A Meta-Analysis of Hydroxyurea Use for β-thalassemia: Implications for Clinical Practice and Medical Education

by

Ali Algiraigri

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN COMMUNITY HEALTH SCIENCES

CALGARY, ALBERTA

NOVEMBER, 2015

© Ali Algiraigri 2015
Abstract

Chronic blood transfusion remains the most feasible therapeutic option for the majority of patients with severe β-thalassemia. However, it is associated with serious risks and complications. An alternative option is desirable and may prevent some of the problems associated with current therapy. Hydroxyurea (HU), an oral chemotherapeutic drug, is expected to increase hemoglobin, thereby minimizing the burden of blood transfusion and its complications. The objective of this study was to conduct a systematic review and meta-analysis to evaluate the clinical efficacy and safety of HU in patients with severe β-thalassemia. HU appears to be effective, well tolerated and associated with mild and transient adverse events; however, large randomized clinical trials (RCTs) should be done to confirm such findings. Nonetheless, based on the results of the present meta-analysis, it is recommended that current practice guidelines for severe β-thalassemia be appended to include a trial of HU.
Acknowledgments

First and foremost I would like to express my deep and sincere gratitude to my supervisor, Dr. Aliya Kassam for the continuous support of my study and research, for her patience, caring, motivation, and immense knowledge. Her guidance helped me in all the time of research and writing of this thesis.

Secondly, I want to thank my committee members, Dr. Nicola Wright and Dr. Elizabeth Oddone Paolucci, for their insightful comments and guidance.

Finally, I would like to recognize the constant support, sacrifice, understanding and love from my wife, Mashael, my daughters, Omaimah and Sarah. They are the major motivation for the success of this work, and the energy to keep accomplishing more and higher goals.
Dedication

To my beloved mother and father, it's impossible to thank you adequately for everything you've done.

To my dear wife, Mashael, who has supported me in all endeavors.

To my beloved daughters; Omaimah and Sarah, who make life fun.
# Table of Contents

<table>
<thead>
<tr>
<th>Content</th>
<th>Page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Acknowledgment</td>
<td>iii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iv</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>v</td>
</tr>
<tr>
<td>List of Tables</td>
<td>viii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>vii</td>
</tr>
<tr>
<td>List of Symbols, Abbreviations, Nomenclatures</td>
<td>ix</td>
</tr>
<tr>
<td><strong>Chapter 1: Literature Review and Study Rationale</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1 Overview of thesis project</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Aim</td>
<td>1</td>
</tr>
<tr>
<td>1.3 Literature review</td>
<td>2</td>
</tr>
<tr>
<td>1.3.1 Description of the condition</td>
<td>2</td>
</tr>
<tr>
<td>1.3.2 Description of the intervention</td>
<td>7</td>
</tr>
<tr>
<td>1.4 Study rationale</td>
<td>8</td>
</tr>
<tr>
<td>1.5 Overview of the methodological approach</td>
<td>9</td>
</tr>
<tr>
<td>1.5.1 Definitions and characteristics of systematic review and meta-analysis</td>
<td>11</td>
</tr>
<tr>
<td>1.5.2 Strengths of systematic review and meta-analysis</td>
<td>11</td>
</tr>
<tr>
<td>1.5.3 Disadvantages of systematic review and meta-analysis</td>
<td>12</td>
</tr>
<tr>
<td>1.6 Clinical practice guideline</td>
<td>14</td>
</tr>
<tr>
<td>**Chapter 2: A Meta-Analysis of Hydroxyurea Use for β-thalassemia:</td>
<td>19</td>
</tr>
<tr>
<td>Implications for Clinical Practice and Medical Education</td>
<td>19</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>20</td>
</tr>
<tr>
<td>2.2 Methods</td>
<td>21</td>
</tr>
<tr>
<td>2.2.1 Data sources and searches</td>
<td>22</td>
</tr>
<tr>
<td>2.2.2 Study selection</td>
<td>22</td>
</tr>
<tr>
<td>2.2.3 Data abstraction and quality assessment</td>
<td>23</td>
</tr>
<tr>
<td>2.2.4 Data synthesis and analysis</td>
<td>23</td>
</tr>
<tr>
<td>2.2.5 Heterogeneity and publication bias</td>
<td>24</td>
</tr>
<tr>
<td>2.3 Results</td>
<td>25</td>
</tr>
<tr>
<td>2.3.1 Study selection</td>
<td>25</td>
</tr>
<tr>
<td>2.3.2 Study characteristics</td>
<td>27</td>
</tr>
<tr>
<td>2.3.3 Study quality</td>
<td>30</td>
</tr>
<tr>
<td>2.3.4 Response rate of HU</td>
<td>30</td>
</tr>
<tr>
<td>2.3.5 Analysis of heterogeneity and publication bias</td>
<td>30</td>
</tr>
<tr>
<td>2.4 Discussion</td>
<td>31</td>
</tr>
<tr>
<td>2.5 Acknowledgments</td>
<td>39</td>
</tr>
<tr>
<td>2.6 Authorship</td>
<td>39</td>
</tr>
<tr>
<td><strong>Chapter 3: Clinical Significance, Implications and Knowledge Translation</strong></td>
<td>40</td>
</tr>
<tr>
<td>3.1 Clinical significance</td>
<td>40</td>
</tr>
<tr>
<td>3.2 Study limitations</td>
<td>41</td>
</tr>
<tr>
<td>3.3 Implications</td>
<td>42</td>
</tr>
<tr>
<td>3.3.1 Implications for clinical practice guidelines</td>
<td>42</td>
</tr>
<tr>
<td>3.3.2 Implications for practice</td>
<td>45</td>
</tr>
<tr>
<td>3.3.3 Implications for research</td>
<td>47</td>
</tr>
<tr>
<td>3.3.4</td>
<td>Implications for EBP and medical education</td>
</tr>
<tr>
<td>-------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>3.4</td>
<td>Knowledge translation</td>
</tr>
<tr>
<td>3.5</td>
<td>References</td>
</tr>
<tr>
<td></td>
<td>Appendix A</td>
</tr>
<tr>
<td></td>
<td>Appendix B</td>
</tr>
<tr>
<td></td>
<td>Appendix C</td>
</tr>
<tr>
<td></td>
<td>Appendix D</td>
</tr>
<tr>
<td></td>
<td>Appendix E</td>
</tr>
</tbody>
</table>
## List of Tables

<table>
<thead>
<tr>
<th>Table #</th>
<th>Title</th>
<th>Page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Genotype and phenotype of β-thalassemia</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>GRADE process in developing recommendation</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>Determination of the strength of GRADE recommendation</td>
<td>18</td>
</tr>
<tr>
<td>4</td>
<td>Summary characteristics of included studies [ordered according to publication year].</td>
<td>26</td>
</tr>
<tr>
<td>5</td>
<td>Study quality assessment by NIH tool</td>
<td>29</td>
</tr>
<tr>
<td>6</td>
<td>Subgroup Analysis of HU complete response rate.</td>
<td>32</td>
</tr>
<tr>
<td>Figure #</td>
<td>Title</td>
<td>Page #</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------------------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>1</td>
<td>Transfusion needs in various β-thalassemia forms.</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>Level of evidence</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>PRISMA flow diagram of literature search.</td>
<td>25</td>
</tr>
<tr>
<td>4</td>
<td>Forest plots of HU responses in lifelong TDβT patients.</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>Meta-regression plot.</td>
<td>33</td>
</tr>
<tr>
<td>6</td>
<td>Funnel plots of HU responses. A. CRR; B. ORR.</td>
<td>34</td>
</tr>
<tr>
<td>7</td>
<td>Current practice approach for lifelong TDβT patients.</td>
<td>44</td>
</tr>
<tr>
<td>8</td>
<td>Proposed management algorithm for lifelong TDβT patients.</td>
<td>44</td>
</tr>
<tr>
<td>9</td>
<td>EBP framework</td>
<td>50</td>
</tr>
<tr>
<td>10</td>
<td>Knowledge-to-action framework.</td>
<td>55</td>
</tr>
</tbody>
</table>
### List of Symbols, Abbreviations, Nomenclatures

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>Alpha</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Beta</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Gamma</td>
</tr>
<tr>
<td>AEs</td>
<td>Adverse events</td>
</tr>
<tr>
<td>Alloimmunization</td>
<td>An immunological response by the recipient against “foreign” non-self-antigens that may follow an erythrocyte transfusion and result in destruction of transfused erythrocytes.</td>
</tr>
<tr>
<td>ASPHO</td>
<td>American Society of Hematology/Oncology</td>
</tr>
<tr>
<td>BJH</td>
<td>British Journal of Haematology</td>
</tr>
<tr>
<td>CBC</td>
<td>Complete blood count</td>
</tr>
<tr>
<td>CENTRAL</td>
<td>Cochrane Central Register of Controlled Trials</td>
</tr>
<tr>
<td>Chelation</td>
<td>Removal of excess iron from the body, using a specific medication called iron chelator</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>CML</td>
<td>Chronic myeloid leukemia</td>
</tr>
<tr>
<td>CRR</td>
<td>Complete response rate</td>
</tr>
<tr>
<td>EBM</td>
<td>Evidence-based medicine</td>
</tr>
<tr>
<td>EBP</td>
<td>Evidence-based practice</td>
</tr>
<tr>
<td>EHA</td>
<td>European Hematology Association</td>
</tr>
<tr>
<td>FDA</td>
<td>Food and Drug Administration</td>
</tr>
<tr>
<td>Hb</td>
<td>Hemoglobin, oxygen-carrying pigment contained in the red blood cells</td>
</tr>
<tr>
<td>HbA</td>
<td>Hemoglobin A, normal hemoglobin.</td>
</tr>
<tr>
<td>HbF</td>
<td>Fetal hemoglobin</td>
</tr>
<tr>
<td>Hemoglobinopathy</td>
<td>A disorder characterized by an abnormality of the structure or function of hemoglobin.</td>
</tr>
<tr>
<td>Hepatitis</td>
<td>Inflammation of the liver</td>
</tr>
<tr>
<td>HIV</td>
<td>Human immunodeficiency virus</td>
</tr>
<tr>
<td>HPLC</td>
<td>High performance liquid chromatography</td>
</tr>
<tr>
<td>HU</td>
<td>Hydroxyurea; also known as Hydroxycarbamide.</td>
</tr>
<tr>
<td>HSCT</td>
<td>Hematopoietic stem cell transplantation</td>
</tr>
<tr>
<td>ICTR</td>
<td>International Clinical Trials Registry Platform</td>
</tr>
<tr>
<td>Iron overload</td>
<td>An excess of iron in the body</td>
</tr>
<tr>
<td>MCV</td>
<td>Mean corpuscular volume; The average volume of red blood cells measured in femtoliters (fL).</td>
</tr>
<tr>
<td>MOOSE</td>
<td>Meta-analysis of Observational Studies in Epidemiology</td>
</tr>
<tr>
<td>MPD</td>
<td>Myeloproliferative disorders</td>
</tr>
<tr>
<td>NIH</td>
<td>National Institutes of Health</td>
</tr>
<tr>
<td>NTD$\beta$T</td>
<td>Non-transfusion dependent $\beta$-thalassemia</td>
</tr>
<tr>
<td>ORR</td>
<td>Overall response rate</td>
</tr>
<tr>
<td>PRISMA</td>
<td>Preferred Reporting Items for Systematic reviews and Meta-Analyses</td>
</tr>
<tr>
<td>PROSPERO</td>
<td>Prospective Register of Systematic Reviews</td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>RBC</td>
<td>Red blood cell</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomized clinical trial</td>
</tr>
<tr>
<td>SCA</td>
<td>Sickle cell anemia</td>
</tr>
<tr>
<td>Splenectomy</td>
<td>Surgical removal of the spleen</td>
</tr>
<tr>
<td>TD</td>
<td>Transfusion dependent</td>
</tr>
<tr>
<td>TDβT</td>
<td>Transfusion dependent β-Thalassemia</td>
</tr>
<tr>
<td>TI</td>
<td>Thalassemia intermedia</td>
</tr>
<tr>
<td>TM</td>
<td>Thalassemia major</td>
</tr>
<tr>
<td>WBC</td>
<td>White blood cell</td>
</tr>
<tr>
<td>Xmn1</td>
<td>Presence of a polymorphism for the enzyme Xmn1 in the GÁ-promoter region</td>
</tr>
</tbody>
</table>
Chapter 1: Literature Review and Study Rationale

1.1 Overview of thesis project

The following thesis research pertained to an in-depth assessment of the use of hydroxyurea for patients with lifelong transfusion-dependent β-thalassemia by applying systematic review and meta-analysis research methods. As a result, this thesis included the following chapters: (1) “The present chapter” which provided the rationale for the study, background review regarding the condition (β-thalassemia), the target intervention (Hydroxyurea, HU), and an overview of the research methods (systematic review and meta-analysis); (2) A comprehensive report on the method, results, and discussion of the meta-analysis comprising a manuscript; (3) Discussion of the implications of this study on current treatment guidelines, clinical practice, medical education and future research.

1.2 Aim

The purpose of this thesis was to examine the effectiveness of hydroxyurea for patients with lifelong transfusion-dependent β-thalassemia. This was accomplished by means of (1) conducting a systematic review of the literature, which included published and unpublished research; (2) determining the effectiveness of hydroxyurea by conducting a meta-analysis of studies that met the inclusionary criteria for patients with lifelong transfusion-dependent β-thalassemia; (3) critically examining the results of the meta-analysis; and (4) making successive medical education, clinical practice and research recommendations for the future.
1.3 Literature review

1.3.1 Description of the condition

Definitions and classifications

The term “thalassemia” is derived from the Greek words “Thalassa” (sea) and “Haema” (blood) and was first reported independently in the United States and Italy in 1925. The name was created under the mistaken belief that these disorders were confined to the Mediterranean region. Thalassemias are inherited disorders of the red blood cell protein, hemoglobin. Hemoglobin is comprised of two essential parts: heme and globin. Globin is comprised of alpha (α) and beta (β) peptide chains. Thalassemia is caused by one or more defects in the genes that are responsible for the production of these alpha and beta peptide chains. The type of thalassemia diagnosed is determined by the defective globin gene involved; patients with affected α-globin genes have α-thalassemia, and those with affected β-globin genes have β-thalassemia.

The fundamental defect in β-Thalassemia is an underproduction of β-globulin chains combined with an excess of free α-globin chains. The direct consequences are a decrease in hemoglobin production (anemia) and imbalance of the globin chains. This imbalance is due to the excess free α-globin chains which are unstable, nonfunctional, and precipitate within the cells, leading to premature cell death - a phenomenon termed ineffective erythropoiesis, the hallmark of this condition.

An estimated 5% of the world’s population (almost 150 million people) carry β-thalassemia genes. Despite the complexity of β-thalassemia genotypes (more than 200 different mutations of β-globin genes), three broad clinical phenotypes of β-
thalassemia are recognized based on the hemoglobin steady state and the need for blood transfusions: major; intermedia; and minor.

1. **β-thalassemia major** represents the most severe form of β-thalassemia, it occurs when the quantity of beta chains produced are drastically reduced or absent. This typically occurs in homozygous state where the patient inherits two mutations of the β-globulin genes. Despite being healthy at birth due to the presence of a high percentage of fetal hemoglobin in their red blood cells, children with β-thalassemia major typically become symptomatic between six and twelve months of age when the normal switch of fetal gamma globulin chains to adult β-globulin chains becomes impaired, leading to profound anemia. Children with β-thalassemia major typically present to the healthcare system within the first two years of life with severe pallor and failure to thrive, which mandate regular blood transfusions to survive.

2. **β-thalassemia minor** represents the mildest form of β-thalassemia. It occurs when the patient inherits only one defective β-globulin gene (heterozygotes state) leading to mild reduction in β-globulin production. Such individuals are asymptomatic and do not usually need any treatment. Studies concerning people with beta-thalassemia minor were not included in this study.

3. **β-thalassemia intermedia** is a term used to define a group of patients with β-thalassemia in whom the clinical severity of the disease is somewhere between the major and minor forms of β-thalassemia. However, it represents a wide spectrum of severity where at one end there are patients suffering from severe symptoms, similar yet milder than, β-thalassemia major patients. While at the other end the individuals may be asymptomatic. The diversity of the clinical manifestations is related to underlying genetic
differences that impact the amount and the function of Hb. The more deleterious gene mutation, the more severe form of thalassemia produced and vice versa.

The following nomenclature describes the production of β-globin according to the genetic mutations of β-globin: β, normal; β+, decreased; and β0, absent. Combination of such mutations yields different severity of β-thalassemia (phenotypes) as depicted in Table 1.

Table 1. Genotype and phenotype of β-thalassemia*

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>β/β</td>
<td>Normal</td>
</tr>
<tr>
<td>β/β+</td>
<td>β-thalassemia minor</td>
</tr>
<tr>
<td>β/β0</td>
<td>β-thalassemia minor</td>
</tr>
<tr>
<td>β+/β+</td>
<td>β-thalassemia intermedia</td>
</tr>
<tr>
<td>β+/β0</td>
<td>β-thalassemia intermedia</td>
</tr>
<tr>
<td>β0/β0</td>
<td>β-thalassemia major</td>
</tr>
</tbody>
</table>

(* Genotype, the genetic characteristics of a disease; phenotype, the clinical/physical characteristics that caused by genotype.

Hemoglobin E/β-thalassemia is another distinct, yet common type of β-thalassemia, where the patient co-inherits a β-thalassemia allele from one parent, and the structural variant, Hb E, from the other parent. Co-inheritance produces remarkable clinical heterogeneity, ranging from mild to severe types.

Figure 1 illustrates the differences and similarities between various types of β-thalassemias.
As shown above, β-thalassemia major and severe E/β-thalassemia represent the most severe forms of β-thalassemia where patients are usually dependent on blood transfusion for their entire lives. Studies concerning people with β-thalassemia major and/or severe E/β-thalassemia were included in this study. For the purposes of this study, lifelong transfusion-dependent β-thalassemia (lifelong TDβT) refers to both diseases (β-thalassemia major and/or severe E/β-thalassemia).

**Management of lifelong TDβT**

The common goal in lifelong TDβT management is to correct the anemia and allow for normal growth and development. To date, blood transfusion is the mainstay of care for patients with lifelong TDβT. However, once started, the transfusion-related complications threaten patient safety and become a major source of morbidity and mortality.

Multiple blood transfusions can lead to accumulation of excessive iron in the body,
a condition known as iron overload. Approximately 1 milliliter of blood contains approximately 1 milligram of iron, thus for every 3–4 units of blood, 1 gram of iron enters the body. The average adult with lifelong TDβT requires at least monthly transfusion, receiving 12-50 units of blood per year. This process is clinically relevant, because the amount of iron increases quickly after repeated transfusions. More importantly, there are no physiologic means to remove excess iron. Ultimately, excess iron accumulates in tissues and disturbs normal body organ functions. Chelation therapy can be used to remove excess iron. Unless treated with iron chelation (therapy used to remove excess iron), iron overload can result in multiple organ damage such as liver failure, heart damage, endocrine failure (such as diabetes mellitus, growth failure and delayed onset of puberty) and early death.\(^8\) Despite the availability of different iron chelation therapies, compliance with chelation drugs is often poor; therefore, problems of iron overload can persist and lead to serious problems in the patient.\(^9\)

Due to the repeated blood transfusions in patients with lifelong TDβT, patients are at higher risk for transfusion-transmitted infections. These infections range from serious chronic viral infections, such as hepatitis and human immunodeficiency virus (HIV), to acute life-threatening bacterial infections. While the risk of infection is fairly low in developed countries, it is much higher in developing countries where thalassemia is most prevalent.\(^10,11\)

Repeated transfusions from different donors leads to the possible development of antibodies to foreign red cells, a phenomenon known as alloimmunization, a serious complication to providing future transfusions that can affect up to 20% of thalassemic patients.\(^12\)
Though transfusions may improve clinical symptoms by correcting the anemia, this intervention is not curative and is associated with numerous potential short and long-term risks and complications as described above. Hematopoietic stem cell transplantation (HSCT) may appear to be a possible curative option; it is a complex procedure and unfortunately has limited applicability and is only available to a minority of patients. Similarly, gene therapy for β-thalassemia is still at the investigational level.

Another treatment option includes fetal hemoglobin (HbF) inducers, such as hydroxyurea, which is described in detail in the next section.

1.3.2 Description of the intervention

Historical background

Hydroxyurea (HU) is an oral antimetabolite chemotherapeutic agent that has been used for several decades to manage a variety of medical disorders, particularly myeloproliferative diseases (MPD) and chronic myelogenous leukemia (CML).\textsuperscript{13,14} The first proof-of-principle experiments of HU for patients with sickle cell anemia (SCA) were reported in 1984, when Platt and colleagues showed a rapid increase in HbF containing reticulocytes without significant bone marrow toxicity.\textsuperscript{15} Since then, clinical studies have been accumulating for more than 30 years regarding the safety and efficacy of HU therapy for patients with SCA.\textsuperscript{16}

In the early 1990s, HU was first shown to raise HbF levels in β-thalassemia.\textsuperscript{17} Subsequently, HU has been tested in β-thalassemia over the last two decades in different cohort studies across different geographical areas that demonstrated encouraging results.\textsuperscript{18-25} There was, however, wide variability in the clinical response to HU in β-thalassemia. Certain clinical (splenectomy) and genetic (Xmn1 polymorphism) factors
have been investigated yet their mechanisms of action and their prediction of the response to HU have remained controversial.\textsuperscript{26}

\textit{How the intervention might work}

HU has been a well-recognized HbF inducer for several years, yet the exact molecular mechanism through which it induces HbF is not fully understood.\textsuperscript{26} HbF has been recognized as an important factor in ameliorating the clinical picture of β-thalassemia by reducing α/β globin chain imbalance, hence, improving chronic anemia and decreasing the need for blood transfusions.\textsuperscript{26} This ultimately leads to a decrease in transfusion-related complications and thereby improves patient quality of life (QoL).

\textit{Safety of the intervention}

One of the greatest concerns regarding the long-term use of HU is the possibility of causing secondary leukemia. This concern is due to HU effects on DNA synthesis and repair, as well as the anecdotal reports in MPD showing an increased risk of secondary leukemia which raises concerns about whether the long-term use of HU can cause leukemia.\textsuperscript{16} However, the long-term experiences with HU in SCA suggest sustained beneficial effects of HU in young people without excessive toxicity, deleterious effects on growth and development, altered fertility, or increased carcinogenicity.\textsuperscript{27-29}

\textbf{1.4 Study rationale}

HU may reduce the burden of blood transfusions in patients with lifelong TDβT, however it may also be associated with several serious adverse effects (AEs) that may offset any benefits. Thus, in-depth assessment of the efficacy and safety of HU in lifelong TDβT is highly warranted. To clarify the role of HU in the treatment of patients with lifelong TDβT and to improve physician awareness regarding its use, a meta-analysis
was conducted to summarize the available data on the efficacy and safety of HU in patients with lifelong TDβT.

A systematic review and meta-analysis can help inform future policy and clinical decision-making about the effectiveness of HU in lifelong TDβT. Knowledge translation of findings from the primary studies and comparison of clinical practice guidelines through systematic review of the current literature can help to provide evidence-based recommendations, which will ultimately improve patient safety and the quality of health care.

1.5 Overview of the methodological approach

In order to understand the methodological approach used in this research, it is important to first describe core concepts of evidence-based medicine (EBM) and how they relate to the meta-analysis approach. In 1992, EBM was introduced by Gordon Guyatt as the new model for practicing medicine and was defined as “ability to assess the validity and importance of evidence before applying it to day-to-day clinical problems”.30 The concept of EBM was further redefined in 1996 by Sackett as “the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients”.31 Subsequently, another revision was introduced in 2005 with the Sicily statement, which stressed the role of the patient and his or her wishes in the decision making process.32,33 Also, the Sicily statement recommended that the name to be changed from EBM to evidence-based practice (EBP) to reflect broader users of the healthcare professionals.90 Based on the latest Sicily statement, EBM can be envisioned as the intersection of patient values and preferences, research evidence and clinical expertise to make clinical decisions.90 For the remaining part of this thesis, the
term “EBP” will be used.

The challenge of attaining consensus among medical experts about EBP led to the creation of the first hierarchy of evidence, also referred to as levels of evidence. 34 Case studies and anecdotal evidence were graded lower on the ranking, while randomized controlled trials held were considered to offer the highest level of evidence. 34 Subsequent developments added meta-analyses of randomized clinical trials (RCTs) above single RCTs and the details of the middle-levels were worked out in greater detail. 35 The evidence hierarchy was designed to reflect the methodological strength of research studies. 34,35 It is assumed that better evidence on this scale is less likely to be affected by bias, more likely to correctly attribute causal powers to a particular treatment, and more generalizable. Figure 2 outlines one of the most updated hierarchy of evidence. 36

Figure 2. Levels of Evidence
Ideally, any clinical decision in medicine should be based upon the guidance that is drawn from a comprehensive review of the literature as part of EBP. Basing a clinical decision on expert opinion can however, be biased.\textsuperscript{37} A single clinical trial is rarely sufficient to provide a confident answer to a clinical question. Indeed, one analysis suggests that most research claims are ultimately proven to be incorrect or inaccurate when additional studies have been performed.\textsuperscript{38} More accurate conclusions need to be reached by inspecting all sources of data as rigorously and objectively as possible. This can be achieved through appropriately conducted systematic reviews and meta-analyses.

While there have been earlier attempts of statistical combination of data from several primary (individual) studies, the term meta-analysis was first defined by Glass in 1976 and since then the practice of using such research methodshas become fundamental in the health and social sciences.\textsuperscript{39}

1.5.1 Definitions and characteristics of systematic review and meta-analysis

Systematic review is a form of research that provides a summary of the literature on a specific question, using systematic and explicit methods to identify, select, critically appraise, and extract and analyze data from relevant research.\textsuperscript{40,41} A meta-analysis is a specific statistical technique used to pool the results from different studies, which generally aims to produce a single estimate of a treatment effect.\textsuperscript{41} These methods use the fundamental core principles of EBP (5As: Ask, Acquire, Appraise, Apply and Assess) to generate recommendations about a specific topic or question.\textsuperscript{40,41}

Several key steps are essential for conducting a systematic review or meta-analysis.\textsuperscript{41} These include:

- Developing a protocol
• Formulating research questions
• Comprehensive Searching of the literature
• Assessing the quality of studies
• Synthesizing the studies and drawing conclusions
• Discussing the strengths and weaknesses of included studies and exploring reasons for the heterogeneity of the results.

1.5.2 Strengths of systematic review and meta-analysis

There are many strengths in conducting systematic reviews and/or meta-analytic research. First, in contrast to narrative review articles or book chapters, most systematic reviews and meta-analyses focus on a narrow, clearly defined topic and include all eligible studies, not just those chosen by the author.\textsuperscript{42} Second, combining studies in meta-analyses increases the sample size and generally produces more precise estimates of the effect size than a single randomized trial.\textsuperscript{41} Third, experts rarely have the time, skills or resources to critically evaluate the body of evidence relevant to a particular clinical topic/question, and a systematic review and meta-analysis can facilitate such investigation. Finally, unique aspects to a single randomized trial, involving the participating patient population, the protocol, the setting in which the trial is performed, or the expertise of the involved clinicians, may limit its generalizability to other settings or individual patients. Thus, the conclusions of systematic reviews and meta-analyses may be more generalizable than single studies.\textsuperscript{41}

1.5.3 Disadvantages of systematic review and meta-analysis

Despite the above-mentioned advantages associated with systematic reviews and meta-analytic research, there are also disadvantages involved with this type of research design. First, the conclusion provided in a systematic review or meta-analysis of the
literature is only as valid and reliable as the methods used to estimate the effect in each of
the primary studies. Conducting a meta-analysis does not resolve issues that were
inherent in the design and implementation of the primary studies. This is also known as
the “garbage in, garbage out” phenomenon. While problematic, this issue can be
effectively mitigated by excluding poor quality studies in the meta-analysis.

The second issue associated with conducting a meta-analysis relates to publication
bias, or the “file drawer effect”. The file drawer effect refers to the tendency for studies
with significant results to be published, while studies with non-significant or inconclusive
results are often left in file drawers. Several reports have shown that studies with
positive results are more likely to be published than studies with negative or inclusive
results. However, this issue is not limited to systematic reviews or meta-analyses but is
common for any type of literature review. To overcome such an issue, a comprehensive
literature search should be broad and not limited to published studies. Grey literature
searches and contacting the field experts may mitigate this issue.

A third disadvantage of systematic reviews and meta-analyses is that conclusions
cannot be drawn by comparing and aggregating studies that include different measuring
techniques, definitions of variables, and participants because they are different; a
phenomenon known as “comparing apples to oranges” or heterogeneity. Subgroup
analyses according to the different studies or patients characteristics could help in
minimizing this issue.

Fourthly, simply labeling a review as a “systematic review” or “meta-analysis”
does not assure that such review was rigorously executed and reported. Poor reporting of
systematic reviews and meta-analysis reduce their usefulness to the decision-makers in
clinical practice. To overcome this issue, different consortiums have published guidelines on how to conduct and report systematic reviews and meta-analyses.\textsuperscript{47,48} For instance, the QUOROM (Quality Of Reporting Of Meta-analyses) Statement developed in 1996, published in 1999 and updated and renamed PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) to standardize and improve reporting of systematic reviews and meta-analyses.\textsuperscript{48} QUOROM and PRISMA emphasize that systematic reviews and meta-analyses should provide the protocol, data, and assessments of risk of bias from individual studies in sufficient detail to allow the reader to verify the results. More recently, PRISMA-P (Preferred reporting items for systematic review and meta-analysis protocols) was developed to guide the development of the most critical steps of conducting systematic reviews and meta-analyses at the protocols stage.\textsuperscript{49}

Finally, systematic reviews and meta-analyses help in bridging the gap between research evidence and decision-makers.\textsuperscript{50} However, the evidence generated by systematic reviews and meta-analyses does not, by itself, establish which clinical decisions need to be made in practice. This is because systematic reviews and meta-analyses may not always provide a comprehensive overview of all of the information that is relevant for decision-making. For instance, patients’ values and preferences as well as resource allocation are beyond the scope of systematic reviews and meta-analyses. Clinical practice guidelines, which are described in detail in the next section, may alleviate such issues by addressing the balance between benefits and harms for patients, patients’ own perspectives and the consideration of resource allocation.\textsuperscript{51}

1.6 Clinical practice guidelines
Clinical practice guidelines have been developed to facilitate the transition of evidence into practice and are important to EBP. Clinical practice guidelines are recommendations, made by expert panels, for clinicians that facilitate clinical decision-making. The Institute of Medicine defines clinical practice guidelines as “...statements that include recommendations, intended to optimize patient care, that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options”. Based on this definition, clinical practice guidelines have two main characteristics: 1) statements that are informed by a systematic review of the evidence, and 2) an assessment of the benefits and harms of alternative care options, as well as cost analysis. Clinical practice guidelines help move the research evidence of systematic reviews and meta-analyses to practical clinical recommendations by: 1) rating the confidence in the effect sizes (family of indices that measure the magnitude of a treatment effect), 2) assessing the benefits and harms of the intended treatment and alternative options, 3) considering patients’ values and preferences, and 4) considering resources for the intervention and its alternatives.

While there are many grading systems for recommendations, the three systems commonly used are: 1) Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) system, 2) the American Heart Association’s (AHA) system, and 3) the US Preventive Services Task Force (USPST) system. Although a detailed discussion of the differences between these systems is beyond the scope of this review, key points will be highlighted.

The GRADE system indicates four levels of confidence in the effect size based on quality of studies, and other factors that will be discussed later, as: high, moderate, low,
and very low, while the AHA and USPSTF systems specify three levels of confidence: A, B, and C in AHA, and high, moderate, and low in USPSTF. Furthermore, the three systems used a similar way to differentiate between recommendations that should be applied (or avoided) in all, or almost all, patients (i.e., strong recommendations), from those that require individualization to the patient’s values, preferences, and circumstances (i.e., weak recommendations). In general, the GRADE system is one of most commonly used system to evaluate or “grade” guidelines worldwide. GRADE is a well-developed formal process used to rate the quality of scientific evidence in systematic reviews and to develop recommendations in guidelines that are evidence-based. For the purpose of our thesis, we used the GRADE system to generate recommendations from the results determined from this study. Table 2 outlines the process of guideline development.

Table 2. GRADE process in developing recommendation*

<table>
<thead>
<tr>
<th>A. Systematic review steps:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formulation of clinical question(s)</td>
</tr>
<tr>
<td>Selection of the outcome(s)</td>
</tr>
<tr>
<td>Searching, retrieval, appraisal of evidence</td>
</tr>
<tr>
<td>Summarization of evidence</td>
</tr>
<tr>
<td>Rate the quality of evidence for outcome(s): High, moderate, low or very low</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. Guideline development steps:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate the overall quality of the evidence</td>
</tr>
<tr>
<td>Decide on the direction (for/against) and grade strength (strong/weak*) of the recommendation considering:</td>
</tr>
<tr>
<td>1. Quality of the evidence</td>
</tr>
<tr>
<td>2. Balance of desirable/undesirable outcomes</td>
</tr>
<tr>
<td>3. Values and preferences of patients</td>
</tr>
<tr>
<td>4. Resource consideration</td>
</tr>
</tbody>
</table>

* Modified from Guyatt et al

As depicted in Table 2, four factors (i.e., quality of the evidence, balance of desirable/undesirable outcomes, values and preferences of patients, and resource consideration) are considered when deciding the direction and strength of a GRADE
recommendation. According to the relationships between these four factors, the recommendations are then labeled as “strong” or “weak” or “for” or “against” as shown in Table 3.

Confidence in the effect size represents the extent to which the effect sizes or estimates are sufficiently trustworthy to support a particular management plan or action. Such confidence varies according to quality of the supporting evidence. More rigorous studies are more likely to provide results that are closer to the truth. For instance, a meta-analysis of well-designed RCTs has more confidence in its findings and conclusions compared to observational studies because of the superiority of an RCT as a study design. In the GRADE process, RCTs represent high-quality evidence, while observational studies are categorized as low-quality evidence to support estimates of intervention effects.

Although the confidence in the effect size is essential for the strength of recommendations, the two entities are distinct: both high or low confidence in effect sizes may be associated with either strong or weak recommendations as depicted in Table 3.
Table 3. Determination of the strength of GRADE recommendation

<table>
<thead>
<tr>
<th>Overall confidence in effect size</th>
<th>Strong recommendation</th>
<th>Weak recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High/moderate confidence (or low/very low in special circumstances)</td>
<td>Low/very low confidence</td>
</tr>
<tr>
<td></td>
<td>AND</td>
<td>OR</td>
</tr>
<tr>
<td>Balance between benefits and harms</td>
<td>The benefits clearly outweigh the harms or vice versa</td>
<td>Not a big difference between benefits and harms</td>
</tr>
<tr>
<td></td>
<td>AND</td>
<td>OR</td>
</tr>
<tr>
<td>Patients’ values and preferences</td>
<td>Almost all informed patients will make the same choice</td>
<td>Variable patient values and preferences; patients chooses</td>
</tr>
<tr>
<td></td>
<td>AND</td>
<td>OR</td>
</tr>
<tr>
<td>Resource considerations</td>
<td>The cost is justified</td>
<td>The cost is not justified</td>
</tr>
</tbody>
</table>

Within the context of this study, a comprehensive evaluation of the role of HU in a patient’s lifelong TDβT use was warranted given the substantial risks and complications associated with current standard treatments (blood transfusions) for patients with lifelong TDβT in this study. To address the efficacy and safety of HU in patients with lifelong TDβT, a systematic review and meta-analysis was conducted.
Chapter 2: A Meta-Analysis of Hydroxyurea Use for β-thalassemia: Implications for Clinical Practice and Medical Education

2.1 Introduction

β-thalassemia is one of the most common inherited diseases worldwide; it is characterized by a reduced ability to produce hemoglobin. Although β-thalassemia is common among people originating from the Mediterranean, Middle East, Central Asia, India, and Southern China, it is no longer limited to those geographical areas due to migration to different regions of the world.59

Lifelong transfusion-dependent β-thalassemia (lifelong TDβT) constitutes the most severe form of β-thalassemia. It includes β-thalassemia major and severe E/β-thalassemia. Despite progress in therapy, such as hematopoietic stem cell transplantation (HSCT), gene therapy and fetal hemoglobin (HbF) inducing agents, chronic blood transfusions remain the standard therapy for the majority of lifelong TDβT.7,60,61 Chronic blood transfusions carry significant risks, such as acute life-threatening events (anaphylaxis, bacterial infection and acute hemolytic reaction), infection, and can result in iron overload that can cause significant multi-system organ damage.62

Recognizing the risks of chronic blood transfusions to patient safety and their non-curative nature, numerous investigators have tried different drugs in order to avoid these risks and improve patient quality of life (QoL).26 Hydroxyurea (HU), an HbF inducer, is expected to decrease the need for blood transfusions and has gained the attention of researchers.26 Over the last two decades, numerous published studies of HU in severe β-thalassemia have shown promising results;18-22,24,25,63-68 however, current practice guidelines of β-thalassemia management overlook it as a treatment option for
The gap in knowledge is partially explained by the absence of rigorous meta-analyses for the use of HU in lifelong TDβT patients, since none currently exist. A meta-analysis investigating the use of HU in lifelong TDβT may change the current practice of how lifelong TDβT patients are treated; therefore, the objective of this study was to conduct a meta-analysis evaluating the efficacy and safety of HU as a potential treatment for patients with lifelong TDβT.

2.2 Methods

2.2.1 Data sources and searches

A comprehensive systematic search of the literature was conducted to evaluate the clinical efficacy and safety of HU in patients with lifelong TDβT of any age. MEDLINE (1946 to April 2015), EMBASE (1974 to April 2015), and Cochrane Central Register of Controlled Trials (CENTRAL) (March 2015) were searched using the following keywords: “Hydroxyurea”, “Hydroxycarbamide”, “Hydrea”, or “Droxia” and “Thalassemia.” Clinical trials registries (ClinicalTrials.gov and World Health Organization International Clinical Trials Registry Platform [ICTRP]) and major conference proceedings (American Society of Hematology [ASH] and European Hematology Association [EHA]) over the last five years were searched. Hand searches were also conducted using reference lists from primary studies. Searches were not restricted by language, publication date or publication type but for human participants only.

Since the expected retrieval studies in this area were primarily of observational study designs, our methods adapted the MOOSE (Meta-analysis Of Observational Studies in Epidemiology) and PRISMA (Preferred Reporting Items for Systematic
reviews and Meta-Analyses) guidelines for meta-analyses. The meta-analysis protocol for this study was registered with Prospective Register of Systematic Reviews (PROSPERO; registration number: CRD42014010138). As this study was based on the systematic review of previously published literature and there was no potential for participant identification, ethical approval was not required to conduct this study.

2.2.2 Study selection

One reviewer (A.A.) screened the citations, first by title and abstract, then by review of the complete article as indicated. RCTs and observational studies (sample size ≥ 10) that assessed the clinical efficacy and/or safety of HU alone, for three months or longer, in patients with lifelong TDβT (β-thalassemia major and/or severe E/β-thalassemia), of any age, were eligible for inclusion. Exclusion criteria included β-thalassemia intermedia, non-transfusion dependent β-thalassemia (NTDβT), mild/moderate E/β-thalassemia, any combination therapy with HU, case reports, case series, or studies with a sample size of less than ten patients. If a study included various types of β-thalassemia patients who were treated with HU, lifelong TDβT patients were included but remaining patients were excluded.

Since transfusion dependency (TD) is a common feature for both lifelong TDβT and β-thalassemia intermedia or NTDβT, this can lead to ambiguity about the type of β-thalassemia. As such, we further categorized TD β-thalassemia to either lifelong TDβT (patients who required lifelong monthly transfusions) or others (β-thalassemia intermedia/NTDβT) according to the primary authors’ definition/labels of the disease. In cases with no clear distinction in the primary study (i.e. only provided as TD
βthalassemia), we asked the primary authors to clarify this issue. If that was not possible, those patients/studies were excluded.

2.2.3 Data abstraction and quality assessment

Two reviewers (A.A. and A.K.) extracted the data independently, by using data extraction forms. Key characteristics were extracted from eligible studies and recorded. These characteristics included: author, year and country of publication, study design, sample size, age of sample, blood transfusion history, HU dose, response rate, types of AEs and follow-up duration. Information was requested from primary authors when it was not available in the published papers.

The quality of the included studies were assessed independently by two reviewers (A.A. and A.K.) by using the National Institutes of Health (NIH) quality assessment tool for before-after (pre-post) studies with no control group. The tool is composed of 12 questions assessing the potential risk for selection bias, information bias, measurement bias, and confounding. Reviewers then used these questions to judge each study to be of “good,” “fair,” or “poor” quality. Disagreements between reviewers were resolved by discussion until consensus was reached.

2.2.4 Data synthesis and analysis

The effect size of our meta-analysis was the response rate of HU in lifelong TDβT patients in decreasing the blood transfusion needs. The response rate of successful treatment with HU was categorized into complete response rate (CRR), where there was a complete cessation of blood transfusion post HU therapy, and overall response rate (ORR), where there was ≥ 50% reduction of transfusion needs post HU therapy. Since all of the included studies were single-arm designs with no control arms, pooled estimates of
the treatment effect for each outcome (CRR and ORR) across the studies were calculated as proportions (the responders over treated sample size with HU), together with their 95% confidence intervals (95% CI). These analyses were performed using a recently published “metaprop” Stata command developed for the use of proportions as effect sizes. Due to the expected heterogeneity within and between studies, we used the random effects model, which takes into consideration between-study and within-study variation and provided a more conservative analysis of the studies than the fixed effect model.

**2.2.5 Heterogeneity and publication bias**

Heterogeneity between studies was assessed by visual inspection of forest plots to detect overlapping 95% CIs, by a formal statistical test of the significance of the heterogeneity (chi-squared test; \( p < 0.05 \)) and by estimation of the percentage of heterogeneity between trials that could not be attributed to sampling error (\( I^2 \)). The \( I^2 \) is expressed as percentage; with value of 25% indicating low, \( I^2 \) value of 50% indicating moderate, and \( I^2 \) value of 75% indicating high heterogeneity, respectively.

To explore heterogeneity among studies, we conducted several sub-analyses, which were as follows: (i) A sensitivity analyses to explore whether a specific study strongly influenced the results, by excluding one study at a time; (ii) Sub-group analysis according to the age at first blood transfusion; (iii) Meta-regression to evaluate the influence of age at first blood transfusion, transfusion thresholds or triggers (hemoglobin level [with or without clinical symptoms] that had to be reached before a blood transfusion was administered), and sample size on the efficacy of HU.

Publication bias was assessed graphically using funnel plots and the Egger test which quantified the asymmetry of the plot. The latter tests the null hypothesis that
small studies give the same results as large studies. An alpha level of $p < 0.05$ was
deemed statistically significant for the statistical calculations in this meta-analysis. All
data were analyzed using Stata, version 13.0.80

2.3 Results

2.3.1 Study selection

The initial literature search yielded 943 references, after duplicates were removed,
and was updated to April 2015. One author (A.A.) then screened the titles and abstracts
of 943 references and 895 were excluded for not meeting the review’s eligibility criteria.
The remaining 48 references were assessed on the basis of their full text for inclusion or
exclusion using the criteria indicated above. Of these, 30 studies were excluded; with the
most common reason being inadequate reported outcome data. Eleven studies met our
inclusion criteria for the meta-analysis; see Figure 3 for PRISMA study flow diagram.
2.3.2 Study characteristics

The studies were published between 2004 and 2014 and conducted in four countries: six in Iran; three in India; one each in Algeria and Pakistan. All of the included studies were single-arm, pre-post design, with no comparison group, all were prospective but one, which was a retrospective study. The pre-study arm was used as historical control and the post-study arm was used as the treatment group where HU was given. All included studies were published in English and as full-text papers. The studies collectively enrolled 859 patients. Sample sizes ranged from n=11 to n=248. Study populations were a mixture of children and adults in all studies. Study characteristics are summarized in Table 4.
### Table 4. Summary characteristics of included studies [ordered according to publication year].

<table>
<thead>
<tr>
<th>Author</th>
<th>Country</th>
<th>Design</th>
<th>Age* [range] in years</th>
<th>Disease</th>
<th>Age at 1st Tx [range] in month</th>
<th>Blood Tx</th>
<th>HU* [mg/kg/d]</th>
<th>N</th>
<th>Overall response: n [%]</th>
<th>Complete response: n [%]</th>
<th>AEs</th>
<th>F/U [month]</th>
<th>Quality [NIH tool]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bordbar 2014</td>
<td>Iran</td>
<td>Pre/Post</td>
<td>20.3 [2-50]</td>
<td>β-TM</td>
<td>All within 1st 24</td>
<td>Q 2-4 weeks</td>
<td>10.5</td>
<td>95</td>
<td>31 [33]</td>
<td>6 [6]</td>
<td>Transient neutropenia and transaminitis</td>
<td>8</td>
<td>Good</td>
</tr>
<tr>
<td>Italia 2010</td>
<td>India</td>
<td>Pre/Post</td>
<td>[8-34]</td>
<td>Severe E/β Thalassemia</td>
<td>66 [NR]</td>
<td>12-15/y</td>
<td>15-20</td>
<td>11</td>
<td>NR</td>
<td>4 [36]</td>
<td>No significant AEs</td>
<td>20</td>
<td>Fair</td>
</tr>
<tr>
<td>Zamani 2009</td>
<td>Iran</td>
<td>Pre/Post</td>
<td>18.4 [10-40]</td>
<td>β-TM</td>
<td>NR</td>
<td>Q 3-4 weeks</td>
<td>10±5</td>
<td>49</td>
<td>44 [90]</td>
<td>12 [24]</td>
<td>No significant AEs.</td>
<td>60</td>
<td>Fair</td>
</tr>
<tr>
<td>Kosaryan 2009</td>
<td>Iran</td>
<td>Pre/Post</td>
<td>17.5 [NR]</td>
<td>β-TM</td>
<td>74.4 [NR]</td>
<td>NR</td>
<td>15.5</td>
<td>248</td>
<td>NR</td>
<td>111 [44.8]</td>
<td>No significant AEs.</td>
<td>89.5</td>
<td>Fair</td>
</tr>
<tr>
<td>Alebouyeh 2004</td>
<td>Iran</td>
<td>Pre/Post</td>
<td>16.3 [6-33]</td>
<td>β-TM</td>
<td>NR</td>
<td>Q 4 weeks</td>
<td>20</td>
<td>36</td>
<td>26 [72]</td>
<td>25 [69]</td>
<td>No significant AEs</td>
<td>48</td>
<td>Fair</td>
</tr>
</tbody>
</table>

AEs, adverse events; β-TM, β-thalassemia major; F/U, follow-up; NIH, National Institutes of Health; N, sample size; NR, not reported; Tx, transfusion; (*), mean or median; (◆), retrospective; (◆), obtained from the primary author.
Most of studies enrolled β-thalassemia major patients only with the exception of two studies; one enrolled severe E/β-thalassemia patients only and another enrolled a mixture of β-thalassemia major & E/β-thalassemia patients. All patients were transfusion dependent and receiving regular blood transfusions, at least once per month. Three studies enrolled patients whose first blood transfusion was given within the first 24 months of age for all participants. Of note, four studies did not report the age at the first blood transfusion.

Although all of included studies but four stated their blood transfusion thresholds, the trigger points were different, ranging from hemoglobin (Hb) of 6 - 8 g/dL. Moreover, four studies based their transfusion decisions solely on the Hb level where the remaining three studies used both Hb level and clinical indications, such as symptomatic anemia and concurrent infection.

In all studies, HU was used as single intervention with no other treatment apart from regular blood transfusions. HU was given as a single daily oral agent in all but one study, four days a week. The dose of HU was similar across all the studies, ranging from 10-20mg/kg/day. The response to HU was measured in similar ways and entailed a decrease in the need for transfusions. The studies used HU for 3 – 6 months before judging its efficacy. The mean or median duration of the studies ranged from 6 - 89 months.

2.3.3 Study quality

Assessment of study quality by using the NIH quality assessment tool for before-after (pre-post) studies with no control group raised several potential methodological limitations as shown in Table 5. For example, the lack of pre-specified
eligibility/selection criteria for the study population was a common issue in most of the included studies. Although all of the participants in the studies represent what would occur in clinical practice, it was not clear if all eligible participants were in fact enrolled; raising concerns about selection bias. Four studies had small sample sizes thereby lacking generalizability of their findings. All studies did not state whether people assessing the outcomes were blinded to the participants' interventions. This could bias the results by affecting the actual outcomes of the participants in the studies.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Was the study question or objective clearly stated?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>2. Were eligibility/selection criteria for the study population prespecified and clearly described?</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>3. Were the participants in the study representative of those who would be eligible for the test/service/intervention in the general or clinical population of interest?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>4. Were all eligible participants that met the prespecified entry criteria enrolled?</td>
<td>Yes</td>
<td>CD</td>
<td>CD</td>
<td>CD</td>
<td>CD</td>
<td>CD</td>
<td>CD</td>
<td>CD</td>
<td>CD</td>
<td>CD</td>
<td>CD</td>
</tr>
<tr>
<td>5. Was the sample size sufficiently large to provide confidence in the findings?</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>6. Was the test/service/intervention clearly described and delivered consistently across the study population?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>7. Were the outcome measures prespecified, clearly defined, valid, reliable, and assessed consistently across all study participants?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>8. Were the people assessing the outcomes blinded to the participants’ exposures/interventions?</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>9. Was the loss to follow-up after baseline 20% or less? Were those lost to follow-up accounted for in the analysis?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>10. Did the statistical methods examine changes in outcome measures from before to after the intervention? Were statistical tests done that provided p values for the pre-to-post changes?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>11. Were outcome measures of interest taken multiple times before the intervention and multiple times after the intervention (i.e., did they use an interrupted time-series design)?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>12. If the intervention was conducted at a group level (e.g., a whole hospital, a community, etc.) did the statistical analysis take into account the use of individual-level data to determine effects at the group level?</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

**Final Quality Rating**

|                   | Good | Fair | Good | Fair | Fair | Fair | Fair | Fair | Fair | Good | Fair | Fair |

CD, cannot determine; NA, not applicable; NIH, National Institute of Health.
2.3.4 Response rate of HU

HU was associated with complete cessation of regular blood transfusion for lifelong TDβT patients i.e. converting them from TD to a transfusion- independent state with a CRR of 26% (95% CI, 13-41%); heterogeneity was considered high with an $I^2$ of 95%, p-value <0.01; as depicted in Figure 4.

**Figure 4.** Forest plots of HU responses in lifelong TDβT patients. A. Complete responserate (complete cessation of blood transfusion); B. Overall response rates (≥50 reduction in blood transfusion need).
In addition, HU was associated with a significant decrease (≥ 50%) in transfusion need among lifelong TDβT patients with an ORR of 60% (95% CI, 41-78%); heterogeneity again was also high with an $I^2$ of 95.2% and p-value <0.01; as shown in Figure 4.

Predictors of clinical response to HU were assessed in some studies. For instance, a splenectomy was associated with good response in two studies \(^{20,21}\) but neutral in three studies. \(^{18,67}\) Although one study did not assess splenectomy as a predictor, all of the enrolled patients were splenectomized and HU overall response was substantially high. \(^{23}\) Similarly, presence of Xmn1 polymorphism was a positive predictor in three studies \(^{18-20}\) but neutral in one study. \(^{68}\)

All studies reported AEs, which were transient and improved with temporary cessation of the drug and/or adjustment of the dose. No long-term AEs like cancer, infertility, or end organ damages were reported.

### 2.3.5 Analysis of heterogeneity and publication bias

Graphical inspection of forest plots for CRR and ORR in Figure 3, revealed non-overlapping 95% CIs, suggesting heterogeneity among studies. This was confirmed by statistically significant high, $I^2 = 95\%$, $p<0.01$ for both CRR and ORR.

To analyze heterogeneity, several sub-analyses were performed. First, sensitivity analyses were conducted by removing one study at a time however the heterogeneity remained high ($I^2 > 75\%$) for both the CRR and ORR effect sizes. Second, subgroup analyses according to the age at the first blood transfusion were conducted. As depicted in Table 6, three categories were generated; (i) Studies where all participants were <24 months at time for their first blood transfusion: three studies. (ii) Studies with patients’
mean age at first blood transfusion ≤ 24 months: two studies (iii) Studies with patients’ mean age at first blood transfusion >24 months: two studies.

Table 6. Subgroup Analysis of HU complete response rate

<table>
<thead>
<tr>
<th>Age at first Transfusion (months)</th>
<th>Study # (Patient #)</th>
<th>CRR (95% CI)</th>
<th>Heterogeneity of included studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1: All ≤ 24</td>
<td>3 (207)</td>
<td>6% (0-12%)</td>
<td>P = 70.4, p = 0.03</td>
</tr>
<tr>
<td>Group 2: mean ≤ 24</td>
<td>2 (164)</td>
<td>25% (19-33%)</td>
<td>P = 0, p = 0.33</td>
</tr>
<tr>
<td>Group 3: mean &gt; 24</td>
<td>2 (259)</td>
<td>44% (38-51%)</td>
<td>P = 0, p = 0.62</td>
</tr>
</tbody>
</table>

As shown in Table 6, the CRR varied according to the group; with the lowest being 6% for group one and highest being 44% for group three, suggesting a significant interaction between the age at first transfusion and the response rate of HU. Of note, the heterogeneity was moderate for group one (70.4%) and zero for groups two and three.

Third, a meta-regression analysis showed significant change in complete response rate (CRR) to HU after controlling for mean age at first blood transfusion in all studies that reported the mean age at first blood transfusion (four studies) (beta co-efficient, 0.18; p = 0.003), adjusted R-squared of 100%, explaining between study variance. The I² was 0.0%; see Figure 5.
Figure 5. Meta-regression plot. HU complete response rate is plotted against the mean age at first blood transfusion, with each circle representing one study in the analysis, and the red line indicating the pooled estimated rates. Circle size is relative to the standard error of the complete response rate.

An analysis using meta-regression, however, did not show significant change in clinical response to HU (CRR) after controlling for transfusion threshold (beta coefficient, 0.09; p = 0.41), adjusted R-squared was 0%; nor for after controlling for sample size (beta co-efficient, 0.001; p = 0.32), adjusted R-squared of 5.72%.

Although intended, it was not possible to perform subgroup analysis for the age of participants (pediatric vs. adult), HU dose (10 vs. 20 mg/kg/day), and type of thalassemia (β-thalassemia major vs. severe E/β-thalassemia) due to an insufficient number of studies that provided data for such analysis. In addition, we were unable to do either subgroup nor meta-regression analyses for the ORR estimate due to the limited number of studies.
Egger’s test showed no evidence of publication bias of studies on CRR ($p = 0.34$) and ORR ($p = 0.7$). This was consistent with the symmetry of the funnel plots as shown in Figure 6.

**Figure 6.** Funnel plots of HU responses. A. CRR; B. ORR
2.4 Discussion

The aim of this meta-analysis was to evaluate the efficacy and safety of HU alone in patients with lifelong TDβT. Eleven studies met our inclusion criteria, but all were observational studies. These studies enrolled 859 patients. The results of the meta-analysis indicated that HU was associated with a complete cessation of blood transfusion in one quarter of study participants. Moreover, 50% reduction of blood transfusion was seen in 60% of the study participants. However, there was high heterogeneity among studies. Such heterogeneity was observed despite stringent inclusion/exclusion criteria and classification of the disease, focusing on a subset of patients requiring lifelong blood transfusion.

The heterogeneity among studies was not unexpected given the poor correlation between the genotype and phenotype of β-thalassemia. Furthermore, blood transfusion practices differ from one region to another depending on institutional policies and blood availability, as depicted in the included studies. In addition, the clinical and genetic predictors of HU response are less known and the controversy will continue given the studies revealed some discrepancy about the role of splenectomy and Xmn1 in predicting the HU response.

Consistent with a previous study, it was observed from the subgroup analyses as well as in this meta-regression that a significant interaction between the age at first blood transfusion and the clinical response to HU existed where the younger the patient, the less responsive they were to HU. This may reflect the differential severity of lifelong TDβT where the most severe form typically presents early and requires transfusion support within the first two years of age. However, even in the severe group, it was observed
that a meaningful response to HU in 28% (95% CI, 21-36%) of patients can occur and patients’ transfusions can be reduced by 50%.

Adverse event profiles of HU were only addressed narratively in most of the included studies, which limited us from performing a meta-analysis on the AEs. However, most patients tolerated the drug well and the majority of the side effects were transient and completely resolved upon decreasing or temporarily stopping the drug. In some studies, the drug was resumed subsequently (with lower dose and titrating up gradually) without significant issues. There were only a few patients who could not tolerate the drug and the decision was made by the authors of the studies to discontinue treatment. The most common reported AEs included transient bone marrow suppression, mild elevation of liver enzymes, nausea, and vomiting. There were no documented long-term AEs like leukemia, any cancer type, or any chronic organ damage such as liver or kidney dysfunction among patients with lifelong TDβT; however, the follow-up duration was not long enough to completely determine the incidence, if any, of the long-term AEs. There was no reported mortality directly related to drug usage in the studies.

Only three published cases of leukemia post HU in β-thalassemia were identified in a comprehensive search of the literature including more than 1500 β-thalassemic patients being treated over 20 years. Two were unlikely related to HU due to a very short interval between the usage of the drug and development of leukemia in one case and the retrospective suspicion of coexistence of chronic leukemia in another case prior to the use of HU. This left only one case of suspected association between HU use and leukemia (chronic myelogenous leukemia after 5 years of HU in β-thalassemia intermedia). This is reassuring and consistent with a comprehensive systematic review.
and meta-analysis that specifically addressed the long-term carcinogenicity of HU among the non-malignant conditions, where no association between long-term use of HU and leukemia was found.  

Although quality of life (QoL) was not assessed in the included studies of this meta-analysis (apart from subjective assessment regarding patient perceptions of HU therapy), it was evident that a decrease or a complete cessation of blood transfusion could be considered a substantial gain, both clinically and for patient QoL.

Chronic blood transfusions can be a significant burden on many, if not all, of the developing countries where most thalassemic patients live. HU is a an economical drug and costs around $95 for one month supply for an average 70kg adult while the cost of one unit of blood is $316\(^{87}\) (minimum of two units per month) and between $1500 – 3760 for one month supply for an average 70kg adult for iron chelators deferoxamine and deferasirox, respectively.\(^{88}\)

The main limitations to this meta-analysis are the following: (1) The conclusions are based on a limited number of observational studies, with a relatively small sample size, (2) There was an absence of control arms in the included studies, (3) There was relatively short follow-up periods for the HU treatment, (an average of 1-2 years, in most studies which limits the conclusion regarding the long-term efficacy and safety), (4) There was also potential of selection bias across most studies due to the lack of prespecified eligibility/selection criteria in most of the included studies.

Strengths of this study include it being the first comprehensive meta-analysis to evaluate the clinical efficacy of HU in lifelong TDβT patients, with rigorous assessment of quality and statistical heterogeneity of the included studies. Furthermore, this meta-
analysis was based on a homogeneous group, focusing on severe and clinically challenging forms of β-thalassemia (lifelong TDβT) as opposed to β-thalassemia intermedia or a mixture of both diseases. Finally the extensive literature searches of the major biomedical databases; ongoing trials, grey literature, conference proceedings and hand-searched journals of potential interest executed could also be considered strengths of this study.

Despite the above-mentioned limitations, and acknowledging the significant need to conduct robust experimental studies in this field, the results of this meta-analysis, suggest a potential usefulness of HU therapy. We could recommend the use of HU in treating adults and children with lifelong TDβT given the results of this meta-analysis as well as the following implications: 1) the serious consequences of chronic blood transfusions and their complications, 2) the growing evidence of the long-term safety of HU, 3) the rapid response to HU in lifelong TDβT (within a few months), and 4) the availability and affordability of HU (especially for the developing countries).

Given these results, we could recommend a pilot trial of HU for a minimum of three to six months in patients with lifelong TDβT to assess the efficacy in individual patients following discussions with the patient and family members along with a structured monitoring plan to ensure efficacy and safety.

In conclusion, HU appears to be potentially effective in the management of lifelong TDβT by decreasing the need for chronic blood transfusions either completely or partially across several studies. It appears to be well tolerated and associated with mild and transient AEs. Patients with lifelong TDβT may benefit from a trial of HU, though
large RCTs assessing efficacy should be carried out to confirm the findings of this meta-analysis.

2.5 Acknowledgments

The authors thank the authors of the studies included in our meta-analysis for providing additional information about their studies at our request.

2.6 Authorship

Contribution: A.A. envisioned and designed the study, wrote the protocol, searched the literature, extracted the data, evaluated the quality of eligible studies, analyzed and interpreted the data, and wrote the manuscript; A.K. edited the protocol, extracted the data, evaluated the quality of eligible studies, analyzed and interpreted the data and edited the manuscript; N.W. edited the protocol, interpreted the data and edited the manuscript; and E.O.P. edited the protocol, interpreted the data and edited the manuscript.

Conflicts of interest disclosure: All authors declare no competing financial interests.

Correspondence: Ali Algiraigri, University of Calgary, 2500 University Drive Northwest, Calgary, AB T2N 1N4; email: aalgiraigri@gmail.com.
Chapter 3: Clinical Significance, Implications and Knowledge Translation

3.1 Clinical significance and summary of main findings

Despite the major progress in the understanding of the molecular biology and genetics of β-thalassemia, the clinical course of the disease remains challenging and the current mainstay of therapy may be suboptimal for the majority of lifelong TDT patients. Since its introduction in 1960 for thalassemia, blood transfusion remains the current standard therapy for the vast majority of lifelong TDT patients. Longitudinal observational follow-up studies have shown that chronic blood transfusion succeeded in prevention of the early deaths of thalassemia in infancy. However, it did not cure but rather transformed this disorder into a chronic disease, with patients usually dying in their second decade of life from the complications of transfusional iron overload. This led to the introduction of iron chelation therapy in the 1970s, which helped in decreasing the burden of the iron overload and its related complications. However, compliance with such therapy, which is a lifelong regimen, remains challenging. Problems of iron overload can be serious and still pose risks to patients today. Although, HSCT is the only available curative option, only a minority of patients are candidates for this procedure.

Recently, a novel therapeutic approach (induction of HbF) was hypothesized for β-thalassemia. Hydroxyurea (HU), an HbF inducer, has gained scientists’ attention and has been studied in different β-thalassemia types. Although, it appeared promising, it has not yet reached the clinical practice guidelines as a potential treatment option. As such, the current meta-analysis was conducted to provide a new insight into the current status and future perspectives of HU for treating lifelong TDT.

The best available evidence that can be drawn from our meta-analysis indicates
that HU is a potential effective drug in the management of lifelong TDβT by decreasing the need for chronic blood transfusions. It appears to be well tolerated and associated with mild and transient AEs. While these results are encouraging, well-designed and executed RCTs are needed to confirm the findings of this meta-analysis.

3.2 Study limitations

Conducting a meta-analysis to answer the question about the efficacy and safety of HU in lifelong TDβT revealed some limitations. These limitations present the opportunity for areas of future research and practice. First, the summary findings of this meta-analysis were based on a limited number of observational studies. To overcome this issue, the literature search was comprehensive by searching multiple databases, by examining ongoing or recently published studies in an abstract forms, by using broad keywords for the systematic search, and by updating the comprehensive search every six months.

One of the most compelling and overarching limitations to the present meta-analysis are the limitations of the individual studies included in the meta-analysis. These limitations include the lack of randomization and blindness, absence of control arms, and relatively small sample sizes. It is well known that the well conducted double-blind, randomized, placebo-controlled trials with large sample sizes represent one of the highest level of research evidence where the results of such trials allow the investigators to draw more strong and generalized conclusions regarding the effectiveness of the interventions.90

Another limitation of this study is that the effect sizes of the meta-analysis (CRR and ORR) are associated with high heterogeneity. Such heterogeneity was observed
Despite the effort of minimizing the clinical heterogeneity by implementing stringent inclusion/exclusion criteria. However, this heterogeneity was explained by subgroup analyses and meta-regression that revealed a significant interaction between the age at first blood transfusion and the clinical response to HU where the younger the patient, the less responsive they were to HU.

Finally, decision-making in clinical practice depends on individual patients and their circumstances, specific factors need to be considered and discussed, such as the availability of the treatment (HU and blood transfusions), care providers’ experience, patients’ values of the treatment, and predictors of the clinical response to HU. Our systematic review and meta-analysis did not provide such data, as this information was not reported consistently in the primary studies.

3.3 Implications

3.3.1 Implications for clinical practice guidelines

The overall confidence in the effect size for the reduction of blood transfusion and serious AEs is low. This is because of the type of studies included in this meta-analysis, which were all observational studies without a control arm. Conversely, the balance between benefits and harms showed a clear advantage for the use HU. The significant reduction of blood transfusion (26% and 60% for complete and overall response rate) and the transient/mild AEs of HU led to this. Furthermore, the growing evidence of the long-term safety of HU is very reassuring where the drug has been used now for more than 17 years in SCA and β-thalassemia without causing any leukemia.85 Regarding the patients’ values and preferences, it is obvious that the net benefits are large, given the complete cessation or even 50% reduction of blood transfusions and its related complications.
However, further research would be needed to confirm this. Additionally, it must be determined whether informed patients would be willing to tolerate the long-term use of HU to gain a significant reduction in the short and long-term risks of blood transfusions. Still, given the results of the present meta-analysis, it is anticipated that informed patients would choose long-term use of HU.

When considering resource allocation, HU is likely associated with minimal expense due to the low cost of the drug. By integrating these factors as suggested by Guyatt, the benefits of HU that could lead to cessation or reduction of blood transfusions outweigh the risks of HU use (i.e., mild AEs and the burdens associated with being on an long-term drug).

Regarding the strength of the recommendation, while the low overall confidence in effect sizes may have necessitated a “weak” recommendation, the benefits of HU including its low cost, as well as the likelihood that most informed patients would make the same choice of using HU, converge in providing a compelling argument to grade the recommendation of HU use in TDT patients as “strong”. As such, we recommend that physicians use HU as a trial treatment for a minimum of three to six months to explore its efficacy. Of course, this needs to be discussed with the individual patient and family members, along with a structured monitoring plan for efficacy and AEs.

Thus, the current practice guidelines for β-thalassemia major, as shown in Figure 7, could be modified based on the results of the present meta-analysis by incorporating a trial of HU as depicted in Figure 8.
Figure 7. Current practice approach for lifelong TDβT patients

- Lifelong TDβT patient
  - Needs or receiving chronic blood transfusions
    - Candidate for HSCT?
      - Yes → HSCT
      - No → Chronic blood transfusions
        - Monitor & manage complications

Figure 8. Proposed management algorithm for lifelong TDβT patients

- Lifelong TDβT patient
  - Needs or receiving chronic blood transfusions
    - Candidate for HSCT?
      - Yes → HSCT
      - No → Trial of HU
        - Effective & safe?
          - Yes → Continue HU & monitor for AEs
          - No → Chronic blood transfusions, monitor & manage AEs
Duration of HU trial therapy

According to our meta-analysis as well as guidelines of sickle cell anemia (SCA), a clinical response to treatment with HU may take three to six months. Therefore, a six-month trial on HU may be required prior to declaring its failure in treating the patient.

3.3.2 Implications for practice

Although the current evidence does not compare different implementation protocols for HU, there is a significant need to implement a standard treatment protocol to ensure adequate dosing and strict monitoring of HU use in patients with lifelong TDβT. In addition, recognizing the need to provide guidance to help physicians and physicians-in-training learn about and potentially use this drug effectively and safely, we recommend a series of clinical practice points for common issues that surround the use of HU in lifelong TDβT. These could be viewed as a suggested treatment protocol. These recommendations were based on the present meta-analysis as well as the published guidelines for NTDβT and SCA.

Pre-HU:

I. All individuals and/or their families should be counseled prior to commencing therapy relating to the known and potential risks of hydroxyurea in β-thalassemia major, as well as the need to avoid pregnancy and continue in long-term follow up.

II. Clinical baseline assessments should include growth parameters, spleen and liver sizes, and any extra medullary enlargements.

III. Laboratory baseline assessments should include complete blood count (CBC) with white blood cell (WBC) differential, reticulocyte count, platelet count, and red
blood cell (RBC) mean corpuscular volume (MCV). Quantitative measurement of HbF (by hemoglobin electrophoresis or high-performance liquid chromatography [HPLC]). Comprehensive metabolic profile, including renal and liver function tests.

IV. Baseline assessment of iron overload status and current iron chelation therapy

V. Pregnancy testing when appropriate

Initiation and escalation of HU

I. HU starting doses for adults and children should be 10 mg/kg/day

II. Dose can be escalated by 5 mg/kg/day every 1-2 month to the maximal tolerated dose, but not exceeding 20 mg/kg/day.

III. If not already on, concurrent folic acid supplementation is recommended.

Post-HU:

I. Response evaluation should be done after 3-6 months of therapy. A decrease in the transfusion needs by at least 30-50% at six months is a meaningful response definition. If no response was seen, HU should be discontinued. Patients showing response should be re-evaluated at 12, 18, and 24 months to ensure maintenance of response. Quality of life and growth parameters could also be monitored.

II. Close monitoring is essential for the success of HU therapy. Monitor CBC with WBC differential, reticulocyte count, and renal and hepatic functions at least once a month when adjusting dosage.

III. If myelosuppression (neutropenia or thrombocytopenia) occurs:
   - Hold hydroxyurea dosing
   - Monitor CBC with WBC differential weekly
• When blood counts have recovered, reinstitute hydroxyurea at a dose 5
  mg/kg/day lower than the dose given before onset of cytopenias.

IV. Safety evaluation should be implemented and may include hepatic and renal
  function studies, once a month. Also, clinical history and physical examination for
  gastrointestinal, neurologic, or dermatologic adverse events, every three months.

V. If a pregnancy is planned, it is prudent to discontinue the drug three months
  before conception.

3.3.3 Implications for research

There seems no reason to justify the absence of RCTs of HU for lifelong TDβT,
since the need for an alternative medical intervention is highly warranted and
encouraging results from several observational studies have been published in the
literature. Also, HU trials in similar conditions such as SCA have been conducted and
showed positive results, which led to Food and Drug Administration (FDA) approval of
HU for SCA in the United States.93,94 An ideal design of a future study would be that of a
randomized, placebo-controlled trial. Double or triple blinding design would be feasible
whereby neither the patient nor the person administering the treatment nor the person
evaluating the response to treatment knows which treatment any particular patient is
receiving. To investigate whether the findings previously reported are consistent and
sustained, these future studies should be larger and longer in duration than those reported
in this review and perhaps in form of large, multicenter RCTs

Additional research is also required to address the many other areas with little or
no evidence that were identified during the writing of this meta-analysis. There remain
many unanswered questions related to the role of HU in lifelong TDBT as discussed below. Particular issues for future trials include:

I. The long-term efficacy and safety of HU in lifelong TDBT is highly important.

Although few observational studies revealed promising sustained responses and a reassuring toxicity profile at longer follow-up periods\textsuperscript{22,95,96}, most of published studies evaluated outcomes at relatively short intervals such as 1-2 years. Longitudinal observational follow-up studies are needed to confirm the sustainability of the response as well as any long-term toxicity of HU.

II. The optimal HU dosage in $\beta$-thalassemia needs to be determined as different investigators used varying doses of HU in their studies. However, most of published studies reported clinical responses pointing toward an effective yet safe dose range of 10-20 mg/kg/day. Whether a start dose of a fixed low-dose regimen versus an escalated dose approach according to toxicity (maximal tolerated dose) is better remains to be answered. Of note, the maximum tolerated dose in $\beta$-thalassemia patients remains lower than those used in patients with SCA.\textsuperscript{26} This dose could be up to a maximum of 20 mg/kg per day. Recent RCT in $\beta$-thalassemia intermedia compared two fixed dose-regimens, 10 versus 20mg/kg/day, and found that the lower dose (10mg/kg/day) was superior.\textsuperscript{97} Whether this low-dose regimen is effective and safe for life-long TDBT needs further investigation.

III. Patients’ socioeconomic status (SES) could potentially play an important role both on physicians’ clinical management decisions as well as on patients’ compliance with treatment. This could impact the response rate when long term
use of HU is needed. Since all of the included studies for our meta-analysis were published in developing countries, it may be useful to study the effects of HU in the developed countries where more health resources are available and where transfusion practices may differ. This may promote a well-structured RCT, which typically needs clinical expertise, human resources and financial support.

IV. Lifelong TDβT care can be complex, multidisciplinary and expensive. This represents an enormous burden for the non-industrialized countries where health-spending budgets may be limited. Studies that investigate the cost/benefit of HU compared to the current conventional therapy are necessary. This could be assessed as secondary objective of experimental or observational studies.

V. Alongside the cost-effectiveness, assessment of patients’ quality of life pre- and post-HU treatment is also crucial. Similarly, this could be assessed as one of the secondary objectives of RCTs or observational studies.

VI. There is great debate about the predictors of effective response to HU in lifelong TDβT patients such as, splenectomy status, β-globin genotype, α-globin genotype, and molecular determinants of increased HbF production. Although randomization according to specific predictors might be challenging, subgroup analysis may reveal certain associations. This will need well-designed and adequately powered trials to answer such questions. For this purpose, multicenter collaboration should be considered and fostered.

3.3.4 Implications for EBP and medical education

The GRADE tool (as discussed in Chapter 1) was used to provide recommendations for or against the use of HU in patients with lifelong TDβT. Figure 9 outlines the EBP
concepts that were used to develop final recommendations about the use of HU in patients with lifelong TDβT found from the results of this research.

![Evidence-based practice framework](image)

**Figure 9.** Evidence-based practice framework used to generate clinical evidence for HU in patients with lifelong TDβT.

Although the presence of clinical practice guidelines is important and helpful in day-to-day practice, they should not restrict the medical community and specifically, medical educators, from examining the best evidence. As shown in this research, five of the current β-thalassemia clinical practice guidelines overlooked the use of HU as a treatment option, when in fact several studies have shown there may be promising evidence in favor of HU. This finding led to the recommendation of HU as a treatment trial for 3 to 6 months.
Well-designed RCTs represent one of the highest levels of evidence. However, absence of such RCTs does not necessarily mean a lack of clinical evidence. The current concept of evidence-based practice involves searching and synthesizing the best “available” evidence that answers clinical questions. This puts significant emphasis on teaching and empowering future physicians with fundamental EBP tools that enable them to ask, acquire, appraise, apply and evaluate the evidence. It is still debatable when EBP teaching should take place and there is little evidence as to the most effective way curriculum should teach EBP in medical school and residency. Nevertheless, researchers have suggested that the longitudinal approach, across all years of medical school and residency programs, provides learners with EBP skills across different levels of their career development, thereby facilitating the building on of each encounter from the previous one.

Although most EBP curricula are incorporated in the clinical years (clerkship or internship) of undergraduate medical education (UME), early introduction in the preclinical years may provide opportunities to teach the fundamentals of EBP. This may ease the transition by increasing students’ comfort of relevant concepts which can assist students with the expected uncertainty of being novices within clinical practice. Similarly, a longitudinal approach would be useful in postgraduate medical education (PGME), where more clinical content could be incorporated into EBP training to increase its relevance and connect EBP to the clinical setting and learning environment. This may potentially generate ongoing opportunities to build on what has been learned in UME curricula, as well as provide a clinical context. For UME and PGME, learning the levels of evidence and utilizing existing resources, such as librarians and available electronic
databases, may be critical steps for the practical application of EBP principles such as searching, acquiring and critically appraising the evidence effectively. In addition, it is important to generate networking opportunities between residents and librarians for future research project collaborations.

While, searching, gathering, appraising and synthesizing evidence are crucial steps in EBP, the evidence alone cannot influence clinical decision making. Each treatment plan must be individualized to reflect the specific characteristics of the patient. As previously stated, in addition to the overall confidence in effect size, three other factors must be considered when determining the direction and strength of a clinical recommendation. These factors include: 1) balance between benefits and harms, 2) consideration of patients’ values and preferences, and 3) resource considerations. The present study on the use of HU in patients with lifelong TDβT illustrates the importance of taking all of these factors into consideration, as well as the value of systematic review/meta-analytic methods, as they relate to EBP.

While it is important to adopt and teach the most recent guideline recommendations, it is also important to realize the time gap between published guidelines and the ongoing research. As research evidence is continuously growing and changing, physicians may not be able to provide optimal patients care if they rely only on the medical knowledge that has been gained during their training. This is why it is so important to teach methods of EBP in medical education as well. With respect to this study, the role of HU in patients with lifelong TDβT revealed that the most recent β-thalassemia guideline (published in 2014) didn’t address the potential role of HU as a treatment option.
3.4 Knowledge translation and implications for medical education

Although knowledge translation is a complex process, its fundamental goal is moving knowledge to action.\textsuperscript{103} The goal then is to translate the results of this meta-analysis into clinical practice. In a busy clinical practice, such methods of research synthesis help practicing hematologists to understand available evidence in a critically appraised and coherent manner. However, the most effective way of knowledge translation is to endorse such results so that they are incorporated into the clinical practice guidelines. Such translational process will need acceptance and willingness of key stakeholders in thalassemia management such as the experts in this field, along with patients who responded to HU to support and endorse our recommendations for changes to current clinical practice guidelines.

In order to achieve such an effect, the first step is to facilitate knowledge dissemination in order for experts to take notice of these findings. Such dissemination will allow for experts in the field to engage in dialogue pertaining to the results from this meta-analysis and its recommendations. This may help to endorse then a new treatment consensus and recommendations about the use of HU in patients with lifelong TDβT.

To achieve such knowledge dissemination, the results of this research have been published, as online abstract form, in different worldwide educational conferences including the second largest hematology conference worldwide, the European Hematology Association (EHA), in Vienna, Austria, June 11-14, 2015. Also, an abstract of such results was published in the preceding book of the American Society of Pediatric Hematology/Oncology (ASPHO) annual meeting, May 6-9, 2015. In addition, the abstract was also published in the American Society of Hematology journal,
Blood. Finally, the meta-analyses manuscript will be submitted for publication in the well-known hematology journal, the British Journal of Haematology (BHJ) that has an impact factor of 4.959.

In order for advances in medical education to take place in the realm of hematology, endorsement by key experts and changes in clinical practice guidelines first need to take place. Knowledge creation has three essential phases as depicted in Figure 10: 1) knowledge inquiry (such as the primary studies that evaluated the role of HU in lifelong TDβT); 2) Knowledge synthesis (such as a meta-analysis addressing the research question for this thesis); 3) Knowledge product (such as a clinical practice guidelines). As knowledge moves through each phase, it becomes more refined and useful to target knowledge users, which in this case would be practicing hematologists as well as hematology residents and medical students. In conducting this meta-analysis, we have sustained the second stage of knowledge synthesis in order to proceed to the next stage of action toward end users in creating a knowledge product that will involve changes in clinical practice guidelines. To facilitate the third stage, we will bring these results alongside with the suggested treatment approach HU trial to national and international experts to promote the addition of a trial of HU in the next edition of the guidelines.
In summary, this study is the first meta-analysis conducted to evaluate the efficacy and safety of HU for patients with lifelong TDBT. While the results of this meta-analysis are encouraging, further work needs to be done in forms of RCTs to consolidate these positive findings. In addition, our meta-analysis revealed substantial need to conduct high quality clinical studies since none exited in this area of study.
3.5 References


80. *StataCorp. 2013. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP* [computer program].


91. Taher A, Vichinsky E, Musallam K, Cappellini MD, Viprakasit V. *Guidelines for the management of non transfusion dependent thalassaemia (NTDT).* Thalassaemia International Federation, Nicosia, Cyprus; 2013.


Appendix A: PRISMA Flow Diagram

PRISMA 2009 Flow Diagram

Identification

- # of records identified through database searching
- # of additional records identified through other sources

Screening

- # of records after duplicates removed
- # of records screened
- # of records excluded

Eligibility

- # of full-text articles assessed for eligibility
- # of full-text articles excluded, with reasons

Included

- # of studies included in qualitative synthesis
- # of studies included in quantitative synthesis (meta-analysis)

For more information, visit [www.prisma-statement.org](http://www.prisma-statement.org).

Appendix B. Search keywords in MEDLINE, EMBASE, and CENTRAL

**Medline (OVID)**
1. exp Hydroxyurea/
2. hydroxyurea*.mp.
3. hydroxycarbamide*.mp.
4. hydrea*.mp.
5. droxia*.mp.
6. 1 or 2 or 3 or 4 or 5
7. exp Thalassemia/
8. thalassemi*.mp.
9. 7 or 8
10. 6 and 9
11. Limit 10 to humans

**EMBASE**
1. exp hydroxyurea/
2. hydroxyurea*.mp.
3. hydroxycarbamide*.mp.
4. hydrea*.mp.
5. droxia*.mp.
6. 1 or 2 or 3 or 4 or 5
7. exp thalassemia/
8. thalassemi*.mp.
9. 7 or 8
10. 6 and 9
11. Limit 10 to human

**COCHRANE**
1. Hydroxyurea/
2. hydroxyurea*.mp.
3. hydroxycarbamide*.mp.
4. hydrea*.mp.
5. droxia*.mp.
6. 1 or 2 or 3 or 4 or 5
7. Thalassemia/
8. thalassemi*.mp.
9. 7 or 8
10. 6 and 9
11. Limit 10 to human
### Table I. Summary characteristics of included studies.

<table>
<thead>
<tr>
<th>Author</th>
<th>Country</th>
<th>Design</th>
<th>Age* [range] in year</th>
<th>Disease</th>
<th>Age at 1st Tx [range] in month</th>
<th>Blood Tx HU* [mg/kg/d]</th>
<th>N</th>
<th>Overall response: n [%]</th>
<th>Complete response: n [%]</th>
<th>AEs</th>
<th>F/U [month]</th>
<th>Quality [NIH tool]</th>
</tr>
</thead>
</table>

*AEs, adverse events; F/U, follow-up; NIH, National Institutes of Health; N, sample size; NR, not reported; Tx, transfusion; *, mean or median*
Appendix D. Quality Assessment Tool for Before-After (Pre-Post) Studies With No Control Group.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Yes</th>
<th>No</th>
<th>Other (CD, NR, NA)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Was the study question or objective clearly stated?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Were eligibility/selection criteria for the study population prespecified and clearly described?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Were the participants in the study representative of those who would be eligible for the test/service/intervention in the general or clinical population of interest?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Were all eligible participants that met the prespecified entry criteria enrolled?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Was the sample size sufficiently large to provide confidence in the findings?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Was the test/service/intervention clearly described and delivered consistently across the study population?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Were the outcome measures prespecified, clearly defined, valid, reliable, and assessed consistently across all study participants?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Were the people assessing the outcomes blinded to the participants' exposures/interventions?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Was the loss to follow-up after baseline 20% or less? Were those lost to follow-up accounted for in the analysis?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Did the statistical methods examine changes in outcome measures from before to after the intervention? Were statistical tests done that provided p values for the pre-to-post changes?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Were outcome measures of interest taken multiple times before the intervention and multiple times</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
after the intervention (i.e., did they use an interrupted time-series design)?

12. If the intervention was conducted at a group level (e.g., a whole hospital, a community, etc.) did the statistical analysis take into account the use of individual-level data to determine effects at the group level?

Quality Rating (Good, Fair, or Poor) (see guidance)

Rater #1 initials:

Rater #2 initials:

Additional Comments (If POOR, please state why):

*CD, cannot determine; NA, not applicable; NR, not reported

Guidance for Assessing the Quality of Before-After (Pre-Post) Studies With No Control Group

The guidance document below is organized by question number from the tool for quality assessment of controlled intervention studies.

**Question 1. Study question**

Did the authors describe their goal in conducting this research? Is it easy to understand what they were looking to find? This issue is important for any scientific paper of any type. Higher quality scientific research explicitly defines a research question.
**Question 2. Eligibility criteria and study population**

Did the authors describe the eligibility criteria applied to the individuals from whom the study participants were selected or recruited? In other words, if the investigators were to conduct this study again, would they know whom to recruit, from where, and from what time period?

Here is a sample description of a study population: men over age 40 with type 2 diabetes, who began seeking medical care at Phoenix Good Samaritan Hospital, between January 1, 2005 and December 31, 2007. The population is clearly described as: (1) who (men over age 40 with type 2 diabetes); (2) where (Phoenix Good Samaritan Hospital); and (3) when (between January 1, 2005 and December 31, 2007). Another sample description is women who were in the nursing profession, who were ages 34 to 59 in 1995, had no known CHD, stroke, cancer, hypercholesterolemia, or diabetes, and were recruited from the 11 most populous States, with contact information obtained from State nursing boards.

To assess this question, reviewers examined prior papers on study methods (listed in reference list) when necessary.

**Question 3. Study participants representative of clinical populations of interest**

The participants in the study should be generally representative of the population in which the intervention will be broadly applied. Studies on small demographic subgroups may raise concerns about how the intervention will affect broader populations of interest. For example, interventions that focus on very young or very old individuals may affect
middle-aged adults differently. Similarly, researchers may not be able to extrapolate study results from patients with severe chronic diseases to healthy populations.

**Question 4. All eligible participants enrolled**

To further explore this question, reviewers may need to ask: Did the investigators develop the I/E criteria prior to recruiting or selecting study participants? Were the same underlying I/E criteria used for all research participants? Were all subjects who met the I/E criteria enrolled in the study?

**Question 5. Sample size**

Did the authors present their reasons for selecting or recruiting the number of individuals included or analyzed? Did they note or discuss the statistical power of the study? This question addresses whether there was a sufficient sample size to detect an association, if one did exist.

An article's methods section may provide information on the sample size needed to detect a hypothesized difference in outcomes and a discussion on statistical power (such as, the study had 85 percent power to detect a 20 percent increase in the rate of an outcome of interest, with a 2-sided alpha of 0.05). Sometimes estimates of variance and/or estimates of effect size are given, instead of sample size calculations. In any case, if the reviewers determined that the power was sufficient to detect the effects of interest, then they would answer "yes" to Question 5.

**Question 6. Intervention clearly described**
Another pertinent question regarding interventions is: Was the intervention clearly defined in detail in the study? Did the authors indicate that the intervention was consistently applied to the subjects? Did the research participants have a high level of adherence to the requirements of the intervention? For example, if the investigators assigned a group to 10 mg/day of Drug A, did most participants in this group take the specific dosage of Drug A? Or did a large percentage of participants end up not taking the specific dose of Drug A indicated in the study protocol?

Reviewers ascertained that changes in study outcomes could be attributed to study interventions. If participants received interventions that were not part of the study protocol and could affect the outcomes being assessed, the results could be biased.

**Question 7. Outcome measures clearly described, valid, and reliable**

Were the outcomes defined in detail? Were the tools or methods for measuring outcomes accurate and reliable—such as, have they been validated or are they objective? This question is important because the answer influences confidence in the validity of study results.

An example of an outcome measure that is objective, accurate, and reliable is death—the outcome measured with more accuracy than any other. But even with a measure as objective as death, differences can exist in the accuracy and reliability of how investigators assessed death. For example, did they base it on an autopsy report, death certificate, death registry, or report from a family member? Another example of a valid study is one whose objective is to determine if dietary fat intake affects blood cholesterol level (cholesterol level being the outcome) and in which the cholesterol level is measured...
from fasting blood samples that are all sent to the same laboratory. These examples would get a "yes."

An example of a "no" would be self-report by subjects that they had a heart attack, or self-report of how much they weight (if body weight is the outcome of interest).

**Question 8. Blinding of outcome assessors**

Blinding or masking means that the outcome assessors did not know whether the participants received the intervention or were exposed to the factor under study. To answer the question above, the reviewers examined articles for evidence that the person(s) assessing the outcome(s) was masked to the participants' intervention or exposure status. An outcome assessor, for example, may examine medical records to determine the outcomes that occurred in the exposed and comparison groups. Sometimes the person applying the intervention or measuring the exposure is the same person conducting the outcome assessment. In this case, the outcome assessor would not likely be blinded to the intervention or exposure status. A reviewer would note such a finding in the comments section of the assessment tool.

In assessing this criterion, the reviewers determined whether it was likely that the person(s) conducting the outcome assessment knew the exposure status of the study participants. If not, then blinding was adequate. An example of adequate blinding of the outcome assessors is to create a separate committee whose members were not involved in the care of the patient and had no information about the study participants' exposure status. Using a study protocol, committee members would review copies of participants' medical records, which would be stripped of any potential exposure information or
personally identifiable information, for prespecified outcomes.

**Question 9. Followup rate**

Higher overall followup rates are always desirable to lower follow up rates, although higher rates are expected in shorter studies, and lower overall followup rates are often seen in longer studies. Usually an acceptable overall follow up rate is considered 80 percent or more of participants whose interventions or exposures were measured at baseline. However, this is a general guideline.

In accounting for those lost to follow up, in the analysis, investigators may have imputed values of the outcome for those lost to follow up or used other methods. For example, they may carry forward the baseline value or the last observed value of the outcome measure and use these as imputed values for the final outcome measure for research participants lost to follow up.

**Question 10. Statistical analysis**

Were formal statistical tests used to assess the significance of the changes in the outcome measures between the before and after time periods? The reported study results should present values for statistical tests, such as p values, to document the statistical significance (or lack thereof) for the changes in the outcome measures found in the study.

**Question 11. Multiple outcome measures**

Were the outcome measures for each person measured more than once during the course of the before and after study periods? Multiple measurements with the same result
increase confidence that the outcomes were accurately measured.

**Question 12. Group-level interventions and individual-level outcome efforts**

Group-level interventions are usually not relevant for clinical interventions such as bariatric surgery, in which the interventions are applied at the individual patient level. In those cases, the questions were coded as "NA" in the assessment tool.

**General Guidance for Determining the Overall Quality Rating of Before-After Studies**

The questions in the quality assessment tool were designed to help reviewers focus on the key concepts for evaluating the internal validity of a study. They are not intended to create a list from which to add up items to judge a study's quality.

Internal validity is the extent to which the outcome results reported in the study can truly be attributed to the intervention or exposure being evaluated, and not to biases, measurement errors, or other confounding factors that may result from flaws in the design or conduct of the study. In other words, what is the ability of the study to draw associative conclusions about the effects of the interventions or exposures on outcomes?

Critical appraisal of a study involves considering the risk of potential for selection bias, information bias, measurement bias, or confounding (the mixture of exposures that one cannot tease out from each other). Examples of confounding include co-interventions, differences at baseline in patient characteristics, and other issues throughout the questions above. High risk of bias translates to a rating of poor quality; low risk of bias translates to a rating of good quality. Again, the greater the risk of bias, the lower the quality rating of the study.
In addition, the more attention in the study design to issues that can help determine if there is a causal relationship between the exposure and outcome, the higher quality the study. These issues include exposures occurring prior to outcomes, evaluation of a dose-response gradient, accuracy of measurement of both exposure and outcome, and sufficient timeframe to see an effect.

Generally, when reviewers evaluate a study, they will not see a "fatal flaw," but instead will find some risk of bias. By focusing on the concepts underlying the questions in the quality assessment tool, reviewers should ask themselves about the potential for bias in the study they are critically appraising. For any box checked "no" reviewers should ask, "What is the potential risk of bias resulting from this flaw in study design or execution?" That is, does this factor lead to doubt about the results reported in the study or doubt about the ability of the study to accurately assess an association between the intervention or exposure and the outcome?

The best approach is to think about the questions in the assessment tool and how each one reveals something about the potential for bias in a study. Specific rules are not useful, as each study has specific nuances. In addition, being familiar with the key concepts will help reviewers be more comfortable with critical appraisal. Examples of studies rated good, fair, and poor are useful, but each study must be assessed on its own.

Last Updated March 2014.
Appendix E. Excluded studies

Table 4. Characteristics of excluded studies [ordered alphabetically]

<table>
<thead>
<tr>
<th>Ref</th>
<th>Study</th>
<th>Reason for exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1]</td>
<td>Ajanta 2002</td>
<td>Not enough data to calculate outcome</td>
</tr>
<tr>
<td>[3]</td>
<td>Amoozgar 2011</td>
<td>Not enough data to calculate outcome</td>
</tr>
<tr>
<td>[4]</td>
<td>Ansari 2013</td>
<td>Abstract which published later as full-text article</td>
</tr>
<tr>
<td>[5]</td>
<td>Ansari 2013</td>
<td>Already included in other study</td>
</tr>
<tr>
<td>[6]</td>
<td>Ansari 2007</td>
<td>Already included in other study</td>
</tr>
<tr>
<td>[9]</td>
<td>Banan 2013</td>
<td>Review</td>
</tr>
<tr>
<td>[10]</td>
<td>Banan 2012</td>
<td>Case series</td>
</tr>
<tr>
<td>[12]</td>
<td>Biswas 2014</td>
<td>Not enough data to calculate outcome</td>
</tr>
<tr>
<td>[13]</td>
<td>Bradai 2003</td>
<td>Already included in other study</td>
</tr>
<tr>
<td>[14]</td>
<td>Choudhry 1997</td>
<td>Different dosage regimen</td>
</tr>
<tr>
<td>[16]</td>
<td>Karimi 2010</td>
<td>Not enough data to calculate outcome</td>
</tr>
<tr>
<td>[17]</td>
<td>Karimi 2010</td>
<td>Abstract which published later as full-text article</td>
</tr>
<tr>
<td>[18]</td>
<td>Karimi 2012</td>
<td>Not enough data to calculate outcome</td>
</tr>
<tr>
<td>[20]</td>
<td>Loukopoulos 1998</td>
<td>Sample size &lt; 10</td>
</tr>
<tr>
<td>[21]</td>
<td>Mtarelidze 2008</td>
<td>Sample size &lt; 10</td>
</tr>
<tr>
<td>[22]</td>
<td>Olivieri 1998</td>
<td>Review</td>
</tr>
<tr>
<td>[23]</td>
<td>Pourfarzad 2013</td>
<td>Not enough data to calculate outcome</td>
</tr>
<tr>
<td>[24]</td>
<td>Rabbani 2014</td>
<td>Not enough data to calculate outcome</td>
</tr>
<tr>
<td>[25]</td>
<td>Ramanan 2013</td>
<td>Combination therapy (HU and Wheatgrass)</td>
</tr>
<tr>
<td>[26]</td>
<td>Rongchi 2013</td>
<td>Editorial</td>
</tr>
<tr>
<td>[27]</td>
<td>Sharma 2010</td>
<td>Sample size &lt; 10</td>
</tr>
<tr>
<td>[28]</td>
<td>Tafrali 2013</td>
<td>Not enough data to calculate outcome</td>
</tr>
<tr>
<td>[29]</td>
<td>Taher 2010</td>
<td>Not enough data to calculate outcome</td>
</tr>
</tbody>
</table>
References


parameters and cardiac function of patients with beta-thalassemia intermedia," European Journal of Haematology, vol. 84, pp. 52-58, Jan 2010.


