Cochrane Library

Cochrane Database of Systematic Reviews

Stem cell transplantation for systemic sclerosis (Review)

Bruera S, Sidanmat H, Molony DA, Mayes MD, Suarez-Almazor ME, Krause K, Lopez-Olivo MA

Bruera S, Sidanmat H, Molony DA, Mayes MD, Suarez-Almazor ME, Krause K, Lopez-Olivo MA.
Stem cell transplantation for systemic sclerosis.
Cochrane Database of Systematic Reviews 2022, Issue 7. Art. No.: CD011819.
DOI: 10.1002/14651858.CD011819.pub2.

www.cochranelibrary.com

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

WILEY

TABLE OF CONTENTS

ABSTRACT 1
PLAIN LANGUAGE SUMMARY 2
SUMMARY OF FINDINGS 5
BACKGROUND 11
OBJECTIVES 11
METHODS 11
RESULTS 15
Figure 1. 17
Figure 2. 18
Figure 3. 19
DISCUSSION 23
AUTHORS' CONCLUSIONS 24
ACKNOWLEDGEMENTS 25
REFERENCES 26
CHARACTERISTICS OF STUDIES 29
DATA AND ANALYSES 37
Analysis 1.1. Comparison 1: Efficacy – survival, Outcome 1: Overall mortality 38
Analysis 1.2. Comparison 1: Efficacy – survival, Outcome 2: Event-free survival 39
Analysis 2.1. Comparison 2: Participant-reported outcomes – HAQ-DI, Outcome 1: HAQ-DI 39
Analysis 2.2. Comparison 2: Participant-reported outcomes – HAQ-DI, Outcome 2: Improvement in HAQ-DI (≥ 0.4 change) 40
Analysis 3.1. Comparison 3: Efficacy – skin thickness, Outcome 1: Modified Rodnan skin scores 41
Analysis 3.2. Comparison 3: Efficacy – skin thickness, Outcome 2: Modified Rodnan skin score improvement (≥ 25% change in mRSS OR ± ≥ 5 if baseline mRSS ≤ 20) 41
Analysis 4.1. Comparison 4: Efficacy – interstitial lung disease, Outcome 1: Predicted forced vital capacity (% predicted) 42
Analysis 4.2. Comparison 4: Efficacy – interstitial lung disease, Outcome 2: Predicted forced vital capacity improvement (10% change in FVC % predicted) 43
Analysis 4.3. Comparison 4: Efficacy – interstitial lung disease, Outcome 3: DLCO (% predicted) 43
Analysis 4.4. Comparison 4: Efficacy – interstitial lung disease, Outcome 4: DLCO improvement (≥ 15% change in DLCO % predicted) 44
Analysis 4.5. Comparison 4: Efficacy – interstitial lung disease, Outcome 5: Predicted total lung capacity 44
Analysis 5.1. Comparison 5: Safety, Outcome 1: Serious adverse events 46
Analysis 5.2. Comparison 5: Safety, Outcome 2: Treatment-related mortality 47
Analysis 5.3. Comparison 5: Safety, Outcome 3: Serious non-lethal infections 48
Analysis 5.4. Comparison 5: Safety, Outcome 4: Study withdrawals (including non-adherence, non-lethal adverse events, organ failure, dropout) 49
Analysis 5.5. Comparison 5: Safety, Outcome 5: Renal failure 50
Analysis 6.1. Comparison 6: Efficacy – renal function (creatinine clearance), Outcome 1: Autologous non-myeloablative selective HSCT 50
Analysis 7.1. Comparison 7: Efficacy – cardiac function, Outcome 1: Autologous non-myeloablative selective HSCT 51
Analysis 8.1. Comparison 8: Participant-reported outcomes – SF-36, Outcome 1: Physical Function 53
Analysis 8.2. Comparison 8: Participant-reported outcomes – SF-36, Outcome 2: Physical Role Limitation 53
Analysis 8.3. Comparison 8: Participant-reported outcomes – SF-36, Outcome 3: Body Pain 53
Analysis 8.4. Comparison 8: Participant-reported outcomes – SF-36, Outcome 4: General Health Perception 54
Analysis 8.5. Comparison 8: Participant-reported outcomes – SF-36, Outcome 5: Vital Energy Fatigue 54
Analysis 8.6. Comparison 8: Participant-reported outcomes – SF-36, Outcome 6: Social Function 54
Analysis 8.7. Comparison 8: Participant-reported outcomes – SF-36, Outcome 7: Emotional Role Limitation 55
Analysis 8.8. Comparison 8: Participant-reported outcomes – SF-36, Outcome 8: Mental Health 55
Analysis 8.9. Comparison 8: Participant-reported outcomes – SF-36, Outcome 9: Physical Component Summary 55
Analysis 8.10. Comparison 8: Participant-reported outcomes – SF-36, Outcome 10: Mental Component Summary 56
Analysis 8.11. Comparison 8: Participant-reported outcomes – SF-36, Outcome 11: PCS SF-36 improvement (≥ 10-point change) 56

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 8.12. Comparison 8: Participant-reported outcomes – SF-36, Outcome 12: MCS SF-36 improvement (≥ 10-point change) 56
Analysis 9.1. Comparison 9: Participant-reported outcomes – EQ-5D VAS score, Outcome 1: Autologous non-myeloablative selective HSCT 57
APPENDICES 57
HISTORY 58
CONTRIBUTIONS OF AUTHORS 59
DECLARATIONS OF INTEREST 59
SOURCES OF SUPPORT 59
DIFFERENCES BETWEEN PROTOCOL AND REVIEW 59

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

[Intervention Review]

Stem cell transplantation for systemic sclerosis

Sebastian Bruera[1], Harish Sidanmat[2], Donald A Molony[3], Maureen D Mayes[4], Maria E Suarez-Almazor[5], Kate Krause[6], Maria Angeles Lopez-Olivo[5]

[1]Department of Internal Medicine, Baylor College of Medicine, Houston, Texas, USA. [2]Department of General Internal Medicine, The University of Texas, MD Anderson Cancer Center, Houston, Texas, USA. [3]Internal Medicine, UT-Houston Health Science Center, Houston, Texas, USA. [4]Division of Rheumatology and Clinical Immunogenetics, The University of Texas at Houston Medical School, Houston, Texas, USA. [5]Department of Health Services Research, The University of Texas, MD Anderson Cancer Center, Houston, USA. [6]Research Medical Library, The University of Texas, MD Anderson Cancer Center, Houston, Texas, USA

Contact: Maria Angeles Lopez-Olivo, amlopezo@mdanderson.org.

Editorial group: Cochrane Musculoskeletal Group.
Publication status and date: New, published in Issue 7, 2022.

Citation: Bruera S, Sidanmat H, Molony DA, Mayes MD, Suarez-Almazor ME, Krause K, Lopez-Olivo MA. Stem cell transplantation for systemic sclerosis. *Cochrane Database of Systematic Reviews* 2022, Issue 7. Art. No.: CD011819. DOI: 10.1002/14651858.CD011819.pub2.

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Systemic sclerosis (SSc) is a chronic autoimmune disease characterized by systemic inflammation, fibrosis, vascular injury, reduced quality of life, and limited treatment options. Autologous hematopoietic stem cell transplantation (HSCT) has emerged as a potential intervention for severe SSc refractory to conventional treatment.

Objectives

To assess the benefits and harms of autologous hematopoietic stem cell transplantation for the treatment of systemic sclerosis (specifically, non-selective myeloablative HSCT versus cyclophosphamide; selective myeloablative HSCT versus cyclophosphamide; non-selective non-myeloablative HSCT versus cyclophosphamide).

Search methods

We searched for randomized controlled trials (RCTs) in CENTRAL, MEDLINE, Embase, and trial registries from database insertion to 4 February 2022.

Selection criteria

We included RCTs that compared HSCT to immunomodulators in the treatment of SSc.

Data collection and analysis

Two review authors independently selected studies for inclusion, extracted study data, and performed risk of bias and GRADE assessments to assess the certainty of evidence using standard Cochrane methods.

Main results

We included three RCTs evaluating: non-myeloablative non-selective HSCT (10 participants), non-myeloablative selective HSCT (79 participants), and myeloablative selective HSCT (36 participants). The comparator in all studies was cyclophosphamide (123 participants). The study examining non-myeloablative non-selective HSCT had a high risk of bias given the differences in baseline characteristics between the two arms. The other studies had a high risk of detection bias for participant-reported outcomes. The studies had follow-up periods of one to 4.5 years. Most participants had severe disease, mean age 40 years, and the duration of disease was less than three years.

Efficacy

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

No study demonstrated an overall mortality benefit of HSCT when compared to cyclophosphamide. However, non-myeloablative selective HSCT showed overall survival benefits using Kaplan-Meier curves at 10 years and myeloablative selective HSCT at six years. We graded our certainty of evidence as moderate for non-myeloablative selective HSCT and myeloablative selective HSCT. Certainty of evidence was low for non-myeloablative non-selective HSCT.

Event-free survival was improved compared to cyclophosphamide with non-myeloablative selective HSCT at 48 months (hazard ratio (HR) 0.34, 95% confidence interval (CI) 0.16 to 0.74; moderate-certainty evidence). There was no improvement with myeloablative selective HSCT at 54 months (HR 0.54 95% CI 0.23 to 1.27; moderate-certainty evidence). The non-myeloablative non-selective HSCT trial did not report event-free survival.

There was improvement in functional ability measured by the Health Assessment Questionnaire Disability Index (HAQ-DI, scale from 0 to 3 with 3 being very severe functional impairment) with non-myeloablative selective HSCT after two years with a mean difference (MD) of −0.39 (95% CI −0.72 to −0.06; absolute treatment benefit (ATB) −13%, 95% CI −24% to −2%; relative percent change (RPC) −27%, 95% CI −50% to −4%; low-certainty evidence). Myeloablative selective HSCT demonstrated a risk ratio (RR) for improvement of 3.4 at 54 months (95% CI 1.5 to 7.6; ATB −37%, 95% CI −18% to −57%; RPC −243%, 95% CI −54% to −662%; number needed to treat for an additional beneficial outcome (NNTB) 3, 95% CI 2 to 9; low-certainty evidence). The non-myeloablative non-selective HSCT trial did not report HAQ-DI results.

All transplant modalities showed improvement of modified Rodnan skin score (mRSS) (scale from 0 to 51 with the higher number being more severe skin thickness) favoring HSCT over cyclophosphamide. At two years, non-myeloablative selective HSCT showed an MD in mRSS of −11.1 (95% CI −14.9 to −7.3; ATB −22%, 95% CI −29% to −14%; RPC −43%, 95% CI −58% to −28%; moderate-certainty evidence). At 54 months, myeloablative selective HSCT at showed a greater improvement in skin scores than the cyclophosphamide group (RR 1.51, 95% CI 1.06 to 2.13; ATB −27%, 95% CI −6% to −47%; RPC −51%, 95% CI −6% to −113%; moderate-certainty evidence). The NNTB was 4 (95% CI 3 to 18). At one year, for non-myeloablative non-selective HSCT the MD was −16.00 (95% CI −26.5 to −5.5; ATB −31%, 95% CI −52% to −11%; RPC −84%, 95% CI −139% to −29%; low-certainty evidence).

No studies reported data on pulmonary arterial hypertension.

Adverse events

In the non-myeloablative selective HSCT study, there were 51/79 serious adverse events with HSCT and 30/77 with cyclophosphamide (RR 1.7, 95% CI 1.2 to 2.3), with an absolute risk increase of 26% (95% CI 10% to 41%), and a relative percent increase of 66% (95% CI 20% to 129%). The number needed to treat for an additional harmful outcome was 4 (95% CI 3 to 11) (moderate-certainty evidence). In the myeloablative selective HSCT study, there were similar rates of serious adverse events between groups (25/34 with HSCT and 19/37 with cyclophosphamide; RR 1.43, 95% CI 0.99 to 2.08; moderate-certainty evidence). The non-myeloablative non-selective HSCT trial did not clearly report serious adverse events.

Authors' conclusions

Non-myeloablative selective and myeloablative selective HSCT had moderate-certainty evidence for improvement in event-free survival, and skin thicknesscompared to cyclophosphamide. There is also low-certainty evidence that these modalities of HSCT improve physical function. However, non-myeloablative selective HSCT and myeloablative selective HSCT resulted in more serious adverse events than cyclophosphamide; highlighting the need for careful risk–benefit considerations for people considering these HSCTs.

Evidence for the efficacy and adverse effects of non-myeloablative non-selective HSCT is limited at this time. Due to evidence provided from one study with high risk of bias, we have low-certainty evidence that non-myeloablative non-selective HSCT improves outcomes in skin scores, forced vital capacity, and safety.

Two modalities of HSCT appeared to be a promising treatment option for SSc though there is a high risk of early treatment-related mortality and other adverse events.

Additional research is needed to determine the effectiveness and adverse effects of non-myeloablative non-selective HSCT in the treatment of SSc. Also, more studies will be needed to determine how HSCT compares to other treatment options such as mycophenolate mofetil, as cyclophosphamide is no longer the first-line treatment for SSc. Finally, there is a need for a greater understanding of the role of HSCT for people with SSc with significant comorbidities or complications from SSc that were excluded from the trial criteria.

PLAIN LANGUAGE SUMMARY

Stem cell transplantation for the treatment of systemic sclerosis

We conducted a review of the medical literature in February 2022to study the benefits and harms of hematopoietic stem cell transplantation (HSCT) in people with systemic sclerosis (SSc).

What is systemic sclerosis and hematopoietic stem cell transplantation?

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

SSc is an autoimmune disease (where the body's natural defense system attacks normal cells) that affects the skin and internal organs (lungs, digestive system, etc.). Symptoms associated with SSc include thickening of the skin, shortness of breath, digestive symptoms, and difficulties with function or mobility that can affect quality of life. There is also an increased risk of death (mortality) with SSc.

Autologous HSCT is a procedure in which people receive their own healthy stem cells (special cells produced by bone marrow that can turn into different types of blood cell) to replace damaged immune cells that might be causing the disease.

There are two HSCT regimens studied for SSc: myeloablative regimens use either radiation therapy or high doses of chemotherapy that do not allow the bone marrow to recover on its own and non-myeloablative regimens use a lower amount of chemotherapy without radiation therapy, but there are residual cells in the bone marrow afterwards.

There are also two different ways to collect stem cells. Selective HSCT involves a process in which specific stem cells (called CD34+ cells) are chosen for re-infusion. The non-selective process does not include this step.

Is autologous hematopoietic stem cell transplantation a safe and effective treatment for people with systemic sclerosis?

Our review includes three research studies. Each study compared a different type (modality) of stem cell transplantation versus cyclophosphamide (a type of chemotherapy): non-myeloablative non-selective HSCT, non-myeloablative selective HSCT, and myeloablative selective HSCT.

The risk of bias was high for all outcomes with non-myeloablative non-selective HSCT, but mostly low for myeloablative and non-myeloablative selective HSCT. Performance bias (participants knowing which treatment they had) was unclear and detection bias (assessors knowing which treatment the participants had) was high for functional ability as the assessors and participants were not blinded. All participants had early and severe SSc with either skin or lung involvement. The average age of participants was 43 to 47 years, and most were white females. The average duration of disease ranged from 1.5 years to 2 years.

The non-myeloablative selective HSCT trial received CD34+ selection columns from the manufacturer.

Key results

We completed our search in February 2022. Outcomes are at two years for non-myeloablative selective HSCT, 4.5 years for myeloablative selective HSCT, and one year for non-myeloablative non-selective HSCT unless otherwise specified. All comparisons are of stem cell transplantation to cyclophosphamide.

Overall mortality

– There was no difference in mortality between any modality of HSCT and cyclophosphamide.

Event-free survival

– The non-myeloablative selective HSCT group had decreased event-free survival (34% lower risk) at four years and 930 per 1000 people will have been event-free with HSCT at four years.

– There was no change in event-free survival with myeloablative selective HSCT.

Functional ability

– People who received non-myeloablative HSCT had a 13% improvement.

– 53 people out of 100 who receive myeloablative selective HSCT may have meaningful improvement compared to 15 out of 100 who receive cyclophosphamide (37% absolute improvement).

Skin thickening

– People who received non-myeloablative selective HSCT had greater improvement in skin scores (−22% absolute improvement).

– 80 people out of 100 who receive myeloablative selective HSCT may have an improvement in skin scores (25% or greater or 5 points or greater improvement) compared to 53 out of 100 who received cyclophosphamide (27% absolute improvement).

– People who received non-myeloablative non-selective HSCT had greater improvement in their skin scores (−31% absolute improvement).

Serious side effects

– 65 people out of 100 who received non-myeloablative selective HSCT had serious side effects compared to 39 out of 100 who received cyclophosphamide (26% increased absolute risk).

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

– 73 people out of 100 who received myeloablative selective HSCT had serious adverse events compared to 51 out of 100 who received cyclophosphamide (22% increased absolute risk).

– There were no serious side effects reported with non-myeloablative non-selective HSCT or cyclophosphamide.

Certainty of the evidence

We rated the certainty of the evidence from trials using four levels: very low, low, moderate, or high. Very low-certainty evidence means that we are uncertain about the results. High-certainty evidence means that we are very confident in the results.

For non-myeloablative selective HSCT and myeloablative selective HSCT we have moderate certainty in the evidence assessing overall mortality, event-free survival, functional ability, skin thickening, lung function, and serious side effects. The certainty of evidence was downgraded to moderate because of the small number of participants enrolled in the studies and the nature of the studies being unblinded for participant-reported outcomes. For non-selective non-myeloablative HSCT, the certainty of evidence was low for all outcomes. This is because of differences between the cyclophosphamide and HSCT groups before treatment along with the small number of participants enrolled in the study.

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Summary of findings 1. Autologous non-myeloablative selective HSCT compared to cyclophosphamide in systemic sclerosis

Autologous non-myeloablative selective HSCT compared to cyclophosphamide in systemic sclerosis

Patient or population: people with systemic sclerosis
Setting: rheumatology clinics
Intervention: autologous non-myeloablative selective HSCT
Comparison: cyclophosphamide

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	What happens
	Cyclophosphamide	Non-myeloablative selective HSCT				
Overall mortality Follow-up: 2 years	17 per 100	15 per 100 (74 to 31)	**RR 0.90** (0.44 to 1.85)	156 (1 RCT)	⊕⊕⊕⊖ **Moderate**[a]	Moderate-certainty evidence that there is probably no difference in overall mortality compared to cyclophosphamide at 2 years. Absolute change: 2% less risk with HSCT (95% CI 13% less to 10% higher). Relative change is 10% less risk with HSCT (95% CI 56% less to 85% higher). NNTB not applicable.
Event-free survival Follow-up: 4 years	902 per 1000	930 per 1000	**HR 0.34** (0.16 to 0.74)	156 (1 RCT)	⊕⊕⊕⊖ **Moderate**[a]	Moderate-certainty evidence that HSCT probably results in a large increase in event-free survival at 2 years. 930 per 1000 people will have event-free survival with HSCT at 2 years compared to 902 per 1000 people who receive cyclophosphamide.[c]
Functional ability –HAQ-DI Scale 0–3, with 0 representing no/mild impairment and 3 representing very severe impairment.	The mean HAQ-DI was 1.25	**MD 0.39 lower** (0.72 lower to 0.06 lower)	—	131 (1 RCT)	⊕⊕⊖⊖ **Low**[a,b]	Low-certainty evidence that HSCT may result in improvement in HAQ-DI scores. Absolute effect: 13% lower with HSCT (95% CI 24% lower to 2% lower). Relative effect: 27% lower in HSCT (95% CI 50% lower to 4% lower).

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Skin thickness – mRSS The scale is from 0 to 51 in which a higher number is worse skin thickening.	The mean mRSS was −8.8	**MD 11.1 lower** (14.9 lower to 7.3 lower)	—	131 (1 RCT)	⊕⊕⊕⊖ **Moderate**[a]	Moderate-certainty evidence that HSCT probably results in a large reduction in skin thickening compared to cyclophosphamide. Absolute effect: 22% lower with HSCT (95% CI 29% lower to 14% lower). Relative improvement: 43% lower with HSCT (95% CI 58% lower to 28% lower).
Interstitial lung disease – FVC % predicted A lower percentage represents worse lung disease	Mean FVC % predicted was −2.8	**MD 9.1 higher** (3.0 higher to 15.2 higher)	—	131 (1 RCT)	⊕⊕⊕⊖ **Moderate**[a]	Moderate-certainty evidence that HSCT probably results in an increase in FVC compared to HSCT. Relative change: 11% higher with HSCT (95% CI 4% higher to 19% higher).
Pulmonary arterial hypertension	—	—	—	—	—	Not reported.
Serious adverse events Follow-up: 2 years	39 per 100	65 per 100 (47 to 89)	**RR 1.7** (1.2 to 2.3)	156 (1 RCT)	⊕⊕⊕⊖ **Moderate**[a]	Moderate-certainty evidence that HSCT probably has a large risk of serious adverse events compared to cyclophosphamide. Absolute change: 26% increased risk with HSCT (95% CI 10% more to 41% more). Relative change: 66% increased risk with HSCT (95% CI 20% more to 129% more). NNTH with HSCT is 4 (95% CI 3 to 11).

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

CI: confidence interval; **FVC:** forced vital capacity; **HAQ-DI:** Health Assessment Questionnaire Disability Index; **HSCT:** hematopoietic stem cell transplantation; **MD:** mean difference; **mRSS:** Modified Rodnan skin scores; **NNTB:** number needed to treat for an additional beneficial outcome; **NNTH:** number needed to treat for an additional harmful outcome; **RCT:** randomized controlled trial; **RR:** risk ratio.

GRADE Working Group grades of evidence
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[a]Downgraded one level as the sample size is below the 'rule of thumb' suggestion of 400.

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

[b]Downgraded one level as the nature of the intervention does not allow blinding by participants or providers. There is an unclear risk of performance bias and a high risk of detection bias.
[c]The absolute risk of event-free survival was derived from the formula = exp[ln(proportion of participants event-free at four years) × HR] × 1000.

Summary of findings 2. Autologous myeloablative selective HSCT compared to cyclophosphamide in systemic sclerosis

Autologous myeloablative selective HSCT compared to cyclophosphamide in systemic sclerosis

Patient or population: people with systemic sclerosis
Setting: rheumatology clinics
Intervention: autologous myeloablative selective HSCT
Comparison: cyclophosphamide

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	What happens
	Cyclophosphamide	Myeloablative selective HSCT				
Overall mortality Follow-up: 4.5 years	28 per 100	17 per 100 (68 to 40)	**RR 0.59** (0.24 to 1.43)	75 (1 RCT)	⊕⊕⊕⊖ **Moderate**[a]	Moderate-certainty evidence that HSCT likely results in little to no difference in overall mortality at 4.5 years. Absolute change: 12% less risk with HSCT (95% CI 30% less to 7% more). Relative change is 41% less risk with HSCT (95% CI 76% less to 43% more). NNTB not applicable.
Event-free survival Follow-up: 4.5 years	—	—	**HR 0.54** (0.23 to 1.27)	75 (1 RCT)	⊕⊕⊕⊖ **Moderate**[a]	Moderate-certainty evidence that HSCT likely does not improve event-free survival at 4.5 years. 837 per 1000 people will have event-free survival at 4.5 years with HSCT compared to 576 per 1000 people who received cyclophosphamide.
Functional ability – HAQ-DI[b] Scale 0–3 with 0 representing no/mild impairment and 3 representing very severe impairment. The study incorpo-	15 per 100	53 per 100	**RR 3.43** (1.54 to 7.62)	75 (1 RCT)	⊕⊕⊖⊖ **Low**[a,c]	Low-certainty evidence that HSCT may result in a large improvement in the HAQ-DI compared to cyclophosphamide. Absolute improvement: 37% better with HSCT (95% CI 18% better to 57% better). Relative improvement: 243% better in HSCT (95% CI 54% better to 662%). The NNTB is 3 (95% CI 2 to 9).

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

rated a ≥ 0.4 threshold.						
Skin thickness – mRSS Scale 0–51 in which a higher number is worse skin thickening. The study incorporated a ≥ 25% improvement threshold.	53 per 100	80 per 100 (56 to 100)	**RR 1.51** (1.06 to 2.13)	75 (1 RCT)	⊕⊕⊕⊖ **Moderate**[a]	Moderate-certainty evidence that HSCT likely results in a large improvement in skin thickening compared to cyclophosphamide. Absolute improvement: 27% better in HSCT (95% CI 6% better to 47% better). Relative improvement is 51% better in HSCT (95% CI 6% better to 113% better). NNTB 4 (95% CI 3 to 18).
Interstitial lung disease – FVC (% predicted) A lower percentage represents worse lung disease. The study incorporated a ≥ 10% chance	21 per 100	33 per 100 (15 to 72)	**RR 1.63** (0.75 to 3.51)	75 (1 RCT)	⊕⊕⊕⊖ **Moderate**[a]	Moderate-certainty evidence that HSCT likely results in little to no difference in FVC compared to cyclophosphamide. Absolute improvement: 13% higher with HSCT (95% CI 7% lower to 33% higher). Relative improvement is 63% higher with HSCT (95% CI 25% lower to 251% higher). NNTB not applicable.
Pulmonary arterial hypertension	—	—	—	—	—	Not reported.
Serious adverse events Follow-up: 4.5 years	51 per 100	73 per 100 (51 to 100)	**RR 1.43** (0.99 to 2.08)	71 (1 RCT)	⊕⊕⊕⊖ **Moderate**[a]	Moderate-certainty evidence that HSCT likely increases the risk of serious adverse events compared to cyclophosphamide. Absolute change: there is a 22% increased risk with HSCT (95% CI 0% to 44% higher). Relative change: there is a 43% increased risk with HSCT (95% CI 1% lower to 108% higher). NNTH is not applicable.

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

CI: confidence interval; **FVC:** forced vital capacity; **HAQ-DI:** Health Assessment Questionnaire Disability Index; **HR:** hazard ratio; **HSCT:** hematopoietic stem cell transplantation; **mRSS:** Modified Rodnan skin scores; **NNTB:** number needed to treat for an additional beneficial outcome; **NNTH:** number needed to treat for an additional harmful outcome; **RCT:** randomized controlled trial; **RR:** risk ratio.

GRADE Working Group grades of evidence
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[a]Downgraded one level as the sample size is below the 'rule of thumb' suggestion of 400.
[b]Data are reported as dichotomous with thresholds as that is how the paper presented results (mean differences not available).
[c]Downgraded one level as the nature of the intervention does not allow blinding by participants or providers. There is unclear risk of performance bias and high risk of detection bias.

Summary of findings 3. Autologous non-myeloablative non-selective HSCT compared to cyclophosphamide in systemic sclerosis

Autologous non-myeloablative non-selective HSCT compared to cyclophosphamide in systemic sclerosis

Patient or population: people with systemic sclerosis
Setting: rheumatology clinics
Intervention: autologous non-myeloablative non-selective HSCT
Comparison: cyclophosphamide

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	What happens
	Cyclophosphamide	Non-myeloablative non-selective HSCT				
Overall mortality Follow-up: 1 year	0 per 100	0 per 100 (0 to 0)	-	19 (1 RCT)	⊕⊕⊖⊖ **Low**[a,b]	There were 0 deaths reported in either group.
Event-free survival	—	—	—	—	—	Not reported.
Functional ability – HAQ-DI Scale 0–3 with 0 representing no/mild impairment and 3 representing very severe impairment. The study incorporated a ≥ 0.4 threshold	—	—	—	—	—	Not reported.
Skin thickness – mRSS Scale 0–51 in which a higher number is worse skin thickening	The mean mRSS score was 3	**MD 16 lower** (26.5 lower to 5.5 lower)	—	19 (1 RCT)	⊕⊕⊖⊖ **Low**[a,b]	Low-certainty evidence that HSCT may result in a large reduction in skin thickening compared to cyclophosphamide.

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

						Absolute improvement: 31% better with HSCT (95% CI 52% better to 11% lower). Relative improvement: 84% better in HSCT (95% CI 139% better to 29% better).
Interstitial lung disease – FVC % predicted A lower percentage represents worse lung disease. The study incorporated a ≥ 15% chance	Mean FVC % predicted was −6	**MD 18 higher** (1.8 higher to 34.2 higher)	—	19 (1 RCT)	⊕⊕⊖⊖ **Low**[a,b]	Low-certainty evidence that HSCT may result in a large reduction of FVC compared to cyclophosphamide. Relative change: 27% higher with HSCT (95% CI 2% higher to 51% higher).
Pulmonary arterial hypertension	—	—	—	—	—	Not reported.
Serious adverse events Follow-up: 1 year	0 per 100	0 per 100 (0 to 0)	-	19 (1 RCT)	⊕⊕⊖⊖ **Low**[a,b]	There were no serious adverse events reported in either arm.

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

CI: confidence interval; **FVC:** forced vital capacity; **HAQ-DI:** Health Assessment Questionnaire Disability Index; **HSCT:** hematopoietic stem cell transplantation; **MD:** mean difference; **mRSS:** Modified Rodnan skin scores; **RCT:** randomized controlled trial; **RR:** risk ratio.

GRADE Working Group grades of evidence
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[a]Downgraded one level as data provided by one study with small number of participants (19).
[b]Downgraded one level as there were significant differences in baseline characteristics between the two study populations.

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

BACKGROUND

Description of the condition

Systemic sclerosis (SSc), or scleroderma, is a rare chronic autoimmune disease characterized by fibrosis and vascular injury. It is estimated that the annual incidence of SSc is 20 cases per million adults in the US (Mayes 2003; Bergamasco 2019). This disease leads to multi-organ damage including fibrosis of the skin and subcutaneous tissue, Raynaud phenomenon, interstitial lung disease, pulmonary hypertension, renal crisis, and other serious complications (Hudson 2014; Jordan 2015; Park 2015; Steen 2000). SSc is divided into different subsets according to clinical features most notably limited and diffuse cutaneous SSc (LeRoy 1988). The limited cutaneous subset is usually distinguished by Raynaud phenomenon preceding skin manifestations over an extended period of time, delayed incidence of pulmonary hypertension with less frequent interstitial lung disease, presence of anticentromere antibodies and a more favorable prognosis than the diffuse cutaneous subset. The characteristics of the diffuse subset include Raynaud phenomenon with early and extensive skin involvement, early systemic involvement (e.g. interstitial lung disease, renal crisis), tendon friction rubs, presence of antitopoisomerase antibodies or antiribonucleic acid (anti-RNA) polymerase III antibodies, and poorer prognosis. Treatment of SSc is predominantly directed toward preventing or decreasing vascular-related injuries (i.e. pulmonary hypertension, Raynaud phenomenon, and renal crisis) and fibrotic disease (i.e. skin manifestations and other internal organ involvement) via vasoactive and immunosuppressive drugs (Kowal-Bielecka 2009). The vasoactive therapies of SSc are utilized to treat Raynaud phenomenon and pulmonary arterial hypertension. Immunomodulators have been used to try to alter the disease course and improve fibrotic disease, including skin sclerosis and pulmonary interstitial fibrosis (Nagaraja 2015).

Description of the intervention

Since the early 2010s, high-dose chemotherapy followed by autologous hematopoietic stem cell transplantation (HSCT) has emerged as a new potential intervention for severe SSc refractory to conventional treatments (Snowden 2012). Hematopoietic stem cells are immature CD34+ cells capable of differentiating into multiple cell types including mature B-lymphocytes, T-lymphocytes, and macrophages. HSCT is a complex process with multiple steps. The first step is the mobilization of hematopoietic stem cells from the peripheral blood. This is accomplished with cyclophosphamide and granulocyte colony-stimulating factor (G-CSF) or G-CSF alone. After mobilization, stem cells are collected via leukopheresis. After collection of peripheral blood stem cells (PBSCs), the stem cells may be cryopreserved or further purified to reduce lymphocytes (select for CD34+ cells) that may be pathogenic in a process termed 'selection'. Several weeks after collection of PBSCs, a person is conditioned with either non-myeloablative or myeloablative techniques. Non-myeloablative techniques are accomplished using chemotherapy (most commonly high-dose cyclophosphamide) with or without antithymocyte globulin (ATG), whereas myeloablative techniques utilize these medications plus total body irradiation. After conditioning, the person is infused with their own hematopoietic stem cells to reconstitute the immune system. The three types of HSCT that have been utilized in the treatment of SSc are: non-myeloablative non-selective HSCT, non-myeloablative selective HSCT, and selective myeloablative HSCT.

How the intervention might work

The key mechanism by which HSCT is thought to work is by 'resetting' the immune system. The conditioning of the person causes severe immune suppression intended to rid the body of proinflammatory, profibrotic and autoreactive cells. The transplanted stem cells in turn differentiate into mature lymphocytes and monocytes that are less likely to be destructive (van Laar 2013).

Why it is important to do this review

SSc (especially the diffuse type) is a disease that carries poor prognosis with few treatment options (mainly cyclophosphamide, azathioprine, and mycophenolate mofetil) that decrease morbidity and mortality. Two meta-analyses examining the mortality of SSc have shown that the disease carries a standardized mortality ratio (SMR) of 3.5 (95% confidence interval (CI) 3.0 to 4.1, and 2.7 to 4.5) (Elhai 2012; Toledano 2012). Immunomodulatory therapy with cyclophosphamide has been shown to improve skin sclerosis and stabilize pulmonary function, though its effects have been modest (Poormoghim 2012; Tashkin 2006). Some immunomodulators (prednisone and cyclosporine, which consequently are not used) may precipitate renal crises in people at risk (Kowal-Bielecka 2009; Steen 1998). Early studies showed that HSCT had significant improvements in modified Rodnan skin scores (mRSS) with a potential benefit in mortality (Binks 2001; Farge 2002; Nash 2007; Oyama 2007; van Laar 2013). However, the effects on pulmonary function are less clear and have ranged from improvement to deterioration of pulmonary function. Therefore, it is imperative to investigate the effects of HSCT in detail to determine the potential benefits and harms of this therapy as it is already being utilized in people with SSc. Thus, we summarized the evidence to investigate the benefits and harms of HSCT. As a result, a systematic review will aid clinicians in determining if HSCT is a viable option in people with SSc.

OBJECTIVES

To assess the benefits and harms of autologous hematopoietic stem cell transplantation for the treatment of systemic sclerosis. Specifically, we wanted to assess the following:

- Non-selective myeloablative HSCT versus cyclophosphamide
- Selective myeloablative HSCT versus cyclophosphamide
- Non-selective non-myeloablative HSCT versus cyclophosphamide

METHODS

Criteria for considering studies for this review

Types of studies

We included randomized controlled trials (RCTs) reported as full-text, as abstract only, and unpublished. There was no language restriction.

Types of participants

We included people ages at least 18 years with a diagnosis of SSc (diffuse or limited type) as defined by the trial authors,

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

with cutaneous or pulmonary involvement, or both. It was not necessary for participants to fulfil the preliminary American College of Rheumatology (ACR) criteria for scleroderma (ACR 1980); or 2013 American College of Rheumatology/European Alliance of Associations for Rheumatology (ACR/EULAR) Classification Criteria for Scleroderma (van den Hoogen 2013). Trials could include people with any subset of scleroderma.

Types of interventions

We included trials comparing all forms of autologous HSCT with immunomodulatory therapy with cyclophosphamide or other immunomodulators alone or in combination.

Types of outcome measures

We based outcome measures on the suggested outcome measures from the Outcome Measures in Rheumatology (OMERACT) initiative (Distler 2008; Khanna 2009; Merkel 2003); and from discussion with experts.

Major outcomes

- Overall mortality – measured as the total number of deaths from the HSCT group and the comparator.
- Event-free survival – defined as survival without significant organ damage where events included any of the following:
  - death;
  - respiratory failure as defined by:
    - a significant decrease of more than 30% in diffusing in liters for carbon monoxide capacity (DLCO) or a decrease in the forced vital capacity (FVC) of more than 20% predicted and
    - resting arterial partial pressure of oxygen (pO_2) less than 60 mmHg or partial pressure of carbon dioxide (pCO_2) more than 50 mm Hg without supplemental oxygen;
  - renal failure as defined by use of dialysis for more than six months or transplantation;
  - occurrence of cardiomyopathy as defined by left ventricular ejection fraction (LVEF) of less than 30%. Typically, event-free survival is measured as the proportion of participants who did not have an 'event' over the time period conducted in the study.
- Functional ability – measured by the Health Assessment Questionnaire Disability Index (HAQ-DI). The HAQ-DI scores are measured on a scale from 0 to 3. Typically, 0 to 1 represents mild to moderate impairment, 1 to 2 represents moderate to severe impairment, and 2 to 3 represents severe to very severe disability. The minimally clinically important difference for improvement in HAQ-DI in people with SSc has been estimated to be around −0.04 whereas the minimally clinically important difference for worsening has been estimated to be 0.14 (Sekhon 2010).
- Skin thickness – measured by the modified Rodnan skin score (mRSS). The scale monitors changes in skin sclerosis from before treatment to after treatment by clinical palpation at 17 different anatomical sites. Each site is graded on a scale from 0 to 3 in which 0 is no thickening, 1 is mild thickening, 2 is moderate thickening, and 3 is severe thickening. The scale is from 0 to 51 in which a higher number typically represents worse skin thickening. The minimal clinically important difference for mRSS has typically been defined as 3- to 5-point improvement in the scale (Khanna 2019).
- Interstitial lung disease – pulmonary function tests including FVC and DLCO % predicted. The change in FVC has typically been used as the primary outcome measure in SSc trials that study interstitial lung disease as the primary outcome. A lower FVC % predicted typically represents worse interstitial lung disease. One study in people with SSc found that an improvement of 3.0% to 5.3% is the minimally clinically important difference for FVC (Kafaja 2017). The DLCO % predicted can be used as both a marker of pulmonary hypertension and interstitial lung disease. A lower DLCO % predicted is associated with worse overall respiratory function.
- Pulmonary arterial hypertension – measured by mean pulmonary arterial pressure, time to diagnosis, or percent of people with diagnosis. People with an increased pulmonary arterial pressure have more severe pulmonary arterial hypertension. Echocardiography is suggestive of pulmonary arterial hypertension, but a diagnosis can only be made with right heart catheterization.
- Serious adverse events including, but not limited to, serious infections, renal failure, and deaths.

Minor outcomes

- Percent predicted total lung capacity (TLC) – maximum amount of air that is present in the lungs after inspiration. A lower TLC% is suggestive of worsening interstitial lung disease.
- Renal function – measured by creatinine clearance, estimated glomerular filtration rate (eGFR), and serum creatinine (percent of people with serum creatinine above normal). People with a lower creatinine clearance, lower eGFR, and high serum creatinine have worse renal function.
- Cardiac function – measured by echocardiogram (to detect changes in LVEF). A lower ejection fraction is correlated with worse cardiac function.
- Health-related quality of life (HRQoL) – including pain measured by a visual analog scale (VAS) and the Medical Outcomes Study Short Form 36 survey (SF-36). Pain measured by the VAS is typically graded from 0 to 100 in which 100 is very severe. The SF-36 is a composed of 36 items in eight domains and is typically presented as a Physical Component Summary and a Mental Component Summary. A lower score is suggestive of greater disability. The clinically important difference for improvement in the SF-36 was 2.18 for the Physical Component Summary and 1.33 for the Mental Component Summary (Sekhon 2010).
- Safety outcomes – measured by withdrawals from study and adverse events reported as defined by the study authors.
- Inflammatory markers – measured by erythrocyte sedimentation rate (ESR) or C-reactive protein (CRP), or both. High inflammatory markers are potentially suggestive of high disease activity.

Search methods for identification of studies

Electronic searches

We searched the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, Embase, and Web of Science on 4 February 2022.

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

We also searched ClinicalTrials.gov (www.clinicaltrials.gov) and the World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) (www.who.int/ictrp/en/). We searched all databases from their inception, and imposed no restriction on language of publication. The specific search strategy was constructed according to the Cochrane Musculoskeletal Group methods used in reviews (Appendix 1).

Searching other resources

We checked the reference list of all primary studies and review articles for additional relevant citations otherwise not found. We did not have to contact authors of studies for additional information that had not been published, or for further clarification of trial information.

Data collection and analysis

We used EndNote X9 software to manage records retrieved from searches of the electronic databases (EndNote X9). Results from resources not compatible with EndNote were managed on a Microsoft Excel spreadsheet.

Selection of studies

Three review authors (SB, HRS, MLO) independently screened titles and abstracts for inclusion of all the potentially relevant citations using DistillerSR software. The authors identified and removed duplicates. Each unique citation was coded as 'retrieve' (eligible or potentially eligible/unclear) or 'do not retrieve'. Reasons for ineligibility were recorded.

We retrieved the full-text reports/publication of the potentially eligible studies and two review authors (SB, MLO) independently screened them for inclusion. Reasons for exclusion of the ineligible studies were recorded. We resolved disagreements through discussion or, when required, with the consultation of a third person (MSA). Multiple publications of the same study were collated so that each study, rather than each publication, was the unit of interest in the review. We recorded the selection process in sufficient detail to complete a PRISMA flow diagram.

Data extraction and management

We created a data collection form for study characteristics and outcome data. Two review authors (SB, MLO) independently extracted the following study characteristics and outcome data from included studies.

- Methods: study design, total duration of study, details of any 'run in' period, number of study centers and location, study setting, withdrawals, and date of study.
- Participants: number, mean age, age range, sex, disease duration, severity of condition, diagnostic criteria, important SSc baseline data; inclusion criteria, and exclusion criteria.
- Interventions: intervention including method of mobilization and conditioning, comparison, concomitant medications, and excluded medications.
- Outcomes: primary and secondary outcomes specified and collected, and time points reported. Number of events and number of participants per treatment group for dichotomous outcomes, and means, standard deviations (SD), and number of participants per treatment group for continuous outcomes.
- Characteristics of the design of the trial as outlined in the Assessment of risk of bias in included studies section.
- Notes: funding for trial, and notable declarations of interest of trial authors.

We resolved disagreements by consensus or by involving a third person (MSA). One review author (SB) transferred data into Review Manager 5 (Review Manager 2014), and other review authors (HRS, MLO) double-checked that the data were entered correctly by comparing the data against the trial reports.

Assessment of risk of bias in included studies

Three review authors (SB, HRS, MLO) independently assessed the risk of bias for each study by following Cochrane's recommendations for assessment (Higgins 2017). Disagreements were resolved through discussion. Any disagreements that persisted after discussion were resolved by authors (MLO; MSA). We summarized the risk of bias assessment for every outcome included in the summary of findings tables within a study. We assessed the risk of bias according to the following domains

- Random sequence generation.
- Allocation concealment.
- Blinding of participants and personnel.
- Blinding of outcome assessment.
- Incomplete outcome data.
- Selective outcome reporting.
- Other bias such as baseline imbalance and blocked randomization.

We graded each potential source of bias as low, unclear, or high, and presented justification for each judgment in the Characteristics of included studies risk of bias tables. We summarized the risk of bias judgments across different studies for each of the domains listed. We considered blinding separately for different key outcomes where necessary (e.g. for unblinded outcome assessment, risk of bias for all-cause mortality may be very different from for a participant-reported pain scale). As well, we considered the impact of missing data by key outcomes.

When considering treatment effects, we took into account the risk of bias for the studies that contributed to that outcome.

We presented the figures generated by the risk of bias tool to provide summary assessments of the risk of bias.

Assessment of bias in conducting the systematic review

We conducted the review according to a published protocol (Bruera 2015), and reported any deviations from it in the Differences between protocol and review section of the systematic review.

Measures of treatment effect

For both efficacy and safety outcomes, we analyzed data based on an intention-to-treat sample. For dichotomous variables, we determined the risk ratio (RR); or Peto odds ratio (Peto OR) in the case of rare events (less than 10%.) We analyzed continuous data as mean difference (MD) and 95% CI for measures of treatment effect. We utilized final values unless not reported. If not reported, we used mean changes from baseline. When studies used different scales to measure the same conceptual outcome (e.g. disability), we calculated standardized mean differences (SMD) instead, with

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

corresponding 95% CI. SMDs were back-translated to a typical scale (e.g. 0 to 10 for pain) by multiplying the SMD by a typical among-person SD (e.g. the SD of the control group at baseline from the most representative trial) (Higgins 2021). For pooling hazard ratios (HR), we calculated the log HR using the inverse variance method and the corresponding 95% CI.

Absolute differences: for dichotomous outcomes, we calculated the absolute risk difference (ARD) using the risk difference statistic in Review Manager 2014 and expressed the result as a percentage. For continuous outcomes, we calculated the absolute benefit as the improvement in the intervention group minus the improvement in the control group, in the original units.

Relative differences: we calculated the relative percent change for dichotomous data as the RR − 1 and expressed as a percentage. For continuous outcomes, we calculated the relative difference in the change from baseline as the absolute benefit divided by the baseline mean of the control group.

For dichotomous outcomes, such as serious adverse events, we calculated the number needed to treat for an additional beneficial outcome (NNTB) or for an additional harmful outcome (NNTH) from the control group event rate and the relative risk using the Visual Rx NNT calculator (Cates 2008).

Unit of analysis issues

Where a trial reported multiple intervention groups, we included only the relevant groups in analyses but listed additional interventions in the Characteristics of included studies table.

Dealing with missing data

We noted in the Characteristics of included studies table if outcome data were not reported in a usable way and when data were transformed or estimated from a graph.

For dichotomous outcomes (e.g. number of withdrawals due to adverse events), we calculated the withdrawal rate using the number of participants randomized in the group as the denominator.

For continuous outcomes (e.g. mean change in skin scores), we calculated the MD or SMD based on the number of participants analyzed at that time point. If the number of participants analyzed was not presented for each time point, we used the number of randomized participants in each group at baseline.

We calculated missing SD from other statistics such as standard errors, CIs, or P values according to the methods recommended in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2021). If SDs could not be calculated, we imputed them (e.g. from other studies in the meta-analysis).

Assessment of heterogeneity

We assessed clinical and methodological diversity in terms of participants, interventions, outcomes, and study characteristics for the included studies to determine whether a meta-analysis was appropriate. This was conducted by observing these data from the data extraction tables. We assessed statistical heterogeneity by visual inspection of the forest plot to assess for obvious differences in results between the studies, and using the I^2 and Chi^2 statistical tests (Deeks 2021).

As recommended in Section 10.10 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Deeks 2021), the interpretation of an I^2 statistic of 0% to 40% might not be important; 30% to 60% may represent moderate heterogeneity; 50% to 90% may represent substantial heterogeneity; and 75% to 100% represents considerable heterogeneity. As noted in the *Cochrane Handbook for Systematic Reviews of Interventions*, we kept in mind that the importance of the I^2 statistic depends on the magnitude and direction of effects; and strength of evidence for heterogeneity.

For the Chi^2 test, a P value less than or equal to 0.10 indicated evidence of statistical heterogeneity.

If we identified substantial heterogeneity, we reported it and investigated possible causes by following the recommendations in Section 10 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Deeks 2021).

Assessment of reporting biases

We planned to create and examine a funnel plot to explore possible small-study biases if we had at least 10 studies in a meta-analysis. In interpreting funnel plots, we would have examined the different possible reasons for funnel plot asymmetry as outlined in Section 13 of the *Cochrane Handbook for Systematic Reviews of Interventions* and relate this to the results of the review (Page 2021). If we are able to pool more than 10 trials, we would have undertaken formal statistical tests to investigate funnel plot asymmetry and followed the recommendations in Section 13.3 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Page 2021). To assess outcome reporting bias, we checked trial protocols against published reports. For studies published after 1 July 2005, we screened ClinicalTrials.gov and WHO ICTRP trial registers for the a priori trial protocol to evaluate whether there was selective reporting of outcomes.

Data synthesis

We used Review Manager 5 for analyses (Review Manager 2014). Meta-analyses were not conducted as the studies were not sufficiently homogeneous due to different interventions. We reported on the clinical significance of the findings; for dichotomous outcomes, an ATB greater than 10% indicated clinical significance. For continuous variables, with no previously reported clinically important threshold we used the SMD interpretation where values greater than 0.8 were considered clinically significant (large effect).

The primary analysis for self-reported outcomes (e.g. HRQoL, function) was restricted to trials with low risk of detection and selection bias. If there had been multiple time points available in the studies, then we would have analyzed them all. However, we used the final assessment reported in each individual trial for the summary of findings tables. We planned to pool all time points (from short- and long-duration studies).

Subgroup analysis and investigation of heterogeneity

We performed no subgroup analyses as data were not available. We had planned the following subgroups.

- Myeloablative and non-myeloablative conditioning techniques.
- Diffuse and limited subsets of SSc.

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

- Disease duration categorized as short duration (12 months or less) and long duration (more than 12 months).

Sensitivity analysis

We conducted no sensitivity analyses as there were insufficient trials.

Interpreting results and reaching conclusions

We followed the guidelines in the *Cochrane Handbook for Systematic Reviews of Interventions* for interpreting results and maintained awareness of distinguishing a lack of evidence of effect from a lack of effect (Schünemann 2021a). We based our conclusions only on findings from the quantitative or narrative synthesis of included studies for this review. We avoided making recommendations for practice and our implications for research suggested priorities for future research and outlined what the remaining uncertainties are in the area.

Summary of findings and assessment of the certainty of the evidence

We created three summary of findings tables based on the comparison of HSCT versus cyclophosphamide using the following outcomes: overall mortality, event-free survival, functional ability, skin thickness, interstitial lung disease, pulmonary arterial hypertension, and serious adverse events.

Two review authors (SB, MLO) independently assessed the certainty of the evidence. We used the five GRADE considerations (study limitations, consistency of effect, imprecision, indirectness, and publication bias) to assess the certainty of the body of evidence as it related to the studies which contributed data to the meta-analyses for the prespecified outcomes. We used GRADEpro GDT software to prepare the summary of findings tables (GRADEpro GDT).

We used methods and recommendations described in Chapter 14 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Schünemann 2021b). We justified all decisions to downgrade or upgrade the certainty of evidence of studies using footnotes and we made comments to aid the reader's understanding of the review where necessary. In the 'What happens' column of the summary of findings tables, we provided the absolute percent difference, the relative percent change from baseline, and the NNTB.

RESULTS

Description of studies

The study characteristics are summarized in the Characteristics of included studies table and below.

Results of the search

We retrieved 4205 citations through database searching and six additional records identified through other sources From these, we analyzed 3024 records after removal of duplicates. We eliminated 2837 records based on title and abstract alone. We assessed the remaining 187 full-text citations for inclusion. Of these 187 citations, 14 records (three studies) met our inclusion criteria. From these three studies, two reported results from the Autologous Stem Cell Transplantation International Scleroderma (ASTIS, van Laar 2014) trial; eight were reports from the Scleroderma: Cyclophosphamide or Transplantation (SCOT) trial; and one from the Autologous non-myeloablative haemopoietic stem-cell transplantation compared with pulse cyclophosphamide once per month for SSc (ASSIST, Burt 2011) trial. Three records were protocols for the ongoing UPfront autologous hematopoietic Stem cell transplantation vs Immunosuppressive medication in early DiffusE cutaneous systemic sclerosis trial (UPSIDE; van Laar 2020).

Included studies

We included three RCTs. A full description of all included trials is provided in the Characteristics of included studies table. The ASSIST trial compared autologous non-myeloablative non-selective HSCT versus cyclophosphamide (Burt 2011); the ASTIS trial compared autologous non-myeloablative selective HSCT versus cyclophosphamide (van Laar 2014); and the SCOT trial compared autologous myeloablative selective HSCT versus cyclophosphamide (Sullivan 2018).

Study design and setting

All studies were RCTs. The ASTIS trial was a multicenter RCT conducted across 10 countries and 29 centers in Europe (van Laar 2014). The trial was supported by the European Group for Blood and Marrow Transplant, the European League Against Rheumatism, the Assistance Publique-Hôpitaux de Paris, French Ministry of Health Programme Hospitaller de Recherche Clinique, Groupe Francophone de Recherche sur la Sclérodermie, the Association des sclérodermiques de France, the National Institute for Health Research, and grants from Imtix-Sangstat and Amgen Europ, and Miltenyi-Biotec. The SCOT trial was conducted at 26 sites in North America (Sullivan 2018). The trial was primarily funded by the National Institute of Allergy and Infectious Diseases and the National Institutes of Health. The ASSIST trial was a single-center trial in Chicago, USA (Burt 2011). This was a cross-over trial design where participants in the cyclophosphamide arm were allowed to switch to HSCT after one year if there was no clinical response. There was no funding source reported.

Participants

The ASTIS trial enrolled 156 participants: 79 in the HSCT arm and 77 in the cyclophosphamide arm (van Laar 2014). The mean age overall was 44 years, 59% of participants were women, 81% were white, and the average disease duration in years was 1.4. Baseline clinical characteristics between the two groups were similar.

The SCOT trial enrolled 75 participants – 36 in the HSCT arm and 39 in the cyclophosphamide group (Sullivan 2018). The mean age was 46 years, 64% were women, 80% were white, and average disease duration was 27.1 months. Baseline clinical characteristics between the two groups were similar.

The ASSIST trial included 19 participants: 10 in the HSCT arm and nine in the cyclophosphamide arm (Burt 2011). The mean age was 45 years, 89% were women, and 79% were white. The mean disease duration ranged from 13.6 months (HSCT) to 18 months (cyclophosphamide). Baseline mRSS was different between the HSCT group (28) and the cyclophosphamide group (19). The DLCO was also different between the HSCT (58%) and cyclophosphamide (75%) groups.

Inclusion and exclusion criteria of included studies were similar (see Characteristics of included studies table). All trials included people aged at least 18 years to about 60 to 69 years, with an SSc diagnosis according to the ACR criteria, disease duration

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

for less than five years, and extensive (diffuse) skin and internal organ involvement. The ASSIST trial excluded people with a mean pulmonary artery pressure greater than 25 mmHg (Burt 2011). The ASTIS trial excluded people with mean pulmonary pressure greater than 50 mmHg (van Laar 2014). The SCOT trial excluded people with a pulmonary artery peak systolic pressure greater than 55 mmHg by echocardiogram or greater than 25 mmHg by right heart catheterization (Sullivan 2018).

Interventions

All trials utilized cyclophosphamide as the comparator.

The ASTIS trial used non-myeloablative selective HSCT (van Laar 2014).

The comparator group received 12 monthly pulses of 750 mg/m^2 intravenously.

The protocol for HSCT was:

- mobilization: intravenous cyclophosphamide 4 g/m^2 over two days and subcutaneous filgrastim 10 µg/kg followed by leukapheresis and enrichment for CD34+ cells using immunomagnetic separation;
- conditioning: intravenous cyclophosphamide 200 mg/kg over four days and intravenous rabbit ATG 7.5 mg/kg administered over three days with intravenous methylprednisolone 1 mg/kg; and
- stem cell infusion of autologous CD34+ stem cells.

The SCOT trial used myeloablative selective HSCT (Sullivan 2018).

The comparator group received an initial intravenous dose of cyclophosphamide 500 mg/m^2 followed by 11 monthly infusions of 750 mg/m^2.

The protocol for HSCT was:

- mobilization with G-CSF and leukapheresis with CD34+ cell enrichment;
- conditioning: fractionated total-body irradiation (with pulmonary and renal shields limiting exposure), cyclophosphamide 120 mg/kg administered over two days, and equine ATG 90 mg/kg over six days with intravenous methylprednisolone 1 mg/kg prior to each dose; and
- stem cell infusion of autologous CD34+ stem cells.

The ASSIST trial used non-myeloablative non-selective HSCT (Burt 2011).

The comparator group received cyclophosphamide 1 g/m^2 monthly for six cycles.

The protocol for HSCT was:

- mobilization: cyclophosphamide 2 g/m^2 over two days and subcutaneous filgrastim 10 µg/kg from day five after cyclophosphamide administration until apheresis and cells were cryopreserved without selection;
- conditioning: intravenous cyclophosphamide 200 mg/kg given in four equal fractions before stem cell infusion. Rabbit ATG initially at 0.5 mg/kg followed by 1.5 mg/kg for four days administered along with methylprednisolone 1000 mg intravenously prior to infusion of stem cells; and
- infusion of non-selected autologous stem cells from mobilization.

Outcomes

The primary endpoint of the ASTIS trial was event-free survival defined as time in days from randomization until the occurrence of death due to any cause or the development of persistent major organ failure defined as: LVEF less than 30%; resting arterial oxygen tension less than 60 mmHg or resting arterial carbon dioxide tension greater than 50 mmHg without oxygen supply (or both); and the need for renal replacement therapy (van Laar 2014). Secondary end points were treatment-related mortality, serious adverse events, changes in mRSS, quality of life (HAQ-DI and SF-36), body weight, creatinine clearance, FVC, TLC, residual volume, DLCO, EQ-5D VAS score, and index-based utility score.

The endpoint in the SCOT trial was a global rank composite score at 54 months with a secondary endpoint of the global rank composite score at 48 months (Sullivan 2018). The global rank composite score combined mortality and other longitudinal outcomes (failure of event-free survival, FVC, HAQ-DI, and mRSS) in a hierarchical order (in the order outcomes were mentioned) and then depending on the participant's improvement, stability, or worsening with those parameters a global rank composite score was calculated. Initially, the study was powered to detect event-free survival at 54 months but because of slow enrollment, the trial was redesigned with the global rank composite score as the primary endpoint. All trial investigators were blinded to data during the redesign. Other secondary outcomes included event-free survival at 48 and 54 months, FVC, quality of life (HAQ-DI, SF-36), skin thickness (mRSS), DLCO, SF-36, serious adverse events, and treatment-related mortality.

The primary outcome of the ASSIST trial primary outcome was improvement at 12 months defined as: a decrease in mRSS (greater than 25%) with an initial mRSS greater than 14 or an increase in FVC greater than 10% (Burt 2011). Other secondary outcomes included mean differences of the following: FVC, TLC, DLCO, volume of diseased lung on computer tomography scan, skin thickness (mRSS), and quality of life (SF-36). Of note, this trial did not use the HAQ-DI.

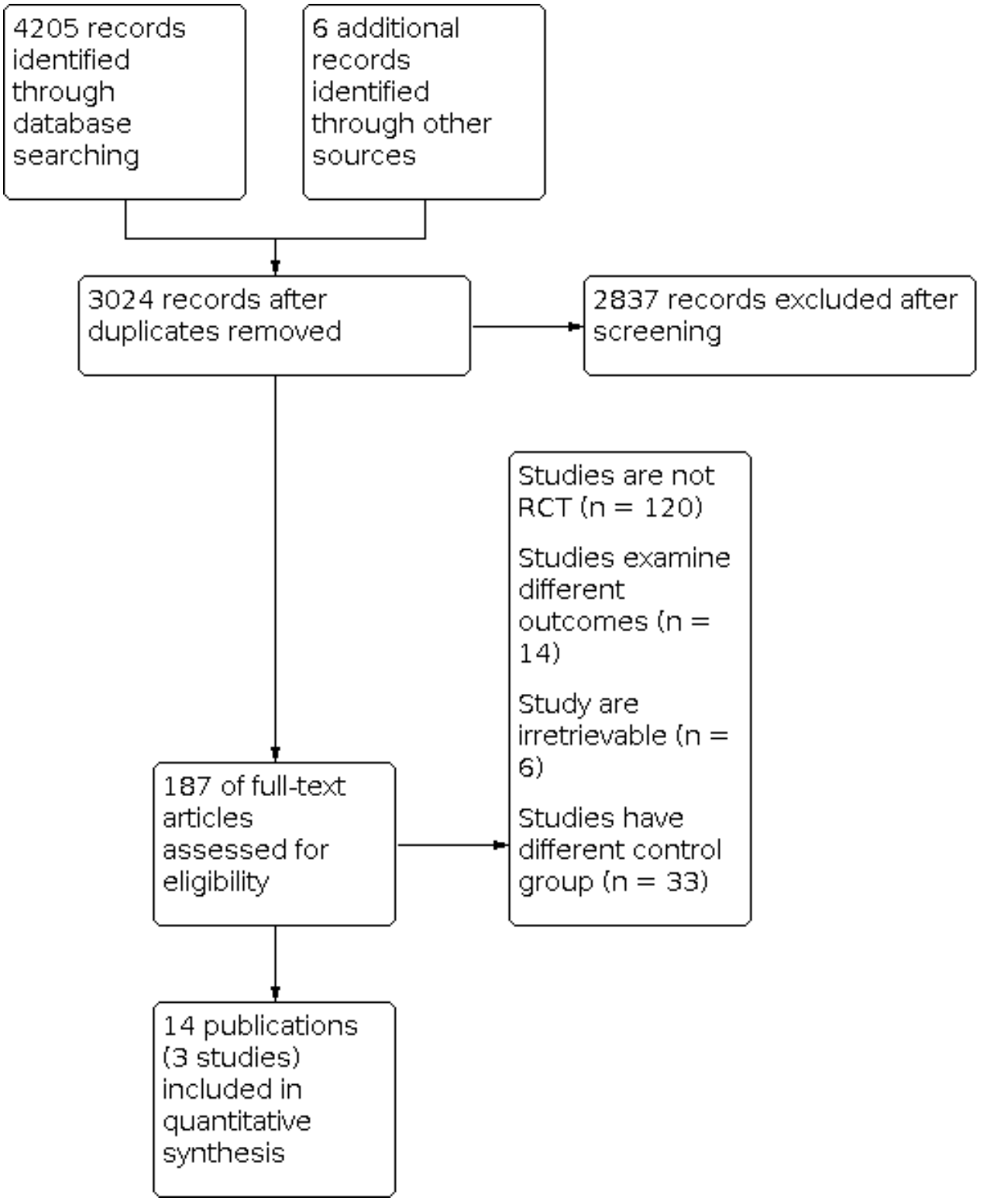
No trial measured pulmonary hypertension or inflammatory markers.

Excluded studies

We excluded 2,837 records on initial citation and abstract screening as they did not meet the inclusion criteria. We excluded 173 citations out of which 120 were excluded because they were not RCTs, 14 were excluded because they examined different outcomes, 6 were irretrievable, and 33 had different control groups. Fourteen citations met our inclusion criteria, and only three studies were included (the other 11 citations were abstracts from these main studies. (Figure 1).

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Figure 1. Study flow diagram.



Studies awaiting classification

There are no studies awaiting classification.

Ongoing studies

There is currently one multicenter, randomized, open-label trial comparing upfront autologous HSCT versus usual care

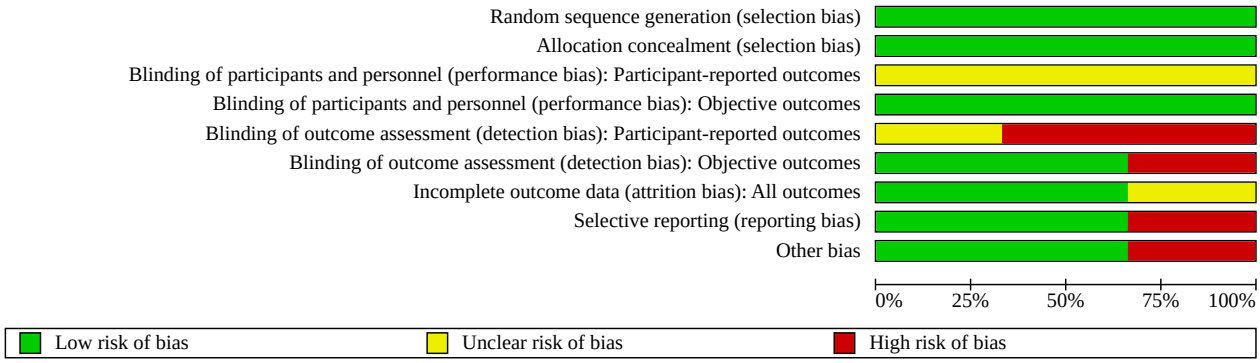
Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

(intravenous cyclophosphamide followed by mycophenolate mofetil with HSCT as a rescue option) (van Laar 2020).

Risk of bias in included studies

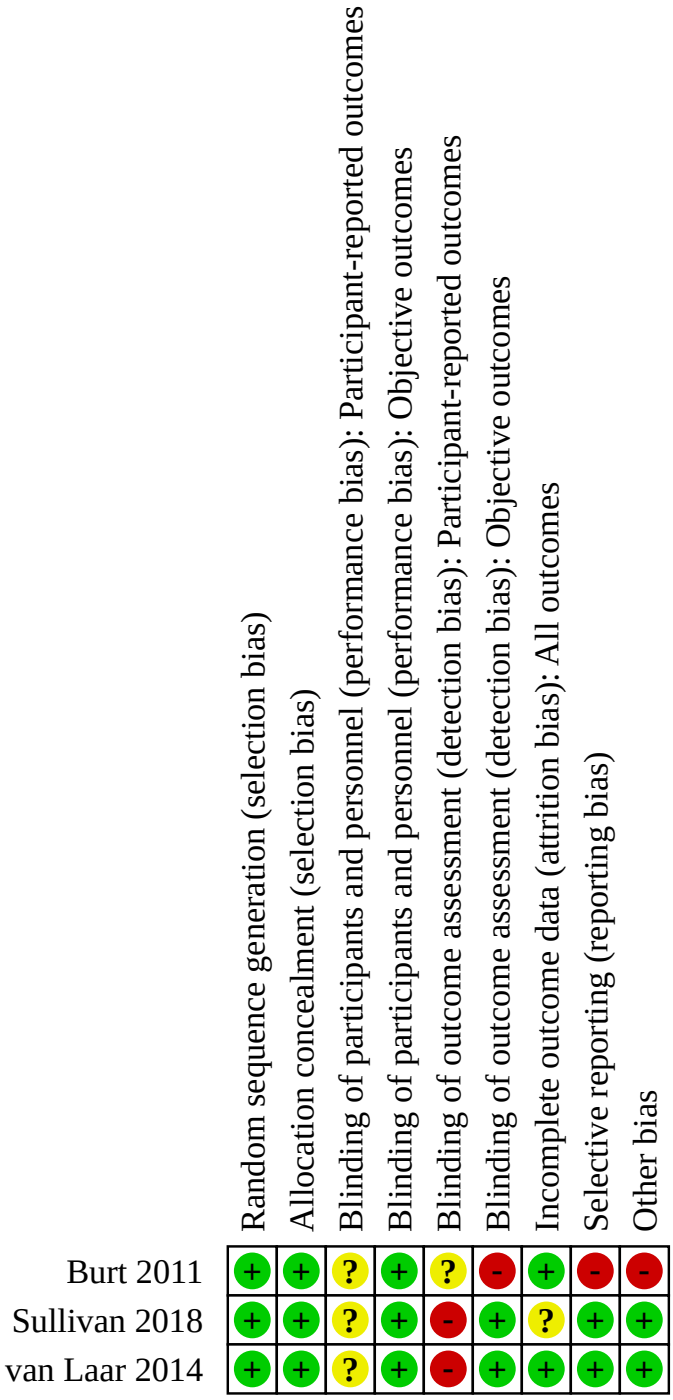
Summaries of risk of bias of the three included trials is shown in the risk of bias figures (Figure 2, Figure 3.)

Figure 2. Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.



Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Figure 3. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.



Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Allocation

All three trials reported adequate sequence generation via computer-generated randomized sequences with participants randomly allocated in a one-to-one ratio with a mixed block design (Burt 2011; Sullivan 2018; van Laar 2014). Therefore, selection bias was low risk.

Blinding

Due to the nature of the treatments (stem cell transplantation versus intravenous cyclophosphamide) blinding was not possible in any study, but was also unlikely to alter non-participant-reported outcomes (such as those measuring end-organ damage). Therefore, the risk of performance bias for outcomes including mortality, event-free survival, mRSS, pulmonary function tests, and serious adverse events were low. The HAQ-DI is a participant-reported outcome. Since these studies were not blinded, the responses to the questionnaire may potentially be influenced by the intervention that participants received. Therefore, performance bias was unclear for this outcome. The SF-36 was a minor outcome in our review, and was participant-reported; therefore, performance bias was unclear.

Detection bias was at low risk for objective outcomes including mortality, event-free survival, mRSS, pulmonary function tests, and serious adverse events in Sullivan 2018 and van Laar 2014, and was high risk in Burt 2011. However, for the HAQ-DI and the SF-36, which are both participant-reported outcomes, the risk of bias was high in Sullivan 2018 and van Laar 2014 due to assessors not being blinded to the intervention that the participant received and could potentially have influence subjective outcomes as well. No methods for blinding outcome assessment was described in Burt 2011 which was at unclear risk.

Incomplete outcome data

The ASTIS trial had a higher rate of withdrawals in the cyclophosphamide (control) group due to non-adherence and protocol violations (9/77 participants) (van Laar 2014). The study utilized an intention-to-treat model and was unclear if it affected the results of the study; therefore, it was judged as unclearrisk. The SCOT trial also had a higher dropout rate in the control group (for the same reasons as van Laar 2014), so had a low risk of affecting the results of the study (Sullivan 2018). There were no dropouts in the ASSIST trial (low risk; Burt 2011).

Selective reporting

All outcome measures from van Laar 2014 and Sullivan 2018 from the protocols were reported in the final publication and each trial was therefore at low risk for reporting bias Serious adverse events were not reported by Burt 2011 and was therefore judged as a high risk of bias.

Other potential sources of bias

In the ASSIST trial, there were significant differences in baseline characteristics (diffusing capacity of lung for carbon monoxide, mRSS, and disease duration) between the HSCT and cyclophosphamide group, possibly because of the small sample size; therefore this was at high risk of other bias (Burt 2011). There were no apparent other biases for Sullivan 2018 or van Laar 2014 (low risk).

Effects of interventions

See: **Summary of findings 1** Autologous non-myeloablative selective HSCT compared to cyclophosphamide in systemic sclerosis; **Summary of findings 2** Autologous myeloablative selective HSCT compared to cyclophosphamide in systemic sclerosis; **Summary of findings 3** Autologous non-myeloablative non-selective HSCT compared to cyclophosphamide in systemic sclerosis

Major outcomes – efficacy

Overall mortality

All three studies reported overall mortality (Analysis 1.1). There was no difference in overall mortality in the three studies when measured as rates. There was moderate-certainty evidence for the comparisons of autologous non-myeloablative HSCT and autologous myeloablative HSCT versus cyclophosphamide (downgraded due to sample size). The certainty of evidence for autologous non-myeloablative non-selective HSCT versus cyclophosphamide was low due to small sample size and baseline differences between groups.

With autologous non-myeloablative HSCT, there was no evidence of a difference in overall mortality between the HSCT group compared to the cyclophosphamide group at two years (RR 0.90, 95% CI 0.44 to 1.85; van Laar 2014). There were 12 deaths (15.2%) in the HSCT group versus 13 deaths (16.9%) in the control group.

With autologous myeloablative selective HSCT, there were six deaths (16.6%) in the HSCT group and 11 deaths (28.2%) in the cyclophosphamide group at 4.5 years (RR 0.59, 95% CI 0.24 to 1.43; Sullivan 2018).

With autologous non-myeloablative non-selective HSCT, there were no deaths in either arm at one year (Burt 2011).

Kaplan-Meier survival curves in autologous non-myeloablative HSCT at 10 years showed a difference in overall survival favoring HSCT (HR 0.29, 95% CI 0.13 to 0.64; van Laar 2014). For autologous myeloablative selective HSCT, there was also a benefit for HSCT with Kaplan-Meier survival curves at six years, though the HRs were not reported (P values; Sullivan 2018).

These results indicate that non-myeloablative selective HSCT and myeloablative selective HSCT may have potential mortality benefits when compared to cyclophosphamide, which are considered clinically important.

Event-free survival

Two trials reported event-free survival (Analysis 1.2; Sullivan 2018; van Laar 2014). The certainty of evidence was moderate due to smaller samples sizes for the comparisons of non-myeloablative selective HSCT and myeloablative selective HSCT versus cyclophosphamide. There were no event-free survival outcomes reported for non-myeloablative non-selective HSCT (Burt 2011).

For non-myeloablative selective HSCT, there was evidence of a difference in event-free survival at four years (defined as death or major organ failure) in favor of HSCT (HR 0.34, 95% CI 0.16 to 0.74; van Laar 2014).

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

There was no evidence of a difference in event-free survival with myeloablative selective HSCT at 4.5 years (HR 0.54, 95% CI 0.23 to 1.27; Sullivan 2018). Of note, they calculated the HR using a Z-test at 54 months. In the SCOT trial publication, a Kaplan-Meier curve demonstrated a log-rank test of 0.06 for the intention-to-treat analysis. The per-protocol analysis demonstrated a significant mortality benefit (P = 0.03). A per-protocol analysis was used as the primary statistical analysis for this trial as several participants had dropped out prior to undergoing stem cell transplantation after being randomized into this arm.

These results indicate that non-myeloablative selective HSCT and myeloablative selective HSCT have better event-free survival when compared to cyclophosphamide and that these are clinically important.

Functional ability

Two trials reported improvement in HAQ-DI for the HSCT arms (Analysis 2.1; Analysis 2.2; Sullivan 2018; van Laar 2014). The certainty of evidence of non-myeloablative selective HSCT and myeloablative selective HSCT was low due to small sample size and the trials could not be blinded, which could lead to inaccuracies in participant-reported outcomes (unclear performance bias and high-risk detection bias). The non-myeloablative non-selective HSCT trial did not report HAQ-DI (Burt 2011).

For non-myeloablative selective HSCT, there was improvement in HAQ-DI at two years in favor of HSCT (MD −0.39, 95% CI −0.72 to −0.06; absolute improvement −13%, 95% CI −24% to −2%; relative improvement −27%, 95% CI −50% to −4%; van Laar 2014).

For myeloablative selective HSCT, there was improvement in HAQ-DI at 4.5 years in favor of HSCT (RR 3.43, 95% CI 1.54 to 7.62; absolute improvement −37%, 95% CI −18% to −57%; relative improvement −243%, 95% CI −54% to −662%; NNTB 3, 95% CI 2 to 9; Sullivan 2018).

These results indicate that non-myeloablative selective HSCT and myeloablative selective HSCT improve functional outcome compared to cyclophosphamide and that these are clinically important.

Skin thickness

All three trials showed improvement in mRSS favoring the HSCT groups (Analysis 3.1; Analysis 3.2). The certainty of evidence for non-myeloablative non-selective HSCT was low due to baseline differences between both groups and small sample size. The certainty of evidence for non-myeloablative selective HSCT and myeloablative selective HSCT was moderate due to smaller sample sizes.

At two years, non-myeloablative selective HSCT probably results in a large reduction in skin thickening (MD −11.1, 95% CI −14.9 to −7.3; absolute improvement −22%, 95% CI −29% to −14%; relative improvement −43%, 95% CI −58% to −28%; van Laar 2014).

At 4.5 years, participants receiving myeloablative selective HSCT had greater improvement with HSCT in skin scores than those receiving cyclophosphamide (RR 1.51, 95% CI 1.06 to 2.13; absolute improvement −27%, 95% CI −47% to −6%; relative improvement −51% (95% CI −113% to −6%); NNTB 4, 95% CI 3 to 18; Sullivan 2018).

Non-myeloablative non-selective HSCT may result in a large reduction in skin thickening compared to cyclophosphamide after one year (MD −16.00 (95% CI −26.5 to −5.5; absolute improvement −31%, 95% CI −52% to −11%; relative improvement −84%, 95% CI −139% to −29%; Burt 2011).

All modalities of HSCT improve skin scores compared to cyclophosphamide; however, the certainty of evidence was higher for non-myeloablative selective HSCT and myeloablative selective HSCT. The improvements in all three modalities are considered clinically important.

Interstitial lung disease by pulmonary function tests

All three trials reported pulmonary function tests, but results varied between trials (Analysis 4.1; Analysis 4.2; Analysis 4.3; Analysis 4.4; Analysis 4.5). The certainty of evidence for non-myeloablative selective HSCT and myeloablative selective HSCT was moderate due to small sample sizes. In the non-myeloablative non-selective HSCT group, the certainty of evidence was low due to small sample sizes and baseline differences.

At two years in the non-myeloablative selective HSCT group, there was improvement in FVC (MD 9.10, 95% CI 3.02 to 15.18; relative improvement 11%, 95% CI 4% to 19%), TLC (MD 6.4, 95% CI 1.0 to 11.8), but not in DLCO (MD −0.60, 95% CI −6.0 to 4.8) (van Laar 2014). van Laar 2014 did not report predicted TLC.

At 4.5 years in the myeloablative selective HSCT group, there was no evidence of a difference in FVC (RR 1.63, 95% CI 0.75 to 3.51) or improvement in DLCO (RR 0.87, 95% CI 0.25 to 2.98) (Sullivan 2018).

At one year in the non-myeloablative non-selective HSCT group, there was improvement in FVC (MD 18.0, 95% CI 1.8 to 34.2; relative improvement 27%, 95% CI 2% to 51%), but not DLCO (MD 12.0, 95% CI −14.8 to 38.8) or predicted TLC (MD 13.0, 95% CI −3.5 to 29.5) (Burt 2011).

These results indicate HSCT has an unclear benefit on lung function. There is some evidence non-myeloablative selective HSCT and non-myeloablative non-selective HSCT could potentially improve outcomes in interstitial lung disease as FVC improved and this was also considered clinically important. However, there was no change in other outcomes (TLC and DLCO).

Pulmonary arterial hypertension

No trials reported pulmonary arterial hypertension.

Major outcomes – safety

Serious adverse events

All three trials reported serious adverse events (Analysis 5.1). The certainty of evidence for non-myeloablative selective HSCT and myeloablative selective HSCT was moderate due to smaller sample sizes. The non-myeloablative non-selective HSCT trial reported no adverse events in either arm (Burt 2011).

At two years in the non-myeloablative selective HSCT, there were significantly more serious adverse events in the HSCT group (64.5%) than the cyclophosphamide group (39%) (RR 1.66, 95% CI 1.20 to 2.29; absolute change 26% increased risk, 95% CI 10% to 41%; relative change 66% increased risk, 95% CI 20% to 129%; NNTH 4, 95% CI 3 to 11; van Laar 2014).

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

At 4.5 years in the myeloablative selective HSCT trial, there was a likely increase in the risk of serious adverse events in the HSCT group (25 events (73.5%)) than the cyclophosphamide group (19 events (51.4%)) (RR 1.43, 95% CI 0.99 to 2.08; Sullivan 2018).

These results indicate that there are more serious adverse events associated with HSCT versus cyclophosphamide.

Minor outcomes – efficacy

Renal function

Only the non-selective myeloablative HSCT trial reported the difference in creatinine between the HSCT and cyclophosphamide (Analysis 6.1; van Laar 2014). There was a greater decrease in creatinine in the HSCT group compared to the cyclophosphamide group at two years (MD -10.90, 95% CI -20.26 to -1.54; relative change −14%, 95% CI −26% to −2%).

Cardiac function

Only the non-myeloablative selective HSCT study reported LVEFs (van Laar 2014). There was no evidence of a difference between groups (RR −0.30, 95% CI −5.18 to 4.58; Analysis 7.1).

Health-related quality of life – 36-item Short Form

All three trials reported SF-36 (Analysis 8.1; Analysis 8.2; Analysis 8.3; Analysis 8.4; Analysis 8.5; Analysis 8.6; Analysis 8.7; Analysis 8.8; Analysis 8.9; Analysis 8.10; Analysis 8.11; Analysis 8.12). The non-myeloablative selective HSCT trial measured SF-36 at two years (van Laar 2014), the myeloablative selective HSCT trial at 4.5 years (Sullivan 2018), and the non-myeloablative non-selective HSCT trial at one year (Burt 2011).

Physical component summary

The non-myeloablative selective HSCT arm demonstrated improvement in the physical component summary of the SF-36 compared to the cyclophosphamide arm (MD 6.1, 95% CI 1.4 to 10.8; absolute improvement 6%, 95% CI 1% to 11%; relative improvement 19%, 95% CI 4% to 34%; Analysis 8.9; van Laar 2014).

The myeloablative selective HSCT trial reported more participants with significant improvement in the HSCT arm compared to the cyclophosphamide arm (RR 3.60, 95% CI 1.64 to 8.00; absolute improvement 40%, 95% CI 20 to 60%; relative improvement 261%, 95% CI 164% to 697%; NNTB 3, 95% CI 2 to 42; Analysis 8.11; Sullivan 2018).

The non-myeloablative non-selective HSCT arms had improvement in the physical component summary of the SF-36 compared to the control arm (MD 26.0, 95% CI 6.2 to 45.9; absolute improvement 26%, 95% CI 6% to 46%; relative improvement 72%, 95% CI 16% to 125%; Analysis 8.9; Burt 2011).

Mental component summary

In the non-myeloablative selective HSCT trial, there was no evidence of a difference in the mental component summary of the SF-36 (MD −0.30, 95% CI −5.98 to 5.38; Analysis 8.10; van Laar 2014).

In the myeloablative selective HSCT trial, more participants had mental component improvement in the HSCT arm compared to the cyclophosphamide arm (RR 3.97, 95% CI 1.20 to 13.10; absolute improvement 23%, 95% CI 6% to 40%; relative improvement 297%, 95% CI 20% to 1210%, NNTB 5, 95% CI 2 to 42; Analysis 8.12; Sullivan 2018).

There were significant improvements in the SF-36 mental component summary for non-myeloablative non-selective HSCT compared to cyclophosphamide (MD 26.0, 95% CI 8.8 to 43.2; absolute improvement 26%, 95% CI 9% to 43%, relative improvement 46%, 95% CI 16 to 77%; Analysis 8.10; Burt 2011).

Improvements in quality of life for both the physical and mental components were clinically important, except for the physical component of the SF-36 with non-myeloablative selective HSCT. These results indicate that the use of HSCT may improve quality of life over time as compared to cyclophosphamide.

Health-related quality of life – EQ-5D visual analog scale scores

There was no evidence of a difference in EQ-5D VAS score between non-myeloablative selective HSCT and cyclophosphamide at two years (MD 6.7, 95% CI −7.7 to 21.1; Analysis 9.1; van Laar 2014). However, the VAS was a component of the European Quality of Life EQ-5D scale. This is a standardized measure of HRQoL that is graded from 0 to 1 (0 being worse quality of life and 1 being better). This showed that non-myeloablative selective HSCT had a highly significant clinical difference between the non-myeloablative selective HSCT and cyclophosphamide arms (MD 0.29, 95% DI 0.12 to 0.45).

The other trials did not report health-related quality of life.

Safety

Inflammatory markers

No studies reported changes in inflammatory markers.

Treatment-related mortality

There were differences in treatment-related mortality among the three trials (Analysis 5.2).

The non-myeloablative selective HSCT trial had eight treatment-related deaths (10.1%) in the HSCT group and no treatment-related deaths in the cyclophosphamide group (Peto OR 7.91, 95% CI 1.91 to 32.67; absolute change 10%, 95% CI 3% to 17%; relative change 1558%, 95% CI −3% to 28,130%; NNTH not calculated with control event rate of zero; van Laar 2014). Treatment-related deaths occurred within the first year of stem cell transplantation.

At 4.5 years the myeloablative selective HSCT trial had one treatment-related death (2.7%) in the HSCT group and no treatment-related deaths in the cyclophosphamide group (Peto OR 8.03, 95% CI 0.16 to 406.02; Sullivan 2018). At 70 months, there was a second treatment-related death in the HSCT arm.

The non-myeloablative non-selective HSCT trial reported no deaths in either arm (Burt 2011).

Serious non-lethal infections

The three trials found no evidence of a difference in risk of non-lethal (grade 3 or 4 adverse event) infections (Analysis 5.3).

At two years, the non-myeloablative selective HSCT study reported eight infections (10%) in the HSCT group and four (5%) in the

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

cyclophosphamide group (OR 2.06, 95% CI 0.59 to 7.13; van Laar 2014).

At 4.5 years, the myeloablative selective HSCT study reported 10 infections (29%) in the HSCT group and seven (18.9%) in the cyclophosphamide group (OR 1.79, 95% CI 0.59 to 5.39; Sullivan 2018).

At one year, the non-myeloablative non-selective HSCT study reported three infections (10%) in the HSCT group and one (9%) in the cyclophosphamide group (OR 3.43, 95% CI 0.29 to 40.95; Burt 2011).

Study withdrawals

The withdrawals for each study are shown in Analysis 5.4.

The non-myeloablative selective HSCT study had four withdrawals (5%) in the HSCT group and 18 (31.6%) in the cyclophosphamide group (RR 0.22, 95% CI 0.08 to 0.61); absolute change 26% decreased risk with HSCT, 95% CI 13% to 39%; relative percent change 38% decreased risk with HSCT, 95% CI 15% to 66%; NNTH with cyclophosphamide was 3, 95% CI 2 to 5; van Laar 2014).

The myeloablative selective HSCT study had more withdrawals in the controls, reporting 6 withdrawals (25%) in the HSCT group and 9 (15.4%) in the cyclophosphamide group (RR 1.63, 95% CI 0.64 to 4.11 decreased with HSCT; absolute change 26%, 95% CI 5% to 47%; relative percent change 54%, 95% CI 6% to 124%; NNTH with cyclophosphamide is 4, 95% CI 3 to 25; Sullivan 2018).

There were no withdrawals in either arm in the non-myeloablative non-selective HSCT study (Burt 2011).

Renal failure

There was no evidence of differences in renal failure between cyclophosphamide and non-myeloablative selective HSCT (Peto OR 1.99, 95% CI 0.62 to 6.45), myeloablative selective HSCT (Peto OR 0.72, 95% CI 0.12 to 4.36), or non-myeloablative non-selective HSCT (Peto OR 6.69, 95% CI 0.13 to 338.79) (Analysis 5.5).

DISCUSSION

This systematic review analyzed and summarized evidence from all RCTs examining the effects of various types of stem cell transplants for treating SSc. Three RCTs including three different modalities for HSCT with a total of 250 participants, comprising 125 that received HSCT, were included. All trials had cyclophosphamide as the control comparator treatment.

Summary of main results

Efficacy

There were no differences in overall mortality between any modality of HSCT or cyclophosphamide as measured by rates. However, survival analysis for non-myeloablative selective HSCT and myeloablative selective HSCT suggests there is a survival benefit for HSCT. However, non-myeloablative selective HSCT demonstrated significant improvement in event-free survival over cyclophosphamide at 48 months (moderate-certainty evidence) (van Laar 2014). There was a significant improvement in event-free survival with myeloablative selective HSCT in the intention-to-treat analysis group, but not in the per-protocol analysis group (Sullivan 2018).

There was low-certainty evidence that both non-myeloablative selective HSCT (van Laar 2014) and myeloablative selective HSCT (Sullivan 2018) provide clinically meaningful benefits in HRQoL metrics including the HAQ-DI and SF-36, when compared to treatment with cyclophosphamide. There was low-certainty evidence that non-myeloablative non-selective HSCT may also improve quality of life as measured by the SF-36 (Burt 2011).

There was moderate-certainty evidence that both non-myeloablative selective HSCT (van Laar 2014) and myeloablative selective HSCT (Sullivan 2018) provide clinically meaningful benefits in skin thickness. There was low-certainty evidence that non-myeloablative non-selective HSCT may also improve skin thickness (Burt 2011).

Pulmonary function tests results were varied across interventions. No study showed improvements in diffusing capacity of lung for carbon monoxide. There were improvements in FVC with non-myeloablative selective HSCT (moderate-certainty evidence; van Laar 2014) and non-myeloablative non-selective HSCT (low-certainty evidence; Burt 2011).

Safety

Non-myeloablative selective HSCT (van Laar 2014) and myeloablative selective HSCT (Sullivan 2018) showed increased risk of adverse events compared to cyclophosphamide. In addition, non-myeloablative selective HSCT also showed increased risk of treatment-related mortality (van Laar 2014). There were increased withdrawals in cyclophosphamide arms with both myeloablative selective HSCT and non-myeloablative selective HSCT. Non-selective non-myeloablative HSCT reported no serious adverse events (Burt 2011).

Overall completeness and applicability of evidence

We included and analyzed all data from RCTs that evaluated the efficacy of autologous HSCT compared to cyclophosphamide in the treatment of SSc. The evidence is applicable and complete as we have included all published data including clinically relevant and patient-centered outcomes. However, one trial evaluated non-myeloablative non-selective HSCT in only 10 participants (Burt 2011). One RCT each evaluated non-myeloablative selective HSCT and myeloablative selective HSCT, though with larger samples (Sullivan 2018; van Laar 2014). Larger trials are likely needed to further determine the effectiveness and harms of HSCT in treating people with SSc, as some of the findings which may have clinical significance.

The inclusion criteria of the included studies were similar. They included diffuse SSc as defined by the ACR and the exclusion of secondary diseases (including chronic infections and cancer). The inclusion criteria for these studies selected people with advanced multi-organ disease, and are not generalizable to people with early disease limited to cutaneous involvement. One differing aspect between the studies was the different thresholds for the inclusion of people with pulmonary arterial hypertension. The SCOT trial for myeloablative selective HSCT excluded people with any evidence of pulmonary hypertension (Sullivan 2018), whereas the ASSIST (Burt 2011) and ASTIS (van Laar 2014) trials excluded people with severe pulmonary arterial hypertension. The participants included in the trials were also likely to have fewer comorbidities than the general population (as the inclusion criteria for enrollment in all three studies were stringent). Since HSCT is an invasive procedure,

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

the evidence should be applicable only to the populations studied and be considered in those with advanced disease but fewer other comorbidities. Prospective longitudinal studies or trials will be needed to evaluate the utility of this intervention in people with milder disease, or in people with other severe comorbidities.

Quality of the evidence

We used the GRADE criteria to assess the certainty of the evidence, which is shown in Summary of findings 1; Summary of findings 2; and Summary of findings 3. In all included trials, the evidence for participant-reported outcomes (HAQ-DI, SF-36, and EQ-5D VAS) was downgraded one level due to possible blinding bias as participants could not be blinded due to the nature of stem cell transplantation. This could cause an unclear performance bias and high risk of detection bias.

We downgraded all outcomes from the non-myeloablative non-selective HSCT trial one level because of differences in baseline characteristics between the HSCT and cyclophosphamide groups (Burt 2011). Furthermore, the ASSIST trial did not clearly report serious adverse events and selection bias risk and was, therefore, rated as high (Burt 2011). Evidence for main outcomes from non-myeloablative selective HSCT and myeloablative HSCT was rated as moderate in the summary of findings tables, but there was imprecision in secondary outcomes such as study withdrawals and treatment-related mortality that should also downgrade the certainty of evidence. The evidence from all modalities of HSCT is limited as there are small sample sizes for each modality that are each represented by one RCT.

There was no evidence of publication bias or unreported/unpublished studies.

Potential biases in the review process

An experienced information specialist developed the search strategy used in this review. Once the search strategy was conducted, two review authors independently analyzed all abstracts and titles and performed bias and quality assessments. We reached consensus by discussion and with a third-party expert. As a result, we minimized errors in selection and abstraction. The main limitations from this review process were the small sample sizes of the RCTs and the data that could be utilized. Furthermore, data from the trials could not be consolidated as all three trials used different types of HSCT that might affect outcomes. Furthermore, two of the three trials included, the ASSIST (Burt 2011) and SCOT (Sullivan 2018) trials, were not powered to detect a decrease in overall mortality or an improvement in event-free survival. Finally, recent studies have demonstrated that mycophenolate mofetil may emerge as the new preferred treatment option given its better safety profile as opposed to cyclophosphamide for various disease manifestations (Tashkin 2016).

Agreements and disagreements with other studies or reviews

Three other systematic reviews have examined HSCT in the treatment of SSc (Host 2017; Puyade 2019; Shouval 2018). Overall, their conclusions were similar to those reported here. Risks of bias were slightly different between Shouval 2018 and our systematic review (not assessed with Host 2017). Blinding of participants was deemed at high risk. In this review, we rated performance bias as low risk for objective outcomes and unclear risk for participant-reported outcomes. We rated detection bias as high risk for participant-reported outcomes but low risk for objective outcomes. All primary outcome measures except for HAQ-DI are objective measures that are unlikely to be affected by the blinding of participants; however, participant outcomes may potentially be affected and therefore the level of evidence was downgraded on the summary of findings table. Another significant difference is Shouval 2018 conducted pooled analysis with all separate modalities of HSCT. This was not done in our review as the different modalities had large differences in mobilization, conditioning, and infusion of stem cells and we do not believe pooling these results would provide clinically meaningful information that would be applicable towards the care of patients. Otherwise, conclusions regarding overall mortality, event-free survival, skin thickness, interstitial lung disease, and safety were similar.

AUTHORS' CONCLUSIONS

Implications for practice

This review provides a summary of evidence of the role of hematopoietic stem cell transplantation (HSCT) in treatment of systemic sclerosis (SSc). As compared to cyclophosphamide, the different modalities of HSCT demonstrated improvement in skin thickening, quality of life, and the forced vital capacity in pulmonary function tests. The ASTIS trial was the only trial powered to meet its primary end point of event-free survival and it demonstrated there is a benefit in this outcome with non-selective myeloablative HSCT (van Laar 2014). However, HSCT in the ASTIS trial was associated with more serious adverse events, and more concerning, with a significant increase in treatment-related mortality (van Laar 2014). The SCOT trial had a better safety profile, but did not show a significant benefit for event-free survival in the intention-to-treat population, although it showed a significant benefit for this outcome in the per-protocol analysis (Sullivan 2018).

Implications for research

More randomized controlled trials, especially evaluating non-myeloablative non-selective HSCT and myeloablative selective HSCT are needed to further determine the efficacy and safety of HSCT. However, these trials are challenging to conduct as SSc is a rare disease, it is difficult to enroll patients, and other therapies are also currently being investigated. There is a need to establish registries comparing different HSCT modalities as it would be very difficult to conduct head-to-head trials given these challenges. It will also be important to determine the safety and efficacy of HSCT when compared to mycophenolate mofetil as this agent is becoming a preferred over cyclophosphamide, and none of the three trials evaluated this drug since this was not the standard of care at the time these trials were done. Further research is needed to determine the benefits of HSCT in populations who were excluded from these trials (from the exclusion criteria) due to disease manifestations or comorbidities (or both), especially those with pulmonary hypertension (Spierings 2022). Finally, although not specifically examined in this review, non-myeloablative selective HSCT and myeloablative selective HSCT have shown worse survival outcomes in people who smoke tobacco products and further research is needed to determine how clinicians may best care for this patient population.

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ACKNOWLEDGEMENTS

We thank Greg F Pratt, DDS, MS (Research Medical Library, The University of Texas MD Anderson Cancer Center) for providing comments on the search strategy.

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

REFERENCES

References to studies included in this review

Burt 2011 *{published data only}*

* Burt RK, Shah SJ, Dill K, Grant T, Gheorghiade M, Schroeder J, et al. Autologous non-myeloablative haemopoietic stem-cell transplantations compared with pulse cyclophosphamide once per months for systemic sclerosis (ASSIST): an open-label, randomised phase 2 trial. *Lancet* 2011;**378**:498-506.

NCT00278525. Cyclophosphamide and rATG with hematopoietic stem cell support in systemic scleroderma. clinicaltrials.gov/ct2/show/NCT00278525 (first received 18 January 2006). [NCT00278525]

NCT01445821. Autologous stem cell systemic sclerosis immune suppression trial (DIScl2011). clinicaltrials.gov/ct2/show/NCT01445821 (first received 4 October 2011). [NCT01445821]

Sullivan 2018 *{published data only}*

Assassi S, Mayes MD, Pedroza C, Chang JT, Furst DE, Crofford LJ, et al. Immunoablation followed by autologous stem cell transplantation in systemic sclerosis patients decreases significantly the interferon signature. 2015 ACR/ARHP Annual Meeting; 6-11 Nov 2015; San Francisco (CA).

Assassi S, Wallace P, Ying J, Keyes-Elstein L, Goldmuntz E, Turner J. Treatment with myeloablation followed by autologous stem cell transplantation normalizes the systemic sclerosis molecular signature. *Journal of Scleroderma and Related Disorders* 2018;**3**:s272-3.

Assassi S, Wang X, Keyes-Elstein L, Goldmuntz E, Turner J, Zheng W. Changes in the systemic sclerosis molecular signatures after myeloablation followed by autologous hematopoietic stem cell transplantation and their clinical correlates. *Arthritis and Rheumatology* 2018;**70**(Suppl 9):s1002-3.

Bellocchi C, Ying J, Goldmuntz E, Keyes-Elstein L, Varga J, Hinchcliff M, et al. Systemic sclerosis has a distinct serum protein profile that correlates with its clinical manifestations. *Arthritis and Rheumatology* 2018;**70**:s1213-4.

Bellochi C, Sullivan K, Goldmuntz E, Ying J, Keyes-Elstein L, Varga J, et al. Myeloablation followed by autologous stem cell rescue leads to normalisation of serum IL-6 levels in patients with systemic sclerosis. *Annals of the Rheumatic Diseases* 2018;**77**:s1100-1.

Franks J, Martyanov V, Wood TA, Crofford L, Keyes-Elstein L, Furst DE. Machine learning classification of peripheral blood gene expression identifies a subset of patients with systemic sclerosis most likely to show clinical improvement in response to hematopoietic stem cell transplant. *Arthritis and Rheumatology* 2018;**70**:s2091-3.

Goldin J, Keyes-Elstein L, Crofford L, Furst DE, Goldmuntz E, Mayes MD, et al. Changes in quantitative scleroderma lung CT measures in patients treated with cyclophosphamide or transplantation. *Arthritis and Rheumatology* 2018;**70**:s1005-6.

NCT00114530. Scleroderma: Cyclophosphamide Or Transplantation (SCOT). clinicaltrials.gov/ct2/show/NCT00114530 (first received 16 June 2005). [NCT00114530]

Sullivan K, Keyes-Elstein L, McSweeney P, Pinckney A, Welch B, Mayes M, et al. Myeloablative autologous transplantation of CD34+ selected hematopoietic stem cells (HSCT) vs monthly intravenous cyclophosphamide (CY) for severe scleroderma with internal organ involvement: outcomes of a randomized North American clinical trial. *Biology of Blood and Marrow Transplantation* 2017;**23**:S118-91.

Sullivan K, Keyes-Elstein L, McSweeney P, Pinckney A, Welch B, Mayes MD, et al. Myeloablative autologous transplantation of CD34+-selected hematopoietic stem cells (HSCT) vs monthly intravenous cyclophosphamide (CYC) for severe scleroderma with internal organ involvement: outcomes of a randomized North American clinical trial [abstract]. *Arthritis and Rheumatology* 2016;**68**(Suppl 10):S21.

* Sullivan KM, Goldmuntz EA, Keyes-Elstein L, McSweeney PA, Pinckney A, Welch B, et al. Myeloablative autologous stem-cell transplantation for severe scleroderma. *New England Journal of Medicine* 2019;**378**:35-47. [PMID: 29298160]

van Laar 2014 *{published data only (unpublished sought but not used)}***ISRCTN54371254**

Abdallah NA, Wang M, Lansiaux P, Puyade M, Berthier S, Terriou L, et al. Long term outcomes of the French ASTIS systemic sclerosis cohort using the global rank composite score. *Bone Marrow Transplantation* 2021;**56**:2259-67.

van Laar J, Farge D, Sont J and the EBMT/EULAR Scleroderma Study Group. LB0002. The ASTIS trial: autologous stem cell transplantation versus IV pulse cyclophosphamide in poor prognosis systemic sclerosis, first results. *Annals of the Rheumatic Diseases* 2013;**71**:151.

* van Laar JM, Farge D, Sont J, Naraghi K, Marjanovic Z, Larghero J, et al. Autologous hematopoietic stem cell transplantation vs intravenous pulse cyclophosphamide in diffuse cutaneous systemic sclerosis. *JAMA* 2014;**311**(24):2490-8.

van Laar JM, Farge D, Sont JK, Naraghi K, Marjanovic Z, Schuerwegh A, et al, EBMT/EULAR Scleroderma Study Group. High dose immunoablation and autologous hematopoietic stem cell transplantation versus monthly intravenous pulse therapy cyclophosphamide in severe systemic sclerosis. *Arthritis & Rheumatism* 2012;**64**(S10):4167.

References to ongoing studies

van Laar 2020 *{unpublished data only}*

NCT04464434. Upfront autologous HSCT versus immunosuppression in early diffuse cutaneous systemic sclerosis (UPSIDE). clinicaltrials.gov/ct2/show/NCT04464434 (first received 9 July 2020).

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Additional references

ACR 1980

American College of Rheumatology. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. *Arthritis and Rheumatism* 1980;**23**(5):581-90. [PMID: 7378088]

Bergamasco 2019

Bergamasco A, Hartmann N, Wallace L, Verpillat P. Epidemiology of systemic sclerosis and systemic sclerosis-associated interstitial lung disease. *Clin Epidemiol* 2019;**11**:257-273.

Binks 2001

Binks M, Passweg JR, Furst D, McSweeney P, Sullivan K, Besenthal C, et al. Phase I/II trial of autologous stem cell transplantation in systemic sclerosis: procedure related mortality and impact on skin disease. *Annals of the Rheumatic Diseases* 2001;**60**(6):577-84. [PMID: 11350846]

Cates 2008 [Computer program]

Visual Rx 2.0 NNT calculator. Cates C. Dr Chris Cates EBM, 2008. Available at www.nntonline.net.

Deeks 2021

Deeks JJ, Higgins JP, Altman DG. Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.2 (updated February 2021). Cochrane, 2021. Available from training.cochrane.org/handbook/archive/v6.2.

Distler 2008

Distler O, Behrens F, Pittrow D, Huscher D, Denton CP, Foeldvari I, et al, EPOSS-Omeract Group. Defining appropriate outcome measures in pulmonary arterial hypertension related to systemic sclerosis: a Delphi consensus study with cluster analysis. *Arthritis and Rheumatism* 2008;**59**(6):867-75. [PMID: 18512721]

Elhai 2012

Elhai M, Meune C, Avouac J, Kahan A, Allanore Y. Trends in mortality in patients with systemic sclerosis over 40 years: a systematic review and meta-analysis of cohort studies. *Rheumatology (Oxford)* 2012;**51**(6):1017-26. [PMID: 21900368]

EndNote X9 [Computer program]

EndNote X9. Philadelphia (PA): Clarivate Analytics, 2019. Available at endnote.com.

Farge 2002

Farge D, Marolleau JP, Zohar S, Marjanovic Z, Cabane J, Mounier N, et al, Intensification et Autogreffe dans les Maladies Auto Immunes Resistantes (ISAMAIR) Study Group. Autologous bone marrow transplantation in the treatment of refractory systemic sclerosis: early results from a French Multicentre phase I-II study. *British Journal of Haematology* 2002;**119**(3):726-39. [PMID: 12437652]

GRADEpro GDT [Computer program]

McMaster University (developed by Evidence Prime) GRADEpro GDT. Hamilton (ON): McMaster University (developed by Evidence Prime), accessed 3 March 2020. Available at gradepro.org.

Higgins 2017

Higgins JP, Altman DG, Sterne JA. Chapter 8: Assessing risk of bias in included studies. In: Higgins JP, Churchill R, Chandler J, Cumpston MS, editor(s), Cochrane Handbook for Systematic Reviews of Interventions Version 5.2.0 (updated June 2017), The Cochrane Collaboration, 2017. Available from training.cochrane.org/handbook/archive/v5.2/.

Higgins 2021

Higgins JP, Li T, Deeks JJ. Chapter 6: Choosing effect measures and computing estimates of effect. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.2 (updated February 2021). Cochrane, 2021. Available from training.cochrane.org/handbook/archive/v6.2.

Host 2017

Host L, Nikpour M, Calderone A, Cannell P, Roddy J. Autologous stem cell transplantation in systemic sclerosis: a systematic review. *Clinical and Experimental Rheumatology* 2017;**35**(106):198-207.

Hudson 2014

Hudson M, Fritzler MJ. Diagnostic criteria of systemic sclerosis. *Journal of Autoimmunity* 2014;**48-49**:38-41. [PMID: 24461384]

Jordan 2015

Jordan S, Maurer B, Toniolo M, Michel B, Distler O. Performance of the new ACR/EULAR classification criteria for systemic sclerosis in clinical practice. *Rheumatology (Oxford)* 2015;**54**(8):1454-8.

Kafaja 2017

Kafaja S, Celements PJ, Wilhalme H, Tseng C, Furst DE, Kim GH, et al. Reliability and minimal clinically important differences of FVC results from the Scleroderma Lung Studies (SLS-I and SLS-II). *Journal of Respiratory and Critical Care Medicine* 2017;**197**(5):971.

Khanna 2009

Khanna D, Distler O, Avouac J, Behrens F, Clements PJ, Denton C, et al, CRISS, EPOSS. Measures of response in clinical trials of systemic sclerosis: the Combined Response Index for Systemic Sclerosis (CRISS) and Outcome Measures in Pulmonary Arterial Hypertension related to Systemic Sclerosis (EPOSS). *Journal of Rheumatology* 2009;**36**(10):2356-61. [PMID: 19820225]

Khanna 2019

Khanna D, Clements PJ, Volkmann ER, Wilhalme H, Tseng C, Furst DE, et al. Minimal clinically important differences for the modified Rodnan Skin Score: results from the Scleroderma Lung Studies (SLS-I and SLS-II). *Arthritis Research and Therapy* 2019;**21**:23.

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kowal-Bielecka 2009

Kowal-Bielecka O, Landewé R, Avouac J, Chwiesko S, Miniati I, Czirjak L, et al, EUSTAR Co-Authors. EULAR recommendations for the treatment of systemic sclerosis: a report from the EULAR Scleroderma Trials and Research group (EUSTAR). *Annals of the Rheumatic Diseases* 2009;**68**(5):620-8. [PMID: 19147617]

LeRoy 1988

LeRoy EC, Black C, Fleischmajer R, Jablonska S, Krieg T, Medsger TA Jr, et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. *Journal of Rheumatology* 1988;**15**(2):202-5. [PMID: 3361530]

Mayes 2003

Mayes MD, Lacey JV Jr, Beebe-Dimmer J, Gillespie BW, Cooper B, Laing TJ, et al. Prevalence, incidence, survival, and disease characteristics of systemic sclerosis in a large US population. *Arthritis and Rheumatology* 2003;**48**(8):2246-55. [PMID: 12905479]

Merkel 2003

Merkel PA, Clements PJ, Reveille JD, Suarez-Almazor ME, Valentini G, Furst DE. Current status of outcome measure development of clinical trials in systemic sclerosis. Report from OMERACT 6. *Journal of Rheumatology* 2003;**30**(7):1630-47. [PMID: 12858472]

Nagaraja 2015

Nagaraja V, Denton CP, Khanna D. Old medications and new targeted therapies in systemic sclerosis. *Rheumatology (Oxford)* 2015;**54**(11):1944-53. [PMID: 25065013]

Nash 2007

Nash RA, McSweeney PA, Crofford LJ, Abidi M, Chen C, Godwin D, et al. High-dose immunosuppressive therapy and autologous hematopoietic cell transplantation for severe systemic sclerosis: long-term follow-up of the US multicenter pilot study. *Blood* 2007;**110**(4):1388-96. [PMID: 17452515]

Oyama 2007

Oyama Y, Barr WG, Statkute L, Corbridge T, Gonda EA, Jovanovic B, et al. Autologous non-myeloablative hematopoietic stem cell transplantation in patients with systemic sclerosis. *Bone Marrow Transplantation* 2007;**40**(6):549-55. [PMID: 17646844]

Page 2021

Page MJ, Higgins JP, Sterne JA. Chapter 13: Assessing risk of bias due to missing results in a synthesis. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.2 (updated February 2021). Cochrane, 2021. Available from training.cochrane.org/handbook/archive/v6.2.

Park 2015

Park JS, Park MC, Song JJ, Park YB, Lee SK, Lee SW. Application of the 2013 ACR/EULAR classification criteria for systemic sclerosis to patients with Raynaud's phenomenon. *Arthritis Research & Therapy* 2015;**17**(1):77. [PMID: 25889905]

Poormoghim 2012

Poormoghim H, Moradi Lakeh M, Mohammadipour M, Sodagari F, Toofaninjed N. Cyclophosphamide for scleroderma lung disease: a systematic review and meta-analysis. *Rheumatology International* 2012;**32**(8):2431-44. [PMID: 21691743]

Puyade 2019

Puyade M, Maltez N, Lansiaux P, Pugnet G, Roblot P, Colmegna I. Health-related quality of life in systemic sclerosis before and after autologous haematopoietic stem cell transplant – a systematic review. *Rheumatology* 2019;**59**(4):779-89.

Review Manager 2014 [Computer program]

The Nordic Cochrane Centre, The Cochrane Collaboration Review Manager 5 (RevMan 5). Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Schünemann 2021a

Schünemann HJ, Vist GE, Higgins JP, Santesso N, Deeks JJ, Glasziou P, et al. Chapter 15: Interpreting results and drawing conclusions. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editor(s). Cochrane Handbook for Systematic Reviews of Interventions. Version 6.2 (updated February 2021). Cochrane, 2021. Available from training.cochrane.org/handbook/archive/v6.2.

Schünemann 2021b

Schünemann HJ, Higgins JP, Vist GE, Glasziou P, Akl EA, Skoetz N, et al. Chapter 14: Completing 'Summary of findings' tables and grading the certainty of the evidence. In Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.2 (updated February 2021). Cochrane, 2021. Available from training.cochrane.org/handbook/archive/v6.2.

Sekhon 2010

Sekhon S, Pope J, Canadian Scleroderma Research Group, Baron M. The minimally important difference in clinical practice for patient-centered outcomes including Health Assessment Questionnaire, fatigue, pain, sleep, global visual analog scale, and SF-36 in scleroderma. *Journal of Rheumatology* 2010;**37**(3):591-8.

Shouval 2018

Shouval R, Furie N, Raanani P, Nagler A, Gafter-Gvili A. Autologous hematopoietic stem cell transplantation for systemic sclerosis: a systematic review and meta-analysis. *Biology of Blood and Marrow Transplantation* 2018;**24**(5):937-44.

Snowden 2012

Snowden JA, Saccardi R, Allez M, Ardizzone S, Arnold R, Cervera R, et al, EBMT Autoimmune Disease Working Party (ADWP), Paediatric Diseases Working Party (PDWP). Haematopoietic SCT in severe autoimmune diseases: updated guidelines of the European Group for Blood and Marrow Transplantation. *Bone Marrow Transplantation* 2012;**47**(6):770-90. [PMID: 22002489]

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Spierings 2022

Spierings J, Nihtyanova SI, Derrett-Smith E, Clark KE, van Laar JM, Ong V, et al. Outcomes linked to eligibility for stem cell transplantation trials in diffuse systemic cutaneous systemic sclerosis. *Rheumatology (Oxford)* 2022;**61**(5):1948-56.

Steen 1998

Steen VD, Medsger TA Jr. Case-control study of corticosteroids and other drugs that either precipitate or protect from the development of scleroderma renal crisis. *Arthritis and Rheumatism* 1998;**41**(9):1613-9. [PMID: 9751093]

Steen 2000

Steen VD, Medsger TA Jr. Severe organ involvement in systemic sclerosis with diffuse scleroderma. *Arthritis and Rheumatism* 2000;**43**(11):2437-44. [PMID: 11083266]

Tashkin 2006

Tashkin DP, Elashoff R, Clements PJ, Goldin J, Roth MD, Furst DE, et al. Cyclophosphamide versus placebo in scleroderma lung disease. *New England Journal of Medicine* 2006;**354**(25):2655-66.

Tashkin 2016

Tashkin MP, Roth MD, Clements PJ, Khanna D, Kleerup EC, Golding K, et al. Mycophenolate mofetil versus oral cyclophosphamide in scleroderma-related interstitial lung disease (SLS II): a randomised controlled, double-blind, parallel group trial. *Lancet Respiratory Medicine* 2016;**4**(9):708-19.

Toledano 2012

Toledano E, Candelas G, Rosales Z, Martínez Prada C, León L, Abásolo L, et al. A meta-analysis of mortality in rheumatic diseases. *Reumatologia Clinica* 2012;**8**(6):334-41. [PMID: 22789463]

van den Hoogen 2013

van den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League against Rheumatism collaborative initiative. *Arthritis & Rheumatism* 2013;**65**(11):2737-47. [PMID: 24122180]

van Laar 2013

van Laar JM, Sullivan K. Stem cell transplantation in systemic sclerosis. *Current Opinion in Rheumatology* 2013;**25**(6):719-25. [PMID: 24047607]

References to other published versions of this review

Bruera 2015

Bruera S, Siddhanamatha HR, Molony DA, Mayes MD, Lopez-Olivo MA. Stem cell transplantation for systemic sclerosis. *Cochrane Database of Systematic Reviews* 2015, Issue 9. Art. No: CD011819. [DOI: 10.1002/14651858.CD011819]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Burt 2011

Study characteristics	
Methods	Study design: single center, randomized controlled trial, cross-over Trial time period: not available Setting: 1 site in Chicago, US Analysis: intention-to-treat
Participants	**Baseline characteristics** *Cyclophosphamide (9 participants)* • Mean age: 44 years (range 26–54 years) • Number male/female: 1/8 • Disease duration: 18 months (range 6–36 months) • mRSS: 19 (range 4–45) • FVC: 67% (range 43–84%) • DLCO: 75% (range 29–111%) • Number of diffuse SSc/limited SSc: 7/2 *HSCT (10 participants)* • Mean age: 45 years (range 32–58 years) • Number male/female: 1/9

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Burt 2011 *(Continued)*

- Disease duration: 13.6 months (range 2–33 months)
- mRSS: 28 (range 6–48)
- FVC: 62% (range 53–70%)
- DLCO: 58% (range 29–82%)
- Number diffuse SSc/limited SSc: 8/2

Inclusion criteria

- Aged < 60 years
- Diffuse SSc (cutaneous involvement proximal to elbow or knee with an mRSS > 14
- Internal-organ involvement, defined as ≥ 1 of:
  - DLCO < 80% or decline in FVC by ≥ 10% in previous 12 months
  - pulmonary fibrosis or ground-glass appearance on high-resolution chest CT
  - abnormal electrocardiogram
  - gastrointestinal tract involvement
- People with restricted skin involvement (mRSS < 14) were eligible only if they had coexistent pulmonary involvement

Exclusion criteria

- Previous treatment with > 6 intravenous treatments of cyclophosphamide
- Total lung capacity < 45% predicted volume
- LVEF < 40% or symptomatic cardiac disease, tricuspid annular plane systolic excursion < 1.8 cm
- Duration of SSc > 4 years
- Renal insufficiency defined as creatinine > 177 µmol/L
- Pulmonary artery systolic pressure > 40 mmHg or mean artery pressure > 25 mmHg

Interventions	Non-myeloablative non-selective HSCT vs IV cyclophosphamide
	• HSCT
	  ○ Mobilization: cyclophosphamide 2 g/m^2 with subcutaneous filgrastim 10 µg/kg/day; leukapheresis without manipulation
	  ○ Conditioning: cyclophosphamide 200 mg/kg (given in 4 equal fractions throughout 4 days). Rabbit ATG at 0.5 mg/kg 5 days prior to stem cell transplantation and then 1.5 mg/kg for 4 days prior to stem cell transplantation. Each dose of rabbit ATG was administered with IV methylprednisolone 1000 mg
	  ○ Infusion of non-selected autologous stem cells
	• IV cyclophosphamide
	  ○ Cyclophosphamide 1 g/m^2 monthly for 6 cycles

Outcomes	Primary outcome: improvement at 12 months defined as: a decrease in mRSS > 25% with an initial mRSS > 14 or an increase in FVC > 10% Secondary outcomes: FVC, total lung capacity, DLCO, volume of diseased lung on CT scan, mRSS, and SF-36
Notes	Source of funding: not reported Conflicts of interest: none

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated randomization sequence. Quote: "use of a computer-generated randomisation sequence" (p.499).

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Burt 2011 *(Continued)*

Allocation concealment (selection bias)	Low risk	Participants randomly allocated in 1:1 ratio with a mixed block design with initial blocks of 10 and subsequent block size of 4. Quote: "randomly allocated patients in a one-to-one ratio to receive HSCT or cyclophosphamide by use of a computer-generated randomisation sequence with a mixed block design, with initial blocks of ten and subsequent block size of four" (p.499).
Blinding of participants and personnel (performance bias) Participant-reported outcomes	Unclear risk	Blinding not possible, reported outcomes (e.g. SF-36, HAQ-DI) may have been influenced.
Blinding of participants and personnel (performance bias) Objective outcomes	Low risk	Blinding not possible, objective outcomes (such as pulmonary function tests, mortality, etc.) are unlikely to be influenced by lack of blinding.
Blinding of outcome assessment (detection bias) Participant-reported outcomes	Unclear risk	No methods for blinding outcome assessment described.
Blinding of outcome assessment (detection bias) Objective outcomes	High risk	No methods for blinding outcome assessment described.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts.
Selective reporting (reporting bias)	High risk	Serious adverse events not clearly reported. Outcome measures from protocol were reported.
Other bias	High risk	Difference in baselines characteristics such as mRSS and pulmonary function tests were significant – likely due to small sample size of study.

Sullivan 2018

Study characteristics	
Methods	Study design: multicenter active controlled parallel group randomized controlled trial Trial time period: 2001–2008 Setting: 26 sites in North America Analysis: intention-to-treat
Participants	**Baseline characteristics** *Cyclophosphamide (39 participants)* • Mean age: 46.9 (SD 10.4) years • Number male/female: 10/29 • Disease duration, mean: 29 (SD 16) months

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Sullivan 2018 *(Continued)*

- mRSS, mean: 30.8 (SD 10.5)
- FVC, mean: 73.8% (SD 17.0%)
- DLCO, mean: 52.7% (SD 14%)
- HAQ-DI, mean: 1.4 (SD 0.9)
- Creatinine clearance, mean: 124.9 (SD 54.3) mL/minute
- Smoking status number current or former/never: 10/29

HSCT (36 participants)

- Mean age: 44.9 (SD 10.9) years
- Number male/female: 17/19
- Disease duration, mean: 25.1 (SD 12.9) months
- mRSS, mean: 28.5 (SD 8.7)
- FVC, mean: 74.5% (SD 14.8%)
- DLCO, mean: 53.9% (SD 7.6%)
- HAQ-DI, mean: 1.2 (SD 0.6)
- Creatinine clearance, mean: 122.8 (SD 41.7) mL/minute
- Smoking status number current or former/never: 14/22

Inclusion criteria

- Aged 18–69 years
- Diagnosis of SSc according to ACR criteria
- Disease duration ≤ 5 years with pulmonary or renal involvement
- Extensive skin and internal organ involvement involving either the lungs or kidneys, that threatens the participant's life
  - DLCO or FVC < 70% of predicted value
  - Renal involvement defined as previous scleroderma-related renal disease

Exclusion criteria

- Respiratory: DLCO < 40% or FVC < 45%
- Renal: creatinine clearance < 40 mL/minute
- Cardiac: LVEF < 50%; atrial fibrillation necessitating oral anticoagulation; uncontrolled ventricular arrhythmias; hemodynamically significant pericardial effusion
- Active gastric antral vascular ectasia
- Pulmonary arterial hypertension
- Treatment with > 6 months of cyclophosphamide
- Concurrent neoplasms or myelodysplasia

Interventions	Myeloablative selective HSCT vs IV cyclophosphamide
	• HSCT
	  ○ Mobilization: G-CSF; leukapheresis and selection of CD34+ stem cells
	  ○ Conditioning: fractionated total-body irradiation (800 cGy with pulmonary and renal shields limiting exposure to 200 cGy), cyclophosphamide 120 mg/kg over 2 days, and equine ATG (90 mg/kg) over 6 days administered with IV methylprednisolone 1 mg/kg
	  ○ Post-transplant care including glucocorticoids, lisinopril, G-CSF, and anti-infective agents
	• Cyclophosphamide
	  ○ IV cyclophosphamide initially 500 mg/m^2 followed by 750 mg/m^2 pulse monthly for 11 months

Outcomes	Primary outcome: GRCS at 54 months Secondary outcomes: event-free survival at 48 and 54 months, FVC, HAQ-DI, mRSS, DLCO, SF-36, serious adverse events, treatment toxicity, and treatment-related mortality

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Sullivan 2018 *(Continued)*

	The GRCS combines mortality and other longitudinal outcomes in a hierarchical ordering. The outcomes in their order of hierarchy was agreed upon by the SCOT Steering Committee and was: death (at any time period), failure of event-free survival, FVC, HAQ-DI, and mRSS Quote: "To derive the GRCS, each subject is first compared to every other subject and assigned a pairwise comparison score of 1 (better off), 0 (no different), or −1 (worse off). As a result of these pair-wise comparisons, each subject will have a set of n−1 subscores (where n equals the number of subjects), which 1 DLCO change = {DLCO at assessment ((% predicted) – DLCO at baseline (% predicted))}/ DLCO at baseline (% predicted). Adjusted DLCO will be used based upon a subject's hemoglobin 17 gm/dL and altitude adjustments. Hemoglobin will be adjusted per the Cotes (1972) formula, and percent predicted per the Crapo Morris equation. 2 FVC change = {FVC at assessment ((% predicted) – FVC at baseline (% predicted))}/ FVC at baseline (% predicted). 6 are then summed to yield the GRCS. To determine subscores for pairs of subjects, subject pairs are compared with respect to each component outcome in the hierarchy, in sequence, until either ties are broken or all components are evaluated." (supplementary appendix, pp.5–6)
Notes	Funding source: NIAID and NIH to Duke University Conflicts of interest: 8 authors reported funding by NIAID or NIH

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated randomization. Quote: "For randomization, sites used a secure interactive web response system developed and maintained at the Statistical and Clinical Coordinating Center." (supplementary appendix, p.11)
Allocation concealment (selection bias)	Low risk	Randomization was 1:1 and stratified by site. Quote: "Randomization was performed on a 1:1 basis and stratified according to site, with the difference in the number of participants in the two treatment groups constrained to two or fewer at each site." (p.36)
Blinding of participants and personnel (performance bias) Participant-reported outcomes	Unclear risk	Blinding not possible due to the nature of the interventions. Therefore, we placed participant-reported outcomes such as the HAQ-DI, SF-36, and health-related quality of life VAS as unclear as participants may have been influenced depending on the treatment that they received. Quote: "Although blinding is not possible in a transplantation trial, objective outcomes were placed higher in the global rank composite score hierarchy to mitigate this limitation." (p.46)
Blinding of participants and personnel (performance bias) Objective outcomes	Low risk	Blinding not possible due to the nature of the interventions; however, objective outcomes (mortality, event-free survival, mRSS, pulmonary function tests, serious adverse events, renal function, cardiac function) measured in this study do not seem likely to be altered by performance bias. Quote: "Although blinding is not possible in a transplantation trial, objective outcomes were placed higher in the global rank composite score hierarchy to mitigate this limitation." (p.46)
Blinding of outcome assessment (detection bias) Participant-reported outcomes	High risk	Participants and assessors were not blinded. For participant-reported outcomes such as the HAQ-DI, SF-36, and health-related quality of life VAS, we placed this as a high risk for potential bias.

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Sullivan 2018 *(Continued)*

Blinding of outcome assessment (detection bias) Objective outcomes	Low risk	Participants and assessors were not blinded; however, the objective outcomes (mortality, event-free survival, mRSS, pulmonary function tests, serious adverse events, renal function, cardiac function) are unlikely to be influenced.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Large dropout, though there was intention-to-treat at randomization. Quote: "The intention-to-treat population was defined as all the participants who had undergone randomization." (p.38)
Selective reporting (reporting bias)	Low risk	Primary outcome changed due to unexpected low enrolment, but all secondary outcomes reported as per protocol.
Other bias	Low risk	None.

van Laar 2014

Study characteristics	
Methods	Study design: multicenter active controlled parallel group randomized controlled trial Trial time period: 2001–2008 Setting: multicenter in 10 countries and 29 centers in Europe Analysis: intention-to-treat
Participants	**Baseline characteristics** *Cyclophosphamide (77 participants)* • Mean age: 43.3 (SD 11.5) years • Number male/female: 28/49 • Disease duration, mean: 1.5 (SD 1.4) years • mRSS, mean: 25.8 (SD 7.9) • FVC, mean: 81.1% (SD 17.6%) • DLCO, mean: 57.7% (SD 14%) • HAQ-DI, mean: 1.44 (SD 0.84) • Creatinine clearance, mean: 76.5 mL/minute (SD 26.0) mL/minute • Smoking status number current/former/never: 13/30/34 *HSCT (79 participants)* • Mean age: 44.2 (SD 11.1) years • Number male/female: 36/43 • Disease duration, mean: 1.4 (SD 1.2) years • mRSS, mean: 24.8 (SD 8.1) • FVC, mean: 81.7% (SD 19.3%) • DLCO, mean: 59.3% (SD 14.3%) • HAQ-DI, mean: 1.25 (SD 0.74) • Creatinine clearance, mean: 76.8 (SD 26.1) mL/minute • Smoking status number current/former/never: 10/31/38 **Inclusion criteria** • Aged 18–65 years • Diagnosis of SSc according to ACR criteria

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

van Laar 2014 *(Continued)*

- Disease duration ≤ 4 years with mRSS ≥ 15 with evidence of organ involvement as characterized by:
  - DLCO or FVC < 80% and evidence of interstitial lung disease
  - renal abnormalities
  - cardiac involvement
- Disease duration < 2 years with mRSS ≥ 20 with involvement of trunk with erythrocyte sedimentation rate > 25 or hemoglobin < 11 g/dL

Exclusion criteria

- Respiratory: mean pulmonary arterial pressure > 50 mmHg, DLCO < 40%
- Renal: creatinine clearance < 40 mL/minute
- Cardiac: clinical evidence of refractory congestive heart failure; LVEF < 45%; atrial fibrillation necessitating oral anticoagulation; uncontrolled ventricular arrhythmias; hemodynamically significant pericardial effusion
- Liver failure
- Concurrent neoplasms or myelodysplasia

Interventions	Non-myeloablative selective HSCT vs IV cyclophosphamide
	• HSCT
	  ◦ Mobilization: cyclophosphamide 4 g/m^2 administered in equal amounts over 2 days with filgrastim 10 mg/kg/day; leukapheresis and selection of CD34+ stem cells using immunomagnetic separation (CliniMACS, Miltenyl Biotec)
	  ◦ Conditioning: cyclophosphamide 200 mg/kg (administered over 4 days) with rabbit ATG 7.5 mg/kg (administered over 3 days) with IV methylprednisolone 1 mg/kg prior to rabbit ATG infusions
	  ◦ Infusion of CD34+ autologous stem cells
	• Cyclophosphamide
	  ◦ Cyclophosphamide 750 mg/m^2 IV pulse monthly for 12 months

Outcomes	Primary outcomes: EFS defined as time in days from randomization until occurrence of death due to any cause or development of persistent major organ failure defined as: • LVEF < 30% • resting arterial oxygen tension < 60 mmHg or resting arterial carbon dioxide tension > 50 mmHg without oxygen supply (or both) and • need for renal replacement therapy Secondary outcomes: treatment-related mortality, serious adverse events, changes in mRSS, quality of life (HAQ-DI, SF-36, health-related quality of life VAS), body weight, creatinine clearance, FVC, total lung capacity, residual volume, DLCO, EuroQol-5D, and index-based utility score
Notes	Funding: funded by the European Group for Blood and Marrow Transplant, the European League Against Rheumatism, the Assistance Publique-Hôpitaux de Paris, French Ministry of Health Programme Hospitaller de Recherche Clinique, Groupe Francophone de Recherche sur la Sclérodermie, the Association des Sclérodermiques de France, the National Institute for Health Research, and grants from Imtix-Sangstat and Amgen Europ, and Miltenyi-Biotec Conflicts of Interest: 11 authors disclosed royalties, consulting fees, research support, or speaker fees from various different companies including Genentech, Roche, Menarini, Bristol-Myers Squibb, Abbott, Novartis, Miltenyl, Tigenix, Pfizer, GlaxoSmithKline, Chugal, and Actellon

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated randomization.

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

van Laar 2014 *(Continued)*

		Quote: "was performed centrally by telephone at the study administration office according to a computer-generated randomization program for each site." (p.2491)
Allocation concealment (selection bias)	Low risk	1:1 ratio by blocked randomization with random block sizes. Quote: "patients were randomly assigned in a 1:1 ratio by blocked randomization to receive HSCT or 12 intravenous pulses of cyclophosphamide (Figure 1). Block randomization was performed centrally by telephone at the study administration office according to a computer-generated randomization program for each site, with random block sizes (2,4, 6)." (p.2491)
Blinding of participants and personnel (performance bias) Participant-reported outcomes	Unclear risk	Blinding not possible. Participant-reported outcomes (e.g. SF-36, HAQ-DI, health-related quality of life VAS) may be influenced. Quote: "Patients and assessors were not blinded." (p.2492)
Blinding of participants and personnel (performance bias) Objective outcomes	Low risk	Blinding not possible; however, objective outcomes (mortality, EFS, mRSS, pulmonary function tests, serious adverse events, renal function, and cardiac function) were unlikely to be influenced. Quote: "Patients and assessors were not blinded." (p.2492)
Blinding of outcome assessment (detection bias) Participant-reported outcomes	High risk	Participants and assessors not blinded. This creates a high risk of bias for participant-reported outcomes (HAQ-DI, SF-36, and health-related quality of life VAS). Quote: "Patients and assessors were not blinded." (p.2492)
Blinding of outcome assessment (detection bias) Objective outcomes	Low risk	Participants and assessors not blinded; however, objective outcomes (mortality, EFS, mRSS, pulmonary function tests, serious adverse events, renal function, and cardiac function) were unlikely to be influenced. Quote: "Patients and assessors were not blinded." (p.2492)
Incomplete outcome data (attrition bias) All outcomes	Low risk	High dropout (32%) in the control group that was largely because of non-adherence (9/77 dropouts) and death. The study incorporated an intention-to-treat model. There was no evidence that the result was not biased by missing outcome data and the missingness in the outcomes could depend on the true value. Though this occurred in the control group and if the missingness is like a result of the true value then it is unlikely to have affected the results that were presented in the study. Quote: "We analyzed all data by intention-to-treat (ITT)." (p.2492)
Selective reporting (reporting bias)	Low risk	Outcome measures from protocol were reported.
Other bias	Low risk	None.

ACR: American College of Rheumatology; ATG: antithymocyte globulin; CT: computer tomography; DLCO: diffusing in liters for carbon monoxide capacity; ECG: electrocardiogram; EFS: event-free survival; FVC: forced vital capacity; G-CSF: granulocyte colony-stimulating factor; GRCS: Global Rank Composite Score; HAQ-DI: Health Assessment Questionnaire Disability Index; HSCT: hematopoietic stem cell transplantation; IV: intravenous; LVEF: left ventricular ejection fraction; mRSS: modified Rodnan skin score; NIAID: National Institute of Allergy and Infectious Disease; NIH: National Institutes of Health; SCOT: Scleroderma: Cyclophosphamide or Transplantation; SD: standard deviation; SF-36: 36-item Short Form; SSC: systemic sclerosis; VAS: visual analog scale.

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Characteristics of ongoing studies *[ordered by study ID]*

van Laar 2020

Study name	Upfront autologous HSCT versus Immunosuppression in Early Diffuse Cutaneous Systemic Sclerosis (UPSIDE)
Methods	Multicenter, randomized, open-label trial
Participants	2013 ACR-EULAR classification criteria for SSc with disease duration ≤ 2 years. Must have had severe skin, pulmonary, renal, or cardiac involvement. Plan to enroll 120 participants
Interventions	Autologous non-myeloablative HSCT vs cyclophosphamide (followed by mycophenolate mofetil)
Outcomes	Primary outcome: event-free survival at 24 months Secondary outcomes: progression-free survival, overall mortality, adverse events, changes in skin involvement, changes in cardiac function, changes in pulmonary function, quality of life changes
Starting date	September 2020
Contact information	Julia Spierings, MD J.Spierings@umcutrecht.nl
Notes	ClinicalTrials.gov identifier: NCT04464434

ACR-EULAR: American College of Rheumatology/European Alliance of Associations for Rheumatology; HSCT: hematopoietic stem cell transplantation; SSC: systemic sclerosis.

DATA AND ANALYSES

Comparison 1. Efficacy – survival

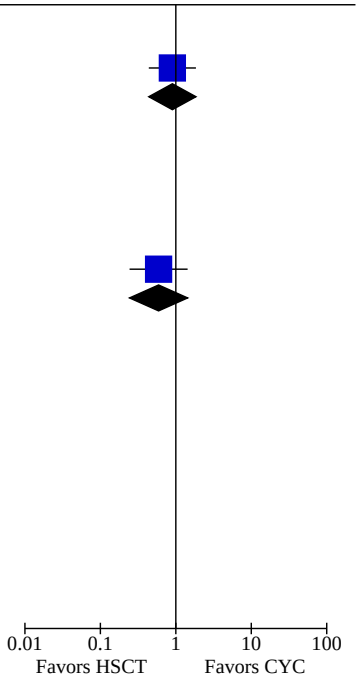
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.1 Overall mortality	3		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
1.1.1 Autologous non-myeloablative selective HSCT	1	156	Risk Ratio (M-H, Random, 95% CI)	0.90 [0.44, 1.85]
1.1.2 Autologous myeloablative selective HSCT	1	75	Risk Ratio (M-H, Random, 95% CI)	0.59 [0.24, 1.43]
1.1.3 Autologous non-myeloablative non-selective HSCT	1	19	Risk Ratio (M-H, Random, 95% CI)	Not estimable
1.2 Event-free survival	2		Hazard Ratio (IV, Random, 95% CI)	Subtotals only
1.2.1 Autologous non-myeloablative selective HSCT	1		Hazard Ratio (IV, Random, 95% CI)	0.34 [0.16, 0.74]

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.2.2 Autologous myeloablative selective HSCT	1		Hazard Ratio (IV, Random, 95% CI)	0.54 [0.23, 1.27]

Analysis 1.1. Comparison 1: Efficacy – survival, Outcome 1: Overall mortality

Study or Subgroup	HSCT Events	HSCT Total	Cyclophosphamide Events	Cyclophosphamide Total	Weight	Risk Ratio M-H, Random, 95% CI
1.1.1 Autologous non-myeloablative selective HSCT						
van Laar 2014 (1)	12	79	13	77	100.0%	0.90 [0.44 , 1.85]
Subtotal (95% CI)		**79**		**77**	**100.0%**	**0.90 [0.44 , 1.85]**
Total events:	12		13			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.29 (P = 0.77)						
1.1.2 Autologous myeloablative selective HSCT						
Sullivan 2018 (2)	6	36	11	39	100.0%	0.59 [0.24 , 1.43]
Subtotal (95% CI)		**36**		**39**	**100.0%**	**0.59 [0.24 , 1.43]**
Total events:	6		11			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.16 (P = 0.24)						
1.1.3 Autologous non-myeloablative non-selective HSCT						
Burt 2011 (3)	0	10	0	9		Not estimable
Subtotal (95% CI)		**10**		**9**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
Test for subgroup differences: $Chi^2 = 0.52$, df = 1 (P = 0.47), $I^2 = 0\%$						

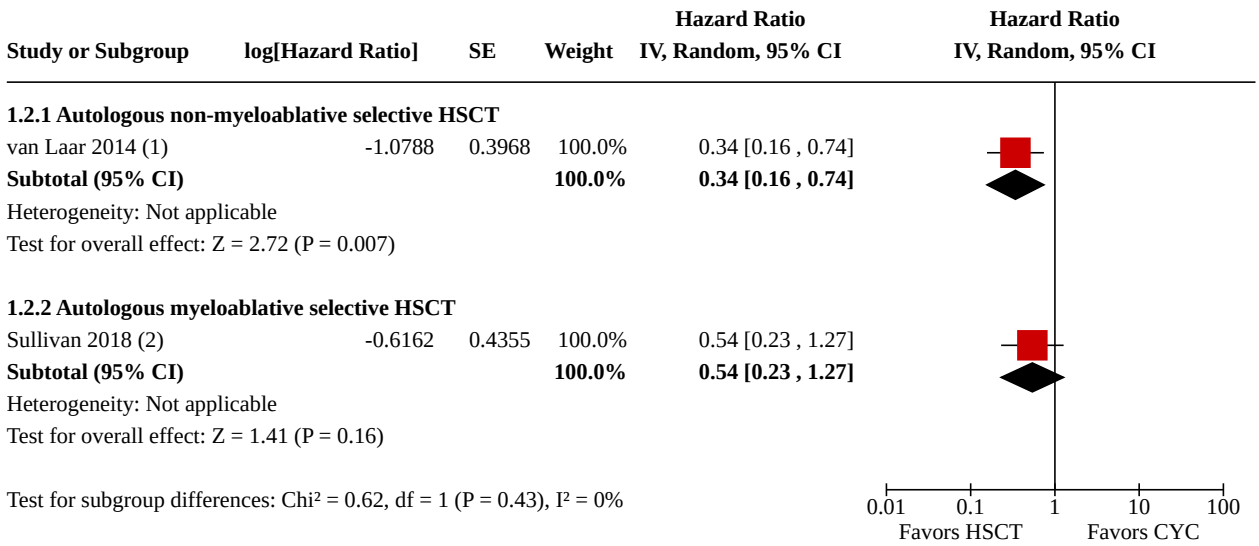


Footnotes

(1) 1-year death rate.
(2) 4.5-year death rate
(3) 2-year death rate

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.2. Comparison 1: Efficacy – survival, Outcome 2: Event-free survival



Footnotes

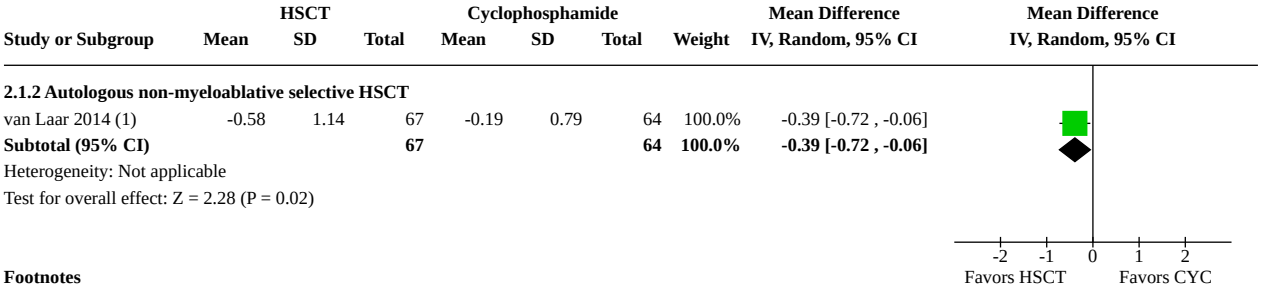
(1) 48 month EFS defined as survival without death or major organ failure.

(2) 54 month EFS defined as survival without significant organ damage or death

Comparison 2. Participant-reported outcomes – HAQ-DI

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.1 HAQ-DI	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.1.2 Autologous non-myeloablative selective HSCT	1	131	Mean Difference (IV, Random, 95% CI)	-0.39 [-0.72, -0.06]
2.2 Improvement in HAQ-DI (≥ 0.4 change)	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
2.2.1 Autologous myeloablative selective HSCT	1	75	Risk Ratio (M-H, Fixed, 95% CI)	3.43 [1.54, 7.62]

Analysis 2.1. Comparison 2: Participant-reported outcomes – HAQ-DI, Outcome 1: HAQ-DI

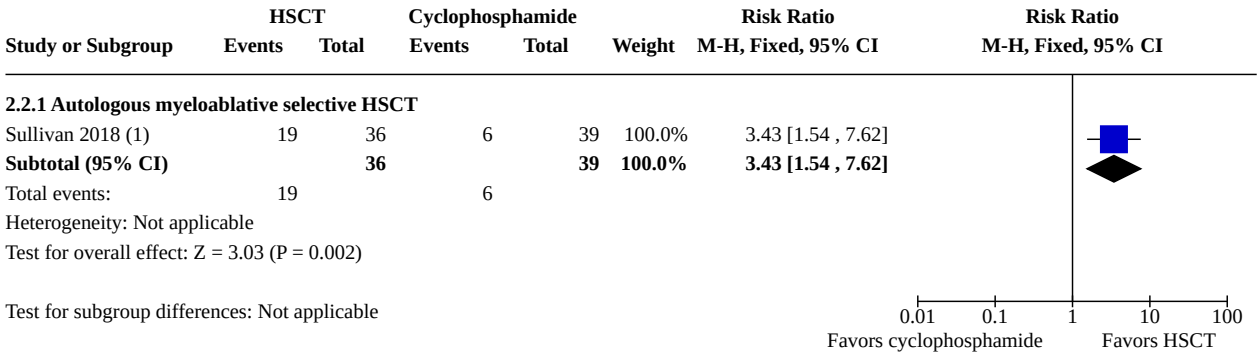


Footnotes

(1) 2 years follow-up; values are mean area under the curve (AUC)

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.2. Comparison 2: Participant-reported outcomes – HAQ-DI, Outcome 2: Improvement in HAQ-DI (≥ 0.4 change)



Footnotes

(1) Data presented in two subsets - EFS survivors and EFS failure. EFS survivor assessments were done at 54 months. EFS failure assessments are from t

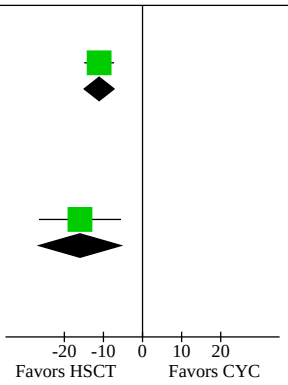
Comparison 3. Efficacy – skin thickness

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.1 Modified Rodnan skin scores	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
3.1.1 Autologous non-myeloablative selective HSCT	1	131	Mean Difference (IV, Random, 95% CI)	-11.10 [-14.92, -7.28]
3.1.2 Autologous non-myeloablative non-selective HSCT	1	19	Mean Difference (IV, Random, 95% CI)	-16.00 [-26.49, -5.51]
3.2 Modified Rodnan skin score improvement (≥ 25% change in mRSS OR ± ≥ 5 if baseline mRSS ≤ 20)	1		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
3.2.1 Autologous myeloablative selective HSCT	1	75	Risk Ratio (M-H, Random, 95% CI)	1.51 [1.06, 2.13]

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.1. Comparison 3: Efficacy – skin thickness, Outcome 1: Modified Rodnan skin scores

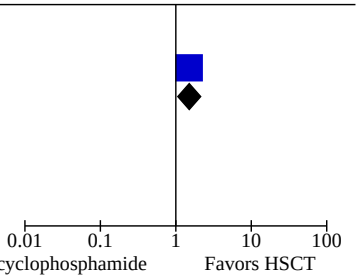
Study or Subgroup	HSCT Mean	HSCT SD	HSCT Total	Cyclophosphamide Mean	Cyclophosphamide SD	Cyclophosphamide Total	Weight	Mean Difference IV, Random, 95% CI
3.1.1 Autologous non-myeloablative selective HSCT								
van Laar 2014 (1)	-19.9	10.2	67	-8.8	12	64	100.0%	-11.10 [-14.92 , -7.28]
Subtotal (95% CI)			**67**			**64**	**100.0%**	**-11.10 [-14.92 , -7.28]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 5.69 (P < 0.00001)								
3.1.2 Autologous non-myeloablative non-selective HSCT								
Burt 2011 (2)	-13	7.9	10	3	14.2	9	100.0%	-16.00 [-26.49 , -5.51]
Subtotal (95% CI)			**10**			**9**	**100.0%**	**-16.00 [-26.49 , -5.51]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.99 (P = 0.003)								
Test for subgroup differences: Chi² = 0.74, df = 1 (P = 0.39), I² = 0%								



Footnotes
(1) 2 years follow-up; values are mean area under the curve (AUC)
(2) Outcome is over 1 year

Analysis 3.2. Comparison 3: Efficacy – skin thickness, Outcome 2: Modified Rodnan skin score improvement (≥ 25% change in mRSS OR ± ≥ 5 if baseline mRSS ≤ 20)

Study or Subgroup	HSCT Events	HSCT Total	Cyclophosphamide Events	Cyclophosphamide Total	Weight	Risk Ratio M-H, Random, 95% CI
3.2.1 Autologous myeloablative selective HSCT						
Sullivan 2018 (1)	31	39	19	36	100.0%	1.51 [1.06 , 2.13]
Subtotal (95% CI)		**39**		**36**	**100.0%**	**1.51 [1.06 , 2.13]**
Total events:	31		19			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.31 (P = 0.02)						



Footnotes
(1) Data presented in two subsets - EFS survivors and EFS failure. EFS survivor assessments were done at 54 months. EFS failure assessments are from the l

Comparison 4. Efficacy – interstitial lung disease

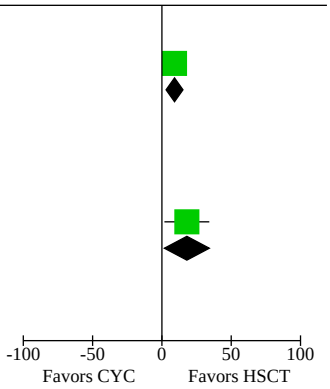
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
4.1 Predicted forced vital capacity (% predicted)	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
4.1.1 Autologous non-myeloablative selective HSCT	1	131	Mean Difference (IV, Random, 95% CI)	9.10 [3.02, 15.18]
4.1.2 Autologous non-myeloablative non-selective HSCT	1	19	Mean Difference (IV, Random, 95% CI)	18.00 [1.81, 34.19]
4.2 Predicted forced vital capacity improvement (10% change in FVC % predicted)	1		Risk Ratio (M-H, Random, 95% CI)	Subtotals only

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
4.2.1 Autologous myeloablative selective HSCT	1	75	Risk Ratio (M-H, Random, 95% CI)	1.63 [0.75, 3.51]
4.3 DLCO (% predicted)	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
4.3.1 Autologous non-myeloablative selective HSCT	1	131	Mean Difference (IV, Fixed, 95% CI)	-0.60 [-6.02, 4.82]
4.3.2 Autologous non-myeloablative non-selective HSCT	1	19	Mean Difference (IV, Fixed, 95% CI)	12.00 [-14.78, 38.78]
4.4 DLCO improvement (≥ 15% change in DLCO % predicted)	1	75	Risk Ratio (M-H, Random, 95% CI)	0.87 [0.25, 2.98]
4.4.1 Autologous myeloablative selective HSCT	1	75	Risk Ratio (M-H, Random, 95% CI)	0.87 [0.25, 2.98]
4.5 Predicted total lung capacity	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
4.5.1 Autologous non-myeloablative selective HSCT	1	131	Mean Difference (IV, Fixed, 95% CI)	6.40 [1.00, 11.80]
4.5.2 Autologous non-myeloablative non-selective HSCT	1	19	Mean Difference (IV, Fixed, 95% CI)	13.00 [-3.50, 29.50]

Analysis 4.1. Comparison 4: Efficacy – interstitial lung disease, Outcome 1: Predicted forced vital capacity (% predicted)

	HSCT			Cyclophosphamide				Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
4.1.1 Autologous non-myeloablative selective HSCT									
van Laar 2014 (1)	6.3	18.3	67	-2.8	17.2	64	100.0%	9.10 [3.02 , 15.18]	
Subtotal (95% CI)			**67**			**64**	**100.0%**	**9.10 [3.02 , 15.18]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 2.93 (P = 0.003)									
4.1.2 Autologous non-myeloablative non-selective HSCT									
Burt 2011 (2)	12	15.7	10	-6	19.8	9	100.0%	18.00 [1.81 , 34.19]	
Subtotal (95% CI)			**10**			**9**	**100.0%**	**18.00 [1.81 , 34.19]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 2.18 (P = 0.03)									
Test for subgroup differences: Chi² = 1.02, df = 1 (P = 0.31), I² = 1.7%									



Footnotes

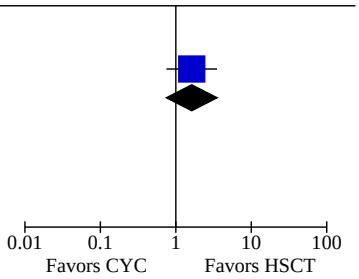
(1) 2 years follow-up; values are mean area under the curve (AUC)

(2) Outcome is over 1 year

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 4.2. Comparison 4: Efficacy – interstitial lung disease, Outcome 2: Predicted forced vital capacity improvement (10% change in FVC % predicted)

Study or Subgroup	HSCT Events	HSCT Total	Cyclophosphamide Events	Cyclophosphamide Total	Weight	Risk Ratio M-H, Random, 95% CI
4.2.1 Autologous myeloablative selective HSCT						
Sullivan 2018 (1)	12	36	8	39	100.0%	1.63 [0.75 , 3.51]
Subtotal (95% CI)		**36**		**39**	**100.0%**	**1.63 [0.75 , 3.51]**
Total events:	12		8			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.23 (P = 0.22)						
Test for subgroup differences: Not applicable						

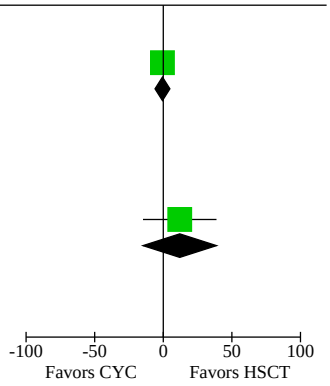


Footnotes

(1) Data presented in two subsets - EFS survivors and EFS failure. EFS survivor assessments were done at 54 months. EFS failure assessments are from the l

Analysis 4.3. Comparison 4: Efficacy – interstitial lung disease, Outcome 3: DLCO (% predicted)

Study or Subgroup	HSCT Mean	HSCT SD	HSCT Total	Cyclophosphamide Mean	Cyclophosphamide SD	Cyclophosphamide Total	Weight	Mean Difference IV, Fixed, 95% CI
4.3.1 Autologous non-myeloablative selective HSCT								
van Laar 2014 (1)	-4.7	13.7	67	-4.1	17.6	64	100.0%	-0.60 [-6.02 , 4.82]
Subtotal (95% CI)			**67**			**64**	**100.0%**	**-0.60 [-6.02 , 4.82]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.22 (P = 0.83)								
4.3.2 Autologous non-myeloablative non-selective HSCT								
Burt 2011 (2)	11	18.6	10	-1	37	9	100.0%	12.00 [-14.78 , 38.78]
Subtotal (95% CI)			**10**			**9**	**100.0%**	**12.00 [-14.78 , 38.78]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.88 (P = 0.38)								
Test for subgroup differences: $Chi^2 = 0.82$, df = 1 (P = 0.37), $I^2 = 0\%$								



Footnotes

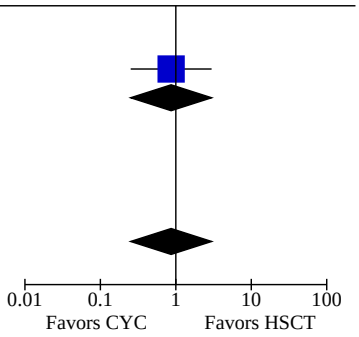
(1) Outcome is over 2 years

(2) Outcome is over 1 year

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 4.4. Comparison 4: Efficacy – interstitial lung disease, Outcome 4: DLCO improvement (≥ 15% change in DLCO % predicted)

Study or Subgroup	HSCT Events	HSCT Total	Cyclophosphamide Events	Cyclophosphamide Total	Weight	Risk Ratio M-H, Random, 95% CI
4.4.1 Autologous myeloablative selective HSCT						
Sullivan 2018 (1)	4	36	5	39	100.0%	0.87 [0.25 , 2.98]
Subtotal (95% CI)		**36**		**39**	**100.0%**	**0.87 [0.25 , 2.98]**
Total events:	4		5			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.23 (P = 0.82)						
Total (95% CI)		**36**		**39**	**100.0%**	**0.87 [0.25 , 2.98]**
Total events:	4		5			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.23 (P = 0.82)						
Test for subgroup differences: Not applicable						

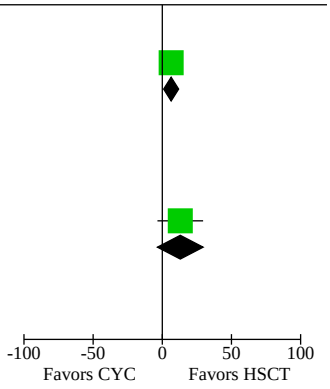


Footnotes

(1) Data presented in two subsets - EFS survivors and EFS failure. EFS survivor assessments were done at 54 months. EFS failure assessments are from the l

Analysis 4.5. Comparison 4: Efficacy – interstitial lung disease, Outcome 5: Predicted total lung capacity

Study or Subgroup	HSCT Mean	HSCT SD	HSCT Total	Cyclophosphamide Mean	Cyclophosphamide SD	Cyclophosphamide Total	Weight	Mean Difference IV, Fixed, 95% CI
4.5.1 Autologous non-myeloablative selective HSCT								
van Laar 2014 (1)	5.1	17.5	67	-1.3	13.9	64	100.0%	6.40 [1.00 , 11.80]
Subtotal (95% CI)			**67**			**64**	**100.0%**	**6.40 [1.00 , 11.80]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.32 (P = 0.02)								
4.5.2 Autologous non-myeloablative non-selective HSCT								
Burt 2011 (2)	4	17.9	10	-9	18.7	9	100.0%	13.00 [-3.50 , 29.50]
Subtotal (95% CI)			**10**			**9**	**100.0%**	**13.00 [-3.50 , 29.50]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.54 (P = 0.12)								
Test for subgroup differences: $Chi^2 = 0.56$, df = 1 (P = 0.46), $I^2 = 0\%$								



Footnotes

(1) Outcome is over 2 years

(2) Outcome is over 1 year

Comparison 5. Safety

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
5.1 Serious adverse events	3		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
5.1.1 Autologous non-myeloablative selective HSCT	1	156	Risk Ratio (M-H, Random, 95% CI)	1.66 [1.20, 2.29]
5.1.2 Autologous myeloablative selective HSCT	1	71	Risk Ratio (M-H, Random, 95% CI)	1.43 [0.99, 2.08]

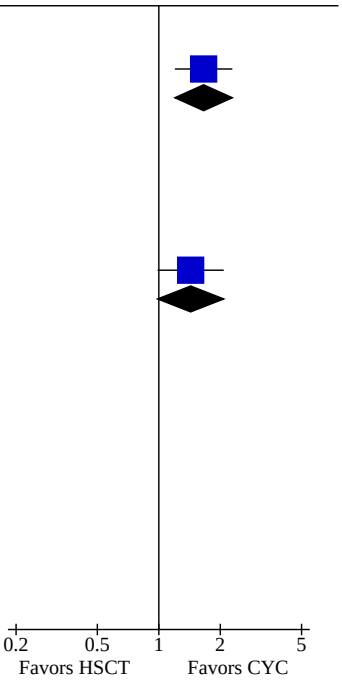
Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
5.1.3 Autologous non-myeloablative non-selective HSCT	1	19	Risk Ratio (M-H, Random, 95% CI)	Not estimable
5.2 Treatment-related mortality	3		Peto Odds Ratio (Peto, Fixed, 95% CI)	Subtotals only
5.2.1 Autologous non-myeloablative selective HSCT	1	156	Peto Odds Ratio (Peto, Fixed, 95% CI)	7.91 [1.91, 32.67]
5.2.2 Autologous myeloablative selective HSCT	1	75	Peto Odds Ratio (Peto, Fixed, 95% CI)	8.03 [0.16, 406.02]
5.2.3 Autologous non-myeloablative non-selective HSCT	1	19	Peto Odds Ratio (Peto, Fixed, 95% CI)	Not estimable
5.3 Serious non-lethal infections	3		Odds Ratio (IV, Random, 95% CI)	Subtotals only
5.3.1 Autologous non-myeloablative selective HSCT	1	156	Odds Ratio (IV, Random, 95% CI)	2.06 [0.59, 7.13]
5.3.2 Autologous myeloablative selective HSCT	1	71	Odds Ratio (IV, Random, 95% CI)	1.79 [0.59, 5.39]
5.3.3 Autologous non-myeloablative non-selective HSCT	1	19	Odds Ratio (IV, Random, 95% CI)	3.43 [0.29, 40.95]
5.4 Study withdrawals (including non-adherence, non-lethal adverse events, organ failure, dropout)	3		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
5.4.1 Autologous non-myeloablative selective HSCT	1	156	Risk Ratio (M-H, Random, 95% CI)	0.22 [0.08, 0.61]
5.4.2 Autologous myeloablative selective HSCT	1	75	Risk Ratio (M-H, Random, 95% CI)	0.72 [0.29, 1.83]
5.4.3 Autologous non-myeloablative non-selective HSCT	1	19	Risk Ratio (M-H, Random, 95% CI)	Not estimable
5.5 Renal failure	3		Peto Odds Ratio (Peto, Fixed, 95% CI)	Subtotals only
5.5.1 Autologous non-myeloablative selective HSCT	1	156	Peto Odds Ratio (Peto, Fixed, 95% CI)	1.99 [0.62, 6.45]
5.5.2 Autologous myeloablative selective HSCT	1	71	Peto Odds Ratio (Peto, Fixed, 95% CI)	0.72 [0.12, 4.36]
5.5.3 Autologous non-myeloablative non-selective HSCT	1	19	Peto Odds Ratio (Peto, Fixed, 95% CI)	6.69 [0.13, 338.79]

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 5.1. Comparison 5: Safety, Outcome 1: Serious adverse events

Study or Subgroup	HSCT Events	HSCT Total	Cyclophosphamide Events	Cyclophosphamide Total	Weight	Risk Ratio M-H, Random, 95% CI
5.1.1 Autologous non-myeloablative selective HSCT						
van Laar 2014 (1)	51	79	30	77	100.0%	1.66 [1.20 , 2.29]
Subtotal (95% CI)		**79**		**77**	**100.0%**	**1.66 [1.20 , 2.29]**
Total events:	51		30			
Heterogeneity: Not applicable						
Test for overall effect: Z = 3.06 (P = 0.002)						
5.1.2 Autologous myeloablative selective HSCT						
Sullivan 2018 (2)	25	34	19	37	100.0%	1.43 [0.99 , 2.08]
Subtotal (95% CI)		**34**		**37**	**100.0%**	**1.43 [0.99 , 2.08]**
Total events:	25		19			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.89 (P = 0.06)						
5.1.3 Autologous non-myeloablative non-selective HSCT						
Burt 2011 (3)	0	10	0	9		Not estimable
Subtotal (95% CI)		**10**		**9**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
Test for subgroup differences: Chi² = 0.34, df = 1 (P = 0.56), I² = 0%						



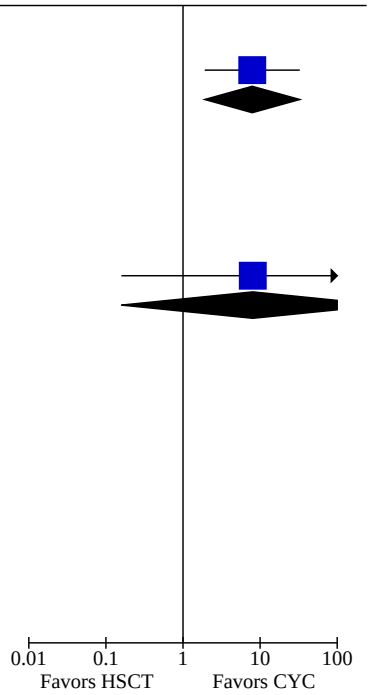
Footnotes

(1) Adverse events within the first two years
(2) 6-year adverse events
(3) Adverse events not clearly reported at one year

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 5.2. Comparison 5: Safety, Outcome 2: Treatment-related mortality

Study or Subgroup	HSCT Events	HSCT Total	Cyclophosphamide Events	Cyclophosphamide Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
5.2.1 Autologous non-myeloablative selective HSCT						
van Laar 2014 (1)	8	79	0	77	100.0%	7.91 [1.91 , 32.67]
Subtotal (95% CI)		**79**		**77**	**100.0%**	**7.91 [1.91 , 32.67]**
Total events:	8		0			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.86 (P = 0.004)						
5.2.2 Autologous myeloablative selective HSCT						
Sullivan 2018 (2)	1	36	0	39	100.0%	8.03 [0.16 , 406.02]
Subtotal (95% CI)		**36**		**39**	**100.0%**	**8.03 [0.16 , 406.02]**
Total events:	1		0			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.04 (P = 0.30)						
5.2.3 Autologous non-myeloablative non-selective HSCT						
Burt 2011	0	10	0	9		Not estimable
Subtotal (95% CI)		**10**		**9**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						



Footnotes

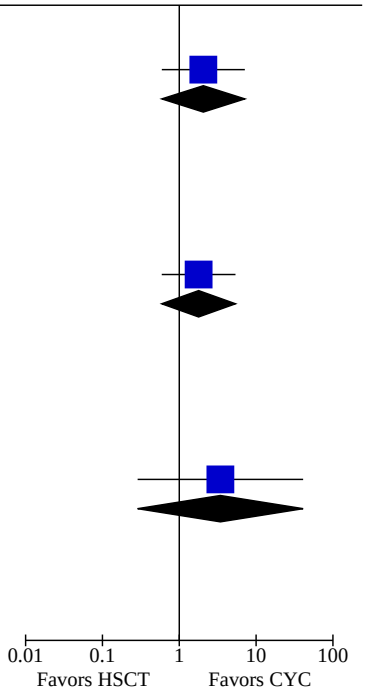
(1) Treatment-related deaths occurred during year 1.

(2) The SCOT trial reported an additional treatment-related death at 70 months. This is not included in the analysis as it occurred after 54 months.

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 5.3. Comparison 5: Safety, Outcome 3: Serious non-lethal infections

Study or Subgroup	HSCT Events	HSCT Total	Cyclophosphamide Events	Cyclophosphamide Total	Weight	Odds Ratio IV, Random, 95% CI
5.3.1 Autologous non-myeloablative selective HSCT						
van Laar 2014	8	79	4	77	100.0%	2.06 [0.59 , 7.13]
Subtotal (95% CI)		**79**		**77**	**100.0%**	**2.06 [0.59 , 7.13]**
Total events:	8		4			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.14 (P = 0.26)						
5.3.2 Autologous myeloablative selective HSCT						
Sullivan 2018	10	34	7	37	100.0%	1.79 [0.59 , 5.39]
Subtotal (95% CI)		**34**		**37**	**100.0%**	**1.79 [0.59 , 5.39]**
Total events:	10		7			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.03 (P = 0.30)						
5.3.3 Autologous non-myeloablative non-selective HSCT						
Burt 2011 (1)	3	10	1	9	100.0%	3.43 [0.29 , 40.95]
Subtotal (95% CI)		**10**		**9**	**100.0%**	**3.43 [0.29 , 40.95]**
Total events:	3		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.97 (P = 0.33)						



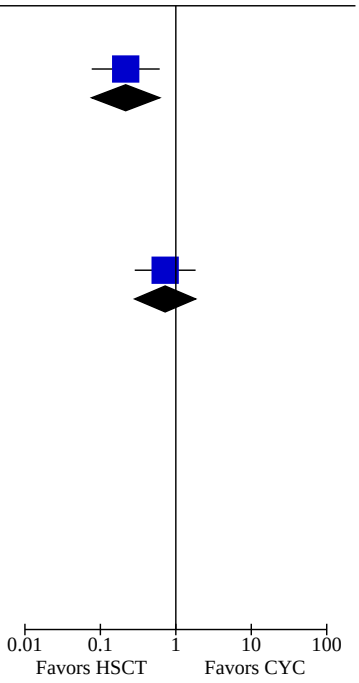
Footnotes

(1) clostridium difficile, micrococcus, cytomegalovirus, cellulitis

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 5.4. Comparison 5: Safety, Outcome 4: Study withdrawals (including non-adherence, non-lethal adverse events, organ failure, dropout)

Study or Subgroup	HSCT Events	HSCT Total	Cyclophosphamide Events	Cyclophosphamide Total	Weight	Risk Ratio M-H, Random, 95% CI
5.4.1 Autologous non-myeloablative selective HSCT						
van Laar 2014 (1)	4	79	18	77	100.0%	0.22 [0.08 , 0.61]
Subtotal (95% CI)		**79**		**77**	**100.0%**	**0.22 [0.08 , 0.61]**
Total events:	4		18			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.89 (P = 0.004)						
5.4.2 Autologous myeloablative selective HSCT						
Sullivan 2018	6	36	9	39	100.0%	0.72 [0.29 , 1.83]
Subtotal (95% CI)		**36**		**39**	**100.0%**	**0.72 [0.29 , 1.83]**
Total events:	6		9			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.69 (P = 0.49)						
5.4.3 Autologous non-myeloablative non-selective HSCT						
Burt 2011	0	10	0	9		Not estimable
Subtotal (95% CI)		**10**		**9**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						

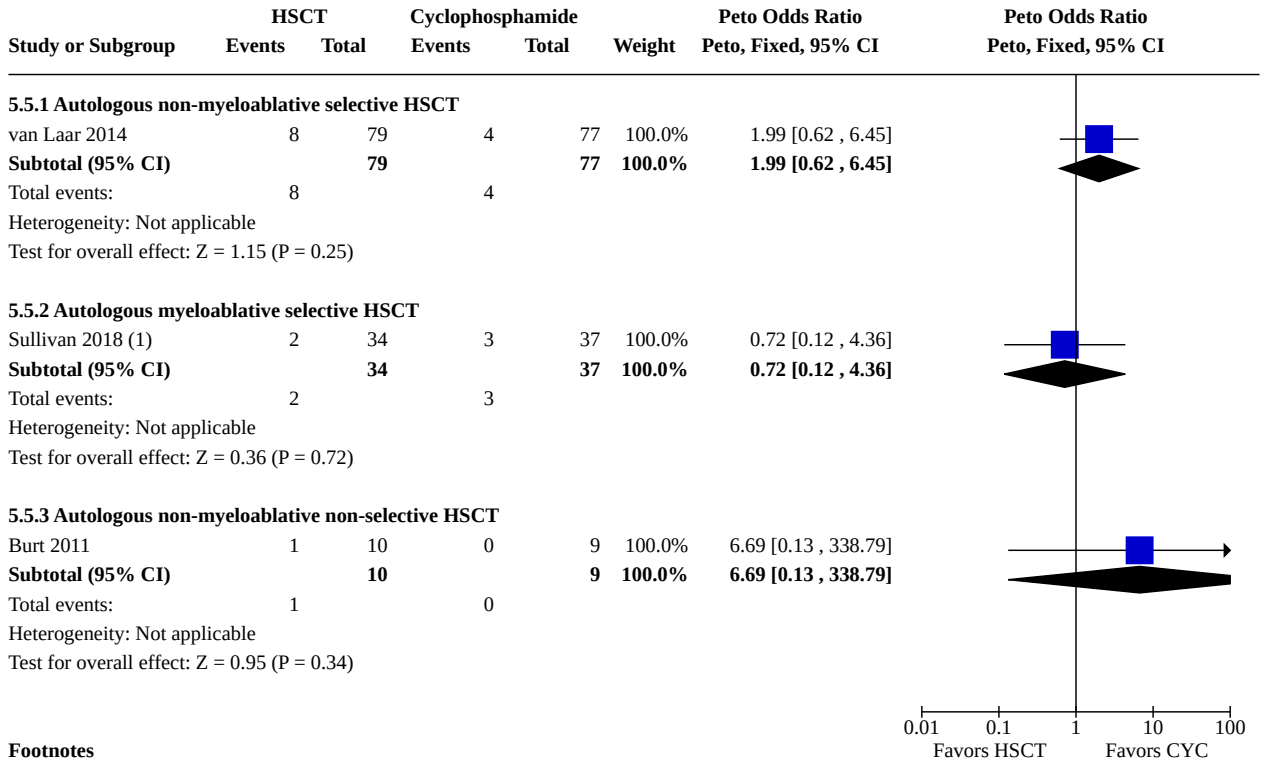


Footnotes

(1) median follow-up of 5.8 years

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 5.5. Comparison 5: Safety, Outcome 5: Renal failure



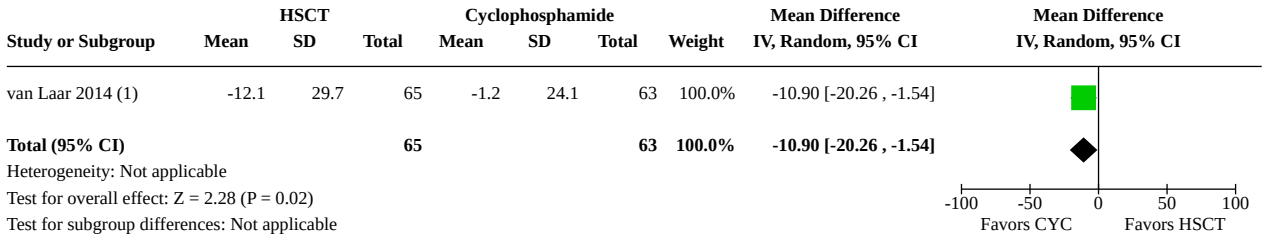
Footnotes

(1) Combining cases of scleroderma renal crisis and renal failure

Comparison 6. Efficacy – renal function (creatinine clearance)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
6.1 Autologous non-myeloablative selective HSCT	1	128	Mean Difference (IV, Random, 95% CI)	-10.90 [-20.26, -1.54]

Analysis 6.1. Comparison 6: Efficacy – renal function (creatinine clearance), Outcome 1: Autologous non-myeloablative selective HSCT



Footnotes

(1) Creatinine clearance, mL/min; values are area under the time response curve from baseline to 2 years follow-up

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

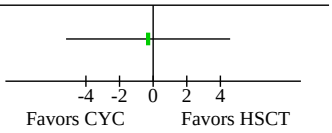
Comparison 7. Efficacy – cardiac function

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
7.1 Autologous non-myeloablative selective HSCT	1		Mean Difference (IV, Random, 95% CI)	Subtotals only

Analysis 7.1. Comparison 7: Efficacy – cardiac function, Outcome 1: Autologous non-myeloablative selective HSCT

Study or Subgroup	HSCT Mean	HSCT SD	HSCT Total	Cyclophosphamide Mean	Cyclophosphamide SD	Cyclophosphamide Total	Mean Difference IV, Random, 95% CI
van Laar 2014 (1)	-2.2	14.7	67	-1.9	13.8	64	-0.30 [-5.18 , 4.58]

Test for subgroup differences: Not applicable



Footnotes

(1) Left ventricular ejection fraction, % by cardiac echocardiography; values are change in the area under the time response curve from baseline to 2 years follow

Comparison 8. Participant-reported outcomes – SF-36

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
8.1 Physical Function	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
8.1.1 Autologous non-myeloablative non-selective HSCT	1	19	Mean Difference (IV, Fixed, 95% CI)	25.00 [-2.55, 52.55]
8.2 Physical Role Limitation	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
8.2.1 Autologous non-myeloablative non-selective HSCT	1	19	Mean Difference (IV, Fixed, 95% CI)	20.00 [-13.14, 53.14]
8.3 Body Pain	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
8.3.1 Autologous non-myeloablative non-selective HSCT	1	19	Mean Difference (IV, Fixed, 95% CI)	27.00 [6.85, 47.15]
8.4 General Health Perception	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
8.4.1 Autologous non-myeloablative non-selective HSCT	1	19	Mean Difference (IV, Fixed, 95% CI)	29.00 [4.43, 53.57]
8.5 Vital Energy Fatigue	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
8.5.1 Autologous non-myeloablative non-selective HSCT	1	19	Mean Difference (IV, Fixed, 95% CI)	11.00 [-10.85, 32.85]

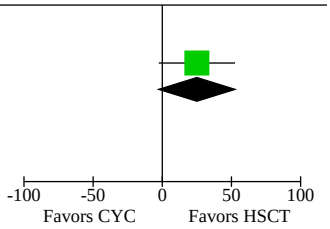
Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
8.6 Social Function	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
8.6.1 Autologous non-myeloablative non-selective HSCT	1	19	Mean Difference (IV, Fixed, 95% CI)	34.00 [7.10, 60.90]
8.7 Emotional Role Limitation	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
8.7.1 Autologous non-myeloablative non-selective HSCT	1	19	Mean Difference (IV, Fixed, 95% CI)	49.00 [8.74, 89.26]
8.8 Mental Health	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
8.8.1 Autologous non-myeloablative non-selective HSCT	1	19	Mean Difference (IV, Fixed, 95% CI)	4.00 [-9.27, 17.27]
8.9 Physical Component Summary	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
8.9.1 Autologous non-myeloablative selective HSCT	1	131	Mean Difference (IV, Fixed, 95% CI)	6.10 [1.43, 10.77]
8.9.2 Autologous non-myeloablative non-selective HSCT	1	19	Mean Difference (IV, Fixed, 95% CI)	26.00 [6.15, 45.85]
8.10 Mental Component Summary	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
8.10.1 Autologous non-myeloablative selective HSCT	1	131	Mean Difference (IV, Fixed, 95% CI)	-0.30 [-5.98, 5.38]
8.10.2 Autologous non-myeloablative non-selective HSCT	1	19	Mean Difference (IV, Fixed, 95% CI)	26.00 [8.80, 43.20]
8.11 PCS SF-36 improvement (≥ 10-point change)	1		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
8.12 MCS SF-36 improvement (≥ 10-point change)	1		Risk Ratio (M-H, Random, 95% CI)	Subtotals only

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 8.1. Comparison 8: Participant-reported outcomes – SF-36, Outcome 1: Physical Function

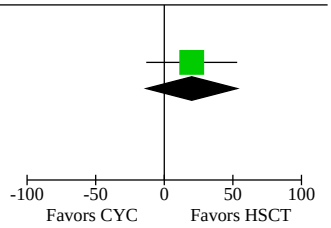
Study or Subgroup	HSCT Mean	HSCT SD	HSCT Total	Cyclophosphamide Mean	Cyclophosphamide SD	Cyclophosphamide Total	Weight	Mean Difference IV, Fixed, 95% CI
8.1.1 Autologous non-myeloablative non-selective HSCT								
Burt 2011 (1)	32	29.62	10	7	31.45	9	100.0%	25.00 [-2.55 , 52.55]
Subtotal (95% CI)			**10**			**9**	**100.0%**	**25.00 [-2.55 , 52.55]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.78 (P = 0.08)								
Test for subgroup differences: Not applicable								



Footnotes
(1) Outcomes is at 1 year

Analysis 8.2. Comparison 8: Participant-reported outcomes – SF-36, Outcome 2: Physical Role Limitation

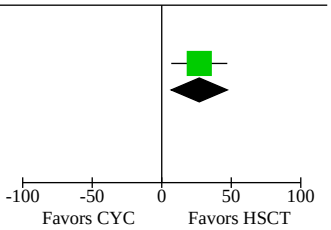
Study or Subgroup	HSCT Mean	HSCT SD	HSCT Total	Cyclophosphamide Mean	Cyclophosphamide SD	Cyclophosphamide Total	Weight	Mean Difference IV, Fixed, 95% CI
8.2.1 Autologous non-myeloablative non-selective HSCT								
Burt 2011 (1)	27	38.9	10	7	34.8	9	100.0%	20.00 [-13.14 , 53.14]
Subtotal (95% CI)			**10**			**9**	**100.0%**	**20.00 [-13.14 , 53.14]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.18 (P = 0.24)								



Footnotes
(1) Outcomes is at 1 year

Analysis 8.3. Comparison 8: Participant-reported outcomes – SF-36, Outcome 3: Body Pain

Study or Subgroup	HSCT Mean	HSCT SD	HSCT Total	Cyclophosphamide Mean	Cyclophosphamide SD	Cyclophosphamide Total	Weight	Mean Difference IV, Fixed, 95% CI
8.3.1 Autologous non-myeloablative non-selective HSCT								
Burt 2011 (1)	21	23	10	-6	21.8	9	100.0%	27.00 [6.85 , 47.15]
Subtotal (95% CI)			**10**			**9**	**100.0%**	**27.00 [6.85 , 47.15]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.63 (P = 0.009)								
Test for subgroup differences: Not applicable								

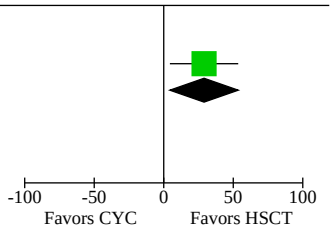


Footnotes
(1) Outcomes is at 1 year

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 8.4. Comparison 8: Participant-reported outcomes – SF-36, Outcome 4: General Health Perception

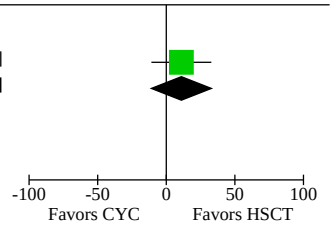
Study or Subgroup	HSCT Mean	HSCT SD	HSCT Total	Cyclophosphamide Mean	Cyclophosphamide SD	Cyclophosphamide Total	Weight	Mean Difference IV, Fixed, 95% CI
8.4.1 Autologous non-myeloablative non-selective HSCT								
Burt 2011 (1)	6	27.38	10	-23	27.2	9	100.0%	29.00 [4.43 , 53.57]
Subtotal (95% CI)			**10**			**9**	**100.0%**	**29.00 [4.43 , 53.57]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.31 (P = 0.02)								
Test for subgroup differences: Not applicable								



Footnotes
(1) Outcomes is at 1 year

Analysis 8.5. Comparison 8: Participant-reported outcomes – SF-36, Outcome 5: Vital Energy Fatigue

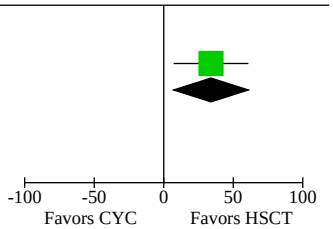
Study or Subgroup	HSCT Mean	HSCT SD	HSCT Total	Cyclophosphamide Mean	Cyclophosphamide SD	Cyclophosphamide Total	Weight	Mean Difference IV, Fixed, 95% CI
8.5.1 Autologous non-myeloablative non-selective HSCT								
Burt 2011 (1)	13	21.5	10	2	26.5	9	100.0%	11.00 [-10.85 , 32.85]
Subtotal (95% CI)			**10**			**9**	**100.0%**	**11.00 [-10.85 , 32.85]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.99 (P = 0.32)								
Test for subgroup differences: Not applicable								



Footnotes
(1) Outcomes is at 1 year

Analysis 8.6. Comparison 8: Participant-reported outcomes – SF-36, Outcome 6: Social Function

Study or Subgroup	HSCT Mean	HSCT SD	HSCT Total	Cyclophosphamide Mean	Cyclophosphamide SD	Cyclophosphamide Total	Weight	Mean Difference IV, Fixed, 95% CI
8.6.1 Autologous non-myeloablative non-selective HSCT								
Burt 2011 (1)	22	31.5	10	-12	28.32	9	100.0%	34.00 [7.10 , 60.90]
Subtotal (95% CI)			**10**			**9**	**100.0%**	**34.00 [7.10 , 60.90]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.48 (P = 0.01)								
Test for subgroup differences: Not applicable								

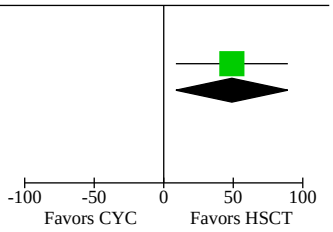


Footnotes
(1) Outcomes is at 1 year

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 8.7. Comparison 8: Participant-reported outcomes – SF-36, Outcome 7: Emotional Role Limitation

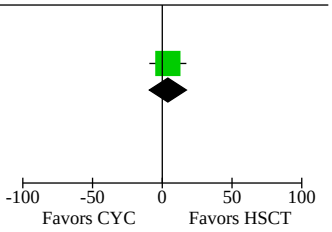
Study or Subgroup	HSCT Mean	HSCT SD	HSCT Total	Cyclophosphamide Mean	Cyclophosphamide SD	Cyclophosphamide Total	Weight	Mean Difference IV, Fixed, 95% CI
8.7.1 Autologous non-myeloablative non-selective HSCT								
Burt 2011 (1)	8	45.6	10	-41	43.88	9	100.0%	49.00 [8.74 , 89.26]
Subtotal (95% CI)			**10**			**9**	**100.0%**	**49.00 [8.74 , 89.26]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.39 (P = 0.02)								
Test for subgroup differences: Not applicable								



Footnotes
(1) Outcomes is at 1 year

Analysis 8.8. Comparison 8: Participant-reported outcomes – SF-36, Outcome 8: Mental Health

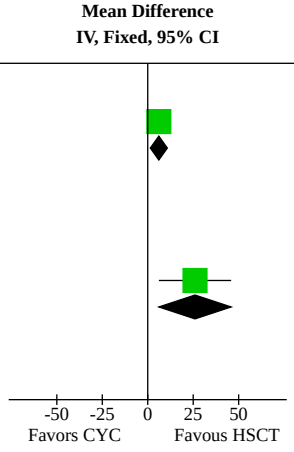
Study or Subgroup	HSCT Mean	HSCT SD	HSCT Total	Cyclophosphamide Mean	Cyclophosphamide SD	Cyclophosphamide Total	Weight	Mean Difference IV, Fixed, 95% CI
8.8.1 Autologous non-myeloablative non-selective HSCT								
Burt 2011 (1)	9	15.7	10	5	13.8	9	100.0%	4.00 [-9.27 , 17.27]
Subtotal (95% CI)			**10**			**9**	**100.0%**	**4.00 [-9.27 , 17.27]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.59 (P = 0.55)								
Test for subgroup differences: Not applicable								



Footnotes
(1) Outcomes is at 1 year

Analysis 8.9. Comparison 8: Participant-reported outcomes – SF-36, Outcome 9: Physical Component Summary

Study or Subgroup	HSCT Mean	HSCT SD	HSCT Total	Cyclophosphamide Mean	Cyclophosphamide SD	Cyclophosphamide Total	Weight	Mean Difference IV, Fixed, 95% CI
8.9.1 Autologous non-myeloablative selective HSCT								
van Laar 2014 (1)	10.1	15.8	67	4	11.2	64	100.0%	6.10 [1.43 , 10.77]
Subtotal (95% CI)			**67**			**64**	**100.0%**	**6.10 [1.43 , 10.77]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.56 (P = 0.01)								
8.9.2 Autologous non-myeloablative non-selective HSCT								
Burt 2011 (2)	20	22.1	10	-6	22	9	100.0%	26.00 [6.15 , 45.85]
Subtotal (95% CI)			**10**			**9**	**100.0%**	**26.00 [6.15 , 45.85]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.57 (P = 0.01)								
Test for subgroup differences: Chi² = 3.66, df = 1 (P = 0.06), I² = 72.7%								

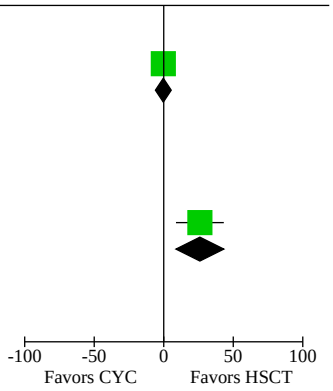


Footnotes
(1) Outcome is at 2 years
(2) Outcome is at 1 year

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 8.10. Comparison 8: Participant-reported outcomes – SF-36, Outcome 10: Mental Component Summary

Study or Subgroup	HSCT Mean	HSCT SD	HSCT Total	Cyclophosphamide Mean	Cyclophosphamide SD	Cyclophosphamide Total	Weight	Mean Difference IV, Fixed, 95% CI
8.10.1 Autologous non-myeloablative selective HSCT								
van Laar 2014 (1)	3.1	16	67	3.4	17.1	64	100.0%	-0.30 [-5.98 , 5.38]
Subtotal (95% CI)			**67**			**64**	**100.0%**	**-0.30 [-5.98 , 5.38]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.10 (P = 0.92)								
8.10.2 Autologous non-myeloablative non-selective HSCT								
Burt 2011 (2)	12	21	10	-14	17.2	9	100.0%	26.00 [8.80 , 43.20]
Subtotal (95% CI)			**10**			**9**	**100.0%**	**26.00 [8.80 , 43.20]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.96 (P = 0.003)								
Test for subgroup differences: Chi^2 = 8.10, df = 1 (P = 0.004), I^2 = 87.7%								



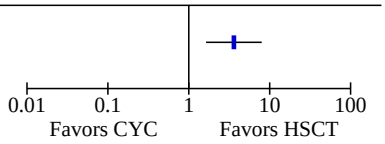
Footnotes

(1) Outcome is at 2 years

(2) Outcome is at 1 year

Analysis 8.11. Comparison 8: Participant-reported outcomes – SF-36, Outcome 11: PCS SF-36 improvement (≥ 10-point change)

Study or Subgroup	HSCT Events	HSCT Total	Cyclophosphamide Events	Cyclophosphamide Total	Risk Ratio M-H, Random, 95% CI
Sullivan 2018 (1)	20	36	6	39	3.61 [1.64 , 7.97]
Test for subgroup differences: Not applicable					

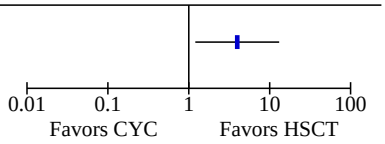


Footnotes

(1) Data presented in two subsets - EFS survivors and EFS failure. EFS survivor assessments were done at 54 months. EFS failure assessments are

Analysis 8.12. Comparison 8: Participant-reported outcomes – SF-36, Outcome 12: MCS SF-36 improvement (≥ 10-point change)

Study or Subgroup	HSCT Events	HSCT Total	Cyclophosphamide Events	Cyclophosphamide Total	Risk Ratio M-H, Random, 95% CI
Sullivan 2018 (1)	11	36	3	39	3.97 [1.20 , 13.10]
Test for subgroup differences: Not applicable					



Footnotes

(1) Data presented in two subsets - EFS survivors and EFS failure. EFS survivor assessments were done at 54 months. EFS failure assessments are

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

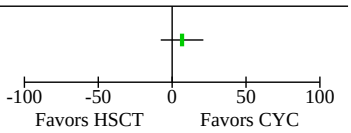
Comparison 9. Participant-reported outcomes – EQ-5D VAS score

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
9.1 Autologous non-myeloablative selective HSCT	1		Mean Difference (IV, Random, 95% CI)	Subtotals only

Analysis 9.1. Comparison 9: Participant-reported outcomes – EQ-5D VAS score, Outcome 1: Autologous non-myeloablative selective HSCT

Study or Subgroup	HSCT Mean	HSCT SD	HSCT Total	Cyclophosphamide Mean	Cyclophosphamide SD	Cyclophosphamide Total	Mean Difference IV, Random, 95% CI
van Laar 2014 (1)	16.9	44.5	67	10.2	39.7	64	6.70 [-7.73 , 21.13]

Test for subgroup differences: Not applicable



Footnotes
(1) Outcome is at 2 years

APPENDICES

Appendix 1. Search strategies

Ovid MEDLINE

1	exp Scleroderma, Systemic/
2	(Systemic adj2 sclerosis).ti,ab.
3	SSc.tw.
4	scleroderma.tw.
5	or/1-4
6	exp Stem Cells/
7	(stem adj3 cell).tw
8	exp Hematopoietic Stem Cell Transplantation/
9	(hematopoietic adj5 transplantation).tw.
10	HSCT.tw.
11	or/6-10
12	5 and 11

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Ovid Embase

1	exp Scleroderma, Systemic/
2	scleroderm*.ti,ab.
3	"systemic scleros*".ti,ab.
4	"crest syndrome*".ti,ab.
5	or/1-4
6	exp Stem Cell Transplantation/
7	(stem cell* adj3 (transplant* or therap*)).ti,ab.
8	HSCT.ti,ab.
9	or/6-8
10	5 and 9
11	HUMANS/
12	NONHUMAN/ or ANIMAL/ or ANIMAL EXPERIMENT/
13	12 not 11
14	10 not 13

Clinical Trials.gov

1. Systemic Sclerosis OR Scleroderma | Stem Cell Transplant OR HSCT OR stem cell therapy

CENTRAL

1. Stem Cell Transplant*:ti,ab,kw
2. Stem Cell Therap*
3. HSCT
4. MeSH descriptor Stem Cell Transplantation explode all trees
5. #1 or #2 or #3 or #4
6. Scleroderm*
7. Systemic Scleros*
8. CREST syndrome
9. MeSH descriptor: [Scleroderma, Systemic] explode all trees
10. #6 or #7 or #8 or #9
11. #5 and #10

HISTORY

Protocol first published: Issue 9, 2015

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Date	Event	Description
30 July 2014	Amended	CMSG ID C280

CONTRIBUTIONS OF AUTHORS

Draft protocol – SB, MLO, DAM, MDM.

Search strategy – SB, DAM, Pratt G (Research Medical Library, The University of Texas MD Anderson Cancer Center), KK.

Identify relevant titles and abstracts from searches – SB, HRS, MLO, MSA

Obtain copies of trials – SB, HRS, MLO

Selection of trials – SB, HRS, MLO, MSA

Extract data from trials – SB, HRS, MLO

Enter data into Review Manager – SB, HRS, MLO

Assessed risk of bias – SB, HRS, MLO, MSA

Conducted analyses – SB, MLO, MSA.

Interpret data – SB, HRS, MLO, DAM, MDM, MSA

Assessed the certainty of the evidence – SB, MLO, MSA

Draft final review – SB, HRS, MAL, DAM, MDM, Suarez-Almazor ME.

Update review – SB, MLO

DECLARATIONS OF INTEREST

SB: none.

HS: none.

DAM: none.

MDM declares the following: Galapagos NV (Independent Contractor - Consultant), Actelion Pharmaceuticals US, Inc. (Independent Contractor - Consultant), Eicos Sciences (Independent Contractor - Consultant), Boehringer Ingelheim (Independent Contractor - Consultant)

KK: none.

MAL: none.

SOURCES OF SUPPORT

Internal sources

- The University of Texas, MD Anderson Cancer Center, USA

  Office supplies, interlibrary loan, etc.

External sources

- None, Other

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We added more descriptions of individual outcomes to the 'Methods' section.

Kate Krause joined the review team.

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.