TECHNICAL REPORT FOR THE DRUG SAFETY AND EFFECTIVENESS NETWORK:
INTERVENTIONS FOR THE PROPHYLAXIS AND TREATMENT OF GRAFT-VERSUS-HOST DISEASE IN PATIENTS UNDERGOING ALLOGENEIC HEMATOPOIETIC STEM CELL TRANSPLANT – A SYSTEMATIC REVIEW WITH NETWORK META-ANALYSES

Authors:
Dianna Wolfe, Brian Hutton, David Moher, Fatemeh Yazdi, Chris Bredeson, Natasha Kekre, David Allan

November 2016

Research Conducted by:
The Knowledge Synthesis Group:
Clinical Epidemiology Program,
Ottawa Hospital Research Institute
501 Smyth Road
The Ottawa Hospital, General Campus
Centre for Practice Changing Research
Ottawa, Ontario
Box 201b, K1H 8L6

EXECUTIVE SUMMARY

OVERVIEW

Graft-versus-host disease (GVHD) is a potentially life-threatening complication that frequently occurs following allogeneic hematopoietic cell transplantation (HSCT). Matching of donor and recipient for major histocompatibility antigens reduces the risk of development of GVHD; however, 35–40% of fully matched recipients will still develop acute GVHD (aGVHD) due to unmatched minor histocompatibility antigens. To further reduce the risk of GVHD, various GVHD prophylactic and treatment strategies have been developed. The agents used in these strategies generally focus on reducing immune responses in the recipient post-transplant, specifically by targeted destruction, reduced production, or inhibition of T cells. Reducing graft immune response post-transplantation reduces the risk of GVHD; however, it also reduces the graft's ability to destroy residual host tumour cells and to re-establish a host immune system, potentially leading to higher risks of relapse, infection, and mortality. The ideal strategy to prevent or treat GVHD would reduce the incidence or health impact of GVHD, with no increase in relapse or mortality. Balancing benefits with harms remains a challenge.

Objectives Addressed in this Review

- *To compare the benefits (i.e., prevention of GVHD) and harms (e.g., risk of relapse, infection, and mortality) of competing regimens for prophylaxis of GVHD in patients undergoing HSCT, and to establish a hierarchy of intervention strategies according to their efficacy and safety.*
- *To compare the benefits (i.e., resolution of GVHD) and harms (e.g., risk of relapse, infection, and mortality) of competing regimens for treatment of GVHD in patients undergoing HSCT, and to establish a hierarchy of intervention strategies according to their efficacy and safety.*

OVERVIEW OF RESEARCH METHODS

Using data from randomized controlled trials, networks of evidence of GVHD prophylaxis regimens and GVHD treatment regimens were developed. For regimens that have never been directly compared in head-to-head trials, their effects can be compared using ***network meta-analysis*** to derive comparisons between therapies and to rank all regimens according to their relative effects on the outcomes of interest (e.g., the incidence of GVHD, relapse, and mortality). This approach allows us to understand the strength of the evidence supporting various interventions for the prevention and treatment of GVHD. This is helpful to generate recommendations for the optimal management of patients with GVHD, and to consider design of future RCTs that leverage the existing foundation of evidence to select the most appropriate control arm.

SYSTEMATIC REVIEW METHODS

The databases Medline, PubMed, Embase, and the Cochrane Register of Controlled Trials were searched for randomized trials of patients undergoing HSCT. Studies were included if patients underwent allogeneic HSCT in the treatment of hematologic neoplasias or benign disease and were randomly allocated to receive a pharmacological intervention for the prophylaxis or treatment of acute or chronic GVHD. Outcomes of interest included overall mortality, relapse of underlying disease, incidence of acute and chronic GVHD (prophylaxis review), resolution of acute and chronic GVHD (treatment review), weaning from GVHD interventions (treatment review), and specific harms. We conducted separate analyses for the prophylaxis and treatment of GVHD, using Bayesian network meta-analysis to compare interventions for outcomes of interest, where feasible. All outcomes were analysed as binary endpoints, with summary comparisons between regimens reported as odds ratios with 95% credible intervals. For outcomes for which network meta-analysis were not possible, detailed narrative summaries were prepared.

RESULTS

Thirty-two trials assessed 19 unique GVHD prophylactic regimens in 3,875 total patients. Overall 7 trials assessed 10 unique treatment strategies for aGVHD (4 studies) and chronic GVHD (cGVHD) (3 studies) in 830 total patients; represented regimens are listed in **Table 1**. Overall, there was substantial variability in patient populations with respect to age, underlying hematologic disease, disease risk of relapse/mortality, and transplant donor status (i.e., related vs. unrelated, matched vs. unmatched). Trial publication dates ranged from 1979–2015.

Table 1: Interventions compared in networks of GVHD prophylaxis and treatment

GVHD prophylaxis	GVHD treatment
• MTX+TAC • MTX • CSA • CSA+MTX • CSA+Steroids • MTX+Steroids • CSA+MTX+Steroids • MTX+TAC+Steroids • MTX+TAC+SIR • MTX+ATG+Steroids • CSA+MTX+ATG+Steroids • CSA+MTX+UDCA+Steroids • CSA+MTX+MSCs • CsA+HCQ • SIR+TAC • MMF+TAC+RTX • MMF+TAC • THAL+standard prophylaxis • Placebo+standard prophylaxis	• Steroids • CsA • CsA+Steroids • ATG+Steroids • AZA+Steroids • MMF+Steroids • ETN+Steroids • THAL+CsA+Steroid • MMF+(CsA or TAC or SIR)+Steroids • PLB+(CsA or TAC or SIR)+Steroids

Abbreviations. ATG = anti-thymocyte globulin; CsA = cyclosporin A; ETN = etanercept; HCQ = hydroxychloroquine; MMF = mycophenolate mofetil; MTX = methotrexate; RITX = rituximab; SIR = sirolimus; TAC = tacrolimus; THAL = Thalidomide; UDCA = ursodeoxycholic acid.

Objective 1: Comparing regimens for GVHD prophylaxis

Most comparisons between treatment regimens were informed only by indirect evidence (i.e. head-to-head trials were not available), and many of the direct comparisons were informed by single studies with small numbers of patients. Thus, a sparse evidence base and considerable between-study heterogeneity in patient populations complicates targeting of findings from meta-analyses. The ideal outcome to assess cGVHD (i.e., incidence of extensive cGVHD in patients alive 100 days post-transplant) was unavailable for many trials. In its place, data for overall cGVHD in randomized patients was used. Regarding disease relapse, only studies following patients for a median duration of between 2–3 years were included for analysis. Clinical interpretation of findings from network meta-analyses were as follows:

- **Prevention of aGVHD and cGVHD**: Compared to the reference treatment MTX+TAC, data suggested that MTX+SIR+TAC and SIR+TAC were superior for prevention of aGVHD, while CsA+MTX+ATG+Steroids was superior for prevention of cGVHD. Five regimens were statistically significantly poorer than MTX+TAC at preventing aGVHD; these were CsA+MTX, CsA+Steroids, CsA, MTX, and MTX+Steroids. Inspection of results used to inform relative rankings of all regimens for the prevention of aGVHD showed that although MTX+SIR+TAC was ranked highest, there was little difference in efficacy in the top 4 ranked interventions

(MTX+SIR+TAC, SIR+TAC, CsA+MTX+MSCs, and MTX+TAC+Steroids), given their relatively similar SUCRA values (range: 0.81–0.92). For the prevention of cGVHD, CsA+MTX+ATG+Steroids was ranked highest, demonstrating significantly improved efficacy compared to all other interventions except the second-ranked intervention, HCQ+CsA.

- **Disease relapse and overall mortality**: The best regimens at preventing acute and chronic GVHD (MTX+SIR+TAC, SIR+TAC, CsA+MTX+ATG+Steroids) and the standard of care (MTX+TAC) were the 4 regimens ranked lowest for prevention of relapse of underlying disease within 2–3 years of transplant compared to 7 other regimens that had available data for use in network meta-analysis. For these same regimens in the analysis of mortality within 1 year of transplant, SIR+TAC and MTX+SIR+TAC were ranked 12th and 13th out of 15 regimens with available data for network meta-analysis, while CsA+MTX+ATG+Steroids and MTX+TAC were ranked 4th and 9th. While the addition of ursodeoxycholic acid to CsA+MTX+Steroids significantly reduced the risk of death within 1 year of transplant, the addition of mesenchymal stem cells (MSCs) to CsA+MTX significantly increased the risk of death compared to all but one other regimen. 100-day mortality was not significantly influenced by any of the regimens evaluated.

- **Additional harms-related findings**: Other adverse events including specific organ toxicity and infections were described in a small number of studies, precluding performance of meta-analyses. Based on available data, the addition of steroids or ATG appeared to increase the risk of CMV reactivation, while the use of SIR appeared to increase the risk of VOD.

- ***Key clinical messages regarding interventions for GVHD prophylaxis were as follows***:

    - Regimens of single agents were less efficacious to prevent GVHD than regimens involving multiple agents; however, single-agent regimens generally were more efficacious at preventing relapse of underlying disease within 2–3 years of transplant.

    - Regimens containing a calcineurin inhibitor (i.e., TAC or CsA) had greater efficacy to prevent aGVHD than regimens without; however, there was no significant difference between TAC- or CsA-containing regimens.

Objective 2: Comparing regimens for GVHD treatment

Network meta-analyses could not be performed to compare GVHD treatment regimens. This was a consequence of two limitations of the evidence base: (1) few studies for treatment were located (n = 7), and these were split between treatment of cGVHD and aGVHD and did not produce robust networks of evidence for analysis; and (2) studies were found to be associated with several important sources of between-study heterogeneity including endpoint definitions and duration of follow-up. As formal meta-analysis was judged inappropriate, within-study comparisons were assessed and described narratively. Primary findings from included studies were as follows:

- **Response to treatment**: MMF+Steroids was superior to Etanercept+Steroids in the treatment of aGVHD after 28 and 56 days of follow-up, and Steroids alone were superior to Alemtuzumab+Steroids in the treatment of cGVHD after 9 months. No other comparisons were associated with statistically significant differences between regimens. No treatment for aGVHD or cGVHD was clearly superior in terms of patients' weaning from steroids.

- **Overall and non-relapse mortality**: In the treatment of aGVHD, when MMF was added to Steroids, increased overall mortality (but not non-relapse mortality) occurred after 1 year of follow-up. In the treatment of cGVHD, when MMF was added to (CsA or TAC or SIR)+Steroids, increased

overall mortality (but not non-relapse mortality) occurred after 4 years. When Alemtuzumab was added to Steroids, greater non-relapse mortality (but not overall mortality) occurred at 4 years.

RECOMMENDATIONS AND FUTURE STUDIES

Numerous drugs are used in a broad variety of single- and multi-agent regimens for both the prophylaxis and treatment of GVHD in HSCT. This review found that comparative evidence from randomized trials is lacking for many comparisons of these regimens, which has led to considerable practice variation between institutions. This systematic review of the evidence, incorporating network meta-analyses where possible, was conducted to address these gaps. Totals of 32 and 7 randomized trials of GVHD prophylactic and treatment regimens, respectively, were identified and studied to inform analyses in this review.

The following key points for clinical practice were identified in this review:

- Prophylactic regimens for GVHD that demonstrated increased benefits (i.e., a reduced risk of acute or chronic GVHD) were associated with increased harms (i.e., increased risk of relapse, infection, or mortality). These regimens included the following combinations: MTX+TAC, MTX+SIR+TAC, SIR+TAC, and CsA+MTX+ATG+Steroids.
- MTX+TAC should be considered the gold standard for GVHD prophylaxis to which other regimens are compared.
- While the addition of ATG in prophylaxis appears to reduce cGVHD, it has not been used in combination with TAC. Use of ATG does not appear to increase overall mortality or relapse but can increase CMV reactivation and possibly other infections.
- In the treatment of aGVHD, response rates with the use of MMF+Steroids may be improved; however, overall survival may be compromised. This regimen warrants further study.
- No therapy appears superior to Steroids alone for the treatment of cGVHD.

Future studies of GVHD prophylaxis should be stratified for competing factors, such as patient age and underlying disease, and donor factors. Prophylactic regimens should be compared with standard therapy, such as MTX+TAC or MTX+CSA+ATG. In the evaluation of cGVHD occurrence, studies should report extensive cGVHD (or use existing grading scales) in patients who are alive at 100 days, with a follow-up of 2 years post-transplant. Treatment interventions for GVHD should be evaluated for standardized outcome responses at consistent follow-up times, and be compared with Steroids alone or MMF+Steroids.

Supplemental Information: Definitions of Relevance for Readers of this Summary

What is a systematic review?

- A systematic review is a type of literature review performed using a structured approach to identify, appraise and analyze research studies that are relevant for answering a question posed by a research team. Reviews of randomized controlled trials are commonly used for the practices of technology assessment and evidence based medicine. Findings from identified research studies can be synthesized in a variety of ways, one of which includes the practice of *meta-analysis*.

What is meta-analysis?

- Meta-analysis is a statistical technique that can be used to quantitatively combine the findings from research studies. For example, the odds ratios from a series of trials comparing rates of cGVHD incidence with two different prophylactic regimens may be combined using meta-analysis to estimate a summary odds ratio based on all available evidence. Meta-analysis uses a weighted average approach to combine study data, and produces a precise estimate of the treatment effect of interest. Studies included in a meta-analysis are traditionally identified by a systematic review, and are used to compare two treatments of interest.

What is a network meta-analysis?

- A network meta-analysis can be thought of as a generalization of traditional meta-analysis where there is an interest in comparing more than two treatments. Studies included in a network meta-analysis are also traditionally identified by a systematic review process.

TABLE OF CONTENTS

ACRONYMS AND ABBREVIATIONS

AA = aplastic anemia
aGVHD = acute graft-versus-host disease
ALL = acute lymphoblastic leukemia
AML = acute myeloid leukemia
ATG = antithymocyte globulin
AZA = alemtuzumab
BM = bone marrow
cGVHD = chronic graft-versus-host disease
CI = confidence interval
CML = chronic myelogenous leukemia
CMV = cytomegalovirus
CrI = credible interval
CsA = cyclosporine A
DIC = deviance information criteria
DNA = deoxyribonucleic acid
ECP = extracorporeal phototherapy
ETN = etanercept
FE = fixed effects
GVHD = graft-versus-host disease
HCQ = hydroxychloroquine
HLA = human leukocyte antigen
HSCT = hematopoietic stem cell transplant
ITT = intention to treat
K-M = Kaplan-Meier
MDS = myelodysplastic syndromes
MM = multiple myeloma
MMF = mycophenolate mofetil
MSCs = mesenchymal stem cells
mTOR = mammalian target of rapamycin
MTX = methotrexate
NHL = non-Hodgkins Lymphoma
NMA = network meta-analysis
NR = not reported
OR = odds ratio
PB = peripheral blood
PCR = polymerase chain reaction
PLB = placebo
PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT = randomized controlled trial
RE = random effects
RITX = rituximab
SIR = sirolimus
SUCRA = surface under the cumulative ranking curve
TAC = tacrolimus
THAL = thalidomide
TNF = tissue necrosis factor
UC = umbilical cord
UDCA = ursodeoxycholic acid
VOD = veno-occlusive disease

1. RESEARCH QUESTION: WHAT ARE THE RELATIVE EFFECTIVENESS AND HARMS OF INTERVENTIONS FOR THE PROPHYLAXIS AND TREATMENT OF GRAFT-VERSUS-HOST DISEASE?

1.1. Introduction

Allogeneic hematopoietic cell transplantation (HSCT) is a multi-faceted, complex health care intervention that can be life-saving for patients with hematologic cancers, other serious blood or immune disorders, and inherited metabolic conditions.[1,2] Optimization of strategies to improve patient outcomes has evolved gradually as transplant activity has increased globally over the past 25 years. With increasing research activity, it is now possible to address questions regarding drug effectiveness and safety in HSCT using evidence available from randomized controlled trials (RCTs). To date, clear recommendations regarding treatment selection within an array of transplant-related interventions remains lacking. In particular, it remains challenging to quantify the strength of evidence that underpins many treatment and prevention strategies used currently by transplant centers. This is due to a lack of direct or indirect comparisons for particular interventions, and has led to wide variation in institutional protocols for many aspects of care in HSCT.[3-6] Rigorous systematic reviews addressing the existing body of evidence are currently needed. Many of the drug interventions in HSCT are costly and are associated with significant potential toxicities, such as prolonged immune suppression, severe mucositis, and risk of renal and liver injury, making a formal analysis of safety and effectiveness highly relevant for the improvement of care and the sustainability of transplantation programs.

Graft-versus-host disease (GVHD) is a potentially life-threatening complication that frequently occurs following allogeneic HSCT. GVHD occurs when T cells (a type of immune cell) in the transplant graft identify recipient cells as foreign and subsequently attack.[7] Classically, patients with GVHD present with maculopapular rash, nausea, vomiting, anorexia, profuse diarrhea, ileus, or cholestatic hepatitis. Biopsy may help to confirm GVHD in clinically unclear cases.

Three requirements are necessary for the development of GVHD following HSCT: (1) the donor transplant must contain immuno-competent T cells, (2) the recipient must not be able to mount an immune response to eliminate the donor T cells, and (3) the transplant recipient must express tissue antigens that are foreign to the transplant donor. The first two requirements are met with any allogeneic HSCT. Although techniques are available to deplete the number of T cells present in donor grafts, invariably some T cells remain in the transplant. In the HSCT recipient, pre-transplant conditioning regimens and post-transplant management cause extreme immunosuppression, meaning the recipient is incapable of mounting an immune response to the remaining donor T cells. This is beneficial for graft survival, but detrimental when donor T cells attack recipient cells.

The final requirement, expression of tissue antigens by the recipient that are foreign to the transplant donor, can be partially controlled through selection of a donor with antigens matching those of the recipient. The most immuno-stimulating of the antigens found on recipient cells are human leukocyte antigens (HLAs). These proteins, also known as the major histocompatibility complex, have high genetic variability, making exact matches between donor and recipient difficult. Despite being fully matched with their donors for the 8 major HLAs, 35–40% of HSCT recipients will develop acute GVHD due to unmatched minor histocompatibility antigens.

1.1.1. Acute versus Chronic GVHD

Myeloablative conditioning regimens are associated with a high risk of GVHD in the first 100 days post-transplant due to their extreme immunosuppressive effects coupled with a high degree of tissue damage and subsequent widespread inflammation. Non-myeloablative techniques reduce the risk of GVHD in the first 100 days; however, the myeloablative consequences of the graft-versus-tumour effect (i.e., overwhelming

of the original cancer cells in the bone marrow by the donor graft) may simply delay the signs of GVHD to >100 days post-transplant. Originally, acute GVHD (aGVHD) was defined having an onset before day 100 post-transplant, with chronic GVHD (cGVHD) arising beyond 100 days. However, this definition may no longer be accurate, as non-myeloablative conditioning regimens gain popularity and late-onset aGVHD and overlap syndrome have become more commonly recognized. Current thought is that clinical manifestations should be used to differentiate aGVHD from cGVHD rather than time of onset.

Diagnostic signs have been identified in the skin, appendages, mouth, eyes, female genitalia, esophagus, lungs, and connective tissues to differentiate cGVHD from aGVHD. The severity of aGVHD can be graded on a scale from I to IV, with higher grading suggestive of greater severity.[8] Patients with Grades III or IV acute GVHD have reduced 5-year survival (25% and 5%, respectively). Chronic GVHD can be manifested in a single organ or may be widespread, leading to debilitating consequences (e.g., joint contractures, loss of sight, end-stage lung disease, or mortality resulting from profound immunosuppression). A 3-level scoring system of cGVHD severity has been proposed (mild, moderate, severe), reflecting the clinical impact of cGVHD on patient function. Incidence rates of cGVHD can reach as high as 80% after HSCT.

1.1.2. Interventions available for the prophylaxis and treatment of GVHD

A variety of approaches to management of GVHD are available, including pharmacologics that cause targeted destruction of T cells (e.g., antithymocyte globulin, cyclophosphamide, alemtuzumab), reduced production of T cells (e.g., everolimus, sirolimus, tyrosine kinase inhibitors, methotrexate, mycophenolate mofetil, azathioprine), or inhibition or immunomodulation of T cells (e.g., cyclosporine, tacrolimus, corticosteroids, mesenchymal stromal cells, etanercept, thalidomide, lenalidomide).

- ***Antithymocyte globulin (ATG)***: Antithymocyte globulin dramatically reduces the number of circulating T-lymphocytes through cell lysis. This reduces host immune response in the immediate post-transplant phase, but also increases the risk of infection.[9]

- ***Cyclophosphamide***: Cyclophosphamide is an alkylating agent that causes intra-strand as well as inter-strand DNA crosslinks, preventing DNA replication and causing cell death. It selectively targets certain T cells.[10] Because of high associated toxicity and side-effects, its use is generally of short duration.

- ***Alemtuzumab***: Alemtuzumab is a monoclonal antibody that binds to a protein present on the surface of mature lymphocytes, which are then targeted for destruction.

- ***Mammalian target of rapamycin (mTOR) inhibitors (everolimus, sirolimus)***: mTOR inhibitors prevent the growth and division of cells in the body by blocking mTOR proteins. Specifically, they inhibit T-cell proliferation and activation, when induced by cytokines, including interleukins and colony- stimulating factors.[11]

- ***Tyrosine kinase inhibitors***: Tyrosine kinase inhibitors prevent the growth and division of cells in the body by blocking enzymes responsible for sending these signals within the body.[12]

- ***Methotrexate***: Methotrexate is a folate antimetabolite that inhibits DNA synthesis, repair, and replication in actively dividing cells. Methotrexate can reduce antigen-stimulated T-cell proliferation.[13] Common side-effects include neutropenia and mucositis.

- ***Mycophenolate mofetil***: Mycophenolate mofetil selectively inhibits the proliferation of T and B lymphocytes through the inhibition of inosine monophosphate dehydrogenase.[14] Mycophenolate

mofetil has fewer toxic effects than methotrexate and may encourage more rapid neutrophil engraftment.

- ***Azathioprine***: Azathioprine is a purine analogue that inhibits DNA synthesis of rapidly proliferating cells, such as T cells and B cells.[15]

- ***Calcineurin inhibitors (cyclosporine, tacrolimus)***: These drugs inhibit calcineurin, a protein involved in the activation of T cells in the immune system.[16] Both cyclosporine and tacrolimus have similar mechanisms of action, clinical effectiveness, and toxic effects. Calcineurin inhibitors are commonly combined with other immunosuppressive drugs (e.g., methotrexate) for prophylaxis.

- ***Corticosteroids***: Corticosteroids suppress a broad range of immune responses mediated by T cells and B cells.[17] T cells may be rapidly depleted from the blood stream due to increased emigration, impairment of growth factor, reduced release from lymphoid tissues, and programmed cell death.

- ***Mesenchymal stromal cells***: Several groups have demonstrated immune-modulatory properties of MSCs based on cell-cell contact mechanisms and the secretion of immune bioactive factors.[18]

- ***Etanercept***: Etanercept is a monoclonal antibody that blocks tumour necrosis factor (TNF), thereby modulating the immune response.

- ***Thalidomide and analogues (thalidomide, lenalidomide)***: Thalidomide analogues have multiple mechanisms of action, including immunomodulation through selective inhibition of pro-inflammatory cytokines such as TNF.

- ***Hydroxychloroquine (Plaquenil)*** and ***extracorporeal phototherapy*** have also been studied for use in management of GVHD and will be considered as additional comparators of interest.[19,20]

In clinical practice, while monotherapy remains an option, it is increasingly common for patients to be managed with regimens involving two or more of these interventions. Balancing effectiveness (control of GVHD) with overall safety (no increase in infections or disease relapse) remains a challenge.

While the combination of short-course methotrexate combined with calcineurin inhibitor therapy has been the mainstay of GVHD prophylaxis in most centers for many years, newer approaches that are less toxic have been investigated, including the use of sirolimus and post-transplant cyclophosphamide to reduce allo-reactive T cells.

1.1.3. Existing literature of relevance

A recent review by Flowers[21] outlined common approaches to treating chronic GVHD using a risk-adapted approach, and concluded that some strategies may work for particular organ involvement but may not work for other organs.

A 2014 systematic review compared multiple interventions for prophylaxis of GVHD in patients undergoing allogeneic HSCT;[22] however, this review had several limitations, including double-counting of several related trials, omission of other relevant trials, and failure to consider several patient-important outcomes including survival, infections, and relapse. Additionally, the review did not discuss clinical considerations (including population and study design variations) which may be relevant for validity and aspects of generalizability, and treatment of GVHD was not at all addressed.

1.1.4. Why this review is important

Numerous agents are used in a broad variety of regimens for both the prophylaxis and treatment of GVHD in HSCT. Generally, comparative evidence is lacking for these interventions, especially head-to-head data, which has led to considerable practice variation between institutions. Inter-institutional variation in clinical practice is a barrier to improving the success of prophylaxis and treatment strategies and improving patient safety. A robust systematic review of the evidence, incorporating network meta-analyses to compare regimens and their impact on key outcomes, will help to explore the relative benefits and harms of competing interventions in this area, ultimately improving patient outcomes.

1.1.5. Objectives of this systematic review

1) To compare the benefits (i.e., prevention of GVHD) and harms (risk of infection and relapse of disease) of competing regimens for prophylaxis of GVHD in patients undergoing HSCT, and to establish a hierarchy of intervention strategies according to their efficacy and safety.

2) To compare the benefits (i.e., resolution of GVHD, and restoration of organ function) and harms (risk of infection and relapse of disease) of competing regimens for treatment of GVHD in patients undergoing HSCT, and to establish a hierarchy of intervention strategies according to their efficacy and safety.

2. REVIEW METHODS

The checklist from the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) Extension Statement for Network Meta-analyses (NMA) was used to guide preparation of this report.[23] A completed checklist is provided in **Appendix 9.**

2.1. Protocol and Registration

The protocol for the review was developed using an integrated framework amongst researchers and knowledge users from the Ottawa Hospital Research Institute (OHRI), the Canadian Bone Marrow Transplant Group, and the pan-Canadian Oncology Drug Review (pCODR). The protocol for the review was subsequently made openly available online in the University of Ottawa Library's online repository, and can be found at the following link: https://www.ruor.uottawa.ca (search for term 'DSEN stem cell transplantation'). The protocol was registered in PROSPERO (CRD42015026738).

2.2. Study Eligibility Criteria

Tables 1 and **2** provide detailed summaries of the eligibility criteria set out for the reviews of the prophylaxis and treatment of GVHD, respectively, using the Population-Interventions-Comparators-Outcomes-Study design (PICOS) framework.

Table 1: Overview of study eligibility criteria for the prophylaxis of GVHD	
PICOS Element	**Description**
Population	Patients undergoing allogeneic HSCT in the treatment of hematologic neoplasias and benign disease, including but not limited to Acute Lymphoblastic Leukemia (ALL), Non-Hodgkins Lymphoma (NHL), Chronic Myelogenous Leukemia (CML), Myelodysplastic Syndromes (MDS), Acute Myeloid Leukemia (AML), Multiple Myeloma (MM), Aplastic Anemia (AA)
Interventions and Comparators	Studies of monotherapy and combination therapy involving several agents, including the following: methotrexate, calcineurin inhibitors (tacrolimus, cyclosporin), corticosteroids (prednisolone, methylprednisolone), antithymocyte antiglobulin (ATG), and mTOR inhibitors (everolimus, sirolimus), mycophenolate mofetil and cyclophosphamide. Inclusion of each regimen was determined in discussion with our clinical experts.
Outcomes	• All-cause mortality • Transplantation related mortality • Relapse of underlying disease • Incidence of acute GVHD (within 100 days post-transplant of grades II-IV) • Incidence of chronic GVHD (>100 days post-transplant of grades II-IV) • Specific harms associated with different regimens
Study Design	Randomized controlled trials of at least 180 days of follow-up

Table 2: Overview of study eligibility criteria for the treatment of GVHD	
PICOS Element	**Description**
Population	Patients diagnosed with acute or chronic GVHD following HSCT in the treatment of hematologic neoplasias and benign disease, including but not limited to Acute Lymphoblastic Leukemia (ALL), Non-Hodgkins Lymphoma (NHL), Chronic Myelogenous Leukemia (CML), Myelodysplastic Syndromes (MDS), Acute Myeloid Leukemia (AML), Multiple Myeloma (MM), Aplastic Anemia (AA)
Interventions and Comparators	Studies of monotherapy and combination therapy involving several agents, including the following: steroids, cyclosporine, tacrolimus, mycophenelate mofetil, etanercept, ATG, azathioprine, lenalidomide, thalidomide, extracorporeal phototherapy (ECP), mesenchymal stem cells, and hydroxychloroquine. Inclusion of each regimen was determined in discussion with our clinical experts.
Outcomes	• All-cause mortality • Resolution of acute and chronic GVHD • Steroid weaning • Weaning from GVHD interventions
Study Design	Randomized controlled trials of any duration

For studies that were associated with multiple publications (e.g., updates of different follow-up durations), we retained the most up-to-date report and made note of all related manuscripts. Only studies published in English were retained for inclusion.

2.3. Search strategies to identify relevant literature

In June 2013, requestors of this DSEN query conducted preliminary work with members of the MAGIC (Methods and Applications Group for Indirect Comparisons) team based at The Ottawa Hospital to explore some of the literature available for developing clinical guidance related to allogeneic HSCT. This was conducted in the form of a scoping review of published RCTs, and involved a systematic search for studies for various aspects of care in the realm of HSCT. The search was developed and conducted with the input of an information specialist and covered the following databases: Medline, PubMed, Embase, and the Cochrane Register of Controlled Trials. The search was also peer reviewed by a second information specialist using PRESS criteria.[24] Study citations from the different databases were de-duplicated using

Reference Manager software. The search strategy is provided in **Appendix 1**. The search was updated in June 2015 and the results of the aggregated database are presented in the Results section.

2.4. Process of study selection

Citations were screened for relevance based on title, keywords, and abstract (Level 1 screening) and full text (Level 2 screening) by two independent reviewers. Level 1 citations deemed potentially relevant or lacking sufficient information to make a decision were carried forward to Level 2. Study selection was conducted using Distiller Systematic Review Software (DSR) (Evidence Partners, Inc; Ottawa, Canada). Where consensus was not achieved following discussion, a third independent party (DA) was consulted to settle disagreements. At both stages of screening, a pilot exercise of a number of abstracts/full texts was performed to establish a baseline amongst the reviewers. Studies were not screened on outcome; however, studies included at the full-text stage that did not have an outcome of interest to the review were not moved forward for data extraction. The process of literature selection has been reported using a flow diagram as recommended by the PRISMA Extension Statement for Network Meta-analyses,[23] and encompasses both the 2013 and 2015 searches performed (**Appendix 1**).

2.5. Data collection from included studies

Primary data collection of included studies was performed independently by two reviewers using a standardized electronic data collection form in Microsoft Excel (version 2010, Microsoft Corporation, Seattle, USA). Collected data was compared for accuracy and agreement, with disagreements settled by discussion. The following elements were collected for each included study when available:

- Study characteristics (authors, year of publication, journal, countries of performance);
- Patient data (eligibility criteria; number per group; key demographics, including age, sex, primary diagnosis, comorbidities, Karnofsky performance score (if available); recipient/donor matching for HLA, sex, and CMV status; co-interventions used, including conditioning regimen and intensity category);
- Treatment arm data (drug(s) used as well as dosage and other aspects of administration);
- Outcome data (number of events and number of patients randomized for binary endpoints, and means with standard deviations for continuous endpoints).

To standardize follow-up times for mortality analyses, overall mortality data were extracted from Kaplan-Meier (K-M) survival curves at 100 days and 1 year of follow-up using data digitization software (Digitizelt, © I. Bormann 2001–2016; http://www.digitizeit.de), as for the majority of studies this information was available only from graphs.

All study characteristics were summarized in tabular form to facilitate inspection and discussion with clinical experts in terms of study heterogeneity and grouping of interventions to inform analysis. While the majority of included studies included a mixture of indications, we also wished to develop an awareness of the severity of study populations; two of the research team members (DA, NK) reviewed study characteristics and categorized populations in terms of their extent of high risk patients based on aspects of the criteria described by Kahl et al[25] as well as application of the Disease Risk Index.[26] These assessments were also considered in assessments of heterogeneity between studies.

All relevant RCTs were evaluated using the revised Cochrane risk-of-bias (RoB) tool.[27] The Cochrane RoB tool evaluates seven domains including sequence generation, allocation concealment, blinding, missing outcome data, selective outcome reporting, attrition, and "other sources of bias." Other sources of bias evaluated for this review included dissimilarities between treatment groups, with respect to demographics, comorbidities, co-interventions, and other factors. An overall assessment of the RoB for each study was

determined for each general outcome category. Any disagreements were resolved through discussion or third-party adjudication. Results from the RoB appraisal were summarized narratively.

2.5.1. Outcomes of interest

2.5.1.1. Acute GVHD

Based on input from participating clinical experts, the aGVHD outcome was limited to Grades II–IV aGVHD, based on current scoring systems. Studies that reported only overall aGVHD, with no limits on grade, were excluded from the review.

2.5.1.2. Chronic GVHD

The development of cGVHD generally occurs after day 100 post-transplant. The risk of mortality in the first 100 days post-transplant is generally high, meaning that many patients initially randomized in a study may not survive to the commencement of the cGVHD follow-up period. To account for the competing risk of death in the first 100 days, the most unbiased denominator for cGVHD analyses was considered to be only the patients alive at 100 days post-transplant (as opposed to the number originally randomized).

Amongst included studies, cGVHD was reported either as overall cGVHD or scored as extensive or limited and reported by these grades. Based on input from participating clinical experts, the incidence of extensive cGVHD was chosen as the preferred outcome for the evaluation of cGVHD prophylaxis. Thus, the preferred data for the cGVHD outcome was considered to be the incidence of extensive cGVHD in patients that survived to day 100 post-transplant.

Based on expert opinion, the risk of cGVHD was considered highest in the first 2 years post-transplant. Thus, where data were available for multiple follow-up times in a study, the follow-up period most inclusive of the 2-year time point was preferred over early follow-up or long-term follow-up data.

2.5.1.3. Mortality

Mortality outcomes of interest included overall mortality at 100 days and 1 year post-transplant. Because these data were extracted from K-M curves at the time points of interest, follow-up time was considered consistent between studies. While transplant related mortality was also of interest, several difficulties were encountered in data extraction that precluded the ability to robustly gather data for this endpoint, and thus a formal analysis could not be performed.

2.5.1.4. Specific harms

Upon consultation with clinical experts, the following specific harms of interest were identified: total adverse events, severe adverse events, sepsis, pneumonia, CMV infection, fungal infection, other infections, renal toxicity, hepatic toxicity, veno-occlusive disease (VOD), and other harms (e.g., multi-organ failure, neurotoxicity).

2.6. Classification of interventions for network meta-analysis

Input from clinical expert opinion deemed that analyses for the current review should be performed at the regimen/drug level, with no additional reflection of specific doses when establishing treatment groups for meta-analyses. Thus, all GVHD prophylactic regimens using the same drugs were grouped into the same node in evidence networks, regardless of drug dosage. The same approach was employed for GVHD treatment regimens.

2.7. Methods for meta-analysis and network meta-analysis

A priori, there was interest to explore the feasibility of network meta-analysis for all clinical endpoints given the presence of multiple interventions of interest. For each outcome, we first assessed whether a connected treatment network of interventions was present (a necessary criteria for NMA). When this was present, we next explored the extent of homogeneity/similarity of patient populations and methodology across included studies, as the performance of valid network meta-analyses requires a similar distribution of effect modifiers across studies; where more than one study was present for pairwise comparisons within the treatment network, we also assessed statistical heterogeneity of corresponding treatment effects using the I^2 statistic. If homogeneity within pairwise comparisons and across comparisons within the network was judged sufficient, network meta-analyses were performed. Where studies could not be included in a network due to poor reporting of specific interventions or extensive between-study heterogeneity was present, a narrative summary of the study was undertaken. Similarly, where networks could not be generated due to fragmentation of interventions and low numbers of studies, pairwise ORs were calculated to identify regimen comparisons with significant associations with the various outcomes, and narrative summaries for each outcome were written.

Traditional pairwise frequentist meta-analyses were conducted using the statistical software Comprehensive Meta-analysis© (Biostat, Inc., Englewood, NJ: https://www.meta-analysis.com/). Both fixed and random effects models were fit and assessments of heterogeneity were made using the I^2 statistic. The results of the traditional pairwise meta-analyses were useful in the evaluation of heterogeneity and informal inspection of consistency in the network meta-analyses discussed below.

Bayesian network meta-analyses were conducted using established methods described by the National Institute for Clinical Excellence.[28,29] All analyses were conducted using WinBUGS software and the algorithms presented in technical guidance documents published by NICE. Both fixed and random effects models were fit to arm-level data. All models were fit using 40,000 or more burn-in iterations and 40,000 sampling iterations. Model convergence was assessed by inspection of trace plots and Monte Carlo standard error of the estimated parameters. Model fit was evaluated by posterior residual deviance values for both FE and RE models, with the model fit assumed to be adequate if the posterior residual deviance was similar to the number of data points in the model. DIC values were used to compare the relative fit of the models—a model was identified as having a relative improvement in fit if its DIC value was 5 points or lower than the DIC of another model. However, if the network contained many single-study connections between interventions, results from the FE model were preferred. All of the networks analysed for this report were comprised of many single-study connections; thus, results from FE models are focused on for summary of findings, with findings from random effects models provided in the report's appendices. This approach has previously been seen in applications in other clinical realms and avoids misleading summary estimates which can be a consequence of the presence of a poorly estimated between-study variance parameter.[30]

2.8. Summarizing measures of effect

2.8.1. Summary Measures and Reference Intervention

All outcomes of interest were assessed as binary endpoints, allowing pairwise comparisons (whether from individual studies or from meta-analysis) to be summarized using odds ratios (ORs) and corresponding 95% CrIs. Odds ratios obtained through traditional pairwise meta-analyses were compared to ORs obtained from network meta-analyses to aid in evaluation of consistency between the direct and indirect evidence of the network.

For each outcome of interest, forest plots were generated to provide a visual display of efficacy of each regimen in the network compared to the current treatment standard. For the prophylaxis of GVHD, input from our clinical experts identified the current treatment standard as methotrexate plus tacrolimus

(MTX+TAC). All ORs <1 in for analyses presented in this report suggest the comparator regimen had greater efficacy than MTX+TAC, while ORs >1 indicated the comparison regimen had lower efficacy than MTX+TAC. For all analyses, 95% credible intervals which included the value of 1 were considered to indicate that there was no statistically significant difference between the comparison regimen and MTX+TAC.

2.8.2. Graphical Presentations of Findings

Forest plots of summary comparisons versus the reference intervention (MTX + TAC) have been presented for all endpoints where network meta-analyses were performed. Additionally, *league tables* presenting all pairwise comparisons estimated from network meta-analysis were prepared; in these tables, regimens in the upper/left-most region of the league table have potentially greater efficacy than regimens appearing lower and further right in the table. For each league table, interventions are ordered from left to right in terms of decreasing SUCRA value, meaning preferred interventions are those presented on the left of the table (as SUCRA values nearer 1 suggest preferred interventions). The SUCRA value[31] for each regimen has also been presented above the regimen, allowing evaluation of relative rankings of regimens. One regimen may appear above another regimen in the league table; however, the effect sizes of regimens may not differ greatly in magnitude. Thus, odds ratios and 95% CrIs should be considered the primary means of assessing the importance of differences between regimens.

League tables can be complex to interpret in the presence of many interventions, a situation encountered in the GVHD prophylaxis analyses presented in this review. Furthermore, when evidence networks consist of many interventions and the comparisons made within trials are both broad, there may be varying degrees of faith in pairwise comparisons dependent upon the number of intermediate treatments between therapies; those with just one intermediate therapy are often called *simple indirect comparisons*, while those with more intermediate therapies are called *compound indirect evidence.* Comparisons of treatments involving direct evidence (i.e. where at least one head-head trial informs the comparison of two therapies) may be considered of greatest validity by researchers, while comparisons informed by simple indirect evidence are typically considered of greater validity than comparisons informed by compound indirect evidence. Due to the diverse nature of the evidence networks in this review which consist of several comparisons based on compound indirect evidence, we have color-coded league tables of summary findings to demonstrate key sources of evidence for each comparison as follows:

- Red squares denote pairwise comparisons with at least one head-to-head trial available;
- Orange squares denote pairwise comparisons with simple indirect evidence available;
- Yellow squares denote pairwise comparisons with compound indirect evidence available.

2.8.3. Assessment of heterogeneity and inconsistency for network meta-analysis

An important step in the practice of systematic reviews that incorporate network meta-analyses is the validation of the assumption that patients in the included trials are 'jointly randomizable,' or in other words, that they are sufficiently homogeneous clinically that a patient in any one of the studies could have been a patient in any of the other included trials.[32] We empirically evaluated this assumption by review of the patient eligibility criteria and pertinent patient demographics, in collaboration with our participating clinical experts (DA, NK). To assess the presence of clinical and methodologic heterogeneity within each pairwise comparison of the treatment network, we inspected trial and population characteristics. This was performed by inspection of tabulated lists of descriptive statistics for the a priori characteristics mentioned below (i.e., means and frequency distributions as appropriate for each characteristic), as well as review of measures of statistical heterogeneity mentioned above.

To ensure homogeneity and similarity across pairwise comparisons in the treatment network, we compared the descriptive statistics of key measures across the different pairwise comparisons in the network to verify

they were similar. To identify covariates necessary for review, we consulted our clinical expert team members and grouped traits that were identified in past studies of prognostic risk factors.[33-36] The following characteristics were considered most important to the establishment of transitivity within and across pairwise comparisons:

- Average patient age, gender distribution, and race distribution;
- % with different diagnoses (AML, ALL, CML, etc.);
- Disease risk (based on clinical experts' assessments);
- Presence of comorbidities (pulmonary disease, rheumatologic disease, renal dysfunction, etc.), if reported;
- Year of study publication (for consideration of changes in co-interventions such as newer antimicrobials, molecular screening tests for infection, and newer conditioning regimens used);
- % of patients receiving an unrelated donor transplant;
- % of patients with full HLA match;
- % of patients with mismatched donor gender;
- % of patients that were cytomegalovirus (CMV)-seronegative and received transplants from CMV-seropositive donors;
- source of donor cells (bone marrow, peripheral blood stem cells, umbilical cord blood).

Another key assumption underlying NMA is that of *consistency*. That is, there is no conflict between direct and indirect evidence[29] that could result from heterogeneity in effect modifiers in the studies contributing to the direct and indirect evidence. To assess inconsistency in cases where networks consisted of one or more closed loops, we compared DIC statistics between consistency and inconsistency models. We also inspected plots of the posterior mean deviance of the individual data points in the inconsistency model against their posterior mean deviance in the consistency model to identify any loops in the network where inconsistency was present.

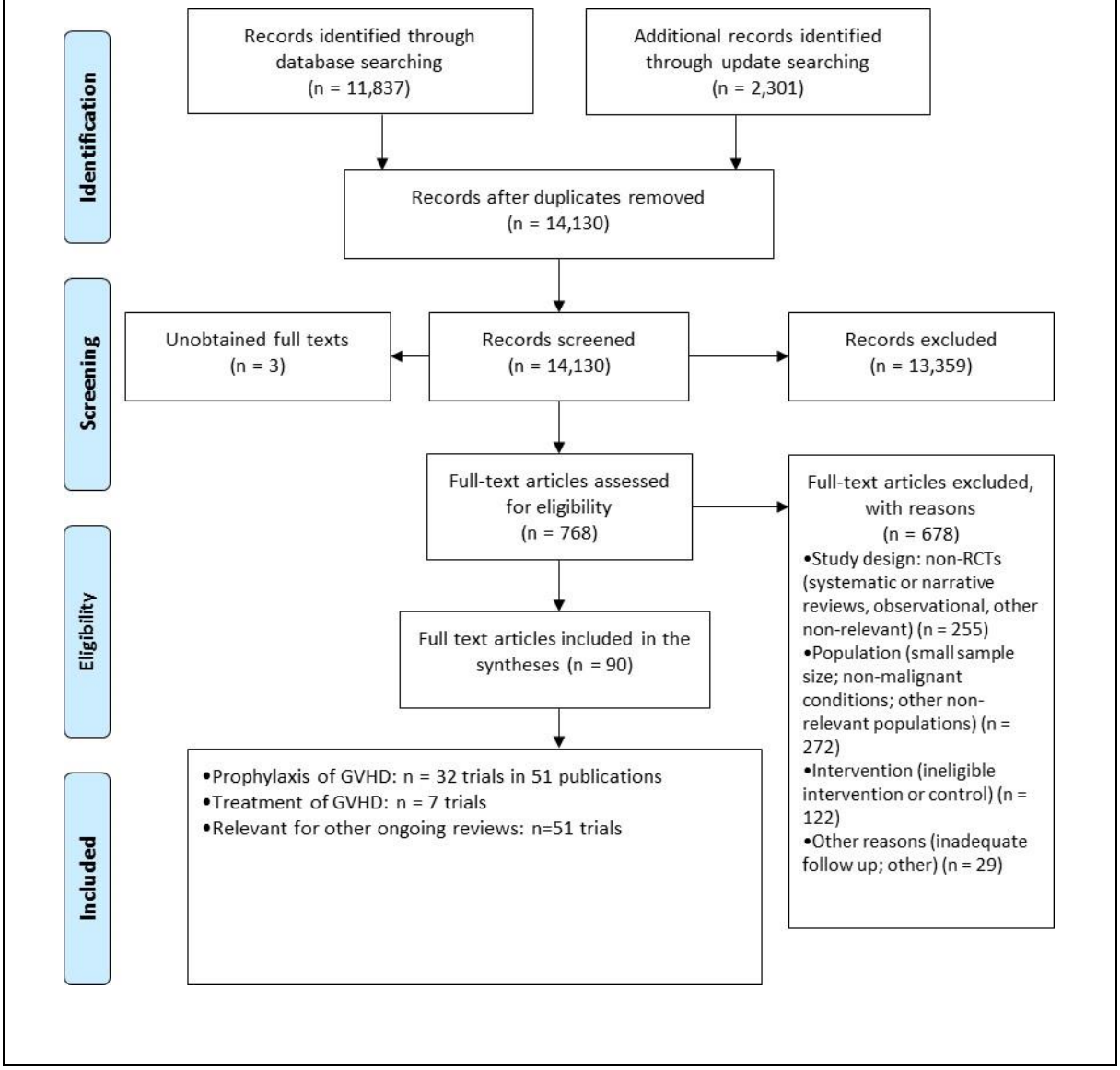
Although planned in the review's protocol to further establish the robustness of findings from primary analyses, subgroup analyses and meta-regression analyses were not feasible in the context of the GVHD prophylaxis review due to the presence of many single-study connections in most evidence networks, as well as a failure of studies to report outcomes in patient subgroups of relevance.

3. REVIEW FINDINGS—GVHD PROPHYLAXIS

3.1. Availability of relevant literature

Figure 1 presents a summary of the process of study selection. The electronic literature search identified a total of 14,130 citations for review following removal of duplicates. Stage 1 screening identified a total of 768 citations that were considered to be potentially eligible, and which were subsequently retrieved in full text for Stage 2 screening. Following screening of full text articles, a total of 109 publications describing 90 unique studies were retained, with 51 publications describing 32 unique trials of GVHD prophylaxis, and 7 trials in 7 unique publications for the review of GVHD treatment. The remaining studies retained were kept to serve as the evidence base for other ongoing reviews related to stem cell transplantation being performed for the knowledge users of the current review.

Figure 1: PRISMA flow diagram of study selection



3.2. Overview of study characteristics

Table 3 presents a detailed summary of study characteristics. A total of 3,875 patients were enrolled in the 32 included studies for GVHD prophylaxis. Median sample size of the included trials was 106 (range 54–329). Median year of publication was 2000 (range 1979–2015); 1 (3.1%), 7 (21.8), 8 (25.0%), 6 (18.8%), and 10 (31.3%) studies were published before 1980, between 1980–1990, between 1991–2000, between 2001–2010 and after 2010, respectively. Fourteen (43.8%) and 15 (46.9%) studies were single-centre and multi-centre studies, respectively, while 3 studies (9.4%) didn't report this information. Three trials (9.4%) were entirely supported by for-profit/industry sources, while 18 (56.3%) were funded by non-profit/government organizations, 6 (18.8%) had support from both private and public sources and no information could be obtained for the remaining 5 trials (15.6%). Study duration varied greatly, with 7 trials (21.9%) following patients for up to one year, 16 (50.0%) for up to 5 years, and 9 (28.1%) following beyond 5 years. One trial followed patients for 21 years.

3.3. Overview of patient characteristics

A total of 15 studies (46.9%) enrolled patients 12 years of age or younger, while the remaining 17 (53.1%) did not. The median proportion of female patients amongst the 23 studies (93.8%) reporting this information was 41.0% (range 12.0%–56.0%). Donor source was 100% marrow in 21 studies (65.6%), 100% peripheral blood in 5 studies (15.6%), mixed in 4 studies (12.5%), and unclear or not reported in 2 studies (6.3%). Regarding donor status, 100% of patients received cells from a related donor in 17 studies (53.1%), no patients received cells from a related donor in 4 studies (12.5%), between 29%–91% patients received cells from a related donor in 8 studies (25.0%), and 3 studies (9.4%) did not report this information. For HLA match status, 21 studies (65.6%) had 100% of patients with full match, while 6 studies (18.8%) had between 68%–96% of patients with a full match and 5 studies (15.6%) did not report this information. Regarding the proportion of male recipients receiving cells from a female donor, 19 studies (59.4%) did not report this information, while in the remaining 13 studies, this proportion ranged from 13%–28%. Conditioning regimens used for patients' transplants include myeloablative regimens in 5 studies and reduced intensity in 2 studies, while intensity was unclear in all remaining studies.

Most studies reported the distribution of patients' underlying hematologic diseases. However, limitations in reporting were identified; some studies reported the proportion of patients with acute leukemia without reporting the distribution of ALL and ANL patients; and some studies reported the proportion of patients with lymphoma, but without reporting the distribution of Hodgkin and non-Hodgkin lymphoma. In these cases, patients with both diseases (i.e., ALL and ANL, or Hodgkin and non-Hodgkin lymphoma) were assumed to have been included in the study sample. Some older studies included patients with CML, however CML is no longer treated with HSCT. Although the population sampled in these studies may no longer be clinically relevant, the effects of the interventions on the GVHD outcomes may be unaffected by the underlying hematologic disease. Thus, these studies were included in the network analyses and where possible, sensitivity analyses were conducted to determine if CML patients had an impact on the results. Similarly, studies that included aplastic anaemia patients were included in the analyses, despite aplastic anaemia not being a hematologic malignancy, and the impact of these patients on the results was evaluated with a sensitivity analysis. The Table in **Appendix 2** reports the distribution of underlying disease in the included studies. Assessment of disease risk based on the methods described earlier found that (1) 16 studies provided insufficient information to establish the proportion of high risk patients; (2) one study[37] enrolled 100% high risk patients while 2 studies[38,39] enrolled none; and (3) amongst the remaining studies, the enrolled proportion of high risk patients ranged between 2% to 69% (**Table 3**). Findings from these evaluations highlight both limited information reported regarding study populations in some cases as well as a diverse range in populations across studies.

Table 3: Summary of characteristics of studies included in the systematic review of regimens for GVHD prophylaxis										
Author (year)	**N total (N = 3875)**	**Regimens compared**	**Intensity of Conditioning Regimen**	**Patients' age range (years)**	**Donor source (%BM, %PB, %CB)**	**% related donor**	**% full HLA match**	**% female donor / male recipient**	**% of study patients considered to be high risk**	**Endpoints reported**
Kuzmina[40-42] (2015)	67	CSA+MTX; CSA+MTX+MSCs	MA & RI	18–63	100,0,0	100%	100%	NR	NR	α, γ
Cutler[38,43] (2014)	304	MTX+TAC; SIR+TAC	MA	13–59	0,100,0	100%	100%	21%	2%	α, β, γ
Glass[37] (2014)	84	MMF+TAC; RITX+TAC+MMF	MA	38–57	4,96,0	29%	68%	13%	100%	α
Ruutu[44] (2014)	242	CSA+MTX+Steroids; CSA+MTX+UDCA+ Steroids	Unclear	1–59	79,21,0	55%	96%	NR	42%	β, γ
Bonifazi[45] (2014)	155	CSA+MTX; CSA+MTX+ATG+ Steroids	MA	18–65	NR	NR	NR	NR	NR	β
Armand[46] (2013)	138	MTX+TAC; MTX+TAC+SIR	RI	23–70	0,100,0	NR	NR	NR	NR	α, β, γ
Martin[39] (2012)	138	MTX+TAC; MTX+TAC+Steroids	MA	8–63	25,75,0	36%	84%	28%	0%	α
Pidala[47] (2012)	74	MTX+TAC; SIR+TAC	MA	16–70	0,100,0	47%	100%	NR	15%	α
Pulsipher[48,49] (2014)	143	MTX+TAC; MTX+TAC+SIR	MA	1–21	Unclear	NR	NR	NR	69%	α, β, γ
Socie[50-54] (2011)	201	CSA+MTX; CSA+MTX+ATG+ Steroids	MA	18–60	18,82,0	0%	100%	13%	47%	α, β, γ
Perkins[55] (2010)	89	MTX+TAC; MMF+TAC	MA, non-MA	23–70	0,100,0	48%	94%	27%	37%	α, β, γ
Fong[56] (2007)	104	HCQ+CsA; CsA	Unclear or other	20–65	0,100,0	100%	100%	25%	NR	β, γ
Lee[57] (2004)	80	CsA; CsA+MTX	MA	15–49	100,0,0	100%	100%	16%	15%	α, β, γ
Hiraoka[58] (2001)	136	CSA+MTX+Steroids; TAC+MTX+Steroids	MA	2–56	100,0,0	47%	95%	NR	24%	α, β, γ
Bacigalupo[59] (2001)-I	54	CSA+MTX; CSA+MTX+ATG+ Steroids	Unclear	13–51	100,0,0	0%	100%	NR	NR	α, β, γ

Table 3: Summary of characteristics of studies included in the systematic review of regimens for GVHD prophylaxis										
Author (year)	**N total (N = 3875)**	**Regimens compared**	**Intensity of Conditioning Regimen**	**Patients' age range (years)**	**Donor source (%BM, %PB, %CB)**	**% related donor**	**% full HLA match**	**% female donor / male recipient**	**% of study patients considered to be high risk**	**Endpoints reported**
Bacigalupo[59] (2001)-II	55	CSA+MTX; CSA+MTX+ATG+ Steroids	MA	14–52	100,0,0	0%	100%	NR	NR	α, β, γ
Chao[60] (2000)	186	CsA+MTX; CsA+MTX+Steroids	MA	1–50	100,0,0	100%	100%	26%	25%	α, γ
Deeg[61] (2000)	122	CsA; CsA+Steroids	MA, non-MA	12–59	100,0,0	100%	100%	NR	NR	α, β
Nash[62] (2000)	180	CSA+MTX; TAC+MTX	MA	12–61	100,0,0	0%	83%	15%	10%	α, β, γ
Ruutu[63] (2000)	108	CsA+MTX; CsA+MTX+ Steroids	MA	14–54	100,0,0	100%	100%	15%	35%	α, β, γ
Ross[64] (1999)	149	CSA+Steroids; CSA+MTX+Steroids	MA	1–50	100,0,0	100%	NR	24%	NR	α, β, γ
Ratanatharathorn[65,66] (1998)	329	CSA+MTX; TAC+MTX	MA	16–63	100,0,0	100%	100%	NR	35%	α, β, γ
Zikos[67] (1998)	60	CSA; CSA+MTX	MA	11–43	100,0,0	100%	100%	NR	18%	α, β
Chao[68] (1996)	59	THAL+Standard prophylaxis[a]; Placebo+Standard prophylaxis[a]	MA	>18	100,0,0	91%	100%	NR	31%	β*
Mrsic[69] (1990)	76	CSA; CSA+MTX	MA	5–43	100,0,0	100%	100%	22%	18%	α, β, γ
Torres[70] (1989)	57	CSA; MTX	MA	18–46	100,0,0	100%	100%	14%	NR	α, γ
Backman[71] (1988)	59	CSA; MTX	MA	1–46	100,0,0	100%	100%	NR	14%	α, β, γ
Forman[72] (1987)	107	MTX+Steroids; CSA+Steroids	Unclear	6–39	100,0,0	NR	NR	NR	46%	α
Storb [73-76] (1986)	93	CsA; CsA+MTX	MA	<1% under 18	100,0,0	100%	100%	NR	NR	α, β, γ
Irle[77] (1985)	56	MTX; CSA	MA	30–47	100,0,0	100%	100%	NR	n/a????	α, β

Table 3: Summary of characteristics of studies included in the systematic review of regimens for GVHD prophylaxis

Author (year)	N total (N = 3875)	Regimens compared	Intensity of Conditioning Regimen	Patients' age range (years)	Donor source (%BM, %PB, %CB)	% related donor	% full HLA match	% female donor / male recipient	% of study patients considered to be high risk	Endpoints reported
Storb [78-80] (1983)	114	CsA; MTX	MA	11–49	100,0,0	100%	100%	NR	NR	α, β, γ
Weiden[81] (1979)	56	MTX; MTX+ATG+Steroids	MA	1–47	100,0,0	100%	100%	NR	54%	α, β, γ

Studies are listed in reverse chronologic order by publication date. Endpoints reported by each study are denoted using symbols in the final column, where α = aGVHD, β = cGVHD, γ = 100-day and 1-year overall mortality.

*Endpoint was not included in a NMA but was summarized narratively.

[a]Patients in both arms received one of the following standard GVHD prophylactic regimens: CsA+Steroids, CsA+MTX, CsA+MTX+Steroids and CsA+MTX+ATG+Steroids.

ATG = anti-thymocyte globulin; BM = bone marrow; CsA = cyclosporin A; HCQ = hydroxychloroquine; HLA = human leukocyte antigen; MA = myeloablative; MMF = mycophenolate mofetil; MTX = methotrexate; NMA = non-myeloablative; NR = not reported; PB = peripheral blood; RI = reduced intensity; RITX = rituximab; SIR = sirolimus; TAC = tacrolimus; THAL = thalidomide; UC = umbilical cord; UDCA = ursodeoxycholic acid

3.4. Overview of network geometry and interventions compared

Generally, the networks of the evidence available for each of the outcomes of interest had high proportions of single-study connections. **Table 4** presents the numbers of interventions, studies, and single-study connections available for each of the outcomes assessed.

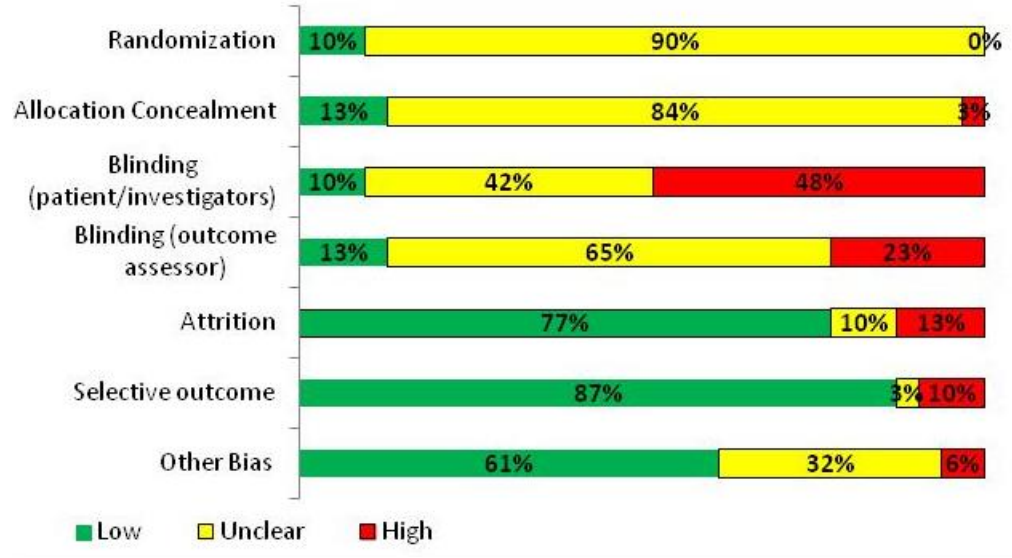
Table 4: Summary of interventions compared, number of studies, and number of single-study connections by outcomes analysed by NMA for GVHD prophylaxis

Clinical Outcome	Number of interventions	Number of studies reporting outcome	Number of single-study connections Number of head-to-head connections Total number of possible connections
Acute GVHD	15	28	10 single-study connections 16 head-to-head connections 105 possible connections
Chronic GVHD	14	23	9 single-study connections 14 head-to-head connections 91 possible connections
1-year overall mortality	15	23	8 single-study connections 14 head-to-head connections 105 possible connections
100-day overall mortality	15	23	8 single-study connections 14 head-to-head connections 105 possible connections
Relapse	11	17	6 single-study connections 10 head-to-head comparisons 55 possible connections

3.5. Findings from Risk of Bias Assessment

Risk of bias assessment was possible for all but one of the 32 studies (96.9%). This study (Bonifazi et al) was reported in abstract format and was not assessed given the insufficient reporting of methods used for the trial. Adequate sequence generation was clearly reported in 3 of the studies (9.7%), while method of random sequence generation was not clearly reported for the remaining 28 studies (90.3%). Four of 31 studies (12.9%) had low risk of bias for allocation concealment, while totals of 1 (3.2%) and 26 (83.9%) studies had high and unclear risk for this domain, respectively. Patients and investigators were blinded from regimen allocation in 3 studies (9.7%) and not blinded in 15 studies (48.4%). Independent and blinded outcome assessment was reported in 4 studies (12.9%), while 7 studies (22.6%) had high risk of bias given lack of blinding, and 20 studies (64.5%) did not report sufficient information regarding this domain. Twenty-four studies (77.4%) used intention-to-treat (ITT) or modified ITT analysis for the main outcomes of aGVHD, mortality, and relapse, and therefore had low risk of attrition bias. However, the outcome of chronic GVHD was not always reported for the entire population given the rate of attrition due to mortality or relapse before 100 days post-transplant; therefore, this particular outcome might be subject to a high risk of attrition bias (data not shown). The majority of studies (87.1%) provided reference to the study protocol and had low risk of bias for selective outcome reporting, while this information was either not reported in 3 (9.7%), or suggested high risk in one study (3.2%). Other potential sources of bias deemed important by the reviewers included imbalances in baseline patient characteristics, including co-interventions administered and other potentially important factors. Data for baseline variables and co-intervention were poorly reported in 10 studies (32.3%), preventing the reviewers from objectively assessing these potential other biases, rendering the risk for this domain unclear. The remaining 19 (61.3), and 2 (6.5%) studies had low and high risk of other biases, respectively. A complete listing of risk of bias assessments for all included studies is presented in **Appendix 3**. A summary of the distribution of risk of bias across studies is provided in **Figure 2**.

Figure 2: Summary of Risk of Bias Assessments of Included Trials, GVHD prophylaxis (n = 31 studies)



3.6. Findings: Acute GVHD

A total of 28 studies reported aGVHD incidence in 3,304 patients and were analysed using NMA. One other study (59 patients) was not included in the analysis and summarized narratively due to ambiguity in the reported prophylactic regimens.[68] Many studies (n = 11) did not cite a reference for the grading of aGVHD, and thus the degree of homogeneity of the outcome definitions was unclear. Of the 17 studies that did cite a grading system, 3 different citations were referenced.[8,82,83] Follow-up for the aGVHD outcome was generally consistent at 100 days in the included studies.

Figure 3 presents a network diagram highlighting the available evidence for the comparison of prophylactic regimens for GVHD, where the incidence of acute GVHD was evaluated as an outcome. Each of the 15 prophylactic regimens is represented by a node and the randomized trials comparing pairs of regimens are shown as links between the nodes. The size of the node is reflective of the total number of patients enrolled in studies evaluating that regimen, while the width of the link between two nodes is reflective of the number of studies that directly compared the two regimens. Data from head-to-head trials were available for 16 of the 105 potential pairwise comparisons in the network, with single studies informing 9 of the 16 head-to-head comparisons.

Figure 3: Network diagram, Acute GVHD (15 interventions, 3.304 patients)

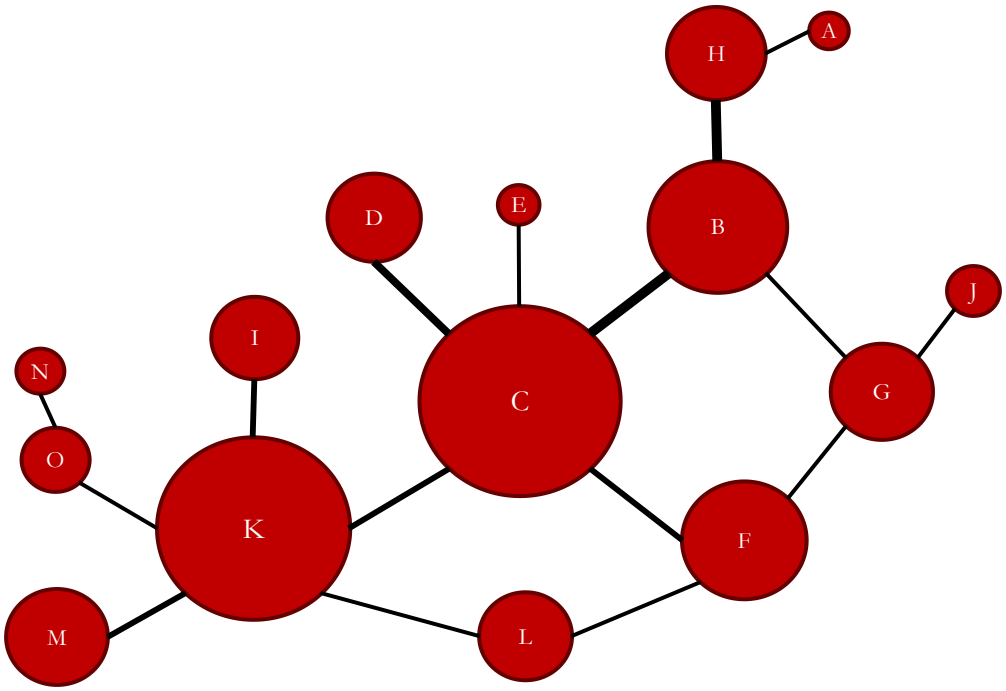


Figure legend, Treatments Shown	H = MTX (n = 177)
A = ATG+MTX+Steroids (n = 29)	I = MTX+SIR+TAC (n = 139)
B = CsA (n = 352)	J = MTX+Steroids (n = 51)
C = CsA+MTX (n = 739)	K = MTX+TAC (n = 681)
D = CsA+MTX+ATG+Steroids (n = 159)	L = MTX+TAC+Steroids (n = 158)
E = CsA+MTX+MSCs (n = 32)	M = SIR+TAC (n = 188)
F = CsA+MTX+Steroids (n = 283)	N = RITX+TAC+MMF (n = 42)
G = CsA+Steroids (n = 190)	O = TAC+MMF (n = 84)

3.6.1. Results from traditional pairwise meta-analyses

Table 5 presents a summary of pairwise estimates that were derived from direct evidence (i.e., head-to-head trials). In the table, these results are summarized alongside the numbers of available trials and patients for each direct comparison, as well as the related summary estimate derived from the network meta-analysis which is described below; this has been done to facilitate comparison of results between analyses.

The direct comparisons with the largest numbers of studies available were MTX vs CsA (5 RCTs, n = 361) and CsA vs CsA+MTX (5 RCTs, n = 374). Pairwise findings suggested a significantly reduced risk of acute GVHD with CsA compared to CsA+MTX, CsA+MTX compared to MTX+TAC, CSA+MTX compared CSA+MTX+ATG+Steroids, MTX+TAC compared to MTX+TAC+SIR, MTX+TAC compared to SIR+TAC, CSA+MTX+Steroids compared to MTX+TAC+Steroids, CSA+MTX compared to CSA+MTX+Steroids, and MTX+Steroids compared to CSA+Steroids. Estimates from traditional pairwise meta-analyses compared well with estimates derived from network meta-analysis.

Table 5: Summary of results from meta-analysis and NMA, Acute GVHD				
Comparison		**# of Trials (patients)**	**Direct Estimate OR (95% CrI)**	**NMA Estimate OR (95% CI)**
			**Values <1 favor comparator*	
Reference	**Comparator**			
MTX	ATG+MTX+Steroids	1 (56)	0.33 (0.06–1.85)	0.30 (0.04–1.66)
MTX	CSA	5 (361)	0.65 (0.42–1.00)	**0.65 (0.42–0.99)**
CSA	CSA+Steroids	1 (122)	0.54 (0.25–1.16)	0.65 (0.34–1.22)
CSA	CSA+MTX	5 (374)	**0.50 (0.32–0.77)**	**0.46 (0.30–0.69)**
CSA+MTX	MTX+TAC	2 (509)	**0.53 (0.37–0.77)**	**0.47 (0.33–0.67)**
CSA+MTX	CSA+MTX+MSCs	1 (66)	0.25 (0.06–1.01)	0.23 (0.06–1.01)
CSA+MTX	CSA+MTX+ATG+Steroids	3 (310)	**0.42 (0.19–0.92)**	**0.43 (0.26–0.68)**
CSA+MTX	CSA+MTX+Steroids	2 (294)	0.56 (0.31–1.02)	0.65 (0.40–1.05)
MTX+TAC	MTX+TAC+SIR	1 (139)	**0.31 (0.11–0.83)**	**0.29 (0.10–0.77)**
MTX+TAC	SIR+TAC	2 (378)	**0.51 (0.32–0.81)**	**0.47 (0.30–0.74)**
MTX+TAC	MTX+TAC+Steroids	1 (138)	0.95 (0.45–2.02)	0.56 (0.30–1.04)
MTX+TAC	MMF+TAC	1 (89)	1.01 (0.37–2.78)	0.99 (0.35–2.85)
MMF+TAC+RTX	MMF+TAC	1 (84)	0.91 (0.38–2.15)	0.91 (0.38–2.17)
CSA+MTX+Steroids	MTX+TAC+Steroids	1 (131)	**0.24 (0.11–0.54)**	**0.41 (0.22–0.74)**
CSA+MTX	CSA+MTX+Steroids	1 (149)	**0.35 (0.13–0.89)**	**0.46 (0.23–0.90)**
MTX+Steroids	CSA+Steroids	1 (105)	**0.43 (0.19–0.97)**	**0.43 (0.18–0.95)**

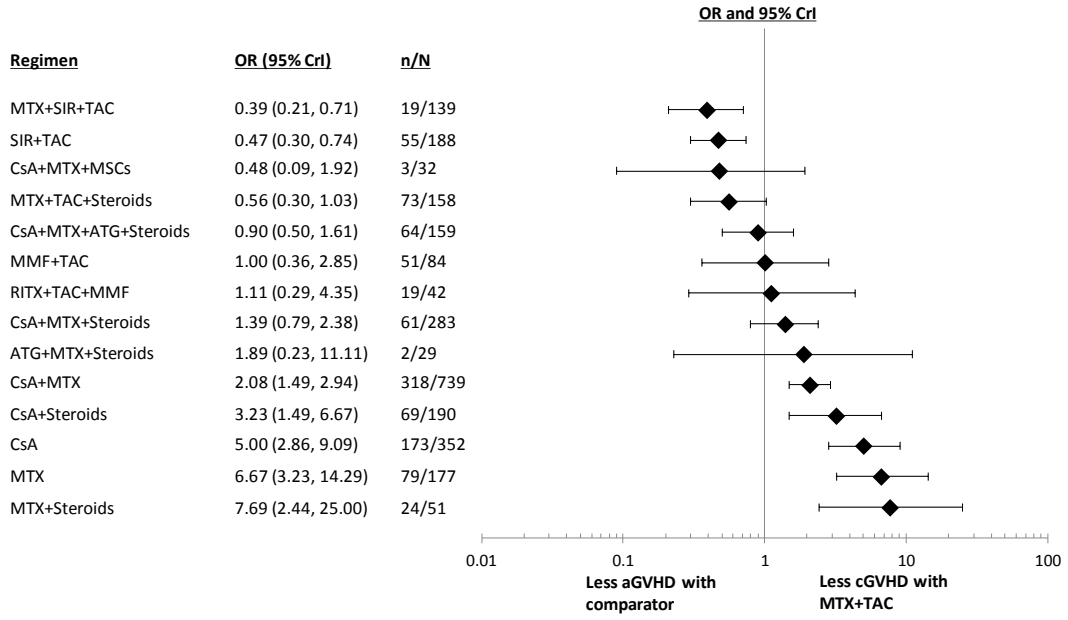
3.6.2. Results from network meta-analysis

Most comparisons (i.e., 89 of 105) were informed only by indirect evidence and many of the comparisons with direct evidence were informed by single studies with limited numbers of patients. Model fit statistics for the NMA indicated that both the fixed-effects (FE) and random-effects (RE) models had an adequate fit. Posterior residual deviance values for the FE and RE models of 78.5 and 57.5, respectively, were obtained, the latter of which was a more desirable fit, given the 56 data points in the model. DIC values (351.874 versus 340.692) suggested some difference in fit favouring the RE model; however, the FE model was preferred due to concerns regarding the potential impact of the high number of single-study connections in the network on the ability to obtain a realistic estimate of the between study variance parameter. Results from the RE model have also been prepared and are presented in full in the appendices to the report.

3.6.2.1. Comparisons versus MTX+TAC

Figure 4 presents a forest plot summarizing comparisons of all treatments in the evidence network to the chosen reference therapy, MTX+TAC, based on the FE analysis. Two regimens—MTX+SIR+TAC and SIR+TAC—were found to have statistically significantly reduced risk of aGVHD relative to MTX+TAC (range of ORs from 0.39–0.47); analogously, 5 regimens (CsA+MTX, CsA+Steroids, CsA, MTX, and MTX+Steroids) were associated with significantly increased risks compared to MTX+TAC (range of ORs 2.08–7.69). Comparisons involving all other regimens were associated with credible intervals that included 1 and, thus, were not significantly different from MTX+TAC in preventing aGVHD. In the RE analysis, only SIR+TAC remained associated with a statistically significant reduction of risk while CsA and MTX remained associated with significantly increased risk.

Figure 4: Forest plot of comparisons from FE NMA versus MTX+TAC - Acute GVHD

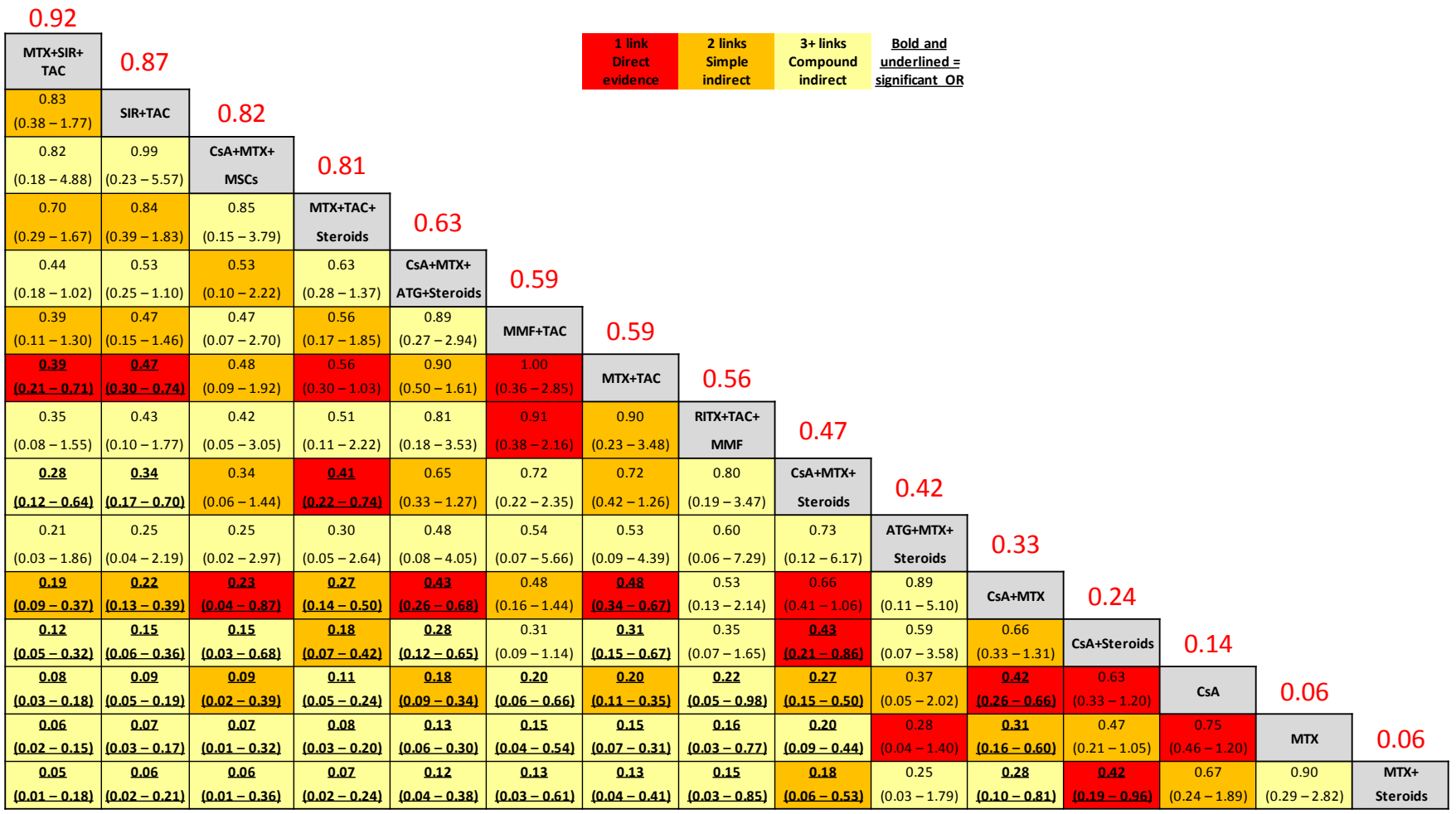


Odds ratios from network meta-analysis comparing each regimen to the reference treatment of MTX+TAC are presented along with the number of events and patients per regimen. Odds ratios <1 suggest more risk of acute GVHD with MTX+TAC while odds ratios >1 suggest less risk of acute GVHD with MTX+TAC. Results from the FE analysis are shown.

3.6.2.2. Comparisons between other prophylactic regimens

Figure 5 presents a league table of the estimates for all pairwise comparisons generated from the FE network meta-analysis. Comparisons of the 6 top-ranked regimens based on SUCRA value (MTX + SIR + TAC, SIR + TAC, CsA + MTX + MSCs, MTX + TAC + Steroids, CsA + MTX + ATG + Steroids, MMF + TAC) were only informed by indirect evidence (i.e., there were no head-to-head trials) and no significant differences in efficacy between these 6 regimens were found. Some of the other regimens were evaluated on very low numbers of patients, potentially reducing the power to detect significant differences in efficacy relative to other regimens (e.g., ATG+MTX+Steroids). Some of the comparisons included in the league table should be interpreted carefully, specifically those which are not informed by head to head trials or where the comparison was based on compound indirect evidence (i.e., comparisons in **Figure 3** where there are 2 or more intermediate treatments between the interventions of interest). Results from the corresponding RE analysis found that while point estimates of pairwise comparisons remained relative unchanged, corresponding 95% credible intervals widened noticeably in many cases, and in some cases differences between therapies no longer were statistically significant.

Figure 5: League table summary of pairwise comparisons from FE NMA, Odds Ratios with 95% CrI, Acute GVHD



Interventions are sorted from left to right in order of decreasing SUCRA value (i.e., preferred treatments appear first). For each comparison shown, the upper/left-most regimen is the comparator group while the lower/right-most treatment is the reference treatment; a value <1 suggests fewer cases of acute GVHD with the comparator than with the reference group. Statistically significant differences (i.e., estimates with a 95% credible interval excluding 1) are bolded and underlined. Comparisons with direct evidence, simple indirect evidence, and compound indirect evidence have been highlighted in red, orange, and yellow, respectively.

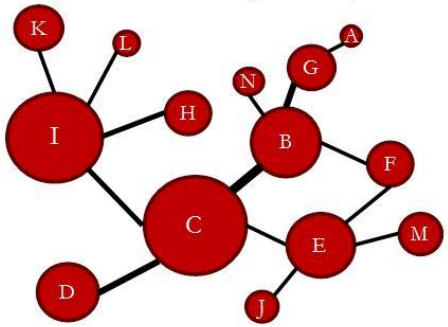
3.7. Findings: Chronic GVHD

Based on input from participating clinical experts, frequency of cGVHD associated with treatment regimens was compared based upon the incidence of extensive cGVHD in patients surviving to day 100 post-transplant. However, not all studies reported the incidence of extensive cGVHD, and not all studies reported the number of patients that were evaluable for cGVHD at day 100 (i.e., survived to day 100). To assess the feasibility of analyses, four potential cGVHD networks were generated; their attributes and network diagrams are presented in **Table 6**. The number of studies reporting data decreased progressively as *overall cGVHD* was changed to *extensive cGVHD*, and as *total randomized patients* was changed to *total evaluable patients*. The evidence network became fragmented when the preferred data—extensive cGVHD in evaluable patients—were considered (**Table 6**). Due to the paucity of data available for extensive cGVHD and for evaluable patients, only the occurrence of overall cGVHD in the total number of randomized patients was analysed using NMA, represented by the first network diagram in **Table 6**. Use of total randomized patients as the denominator may reduce the apparent incidence of cGVHD for regimens with relatively high mortality in the first 100 days post-transplant, compared to the true incidence of cGVHD had the number of evaluable patients been used as the denominator. The level and direction of bias introduced into the NMA is unclear. The results of NMAs for extensive cGVHD in evaluable patients and overall cGVHD in evaluable patients have been summarized in **Appendix 7**.

For the analysis of overall cGVHD in randomized patients, 14 prophylactic regimens had available data from a total of 23 studies that enrolled a total of 3,090 patients (**Table 6**). Data from head-to-head trials were available for 14 of the 91 potential pairwise comparisons in the network, with single studies informing 9 of the 14 head-to-head comparisons. Most studies had a median follow-up of at least 2 years; however, minimum follow-up in these studies may have been considerably less than 2 years. Three of 23 studies followed patients for less than 2 years (i.e., 0.5-year minimum, 1.2-year median, and 1.8-year median).[57,73,81]

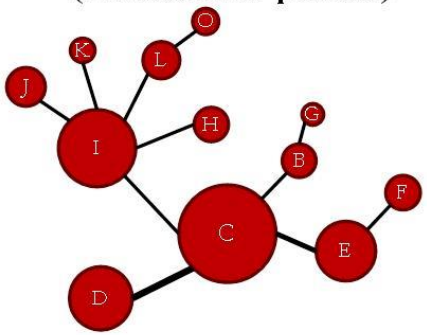
Table 6: Summary of number of interventions, studies, and patients in four potential networks for cGVHD

Overall cGVHD
(# randomized patients)



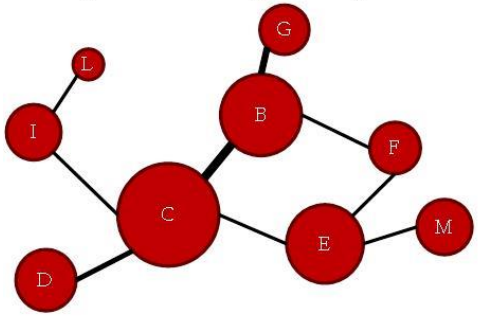
23 studies, 3,090 patients

Extensive cGVHD
(# randomized patients)



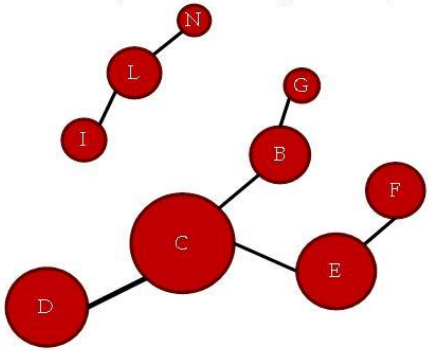
14 studies, 1,904 patients

Overall cGVHD
(# evaluable patients)



15 studies, 1,344 patients

Extensive cGVHD
(# evaluable patients)



8 studies, 714 patients

Figure legend
A = ATG+MTX+Steroids
B = CsA
C = CsA+MTX
D = CsA+MTX+ATG+Steroids
E = CsA+MTX+Steroids
F = CsA+Steroids
G = MTX
H = MTX+SIR+TAC
I = MTX+TAC
J = MTX+TAC+Steroids
K = SIR+TAC
L = TAC+MMF
M = UDCA+CsA+MTX+Steroids
N = HCQ+CsA

3.7.1. Results from traditional pairwise meta-analyses

Table 7 presents a summary of pairwise estimates that were derived from direct evidence (i.e., the head-to-head trials), summarized alongside the numbers of available trials and patients for each direct comparison, as well as the related summary estimate derived from network meta-analysis.

The direct comparisons with the largest numbers of studies available were CsA vs CsA+MTX (4 RCTs, n = 309), CsA+MTX+ATG+Steroids vs CsA+MTX (3 RCTs, n = 465), and CsA vs MTX (3 RCTs, n = 229). Pairwise meta-analysis suggested a significantly reduced risk of chronic GVHD with CsA+MTX+ATG+Steroids compared to CsA+MTX, and with CsA when compared to CsA+Steroids. Overall, estimates from traditional pairwise meta-analyses compared well with estimates derived from network meta-analysis, except for one comparison—CsA vs CsA+Steroids. When indirect evidence was included by the NMA to estimate the OR, CsA was no longer more efficacious than CsA+Steroids.

Table 7: Summary of results from pairwise meta-analysis and NMA, Overall chronic GVHD in randomized patients

Comparison		# of Trials (patients)	Direct Estimate OR (95% CrI)	NMA Estimate OR (95% CI)
			**Values <1 favor comparator*	
Reference	**Comparator**			
MTX+TAC	SIR+TAC	1/304	0.73 (0.46 – 1.15)	0.73 (0.47 – 1.14)
UDCA+CsA+MTX+Steroids	CsA+MTX+Steroids	1/242	0.97 (0.58 – 1.60)	0.96 (0.58 – 1.61)
CsA+MTX+ATG+Steroids	CsA+MTX	3/465	**0.33 (0.22 – 0.49)**	**0.32 (0.22 – 0.47)**
MTX+SIR+TAC	MTX+TAC	2/282	0.94 (0.57 – 1.56)	0.94 (0.57 – 1.56)
CsA	CsA+Steroids	1/122	**0.37 (0.16 – 0.83)**	0.69 (0.36 – 1.31)
MTX+TAC+Steroids	CsA+MTX+Steroids	1/131	0.97 (0.49 – 1.93)	0.97 (0.49 – 1.94)
CsA	MTX	3/229	0.95 (0.55 – 1.64)	0.95 (0.54 – 1.66)
CsA+MTX	CsA	4/309	0.65 (0.40 – 1.06)	0.81 (0.51 – 1.27)
CsA+MTX	MTX+TAC	2/509	0.99 (0.69 – 1.42)	0.99 (0.69 – 1.42)
MTX+TAC	TAC+MMF	1/89	0.91 (0.35 – 2.40)	0.90 (0.33 – 2.40)
CsA+MTX+Steroids	CsA+Steroids	1/149	0.97 (0.51 – 1.85)	0.63 (0.35 – 1.12)
CsA+MTX	CsA+MTX+Steroids	1/108	1.62 (0.74 – 3.53)	0.88 (0.46 – 1.71)
ATG+MTX+Steroids	MTX	1/56	0.24 (0.03 – 2.30)	0.21 (0.00 – 2.03)
HCQ+CsA	CsA	1/95	0.53 (0.23 – 1.21)	0.53 (0.22 – 1.21)

3.7.2. Results from network meta-analysis

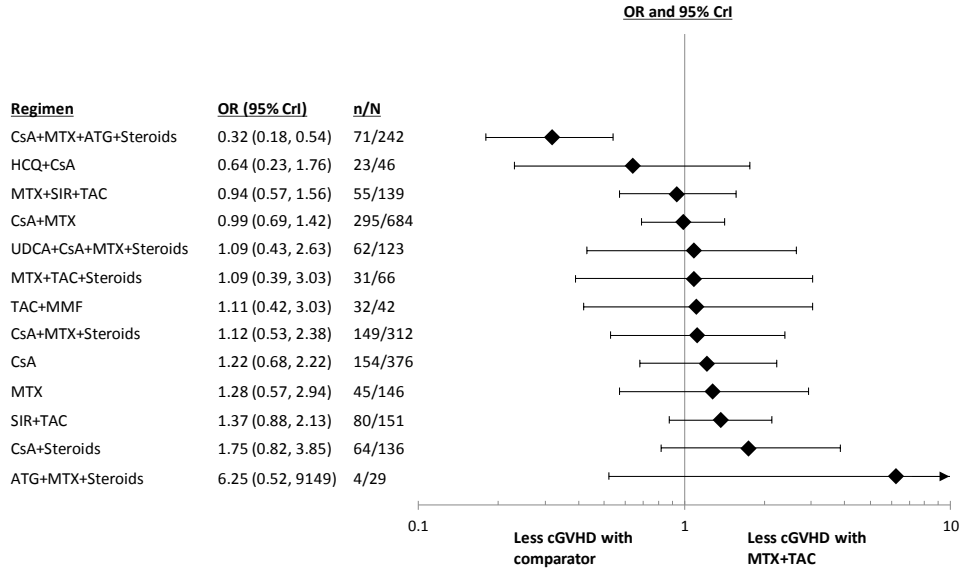
Most comparisons (i.e., 77 of 91) were informed only by indirect evidence and many of the comparisons with direct evidence were informed by single studies with limited numbers of patients. Model fit statistics

for the NMA indicated that both the fixed-effects (FE) and random-effects (RE) models had an adequate fit. Posterior residual deviance values for the FE and RE models of 55.85 and 47.29, respectively, were obtained, the latter of which was a more desirable fit, given the 46 data points in the model. DIC values (292.873 versus 289.693) suggested some difference in fit favouring the RE model; however, a FE model was preferred, regarding the potential impact of the high number of single-study connections in the network in terms of the ability to obtain a realistic estimate of the between study variance parameter. Results from the RE model have also been presented in full in the appendices to the report.

3.7.2.1. Comparisons versus MTX+TAC

Figure 6 presents a forest plot summarizing comparisons of all treatments in the evidence network to the chosen reference therapy, MTX+TAC. One regimen—CsA+MTX+ATG+Steroids—was found to have statistically significantly reduced risk of cGVHD relative to MTX+TAC (OR = 0.32); no regimens were associated with significantly increased risk compared to MTX+TAC. Comparisons with all regimens except CsA+MTX+ATG+Steroids were associated with credible intervals that included 1 and, thus, were not significantly different from MTX+TAC in preventing cGVHD. Many estimates were associated with wide credible intervals, as was also found in the corresponding RE analysis where findings were analogous.

Figure 6: Forest plot of comparisons versus MTX+TAC, Overall chronic GVHD in randomized patients (FE model)

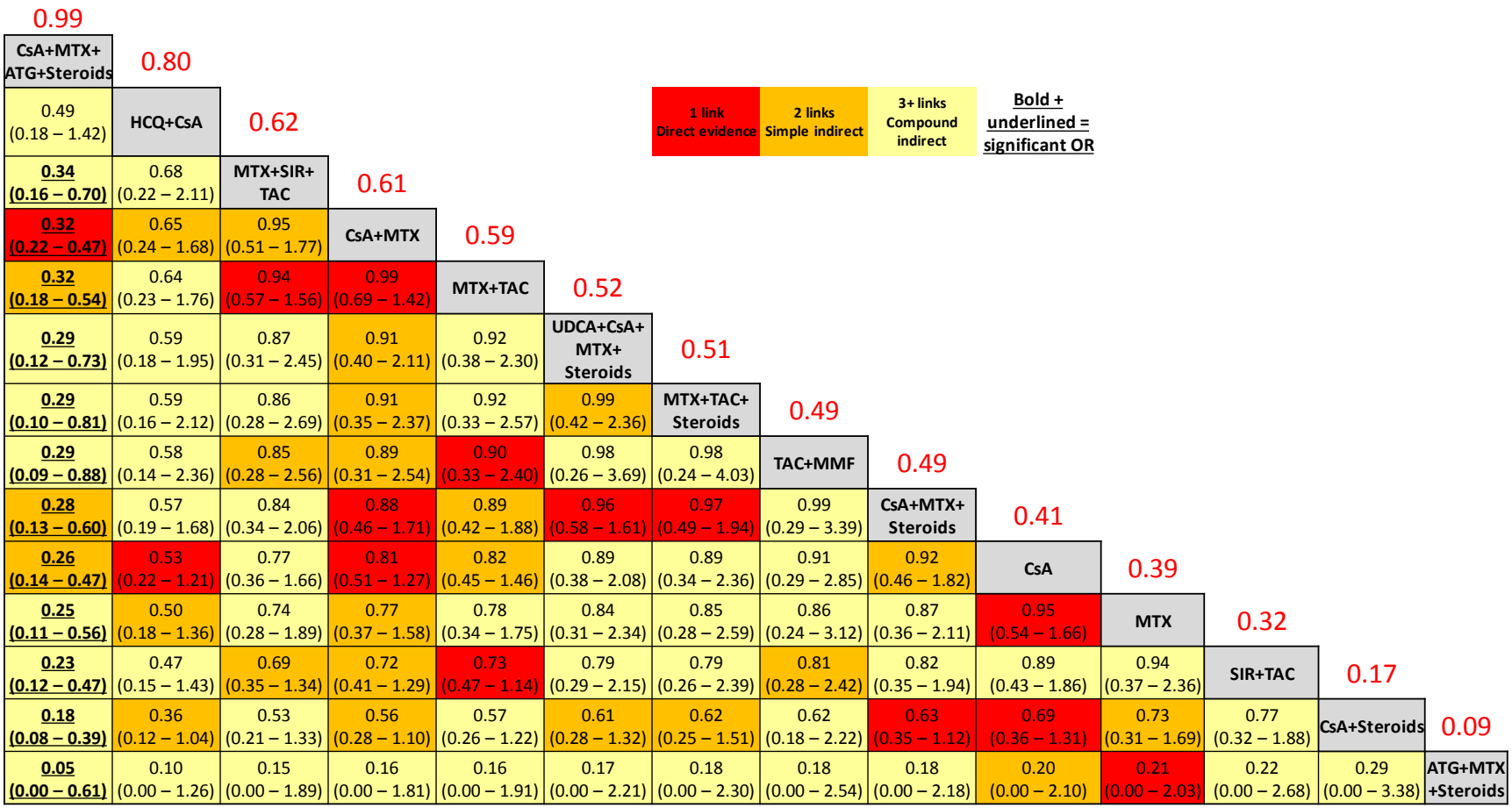


Odds ratios from network meta-analysis comparing each regimen to the reference treatment of MTX+TAC are presented along with the number of events and patients per regimen. Odds ratios <1 suggest more risk of chronic GVHD with MTX+TAC while odds ratios >1 suggest less risk of acute GVHD with MTX+TAC.

3.7.2.2. Comparisons between all prophylactic regimens

Figure 7 presents a league table of the estimates for all 91 pairwise comparisons generated from the fixed effects network meta-analysis. The regimen CsA+MTX+ATG+Steroids was found to have significantly greater efficacy to prevent cGVHD than all other regimens in the network. Some regimens were evaluated on very low numbers of patients, reducing the power to detect significant differences in efficacy relative to other regimens (e.g., ATG+MTX+Steroids). Some of the comparisons included in the league table should be interpreted carefully, specifically those which are not informed by head to head trials or where the comparison was based on compound indirect evidence. Results from the corresponding RE analysis found that while point estimates of pairwise comparisons remained relative unchanged, corresponding 95% credible intervals widened noticeably in several cases.

Figure 7: League table summary of pairwise comparisons from FE NMA, Odds Ratios with 95% CrI, Overall cGVHD in randomized patients



Interventions are sorted from left to right in order of decreasing SUCRA value (i.e., preferred treatments appear first). For each comparison shown, the upper/left-most regimen is the comparator group while the lower/right-most treatment is the reference treatment; a value <1 suggests fewer cases of chronic GVHD with the comparator than with the reference group. Statistically significant differences (i.e., estimates with a 95% credible interval excluding 1) are bolded and underlined. Comparisons with direct evidence, simple indirect evidence, and compound indirect evidence have been highlighted in red, orange, and yellow, respectively.

3.8. Findings: Overall mortality

Twenty-three studies, evaluating 3,006 patients were included in both the 100-day and the 1-year mortality networks. **Figure 8** presents a network diagram highlighting the extent of available evidence for the comparison of prophylactic regimens for GVHD, where 100-day and 1-year mortality were evaluated as outcomes. Data from head-to-head trials were available for 14 of the 105 potential pairwise comparisons in the network, with single studies informing 8 of the 14 head-to-head comparisons.

Figure 8: Network diagram, Overall mortality (15 interventions, 3,006 patients)

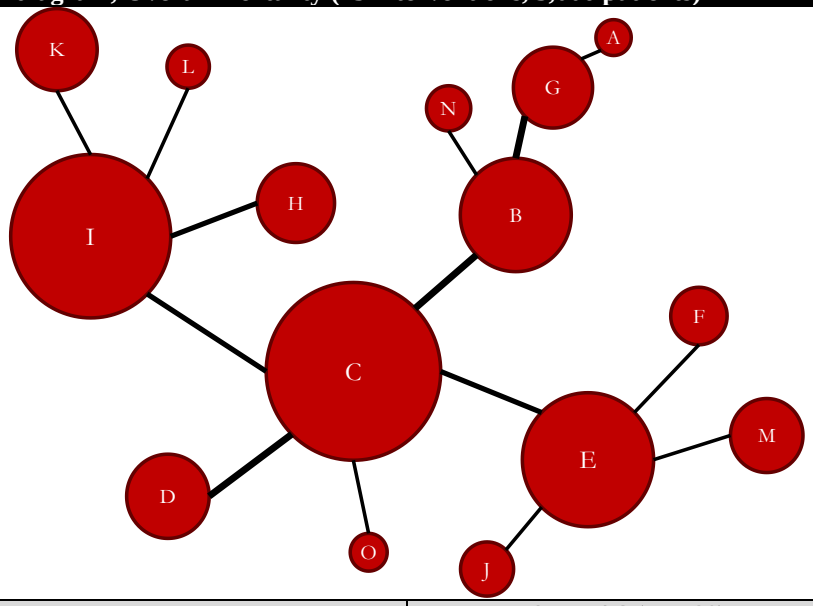


Figure legend	
A = ATG+MTX+Steroids (n = 29)	H = MTX+SIR+TAC (n = 139)
B = CsA (n = 288)	I = MTX+TAC (n = 598)
C = CsA+MTX (n = 710)	J = MTX+TAC+Steroids (n = 66)
D = CsA+MTX+ATG+Steroids (n = 159)	K = SIR+TAC (n = 151)
E = CsA+MTX+Steroids (n = 402)	L = TAC+MMF (n = 42)
F = CsA+Steroids (n = 74)	M = UDCA+CsA+MTX+Steroids (n = 123)
G = MTX (n = 147)	N = HCQ+CsA (n = 46)
	O = CsA+MTX+MSCs (n = 32)

3.8.1. Results from traditional pairwise meta-analyses

3.8.1.1. 100-day mortality

Table 8 presents a summary of pairwise estimates that were derived from direct evidence (i.e., the head-to-head trials), summarized alongside the numbers of available trials and patients for each direct comparison, as well as the related summary estimate derived from the network meta-analysis. The direct comparisons with the largest numbers of studies available were CsA+MTX+ATG+Steroids vs CsA+MTX (3 RCTs, n = 310), CsA vs CsA+MTX (3 RCTs, n = 249), and CsA vs MTX (3 RCTs, n = 230). None of the comparisons demonstrated statistically significant differences between therapies in the risk of mortality in pairwise meta-analyses. Estimates from traditional pairwise meta-analyses compared well with estimates derived from network meta-analysis.

Table 8: Summary of results from pairwise meta-analysis and FE NMA, 100-day overall mortality				
Comparison		**# of Trials (patients)**	**Direct Estimate OR (95% CrI)**	**NMA Estimate OR (95% CI)**
			**Values <1 favor comparator*	
Reference	**Comparator**			
MTX+TAC	SIR+TAC	1 (304)	0.94 (0.44 – 2.02)	0.94 (0.43 – 2.02)
CsA+MTX+Steroids	UDCA+CsA+MTX+Steroids	1 (242)	0.49 (0.24 – 1.03)	0.48 (0.23 – 1.01)
CsA+MTX	CsA+MTX+ATG+Steroids	3 (310)	0.87 (0.48 – 1.56)	0.87 (0.48 – 1.56)
MTX+SIR+TAC	MTX+TAC	2 (282)	0.56 (0.21 – 1.50)	0.55 (0.20 – 1.43)
CSA+MTX	CSA+MTX+MSCs	1 (66)	0.18 (0.01 – 3.83)	0.29 (0.01 – 3.13)
CsA+MTX+Steroids	MTX+TAC+Steroids	3 (230)	0.98 (0.38 – 2.55)	0.98 (0.36 – 2.58)
CsA	MTX	3 (249)	0.55 (0.34 – 1.25)	0.53 (0.23 – 1.20)
CsA	CsA+MTX	2 (509)	0.60 (0.31 – 1.18)	0.62 (0.31 – 1.20)
MTX+TAC	CsA+MTX	1 (89)	0.80 (0.54 – 1.20)	0.80 (0.54 – 1.19)
MTX+TAC	TAC+MMF	1 (149)	0.92 (0.26 – 3.28)	0.92 (0.24 – 3.45)
CsA+Steroids	CsA+MTX+Steroids	2 (294)	0.63 (0.21 – 1.86)	0.62 (0.19 – 1.83)
CsA+MTX+Steroids	CsA+MTX	1 (56)	0.94 (0.43 – 2.05)	0.93 (0.43 – 2.06)
MTX	ATG+MTX+Steroids	1 (95)	0.38 (0.12 – 1.24)	0.36 (0.10 – 1.15)
CsA	HCQ+CsA	1 (66)	0.82 (0.29 – 2.30)	0.82 (0.28 – 2.36)

3.8.1.2. 1-year mortality

Table 9 presents a summary of pairwise estimates that were derived from direct evidence (i.e., the head-to-head trials), summarized alongside the numbers of available trials and patients for each direct comparison, as well as the related summary estimate derived from the network meta-analysis. The direct comparisons with the largest numbers of studies available were CsA+MTX+ATG+Steroids vs CsA+MTX (3 RCTs, n = 310), CsA vs CsA+MTX (3 RCTs, n = 249), and CsA vs MTX (3 RCTs, n = 230). Pairwise meta-analysis suggested a significantly reduced risk of mortality with UDCA+CsA+MTX+Steroids compared to CsA+MTX+Steroids, and with CsA+MTX+MSCs when compared to CsA+MTX. Estimates from traditional pairwise meta-analyses again compared well with estimates derived from network meta-analysis.

Table 9: Summary of results from pairwise meta-analysis and FE NMA, 1-year overall mortality				
Comparison		**# of Trials (patients)**	**Direct Estimate OR (95% CrI)**	**NMA Estimate OR (95% CI)**
			**Values <1 favor comparator*	
Reference	**Comparator**			
SIR+TAC	MTX+TAC	1 (304)	0.86 (0.52 – 1.42)	0.86 (0.52 – 1.42)
CsA+MTX+Steroids	UDCA+CsA+MTX+Steroids	1 (242)	**0.53 (0.31 – 0.91)**	**0.53 (0.31 – 0.91)**
CsA+MTX	CsA+MTX+ATG+Steroids	3 (310)	0.93 (0.58 – 1.48)	0.93 (0.58 – 1.49)
MTX+SIR+TAC	MTX+TAC	2 (282)	0.73 (0.42 – 1.28)	0.73 (0.41 – 1.29)
CSA+MTX	CSA+MTX+MSCs	1 (66)	**0.09 (0.01 – 0.78)**	**0.06 (<0.01 – 0.46)**
CsA+MTX+Steroids	MTX+TAC+Steroids	1 (131)	0.84 (0.40 – 1.80)	0.84 (0.39 – 1.79)
CsA	MTX	3 (230)	0.64 (0.37 – 1.12)	0.65 (0.37 – 1.13)
CsA	CsA+MTX	3 (249)	0.67 (0.37 – 1.23)	0.76 (0.43 – 1.33)
MTX+TAC	CsA+MTX	2 (509)	0.82 (0.58 – 1.16)	0.82 (0.57 – 1.17)
MTX+TAC	TAC+MMF	1 (89)	0.82 (0.35 – 1.93)	0.81 (0.34 – 1.94)
CsA+Steroids	CsA+MTX+Steroids	1 (149)	0.85 (0.40 – 1.80)	0.85 (0.39 – 1.82)
CsA+MTX+Steroids	CsA+MTX	2 (294)	0.82 (0.42 – 1.58)	0.90 (0.53 – 1.52)
MTX	ATG+MTX+Steroids	1 (56)	0.83 (0.28 – 2.44)	0.89 (0.31 – 2.51)
HCQ+CsA	CsA	1 (95)	0.58 (0.26 – 1.30)	0.57 (0.25 – 1.30)

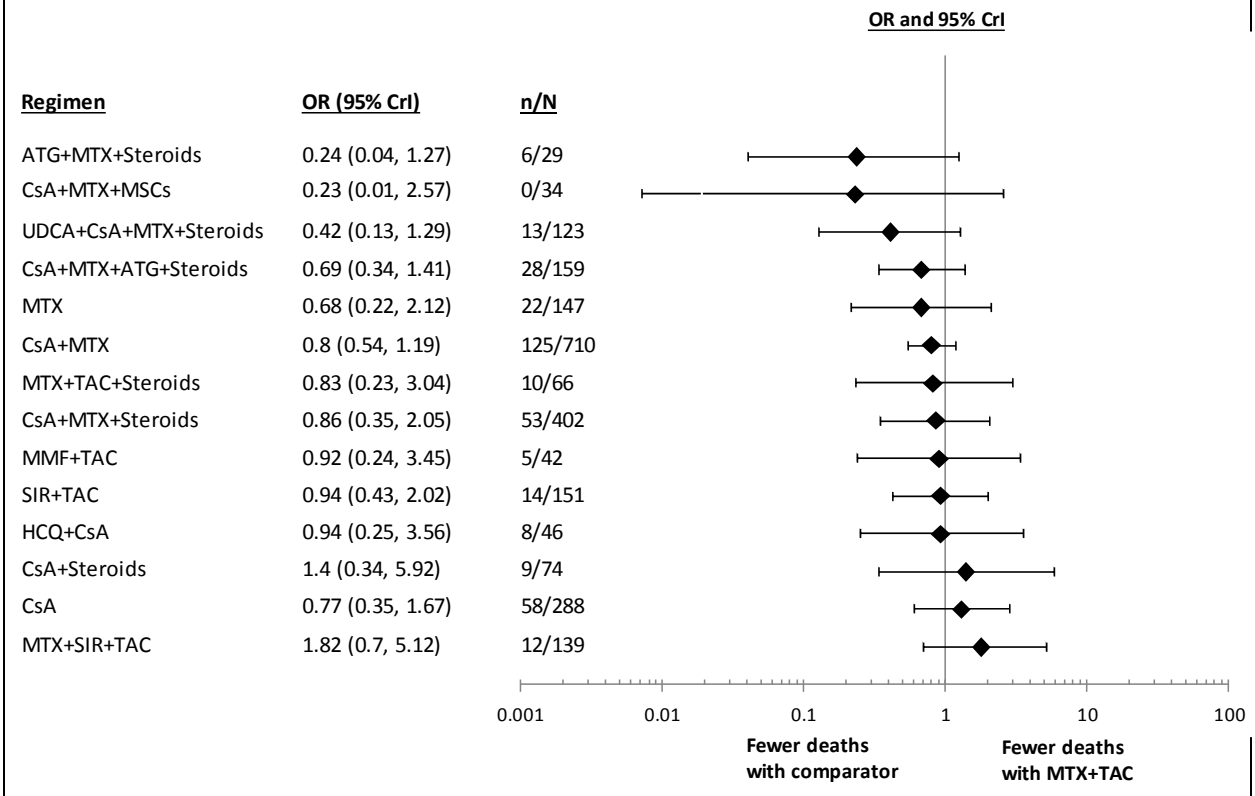
3.8.2. Results from network meta-analysis

For both 100-day and 1-year mortality, most comparisons (i.e., 91 of 105) were informed only by indirect evidence and many of the comparisons with direct evidence were informed by single studies with limited numbers of patients. Model fit statistics for NMAs for both 100-day and 1-year mortality indicated that both the fixed-effects (FE) and random-effects (RE) models had an adequate fit. Posterior residual deviance values for the FE and RE models of 45.75 and 44.75, respectively, for 100-day mortality and of 55.39 and 48.73, respectively, for 1-year mortality were obtained, with the 46 data points in the models. DIC values (251.279 versus 252.823 for 100-day mortality; 288.402 versus 287.395 for 1-year mortality) suggested no difference in fit between FE and RE models; however, a FE model was considered preferable due to concerns regarding adequate estimation of the between study variance parameter, as mentioned earlier.

3.8.2.1. 100-day mortality: Comparisons versus MTX+TAC

Figure 9 presents a forest plot summarizing comparisons of all treatments in the evidence network to the chosen reference therapy, MTX+TAC. All of the interventions were associated with credible intervals that included 1 and, thus, were not significantly different from MTX+TAC in reducing 100-day overall mortality. This was also the case when analyses were run using an RE model for NMA.

Figure 9: Forest plot of comparisons versus MTX+TAC, 100-day overall mortality (FE model)



Odds ratios from network meta-analysis comparing each regimen to the reference treatment of MTX+TAC are presented along with the number of events and patients per regimen. Odds ratios <1 suggest more risk of 100-day death with MTX+TAC while odds ratios >1 suggest less risk of 100-day death with MTX+TAC.

3.8.2.2. 100-day mortality: Comparisons between all prophylactic regimens

Figure 10 presents a league table of the estimates for all 105 pairwise comparisons generated from the network meta-analysis. Some of the comparisons included in the league table should be interpreted carefully, specifically those which are not informed by head to head trials or where the comparison was based on compound indirect evidence. Only two comparisons (ATG+MTX+Steroids vs. CsA and ATG+MTX+Steroids vs. MTX+SIR+TAC) were associated with statistically significant differences in overall mortality at 100 days, with ATG+MTX+Steroids being associated with improved efficacy in both comparisons. The majority of other comparisons were associated with considerable uncertainty. The corresponding NMA using an RE model produced findings which showed summary estimates with even greater uncertainty, and no differences in any comparison of therapies achieved statistical significance.

Figure 10: League table summary of pairwise comparisons from FE NMA, Odds Ratios with 95% CrI, 100-day overall mortality

ATG+MTX+ Steroids 0.89													
1.07 (0.06 – 46.49)	**CsA+MTX+ MSCs** 0.82												
0.58 (0.08 – 4.01)	0.55 (0.01 – 7.84)	**UDCA+CsA+ MTX+/- Steroids** 0.81											
0.35 (0.06 – 1.89)	0.33 (0.01 – 3.97)	0.60 (0.17 – 2.02)	**CsA+MTX+ ATG+Steroids** 0.61										
0.36 (0.10 – 1.15)	0.33 (0.01 – 4.66)	0.61 (0.13 – 2.78)	1.01 (0.30 – 3.49)	**MTX** 0.59									
0.30 (0.05 – 1.50)	0.29 (0.01 – 3.13)	0.52 (0.17 – 1.50)	0.87 (0.48 – 1.56)	0.86 (0.29 – 2.46)	**CsA+MTX** 0.54								
0.29 (0.04 – 2.18)	0.27 (0.01 – 4.28)	0.50 (0.15 – 1.68)	0.83 (0.22 – 3.30)	0.82 (0.16 – 4.14)	0.96 (0.28 – 3.34)	**MTX+TAC+ Steroids** 0.49							
0.28 (0.04 – 1.68)	0.27 (0.01 – 3.46)	0.48 (0.23 – 1.01)	0.81 (0.31 – 2.20)	0.80 (0.21 – 3.01)	0.93 (0.43 – 2.06)	0.98 (0.36 – 2.58)	**CsA+MTX+ Steroids** 0.48						
0.27 (0.03 – 2.23)	0.25 (0.01 – 4.07)	0.45 (0.08 – 2.61)	0.76 (0.17 – 3.51)	0.75 (0.13 – 4.36)	0.87 (0.22 – 3.52)	0.92 (0.15 – 5.70)	0.94 (0.19 – 4.50)	**TAC+MMF** 0.45					
0.26 (0.04 – 1.58)	0.25 (0.01 – 3.17)	0.44 (0.11 – 1.76)	0.74 (0.26 – 2.12)	0.73 (0.18 – 2.86)	0.85 (0.36 – 2.05)	0.90 (0.20 – 3.93)	0.91 (0.28 – 2.92)	0.98 (0.20 – 4.49)	**SIR+TAC** 0.43				
0.23 (0.03 – 1.37)	0.22 (0.01 – 3.27)	0.39 (0.07 – 2.07)	0.65 (0.16 – 2.66)	0.64 (0.17 – 2.51)	0.75 (0.22 – 2.69)	0.80 (0.14 – 4.54)	0.80 (0.18 – 3.59)	0.86 (0.13 – 5.71)	0.88 (0.19 – 4.04)	**HCQ+CsA** 0.38			
0.24 (0.04 – 1.27)	0.23 (0.01 – 2.57)	0.42 (0.13 – 1.29)	0.69 (0.34 – 1.41)	0.68 (0.22 – 2.12)	0.80 (0.54 – 1.19)	0.83 (0.23 – 3.04)	0.86 (0.35 – 2.05)	0.92 (0.24 – 3.45)	0.94 (0.43 – 2.02)	1.06 (0.28 – 3.95)	**MTX+TAC** 0.26		
0.17 (0.02 – 1.41)	0.16 (0.00 – 2.68)	0.30 (0.07 – 1.11)	0.50 (0.11 – 2.21)	0.49 (0.08 – 2.73)	0.57 (0.14 – 2.22)	0.60 (0.13 – 2.55)	0.62 (0.19 – 1.83)	0.65 (0.09 – 4.45)	0.67 (0.13 – 3.33)	0.76 (0.11 – 4.94)	0.71 (0.17 – 2.93)	**CsA+ Steroids** 0.24	
0.19 (0.04 – 0.78)	0.18 (0.01 – 2.13)	0.32 (0.09 – 1.14)	0.53 (0.22 – 1.31)	0.53 (0.23 – 1.20)	0.62 (0.31 – 1.20)	0.64 (0.16 – 2.62)	0.66 (0.23 – 1.84)	0.70 (0.15 – 3.29)	0.72 (0.24 – 2.18)	0.82 (0.28 – 2.36)	0.77 (0.35 – 1.67)	1.09 (0.23 – 5.08)	**CsA** 0.14
0.13 (0.02 – 0.90)	0.12 (0.00 – 1.73)	0.23 (0.05 – 1.00)	0.38 (0.11 – 1.27)	0.37 (0.08 – 1.64)	0.44 (0.15 – 1.24)	0.46 (0.09 – 2.26)	0.47 (0.12 – 1.69)	0.50 (0.09 – 2.54)	0.51 (0.14 – 1.76)	0.58 (0.11 – 2.95)	0.55 (0.20 – 1.43)	0.77 (0.13 – 4.28)	0.71 (0.20 – 2.47)	**MTX+ SIR+TAC**

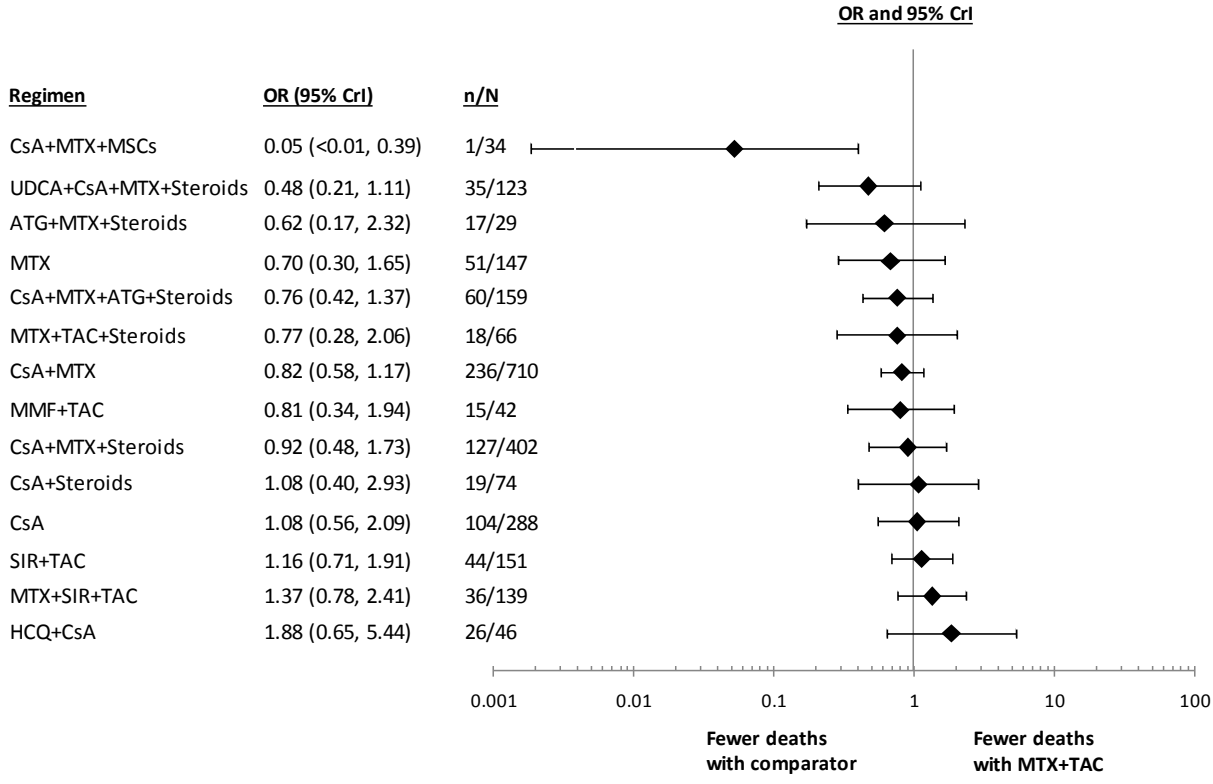
1 link Direct evidence | 2 links Simple indirect | 3+ links Compound indirect | **Bold and underlined = significant OR**

Interventions are sorted from left to right in order of decreasing SUCRA value (i.e., preferred treatments appear first). For each comparison shown, the upper/left-most regimen is the comparator group while the lower/right-most treatment is the reference treatment; a value <1 suggests fewer deaths overall at 100 days with the comparator than with the reference group. Statistically significant differences (i.e., estimates with a 95% credible interval excluding 1) are bolded and underlined. Comparisons with direct evidence, simple indirect evidence, and compound indirect evidence have been highlighted in red, orange, and yellow, respectively.

3.8.2.3. 1-year mortality: Comparisons versus MTX+TAC

Figure 11 presents a forest plot summarizing comparisons of all treatments in the evidence network to the chosen reference therapy, MTX+TAC. One regimen—CsA+MTX+MSCs—was found to have statistically significantly reduced risk of mortality at 1 year relative to MTX+TAC (OR = 0.05; 95% CrI = <0.01–0.39). Comparisons involving all other regimens were associated with credible intervals that included 1 and, thus, were not significantly different from MTX+TAC in reducing 1-year overall mortality. Analogous interpretations were drawn from the RE analysis reported in the appendices.

Figure 11: Forest plot of comparisons versus MTX+TAC, 1-year overall mortality (FE model)

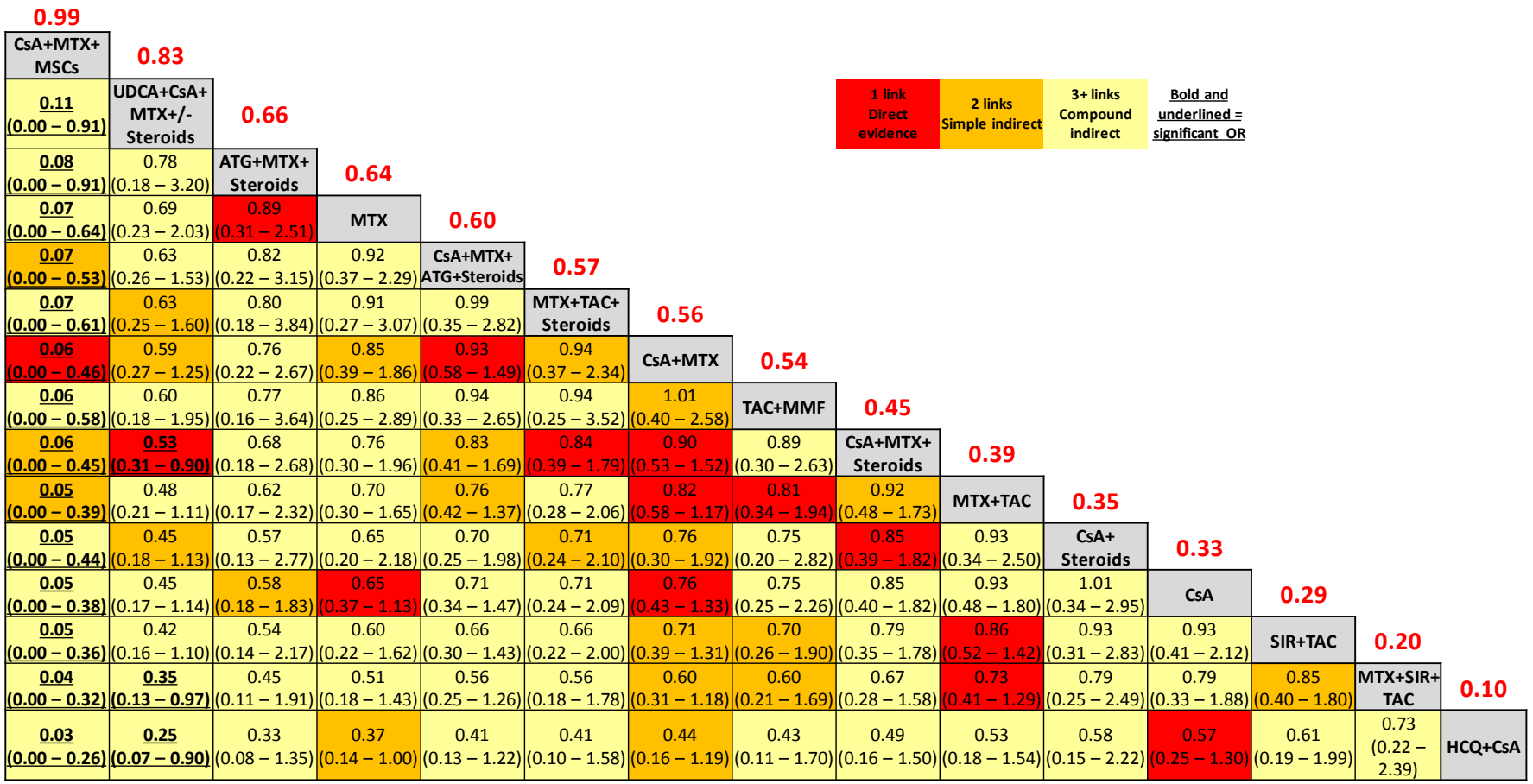


Odds ratios from network meta-analysis comparing each regimen to the reference treatment of MTX+TAC are presented along with the number of events and patients per regimen. Odds ratios <1 suggest more risk of 100-day death with MTX+TAC while odds ratios >1 suggest less risk of 100-day death with MTX+TAC.

3.8.2.4. 1-year mortality: Comparisons Between all Prophylactic Regimens

Figure 12 presents a league table of the estimates for all 105 pairwise comparisons generated from the FE network meta-analysis. Some of the comparisons included in the league table should be interpreted carefully, specifically those which are not informed by head to head trials or where the comparison was based on compound indirect evidence. CsA+MTX+MSCs was found to have statistically significantly reduced risk of mortality at 1 year relative to all other interventions included in the network; all other pairwise comparisons between interventions were associated with 95% credible intervals that included 1. In the related RE analysis, uncertainty of all comparisons increased, and significant advantages of CsA+MTX+MSCs over some regimens remained while in some cases the credible interval widened to the extent of including the value of 1.

Figure 12: League table summary of pairwise comparisons, 1-year overall mortality



0.99														
CsA+MTX+ MSCs	0.83													
0.11 (0.00 – 0.91)	UDCA+CsA+ MTX+/- Steroids	0.66												
0.08 (0.00 – 0.91)	0.78 (0.18 – 3.20)	ATG+MTX+ Steroids	0.64											
0.07 (0.00 – 0.64)	0.69 (0.23 – 2.03)	0.89 (0.31 – 2.51)	MTX	0.60										
0.07 (0.00 – 0.53)	0.63 (0.26 – 1.53)	0.82 (0.22 – 3.15)	0.92 (0.37 – 2.29)	CsA+MTX+ ATG+Steroids	0.57									
0.07 (0.00 – 0.61)	0.63 (0.25 – 1.60)	0.80 (0.18 – 3.84)	0.91 (0.27 – 3.07)	0.99 (0.35 – 2.82)	MTX+TAC+ Steroids	0.56								
0.06 (0.00 – 0.46)	0.59 (0.27 – 1.25)	0.76 (0.22 – 2.67)	0.85 (0.39 – 1.86)	0.93 (0.58 – 1.49)	0.94 (0.37 – 2.34)	CsA+MTX	0.54							
0.06 (0.00 – 0.58)	0.60 (0.18 – 1.95)	0.77 (0.16 – 3.64)	0.86 (0.25 – 2.89)	0.94 (0.33 – 2.65)	0.94 (0.25 – 3.52)	1.01 (0.40 – 2.58)	TAC+MMF	0.45						
0.06 (0.00 – 0.45)	**0.53 (0.31 – 0.90)**	0.68 (0.18 – 2.68)	0.76 (0.30 – 1.96)	0.83 (0.41 – 1.69)	0.84 (0.39 – 1.79)	0.90 (0.53 – 1.52)	0.89 (0.30 – 2.63)	CsA+MTX+ Steroids	0.39					
0.05 (0.00 – 0.39)	0.48 (0.21 – 1.11)	0.62 (0.17 – 2.32)	0.70 (0.30 – 1.65)	0.76 (0.42 – 1.37)	0.77 (0.28 – 2.06)	0.82 (0.58 – 1.17)	0.81 (0.34 – 1.94)	0.92 (0.48 – 1.73)	MTX+TAC	0.35				
0.05 (0.00 – 0.44)	0.45 (0.18 – 1.13)	0.57 (0.13 – 2.77)	0.65 (0.20 – 2.18)	0.70 (0.25 – 1.98)	0.71 (0.24 – 2.10)	0.76 (0.30 – 1.92)	0.75 (0.20 – 2.82)	0.85 (0.39 – 1.82)	0.93 (0.34 – 2.50)	CsA+ Steroids	0.33			
0.05 (0.00 – 0.38)	0.45 (0.17 – 1.14)	0.58 (0.18 – 1.83)	0.65 (0.37 – 1.13)	0.71 (0.34 – 1.47)	0.71 (0.24 – 2.09)	0.76 (0.43 – 1.33)	0.75 (0.25 – 2.26)	0.85 (0.40 – 1.82)	0.93 (0.48 – 1.80)	1.01 (0.34 – 2.95)	CsA	0.29		
0.05 (0.00 – 0.36)	0.42 (0.16 – 1.10)	0.54 (0.14 – 2.17)	0.60 (0.22 – 1.62)	0.66 (0.30 – 1.43)	0.66 (0.22 – 2.00)	0.71 (0.39 – 1.31)	0.70 (0.26 – 1.90)	0.79 (0.35 – 1.78)	0.86 (0.52 – 1.42)	0.93 (0.31 – 2.83)	0.93 (0.41 – 2.12)	SIR+TAC	0.20	
0.04 (0.00 – 0.32)	**0.35 (0.13 – 0.97)**	0.45 (0.11 – 1.91)	0.51 (0.18 – 1.43)	0.56 (0.25 – 1.26)	0.56 (0.18 – 1.78)	0.60 (0.31 – 1.18)	0.60 (0.21 – 1.69)	0.67 (0.28 – 1.58)	0.73 (0.41 – 1.29)	0.79 (0.25 – 2.49)	0.79 (0.33 – 1.88)	0.85 (0.40 – 1.80)	MTX+SIR+ TAC	0.10
0.03 (0.00 – 0.26)	**0.25 (0.07 – 0.90)**	0.33 (0.08 – 1.35)	0.37 (0.14 – 1.00)	0.41 (0.13 – 1.22)	0.41 (0.10 – 1.58)	0.44 (0.16 – 1.19)	0.43 (0.11 – 1.70)	0.49 (0.16 – 1.50)	0.53 (0.18 – 1.54)	0.58 (0.15 – 2.22)	0.57 (0.25 – 1.30)	0.61 (0.19 – 1.99)	0.73 (0.22 – 2.39)	HCQ+CsA

Interventions are sorted from left to right in order of decreasing SUCRA value (i.e., preferred treatments appear first). For each comparison shown, the upper/left-most regimen is the comparator group while the lower/right-most treatment is the reference treatment; a value <1 suggests fewer deaths overall at 1 year with the comparator than with the reference group. Statistically significant differences (i.e., estimates with a 95% credible interval excluding 1) are bolded and underlined. Comparisons with direct evidence, simple indirect evidence, and compound indirect evidence have been highlighted in red, orange, and yellow, respectively.

3.9. Overall mortality—Other follow-up times

Several studies reported overall mortality data at other follow-up times and have been summarized narratively and presented in **Table 10**. Only one comparison demonstrated a statistically significant difference in overall mortality. At 1.5 years post-transplant, CsA+MTX was associated with significantly less mortality than CsA; however, at longer follow-up (3.5–4 years), the difference was no longer significant.[73]

Table 10: Summary of results: Risk of overall mortailty in the evaluation of GVHD prophylaxis at any follow-up time other than 100 days and 1 year post-transplant

Author	Follow-up	Treatment regimen	Group risk	OR (95% CI)
Kuzmina[40] (2015)	0.1–6.1 years	CsA+MTX+MSCs	6/32 (19%)	0.37 (0.12–1.15)
		CsA+MTX	13/34 (38%)*	
Cutler[43] (2014)	2 years	SIR+TAC	62/151 (41%)*	0.85 (0.54–1.35)
		MTX+TAC	57/153 (37%)	
Bonifazi[45] (2014)	2 years	CSA+MTX	21/83 (25%)*	0.78 (0.37–1.65)
		CSA+MTX+ATG+Steroids	15/72 (21%)	
Ruutu[44] (2014)	10 years	UDCA+CsA+MTX+Steroids	64/123 (52%)	0.66 (0.40–1.10)
		CsA+MTX	74/119 (62%)*	
Armand (2013)	2 years	MTX+SIR+TAC	22/66 (33%)*	0.81 (0.39–1.66)
		MTX+TAC	21/73 (29%)	
Martin[39] (2012)	200 days	MTX+TAC+Steroids	22/92 (24%)	1.00 (0.44–2.29)
		MTX+TAC	11/46 (24%)*	
Pidala[47] (2012)	2 years	MTX+TAC	11/37 (31%)	0.70 (0.26–1.83)
		SIR+TAC	14/37 (39%)*	
Socie[50] (2011)	3 years	ATG+CsA+MTX	49/103 (48%)	0.74 (0.42–1.29)
		CsA+MTX	54/98 (55%)*	
Perkins[55] (2010)	3 years	MTX+TAC	27/47 (58%)*	0.61 (0.26–1.42)
		MMF+TAC	19/42 (46%)	
Fong[56] (2007)	2.9 years	HCQ+CsA	34/46 (74%)*	0.98 (0.39–2.44)
		CsA	36/49 (74%)	
Lee[57] (2004)	1.8 (0.4–4.0) years	CsA	5/40 (13%)	0.38 (0.12–1.21)
		CsA+MTX	11//40 (28%)*	
Bacigalupo[59] (2001)	3 years	CsA+MTX+ATG+Steroids	13/29 (45%)*	0.97 (0.33–2.84)
		CsA+MTX	11/25 (44%)	
Deeg[84] (2000)	3.4 (1.4–4.9) years	CsA	43/60 (72%)*	0.89 (0.41–1.95)
		CsA+Steroids	43/62 (69%)	
Deeg[84] (2000)	6.1 (3.6–8.0) years	CsA	46/60 (77%)*	0.88 (0.38–2.00)
		CsA+Steroids	46/62 (74%)	
Nash[62] (2000)	2 years minimum	CsA+MTX	45/90 (50%)*	0.84 (0.47–1.50)
		MTX+TAC	41/90 (46%)	
Ruutu[63] (2000)	6 (4.2–9.1) years	CsA+MTX	27/55 (49%)*	68 (0.32–1.46)
		CsA+MTX+Steroids	21/53 (40%)	

Table 10: Summary of results: Risk of overall mortailty in the evaluation of GVHD prophylaxis at any follow-up time other than 100 days and 1 year post-transplant

Author	Follow-up	Treatment regimen	Group risk	OR (95% CI)
Ratanatharathorn[65] (1998)	2 years	CSA+MTX	71/164 (43%)	0.65 (0.42–1.01)
		TAC+MTX	89/165 (54%)*	
Zikos[67] (1998)	5.8 (2.5–6.9) years	CsA	9/28 (32%)*	0.70 (0.23–2.17)
		CsA+MTX	8/32 (25%)	
Chao[64] (1993)	5.3 years	CsA+Steroids	34/74 (46%)*	0.66 (0.34–1.28)
		CsA+MTX+Steroids	27/75 (36%)	
Mrsic[69] (1990)	3 years	CsA	23/39 (58%)*	0.42 (0.17–1.06)
		CsA+MTX	14/37 (37%)	
Torres[70] (1989)	1.2 (0.1–6.3) vs. 1.7 (0.2–3.6) years	CsA	11/26 (42%)*	0.98 (0.34–2.83)
		MTX	13/31 (42%)	
Backman[71] (1988)	2 years	CsA	13/30 (43%)*	0.69 (0.24–1.97)
		MTX	10/29 (35%)	
Backman[71] (1988)	4.1 (2.7–5.8) years	CsA	14/30 (47%)*	0.81 (0.29–2.26)
		MTX	12/29 (41%)	
Storb[73] (1986)	1.5 years	CsA	23/50 (45%)*	**0.31 (0.12–0.78)**
		CsA+MTX	9/43 (20%)	
Storb[73] (1986)	3–4.5 years	CsA	24/50 (48%)*	0.58 (0.25–1.34)
		CsA+MTX	15/43 (35%)	
Irle[77] (1985)	2.3 years maximum	CsA	21/26 (81%)	0.84 (0.21–3.30)
		MTX	25/30 (83%)*	
Storb[78] (1983)	0.2–2.5 years	CsA	10/36 (28%)	0.57 (0.22–1.47)
		MTX	17/42 (40%)*	
Storb[78] (1983)	2 years	CsA	10/36 (29%)	0.47 (0.18–1.20)
		MTX	19/42 (45%)*	
Storb[78] (1983)	2.9 (1.6–3.9) years	CsA	14/36 (39%)	0.74 (0.30–1.86)
		MTX	18/39 (46%)*	
Weiden[81] (1979)	≤2 years	ATG+MTX	16/29 (55%)	0.62 (0.21–1.82)
		MTX	18/27 (67%)*	

*denotes the reference group for each pairwise comparison
ATG = anti-thymocyte globulin; CsA = cyclosporin A; MMF = mycophenolate mofetil; MSCs = mesenchymal stem cells; MTX = methotrexate; SIR = sirolimus; TAC = tacrolimus; UDCA = ursodeoxycholic acid

3.10. Non-relapse mortality

Many studies reported data for the risk of non-relapse mortality (NRM); however, the competing risk of relapse and lack of reporting of patient time at risk prevented NMA. These studies have been summarized narratively and presented in **Table 11** according to follow-up time. Three studies identified significant differences in non-relapse mortality between treatment groups. Pidala et al[47] found no significant difference between MTX+TAC and SIR+TAC after a median of approximately 1.5 years of follow-up; however, at 2 years of follow-up, the odds of NRM were significantly lower for patients in the MTX+TAC group than

for those in the SIR+TAC group. Ruutu et al[44] (2014) demonstrated a significantly lower odds of NRM in the UDCA+CsA+MTX+Steroids group compared to the CsA+MTX+Steroids group at both 1 and 10 years of follow-up. Storb et al[73] (1986) found at long-term follow-up at 20 years post-transplant there were significantly reduced odds of NRM in the CsA+MTX group compared to the CsA group; however, at earlier time points, there were no significant differences.

Table 11 - Findings: non-relapse mortality in the evaluation of GVHD prophylaxis at any follow-up time

Author	Follow-up	Treatment regimen	Group risk	OR (95% CI)
Socie[50] (2011)	100 days	CsA+MTX+ATG	11/103 (11%)	0.78 (0.33–1.84)
		CsA+MTX	13/98 (13%)*	
Cutler[43] (2014)	100 days	SIR+TAC	11/151 (7%)*	0.99 (0.41–2.35)
		MTX+TAC	11/153 (7%)	
Bacigalupo[59] (2001)	1 year	CsA+MTX+ATG+Steroids	13/27 (47%)	0.93 (0.32–2.67)
		CsA+MTX	14/28 (49%)*	
Ruutu ([44]2014)	1 year	UDCA+CsA+MTX+Steroids	23/123 (19%)	**0.45 (0.25–0.82)**
		CsA+MTX+Steroids	40/119 (34%)*	
Glass[37] (2014)	1 year	MMF+TAC	16/42 (37%)*	0.81 (0.33–1.99)
		RITX+TAC+MMF	14/42 (34%)	
Perkins[55] (2010)	1 year	MTX+TAC	12/47 (26%)	0.76 (0.30–1.93)
		TAC+MMF	13/42 (31%)*	
Storb[73] (1986)	1.15 (0.36–1.98) vs. 1.23 (0.3–1.95) years	CsA	18/50 (36%)*	0.38 (0.14–1.02)
		CsA+MTX	7/43 (16%)	
Pidala[47] (2012)	1.4 (0.3–2.7) vs 1.7 (0.3–2.7)	MTX+TAC	2/37 (5%)	0.21 (0.04–1.05)
		SIR+TAC	8/37 (22%)*	
Lee[57] (2004)	1.8 (0.4–4.0)	CsA	4/40 (10%)	1.00 (0.23–4.31)
		CsA+MTX	4/40 (10%)*	
Torres[70] (1989)	1.7 (0.2–3.6) vs. 1.2 (0.1–6.3) years	CsA	5/26 (19%)	0.82 (0.23–2.96)
		MTX	7/31 (23%)*	
Weiden[81] (1979)	≤2 years	MTX+ATG	8/29 (28%)	0.55 (0.18–1.70)
		MTX	11/27 (41%)*	
Ratanatharathorn[65] (1998)	2 years	CsA+MTX	47/164 (29%)	0.72 (0.45–1.15)
		MTX+TAC	59/165 (36%)*	
Lee[57] (2004)	2 years	CsA	5/40 (13%)*	0.78 (0.19–3.14)
		CsA+MTX	4/40 (11%)	
Pidala[47] (2012)	2 years	MTX+TAC	3/37 (8%)	**0.34 (0.06–0.95)**
		SIR+TAC	10/37 (28%)*	
Armand (2013)	2 years	MTX+SIR+TAC	9/66 (14%)	0.99 (0.38–2.62)
		MTX+TAC	10/73 (14%)*	
Nash[62] (2000)	2 years minimum	CsA+MTX	38/90 (42%)*	0.68 (0.37–1.25)
		MTX+TAC	30/90 (33%)	
Irle[77] (1985)	2.3 years	CsA	10/26 (38%)	0.42 (0.14–1.22)
		MTX	18/30 (60%)*	

Table 11 - Findings: non-relapse mortality in the evaluation of GVHD prophylaxis at any follow-up time				
Author	**Follow-up**	**Treatment regimen**	**Group risk**	**OR (95% CI)**
Bacigalupo[59] (2001)	2.3 vs. 2.8 years	CsA+MTX+ATG+Steroids	12/29 (42%)*	0.94 (0.32–2.81)
		CsA+MTX	10/25 (40%)	
Fong[56] (2007)	2.9 years	HCQ+CsA	21/46 (46%)*	0.82 (0.36–1.85)
		CsA	20/49 (41%)	
Perkins[55] (2010)	3 years	MTX+TAC	14/47 (30%)	0.95 (0.38–2.34)
		TAC+MMF	13/42 (31%)*	
Socie[50] (2011)	3 years	CsA+MTX+ATG	21/103 (20%)	0.55 (0.29–1.05)
		CsA+MTX	31/98 (32%)*	
Deeg[84] (2000)	3.4 (1.4–4.9) years	CsA	31/60 (52%)*	0.94 (0.46–1.90)
		CsA+Steroids	31/62 (50%)	
Storb[73] (1986)	3–4.5 years	CsA	20/50 (40%)*	0.45 (0.18–1.12)
		CsA+MTX	10/43 (23%)	
Pidala[47] (2012)	4 years	MTX+TAC	4/26 (16%)*	0.57 (0.12–2.81)
		SIR+TAC	3/32 (10%)	
Glass[37] (2014)	4 years	MMF+TAC	16/42 (38%)	0.90 (0.38–2.17)
		RITX+TAC+MMF	17/42 (40%)*	
Kuzmina[42] (2015)	5 years	CsA+MTX+MSCs	4/39 (10%)*	0.75 (0.16–3.60)
		CsA+MTX	3/38 (8%)	
Ross[64] (1999)	5.3 years	CsA+Steroids	26/74 (35%)*	0.77 (0.38–1.53)
		CsA+MTX+Steroids	22/75 (29%)	
Bacigalupo[59] (2001)	5.7 years	CsA+MTX+ATG+Steroids	11/37 (30%)*	0.44 (0.14–1.36)
		CsA+MTX	6/38 (16%)	
Zikos[67] (1998)	5.8 (2.5–6.9) vs. 4.7 (2.5–6.8) years	CsA	3/28 (11%)	0.65 (0.14–3.00)
		CsA+MTX	5/32 (16%)*	
Ruutu[63] (2000)	6 (4.2–9.1) years	CsA+MTX	12/55 (22%)*	0.64 (0.24–1.71)
		CsA+MTX+Steroids	8/53 (15%)	
Bacigalupo[59] (2001)	9 years	CsA+MTX+ATG+Steroids	23/56 (41%)	0.67 (0.31–1.43)
		CsA+MTX	27/53 (51%)*	
Ruutu[44] (2014)	10 years	CsA+MTX+UDCA+Steroids	34/123 (28%)	**0.55 (0.32–0.93)**
		CsA+MTX	49/119 (41%)*	
Storb[73] (1986)	20 (10–21) years	CsA	23/50 (46%)*	**0.40 (0.17–0.98)**
		CsA+MTX	11/43 (26%)	
Storb[78] (1983)	Not reported	CsA	9/36 (25%)	0.74 (0.27–2.02)
		MTX	13/42 (31%)*	
Pulsipher[48] (2011)	Not reported	MTX+TAC	8/69 (12%)	0.83 (0.31–2.23)
		MTX+SIR+TAC	10/73 (14%)*	

[a]Patients in both arms received one of the following standard GVHD prophylactic regimens: CsA+Steroids, CsA+MTX, CsA+MTX+Steroids and CsA+MTX+ATG+Steroids. *denotes the reference group for each pairwise comparison

ATG = anti-thymocyte globulin; CsA = cyclosporin A; MMF = mycophenolate mofetil; MSCs = mesenchymal stem cells; MTX = methotrexate; RITX = rituximab; SIR = sirolimus; TAC = tacrolimus; THAL = thalidomide; UDCA = ursodeoxycholic acid

3.11. Relapse of underlying disease—2–3-year follow-up

A total of 17 studies reported the incidence of disease relapse in 2,188 patients and were analysed using NMA. Data for disease relapse were reported in several different formats in the included studies, including frequency data, and cumulative incidence and Kaplan-Meier estimates. Regardless of type, all data were pooled to allow a more robust analysis. In 7 studies, cumulative incidence and Kaplan-Meier estimates were converted to frequency data by multiplying the reported estimate by the initial group sample size. The accuracy of these calculated frequency estimates was tested in a sensitivity analysis in which the results of the NMA of all 17 pooled studies were compared to the results of an NMA of the 10 studies reporting frequency data (see report appendices). The point estimates and CrIs of the comparisons in common between the full network of all data types (including estimated frequency data) and the smaller network of only reported frequency data were comparable; thus, the use of estimated frequency data appeared robust, and we report the results of the analysis including all 17 studies below. Follow-up for the disease relapse outcome was set at 2–3 years and studies reporting data outside of this follow-up period were not included in the analysis.

Figure 13 presents a network diagram highlighting the available evidence for the comparison of prophylactic regimens for GVHD, where the incidence of disease relapse was evaluated as an outcome. Data from head-to-head trials were available for 10 of the 55 potential pairwise comparisons in the network, with single studies informing 6 of the 10 head-to-head comparisons.

Figure 13: Network diagram, Relapse of underlying disease (11 interventions, 2,188 patients)

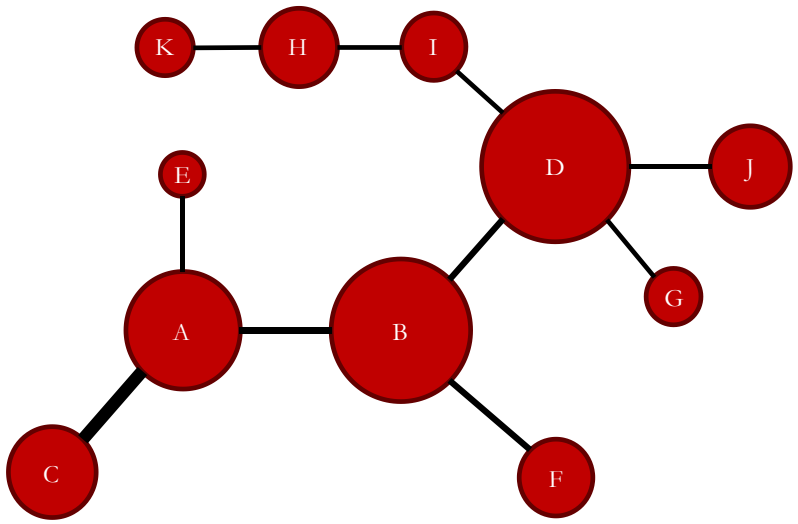


Figure legend, Treatments Shown	
A = CsA (n = 310)	F = CsA+MTX+ATG+Steroids (n = 132)
B = CsA+MTS (n = 460)	G = MTX+SIR+TAC (n = 70)
C = MTX (n = 187)	H = CsA+MTX+Steroids (n = 140)
D = MTX+TAC (n = 515)	I = MTX+TAC+Steroids (n = 103)
E = HCQ+CsA (n = 46)	J = SIR+TAC (n = 151)
	K = CsA+Steroids (n = 74)

3.11.1. Results from traditional pairwise meta-analyses

Table 12 presents a summary of pairwise estimates that were derived from direct evidence (i.e., head-to-head trials). In the table, these results are summarized alongside the numbers of available trials and patients for each direct comparison, as well as the related summary estimate derived from the network meta-analysis which is described below; this has been done to facilitate comparison of results between analyses. The direct comparison with the largest number of studies available was CsA vs MTX (5 RCTs, n = 358). Pairwise findings suggested a significantly reduced risk of disease relapse with CsA+MTX compared to CsA. Estimates from traditional pairwise meta-analyses compared well with estimates derived from network meta-analysis.

Table 12: Summary of results from meta-analysis and NMA, Relapse of underlying disease

Comparison		# of Trials (patients)	Direct Estimate OR (95% CrI)	NMA Estimate OR (95% CI)
			**Values <1 favor comparator*	
Reference	**Comparator**			
CsA+MTX	CsA	2 (173)	**0.29 (0.11–0.73)**	**0.26 (0.10–0.65)**
HCQ+CsA	CsA	1 (95)	0.75 (0.33–1.71)	0.75 (0.33–1.72)
CsA+MTX+ATG+Steroids	CsA+MTX	2 (255)	0.85 (0.48–1.49)	0.84 (0.47–1.49)
MTX+TAC	CsA+MTX	2 (509)	0.81 (0.53–1.25)	0.81 (0.53–1.24)
CsA+Steroids	CsA+MTX+Steroids	1 (149)	0.80 (0.32–1.97)	0.79 (0.31–1.96)
MTX+TAC+Steroids	CsA+MTX+Steroids	1 (131)	0.49 (0.18–1.33)	0.48 (0.17–1.30)
CsA	MTX	5 (358)	0.65 (0.37–1.16)	0.64 (0.36–1.14)
MTX+SIR+TAC	MTX+TAC	1 (140)	0.78 (0.39–1.56)	0.77 (0.38–1.57)
MTX+TAC	MTX+TAC+Steroids	1 (74)	0.55 (0.19–1.63)	0.54 (0.17–1.62)
MTX+TAC	SIR+TAC	1 (304)	0.96 (0.58–1.57)	0.95 (0.58–1.57)

3.11.2. Results from network meta-analysis

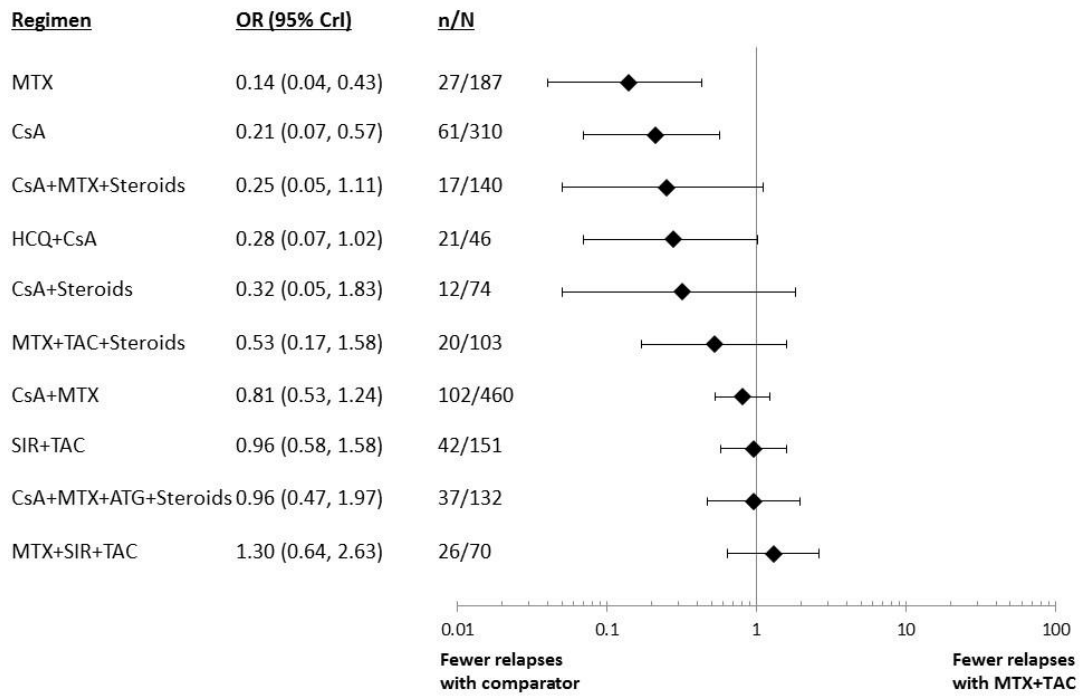
Most comparisons (i.e., 45 of 55) were informed only by indirect evidence and many of the comparisons with direct evidence were informed by single studies with limited numbers of patients. Model fit statistics for the NMA indicated that both the fixed-effects (FE) and random-effects (RE) models had an adequate fit. Posterior residual deviance values for the FE and RE models of 31.94 and 32.04, respectively, indicating adequate fit of both models, given the 34 data points in the model. DIC values (192.709 versus 194.545) suggested similar fit for both models. The FE analysis was preferred for reasons described earlier. Results from the RE model are provided in the appendices to the report.

3.11.2.1. Comparisons versus MTX+TAC

Figure 14 presents a forest plot summarizing comparisons of all treatments in the evidence network to the chosen reference therapy, MTX+TAC, based on the FE analysis. Two regimens—MTX and CsA—were found to have statistically significantly reduced risk of disease relapse relative to MTX+TAC (range of ORs from 0.14 to 0.21). No regimens were associated with significantly increased risk of disease relapse

compared to MTX+TAC, as all others were associated with credible intervals that included 1. While 95% credible intervals were wider in the corresponding RE analysis, clinical interpretations remained the same.

Figure 14: Forest plot of comparisons from NMA versus MTX+TAC—Relapse of underlying disease (FE model)

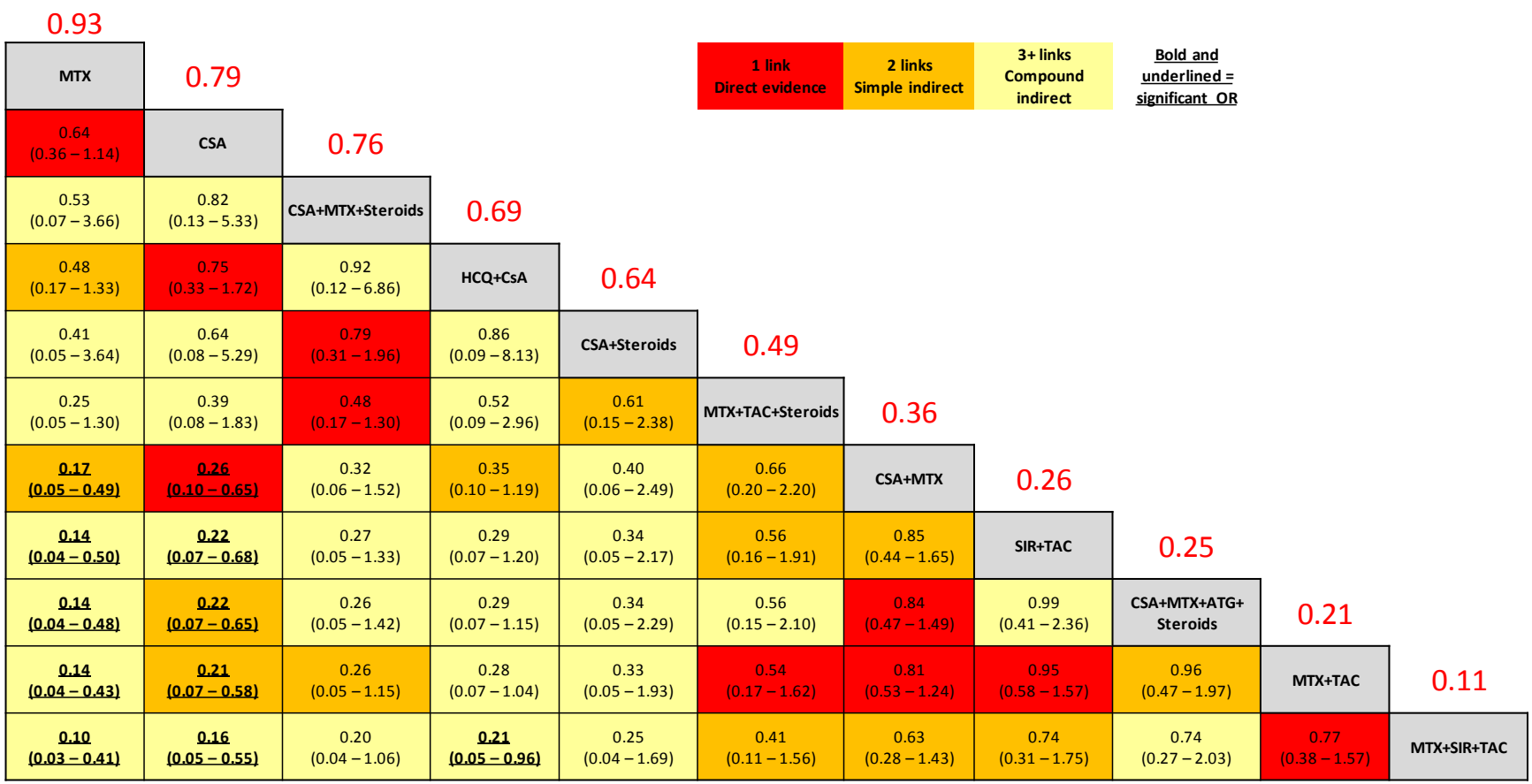


Odds ratios from network meta-analysis comparing each regimen to the reference treatment of MTX+TAC are presented along with the number of events and patients per regimen. Odds ratios <1 suggest more risk of relapse with MTX+TAC while odds ratios >1 suggest less risk of relapse with MTX+TAC. Results from the FE analysis are shown.

3.11.2.2. Comparisons between other prophylactic regimens

Figure 15 presents a league table of the estimates for all pairwise comparisons generated from the FE network meta-analysis. Comparisons of the 6 top-ranked regimens based on SUCRA value (MTX, CsA, CsA+MTX+Steroids, HCQ+CsA, CsA+Steroids, and MTX+TAC+Steroids) demonstrated no significant differences in their risk of disease relapse. One of these regimens (HCQ+CsA) was evaluated on a very low number of patients, reducing the power to detect significant differences in disease relapse risk relative to other regimens. Some of the comparisons included in the league table should be interpreted carefully, specifically those that are based on compound indirect evidence (i.e., comparisons in **Figure 15** where there are 2 or more intermediate treatments between the interventions of interest). Results from the corresponding RE analysis presented in the appendices found increased uncertainty associated with all comparisons, and again there were no significant differences identified amongst the top ranked treatments as based upon ordering by SUCRA. As was seen in findings from the FE analysis, significantly reduced risk of relapse with MTX (relative to CSA+MTX, CSA+MTX+ATG+Steroids and MTX+SIR+TAC) and with CsA (relative to CSA+MTX) remained.

Figure 15: League table summary of pairwise comparisons from FE NMA, Odds Ratios with 95% CrI—Relapse of underlying disease



Interventions are sorted from left to right in order of decreasing SUCRA value (i.e., preferred treatments appear first). For each comparison shown, the upper/left-most regimen is the comparator group while the lower/right-most treatment is the reference treatment; a value <1 suggests fewer cases of relapse with the comparator than with the reference group. Statistically significant differences (i.e., estimates with a 95% credible interval excluding 1) are bolded and underlined. Comparisons with direct evidence, simple indirect evidence, and compound indirect evidence have been highlighted in red, orange, and yellow, respectively.

3.12. Relapse of underlying disease—Other follow-up times

Many studies reported disease relapse data at other follow-up times and have been summarized narratively and presented in **Table 13**. Only one study identified a significant difference in the risk of relapse between treatment groups. Backman et al[71] (1988) found that patients taking MTX as prophylaxis for GVHD were at significantly lower risk of disease relapse than patients taking CsA, at 4.1 years post-transplant (OR = 0.17; 95% CrI = 0.03–0.89).

Table 13: Summary of results: Risk of disease relapse in the evaluation of GVHD prophylaxis at any follow-up time other than 2–3 years post-transplant

Author	Follow-up	Treatment regimen	Group risk	OR (95% CI)
Kuzmina[41] (2015)	5 years	CsA+MTX+MSCs	9/39 (23%)	0.74 (0.26–2.04)
		CsA+MTX	11/38 (29%)*	
Kuzmina[40] (2015)	0.1–6.1 years	CsA+MTX+MSCs	8/32 (25%)	0.80 (0.27–2.38)
		CsA+MTX	10/34 (29%)*	
Ruutu[44] (2014)	1 year	UDCA+CsA+MTX+Steroids	17/115 (15%)	0.79 (0.39–1.59)
		CsA+MTX+Steroids	21/117 (18%)*	
Ruutu[44] (2014)	10 years	UDCA+CsA+MTX+Steroids	41/115 (36%)	0.95 (0.56–1.63)
		CsA+MTX	43/117 (37%)*	
Glass[37] (2014)	4 years	MMF+TAC	10/42 (24%)*	1.00 (0.37–2.73)
		RITX+TAC+MMF	10/42 (24%)	
Martin[39] (2012)	200 days	MTX+TAC+Steroids	19/92 (21%)	1.36 (0.59–3.11)
		MTX+TAC	12/46 (26%)*	
Pidala[47] (2012)	3.4 (2.3–5.0) vs. 4.1 (2.4–5.3) years	MTX+TAC	14/37 (39%)*	0.38 (0.13–1.10)
		SIR+TAC	7/37 (19%)	
Pulsipher[48] (2011)	Not reported	MTX+TAC	19/69 (28%)	0.94 (0.45–1.96)
		MTX+SIR+TAC	21/73 (29%)*	
Bacigalupo[59] (2001)	1 year	CsA+MTX+ATG+Steroids	10/27 (36%)*	0.37 (0.11–1.28)
		CsA+MTX	5/28 (18%)	
Deeg[84] (2000)	6.1 (3.6–8.0) years	CsA	17/60 (28%)*	0.67 (0.29–1.54)
		CsA+Steroids	13/62 (21%)	
Ruutu[63] (2000)	6 (4.2–9.1) years	CsA+MTX	16/55 (29%)	0.87 (0.38–1.97)
		CsA+MTX+Steroids	17/53 (32%)*	
Chao[68] (1996)	5.8 (2.8–6.9) vs. 4.7 (2.5–6.8) years	THAL+Standard prophylaxis[a]	8/28 (29%)*	0.45 (0.12–1.74)
		Placebo+Standard prophylaxis[a]	4/26 (13%)	
Chao[64] (1993)	Not reported	CsA+Steroids	9/74 (12%)*	0.74 (0.26–2.11)
		CsA+MTX+Steroids	7/75 (9%)	
Backman[71] (1988)	4.1 (2.7–5.8) years	CsA	9/30 (30%)*	**0.17 (0.03–0.89)**
		MTX	2/29 (7%)	
Forman[72] (1987)	Not reported	CsA+Steroids	9/54 (17%)*	0.64 (0.21–1.94)
		MTX+Steroids	6/53 (11%)	
Storb[73] (1986)	1.15 (0.36–1.98) vs. 1.23 (0.3–1.95) years	CsA	3/50 (6%)	0.49 (0.11–2.16)
		CsA+MTX	5/43 (12%)*	

Table 13: Summary of results: Risk of disease relapse in the evaluation of GVHD prophylaxis at any follow-up time other than 2–3 years post-transplant

Author	Follow-up	Treatment regimen	Group risk	OR (95% CI)
Storb[73] (1986)	4 years	CsA	5/50 (10%)	0.57 (0.17–1.95)
		CsA+MTX	7/43 (16%)*	
Storb[74] (1986)	20 (10–21) years	CsA	6/50 (12%)	0.84 (0.25–2.83)
		CsA+MTX	6/43 (14%)*	

[a]Patients in both arms received one of the following standard GVHD prophylactic regimens: CsA+Steroids, CsA+MTX, CsA+MTX+Steroids and CsA+MTX+ATG+Steroids.
*denotes the reference group for each pairwise comparison
ATG = anti-thymocyte globulin; CsA = cyclosporin A; MMF = mycophenolate mofetil; MSCs = mesenchymal stem cells; MTX = methotrexate; RITX = rituximab; SIR = sirolimus; TAC = tacrolimus; THAL = thalidomide; UDCA = ursodeoxycholic acid

3.13. Findings: Specific harms

Few studies reported harm-related data. As a result, these data have been summarized narratively under the headings for each specific harm of interest. Where possible, ORs have been calculated to identify significant differences between pairs of interventions.

3.13.1. Total adverse events

One study (Hiraoka 2001) reported a greater proportion of patients experiencing total adverse events when given MTX+TAC+Steroids (31%) compared to patients given CsA+MTX+Steroids (18%) at 900 days post-transplant. However, the results were not statistically significant (OR = 2.06; 95% CI = 0.91–4.65).

3.13.2. Serious adverse events

One study (Socie[50] 2011) reported a significantly greater proportion of patients experiencing serious adverse events when given CsA+MTX+ATG+Steroids (76%) compared to patients given CsA+MTX (65%) at 2 years post-transplant. These results were statistically significant (OR = 2.14; 95% CI = 1.14–3.99).

3.13.3. Infection—Sepsis

One study (Forman[72] 1987) reported a significantly greater proportion of septic patients given MTX+Steroids (98%) than patients given CsA+Steroids (72%)(OR = 20.0; 95% CI = 2.53–157.92). The follow-up period was not reported.

3.13.4. Infection—Pneumonia

Six studies reported various types of pneumonia occurring during follow-up, including interstitial pneumonia of many causes (idiopathic, CMV, pneumocystis, unknown cause); CMV-induced interstitial pneumonia; pneumonia diagnosed as alveolar infiltrates due to aspiration, bacterial, fungal, or unknown causes; and "pneumonia" with no definition. The findings from these studies are presented in **Table 14**.

Weiden et al[81] observed that MTX was associated with a significantly increased incidence of transient interstitial pneumonia, but not fatal interstitial pneumonia, when compared to ATG+MTX+Steroids at ≤2 years of follow-up. At 2.3 years of follow-up, the incidence of non-CMV pneumonia was significantly increased in MTX patients compared to CsA patients; however, CMV-induced interstitial pneumonitis was not increased (Irle 1985). Finally, Ruutu et al. found that CsA+MTX was associated with a significantly

increased incidence of pneumonia compared to CsA+MTX+Steroids at 4 months and at 1 year post-transplant. No other significant findings were found.

Table 14: Summary of results - Incidence of pneumonia in the evaluation of GVHD prophylaxis

Author	Follow-up	Outcome definition	Treatment regimen	Group incidence	OR
Weiden[81] (1979)	≤2 years	All interstitial pneumonia[a], transient or fatal	ATG+MTX+Steroids	8/29 (28%)*	**3.28 (1.08–9.99)**
			MTX	15/27 (56%)	
Weiden[81] (1979)	≤2 years	Transient interstitial pneumonia[a]	ATG+MTX+Steroids	2/29 (7%)*	**5.68 (1.08–29.80)**
			MTX	8/27 (30%)	
Weiden[81] (1979)	≤2 years	Fatal interstitial pneumonia[a]	ATG+MTX+Steroids	6/29 (21%)*	1.34 (0.39–4.66)
			MTX	7/27 (26%)	
Irle[77] (1985)	2.3 years	Pneumonia—CMV induced interstitial pneumonitis	CsA	4/26 (15%)*	1.67 (0.43–6.52)
			MTX	7/30 (23%)	
Irle[77] (1985)	2.3 years	Pneumonia—other causes[b]	CsA	4/26 (15%)*	**4.81 (1.33–17.38)**
			MTX	14/30 (47%)	
Storb[73] (1986)	3-4.5 years	Interstitial pneumonia	CsA	14/50 (28%)	2.40 (0.83–6.93)
			CsA+MTX	6/43 (14%)*	
Forman[72] (1987)	NR	Interstitial pneumonia	CsA+Steroids	14/54 (26%)*	1.24 (0.53–2.88)
			MTX+Steroids	16/53 (30%)	
Ruutu[63] (2000)	4 months	Pneumonia	CsA+MTX	13/55 (24%)	**3.79 (1.15–12.51)**
			CsA+MTX+Steroids	4/53 (8%)*	
Ruutu[63] (2000)	1 year	Pneumonia	CsA+MTX	27/55 (50%)	**3.29 (1.43–7.58)**
			CsA+MTX+Steroids	12/53 (23%)*	
Glass[37] (2014)	4 years	Pneumonia	RITX+TAC+MMF	11/42 (26%)	1.14 (0.42–3.05)
			MMF+TAC	10/42 (24%)*	

[a]includes idiopathic, CMV, pneumocystis, and unknown causes
[b]pneumonia diagnosed by alveolar infltrates due to aspiration, bacterial, fungal, or unknown causes
*denotes the reference group for each pairwise comparison
ATG = anti-thymocyte globulin; CsA = cyclosporin A; MMF = mycophenolate mofetil; MTX = methotrexate; RITX = rituximab

3.13.5. Infection—CMV

Seven studies reported the incidence of CMV reactivation, viremia, or disease, with the findings presented in **Table 15**. Two studies identified significant differences in CMV-related outcomes. Ruutu et al[63] found that CsA+MTX was associated with significantly greater incidence of CMV viremia, signs, and symptoms than CsA+MTX+Steroids at 1 year of follow-up but not at 4 months of follow-up. However, CsA+MTX was associated with a significantly lower proportion of CMV-positive patients diagnosed by PCR than ATG+CsA+MTX+Steroids at 2 years of follow-up (Socie[50] 2011). No other comparisons demonstrated significant differences in CMV-related outcomes.

Table 15: Summary of results-Incidence of CMV-related outcomes in the evaluation of GVHD prophylaxis					
Author	**Follow-up**	**Outcome definition**	**Treatment regimen**	**Group incidence**	**OR**
Nash[62] (2000)	2 years minimum	CMV[a]	CsA+MTX	14/90 (16%)*	1.17 (0.54–2.57)
			MTX+TAC	16/90 (18%)	
Ruutu[63] (2000)	4 months	CMV[b]	CsA+MTX	18/55 (33%)	2.38 (0.96–5.92)
			CsA+MTX+Steroids	9/53 (17%)*	
Ruutu[63] (2000)	1 year	CMV[b]	CsA+MTX	27/55 (49%)	**2.69 (1.20–6.02)**
			CsA+MTX+Steroids	14/53 (26%)*	
Bacigalupo-I[59] (2001)	NR	CMV reactivation	ATG+CsA+MTX+Steroids	24/29 (83%)*	1.53 (0.33–7.15)
			CsA+MTX	22/25 (88%)	
Bacigalupo-II[59] (2001)	NR	CMV reactivation	ATG+CsA+MTX+Steroids	21/27 (78%)	1.66 (0.50–5.53)
			CsA+MTX	19/28 (68%)*	
Pulsipher[48] (2011)	NR	CMV reactivation	MTX+TAC	13/69 (19%)	1.65 (0.66–4.15)
			MTX+TAC+SIR	9/73 (12%)*	
Socie[50] (2011)	2 years	CMV—diagnosed by PCR	ATG+CsA+MTX+Steroids	57/103 (54%)	**3.10 (1.72–5.56)**
			CsA+MTX	28/98 (30%)*	
Martin[39] (2012)	200 days	CMV viremia or disease	MTX+TAC+Steroids	23/92 (25%)*	1.06 (0.47–2.38)
			MTX+TAC	12/46 (26%)	
Cutler[43] (2014)	100 days	CMV reactivation	SIR+TAC	20/151 (13%)*	1.16 (0.61–2.21)
			MTX+TAC	23/153 (15%)	

[a]method of diagnosis not reported
[b] viremia (early antigen-positive) and signs and symptoms
*denotes the reference group for each pairwise comparison
ATG = anti-thymocyte globulin; CMV = cytomegalovirus; CsA = cyclosporin A; MTX = methotrexate; PCR = polymerase chain reaction; SIR = sirolimus; TAC = tacrolimus

3.13.6. Infection—Fungal

Two studies reported outcomes related to fungal infection. Deeg et al[84] demonstrated that fewer fungal infections occurred in patients receiving CsA (20%) compared to CsA+Steroids (26%), although the difference was not statistically significant (OR = 1.39; 95% CI = 0.59–3.26). Martin et al[39] found very few occurrences of invasive yeast or mold infections in patients receiving either MTX+TAC (1/46) or MTX+TAC+Steroids (1/92). There was no significant difference in the two patient groups (OR = 2.02; 95% CI = 0.12–33.08).

3.13.7. Infection—Other

Four studies reported other infectious outcomes, including an aggregation of all infections, all bacterial infections, grade 3–5 adverse infectious events, "infectious complications," and gastrointestinal infections. These data are summarized in **Table 16**. Only Glass et al[37] identified a significant difference in interventions: MMF+TAC was associated with a significantly greater incidence of Grade 3–5 infectious adverse events compared to RITX+TAC+MMF at 4 years.

Table 16: Summary of results: Incidence of other infections in the evaluation of GVHD prophylaxis					
Author	**Follow-up**	**Outcome definition**	**Treatment regimen**	**Group incidence**	**OR**
Deeg[84] (2000)	6.1 years	All infections	CsA	56/60 (93%)*	4.36 (0.47–40.16)
			CsA+Steroids	61/62 (98%)	
Deeg[84] (2000)	6.1 years	All bacterial infections	CsA	40/60 (67%)	1.10 (0.52–2.32)
			CsA+Steroids	40/62 (65%)*	
Martin[39] (2012)	200 days	Gastrointestinal infection	MTX+TAC+Steroids	22/92 (24%)*	1.39 (0.63–3.07)
			MTX+TAC	14/46 (30%)	
Glass[37] (2014)	4 years	Grade 3–5 infectious adverse events	RITX+TAC+MMF	29/42 (69%)*	**5.83 (1.52–22.35)**
			MMF+TAC	39/42 (93%)	
Bonifazi[45] (2014)	2 years	Infectious complications	CSA+MTX+ATG+Steroids	48/83 (58%)	1.16 (0.61–2.19)
			CSA+MTX	39/72 (54%)*	
*denotes the reference group for each pairwise comparison ATG = anti-thymocyte globulin; CsA = cyclosporin A; MMF = mycophenolate mofetil; MTX = methotrexate; RITX = rituximab; TAC = tacrolimus					

3.13.8. Renal toxicity

Seven studies reported on the incidence of renal toxicity within the patient groups (**Table 17**). Renal dysfunction was defined as an elevation in creatinine in all studies. CsA was found to significantly increase the incidence of renal dysfunction at 1.5 years of follow-up compared to MTX (Torres[70] 1989), while two studies demonstrated that MTX+TAC was associated with a significantly greater incidence of renal dysfunction at 60–100 days, when renal dysfunction was defined as >2 mg/dl but not when defined as >2 × baseline (Ratanatharathorn[65], Nash[62]). Ratanatharathorn found no significant difference in these treatment groups at 1 year of follow-up.[65]

Table 17: Summary of results: Incidence of renal toxicity outcomes in the evaluation of GVHD prophylaxis					
Author	**Follow-up**	**Outcome definition**	**Treatment regimen**	**Group incidence**	**OR**
Forman[72] (1987)	NR	Creatinine >2.0 mg/dl	CsA+Steroids	5/54 (9%)	1.25 (0.32–4.93)
			MTX+Steroids	4/53 (8%)*	
Torres[70] (1989)	1.5 years	Creatinine >2.0 mg/dl or 2 × baseline	CsA	16/26 (62%)	**3.91 (1.29–11.83)**
			MTX	9/31 (29%)*	
Mrsic[69] (1990)	30 months	Creatinine >150 mmol/l	CsA	14/39 (36%)*	1.36 (0.54–3.42)
			CsA+MTX	16/37 (43%)	
Ratanatharathorn[65] (1998)	60 days	Creatinine >2.0 mg/dl	CsA+MTX	79/164 (48%)*	**1.61 (1.04–2.50)**
			MTX+TAC	99/165 (60%)	
Ratanatharathorn[65] (1998)	1 year	Creatinine >2.0 mg/dl	CsA+MTX	98/164 (60%)*	1.38 (0.88–2.17)
			MTX+TAC	111/165 (67%)	
Ratanatharathorn[65] (1998)	60 days	Creatinine >2 × baseline	CsA+MTX	119/164 (73%)*	1.51 (0.91–2.53)

Table 17: Summary of results: Incidence of renal toxicity outcomes in the evaluation of GVHD prophylaxis					
Author	**Follow-up**	**Outcome definition**	**Treatment regimen**	**Group incidence**	**OR**
			MTX+TAC	132/165 (80%)	
Ratanatharathorn[65] (1998)	1 year	Creatinine >2 × baseline	CsA+MTX	131/164 (80%)*	1.35 (0.76–2.37)
			MTX+TAC	139/165 (84%)	
Nash[62] (2000)	100 days	Creatinine >2.0 mg/dl	CsA+MTX	36/90 (40%)*	**1.88 (1.04–3.39)**
			MTX+TAC	50/90 (46%)	
Nash[62] (2000)	100 days	Creatinine >2 × baseline	CsA+MTX	69/90 (77%)*	1.41 (0.68–2.92)
			MTX+TAC	74/90 (82%)	
Socie[50] (2011)	2 years	Renal failure	CsA+MTX+ATG+Steroids	23/103 (22%)	1.37 (0.68–2.76)
			CsA+MTX	17/98 (18%)*	
Glass[37] (2014)	4 years	"Abnormal" creatinine	RITX+TAC+MMF	7/42 (17%)	1.00 (0.32–3.14)
			MMF+TAC	7/42 (17%)*	

*denotes the reference group for each pairwise comparison
ATG = anti-thymocyte globulin; CsA = cyclosporin A; MMF = mycophenolate mofetil; MTX = methotrexate; RITX = rituximab; TAC = tacrolimus

3.13.9. Hepatic toxicity

Three studies reported the incidence of hepatic toxicity (**Table 18**). Only one study identified a significant difference between the interventions with respect to hepatic toxicity (Ruutu, 2013). Regardless how it was defined, the incidence of hepatic dysfunction was significantly greater in CsA+MTX compared to UDCA+CsA+MTX at 90 days and 1 year of follow-up.

Table 18: Summary of results: Incidence of hepatic toxicity in the evaluation of GVHD prophylaxis					
Author	**Follow-up**	**Outcome definition**	**Treatment regimen**	**Group incidence**	**OR**
Irle[77] (1985)	2.3 years	Hepatic dysfunction[a]	CsA	24/26 (92%)	1.85 (0.31–11.01)
			MTX	26/30 (87%)*	
Ruutu[44] (2013)	90 days	Bilirubin >2.5 × ULN	UDCA+CsA+MTX	18/123 (15%)*	**2.05 (1.08–3.92)**
			CsA+MTX	31/119 (26%)	
Ruutu[44] (2013)	90 days	ALT >2.5 × ULN	UDCA+CsA+MTX	64/123 (52%)*	**1.89 (1.22–3.18)**
			CsA+MTX	80/119 (67%)	
Ruutu[44] (2013)	1 year	Severe hepatic disease[b]	UDCA+CsA+MTX	6/123 (5%)*	**3.71 (1.42–9.64)**
			CsA+MTX	19/119 (16%)	
Glass[37] (2014)	4 years	Abnormal bilirubin	RITX+TAC+MMF	9/42 (21%)	2.59 (0.73–9.20)
			MMF+TAC	4/42 (10%)*	

[a]categorized as follows: no liver function abnormality; liver function abnormality from causes other than VOD; mild and subclinical VOD with maximum bilirubin <257 umol/l, maximum ALT <200 IU, weight gain >10% of baseline weight; severe VOD with bilirubin >257 gmol/l, maximum ALT >200 IU, weight gain >10% of baseline and/or encephalopathy; fatal outcome with VOD being a significant contributing factor to death.
[b]stage III-IV liver GVHD, VOD (Jones criteria), or fatal non-GVHD liver failure not fulfilling Jones' criteria for VOD
*denotes the reference group for each pairwise comparison
ALT = alanine transferase; CsA = cyclosporin A; MMF = mycophenolate mofetil; MTX = methotrexate; RITX = rituximab; TAC = tacrolimus; UDCA = ursodeoxycholic acid; VOD = veno-occlusive disease

3.13.10. Veno-occlusive disease

Four studies evaluated the incidence of veno-occlusive disease in patient groups (**Table 19**). Pulsipher et al. demonstrated a borderline significant increase in VOD in patients receiving MTX+TAC+SIR compared to patients receiving MTX+TAC. No other comparisons were found to be significantly different with respect to incidence of VOD.

Table 19: Summary of results: Incidence of veno-occlusive disease in the evaluation of GVHD prophylaxis

Author	Follow-up	Outcome definition	Treatment regimen	Group incidence	OR
Ruutu[44] (2014)	90 days	VOD[a]	UDCA+CsA+MTX	3/123 (2%)*	1.75 (0.41–7.51)
			CsA+MTX	5/119 (4%)	
Ruutu[44] (2014)	90 days	VOD[b]	UDCA+CsA+MTX	14/123 (11%)*	1.04 (0.47–2.28)
			CsA+MTX	14/119 (12%)	
Cutler[43] (2014)	100 days	VOD[c]	SIR+TAC	17/151 (11%)	2.30 (0.96–5.50)
			MTX+TAC	8/153 (5%)*	
Pulsipher[48] (2014)	NR	VOD[c]	MTX+TAC	6/70 (9%)*	**2.76 (1.004–7.58)**
			MTX+TAC+SIR	15/73 (21%)	

[a]diagnosed according to the criteria of Jones et al.
[b]diagnosed according to the criteria of McDonald et al.
[c]no definition provided
*denotes the reference group for each pairwise comparison
CsA = cyclosporin A; MTX = methotrexate; NR = not reported; SIR = sirolimus; TAC = tacrolimus; UDCA = ursodeoxycholic acid; VOD = veno-occlusive disease

3.13.11. Other harms

Two studies evaluated harms that did not fit into the categories of specific harms identified by our content experts. Mrsic et al[69] found no significant difference in the incidence of neurotoxicity after 2.5 years in patients receiving CsA compared to patients receiving CsA+MTX (15% vs. 11%; OR = 2.48; 95% CI = 0.69–8.88). As well, no significant difference in the incidence of multi-organ failure was found in patients receiving CsA+MTX+ATG+Steroids compared to patients receiving CsA+MTX (5% vs. 5%; OR = 1.05; 95% CI = 0.30–3.76) (Socie[50] 2011).

3.14. Other studies not considered in network meta-analyses

One study evaluated prophylaxis of cGVHD, comparing thalidomide vs. placebo. In this study, patients were assigned to 3 different standard prophylactic regimens for aGVHD, to which thalidomide or placebo were added at day 80 post-transplant, resulting in a total of 6 different prophylactic regimens. The results were not presented by subgroup of standard prophylaxis regimen, only overall by thalidomide and placebo; thus, the data could not be included in NMAs and have been summarized narratively here. Study characteristics are presented in **Table 3**; a total of 54 patients were included in the study, 28 receiving thalidomide and 26 receiving placebo. The three standard prophylactic regimens were CsA+MTX (n = 13), CsA+MTX+Steroids+(ATG to some patients) (n = 34), and CsA+Steroids (n = 7).

Two outcomes of interest were reported: the incidence of cGVHD at one year post-transplant and 1-year overall mortality. Compared to the placebo group, patients in the thalidomide group were significantly more likely to develop cGVHD within one year of transplant (64% vs. 38%; OR = 3.38; 95% CI = 1.10–10.35) and were significantly more likely to die within one year of transplant (39% vs. 8%; OR = 7.76; 95% CI = 1.52–39.62).

3.15. Assessment of between-study heterogeneity and inconsistency

As described earlier, several forms of variation in patient and study characteristics were noted across the included studies in terms of types of patients enrolled, year of study publication and transplantation methods. Unfortunately, subgroup analyses and meta-regression analyses were not feasible in the context of this review due to the presence of many single-study connections across networks (which precludes performance of meaningful meta-regression analyses), as well as a failure of studies to report outcomes in subgroups of relevance which would enable the performance of more focused analyses. Given the high extent heterogeneity present and the highly limited ability to evaluate its corresponding impact on study findings, the validity and applicability of findings to specific populations must be cautiously considered.

The networks of both the aGVHD and cGVHD endpoints included a closed loop; thus, fixed effect inconsistency models were also fit to the data. A table provided in **Appendix 7** provides a summary of posterior residual deviance and DIC values which enabled inspection of inconsistency between the direct and indirect evidence. Overall, these analyses suggested that the findings derived from direct and indirect evidence were not significantly different for the aGVHD outcome. For cGVHD, comparison of DIC values suggested some possible evidence of a differences in the direct and indirect evidence, however a clear cause of this potential difference could not be found.

4. REVIEW FINDINGS—TREATMENT OF GVHD

4.1. Availability of relevant literature

Figure 1 presented a summary of the process of study selection for both the review of GVHD prophylaxis and the review of GVHD treatment. Following screening of full text articles, there were 7 trials in 7 unique publications available for the review of GVHD treatment.[85-91]

4.2. Overview of study characteristics

A total of 830 patients were enrolled in the 7 included studies.[85-91] Median sample size of the included trials was 96 patients (range 54–235). Median year of publication was 2001 (range 1985–2014). Totals of 2 (28.5%), 1 (14.3%), 3 (42.9%), and 1 studies (14.3%) were published between 1980–1990, 1991–2000, 2001–2010, and after 2010, respectively. Of the 5 studies reporting setting, 3 (42.9%) were multi-centre[87,88,90] and 2 (28.6%) were single-centre.[86,91] One trial (14.3%) was entirely supported by for-profit/industry sources,[86] while 3 (42.9%) were funded by non-profit/government organizations, and 1 (14.3%) had support from both private and public sources. No information could be obtained for one trial regarding the source of funds.[91] Study follow-up duration varied from 9–12 months in 2 studies (28.5%),[87,88] and 2–5 years in 5 studies (57.1%).[85,86,89-91]

One study[86] allowed crossover between groups for patients non-responsive to first-line therapy in the first 7 days. This study's data were analysed using an intention-to-treat approach.

4.3. Overview of patient characteristics

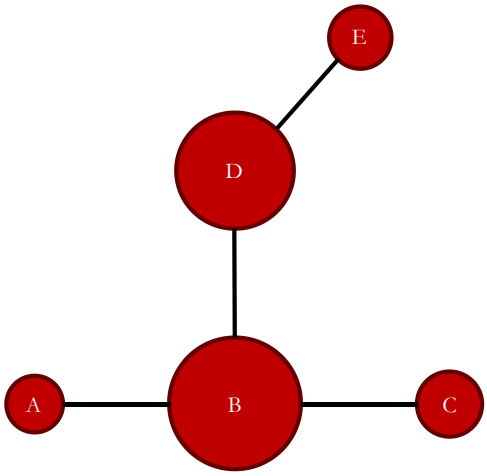
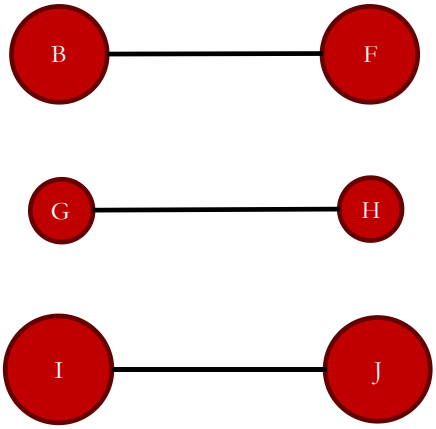
Five studies (71.4%) enrolled patients 12 years of age or younger as part of the inclusion criteria.[86-89,91] One study restricted eligibility to those of ages >12 years,[85] and the age range of participants in the remaining trial could not be ascertained.[90] The median proportion of female patients was 40% (range 28.0–46.0%) in the 6 studies reporting this information, while data for one study was not provided.[91] Donor source was 100% bone marrow (BM) in 4 studies (57.1%).[85,86,89,91] In the remaining 3 studies (42.9%),[87,88,90] BM, umbilical cord blood (UC), and peripheral blood (PB) stem cells were used in varying proportions, with UC, PB, and BM consisting of between 3.0–9.0%, 71.0–84.0%, and 11.0–20.0% of the transplants,

respectively. Three studies (42.9%) reported full HLA matching for 83.0%,[90] 91.0%,[85] and 94.0% of their participants;[89] HLA matching information was not reported in the remaining 4 studies (57.1%). Two studies (28.5%) reported information about the proportion of gender matched donor/recipients.[90,91] The proportion of patients that received female-to-male transplants was 17.0% and 22.0% in these two studies. The remaining five studies (71.4%) did not report gender matching (**Table 20, 21**).

4.4. Overview of network geometry and interventions compared

Figures 16a and 16b present network diagrams highlighting the extent of available evidence for the comparison of treatment regimens for acute and chronic GVHD for all outcomes. Each of the 10 distinct treatment regimens present across the two networks is represented by a node and the randomized trials comparing pairs of regimens are shown as links between the nodes. The size of each node is reflective of the total number of patients enrolled in studies evaluating that regimen, while the width of the link between two nodes is reflective of the number of studies that directly compared the two regimens. Due to the small number of relevant studies identified, heterogeneity in eligibility criteria in terms of type of GVHD being treated, variability in outcome definitions and follow-up periods and fragmentation of evidence networks, primary focus in terms of synthesizing findings used a narrative approach for this review.

Figures 16a and 16b: Network diagrams of evidence base, acute and chronic GVHD treatment

Treatment of aGVHD	Treatment of cGVHD
16a 	**16b** 
Figure legend A = CsA B = Steroids C = ATG+Steroids D = MMF+Steroids E = ETN +Steroids	F = AZA+Steroids G = THAL+CsA+Steroids H = CsA+Steroids I = MMF+(CsA or TAC or SIR)+Steroids J = (CsA or TAC or SIR)+Steroids

4.5. Outcomes reported

The outcomes of interest in the assessment of GVHD treatment included response to treatment, weaning from steroids and other GVHD treatments, mortality, and incidence of infection. Availability of this data as well as variations in assessment methods between studies are described next.

4.5.1. Response to treatment

All seven studies reported response to treatment in 830 patients. All studies included findings for "complete response" to treatment, which was generally defined as alive with resolution of clinical signs and symptoms of GVHD, without evidence of relapse. Two studies[87,88] excluded patients that had received additional

salvage therapies from the definition of complete response. One study[90] included only patients that had resolution of all clinical signs, had discontinued treatment and received no salvage therapies. All studies except two[89,90] reported partial response to treatment, either alone or in combination with complete response. Partial response was generally defined as improvement in signs in at least one organ system, without deterioration in another. Follow-up times ranged from 14–56 days in studies evaluating treatment of aGVHD and from 2 months to >2 years in studies evaluating cGVHD.

4.5.2. Weaning from steroids or other GVHD treatments

Three studies reported weaning from steroids or other GVHD treatments in 477 patients.[87,88,90] Definitions ranged from "reduction of dose or withholding medication" to "discontinuation of all immunosuppressive medication," and follow-up times ranged from 28 days to 9 months in studies evaluating treatment of aGVHD,[87,88] while the single study evaluating cGVHD treatment followed patients for 2 years.[90]

4.5.3. Overall mortality

Six studies reported data for overall mortality in 739 patients.[85,86,88-91] Follow-up ranged from 56 days to >3.8 years in studies evaluating the treatment of aGVHD and from 1–4 years in studies of cGHVD treatment.

4.5.4. Non-relapse mortality

Five studies reported data for non-relapse mortality in 504 patients.[85,86,89-91] Follow-up ranged from 100 days to 4 years in studies evaluating the treatment of aGVHD and up to 4 years in studies of cGHVD treatment, although one cGVHD study followed patients for an unspecified period of time.

4.5.5. Incidence of infection

Six studies reported data regarding the incidence of infection in 753 patients.[86-91] Infection definitions were extremely variable, including incidence of at least one infection of any kind, varicella zoster, cytomegalovirus, Epstein-Barr Virus, any viral infection, bacteremia, fungal infections, pneumonitis, conjunctivitis, etc. Follow-up times were often unclear; however, when reported they ranged from 56 days to 4 years for aGVHD treatment and from 9 months to 4 years in studies of cGVHD treatment.

4.5.6. Findings from Risk of Bias Assessment

Risk of bias assessment was possible for all 7 included studies. Adequate sequence generation was clearly reported in 2 studies (28.6%).[85,89] The remaining 5 studies (71.4%) did not provide sufficient detail on the method of random sequence generation. None of the 7 studies reported adequate information about allocation concealment and, therefore, this domain was coded as unclear risk of bias. Patients and investigators were blinded from the randomized allocation in 3 studies (42.9%),[88-90] and un-blinded in 2 (28.6%).[86,91] Independent and blinded outcome assessment was reported in 4 studies[87-90] (57.1%) while 2 studies (28.6%) were at high risk of bias for this domain given lack of blinding.[86,91] Five studies (571.4%) used intention-to-treat (ITT) or modified ITT analysis for the main outcomes of mortality, treatment success, and weaning of medication (GVHD treatment or steroids) and, therefore, exposed to low risk for attrition bias.[85-88,90] Four studies (57.1%) provided reference to the study protocol and were at low risk of bias for selective outcome reporting.[86,88,90,91] Data for baseline variables and co-intervention were poorly reported in 1 study (14.3%), which prevented the reviewers from objectively assessing other biases caused by imbalanced patient characteristics at baseline, co-interventions, or other possibly important variables. This domain was thus rated as unclear.[85,86,91] The remaining 2 (28.6%) and 2 (28.6%) studies were rated as low[89,90] and high risk[87,88] of other biases, respectively. **Figure 17** provies a summary of the distribution of risk of bias assessments across domains and studies, while **Appendix 4** presents a detailed breakdown of assessments by study.

Figure 17: Risk of Bias Summary, Studies of GVHD treatment

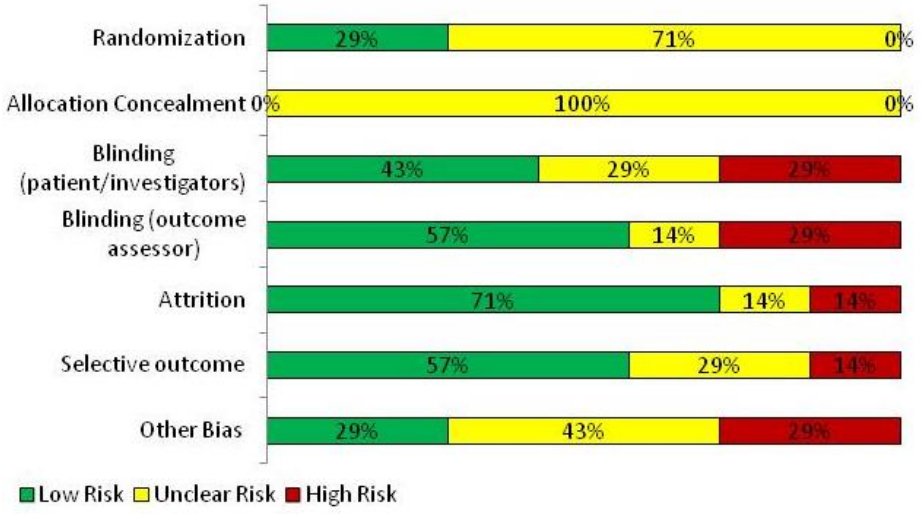


Table 20: Summary of characteristics of studies evaluating the treatment of acute GVHD										
Author (year)	**N total**	**Prophylaxis of aGVHD (%)**	**Treatment regimens**	**Patients' age range (years)**	**Donor source (%BM, %PB, %UC)**	**% related donor**	**% full HLA match**	**% female donor / male recipient**	**% of study patients considered to be high risk**	**Endpoints reported**
Kennedy[85] (1985)	77	MTX (100%)	a) Steroids for 14 days then tapered; b) CsA for 50 days then tapered	12–46	100,0,0	100	91	NR	30%	α, γ, ε
Cragg[86] (2000)	96	MTX and CsA or MTX+ATG+prednisone (NR).	a) ATG for 5 days +Steroids for 7 days then tapered; b) Steroids for 14 days then tapered	0–51	100,0,0	49	NR	NR	47%	α, γ, ε, ϕ
Alousi[87] (2009)	91	MMF (30%); those given MMF prophylaxis were allocated to ETN group or other regimens not presented here	a) ETN for 28 days+Steroids for 7 days then tapered; b) MMF until 28 days after steroids tapered+Steroids for 7 days then tapered	7.5–70	20,70,10	50	NR	NR	NR	α, β, ϕ
Bolanos-Meade[88] (2014)	235	Steroids (100%)	a) MMF for 56 days or when steroids discontinued (whichever earlier)+Steroids for at least 3 days then tapered; b) PLB+Steroids for at least 3 days then tapered	9.1–76.3	17,82,3	41	NR	NR	NR	α, β, γ, ϕ

ATG = anti-thymocyte globulin; BM = bone marrow; CB = cord blood; CsA = cyclosporin A; ETN = etanercept; HLA = human leukocyte antigen; MMF = mycophenolate mofetil; NR = not reported; PLB = placebo; TAC = tacrolimus; THAL = Thalidomide; SIR = sirolimus; UC = umbilical cord

α = response to treatment; β = weaning from all GVHD treatments or steroids; γ = overall mortality; ε = non-relapse mortality; ϕ = incidence of infection

Table 21: Summary of characteristics of studies evaluating the treatment of chronic GVHD										
Author (year)	**N total**	**Proportion with previous aGVHD II-IV; Prophylaxis of aGVHD (%)**	**Treatment regimens compared**	**Patients' age range (years)**	**Donor source (%BM, %PB, %CB)**	**% related donor**	**% full HLA match**	**% female donor / male recipient**	**% of study patients considered to be high risk**	**Endpoints reported**
Sullivan[89] (1988)	126	44.4%; trimethoprim/sulphoxazole (100%)	a) AZA for 9 months+Steroids for 9 months; b) PLB+Steroids for 9 months	1–48	100,0,0	94	94	NR	47%	α, γ, ε, ϕ
Arora[91] (2001)	54	70.4%; MTX+CsA or T-cell depletion or others (NR), methylprednisolone at randomization (100%)	a) (THAL+CsA+Steroids) for 9 months after clinical resolution of cGVHD; b) (CsA+Steroids) for 9 months after clinical resolution of cGVHD	2–60	100,0,0	63	NR	17	NR	α, γ, ε, ϕ
Martin[90] (2009)	151	31.1%; Steriods (95%), CsA (30%), TAC (52%), SIR (1%)	a) [MMF+(CsA or TAC or SIR)+Steroids] for unknown duration; b) [PLB+(CsA or TAC or SIR)+Steroids] for unknown duration	NR (55% older than 50 years)	11,84,5	51	83	22	32%	α, β, γ, ε, ϕ

AZA = alemtuzumab; BM = bone marrow; CB = cord blood; CsA = cyclosporin A; ETN = etanercept; HLA = human leukocyte antigen; PLB = placebo; MMF = mycophenolate mofetil; NR = not reported; TAC = tacrolimus; THAL = Thalidomide; SIR = sirolimus; UC = umbilical cord

α = response to treatment; β = weaning from all GVHD treatments or steroids; γ = overall mortality; ε = non-relapse mortality; ϕ = incidence of infection

4.6. Findings: Response to treatment

All seven studies presented data for 10 treatment regimens evaluating complete or partial response to treatment.

4.6.1. Response to treatment of acute GVHD

Three of four studies evaluated complete response,[85,87,88] and three of four studies evaluated complete+partial response.[85-87] **Table 22** presents a summary of the results for both endpoints.

Based on raw data reported across studies, calculation of odds ratios and corresponding confidence intervals identified statistically significant differences in the following treatment comparisons:

- ETN+Steroids demonstrated a significantly lower likelihood of complete response to treatment at both 28 days (OR 0.24, 95% CI 0.10–0.57) and 56 days (OR 0.28, 95% CI 0.12–0.68) compared to MMF+Steroids.

- ETN+Steroids had a significantly lower likelihood of complete or partial response at 28 days compared to MMF+Steroids (OR 0.26, 95% CI 0.11–0.65).

- In patients receiving transplants from unrelated donors, ATG+Steroids had a significantly lower likelihood of complete or partial response at 42 days compared to steroids alone (OR = 0.25; 95% CI = 0.07–0.90) (**Appendix 4**).

Table 22: Summary of results, response to treatment of aGVHD					
Author	**Follow-up**	**Outcome**	**Treatment regimen**	**Group incidence**	**OR**
Kennedy[85] (1985)	14 days	Complete response[a]	Steroids	5/39 (13%)	0.65 (0.19–2.27)
			CsA	7/38 (18%)	
Kennedy[85] (1985)	14 days	Complete+partial response	Steroids	16/39 (41%)	0.45 (0.18–1.13)
			CsA	23/38 (61%)	
Cragg[86] (2000)	42 days	Complete+partial response[c]	ATG+Steroids	34/45 (76%)	0.97 (0.36–2.58)
			Steroids	32/42 (76%)	
Alousi[87] (2009)	28 days	Complete response[d]	ETN+Steroids	12/46 (26%)[e]	**0.24 (0.10–0.57)**
			MMF+Steroids	27/45 (60%)	
Alousi[87] (2009)	28 days	Complete+partial response[f]	ETN+Steroids	22/46 (48%)	**0.26 (0.11–0.65)**
			MMF+Steroids	35/45 (78%)	
Alousi[87] (2009)	56 days	Complete response	ETN+Steroids	20/46 (44%)[g]	**0.28 (0.12–0.68)**
			MMF+Steroids	33/46 (73%)	
Bolanos-Meade[88] (2014)	14 days	Complete response[h]	MMF+Steroids	51/116 (44%)	0.80 (0.48–1.33)
			PLB+Steroids	59/119 (50%)	
Bolanos-Meade[88] (2014)	28 days	Complete response	MMF+Steroids	54/116 (47%)	1.08 (0.65–1.81)
			PLB+Steroids	53/119 (45%)	
Bolanos-Meade[88] (2014)	56 days	Complete response	MMF+Steroids	70/116 (60%)	1.31 (0.78–2.20)
			PLB+Steroids	64/119 (54%)	
Bolanos-Meade[88] (2014)	56 days	Alive and free of aGVHD and cGVHD	MMF+Steroids	69/116 (60%)	1.44 (0.86–2.42)
			PLB+Steroids	60/119 (50%)	

[a] resolution of clinical evidence of GVHD
[b] resolution of clinical evidence of GVHD or improvement in at least one organ without deterioration in the others
[c] resolution of all symptoms in all organs or improvement in ≥1 organ without any worsening in other organs
[d] resolution of all signs and symptoms of GVHD in all organs without intervening salvage therapies
[e] cumulative incidence 44% vs. 67%; significance not reported
[f] resolution of all signs and symptoms of GVHD in all organs without intervening salvage therapies or improvement of one stage in one or more organs without progression in any organ
[g] cumulative incidence 61% vs. 82%; significance not reported
[h] resolution of all signs and symptoms of aGVHD in all evaluable organs without additional therapies (i.e., alive and free of aGVHD but not necessarily cGVHD)
ATG = anti-thymocyte globulin; CsA = cyclosporin A; ETN = etanercept; MMF = mycophenolate mofetil; PLB = placebo

4.6.2. Response to treatment of chronic GVHD

All three studies evaluated complete response and reported findings.[89-91] **Table 23** presents a summary of findings by study. After pairwise analysis based on the available raw data, only one statistically significant difference was identified, showing that AZA+Steroids had a significantly lower likelihood of complete response to treatment than steroids alone after 9 months of treatment (OR 0.33, 95% CI 0.16-0.68). All other comparisons were associated with wide 95% confidence intervals and failed to achieve statistical significance.

Table 23: Summary of results, response to treatment of cGVHD					
Author	**Follow-up**	**Outcome**	**Treatment regimen**	**Group incidence**	**OR**
Sullivan[89] (1988)	9 months	Complete response[a]	AZA+Steroids	22/63 (35%)	**0.33 (0.16–0.68)**
			PLB+Steroids	39/63 (62%)	
Arora[91] (2001)	2 months	Complete response[b]	THAL+CsA+Steroid	2/24 (8%)	1.14 (0.15–8.76)
			CsA+Steroid	2/27 (7%)	
Arora[91] (2001)	6 months	Complete response	THAL+CsA+Steroid	4/24 (17%)	0.51 (0.13–2.05)
			CsA+Steroid	7/25 (28%)	
Arora[91] (2001)	1 year	Complete response	THAL+CsA+Steroid	10/20 (50%)	0.83 (0.25–2.80)
			CsA+Steroid	12/22 (55%)	
Martin[90] (2009)	<2 years	Treatment success[c]	MMF+(CsA or TAC or SIR)+Steroid	11/74 (15%)[d]	1.17 (0.46–2.94)
			PLB+(CsA or TAC or SIR)+Steroid	10/77 (13%)	
Martin[90] (2009)	>2 years	Treatment success	MMF+(CsA or TAC or SIR)+Steroid	0/74 (0%)	NE[e]
			PLB+(CsA or TAC or SIR)+Steroid	377 (4%)	

[a] alive without relapse or clinically active cGVHD, with or without histologic evidence of cGVHD
[b] resolution of all signs and symptoms of cGVHD
[c] resolution of all reversible manifestations of cGVHD with no secondary systemic therapy and withdrawal of all systemic treatment, including the study drug
[d] cumulative incidence 23% vs. 18%; 2-year adjusted HR = 1.66 (0.7–3.7)
[e] Not estimable
AZA = alemtuzumab; CsA = cyclosporin A; MMF = mycophenolate mofetil; PLB = placebo; SIR = sirolimus; TAC = tacrolimus; THAL = Thalidomide

4.7. Findings: Weaning from all GVHD treatments or steroids

Three studies (2 for aGVHD,[87,88] 1 for cGVHD[90]) presented data for 5 treatment regimens evaluating the occurrence of weaning from GVHD treatments and steroids as an outcome. There were no statistically significant differences identified for any of the comparisons during any of the follow-up periods. A summary of all study findings is provided in **Table 24**.

Table 24: Summary of results: Weaning from all GVHD treatments or steroids in the treatment of acute or chronic GVHD					
Author	**Follow-up**	**Outcome**	**Treatment regimen**	**Group incidence**	**OR**
Evaluation of treatment of aGVHD					
Alousi[87] (2009)	9 months	Discontinuation of all immunosuppressive therapy	ETN+Steroids	N = 46[a]	NA[a]
			MMF+Steroids	N = 45[a]	
Bolanos-Meade[88] (2014)	0–28 days	Reduction of dose or withholding medication	MMF+Steroids	58/113 (51%)	1.05 (0.63–1.77)
			PLB+Steroids	58/116 (50%)	
			MMF+Steroids	26/61 (43%)	0.87 (0.43–1.77)

Table 24: Summary of results: Weaning from all GVHD treatments or steroids in the treatment of acute or chronic GVHD

Author	Follow-up	Outcome	Treatment regimen	Group incidence	OR
Bolanos-Meade[88] (2014)	28–56 days	Reduction of dose or withholding medication	PLB+Steroids	29/63 (46%)	
Evaluation of treatment of aGVHD					
Martin[90] (2009)	2 years	End of systemic treatment	MMF+(CsA or TAC or SIR)+Steroid	15/74 (20%)[b]	1.05 (0.47–2.34)
			PLB+(CsA or TAC or SIR)+Steroid	15/77 (19%)	
Martin[90] (2009)	2 years	Withdrawal of prednisone	MMF+(CsA or TAC or SIR)+Steroid	30/74 (41%)[c]	0.91 (0.48–1.74)
			PLB+(CsA or TAC or SIR)+Steroid	33/77 (43%)	

[a] cumulative incidence: 35% (24–52) vs. 38% (21–49); not statistically significant
[b] cumulative incidence: 23% vs. 18%; 2-year adjusted HR = 1.51 (0.7–3.2)
[c] cumulative incidence: 44% vs. 41%; 2-year adjusted HR = 1.28 (0.8–2.1)
CsA = cyclosporin A; ETN = etanercept; MMF = mycophenolate mofetil; PLB = placebo; SIR = sirolimus; TAC = tacrolimus

4.8. Findings: Overall mortality

Six studies presented data for 9 treatment regimens evaluating overall mortality as an outcome; 3 studies evaluated the treatment of aGVHD[85,86,88] and 3 evaluated the treatment of cGVHD.[89-91]

4.8.1. Overall mortality in the evaluation of treatment of aGVHD

Table 25 provides a summary of findings related to overall mortality in studies for treatment of acute GVHD. Data from multiple follow-up durations was available from the included studies are shown. Inspection of between-group comparisons across the studies identified only one statistically significant difference which suggested that MMF+Steroids was associated with a reduced risk of death at 1 year compared to steroids alone (OR 0.40, 95% CI 0.24-0.68).

Table 25: Summary of results: Overall mortality in the evaluation of treatment of aGVHD				
Author	**Follow-up**	**Treatment regimen**	**Group incidence**	**OR**
Kennedy[85] (1985)	1 year	Steroids	26/39 (67%)	1.62 (0.64–4.08)
		CsA	21/38 (55%) *	
Kennedy[85](1985)	3 years	Steroids	28/39 (72%)	1.17 (0.44–3.12)
		CsA	26/38 (68%) *	
Cragg[86] (2000)	>1 year	ATG+Steroids	29/50 (58%)	1.51 (0.67–3.37)
		Steroids	22/46 (48%) *	
Bolanos-Meade[88] (2014)	56 days	MMF+Steroids	11/116 (9%)	0.85 (0.36–1.99)
		PLB+Steroids	13/119 (11%) *	
Bolanos-Meade[88] (2014)	6 months	MMF+Steroids	32/116 (28%)	1.04 (0.58–1.84)
		PLB+Steroids	32/119 (27%) *	
Bolanos-Meade[88] (2014)	1 year	MMF+Steroids	49/116 (42%)	**0.40 (0.24–0.68)**
		PLB+Steroids	77/119 (65%) *	
ATG = anti-thymocyte globulin; CsA = cyclosporin A; MMF = mycophenolate mofetil; PLB = placebo *denotes the reference group for each pairwise comparison				

4.8.2. Overall mortality in the evaluation of treatment of cGVHD

Table 26 provides a summary of findings related to overall mortality in studies for treatment of chronic GVHD. Data from variable follow-up durations was available from the included studies are shown. Inspection of between-group comparisons across the studies identified two statistically significant differences: the first suggested that MMF in combination with CsA+TAC or SIR+steroids significantly increased the odds of death at 4 years compared to CsA+TAC or SIR+Steroid alone (OR 2.40, 95% CI 1.03-5.60); the second suggested that AZA+Steroids increased the odds of death at ≥3.8 years compared to steroids alone in the treatment of clinical cGVHD (patients with subclinical cGVHD or cGVHD that progressed from subclinical to clinical were not included).

Table 26: Summary of results: Overall mortality in the evaluation of treatment of cGVHD				
Author	**Follow-up**	**Treatment regimen**	**Group incidence**	**OR**
Sullivan[89] (1988)	≥3.8 years	AZA+Steroids	33/63 (52%)	1.91 (0.94–3.90)
		PLB+Steroids	23/63 (37%) *	
Arora[91] (2001)	>2 years	THAL+CsA+Steroid	0/27 (0%)	NE[a]
		CsA+Steroid	0/27 (0%)	
Martin[90] (2009)	4 years	MMF+(CsA or TAC or SIR)+Steroid	19/74 (26%)	**2.40 (1.03–5.60)**
		PLB+(CsA or TAC or SIR)+Steroid	10/77 (13%) *	
[a] not estimable AZA = alemtuzumab; CsA = cyclosporin A; MMF = mycophenolate mofetil; PLB = placebo; SIR = sirolimus; TAC = tacrolimus; THAL = Thalidomide *denotes the reference group for each pairwise comparison				

4.9. Findings: Non-relapse mortality

Five studies presented data for 8 treatment regimens evaluating non-relapse mortality as an outcome; 2 studies evaluated the treatment of aGVHD[85,86] and 3 evaluated the treatment of cGVHD.[89-91]

4.9.1. Non-relapse mortality in the evaluation of treatment of aGVHD

Table 27 provides a summary of findings related to non-relapse mortality in studies for treatment of acute GVHD. Duration of follow-up for between group comparisons varied from 100 days to 4 years across studies. None of the comparisons between regimens reached conventional statistical significance based on inspection of odds ratios and their corresponding 95% confidence intervals.

Table 27: Summary of results: Non-relapse mortality in the evaluation of treatment of aGVHD

Author	Follow-up	Treatment regimen	Group incidence	OR
Kennedy[85] (1985)	100 days	Steroids	20/39 (51%)	1.80 (0.73–4.48)
		CsA	14/38 (37%) *	
Kennedy[85] (1985)	151–365 days	Steroids	4/39 (10%)	2.06 (0.35–11.96)
		CsA	2/38 (5%) *	
Kennedy[85] (1985)	1–3 years	Steroids	24/39 (62%)	2.20 (0.88–5.47)
		CsA	16/38 (42%) *	
Cragg[86] (2000)	4 years (median)	ATG+Steroids	27/50 (54%)	2.20 (0.97–5.01)
		Steroids	16/46 (35%) *	

ATG = anti-thymocyte globulin; CsA = cyclosporin A
*denotes the reference group for each pairwise comparison

4.9.2. Non-relapse mortality in the evaluation of treatment of cGVHD

Table 28 provides a summary of findings related to non-relapse mortality in studies for treatment of chronic GVHD. Data from variable follow-up durations was available. Inspection of between-group comparisons across the studies identified only one statistically significant difference which suggested that AZA+Steroids had a significantly higher likelihood of non-relapse mortality at ≥3.8 years compared to steroids alone (OR 3.45, 95% CI 1.44-8.25). The actuarial probabilities of non-relapse mortality for these interventions were also found to be significantly different in the original paper (p = 0.003).

Table 28: Summary of results: Non-relapse mortality in the evaluation of treatment of cGVHD

Author	Follow-up	Treatment regimen	Group incidence	OR
Sullivan[89] (1988)	≥3.8 years	AZA+Steroids	23/63 (37%)[a]	**3.45 (1.44–8.25)**
		PLB+Steroids	9/63 (14%) *	
Arora[91] (2001)	Unclear	THAL+CsA+Steroid	9/27 (33%)	1.19 (0.38–3.75)
		CsA+Steroid	8/27 (30%) *	
Martin[90] (2009)	4 years	MMF+(CsA or TAC or SIR)+Steroid	8/74 (14%)	1.75 (0.54–5.60)
		PLB+(CsA or TAC or SIR)+Steroid	5/77 (7%) *	

[a] actuarial probability: 40% vs. 21%; p = 0.003.AZA = alemtuzumab; CsA = cyclosporin A; MMF = mycophenolate mofetil; PLB = placebo; TAC = tacrolimus; THAL = thalidomide; SIR = sirolimus
*denotes the reference group for each pairwise comparison

4.10. Findings: Infection

Six studies presented data for 9 treatment regimens evaluating the incidence of infections; 3 studies evaluated the treatment of aGVHD[86-88] and 3 evaluated the treatment of cGVHD.[89-91] There were variations in outcome definition and follow-up time across studies.

4.10.1. Infection in the evaluation of treatment of aGVHD

Table 29 provides a summary of findings related to the incidence of infections in studies for treatment of acute GVHD. Data for multiple forms of infections and from variable follow-up durations across studies were available. Inspection of between-group comparisons across the studies identified two differences that achieved statistical significance: ATG+Steroids was associated with significantly more CMV infections than steroids alone by 48 months (OR 2.83, 95% CI 1.15-6.93), and ATG+Steroids was associated with significantly more pneumonitis than steroids alone over an unknown follow-up period (OR 3.18, 95% CI 1.33–7.64).[86]

Table 29: Summary of results: Incidence of infection in the evaluation of treatment of aGVHD

Author	Follow-up	Outcome definition	Treatment regimen	Group incidence	OR
Cragg[86] (2000)	4 years	CMV	ATG+Steroids	22/50 (44%)	**2.83 (1.15–6.93)**
			Steroids	10/46 (22%) *	
Cragg[86] (2000)	Unclear	Gram-negative infections	ATG+Steroids	20/50 (40%)	2.12 (0.88–5.13)
			Steroids	11/46 (24%) *	
Cragg[86] (2000)	Unclear	Fungal infection[a]	ATG+Steroids	7 /50 (14%)	2.33 (0.57–9.62)
			Steroids	3/46 (7%) *	
Cragg[86] (2000)	Unclear	Pneumonitis[b]	ATG+Steroids	25/50 (50%)	**3.18 (1.33–7.64)**
			Steroids	11/46 (24% *)	
Cragg[86] (2000)	Unclear	Epstein-Barr virus lymphoproliferative disease.	ATG+Steroids	3/50 (6%)	2.87 (0.29–28.64)
			Steroids	1/46 (2%) *	
Alousi[87] (2009)	9 months	Severe/life-threatening/fatal infection	ETN+Steroids	N = 46[c]	NA[c]
			MMF+Steroids	N = 45[c]	
Bolanos-Meade[88] (2014)	56 days	At least one infection	MMF+Steroids	81/116 (70%)	1.26 (0.73–2.18)
			PLB+Steroids	77/119 (65%) *	
Bolanos-Meade[88] (2014)	1 year	Severe/life-threatening infection	MMF+Steroids	N = 116[d]	NA[d]
			PLB+Steroids	N = 119[d]	
Bolanos-Meade[88] (2014)	Unclear	Epstein-Barr virus reactivation	MMF+Steroids	6/116 (5%)	1.57 (0.43–5.71)
			PLB+Steroids	4/119 (3%) *	

[a] includes *Candida* and *Aspergillus* infections
[b] includes pulmonary syndromes defined as infectious based on pathology, microbiology, and virology from bronchoalveolar lavage and non-infectious pneumonitis, where such studies excluded infectious pathogens
[c] cumulative incidence: 47% (33–63) vs. 44% (30–59); not statistically significant
[d] cumulative Incidence: 45% (35–54) vs. 43% (34–52); not statistically significant
*denotes the reference group for each pairwise comparison
ATG = anti-thymocyte globulin; CsA = cyclosporin A; ETN = etanercept; MMF = mycophenolate mofetil; PLB = placebo

4.10.2. Infection in the evaluation of treatment of cGVHD

Table 30 provides a summary of findings related to the incidence of infections in studies for treatment of chronic GVHD. Data for multiple forms of infections and from variable follow-up durations across studies were available. Inspection of between-group comparisons across the studies identified one statistically significant difference between therapies: Thalidomide+CsA+Steroids was associated with significantly fewer bacterial infections over 16 months compared to CsA+Steroids (1.7 infections per thousand person-days versus 3.0 infections per thousand person-days).[91]

Table 30: Summary of results: Incidence of infection in the evaluation of treatment of cGVHD

Author	Follow-up	Outcome definition	Treatment regimen	Group incidence	OR
Sullivan[89] (1988)	9 months	At least one infection	AZA+Steroids	39/63 (62%)	1.91 (0.94–3.87)
			PLB+Steroids	29/63 (46%) *	
Sullivan[89] (1988)	9 months	Varicella zoster	AZA+Steroids	25/63 (40%)	1.64 (0.78–3.46)
			PLB+Steroids	18/63 (29%) *	
Sullivan[89] (1988)	9 months	Bacteremia	AZA+Steroids	7/63 (77%)	1.84 (0.51–6.64)
			PLB+Steroids	4/63 (6%) *	
Sullivan[89] (1988)	9 months	Interstitial pneumonia	AZA+Steroids	9/63 (14%)	3.33 (0.86–12.95)
			PLB+Steroids	3/63 (5%) *	
Sullivan[89] (1988)	9 months	Non-interstitial pneumonia	AZA+Steroids	10/63 (16%)	1.00 (0.38–2.60)
			PLB+Steroids	10/63 (16%) *	
Arora[91] (2001)	16 months	All infections	THALCsA+Steroid	4.6/1000-person days[a]	NA[a]
			CsA+Steroid	6.6/1000-person days[a]	
Arora[91] (2001)	>16 months	Bacterial infection	THAL+CsA+Steroid	**1.7/1000-person days[b]**	**NA[b]**
			CsA+Steroid	**3.0/1000-person days[b]**	
Arora[91] (2001)	>16 months	Viral infection	THAL+CsA+Steroid	1.8/1000-person days[c]	NA[c]
			CsA+Steroid	2.2/1000-person days[c]	
Arora[91] (2001)	>16 months	Fungal infection	THAL+CsA+Steroid	1.1/1000-person days[d]	NA[d]
			CsA+Steroid	1.2/1000-person days[d]	
Martin[90] (2009)	4 years	Any infection[e]	MMF+(CsA or TAC or SIR)+Steroid	55/74 (74%)	0.43 (0.19–1.01)
			PLB+(CsA or TAC or SIR)+Steroid	67/77 (87%) *	
Martin[90] (2009)	4 years	Bacterial infection[e]	MMF+(CsA or TAC or SIR)+Steroid	18/74 (24%)	1.23 (0.57–2.63)
			PLB+(CsA or TAC or SIR)+Steroid	16/77 (21%) *	
Martin[90] (2009)	4 years	Viral infection[e]	MMF+(CsA or TAC or SIR)+Steroid	27/74 (36%)	1.06 (0.55–2.07)
			PLB+(CsA or TAC or SIR)+Steroid	27/77 (35%) *	
Martin[90] (2009)	4 years	Fungal infection[e]	MMF+(CsA or TAC or SIR)+Steroid	18/74 (24%)	0.71 (0.35–1.45)
			PLB+(CsA or TAC or SIR)+Steroid	24/77 (31%) *	

Table 30: Summary of results: Incidence of infection in the evaluation of treatment of cGVHD

Author	Follow-up	Outcome definition	Treatment regimen	Group incidence	OR
Martin[90] (2009)	4 years	Conjunctivitis[e]	MMF+(CsA or TAC or SIR)+Steroid	4/74 (5%)	0.34 (0.10–1.13)
			PLB+(CsA or TAC or SIR)+Steroid	1/77 (14%) *	

[a] p = 0.07
[b] p = 0.04
[c] not significant
[d] not significant
[e] up to 30 days after discontinuation of study drug
*denotes the reference group for each pairwise comparison

AZA = alemtuzumab; CsA = cyclosporin A; MMF = mycophenolate mofetil; PLB = placebo; TAC = tacrolimus; THAL = Thalidomide; SIR = sirolimus

5. DISCUSSION

5.1. Summary of main findings

Numerous drugs are used in a broad variety of single- and multi-agent regimens for both the prophylaxis and treatment of GVHD in HSCT. The current review found that comparative evidence from randomized trials is lacking for many comparisons of these regimens, which has led to considerable practice variation between both physicians and institutions. This systematic review of the evidence, incorporating network meta-analyses where possible, was conducted to address these gaps in relation to interventions for both the prophylaxis and treatment of GVHD. Totals of 32 and 7 randomized trials of GVHD prophylactic and treatment regimens, respectively, were identified and studied to inform analyses in this review. Included studies were published between 1979–2015, and considerable heterogeneity between study populations was noted regarding patient age, underlying hematologic disease, disease risk of relapse/mortality, and transplant donor status (i.e., related vs. unrelated, matched vs. unmatched). Overall 19 different GVHD prophylaxis regimens were assessed amongst the included 32 prophylaxis trials in a total of nearly 4,000 patients, while 10 treatment regimens were assessed in a total of 7 studies and 830 patients. Most comparisons between treatments in our analyses were based on indirect evidence only. Further discussion of findings from the completed work follows below. *We emphasize for readers that findings from network meta-analysis described in the report should be interpreted with considerable caution given the high degree of heterogeneity present between studies which could not be well addressed based upon the evidence structure as well as the mixed reporting and eligibility criteria of studies.*

5.1.1. Findings, GVHD Prophylaxis

In the interpretation of the NMAs presented in this report, it should be noted that the relative ranking of regimens by NMA demonstrates trend, and does not necessarily indicate sufficient evidence existed to confirm any of the regimens were significantly different from each other. In our NMA evaluating efficacy for the prevention of aGVHD, the 4 top-ranked regimens were not found to be statistically significantly different from each other and demonstrated little difference in their SUCRA values (MTX+SIR+TAC, SIR+TAC, CsA+MTX+MSCs, MTX+TAC+Steroids). When compared to the chosen standard of care for GVHD prevention—MTX+TAC—two regimens were found to be associated with significantly fewer cases of aGVHD (MTX+SIR+TAC and SIR+TAC), while five regimens were found to be associated with significantly more cases of aGVHD (CsA+MTX, CsA+Steroids, CsA, MTX, and MTX+Steroids).

Regarding the prevention of cGVHD, CsA+MTX+ATG+Steroids was ranked the highest, demonstrating significantly higher efficacy than all regimens other than HCQ+CsA; however, there were no other statistically significant differences amongst the 14 regimens compared. The paucity of significant findings may be a partially due to patient sample and outcome definition used in the analysis. Most studies did not report the number patients alive at 100 days post-transplant, the time after which cGVHD, by definition, occurs. Thus, the number of patients randomized per group was used as denominators in our analyses, which did not account for the competing risk of death prior to 100 days. Regimens with high 100-day mortality would have fewer patients alive to potentially develop cGVHD, and therefore, even with low true efficacy, they may appear as efficacious as regimens with more patients alive at day 100 and a greater true efficacy to prevent cGVHD. Our analysis would not have detected a difference between these regimens, despite their significant differences in true efficacy. As well, most studies reported overall cGVHD rather than extensive cGVHD. Overall cGVHD is less rigorously defined and encompasses a broader range of symptoms, including mild cases that may or may not require treatment. Extensive cGVHD is associated with an increased risk of death, and includes only cases with internal organ involvement that require treatment. Use of the less specific outcome definition "overall cGVHD" potentially increased heterogeneity in our analysis and may have diluted the actual effects of the regimens to the point where significant differences could not be detected.

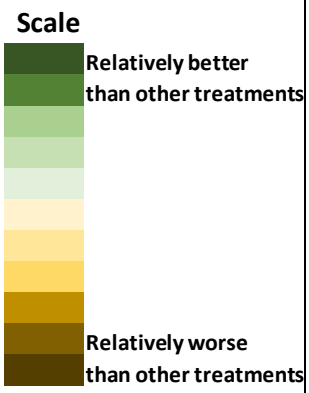
Few significant differences were found between the 15 prophylactic regimens evaluated with respect to overall mortality. At 100 days post-transplant, only ATG+MTX+Steroids, the highest ranked regimen, was associated with significantly fewer deaths than CsA and MTX+SIR+TAC, the lowest ranked regimens. No other regimens demonstrated significant differences. At 1 year post-transplant, the addition of mesenchymal stem cells (MSCs) to CsA+MTX was associated with a significantly lower risk of overall mortality than all other regimens in our NMA; however, at 5 years post-transplant, there were no significant differences between the regimens with respect to overall and non-relapse mortality. This promising therapy was also one of the higher ranked regimens for prevention of aGVHD in our analyses. MSCs can be potent inhibitors of the immune response through the secretion of cytokines[40]; however, the authors comment that further study is required to establish the characteristics of MSCs that are most suitable for the prevention of aGVHD. Also at 1 year of follow-up, the addition of ursodeoxycholic acid (UDCA) to CsA+MTX+Steroids was associated with significantly fewer deaths than CsA+MTX+Steroids alone. GVHD often impairs liver function and UDCA acts to improve liver function; thus, these findings suggest that the beneficial hepatic effects of UDCA may be sufficient to reduce death in HSCT patients. UDCA also was associated with a significant lower risk of non-relapse mortality at 1 and 10 years post-transplant (but not overall mortality at 10 years post-transplant), when compared with CsA+MTX+Steroids in the single study that evaluated that comparison[44]. This also suggests that the beneficial effects of UDCA on mortality are not a result of reduced risk of relapse but more systemic effects.

Network meta-analysis of disease relapse data did not detect significant differences between the 6 top-ranked regimens at 2–3 years post-transplant. Single-agent regimens (i.e., MTX and CsA) trended to be associated with reduced occurrence of relapse, likely because of their lesser immunosuppressive effects than multi-agent regimens. These two agents were the only regimens that were significantly different from the standard of care, MTX+TAC.

The ideal prophylaxis for GVHD would be highly efficacious at preventing both acute and chronic GVHD, while not increasing the risk of harms, such as relapse of underlying disease, non-relapse mortality, and infection. Our analyses demonstrate that this ideal regimen has yet to be identified. High levels of immunosuppression are necessary to prevent graft T cells from identifying the transplant recipient's somatic cells as foreign and attacking them. However, suppressing the graft cells also reduces their abilities to achieve engraftment and overwhelm the recipient's original tumour cells in a graft-versus-tumour effect. Without engraftment and a sufficient graft-versus-tumour effect, the recipient's underlying disease will return. As well, delay in the time to engraftment puts the recipient at a high risk of infection due to the lack of a functioning immune system. Thus, in the quest to reduce GVHD, the risks of disease relapse, infection, and potentially death are increased. In our analyses, 10 prophylactic regimens were common to the aGVHD prevention, disease relapse, and overall mortality networks. When these regimens are listed by their ranking to prevent aGVHD, and their ranking for disease relapse is reported alongside, we demonstrate that those regimens that are relatively better at preventing aGVHD tend to be relatively worse with respect to disease relapse (**Figure 18**). Similarly, the top- and bottom-most ranked regimens for aGVHD prevention (MTX+SIR+TAC and MTX, respectively) flip-flop in their rankings with respect to overall mortality. When regimen rankings are compared for the cGVHD outcome, the flip-flop of rankings is not as striking, likely because of greater uncertainty with respect to the ranking of regimens for cGVHD prevention, due to limitations with the cGVHD data (e.g., the use of overall cGVHD in the total number of patients randomized rather than extensive cGVHD in the number of patients alive at 100 days post-transplant in analyses) (**Figure 19**).

Figure 18: Comparison of relative SUCRA rankings of GVHD prophylaxis regimens in common between the aGVHD and disease relapse outcomes

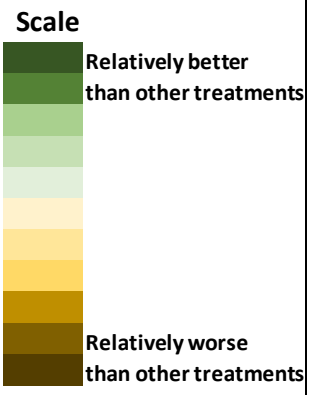
Regimen (ordered by aGVHD SUCRA ranking)	aGVHD ranking	Disease relapse ranking	100-d overall mortality ranking	1-year overall mortality ranking
MTX+SIR+TAC	1	10	10	10
SIR+TAC	2	7	6	9
MTX+TAC+Steroids	3	5	4	3
CsA+MTX+ATG+Steroids	4	8	1	2
MTX+TAC	5	9	7	6
CsA+MTX+Steroids	6	3	5	5
CsA+MTX	7	6	3	4
CsA+Steroids	8	4	8	7
CsA	9	2	9	8
MTX	10	1	2	1



ATG = anti-thymocyte globulin; CsA = cyclosporin A; MTX = methotrexate; SIR = sirolimus; TAC = tacrolimus

Figure 19: Comparison of relative SUCRA rankings of GVHD prophylaxis regimens in common between the cGVHD and disease relapse outcomes

Regimen (ordered by cGVHD SUCRA ranking)	cGVHD ranking	Relapse ranking	100-d overall mortality ranking	1-year overall mortality ranking
CsA+MTX+ATG+Steroids	1	9	1	2
HCQ+CsA	2	4	8	11
MTX+SIR+TAC	3	11	11	10
CsA+MTX	4	7	3	4
MTX+TAC	5	10	7	6
MTX+TAC+Steroids	6	6	4	3
CsA+MTX+Steroids	7	3	5	5
CsA	8	2	10	8
MTX	9	1	2	1
SIR+TAC	10	8	6	9
CsA+Steroids	11	5	9	7



ATG = anti-thymocyte globulin; CsA = cyclosporin A; HCQ = hydroxychloroquine; MTX = methotrexate; SIR = sirolimus; TAC = tacrolimus

Despite our analyses not identifying a clear optimal regimen, three generalizations can be made. First, single-agent regimens appear to be less efficacious to prevent GVHD that multi-agent regimens; however, single-agent regimens generally were more efficacious to prevent disease relapse. Second, regimens containing a calcineurin inhibitor (i.e., TAC or CsA) appeared to have greater efficacy to prevent aGVHD than regimens without; however, there was no significant difference between TAC- or CsA-containing regimens. Third, the addition of ATG to prophylaxis regimens may be worth further investigation. Overall, the two regimens that included ATG were consistently ranked amongst the highest with respect to both 100-day and 1-year overall mortality, and of these regimens, CsA+MTX+ATG+Steroids outperformed ATG+MTX+Steroids with respect to prevention of both acute and chronic GVHD. The risk of harms associated with ATG-containing regimens appears to be mixed. The risk of serious adverse events and the risk of CMV infection diagnosed by PCR were reported to be significantly increased for CsA+MTX+ATG+Steroids compared to CsA+MTX[50], but not the risks of renal or multi-organ failure[50], CMV reactivation[59], or infectious complications[45]. As well, ATG+MTX+Steroids had a significant protective effect against transient interstitial pneumonia, but not significant against fatal interstitial pneumonia, relative to MTX[81]. Further high quality research is needed regarding the use of ATG, potentially with exploration of ATG in combination with TAC in prophylaxis regimens.

The current evidence base for GVHD prophylaxis is broad but not deep, as while many regimens have been evaluated in a range of studies, there has been minimal replication of direct comparisons between active regimens. As a result, in our analyses, most networks contained many regimens but were comprised mainly of single-study connections, with no closed loops. The robustness of findings in such NMAs is reduced because the evidence upon which most comparisons are based is indirect (i.e., there are no head-to-head trials). As well, meta-regression to evaluate potentially significant covariates (e.g., underlying disease type) is not possible with sparse networks. To deepen the evidence base, future studies in GVHD prophylaxis should compare regimens to standard therapy, such as MTX+TAC or MTX+CsA+ATG+Steroids. As well, future studies should stratify patients on potentially significant covariates such as underlying disease, patient age, and donor factors to ensure between-group homogeneity. Finally, to improve future syntheses in the prevention of cGVHD, future studies should standardize methods and report extensive cGVHD in patients who are alive at 100 days, with a follow-up of 2 years post-transplant. Such improvements in choice of comparators and planning and reporting of analyses will enable the performance of more robust analyses. As noted earlier, findings from the network meta-analyses presented in this review should be viewed cautiously given the high degree of heterogeneity present between studies.

5.1.2. Findings, GVHD Treatment

The 7 studies included in the review of GVHD treatment evaluated either the treatment of aGVHD (4 studies, 499 patients, 5 interventions) or cGVHD (3 studies, 331 patients, 6 interventions). Due to the sparseness of the evidence base; heterogeneity in terms of type of GVHD being treated, outcome definitions, and follow-up periods; and fragmentation of evidence networks, NMAs could not be conducted and outcomes were summarized narratively.

In the treatment of aGVHD, MMF+Steroids significantly outperformed ETN+Steroids after 28 and 56 days of treatment in one study; however, there was no difference in the discontinuation of immunosuppressive therapy or in the risk of severe, life-threatening, or fatal infections after 9 months[87]. MMF+Steroids also demonstrated significantly reduced risk of mortality at 1 year compared to steroids alone[88] in one study, but no differences were found with respect to infection outcomes. Exploratory NMAs of the limited evidence base demonstrated no significant difference in treatment efficacy between any of the included regimens; however, MMF+Steroids appeared to be associated with significantly reduced overall mortality compared to ATG+Steroids or steroids alone, and CsA appeared to be associated with significantly reduced non-relapse mortality compared to ATG+Steroids. Treatment with MMF+Steroids shows some promise and larger clinical trials of this regimen are warranted.

Regarding the treatment of cGVHD, only steroids alone demonstrated significantly higher response to treatment than AZA+Steroids.

In addition to synthesis of existing evidence, the current review also provided an opportunity for identification of potential approaches to improving the design of future trials of interventions for prophylaxis and treatment of GVHD. Based upon discussions with our collaborating clinical experts, several considerations of this nature were identified. First, future studies of GVHD prophylaxis should consider stratification of patients for competing factors, such as patient age and underlying disease, as well as donor factors. In terms of comparator selection, prophylactic regimens should be compared with standard therapies such as MTX+TAC or MTX+CSA+ATG. In terms of endpoint selection, the evaluation of cGVHD occurrence should include an assessment of the incidence of extensive cGVHD (or use existing grading scales) in patients who are alive at 100 days, with a follow-up of 2 years post-transplant. Regarding future studies for interventions of GVHD treatment, patients should be evaluated for standardized outcome responses at consistent follow-up times, and be compared with Steroids alone or MMF+Steroids. These changes will help to further the future development of informative clinical trials to assess interventions for the prevention and treatment of GVHD in patients undergoing allogeneic HSCT.

5.2. Strengths and limitations

There are several limitations to be noted regarding the current review regarding deficiencies related to the evidence available and its implications for analyses performed in this review. First, there was a broad time frame during which the included studies were performed, with publication dates ranging from the late 1970s until 2015. While clinical experts could not identify specific aspects of care that have changed over time, this remains an important consideration (including changes in cell source and conditioning regimen intensity). Diversity in patients over time remains a factor, in addition to the clinical heterogeneity identified across studies in terms of the patient populations enrolled. This represents an ongoing difficulty of evidence syntheses in such transplant populations given the challenges of enrolling such patients. Second, as described throughout the review, the high number of regimens to be compared and the relatively small number of related studies resulted in network structures which consisted of high numbers of comparisons informed by indirect evidence only (often of a compound nature) and networks formed primarily of single study connections between interventions; while fixed effects NMAs were thus a primary focus for our analyses, we have also presented all findings from random effects analyses in appendices to the report to respect the extent of heterogeneity present between studies and to provide readers with this information. Several additional limitations of the studies themselves were also noted during the review; this included high levels of 'unclear' judgements during risk of bias evaluation due to missing information, poor reporting of harms endpoints, and inclusion of certain subgroups of patients of limited relevance to current care (including CML patients in some cases, as noted earlier). The failure of studies to report subgroup findings for specific types of patients paired with the limited geometry of the evidence networks also precluded the use of subgroup analysis or meta-regression to formally assess the impact of clinical heterogeneity in our analyses; this is unfortunate, as variation in event rates across studies amongst the different interventions was commonly observed, and further increases the likelihood of the presence of clinical heterogeneity amongst the included studies. Lastly, it is currently unclear to what degree the underlying biology of graft versus host disease may have changed over time; changes in relation to HSCT administration (including aspects such as the involvement of related versus unrelated donors) may thus play an important role in GVHD prophylaxis and treatment, however this is not yet well understood.

There are also several strengths of note for the current review. To our knowledge, the current review represents the most inclusive knowledge synthesis at this time prepared to evaluate the current state of evidence for different interventions for prophylaxis as well as treatment against GVHD in this population; while one past systematic review was reported in 2014 which involved indirect treatment comparisons, the review was limited by several methodologic choices, a failure to describe and assess clinical heterogeneity

of the study populations and methods it included, and it did not address the aspect of treatment of GVHD[22]. The current review lends considerable attention to these important details to clarify the evidence base for readers. Several additional clinically relevant endpoint measures not addressed in the past review (including mortality and relapse) were also assessed in the current review. Robust Bayesian methods for network meta-analysis were employed for all analyses, and reporting was guided by the PRISMA extension statement for network meta-analysis. We also identified additional trials of relevance not included in the 2014 systematic review, as well as duplicate publications that required careful evaluation.

5.3. Conclusions

The current review found that comparative evidence from randomized trials is lacking for many comparisons of single-agent and multi-agent regimens currently used for prophylaxis and treatment of graft versus host disease in patients undergoing allogeneic HSCT. This systematic review of the evidence, incorporating network meta-analyses where possible, was conducted to address these gaps. ***The following key points for clinical practice were identified in this review***:

- Prophylactic regimens for GVHD that demonstrated increased benefits (i.e., a reduced risk of acute or chronic GVHD) were associated with increased harms (i.e., increased risk of relapse, infection, or mortality). These regimens included the following combinations: MTX+TAC, MTX+SIR+TAC, SIR+TAC, and CsA+MTX+ATG+Steroids.
- MTX+TAC should be considered the standard of care for GVHD prophylaxis to which other regimens are compared in future studies.
- While the addition of ATG in prophylaxis appears to reduce cGVHD, it has not been used in combination with TAC. Use of ATG does not appear to increase overall mortality or relapse but can increase CMV reactivation and possibly other infections.
- In the treatment of aGVHD, response rates with the use of MMF+Steroids may be improved; however, overall survival may be compromised. This regimen warrants further study.
- No therapy currently appears superior to steroids alone for the treatment of cGVHD.

In collaboration with participating clinical experts, suggestions to further the evidence base in terms of the conduct and reporting of future trials were put forth. Selection of prophylactic and treatment regimens for GVHD remains complex, and further development of the evidence base can provide additional valuable information for future reviews of the literature.

<u>Notes regarding this work</u>:

- Funding for this work was provided by a team grant from the Canadian Institutes of Health Research and the Drug Safety and Effectiveness Network.
- BH is funded by a New Investigator Award from the Canadian Institutes of Health Research and the Drug Safety and Effectiveness Network.

Reference List

(1) Copeland E. Hematopoietic stem cell transplantation. New England Journal of Medicine 354, 1813-1826. 2006.

(2) Leger C and Nevill T. Hematopoietic stem cell transplantation: a primer for the primary care physician. CMAJ 170[10], 1569-1577. 2004.

(3) Lee S, Joffe S, Artz A, Champlin R, Davies S, and et al. Individual physician practice variation in hematopoietic cell transplantation. J Clin Oncol 26[13], 2162-2170. 2008.

(4) Lee S, Astigarraga C, Eapen M, Artz A, Davies S, and et al. Variation in supportive care practices in hematopoietic cell transplantation. Biol Blood Marrow Transplant 14[11], 1231-1238. 2008.

(5) Rasheed W, Ghavamzadeh A, Hamladji R, Ben Othman T, and et al. Hematopoietic stem cell transplantation practice variation among centers in the Eastern Mediterranean Region (EMRO): Eastern Mediterranean Bone Marrow Transplantation (EMBMT) group survey. Hematol Oncol Stem Cell Ther 6[1], 14-19. 2013.

(6) Pidala J, Lee S, Quinn G, Jim H, Kim J, and Anasetti C. Variation in management of immune suppression after allogeneic hematopoietic cell transplantation. Biol Blood Marrow Transplant 17[10], 1528-1536. 2011.

(7) Ferrara J, Levine J, Reddy P, and Holler E. Graft versus host disease. Lancet 373, 1550-1561. 2009.

(8) Przepiorka D, Weisdorf D, Martin P et al. 1994 Consensus Conference on Acute GVHD Grading. 1995;15:825-828.

(9) Mohty M. Mechanisms of action of antithymocyte globulin: T-cell depletion and beyond. Leukemia 21[7], 1387-1394. 2007.

(10) Leventhal J, Huang Y, Xu H, Goode I, and Ildstad S. Novel regulatory therapies for prevention of Graft-versus-host disease. BMC Medicine 10[48]. 2012.

(11) Faivre S, Kroemer G, and Raymond E. Current development of mTOR inhibitors as anticancer agents. Nat Rev Drug Discov 5[8], 671-688. 2006.

(12) Arora A and Scholar E. Role of Tyrosine Kinase Inhibitors in Cancer Therapy. Journal of Pharmacology and Experimental Therapeutics 315[3], 971-979. 2005.

(13) Johnston A, Gudjonsson J, Sigmundsdottir H Runar Ludviksson B, and et al. The anti-inflammatory action of methotrexate is not mediated by lymphocyte apoptosis, but by the suppression of activation and adhesion molecules. Clinical Immunology 114[2], 154-163. 2005.

(14) Allison A and Euqui M. Mycophenolate mofetil and its mechanisms of action. Immunopharmacology 47[2-3], 85-118. 2000.

(15) Maltzman J and Koretzky G. Azathioprine: Old Drug, New Actions. Journal of Clinical Investigation 111, 1122-1124. 2003.

(16) Liu J, Farmer J, Lane W, Friedman J, Weissman I, and Schreiber S. Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. Cell 66[4], 807-815. 1991.

(17) Coutinho A and Chapman K. The anti-inflammatory and immunosuppressive effects of glucocorticoids, recent developments and mechanistic insights. Mol Cell Endocrinol 335[1], 2-13. 2011.

(18) Keating A. Mesenchymal stromal cells: new directions. Stem Cell 10[6], 709-716. 2012.

(19) Gilman A, Chan K, Mogul M, Morris C, and et al. Hydroxychloroquine for the treatment of chronic graft-versus-host disease. Biol Blood Marrow Transplant 6[3A], 327-334. 2000.

(20) Klassen J. The role of photopheresis in the treatment of graft-versus-host disease. Curr Oncol 17[2], 55-58. 2010.

(21) Flowers M and Martin P. How we treat chronic graft-versus-host disease. Blood 125[4], 606-615. 2015.

(22) Ziakis P, Zervou F, Zacharioudakis I, and Mylonakis E. Graft-Versus-Host Disease Prophylaxis after Transplantation: A Network Meta-Analysis. PLOS ONE DOI:10.1371/journal.pone.0114735. 2014.

(23) Hutton B, Salanti G, Caldwell D, Schmid C, Chaimani A, Cameron C, Ioannidis J, and et al. The PRISMA Extension Statement for Reporting of Systematic Reviews Incorporating Network Meta-Analyses of Healthcare Interventions: Checklist and Explanations. Annals of Internal Medicine In press. 2015.

(24) Sampson M, McGowan J, Cogo E, Grimshaw J, Moher D, and Lefebvre C. An evidence-based practice guideline for the peer review of electronic search strategies. Journal of Clinical Epidemiology 62[9], 944-952. 2009.

(25) Kahl C, Storer B, Sandmaier B, Mielcarek M, and et al. Relapse risk in patients with malignant diseases given allogeneic hematopoietic cell transplantation after non-myeloablative conditioning. Blood 110[7], 2744-2748. 2007.

(26) Armand P, Gibson CJ, Cutler C, Ho V, and et al. A disease risk index for patients undergoing allogeneic stem cell transplantation. Blood 120[4], 9905-913. 2012.

(27) Higgins J, Altman D, Gotzsche P, and et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 343. 2011.

(28) Dias S, Welton N, Sutton A, and Ades A. NICE DSU Technical Support Document 2:A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials. http://www.nicedsu.org.uk . 2011.

(29) Dias S, Welton N, Sutton A, Caldwell D, Lu G, and Ades A. NICE DSU Technical Support Document 4: Inconsistency in Networks of Evidence Based on Randomised Controlled Trials. http://www.nicedsu.org.uk . 2011.

(30) Cameron C, Coyle D, Richter T, and et al. Systematic review and network meta-analysis comparing antithrombotic agents for the prevention of stroke and major bleeding in

patients with atrial fibrillation. BMJ Open 4[6]. 2014.

(31) Salanti G, Ades A, and Ioannidis J. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. Journal of Clinical Epidemiology 64[2], 163-171. 2011.

(32) Salanti G. Indirect and mixed-treatment comparison, network, or multipletreatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. Research Synthesis Methods 3[2]. 2012.

(33) Sorror M, Storer B, Sandmaier B, Maloney D, Chauncey T, and et al. Hematopoietic cell transplantation-comorbidity index and Karnofsky performance status are independent predictors of morbidity and mortality after allogeneic nonmyeloablative hematopoietic cell transplantation. Cancer 112[9], 1992-2001. 2008.

(34) Vaughn J, Storer B, Armand P, Raimondi R, Gibson C, and et al. Design and Validation of an Augmented Hematopoietic Cell Transplantation-Comorbidity Index Comprising Pretransplant Ferritin, Albumin, and Platelet Count for Prediction of Outcomes after Allogeneic Transplantation. Biol Blood Marrow Transplant pii: S1083-8791(15)00231-1. doi: 10.1016/j.bbmt.2015.04.002. 2015.

(35) Nakaya A, Mori T, Tanaka M, Tomita M, and et al. Does the hematopoietic cell transplantation specific comorbidity index (HCT-CI) predict transplantation outcomes? A prospective multicenter validation study of the Kanto Study Group for Cell Therapy. Biol Blood Marrow Transplant 20[10], 1553-1559. 2014.

(36) Guilfoyle R, Demers A, Bredeson C, Richardson E, and et al. Performance status, but not the hematopoietic cell transplantation comorbidity index (HCT-CI), predicts mortality at a Canadian transplant center. Bone Marrow Transplant 43[2], 133-139. 2009.

(37) Glass B, Hasenkamp J, Wulf G, Dreger P, and et al. Rituximab after lymphoma-directed conditioning and allogeneic stem-cell transplantation for relapsed and refractory aggressive non-Hodgkin lymphoma (DSHNHL R3): an open-label, randomised, phase 2 trial. Lancett Oncology 15, 757-766. 2014.

(38) Cutler C, Logan BR, Nakamura R et al. Tacrolimus/sirolimus vs. tacrolimus/methotrexate for graft-vs.-host disease prophylaxis after HLA-matched, related donor hematopoietic stem cell transplantation: Results of blood and marrow transplant clinical trials network trial 0402. 2012;120:739.

(39) Martin P, Furlong T, Rowley S, Pergam S, and et al. Evaluation of Oral Beclomethasone Dipropionatefor Prevention of Acute Graft-versus-Host Disease. Biol Blood Marrow Transplant 18, 922-929. 2012.

(40) Kuzmina L, Petinati N, Shipounova I, Sats N, and et al. Analysis of multipotent mesenchymal stromal cells used foracute graft-versus-host disease prophylaxis. European Journal of Haematology 96[4], 425-434. 2015.

(41) Shipounova IN, Petinati NA, Bigildeev AE et al. Analysis of results of acute graft-versus-host disease prophylaxis with donor multipotent mesenchymal stromal cells in patients with hemoblastoses after allogeneic bone marrow transplantation. 2014;79:1363-1370.

(42) Kuzmina L, Parovichnikova E, Drokov M et al. Multipotent mesenchimal stromal cells as graft-versushost disease prophylaxis: A prospective randomised clinical trial. 2015;50:S184.
(43) Cutler C, Logan B, Nakamura R, Johnston L, and et al. Tacrolimus/sirolimus vs tacrolimus/methotrexate as GVHD prophylaxisafter matched, related donor allogeneic HCT. Blood 124[8], 1372-1377. 2014.

(44) Ruutu T, Juvonen E, Remberger M et al. Improved survival with ursodeoxycholic acid prophylaxis in allogeneic stem cell transplantation: long-term follow-up of a randomized study. 2014;20:135-138.
(45) Bonifazi F, Solano C, Wolschke C, and Patriarca F. Prevention of Chronic GvHD after HLA-Identical Sibling Peripheral Hematopietic Stem Cell Transplantation with or without Anti-Lymphocyte Globulin (ATG). Results from a Prospective, Multicenter Randomized Phase III Trial. Blood 124[21], 37. 2014.

(46) Armand P, Kim H, Sainvil M, Bachanova V, and et al. The Addition Of Sirolimus To The Gvhd Prophylaxis Regimen In Reduced Intensity Allogeneic Stem Cell Transplantation For Lymphoma:A Multicenter Randomized Trial. Blood 122[21], 704. 2013.

(47) Pidala J, Kim J, Jim H, Kharfan-Dabaja M, and et al. A randomized phase II study to evaluate tacrolimus in combination with sirolimusor methotrexate after allogeneic hematopoietic cell transplantation. Haematologica 97[12], 1882-1889. 2012.

(48) Pulsipher M, Langholz B, Wall D, Schultz K, and et al. The addition of sirolimus to tacrolimus/methotrexate GVHD prophylaxis inchildren with ALL: a phase 3 Children's Oncology Group/Pediatric Blood andMarrow Transplant Consortium trial. Blood 123[13], 2017-2025. 2014.

(49) Hamilton B, Bolwell B, Kalaycio M, Rybicki L, and et al. Long-Term Follow-up of a Randomized Trial Comparing Cyclosporine and Short Course Methotrexate with Cyclosporine and Mycophenolate Mofetil for Graft Versus Host Disease Prophylaxis in Myeloablative HLA-Identical Sibling Hematopoietic Cell Transplantation. Blood 120, Abstract 3073. 2012.

(50) Socie G, Schmoor C, Bethge W, Ottinger H, and et al. Chronic graft-versus-host disease: long-term results from a randomized trial ongraft-versus-host disease prophylaxis with or without anti-T-cell globulin ATG-Fresenius. Blood 117[23], 6375-6382. 2011.

(51) Finke J, Schmoor C, Bethge W, Ottinger H, and et al. Randomized Trial on GvHD Prophylaxis with or without Anti-T-Cell Globulin ATG-Fresenius (ATG-F) In Allogeneic Hematopoietic Cell Transplantation From Matched Unrelated Donors: Final Results and Analysis of Prognostic Factors. Blood, Abstract 1249 116. 2010.

(52) Finke J, Bethge W, Schmoor C, Ottinger H, and et al. Standard graft-versus-host disease prophylaxis with or without anti-T-cell globulin in haematopoietic cell transplantation from matched unrelated donors: a randomised, open-label, multicentre phase 3 trial. Lancet Oncology 10, 855-864. 2009.

(53) Finke J, Schmoor C, Bethge W, Ottinger H, and et al. Prognostic Factors Affecting Outcome after AllogeneicTransplantation for Hematological Malignancies fromUnrelated Donors: Results from a Randomized Trial. Biol Blood Marrow Transplant 18, 1716-1726. 2012.

(54) Socie G, Schmoor C, Bethge W, Ottinger H, and et al. Chronic Graft-Versus-Host Disease: Lessons From a Randomized Trial on GvHD Prophylaxis with or without Anti-T-Cell Globulin ATG-Fresenius (ATG-F) In Allogeneic Hematopoietic Cell Transplantation (HSCT) From Matched Unrelated Donors. Blood, Abstract 212 116. 2010.

(55) Perkins J, Field T, Kim J, Kharfan-Dabaja M, and et al. A Randomized Phase II Trial Comparing Tacrolimusand Mycophenolate Mofetil to Tacrolimusand Methotrexate for Acute Graft-versus-HostDisease Prophylaxis. Biol Blood Marrow Transplant 16, 937-947. 2010.

(56) Fong T, Trinkaus K, Adkins D, Vij R, and et al. A Randomized Double-Blind Trial ofHydroxychloroquine for the Prevention of ChronicGraft-versus-Host Disease after Allogeneic PeripheralBlood Stem Cell Transplantation. Biol Blood Marrow Transplant 13, 1201-1206. 2007.

(57) Lee K, Choi S, Lee J, Kim S, and et al. Cyclosporine alonevscyclosporine plus methotrexate for post-transplantimmunosuppression after HLA-identical sibling bone marrowtransplantation: a randomized prospective study. Bone Marrow Transplant 34, 627-636. 2004.

(58) Hiraoka A, Ohashi Y, Okamoto S, Moriyama Y, and et al. Phase III study comparing tacrolimus (FK506) with cyclosporine forgraft-versus-host disease prophylaxis after allogeneic bone marrow transplantation. Bone Marrow Transplant 28, 181-185. 2001.

(59) Bacigalupo A, Lamparelli T, Bruzzi P, Guidi S, and et al. Antithymocyte globulin for graft-versus-host disease prophylaxis in transplantsfrom unrelated donors: 2 randomized studies fromGruppo Italiano Trapianti Midollo Osseo (GITMO). Blood 98, 2942-2947. 2001.

(60) Chao N, Snyder D, Jain M, Wong R, and et al. Equivalence of 2 Effective Graft Versus Host Disease Prophylaxis Regimens: Results of a Prospective Double-Blind Randomized Trial. Biol Blood Marrow Transplant 6[3], 254-261. 2000.

(61) Deeg H, Lin D, Leisenring W, Boeckh M, and et al. Cyclosporine or Cyclosporine Plus Methylprednisolone for Prophylaxis ofGraft-Versus-Host Disease: A Prospective, Randomized Trial. Blood 89[10], 3880-3887. 1997.

(62) Nash R, Antin J, Karanes C, Fay J, and et al. Phase 3 study comparing methotrexate and tacrolimus with methotrexate andcyclosporine for prophylaxis of acute graft-versus-host disease after marrowtransplantation from unrelated donors. Blood 96[6], 2062-2068. 2000.

(63) Ruutu T, Volin L, Parkkali T, Juvonen E, and Elonen E. Cyclosporine, methotrexate, and methylprednisolone compared with cyclosporineand methotrexate for the prevention of graft-versus-host disease in bone marrowtransplantation from HLA-identical sibling donor: a prospective randomized study. Blood 96[7], 2391-2398. 2000.

(64) Ross M, Schmidt G, Niland J, Amylon M, and et al. Cyclosporine, methotrexate, and prednisone compared with cyclosporine and prednisone for prevention of acute graft-vs.-host disease: effect on chronic graft-vs.-host disease and long-term survival. Biol Blood Marrow Transplant 5[5], 285-291. 1999.

(65) Ratanatharathtorn V, Nash R, Przepiorka P, Devine S, and et al. Phase III Study Comparing Methotrexate and Tacrolimus (Prograf, FK506)With Methotrexate and Cyclosporine for Graft-Versus-Host DiseaseProphylaxis After HLA-Identical Sibling Bone Marrow Transplantation.

Blood 92[7], 2303-2314. 1998.

(66) Wingard J, Nash R, Przepiorka D, Klein J, and et al. Relationship of tacrolimus (FK506) whole bloodconcentrations and efficacy and safety after HLA-identical sibling bone marrow transplantation. Biol Blood Marrow Transplant 4, 157-163. 1998.

(67) Zikos P, Van Lint M, Frassoni F, Lamparelli T, and et al. Low Transplant Mortality in Allogeneic Bone Marrow Transplantation for AcuteMyeloid Leukemia: A Randomized Study of Low-Dose Cyclosporin VersusLow-Dose Cyclosporin and Low-Dose Methotrexate. Blood 91[9], 3503-3508. 1998.

(68) Chao N, Parker P, Niland J, Wong R, and et al. Paradoxical effect of thalidomide prophylaxis on chronic graft versus host disease. Biol Blood Marrow Transplant 2, 86-92. 1996.

(69) Mrsic M, Labar B, Bogdanic V, Nemet D, and et al. Combination of cyclosporin and methotrexate for prophylaxis of acute graft-versus-host disease after allogeneic bone marrow transplantation for leukemia. Bone Marrow Transplant 6[2], 137-141. 1990.

(70) Torres A, Martinez F, Gomez P, Herrera C, and et al. Cyclosporin A versus methotrexate, followed by rescue with folinic acid as prophylaxis of acute graft-versus-host disease after bone marrow transplantation. Blut. 58[2], 63-68. 1989.

(71) Backman L, Ringden O, Tollemar J, and Lonnqvist B. An increased risk of relapse in cyclosporin-treated compared with methotrexate-treated patients: long-term follow-up of a randomized trial. Bone Marrow Transplantation 3, 463-471. 1988.

(72) Forman S, Blume K, Krance R, Miner P, and et al. A prospective randomized study of acute graft versus host disease in 107 patients with leukemia: methotrexate/prednisone versus cyclosporine/prednisone. Transplantation Proceedings 19[1], 2605-2607. 1987.

(73) Storb R, Deeg HJ, Whitehead J et al. Methotrexate and cyclosporine compared with cyclosporine alone for prophylaxis of acute graft versus host disease after marrow transplantation for leukemia. 1986;314:729-735.

(74) Sorror ML, Leisenring W, Deeg HJ, Martin PJ, Storb R. Twenty-year follow-up of a controlled trial comparing a combination of methotrexate plus cyclosporine with cyclosporine alone for prophylaxis of graft-versus-host disease in patients administered HLA-identical marrow grafts for leukemia. 2005;11:814-815.

(75) Storb R, Deeg HJ, Whitehead J et al. Marrow transplantation for leukemia and aplastic anemia: two controlled trials of a combination of methotrexate and cyclosporine v cyclosporine alone or methotrexate alone for prophylaxis of acute graft-v-host disease. 1987;19:2608-2613.

(76) Storb R, Deeg HJ, Pepe M et al. Methotrexate and cyclosporine versus cyclosporine alone for prophylaxis of graft-versus-host disease in patients given HLA-identical marrow grafts for leukemia: long-term follow-up of a controlled trial. 1989;73:1729-1734.

(77) Irle C, Deeg HJ, Buckner CD et al. Marrow transplantation for leukemia following fractionated total body irradiation. A comparative trial of methotrexate and cyclosporine. 1985;9:1255-1261.

(78) Storb R, Deeg HJ, Thomas ED. Preliminary results of prospective randomized trials comparing methotrexate and cyclosporine for prophylaxis of graft-vs.host-disease after HLA-identical marrow transplantations. 1983;15:2620-2623.

(79) Deeg HJ, Storb R, Thomas ED et al. Cyclosporine as prophylaxis for graft-versus-host disease: a randomized study in patients undergoing marrow transplantation for acute nonlymphoblastic leukemia. 1985;65:1325-1334.

(80) Storb R, Deeg HJ, Thomas ED et al. Marrow transplantation for chronic myelocytic leukemia: a controlled trial of cyclosporine versus methotrexate for prophylaxis of graft-versus-host disease. 1985;66:698-702.

(81) Weiden P, Doney K, Storb R, and Thomas ED. Antihuman thymocyte globulin for prophylaxis of graft versus host disease: a randomized trial in patients with leukemia treated with HLA-identical sibling marrow grafts. Transplantation 27[4], 227-230. 1979.

(82) Glucksberg H, Storb R, Fefer A et al. Clinical manifestations of graft-versus-host disease in human recipients of marrow from HL-A-matched sibling donors. 1974;18:295-304.

(83) Thomas ED, Storb R, Clift RA et al. Bone-marrow transplantation (second of two parts). 1975;292:895-902.

(84) Deeg HJ, Flowers ME, Leisenring W, Appelbaum FR, Martin PJ, Storb RF. Cyclosporine (CSP)or CSP plus methylprednisolone for graft-versus-host disease prophylaxis in patients with high-risk lymphohemopoietic malignancies: long-term follow-up of a randomized trial. 2000;96:1194-1195.

(85) Kennedy M, Deeg J, Storb R, Doney K, and et al. Treatment of Acute Graft Versus Host Disease After AIIogeneic Marrow Transplantation: Randomized Study Comparing Corticosteroids and Cyclosporin. American Journal of Medicine 78, 978-983. 1985.

(86) Cragg L, Blazar B, DeFor T, Kolatker N, Miller W, and et al. A Randomized Trial Comparing Prednisone WithAntithymocyte Globulin/Prednisone as an InitialSystemic Therapy for Moderately Severe Acute Graft-Versus-Host Disease. Biol Blood Marrow Transplant 6, 441-447. 2000.

(87) Alousi A, Weisdorf D, Logan B, Bolanos-Meade J, and et al. Etanercept, mycophenolate, denileukin, or pentostatin plus corticosteroids foracute graft-versus-host disease: a randomized phase 2 trial from the Blood andMarrow Transplant Clinical Trials Network. Blood 114, 511-517. 2009.

(88) Bolanos-Meade J, Logan B, Alousi A, Antin J, and et al. Phase 3 clinical trial of steroids/mycophenolate mofetil vs steroids/placebo as therapy for acute GVHD: BMT CTN 0802. Blood 124[22], 3221-3227. 2014.

(89) Sullivan K, Witherspoon R, Storb R, Deeg H, and et al. Alternating-day cyclosporine and prednisone for treatment of high-risk chronic graft-v-host disease. Blood 72[2], 555-561. 1988.

(90) Martin P, Storer B, Rowley S, Flowers M, and et al. Evaluation of mycophenolate mofetil for initial treatment of chronic graft-versus-host disease. Blood 113[21], 5074-5082. 2009.

(91) Arora M, Wagner J, Davies M, and Blazar B. Randomized clinical trial of thalidomide, cyclosporine, and prednisone versus cyclosporine and prednisone as initial therapy for chronic graft-versus-host disease. Biol Blood Marrow Transplant 7[5], 265-273. 2001.

6. REPORT APPENDICES

Cite this report as: Wolfe D, Yazdi F, Hutton B, Moher D, Bredeson C, Kekre N, Allan D. Interventions for the Prophylaxis and Treatment of Graft versus Host Disease in Patients Undergoing Allogeneic Stem Cell Transplant (2017).

Appendix 1. Literature Search Strategies

Appendix 2. Additional Characteristics of Included Studies

Appendix 3. Risk of Bias Assessments of Included Studies

Appendix 4. Summary of results for subgroups presented in studies evaluating treatment of GVHD

Appendix 5. Findings from Secondary Analyses

Appendix 6. Summary of Inconsistency Analyses

Appendix 7. Supplemental network meta-analyses, prevention of chronic GVHD

Appendix 8. Prophylaxis of GVHD - Findings from Random Effects NMAs

Appendix 9. PRISMA NMA Checklist

6.1. Appendix 1: Literature search strategies

Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1946 to Present> **June 12, 2013; updated June 2015**

---

1 Hematopoietic Stem Cell Transplantation/
2 ((h?ematopoietic adj3 transplant$) or hsct$).tw.
3 peripheral blood cell transplant$.tw.
4 peripheral blood stem cell transplant$.tw.
5 Bone Marrow Transplantation/
6 bone marrow transplant$.tw.
7 STEM CELL TRANSPLANTATION/
8 stem cell transplant$.tw.
9 stem cell therap$.tw.
10 PERIPHERAL BLOOD STEM CELL TRANSPLANTATION/
11 peripheral stem cell transplant$.tw.
12 H?ematopoietic peripheral blood stem cell transplant$.tw.
13 or/1-12
14 randomized controlled trial.pt.
15 controlled clinical trial.pt.
16 randomized.ab.
17 placebo.ab.
18 clinical trials as topic.sh.
19 randomly.ab.
20 trial.ti.
21 or/14-20
22 exp animals/ not humans.sh.
23 21 not 22
24 13 and 23

6.2. Appendix 2: Additional Characteristics of included studies

Distribution of underlying disease across studies of regimens for GVHD prophylaxis		
Underlying disease	**# of studies that included any patients with the underlying disease (32 studies overall)**	**Number of patients with the underlying disease (3,875 patients overall)**
Acute lymphoblastic leukemia	21	770
Acute myeloid leukemia	23	1165
Chronic myeloid leukemia	23	747
Non-Hodgkin lymphoma	10	248
Myelodysplastic syndrome	13	232
Multiple myeloma	7	52
Other malignant disease[a]	13	77
Any acute leukemia[b]	28	2269
Any lymphoma[c]	14	434
Aplastic anemia	3	12
Other non-malignant disease[d]	2	15

[a] includes Hodgkin's lymphoma, chronic lymphocytic leukemia, acute mixed-lineage leukemia, myeloproliferative disease, NK-cell leukemia, and others

[b] includes ALL and AML patients listed separately in table as well as patients reported as "acute leukemia"

[c] includes non-Hodgkin lymphoma patients listed separately in table as well as patients reported as "lymphoma" or Hodgkin's lymphoma

[d] includes myelofibrosis, essential thrombocythemia, hypereosinophilic syndrome, Fanconi disease, amegakaryocytic thrombocytopenia, chronic granulomatous disease, familial hemophagocytic lymphohistiocytosis, and aspartylglucosaminuria

6.3. Appendix 3: Cochrane risk of bias assessments of included studies

The table provided below provides a detailed breakdown of findings from risk of bias assessments using the Cochrane Risk of Bias tool. Findings for the GVHD prophylaxis evidence base are presented first, followed by the GVHD treatment evidence base. Detailed rationale for assessments are available upon request.

Risk of bias assessments of included studies, GVHD prophylaxis

REFID	Domain-Specific Assessments							Overall Evaluations	
	Random sequence generation	Allocation concealment	Blinding of patients & personnel	Blinding of outcome assessor	Attrition: Incomplete outcome data	Selective outcome reporting	Other bias	Outcome I: MORTALITY/ SURVIVAL	Outcome II: aGVHD/cGVHD
Cutler[38,43] (2014)									
Glass[37] (2014)									
Ruutu[44] (2014)									
Armand[46] (2013)									
Bacigalupo[59] (2001)-I									
Bacigalupo[59] (2001)-II									
Chao[68] (1996)									
Chao[60] (2000)									
Deeg[84] (2000)									
Fong[56] (2007)									
Forman[72] (1987)									
Hiraoka[58] (2001)									
Irle[77] (1985)									
Lee[57] (2004)									
Martin[39] (2012)									
Mrsic[69] (1990)									
Nash[62] (2000)									
Perkins[55] (2010)									

Risk of bias assessments of included studies, GVHD prophylaxis									
REFID	Domain-Specific Assessments							Overall Evaluations	
	Random sequence generation	Allocation concealment	Blinding of patients & personnel	Blinding of outcome assessor	Attrition: Incomplete outcome data	Selective outcome reporting	Other bias	Outcome I: MORTALITY/ SURVIVAL	Outcome II: aGVHD/cGVHD
Pidala[47] (2012)									
Pulsipher[48] (2014)									
Ratanatharathorn[65] (1998)									
Backman[71] (1988)									
Ross[64] (1999)									
Ruutu[63] (2000)									
Socie[50] (2011)									
Storb[78] (1983)									
Storb[73] (1986)									
Torres[70] (1989)									
Weiden[81] (1979)									
Zikos[67] (1998)									
Kuzmina[40] (2015)									
Green = low risk, red = high risk, yellow = unclear risk; black = no study data for this endpoint.									

Risk of bias assessments of included studies, GVHD treatment									
REFID	Domain-Specific Assessments							Overall Evaluations	
	Random sequence generation	Allocation concealment	ROB due to blinding: patients & personnel	ROB due to blinding: outcome assessor	Attrition: Incomplete outcome data	Selective outcome reporting	Other bias	Outcome I: MORTALITY/ SURVIVAL	Outcome II: GVHD (ACUTE/CHRONIC)
Kennedy[85] (1985)									
Cragg[86] (2000)									
Alousi[87] (2009)									
Bolanos-Meade[88] (2014)									
Sullivan[89] (1988)									
Arora[91] (2001)									
Martin[90] (2009)									

6.4. Appendix 4: Summary of results for subgroups presented in studies evaluating treatment of GVHD

Summary of results for subgroups presented in studies evaluating treatment of GVHD						
Author (Year)	**Patient sub-group**	**Post-transplant follow-up**	**Outcome definition**	**Comparison (sample size)**	**Group incidence n (%)**	**OR**
Response to treatment						
Kennedy[85] (1985)	Patients requiring secondary therapy	100 days	Free of cGVHD	Steroids (20) vs. CsA (18)	6 (30%) vs. 11 (61%)	0.27 (0.07–1.05)
Kennedy[85] (1985)	Patients not requiring secondary therapy	100 days	Free of cGVHD	Steroids (19) vs. CsA (20)	3 (16%) vs. 5 (25%)	0.56 (0.11–2.77)
Cragg[86] (2000)	age < 18	42 days	Complete+partial response[b]	ATG+Steroids (19) vs. Steroids (19)	15 (79%) vs. 16 (84%)	0.70 (0.13–3.68)
	age ≥ 18			ATG+Steroids (26) vs. Steroids (23)	19 (73%) vs. 16 (70%)	1.19 (0.34–4.11)
	Related donors			ATG+Steroids (21) vs. Steroids (22)	17 (81%) vs. 16 (73%)	1.16 (0.38–6.71)
	Unrelated donors			ATG+Steroids (34) vs. Steroids (20)	17 (71%) vs. 16 (80%)	**0.25 (0.07–0.90)**
	HLA match[a]			ATG+Steroids (19) vs. Steroids (10)	14 (74%) vs. 9 (90%)	0.31 (0.03–3.12)
	HLA match: Other			ATG+Steroids (26) vs. Steroids (32)	20 (77%) vs. 23 (72%)	1.13 (0.40–4.31)
	Donor recipient sex: match			ATG+Steroids (22) vs. Steroids (23)	16 (73%) vs. 18 (78%)	0.74 (0.19–2.90)
	Donor recipient sex: mismatch			ATG+Steroids (23) vs. Steroids (19)	18 (78%) vs. 14 (74%)	1.29 (0.31–5.33)
	GVHD PRX: CsA + MTX			ATG+Steroids (24) vs. Steroids (22)	18 (75%) vs. 18 (82%)	0.67 (0.16–2.77)
	GVHD PRX: Other treatment			ATG+Steroids (21) vs. Steroids (20)	16 (76%) vs. 14 (70%)	1.37 (0.34–5.49)
	conditioning: CY/TBI			ATG+Steroids (33) vs. Steroids (34)	25 (76%) vs. 26 (76%)	0.96 (0.31–2.96)

	conditioning: Other			ATG+Steroids (12) vs. Steroids (8)	9 (75%) vs. 6 (75%)	1.00 (0.13–7.89)
Arora[91] (2001)	High risk[c]	2 months	Complete response[d]	THAL+CsA+Steroid (11) vs. CsA+Steroid (13)	1 (9%) vs. 1 (8%)	1.20 (0.07–21.72)
Arora[91] (2001)	High risk	6 months	Complete response	THAL+CsA+Steroid (11) vs. CsA+Steroid (12)	2 (18%) vs. 1 (8%)	2.44 (0.19–31.53)
Arora[91] (2001)	High risk	1 year	Complete response	THAL+CsA+Steroid (8) vs. CsA+Steroid (9)	3 (38%) vs. 3 (33%)	1.20 (0.16–8.80)
Alousi[87] (2009)	Not receiving MMF PRX prior to randomization	28 days	Complete response[e]	ETN+Steroids (32) vs. MMF+Steroids (45)	9 (28%) vs. 27 (60%)	**0.26 (0.10–0.69)**
Alousi[87] (2009)	Not receiving MMF PRX prior to randomization	56 days	Complete response	ETN+Steroids (32) vs. MMF+Steroids (45)	17 (53%) vs. 33 (73%)	0.76 (0.31–1.82)
Bolanos-Meade [88](2014)	Grade III–IV aGVHD patients only	56 days	Alive and free of aGVHD **and** cGVHD	MMF+Steroids (37) vs. PLB+Steroids (41)	20 (54%) vs. 21 (51%)	1.06 (0.50–2.25)
Bolanos-Meade[88] (2014)	Liver/gut aGVHD patients only	56 days	Alive and free of aGVHD **and** cGVHD	MMF+Steroids (62) vs. PLB+Steroids (58)	32 (52%) vs. 35 (60%)	0.70 (0.34–1.45)
Overall mortality						
Sullivan[89] (1988)	age </=20 years	≥3.8 years	—	AZA+Steroids (17) vs. PLB+Steroids (31)	6 (35%) vs. 8 (26%)	1.57 (0.44–5.64)
Sullivan[89] (1988)	age > 20 years	≥3.8 years	—	AZA+Steroids (46) vs. PLB+Steroids (32)	27 (59%) vs. 15 (47%)	1.61 (0.65–4.00)
Sullivan [89](1988)	Previous aGVHD II-IV	≥3.8 years	—	AZA+Steroids (26) vs. PLB+Steroids (29)	16 (62%) vs. 12 (41%)	2.27 (0.77–6.69)
Sullivan[89] (1988)	Clinical cGVHD[f]	≥3.8 years	—	AZA+Steroids (38) vs. PLB+Steroids (39)	23 (61%) vs. 13 (33%)	**3.07 (1.21–7.78)**

[a] unrelated mismatched donor
[b] resolution of all symptoms in all organs or improvement in one or more organ without any worsening in other organs.
[c] presence of 1 or more of 3 risk factors: progressive onset, platelet count of <100,000µL, and bilirubin level of >/3 mg/dL
[d] resolution of all signs and symptoms of cGVHD
[e] resolution of all signs and symptoms of GVHD in all organs without intervening salvage therapies
[f] patients with subclinical cGVHD or with subclinical that progressed to clinical cGVHD were not included.
ATG = anti-thymocyte globulin; AZA = alemtuzumab; CsA = cyclosporin A; ETN = etanercept; MMF = mycophenolate mofetil; PLB = placebo; SIR = sirolimus; TAC = tacrolimus; THAL = thalidomide

6.5. Appendix 5: Findings from secondary analyses

6.5.1. Sensitivity analysis to determine robustness of estimated frequency data for disease relapse at 2–3 years post-transplant

Frequency data were estimated from cumulative incidence or Kaplan-Meier estimates for 7 of the 17 studies reporting disease relapse at 2–3 years post-transplant.[47,48,50,58,59,64] To determine the robustness of these data, a sensitivity analysis was conducted by removing those 7 studies from the NMA. The comparisons in common between the full NMA (with all 17 studies) and the reduced NMA (with 10 studies that reported frequency data[56,57,62,70,71,73,77,78]) are compared in the table below, using a fixed effect model. All effect estimates and CrIs were comparable between networks, indicating that the estimated frequency data were robust. Model fit was good for both networks (full network: total deviance residual = 31.94 vs. 32.04 and DIC = 192.709 vs. 194.545 for FE and RE models, respectively, with 34 data points; reduced network: total deviance residual = 18.57 vs. 18.63 and DIC = 107.426 vs. 109.400 for FE and RE models, respectively, with 20 data points).

Sensitivity analysis—robustness of estimated frequency data for disease relapse at 2–3 years post-transplant

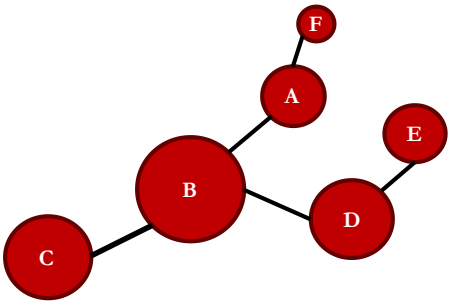
Comparison	OR (95% CrI)	
	Full network	Reduced network
MTX vs. CsA	0.64 (0.36–1.14)	0.64 (0.36–1.14)
MTX vs. HCQ+CsA	0.48 (0.17–1.33)	0.48 (0.17–1.31)
MTX vs. CsA+MTX	**0.17 (0.05–0.49)**	**0.17 (0.06–0.49)**
MTX vs. MTX+TAC	**0.14 (0.04–0.43)**	**0.14 (0.04–0.43)**
CsA vs. HCQ+CsA	0.75 (0.33–1.72)	0.75 (0.33–1.71)
CsA vs. CsA+MTX	**0.26 (0.10–0.65)**	**0.27 (0.10–0.64)**
CsA vs. MTX+TAC	**0.21 (0.07–0.58)**	**0.22 (0.08–0.57)**
HCQ+CsA vs. CsA+MTX	0.35 (0.10–1.19)	0.36 (0.10–1.19)
HCQ+CsA vs. MTX+TAC	0.28 (0.07–1.04)	0.29 (0.08–1.03)
CsA+MTX vs. MTX+TAC	0.81 (0.53–1.24)	0.81 (0.53–1.24)

6.6. Appendix 6: Summary of inconsistency analyses

Assessment of inconsistency assumption: prophylaxis of GVHD outcomes		
Outcome	**Model type**	**DIC**
Acute GVHD	Fixed effects	351.874
	Fixed effects inconsistency	351.487
Chronic GVHD	Fixed effects	292.873
	Fixed effects inconsistency	286.594
A reduction in DIC value of 5 or more indicates a better model fit.		

6.7. Appendix 7: Supplemental network meta-analyses, prevention of chronic GVHD

Network diagram, GVHD prophylaxis, extensive cGVHD in evaluable patients



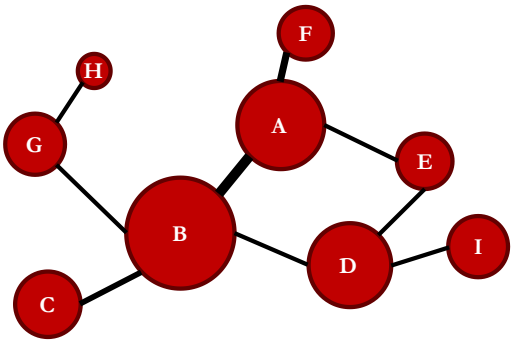
Legend	
Drug	**Abbreviation**
CsA	A
CsA+MTX	B
CsA+MTX+ATG+Steroids	C
CsA+MTX+Steroids	D
CsA+Steroids	E
MTX	F

Network Characteristics

Characteristic	Number
Number of Interventions	6
Number of Studies	6
Total Number of Patients in Network	603
Total Number of Events in Network	155
Total Possible Pairwise Comparisons	15
Total Number Pairwise Comparisons With Direct Data	5
Number of Two-arm Studies	6
Number of Multi-Arms Studies	0
Number of Studies With No Zero Events	6
Number of Studies With At Least One Zero Event	0
Number of Studies with All Zero Events	0

0.97 **CsA+MTX+ATG+Steroids**	0.66	0.63	0.47	0.14	0.13
0.33 (0.09 – 1.26)	**CsA+MTX+Steroids**				
0.32 (0.07 – 1.46)	0.97 (0.47 – 1.98)	**CsA+Steroids**			
0.21 **(0.11 – 0.38)**	0.63 (0.19 – 1.93)	0.65 (0.16 – 2.45)	**CsA+MTX**		
0.07 **(0.02 – 0.23)**	**0.20** **(0.04 – 0.98)**	0.21 (0.03 – 1.18)	**0.32** **(0.11 – 0.92)**	**CsA**	
0.05 **(0.01 – 0.46)**	0.16 (0.01 – 1.74)	0.16 (0.01 – 1.99)	0.25 (0.03 – 2.07)	0.77 (0.11 – 4.94)	**MTX**

Network diagram, GVHD prophylaxis, extensive cGVHD in evaluable patients



Legend	
Drug	**Abbreviation**
CsA	A
CsA+MTX	B
CsA+MTX+ATG+Steroids	C
CsA+MTX+Steroids	D
CsA+Steroids	E
MTX	F
MTX+TAC	G
TAC+MMF	H
UDCA+CsA+MTX+/-Steroids	I

Network Characteristics

Characteristic	Number
Number of Interventions	12
Number of Studies	14
Total Number of Patients in Network	1,904
Total Number of Events in Network	473
Total Possible Pairwise Comparisons	66
Total Number Pairwise Comparisons With Direct Data	11
Number of Two-arm Studies	14
Number of Multi-Arms Studies	0
Number of Studies With No Zero Events	14
Number of Studies With At Least One Zero Event	0
Number of Studies with All Zero Events	0

0.96											
CsA+MTX+ATG+ Steroids	0.86										
0.66 (0.19 – 2.17)	SIR+TAC	0.71									
0.37 (0.18 – 0.78)	0.57 (0.22 – 1.47)	MTX+TAC	0.66								
0.38 (0.07 – 2.07)	0.57 (0.10 – 3.43)	1.01 (0.24 – 4.62)	RITX+TAC+MMF	0.55							
0.29 (0.10 – 0.84)	0.44 (0.13 – 1.50)	0.77 (0.34 – 1.65)	0.76 (0.14 – 3.83)	MTX+TAC+ Steroids	0.50						
0.26 (0.08 – 0.81)	0.40 (0.11 – 1.44)	0.71 (0.30 – 1.65)	0.70 (0.20 – 2.24)	0.92 (0.29 – 2.99)	TAC+MMF	0.43					
0.22 (0.08 – 0.60)	0.33 (0.08 – 1.37)	0.58 (0.21 – 1.62)	0.58 (0.09 – 3.38)	0.76 (0.21 – 2.79)	0.83 (0.21 – 3.17)	CsA+Steroids	0.41				
0.22 (0.13 – 0.36)	0.33 (0.11 – 1.01)	0.58 (0.33 – 1.00)	0.58 (0.11 – 2.67)	0.76 (0.29 – 2.02)	0.82 (0.30 – 2.28)	1.00 (0.41 – 2.41)	CsA+MTX	0.39			
0.21 (0.10 – 0.43)	0.32 (0.09 – 1.08)	0.56 (0.26 – 1.19)	0.55 (0.10 – 2.77)	0.73 (0.24 – 2.19)	0.78 (0.25 – 2.51)	0.96 (0.47 – 1.93)	0.95 (0.56 – 1.63)	CsA+MTX+ Steroids	0.30		
0.15 (0.03 – 0.63)	0.23 (0.04 – 1.10)	0.41 (0.10 – 1.37)	0.40 (0.05 – 2.69)	0.54 (0.11 – 2.25)	0.58 (0.11 – 2.55)	0.70 (0.13 – 3.53)	0.71 (0.16 – 2.70)	0.74 (0.15 – 3.14)	MTX+SIR+TAC	0.14	
0.06 (0.01 – 0.55)	0.10 (0.01 – 1.04)	0.17 (0.02 – 1.49)	0.16 (0.01 – 2.22)	0.22 (0.02 – 2.28)	0.24 (0.02 – 2.42)	0.29 (0.03 – 2.82)	0.29 (0.03 – 2.41)	0.30 (0.03 – 2.68)	0.41 (0.03 – 5.54)	MTX	0.08
0.07 (0.02 – 0.20)	**0.10 (0.02 – 0.44)**	**0.18 (0.05 – 0.54)**	0.17 (0.02 – 1.11)	**0.23 (0.05 – 0.91)**	0.25 (0.05 – 1.03)	0.31 (0.07 – 1.13)	**0.31 (0.09 – 0.81)**	**0.32 (0.09 – 0.97)**	0.43 (0.07 – 2.48)	1.04 (0.16 – 6.48)	CsA

1 link Direct evidence | 2 links Simple indirect | 3+ links Compound indirect | Bold + underline = significant OR

6.8. Appendix 8: Prophylaxis of GVHD: Findings from Random Effects NMAs

Acute GVHD

SIR+TAC														
0.85 (0.16 – 4.03)	MTX+SIR+TAC													
0.56 (0.09 – 2.88)	0.66 (0.12 – 3.65)	MTX+TAC+Steroids												
0.59 (0.05 – 7.12)	0.70 (0.06 – 9.01)	1.05 (0.10 – 12.41)	CsA+MTX+MSCs											
0.33 (0.05 – 1.78)	0.39 (0.06 – 2.29)	0.59 (0.11 – 3.07)	0.56 (0.06 – 4.98)	CsA+MTX+ATG or ALZ+Steroids										
0.32 (0.09 – 0.92)	0.38 (0.12 – 1.18)	0.57 (0.16 – 2.05)	0.54 (0.06 – 4.76)	0.97 (0.25 – 3.86)	MTX+TAC									
0.32 (0.04 – 2.37)	0.38 (0.05 – 3.03)	0.57 (0.07 – 4.90)	0.54 (0.03 – 8.91)	0.98 (0.11 – 8.87)	1.01 (0.18 – 5.66)	MMF+TAC								
0.29 (0.02 – 3.79)	0.34 (0.02 – 4.82)	0.52 (0.03 – 7.86)	0.49 (0.02 – 12.58)	0.89 (0.06 – 13.90)	0.92 (0.08 – 9.87)	0.90 (0.18 – 4.63)	RITX+TAC+MMF							
0.23 (0.04 – 1.11)	0.26 (0.05 – 1.41)	0.40 (0.11 – 1.44)	0.38 (0.04 – 3.34)	0.68 (0.17 – 2.71)	0.70 (0.20 – 2.39)	0.69 (0.08 – 5.78)	0.77 (0.05 – 11.19)	CsA+MTX+Steroids						
0.17 (0.01 – 3.89)	0.20 (0.01 – 4.84)	0.30 (0.02 – 6.84)	0.28 (0.01 – 8.89)	0.51 (0.03 – 10.40)	0.52 (0.04 – 10.49)	0.53 (0.02 – 16.36)	0.59 (0.02 – 25.85)	0.75 (0.05 – 14.70)	ATG+MTX+ Steroids					
0.13 (0.03 – 0.54)	**0.16 (0.03 – 0.69)**	**0.24 (0.06 – 0.89)**	0.23 (0.03 – 1.57)	0.41 (0.15 – 1.07)	0.42 (0.16 – 1.07)	0.41 (0.06 – 2.94)	0.46 (0.03 – 5.93)	0.60 (0.22 – 1.54)	0.80 (0.05 – 9.66)	CsA+MTX				
0.09 (0.01 – 0.59)	**0.11 (0.01 – 0.75)**	**0.16 (0.03 – 0.90)**	0.15 (0.01 – 1.62)	0.28 (0.05 – 1.45)	0.29 (0.06 – 1.38)	0.28 (0.03 – 2.93)	0.31 (0.02 – 5.37)	0.41 (0.11 – 1.48)	0.55 (0.03 – 8.01)	0.69 (0.18 – 2.61)	CsA+Steroids			
0.06 (0.01 – 0.29)	**0.07 (0.01 – 0.37)**	**0.10 (0.02 – 0.47)**	**0.10 (0.01 – 0.81)**	**0.18 (0.05 – 0.62)**	**0.18 (0.05 – 0.62)**	0.18 (0.02 – 1.51)	0.20 (0.01 – 2.89)	0.26 (0.08 – 0.82)	0.35 (0.02 – 3.67)	**0.44 (0.19 – 0.98)**	0.64 (0.18 – 2.27)	CsA		
0.04 (0.00 – 0.44)	**0.05 (0.00 – 0.56)**	**0.07 (0.01 – 0.71)**	0.07 (0.00 – 1.13)	0.12 (0.01 – 1.16)	0.12 (0.01 – 1.14)	0.12 (0.01 – 2.07)	0.13 (0.00 – 3.52)	0.17 (0.02 – 1.34)	0.23 (0.01 – 5.29)	0.29 (0.04 – 2.34)	0.43 (0.08 – 2.12)	0.67 (0.08 – 5.20)	MTX+Steroids	
0.05 (0.01 – 0.29)	**0.06 (0.01 – 0.37)**	**0.08 (0.01 – 0.49)**	**0.08 (0.01 – 0.79)**	**0.14 (0.03 – 0.67)**	**0.15 (0.03 – 0.68)**	0.15 (0.01 – 1.45)	0.16 (0.01 – 2.71)	**0.21 (0.05 – 0.91)**	0.28 (0.02 – 2.59)	0.36 (0.11 – 1.18)	0.52 (0.11 – 2.47)	0.82 (0.35 – 1.96)	1.22 (0.14 – 11.76)	MTX

Chronic GVHD

CsA+MTX+ ATG+ Steroids												
0.47 (0.09 – 2.46)	HCQ+CsA											
0.34 (0.08 – 1.42)	0.72 (0.11 – 4.99)	MTX+SIR+TAC										
0.32 (0.16 – 0.67)	0.69 (0.16 – 3.11)	0.95 (0.28 – 3.21)	CsA+MTX									
0.31 (0.11 – 0.95)	0.66 (0.12 – 3.72)	0.93 (0.38 – 2.23)	0.97 (0.42 – 2.21)	MTX+TAC								
0.30 (0.05 – 1.71)	0.63 (0.08 – 5.23)	0.87 (0.12 – 6.35)	0.93 (0.20 – 4.53)	0.94 (0.17 – 5.65)	UDCA+CsA+ MTX +/- Steroids							
0.30 (0.05 – 1.75)	0.62 (0.08 – 5.47)	0.86 (0.12 – 6.66)	0.91 (0.18 – 4.67)	0.93 (0.15 – 5.95)	0.99 (0.18 – 5.24)	MTX+TAC+ Steroids						
0.29 (0.08 – 1.04)	0.61 (0.11 – 3.52)	0.85 (0.17 – 4.26)	0.89 (0.31 – 2.60)	0.92 (0.24 – 3.55)	0.96 (0.30 – 3.03)	0.98 (0.28 – 3.42)	CsA+MTX+ Steroids					
0.29 (0.05 – 1.68)	0.61 (0.07 – 5.63)	0.85 (0.16 – 4.39)	0.89 (0.17 – 4.44)	0.92 (0.22 – 3.70)	0.97 (0.10 – 8.86)	0.97 (0.10 – 9.82)	0.99 (0.14 – 6.98)	TAC+MMF				
0.25 (0.09 – 0.66)	0.53 (0.14 – 1.95)	0.73 (0.18 – 2.86)	0.77 (0.38 – 1.49)	0.79 (0.26 – 2.29)	0.83 (0.16 – 4.04)	0.85 (0.15 – 4.44)	0.86 (0.27 – 2.59)	0.86 (0.14 – 5.03)	CsA			
0.23 (0.05 – 1.11)	0.48 (0.06 – 3.89)	0.67 (0.16 – 2.84)	0.71 (0.17 – 2.87)	0.73 (0.23 – 2.29)	0.77 (0.09 – 6.12)	0.78 (0.09 – 6.65)	0.80 (0.14 – 4.56)	0.80 (0.13 – 4.94)	0.92 (0.20 – 4.49)	SIR+TAC		
0.22 (0.06 – 0.80)	0.47 (0.10 – 2.21)	0.66 (0.12 – 3.22)	0.70 (0.23 – 1.97)	0.71 (0.18 – 2.70)	0.75 (0.12 – 4.32)	0.76 (0.11 – 4.79)	0.78 (0.18 – 3.07)	0.77 (0.11 – 5.44)	0.90 (0.39 – 2.03)	0.97 (0.16 – 5.55)	MTX	
0.17 (0.04 – 0.63)	0.36 (0.06 – 1.95)	0.50 (0.09 – 2.55)	0.53 (0.16 – 1.57)	0.54 (0.13 – 2.15)	0.57 (0.12 – 2.52)	0.58 (0.11 – 2.76)	0.59 (0.20 – 1.58)	0.59 (0.08 – 4.20)	0.68 (0.23 – 1.97)	0.74 (0.12 – 4.33)	0.76 (0.19 – 2.91)	CsA+Steroids
0.03 (0.00 – 0.52)	0.07 (0.00 – 1.34)	0.10 (0.00 – 1.91)	0.11 (0.00 – 1.45)	0.11 (0.00 – 1.77)	0.11 (0.00 – 2.34)	0.11 (0.00 – 2.47)	0.12 (0.00 – 1.94)	0.11 (0.00 – 2.89)	0.14 (0.00 – 1.80)	0.15 (0.00 – 3.08)	0.16 (0.00 – 1.75)	0.20 (0.00 – 3.31)

100-day Mortality

ATG + MTX + Steroids														
1.29 (0.05 – 52.27)	CsA + MTX + MSCs													
0.72 (0.06 – 9.04)	0.54 (0.01 – 10.77)	UDCA + CsA + MTX + /-Steroids												
0.41 (0.05 – 3.46)	0.30 (0.01 – 4.27)	0.57 (0.10 – 2.98)	CsA + MTX + ATG + Steroids											
0.40 (0.08 – 1.65)	0.30 (0.01 – 5.13)	0.54 (0.07 – 4.16)	0.96 (0.20 – 4.52)	MTX										
0.36 (0.05 – 2.69)	0.27 (0.01 – 3.36)	0.49 (0.10 – 2.25)	0.88 (0.42 – 1.94)	0.92 (0.24 – 3.76)	CsA + MTX									
0.36 (0.03 – 5.27)	0.28 (0.01 – 5.08)	0.50 (0.09 – 2.70)	0.89 (0.14 – 5.44)	0.91 (0.11 – 7.43)	1.00 (0.18 – 5.25)	MTX + TAC + Steroids								
0.35 (0.04 – 3.56)	0.26 (0.01 – 4.02)	0.48 (0.15 – 1.48)	0.84 (0.24 – 3.20)	0.89 (0.16 – 4.62)	0.96 (0.33 – 2.78)	0.97 (0.26 – 3.54)	CsA + MTX + Steroids							
0.34 (0.03 – 4.78)	0.26 (0.01 – 5.59)	0.47 (0.05 – 4.70)	0.83 (0.12 – 5.86)	0.87 (0.10 – 7.90)	0.93 (0.15 – 5.46)	0.99 (0.07 – 10.84)	0.99 (0.14 – 7.58)	TAC + MMF						
0.32 (0.03 – 3.82)	0.24 (0.01 – 4.42)	0.43 (0.06 – 3.71)	0.77 (0.17 – 4.21)	0.79 (0.13 – 5.88)	0.87 (0.24 – 3.81)	0.88 (0.11 – 8.16)	0.90 (0.17 – 5.50)	0.94 (0.13 – 6.59)	SIR + TAC					
0.31 (0.04 – 2.59)	0.23 (0.01 – 3.16)	0.42 (0.08 – 2.37)	0.74 (0.27 – 2.31)	0.75 (0.18 – 4.02)	0.84 (0.41 – 1.84)	0.85 (0.14 – 5.35)	0.87 (0.25 – 3.28)	0.89 (0.19 – 4.20)	0.95 (0.29 – 2.95)	MTX + TAC				
0.28 (0.03 – 2.74)	0.20 (0.01 – 4.46)	0.37 (0.04 – 3.82)	0.66 (0.12 – 4.31)	0.68 (0.13 – 3.94)	0.73 (0.16 – 3.98)	0.76 (0.08 – 8.00)	0.76 (0.12 – 5.68)	0.80 (0.07 – 8.35)	0.82 (0.10 – 6.96)	0.87 (0.16 – 5.32)	HCQ + CsA			
0.23 (0.01 – 3.17)	0.17 (0.00 – 3.47)	0.30 (0.05 – 1.82)	0.53 (0.08 – 3.53)	0.56 (0.06 – 4.70)	0.61 (0.10 – 3.38)	0.61 (0.09 – 4.15)	0.62 (0.15 – 2.51)	0.66 (0.05 – 7.04)	0.72 (0.07 – 5.73)	0.75 (0.10 – 4.54)	0.85 (0.07 – 8.10)	CsA + Steroids		
0.22 (0.04 – 1.42)	0.17 (0.01 – 2.44)	0.31 (0.06 – 1.88)	0.55 (0.18 – 1.84)	0.57 (0.22 – 1.60)	0.62 (0.28 – 1.51)	0.64 (0.10 – 4.24)	0.65 (0.17 – 2.63)	0.65 (0.10 – 4.53)	0.71 (0.14 – 3.44)	0.75 (0.25 – 2.31)	0.84 (0.21 – 3.24)	1.01 (0.16 – 7.68)	CsA	
0.17 (0.02 – 2.01)	0.12 (0.00 – 2.33)	0.23 (0.03 – 1.84)	0.40 (0.08 – 2.08)	0.42 (0.07 – 3.11)	0.46 (0.11 – 1.92)	0.47 (0.05 – 4.16)	0.47 (0.08 – 2.76)	0.48 (0.07 – 3.29)	0.51 (0.10 – 2.56)	0.54 (0.17 – 1.73)	0.62 (0.07 – 5.02)	0.74 (0.08 – 7.12)	0.73 (0.14 – 3.94)	MTX + SIR + TAC

1-year Mortality

CsA + MTX + MSCs														
0.10 (0.00 – 1.83)	UDCA + CsA + MTX + /-Steroids													
0.10 (0.00 – 2.37)	0.92 (0.07 – 15.46)	ATG + MTX + Steroids												
0.08 (0.00 – 1.23)	0.76 (0.10 – 7.05)	0.84 (0.16 – 3.95)	MTX											
0.06 (0.00 – 1.24)	0.62 (0.09 – 4.32)	0.68 (0.04 – 8.97)	0.81 (0.08 – 6.63)	MTX + TAC + Steroids										
0.06 (0.00 – 1.17)	0.61 (0.06 – 6.76)	0.66 (0.04 – 8.65)	0.80 (0.09 – 6.42)	0.98 (0.09 – 11.86)	TAC + MMF									
0.06 (0.00 – 0.62)	0.57 (0.11 – 3.00)	0.63 (0.07 – 4.33)	0.75 (0.19 – 2.47)	0.93 (0.16 – 5.37)	0.94 (0.16 – 5.20)	CsA + MTX								
0.06 (0.00 – 0.72)	0.58 (0.08 – 3.58)	0.63 (0.06 – 4.87)	0.75 (0.14 – 3.16)	0.93 (0.13 – 6.44)	0.95 (0.13 – 6.22)	1.00 (0.41 – 2.32)	CsA + MTX + ATG + Steroids							
0.06 (0.00 – 0.72)	0.52 (0.14 – 1.98)	0.57 (0.05 – 5.20)	0.68 (0.12 – 3.28)	0.84 (0.21 – 3.53)	0.86 (0.11 – 6.26)	0.91 (0.33 – 2.53)	0.91 (0.25 – 3.62)	CsA + MTX + Steroids						
0.05 (0.00 – 0.65)	0.49 (0.08 – 3.36)	0.54 (0.05 – 4.46)	0.65 (0.12 – 2.90)	0.80 (0.11 – 5.88)	0.82 (0.19 – 3.52)	0.86 (0.34 – 2.21)	0.86 (0.25 – 3.25)	0.95 (0.24 – 3.79)	MTX + TAC					
0.05 (0.00 – 0.68)	0.49 (0.08 – 3.63)	0.54 (0.08 – 3.21)	0.65 (0.27 – 1.56)	0.79 (0.12 – 6.41)	0.81 (0.12 – 6.16)	0.86 (0.37 – 2.34)	0.86 (0.27 – 3.41)	0.94 (0.26 – 4.01)	0.99 (0.29 – 3.89)	CsA				
0.05 (0.00 – 0.94)	0.45 (0.07 – 3.17)	0.49 (0.03 – 6.71)	0.58 (0.06 – 4.77)	0.73 (0.10 – 5.20)	0.74 (0.06 – 8.26)	0.78 (0.14 – 4.44)	0.77 (0.12 – 5.63)	0.85 (0.21 – 3.49)	0.91 (0.13 – 6.38)	0.91 (0.12 – 6.01)	CsA + Steroids			
0.05 (0.00 – 0.77)	0.43 (0.04 – 4.39)	0.47 (0.03 – 5.56)	0.56 (0.07 – 3.99)	0.69 (0.06 – 7.63)	0.70 (0.10 – 4.95)	0.74 (0.15 – 3.78)	0.74 (0.13 – 4.85)	0.81 (0.12 – 5.52)	0.86 (0.23 – 3.19)	0.87 (0.12 – 5.10)	0.95 (0.09 – 10.25)	SIR + TAC		
0.04 (0.00 – 0.58)	0.36 (0.04 – 3.09)	0.40 (0.03 – 4.09)	0.47 (0.07 – 2.80)	0.58 (0.07 – 5.43)	0.59 (0.10 – 3.46)	0.62 (0.16 – 2.52)	0.62 (0.13 – 3.41)	0.69 (0.13 – 3.75)	0.73 (0.27 – 2.01)	0.73 (0.13 – 3.51)	0.81 (0.09 – 7.36)	0.85 (0.16 – 4.42)	MTX + SIR + TAC	
0.03 (0.00 – 0.57)	0.28 (0.03 – 3.30)	0.31 (0.03 – 2.93)	0.36 (0.07 – 1.96)	0.45 (0.04 – 5.59)	0.46 (0.04 – 5.47)	0.49 (0.10 – 2.91)	0.48 (0.08 – 3.73)	0.53 (0.08 – 4.29)	0.56 (0.09 – 4.24)	0.57 (0.13 – 2.40)	0.63 (0.06 – 7.68)	0.66 (0.07 – 7.34)	0.77 (0.09 – 7.44)	HCQ + CsA

Disease Relapse

MTX										
0.64 (0.32 – 1.26)	CSA									
0.58 (0.06 – 7.47)	0.90 (0.10 – 10.34)	CsA + MTX + Steroids								
0.47 (0.12 – 1.88)	0.74 (0.23 – 2.46)	0.82 (0.05 – 10.52)	HCQ + CsA							
0.45 (0.03 – 7.81)	0.69 (0.06 – 11.38)	0.79 (0.22 – 2.78)	0.96 (0.05 – 19.67)	CsA + Steroids						
0.27 (0.04 – 2.32)	0.42 (0.06 – 3.14)	0.47 (0.12 – 1.69)	0.56 (0.06 – 5.97)	0.59 (0.09 – 3.67)	MTX + TAC + Steroids					
0.18 (0.05 – 0.72)	**0.27 (0.08 – 0.91)**	0.30 (0.04 – 2.24)	0.37 (0.07 – 2.03)	0.39 (0.03 – 4.10)	0.66 (0.13 – 3.05)	CSA + MTX				
0.15 (0.02 – 1.01)	0.23 (0.04 – 1.39)	0.25 (0.03 – 2.23)	0.31 (0.04 – 2.77)	0.32 (0.03 – 3.89)	0.56 (0.10 – 3.04)	0.85 (0.24 – 2.96)	SIR + TAC			
0.15 (0.03 – 0.87)	0.23 (0.06 – 1.13)	0.26 (0.03 – 2.44)	0.32 (0.05 – 2.35)	0.33 (0.02 – 4.41)	0.57 (0.09 – 3.43)	0.86 (0.36 – 2.22)	1.00 (0.22 – 4.99)	CSA + MTX + ATG + Steroids		
0.14 (0.03 – 0.72)	**0.22 (0.05 – 0.94)**	0.24 (0.04 – 1.62)	0.30 (0.05 – 2.08)	0.31 (0.03 – 2.99)	0.54 (0.13 – 2.12)	0.81 (0.38 – 1.71)	0.95 (0.35 – 2.65)	0.94 (0.28 – 2.92)	MTX + TAC	
0.11 (0.02 – 0.80)	0.17 (0.03 – 1.12)	0.19 (0.02 – 1.71)	0.23 (0.03 – 2.16)	0.25 (0.02 – 3.03)	0.42 (0.07 – 2.39)	0.63 (0.17 – 2.45)	0.74 (0.17 – 3.42)	0.73 (0.14 – 3.56)	0.78 (0.26 – 2.41)	MTX + SIR + TAC

6.9. Appendix 9: PRISMA NMA checklist

PRISMA NMA Checklist of Items to Include When Reporting a Systematic Review Involving a Network Meta-analysis

Section/Topic	Item #	Checklist Item	Reported on Page #
TITLE			
Title	1	Identify the report as a systematic review *incorporating a network meta-analysis (or related form of meta-analysis).*	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: **Background:** main objectives **Methods:** data sources; study eligibility criteria, participants, and interventions; study appraisal; and *synthesis methods, such as network meta-analysis.* **Results:** number of studies and participants identified; summary estimates with corresponding confidence/credible intervals; *treatment rankings may also be discussed. Authors may choose to summarize pairwise comparisons against a chosen treatment included in their analyses for brevity.* **Discussion/Conclusions:** limitations; conclusions and implications of findings. **Other:** primary source of funding; systematic review registration number with registry name.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known, *including mention of why a network meta-analysis has been conducted.*	12-13
Objectives	4	Provide an explicit statement of questions being addressed, with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	13

METHODS			
Protocol and registration	5	Indicate whether a review protocol exists and if and where it can be accessed (e.g., Web address); and, if available, provide registration information, including registration number.	13
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale. *Clearly describe eligible treatments included in the treatment network, and note whether any have been clustered or merged into the same node (with justification).*	13-14
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	14
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	88
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	15
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	15
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	15-16
Geometry of the network	**S1**	Describe methods used to explore the geometry of the treatment network under study and potential biases related to it. This should include how the evidence base has been graphically summarized for presentation, and what characteristics were compiled and used to describe the evidence base to readers.	16
Risk of bias within individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	15

Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means). *Also describe the use of additional summary measures assessed, such as treatment rankings and surface under the cumulative ranking curve (SUCRA) values, as well as modified approaches used to present summary findings from meta-analyses.*	17-18
Planned methods of analysis	14	Describe the methods of handling data and combining results of studies for each network meta-analysis. This should include, but not be limited to: • *Handling of multi-arm trials;* • *Selection of variance structure;* • *Selection of prior distributions in Bayesian analyses; and* • *Assessment of model fit.*	17-19
Assessment of Inconsistency	**S2**	Describe the statistical methods used to evaluate the agreement of direct and indirect evidence in the treatment network(s) studied. Describe efforts taken to address its presence when found.	18
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	NA
Additional analyses	16	Describe methods of additional analyses if done, indicating which were pre-specified. This may include, but not be limited to, the following: • Sensitivity or subgroup analyses; • Meta-regression analyses; • *Alternative formulations of the treatment network; and* • *Use of alternative prior distributions for Bayesian analyses (if applicable).*	18-19
RESULTS†			

Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	20
Presentation of network structure	**S3**	Provide a network graph of the included studies to enable visualization of the geometry of the treatment network.	Throughout sections 3 and 4
Summary of network geometry	**S4**	Provide a brief overview of characteristics of the treatment network. This may include commentary on the abundance of trials and randomized patients for the different interventions and pairwise comparisons in the network, gaps of evidence in the treatment network, and potential biases reflected by the network structure.	Throughout sections 3 and 4
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	21-26; 58-63
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment.	25-26; 60-61; 90-92
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: 1) simple summary data for each intervention group, and 2) effect estimates and confidence intervals. *Modified approaches may be needed to deal with information from larger networks.*	Upon request
Synthesis of results	21	Present results of each meta-analysis done, including confidence/credible intervals. *In larger networks, authors may focus on comparisons versus a particular comparator (e.g. placebo or standard care), with full findings presented in an appendix. League tables and forest plots may be considered to summarize pairwise comparisons.* If additional summary measures were explored (such as treatment rankings), these should also be presented.	Throughout sections 3 and 4, beginning on page 26
Exploration for inconsistency	**S5**	Describe results from investigations of inconsistency. This may include such information as measures of model fit to compare consistency and inconsistency models, *P* values from statistical tests, or summary of inconsistency estimates from different parts of the treatment network.	58
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies for the evidence base being studied.	NA
Results of additional analyses	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression analyses, *alternative network geometries studied, alternative choice of prior distributions for Bayesian analyses*, and so forth).	Not feasible based upon evidence structures

DISCUSSION			
Summary of evidence	24	Summarize the main findings, including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy-makers).	73-77
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review level (e.g., incomplete retrieval of identified research, reporting bias). *Comment on the validity of the assumptions, such as transitivity and consistency. Comment on any concerns regarding network geometry (e.g., avoidance of certain comparisons).*	77
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	78
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review. This should also include information regarding whether funding has been received from manufacturers of treatments in the network and/or whether some of the authors are content experts with professional conflicts of interest that could affect use of treatments in the network.	78

PICOS = population, intervention, comparators, outcomes, study design.

* Text in italics indicates wording specific to reporting of network meta-analyses that has been added to guidance from the PRISMA statement.

† Authors may wish to plan for use of appendices to present all relevant information in full detail for items in this section.