IV iron doses among long-term HD patients have likely been titrated based on individual patient responsiveness to treatment, but doses in the early HD period may more likely be driven by nephrologist practice patterns and regional guidelines.[5,51–53]

Some nephrologists may choose to administer very high doses of ESA and/or large bolus doses of IV iron to quickly achieve hemoglobin increases and avoid the symptoms and potential consequences of severe anemia. However, ESA labels warn that rapid hemoglobin increases may increase risk of adverse cardiovascular events.[75,93] It is also possible that patients who start HD with low hemoglobin experience worse outcomes on HD irrespective of HD treatment strategy.

We hypothesize that management of anemia before the start of dialysis improves survival after HD start by avoiding the potential harms of chronic anemia, high doses of ESA and IV iron in the early months of HD, and/or a rapid hemoglobin rise. To test this hypothesis, we will investigate: (1) the association between anemia at HD start and all-cause mortality through 1 year of HD; and (2) mortality rates for different doses of IV iron and ESA during the early dialysis period.

METHODS

Data Source

The DOPPS is a large, international prospective cohort study of patients age ≥18 years treated with in-center HD in 21 countries. Maintenance HD patients were randomly selected from national samples of dialysis facilities in each country; detailed information is included in prior publications[94,95] and at http://www.dopps.org. Study approval and patient consent were obtained as required by national and local ethics committee regulations. This analysis included DOPPS phase 4 (2009-2011) and phase 5 (2012-2015) patients who enrolled in the DOPPS within 30 days after initiating HD.

Variables

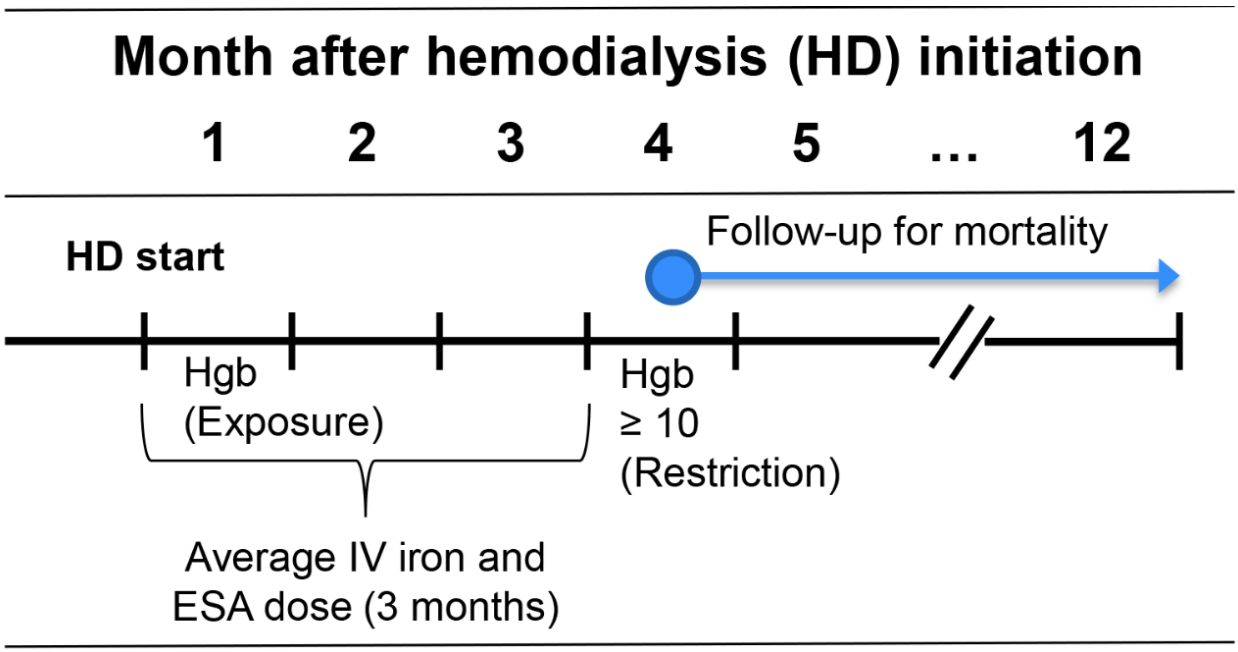
Information on age, sex, post-dialysis weight, body mass index (BMI), comorbid conditions (diabetes, hypertension, congestive heart failure [CHF], peripheral vascular disease [PVD], cancer), and catheter use were abstracted from medical records at DOPPS enrollment. Monthly data on medication prescriptions (ESA and IV iron) were also abstracted from medical records. To convert ESA doses to units of IV epoetin, we used conversion factors of 250:1 for darbepoetin,[96] 208:1 (250/1.2) for pegylated epoetin-β,[97] and 1.15:1 for subcutaneous injections.[98] Laboratory values (hemoglobin, TSAT, ferritin, albumin, phosphorus) are described as "month 1" when measured 0-30 days after starting HD (first measurement recorded in DOPPS) and "month 4" when measured 91-120 days after starting HD.

Study Design

In our primary analysis, we estimated the effect of month 1 hemoglobin on all-cause mortality from month 4 through month 12. We restricted to patients with hemoglobin ≥10.0 g/dL in month 4 to facilitate a comparison of mortality rates between patients with similar hemoglobin levels 4 months after HD initiation (**Figure 1**). This allows us to investigate whether the hypothesized harm of starting HD with low (vs. higher) hemoglobin is sustained even after a rapid hemoglobin increase into target range. While hemoglobin targets, specifically upper targets, are controversial[86] and have varied over time and by region,[5,51–53] we chose a restriction of ≥10 g/dL at month 4. Patients with low hemoglobin at HD start due to severe comorbidity or a lack of responsiveness to pre-ESKD anemia therapy are less likely to experience an hemoglobin increase to ≥10 g/dL in the subsequent months; those with hemoglobin ≥10 g/dL after 4 months on HD would have theoretically been able to start HD with a higher hemoglobin, with active

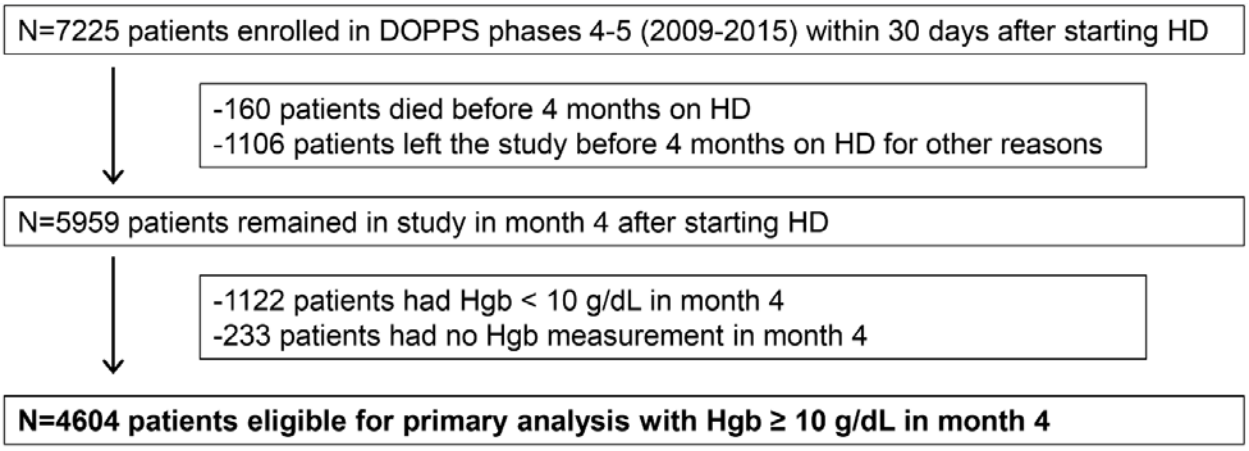
anemia management prior to starting HD. By excluding patients poorly responsive to anemia treatment after dialysis start, our study is thus designed to include an exposure variable that should mostly reflect differences in pre-ESKD anemia care. Among 5726 patients with hemoglobin measured in month 4, we excluded 1122 (20%) with hemoglobin <10 g/dL. The remaining 4604 patients were eligible for the primary analysis (**Figure 2**); 447 from DOPPS phase 4 and 4157 from DOPPS phase 5, when more emphasis was placed on enrolling incident HD patients.

Figure 1. Illustration of study design and timing of data collection



Hemoglobin (Hgb) values in month 1 were measured 0-30 days after starting HD. Hgb values in month 4 were measured 91-120 days after starting HD. The primary exposure is month 1 Hgb. The primary analysis was restricted to 4604 patients with Hgb ≥10 g/dL in month 4, with mortality follow-up beginning after Hgb measurement in month 4, and ending after 12 months on HD. A secondary analysis of ESA and IV iron dose and mortality was restricted to 5959 patients with no Hgb restriction in month 4.

Figure 2. Flow diagram of patient selection with inclusion/exclusion criteria



Statistical Analysis

Cox regression was used to estimate the effect of month 1 hemoglobin on all-cause mortality, restricted to patients with month 4 hemoglobin ≥10.0 g/dL. Left-truncated Cox models were used, with time on dialysis as the time scale; time at risk started after the hemoglobin measurement in month 4 and ended at the time of death, 7 days after leaving the facility due to transfer or change in kidney replacement therapy modality, loss to follow-up, end of study phase, or 1 year after initiating HD (whichever event occurred first). Hemoglobin was categorized to explore the functional form of association with mortality. Cox models were stratified by DOPPS phase and country, and by dialysis chain affiliation in the United States (US). Within-facility clustering was accounted for by using robust sandwich covariance estimators. Covariates selected for adjustment included age, sex, BMI, history of five comorbidities (diabetes, hypertension, CHF, PVD, cancer), catheter use at study enrollment, serum albumin and phosphorus in month 1 and month 4, and hemoglobin in month 4.

We performed several sensitivity analyses to test the robustness of the findings. First, we additionally adjusted for ESA and IV iron dose over the first 3 months of HD to investigate the role of these potential mediators on causal pathways between low hemoglobin at HD initiation and mortality. Second, we used hemoglobin measured in the month prior to initiating HD as the exposure rather than month 1 hemoglobin. Third, we excluded patients dialyzing with a catheter at study enrollment as a proxy for minimal pre-ESKD nephrology care. Fourth, we additionally adjusted for an indicator of >1 month of pre-ESKD nephrology care in a subset of 29% of patients for whom the data were available. Fifth, we varied the month 4 hemoglobin restriction from ≥10.0 to ≥10.5 and ≥11.0 g/dL because some ESA hyporesponsive patients could be treated to hemoglobin ≥10.0 g/dL, but not readily to higher hemoglobin levels. Sixth, we excluded patients not receiving any ESA therapy during the first 3 months of HD, an indicator of likely endogenous erythropoietin production in most patients.

In a secondary analysis not restricted to month 4 hemoglobin ≥10 g/dL, we illustrated the distribution of ESA and IV iron dose over the first 3 months of HD therapy by month 1 hemoglobin, TSAT, and ferritin across three regions (Europe, Japan, US). Countries outside of these three regions were included in the mortality models but were not shown in these descriptive analyses due to small sample sizes. In this population, we also estimated the effects of average ESA and IV iron dose administered over the first 3 months of HD on all-cause mortality. Cox regression models were implemented as in the primary analysis, with additional adjustment for hemoglobin, TSAT, and ferritin in month 1, but without adjustment for month 4 hemoglobin, a potential mediator. We repeated this analysis in a subgroup of patients with hemoglobin <10.0 g/dL in month 1

to better inform optimal treatment practices for patients who initiated HD with low hemoglobin.

We used multiple imputation, assuming data were missing at random, to impute missing covariate values using the Sequential Regression Multiple Imputation Method by IVEware.[99] Results from 20 such imputed data sets were combined for the final analysis using Rubin's formula.[100] The proportion of missing data was below 10% for all covariates, with the exception of month 1 laboratory measures (ferritin 28%; TSAT 27%; albumin 16%; phosphorus 11%) and the five comorbidities (10%-21%). All analyses were conducted using SAS software, version 9.4 (SAS institute, Cary, NC).

RESULTS

Patient Characteristics

Our primary analysis included 4604 patients with hemoglobin ≥10.0 g/dL in month 4 after starting HD. Among these patients, 53% had hemoglobin <10.0 g/dL in month 1, including 6% with hemoglobin <8.0 g/dL, and 20% had hemoglobin ≥11.0 g/dL in month 1, including 7% with hemoglobin ≥12.0 g/dL and 2% ≥13.0 g/dL. The mean hemoglobin in month 4 was between 11.3 and 11.6 g/dL across the five groups of month 1 hemoglobin (from <8.0 to ≥11.0 g/dL). Patients with lower hemoglobin in month 1 were younger, more likely to dialyze with a catheter, and had higher ESA doses over the subsequent 3 months, but only marginally higher IV iron doses (**Table 1**). Patients with lower hemoglobin in month 1 also had lower TSAT, higher serum ferritin, and lower serum albumin in month 1, but all of these differences were minimized by month 4. The prevalence of comorbidities varied minimally by month 1 hemoglobin.

Table 1. Patient characteristics by hemoglobin in month 1 after starting HD, restricted to patients with hemoglobin ≥10.0 g/dL in month 4 after starting HD

	All	Hemoglobin (g/dL) in month 1 after HD start				
Patient characteristic		**< 8.0**	**8.0-8.9**	**9.0-9.9**	**10.0-10.9**	**≥ 11.0**
N patients	4604	283 (6%)	822 (18%)	1260 (28%)	1209 (27%)	897 (20%)
Baseline characteristics and treatments						
Age (years)	64.1 ± 14.5	59.2 ± 16.1	62.7 ± 14.9	64.4 ± 14.7	65.1 ± 14.2	64.9 ± 13.7
Sex (% male)	60%	56%	60%	57%	60%	66%
Post-dialysis weight (kg)	80.4 ± 22.8	75.9 ± 22.9	78.3 ± 22.1	80.7 ± 23.3	80.6 ± 21.3	81.8 ± 23.5
BMI (kg/m^2)	28.1 ± 6.9	26.8 ± 6.3	27.2 ± 6.5	28.4 ± 7.2	28.3 ± 6.6	28.4 ± 7.0
Catheter use (%)	59%	65%	63%	60%	59%	51%
Hemodiafiltration (%)	3%	2%	3%	3%	3%	4%
Single pool Kt/V	1.33 ± 0.35	1.28 ± 0.37	1.33 ± 0.37	1.35 ± 0.35	1.33 ± 0.35	1.33 ± 0.35
Laboratory values < 30 days after starting HD						
Hemoglobin (g/dL)	9.9 ± 1.3	7.4 ± 0.5	8.5 ± 0.3	9.5 ± 0.3	10.4 ± 0.3	11.8 ± 0.8
TSAT (%)	20.5 ± 9.8	18.9 ± 9.3	19.4 ± 10.1	20.2 ± 9.5	20.4 ± 9.4	22.6 ± 10.2
Ferritin (ng/mL)	337 ± 334	402 ± 365	366 ± 368	336 ± 325	331 ± 333	305 ± 305
Serum albumin (g/dL)	3.4 ± 0.5	3.1 ± 0.5	3.3 ± 0.6	3.4 ± 0.5	3.5 ± 0.5	3.6 ± 0.5
Serum phosphorus (mg/dL)	4.7 ± 1.5	5.0 ± 2.0	4.7 ± 1.5	4.6 ± 1.4	4.6 ± 1.4	4.8 ± 1.5
Laboratory values 91-120 days after starting HD						
Hemoglobin (g/dL)	11.5 ± 1.0	11.3 ± 1.0	11.5 ± 1.0	11.4 ± 1.0	11.4 ± 0.9	11.6 ± 1.1
TSAT (%)	26.6 ± 11.7	26.6 ± 12.5	25.7 ± 12.4	26.4 ± 11.4	26.7 ± 11.4	28.0 ± 11.9
Ferritin (ng/mL)	428 ± 377	388 ± 361	402 ± 352	439 ± 398	454 ± 372	424 ± 398
Serum albumin (g/dL)	3.7 ± 0.4	3.6 ± 0.5	3.6 ± 0.5	3.6 ± 0.5	3.7 ± 0.4	3.7 ± 0.4
Serum phosphorus (mg/dL)	5.2 ± 1.5	5.5 ± 1.6	5.4 ± 1.7	5.2 ± 1.5	5.2 ± 1.4	4.9 ± 1.3
Anemia treatment during first 3 months of HD						
ESA use (%, any during 3 months)	92%	98%	98%	98%	91%	77%
ESA dose (1000 units/week)	12.7 ± 10.4	19.6 ± 13.6	16.7 ± 11.0	13.8 ± 10.1	10.8 ± 8.3	8.0 ± 8.9

IV iron use (%, any during 3 months)	86%	86%	87%	87%	86%	82%
IV iron dose (mg/month)	416 ± 292	445 ± 308	449 ± 301	428 ± 286	412 ± 281	364 ± 291
Comorbid conditions (%)						
Diabetes	62%	60%	61%	63%	63%	59%
Hypertension	87%	87%	88%	85%	88%	88%
CHF	27%	25%	25%	30%	27%	26%
Peripheral vascular disease	16%	16%	14%	18%	16%	17%
Cancer (non-skin)	11%	9%	12%	12%	12%	9%

Mean ± SD or % shown; mean ESA and IV iron doses averaged over 3 months and treat non-users as 0 dose; note columns do not sum to the total due to 133 (3%) patients missing month 1 Hgb data. Abbreviations: BMI, body mass index; CHF, congestive heart failure; ESA, erythropoiesis-stimulating agents; HD, hemodialysis; Hgb, hemoglobin; IV, intravenous; SD, standard deviation; TSAT, transferrin saturation.

Hemoglobin and Mortality

Among the 4604 patients, we observed 277 deaths and a mortality rate of 0.105/year during follow-up. Patients with lower month 1 hemoglobin – those who experienced a rapid hemoglobin increase to ≥10.0 g/dL during the subsequent 3 months – had higher mortality rates than patients who started HD with higher hemoglobin (**Table 2**). Compared with hemoglobin ≥11.0 g/dL, the adjusted hazard ratio (HR; 95% confidence interval [CI]) was 1.99 (1.18-3.38) for hemoglobin <8.0 g/dL and ranged from 1.18 to 1.35 for hemoglobin 8.0-10.9 g/dL (**Table 2, Model 3**). As a continuous variable, month 1 hemoglobin was inversely associated with mortality (adjusted HR for 1 g/dL higher hemoglobin = 0.89; 95% CI: 0.81-0.97; p for trend = 0.01). In sensitivity analyses, the HR for 1 g/dL higher hemoglobin was generally consistent when: (1) adjusting for potential mediators, ESA and IV iron dose over the subsequent 3 months (HR=0.89); (2) using hemoglobin measured in the month prior to initiating HD as the exposure (HR=0.90); (3) excluding catheter users (HR=0.84); (4) adjusting for >1 month of pre-ESKD nephrology care (HR=0.87); (5) varying the month 4 hemoglobin restriction of ≥10.0 to ≥10.5 (HR=0.87) and to ≥11.0 (HR=0.89) g/dL; and (6) excluding patients not treated with ESA during the first 3 months of HD (HR=0.89).

Table 2. HR of mortality for hemoglobin measured in month 1 after starting HD, by level of covariate adjustment, among patients with hemoglobin ≥10.0 g/dl in month 4 after starting HD

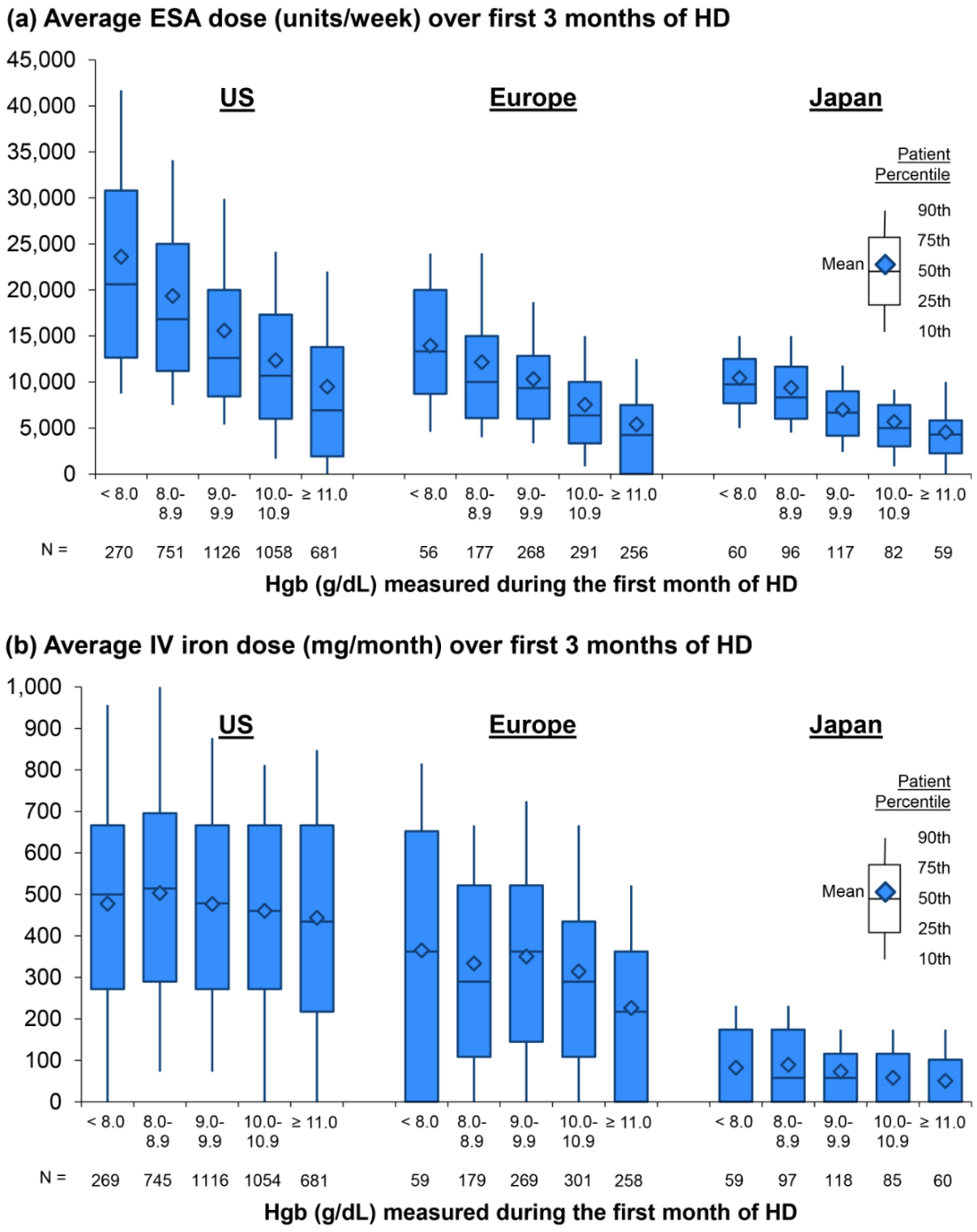
Exposure	N (%)	Model 1	Model 2	Model 3
Hemoglobin (g/dL) in month 1 after HD start, categories				
< 8.0	283 (6%)	1.81 (1.07-3.05)	2.10 (1.25-3.54)	1.99 (1.18-3.38)
8.0-8.9	822 (18%)	1.23 (0.81-1.86)	1.30 (0.87-1.95)	1.23 (0.83-1.84)
9.0-9.9	1260 (28%)	1.52 (1.06-2.18)	1.39 (0.96-2.01)	1.35 (0.93-1.95)
10.0-10.9	1209 (27%)	1.28 (0.89-1.85)	1.26 (0.86-1.84)	1.18 (0.81-1.73)
≥ 11.0	897 (20%)	1 (Ref.)	1 (Ref.)	1 (Ref.)
Hemoglobin (g/dL) in month 1 after HD start, continuous				
per 1 g/dL	--	0.91 (0.83-0.99)	0.88 (0.81-0.97)	0.89 (0.81-0.97)

HR (95% CI) of all-cause mortality in left-truncated Cox model (vintage time scale) from month 4 through month 12 of HD; all Cox models stratified by DOPPS phase, country, and US dialysis chain affiliation; Model 1: unadjusted; Model 2: adjusted for age, sex, BMI, 5 comorbidities (diabetes, hypertension, CHF, PVD, cancer) and catheter use; Model 3: further adjusted for serum albumin and phosphorus in month 1 and month 4, and hemoglobin in month 4. Abbreviations: BMI, body mass index; CHF, congestive heart failure; CI, confidence interval; DOPPS, Dialysis Outcomes and Practice Patterns Study; HD, hemodialysis; HR, hazard ratio; PVD, peripheral vascular disease.

Description of ESA and IV Iron Dosing

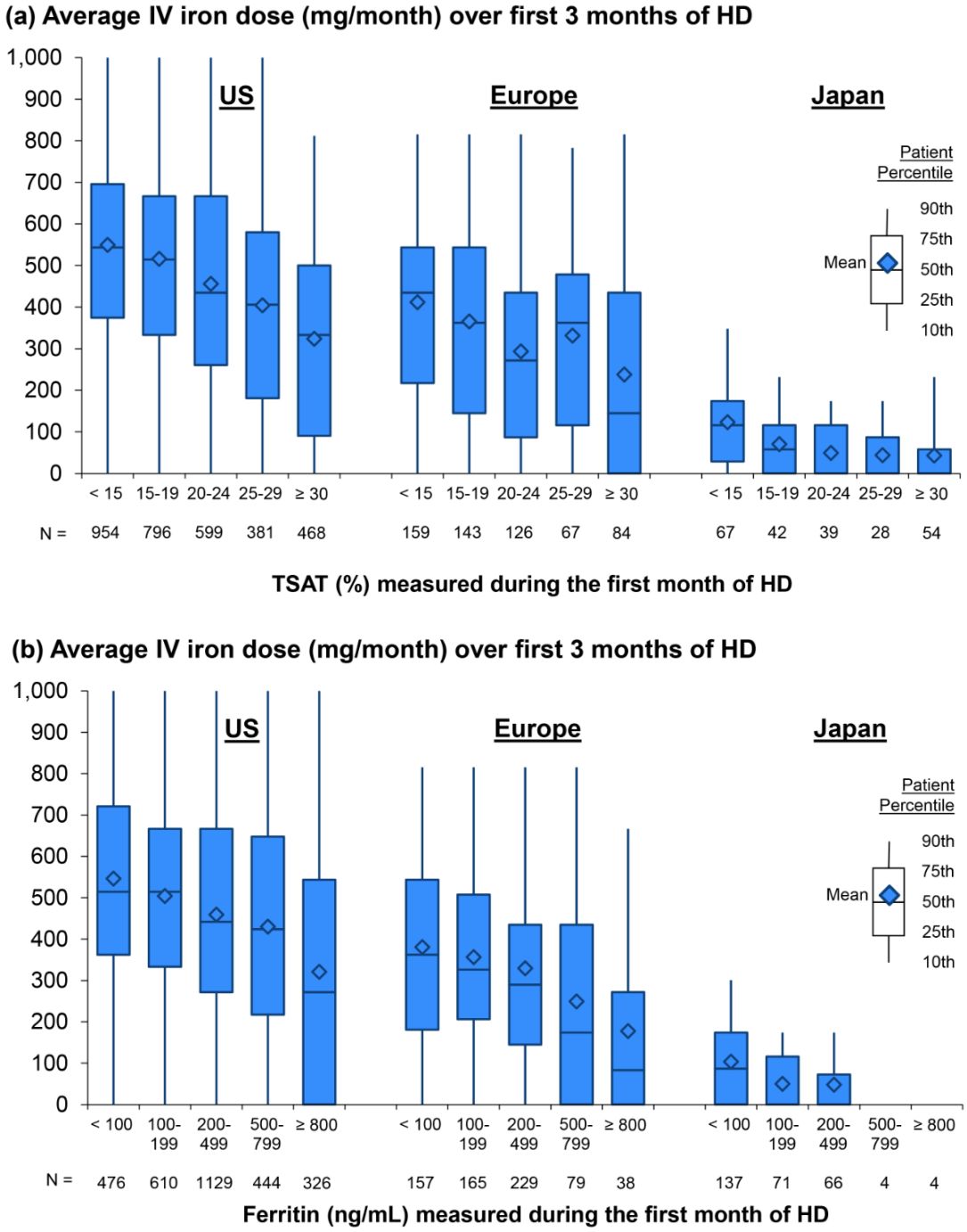
We also studied 5959 patients who survived to month 4 (to measure ESA and IV iron dose administered over the first 3 months of HD), but not restricting to patients with hemoglobin ≥10.0 g/dL in month 4, to illustrate the variation in dosing patterns. Patients with lower month 1 hemoglobin had greater ESA doses over the subsequent 3 months, reflecting indication for the treatment and/or pre-ESKD ESA hyporesponsiveness (**Figure 3a**). Patients with lower month 1 hemoglobin had greater IV iron doses over the subsequent 3 months in Europe, but not in the US, where a median dose of 450-500 mg/month was observed regardless of hemoglobin level (**Figure 3b**). IV iron doses in the first 3 months of HD were higher among patients initiating HD with lower TSAT and lower ferritin in all regions (**Figure 4**). However, relatively high IV iron doses were still observed among many patients with ferritin levels ≥800 ng/mL. Overall, ESA and IV iron doses in the first 3 months of HD were highest in the US, followed by Europe, and lowest in Japan across all strata.

Figure 3. Distribution of: (a) ESA dose; and (b) IV iron dose administered over the first 3 months of HD therapy, by region and hemoglobin during the first month of HD



In contrast to primary analysis, these data were not restricted to patients with Hgb ≥10 g/dL in month 4 after starting HD. Europe: Belgium, France, Germany, Italy, Spain, Sweden, United Kingdom. Data from other regions (Canada, Australia, New Zealand, China, Russia, Turkey, Gulf Cooperation Council) were suppressed in this figure due to small sample sizes, but were included in other analyses. Abbreviations: ESA, erythropoiesis-stimulating agents; HD, hemodialysis; Hgb, hemoglobin; IV, intravenous

Figure 4. Distribution of IV iron dose administered over the first 3 months of HD therapy, by region and (a) TSAT and (b) ferritin measured during the first month of HD



In contrast to primary analysis, these data were not restricted to patients with Hgb ≥10 g/dL in month 4 after starting HD. Europe: Belgium, France, Germany, Italy, Spain, Sweden, United Kingdom. Japanese data were suppressed in subgroups with ferritin ≥500 ng/mL due to insufficient sample size. Data from other regions (Canada, Australia, New Zealand, China, Russia, Turkey, Gulf Cooperation Council) were suppressed in this figure due to small sample sizes, but were included in other analyses. Abbreviations: HD, hemodialysis; Hgb, hemoglobin; IV, intravenous; TSAT, transferrin saturation.

Anemia Treatment and Mortality

Among this population of 5959 patients who survived to month 4, 92% were prescribed an ESA, and 83% were prescribed IV iron at some point during the first 3 months on HD. The associations of both ESA dose and IV iron dose over the first 3 months of HD on all-cause mortality from month 4 through month 12 are shown in **Table 3**. We observed elevated mortality at only very high ESA doses; the adjusted HR (95% CI) was 1.43 (1.02-2.01) for >25,000 (12% of patients) versus 5000-10,000 units/week. In a subgroup analysis of 3378 patients with month 1 hemoglobin <10.0 g/dL, the pattern was similar. The adjusted association between IV iron dose and mortality was U-shaped, with the lowest mortality rate observed for patients receiving 200-399 mg/month in both the whole sample and subgroup with month 1 hemoglobin <10.0 g/dL (**Table 3**).

Table 3. HR of mortality for ESA and IV iron dose over the first 3 months of HD without restricting to patients with hemoglobin ≥10 g/dL in month 4, (a) overall, and (b) among a subset of patients with hemoglobin <10.0 g/dL in month 1 of HD

	(a) All HD patients		(b) Restricted to subset with Hgb < 10.0 g/dL in first month of HD	
Exposure	**N (%)**	**Adjusted HR (95% CI)**	**N (%)**	**Adjusted HR (95% CI)**
Average ESA dose (units/week) over first 3 months on HD				
None	479 (8%)	0.94 (0.59-1.51)	103 (3%)	0.63 (0.22-1.82)
< 5K	711 (12%)	0.84 (0.54-1.29)	253 (8%)	0.56 (0.25-1.27)
5K-10K	1543 (26%)	1 (Ref.)	836 (25%)	1 (Ref.)
10K-15K	1203 (21%)	0.89 (0.64-1.24)	773 (23%)	0.85 (0.56-1.27)
15K-25K	1223 (21%)	0.97 (0.71-1.33)	818 (25%)	0.80 (0.53-1.20)
> 25K	699 (12%)	1.43 (1.02-2.01)	537 (16%)	1.54 (1.01-2.36)
Average IV iron dose (mg/month) over first 3 months on HD				
None	1008 (17%)	1.50 (1.05-2.15)	549 (17%)	1.74 (1.08-2.79)
< 200	741 (13%)	1.08 (0.72-1.62)	407 (12%)	1.25 (0.74-2.10)
200-399	1157 (20%)	1 (Ref.)	605 (18%)	1 (Ref.)
400-599	1359 (23%)	1.11 (0.78-1.57)	783 (24%)	1.39 (0.88-2.18)
≥ 600	1604 (27%)	1.24 (0.87-1.76)	975 (29%)	1.29 (0.80-2.07)

HR (95% CI) of all-cause mortality in left-truncated Cox model (vintage time scale) from month 4 through month 12 of HD; Cox models stratified by DOPPS phase, country, and US dialysis chain affiliation; Adjustments: age, sex, BMI, 5 comorbidities (diabetes, hypertension, CHF, PVD, cancer), catheter use, serum albumin and phosphorus in month 1 and month 4, Hgb, TSAT, and ferritin in month 1; Note this analysis was not restricted to patients with Hgb ≥10 g/dL in month 4. Abbreviations: BMI, body mass index; CHF, congestive heart failure; CI, confidence interval; DOPPS, Dialysis Outcomes and Practice Patterns Study; ESA, erythropoiesis-stimulating agents; HD, hemodialysis; Hgb, hemoglobin; HR, hazard ratio; IV, intravenous; PVD, peripheral vascular disease; TSAT, transferrin saturation.

DISCUSSION

Among patients with hemoglobin levels ≥10.0 g/dL in month 4 of HD, 53% of patients had hemoglobin <10.0 g/dL in month 1 after HD initiation, and lower hemoglobin in month 1 of HD was associated with a higher mortality rate in months 4-12. While we expected to observe a greater mortality rate for patients who initiated HD with low hemoglobin, this elevated rate was still observed among patients achieving hemoglobin ≥10.0 g/dL in the early HD period. In secondary analyses not restricted to hemoglobin ≥10.0 g/dL in month 4 of HD, we found a U-shaped association between IV iron dose over the first 3 months of HD and mortality, and elevated mortality for patients receiving the largest doses of ESA (>25,000 units/week).

Our primary result (**Table 2, Model 3**) is consistent with our hypothesis that management of anemia before the start of dialysis improves survival after HD start, although there are many possible explanations for these findings. One potential explanation is that intense anemia treatment and/or a rapid hemoglobin rise in the early HD period may be responsible for the elevated mortality rate. Patients who started HD with the lowest hemoglobin levels received the largest ESA doses over the first 3 months of HD. However, additional adjustment for ESA and IV iron doses (potential mediators) had minimal impact on the association between hemoglobin at HD initiation and mortality, thus making this explanation unlikely, though we did not conduct a formal mediation analysis.[101] Another possibility is that anemia treatment prior to ESKD has long-term benefits compared with initiating anemia therapy after HD start. Among patients with hemoglobin ≥10.0 g/dL in month 4 of HD, mean hemoglobin levels in month 4 were similar for patients who initiated HD with hemoglobin ≥11.0 (11.6 g/dL)

versus hemoglobin <8.0 (11.3 g/dL) g/dL. The subsequent mortality rate was twice as high for patients who started HD with hemoglobin <8.0 vs. ≥11.0 g/dL, despite also having a similar comorbidity profile (**Table 1**). This striking difference in mortality may point to a lingering effect of untreated anemia in CKD. We also found an 18%-35% greater mortality rate between hemoglobin 8.0-10.9 versus ≥11.0 g/dL, though surprisingly minimal difference within the range of 8.0-10.9 g/dL. It's possible that a "step function" exists, with low-, medium-, and high-risk groups, though our study is not powered to detect the exact hemoglobin cutpoints. In addition, a non-causal explanation is that the association may be biased due to residual confounding, possibly because patients with higher hemoglobin at HD start adhered more to prescribed treatments and recommendations and/or received better overall quality of care before HD initiation. We adjusted for catheter use at study enrollment as a proxy for this latter potential confounder, and also excluded catheter users in a sensitivity analysis.

The relation between anemia management during the transition period to ESKD and post-dialysis outcomes is challenging to assess and often requires an innovative study design. McCausland et al.[88] performed a post-hoc follow-up analysis of TREAT among the subset (only 15%) of randomized patients who reached ESKD.[20] Mean hemoglobin at dialysis start was higher in the darbepoetin intervention (11.3 g/dL) versus the control (9.5 g/dL) group, but no all-cause mortality benefit (HR=1.16; 95% CI: 0.69-1.93) was observed from HD start (when anemia therapy switched from protocol-driven to physician-driven) through 6 months; differences with our study may be explained by the study design and/or selection criteria. An increased stroke rate was observed in the intervention group, although paradoxically, it is likely ESA doses were much higher (by

indication) during the at-risk period in the control group. The authors concluded there is no apparent benefit from starting dialysis therapy with a higher hemoglobin level, but acknowledged other factors (e.g., avoiding transfusions) should help inform whether to treat mild-to-moderate anemia with ESA in CKD patients preparing for dialysis. A recent study by Wetmore et al.,[87] using United States Renal Data System (USRDS) data, in contrast to McCausland et al.,[88] drew similar conclusions as our study when analyzing US Medicare patients over 66 years old at dialysis initiation. The 1-year mortality rate was slightly lower (HR=0.88) among patients treated with ESA prior to dialysis start and who maintained hemoglobin ≥9.0 g/dL during the 3 months before and after dialysis start versus patients with hemoglobin <9.0 g/dL before dialysis start who then received ESA after HD start and experienced an hemoglobin increase to ≥9.0 g/dL. It is encouraging that the current DOPPS study achieved similar findings as the USRDS study while using a different data source (international sample with no age restriction) and a different analytical approach to address the question of interest.

Results from large observational studies[102,103] are consistent with our finding that patients receiving the highest ESA doses in the early HD period had worse survival. Following patients soon after HD initiation may better capture dosing patterns (practice preferences) before patients transition to a more individualized steady-state dosing protocol, at which point observed associations between ESA dose and mortality are more likely attributable to confounding factors that drive ESA dose requirements than a causal effect.[104,105]

While our main focus was on hemoglobin levels, we also investigated IV iron dosing. IV iron dose over the first 3 months of HD was generally greater in patients who initiated

HD with low levels of ferritin and TSAT, as expected. Patients starting HD with lower (vs. higher) hemoglobin received more IV iron in the next 3 months, but this pattern was not observed in the US, where a high median dose of 450-500 mg/month was observed regardless of hemoglobin at HD start. This may reflect the practice of bolus IV iron dosing in many US facilities, where doses of at least 100 mg are administered in consecutive HD sessions.[106,107] Michels et al.[108] found that a low-dose (vs. bolus dose) IV iron strategy in the early HD period may be beneficial by reducing ESA doses and risk of mortality. While we did not assess specific dosing patterns, we similarly observed a greater mortality rate at high doses of IV iron that would generally characterize a bolus dosing strategy, with the lowest mortality rate observed at 200-399 mg/month. Our results were also consistent with Kuo et al.,[109] who demonstrated that a low-dose IV iron strategy was optimal for incident HD patients. The recent PIVOTAL trial randomized patients to a proactive high IV iron dose (400 mg/month IV iron; discontinue if ferritin >700 ng/mL or TSAT >40%) or reactive low IV iron dose (100-400 mg/month IV iron; discontinue if ferritin >200 ng/mL and TSAT >20%) and demonstrated non-inferiority of the high IV iron dose strategy;[65] however, these “high” doses were lower than the median dose of 450-500 mg/month we observed in the US, and the upper ferritin threshold of 700 ng/mL in the “proactive” arm was lower than the median serum ferritin levels observed in the US.[29] Generalizability of these findings in the context of the high levels of serum ferritin and IV iron dosing observed in the US remains an open question.

This analysis had some limitations. First, to address potential residual confounding due to better pre-ESKD care (unrelated to anemia), we adjusted for catheter use at study enrollment, a proxy for lack of pre-ESKD nephrology care. Regarding potential residual

confounding due to health status, it is reassuring that the distribution of key risk factors (e.g., serum albumin at month 4) and prevalence of comorbidities was similar by hemoglobin level at HD initiation (**Table 1**). Second, we were unable to determine history of anemia therapy – including usage or dosage – prior to HD start, which would have helped inform the causes(s) of initiating HD with low hemoglobin (e.g., lack or treatment or hyporesponsiveness); this limitation prompted us to restrict to patients with hemoglobin ≥10 g/dL in month 4, excluding patients whose low month 1 hemoglobin was likely due to poor general health or ESA hyporesponsiveness, thus using low hemoglobin at HD start as a proxy for lack of pre-ESKD anemia treatment. Third, data on potentially important variables were not available for a majority of patients, including transfusions, C-reactive protein (not measured in North America), and residual kidney function. Finally, our exposure variable was hemoglobin measured within 30 days after initiating HD (median=13 days), but hemoglobin could have changed in the days immediately after starting HD. We thus performed a sensitivity analysis in a subset of patients for which data was available on hemoglobin levels immediately prior to starting HD.

This analysis also had some key strengths. We considered mortality as the primary outcome rather than a surrogate outcome, such as change in hemoglobin, which is often used in smaller RCTs with short follow-up to demonstrate effectiveness, but not safety, of pre-ESKD anemia therapies. Further, the large international DOPPS sample reflected a wide range of anemia management practices in the early HD period. In particular, we observed that IV iron prescription patterns did not appear to be driven by

hemoglobin levels in the US, where doses were very high in the first 3 months of HD (median 450-500 mg/month).

In this study, we found that, even among patients who achieved hemoglobin ≥10.0 g/dL within 4 months of starting HD, low hemoglobin at HD initiation was common (53% below 10.0 g/dL) and was associated with elevated mortality. Compared with hemoglobin ≥11.0 g/dL at HD start, we observed a two-fold greater mortality rate for hemoglobin <8.0 g/dL and an 18%-35% greater mortality rate for hemoglobin 8.0-10.9 g/dL. A more proactive approach to anemia management in advanced CKD may thus improve first-year survival on HD, though long-term prospective studies examining anemia treatments and adverse events starting in the non-dialysis CKD setting are needed.

CHAPTER III

The Effect of New Inflammation on Hyporesponsiveness to Erythropoiesis-Stimulating Agent Therapy in Hemodialysis Patients: A Self-Matched Longitudinal Study of Anemia Management in the DOPPS

Coauthors who contributed to this work include: Hal Morgenstern, Nancy L Fleischer, Raymond C Vanholder, Nafeesa N Dhalwani, Elke Schaeffner, Douglas E Schaubel, Tadao Akizawa, Glen James, Marvin V Sinsakul, Ronald L Pisoni, and Bruce M Robinson.

INTRODUCTION

Best practice guidelines for anemia management in hemodialysis (HD) patients, including use of erythropoiesis-stimulating agents (ESA), have varied over time and by international region, but physicians have now generally agreed on a lower hemoglobin target of 10 g/dL.[5,51–53,67,86] Randomized trials have demonstrated cardiovascular harm of targeting higher hemoglobin levels in anemic chronic kidney disease (CKD) patients,[17–20] but the mechanism remains unclear. Financial incentives and clinical risk mitigation strategies to reduce ESA doses also motivate the need to distinguish between HD patients who may require higher ESA doses to achieve hemoglobin ≥10

g/dL and those who may be over-treated. ESA hyporesponsiveness is thought to be present in about 10% of HD patients[110,111] and is commonly defined as one of the following: (1) a decrease in hemoglobin level at constant ESA dose; (2) an increase in ESA dose to preserve a similar hemoglobin level; or (3) a failure to raise hemoglobin into the target range despite large ESA doses (2).

Inflammation, easily identified clinically by a high C-reactive protein (CRP) level, is common in HD patients and associated with increased mortality.[47,69,112,113] Inflammation may also blunt the haematopoetic response of ESA therapy to produce hemoglobin by decreasing bone marrow response to ESA, altering iron regulation through hepcidin, and/or by causing hemolysis of red cells/erythrocytes.[69,110,114,115] It is possible that inflamed HD patients could benefit from proactive adjustment of anemia medications or, in the future, from anemia therapies such as hypoxia-inducible factor prolyl hydroxylase inhibitors (HIF-PHI) that may be effective despite inflammation.[116–118]

Several cross-sectional analyses have shown a positive correlation between CRP and ESA dose, but may be confounded by patient health status.[66,76–78,119] Longitudinal studies have been less frequent and have yielded mixed results using a variety of analytic approaches.[79–82] In this study, we focus on newly developed inflammation and aim to quantify the magnitude of within-patient changes in hemoglobin and ESA dose relative to pre-inflammation levels. We hypothesized that patients are more likely to be ESA hyporesponsive, with lower hemoglobin levels and/or larger ESA doses in the 3 months following an increase in CRP (from ≤5 to >10 mg/L), compared to the 3 months preceding this increase.

METHODS

Data source

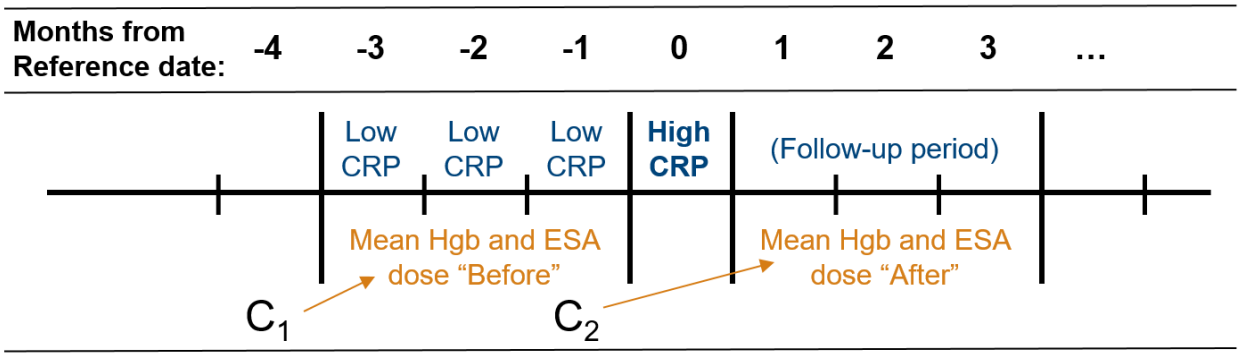
The Dialysis Outcomes and Practice Patterns Study (DOPPS) is an international, multiphase prospective cohort study of patients ≥18 years old treated with in-center HD in 21 countries. Maintenance HD patients were randomly selected from national samples of HD facilities in each country; detailed information is included in prior publications[94,95] and at http://www.dopps.org. Study approval and patient consent were obtained as required by national and local ethics committee regulations. Information on patient demographics and comorbidity history was abstracted from medical records at DOPPS enrollment in each phase. Monthly data on measured laboratory values and medication prescriptions were abstracted from medical records at baseline and during follow-up.

This analysis included HD patients from ten DOPPS countries where monthly CRP data were widely available: Japan, Australia, and New Zealand (ANZ), and seven countries in Europe: Belgium, France, Germany, Italy, Spain, Sweden, and the UK. No data from the US or Canada were used because routine measurement of CRP in the HD setting remains rare in North America.[120] Countries were included only during phases when data on laboratory values and medications were collected monthly: all ten countries in DOPPS phase 4 (2009-2011), all countries except Belgium and Sweden in phase 5 (2012-2015), and only Japan in phase 6 (2015-2018).

Study design

Our goal was to assess whether newly developed inflammation led to increased ESA resistance. To operationalize this hypothesis, we used a self-matched before-after study design as illustrated in **Figure 5** to assess within-patient changes in hemoglobin, ESA dose, and ESA hyporesponsiveness from the "before" period (little or no inflammation) to the "after" period (following the onset of inflammation). This self-matched design prevents confounding due to fixed patient characteristics, e.g., sex, baseline age, and comorbidity history, as well as other unmeasured confounders such as genetic or environmental factors.

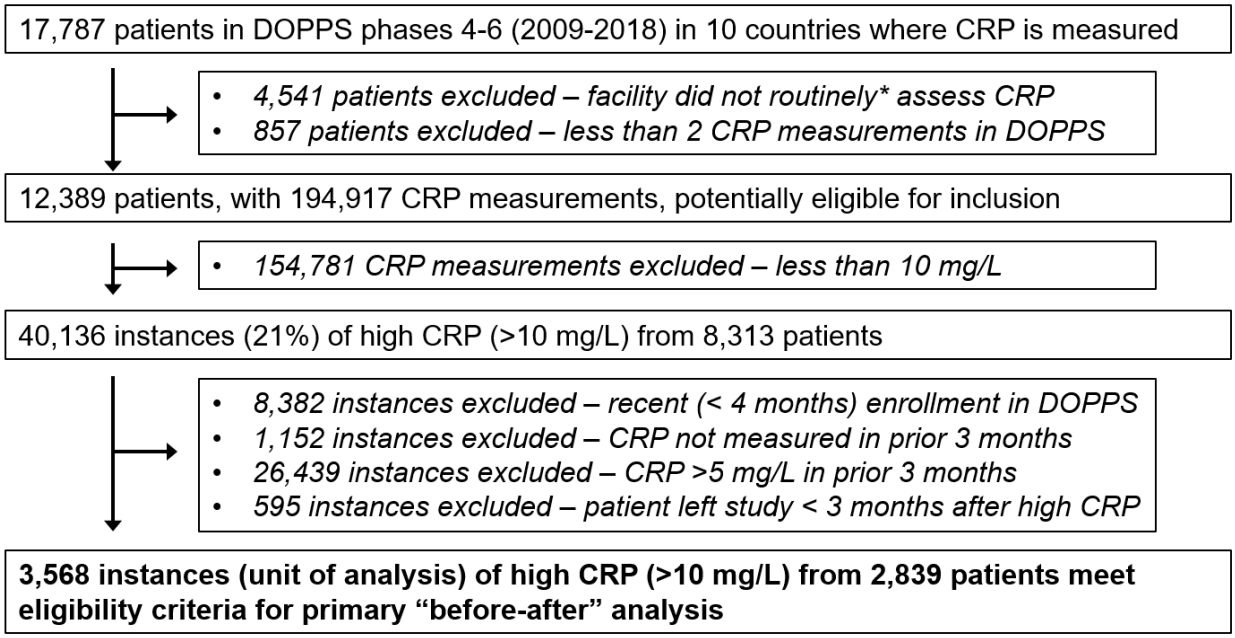
Figure 5. Illustration of before-after study design



For a given patient, average hemoglobin and ESA dose was observed during the 3 months following an increase in CRP from low (≤5 mg/L) to high (>10 mg/L). Time-varying confounders were included during the month preceding the "Before" period (C_1) and the month preceding the CRP increase (C_2).

Because longitudinal ascertainment of CRP was required, we excluded patients dialyzing in facilities that did not routinely assess CRP (measured less than once every 4 months on average) and patients with fewer than two CRP measurements during DOPPS follow-up. The remaining patients were potentially eligible for inclusion. We then identified instances of high CRP (>10 mg/L), considered "month 0." These instances needed to meet four additional criteria to be included in the matched analysis: (1) the patient was enrolled in DOPPS ≥4 months prior to month 0; (2) CRP was measured at least once during the 3 months prior to month 0; (3) all available CRP values were low (≤5 mg/L) during the 3 months prior to month 0; and (4) the patient remained in DOPPS ≥3 months following month 0. Detailed information on the number excluded for various reasons is shown in the flow diagram (**Figure 6**).

Figure 6. Flow chart illustrating inclusion/exclusion criteria



*Routine measurement of CRP by a facility defined as ≥25% of patient-months with a CRP measurement, i.e., CRP measured at least once every 4 months on average. Note that the 194,917 CRP measurements from 12,389 patients were used as the basis to report the distribution of CRP in Figure 7.

Statistical analyses

We first summarized the distribution of CRP levels by country. After applying the inclusion/exclusion criteria above, we summarized both time-fixed and time-varying patient characteristics of the study sample used in the matched analysis. In descriptive analyses to illustrate trends in hemoglobin, ESA dose, and ESA hyporesponsiveness over the 3 months before and after the CRP increase, the mean and standard error (SE) were calculated in each month. To convert ESA doses to units of IV epoetin, we used conversion factors of 250:1 for darbepoetin,[96] 208:1 (250/1.2) for pegylated epoetin-β ,[97] and 1.15:1 for subcutaneous injections.[98]

ESA hyporesponsiveness, the main binary outcome, was defined in each 3-month period as low hemoglobin (<10 g/dL) plus high ESA dose, where the threshold for high ESA dose was lower in Japan (>6000 units/week) than in Europe/ANZ (>8000 units/week) due to generally lower ESA doses in Japan. Hemoglobin levels and ESA doses were averaged over each 3-month period. To estimate the unadjusted prevalence ratio (PR) of ESA hyporesponsiveness in the after vs. before period, we used Mantel-Haenszel methods for matched designs[121] to analyze the 2x2 table. To incorporate potential time-varying confounders, we used an extension of the modified Poisson regression approach for correlated binary data.[122]

The two secondary outcomes were hemoglobin level and ESA dose, each averaged over the 3 months before and after the increase in CRP. We used a natural log transformation of ESA dose due to skewness of the distribution, but we also modeled the untransformed ESA dose. For these continuous outcomes, we used mixed-effects linear regression with an indicator variable for “after” (vs. “before”) as the exposure

contrast of interest. Because multiple inflammation events per patient could be eligible, we used a random intercept to account for within-facility and within-patient clustering.

Factors that are constant within patients (e.g., sex) or change uniformly over time (e.g., age) cannot be confounders in this analysis because they are "matched" perfectly within patients. Within-patient factors that changed between the "before" and "after" periods (e.g., laboratory values, medications) could plausibly confound the estimated effect of rising CRP on each outcome. We adjusted for several of these potential confounders to exclude alternative sources of changes in hemoglobin or ESA dose; we included a set of covariates measured at two time points: 4 months prior to the high CRP and 1 month prior to the high CRP (C_1 and C_2 as illustrated in **Figure 5**). By measuring potential confounders before high CRP was observed, the covariates cannot be affected by the new inflammation (thus avoiding controlling for a mediator on the causal pathway), while they may still plausibly impact hemoglobin and ESA dose during the "before" (C_1) and "after" (C_2) periods. Our models thus included adjustment for DOPPS phase, country, age, sex, vintage (time since HD initiation), BMI, and history of 13 comorbidities (listed in **Table 4**), plus the following time-dependent variables measured at 4 months and 1 month prior to the observed high CRP: serum albumin, white blood cell count, serum phosphorus, cinacalcet use, IV iron dose, hospitalization, and catheter use.

We performed subgroup analyses to assess heterogeneity between Japan and the other countries (due to population differences in CRP levels) and effect modification by patient characteristics. We also performed sensitivity analyses to assess the robustness of our results: (a) varying the number of CRP measurements during the 3-month "before" period; (b) varying the thresholds used to define "low" and "high" CRP; (c)

varying the length of the outcome assessment period; (d) varying the longevity of the CRP increase as sustained (CRP >10 mg/L throughout 3-month “after” period) vs. transient (CRP ≤5 mg/L throughout 3-month “after” period); and (e) varying the ESA dose threshold used to define ESA hyporesponsiveness.

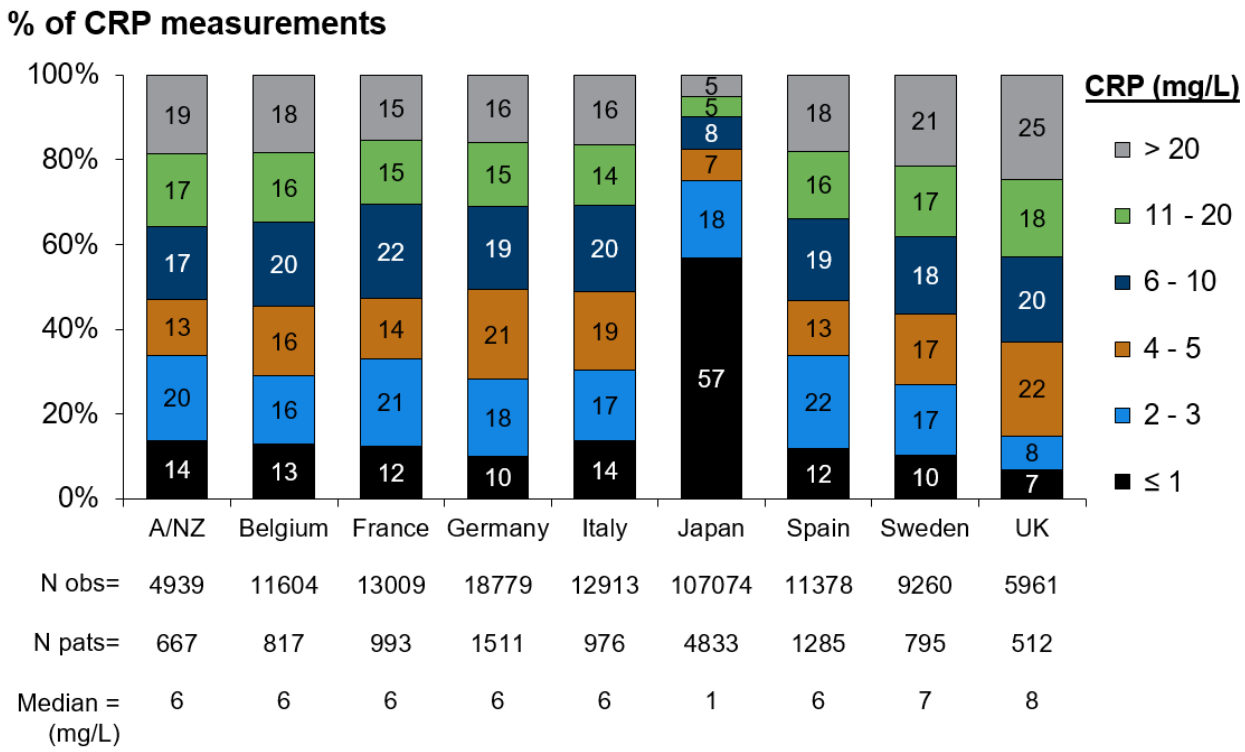
We used multiple imputation, assuming data were missing at random, to impute missing covariate values using the Sequential Regression Multiple Imputation Method by IVEware.[99] Results from 20 such imputed data sets were combined for the final analysis using Rubin’s formula.[100] The proportion of missing data was below 10% for all covariates, with the exception of white blood cell count (12%). All analyses were conducted using SAS software, version 9.4 (SAS institute, Cary, NC).

RESULTS

Prevalence of high CRP, by country

As shown in **Figure 6**, 12,389 patients potentially eligible for inclusion had a total of 194,917 CRP measurements; the median number of measurements was 13 (interquartile range [IQR]: 6, 24). The CRP distribution in this population is reported by country in **Figure 7**. The prevalence of high CRP (>10 mg/L) was greatest in the UK (43%) and 30%-40% across other Europe/ANZ countries; median CRP was 6-8 mg/L across Europe/ANZ. The prevalence of CRP >10 mg/L was much lower in Japan (10%), where 57% of CRP measurements were ≤1 mg/L.

Figure 7. CRP distribution, by country



N pats = number of patients with CRP measurements who are potentially eligible for inclusion (as described in Figure 6); N obs = number of monthly CRP measurements available from these patients (denominators for the figure). A/NZ = Australia / New Zealand.

Self-matched analysis: Patient characteristics

After applying the inclusion/exclusion criteria as in **Figure 6**, we identified 3,568 instances of high CRP (month 0) from 2,839 patients eligible for the primary analysis: 1,659 from DOPPS phase 4, 1,316 from phase 5, and 593 from phase 6. Baseline patient characteristics treated as time-fixed are shown in **Table 4** for patients eligible for the before-after analysis, by region. Compared to Europe/ANZ, patients in Japan tended to have longer vintage, lower BMI, and were less likely to have several comorbid conditions. Time-varying patient characteristics collected longitudinally are shown in **Table 5**. In the 3 months after vs. before an increase in CRP, patients in both regions tended to experience modest decreases in transferrin saturation (TSAT) and serum albumin and modest increases in ferritin and white blood cell (WBC) count. The proportion of patients prescribed IV iron and their respective doses changed minimally. In Europe/ANZ, patients were more likely to receive a red blood cell transfusion (6% vs. 3%) or experience an inpatient hospitalization (26% vs. 19%) in the 3 months after vs. before an increase in CRP, but differences were minimal in Japan. The median (IQR) CRP was 19 (14, 37) in Europe/ANZ and 20 (14, 36) in Japan during the reference month, then dropped to 6 (3, 14) in Europe/ANZ and 3 (1, 7) in Japan during the after period, illustrating that in most cases the rise in CRP to >10 mg/L was not sustained.

Table 4. Summary statistics for time-fixed patient characteristics, by region

	Europe/ANZ	**Japan**
N eligible instances of high CRP	1530	2038
N patients	1293	1546
Age (years)	67.3 (14.5)	68.1 (11.5)
Sex (% male)	62%	71%
Time on dialysis (years)	3.5 (1.7, 6.9)	6.5 (3.0, 12.5)
Body mass index (kg/m^2)		
< 18	4%	16%
18-25	45%	71%
25-30	34%	12%
≥ 30	17%	2%
Comorbidities (%)		
Coronary artery disease	38%	33%
Heart failure	21%	23%
Cerebrovascular disease	17%	15%
Other cardiovascular disease	33%	30%
Cancer (non-skin)	16%	13%
Diabetes	38%	45%
Gastrointestinal bleeding	5%	5%
Hypertension	87%	84%
Lung disease	13%	3%
Neurologic disease	13%	7%
Psychiatric disorder	19%	5%
Peripheral vascular disease	31%	17%
Recurrent cellulitis, gangrene	7%	5%

Mean (standard deviation), median (interquartile range), or % shown; Age, time on dialysis, and body mass index were captured at the time of the CRP increase (month 0, as defined in Figure 5); Comorbidities were captured at DOPPS enrollment; A/NZ = Australia / New Zealand

Table 5. Summary statistics for time-varying patient characteristics before and after the CRP increase from ≤5 to >10 mg/L, by region

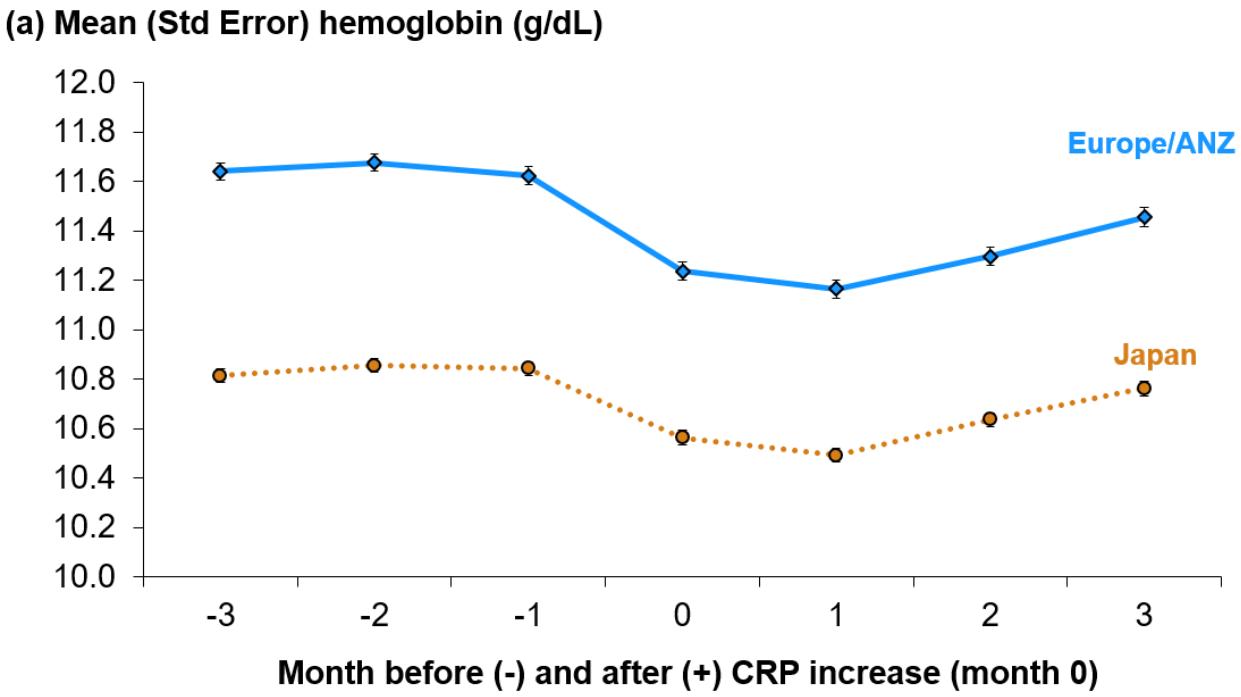
	Europe/ANZ		Japan	
N eligible instances of high CRP	1530		2038	
N patients	1293		1546	
Time-varying characteristic	**3 months "before"**	**3 months "after"**	**3 months "before"**	**3 months "after"**
CRP (mg/L)	3 (2, 4)	6 (3, 14)	2 (1, 3)	3 (1, 7)
TSAT (%)	29.6 (12.3)	27.7 (11.9)	25.2 (11.1)	24.1 (11.1)
Serum ferritin (ng/mL)	402 (223, 613)	452 (259, 689)	77 (37, 147)	83 (42, 172)
Serum albumin (g/dL)	3.8 (0.4)	3.7 (0.5)	3.7 (0.4)	3.6 (0.4)
Serum phosphorus (mg/dL)	4.9 (1.4)	4.9 (1.5)	5.2 (1.2)	5.1 (1.2)
Mean WBC count (10^3 cells/mm^3)	6.7 (1.9)	6.9 (2.1)	5.9 (1.8)	6.1 (1.9)
IV iron use (%, any during 3 mo)	75%	74%	31%	33%
IV iron dose (mg/month)	261 (145, 435)	272 (145, 435)	116 (58, 174)	116 (58, 174)
Cinacalcet use (%, any during 3 mo)	18%	19%	26%	27%
Catheter use (%, any during 3 mo)	25%	25%	1%	1%
Transfused (%, any during 3 mo)	3%	6%	2%	3%
Hospitalized (%, any during 3 mo)	19%	26%	16%	17%

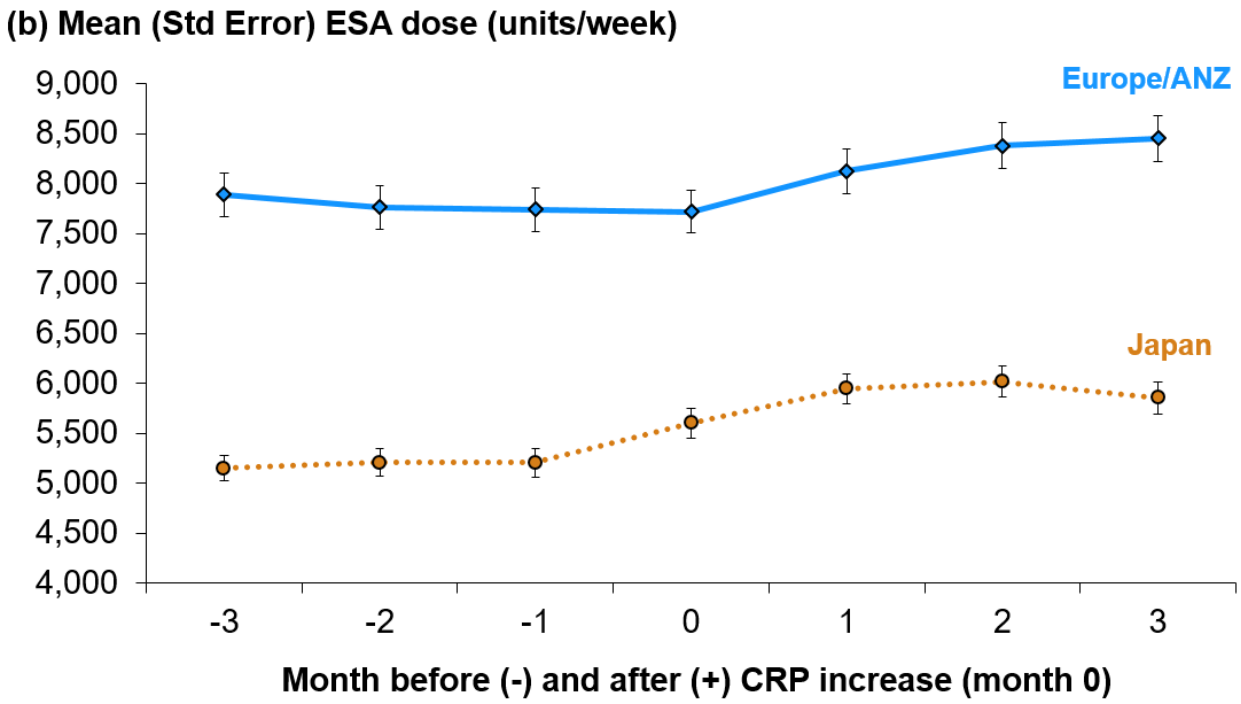
Mean (standard deviation), median (interquartile range), or % shown among all eligible instances of high CRP; "before" = 3 months before the CRP increase and "after" = 3 months after the CRP increase; Monthly lab measures averaged over 3 months; IV iron dose averaged over 3 months among users; A/NZ = Australia / New Zealand, TSAT=transferrin saturation, WBC=white blood cell, IV=intravenous

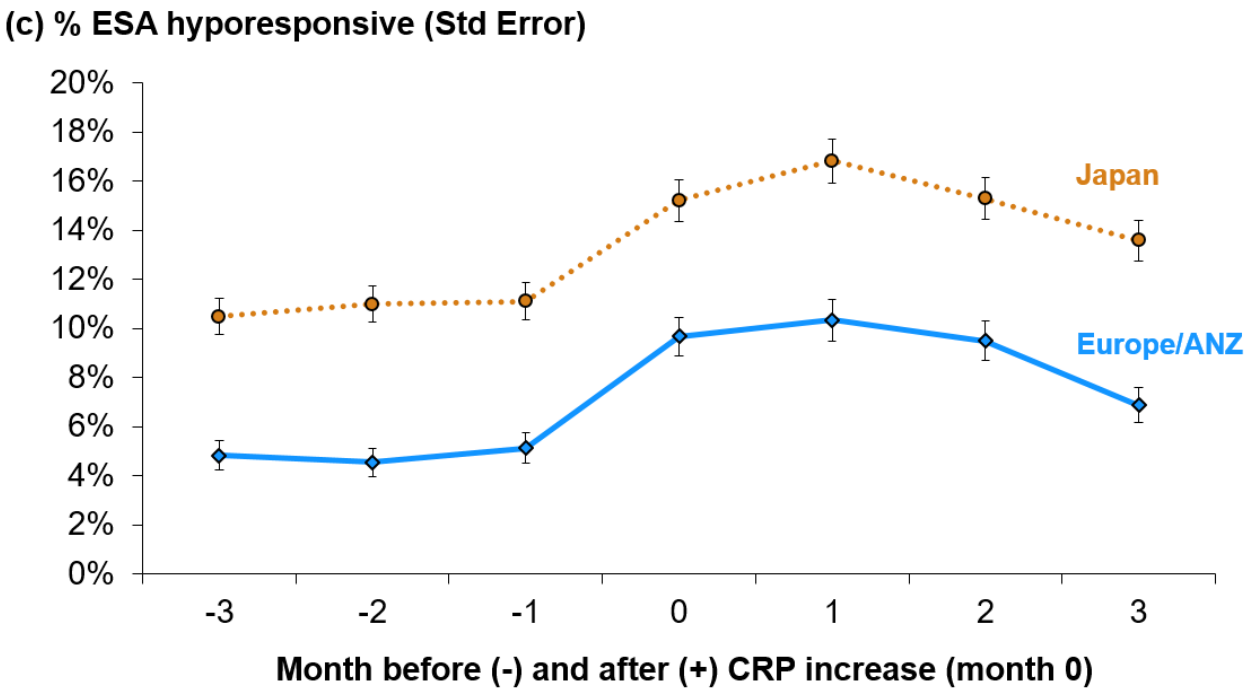
Self-matched analysis: Descriptive results

In **Figure 8**, we present unadjusted monthly (a) mean hemoglobin, (b) mean ESA dose, and (c) proportion ESA hyporesponsive over the 3 months “before” and “after” the high CRP was observed (month 0), by region. In the two regions, hemoglobin changes paralleled each other during the 7-month study period. In Europe/ANZ, mean hemoglobin was 11.6-11.7 g/dL in the 3 months prior to the CRP increase, decreased to 11.2 g/dL in month 0 (concurrent with the CRP increase), then rebounded to 11.5 g/dL 3 months later. In Japan, mean hemoglobin was about 10.8 g/dL in the 3 months prior to the CRP increase, decreased to 10.6 g/dL in month 0, then rebounded back to 10.8 g/dL 3 months later. Mean ESA dose in Europe/ANZ was about 7,800 units/week in the 3 months prior to the CRP increase, then steadily increased to about 8,500 units/week, starting 1 month following the CRP increase. In Japan, mean ESA dose was about 5,200 units/week in the 3 months prior to the CRP increase; in contrast to Europe/ANZ, ESA dose started to increase in month 0 (immediately following the CRP increase) and rose to over 6,000 units/week 2 months after the CRP increase. ESA hyporesponsiveness in both regions increased in month 0, peaked in month 1, and then started to decline towards pre-inflammation levels by month 3.

Figure 8. (a) Mean monthly hemoglobin, (b) mean monthly ESA dose, and (c) % ESA hyporesponsive in the 3 months before and after a CRP increase from ≤5 to >10 mg/L, by region







Mean hemoglobin and ESA dose were calculated as the average across all patients at each time point. Months during which ESA was not prescribed are considered 0 units/week. ESA hyporesponsive defined as hemoglobin <10 g/dL and ESA dose > 6000 (Japan) or > 8000 (Europe/ANZ) units/week. A/NZ = Australia / New Zealand.

Self-matched analysis: Model results

The main findings of this self-matched analysis are shown in the top row of **Table 6**. The adjusted prevalence ratio of ESA hyporesponsivenes of 1.68 (95% CI: 1.48, 1.91) indicates that patients were much more likely to be hyporesponsive during the 3 months after vs. before the rise in CRP. The unadjusted prevalence ratio was also 1.68 (95% CI: 1.48, 1.91), providing strong evidence that our self-matched design adequately accounted for time-fixed and time-varying confounders. Results from the adjusted mixed-effects linear regression models showed that hemoglobin levels were on average 0.26 g/dL lower (95% confidence interval [CI]: 0.22, 0.30) in the 3 months “after” vs. “before” the rise in CRP. The average within-patient change in log(ESA dose) was

0.080 (95% CI: 0.057, 0.104), which, after exponentiating, can be interpreted as an approximate 8.4% (95% CI: 5.8%, 11.0%) increase in ESA dose. In absolute terms, the average within-patient increase in ESA dose was 588 units/week (95% CI: 403, 773). **Table 6** also shows that results from several subgroup analyses by region, catheter use, sex, and age were all directionally consistent with the primary analysis.

Table 7 illustrates the robustness of our results to several sensitivity analyses. Results were consistent when requiring three CRP measurements during the “before” period (a). Increasing the contrast when defining low and high CRP (e.g., from ≤3 to >20 mg/L) (b) resulted in a similar decrease in hemoglobin but a larger increase in ESA dose (14.7%). Reducing the length of the “after” period (e.g., from 3 to 1 month) (c) resulted in a larger hemoglobin decrease (0.42 g/dL) but smaller ESA dose increase (4.4%), as also reflected in the descriptive results (illustrated in **Figure 8**). We observed much larger changes among patients whose CRP increase was sustained >10 mg/L (0.70 g/dL decrease in hemoglobin and 14.2% increase in ESA dose) throughout the 3-month "after" period, compared to those with a transient CRP increase (d). Finally, the adjusted prevalence ratio for ESA hyporesponsiveness was consistent (1.71 vs. 1.68) when increasing the ESA dose thresholds from 6000 to 7500 units/week in Japan and 8000 to 10000 units/week in Europe/ANZ (e).

Table 6. Within-patient changes (95% CI) in hemoglobin, ESA dose, and ESA hyporesponsiveness from the 3 months before vs. after the CRP increase from ≤5 to >10 mg/L, overall and by subgroup

Subgroup	Instances of high CRP	Change in hemoglobin (g/dL)	Relative change in ESA dose (%)	Prevalence ratio of ESA hyporesponsiveness
Overall	3568	-0.26 (-0.30, -0.22)	8.4 (5.8, 11.0)	1.68 (1.48, 1.91)
By region				
Europe/ANZ	1530	-0.34 (-0.41, -0.27)	5.2 (1.5, 9.0)	2.09 (1.60, 2.74)
Japan	2038	-0.20 (-0.25, -0.16)	10.8 (7.4, 14.3)	1.54 (1.34, 1.78)
By catheter use during "before" period				
Any catheter use	372	-0.50 (-0.62, -0.37)	9.8 (2.0, 18.3)	3.16 (1.74, 5.76)
No catheter use	2908	-0.23 (-0.27, -0.19)	8.0 (5.4, 10.8)	1.54 (1.35, 1.77)
By sex				
Male	2389	-0.25 (-0.30, -0.20)	7.2 (4.1, 10.4)	1.63 (1.40, 1.90)
Female	1177	-0.29 (-0.36, -0.22)	10.8 (6.3, 15.4)	1.81 (1.45, 2.27)
By age (years)				
< 60	874	-0.23 (-0.31, -0.15)	9.2 (4.1, 14.4)	1.51 (1.14, 1.98)
60-75	1559	-0.27 (-0.33, -0.21)	7.8 (3.9, 11.8)	1.79 (1.47, 2.17)
> 75	1135	-0.28 (-0.35, -0.21)	8.9 (4.7, 13.3)	1.66 (1.35, 2.04)

Linear mixed model with random facility and patient intercepts to calculate mean changes in hemoglobin and ESA dose, and modified Poisson regression to calculate prevalence ratio of ESA hyporesponsiveness; baseline adjustment for DOPPS phase, country, age, sex, vintage, BMI, and 13 comorbidities, and adjustment for serum albumin, WBC count, serum phosphorus, cinacalcet use, IV iron dose, hospitalization, and catheter use at 4 months and 1 month prior to the CRP increase. ESA hyporesponsiveness defined as hemoglobin <10 g/dL and ESA dose >6000 (Japan) or >8000 (Europe/ANZ) units/week; primary analysis includes patients with CRP ≤ 5 mg/L during 3 month "before" period, increased to > 10 mg/L, then followed up during 3 month "after" period; ANZ = Australia / New Zealand

Table 7. Within-patient changes (95% CI) in hemoglobin, ESA dose, and ESA hyporesponsiveness from the 3 months before vs. after the CRP increase: Sensitivity analyses

Sensitivity analysis	Instances of high CRP	Change in hemoglobin (g/dL)	Relative change in ESA dose (%)	Prevalence ratio of ESA hyporesponsiveness
Primary analysis	3568	-0.26 (-0.30, -0.22)	8.4 (5.8, 11.0)	1.68 (1.48, 1.91)
(a) CRP measurements during 3 month "before" period				
CRP measured all 3 months	2312	-0.26 (-0.30, -0.21)	9.9 (6.7, 13.1)	1.65 (1.42, 1.93)
CRP measured in 1 or 2 of the 3 months	1256	-0.28 (-0.36, -0.21)	5.9 (1.8, 10.2)	1.76 (1.41, 2.20)
(b) Varying thresholds for "low" and "high" CRP				
CRP increase from ≤ 10 to > 20 mg/L	3008	-0.27 (-0.32, -0.23)	10.2 (7.4, 13.1)	1.56 (1.35, 1.80)
CRP increase from ≤ 5 to > 20 mg/L	1703	-0.30 (-0.36, -0.24)	11.7 (8.0, 15.6)	1.70 (1.39, 2.07)
CRP increase from ≤ 3 to > 20 mg/L	1053	-0.27 (-0.34, -0.19)	14.7 (9.7, 19.8)	1.66 (1.29, 2.15)
CRP increase from ≤ 3 to > 10 mg/L	2178	-0.27 (-0.32, -0.22)	11.1 (7.8, 14.5)	1.73 (1.46, 2.05)
CRP increase from ≤ 3 to > 5 mg/L	4230	-0.18 (-0.22, -0.15)	6.4 (4.2, 8.7)	1.43 (1.25, 1.63)
CRP increase from ≤ 1 to > 5 mg/L	1624	-0.18 (-0.23, -0.12)	10.9 (7.2, 14.7)	1.36 (1.10, 1.68)
(c) Vary length of "after" period for assessing outcome				
1 month "after" period	3958	-0.42 (-0.46, -0.37)	4.4 (1.9, 6.9)	1.91 (1.68, 2.17)
2 month "after" period	3755	-0.34 (-0.39, -0.30)	7.7 (5.3, 10.2)	1.82 (1.60, 2.08)
(d) By longevity of CRP increase in "after" period				
Sustained: CRP > 10 mg/L in "after" period	352	-0.70 (-0.85, -0.55)	14.2 (4.7, 24.5)	2.89 (1.97, 4.24)
Transient: CRP ≤ 5 mg/L in "after" period	1652	-0.14 (-0.19, -0.09)	5.6 (2.1, 9.1)	1.22 (1.00, 1.48)
(e) Vary thresholds for ESA hyporesponsiveness ESA dose >7500 (Japan) or >10000 (Europe/ANZ) units/week	3568	N/A	N/A	1.71 (1.49, 1.96)

Linear mixed model with random facility and patient intercepts to calculate mean changes in hemoglobin and ESA dose, and modified Poisson regression to calculate prevalence ratio of ESA hyporesponsiveness; baseline adjustment for DOPPS phase, country, age, sex, vintage, BMI, and 13 comorbidities, and adjustment for serum albumin, WBC count, serum phosphorus, cinacalcet use, IV iron dose, hospitalization, and catheter use at 4 months and 1 month prior to the CRP increase. ESA hyporesponsiveness defined as hemoglobin <10 g/dL and ESA dose >6000 (Japan) or >8000 (Europe/ANZ) units/week unless otherwise specified; primary analysis includes patients with CRP ≤ 5 mg/L during 3 month "before" period, increased to > 10 mg/L, then followed up during 3 month "after" period; ANZ = Australia / New Zealand

DISCUSSION

This self-matched longitudinal ("before-after") design and analysis tracked real-world changes in anemia control and ESA dosing in an international sample of HD patients over the 3 months before and after detection of new inflammation by routine CRP measurement. The results supported our hypothesis of a hemoglobin decrease and ESA dose increase, resulting in greater ESA resistance and exposing patients to the potential risks of larger ESA doses.[17–20] The associations were particularly strong among patients whose CRP increase was sustained over the subsequent 3 months, further supporting a causal relation between inflammation and ESA hyporesponsiveness.

Our results were generally consistent with other longitudinal studies,[79,80,82] but differed from an observational extension of the Dialysis Patients' Response to IV Iron with Elevated Ferritin (DRIVE) trial.[81] The DRIVE authors found that CRP was not associated with hemoglobin response, though they did observe a lower likelihood of hemoglobin response to IV iron therapy when CRP was >14.1 mg/L.[81] These divergent results could be explained by the restricted selection criteria for DRIVE (e.g., hemoglobin ≤11 g/dl, TSAT <25%, ferritin 500-1200 ng/ml, 22,500 units/week ESA),[81] which likely included many participants who had previously experienced a hemoglobin drop in response to inflammation. Bradbury et al.[79] observed that elevated CRP led to larger ESA doses at the same hemoglobin levels. However, the authors acknowledged the potential for selection bias in their US HD sample because only about 1% of patients had CRP measured. These patients were likely selected for CRP measurement due to suspicion of inflammation, as the median CRP (20 mg/L) was much higher than

reported in other HD populations.[47,77,78] Gillespie et al.[82] conducted a case-crossover study of ESA hyporesponsiveness defined as hemoglobin <10 g/dL and ESA dose >median dose of 80 units/kg/week, which they observed in 672 European HD patients. Among the many exposures Gillespie et al. examined, they found a positive association with CRP (adjusted odds ratio for highest vs. lowest quartile [no values provided] = 2.02, 95% CI: 1.20-3.38). Kimachi et al.[80] used Japanese DOPPS data from phases 2-4 (2002-2011) to evaluate the cumulative incidence of ESA hyporesponsiveness (hemoglobin <10 g/dL and >9000 units/week ESA) by baseline CRP. Those authors found that the risk of ESA hyporesponsiveness was highest at CRP >10 mg/L, but also elevated at CRP 3-10 mg/L (vs. CRP <1 mg/L).

The proportion of CRP measurements >10 mg/L was much lower in Japan (10%) than in Europe/ANZ (34%), consistent with prior research.[47,77,80] Japanese HD practices may help explain this discrepancy, including the use of ultrapure dialysate fluid to keep endotoxins low, and the avoidance of central venous catheters, which can cause infections and inflammatory reactions.[120] However, CRP levels are also lower in Asians than in whites outside the HD setting,[123] suggesting dietary, environmental, and/or genetic factors as likely contributors to differences in CRP levels.[112]

In our analysis, hemoglobin began to decline in the same month that CRP rose. Elevated CRP is generally considered a marker of inflammation, so new inflammation may alter hemoglobin levels roughly concurrent with its effect on CRP. In Japan the increase in ESA dose occurs in the same month that hemoglobin began to decline (i.e., month 0), but in Europe/ANZ the initial ESA dose increase lags by 1 month. Increases in CRP from ≤5 to >10 mg/L were also less likely to be sustained for 3+ months in

Japan (5%) than in Europe/ANZ (18%). These findings suggest either that clinicians in Japan respond more rapidly to decline in hemoglobin than in Europe/ANZ or that Japanese providers are reacting proactively to rises in CRP below the 10 mg/L threshold used in this analysis.

CRP is relatively inexpensive and convenient to routinely measure in the HD setting;[124] this is generally done in Europe and Japan, but not in North America. Routine measurement of CRP can potentially help to better identify causes of, and inform targeted strategies to reduce, inflammation in HD patients. For example, a rise in CRP may prompt examination for source of infection (e.g., dental and diabetic foot exams) and timely initiation of antimicrobial therapy when indicated. Other longer-term strategies to limit or reduce inflammation include removing old non-functioning arteriovenous grafts,[125,126] transplant nephrectomy,[127] using ultrapure dialysate fluid,[110,112] and improving diet and exercise.[113] Specific to anemic patients, quicker recognition of new inflammation can help identify the cause of worsening anemia and guide reactive ESA and IV iron dosing decisions. Further, frequent assessment of inflammation can help identify patients who may be candidates for new alternative anemia therapies, such as HIF-PHIs, that may be less susceptible to the effects of inflammation than current ESA and IV iron-based treatment regimens.[116–118]

This study had some limitations. First, because measurement frequency of CRP varies across HD facilities, patients in facilities that measure CRP only every 3 months may experience a transiently high CRP that is not detected; however, a sensitivity analysis limited to patients with CRP measured at least monthly during the 3-month “before” period was consistent with the primary analysis, suggesting that this may not be a

concern. Second, if patients excluded due to <3 months of data post-inflammation were more likely to have experienced ESA hyporesponsiveness following their high CRP, the true effect may be underestimated; however, that bias is likely minimal because of the small proportion (14%) of excluded patients. Third, while we adjusted for several time-varying confounders, it is possible that our estimates suffered from residual confounding by unmeasured time-varying risk factors for ESA hyporesponsiveness. However, because the extensive covariate adjustments in our models had little impact on our estimates in this self-matched study, the likelihood of bias due to unmeasured confounding is low.

Several strengths distinguish this analysis from other longitudinal studies of inflammation and ESA hyporesponsiveness.[79,80,82] First, the longitudinal study design focuses on incident inflammation to avoid the temporal ambiguity of cross-sectional designs. By matching patients to themselves and measuring outcomes before and after the detection of elevated CRP, this design does not require a comparison group of patients who did not experience an increase in CRP. Indeed, the self-matching seems to have controlled adequately for potential confounders—both fixed and time-varying factors—as evidenced by the unchanged estimates after additional adjustment. Future studies, however, could perform between-patient comparisons that we did not investigate. Second, we utilized a large international sample of HD patients from facilities that routinely measured CRP, the best available marker of inflammation, to avoid bias in which a clinical indication for measuring CRP also affects the outcome (a phenomenon we call “measurement-by-indication bias”). Third, in addition to a single ESA hyporesponse outcome, we treated the two components of that outcome,

hemoglobin and ESA dose, as separate continuous outcomes, allowing us to better explore relative changes in ESA sensitivity without relying on the flawed erythropoietin resistance index.[104,128]

This study demonstrates that new inflammation, as detected by an increase in CRP, is associated with development of ESA resistance and reduction in hemoglobin levels under current anemia treatment paradigms. These findings speak to a potentially important role for anemia therapies that are less susceptible to the effects of inflammation.

CHAPTER IV

Replicating Randomized Trial Results with Observational Data using the Parametric g-formula: An Application to Intravenous Iron Treatment in Hemodialysis Patients

Coauthors who contributed to this work include: Hal Morgenstern, Nancy L Fleischer, and Bruce M Robinson.

INTRODUCTION

Large high-quality randomized trials are costly, time-consuming, and inflexible to different selection criteria and intervention protocols. While a practical alternative is to utilize observational data, confounding can be difficult to overcome, particularly in the presence of treatment-confounder feedback loops. Marginal structural models can handle these time-dependent confounders, but are more efficient for estimating the effect of static treatment regimens.[129–131] The g-formula, an extension of standardization to longitudinal data, is better suited to evaluate the complex and dynamic treatment strategies in a "target trial".[83,132,133] The g-formula can consistently estimate the probability of the outcome under a hypothetical intervention by estimating a weighted sum of the risk across all risk factor histories, as in ordinary non-parametric standardization with categorical variables.[121] The parametric version of the g-formula is an appealing approach that utilizes modeling to avoid sparse cells,[131] but

implementation in practice has lagged for reasons ranging from software and data limitations to unfamiliarity or lack of confidence in the method.

Cohort studies of patients with end-stage kidney disease undergoing hemodialysis 3 times/week in hemodialysis centers with standardized treatment protocols are particularly suitable for g-methods that depend on rich longitudinal data. Most hemodialysis patients are anemic and require erythropoiesis stimulating agent (ESA) therapy to maintain hemoglobin levels in target range – generally 10.0-11.5 g/dL.[86] Intravenous (IV) iron is often administered to complement ESA treatment and avoid iron deficiency by replacing the iron utilized for erythropoiesis.[9] Conflicting evidence from observational data exists regarding the safety of high-dose IV iron supplementation in hemodialysis patients.[48,62–64,91,134] IV iron dosing decisions are, in the context of hemoglobin level, guided primarily by serum ferritin, a marker of iron stores, and transferrin saturation (TSAT), a marker of circulating iron.[15] Many hemodialysis patients with high ferritin but low TSAT levels may still be functionally iron deficient and experience a hemoglobin increase in response to IV iron;[61] discontinuing IV iron in these patients may have deleterious effects, including need for higher doses of expensive[135] and potentially harmful[17–20] ESA therapy. On the other hand, continuing high-dose IV iron therapy despite elevated serum ferritin may improve surrogate outcomes such as hemoglobin and ESA dose, but result in greater risk of mortality. Investigators of the recently published Proactive IV Iron Therapy in Haemodialysis Patients (PIVOTAL) study, a large UK-based randomized controlled trial, concluded that a proactive high-dose (vs. reactive low-dose) IV iron treatment regime was superior.[65]

The first goal of our study is to replicate findings from the PIVOTAL trial by applying the parametric g-formula to hemodialysis patients in the European arm of the Dialysis Outcomes and Practice Patterns Study (DOPPS), where anemia management practices are relatively similar to the UK (and much different than in North America or Japan).[136] The second goal is to simulate the PIVOTAL study in a similar trial population by applying the parametric g-formula to DOPPS patients restricted according to PIVOTAL inclusion criteria. We aim to provide a proof-of-concept for extending the parametric g-formula target-trial approach to research questions that remain unaddressed by randomized trials, which are generally designed to test only one or two specific hypotheses. If the hypothetical target trial we emulate is similar enough to the actual trial, the PIVOTAL findings should be replicable in our simulation. The potential to evaluate many variations of complex intervention strategies across different populations could prove to be enormously informative in the age of big data.

METHODS

Data source

The DOPPS is a prospective cohort study of center-based, adult chronic hemodialysis patients in 21 countries, ongoing since 1996 in multiple phases. Study sites and patients are randomly selected to achieve nationally representative samples in each country. Details on study design and objectives are included in prior publications[94,95] and at https://www.DOPPS.org/. This analysis included a cohort of hemodialysis patients from 7 European countries (Belgium, France, Germany, Italy, Spain, Sweden, UK) in DOPPS phase 4 (2009-2011) and phase 5 (2012-2015). Information on patient demographics and comorbidity history was abstracted from medical records at DOPPS enrollment.

Measured laboratory values and medication prescriptions were abstracted from medical records at baseline and monthly during follow-up.

Treatment strategies

As in the PIVOTAL trial,[65] the IV iron dose assigned each month depended on the most recent values of ferritin and TSAT (**Table 8**). In the proactive high-dose arm, 400 mg IV iron was administered monthly unless upper thresholds of ferritin (>700 ng/mL) or TSAT (>40%) were reached. In the reactive low-dose arm, lower IV iron doses (100 or 200 mg/month) were administered with progressively lower thresholds for ferritin.

Table 8. Summary of treatment strategies per PIVOTAL[65] trial protocol

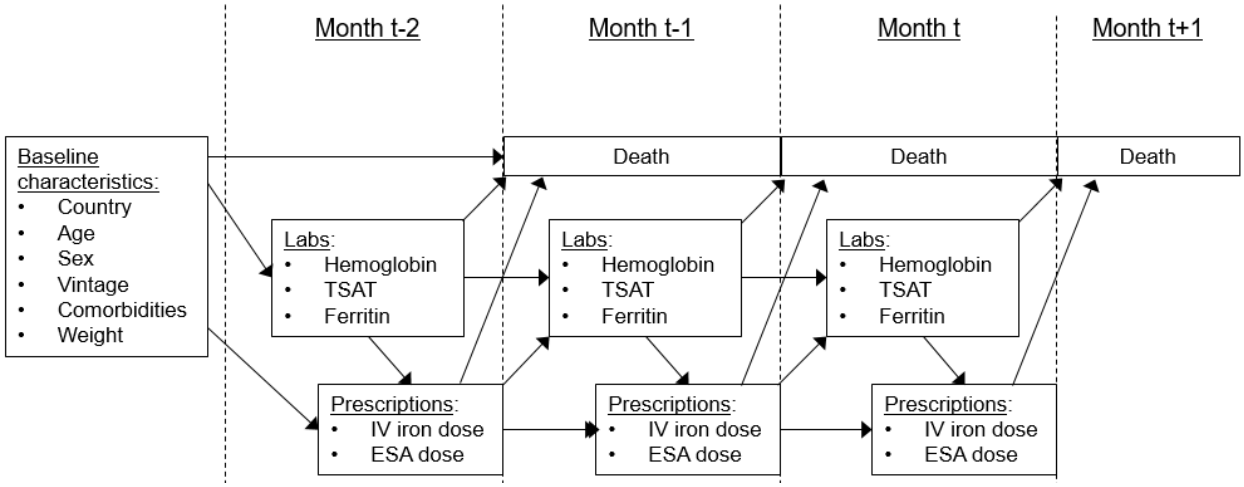
	TSAT (%)		
Ferritin (ng/mL)	**< 20**	**20-40**	**> 40**
< 100	400	400	0
100-200	Low: 200 High: 400	Low: 200 High: 400	0
200-700	Low: 100 High: 400	Low: 0 High: 400	0
> 700	0	0	0

Intravenous (IV) iron dose (mg) administered in the following month based on most recent value of serum ferritin and transferrin saturation (TSAT) under the proactive “High” dose vs. reactive “Low” dose treatment strategy; cells without a Low/High designation indicates that the dose was the same for both strategies.

Statistical analysis

To test the high vs. low-dose IV iron treatment strategies, we implemented the parametric g-formula to account for the treatment (IV iron) – confounder (ferritin, TSAT) feedback loop (**Figure 9**). The two primary steps of the parametric g-formula are (1) modeling the joint distribution of all variables and (2) simulating variables over the follow-up period using the estimates from Step 1. These steps are described in more detail below. Additional details related to the formulae and assumptions have been previously reported.[132,137]

Figure 9. Illustration of longitudinal data collection and hypothesized relationships



Baseline characteristics captured at study enrollment; Laboratory values measured monthly; updated prescriptions captured at the end of each month; Patients followed up continuously for mortality events

Step 1: Parametric models

To estimate the coefficients used in the Step 2 simulation, we first modeled the joint distribution of all relevant variables, specified as either time-fixed (captured once at study enrollment) or time-updated (measured monthly). We fit separate regression models for each time-updated variable – including treatments, potential confounders, and outcomes – for all patient-months with available data. All models in month k included the following variables:

(1) Time-fixed covariates: indicators for DOPPS phase and country; sex; age and time since hemodialysis initiation (updated monthly); weight; and 14 indicators for summary comorbidities.

(2) Values from months k-1 and k-2 for the following time-updated covariates: indicators for hospitalization (any during month) and catheter use for vascular access (status at month end); serum phosphorus; hemoglobin; ESA dose; and IV iron dose.

(3) Value from month k-1 (carried forward up to 2 months) for the following time-updated covariates: C-reactive protein (CRP); serum albumin; ferritin; and TSAT. Because these variables were only assessed every 3 months in some hemodialysis facilities, we carry forward to represent the most recent chart information available to the prescribing physician; to help counter potential biases with this approach, models also included indicator variables for whether each lab was most recently measured 1, 2, or 3 months prior. DOPPS facilities that did not routinely measure ferritin or TSAT (in <25% of patient-months) were excluded.

Table 9 provides details on how each time-updated variable was parameterized when used as the outcome and when used as an exposure. Model choices were intuitively linked to the distribution (e.g., logistic regression for binary, linear regression for normally distributed, log-transform if skewed). IV iron doses were primarily limited to a small number of discrete prescriptions (e.g., 25, 50, 62.5, or 100 mg/week). We thus categorized doses into 4 groups (0, 1-35, 35-63, >63 mg/week), with each category roughly corresponding to 1 of the 4 doses assigned per clinical protocol in PIVOTAL (0, 100, 200, 400 mg/month).

Table 9 also describes the sequential model ordering. Model 1 (hospitalization as the outcome) includes all predictors described in the previous paragraph (1-3), then Model 2 (CRP) includes all variables described above, plus an indicator for hospitalization in the current month, then Model 3 adds CRP in the current month, etc. The model numbering indicates a natural ordering within each month: hospitalizations occurring during the month, then laboratory measurements, then catheter use and prescription of medications at month end. Laboratory values are generally measured simultaneously and thus the lab ordering is largely arbitrary. However, medication prescriptions at month end depend on the current month's lab values (not vice versa), illustrating the importance of a "causal" ordering.

Table 9. Summary of Step 1 models and covariates

Model	Variable	Regression model when used as outcome	Functional form when used as predictor
1	Hospitalization	Logistic	Binary (yes/no)
2	C-reactive protein	Linear (log-scale)	Log-linear
3	Serum albumin	Linear	Linear
4	Serum phosphorus	Linear (log-scale)	Categories (3.5, 5.5, 7.0 mg/dL)
5	Hemoglobin	Linear	Categories (9, 10, 11, 12, 13 g/dL)
6	Serum ferritin	Linear (log-scale)	Categories (100, 200, 400, 700, 1000 ng/mL)
7	TSAT	Linear (log-scale)	Categories (15, 20, 25, 30, 35, 40%)
8	Catheter use	Logistic	Binary (yes/no)
9	IV iron dose	Multinomial logistic	Categories (0, 35, 63 mg/week)
10+11	ESA dose	Logistic, linear (log-scale)	Categories (0, 3000, 6000, 9000, 15000 units/week)
12	Died next month	Logistic	N/A

Categories: indicates cut-points used; IV iron doses were largely discrete, and so the 3 non-zero categories of 1-35 (mostly 25), 35-63 (mostly 50 or 62.5), and >63 (mostly 100) mg/week generally correspond to 100, 200, and 400 mg/month, respectively; For ESA dose, separate models were used to first model use (yes/no), and then the dosage among the users

Step 2: Simulation

In the next step, we simulated variables and outcomes using output from the Step 1 models. We began by designating each patient's third month of follow-up as month 1; this is the month we began the simulation because data from the prior 2 months are needed to serve as inputs. We augmented our data by re-sampling patients with replacement (N=10,000) to obtain more stable estimates as recommended by Keil et al.[138]

Starting with Model 1 (hospitalization) in **Table 9**, we calculated the predicted probability ($\hat{p}$) of hospitalization for each patient by applying model coefficients from Step 1 to the observed covariates. We then moved to Model 2 (CRP) and calculated the predicted value for CRP by similarly applying model coefficients from Step 1 to the observed covariates and the newly predicted binary hospitalization indicator. This process was repeated iteratively for each modeled outcome in **Table 9** through Model 10, when we calculated the predicted probability that the patient died in the following month, based on all of the newly predicted values plus patient characteristics and 2-month risk factor history.

Rather than assign values deterministically, error terms were added to every predicted value. For continuous outcomes, we added the square root of the model mean squared error (root MSE) multiplied by a randomly drawn value from a normal (0,1) distribution. If the updated predicted value falls outside the range of observed data, we re-drew the random value until it fell within the boundaries. For binary outcomes, we compared $\hat{p}$ with a randomly drawn value (U) from a uniform (0,1) distribution, assigning a value of 1 if $\hat{p} > U$ and 0 if $\hat{p} < U$. For multinomial outcomes, we divided the space between 0 and

1 proportionally based on the predicted probability of each category, then randomly drew a value from a uniform (0,1) distribution, assigning the value to whichever “bucket” the random number falls.

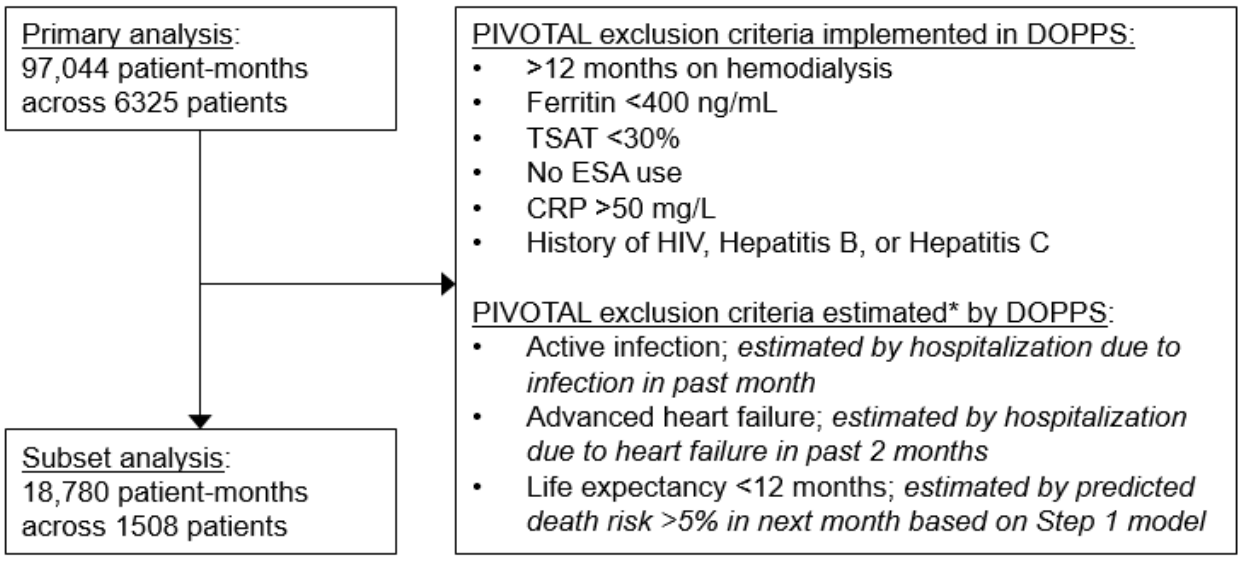
After predicting values for all modeled variables in month 1, we removed patients who were simulated to have died in that month and replaced the observed k-2 values with observed k-1 values, and replaced the k-1 values with the newly predicted month k values. Starting again with Model 1, we predicted values in month 2 for hospitalization, CRP, etc. This process was iterated for 12 months to simulate 1 year of follow up. This simulation is defined as the “natural course” because all variables – including the exposure of interest (IV iron dose) – were assigned probabilistically. Assuming independent censoring, the simulated risk factor histories and risk of death under the natural course should match the observed data because the joint distribution for the simulated population should be equal to the study population; any departures may signify potential model misspecification.

We performed this 12-month simulation three times: once under the “natural course” and then again under the proactive high-dose and reactive low-dose IV iron treatment strategies to simulate what would occur if the entire study population were treated under each protocol. For the latter approaches, we assigned IV iron deterministically, under the protocols detailed in **Table 8**, based on the most recent values of each patient’s ferritin and TSAT levels.

Subset analysis

For our second objective, we attempt to more closely emulate results from the PIVOTAL trial by restricting our DOPPS sample based on PIVOTAL inclusion criteria. Prior to Step 1, for each patient, we select as the new "baseline" the first month that the patient met PIVOTAL eligibility criteria.[139] **Figure 10** summarizes these criteria and how we attempted to replicate each criterion in DOPPS. Step 1 models included the baseline month and all subsequent patient-months for eligible patients; Step 2 was then carried out as in the primary analysis.

Figure 10. Flow diagram with PIVOTAL[65] exclusion criteria



PIVOTAL exclusion criteria derived from Table 2 in Macdougall et al.[139] and divided into 3 groups: (1) objective criteria we could implement directly in DOPPS; (2) criteria for which we did not capture enough information in DOPPS (living-donor transplant scheduled within 12 months; current active malignancy; chronic liver disease; pregnancy or breast feeding; history of acquired iron overload; previous severe hypersensitivity reactions to IV iron sucrose; compromised ability to give written informed consent and/or comply with study procedures); and (3) *criteria that did not correspond directly with DOPPS data, but were estimated using available data as follows: (i) active infection: hospitalized with infection in previous month; (ii) advanced heart failure: hospitalized with heart failure in previous 2 months; (iii) life expectancy <12 months per judgement of the investigator: predicted death risk >5% (>96th percentile) in next month in Step 1 models

Summarizing results

We reported 12-month trajectories for all modeled variables and cumulative risk of mortality for (1) observed DOPPS data; (2) natural course simulation; (3) PIVOTAL high-dose simulation; and (4) PIVOTAL low-dose simulation. We sought to make three comparisons: observed data vs. natural course simulation (1 vs. 2) to check for model misspecification; PIVOTAL high vs. low-dose (3 vs. 4) simulations to assess the treatment strategies; and simulated PIVOTAL strategies vs. the published PIVOTAL trial data[65] to assess how closely our parametric g-formula results matched a real randomized trial. From our simulations, we reported the 1-year mortality risk ratio (RR) and risk difference (RD) comparing the two PIVOTAL strategies.

Confidence intervals (CIs) were estimated by combining multiple imputation with bootstrapping.[140] We first applied multiple imputation to deal with missing data using the Sequential Regression Multiple Imputation Method by IVEware.[99] Information from both between-patient and within-patient (using each of the 3 prior months) was incorporated into the imputation. We then resampled patients with replacement 100 times and performed Step 1 and Step 2 on each resample. This process was repeated for 10 imputations, resulting in 1000 datasets. We then derived the 95% CI as the 2.5th and 97.5th percentile of the estimator distribution from these datasets based on the “MI boot (pooled sample)” procedure described by Schomaker and Heumann.[140]

In general, we relied on published g-formula analyses by Taubman et al.[141] and others,[132,137,138,142,143] following their step-by-step approach to help guide our analysis. Although sample code was provided in this previous research, we coded from scratch

using SAS version 9.4 (SAS institute, Cary, NC) to avoid a “black box” implementation without a full understanding of the mechanics.

RESULTS

Study sample

Models in Step 1 utilized data from 97,044 patient-months across 6325 patients; the median (interquartile range [IQR]) number of months contributed by each patient was 15 (9, 26). **Table 10** shows baseline patients characteristics for (1) the full DOPPS sample used in our primary analysis (N=6325); (2) the DOPPS subset after restricting based on PIVOTAL eligibility criteria (N=1508); and (3-4) PIVOTAL patients randomized to the high-dose and low-dose IV iron treatment protocols. Note that blank cells in the PIVOTAL columns represent variables not reported in the PIVOTAL Table 1.[65] There were several key differences between DOPPS patients and PIVOTAL participants: DOPPS patients were older, had been on hemodialysis for a longer period, weighed less, had higher levels of serum ferritin, TSAT, and hemoglobin, and were more likely to have a history of heart failure, hypertension, and peripheral vascular disease. Some of these differences were neutralized by further restriction of the DOPPS data based on PIVOTAL eligibility criteria (e.g., time since hemodialysis start, ferritin, TSAT), but others were not (e.g., age, weight, hemoglobin, comorbidity history).

Table 10. Summary of baseline patient characteristics in the DOPPS (by type of analysis) and the PIVOTAL trial[65] (by treatment group)

	DOPPS observational data		PIVOTAL trial data	
Patient characteristics	**Primary analysis**	**PIVOTAL-restricted**	**Proactive high dose arm**	**Reactive low dose arm**
N patients	6325	1508	1093	1048
Time-fixed variables				
Age (years)	65.9 ± 15.1	66.0 ± 15.4	62.7 ± 14.9	62.9 ± 15.1
Sex (% male)	61%	62%	65%	66%
Time since HD start (months)	23.0 (6.1, 62.0)	4.1 (2.5, 6.5)	4.9 (2.8, 8.4)	4.8 (2.8, 8.1)
Weight (kg)	72.0 ± 16.8	75.3 ± 17.6	81.3 ± 21.0	82.9 ± 20.9
Anemia-related variables				
Serum ferritin (ng/mL)	357 (183, 581)	179 (99, 276)	214 (132, 305)	217 (137, 301)
TSAT (%)	24 (18, 33)	19 (15, 24)	20 (16, 24)	20 (16, 24)
Hemoglobin (g/dL)	11.4 ± 1.4	11.2 ± 1.4	10.6 ± 1.4	10.5 ± 1.4
ESA use (%)	88%	100%	100%	100%
ESA dose (1000 units/week)	7.8 (4.8, 12.5)	8.6 (5.0, 13.0)	8.0 (5.0, 10.0)	8.0 (5.0, 12.0)
IV iron use (%)	70%	81%	--	--
IV iron dose (mg/month)	383 ± 232	439 ± 245	--	--
Other time-updated variables				
Serum albumin (g/dL)	3.7 ± 0.5	3.7 ± 0.5	--	--
Serum phosphorus (mg/dL)	4.9 ± 1.6	5.1 ± 1.5	--	--
C-reactive protein (mg/L)	6.0 (2.9, 13.4)	5.0 (2.9, 10.5)	6.0 (3.3, 13.9)	7.0 (4.0, 15.0)
Hospitalized in last month (%)	10%	10%	--	--
Catheter use (%)	28%	35%	41%	41%
Comorbidity history (%)				
Coronary artery disease	34%	33%	--	--
Heart failure	21%	21%	4%	4%
Cerebrovascular disease	16%	15%	--	--

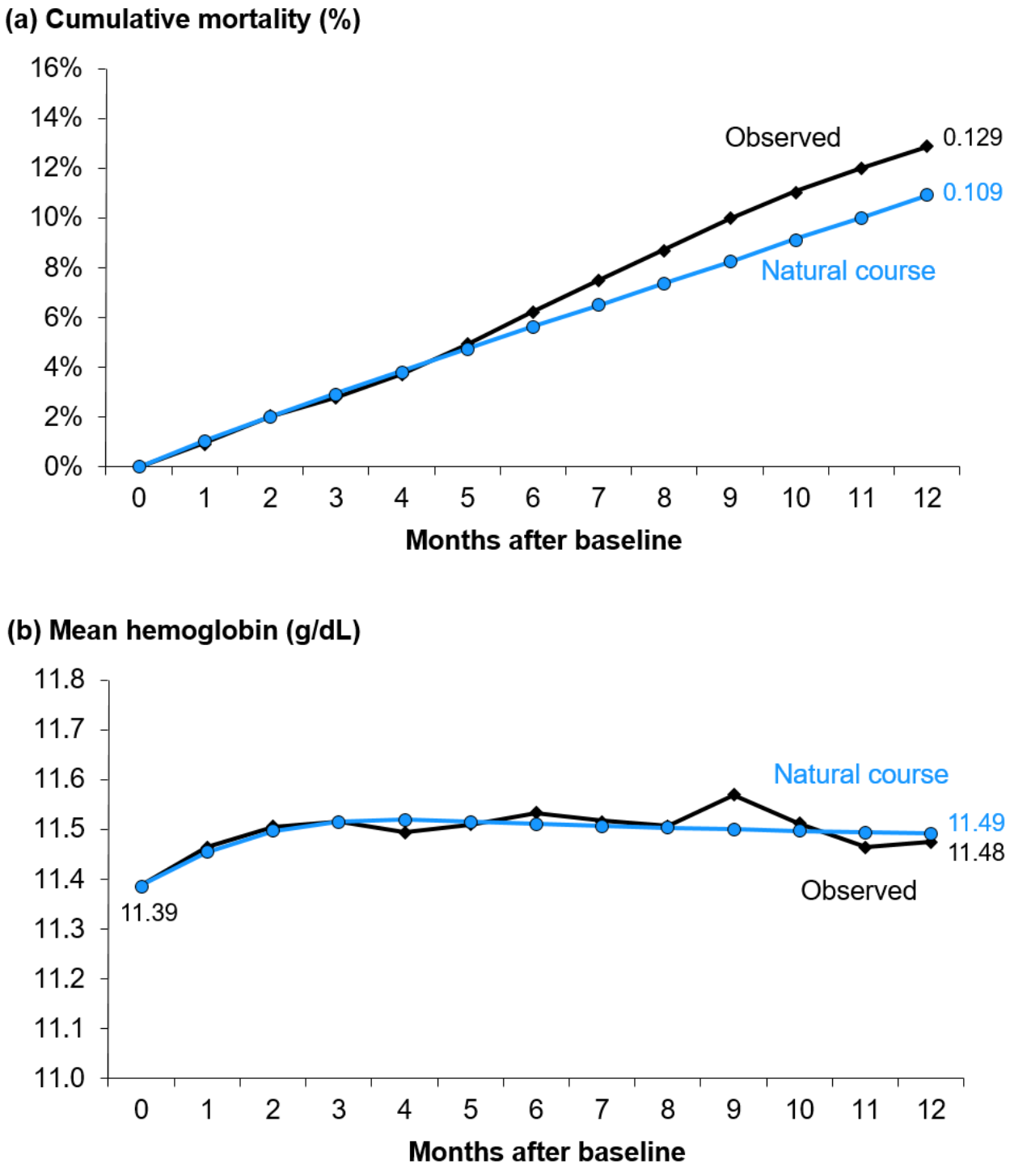
Other cardiovascular disease	31%	28%	--	--
Cancer (non-skin)	17%	17%	--	--
Diabetes	36%	42%	45%	44%
Hepatitis B or C	5%	0%	0%	0%
Gastrointestinal bleeding	5%	6%	--	--
Hypertension	87%	89%	74%	72%
Lung disease	14%	14%	--	--
Neurologic disease	12%	11%	--	--
Psychiatric disorder	17%	14%	--	--
Peripheral vascular disease	30%	28%	8%	9%
Recurrent cellulitis, gangrene	9%	7%	--	--

Mean ± standard deviation, median (IQR), or % shown; Median ESA dose restricted to users; PIVOTAL trial data derived from Table 1 in Macdougall et al.,[65] with variables shown as "--" if not reported; PIVOTAL-restricted DOPPS patients are a subset of the patients included in the primary analysis, but further restricted to emulate PIVOTAL exclusion criteria.

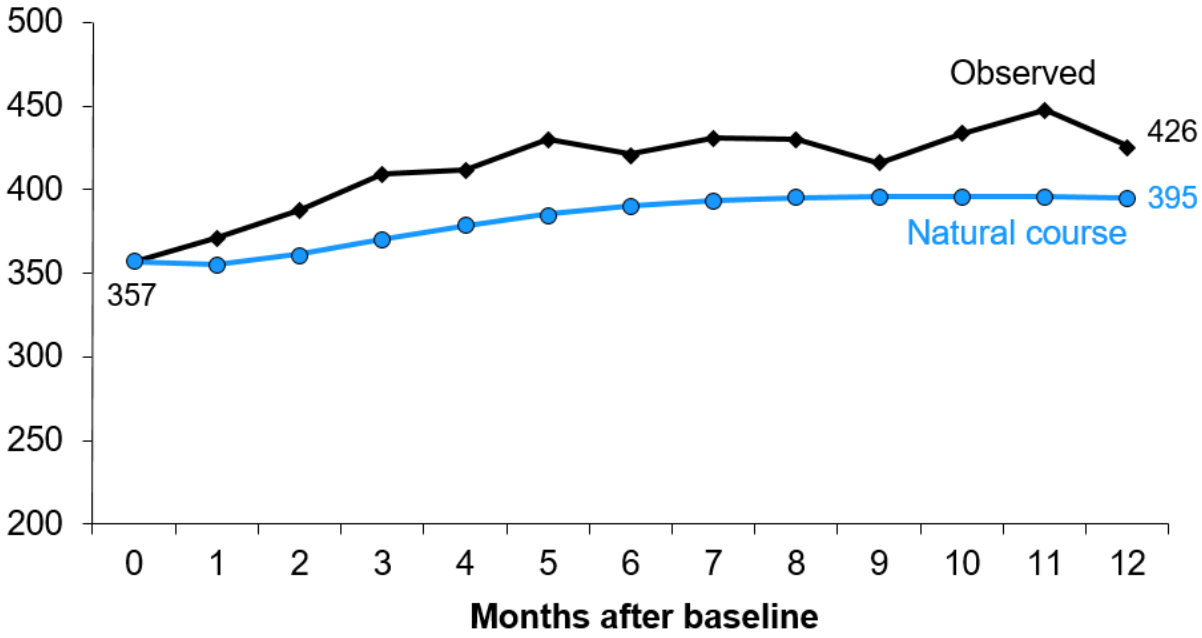
Natural course vs. Observed data

For the 6325 DOPPS patients included in our primary analysis, we first compared observed data (i.e., mean or median levels for up to 12 months of DOPPS follow-up) with our natural course simulation. Trends across the 12 months for key variables were generally similar for the observed data vs. natural course (**Figure 11**).

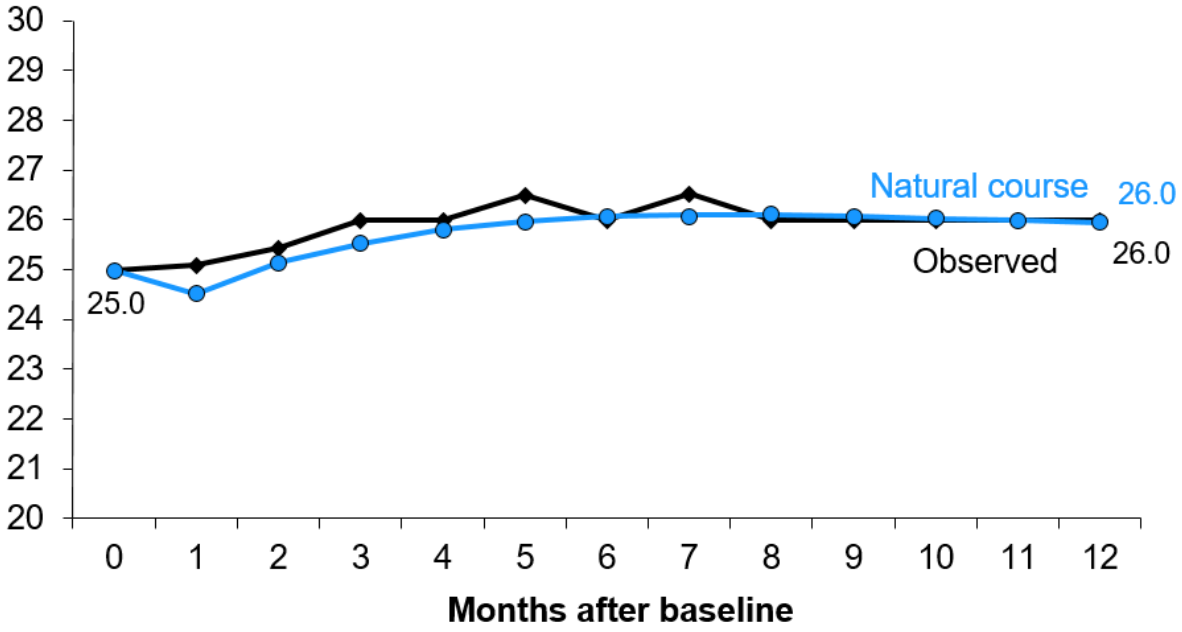
Figure 11. Comparison of observed DOPPS data vs. natural course simulation



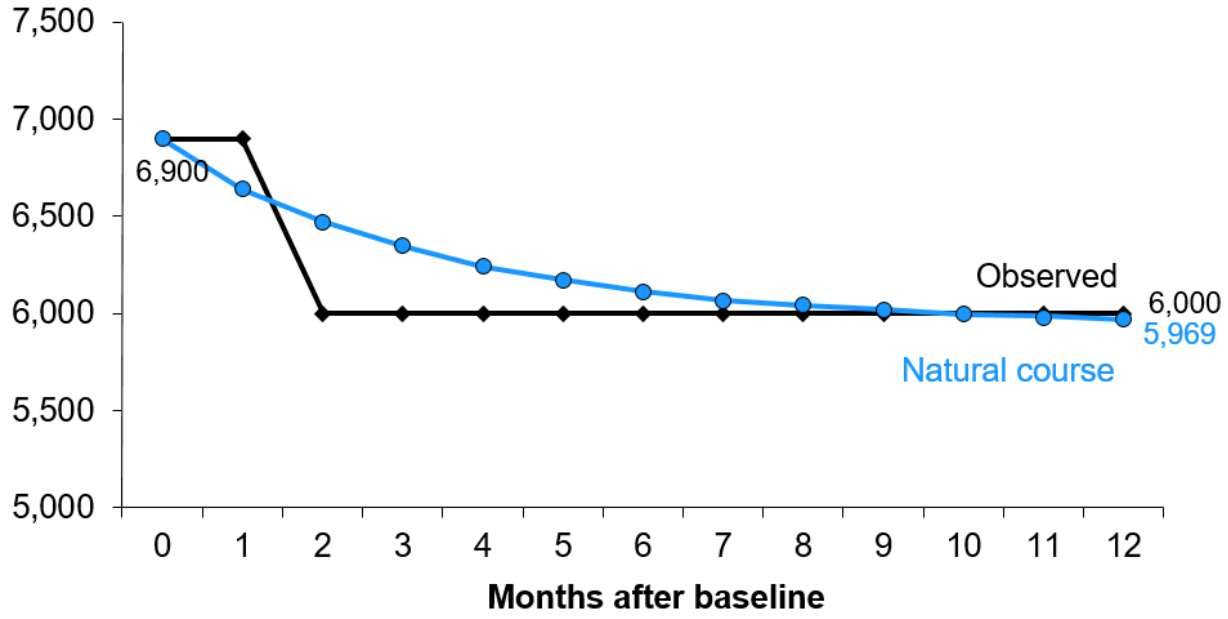
(c) Median ferritin (ng/mL)



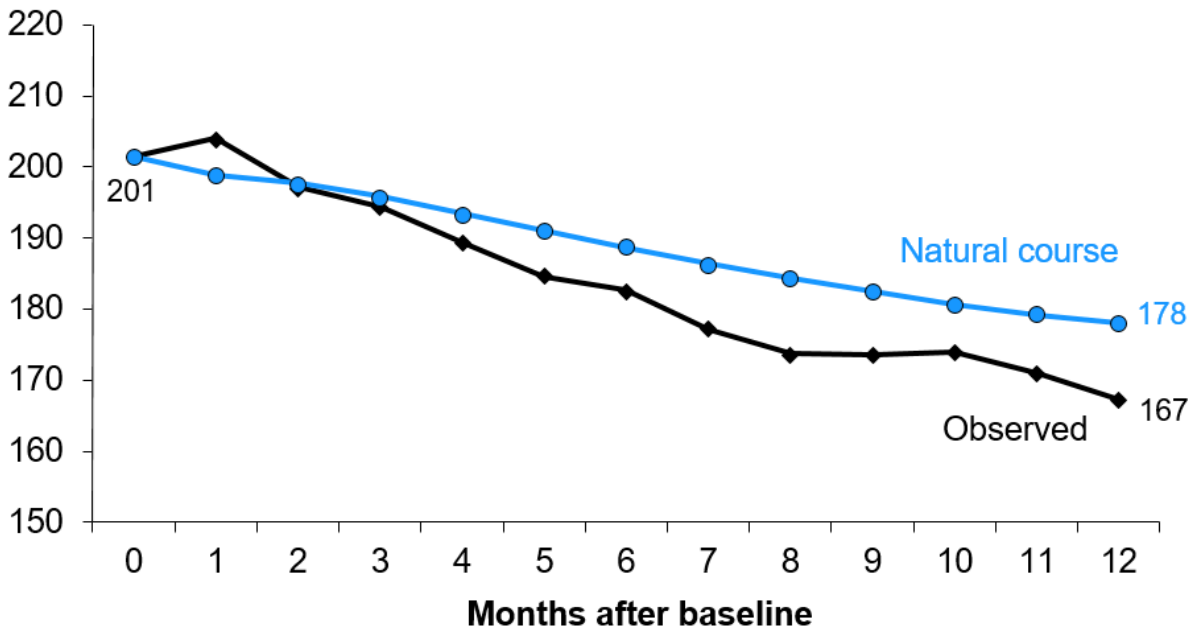
(d) Median TSAT (%)



(e) Median ESA dose (units/week)



(f) Mean IV iron dose (mg/month)

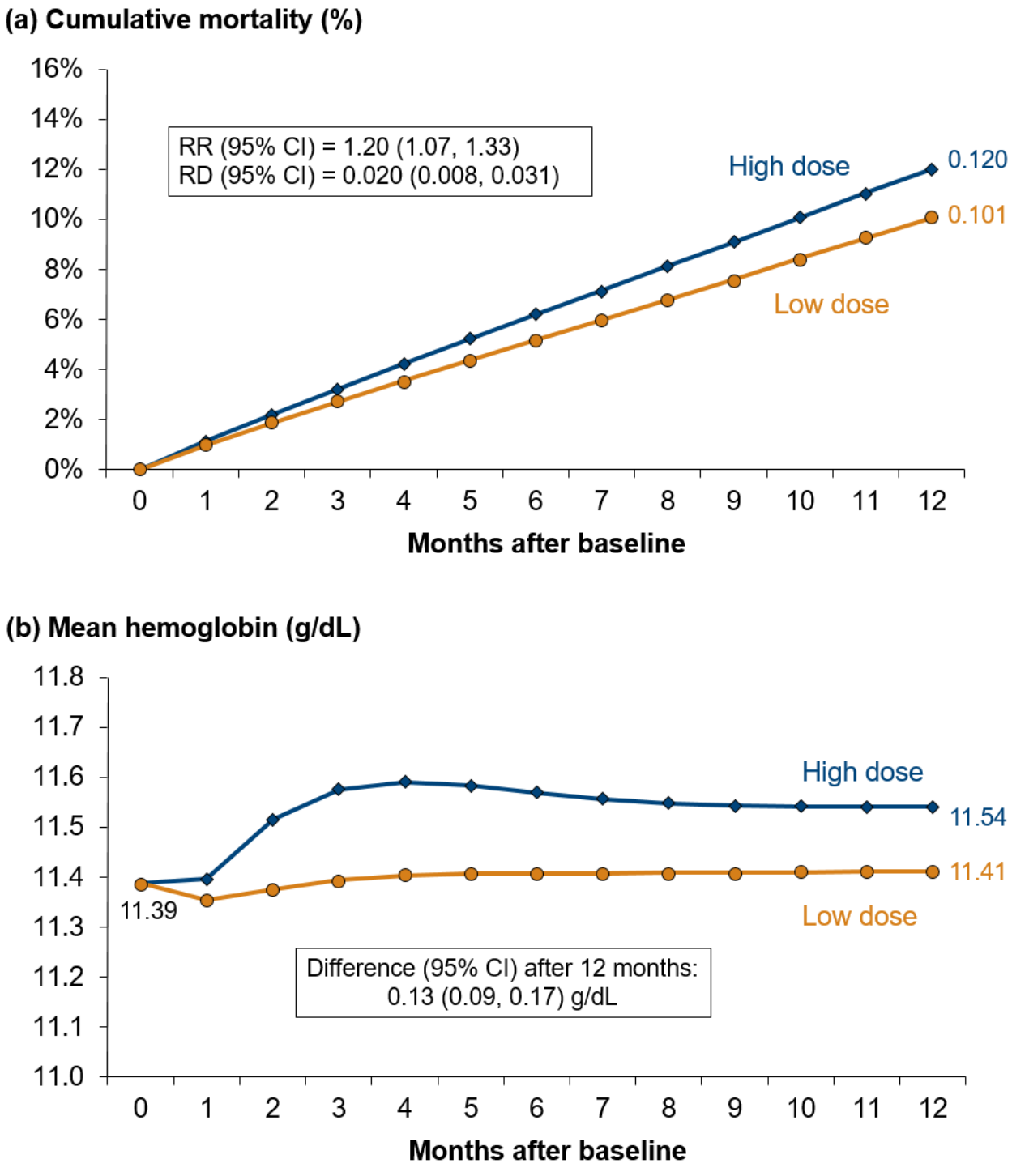


DOPPS data observed over the 12 months after baseline; Natural course simulation based on 12 simulated months using the parametric g-formula; Outcomes: (a) all-cause mortality, (b) hemoglobin, (c) serum ferritin, (d) TSAT, (e) ESA dose, (f) IV iron dose.

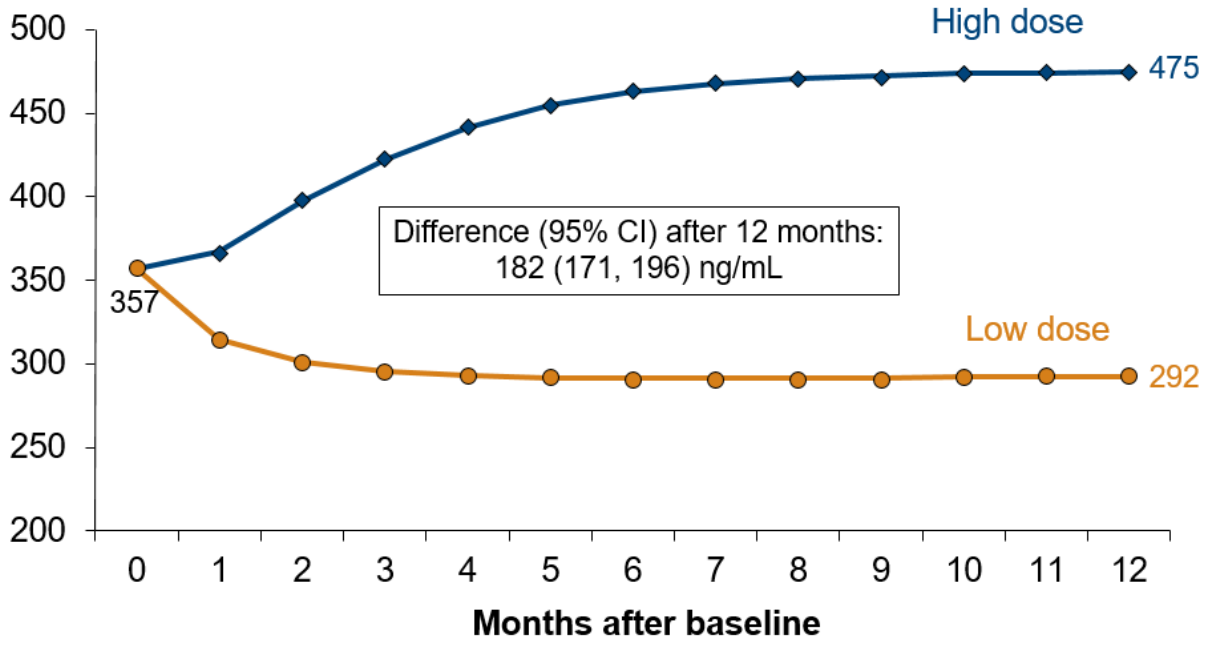
Comparing simulated interventions

The 1-year mortality risk in parametric g-formula simulations was 0.120 vs. 0.101 under the high vs. low IV iron dose simulated interventions (**Figure 12a**); the corresponding RR was 1.20 (95% CI: 1.07, 1.33), and the RD was 0.020 (95% CI: 0.008, 0.031). Differences in secondary outcomes under the two interventions over the 12-month simulation were as follows: Mean hemoglobin was 0.13 (95% CI: 0.09, 0.17) g/dL higher for the high vs. low dose strategy (**Figure 12b**). Median ferritin was 357 ng/mL at baseline and increased to 475 ng/mL under the high-dose strategy while decreasing to 292 ng/mL under the low-dose strategy, a difference at 12 months of 182 (95% CI: 171, 196) ng/mL (**Figure 12c**). Median TSAT was 25% and decreased slightly to 23.9% under the low-dose strategy, and gradually increased to 27.5% under the high-dose strategy, a difference of 3.6% (95% CI: 3.2%, 4.0%) (**Figure 12d**). Median ESA dose was 506 (95% CI: 287, 718) units/week lower (6.7% lower) under the high vs. low dose strategy at 12 months (**Figure 12e**). Mean assigned IV iron dose (including 0 doses) was much greater under the high vs. low IV iron dose strategy (253 vs. 80 mg/month) at 12 months (**Figure 12f**). Comparing cumulative dosing over the 12-month period, patients assigned to the high vs. low dose strategy received 5.8% (95% CI: 3.3%, 8.1%) less ESA and three times as much IV iron (3166 vs. 981 mg) (**Figure 13**).

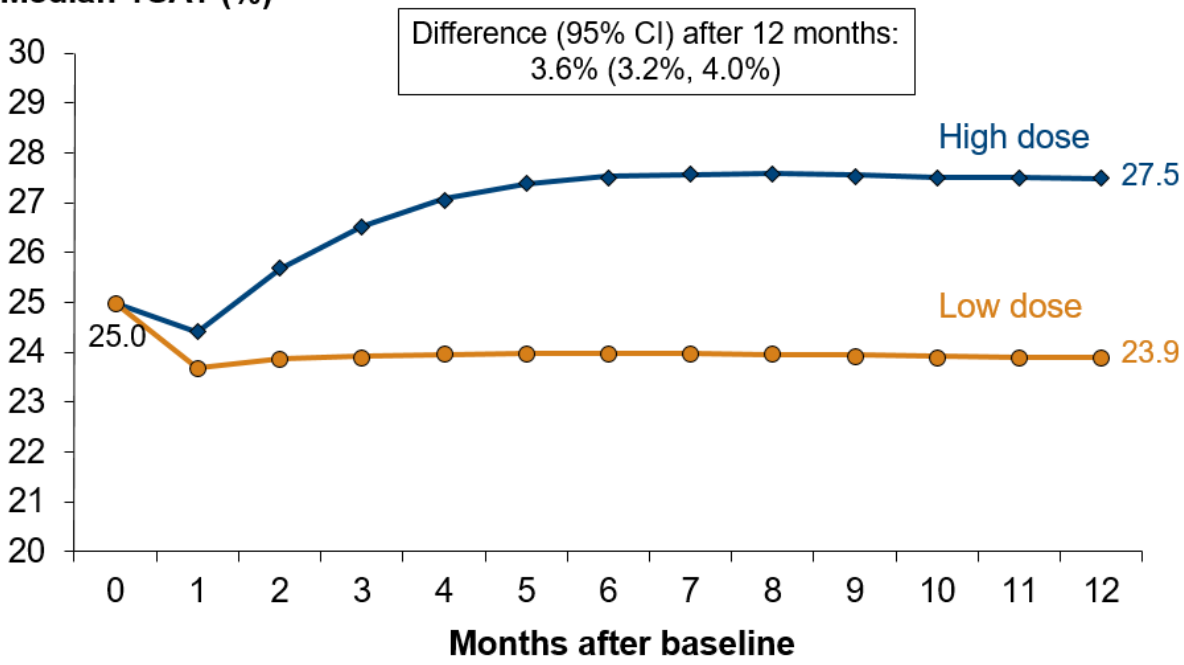
Figure 12. Comparison of proactive high-dose vs. reactive low-dose IV iron treatment strategy over 12 months using the parametric g-formula



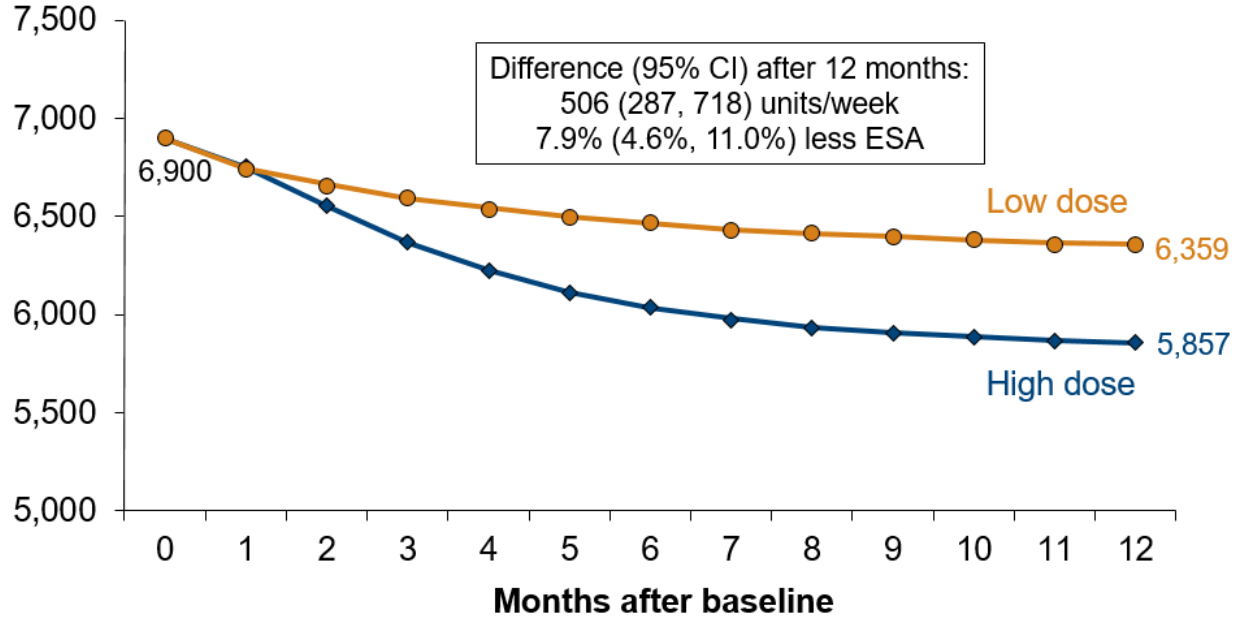
(c) Median ferritin (ng/mL)



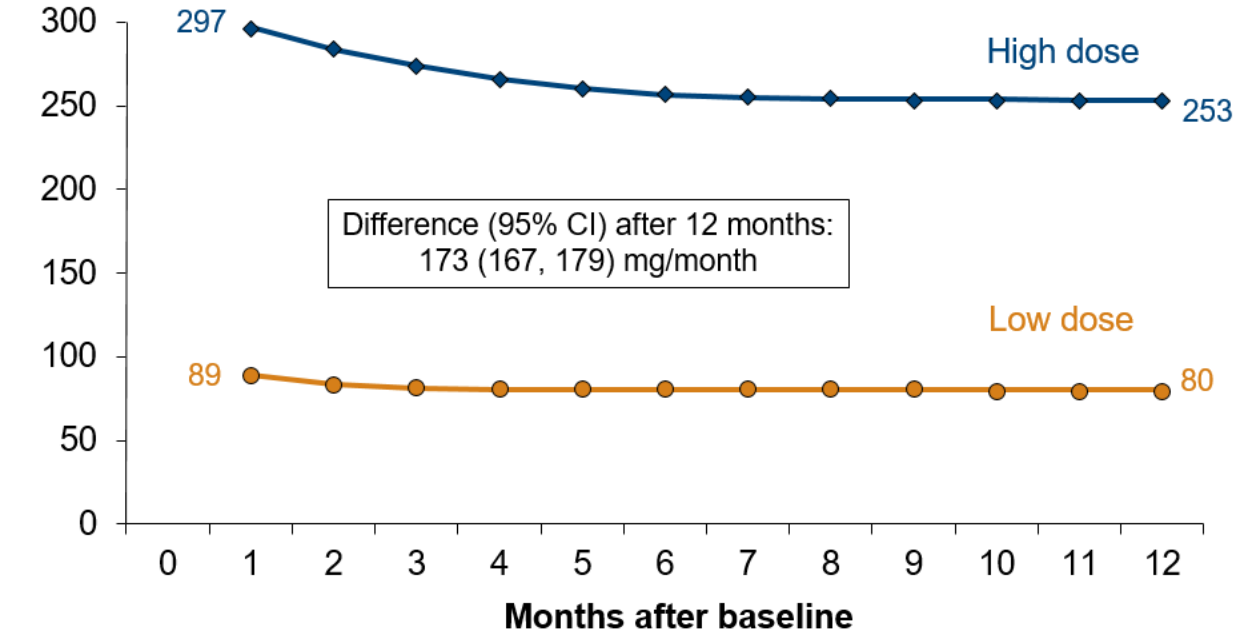
(d) Median TSAT (%)



(e) Median ESA dose (units/week)

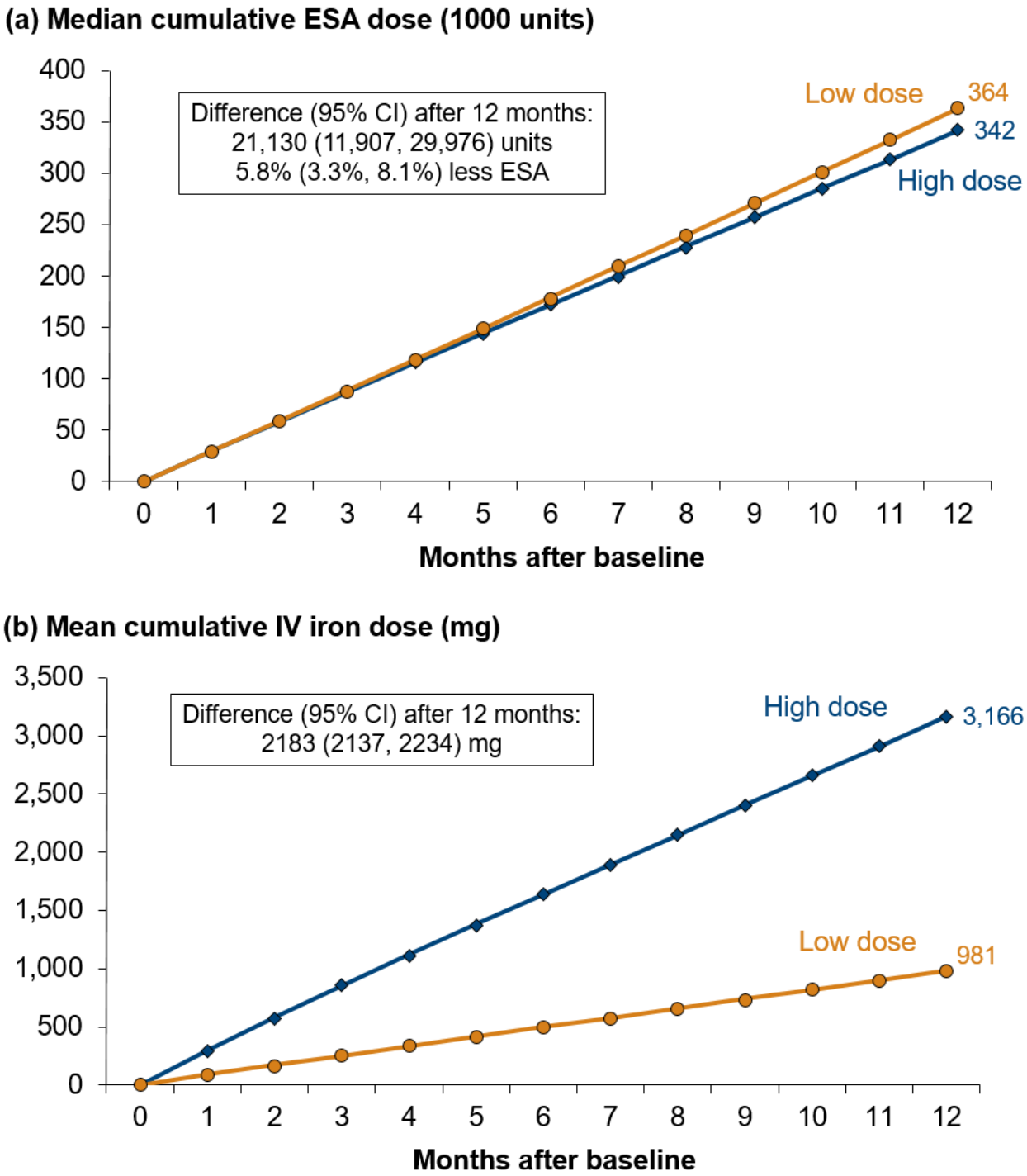


(f) Mean IV iron dose (mg/month)



RR=Risk ratio; RD=Risk difference; High dose and low dose strategies defined by PIVOTAL[65] trial protocol as described in Table 8; Outcomes: (a) all-cause mortality, (b) hemoglobin, (c) serum ferritin, (d) TSAT, (e) ESA dose, (f) IV iron dose.

Figure 13. Proactive high-dose vs. reactive low-dose IV iron treatment strategy: Comparison of cumulative doses of ESA and IV iron over 12 months using the parametric g-formula



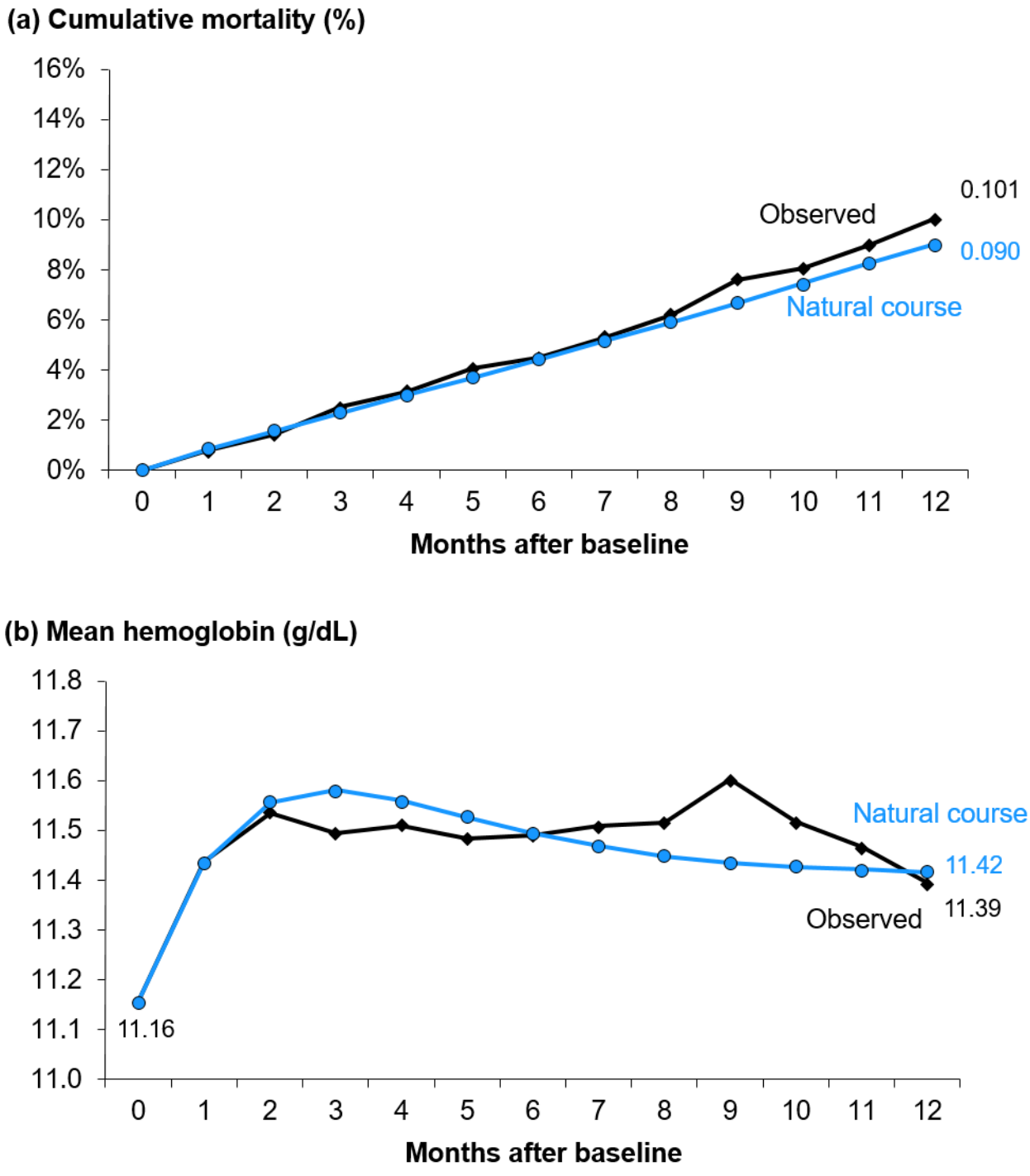
High dose and low dose strategies defined by PIVOTAL[65] trial protocol as described in Table 8; Outcomes: (a) cumulative ESA dose; (b) cumulative IV iron dose

Restricting to a PIVOTAL-like subset

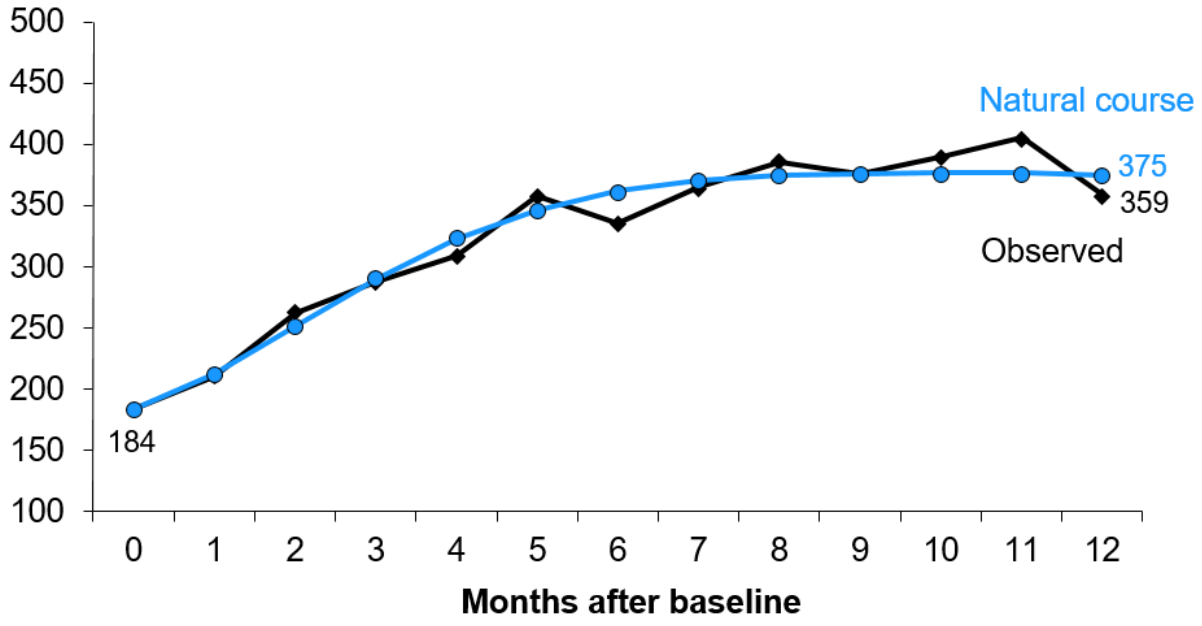
In our second objective attempting to emulate the PIVOTAL population by further restriction of the DOPPS data as illustrated in **Figure 10**, our sample size was reduced from 6325 to 1508 patients. In this subset, we found no major departures from the observed data in our natural-course simulation (**Figure 14**).

The 1-year mortality risk was 0.098 vs. 0.083 under the high vs. low IV iron dose simulated interventions (**Figure 15a**); the corresponding RR and RD was 1.19 (95% CI: 0.84, 1.59) and 0.015 (95% CI: -0.015, 0.041) – very similar to the primary result (**Figure 12a**), albeit with less precision. Baseline levels of hemoglobin, ferritin, and TSAT were much lower in this subset compared to the primary analysis (**Table 10**); subsequent rises are illustrated under both treatment strategies – though more pronounced under the high-dose strategy – and after 12 months, the differences between strategies (**Figure 15b-d**) were comparable to those observed in the primary analysis (**Figure 12b-d**). Median ESA dose was higher at baseline in this subset; doses declined under both treatment strategies over the 12-month period but more precipitously under the high-dose strategy (**Figure 15e**). Mean assigned IV iron dose also started higher in this subset – per protocol to more proactively treat the lower ferritin and TSAT levels – and then eventually reached a steady-state with doses under the two strategies (**Figure 15f**) similar to the primary analysis (**Figure 12f**).

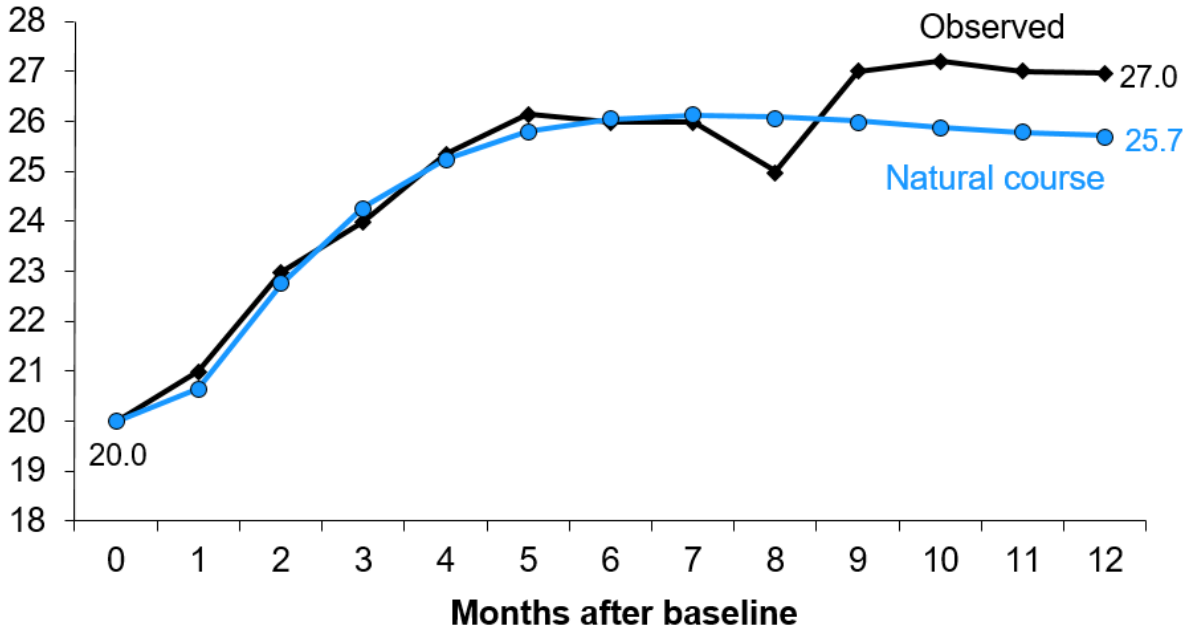
Figure 14. Comparison of observed DOPPS data vs. natural course simulation, restricted to PIVOTAL-like patients



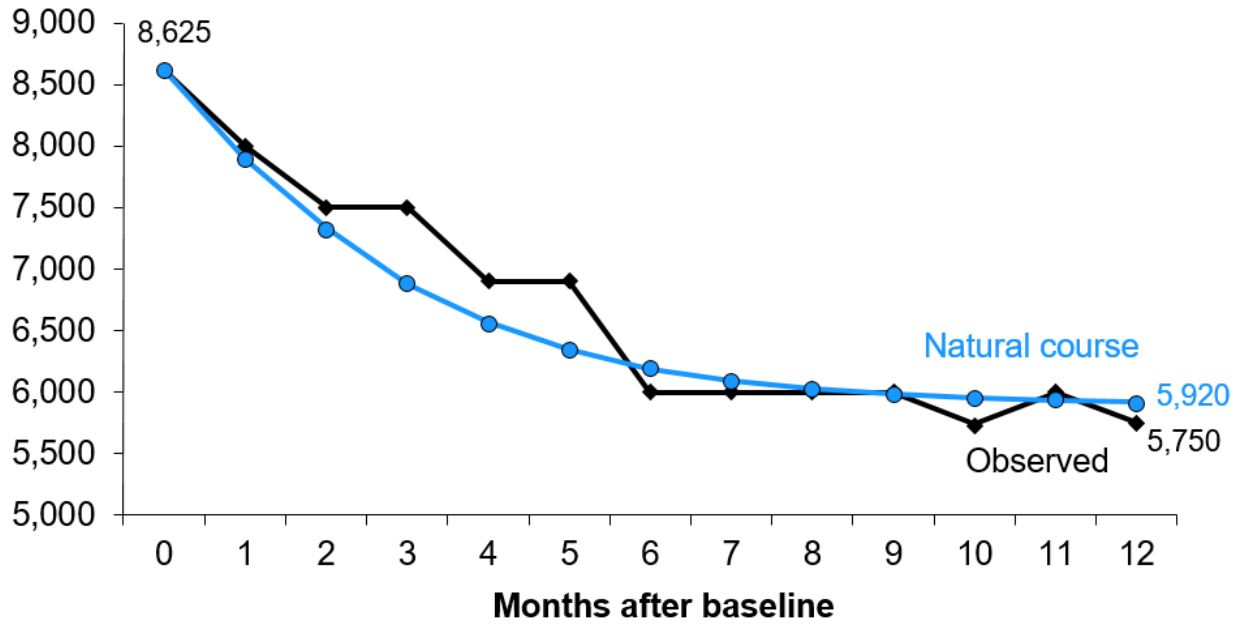
(c) Median ferritin (ng/mL)



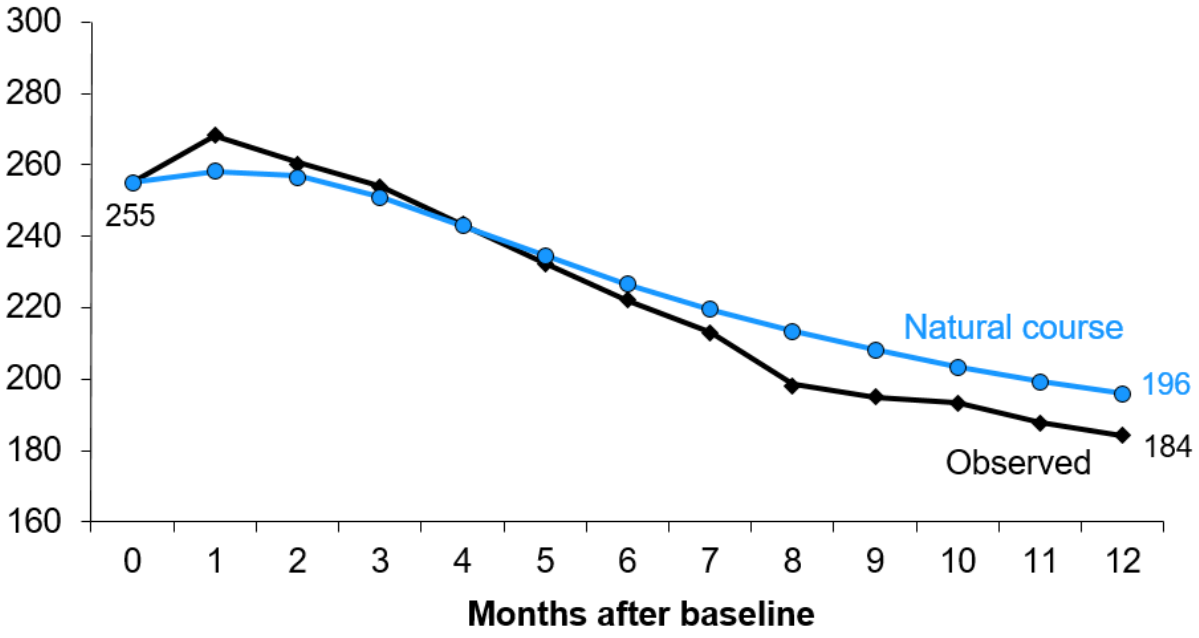
(d) Median TSAT (%)



(e) Median ESA dose (units/week)

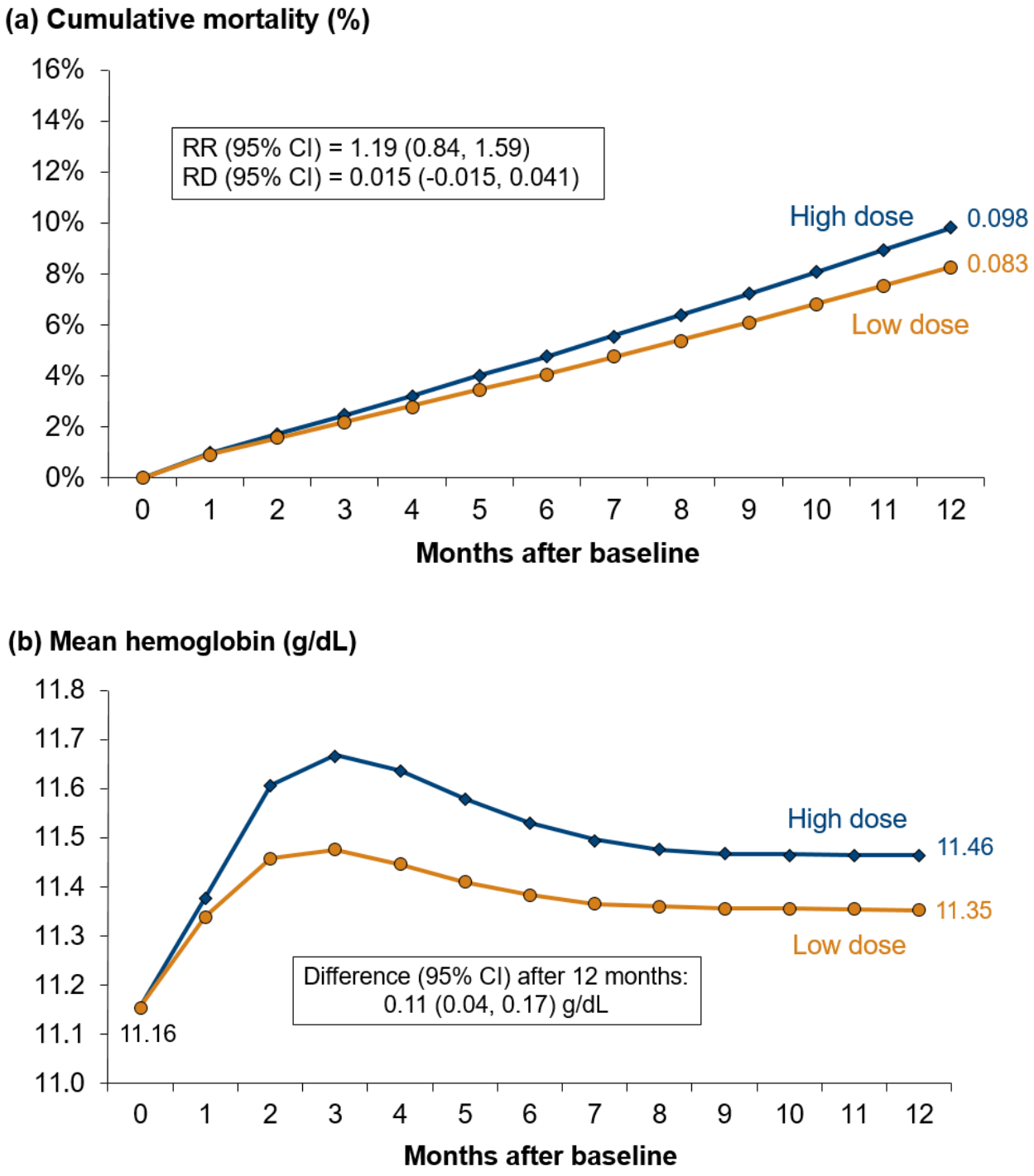


(f) Mean IV iron dose (mg/month)

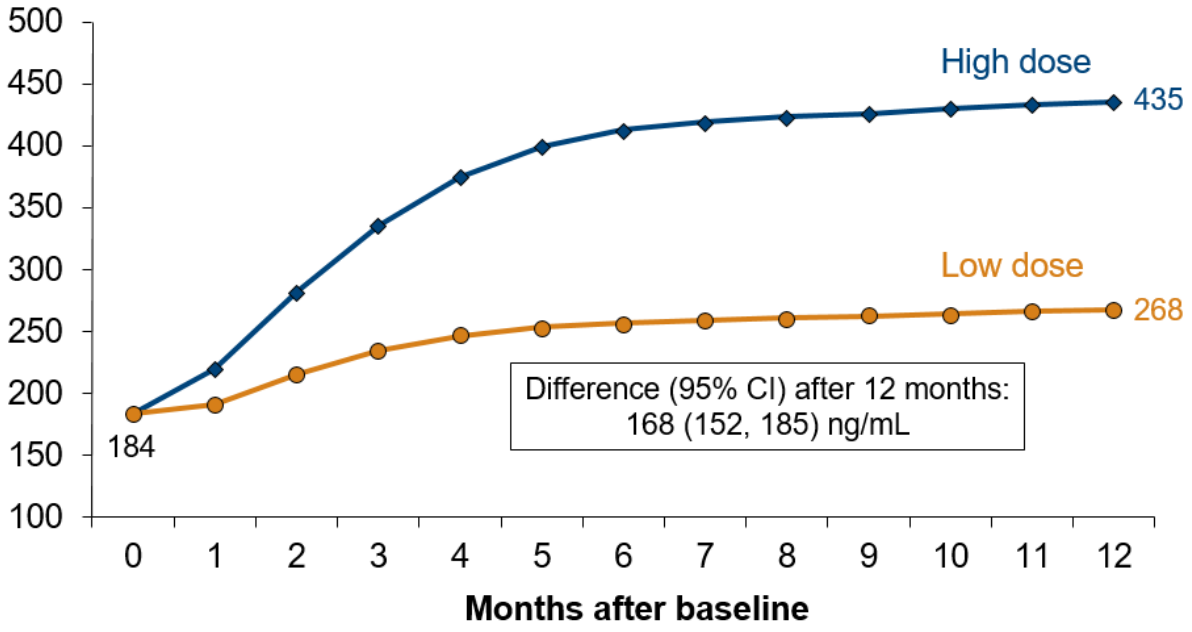


DOPPS data observed over the 12 months after baseline; Natural course simulation based on 12 simulated months using the parametric g-formula; N=1508 PIVOTAL-like DOPPS patients restricted to emulate PIVOTAL exclusion criteria; Outcomes: (a) all-cause mortality, (b) hemoglobin, (c) serum ferritin, (d) TSAT, (e) ESA dose, (f) IV iron dose

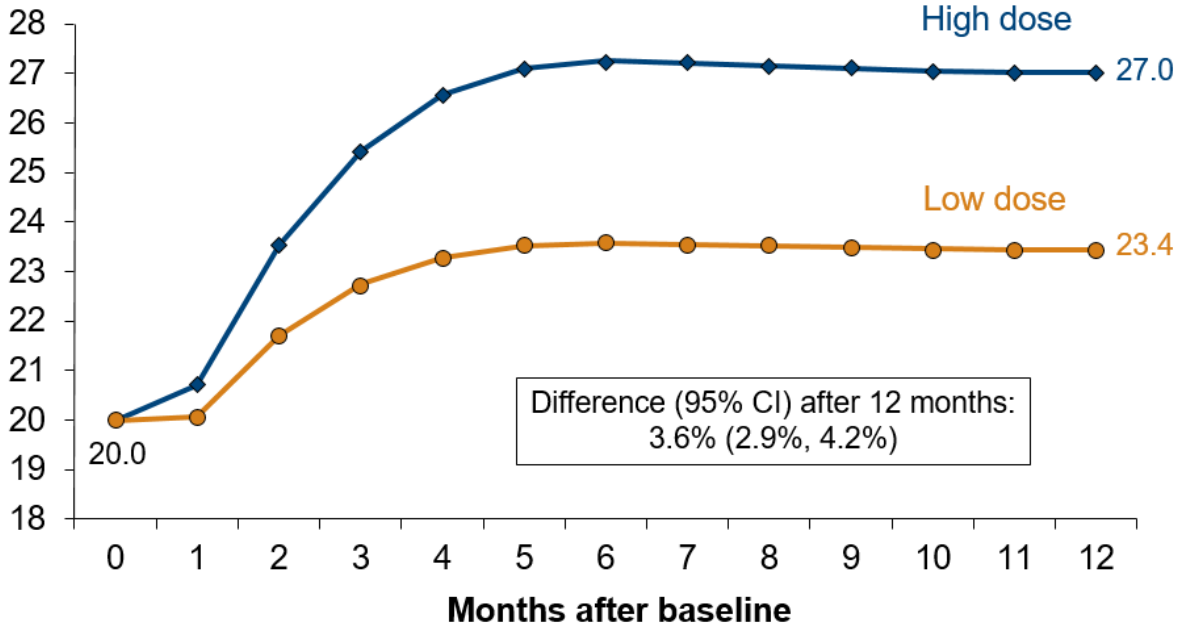
Figure 15. Comparison of proactive high-dose vs. reactive low-dose IV iron treatment strategy over 12 months using the parametric g-formula, restricted to PIVOTAL-like patients



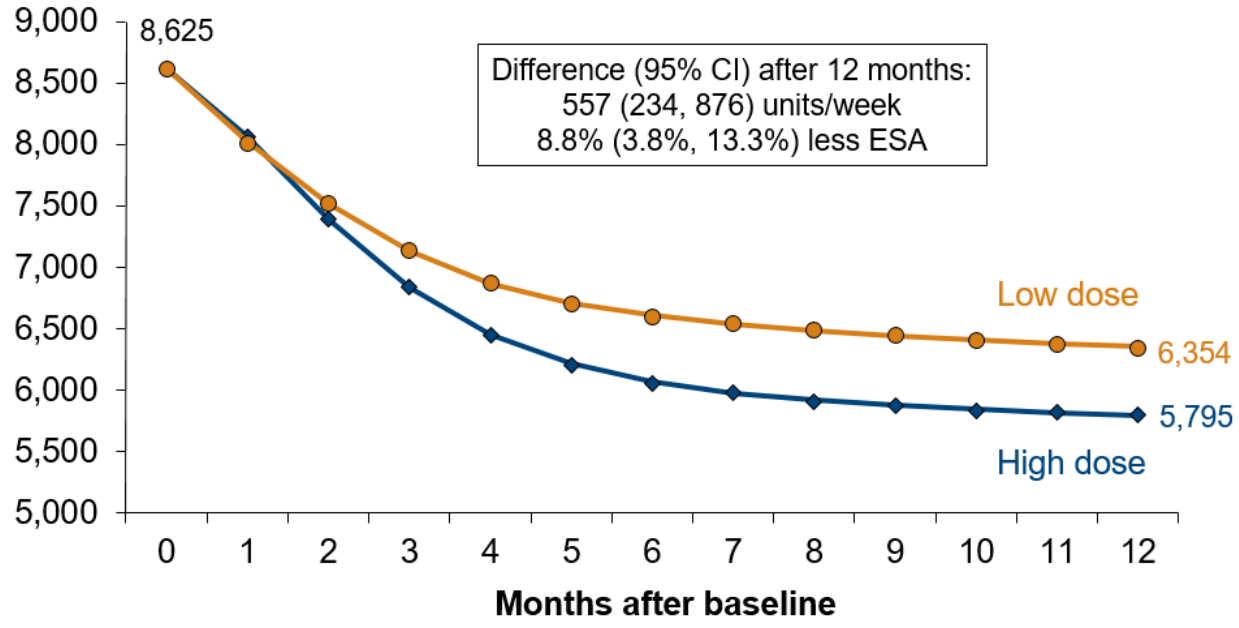
(c) Median ferritin (ng/mL)



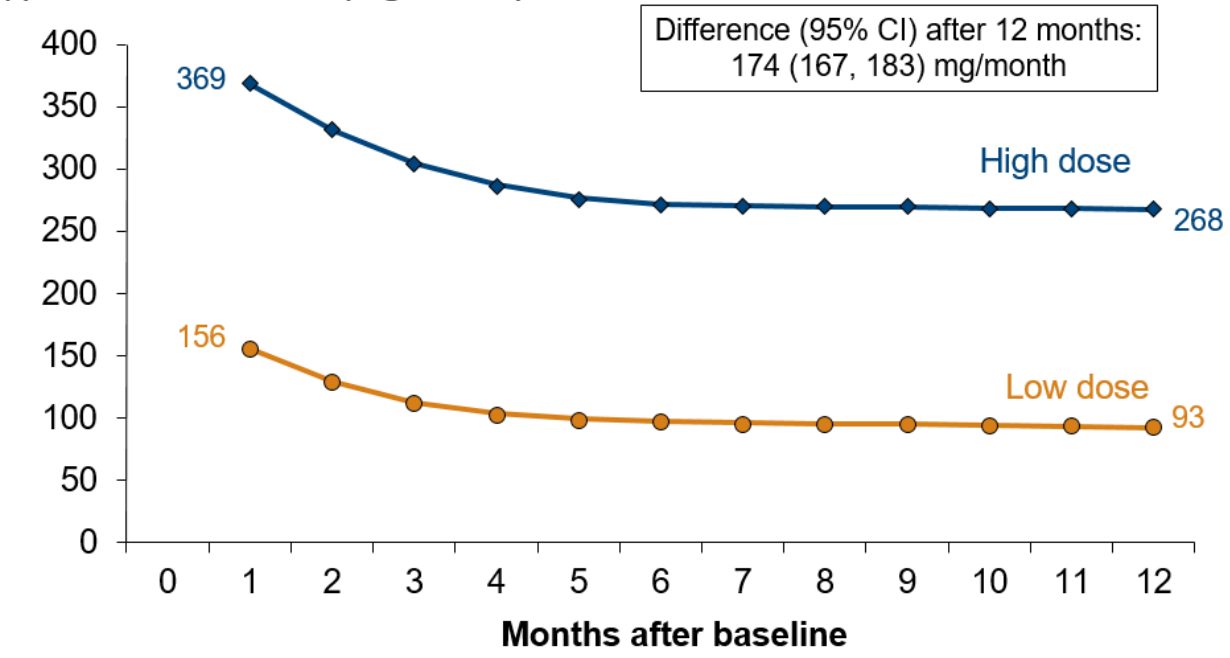
(d) Median TSAT (%)



(e) Median ESA dose (units/week)



(f) Mean IV iron dose (mg/month)



RR=Risk ratio; RD=Risk difference; High dose and low dose strategies defined by PIVOTAL[65] trial protocol as described in Figure 10; N=1508 PIVOTAL-like DOPPS patients restricted to emulate PIVOTAL exclusion criteria; Outcomes: (a) all-cause mortality, (b) hemoglobin, (c) serum ferritin, (d) TSAT, (e) ESA dose, (f) IV iron dose.

Comparisons with PIVOTAL

Table 11 summarizes our parametric g-formula results (primary and restricted) in comparison to the PIVOTAL randomized trial. The 1-year mortality risk was about 0.08 in both PIVOTAL arms, whereas we observed a risk difference of 0.019 (primary) and 0.015 (restricted) under the high vs. low dose simulation. After 12 months, the difference in the mean cumulative IV iron dose assigned under the high vs. low-dose strategy was ~2000 mg in the PIVOTAL trial and in both our simulations. We found that median cumulative ESA dose was 20,000-30,000 units lower under the high vs. low dose strategy after 12 months; this difference was smaller than the 90,000 units lower median cumulative ESA dose reported in the PIVOTAL trial. Similarly, differences in laboratory values after 12 months under the high vs. low-dose strategy in the full DOPPS simulations (0.13 g/dL higher mean hemoglobin, 183 ng/mL higher median ferritin, 3.6% higher median TSAT) were directionally consistent with PIVOTAL findings, but smaller in magnitude (as estimated from Figure S5, S7, S8 in Macdougall et al.[65]: ~0.2 g/dL higher mean hemoglobin, ~450 ng/mL higher median ferritin, ~7% higher median TSAT).

Table 11. Summary of findings: Comparing PIVOTAL trial with DOPPS simulation

	PIVOTAL trial: Randomized results			DOPPS simulation: Full sample			DOPPS simulation: PIVOTAL-Restricted		
Outcomes	**High-dose**	**Low-dose**	**Difference**	**High-dose**	**Low-dose**	**Difference**	**High-dose**	**Low-dose**	**Difference**
N patients	1093	1048	--	6325	6325	--	1508	1508	--
Laboratory values at baseline									
Mean hemoglobin (g/dL)	10.6	10.5	--	11.39	11.39	--	11.16	11.16	--
Median ferritin (ng/mL)	214	217	--	357	357	--	184	184	--
Median TSAT (%)	20.0	20.0	--	25.0	25.0	--	20.0	20.0	--
Laboratory values after 12 months*									
Mean hemoglobin (g/dL)	11.1	10.9	0.2	11.54	11.41	0.13	11.46	11.35	0.11
Median ferritin (ng/mL)	580	130	450	475	292	183	435	268	167
Median TSAT (%)	26	19	7	27.5	23.9	3.6	27.0	23.4	3.6
Cumulative dose through 12 months*									
Median ESA dose (100K units)	380	470	-90	342	364	-22	353	379	-26
Mean IV iron dose (mg)	3800	1800	2000	3166	981	2185	3460	1267	2193
All-cause mortality*									
1-year cumulative risk	0.08	0.08	0	0.120	0.101	0.019	0.098	0.083	0.015

PIVOTAL trial data derived from Macdougall et al.;[65] *indicates numbers were approximated from figures; PIVOTAL-restricted DOPPS patients are a subset of the patients included in the primary analysis, but further restricted to emulate PIVOTAL exclusion criteria.

DISCUSSION

In the DOPPS cohort of hemodialysis patients, we implemented the parametric g-formula to compare patient outcomes under two simulated IV iron treatment regimens defined by the protocol used in the recently published PIVOTAL randomized trial.[65] We found that after 12 months, the proactive high-dose vs. reactive low-dose strategy resulted in much higher serum ferritin levels, slightly higher levels of hemoglobin and TSAT, and slightly lower ESA doses, but a higher risk of mortality. Thus, our simulated findings in both the main and restricted analyses (**Figures 12 and 15**) do not suggest a preference for the proactive high IV iron dose in hemodialysis patients.

These findings, summarized in **Table 11**, were directionally consistent with PIVOTAL results with the critical exception of all-cause mortality. PIVOTAL authors observed a hazard ratio (HR) of 0.85 (95% CI: 0.73, 1.00) for their primary composite outcome over the full 42-month follow-up period for the high-dose vs. low-dose arm, although Figure 2b in Macdougall et al.[65] appears to show no difference (HR~1) in all-cause mortality after the first 12 months of follow-up. Our simulated differences in laboratory values after 12 months under the high vs. low-dose strategy were directionally consistent with the PIVOTAL trial, but smaller in magnitude.

One possibility as to why our results did not match PIVOTAL more closely, in terms of both all-cause mortality and secondary outcomes, is that incident hemodialysis patients could be immediately randomized to a treatment protocol in the PIVOTAL trial, while the parametric g-formula requires 2 previous months of data to inform the models and simulations; this functionally limits us to patients with 3+ months on hemodialysis therapy, after low hemoglobin levels are likely to have been mostly corrected.[144] If

anemia treatments provide an initial boost to levels of hemoglobin, ferritin, and TSAT in previously untreated incident hemodialysis patients that dissipates once patients enter more of a steady-state, this may help explain why the effect sizes we found for these laboratory measures were smaller than in the PIVOTAL trial. Another possibility is model misspecification in Step 1. Similarities in the trajectories of our natural course simulation vs. the observed data (**Figure 11**) were encouraging, as any departures may signify potential model misspecification. However, if we are consistently underestimating the effect of IV iron on intermediate outcomes (i.e., hemoglobin, ferritin, TSAT), any biases in Step 1 models may affect predictions of ESA dose and mortality risk in the Step 2 simulation. We adjusted for numerous time-fixed and time-updated confounders, but we acknowledge the potential remains for residual confounding due to unmeasured or misspecified risk factors. A third possibility is that IV iron may have different effects on iron measures and survival in the generally healthier patients selected for the trial.[139] There were clear differences in the DOPPS cohort vs. PIVOTAL participants, many of which remained even after we attempted to restrict our sample to PIVOTAL-like patients (**Table 10**). While we were able to restrict on unambiguous lab cut-offs (e.g., ferritin <400 ng/mL), we were limited in our ability to restrict on other more subjective criteria (e.g., “life expectancy <12 months per the judgement of the investigator”).

To address the generalizability of the PIVOTAL trial to a broader target population, we performed the g-formula analysis in two samples. In the subset analysis, we attempted to mimic the PIVOTAL trial exclusion criteria as closely as possible to demonstrate the proof-of-concept. While results were mixed, we also simulated the comparison of strategies in a more representative sample of DOPPS hemodialysis patients in Europe.

This latter analysis is more relevant for generalizing PIVOTAL findings and was thus considered the primary analysis because the IV iron treatment strategies evaluated in PIVOTAL would in practice be implemented across all hemodialysis patients.

Some observational studies found that higher IV iron doses were associated with elevated risk of adverse events,[48,62,134] and some did not.[63,64,91] However, all of these studies considered IV iron as a static rather than dynamic treatment strategy; thus, we cannot quantitatively compare those effect estimates to ours. Most of these studies were conducted in the US, where ferritin levels are much higher than in Europe.[136] In PIVOTAL,[65] the ferritin threshold at which to discontinue IV iron in the proactive high-dose arm was 700 ng/mL, lower than the median value observed in the US,[29] limiting generalizability of our analysis – and PIVOTAL itself – regarding optimal treatment for patients with ferritin >700 ng/mL.

A key strength of our study is that utilization of the parametric g-formula allowed us to emulate a "target trial"[133] that compares *well-defined* dynamic treatment strategies. Rather than ask whether patients who received >400 vs. 200-399 mg of iron over a specified time period had better outcomes, our research question is more consistent with the complexities of clinical practice, where IV iron prescriptions depend on time-updated ferritin and TSAT levels. This study design, in contrast to a randomized trial, is flexible to many potential interventions (e.g., altering the ferritin/TSAT criteria) and inclusion criteria. Second, this method properly accounts for a treatment-confounder feedback loop (e.g., ferritin → IV iron → ferritin),[61] but without the possibility that unstable weights will drive results, as with inverse probability weighting methods.[132,142,143] Third, using a European cohort has two advantages: (1) we were able

to adjust for CRP, a marker of inflammation with a strong positive association with both ferritin and mortality that is not routinely measured in the US;[47] and (2) we avoided violations of positivity (when certain subgroups always or never receive the treatment), which would occur in other regions where IV iron dosing strategies are either more aggressive than the high-dose arm (US) or more conservative than the low-dose arm (Japan).[136] Finally, while a small sample size can be augmented in Step 2,[138] our large sample allows for improved precision of the Step 1 coefficient estimation.

Our study had some limitations shared by all parametric g-formula analyses. First, the parametric g-formula can account for time-dependent confounders, but only to the extent they are measured accurately. Second, under the "g-null paradox,"[132] we may still observe an association seemingly due to a treatment effect when the causal null hypothesis is true, given a large enough sample size; however, there is no evidence this occurs in practice.[145] Lastly, reliance on many parametric models creates more opportunity for bias, as misspecification in one model may reverberate throughout the simulation.

We could not replicate all PIVOTAL findings, and the following limitations and obstacles to using large databases to mimic randomized trials should be appreciated. We were unable to narrow our cohort to a PIVOTAL-like population through restriction alone, despite attempts to implement the trial exclusion criteria. Our analysis assumed perfect adherence with the treatment strategy; while this is reasonable from one perspective (IV medications routinely and conveniently administered at each hemodialysis treatment 3x/week), participating clinics may not have adhered to the assigned strategies and patients sometimes miss treatment sessions. However, PIVOTAL findings were very

similar when analyzed per-protocol vs. intent to treat,[65] making this limitation unlikely. Some model misspecification likely played a role in the discrepancies. The basic principal of the non-parametric g-formula as an extension of standardization is appealing for many reasons; but standard nonparametric standardization through stratification is not practical in most multivariable datasets, and extension to the parametric g-formula requires extensive modeling that, in practice, is unlikely to fully account for the many unknown associations and interactions between variables.

While we had mixed success with our goal of mimicking the PIVOTAL trial,[65] the ability to evaluate, rather than only speculate, on how closely our simulated results mirrored actual trial results was extremely valuable and distinguishes this study from prior applications of the parametric g-formula. Because the hypothetical target trial we emulated was not identical to the published PIVOTAL trial, we may not necessarily expect the same answer to these slightly different research questions.[133] This gap can be minimized, but will always exist due to the limits of both observational data (and methods) and trial data, which often represent a highly specialized population that may not be generalizable to the target population of interest. The greatest advantage of the target trial approach using the parametric g-formula is the ability to evaluate many variations of complex intervention strategies within a single cohort study. We encountered some obstacles in implementation and other general limitations in using observational data to mimic a randomized trial, including discrepancies in study populations and potential model misspecification. On balance, however, our results were promising and illustrate the potential of the parametric g-formula to efficiently evaluate multiple dynamic treatment strategies across different populations.

CHAPTER V

Conclusion

Overview

There are fewer randomized trials in nephrology than many other medical specialties,[146,147] leading to a weaker evidence base that results in conflicting treatment guidelines.[55] However, this can also be considered an opportunity for high-quality observational research to make an impact; in particular, the g-methods[131] and other epidemiologic designs that mimic a "target trial"[133] can best address causal questions within cohort studies. While observational data certainly has limitations, proper analytic methods can mitigate a great deal of the potential biases. In the hemodialysis (HD) setting, patients typically travel to the HD facility three times per week for four hours. Medications administered intravenously are convenient, promote patient compliance, and are easily recorded by the HD facility. The regular HD schedule also provides a great opportunity for creating and updating vast databases of longitudinal information on patient condition, treatments, and laboratory measurements, making the Dialysis Outcomes and Practice Patterns Study (DOPPS) well-suited to address the three aims of this dissertation.

A key advantage of the DOPPS is the inclusion of international data from three major regions: North America, Europe, and Japan. Anemia management strategies for HD patients vary widely across regions, as illustrated by the descriptive data from Aim 1

(**Figures 3-4**). Both ESA and IV iron doses are highest in North America, followed by Europe, and lowest in Japan, even within strata of hemoglobin, ferritin, and TSAT. Any recommendations should thus be viewed through a region-specific lens. In Aim 2, we observed that acute inflammation, as measured by a rise in C-reactive protein (CRP), was more quickly controlled in Japan than in Europe; while there are opportunities for improvement in Europe, providers in North America must take the first step of routinely measuring CRP. In Aim 3, the contrast between PIVOTAL strategies is mainly relevant in Europe; in Japan, IV iron dosing is more conservative than the PIVOTAL low-dose arm, while in the US, IV iron dosing is more aggressive than the PIVOTAL high-dose arm. Future applications of the g-formula using Japan-only or US-only data would ideally assess region-specific protocols to reflect observed treatment strategies.

Summary of findings: Aim 1

Aim 1 focuses on anemia management strategies before (treated vs. untreated) and after (intensity of treatment) the transition period to HD. The relation between anemia management during the transition period to HD and post-dialysis outcomes is challenging to assess and often requires an innovative study design. Our primary analysis examined the association between hemoglobin levels at HD start and first-year HD mortality, *among patients who had hemoglobin levels ≥10 g/dL after 4 months on HD*. While at first it may be counter-intuitive to make this restriction after the exposure variable is defined, this strategy of measuring month 1 exposure status retrospectively and following patients from month 4 through month 12 allows for a fairer comparison of hemoglobin levels at HD start. Patients with low hemoglobin at HD start due to severe comorbidity or a lack of responsiveness to anemia therapy will likely not experience a

hemoglobin increase to target range (≥10 g/dL) in the subsequent months, while those who did experience an increase to target range would have theoretically been able to start HD with a higher hemoglobin with proper anemia management prior to starting HD. Our study thus includes an exposure variable that should largely reflect differences in pre-dialysis anemia care.

Our results showed that even among patients with hemoglobin ≥10 g/dL four months later, anemia at HD initiation was common and associated with elevated mortality. While we expected to observe a greater mortality rate for patients who initiated HD with low hemoglobin, this elevated rate was still observed among patients achieving hemoglobin ≥10.0 g/dL in the early HD period, a subset of patients whose low hemoglobin levels at HD start could have theoretically been corrected by pre-dialysis treatment. This is consistent with our hypothesis that a more proactive approach to anemia management in advanced CKD may improve survival on HD, though there are other possible explanations for these results. It is possible that intense anemia treatment in the early HD period may be responsible for the elevated mortality rate, though additional adjustment for erythropoiesis-stimulating agent (ESA) and intravenous (IV) iron doses (potential mediators) had minimal impact on the primary result, making this explanation unlikely. It is also possible that the association may be biased due to residual confounding because patients with higher hemoglobin at HD start are generally healthier and/or have likely received better overall quality of care before HD initiation.

Also in this first aim, we compared mortality rates by intensity of anemia treatment approaches in the early dialysis period. This study provides a unique perspective by focusing on patients who start HD with severe anemia (hemoglobin <10 g/dL), a group

that is likely treated based on nephrologist practice patterns and regional guidelines[5,51–53] because transition to a more individualized steady-state ESA dose, when the association between ESA dose and mortality is more likely reflective of confounding factors that drive ESA dose requirements,[104,105] has not yet occurred.

We observed a U-shaped association between IV iron dose over the first three months of HD and mortality, with the lowest mortality rate observed at 200-399 mg/month. Patients initiating HD with lower hemoglobin generally received more IV iron, but this pattern was not observed in the US, where a high median dose of 450-500 mg/month was observed regardless of hemoglobin at HD start. While the recent Proactive IV Iron Therapy in Hemodialysis Patients (PIVOTAL) trial[65] demonstrated the superiority of a high vs. low dose IV iron treatment strategy, the “high” doses were lower than the median doses administered in the US, and the upper ferritin threshold of 700 ng/mL in the “proactive” arm was lower than the median serum ferritin levels observed in the US.[29] Generalizability of the PIVOTAL findings in the context of the high levels of serum ferritin and IV iron dosing observed in the US thus remains an open question.

We also observed elevated mortality for patients receiving the largest doses of ESA (>25,000 units/week) during the first three months of HD. Following patients soon after HD initiation may better capture dosing patterns (practice preferences) before patients transition to a more individualized steady-state dosing protocol, at which point observed associations between ESA dose and mortality are more likely attributable to confounding factors that drive ESA dose requirements than a causal effect.[104,105]

Summary of findings: Aim 2

Aim 2 focuses on hemoglobin response to ESA therapy, and the degree to which it may be blunted by inflammation, as measured by CRP. Rather than attempt to estimate a simple, but potentially biased, cross-sectional association between CRP and ESA responsiveness, we investigated how within-patient changes in inflammation status lead to changes in hemoglobin and ESA dose. This innovative self-matched longitudinal ("before-after") design and analysis improves upon previous between-patient analyses of CRP and ESA resistance by eliminating confounding due to baseline patient characteristics (e.g., age, sex, comorbidity history). We tracked real-world changes in anemia control and ESA dosing in an international sample of HD patients over the 3 months before and after detection of new inflammation by routine CRP measurement. Either greater ESA doses at the same hemoglobin level, or lower hemoglobin levels at the same ESA dose, would be indicative of ESA resistance.

Results showed that patients experiencing new inflammation had *both* higher ESA doses and lower hemoglobin levels, doubly supporting the hypothesis that inflammation increases resistance to ESA treatment and thus exposes patients to the potential cardiovascular risks of larger ESA doses.[17–20] While CRP, and the underlying inflammation, is not easily modifiable, a better understanding of how ESA dose requirements may change in response to inflammation occurrences could be valuable for clinicians. This information may help optimize ESA utilization by achieving hemoglobin targets while using less ESA. These findings also speak to a potentially important role for anemia therapies such as hypoxia-inducible factor prolyl hydroxylase inhibitors (HIF-PHIs) that may be effective despite inflammation.[116–118]

These implications point to the importance of measuring and monitoring inflammation through routine CRP assessment; this is done in most European and Japanese HD centers, but not in North America. Routine CRP measurement is relatively inexpensive and convenient to measure,[124] and can help better identify causes of, and inform targeted strategies to reduce, inflammation. Quicker recognition of new inflammation in HD patients could also help identify the cause of worsening anemia and guide reactive ESA and IV iron dosing decisions.

Additional results and study design characteristics all further support a causal relation between inflammation and ESA hyporesponsiveness. First, we restricted to HD facilities that routinely measured CRP to avoid bias in which a clinical indication for measuring CRP also affects the outcome. Second, the longitudinal study design focuses on incident inflammation to avoid the temporal ambiguity of cross-sectional designs. Third, by matching patients to themselves and measuring outcomes before and after the detection of elevated CRP, this design does not require a comparison group of patients who did not experience an increase in CRP. Fourth, while residual confounding by unmeasured time-varying risk factors for ESA hyporesponsiveness is possible, the likelihood of this bias is low when considering how the self-matching seems to have controlled adequately for both time-fixed and time-varying confounders, as evidenced by minimal changes in the effect estimates after adjustment for time-varying confounders. Finally, the associations were particularly strong among patients whose CRP increase was sustained over the subsequent 3 months, consistent with the hypothesized mechanism.

Summary of findings: Aim 3

Aim 3 focuses on the controversy[37–39,56,59,60,148] regarding IV iron supplementation and upper ferritin targets in the context of the recently published PIVOTAL trial,[65] and more generally investigates how closely observational data can mimic an actual randomized trial using the parametric g-formula. The PIVOTAL trial[65] randomized HD patients to a proactive high dose (400 mg/month; discontinue if ferritin > 700 ng/mL or TSAT >40%) vs. reactive low dose (100-200 mg; discontinue if ferritin <200 ng/mL and TSAT >20%) IV iron strategy. After 42 months of follow-up, the hazard ratio for the composite primary end point of nonfatal myocardial infarction, nonfatal stroke, hospitalization for heart failure, or all-cause mortality in the PIVOTAL trial was 0.85 (95% CI: 0.73, 1.00) for the high vs. low dose arm. PIVOTAL results for all-cause mortality were similar, although no differences between the high and low dose arm were observed during the first year of follow-up. PIVOTAL authors also found that the high-dose group had lower ESA doses and higher levels of hemoglobin, TSAT, and ferritin over the first 12 months.

We performed the parametric g-formula analysis using DOPPS data to mimic two hypothetical target trials. In the primary analysis, we compared PIVOTAL treatment strategies in the full DOPPS sample (i.e., nationally representative samples of HD patients in seven European countries), with the goal of extending PIVOTAL findings outside of the narrow trial inclusion criteria. In the subset analysis, we attempted to simulate the PIVOTAL trial as closely as possible, by implementing the marginal exclusion criteria applied to the PIVOTAL trial,[139] to demonstrate the proof-of-concept. In both analyses, we found that the proactive high-dose vs. reactive low-dose strategy resulted in much higher serum ferritin levels, slightly higher levels of hemoglobin and

TSAT, and slightly lower ESA doses, but a higher risk of mortality. Our simulated differences in laboratory values after 12 months under the high vs. low-dose strategy were directionally consistent with PIVOTAL, but smaller in magnitude. Our mortality findings (RR=1.20 for high vs. low dose strategy), however, were not consistent with PIVOTAL. Potential explanations for these divergent results include model misspecification and/or differences in the study populations.

Results from some observational studies have shown that a higher IV iron dose was associated with an elevated risk of adverse events[48,62,134] and some have not.[63,64,91] However, all of these studies considered IV iron as a static rather than dynamic treatment strategy, and thus we cannot quantitatively compare our effect estimates. Rather than ask whether patients who received >400 vs. 200-399 mg of iron over a specified time period had better outcomes, our research question is more consistent with the complexities of clinical practice, where IV iron prescriptions depend on time-updated ferritin and TSAT levels. Further, most of these other observational studies were conducted in the US, where ferritin levels and cumulative IV iron doses are much higher than in Europe.[136] In the PIVOTAL trial,[65] the ferritin threshold at which to discontinue IV iron in the proactive high-dose arm was 700 ng/mL, lower than the median value observed in the US,[29] limiting generalizability of our analysis – and the PIVOTAL trial itself – regarding optimal treatment for patients with ferritin >700 ng/mL to US populations.

We implemented the parametric g-formula to properly account for the treatment-confounder feedback loop (i.e., ferritin → IV iron → ferritin)[61] that would cause standard analytic methods to fail.[129] Time-dependent confounding is challenging because models

may be biased with or without adjustment for these confounders (because they are also mediators), and information on exposures and key confounders must be measured frequently during follow-up. A marginal structural model was another option to handle this time-dependent confounding, but unstable weights can often drive results when using inverse probability weighting methods;[132,142,143] further, marginal structural models tend to be inefficient when comparing dynamic treatment strategies.[130,149,150] Nonetheless, we acknowledge encountering some obstacles in implementation of the g-formula. This multivariate longitudinal extension of standardization to the parametric g-formula requires extensive modeling that, in practice, is unlikely to fully account for the many unknown associations and interactions between variables. Our stated goal of simulating PIVOTAL[65] findings was also hampered by our inability to accurately duplicate the trial population even after restricting the DOPPS sample based on marginal exclusion criteria applied to the PIVOTAL trial.[139] This ability to evaluate, rather than only speculate, on how closely our simulated results mirrored actual trial results was a strength and distinguishes this study from prior applications of the parametric g-formula.

Future directions

This dissertation research comes at a key, and potentially transitionary, time with respect to anemia management strategy in dialysis. For many years, ESA and IV iron have been the dominant drugs used to reach and maintain target hemoglobin levels in HD patients. New therapies, however, that may compliment or even replace ESA and IV iron have either been recently approved or will be available soon.[57] Rather than receive iron intravenously, iron deficient patients may be treated by various oral iron

formulations, such as iron-containing phosphate binders. These medications may reduce IV iron and/or ESA doses by including iron in a phosphate binder,[151,152] a medication that over 80% of HD patients already require.[29] These new iron formulations, already on the market, are promising and have the potential to lower costs, but should be used cautiously until they prove to be safe and effective.[148] Additionally, a new class of anemia drugs, HIF-PHIs, is currently in phase 3 trials in many countries, with recent approval for use in China. HIF-PHIs stimulate endogenous erythropoietin production while simultaneously coordinating iron bioavailability, thus raising hemoglobin levels.[116,117] Early trials[153,154] have shown HIF-PHIs to be effective at raising hemoglobin without causing serious adverse events, but there are theoretical safety concerns, primarily in tumor growth, that the relatively short follow-up periods may not have been able to capture.[116,117] Oral administration of HIF-PHIs, compared to the injectable administration of all current ESA's, may provide a major advantage in their effectiveness in non-dialysis CKD patients who do not visit the HD facility three times per week, particularly in low- and middle-income countries where access to IV or subcutaneous medications may be limited. Another potential advantage of these drugs is to better prepare patients for the transition from non-dialysis CKD to dialysis; as shown in the Aim 1 research, the consequences of initiating HD with severe anemia are sustained even among patients with hemoglobin treated into target range in the early HD period. Further, while patients tend to become ESA hyporesponsive when inflamed, as shown in the Aim 2 research, HIF-PHIs may remain effective in the presence of inflammation.[155] With HIF-PHIs not yet on the market in most countries, and only very recent uptake of alternative iron supplementation routes, only data on more traditional

anemia therapies, ESA and IV iron, were available for the research questions addressed in this dissertation.

A theme across all three aims was the importance of clearly defining the research question, and choosing the appropriate study design. We encountered challenges along the way but were able to consistently develop appropriate solutions. In **Aim 1**, we attempted to assess the impact of pre-dialysis anemia management while limited to a cohort of HD patients with no information on pre-dialysis treatment history. We developed an approach – that might at first seem counterintuitive and lead to selection bias– of restricting the data based on information four months *after* the exposure variable was defined to better align the analysis with the hypothesis. This restriction was not done after exposure status was measured in the first month of HD; but rather, exposure status was measured retrospectively at HD initiation after eligible patients were selected after four months of HD treatment. In **Aim 2**, we sought to assess the impact of inflammation on ESA hyporesponsiveness, and developed an elegant self-matched longitudinal design and analysis to minimize confounding due to differences in patient characteristics. In **Aim 3**, we attempted to compare dynamic treatment strategies in the presence of potentially severe time-dependent confounding – causing standard methods to fail whether or not the confounders are included in the model– and identified the parametric g-formula, which is known to many epidemiologists but rarely implemented, as the best approach. Using observational data to emulate a "target trial"[133] comparing well-defined dynamic treatment strategies that are consistent with the complexities of clinical practice is very appealing, and becoming more feasible in the era of big data. The advantages are clear – including the ability to evaluate many

variations of complex intervention strategies across differing inclusion criteria within a single cohort study – particularly in settings where randomized trials may be impractical or unethical. However, the limitations to doing so in practice should be appreciated, including residual confounding and discrepancies between the study cohort and target trial populations. When the target trial emulated is not identical to the idealized target trial, we do not necessarily expect the same answer to these slightly different research questions.[133] This gap can be minimized, but will always exist due to the limits of both observational data (and methods) and trial data, which often represent a highly specialized population that may not be generalizable to the target population of interest.

Future studies of interest related to anemia management in CKD and dialysis would ideally apply similarly innovative analytic techniques to assess the real-world safety and effectiveness of some of the new classes of medications, including iron-containing phosphate binders and HIF-PHIs. New-user designs would best align the conceptual hypothesis with the operational hypothesis in most cases, leading to methods that address clearly defined research questions. While the advantages of new-user designs to avoid biases are clear, the number of initiators available often limits statistical power. However, when analyzing treatments new to the market, most users will be new users, leading to increased feasibility of such analyses if longitudinal data are collected frequently enough to accurately model predictors of treatment initiation and avoid confounding by indication bias. With sufficient data collected, the g-methods,[131] including inverse probability weighting and the parametric g-formula, could then be used to emulate hypothetical trials of these treatments in a real-world setting. The largest obstacle to implementing these approaches may be the barrier to entry; the g-formula

knowledge base and SAS code developed for this dissertation will greatly reduce the amount of effort required to implement these analyses in practice. Potential projects may include evaluating alternative non-PIVOTAL protocols within the Aim 3 framework, or designing a comparative effectiveness target trial around treatment strategies with new anemia medications vs. the traditional therapies (ESA and IV iron). Finally, while dialysis data are uniquely positioned to facilitate implementation of these methods for reasons described above, researchers outside of nephrology could and should move toward applying these methodological principles when the observational data structure, size, and availability allow.

Conclusions

This dissertation outlines critical gaps in the literature on anemia management in HD patients, and describes three studies that utilize innovative designs and complex statistical analyses to address these gaps. The resultant manuscripts represent a meaningful contribution to the literature, and attempt to advance both the optimization of anemia management strategies in HD patients and the use of causal inferences principles to guide epidemiologic research using observational data.

BIBLIOGRAPHY

1. Mayo Clinic. Chronic Kidney Disease. https://www.mayoclinic.org/diseases-conditions/chronic-kidney-disease/symptoms-causes/syc-20354521. Published 2017. Accessed March 2, 2018.
2. National Kidney Foundation. A to Z Health Guide: Hemodialysis. https://www.kidney.org/atoz/content/hemodialysis. Published 2015. Accessed March 3, 2018.
3. Babitt JL, Lin HY. Mechanisms of anemia in CKD. *J Am Soc Nephrol*. 2012;23:1631-1634.
4. Lopez A, Cacoub P, Macdougall IC, Peyrin-Biroulet L. Iron deficiency anaemia. *Lancet*. 2016;387(10021):907-916.
5. Kidney Disease: Improving Global Outcomes (KDIGO) Anemia Work Group. KDIGO Clinical Practice Guideline for Anemia in Chronic Kidney Disease. *Kidney Int Suppl*. 2012;2(4):279–335.
6. National Institute of Diabetes and Digestive and Kidney Diseases. Anemia in Chronic Kidney Disease. https://www.niddk.nih.gov/health-information/kidney-disease/chronic-kidney-disease-ckd/anemia. Published 2014. Accessed March 3, 2018.
7. Macdougall IC. Iron Treatment Strategies in Nondialysis CKD. *Semin Nephrol*. 2016;36(2):99-104.
8. Ramanathan G, Olynyk JK, Ferrari P. Diagnosing and preventing iron overload. *Hemodial Int*. 2017;21:S58-S67.
9. Kalantar-Zadeh K, Streja E, Miller JE, Nissenson AR. Intravenous Iron Versus Erythropoiesis-Stimulating Agents: Friends or Foes in Treating Chronic Kidney Disease Anemia? *Adv Chronic Kidney Dis*. 2009;16(2):143-151.
10. Gaweda AE. Markers of iron status in chronic kidney disease. *Hemodial Int*. 2017;21:S21-S27.
11. Eschbach JW, Egrie JC, Downing MR, Browne JK, Adamson JW. Correction of the anemia of end-stage renal disease with recombinant human erythropoietin. *N Engl J Med*. 1987;316(2):73-78.
12. Goodkin DA, Fuller DS, Robinson BM, et al. Naturally Occurring Higher Hemoglobin Concentration Does Not Increase Mortality among Hemodialysis Patients. *J Am Soc Nephrol*. 2011;22(2):358-365.
13. Van Wyck DB, Stivelman JC, Ruiz J, Kirlin LF, Katz MA, Ogden DA. Iron status in patients receiving erythropoietin for dialysis-associated anemia. *Kidney Int*.

1989;35(2):712-716.

14. Ibrahim HN, Ishani A, Foley RN, Guo H, Liu J, Collins AJ. Temporal Trends in Red Blood Transfusion Among US Dialysis Patients, 1992-2005. *Am J Kidney Dis*. 2008;52(6):1115-1121.
15. Coyne DW. Iron indices: What do they really mean? *Kidney Int*. 2006;69(Suppl 101):S4-S8.
16. Cody JD, Hodson EM. Recombinant human erythropoietin versus placebo or no treatment for the anaemia of chronic kidney disease in people not requiring dialysis (Review). *Cochrane Database Syst Rev*. 2016;(1):59.
17. Besarab A, Bolton WK, Browne JK, et al. The Effects of Normal as Compared with Low Hematocrit Values in Patients with Cardiac Disease Who Are Receiving Hemodialysis and Epoetin. *N Engl J Med*. 1998;339(9):584-590.
18. Drueke TB, Locatelli F, Clyne N, et al. Normalization of Hemoglobin Level in Patients with Chronic Kidney Disease and Anemia. *N Engl J Med*. 2006;355(20):2071-2084.
19. Singh AK, Szczech L, Tang KL, et al. Correction of Anemia with Epoetin Alfa in Chronic Kidney Disease. *N Engl J Med*. 2006;355(20):2085-2098.
20. Pfeffer MA, Burdmann EA, Chen C-Y, et al. A Trial of Darbepoetin Alfa in Type 2 Diabetes and Chronic Kidney Disease. *N Engl J Med*. 2009;361(21):2019-2032.
21. Centers for Medicare & Medicaid Services. Medicare Program; end stage renal disease prospective payment system: final rule. *Fed Regist*. 2010;75:49029-49214.
22. Food and Drug Administration. *FDA Drug Safety Communication: Modified Dosing Recommendations to Improve the Safe Use of Erythropoiesis-Stimulating Agents (ESAs) in Chronic Kidney Disease*.; 2011.
23. Centers for Medicare & Medicaid Services. Medicare program; end-stage renal disease prospective payment system and quality incentive program; ambulance fee schedule; durable medical equipment; and competitive acquisition of certain durable medical equipment prosthetics, orthotics and supplies. *Fed Regist*. 2011;76:70228–70316.
24. Manns BJ, Tonelli M. The new FDA labeling for ESA-implications for patients and providers. *Clin J Am Soc Nephrol*. 2012;7(2):348-353.
25. Iglehart JK. Bundled Payment for ESRD - Including ESAs in Medicare's Dialysis Package. *N Engl J Med*. 2011;364(7):593-595.
26. Chertow GM, Liu J, Monda KL, et al. Epoetin Alfa and Outcomes in Dialysis amid Regulatory and Payment Reform. *J Am Soc Nephrol*. 2016;27(10):3129-3138.
27. Fuller DS, Pisoni RL, Bieber BA, Port FK, Robinson BM. The DOPPS practice monitor for US dialysis care: Update on trends in anemia management 2 years into the bundle. *Am J Kidney Dis*. 2013;62(6):1213-1216.
28. Wang C, Kane R, Levenson M, et al. Association between changes in CMS

reimbursement policy and drug labels for erythrocyte-stimulating agents with outcomes for older patients undergoing hemodialysis covered by fee-for-service medicare. *JAMA Intern Med*. 2016;176(12):1818-1825.

29. US-DOPPS Practice Monitor (DPM), August 2017. http://www.dopps.org/dpm. Accessed March 1, 2018.

30. Cotter DJ, Stefanik K, Zhang Y, Thamer M, Scharfstein D, Kaufman J. Hematocrit was not validated as a surrogate end point for survival among epoetin-treated hemodialysis patients. *J Clin Epidemiol*. 2004;57(10):1086-1095.

31. Steinbrook R. Lower Erythropoietin Doses and Medicare Payment Reform: Win-Wins for Patients with End-stage Renal Disease. *J Am Med Assoc*. 2016;176(12):1749-1751.

32. Fuller DS, Zepel L, Bieber BA, Robinson BM, Pisoni RL. Hemodialysis Facility Variation in Hospitalization and Transfusions Using Medicare Claims: The DOPPS Practice Monitor for US Dialysis Care. *Am J Kidney Dis*. 2016;67(2):337-340.

33. Hörl WH. Anaemia management and mortality risk in chronic kidney disease. *Nat Rev Nephrol*. 2013;9(5):291-301.

34. Kapoian T, O'Mara NB, Singh AK, et al. Ferric Gluconate Reduces Epoetin Requirements in Hemodialysis Patients with Elevated Ferritin. *J Am Soc Nephrol*. 2008;19(2):372-379.

35. Robinson BM, Larkina M, Bieber B, et al. Evaluating the effectiveness of IV iron dosing for anemia management in common clinical practice: Results from the Dialysis Outcomes and Practice Patterns Study (DOPPS). *BMC Nephrol*. 2017;18(330).

36. Karaboyas A, Zee J, Morgenstern H, et al. Understanding the Recent Increase in Ferritin Levels in United States Dialysis Patients: Potential Impact of Changes in Intravenous Iron and Erythropoiesis-Stimulating Agent Dosing. *Clin J Am Soc Nephrol*. 2015;10(6):1814-1821.

37. Slotki I, Cabantchik ZI. The Labile Side of Iron Supplementation in CKD. *J Am Soc Nephrol*. 2015;26:2612-2619.

38. Dwyer JP. We Give Too Much Intravenous Iron. *Semin Dial*. 2016;29(4):309-311.

39. Vaziri ND. Safety Issues in Iron Treatment in CKD. *Semin Nephrol*. 2016;36(2):112-118.

40. Rostoker G, Griuncelli M, Loridon C, et al. Reassessment of Iron biomarkers for prediction of dialysis iron Overload: An MRI Study. *PLoS One*. 2015;10(7):1-16.

41. Kalantar-Zadeh K, Kalantar-Zadeh K, Lee GH. The fascinating but deceptive ferritin: to measure it or not to measure it in chronic kidney disease? *Clin J Am Soc Nephrol*. 2006;1 Suppl 1:9-18.

42. Ferrari P, Kulkarni H, Dheda S, et al. Serum iron markers are inadequate for guiding iron repletion in chronic kidney disease. *Clin J Am Soc Nephrol*.

2011;6(1):77-83.

43. Richardson D, Hodsman A, Van Schalkwyk D, Tomson C, Warwick G. Management of anaemia in haemodialysis and peritoneal dialysis patients (Chapter 8). *Nephrol Dial Transplant*. 2007;22(Suppl 7).

44. Kalantar-Zadeh K, Don BR, Rodriguez RA, Humphreys MH. Serum ferritin is a marker of morbidity and mortality in hemodialysis patients. *Am J Kidney Dis*. 2001;37(3):564-572.

45. Kalantar-Zadeh K, Rodriguez RA, Humphreys MH. Association between serum ferritin and measures of inflammation, nutrition and iron in haemodialysis patients. *Nephrol Dial Transplant*. 2004;19(1):141-149.

46. Rambod M, Kovesdy CP, Kalantar-Zadeh K. Combined High Serum Ferritin and Low Iron Saturation in Hemodialysis Patients: The Role of Inflammation. *Clin J Am Soc Nephrol*. 2008;3(11):1691-1701.

47. Bazeley J, Bieber BA, Li Y, et al. C-reactive protein and prediction of 1-year mortality in prevalent hemodialysis patients. *Clin J Am Soc Nephrol*. 2011;6(10):2452-2461.

48. Kalantar-Zadeh K, Regidor DL, McAllister CJ, Beckie M, Warnock DG. Time-Dependent Associations between Iron and Mortality in Hemodialysis Patients. *J Am Soc Nephrol*. 2005;16(10):3070-3080.

49. Ford BA, Coyne DW, Eby CS, Scott MG. Variability of ferritin measurements in chronic kidney disease; implications for iron management. *Kidney Int*. 2009;75(1):104-110.

50. Van Wyck DB, Alcorn H, Gupta R. Analytical and biological variation in measures of anemia and iron status in patients treated with maintenance hemodialysis. *Am J Kidney Dis*. 2010;56(3):540-546.

51. Kliger AS, Foley RN, Goldfarb DS, et al. KDOQI US Commentary on the 2012 KDIGO Clinical Practice Guideline for Anemia in CKD The 2012 KDIGO (Kidney Disease: Improving Global Outcomes) Clinical Practice Guideline for Anemia in Chronic Kidney. *Am J Kidney Dis*. 2013;62(5):849-859.

52. Locatelli F, Bárány P, Covic A, et al. Kidney Disease: Improving Global Outcomes guidelines on anaemia management in chronic kidney disease: A European Renal Best Practice position statement. *Nephrol Dial Transplant*. 2013;28(6):1346-1359.

53. Tsubakihara Y, Nishi S, Akiba T, et al. 2008 Japanese society for dialysis therapy: Guidelines for renal anemia in chronic kidney disease. *Ther Apher Dial*. 2010;14(3):240-275.

54. Phrommintikul A, Haas SJ, Elsik M, Krum H. Mortality and target haemoglobin concentrations in anaemic patients with chronic kidney disease treated with erythropoietin: a meta-analysis. *Lancet*. 2007;369(9559):381-388.

55. Del Vecchio L, Locatelli F. Clinical practice guidelines on iron therapy: A critical evaluation. *Hemodial Int*. 2017;21:S125-S131.

56. Stivelman JC. Target-based Anemia Management with Erythropoiesis Stimulating Agents (Risks and Benefits Relearned) and Iron (Still More to Learn). *Semin Dial*. 2017;30(2):142-148.

57. Collister D, Rigatto C, Tangri N. Anemia management in chronic kidney disease and dialysis: A narrative review. *Curr Opin Nephrol Hypertens*. 2017;26(3):214-218.

58. Macdougall IC, Bircher AJ, Eckardt K-U, et al. Iron management in chronic kidney disease: conclusions from a “Kidney Disease: Improving Global Outcomes” (KDIGO) Controversies Conference. *Kidney Int*. 2016;89(1):28-39.

59. Hung S-C, Tarng D-C. ESA and iron therapy in chronic kidney disease: a balance between patient safety and hemoglobin target. *Kidney Int*. 2014;86(4):676-678.

60. Charytan DM, Pai AB, Chan CT, et al. Considerations and Challenges in Defining Optimal Iron Utilization in Hemodialysis. *J Am Soc Nephrol*. 2015;26(6):1238-1247.

61. Coyne DW, Kapoian T, Suki W, et al. Ferric Gluconate Is Highly Efficacious in Anemic Hemodialysis Patients with High Serum Ferritin and Low Transferrin Saturation: Results of the Dialysis Patients’ Response to IV Iron with Elevated Ferritin (DRIVE) Study. *J Am Soc Nephrol*. 2007;18(3):975-984.

62. Bailie GR, Larkina M, Goodkin DA, et al. Data from the Dialysis Outcomes and Practice Patterns Study validate an association between high intravenous iron doses and mortality. *Kidney Int*. 2015;87(1):162-168.

63. Miskulin DC, Tangri N, Bandeen-Roche K, et al. Intravenous iron exposure and mortality in patients on hemodialysis. *Clin J Am Soc Nephrol*. 2014;9(11):1930-1939.

64. Feldman HI, Joffe M, Robinson BM, et al. Administration of Parenteral Iron and Mortality among Hemodialysis Patients. *J Am Soc Nephrol*. 2004;15(6):1623-1632.

65. Macdougall IC, White C, Anker SD, et al. Intravenous Iron in Patients Undergoing Maintenance Hemodialysis. *N Engl J Med*. 2018;380(5):447-458.

66. Macdougall IC, Cooper AC. Erythropoietin resistance: the role of inflammation and pro-inflammatory cytokines. *Nephrol Dial Transplant*. 2002;17(Suppl 11):39-43.

67. National Kidney Foundation. KDOQI Clinical Practice Guidelines and Clinical Practice Recommendations for Anemia in Chronic Kidney Disease. *Am J Kidney Dis*. 2006;47(5 Suppl 3):S11-145.

68. Kaysen GA, Dubin JA, Müller HG, Rosales LM, Levin NW. The acute-phase response varies with time and predicts serum albumin levels in hemodialysis patients. *Kidney Int*. 2000;58(1):346-352.

69. Chawla LS, Krishnan M. Causes and consequences of inflammation on anemia management in hemodialysis patients. *Hemodial Int*. 2009;13(2):222-234.

70. Hamano T, Fujii N, Hayashi T, Yamamoto H, Iseki K, Tsubakihara Y. Thresholds of iron markers for iron deficiency erythropoiesis-finding of the Japanese nationwide dialysis registry. *Kidney Int Suppl*. 2015;5(1):23-32.

71. Bailie GR, Larkina M, Goodkin DA, et al. Variation in intravenous iron use internationally and over time: The Dialysis Outcomes and Practice Patterns Study (DOPPS). *Nephrol Dial Transplant*. 2013;28(10):2570-2579.

72. Birnie K, Caskey F, Ben-Shlomo Y, et al. Erythropoiesis-stimulating agent dosing , haemoglobin and ferritin levels in UK haemodialysis patients 2005 – 13. *Nephrol Dial Transplant*. 2016;(1):1-7.

73. United States Renal Data System. *2017 USRDS Annual Data Report: Epidemiology of Kidney Disease in the United States*. Bethesda, MD; 2017.

74. Karaboyas A, Robinson BM, Meier Y, et al. Anemia and Iron Management over the First 5 Years after Dialysis Start: Results from the DOPPS. In: *American Society of Nephrology Kidney Week*. Chicago, IL; 2016:SA-OR110.

75. Unger EF, Thompson AM, Blank MB, Temple R. Erythropoiesis-Stimulating Agents - Time for a Reevaluation. *N Engl J Med*. 2010;362(3):189-192.

76. Bárány P, Divino Filho JC, Bergström J. High C-reactive protein is a strong predictor of resistance to erythropoietin in hemodialysis patients. *Am J Kidney Dis*. 1997;29(4):565-568.

77. Locatelli F, Andrulli S, Memoli B, et al. Nutritional-inflammation status and resistance to erythropoietin therapy in haemodialysis patients. *Nephrol Dial Transplant*. 2006;21(4):991-998.

78. Rattanasompattikul M, Molnar MZ, Zaritsky JJ, et al. Association of malnutrition-inflammation complex and responsiveness to erythropoiesis-stimulating agents in long-term hemodialysis patients. *Nephrol Dial Transplant*. 2013;28(7):1936-1945.

79. Bradbury BD, Critchlow CW, Weir MR, Stewart R, Krishnan M, Hakim RH. Impact of elevated C-reactive protein levels on erythropoiesis- stimulating agent (ESA) dose and responsiveness in hemodialysis patients. *Nephrol Dial Transplant*. 2009;24(3):919-925.

80. Kimachi M, Fukuma S, Yamazaki S, et al. Minor Elevation in C-Reactive Protein Levels Predicts Incidence of Erythropoiesis-Stimulating Agent Hyporesponsiveness among Hemodialysis Patients. *Nephron Clin Pract*. 2015;131(2):123-130.

81. Singh AK, Coyne DW, Shapiro W, Rizkala AR. Predictors of the response to treatment in anemic hemodialysis patients with high serum ferritin and low transferrin saturation. *Kidney Int*. 2007;71(11):1163-1171.

82. Gillespie IA, Macdougall IC, Richards S, et al. Factors precipitating erythropoiesis-stimulating agent responsiveness in a European haemodialysis cohort: case-crossover study. *Pharmacoepidemiology*. 2015;24:414-425.

83. Robins J. A new approach to causal inference in mortality studies with a sustained exposure period-application to control of the healthy worker survivor

effect. *Math Model*. 1986;7(9-12):1393-1512.

84. National Institute of Diabetes and Digestive and Kidney Diseases. Kidney Disease Statistics for the United States. https://www.niddk.nih.gov/health-information/health-statistics/kidney-disease. Published 2016. Accessed March 3, 2018.

85. Besarab A, Amin N, Ahsan M, et al. Optimization of epoetin therapy with intravenous iron therapy in hemodialysis patients. *J Am Soc Nephrol*. 2000;11(3):530-538.

86. Shah HH, Fishbane S. Is there an established hemoglobin target range for patients undergoing chronic dialysis? *Semin Dial*. 2018;31(4):415-419.

87. Wetmore JB, Li S, Yan H, et al. Predialysis anemia management and outcomes following dialysis initiation: A retrospective cohort analysis. *PLoS One*. 2018;13(9).

88. Mc Causland FR, Claggett B, Burdmann EA, et al. Treatment of Anemia With Darbepoetin Prior to Dialysis Initiation and Clinical Outcomes: Analyses From the Trial to Reduce Cardiovascular Events With Aranesp Therapy (TREAT). *Am J Kidney Dis*. 2019;73(3):309-315.

89. Macdougall IC. Intravenous iron therapy in patients with chronic kidney disease: recent evidence and future directions. *Clin Kidney J*. 2017;10(suppl_1):i16-i24.

90. Shepshelovich D, Rozen-Zvi B, Avni T, Gafter U, Gafter-Gvili A. Intravenous Versus Oral Iron Supplementation for the Treatment of Anemia in CKD: An Updated Systematic Review and Meta-analysis. *Am J Kidney Dis*. 2016;68(5):677-690.

91. Tangri N, Miskulin DC, Zhou J, et al. Effect of intravenous iron use on hospitalizations in patients undergoing hemodialysis: A comparative effectiveness analysis from the DEcIDE-ESRD study. *Nephrol Dial Transplant*. 2015;30(4):667-675.

92. Hougen I, Collister D, Bourrier M, et al. Safety of intravenous iron in dialysis: A systematic review and meta-analysis. *Clin J Am Soc Nephrol*. 2018;13(3):457-467.

93. Singh AK. Editorial: The FDA's perspective on the risk for rapid rise in hemoglobin in treating CKD anemia: Quo vadis. *Clin J Am Soc Nephrol*. 2010;5(4):553-556.

94. Young EW, Goodkin DA, Mapes DL, et al. The Dialysis Outcomes and Practice Patterns Study (DOPPS): An international hemodialysis study. *Kidney Int*. 2000;57(Suppl 74):S74-S81.

95. Pisoni RL, Gillespie BW, Dickinson DM, Chen K, Kutner MH, Wolfe RA. The Dialysis Outcomes and Practice Patterns Study (DOPPS): Design, Data elements, and Methodology. *Am J Kidney Dis*. 2004;44(No. 5, Suppl 2 (November)):S7-S15.

96. Bock HA, Hirt-Minkowski P, Brünisholz M, Keusch G, Rey S, Von Albertini B. Darbepoetin alpha in lower-than-equimolar doses maintains haemoglobin levels in

stable haemodialysis patients converting from epoetin alpha/beta. *Nephrol Dial Transplant*. 2008;23(1):301-308.

97. Choi P, Farouk M, Manamley N, Addison J. Dose conversion ratio in hemodialysis patients switched from darbepoetin alfa to PEG-epoetin beta: AFFIRM study. *Adv Ther*. 2013;30(11):1007-1017.

98. McFarlane PA, Pisoni RL, Eichleay MA, Wald R, Port FK, Mendelssohn D. International trends in erythropoietin use and hemoglobin levels in hemodialysis patients. *Kidney Int*. 2010;78(2):215-223.

99. Raghunathan TE, Solenberger PW, Hoewyk J Van. IVEware : Imputation and Variance Estimation Software User Guide. *Ann Arbor, MI Surv Methodol Program, Surv Res Center, Inst Soc Res Univ Michigan*. 2002;(March).

100. Little RJ, Rubin DB. *Statistical Analysis with Missing Data*. New York, NY: Wiley; 1987.

101. VanderWeele TJ. *Explanation in Causal Inference: Methods for Mediation and Interaction*. New York: Oxford University Press; 2015.

102. Zhang Y, Thamer M, Kaufman JS, Cotter DJ, Hernán MA. High doses of epoetin do not lower mortality and cardiovascular risk among elderly hemodialysis patients with diabetes. *Kidney Int*. 2011;80(6):663-669.

103. Suttorp MM, Hoekstra T, Mittelman M, et al. Treatment with high dose of erythropoiesis-stimulating agents and mortality: analysis with a sequential Cox approach and a marginal structural model. *Pharmacoepidemiology*. 2015;24:1068-1075.

104. Zhang Y, Thamer M, Stefanik K, Kaufman J, Cotter DJ. Epoetin requirements predict mortality in hemodialysis patients. *Am J Kidney Dis*. 2004;44(5):866-876.

105. Bradbury BD, Danese MD, Gleeson M, Critchlow CW. Effect of epoetin alfa dose changes on hemoglobin and mortality in hemodialysis patients with hemoglobin levels persistently below 11 g/dL. *Clin J Am Soc Nephrol*. 2009;4(3):630-637.

106. Brookhart MA, Freburger JK, Ellis AR, Wang L, Winkelmayer WC, Kshirsagar A V. Infection Risk with Bolus versus Maintenance Iron Supplementation in Hemodialysis Patients. *J Am Soc Nephrol*. 2013;24:1151-1158.

107. Kshirsagar A V., Freburger JK, Ellis AR, Wang L, Winkelmayer WC, Brookhart MA. Intravenous iron supplementation practices and short-term risk of cardiovascular events in hemodialysis patients. *PLoS One*. 2013;8(11):1-8.

108. Michels WM, Jaar BG, Ephraim PL, et al. Intravenous iron administration strategies and anemia management in hemodialysis patients. *Nephrol Dial Transplant*. 2017;32(1):173-181.

109. Kuo K-L, Hung S, Liu J-S, Chang Y-K, Hsu C-C, Tarng D-C. Iron supplementation associates with low mortality in pre-dialyzed advanced chronic kidney disease patients receiving erythropoiesis-stimulating agents: A nationwide database analysis. *Nephrol Dial Transplant*. 2015;30(9):1518-1525.

110. Kovesdy CP. Can Reduction of Inflammation Improve ESA Dose Response? *Semin Dial*. 2013;26(5):540-542.

111. Kanbay M, Perazella MA, Kasapoglu B, Koroglu M, Covic A. Erythropoiesis stimulatory agent- resistant anemia in dialysis patients: Review of causes and management. *Blood Purif*. 2010;29(1):1-12.

112. Kalantar-Zadeh K, Ikizler TA, Block G, Avram MM, Kopple JD. Malnutrition-Inflammation Complex Syndrome in Dialysis Patients: Causes and Consequences. *Am J Kidney Dis*. 2003;42(5):864-881.

113. Nowak KL, Chonchol M. Does inflammation affect outcomes in dialysis patients? *Semin Dial*. 2018;31(4):388-397.

114. Lowrie EG. Acute-phase inflammatory process contributes to malnutrition, anemia, and possibly other abnormalities in dialysis patients. *Am J Kidney Dis*. 1998;32(6 Suppl 4):S105-12.

115. Coyne DW. Hepcidin: Clinical utility as a diagnostic tool and therapeutic target. *Kidney Int*. 2011;80(3):240-244.

116. Gupta N, Wish JB. Hypoxia-Inducible Factor Prolyl Hydroxylase Inhibitors: A Potential New Treatment for Anemia in Patients With CKD. *Am J Kidney Dis*. 2017;69(6):815-826.

117. Sugahara M, Tanaka T, Nangaku M. Prolyl hydroxylase domain inhibitors as a novel therapeutic approach against anemia in chronic kidney disease. *Kidney Int*. 2017;92(2):306-312.

118. Del Vecchio L, Locatelli F. Investigational hypoxia-inducible factor prolyl hydroxylase inhibitors (HIF-PHI) for the treatment of anemia associated with chronic kidney disease. *Expert Opin Investig Drugs*. 2018;27(7):613-621.

119. Kalantar-Zadeh K, McAllister CJ, Lehn RS, Lee GH, Nissenson AR, Kopple JD. Effect of malnutrition-inflammation complex syndrome on EPO hyporesponsiveness in maintenance hemodialysis patients. *Am J Kidney Dis*. 2003;42(4):761-773.

120. Robinson BM, Akizawa T, Jager KJ, Kerr PG, Saran R, Pisoni RL. Factors affecting outcomes in patients reaching end-stage kidney disease worldwide: differences in access to renal replacement therapy, modality use, and haemodialysis practices. *Lancet*. 2016;388(10041):294-306.

121. Rothman KJ, Greenland S, Lash TJ. *Modern Epidemiology*. Third.; 2014.

122. Zou GY, Donner A. Extension of the modified Poisson regression model to prospective studies with correlated binary data. *Stat Methods Med Res*. 2013;22(6):661-670.

123. Kelley-Hedgepeth A, Lloyd-Jones DM, Colvin A, et al. Ethnic differences in C-reactive protein concentrations. *Clin Chem*. 2008;54(6):1027-1037.

124. Barany P. Inflammation, serum C-reactive protein, and erythropoietin resistance. *Nephrol Dial Transplant*. 2001;16(2):224-227.

125. Achinger SG, Ayus JC. When the source of inflammation is hiding in plain sight: Failed kidney transplants, clotted arteriovenous grafts, and central venous catheters. *Semin Dial*. 2019;32(1):15-21.

126. de Francisco ALM, Stenvinkel P, Vaulont S. Inflammation and its impact on anaemia in chronic kidney disease: from haemoglobin variability to hyporesponsiveness. *NDT Plus*. 2009;2(Supplement 1):i18-i26.

127. López-Gómez JM, Pérez-Flores I, Jofré R, et al. Presence of a failed kidney transplant in patients who are on hemodialysis is associated with chronic inflammatory state and erythropoietin resistance. *J Am Soc Nephrol*. 2004;15(9):2494-2501.

128. Chait Y, Kalim S, Horowitz J, et al. The greatly misunderstood erythropoietin resistance index and the case for a new responsiveness measure. *Hemodial Int*. 2016;20(3):392-398.

129. Daniel RM, Cousens SN, De Stavola BL, Kenward MG, Sterne JAC. Methods for dealing with time-dependent confounding. *Stat Med*. 2013;32(9):1584-1618.

130. Hernán MA, Lanoy E, Costagliola D, Robins JM. Comparison of dynamic treatment regimes via inverse probability weighting. *Basic Clin Pharmacol Toxicol*. 2006;98(3):237-242.

131. Hernán MA, Robins JM. *Causal Inference*. Boca Raton: Chapman & Hall/CRC, forthcoming; 2019.

132. Young JG, Cain LE, Robins JM, O'Reilly EJ, Hernán MA. Comparative Effectiveness of Dynamic Treatment Regimes: An Application of the Parametric G-Formula. *Stat Biosci*. 2011;3(1):119-143.

133. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol*. 2016;183(8):758-764.

134. Brookhart MA, Schneeweiss S, Avorn J, Bradbury BD, Liu J, Winkelmayer WC. Comparative mortality risk of anemia management practices in incident hemodialysis patients. *Jama*. 2010;303(9):857-864.

135. Clement FM, Klarenbach S, Tonelli M, Wiebe N, Hemmelgarn B, Manns BJ. An economic evaluation of erythropoiesis-stimulating agents in CKD. *Am J Kidney Dis*. 2010;56(6):1050-1061.

136. Karaboyas A, Morgenstern H, Pisoni RL, et al. Association between serum ferritin and mortality: Findings from the USA, Japan and European Dialysis Outcomes and Practice Patterns Study. *Nephrol Dial Transplant*. 2018;33(12):2234-2244.

137. Cole SR, Richardson DB, Chu H, Naimi AI. Analysis of occupational asbestos exposure and lung cancer mortality using the G formula. *Am J Epidemiol*. 2013;177(9):989-996.

138. Keil AP, Edwards JK, Richardson DB, Naimi AI, Cole SR. The parametric g-formula for time-to-event data: Intuition and a worked example. *Epidemiology*. 2014;25(6):889-897.

139. Macdougall IC, White C, Anker SD, et al. Randomized trial comparing proactive, high-dose versus reactive, low-dose intravenous iron supplementation in hemodialysis (PIVOTAL): Study design and baseline data. *Am J Nephrol*. 2018;48(4):260-268.

140. Schomaker M, Heumann C. Bootstrap inference when using multiple imputation. *Stat Med*. 2018;37(14):2252-2266.

141. Taubman SL, Robins JM, Mittleman MA, Hernán MA. Intervening on risk factors for coronary heart disease: An application of the parametric g-formula. *Int J Epidemiol*. 2009;38(6):1599-1611.

142. Zhang Y, Young JG, Thamer M, Hernán MA. Comparing the Effectiveness of Dynamic Treatment Strategies Using Electronic Health Records: An Application of the Parametric g-Formula to Anemia Management Strategies. *Health Serv Res*. 2017:1-19.

143. Westreich D, Cole SR, Young JG, et al. The parametric g-formula to estimate the effect of highly active antiretroviral therapy on incident AIDS or death. *Stat Med*. 2012;31(18):2000-2009.

144. System USRD. *2018 USRDS Annual Data Report: Epidemiology of Kidney Disease in the United States*. Bethesda, MD; 2018.

145. Newsome SJ, Keogh RH, Daniel RM. Estimating long-term treatment effects in observational data: A comparison of the performance of different methods under real-world uncertainty. *Stat Med*. 2018;37(15):2367-2390.

146. Palmer SC, Sciancalepore M, Strippoli GFM. Trial quality in nephrology: How are we measuring up? *Am J Kidney Dis*. 2011;58(3):335-337.

147. Strippoli GFM, Craig JC, Schena FP. The Number, Quality, and Coverage of Randomized Controlled Trials in Nephrology. *J Am Soc Nephrol*. 2004;15(2):411-419.

148. Roger SD. Practical considerations for iron therapy in the management of anaemia in patients with chronic kidney disease. *Clin Kidney J*. 2017;10(Suppl 1):i9-i15.

149. Zhang Y, Thamer M, Kaufman J, Cotter D, Hernan MA. Comparative effectiveness of two anemia management strategies for complex elderly dialysis patients. *Med Care*. 2014;52(3):132-139.

150. Cain LE, Saag MS, Petersen M, et al. Using observational data to emulate a randomized trial of dynamic treatmentswitching strategies: An application to antiretroviral therapy. *Int J Epidemiol*. 2016;45(6):2038-2049.

151. Locatelli F, Del Vecchio L. Iron-based phosphate binders: a paradigm shift in the treatment of hyperphosphatemic anemic CKD patients? *J Nephrol*. 2017;30(6):755-765.

152. Umanath K, Jalal DI, Greco BA, et al. Ferric Citrate Reduces Intravenous Iron and Erythropoiesis-Stimulating Agent Use in ESRD. *J Am Soc Nephrol*. 2015;26(10):2578-2587.

153. Besarab A, Chernyavskaya E, Motylev I, et al. Roxadustat (FG-4592): Correction of Anemia in Incident Dialysis Patients. *J Am Soc Nephrol*. 2016;27(4):1225-1233.

154. Holdstock L, Meadowcroft AM, Maier R, et al. Four-Week Studies of Oral Hypoxia-Inducible Factor-Prolyl Hydroxylase Inhibitor GSK1278863 for Treatment of Anemia. *J Am Soc Nephrol*. 2016;27(4):1234-1244.

155. Provenzano R, Besarab A, Sun CH, et al. Oral hypoxia–inducible factor prolyl hydroxylase inhibitor roxadustat (FG-4592) for the treatment of anemia in patients with CKD. *Clin J Am Soc Nephrol*. 2016;11(6):982-991.