PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Stem cell transplantation of matched sibling donors compared with immunosuppressive therapy for acquired severe aplastic anemia – a Cochrane Systematic Review
AUTHORS	Peinemann, Frank; Labeit, Alexander

VERSION 1 - REVIEW

REVIEWER	Seiji Kojima MD,PhD Department of Pediatrics, Nagoya University Graduate School of Medicine, Nagoya, Japan
REVIEW RETURNED	28-Mar-2014

GENERAL COMMENTS	The authors tried to evaluate the effectiveness and adverse events of first alloSCT and IST in patients with severe aplastic anemia by metaanalysis. However, only 3 trials satisfied the criteria and total number of patients were too small ,only 302 patients. Moreover, all data were collected 15 to 30 years ago. The treatments were out of date. The results may cause misleading for modern practice.

REVIEWER	Robert Hills Cardiff University, Cardiff UK
REVIEW RETURNED	24-Apr-2014

GENERAL COMMENTS	Thank you for giving me the opportunity to review this paper. I have answered "No" to a couple of the questions given above - a better response would be "not sure". This is a well-performed analysis of the effect of sibling allograft for aplastic anaemia. The analysis is using the "Donor/No Donor" analysis proposed by Gray and Wheatley. This method, and its limitations, could be better described here. It compares patients with a donor with those with no donor. It doesn't compare sibling allograft with no sibling allograft. It is important that the compliance with transplant is given here. How many of the "donor" patients were transplanted - some may not have been for various reasons. Additionally, this method looks only at matched sibling allograft - matched unrelated donor would be in the no donor group, and the assumption here is that most patients with donors get allografts now, rather than after progression, and that there are very few alternative transplants performed, or those that are are of limited value. The limitations in the modern era of this approach in AML (its original

	setting) are well known and require some rehearsing here. In particular, the arms should be called donor vs no donor not SCT vs not. Additionally, given that the outcome is OS, I'm not sure how high the risk of bias is from not blinding - death is a pretty objective endpoint. Please therefore analyse as donor vs no donor (to remove selection and zero timeshift biases) and give compliance levels - what transplants were performed.

REVIEWER	Jacqueline Milton Boston University USA
REVIEW RETURNED	25-Apr-2014

GENERAL COMMENTS	1. This study examined subjects with ages ranging from early childhood to adulthood. The discussion describes a variety to studies in which patients younger than 40 years of age had 75% to 90% change of long-term cure with allogeneic bone marrow transplantation and a three to five year survival rate between 75% and 95% in younger patients. Is there reason to believe that age may be confounding the association treatment (transplantation vs. immunosuppressive therapy) and mortality? 2. The authors combined those undergoing immunosuppressive therapy as comparator with either 1)antithymocyte, 2)antilyphocyte or 3) a combination of the two. Is there any evidence of a difference in outcome between these 3 methods?

VERSION 1 – AUTHOR RESPONSE

Reviewer #1: comment to Author **Seiji Kojima MD,PhD; School of Medicine, Nagoya, Japan**	**Response by author**
The authors tried to evaluate the effectiveness and adverse events of first alloSCT and IST in patients with severe aplastic anemia by metaanalysis. However, only 3 trials satisfied the criteria and total number of patients were too small ,only 302 patients. Moreover, all data were collected 15 to 30 years ago. The treatments were out of date. The results may cause misleading for modern practice.	The findings result from the available study data. The time period in which the data were collected and the consequences were addressed in the manuscript. Quoted from the section 'Strengths and limitations': *All data were collected about 15 up to more than 30 years ago. Thus, the results may not be applicable to current modern standard care.*

Reviewer #2: comment to Author **Robert Hills; Cardiff University, Cardiff UK**	**Response by author**
This is a well-performed analysis of the effect of sibling allograft for aplastic anaemia. The analysis is using the "Donor/No Donor" analysis proposed by Gray and Wheatley. This method, and its limitations, could be better described here. It compares patients with a donor with those with no donor. It doesn't compare sibling allograft with no sibling allograft. It is important that the compliance with transplant is given here. How many of the "donor" patients were transplanted - some may not have been for various reasons. Additionally, this method looks only at matched sibling allograft - matched unrelated donor would be in the no donor group, and the assumption here is that most patients with donors get allografts now, rather than after progression, and that there are very few alternative transplants performed, or those that are are of limited value. The limitations in the modern era of this approach in AML (its original setting) are well known and require some rehearsing here. In particular, the arms should be called donor vs no donor not SCT vs not.	First, sibling, related, or unrelated donors constitute a heterogeneous pool. In the past, survival differed considerably among those groups. We intended from the start to reduce heterogeneity by confining to sibling donors. Second, we tried to keep information about 'Mendelian randomization' as short as possible and we refered to two papers that are best suited to explain the principles behind this specific concept. We want to take up the suggestion of the reviewer to broaden the text by providing more explanagory information about this method. Therefore, we added the following section in the method chapter: *Principle of 'Mendelian randomization'* *There are ethical concerns around randomization of patients with severe aplastic anemia to transplantation versus non-transplantation. In general, MSD-HSCT is a life-threatening treatment that can lead to early severe adverse events including death. Gray 1991 and Wheatley 2004 described the potential of 'Mendelian randomization' to minimize bias when comparing MSD-HSCT with an alternative therapy. The base concept has been ascribed to Katan 1986. 'Mendelian randomization' means the view that nature itself has already 'randomized' the paternal and maternal part of a gene given that donor and recipient are siblings. Thus, 'Mendelian randomization' by definition accepts only siblings as transplant donors and these sibling donors are required to have 'identical' or matched features of specific transplant-relevant HLA sites when compared with the transplant recipient. Therefore, patients with an HLA-matched sibling will be allocated to the MSD-HSCT group. On the other hand, patients with siblings that are not HLA compatible will be allocated to the immunosuppressive therapy group. The term 'Mendelian randomization' refers to the fact that the genetic distribution of paternal and maternal alleles follows a random process and is determined before birth. This concept takes advantage of an instrumental variable for allocating the patients to treatment groups and, at the same time, this variable is neither associated with the treatment nor associated with the outcome.* We added the following sentences with respect to 'Mendelian randomization' to the section 'Characteristics of included articles': *The authors of all included studies did not report that 'Mendelian randomization' was planned and the authors did not report the size of the involved families. The authors did not report the numbers of siblings and the results of the individual genetic analyses.* We added the following sentences with respect to

	'Mendelian randomization' to the section 'Strengths and limitations' of the discussion chapter: *Use of 'Mendelian randomization' is no guarantee that bias is minimized.* *This may be because tissue typing data may not be accurate. Patients may have only one sib-ling either in the donor or in the no donor group. Large families have a greater chance of find-ing a donor. Therefore, designing a non-randomized controlled trial by applying 'Mendelian randomization' requires careful thought to effectively reduce bias and control for potential confounders. There is a time lag in patients with siblings because tissue typing and readiness for assignment to treatment group may possibly take several months (Wheatley 2004). On the other hand, patients with no siblings can be assigned immediately and are at earlier risk for adverse events.* *Nitsch 2006 described the limits to causal inference based on 'Mendelian randomization' [29].* We want to address a quote from the reference: Gray R, Wheatley K. How to avoid bias when comparing bone marrow transplantation with chemotherapy. Bone Marrow Transplant 1991;7(Suppl 3):9-12. "The value of allogeneic BMT can, however, be assessed unbiasedly using 'Mendelian randomisation', i.e. comparing patients whose siblings are HLA-compatible with those who are not." We also want to address a quote from reference: Wheatley K, Gray R. Commentary: Mendelian randomization: an update on its use to evaluate allogeneic stem cell transplantation in leukaemia. Int J Epidemiol 2004;33:15-17. "Thus, we need an alternative, unbiased method for evaluating SCT. The answer is provided by Mendelian, or genetic, randomization. [...] In the haematological context, for a patient's sibling to be a suitable donor, he/she must have inherited the same tissue type as the patient from their mother and father. Since the chances of there being a match depend on the random assortment of genes at fertilization, only one in four siblings will be expected to have the same tissue type as the patient. Thus, whether or not a patient has a matched sibling donor available is essentially a random process and the presence or absence of a donor can be used as a surrogate for randomization."
Additionally, given that the outcome is OS, I'm not sure how high the risk of bias is from not blinding - death is a pretty objective endpoint.	We agree with the reviewer. We added the following sentences with respect to risk of bias to the section 'Characteristics of included articles': *We judged a low risk of bias for blinding the assessment of overall mortality. Blinding or lack of blinding is not expected to make a difference concerning overall mortality.*
Please therefore analyse as donor vs no donor (to remove selection and zero timeshift biases) and give compliance levels - what transplants were	We would like to refer to our response expressed in the first row.

performed.	

Reviewer #3: comment to Author Jacqueline Milton, Boston University, USA	**Response by author**
1. This study examined subjects with ages ranging from early childhood to adulthood. The discussion describes a variety to studies in which patients younger than 40 years of age had 75% to 90% change of long-term cure with allogeneic bone marrow transplantation and a three to five year survival rate between 75% and 95% in younger patients. Is there reason to believe that age may be confounding the association treatment (transplantation vs. immunosuppressive therapy) and mortality?	Some guidelines recommended to offer transplantation only to patients younger than 40 years of age. Allogeneic transplantation is a physical demanding treatment and it was assumed that the physical fitness in younger patients is better in older patients and that the survival depends upon physical resistance against the physical stress. Guidelines set a cut-off at 40 years to specify a decision point in an decision tree to find the optimal individual treatment. Some have questioned a strict age cut-off and recommended to acknowledge whether a patient has the required physical fitness to survive the treatment including also patients older than 40 years. According to Gupta et al. Impact of age on outcomes after bone marrow transplantation for acquired aplastic anemia using HLA-matched sibling donors. Haematologica. 2010 Dec;95(12):2119-25: "Mortality risks increased with age. Risks were also higher in patients with a poor performance score." Age is certainly a confounder especially concerning overall mortality as the primary outcome. We want to indicate to the following text in the manuscript: *We judged studies as consistent with the principle of 'Mendelian randomization' if all transplant donors were clearly siblings and if the allocation of patients to treatment groups was not based on age. We regarded studies as not consistent with the principle of 'Mendelian randomization' if age was not balanced between groups, indicating that age played a role in the group assignment. Example for imbalance: distribution of age categories was statistically not comparable (P value less than 0.05).*
2. The authors combined those undergoing immunosuppressive therapy as comparator with either 1)antithymocyte, 2)antilyphocyte or 3) a combination of the two. Is there any evidence of a difference in outcome between these 3 methods?	We included IST as comparator with either antithymocyte/¬antilymphocyte globulin or ciclosporin or a combination of the two. That means ATG alone or ATG in combination with ciclosporin or ALG alone or ALG in combination with ciclosporin. We have tried to clarify the composition of immunosuppressive therapy and we added explanations of each component. We added or clarified the following text to the introduction chapter: *According to the 2009 Guidelines for the diagnosis and management of aplastic anaemia of the British Committee for Standards in Haematology [4], first-line immunosuppressive therapy (IST) is a combination of antithymocyte globulin (ATG) and ciclosporin. First-line IST is indicated for patients where no MSD is available, which can be expected for 70% of patients with SAA [3]. Ciclosporin is an immunosuppressant drug that is not lymphocytotoxic but has specific inhibitory effects on T-lymphocyte function [5]. In the past, antilymphocyte globulin (ALG) was reported interchangeably in the literature alongside ATG, therefore, ALG is reported in the present study on equal terms. ATG as well as ALG are polyclonal antibodies that recognize a variety of human lymphocyte cell surface antigens, reduce the number of lymphocytes and induce an immunosuppressive effect. They originate in animals*

	immunized with either normal human thymocytes, collected at pediatric cardiac surgery or thoracic duct lymphocytes, collected during therapeutic cannulation [5]. Concerning the hematologic response and the survival of patients after a first treatment for severe aplastic anemia, it may be crucial in what type of animal ATG originates, as a randomized study showed that rabbit ATG was inferior in this respect to horse ATG [6]. The currently recommended combination of ciclosporin with ATG in the treatment of severe aplastic anemia is based on their separate and potentially complementary modes of action[5]. We added or clarified the following text to the methods chapter: *We included IST as comparator with ciclosporin combined with ATG as the current mode of IST [10]. To accommodate also former modes of IST, we also included ciclosporin combined with ALG, cyclosporine alone, ATG alone, and ALG. Other agents such as corticosteroids and androgens were not considered.* We added or clarified the following text to the section 'Strengths and limitations' of the discussion chapter: *We did not separate horse ATG from rabbit ATG, although the type of animal as the origin of ATG was reported as a serious effect modifier [6].* We would like to point to some statements in seminal articles. Young NS, Shimamura A. Chapter 9: Acquired bone marrow failure syndromes. Pages 273 to 318. Blood: Principles and Practice of Hematology. Edited by Robert I. Handin, Samuel E. Lux, Thomas P. Stossel. Second edition. Philadelphia: Lippincott Williams & Wilkins, 2003. P 273-318. "Neither ATG nor ALG is an immunologically specific reagent. The immunogen for ATG is normal human thymocytes, collected at pediatric cardiac surgery; for ALG the immunogen is thoracic duct lymphocytes, collected during therapeutic cannulation. The immunized animals respond with production of antibodies that recognize a variety of human cell surface antigens." Marsh JCW, Ball SE, Cavenagh J, et al. Guidelines for the diagnosis and management of acquired aplastic anaemia. Br J Haematol 2009;147:43-70: "The standard immunosuppressive regimen is a combination of antithymocyte globulin (ATG) and ciclosporin." Scheinberg P, Nunez O, Weinstein B, Scheinberg P, Biancotto A, Wu CO, et al. Horse versus rabbit antithymocyte globulin in acquired aplastic anemia. New England Journal of Medicine 2011;365(5):430-8: "In a randomized study, rabbit ATG was inferior to horse ATG as a first treatment for severe aplastic anemia, as indicated by hematologic response and survival."

VERSION 2 – REVIEW

REVIEWER	Robert K Hills Cardiff University, Cardiff UK
REVIEW RETURNED	12-Jun-2014

GENERAL COMMENTS	The Mendelian approach of Gray & Wheatley (1991) clearly refers to a donor vs no donor comparison, and indeed if you look at the work of both authors, especially Keith Wheatley and the AML cooperative group, the comparison of transplant is in a donor vs no donor fashion. This shows the concept of compliance is important here - what proportion of patients with a matched sibling donor get a sib allograft. This is missing from the paper, and is absolutely crucial. As Wheatley points out in another paper, low compliance gives an artificially precise estimate of a diluted treatment effect. If only 10% of patients get SCT, for example, then o-e is divided by 10, but V is unaltered meaning that the confidence intervals are tight around no effect. We also need to know how many patients in the no donor group got a MUD SCT; and whether sibling allograft was given upfront or reserved for salvage. Please also call the arms donor vs no donor as they are not allograft vs not, because of the compliance issue addressed above. I am unclear of the ethical reasons surrounding randomisation to SCT vs not. It is done for autograft in other diseases with poor prognosis, and has been tried for allograft as well. The ethical issues peculiar to this condition need to be made explicit here. The tweaks requested here will address the shortcomings in the bullet point questions above.

VERSION 2 – AUTHOR RESPONSE

Reviewer #1: comment to Author **Seiji Kojima MD,PhD; School of Medicine, Nagoya, Japan**	**Response by author**
The authors tried to evaluate the effectiveness and adverse events of first alloSCT and IST in patients with severe aplastic anemia by metaanalysis. However, only 3 trials satisfied the criteria and total number of patients were too small ,only 302 patients. Moreover, all data were collected 15 to 30 years ago. The treatments were out of date. The results may cause misleading for modern practice.	The findings result from the available study data. The time period in which the data were collected and the consequences were addressed in the manuscript. Quoted from the section 'Strengths and limitations': *All data were collected about 15 up to more than 30 years ago. Thus, the results may not be applicable to current modern standard care.*

Reviewer #2: comment to Author **Robert Hills; Cardiff University, Cardiff UK**	**Response by author**
This is a well-performed analysis of the effect of sibling allograft for aplastic anaemia. The analysis is using the "Donor/No Donor" analysis proposed by Gray and Wheatley. This method, and its limitations, could be better described here. It compares patients with a donor with those with no donor. It doesn't compare sibling allograft with no sibling allograft. It is important that the compliance with transplant is given here. How many of the "donor" patients were transplanted - some may not have been for various reasons. Additionally, this method looks only at matched sibling allograft - matched unrelated donor would be in the no donor group, and the assumption here is that most patients with donors get allografts now, rather than after progression, and that there are very few alternative transplants performed, or those that are are of limited value. The limitations in the modern era of this approach in AML (its original setting) are well known and require some rehearsing here. In particular, the arms should be called donor vs no donor not SCT vs not.	First, sibling, related, or unrelated donors constitute a heterogeneous pool. In the past, survival differed considerably among those groups. We intended from the start to reduce heterogeneity by confining to sibling donors. Second, we tried to keep information about 'Mendelian randomization' as short as possible and we refered to two papers that are best suited to explain the principles behind this specific concept. We want to take up the suggestion of the reviewer to broaden the text by providing more explanagory information about this method. Therefore, we added the following section in the method chapter: *Principle of 'Mendelian randomization'* *There are ethical concerns around randomization of patients with severe aplastic anemia to transplantation versus non-transplantation. In general, MSD-HSCT is a life-threatening treatment that can lead to early severe adverse events including death. Gray 1991 and Wheatley 2004 described the potential of 'Mendelian randomization' to minimize bias when comparing MSD-HSCT with an alternative therapy. The base concept has been ascribed to Katan 1986. 'Mendelian randomization' means the view that nature itself has already 'randomized' the paternal and maternal part of a gene given that donor and recipient are siblings. Thus, 'Mendelian randomization' by definition accepts only siblings as transplant donors and these sibling donors are required to have 'identical' or matched features of specific transplant-relevant HLA sites when compared with the transplant recipient. Therefore, patients with an HLA-matched sibling will be allocated to the MSD-HSCT group. On the other hand, patients with siblings that are not HLA compatible will be allocated to the immunosuppressive therapy group. The term 'Mendelian randomization' refers to the fact that the genetic distribution of paternal and maternal alleles follows a random process and is determined before birth. This concept takes advantage of an instrumental variable for allocating the patients to treatment groups and, at the same time, this variable is neither associated with the treatment nor associated with the outcome.* We added the following sentences with respect to 'Mendelian randomization' to the section 'Characteristics of included articles': *The authors of all included studies did not report that 'Mendelian randomization' was planned and the authors did not report the size of the involved families. The authors did not report the numbers of siblings and the results of the individual genetic analyses.* We added the following sentences with respect to

	'Mendelian randomization' to the section 'Strengths and limitations' of the discussion chapter: *Use of 'Mendelian randomization' is no guarantee that bias is minimized.* *This may be because tissue typing data may not be accurate. Patients may have only one sib-ling either in the donor or in the no donor group. Large families have a greater chance of find-ing a donor. Therefore, designing a non-randomized controlled trial by applying 'Mendelian randomization' requires careful thought to effectively reduce bias and control for potential confounders. There is a time lag in patients with siblings because tissue typing and readiness for assignment to treatment group may possibly take several months (Wheatley 2004). On the other hand, patients with no siblings can be assigned immediately and are at earlier risk for adverse events.* *Nitsch 2006 described the limits to causal inference based on 'Mendelian randomization' [29].* We want to address a quote from the reference: Gray R, Wheatley K. How to avoid bias when comparing bone marrow transplantation with chemotherapy. Bone Marrow Transplant 1991;7(Suppl 3):9-12. "The value of allogeneic BMT can, however, be assessed unbiasedly using 'Mendelian randomisation', i.e. comparing patients whose siblings are HLA-compatible with those who are not." We also want to address a quote from reference: Wheatley K, Gray R. Commentary: Mendelian randomization: an update on its use to evaluate allogeneic stem cell transplantation in leukaemia. Int J Epidemiol 2004;33:15-17. "Thus, we need an alternative, unbiased method for evaluating SCT. The answer is provided by Mendelian, or genetic, randomization. [...] In the haematological context, for a patient's sibling to be a suitable donor, he/she must have inherited the same tissue type as the patient from their mother and father. Since the chances of there being a match depend on the random assortment of genes at fertilization, only one in four siblings will be expected to have the same tissue type as the patient. Thus, whether or not a patient has a matched sibling donor available is essentially a random process and the presence or absence of a donor can be used as a surrogate for randomization."
Additionally, given that the outcome is OS, I'm not sure how high the risk of bias is from not blinding - death is a pretty objective endpoint.	We agree with the reviewer. We added the following sentences with respect to risk of bias to the section 'Characteristics of included articles': *We judged a low risk of bias for blinding the assessment of overall mortality. Blinding or lack of blinding is not expected to make a difference concerning overall mortality.*
Please therefore analyse as donor vs no donor (to remove selection and zero timeshift biases) and give compliance levels - what transplants were	We would like to refer to our response expressed in the first row.

performed.	

Reviewer #3: comment to Author **Jacqueline Milton, Boston University, USA**	**Response by author**
1. This study examined subjects with ages ranging from early childhood to adulthood. The discussion describes a variety to studies in which patients younger than 40 years of age had 75% to 90% change of long-term cure with allogeneic bone marrow transplantation and a three to five year survival rate between 75% and 95% in younger patients. Is there reason to believe that age may be confounding the association treatment (transplantation vs. immunosuppressive therapy) and mortality?	Some guidelines recommended to offer transplantation only to patients younger than 40 years of age. Allogeneic transplantation is a physical demanding treatment and it was assumed that the physical fitness in younger patients is better in older patients and that the survival depends upon physical resistance against the physical stress. Guidelines set a cut-off at 40 years to specify a decision point in an decision tree to find the optimal individual treatment. Some have questioned a strict age cut-off and recommended to acknowledge whether a patient has the required physical fitness to survive the treatment including also patients older than 40 years. According to Gupta et al. Impact of age on outcomes after bone marrow transplantation for acquired aplastic anemia using HLA-matched sibling donors. Haematologica. 2010 Dec;95(12):2119-25: "Mortality risks increased with age. Risks were also higher in patients with a poor performance score." Age is certainly a confounder especially concerning overall mortality as the primary outcome. We want to indicate to the following text in the manuscript: *We judged studies as consistent with the principle of 'Mendelian randomization' if all transplant donors were clearly siblings and if the allocation of patients to treatment groups was not based on age. We regarded studies as not consistent with the principle of 'Mendelian randomization' if age was not balanced between groups, indicating that age played a role in the group assignment. Example for imbalance: distribution of age categories was statistically not comparable (P value less than 0.05).*
2. The authors combined those undergoing immunosuppressive therapy as comparator with either 1)antithymocyte, 2)antilyphocyte or 3) a combination of the two. Is there any evidence of a difference in outcome between these 3 methods?	We included IST as comparator with either antithymocyte/¬antilymphocyte globulin or ciclosporin or a combination of the two. That means ATG alone or ATG in combination with ciclosporin or ALG alone or ALG in combination with ciclosporin. We have tried to clarify the composition of immunosuppressive therapy and we added explanations of each component. We added or clarified the following text to the introduction chapter: *According to the 2009 Guidelines for the diagnosis and management of aplastic anaemia of the British Committee for Standards in Haematology [4], first-line immunosuppressive therapy (IST) is a combination of antithymocyte globulin (ATG) and ciclosporin. First-line IST is indicated for patients where no MSD is available, which can be expected for 70% of patients with SAA [3]. Ciclosporin is an immunosuppressant drug that is not lymphocytotoxic but has specific inhibitory effects on T-lymphocyte function [5]. In the past, antilymphocyte globulin (ALG) was reported interchangeably in the literature alongside ATG, therefore, ALG is reported in the present study on equal terms. ATG as well as ALG are polyclonal antibodies that recognize a variety of human lymphocyte cell surface antigens, reduce the number of lymphocytes and induce an immunosuppressive effect. They originate in animals*

	immunized with either normal human thymocytes, collected at pediatric cardiac surgery or thoracic duct lymphocytes, collected during therapeutic cannulation [5]. Concerning the hematologic response and the survival of patients after a first treatment for severe aplastic anemia, it may be crucial in what type of animal ATG originates, as a randomized study showed that rabbit ATG was inferior in this respect to horse ATG [6]. The currently recommended combination of ciclosporin with ATG in the treatment of severe aplastic anemia is based on their separate and potentially complementary modes of action[5]. We added or clarified the following text to the methods chapter: *We included IST as comparator with ciclosporin combined with ATG as the current mode of IST [10]. To accommodate also former modes of IST, we also included ciclosporin combined with ALG, cyclosporine alone, ATG alone, and ALG. Other agents such as corticosteroids and androgens were not considered.* We added or clarified the following text to the section 'Strengths and limitations' of the discussion chapter: *We did not separate horse ATG from rabbit ATG, although the type of animal as the origin of ATG was reported as a serious effect modifier [6].* We would like to point to some statements in seminal articles. Young NS, Shimamura A. Chapter 9: Acquired bone marrow failure syndromes. Pages 273 to 318. Blood: Principles and Practice of Hematology. Edited by Robert I. Handin, Samuel E. Lux, Thomas P. Stossel. Second edition. Philadelphia: Lippincott Williams & Wilkins, 2003. P 273-318. "Neither ATG nor ALG is an immunologically specific reagent. The immunogen for ATG is normal human thymocytes, collected at pediatric cardiac surgery; for ALG the immunogen is thoracic duct lymphocytes, collected during therapeutic cannulation. The immunized animals respond with production of antibodies that recognize a variety of human cell surface antigens." Marsh JCW, Ball SE, Cavenagh J, et al. Guidelines for the diagnosis and management of acquired aplastic anaemia. Br J Haematol 2009;147:43-70: "The standard immunosuppressive regimen is a combination of antithymocyte globulin (ATG) and ciclosporin." Scheinberg P, Nunez O, Weinstein B, Scheinberg P, Biancotto A, Wu CO, et al. Horse versus rabbit antithymocyte globulin in acquired aplastic anemia. New England Journal of Medicine 2011;365(5):430-8: "In a randomized study, rabbit ATG was inferior to horse ATG as a first treatment for severe aplastic anemia, as indicated by hematologic response and survival."

VERSION 3 - REVIEW

REVIEWER	Robert Hills Cardiff University, Cardiff UK
REVIEW RETURNED	17-Jun-2014

GENERAL COMMENTS	Thank you for addressing most of my comments here. However, the manuscript still states: There are ethical concerns around randomization of patients with severe aplastic anemia to transplantation versus non-transplantation. But the authors do not tell us what they are - we know that people are generally unwilling to randomise (although this is not stated in the M/S) - but if there is uncertainty, as there clearly is to do this, why is it unethical to randomise? The authors need to justify this statement and not just rely on proof by vigorous assertion.

VERSION 3 – AUTHOR RESPONSE

Reviewer #2: comment to Author **Robert Hills; Cardiff University, Cardiff UK**	**Response by author**
Thank you for addressing most of my comments here. However, the manuscript still states: There are ethical concerns around randomization of patients with severe aplastic anemia to transplantation versus non-transplantation. But the authors do not tell us what they are - we know that people are generally unwilling to randomise (although this is not stated in the M/S) - but if there is uncertainty, as there clearly is to do this, why is it unethical to randomise? The authors need to justify this statement and not just rely on proof by vigorous assertion.	Thank you for accepting most of our changes. I know leading experts in stem cell transplantation who have opted against randomisation because of ethical concerns. We do not know all reasons why persons have these concerns and we did not evaluated it. We just wanted to address that these concerns are expressed by some. We did not state that randomisation should be prohibited or is not possible. We did not state that randomistion is unethical. Both treatments may not work and the patient consequently might die. It is called graft failure with transplantation and no response or refractory with immunosuppressive therapy. Both treatments may work and cellular function may be compatible with life. Then, other problems may arise that are different between the treatment groups. Immunosuppressive therapy with a good response and transplantation may result in secondary neoplasia after long follow-up. However, transplantation may also result in graft-versus-host disease within a short time period or after a long follow-up. This disease may cause early death. This disease is not present in the control group. The difference in expecting an early loss of life between the two treatment groups and the reservation not to expose patients to this possible early life threat by random allocation might be a reason. I tried to explain shortly in the following text: *In general, MSD-HSCT is a life-threatening treatment that can lead to early severe adverse events including death.*

	I suggest to remove this sentence and extend the first sentence of the section as follows: *There are ethical concerns around randomization of patients with severe aplastic anemia to transplantation versus non-transplantation* *because the risk of early death is expected to be higher in the transplantation group than in the non-transplantation group. The reason is the potentially life-threatening graft-versus-host disease occuring only in the transplanted patients.* ~~*In general, MSD-HSCT is a life-threatening treatment that can lead to early severe adverse events including death.*~~