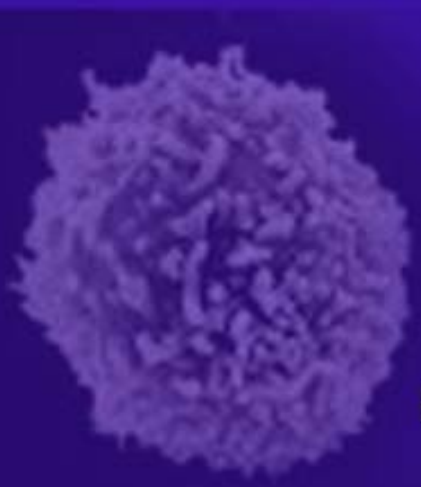
UNIVERSITAT AUTÒNOMA DE BARCELONA
Facultad de Medicina
Departamento de Pediatría, Obstetricia y Ginecología,
y de Medicina Preventiva

Agentes Biológicos en el Tratamiento de Enfermedades Hematológicas Malignas: Revisiones Sistemáticas

TESIS DOCTORAL

Arturo José Martí Carvajal



Director de Tesis
Dr. Xavier Bonfill Cosp

Barcelona, diciembre de 2015

Resumen de los Resultados

Segunda Publicación

Martí-Carvajal AJ, Anand V, Solà I. Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis. Cochrane Database of Systematic Reviews 2015, Issue 4. Art. No.: CD010298. DOI: 10.1002/14651858.CD010298.pub2.

Factor de impacto (2014): 6,032

Principales resultados: se identificaron dos ensayos que aleatorizaron a 528 participantes, comparando ruxolitinib con placebo o mejor terapia disponible (BAT, por sus siglas en inglés). Como los dos ensayos incluidos tuvieron diferentes comparadores no hubo metanálisis. La confianza en los resultados de las estimaciones de estos ensayos fue baja debido al sesgo en su diseño, y sus tamaños de muestra limitados que originó resultados imprecisos.

Existen evidencias de baja calidad para el efecto de ruxolitinib sobre la supervivencia en comparación con el placebo, a las 51 semanas de seguimiento (HR 0,51; IC de 95 % entre 0,27 y 0,98), y en comparación con la BAT a las 48 semanas de seguimiento (HR 0,70; IC 95 % entre 0,20 y 2,47). Del mismo modo, la evidencia es de muy baja calidad para el efecto del ruxolitinib sobre la supervivencia libre de progresión en comparación con la BAT (HR 0,81; IC del 95 % entre 0,47 y 1,39).

Existen evidencias de baja calidad para el efecto del ruxolitinib en términos de calidad de vida. En comparación con el placebo, el fármaco logra una mayor proporción de pacientes con una reducción significativa de las puntuaciones de síntomas (RR 8,82, IC del 95 % entre 4,40 y 17,69), y los pacientes tratados con ruxolitinib obtuvieron mayores puntuaciones del MFSAF (por sus siglas en inglés) al final del seguimiento (DM – 87,90; IC 95 % CI entre -139,58 y -36,22). Un ensayo mostró diferencias significativas en las puntuaciones de la EORTC QLQ-C30 comparando ruxolitinib con BAT (DM 7,60; IC del 95 % entre 0,35 y 14,85).

El efecto del ruxolitinib en la reducción en el tamaño del bazo de los participantes en comparación con placebo o BAT fue incierto y con evidencia de baja calidad (contra el placebo: RR 64,58; IC del 95 %: 9,08 a 459,56; frente a BAT: RR 41,78; IC 95 %: 2,61 a 669,75).

Existen evidencias de baja calidad para el efecto del fármaco en comparación con el placebo sobre la anemia (RR 2,35, IC 95 % entre 1,62 y 3,41), neutropenia (RR 3,57, IC del 95 % entre 1,02 y 12,55) y trombocitopenia (RR 9,74, IC del 95 % entre 2,32 y 40,96). No hubo diferencias entre el ruxolitinib y BAT, en términos de riesgo de anemia (RR 1,35; IC del 95 %: 0,91 a 1,99; evidencia de baja calidad) o trombocitopenia (RR 1,20; IC del 95 %: 0,44 a 3,28, evidencia de baja calidad). El riesgo de eventos adversos no hematológicos de grados 3 o 4 (incluyendo fatiga, artralgias, náuseas, diarrea, dolor en las extremidades y pirexia) fue similar cuando ruxolitinib se comparó con placebo o BAT. La tasa de neutropenia, comparando ruxolitinib con BAT, no se informó en el ensayo respectivo.

Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis (Review)

Martí-Carvajal AJ, Anand V, Solà I

This is a reprint of a Cochrane review, prepared and maintained by The Cochrane Collaboration and published in *The Cochrane Library* 2015, Issue 4

http://www.thecochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis

Arturo J Martí-Carvajal[1], Vidhu Anand[2], Ivan Solà[3]

[1]Iberoamerican Cochrane Network, Valencia, Venezuela. [2]Department of Medicine, University of Minnesota, Minneapolis, MN, USA. [3]Iberoamerican Cochrane Centre, Biomedical Research Institute Sant Pau (IIB Sant Pau), CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain

Contact address: Arturo J Martí-Carvajal, Iberoamerican Cochrane Network, Valencia, Venezuela. arturo.marti.carvajal@gmail.com.

Editorial group: Cochrane Haematological Malignancies Group.
Publication status and date: New, published in Issue 4, 2015.
Review content assessed as up-to-date: 20 November 2014.

Citation: Martí-Carvajal AJ, Anand V, Solà I. Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis. *Cochrane Database of Systematic Reviews* 2015, Issue 4. Art. No.: CD010298. DOI: 10.1002/14651858.CD010298.pub2.

ABSTRACT

Background

Myelofibrosis is a bone marrow disorder characterized by excessive production of reticulin and collagen fiber deposition caused by hematological and non-hematological disorders. The prognosis of myelofibrosis is poor and treatment is mainly palliative. Janus kinase inhibitors are a novel strategy to treat people with myelofibrosis.

Objectives

To determine the clinical benefits and harms of Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis secondary to hematological or non-hematological conditions.

Search methods

We searched the Cochrane Central Register of Controlled Trials (CENTRAL, the Cochrane Library 2014, Issue 11), Ovid MEDLINE (from 1946 to 13 November 2014), EMBASE (from 1980 to 12 January 2013), and LILACS (from 1982 to 20 November 2014). We searched WHO International Clinical Trials Registry Platform and The metaRegister of Controlled Trials. We also searched for conference proceedings of the American Society of Hematology (from 2009 to October 2013), European Hematology Association (from 2009 to October 2013), American Society of Clinical Oncology (from 2009 to October 2013), and European Society of Medical Oncology (from 2009 to October 2013). We included searches in FDA, European Medicines Agency, and Epistemonikos. We handsearched the references of all identified included trials, and relevant review articles. We did not apply any language restrictions. Two review authors independently screened search results.

Selection criteria

We included randomized clinical trials comparing Janus kinase-1 and Janus kinase-2 inhibitors with placebo or other treatments. Both previously treated and treatment naive patients were included.

Data collection and analysis

We used the hazard ratio (HR) and 95% confidence interval (95% CI) for overall survival, progression-free survival and leukemia-free survival, risk ratio (RR) and 95% CI for reduction in spleen size and adverse events binary data, and standardized mean differences (SMD) and 95% CI for continuous data (health-related quality of life). Two review authors independently extracted data and assessed the risk of bias of included trials. Primary outcomes were overall survival, progression-free survival and adverse events.

Main results

We included two trials involving 528 participants, comparing ruxolitinib with placebo or best available therapy (BAT). As the two included trials had different comparators we did not pool the data. The confidence in the results estimates of these trials was low due to the bias in their design, and their limited sample sizes that resulted in imprecise results.

There is low quality evidence for the effect of ruxolitinib on survival when compared with placebo at 51 weeks of follow-up (HR 0.51, 95% CI 0.27 to 0.98) and compared with BAT at 48 weeks of follow-up (HR 0.70, 95% CI 0.20 to 2.47). Similarly there was very low quality evidence for the effect of ruxolitinib on progression free survival compared with BAT (HR 0.81, 95% CI 0.47 to 1.39).

There is low quality evidence for the effect of ruxolitinib in terms of quality of life. Compared with placebo, the drug achieved a greater proportion of patients with a significant reduction of symptom scores (RR 8.82, 95% CI 4.40 to 17.69), and treated patients with ruxolitinib obtained greater MFSAF scores at the end of follow-up (MD -87.90, 95% CI -139.58 to -36.22). An additional trial showed significant differences in EORTC QLQ-C30 scores when compared ruxolitinib with best available therapy (MD 7.60, 95% CI 0.35 to 14.85).

The effect of ruxolitinib on reduction in the spleen size of participants compared with placebo or BAT was uncertain (versus placebo: RR 64.58, 95% CI 9.08 to 459.56, *low quality evidence*; versus BAT: RR 41.78, 95% CI 2.61 to 669.75, *low quality evidence*).

There is low quality evidence for the effect of the drug compared with placebo on anemia (RR 2.35, 95% CI 1.62 to 3.41), neutropenia (RR 3.57, 95% CI 1.02 to 12.55) and thrombocytopenia (RR 9.74, 95% CI 2.32 to 40.96). Ruxolitinib did not result in differences versus BAT in the risk of anemia (RR 1.35, 95% CI 0.91 to 1.99, *low quality evidence*) or thrombocytopenia (RR 1.20; 95% CI 0.44 to 3.28, *low quality evidence*). The risk of non-hematologic grade 3 or 4 adverse events (including fatigue, arthralgia, nausea, diarrhea, extremity pain and pyrexia) was similar when ruxolitinib was compared with placebo or BAT. The rate of neutropenia comparing ruxolitinib with standard medical treatment was not reported by the trial.

Authors' conclusions

Currently, there is insufficient evidence to allow any conclusions regarding the efficacy and safety of ruxolitinib for treating myelofibrosis. The findings of this Cochrane review should be interpreted with caution as they are based on trials sponsored by industry, and include a small number of patients. Unless powered randomized clinical trials provide strong evidence of a treatment effect, and the trade-off between potential benefits and harms is established, clinicians should be cautious when administering ruxolitinib for treating patients with myelofibrosis.

PLAIN LANGUAGE SUMMARY

Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis

Review question

We reviewed the effects of Janus kinase-1 and Janus kinase-2 inhibitors for treating people with myelofibrosis.

Background

Myelofibrosis is a disorder of the bone marrow in which the bone marrow is replaced by fibrous tissue. The symptoms depend on the degree of anemia and enlargement of the spleen. This condition has a poor prognosis and generally its treatment is palliative.

Ruxolitinib is a drug in the class of Janus kinase inhibitors that tries to block the enzyme that derives in the scar tissue.

Study characteristics

We identified two clinical trials that included a limited number of patients comparing ruxolitinib to placebo or standard medical treatment. Both studies were published in 2012, and were conducted in the United States of America (USA) and the United Kingdom (UK). Drug companies sponsored both trials.

Key results

Although the results of the studies only showed a moderate improvement of patients treated with ruxolitinib in terms of their quality of life and a reduction in their spleen size, we could not be sure whether these effects were reliable because of the limitations of the studies and the low number of people they recruited. We also could not be sure whether the drug has an effect on overall survival compared

with a placebo, or when it was compared with an active treatment. The effect of ruxolitinib in terms of progression-free survival was also uncertain. In addition, people treated with this drug showed higher rates of anemia, thrombocytopenia and neutropenia compared with patients treated with a placebo, but the rate of adverse effects was similar to those treated with a medical treatment.

Quality of evidence

The confidence in the results of this review is very low. The studies have limitations in the way they were designed and executed. Moreover, the limited number of patients included in the studies led to imprecise results. Larger studies should provide more information about the effect of ruxolitinib.

Researchers from Cochrane searched all available literature up to 13 November 2014.

SUMMARY OF FINDINGS FOR THE MAIN COMPARISON *[Explanation]*

Ruxolitinib compared with placebo for treating MF

Patient or population: Patients with treating MF
Settings: Ambulatory
Intervention: Ruxolitinib
Comparison: Placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (trials)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Placebo	Ruxolitinib				
Overall survival (number of deaths at follow-up (24 weeks[1]))	**91 per 1000**[2]	**47 per 1000** (25 to 89)[3]	**HR 0.51** (0.27 to 0.98)[4]	309 (1 trial[5])	⊕⊕◯◯ **low**[6,7]	-
Progression-free survival - not measured	See comment	See comment	Not estimable	309 (1 trial[5])	See comment	This outcome was not measured in the included study
Safety (AE, adverse drug reaction): thrombocytopenia Grades 3 or 4 according to National Cancer Institute Follow-up: 52 weeks	**13 per 1000**	**129 per 1000** (31 to 543)	**RR 9.74** (2.32 to 40.96)	306 (1 trial[5])	⊕⊕◯◯ **low**[6,7]	-
Safety (AE, adverse drug reaction): neutropenia Grades 3 or 4 according to National Cancer Institute Follow-up: 52 weeks	**20 per 1000**	**71 per 1000** (20 to 249)	**RR 3.57** (1.02 to 12.55)	306 (1 trial[5])	⊕⊕◯◯ **low**[6,7]	-

Health-related quality of life Patients that achieved a reduction of 50% or more in the total MF Symptom Assessment Form Follow-up: 24 weeks	**52 per 1000**	**458 per 1000** (229 to 919)	**RR 8.82** (4.40 to 17.69)	309 (1 trial[5])	⊕⊕◯◯ **low**[6,7]	-
Reduction in spleen size Magnetic resonance imaging or computed tomography Follow-up: 48 weeks	**6 per 1000**	**419 per 1000** (59 to 1000)	**RR 64.58** (9.08 to 459.56)	309 (1 trial[5])	⊕⊕◯◯ **low**[6,7]	-

*The basis for the **assumed risk** is provided in footnote #2. The **corresponding risk** (and its 95% CI) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: confidence interval; **RR:** risk ratio; **HR:** hazard ratio. **AE:** adverse event; **MF:** myelofibrosis.

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] COMFORT-I 2012 trial established the time of data cut-off for its main outcome at 24 weeks.
[2] Data obtained from deaths in the placebo group at the time of data cut-off (24 weeks).
[3] Data obtained from deaths in the intervention group at the time of data cut-off (24 weeks).
[4] See Analysis 1.1.
[5] COMFORT-I 2012 trial.
[6] Downgraded one level due to limitations in the trial design or execution (high attrition bias).
[7] Downgraded one level due to imprecision (low sample and number of events with an impact in the precision of the effect estimates).

BACKGROUND

Description of the condition

Myelofibrosis (MF) is a bone marrow disorder characterized by excessive production of reticulin and collagen fibers (Ostojic 2012). It implies an increase in the bone marrow fiber content without referring to quantity or quality (reticulin versus collagen) (Thiele 2007). MF can be the outcome of several hematological conditions (i.e., as evolution of a previously known myeloproliferative neoplasm, chronic myeloid leukemia, polycythemia vera, or essential thrombocythemia) (Cervantes 2011; Hoffman 2008) and non-hematological conditions (metastatic cancer, infections such as tuberculosis, fungal infections and HIV, metabolic disorders, radiation, toxins, etc.) (Hoffman 2008).

Primary myelofibrosis (PMF) is a chronic, malignant hematological disorder characterized by splenomegaly, leukoerythroblastosis, teardrop poikilocytosis (i.e., dacryocytes), some degree of marrow fibrosis, increased marrow microvessel density, and extramedullary hematopoiesis (Hoffman 2008). PMF is associated with osteosclerosis, angiogenesis, and an abnormal cytokine expression (Tefferi 2011b).

PMF is an infrequent disease, with an estimated incidence in Western countries that ranges from 0.4 to 1.4 new cases per 100,000 people/year (Barosi 2011b). The average age at diagnosis of PMF is approximately 65 years, and most patients are diagnosed between 50 and 69 years of age (Hoffman 2008). In several case series, men have been affected more frequently than women, but other trials have failed to confirm this male predominance (Hoffman 2008). PMF has rarely been reported in the pediatric age group (Hoffman 2008).

The clinical features of PMF include severe anemia, marked hepatosplenomegaly, constitutional symptoms (e.g., fatigue, night sweats, fever), cachexia, bone pain, splenic infarct, pruritus, thrombosis, and bleeding (Tefferi 2011b). The main causes of the anemia and organomegaly are ineffective erythropoiesis and extramedullary hematopoiesis, respectively (Tefferi 2011b). Other disease complications include symptomatic portal hypertension, which may lead to variceal bleeding or ascites, and non-hepatosplenic extramedullary hematopoiesis, which may lead to cord compression, ascites, pleural effusion, pulmonary hypertension, or diffuse extremity pain (Tefferi 2011b). These other complications are caused by aberrant cytokine production by clonal cells and host immune reaction contributing to PMF-associated bone marrow stromal changes, ineffective erythropoiesis, extramedullary hematopoiesis, cachexia, and constitutional symptoms (Tefferi 2011b). PMF is associated with cytogenetic abnormalities such as deletion of the long arm of chromosome 20 (20q-), deletion of chromosome 13q (13q), trisomy 8 and 9, and abnormalities of chromosome 1 including duplication 1q (Hussein 2009).

Current diagnosis of MF is based on the World Health Organization (WHO) criteria and involves a composite assessment of clinical and laboratory features (Tefferi 2011a). These criteria include major criteria (megakaryocyte proliferation and atypia accompanied by either reticulin, or collagen fibrosis, or both, or not meeting WHO criteria for chronic myelogenous leukemia, polycythemia vera, myelodysplastic syndromes), or other myeloid neoplasm, and demonstration of Janus kinase-2 (guanine-to-thymidine substitution, which results in a change of valine for phenylalanine at codon 617), or a myeloproliferative leukemia virus oncogene mutation, occurring in 60% and 5% to 10% of the patients, respectively, and other myeloproliferative neoplasm-associated molecular abnormalities (i.e. CBL, ASXL1, TET2, and EZH2), or clonal markers (particularly trisomy 9 or 13q−) that distinguish PMF from reactive marrow fibrosis (Barosi 2011b; Wen 2011). Additionally, the presence of minor criteria (leukoerythroblastosis, increased serum lactate dehydrogenase level, anemia, and palpable splenomegaly) (Tefferi 2011a). PMF diagnosis requires meeting all three major criteria and two minor criteria (Tefferi 2011a). Appendix 1 shows the International Working Group for Myeloproliferative Neoplasms Research and Treatment Recommended Criteria for Post-Polycythemia Vera and Post-Essential Thrombocythemia Myelofibrosis (Barosi 2008; Tefferi 2011a).

European Consensus Criteria for grading of MF is based on subjective evaluation of amount and distribution of reticulin and collagen in bone marrow. It is as follows: *MF-0* (scattered linear reticulin with no intersections (crossovers) corresponding to normal bone marrow), MF-*1* (loose network of reticulin with many intersections, especially in perivascular areas), MF-*2* (diffuse and dense increase in reticulin with extensive intersections, occasionally with focal bundles of collagen, or focal osteosclerosis, or both) and MF-*3* (diffuse and dense increase in reticulin with extensive intersections and coarse bundles of collagen, often associated with osteosclerosis) (Hoffman 2008; Thiele 2005; Thiele 2007).

PMF is associated with a low quality of life (Mesa 2009b). The prognosis at the time of diagnosis is based on the International Prognostic Scoring System developed by International Working Group for Myeloproliferative Neaplasms Research and Treatment (Cervantes 2009). The International Prognostic Scoring System includes age (> 65 years), constitutional symptoms, hemoglobin (< 10 g/dL), white blood cell count (> 25 x 10^9/L), and blood blasts (≥ 1%) (Passamonti 2010). International Working Group for Myeloproliferative Neoplasms Research and Treatment developed the Dynamic Prognostic Model with the same prognostic variables generated by International Prognostic Scoring System which can be applied at any time during the disease course (Passamonti 2010). PMF progress to leukemia in 20% of patients, while others die because of comorbid conditions, including cardiovascular events (Barbui 2010), infection, or bleeding (Tefferi 2011b). The median overall survival is 11.3 years for low risk, 7.9 years for intermediate-1 risk, 4.0 years for intermediate-2 risk, and 2.3 years for high-risk MF (Cervantes 2009).

In this Cochrane Review we included MF as a result of both hematological and non-hematological conditions. We have provided a

glossary of medical terms in Appendix 2.

Description of the intervention

Allogeneic stem cell transplantation is the only curative option for patients with PMF who have an appropriate donor available (Hoffman 2008; Ostojic 2011b). However, allogeneic stem cell transplantation is a reasonable option for only a small percentage of eligible patients (i.e., those who are young and unburdened by other co-morbidities) (Ostojic 2011b). A conservative approach is generally accepted with observation of asymptomatic patients and therapeutic intervention for those who have symptoms (Hoffman 2008). Current treatment regimens are mainly palliative and have not demonstrated a major benefit in overall survival (Ostojic 2011b).

Therapy for treating anemia

1. Androgens (nandrolone, fluoxymesterone, methandrostenolone, oxymetholone, methenolone, and danazol) stimulate hematopoietic system by various mechanisms including stimulation of erythropoietin release, increasing bone marrow activity and iron incorporation into the red cells (Shahani 2009). It has been reported to improve the anemia in patients with MF in 30 to 60% of cases (Barosi 2011a).
2. Recombinant human erythropoietin and darbepoetin alfa are growth factors with similar mechanisms of action as erythropoietin (Donnelly 2001). Darbepoetin alfa is an analog of recombinant human erythropoietin with a long half-life that requires less frequent administration (once weekly or every other week) (Cases 2003). The response rates of these drugs ranged from 16% to 60% (Barosi 2011b). However, there is an unexpected association between erythropoietin-stimulating agents and danazol with leukemic transformation in MF (Barosi 2011b).
3. Thalidomide and its analogs, lenalidomide and pomalidomide, have anti-angiogenetic and immunomodulatory activities and have been used previously for MF (Barosi 2011b; Tefferi 2009).

Splenomegaly and myeloproliferation treatment

Overall, this approach decreases the immature circulating myeloid pool accumulating in the spleen. The following interventions have been previously described for these purposes:

1. Hydroxyurea limits the deoxyribonucleic acid biosynthesis. The studies using hydroxycarbamide have reported a response on splenomegaly in up to 40% of treated cause (Barosi 2011b).
2. Chlorambucil, 6-thioguanine, melphalan, and busulfan are oral alkylating agents. Use of the last two drugs is limited by the increased risk of blast transformation and unfavorable toxicity profile (Barosi 2011b).
3. Interferon biologic response modifier inhibits in vitro proliferation of hematopoietic progenitors, particularly of the megakaryocytic lineage (Barosi 2011b). It can be useful in suppressing thrombocytosis and inhibiting the activity of platelet derived growth factor which stimulates the proliferation of fibroblasts (Hoffman 2008).
4. Thalidomide analogs are immunomodulator drugs. Two studies conducted for assessing lenalidomide in patients with MF have reported a reduction of spleen size ranging between 10% and 42% (Mesa 2010b; Quintás-Cardama 2009). Two studies on pomalidomide showed a poor response on spleen size (Begna 2011; Mesa 2010a).
5. Janus kinase inhibitors have been reported for treating MF (Barosi 2011b; Geyer 2014; Mesa 2012a; Randhawa 2012). In November 2011, the U.S. Food and Drug Administration (FDA) approved the use of ruxolitinib, a JAK-1- and JAK-2-selective inhibitor, for the treatment of patients with intermediate or high-risk MF (Deisseroth 2012; Mascarenhas 2012; Randhawa 2012). Two randomized controlled trials (RCTs) have been conducted for assessing the efficacy and safety of this drug in patients with MF (COMFORT-I 2012; COMFORT-II 2012). These RCTs have emphasized on surrogate end point (reduction in spleen volume) and quality of life. Both RCTs have shown significant reduction in spleen volume and enhancement in quality of life. Other JAK inhibitors, such as SAR302503, CYT387, SB1518, and TG101348 may become commercially available in the near future (Geyer 2014; Mesa 2012a).

How the intervention might work

The Janus family includes a cytoplasmic tyrosine kinases (JAK-1, JAK-2, JAK-3, and TYK2) which mediates the signaling of a number of cytokines and growth factors that are important for hematopoiesis and immune function (Pastore 2012; Seavey 2012; Stein 2011; Thompson 2005). JAK-1 plays a major role in the signaling of a number of pro-inflammatory cytokines; JAK-2 is used primarily by receptors for hematopoietic growth factors; JAK-3 primarily mediates immune function, whereas TYK2 functions in association with JAK-2 or JAK-3 to transduce signaling of cytokines such as interleukin 12 (Barosi 2011b; Pastore 2012; Seavey 2012; Stein 2011; Thompson 2005).

Several reviews on JAK inhibitor therapy for MF have been published (Ostojic 2011b; Ostojic 2011a; Pardanani 2011a; Pardanani 2011b; Passamonti 2012; Stein 2011; Tefferi 2011e; Tefferi 2012). Ruxolitinib, which modulates the abnormal cytokine production and signaling, plays a major role in pathogenesis of MF (Vannucchi 2009). Nevertheless, the clinical effect of JAK-2-inhibitors seen in people with MF seems to reflect the effect of the drug over normal hematopoiesis with an unmutated JAK-2-allele rather than on the MPN-clone (myeloproliferative neoplasm) with the mutated JAK-2-allele (Mesa 2009a). Ruxolitinib was initially used in a phase 1/2 trial including 153 patients with MF (Barosi 2011b; Tefferi

2011b; Ostojic 2012). In this trial, treatment was well tolerated, with dose-limiting toxicity represented by reversible thrombocytopenia. Ruxolitinib has shown significant clinical response with a ≥ 50% reduction of splenomegaly in half of the patients and rapid improvement of constitutional symptoms, cachexia, and exercise tolerance (Barosi 2011b; Tefferi 2011b; Ostojic 2012). After a single oral dose, > 95% of the ruxolitinib is absorbed, and > 97% becomes available bounding to plasma proteins. The terminal half-life is two to three hours (Ostojic 2012). Ruxolitinib is metabolized predominantly in the liver, and its metabolites are mainly excreted in urine (Ostojic 2012). Its adverse events include thrombocytopenia, anemia, and a 'cytokine rebound reaction' upon drug discontinuation, characterized by acute relapse of symptoms and splenomegaly (Barosi 2011b; Tefferi 2011c; Tefferi 2011d).

Why it is important to do this review

We conducted this Cochrane Review for several reasons. The primary goals of the therapy for MF are to alleviate the symptoms and to achieve an improvement in the patients' quality of life, but it lacks any real impact on overall survival and progression-free survival (Barosi 2011b; Qureshi 2011). Controversy exists if the current trial endpoints capture a tangible benefit for MF patients (Pardanani 2012). Therefore, we need to perform a critical appraisal of the RCTs conducted to assess ruxolitinib in patients with MF (COMFORT-I 2012; COMFORT-II 2012). Furthermore, this drug has been associated with serious adverse events (anemia and thrombocytopenia) (Tefferi 2011c; Tefferi 2011d). Ruxolitinib is expensive and costs USD7,000 per month of treatment, or USD84,000 per year, for the insured patient (Mesa 2012b).
In this Cochrane Review we have included a rigorous assessment of the risk of bias, using most up-to-date evidence to help clinicians making informed decisions on the use of Janus kinase-1 and Janus kinase-2 inhibitors for treating patients with MF due to hematological or non-hematological conditions.

OBJECTIVES

To determine the clinical benefits and harms of Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis secondary to hematological or non-hematological conditions.

METHODS

Criteria for considering studies for this review

Types of studies

We included RCTs irrespective of their publication status (unpublished or published as an article, an abstract, or a letter) and language. No limits were applied with respect to period of follow-up. We excluded quasi-RCTs.

Types of participants

We included patients with a confirmed diagnosis of MF caused by hematological and non-hematological conditions, irrespective of their age, gender, or ethnicity.

Types of interventions

Intervention

We compared ruxolitinib with placebo or best available therapy in this review. In future updates, we will also include trials assessing the following JAK-1 and 2 inhibitors:

- SAR302503.
- CYT387.
- SB1518.
- TG101348.

Comparisons

- Placebo.
- Other treatments.
- Head-to-head comparisons of JAK inhibitors.

Types of outcome measures

Primary outcomes

1. Overall survival: the time from randomization until death from any cause (FDA 2007).
2. Progression-free survival: the time from randomization until objective tumor progression or death (FDA 2007).
3. Safety:
    - Adverse event: "any untoward medical occurrence that may present during treatment with a pharmaceutical product but which does not necessarily have a causal relationship with this treatment" (Nebeker 2004).
    - Adverse drug reaction: "a response to a drug which is noxious and uninitiated and which occurs at doses normally used in man for prophylaxis, diagnosis, or therapy of disease, or for the modification of physiologic functions" (Nebeker 2004).

Secondary outcomes

1. Health-related quality of life assessed by MF Symptom Assessment Form (MFSAF) (Mesa 2009b) or any other validated scale.
2. Leukemia-free survival.
3. Reduction in spleen size
4. Anemia response defined as an increasing of ≥ 1 g/L at the end of follow-up.

Search methods for identification of studies

We developed the search strategy as indicated in the*Cochrane Handbook for Systematic Reviews of Interventions* (Lefevbre 2011). We conducted this process with the support of the Cochrane Haematological Malignancies Group Trials Search Co-ordinator and adjusted it for each database we searched.

Electronic searches

We searched the following electronic databases:

- Cochrane Central Register of Controlled Trials (CENTRAL, the Cochrane Library 2014, Issue 11).
- MEDLINE (Ovid) and Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations (from 1946 to 13 November 2014).
- EMBASE (OVID) (from 1980 to 12 January 2013).
- LILACS (from 1982 to 20 November 2014).

See Appendix 3; Appendix 4; Appendix 5; Appendix 6 for details.

Searching other resources

We searched the following trial databases for ongoing and unpublished trials:

1. WHO International Clinical Trials Registry Platform (WHO ICTRP) search portal (http://apps.who.int/trialsearch/).
2. The *metaR*egister of Controlled Trials (*m*RCT) (http://www.controlled-trials.com/mrct/) (Appendix 7).

We also searched conference proceedings:

1. American Society of Hematology (ASH) (http://www.hematology.org/) (from 2009 to October 2013).
2. European Hematology Association (EHA) (http://www.ehaweb.org/) (from 2009 to October 2013).
3. American Society of Clinical Oncology (ASCO) (http://www.asco.org/) (from 2009 to October 2013).
4. European Society of Medical Oncology (ESMO) (http://www.esmo.org/) (from 2009 to October 2013).

We also searched the following websites:

1. FDA (http://www.fda.gov/).
2. European Medicines Agency (http://www.ema.europa.eu/ema/).
3. http://www.epistemonikos.org/ (Appendix 8).

We handsearched the references of all identified included trials, relevant review articles, and current treatment guidelines. We did not apply any language restrictions. We used codes of pharmaceutical companies such as INCB018424, SAR302503, CYT387, SB1518, and TG101348 especially in abstract and trial register searches to identify closed or stopped studies, and brand names in search only if available.

Data collection and analysis

We summarized data using standard Cochrane methodologies (Higgins 2011d).

Selection of studies

Regarding methods for study selection, we followed the steps delineated by the*Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011b).

Two authors (AMC and VA) screened the titles and abstracts identified from the above sources to identify potential studies for inclusion. If this could not be done satisfactorily from the title and abstract, a full text version was sought for assessment. We presented the results of the study selection as a flowchart according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement (Moher 2009).

We resolved any disagreement through discussion and consensus. We also contacted the authors of the trials to resolve any doubts about available information or in case of disagreements.

Data extraction and management

We extracted data adequately by collecting the following items: review, reviewer and study information, eligibility criteria, characteristics of participants (age, gender, country), trial design and funding, intervention duration and dosages, and outcomes. We assessed quality criteria according to risk of bias using the Cochrane's 'Risk of bias' assessment tool: random sequence generation; allocation concealment; blinding of participants, personnel, and outcome assessors; incomplete outcome data; selective reporting; and other bias (Higgins 2011b).

For eligible trials, two review authors (AMC and VA) independently extracted the data using the agreed form. We resolved discrepancies through discussion. One review author (AMC) entered data into Cochrane's statistical software, RevMan 2014 and two review authors (VA and IS) independently checked it for accuracy. We also contacted the corresponding trial authors to provide further details.

Assessment of risk of bias in included studies

Three review authors (AMC, VA and IS) independently assessed the risk of bias in pairs of each trial using a simple form, and followed the domain-based evaluation as described in the *Cochrane*

Handbook for Systematic Reviews of Interventions (Higgins 2011a). We resolved any discrepancies through discussion.

We assessed the following domains as at low, unclear, or high risk of bias:

1. Generation of allocation sequence.
2. Allocation concealment.
3. Blinding (of participants, personnel, and outcome assessors).
4. Incomplete outcome data.
5. Selective reporting.
6. Other sources.

Overall risk of bias

We made explicit judgements about whether trials were at low, unclear, or high risk of bias, according to the criteria given in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011d). With reference to (1) to (6) above, we assessed the likely magnitude and direction of the bias and whether we considered it likely to have impact on the findings. As it is unlikely to find trials at low risk of bias in all items, we chose three core domains instead of all: generation of allocation sequence, incomplete outcome data, and selective reporting bias in order to classify a trial as at low, unclear, or high risk of bias. We would have conducted a sensitivity analyses for exploring the impact of the level of bias, if feasible (see Sensitivity analysis).

Measures of treatment effect

1. For the time-to-event data, such as overall survival, progression-free survival, leukemia-free survival, we calculated hazard ratios (HRs) and 95% confidence intervals (95% CIs). We determined HRs for published data according to CMA 2005.
2. For the binary outcomes, such as safety and spleen size reduction (≥ 35%), we calculated the relative risk (RR) with 95% CIs.
3. For the continuous outcomes, such as spleen size, we calculated the mean difference (MD) with 95% CIs.

Dealing with missing data

COMFORT-I 2012 assessed reduction in spleen size as a continuous variable using 79% of the original participants. However, the trial authors used all participants when they measured reduction in spleen size as a binary variable. We reported results using both approaches. We contacted the corresponding trial author for the missing continuous data on reduction in spleen size.

In future updates, in case of missing data on participants or missing statistics (such as standard deviations), we will contact the trial authors. If unsuccessful, we will base our main analysis on completers but we will perform sensitivity analysis for worse and best case scenarios according to the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011c).

Assessment of heterogeneity

We did not conduct a meta-analysis because the comparison controls were different. If we had more than two included trials for each comparison, we would also have assessed statistical heterogeneity in each meta-analysis using the T^2, I^2 and Chi^2 statistics. We would have regarded heterogeneity as substantial if the I^2 statistic value was > 30% and either T^2 was > zero, or there was a low P value (< 0.10) in the Chi^2 test for heterogeneity (Deeks 2011). In future updates we will measure heterogeneity if three or more trials are included.

Assessment of reporting biases

Only two trials were available, so we did not explore publication bias.

We would also have attempted to assess whether this Cochrane Review is subject to publication bias by using a funnel plot to graphically illustrate variability between trials. If we had detected asymmetry, we would have explored causes other than publication bias (e.g., selective outcome reporting, poor methodological quality in smaller studies, true heterogeneity) (Higgins 2011d). In future updates we will construct a funnel plot, provided we have ten or more RCTs for each comparison (Sterne 2011).

Data synthesis

Although we planned to conduct meta-analyses, we ultimately only conducted a qualitative synthesis of the results from the two included trials. We did not pool data due to the huge differences between the control groups in the two included trials and follow-up duration.

In future versions of the review we plan to carry out statistical analyses using RevMan 2014 software using random-effects models according to the *Cochrane Handbook for Systematic Reviews of Interventions* (Deeks 2011).

Summary of findings

We used the principles of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system to assess the quality of the body of evidence associated with all main outcomes (overall survival, progression-free survival, safety (hematological adverse events), health-related quality of life) (Guyatt 2011c), and we constructed a 'Summary of findings' table using GRADEpro 2014 software. The GRADE approach appraises the quality of a body of evidence based on the extent to which one can be confident that an estimate of effect or association reflects the item being assessed. Evaluation of the quality of a body of evidence considers within study risk of bias, the directness of the evidence, heterogeneity in the data, precision of effect estimates, and risk of publication bias (Balshem 2011; Guyatt 2011a; Guyatt 2011b; Guyatt 2011d; Guyatt 2011e; Guyatt 2011f; Guyatt 2011g; Guyatt 2011h; Guyatt 2011i; Guyatt 2013).

We only included hematological adverse events because these are more relevant than non-hematological adverse events (Summary of findings for the main comparison; Summary of findings 2).
We used the number of deaths reported in COMFORT-I 2012; COMFORT-II 2012 to estimate mortality in the 'Summary of findings' table as an approach of overall survival. However, it only was reported in the 'Summary of findings' table rather than in the review text. GRADEpro 2014 does not allow estimation of either assumed or corresponding risks.

Subgroup analysis and investigation of heterogeneity

We would have used the following procedures (and will apply these for future updates, if possible). We had anticipated clinical heterogeneity in the intervention effect and we had proposed to conduct the following subgroup analyses:

- MF subtype:
  - PMF.
  - Philadelphia-chromosome-negative myeloproliferative disorders: post-polycythemia vera MF and post-essential thrombocythemia MF.
  - Secondary MF (such as cancer, tuberculosis, and radiation).
- JAK-2 V617F mutation status at screening.
- Previous MF therapy.
- Ruxolitinib versus other Janus kinase-1 and Janus kinase-2 inhibitors.

These different variables justify subgroup analyses. In future updates we will perform subgroup analyses only for primary outcomes.

Sensitivity analysis

We would also have conducted sensitivity analysis according to Higgins 2011d. In future updates, if we identify sufficient trials, we will conduct sensitivity analyses excluding:

1. Trials at high risk of bias (i.e., trials that do not meet at least one of the criteria for assessing risk of bias as outlined earlier). We will not remove trials at high risk of bias from the main analysis but will analyze them separately.
2. Trials with a total attrition of > 30%, or where baseline differences between the groups exceed 10%, or both.
3. Unpublished studies, since these may not have been subjected to the peer review process and may have intrinsic bias issues.

Trial sequential analysis

If a sufficient number of trials had met the inclusion criteria we would also have conducted a trial sequential analysis, which is a methodology that combines an information size calculation (cumulated sample sizes of included trials) for meta-analysis with the threshold of statistical significance. Trial sequential analysis is a tool for quantifying the statistical reliability of data in a cumulative meta-analysis adjusting P values for repetitive testing on accumulating data (Brok 2009; Pogue 1997; Pogue 1998; Thorlund 2009; Wetterslev 2008).
Meta-analysis may result in type I errors due to sparse data or due to repeated significance testing when updating meta-analysis with new trials (Brok 2009; Higgins 2011e; Wetterslev 2008). In a single trial, interim analysis increases the risk of type I errors. To avoid type I errors, group sequential monitoring boundaries are applied to decide whether a trial could be terminated early because of a sufficiently small P value that is the cumulative Z-curve crosses the monitoring boundaries (Lan 1983). Sequential monitoring boundaries can be applied to meta-analysis as well, called trial sequential monitoring boundaries (Wetterslev 2008). In trial sequential analysis, the addition of each trial in a cumulative meta-analysis is regarded as an interim meta-analysis and helps to clarify whether additional trials are needed.
The idea in trial sequential analysis is that if the cumulative Z-curve crosses the boundary, a sufficient level of evidence is reached and no further trials may be needed. If the Z-curve does not cross the boundary then there is insufficient evidence to reach a conclusion. To construct the trial sequential monitoring boundaries the required information size is needed and is calculated as the least number of participants needed in a well-powered single trial (Brok 2009; Pogue 1997; Pogue 1998; Wetterslev 2008). We would applied trial sequential analysis since it prevents an increase of the risk of type I error (< 5%) due to potential multiple updating in a cumulative meta-analysis and provides us with important information in order to estimate the level of evidence of the experimental intervention. Additionally, trial sequential analysis provides us with important information regarding the need for additional trials and the required sample size of such trials.
We would have applied trial sequential monitoring boundaries according to a heterogeneity-adjusted required information size based on an a priori 10% relative risk reduction employing α = 0.05 and ß = 0.20.
We would have conducted trial sequential analysis using available statistical software (CTU 2011; Thorlund 2011a).

RESULTS

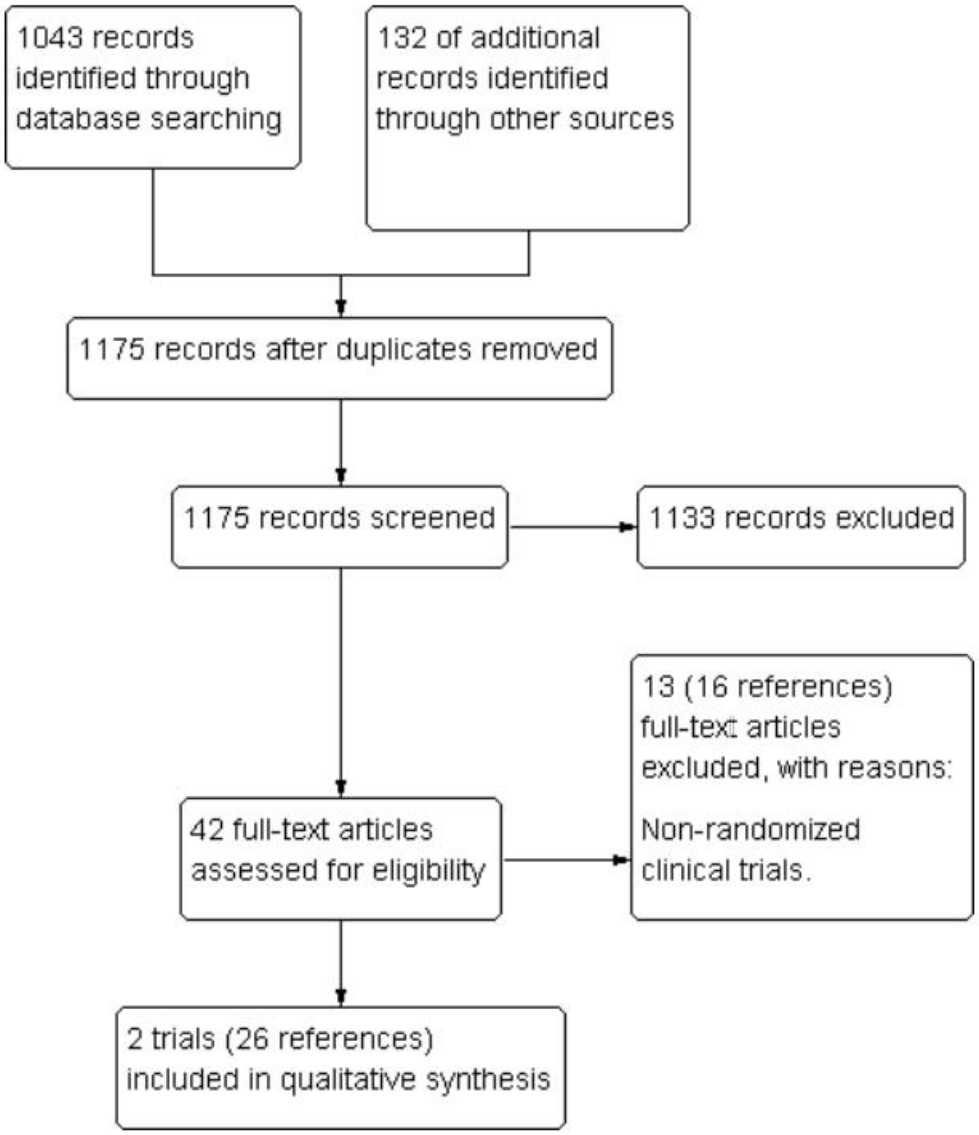
Description of studies

Results of the search

We identified 1175 references using the previously described strategy. Two trials (26 publications) with a total of 528 participants met our inclusion criteria (COMFORT-I 2012; COMFORT-II 2012). Figure 1 shows the flowchart results of the study selection

as a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Moher 2009).

Figure 1. Study flow diagram.



Included studies

Interventions and populations assessed in the trials

One trial compared ruxolitinib with placebo (COMFORT-I 2012) and the other compared ruxolitinib with best available therapy (COMFORT-II 2012). Both trials initiated ruxolitinib following an oral schema of administration (15 mg twice daily or 20 mg twice daily), based on baseline peripheral blood platelet count. We did not identify trials comparing JAK inhibitors head to head. Best available therapy was composed any commercially available as monotherapy or in combination such as: antineoplastic agents, glucocorticoids, anti-anemia preparations, immunomodulatory agents, purine analogs, antigonadotropins and similar agents, interferons, nitrogen mustard analogs, pyrimidine analogs, or no therapy at all and which could be changed during the treatment phase (COMFORT-II 2012).

COMFORT-I 2012 and COMFORT-II 2012 included patients diagnosed with PMF, post-polycythemia vera-MF or post-essential thrombocythemia-MF according to the 2008 WHO criteria. The mean percentage of male participants was 56% (± 2.83), with a median age of 68 years.

Location and timing of trials

Both trials were published in 2012 and were conducted in USA, Canada, Australia (COMFORT-I 2012), and several European countries (Austria, Belgium, France, Germany, Italy, Netherlands, Spain, Sweden, and UK) (COMFORT-II 2012).

Trial methods

The two trials were conducted using a parallel study design. The trials had a sample size of 219 (COMFORT-II 2012) and 309 (COMFORT-I 2012) patients. Both trials were conducted with a priori sample size estimation and the follow-ups ranged from 32 to 61 weeks (COMFORT-I 2012; COMFORT-II 2012).

We have given a detailed description of the trials in the Characteristics of included studies tables (COMFORT-I 2012; COMFORT-II 2012).

Excluded studies

We excluded 13 studies (Geyer 2014; Gisslinger 2012; Guglielmelli 2011; le Coutre 2012; Mesa 2007; Mesa 2014; Pardanani 2011a; Pardanani 2013; Santos 2010; Talpaz 2013; Verstovsek 2010; Verstovsek 2011; Verstovsek 2014). The excluded studies were non-RCTs (see Characteristics of excluded studies for details).

Ongoing trials

We identified one ongoing trial (NCT01437787) entitled "Phase III study of SAR302503 in intermediate-2 and high risk patients with myelofibrosis (JAKARTA)". It is a phase 3, multicenter, randomized, double-blind, placebo-controlled, three-arm study of SAR302503 in patients with intermediate-2 or high-risk PMF, post-polycythemia vera MF, or post-essential thrombocythemia MF with splenomegaly. This RCT will assess the efficacy of daily oral doses of 400 mg or 500 mg of SAR302503 (Investigational Medicinal Product, IMP) compared with placebo in the reduction of spleen volume.

Risk of bias in included studies

We have summarized the risks of bias in the included trials in Figure 2 and Figure 3, and are detailed in the Characteristics of included studies table.

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

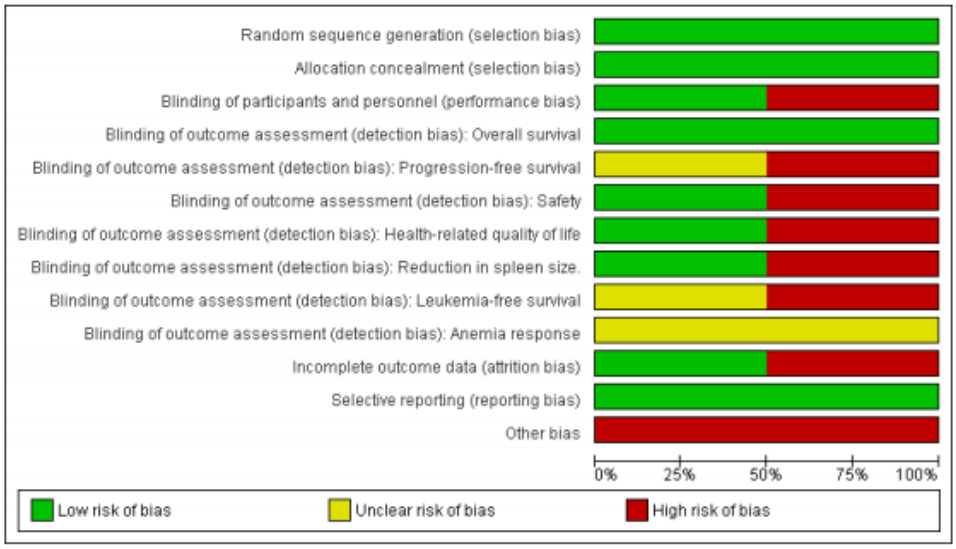


Figure 3. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.

Allocation

Random sequence generation

Both trials randomized participants by an interactive voice response system. The risk of bias arising from the method of generation of the allocation sequence was low in both trials (COMFORT-I 2012; COMFORT-II 2012).

Allocation concealment

Both trials randomized participants by an interactive voice response system. The risk of bias arising from the method of allocation concealment was low in both trials (COMFORT-I 2012; COMFORT-II 2012).

Blinding

Drug preparations in COMFORT-I 2012 were prepared to be indistinguishable, thus avoiding risk of performance or detection bias. On the other hand, COMFORT-II 2012 had a open design

and had a high risk of performance or detection bias for most of the outcomes assessed.

Blinding of outcome assessment (detection bias)

1. Primary outcomes

Overall survival

We judged the risk of bias as low in this domain in both the trials (COMFORT-I 2012; COMFORT-II 2012).

Progression-free survival

The risk of bias of this domain was judged as low in COMFORT-I 2012. We rated the COMFORT-II 2012 trial as at high risk of bias because it is an open trial.

Safety

We reported COMFORT-I 2012 as at low risk of bias for safety outcomes. COMFORT-II 2012 was at high risk of bias.

2. Secondary outcomes

Health-related quality of life

We reported a low risk of bias in this outcome in COMFORT-I 2012. However, the risk of bias was high in COMFORT-II 2012 because it is an open trial.

Leukemia-free survival

One trial did not assess this end point, therefore we judged it as at unclear risk of bias (COMFORT-I 2012). The other trial (COMFORT-II 2012) was at high risk of bias.

Reduction in spleen size

We judged the quality of COMFORT-I 2012 as at low risk of bias. However, COMFORT-II 2012 was at high risk of bias.

Anemia response

We reported the quality of the included trials (COMFORT-I 2012; COMFORT-II 2012) as at unclear risk of bias.

Incomplete outcome data

We judged a high risk of bias for COMFORT-I 2012 due to reporting of the primary outcome (reduction in spleen size) using only 79.2% (245/309) of the initially randomized participants (ruxolitinib group (89.6% (139/155)) versus placebo group (68.8% (106/154)). Furthermore, this trial shows an imbalance of 20.8% between the comparison groups. We considered COMFORT-II 2012 as at low risk of bias.

Selective reporting

Both trials were at low risk of reporting bias (COMFORT-I 2012; COMFORT-II 2012).

Other potential sources of bias

A pharmaceutical company funded COMFORT-I 2012 and COMFORT-II 2012. Therefore, we rated both trials at high risk of industry bias (Lundh 2012).

Effects of interventions

See: **Summary of findings for the main comparison** Ruxolitinib compared with placebo for treating myelofibrosis; **Summary of findings 2** Ruxolitinib compared with best available therapy for treating myelofibrosis

The results of this Cochrane review are based on two included trials (COMFORT-I 2012; COMFORT-II 2012). See Summary of findings for the main comparison and Summary of findings 2 for evidence reported by the trials on outcomes.

1. Primary outcomes

Overall survival

Ruxolitinib versus placebo

Ruxolitinib significantly improved overall survival at 51 weeks of follow-up, when compared with placebo (HR 0.51, 95% CI 0.27 to 0.98; one trial, 309 participants, *low quality evidence*; Analysis 1.1).

Ruxolitinib versus best available therapy

There was no significant difference between ruxolitinib and best available therapy in overall survival at 48 weeks of follow-up (HR 0.70; 95% CI 0.20 to 2.47; one trial, 219 participants; P = 0.58, *low quality evidence*; Analysis 2.1).

Progression-free survival

Ruxolitinib versus placebo

COMFORT-I 2012 did not report results on progression-free survival.

Ruxolitinib versus best available therapy

The comparison between ruxolitinib and best available therapy showed no statistically significant difference in progression-free survival at 48 weeks of follow-up (HR 0.81, 95% CI 0.47 to 1.39; P = 0.45, *low quality evidence*; Analysis 2.2).

Safety

We report data on adverse events (grades 3 or 4 according to the National Cancer Institute Common Terminology Criteria for Adverse Events) observed in 10% or more of patients who received ruxolitinib.

Ruxolitinib versus placebo

Hematological adverse events

Ruxolitinib compared with placebo showed a statistically significant increase in risk of anemia (70/155 (45.16%) versus 29/151 (19.20%); RR 2.35, 95% CI 1.62 to 3.41, *low quality evidence*), thrombocytopenia (20/155 (12.90%) versus 2/151 (1.32%); RR 9.74, 95% CI 2.32 to 40.96, *low quality evidence*), and neutropenia (11/155 (7.09%) versus 3/151 (1.98%); RR 3.57; 95% CI 1.02 to 12.55, *low quality evidence*; Analysis 1.2).

Non-hematological adverse events

Patients treated with ruxolitinib, compared with placebo, had a statistically significant reduction in abdominal pain (4/155 (2.58%) versus 17/151 (11.25%); RR 0.23, 95% CI 0.08 to 0.67, *low quality evidence*), and dizziness (1/155 (0.64%) versus 10/151 (6.62%); RR 0.10, 95% CI 0.01 to 0.75, *low quality evidence*).

Comparing ruxolitinib with placebo, there was not a statistically significant difference in terms of fatigue (8/155 (5.16%) versus 10/151 (6.62%); RR 0.78, 95% CI 0.32 to 1.92; P = 0.59, *low quality evidence*), dyspnea (2/155 (1.29%) versus 10/151 (6.62%); RR 0.32, 95% CI 0.07 to 1.58; P = 0.16, *low quality evidence*), arthralgia (3/155 (1.93%) versus 1/151 (0.66%); RR 2.92, 95% CI 0.31 to 27.79; P = 0.35, *low quality evidence*), nausea (0/155 (0%) versus 1/151 (0.66%); RR 0.32, 95% CI 0.01 to 7.91; P = 0.49, *low quality evidence*), vomiting (1/155 (0.64%) versus 1/151 (0.66%); RR 0.97, 95% CI 0.06 to 15.43; P = 0.99, *low quality evidence*), diarrhea (2/155 (1.29%) versus 0/151 (0%); RR 4.87, 95% CI 0.24 to 100.64; P = 0.39, *low quality evidence*), pain in extremity (2/155 (1.29%) versus 0/151 (0%); RR 4.87, 95% CI 0.24 to 100.64; P = 0.31, *low quality evidence*), or pyrexia (1/155 (0.64%) versus 1/151 (0.66%); RR 0.97; 95% CI 0.06 to 15.43; P = 0.99, *low quality evidence*) (see Analysis 1.3).

COMFORT-I 2012 did not provide details on adverse drug reactions.

Ruxolitinib versus best available therapy

Hematological adverse events

Comparing ruxolitinib versus best available therapy, there was not a statistically significantly difference in terms of anemia (62/146 (4.10%) versus 23/73 (31.50%); RR 1.35, 95% CI 0.91 to 1.99; P = 0.13, *low quality evidence*) and thrombocytopenia (12/146 (8.21%) versus 5/73 (6.84%); RR 1.20, 95% CI 0.44 to 3.28; P = 0.72, *low quality evidence*) (Analysis 2.3).

COMFORT-II 2012 did not report results regarding neutropenia.

Non-hematological adverse events

Ruxolitinib compared with best available therapy showed no statistically significant difference in terms of abdominal pain (5/146 (3.42%) versus 2/73 (2.73%); RR 1.25, 95% CI 0.25 to 6.29; P = 0.80, *low quality evidence*), fatigue (1/146 (0.68%) versus 0/73 (0%); RR 1.51, 95% CI 0.06 to 36.62; P = 0.80, *low quality evidence*), dyspnea (1/146 (0.68%) versus 3/73 (4.10%); RR 0.17, 95% CI 0.02 to 1.57; P = 0.12, *low quality evidence*), arthralgia (1/146 (0.68%) versus 0/73 (0%); RR 1.51, 95% CI 0.06 to 36.62; P = 0.80, *low quality evidence*), nausea (1/146 (0.68%) versus 0/73 (0%); RR 1.51, 95% CI 0.06 to 36.62; P = 0.80, *low quality evidence*), diarrhea (2/146 (1.36%) versus 0/73 (0%); RR 2.52, 95% CI 0.12 to 51.76; P = 0.55, *low quality evidence*), pain in extremity (1/146 (0.68%) versus 0/73 (0%); RR 1.51, 95% CI 0.06 to 36.62; P = 0.80, *low quality evidence*), pyrexia (3/146 (2.05%) versus 0/73 (0%); RR 3.52, 95% CI 0.18 to 67.32; P = 0.40, *low quality evidence*), and headache (2/146 (1.36%) versus 0/73 (0%); RR 2.52, 95% CI 0.12 to 51.76; P = 0.55, *low quality evidence*) (Analysis 2.4).

COMFORT-II 2012 did not provide details on adverse drug reactions.

2. Secondary outcomes

Health-related quality of life

Ruxolitinib versus placebo

COMFORT-I 2012 assessed health-related quality of life using the modified symptom score MFSAF. It measured the symptoms of night sweats, itching, abdominal discomfort, pain under the ribs on the left side, a feeling of fullness (early satiety), muscle or bone pain, and inactivity (COMFORT-I 2012). Each symptom score ranged from 0 (absent symptoms) to 10 (worst imaginable symptoms). The total MFSAF score is the sum of the individual scores, excluding inactivity.

COMFORT-I 2012 reported a higher proportion of patients receiving ruxolitinib that achieved a reduction of 50% or more in the total MFSAF score (RR 8.82, 95% CI 4.40 to 17.69; one trial, 309 participants; Analysis 1.4). The trial found a statistically significant improvement in score in the ruxolitinib treated group compared with placebo at 24 weeks follow-up. It reported that patients receiving ruxolitinib had a mean improvement of 46.1% (median 56.2%) while patients receiving placebo had a mean worsening of 41.8% (median 14.6%) (MD -87.90; 95% CI -139.58 to -36.22; one trial, 232 participants; P = 0.0009, *low quality evidence*; Analysis 1.5).

Ruxolitinib versus best available therapy

COMFORT-II 2012 used the European Organization for Research and Treatment of Cancer quality of life questionnaire core model, whose scale has a range of 0 a 100. This trial assessed this outcome at 48 weeks follow-up. It showed a statistically significant difference comparing ruxolitinib with best available therapy (MD 7.60, 95% CI 0.35 to 14.85; one trial, 96 participants; P = 0.04, *low quality evidence*; Analysis 2.5).

Leukemia-free survival

Ruxolitinib versus placebo

COMFORT-I 2012 showed no statistically significant difference between ruxolitinib and placebo (HR 5.0, 95% CI 0.52 to 48.07; one trial, 309 participants; P = 0.16, *low quality evidence*; Analysis 1.7), regarding leukemia-free survival.

Ruxolitinib versus best available therapy

COMFORT-II 2012 found no statistically significant difference comparing ruxolitinib versus best available therapy on leukemia-free survival (HR 0.65, 95% CI 0.18 to 2.33; one trial, 219 participants; P = 0.51, *low quality evidence*; Analysis 2.8).

Reduction in spleen volume

The primary end point for both included trials was the proportion of patients with at least a reduction of 35% in spleen volume from baseline to the end of follow-up (COMFORT-I 2012; COMFORT-II 2012).

Ruxolitinib versus placebo

At week 24, patients receiving ruxolitinib, 139 participants, had a mean reduction in spleen volume of 31.6% (median 33%) compared with 106 participants on placebo who had a mean increase of 8.1% (median 8.5%) (P = no reported) (COMFORT-I 2012). Ruxolitinib treatment significantly increased the proportion of patients with reduction in spleen volume of $\geq$ 35% as assessed by magnetic resonance imaging (MRI) or computed tomography (CT) (65/155 (41.94%) versus 1/154 (0.65%); RR 64.58, 95% CI 9.08 to 459.56; one trial, 309 participants; P = 0.0001, *low quality evidence*; Analysis 1.6).

Ruxolitinib versus best available therapy

COMFORT-II 2012 showed a statistically significant reduction in spleen size in ruxolitinib group compared with best available therapy. This effect was either at 24 weeks (MD -31.9, 95% CI -53.85 to -9.95; one trial, 216 participants; P = 0.004) or 48 weeks of follow-up (MD -37.4; 95% CI -65.41 to -9.39; P = 0.004; one trial, 216 participants, *low quality evidence*; Analysis 2.6).

Ruxolitinib treatment significantly increased the proportion of patients with reduction in spleen volume of $\geq$ 35% as assessed by MRI or CT (41/146 (71.92%) versus 0/73 (0%); RR 41.78, 95% CI 2.61 to 669.75; P = 0.008, *low quality evidence*; Analysis 2.7).

Anemia response

No trial assessed this outcome.

ADDITIONAL SUMMARY OF FINDINGS *[Explanation]*

Ruxolitinib compared to best available therapy for treating MF

Patient or population: Patients with treating MF
Settings: Ambulatory
Intervention: Ruxolitinib
Comparison: Best available therapy[1]

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (trials)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Best available therapy	Ruxolitinib			
Overall survival (number of deaths at follow-up (48 weeks[2]))	**55 per 1000**[3]	**39 per 1000** (11 to 130)[4]	**HR 0.70** (0.20 to 2.47)[5]	219 (1 trial[1])	⊕⊕◯◯ **low**[6,7]
Progression Free Survival (number of patients who had progression at follow-up (48 weeks[2]))	**260 per 1000**[3]	**217 per 1000** (132 to 342)[4]	**HR 0.81** (0.47 to 1.39)[8]	219 (1 trial[1])	⊕⊕◯◯ **low**[6,7]
Safety (AE, adverse drug reaction): anemia Grades 3 or 4 according to National Cancer Institute Follow-up: 48 weeks	**315 per 1000**	**425 per 1000** (287 to 627)	**RR 1.35** (0.91 to 1.99)	219 (1 trial[1])	⊕⊕◯◯ **low**[6,7]
Safety (AE, adverse drug reaction): thrombocytopenia Grades 3 or 4 according to National Cancer Institute Follow-up: 48 weeks	**68 per 1000**	**82 per 1000** (30 to 225)	**RR 1.20** (0.44 to 3.28)	219 (1 trial[1])	⊕⊕◯◯ **low**[6,7]

Health-related quality of life European Organization for Research and Treatment of Cancer quality of life questionnaire core model. Scores ranges from 0 to 100. Scale from: 0 to 100. Follow-up: 48 weeks		The mean health-related quality of life in the intervention groups was **7.6 higher** (0.35 to 14.85 higher)		96 (1 trial[1])	⊕⊕◯◯ **low**[6,7]
Reduction in spleen size Magnetic resonance imaging or computed tomography Follow-up: 48 weeks			**RR 41.78** (2.61 to 669.75)	219 (1 trial[1])	⊕⊕◯◯ **low**[6,7]

*The basis for the **assumed risk** is provided in footnote #3. The **corresponding risk** (and its 95% CI) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: confidence interval; **RR:** risk ratio; **HR:** hazard ratio.

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1]COMFORT-II 2012.
[2]COMFORT-II 2012 trial established the time of data cut-off for its main outcome at 48 weeks.
[3]Data obtained from deaths in the best available therapy group at the time of data cut-off (48 weeks).
[4]Data obtained from deaths in the intervention group at the time of data cut-off (48 weeks).
[5]See Analysis 2.1.
[6]Downgraded one level due to limitations in the trial design or execution (open design).
[7]Downgraded one level due to imprecision (low sample and number of events with an impact in the precision of the effect estimates).
[8]See Analysis 2.2.

DISCUSSION

Summary of main results

We performed a systematic review with the aim of obtaining the core evidence regarding clinical benefits and harms of Janus kinase inhibitors in MF. This Cochrane Review included two small trials with 528 participants. The trials compared ruxolitinib with placebo (COMFORT-I 2012) and best available therapy (COMFORT-II 2012). Both trials were sponsored by a pharmaceutical company. See Summary of findings for the main comparison and Summary of findings 2 for the grading recommendations for each of the variables assessed by both trials.

The following findings emerged from this Cochrane Review:

- Included trials reported overall survival at 51 weeks (COMFORT-I 2012) and at 48 weeks (COMFORT-II 2012). The analysis showed a significant improvement in overall survival with ruxolitinib compared with placebo but a non-significant change compared with best available therapy.
- One trial assessed progression-free survival at 48 weeks follow-up, which was found to be non-statistically different between ruxolitinib and best available therapy (COMFORT-II 2012).
- Leukemia-free survival was reported by COMFORT-I 2012 and COMFORT-II 2012 and there was no significant difference between ruxolitinib and placebo or best available therapy treated patients.
- With respect to the hematological adverse events, the risk of anemia and thrombocytopenia was increased with ruxolitinib compared with placebo (COMFORT-I 2012), but was similar compared with best available therapy (COMFORT-II 2012). COMFORT-I 2012 reported neutropenia and analysis showed an increased risk, which is statistically significant in the ruxolitinib group compared with placebo.
- Compared with placebo, ruxolitinib showed a significant reduction in abdominal pain and dizziness (COMFORT-I 2012).
- There was not a statistical significant difference regarding non-hematological grade 3 to 4 adverse events including fatigue, arthralgia, nausea, diarrhea, extremity pain and pyrexia, between ruxolitinib and placebo or best available therapy (COMFORT-I 2012; COMFORT-II 2012). Only COMFORT-I 2012 reported vomiting and showed similar risk in ruxolitinib treated patients compared with placebo. The risk of headache was also not statistically different between ruxolitinib and best available therapy.
- Analysis of the included trials reported a statistically significant improvement in health-related quality of life compared with placebo (COMFORT-I 2012) or with best-available therapy (COMFORT-II 2012).
- Improvement in splenomegaly was reported both as dichotomous and continuous approaches. Ruxolitinib reduced the proportion of patients with reduction in spleen volume of ≥ 35% as assessed by MRI or CT compared with placebo or best available therapy at 24 and 48 week of follow-up. COMFORT-II 2012 also showed a statistically significant reduction in spleen volume reported as continuous variable, in ruxolitinib group compared with best available therapy. Continuous data of spleen size reduction in COMFORT-I 2012 trial could not be analyzed because of lack of reporting of the dispersion measures.

Overall completeness and applicability of evidence

We found weak evidence suggesting that ruxolitinib increases overall survival compared with placebo or best available therapy in patients with MF. However, this conclusion is based on two small RCTs which were sponsored by a drug company (COMFORT-I 2012; COMFORT-II 2012). Both trials were not powered for finding significant difference in overall survival. Accordingly, both included trials have risk of random error (Savović 2012; Thorlund 2011b).

The results in this Cochrane Review are based on data from two trials that included a broad range of patients with both primary and secondary MF who received different treatment approaches. Although these aspects could be considered as a threat to applicability, the consistency in the results derived from our analyses shows that the included trials may represent a broad picture of patients with MF. We tried to identify all the published and unpublished data, and ongoing studies to warrant confidence in the completeness of the data gathered in the review. However, we cannot rule out that the calculated effects are overestimated due to potential industry bias, unblinding for assessing health-related quality of life in one included trial (COMFORT-II 2012), and small sample size of the included clinical trials. Furthermore, we do not preclude an underestimation of safety findings (Savović 2012; Thorlund 2011b; Wood 2008). COMFORT-II 2012 did not report data on neutropenia which prevented analysis for this relevant adverse event on comparison with best available therapy. In terms of overall survival, the duration of follow-up in COMFORT-I 2012 and COMFORT-II 2012 was 51 weeks and 48 weeks, respectively. Both periods were very short regarding the reported median overall survival of MF: 11.3 years for low risk, 7.9 years for intermediate-1 risk, 4.0 years for intermediate-2 risk, and 2.3 years for high-risk MF (Cervantes 2009).

When dealing with such neutral results, we need to keep in mind that 'absence of evidence' is not 'evidence of absence' (Altman 1995; Fermi Paradox). The fact that this review did not detect strong differences between comparison groups does not imply that placebo or best available therapy and ruxolitinib have the same overall survival risk. The first possible explanation is failure to determine an appropriate sample size (Green 2002; Schulz 1995). Furthermore, we would like to point out a form of bias known as dichotomization. Dichotomization is the transformation of a

continuous outcome (response) to a binary outcome (Fedorov 2009). There are several publications reporting the negative consequences when a continuous variable is dichotomized i.e., loss of information which leads to loss of power or conversely a sample size increase to maintain power (Altman 2006; Fedorov 2009; MacCallum 2002). Power is reduced and relationships may be obscured or changed (Peacock 2012). However, not only are differences in means difficult for clinicians to interpret, but thresholds also occur in many areas of medical practice and cannot be ignored (Peacock 2012). Dichotomization may also increase the risk of a positive result being a false positive (Altman 2006; Austin 2004). Therefore, it has been strongly recommended to avoid, as much as possible, the categorization of variables when doing analyses (Altman 2006; Cumsille 2000; MacCallum 2002; Streiner 2002). In this Cochrane Review we identified dichotomization in both trials for measuring the clinical benefit of ruxolitinib compared with control on the main outcome of these trials, spleen size reduction (COMFORT-I 2012; COMFORT-II 2012). Both trials assessed the clinical benefit of ruxolitinib on the basis of spleen size measurements and on the basis of the proportions of patients with spleen size reduction in prespecified range (≥ 35%). Also, COMFORT-I 2012 reported only mean and median of spleen size without any data of standard deviation, standard error, 95% CI, and interquartile range, respectively. In consequence, the true precision of the clinical benefit of ruxolitinib compared with placebo on reduction in spleen size is unknown.

COMFORT-I 2012 and COMFORT-II 2012 used a surrogate outcome for assessing a benefit effect of the ruxolitinib on MF. This review could adapt the comments and points of view from Yudkin 2011 and ratified by Godlee 2012 which are related with diabetes world. From their perspective within the world of malignant hematological disorder we would warn that surrogates like spleen size or spleen volume could show much larger responses to treatment than "hard" outcomes that matter to patients, such as overall survival, progression-free survival impairment or quality of life. Furthermore, surrogate outcomes also respond sooner, which makes them popular with drug companies and others doing clinical trials. Moreover, these "hard" end points generally show much smaller responses to interventions than surrogate markers. As it has happened with other medical disorders, to adopt ruxolitinib for treating patients with MF may be based on artificially inflated expectations. Outcome events that are more frequent in occurrence and more proximate in time, compared with customary disease-specific mortality or incidence outcomes, could give answers that are based on smaller trials of shorter duration (Prentice 2009).

Summing up, the clinical meaning of spleen size reduction in prespecified range that is ≥ 35%, is unknown. It can be suggested to adopt ≥ 50% as the cut-off, on the basis of the international response criteria of a reduction of 50% or more in spleen length as assessed by palpation. Since spleen size reduction was the primary outcome used by COMFORT-I 2012 and COMFORT-II 2012 for assessing efficacy of ruxolitinib, from our point of view it would have been more relevant to show the spleen size reduction using continuous measure rather than as binary data, with all dispersion measures.

Quality of the evidence

We did not grade any the results as high quality evidence primarily because of small sample sizes and the high risk of bias due to a lack of blinding and high attrition (Summary of findings for the main comparison; Summary of findings 2).

We found many sources of bias in both included trials (COMFORT-I 2012; COMFORT-II 2012). First, we detected performance bias in COMFORT-II 2012. Second, there was suspicion of detection bias regarding progression-free survival in COMFORT-I 2012. Third, COMFORT-II 2012 was described as open trial, therefore it had a high risk of detection bias (blinding of outcome assessor) regarding health-related quality of life and reduction in spleen size. Fourth, COMFORT-I 2012 showed a high risk of attrition bias (Porta 2008). Fifth, dichotomization increases the risk of a positive result being a false positive (Altman 2006; Austin 2004). Sixth, a pharmaceutical company sponsored both trials and are potentially at high risk of industry bias. Significant evidence supports a clear association between pharmaceutical industry funding of clinical trials and pro-industry results (Als-Nielsen 2003; Djulbegovic 2013; Doucet 2008; Golder 2008; Jørgensen 2008; Lexchin 2003; Lundh 2012; Schott 2010a; Schott 2010b). Industry bias results in publication of scientific research which is in favor of the commercial interests of the sponsors. COMFORT-I 2012 did not report all the data regarding its primary end point, and COMFORT-II 2012 did not report complete health related quality of life data at 48 weeks follow-up. Many recommendations have been suggested, such as a public access to trial protocols and results, and more effort should be made to carry out drug trials independently, without the financial support of pharmaceutical companies (Schott 2010a; Schott 2010b).

COMFORT-I 2012 and COMFORT-II 2012 were small trials which are at potential risk of industry bias (Twombly 2007). A small study could cause a small study effect bias (Hemming 2009). It has been described as "decline effect", by which drugs appear to yield a lower effect size over time (Lauer 2012). The decline effect is due, at least in part, to over interpretation of small studies (Lauer 2012). Therefore, meta-analyses and systematic reviews should always consider the impact of attrition on baseline imbalances and where possible any baseline imbalances in the analyzed data set and their impact on the outcomes reported (Hewitt 2010). In this Cochrane Review, we found an imbalance between ruxolitinib and placebo of 20.8% and 16.4% regarding spleen size reduction and health-related quality, respectively (COMFORT-I 2012).

Potential biases in the review process

In the process for performing a systematic review, there is a group of biases called 'significance-chasing biases' (Ioannidis 2010). This group includes publication bias, selective outcome reporting bias, selective analysis reporting bias, and fabrication bias (Ioannidis 2010).

Publication bias represents a major threat to the validity of systematic reviews, particularly in reviews that include small trials as this Cochrane Review. We included two small trials involving 528 patients. However, this Cochrane Review has a low risk of publication bias due to the meticulous trial search.

Selective outcome reporting bias operates through suppression of information on specific outcomes and has similarities to publication bias in sense that 'negative' results remain unpublished (Ioannidis 2010). We found two trials at low risk of selective outcome reporting bias.

The major limitation of this review is associated with the small sample size of the included trials. A study with low statistical power has a reduced chance of detecting a true effect (power failure), which overestimates the effect size and low reproducibility of results (Button 2013; Freiman 1978; Kirby 2002; Moher 1998). The potential consequences are generation of excess significance, winner's curse, and vibration of effects (Ioannidis 2005; Ioannidis 2008; Pereira 2011). We have provided definitions for excess significance, winner's curse, and vibration of effects definitions in Appendix 9 (Button 2013; Ioannidis 2008).

The main strength of this Cochrane Review is that we have found a need of new powered trials based on main clinical outcomes, such as overall survival and progression-free survival as primary outcomes.

We tried to avoid any bias by having two review authors conduct the steps of study selection, data extraction and analysis, and risk of bias assessment in duplicate with suggestions from other review authors and correspondence with the trial authors when needed. We are not aware of any obvious biases in our review process.

Finally, in the 'Summary of findings' tables we present measures of absolute effect for time-to-event outcomes (overall survival and progression free survival). In generating such estimates we assumed the control group rate that could be valid for RR estimation as a close approximation to HR. As this approach does not reflect cumulative risk our reported estimate might differ to what would be observed in practice. To our knowledge there is no more robust approaches to this, so the illustrative risk included in the 'Summary of findings' tables should be interpreted with caution.

Agreements and disagreements with other studies or reviews

Although this is the first systematic review on the effects of Janus kinase inhibitors for MF, the results are in concordance with other narrative reviews published to date (Mascarenhas 2013; Tefferi 2012). In general terms these reviews acknowledge the role of ruxolitinib in the control of symptoms and the capacity to reduce splenomegaly, but also highlight the contradictory results regarding other relevant outcomes like survival. Despite the survival benefit observed in the COMFORT-I 2012 and the promising results from COMFORT-II 2012, the treatment of MF with ruxolitinib with the intention of prolonging survival would be premature (Mascarenhas 2013), especially having in mind that an advantage over best supportive care has still to be shown (Tefferi 2012).

AUTHORS' CONCLUSIONS

Implications for practice

Currently, there is insufficient evidence to allow any conclusions regarding the efficacy and safety of ruxolitinib for treating MF. The results for the efficacy and safety outcomes in this Cochrane Review come from two small trials. The findings need to be interpreted with caution as they are based on trials sponsored by pharmaceutical company, including small number of patients. Unless RCTs provide strong evidence of a treatment effect, and the trade off between potential benefits and harms is established clinicians should be cautious when recommending and administering ruxolitinib for treating patients with MF.

Implications for research

There is a need for powered RCTs which assess the effect of ruxolitinib in patients with MF. The potential trial should be based on patient-centered outcomes, such as overall survival, progression-free survival, quality of life measures, safety and esophogeal varices and splanchnic vein thromboses rather than spleen measure. Furthermore, the potential trial should have an adequate duration of follow-up. It should be based on median overall survival for low risk, intermediate-1 risk, intermediate-2 risk, and high-risk MF patients according to the International Working Group for Myelofibrosis Research and Treatment report (Cervantes 2009). Due to risk of myelosuppression there is need for more well-conducted trials with risk-stratification to study the independent risk-benefit ratio in different risk-groups.

Potential trials should be conducted with appropriate blinding of outcome assessment for subjective end points (quality of life measures). They should be planned using Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) (Chan 2013a; Chan 2013b; Chan 2013c) and reported according to the Consolidated Standards of Reporting Trials (CONSORT) statement for improving the quality of reporting of efficacy and of harms in clinical research (Calvert 2013; Ioannidis 2004; Moher 2010). The trials should be conducted according to the Patient-Centered Outcomes Research Institute (PCORI) recommendations (Basch 2012; Gabriel 2012; PCORI 2012).

ACKNOWLEDGEMENTS

We thank the editors and editorial staff of the Cochrane Hematological Malignancies Group for their comments.

We acknowledge Ina Monsef for significant help in conducting this Cochrane Review.

We are grateful to Dr. Srdan Verstovsek for providing complementary data for COMFORT-I 2012.

Arturo Martí-Carvajal is a PhD student at the Department of Pediatrics, Obstetrics and Gynecology, and Preventive Medicine of the Universitat Autònoma de Barcelona.

REFERENCES

References to studies included in this review

COMFORT-I 2012 *{published data only}*

Mesa R, Verstovsek S, Gupta V, Mascarenhas J, Atallah E, Sun W, et al. Improvement in weight and total cholesterol and their association with survival in ruxolitinib-treated patients with myelofibrosis from COMFORT-I. Blood (ASH Annual Meeting Abstracts). 2012; Vol. 120: 1733.

Mesa RA, Gotlib J, Gupta V, Catalano JV, Deininger MW, Shields AL, et al. Effect of ruxolitinib therapy on myelofibrosis-related symptoms and other patient-reported outcomes in COMFORT-I: a randomized, double-blind, placebo-controlled trial. *Journal of Clinical Oncology* 2013; **31**(10):1285–92. [PUBMED: 23423753]

Mesa RA, Kantarjian H, Tefferi A, Dueck A, Levy R, Vaddi K, et al. Evaluating the serial use of the Myelofibrosis Symptom Assessment Form for measuring symptomatic improvement: performance in 87 myelofibrosis patients on a JAK1 and JAK2 inhibitor (INCB018424) clinical trial. *Cancer* 2011;**117**(21):4869–77. [PUBMED: 21480207]

Mesa RA, Shields A, Hare T, Erickson-Viitanen S, Sun W, Sarlis NJ, et al. Progressive burden of myelofibrosis in untreated patients: assessment of patient-reported outcomes in patients randomized to placebo in the COMFORT-I study. *Leukemia Research* 2013;**37**(8):911–6. [PUBMED: 23684482]

Talpaz M, Paquette R, Afrin L, Hamburg S, Jamieson K, Terebelo H, et al. Efficacy, hematologic effects, and dose of ruxolitinib in myelofibrosis patients with low starting platelet counts (50-100 x 10^9/L): a comparison to patients with normal or high starting platelet counts. Blood (ASH Annual Meeting Abstracts). 2012; Vol. 120: 176.

Verstovsek S, Kantarjian HM, Estrov Z, Cortes JE, Thomas DA, Kadia T, et al. Long-term outcomes of 107 patients with myelofibrosis receiving JAK1/JAK2 inhibitor ruxolitinib: survival advantage in comparison to matched historical controls. *Blood* 2012;**120**(6):1202–9. [PUBMED: 22718840]

Verstovsek S, Mesa R, Gotlib J, Levy RS, Gupta V, DiPersio JF, et al. Long-term outcome of ruxolitinib treatment in patients with myelofibrosis: durable reductions in spleen volume, improvements in quality of life, and overall survival advantage in COMFORT-I. Blood (ASH Annual Meeting Abstracts). 2012; Vol. 120: 800.

Verstovsek S, Mesa RA, Gotlib J, Levy R, Gupta V, DiPersio JF, et al. Clinical burden and progression of myelofibrosis in a controlled study population of placebo-treated patients (COMFORT-I). Blood (ASH Annual Meeting Abstracts). 2011; Vol. 118: 5146.

* Verstovsek S, Mesa RA, Gotlib J, Levy RS, Gupta V, DiPersio JF, et al. A double-blind, placebo-controlled trial of ruxolitinib for myelofibrosis. *The New England Journal of Medicine* 2012;**366**(9):799–807. [PUBMED: 22375971]

Verstovsek S, Mesa RA, Gotlib J, Levy RS, Gupta V, DiPersio JF, et al. Consistent benefit of ruxolitinib over placebo in spleen volume reduction and symptom improvement across subgroups and overall survival advantage: Results from COMFORT-I. Blood. 2011; Vol. 118: 278.

Verstovsek S, Mesa RA, Gotlib J, Levy RS, Gupta V, DiPersio JF, et al. The clinical benefit of ruxolitinib across patient subgroups: analysis of a placebo-controlled, Phase III study in patients with myelofibrosis. *British Journal of Haematology* 2013;**161**(4):508–16. [PUBMED: 23480528]

Verstovsek S, Mesa RA, Gotlib J, Levy RS, Gupta V, Dipersio JF, et al. Efficacy, safety and survival with ruxolitinib in patients with myelofibrosis: results of a median 2-year follow-up of COMFORT-I. *Haematologica* 2013;**98**(12):1865–71. [PUBMED: 24038026]

Verstovsek S, Mesa RA, Gotlib JR, Gupta V, DiPersio JF, Catalano JV for the COMFORT-I Investigators. Adverse events (AEs) and the return of myelofibrosis (MF)-related symptoms after interruption or discontinuation of ruxolitinib (RUX) therapy. Journal of Clinical Oncology, 2012 ASCO Annual Meeting Proceedings (Post-Meeting Edition). 2012; Vol. 30(15 Suppl): 6624.

Verstovsek S, Mesa RA, Gotlib JR, Levy RS, Gupta V, DiPersio JF, et al. Results of COMFORT-I, a randomized double-blind phase III trial of JAK 1/2 inhibitor INCB18424 (424) versus placebo (PB) for patients with

myelofibrosis (MF). Journal of Clinical Oncology. 2011; Vol. 29 (15 Suppl).

COMFORT-II 2012 *{published data only}*

Cervantes F, Kiladjian JJ, Niederwieser D, Sirulnik A, Stalbovskaya V, McQuity M, et al. Long-term safety, efficacy, and survival findings from Comfort-II, a phase 3 study comparing ruxolitinib with best available therapy (BAT) for the treatment of myelofibrosis (MF). Blood (ASH Annual Meeting Abstracts). 2012; Vol. 120: 801.

Cervantes F, Vannucchi AM, Kiladjian JJ, Al-Ali HK, Sirulnik A, Stalbovskaya V, et al. Three-year efficacy, safety, and survival findings from COMFORT-II, a phase 3 study comparing ruxolitinib with best available therapy for myelofibrosis. *Blood* 2013;**122**(25):4047–53. [PUBMED: 24174625]

Guglielmelli P, Biamonte F, Rotunno G, Artusi V, Artuso L, Bernardis I, et al. Impact of mutational status on outcomes in myelofibrosis patients treated with ruxolitinib in the COMFORT-II study. *Blood* 2014;**123**(14):2157–60. [PUBMED: 24458439]

* Harrison C, Kiladjian JJ, Al-Ali HK, Gisslinger H, Waltzman R, Stalbovskaya V, et al. JAK inhibition with ruxolitinib versus best available therapy for myelofibrosis. *New England Journal of Medicine* 2012;**366**(9):787–98. [PUBMED: 22375970]

Harrison C, Kiladjian JJ, Al-Ali HK, Gisslinger H, Waltzman R, Stalbovskaya V, et al. Results of a randomized study of the JAK Inhibitor INC424 (Ruxolitinib) compared with best available therapy (BAT) in primary myelofibrosis (PMF), post-polycythemia vera myelofibrosis (PPV-MF) or post-essential thrombocythemia myelofibrosis (PET-MF). http://www.oncuview.tv/portals/0/linkedfiles/COMFOR-TII˙ASCO˙2011˙Highlight˙Slides˙Harrison.pdf (accessed 27 February 2013).

Harrison CN, Kiladjian J, Al-Ali HK, Gisslinger H, Waltzman RJ, Stalbovskaya V, et al. Results of a randomized study of the JAK inhibitor ruxolitinib (INC424) versus best available therapy (BAT) in primary myelofibrosis (PMF), post-polycythemia vera-myelofibrosis (PPV-MF) or post-essential thrombocythemia-myelofibrosis (PET-MF). Journal of Clinical Oncology. 2011; Vol. 29:LBA6501.

Harrison CN, Kiladjian J-J, Al-Ali HK, Gisslinger H, Knoops L, Waltzman RJ, et al. Health-related quality of life and symptoms in myelofibrosis patients treated with ruxolitinib versus best available therapy. Blood. 2011; Vol. 118: 795.

Harrison CN, Kiladjian J-J, Gisslinger H, Niederwieser D, Passamonti F, Waltzman RJ, et al. Ruxolitinib provides reductions in splenomegaly across subgroups: An analysis of spleen response in the COMFORT-II study. Blood. 2011; Vol. 118: 279.

Harrison CN, Mesa RA, Kiladjian JJ, Al-Ali HK, Gisslinger H, Knoops L, et al. Health-related quality of life and symptoms in patients with myelofibrosis treated with ruxolitinib versus best available therapy. *British Journal of Haematology* 2013;**162**(2):229–39. [PUBMED: 23672349]

Kiladjian JJ, Gisslinger H, Passamonti F, Niederwieser D, Mendelson E, Sirulnik LA, et al. Health-related quality of life (HRQoL) and symptom burden in patients (Pts) with myelofibrosis (MF) in the COMFORT-II study. Journal of Clinical Oncology (ASCO Annual Meeting Proceedings). 2012; Vol. 30: 6626.

McMullin MF, Harrison CN, Niederwieser D, Demuynck H, Jakel N, Sirulnik A, et al. The use of erythropoietic-stimulating agents (ESAs) with ruxolitinib in patients with primary myelofibrosis (PMF), post-polycythemia vera myelofibrosis (PPV-MF), and post-essential thrombocythemia myelofibrosis (PET-MF). Blood (ASH Annual Meeting Abstracts). 2012; Vol. 20: 2838.

McMullin MF, Harrison CN, Niederwieser D, Demuynck H, Jäckel N, Sirulnik A, et al. Anemia and the use of erythropoietic-stimulating agents with ruxolitinib in the COMFORT-II study. Blood (ASH Annual Meeting Abstracts). 2011; Vol. 118: 5147.

References to studies excluded from this review

Geyer 2014 *{published data only}*

* Geyer H, Cannon K, Knight E, Fauble V, Camoriano J, Emanuel R, et al. Ruxolitinib in clinical practice for therapy of myelofibrosis: single USA center experience following Food and Drug Administration approval. *Leukemia & Lymphoma* 2014;**55**(1):195–7. [PUBMED: 23647081]

Verstovsek S. Tips on using ruxolitinib in everyday practice as therapy for myelofibrosis. *Leukemia & Lymphoma* 2014; **55**(1):5–6. [PUBMED: 23829281]

Gisslinger 2012 *{published data only}*

Gisslinger H, McMullin MF, Jaekel N, Miller CB, Verstovsek S, Harrison CN, et al. A phase Ib, open-label, dose-finding study of ruxolitinib in patients (pts) with primary myelofibrosis (PMF), post-polycythemia vera-myelofibrosis (PPV-MF), or post-essential thrombocythemia-myelofibrosis (PET-MF) and baseline platelets (PLTs) 50 to <100 x 10^9/l. *Journal of Clinical Oncology* 2012;**30**(15 Suppl):TPS6642.

Guglielmelli 2011 *{published data only}*

Guglielmelli P, Barosi G, Rambaldi A, Marchioli R, Masciulli A, Tozzi L, et al. Safety and efficacy of everolimus, a mTOR inhibitor, as single agent in a phase 1/2 study in patients with myelofibrosis. *Blood* 2011;**118**(8):2069–76. [PUBMED: 21725052]

le Coutre 2012 *{published data only}*

le Coutre PD, Gisslinger H, Zachee P, Gupta V, Perez JR, Schenk N, et al. An open-label, multicenter, expanded access study assessing the safety and efficacy of oral ruxolitinib administered to patients with primary myelofibrosis (PMF), post-polycythemia myelofibrosis (PPV MF) or post-essential thrombocythemia myelofibrosis (PET-MF). *Journal of Clinical Oncology* 2012;**30**(15 Suppl): TPS6640.

Mesa 2007 *{published data only}*

Mesa RA, Camoriano JK, Geyer SM, Wu W, Kaufmann SH, Rivera CE, et al. A phase II trial of tipifarnib in

myelofibrosis: primary, post-polycythemia vera and post-essential thrombocythemia. *Leukemia* 2007;**21**(9): 1964–70. [PUBMED: 17581608]

Mesa 2014 *{published data only}*
Mesa RA, Kiladjian JJ, Verstovsek S, Al-Ali HK, Gotlib J, Gisslinger H, et al. Comparison of placebo and best available therapy for the treatment of myelofibrosis in the phase 3 COMFORT studies. *Haematologica* 2014;**99**(2): 292–8. [PUBMED: 23911705]

Pardanani 2011a *{published data only}*
* Pardanani A, Gotlib JR, Jamieson C, Cortes JE, Talpaz M, Stone RM, et al. Safety and efficacy of TG101348, a selective JAK2 inhibitor, in myelofibrosis. *Journal of Clinical Oncology* 2011;**29**(7):789–96. [PUBMED: 21220608]
Pardanani AD, Gotlib J, Jamieson C, Cortes J, Talpaz M, Stone RM, et al. A phase I study of TG101348, an orally bioavailable JAK2-selective inhibitor, in patients with myelofibrosis: clinical response is accompanied by significant reduction in *JAK2*V617F allele burden. Blood (ASH Annual Meeting Abstracts). 2009; Vol. 112: 755.

Pardanani 2013 *{published data only}*
Pardanani A, Laborde RR, Lasho TL, Finke C, Begna K, Al-Kali A, et al. Safety and efficacy of CYT387, a JAK1 and JAK2 inhibitor, in myelofibrosis. *Leukemia* 2013;**27**(6): 1322–7. [PUBMED: 23459451]

Santos 2010 *{published data only}*
Santos FPS, Kantarjian HM, Jain N, Manshouri T, Thomas DA, Garcia-Manero G, et al. Phase 2 study of CEP-701, an orally available JAK2 inhibitor, in patients with primary or post-polycythemia vera/essential thrombocythemia myelofibrosis. *Blood* 2010;**115**(6):1131–6. [PUBMED: 20008298]

Talpaz 2013 *{published data only}*
Talpaz M, Hamburg SI, Jamieson K, Terebelo HR, Afrin L, Winton EF, et al. Preliminary safety and efficacy of ruxolitinib in patients (pts) with primary and secondary myelofibrosis (MF) with platelet counts (PC) of 50-100x10 ^{9}/L. *Journal of Clinical Oncology* 2012;**30**(15 Suppl): TPS6630.
* Talpaz M, Paquette R, Afrin L, Hamburg SI, Prchal JT, Jamieson K, et al. Interim analysis of safety and efficacy of ruxolitinib in patients with myelofibrosis and low platelet counts. *Journal of Hematology & Oncology* 2013;**6**(1):81. [: 24283202]

Verstovsek 2010 *{published data only}*
Verstovsek S, Kantarjian H, Mesa RA, Pardanani AD, Cortes-Franco J, Thomas DA, et al. Safety and efficacy of INCB018424, a JAK1 and JAK2 inhibitor, in myelofibrosis. *New England Journal of Medicine* 2010;**363**(12):1117–27. [PUBMED: 20843246]

Verstovsek 2011 *{published data only}*
Verstovsek S, Kantarjian HM, Estrov Z, Cortes JE, Thomas DA, Kadia T, et al. Comparison of outcomes of advanced myelofibrosis patients treated with ruxolitinib (INCB018424) to those of a historical control group: survival advantage of ruxolitinib therapy. Blood (ASH Annual Meeting Abstracts). 2011; Vol. 118: 793.

Verstovsek 2014 *{published data only}*
Verstovsek S, Tam CS, Wadleigh M, Sokol L, Smith CC, Bui LA, et al. Phase I evaluation of XL019, an oral, potent, and selective JAK2 inhibitor. *Leukemia Research* 2014;**38** (3):316–22. [PUBMED: 24374145]

References to ongoing studies

NCT01437787 *{published data only}*
Pardanani AD, Cortes JE, Cervantes F, Harrison CN, Passamonti F, Lebedinsky C, et al. JAKARTA: A phase III, multicenter, randomized, double-blind, placebo-controlled, three-arm study of SAR302503 in patients with intermediate-2 or high-risk primary myelofibrosis (MF), post-polycythemia vera (PV) MF, or post-essential thrombocythemia (ET) MF with splenomegaly. *Journal of Clinical Oncology* 2012;**30**(15 Suppl):TPS6639.

Additional references

Als-Nielsen 2003
Als-Nielsen B, Chen W, Gluud C, Kjaergard LL. Association of funding and conclusions in randomized drug trials: a reflection of treatment effect or adverse events?. *JAMA* 2003;**290**(7):921–8. [PUBMED: 12928469]

Altman 1995
Altman DG, Bland JM. Absence of evidence is not evidence of absence. *BMJ* 1995;**311**(7003):486. [PUBMED: 7647644]

Altman 2006
Altman DG, Royston P. The cost of dichotomising continuous variables. *BMJ* 2006;**332**(7549):1080. [PUBMED: 16675816]

Austin 2004
Austin PC, Brunner LJ. Inflation of the type I error rate when a continuous confounding variable is categorized in logistic regression analyses. *Statistics in Medicine* 2004;**23** (7):1159–78. [PUBMED: 15057884]

Bain 2012
Bain BJ. The peripheral blood smear. In: Goldman L, Schafer AS editor(s). *Goldman's Cecil Medicine*. 24th Edition. Philadelphia, PA: Elsevier & Saunders, 2012: 1024–31. [: ISBN 978–0–8089–2437–1]

Balshem 2011
Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. *Journal of Clinical Epidemiology* 2011; **64**(4):401–6. [PUBMED: 21208779]

Barbui 2010
Barbui T, Carobbio A, Cervantes F, Vannucchi AM, Guglielmelli P, Antonioli E, et al. Thrombosis in primary myelofibrosis: incidence and risk factors. *Blood* 2010;**115** (4):778–82. [PUBMED: 19965680]

Barosi 2008
Barosi G, Mesa RA, Thiele J, Cervantes F, Campbell PJ, Verstovsek S, et al. Proposed criteria for the diagnosis of

post-polycythemia vera and post-essential thrombocythemia myelofibrosis: a consensus statement from the International Working Group for Myelofibrosis Research and Treatment. *Leukemia* 2008;**22**(2):437–8. [PUBMED: 17728787]

Barosi 2011a

Barosi G, Cervantes F, Ben-Yehuda D, Panagiotidis P, Perez JR, Orlando-Harper ND, et al. A ruxolitinib individual supply program for patients with primary myelofibrosis, post-polycythemia vera myelofibrosis, or post-essential thrombocythemia myelofibrosis. Blood (ASH Annual Meeting Abstracts). 2011; Vol. 118:5170.

Barosi 2011b

Barosi G, Rosti V, Vannucchi AM. Therapeutic approaches in myelofibrosis. *Expert Opinion on Pharmacotherapy* 2011; **12**(10):1597–611. [PUBMED: 21457082]

Basch 2012

Basch E, Aronson N, Berg A, Flum D, Gabriel S, Goodman SN, et al. Methodological standards and patient-centeredness in comparative effectiveness research: the PCORI perspective. *JAMA* 2012;**307**(15):1636–40. [PUBMED: 22511692]

Begna 2011

Begna KH, Mesa RA, Pardanani A, Hogan WJ, Litzow MR, McClure RF, et al. A phase-2 trial of low-dose pomalidomide in myelofibrosis. *Leukemia* 2011;**25**(2): 301–4. [PUBMED: 21052089]

Brok 2009

Brok J, Thorlund K, Wetterslev J, Gluud C. Apparently conclusive meta-analyses may be inconclusive - Trial sequential analysis adjustment of random error risk due to repetitive testing of accumulating data in apparently conclusive neonatal meta-analyses. *International Journal of Epidemiology* 2009;**38**(1):287–98. [PUBMED: 18824466]

Button 2013

Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, et al. Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience* 2013;**14**(5):365–76. [DOI: 23571845]

Calvert 2013

Calvert M, Blazeby J, Altman DG, Revicki DA, Moher D, Brundage MD. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. *JAMA* 2013;**309**(8):814–22. [PUBMED: 23443445]

Cases 2003

Cases A. Darbepoetin alfa: a novel erythropoiesis-stimulating protein. *Drugs of Today* 2003;**39**(7):477–95. [PUBMED: 12973399]

Cervantes 2009

Cervantes F, Dupriez B, Pereira A, Passamonti F, Reilly JT, Morra E, et al. New prognostic scoring system for primary myelofibrosis based on a study of the International Working Group for Myelofibrosis Research and Treatment. *Blood* 2009;**113**(13):2895–901. [PUBMED: 18988864]

Cervantes 2011

Cervantes F, Pereira A. Advances in the understanding and management of primary myelofibrosis. *Current Opinion in Oncology* 2011;**23**(6):665–71. [PUBMED: 21892083]

Chan 2013a

Chan AW, Tetzlaff JM, Altman DG, Dickersin K, Moher D. SPIRIT 2013: new guidance for content of clinical trial protocols. *Lancet* 2013;**381**(9861):91–2. [PUBMED: 23305999]

Chan 2013b

Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Annals of Internal Medicine* 2013;**158**(3):200–7. [PUBMED: 23295957]

Chan 2013c

Chan AW, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ* 2013;**346**: e7586. [PUBMED: 23303884]

CMA 2005

Biostat, Inc. Comprehensive Meta-Analysis. 2.0. Englewood, NJ: Biostat, Inc, 2005.

CTU 2011

Copenhagen Trial Unit. TSA - Trial Sequential Analysis. http://ctu.dk/tsa/ (accessed 29 September 2011). Copenhagen: Central Trial Unit, 2011.

Cumsille 2000

Cumsille F, Bangdiwala SI. [Categorizing variables in the statistical analysis of data: consequences for interpreting the results] [Categorización de variables en el análisis estadístico de datos: consecuencias sobre la interpretación de resultados]. *Revista Panamericana de Salud Pública* 2000; **8**(5):348–54. [PUBMED: 11190972]

Deeks 2011

Deeks JJ, Higgins JPT, Altman DG. Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration. Available from www.cochrane-handbook.org.

Deisseroth 2012

Deisseroth A, Kaminskas E, Grillo J, Chen W, Saber H, Lu HL, et al. U.S. Food and Drug Administration approval: Ruxolitinib for the treatment of patients with intermediate and high-risk myelofibrosis. *Clinical Cancer Research* 2012; **18**(12):3212–7. [PUBMED: 22544377]

Djulbegovic 2013

Djulbegovic B, Kumar A, Miladinovic B, Reljic T, Galeb S, Mhaskar A, et al. Treatment success in cancer: industry compared to publicly sponsored randomized controlled trials. *PLoS One* 2013;**8**(3):e58711. [PUBMED: 23555593]

Donnelly 2001
Donnelly S. Why is erythropoietin made in the kidney? The kidney functions as a critmeter. *American Journal of Kidney Diseases* 2001;**38**(2):415–25. [PUBMED: 11479173]

Doucet 2008
Doucet M, Sismondo S. Evaluating solutions to sponsorship bias. *Journal of Medical Ethics* 2008;**34**(8):627–30. [PUBMED: 18667655]

FDA 2007
Food, Drug Administration. Guidance for industry. Clinical trial endpoints for the approval of cancer drugs and biologics. 2007. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071590.pdf (accessed 22 March 2012).

Fedorov 2009
Fedorov V, Mannino F, Zhang R. Consequences of dichotomization. *Pharmaceutical Statistics* 2009;**8**(1): 50–61. [PUBMED: 18389492]

Fermi Paradox
Fermi Paradox. www.crystalinks.com/fermiparadox.html (accessed 12 April 2013).

Finazzi 2008
Finazzi G, Xu M, Barbui T, Hoffman R. Essential thrombocythemia. In: Hoffman, Ronald editor(s). *Hoffman: Hematology: Basic Principles and Practice*. 5th Edition. Philadelphia: Churchill Livingstone Elsevier, 2008:1149–66. [: ISBN 978–0–443–06715–0]

Freiman 1978
Freiman JA, Chalmers TC, Smith H Jr, Kuebler RR. The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial. Survey of 71 "negative" trials. *New England Journal of Medicine* 1978;**299**(13):690–4. [PUBMED: 355881]

Gabriel 2012
Gabriel SE, Normand SL. Getting the methods right--the foundation of patient-centered outcomes research. *New England Journal of Medicine* 2012;**367**(9):787–90. [PUBMED: 22830434]

Godlee 2012
Godlee F. Outcomes that matter to patients. *BMJ* 2012; **344**:e318. [DOI: http://dx.doi.org/10.1136/bmj.e318]

Golder 2008
Golder S, Loke YK. Is there evidence for biased reporting of published adverse effects data in pharmaceutical industry-funded studies?. *British Journal of Clinical Pharmacology* 2008;**66**(6):767–73. [PUBMED: 18754841]

GRADEpro 2014
McMaster University. GRADEpro. (19 March 2015). McMaster University, 2014.

Green 2002
Green SB. Design of randomized trials. *Epidemiologic Reviews* 2002;**24**(1):4–11. [PUBMED: 12119855]

Greer 2009
Greer JP, Foerster J, Rodgers GM, Paraskevas F, Glader B, Arber DA, et al. *Wintrobe's Clinical Hematology*. 12th Edition. Philadelphia, PA: Lippincott Williams & Wilkins, 2009. [: ISBN 978–0–7817–6507–7]

Guyatt 2011a
Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction - GRADE evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology* 2011;**64**(4):383–94. [PUBMED: 21195583]

Guyatt 2011b
Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *Journal of Clinical Epidemiology* 2011;**64**(4):395–400. [PUBMED: 21194891]

Guyatt 2011c
Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines: 6. Rating the quality of evidence - imprecision. *Journal of Clinical Epidemiology* 2011;**64**(12):1283–93. [PUBMED: 21839614]

Guyatt 2011d
Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 7. Rating the quality of evidence - inconsistency. *Journal of Clinical Epidemiology* 2011;**64**(12):1294–302. [PUBMED: 21803546]

Guyatt 2011e
Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 8. Rating the quality of evidence - indirectness. *Journal of Clinical Epidemiology* 2011;**64**(12):1303–10. [PUBMED: 21802903]

Guyatt 2011f
Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, et al. GRADE guidelines: 5. Rating the quality of evidence - publication bias. *Journal of Clinical Epidemiology* 2011;**64**(12):1277–82. [PUBMED: 21802904]

Guyatt 2011g
Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *Journal of Clinical Epidemiology* 2011;**64**(4):380–2. [PUBMED: 21185693]

Guyatt 2011h
Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, et al. GRADE guidelines: 9. Rating up the quality of evidence. *Journal of Clinical Epidemiology* 2011;**64**(12):1311–6. [PUBMED: 21802902]

Guyatt 2011i
Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. GRADE guidelines: 4. Rating the quality of evidence - study limitations (risk of bias). *Journal of Clinical Epidemiology* 2011;**64**(4):407–15. [PUBMED: 21247734]

Guyatt 2013
Guyatt GH, Oxman AD, Santesso N, Helfand M, Vist G, Kunz R, et al. GRADE guidelines 12. Preparing Summary of findings tables - binary outcomes. *Journal of Clinical Epidemiology* 2013;**66**(2):158–72. [PUBMED: 22609141]

Hemming 2009
Hemming K, Hutton JL. Intrapartum amnioinfusion for meconium-stained amniotic fluid: evidence for small study effect bias?. *BJOG* 2009;**116**(1):128–9. [PUBMED: 19087088]

Hewitt 2010
Hewitt CE, Kumaravel B, Dumville JC, Torgerson DJ, Trial attrition study group. Assessing the impact of attrition in randomized controlled trials. *Journal of Clinical Epidemiology* 2010;**63**(11):1264–70. [PUBMED: 20573482]

Higgins 2011a
Higgins JPT, Altman DG, Sterne JAC. Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Higgins 2011b
Higgins JPT, Deeks JJ. Chapter 7: Selecting studies and collecting data. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Higgins 2011c
Higgins JPT, Deeks JJ, Altman DG. Chapter 16: Special topics in statistics. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Higgins 2011d
Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Higgins 2011e
Higgins JP, Whitehead A, Simmonds M. Sequential methods for random-effects meta-analysis. *Statistics in Medicine* 2011;**30**(9):903–21. [PUBMED: 21472757]

Hoffman 2008
Hoffman R, Xu M, Barosi G. Primary myelofibrosis. In: Hoffman, Ronald editor(s). *Hoffman: Hematology: Basic Principles and Practice*. 5th Edition. Philadelphia: Churchill Livingstone Elsevier, 2008:1125–47. [: ISBN 978–0–443–06715–0]

Hussein 2009
Hussein K, Van Dyke DL, Tefferi A. Conventional cytogenetics in myelofibrosis: literature review and discussion. *European Journal of Haematology* 2009;**82**(5): 329–38. [PUBMED: 19141119]

Ioannidis 2004
Ioannidis JP, Evans SJ, Gøtzsche PC, O'Neill RT, Altman DG, Schulz K, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Annals of Internal Medicine* 2004;**141**(10): 781–8. [PUBMED: 15545678]

Ioannidis 2005
Ioannidis JP. Why most published research findings are false. *PLoS Medicine* 2005;**2**(8):e124. [PUBMED: 16060722]

Ioannidis 2008
Ioannidis JP. Why most discovered true associations are inflated. *Epidemiology* 2008;**19**(5):640–8. [PUBMED: 18633328]

Ioannidis 2010
Ioannidis JP. The art of getting it wrong. *Research Synthesis Methods* 2010;**1**(3-4):169–84.

Jørgensen 2008
Jørgensen AW, Maric KL, Tendal B, Faurschou A, Gøtzsche PC. Industry-supported meta-analyses compared with meta-analyses with non-profit or no support: differences in methodological quality and conclusions. *BMC Medical Research Methodology* 2008;**8**:60. [PUBMED: 18782430]

Kirby 2002
Kirby A, Gebski V, Keech AC. Determining the sample size in a clinical trial. *Medical Journal of Australia* 2002;**177**(5): 256–7. [PUBMED: 12197821]

Lan 1983
Lan GKK, Demets DL. Discrete sequential boundaries for clinical trials. *Biometrika* 1983;**70**(3):659–63.

Lauer 2012
Lauer MS. From hot hands to declining effects: the risks of small numbers. *Journal of the American College of Cardiology* 2012;**60**(1):72–4. [PUBMED: 22742403]

Lefevbre 2011
Lefebvre C, Manheimer E, Glanville J. Chapter 6: Searching for studies. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Lexchin 2003
Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ* 2003;**326**(7400):1167–70. [PUBMED: 12775614]

Lundh 2012
Lundh A, Sismondo S, Lexchin J, Busuioc OA, Bero L. Industry sponsorship and research outcome. *Cochrane Database of Systematic Reviews* 2012, Issue 12. [DOI: 10.1002/14651858.MR000033.pub2]

MacCallum 2002
MacCallum RC, Zhang S, Preacher KJ, Rucker DD. On the practice of dichotomization of quantitative variables. *Psychological Methods* 2002;**7**(1):19–40.

Mascarenhas 2012
Mascarenhas J, Hoffman R. Ruxolitinib: the first FDA approved therapy for the treatment of myelofibrosis. *Clinical Cancer Research* 2012;**18**(11):3008–14. [PUBMED: 22474318]

Mascarenhas 2013
Mascarenhas J, Hoffman R. A comprehensive review and analysis of the effect of ruxolitinib therapy on the survival of patients with myelofibrosis. *Blood* 2013;**121**(24):4832–7. [PUBMED: 23570800]

Mesa 2009a
Mesa R, Gale RP. Hypothesis: how do JAK2-inhibitors work in myelofibrosis. *Leukemia Research* 2009;**33**(9): 1156–7. [PUBMED: 19450878]

Mesa 2009b
Mesa RA, Schwager S, Radia D, Cheville A, Hussein K, Niblack J, et al. The Myelofibrosis Symptom Assessment Form (MFSAF): an evidence-based brief inventory to measure quality of life and symptomatic response to treatment in myelofibrosis. *Leukemia Research* 2009;**33**(9): 1199–203. [PUBMED: 19250674]

Mesa 2010a
Mesa RA, Pardanani AD, Hussein K, Wu W, Schwager S, Litzow MR, et al. Phase1/-2 study of Pomalidomide in myelofibrosis. *American Journal of Hematology* 2010;**85**(2): 129–30. [PUBMED: 20052748]

Mesa 2010b
Mesa RA, Yao X, Cripe LD, Li CY, Litzow M, Paietta E, et al. Lenalidomide and prednisone for myelofibrosis: Eastern Cooperative Oncology Group (ECOG) phase 2 trial E4903. *Blood* 2010;**116**(22):4436–8. [PUBMED: 20651074]

Mesa 2012a
Mesa RA. The evolving treatment paradigm in myelofibrosis. *Leukemia & Lymphoma* 2013;**54**(2):242–51. [PUBMED: 22793267]

Mesa 2012b
Mesa RA, Yasothan U, Kirkpatrick P. Ruxolitinib. *Nature Reviews. Drug Discovery* 2012;**11**(2):103–4. [PUBMED: 22293561]

Moher 1998
Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? . *Lancet* 1998;**352**(9128):609–13. [PUBMED: 9746022]

Moher 2009
Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Journal of Clinical Epidemiology* 2009;**62**(10):1006–12. [PUBMED: 19631508]

Moher 2010
Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;**340**:c869. [PUBMED: 20332511]

Nebeker 2004
Nebeker JR, Barach P, Samore MH. Clarifying adverse drug events: a clinician's guide to terminology, documentation, and reporting. *Annals of Internal Medicine* 2004;**140**(10): 795–801. [PUBMED: 15148066]

Oken 1982
Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *American Journal of Clinical Oncology* 1982;**5**(6):649–55. [PUBMED: 7165009]

Ostojic 2011a
Ostojic A, Vrhovac R, Verstovsek S. Ruxolitinib: a new JAK1/2 inhibitor that offers promising options for treatment of myelofibrosis. *Future Oncology* 2011;**7**(9): 1035–43. [PUBMED: 21919691]

Ostojic 2011b
Ostojic A, Vrhovac R, Verstovsek S. Ruxolitinib for the treatment of myelofibrosis. *Drugs of Today* 2011;**47**(11): 817–27. [PUBMED: 22146225]

Ostojic 2012
Ostojic A, Vrhovac R, Verstovsek S. Ruxolitinib for the treatment of myelofibrosis: its clinical potential. *Therapeutics and Clinical Risk Management* 2012;**8**:95–103. [PUBMED: 22399854]

Pardanani 2011b
Pardanani A, Tefferi A. Targeting myeloproliferative neoplasms with JAK inhibitors. *Current Opinion in Hematology* 2011;**18**(2):105–10. [PUBMED: 21245760]

Pardanani 2012
Pardanani A. Ruxolitinib for myelofibrosis therapy: current context, pros and cons. *Leukemia* 2012;**26**(7):1449–51. [PUBMED: 22285996]

Passamonti 2010
Passamonti F, Cervantes F, Vannucchi AM, Morra E, Rumi E, Pereira A, et al. A dynamic prognostic model to predict survival in primary myelofibrosis: a study by the IWG-MRT (International Working Group for Myeloproliferative Neoplasms Research and Treatment). *Blood* 2010;**115**(9): 1703–8. [PUBMED: 20008785]

Passamonti 2012
Passamonti F, Maffioli M, Caramazza D. New generation small-molecule inhibitors in myeloproliferative neoplasms. *Current Opinion in Hematology* 2012;**19**(2):117–23. [PUBMED: 22227528]

Pastore 2012
Pastore A, Temussi PA. The two faces of Janus: functional interactions and protein aggregation. *Current Opinion in Structural Biology* 2012;**22**(1):30–7. [PUBMED: 22155180]

PCORI 2012
Patient-Centered Outcomes Research Institute (PCORI). Preliminary draft methodology report: "Our questions, our decisions: Standards for patient-centered outcomes research", 2012. http://www.pcori.org/assets/Preliminary-Draft-Methodology-Report.pdf (accessed 1 April 2013): 1–61.

Peacock 2012
Peacock JL, Sauzet O, Ewings SM, Kerry SM. Dichotomising continuous data while retaining statistical

power using a distributional approach. *Statistics in Medicine* 2012;**31**(26):3089–103. [PUBMED: 22865598]

Pereira 2011
Pereira TV, Ioannidis JP. Statistically significant meta-analyses of clinical trials have modest credibility and inflated effects. *Journal of Clinical Epidemiology* 2011;**64**(10): 1060–9. [PUBMED: 21454050]

Pogue 1997
Pogue JM, Yusuf S. Cumulating evidence from randomized trials: utilizing sequential monitoring boundaries for cumulative meta-analysis. *Controlled Clinical Trials* 1997; **18**(6):580-93, 661-6. [PUBMED: 9408720]

Pogue 1998
Pogue J, Yusuf S. Overcoming the limitations of current meta-analysis of randomised controlled trials. *Lancet* 1998; **351**(9095):47–52. [PUBMED: 9433436]

Porta 2008
Porta M. *A Dictionary of Epidemiology*. 5th Edition. New York: Oxford University Press, 2008.

Prentice 2009
Prentice RL. Surrogate and mediating endpoints: current status and future directions. *Journal of the National Cancer Institute* 2009;**101**(4):216–7. [PUBMED: 19211455]

Quintás-Cardama 2009
Quintás-Cardama A, Kantarjian HM, Manshouri T, Thomas D, Cortes J, Ravandi F, et al. Lenalidomide plus prednisone results in durable clinical, histopathologic, and molecular responses in patients with myelofibrosis. *Journal of Clinical Oncology* 2009;**27**(28):4760–6. [PUBMED: 19720904]

Qureshi 2011
Qureshi M, Harrison C. Management of myelofibrosis - where next?. *Expert Opinion on Pharmacotherapy* 2011;**12** (10):1453–5. [PUBMED: 21651445]

Randhawa 2012
Randhawa J, Ostojic A, Vrhovac R, Atallah E, Verstovsek S. Splenomegaly in myelofibrosis-new options for therapy and the therapeutic potential of Janus kinase 2 inhibitors. *Journal of Hematology & Oncology* 2012;**5**:43. [PUBMED: 22852872]

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

Savović 2012
Savović J, Jones HE, Altman DG, Harris RJ, Jüni P, Pildal J, et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. *Annals of Internal Medicine* 2012;**157**(6):429–38. [PUBMED: 22945832]

Schott 2010a
Schott G, Pachl H, Limbach U, Gundert-Remy U, Lieb K, Ludwig WD. The financing of drug trials by pharmaceutical companies and its consequences: part 2: a qualitative, systematic review of the literature on possible influences on authorship, access to trial data, and trial registration and publication. *Deutsches Arzteblatt International* 2010;**107** (17):295–301. [PUBMED: 20490338]

Schott 2010b
Schott G, Pachl H, Limbach U, Gundert-Remy U, Ludwig WD, Lieb K. The financing of drug trials by pharmaceutical companies and its consequences. Part 1: a qualitative, systematic review of the literature on possible influences on the findings, protocols, and quality of drug trials. *Deutsches Arzteblatt International* 2010;**107**(16):279–85. [PUBMED: 20467553]

Schulz 1995
Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;**273**(5):408–12. [PUBMED: 7823387]

Seavey 2012
Seavey MM, Dobrzanski P. The many faces of Janus kinase. *Biochemical Pharmacology* 2012;**83**(9):1136–45. [PUBMED: 22209716]

Shahani 2009
Shahani S, Braga-Basaria M, Maggio M, Basaria S. Androgens and erythropoiesis: past and present. *Journal of Endocrinological Investigation* 2009;**32**(8):704–16. [PUBMED: 19494706]

Stein 2011
Stein BL, Crispino JD, Moliterno AR. Janus kinase inhibitors: an update on the progress and promise of targeted therapy in the myeloproliferative neoplasms. *Current Opinion in Oncology* 2011;**23**(6):609–16. [PUBMED: 21993415]

Sterne 2011
Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011;**343**:d4002. [PUBMED: 21784880]

Streiner 2002
Streiner DL. Breaking up is hard to do: The heartbreak of dichotomizing continuous data. *Canadian Journal of Psychiatry* 2002;**47**(3):262–6.

Tefferi 2009
Tefferi A, Verstovsek S, Barosi G, Passamonti F, Roboz GJ, Gisslinger H, et al. Pomalidomide is active in the treatment of anemia associated with myelofibrosis. *Journal of Clinical Oncology* 2009;**27**(27):4563–9. [PUBMED: 19652059]

Tefferi 2011a
Tefferi A. How I treat myelofibrosis. *Blood* 2011;**117**(13): 3494–504. [PUBMED: 21200024]

Tefferi 2011b
Tefferi A. Primary myelofibrosis: 2012 update on diagnosis, risk stratification, and management. *American Journal of Hematology* 2011;**86**(12):1017–26. [PUBMED: 22086865]

Tefferi 2011c
Tefferi A, Litzow MR, Pardanani A. Long-term outcome of treatment with ruxolitinib in myelofibrosis. *New England Journal of Medicine* 2011;**365**(15):1455–7. [PUBMED: 21995409]

Tefferi 2011d
Tefferi A, Pardanani A. Serious adverse events during ruxolitinib treatment discontinuation in patients with myelofibrosis. *Mayo Clinic Proceedings* 2011;**86**(12): 1188–91. [PUBMED: 22034658]

Tefferi 2011e
Tefferi A, Pardanani A. JAK inhibitors in myeloproliferative neoplasms: rationale, current data and perspective. *Blood Reviews* 2011;**25**(5):229–37. [PUBMED: 21742423]

Tefferi 2012
Tefferi A. JAK inhibitors for myeloproliferative neoplasms: clarifying facts from myths. *Blood* 2012;**119**(12):2721–30. [PUBMED: 22279053]

Thiele 2005
Thiele J, Kvasnicka HM, Facchetti F, Franco V, van der Walt J, Orazi A. European consensus on grading bone marrow fibrosis and assessment of cellularity. *Haematologica* 2005; **90**(8):1128–32. [PUBMED: 16079113]

Thiele 2007
Thiele J, Kvasnicka HM. Myelofibrosis - what's in a name? Consensus on definition and EUMNET grading. *Pathobiology* 2007;**74**(2):89–96. [PUBMED: 17587880]

Thompson 2005
Thompson JE. JAK protein kinase inhibitors. *Drug News & Perspectives* 2005;**18**(5):305–10. [PUBMED: 16193102]

Thorlund 2009
Thorlund K, Devereaux PJ, Wetterslev J, Guyatt G, Ioannidis JP, Thabane L, et al. Can trial sequential monitoring boundaries reduce spurious inferences from meta-analyses?. *International Journal of Epidemiology* 2009; **38**(1):276–86. [PUBMED: 18824467]

Thorlund 2011a
Thorlund K, Engstrøm J, Wetterslev J, Brok J, Imberger G, Gluud C. User manual for Trial Sequential Analysis (TSA). 2011. http://ctu.dk/tsa/files/tsa˙manual.pdf (accessed 30 April 2012).

Thorlund 2011b
Thorlund K, Imberger G, Walsh M, Chu R, Gluud C, Wetterslev J, et al. The number of patients and events required to limit the risk of overestimation of intervention effects in meta-analysis--a simulation study. *PLoS One* 2011;**6**(10):e25491. [PUBMED: 22028777]

Twombly 2007
Twombly R. Small study on industry trial sponsorship leads to big questions about quality and bias. Journal of the National Cancer Institute 2007; Vol. 99, issue 13:988–90. [PUBMED: 17596566]

Vannucchi 2009
Vannucchi AM. How do JAK2-inhibitors work in myelofibrosis: an alternative hypothesis. *Leukemia Research* 2009;**33**(12):1581–3. [PUBMED: 19573914]

Wen 2011
Wen Q, Goldenson B, Crispino JD. Normal and malignant megakaryopoiesis. *Expert Reviews in Molecular Medicine* 2011;**13**:e32. [PUBMED: 22018018]

Wetterslev 2008
Wetterslev J, Thorlund K, Brok J, Gluud C. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. *Journal of Clinical Epidemiology* 2008;**61**(1):64–75. [PUBMED: 18083463]

Wood 2008
Wood L, Egger M, Gluud LL, Schulz KF, Jüni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008;**336** (7644):601–5. [PUBMED: 18316340]

Yudkin 2011
Yudkin JS, Lipska KJ, Montori VM. The idolatry of the surrogate. *BMJ* 2011;**343**:d7995. [PUBMED: 22205706]

References to other published versions of this review

Martí-Carvajal 2013
Martí-Carvajal AJ, Cardona AF, Anand V, Solà I. Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis (Protocol). *Cochrane Database of Systematic Reviews* 2013, Issue 1. [DOI: 10.1002/ 14651858.CD010298]

* *Indicates the major publication for the study*

CHARACTERISTICS OF STUDIES

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

COMFORT-I 2012

Methods	Design: Parallel. Number of arms: Two. Countries: 3 (89 sites in the USA, Canada, and Australia). Phase: 3. Follow-up (median, weeks): 32. Screening: Up to 35 days. Data were gathered from protocol (Page 37/180) On-treatment phase: Participation may continue for subjects receiving benefit on study drug or open label INCB018424 until the later of: marketing approval or when the last randomized subject remaining in the study has completed week 144 (36 months). Data were gathered from protocol (Page 37/180) Follow-up: 28 ± 7 days following the last dose of study drug or open label INCB018424. Data were gathered from protocol (Page 37/180) Recruitment period: September 2009 to April 2010.
Participants	Randomized: • Ruxolitinib: 155. • Placebo: 154. Age (median, range): • Ruxolitinib: 66 (43 to 91). • Placebo: 70 (40 to 86). Gender (% of males): • Ruxolitinib: 51. • Placebo: 57.1. MF subtype: • PMF: ruxolitinib (45.2%) and placebo (54.5%). • Post-polycythemia vera MF: ruxolitinib (32.3%) and placebo (30.5%). • Post-essential thrombocythemia MF: ruxolitinib (22.6%) and placebo (14.3%). IPSS risk status: • High: ruxolitinib (58.1%) and placebo (64.3%). • Intermediate 2: ruxolitinib (41.3%) and placebo (35.1%). Inclusion criteria: • Age 18 years or older. • PMF, post-polycythemia vera-MF or post-essential thrombocythemia-MF according to the 2008 WHO criteria. • MF requiring therapy must be classified as high risk or intermediate risk level 2 according to the prognostic factors defined by the International Working Group. • Eastern Cooperative Oncology Group performance status of 0, 1, 2, or 3; International Prognostic Scoring System score of 2 (intermediate-2 risk) or 3 or more (high risk). • No previously received treatment with a Janus kinase inhibitor. • Less than 10% peripheral-blood blasts, an absolute peripheral-blood $CD34^+$ cell count of more than 20×10^6/L, a platelet count of 100×10^9/L or more, and palpable splenomegaly (≥ 5 cm below the left costal margin). Exclusion criteria:

COMFORT-I 2012 (*Continued*)

	• Life expectancy of < 6 months. • Inadequate bone marrow reserve as demonstrated by specific clinical laboratory counts. • Inadequate liver or renal function. • Clinically significant bacterial, fungal, parasitic or viral infection which require therapy. • An active malignancy over the previous 5 years except specific skin cancers. • Severe cardiac conditions. • Splenic irradiation within 12 months.
Interventions	1. Oral ruxolitinib phosphate tablets: 15 mg twice daily (platelet count of 100×10^9 to 200×10^9/L) or 20 mg twice daily (count that exceeded 200×10^9/L The dose was adjusted for lack of efficacy or excess toxicity as specified in the protocol. 2. Matching placebo: 5 mg (tablets).
Outcomes	Primary outcomes: • Proportion of patients with a reduction of 35% or more in spleen volume from baseline to week 24, measured by means of MRI or CT. Secondary outcomes: • Duration of maintenance of a ≥ 35% reduction from baseline in spleen volume. • The proportion of patients with a reduction in the total symptom score of 50% or more from baseline to week 24, as assessed with the modified MFSAF. • The change in the total symptom score from baseline to week 24. • Overall survival. • Leukemia free survival. • Safety and tolerability.
Notes	Identifier number: NCT00952289 (ClinicalTrials.gov). Official title: A Randomized, Double-blind, Placebo-controlled Study of the JAK Inhibitor INCB018424 Tablets Administered Orally to Subjects With Primary Myelofibrosis, Post-Polycythemia Vera Myelofibrosis or Post-Essential Thrombocythemia Myelofibrosis Conducted date: September 2009 through April 2010. A priori sample estimation: Yes (Page 141 of protocol). Sponsor: Incyte Corporation. Role of sponsor: Data were collected by the academic investigators and analyzed by the sponsor of the study, Incyte. The sponsor, in collaboration with the academic investigators, interpreted the data. The first author and an author who was an employee of the sponsor wrote the initial draft of the manuscript, with assistance from a medical writer who was paid by the sponsor (Page 800) Disclosure statement: Declared.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Randomization will occur centrally, using an Interactive Voice Response System (IVRS)."

COMFORT-I 2012 *(Continued)*

		Comment: Mentioned in protocol.
Allocation concealment (selection bias)	Low risk	"Randomization will occur centrally by an Interactive Voice Response System (IVRS). Site staff will contact an IVRS to obtain the subject study drug assignment. The investigator or designee will select the assigned bottles from their stock that corresponds to the number provided by IVRS and dispense the medication." (Page 117 of protocol)
Blinding of participants and personnel (performance bias) All outcomes	Low risk	"Kits (blinded) of 15, 20 and 25 mg BID doses will be provided" (Protocol Page 14 of 180) We contacted Dr. Srdan Verstovsek (Corresponding author, 16 January 2013) who sent an additional information "...once the blinded part of the study finished, all patients were unblinded and those on placebo were given an option of taking ruxolitinib" Comment: patients, outcome assessors, company, and Biostats personnel were blinded (Dr. Srdan Verstovsek, corresponding author, 28 January 2013)
Blinding of outcome assessment (detection bias) Overall survival	Low risk	Patients, outcome assessors, company, and Biostats personnel were blinded (Dr. Srdan Verstovsek, corresponding author, 28 January 2013)
Blinding of outcome assessment (detection bias) Progression-free survival	Unclear risk	Insufficient information to permit judgment of 'low risk' or 'high risk'
Blinding of outcome assessment (detection bias) Safety	Low risk	Patients, outcome assessors, company, and Biostats personnel were blinded (Dr. Srdan Verstovsek, corresponding author, 28 January 2013)
Blinding of outcome assessment (detection bias) Health-related quality of life	Low risk	Blinding of participants and key study personnel ensured, and unlikely that the blinding could have been broken Comment: patients, outcome assessors, company, and Biostats personnel were blinded (Dr. Srdan Verstovsek, corresponding author, 28 January 2013)

COMFORT-I 2012 *(Continued)*

Blinding of outcome assessment (detection bias) Reduction in spleen size.	Low risk	Patients, outcome assessors, company, and Biostats personnel were blinded (Dr. Srdan Verstovsek, corresponding author, 28 January 2013)
Blinding of outcome assessment (detection bias) Leukemia-free survival	Unclear risk	Insufficient information to permit judgment of 'low risk' or 'high risk'
Blinding of outcome assessment (detection bias) Anemia response	Unclear risk	Insufficient information to permit judgment of 'low risk' or 'high risk'
Incomplete outcome data (attrition bias) All outcomes	High risk	The primary outcome was reduction in spleen size. This end point was reported as follows: 1. Mean reduction in spleen volume at 24 weeks: 79.2% (245/309) of the allocated participants to both comparison groups. Ruxolutinib group (89.6% (139/155)) versus placebo group (68.8% (106/154)). 2. Imabalance between comparison group: 20.8%.
Selective reporting (reporting bias)	Low risk	Primary and secondary end points were measured and reported in the trial results as mentioned in protocol
Other bias	High risk	Bias in the presentation of data: Trial authors did not reported complete data regarding reduction in spleen volume of spleen. They only reported mean and median nor standard deviation, standard error, or 95% CI. Therefore, it was not able to pool COMFORT I data with COMFORT II data regarding that end point. We contacted to the COMFORT I corresponding author (E-mail: Friday, April 12, 2013 5:54 AM). We received his reply on 12 April 2013 08:48 Industry bias: The trial was sponsored by Incyte Corporation. Dr. Verstovsek reports receiving grant support through his institution from Incyte, Exelixis, Celgene, NS Pharma, Infinity Pharmaceuticals, SBIO, Lilly Oncology, AstraZeneca, Geron, Bristol- My-

COMFORT-I 2012 *(Continued)*

		ers Squibb, YM BioSciences, Gilead, and Roche; Dr. Levy, being an employee of Incyte and receiving stock options as part of his compensation; and Dr. Kantarjian, receiving grant support through his institution from Incyte (Page 807)

COMFORT-II 2012

Methods	Design: Parallel. Number of arms: Two Country: 9 European countries (Austria, Belgium, France, Germany, Italy, Netherlands, Spain, Sweden, and UK) in 56 sites Phase: 3 Follow-up (median, weeks): 61.1 Recruitment period: 1 July 2009 to 22 January 2010.
Participants	Randomized: • Ruxolitinib: 146. • best available therapy: 73. Age (median, range): • Ruxolitinib: 67 (35 to 83). • best available therapy: 66 (35 to 85). Gender (% of males): • Ruxolitinib: 57. • Best available therapy: 58. JAK-2 V617F mutation status at screening (%): 1. Ruxolitinib: • Positive: 75. • Negative: 24. • Unknown: 1. 2. Best available therapy: • Positive: 67. • Negative: 27. • Unknown: 6. MF subtype: • Primary: ruxolitinib (53%), placebo (53%). • Post-polycythemia vera: ruxolitinib (33%), placebo (27%). • Post-essential thrombocythemia: ruxolitinib (14%), placebo (19%). Risk category: • Intermediate-2: ruxolitinib (40%), placebo (40%). • High: ruxolitinib (60%), placebo (59%). • Not determined: ruxolitinib (0%), placebo (1%). Inclusion criteria: • 18 years of age or older. • PMF, post-polycythemia vera MF or post- essential thrombocythemia MF according to the 2008 WHO criteria. • MF requiring therapy must be classified as high risk or intermediate risk level 2

	according to the prognostic factors defined by the International Working Group. • Eastern Cooperative Oncology Group performance status of 0, 1, 2, or 3. • Peripheral blood blast count of < 10%. • No previously received treatment with a Janus kinase inhibitor. Exclusion criteria: • Life expectancy of < 6 months. • Pregnant or breastfeeding. • Inadequate bone marrow reserve. • History of platelet counts < 50,000/μL or absolute neutrophil count < 500/μL except during treatment with cytotoxic therapy for any other reason. • Inadequate liver or renal function. • Clinically significant infection which requires therapy. • Option of stem cell transplantation. • History of malignancy in past 5 years except for treated, early-stage squamous or basal cell carcinoma of the skin. • Cardiac disease which may jeopardize the safety of the subject or the compliance with the protocol. • Uncontrolled or unstable angina. • Ongoing or recent treatment with another investigational medication. • Dose or dose-regimen of any therapies used to treat MF has been modified at any time from 2 weeks prior to the start of screening through the beginning of baseline evaluations. • Splenic irradiation within 12 months prior to screening. • Rapid or paroxysmal atrial fibrillation. • Treatment with hematopoietic growth factor receptor agonists at any time within 2 weeks prior to screening or 4 weeks prior to baseline.
Interventions	Ruxolitinib: Starting dose of 15 mg or 20 mg twice daily were selected with starting dose based on baseline platelet count. Dose titration ranging from 5 mg twice daily to a maximum dose of 25 mg twice daily was permitted during the study based on safety and efficacy Best available therapy (oral, parenteral, or no therapy): Selected therapy included a combination of available agents to treat the disease, or its symptoms, or both, and was selected by the investigator for each subject. Therapy changed at different times during the treatment phase. No experimental agents (e.g., those not approved for the treatment of any indication) were allowed Route of administration: Oral (tablets) • Co-interventions. Restricted and prohibited therapies include systemic corticosteroid doses greater than the equivalent of 10 mg prednisolone per day, hematopoietic growth factor receptor agonists, aspirin in doses exceeding 150 mg per day, potent CYP3A4 inducers and inhibitors, etc (page 84-5 of protocol)
Outcomes	Primary outcomes: • Reduction of 35% or more in spleen volume from baseline at week 48. Spleen volume was assessed by MRI or by CT (in the case of patients who were not suitable candidates for MRI) every 12 weeks. Secondary outcomes:

COMFORT-II 2012 (*Continued*)

	• Reduction of 35% or more in spleen volume from baseline at week 24. • The length of time that a reduction in spleen volume of at least 35% was maintained. • The time to a reduction in spleen volume of 35% or more from baseline. • Progression-free survival. • Leukemia-free survival. • Overall survival. • Change in marrow histomorphologic features. • Safety (according to National Cancer Institute's Common Toxicity Criteria) and tolerability. • Quality of life.
Notes	Identifier number: NCT00934544 (ClinicalTrials.gov). Official title: A randomized study of INC424 (INCB018424) tablets compared with best available therapy in subjects with primary MF, post-polycythemia vera-MF or post-essential thrombocythemia MF Conducted date: July 1, 2009 to January 22, 2010. A priori sample estimation: Yes. (Page 100 of protocol). Sponsor: Novartis Pharmaceuticals, and designed by Incyte. Role of sponsor: Data were analyzed and interpreted by the sponsor's clinical and statistical teams in collaboration with authors who were not affiliated with the sponsor. The first author prepared the first draft of the manuscript, with assistance from a medical writer who was funded by Novartis Pharmaceuticals, and made the final decision to submit the manuscript for publication (page 789) Disclosure statement: Declared.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Randomization will occur centrally by an Interactive Voice Response System (IVRS) . Randomization will be stratified by Baseline prognostic risk level, as defined by Cervantes, et al (2009) into two strata" (Page 85 of protocol) Comment: mentioned in protocol.
Allocation concealment (selection bias)	Low risk	"Randomization will occur centrally by an Interactive Voice Response System (IVRS). ..The investigator or designee will select the assigned bottles from their stock that corresponds to the number provided by IVRS and dispense the medication. The investigator will enter the bottle numbers in the eCRF." (Page 85 of Protocol) Comment: mentioned in protocol.

Blinding of participants and personnel (performance bias) All outcomes	High risk	Open label trial (Protocol No. CINC424A2352 page 16/125).
Blinding of outcome assessment (detection bias) Overall survival	Low risk	No blinding, but we judge that the outcome measurement is not likely to be influenced by lack of blinding
Blinding of outcome assessment (detection bias) Progression-free survival	High risk	No blinding, but we judge that the outcome measurement is likely to be influenced by lack of blinding
Blinding of outcome assessment (detection bias) Safety	High risk	No blinding, but we judge that the outcome measurement is likely to be influenced by lack of blinding
Blinding of outcome assessment (detection bias) Health-related quality of life	High risk	No blinding, but we judge that the outcome measurement is likely to be influenced by lack of blinding
Blinding of outcome assessment (detection bias) Reduction in spleen size.	High risk	No blinding, but we judge that the outcome measurement is likely to be influenced by lack of blinding
Blinding of outcome assessment (detection bias) Leukemia-free survival	High risk	No blinding, but we judge that the outcome measurement is likely to be influenced by lack of blinding
Blinding of outcome assessment (detection bias) Anemia response	Unclear risk	Insufficient information to permit judgment of 'low risk' or 'high risk'
Incomplete outcome data (attrition bias) All outcomes	Low risk	"The full analysis set (FAS) consisted of all patients randomized and was analyzed following the intent-to-treat principle" (Protocol No. CINC424A2352 page 947125) "Safety set: consists of all subjects in the FAS and if randomized to the active group who have taken at least 1 dose of study medication. Subjects will be analyzed according to the treatment actually received" (Protocol No. CINC424A2352 page 947125)
Selective reporting (reporting bias)	Low risk	Trialists did not report P value of the Changes in Quality-of-Life and Symptom-Assessment Scores, According to Treatment Group for pooling data

COMFORT-II 2012 (*Continued*)

Other bias	High risk	Industry bias: The trial was sponsored by Novartis Pharmaceuticals

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

Study	Reason for exclusion
Geyer 2014	Single clinical trial.
Gisslinger 2012	Single clinical trial.
Guglielmelli 2011	Not a RCT.
le Coutre 2012	Single trial-phase 4.
Mesa 2007	Not a RCT.
Mesa 2014	Study comparing placebo with best available therapy used as control groups in COMFORT-I 2012; COMFORT-II 2012, respectively.
Pardanani 2011a	Single clinical trial.
Pardanani 2013	Not a RCT.
Santos 2010	Not a RCT.
Talpaz 2013	Single clinical trial.
Verstovsek 2010	Phase 1/2 study.
Verstovsek 2011	Comparison with historical control.
Verstovsek 2014	Phase I trial assessing XL019 a selective JAK2 inhibitor.

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

NCT01437787

Trial name or title	Phase III study of SAR302503 in intermediate-2 and high risk patients with myelofibrosis (JAKARTA)
Methods	Allocation: randomized, endpoint classification: efficacy study, intervention model: crossover assignment, masking: double blind (subject, investigator), primary purpose: treatment

Participants	Age: From 18 years or older. Gender: Both. Inclusion criteria: 1. Diagnosis of PMF or post-polycythemia vera MF or post-essential thrombocythemia MF, according to the 2008 WHO and International Working Group of Myelofibrosis Research and Treatment (IWG-MRT) criteria. 2. MF classified as high-risk or intermediate-risk level 2, as defined by modified IWG-MRT criteria. 3. Enlarged spleen, palpable at least 5 cm below costal margin. 4. At least 18 years of age. 5. Eastern Cooperative Oncology Group performance status of 0, 1, or 2 at study entry. 6. The following laboratory values within 14 days prior to the initiation of IMP or placebo: absolute neutrophil count (ANC) 1.0 x 10^9/L, platelet count 50 x 10^9/L, serum creatinine 1.5 x upper limit of normal (ULN), serum amylase and lipase 1.5 x ULN, direct bilirubin 2.0 x ULN, aspartate aminotransferase or alanine aminotransferase 3 x ULN; higher values (i.e., 5 x ULN) are allowed if clinically compatible with hepatic extramedullary hematopoiesis. Exclusion criteria: 1. Splenectomy. 2. Any chemotherapy, immunomodulatory drug therapy, anagrelide, immunosuppressive therapy, corticosteroids >10 mg/day prednisone or equivalent, or growth factor treatment, or hormones within 14 days prior to initiation of IMP or placebo; darbepoetin use within 28 days prior to initiation of IMP or placebo. 3. Patients who have had exposure to hydroxyurea in the past may be enrolled into the study as long as it has not been administered within 14 days prior to initiation of IMP or placebo. 4. Major surgery within 28 days or radiation within 6 months prior to initiation of IMP or placebo. 5. Prior treatment with a Janus Kinase 2 (JAK-2) inhibitor. The above information is not intended to contain all considerations relevant to a patient's potential participation in a clinical trial
Interventions	Drug: SAR302503 1. SAR302503 400 mg once daily X 28 days, orally, empty stomach, approximately same time each day. 2. SAR302503 500 mg once daily X 28 days, orally, empty stomach, approximately same time each day. Drug: Placebo (Placebo comparator once daily X 28 days, orally, empty stomach, approximately same time each day)
Outcomes	Primary: 1. Response rate (RR), defined as the proportion of patients who have a ≥35% reduction in volume of spleen size at the end of cycle 6, and confirmed 4 weeks thereafter. Secondary outcomes: 1. Symptom response rate (SRR): Proportion of patients with 50% reduction from baseline to the end of cycle 6 in the total symptom score. 6 months No. this assessment will be conducted through the modified MFSAF diary, which will be completed during the week prior to Day 1 of each treatment cycle up to cycle 6, at end of cycle 6, the EOT visit, and the 30-day follow-up visit. 2. Overall survival of either 400 mg/day or 500 mg/day of IMP as compared with placebo. 3. Progression free survival of either 400 mg/day or 500 mg/day of IMP as compared with placebo. 4. Proportion of patients who have 25% reduction in volume of spleen size at end of cycle 6, and confirmed 4 weeks thereafter. 5. Duration of spleen response, measured by MRI (or CT scan in patients with contraindications for MRI). 6. Clinical and laboratory events graded by the NCI CTCAE v4.03.

NCT01437787 (*Continued*)

Starting date	October 2011
Contact information	Contact-Us@sanofi-aventis.com
Notes	Official title: A phase 3, multicenter, randomized, double-blind, placebo-controlled, 3-arm study of SAR302503 in patients with intermediate-2 or high-risk PMF, post-polycythemia vera MF, or post-essential thrombocythemia MF with splenomegaly Sponsors and collaborators: Sanofi Aventis. Study chairs or principal investigators: Clinical Sciences & Operations, Study Director, Sanofi-Aventis Primary objective: 1. To evaluate the efficacy of daily oral doses of 400 mg or 500 mg of SAR302503 (Investigational Medicinal Product, IMP) compared with placebo in the reduction of spleen volume as determined by MRI (or CT scan in patients with contraindications for MRI). Secondary objectives: 1. To evaluate the effect on MF-associated symptoms (key MF symptoms) as measured by the modified MFSAF diary. 2. To evaluate the overall survival of patients treated with either 400 mg/day or 500 mg/day of IMP as compared with placebo. 3. To evaluate the progression free survival of patients treated with either 400 mg/day or 500 mg/day of IMP as compared with placebo. 4. To evaluate the durability of splenic response. 5. To evaluate the safety of IMP.

DATA AND ANALYSES

Comparison 1. Ruxolitinib versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Overall survival	1	309	Hazard Ratio (Random, 95% CI)	0.51 [0.27, 0.98]
2 Hematological adverse events (Adverse events observed in 10% or more of patients who received ruxolitinib. Harm (Grades 3 or 4). According to National Cancer Institute Common terminology criteria for adverse events)	1		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
2.1 Anemia	1	306	Risk Ratio (M-H, Random, 95% CI)	2.35 [1.62, 3.41]
2.2 Thrombocytopenia	1	306	Risk Ratio (M-H, Random, 95% CI)	9.74 [2.32, 40.96]
2.3 Neutropenia	1	306	Risk Ratio (M-H, Random, 95% CI)	3.57 [1.02, 12.55]
3 Non-hematological adverse events (Adverse events observed in 10% or more of patients who received ruxolitinib. Harm (Grades 3 or 4). According to National Cancer Institute common terminology criteria for adverse events)	1		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
3.1 Fatigue	1	306	Risk Ratio (M-H, Random, 95% CI)	0.78 [0.32, 1.92]
3.2 Abdominal pain	1	306	Risk Ratio (M-H, Random, 95% CI)	0.23 [0.08, 0.67]
3.3 Dyspnea	1	306	Risk Ratio (M-H, Random, 95% CI)	0.32 [0.07, 1.58]
3.4 Dizziness	1	306	Risk Ratio (M-H, Random, 95% CI)	0.10 [0.01, 0.75]
3.5 Arthralgia	1	306	Risk Ratio (M-H, Random, 95% CI)	2.92 [0.31, 27.79]
3.6 Vomiting	1	306	Risk Ratio (M-H, Random, 95% CI)	0.97 [0.06, 15.43]
3.7 Diarrhea	1	306	Risk Ratio (M-H, Random, 95% CI)	4.87 [0.24, 100.64]
3.8 Nausea	1	306	Risk Ratio (M-H, Random, 95% CI)	0.32 [0.01, 7.91]
3.9 Pain in extremity	1	306	Risk Ratio (M-H, Random, 95% CI)	4.87 [0.24, 100.64]
3.10 Pyrexia	1	306	Risk Ratio (M-H, Random, 95% CI)	0.97 [0.06, 15.43]
4 Health-related quality of life: proportion of patients with a reduction of 50% or more in MFSAF scores at 24 weeks	1	309	Risk Ratio (M-H, Random, 95% CI)	8.82 [4.40, 17.69]
5 Health-related quality of life: Mean difference in MFSAF at follow-up scores at 24 weeks	1	232	Mean Difference (Random, 95% CI)	-87.90 [-139.58, -36.22]
6 Reduction in spleen size (≥ 35%) (at 48 weeks follow-up)	1	309	Risk Ratio (M-H, Random, 95% CI)	64.58 [9.08, 459.56]
7 Leukemia-free survival	1	309	Hazard Ratio (Random, 95% CI)	5.00 [0.52, 48.07]

Comparison 2. Ruxolitinib versus best available therapy

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Overall survival	1	219	Hazard Ratio (Random, 95% CI)	0.70 [0.20, 2.47]
2 Progression-free survival (at 48 weeks)	1	219	Hazard Ratio (Random, 95% CI)	0.81 [0.47, 1.39]
3 Hematological adverse events	1		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
3.1 Anemia	1	219	Risk Ratio (M-H, Random, 95% CI)	1.35 [0.91, 1.99]
3.2 Thrombocytopenia	1	219	Risk Ratio (M-H, Random, 95% CI)	1.2 [0.44, 3.28]
4 Non-hematological adverse events	1		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
4.1 Fatigue	1	219	Risk Ratio (M-H, Random, 95% CI)	1.51 [0.06, 36.62]
4.2 Abdominal pain	1	219	Risk Ratio (M-H, Random, 95% CI)	1.25 [0.25, 6.29]
4.3 Dyspnea	1	219	Risk Ratio (M-H, Random, 95% CI)	0.17 [0.02, 1.57]
4.4 Arthalgia	1	219	Risk Ratio (M-H, Random, 95% CI)	1.51 [0.06, 36.62]
4.5 Diarrhea	1	219	Risk Ratio (M-H, Random, 95% CI)	2.52 [0.12, 51.76]
4.6 Nausea	1	219	Risk Ratio (M-H, Random, 95% CI)	1.51 [0.06, 36.62]
4.7 Pain in extremity	1	219	Risk Ratio (M-H, Random, 95% CI)	1.51 [0.06, 36.62]
4.8 Pyrexia	1	219	Risk Ratio (M-H, Random, 95% CI)	3.52 [0.18, 67.32]
4.9 Headache	1	219	Risk Ratio (M-H, Random, 95% CI)	2.52 [0.12, 51.76]
5 Health-related quality of life	1	96	Mean Difference (Random, 95% CI)	7.60 [0.35, 14.85]
6 Reduction in spleen size	1		Mean Difference (Random, 95% CI)	Subtotals only
6.1 At 24 weeks follow-up	1	216	Mean Difference (Random, 95% CI)	-31.90 [-53.85, -9.95]
6.2 At 48 weeks follow-up	1	216	Mean Difference (Random, 95% CI)	-37.4 [-65.41, -9.39]
7 Reduction in spleen volume (≥ 35%) (at 24 and 48 weeks follow-up)	1	219	Risk Ratio (M-H, Random, 95% CI)	41.78 [2.61, 669.75]
8 Leukemia-free survival	1	219	Hazard Ratio (Random, 95% CI)	0.65 [0.18, 2.33]

Analysis 1.1. Comparison 1 Ruxolitinib versus placebo, Outcome 1 Overall survival.

Review: Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis

Comparison: 1 Ruxolitinib versus placebo

Outcome: 1 Overall survival

Study or subgroup	Ruxolitinb N		Placebo N	log [Hazard Ratio] (SE)	Hazard Ratio IV,Random,95% CI	Weight	Hazard Ratio IV,Random,95% CI
COMFORT-1 2012	155	-0.66717169 (0.32891813)	154			100.0 %	0.51 [0.27, 0.98]
Total (95% CI)	**155**		**154**			**100.0 %**	**0.51 [0.27, 0.98]**

Heterogeneity: not applicable

Test for overall effect: Z = 2.03 (P = 0.043)

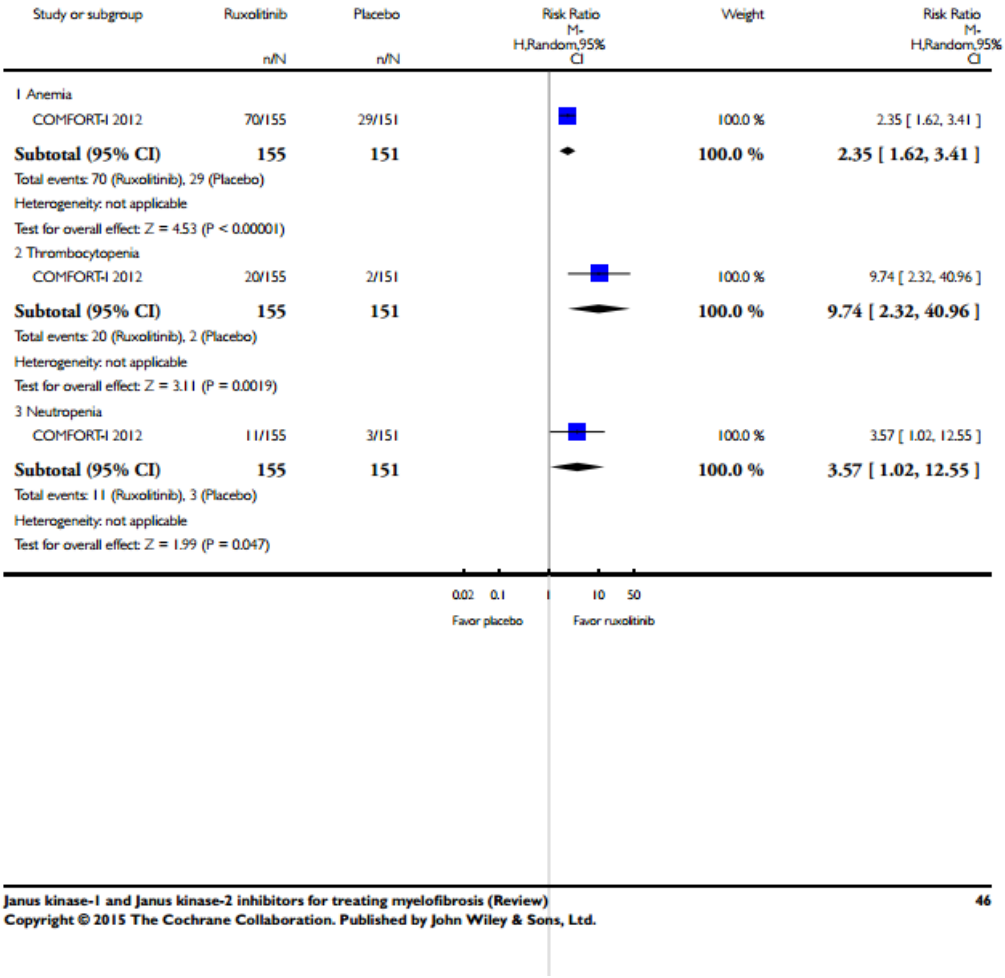
Test for subgroup differences: Not applicable

Analysis 1.2. Comparison 1 Ruxolitinib versus placebo, Outcome 2 Hematological adverse events (Adverse events observed in 10% or more of patients who received ruxolitinib. Harm (Grades 3 or 4). According to National Cancer Institute Common terminology criteria for adverse events).

Review: Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis

Comparison: 1 Ruxolitinib versus placebo

Outcome: 2 Hematological adverse events (Adverse events observed in 10% or more of patients who received ruxolitinib. Harm (Grades 3 or 4). According to National Cancer Institute Common terminology criteria for adverse events)

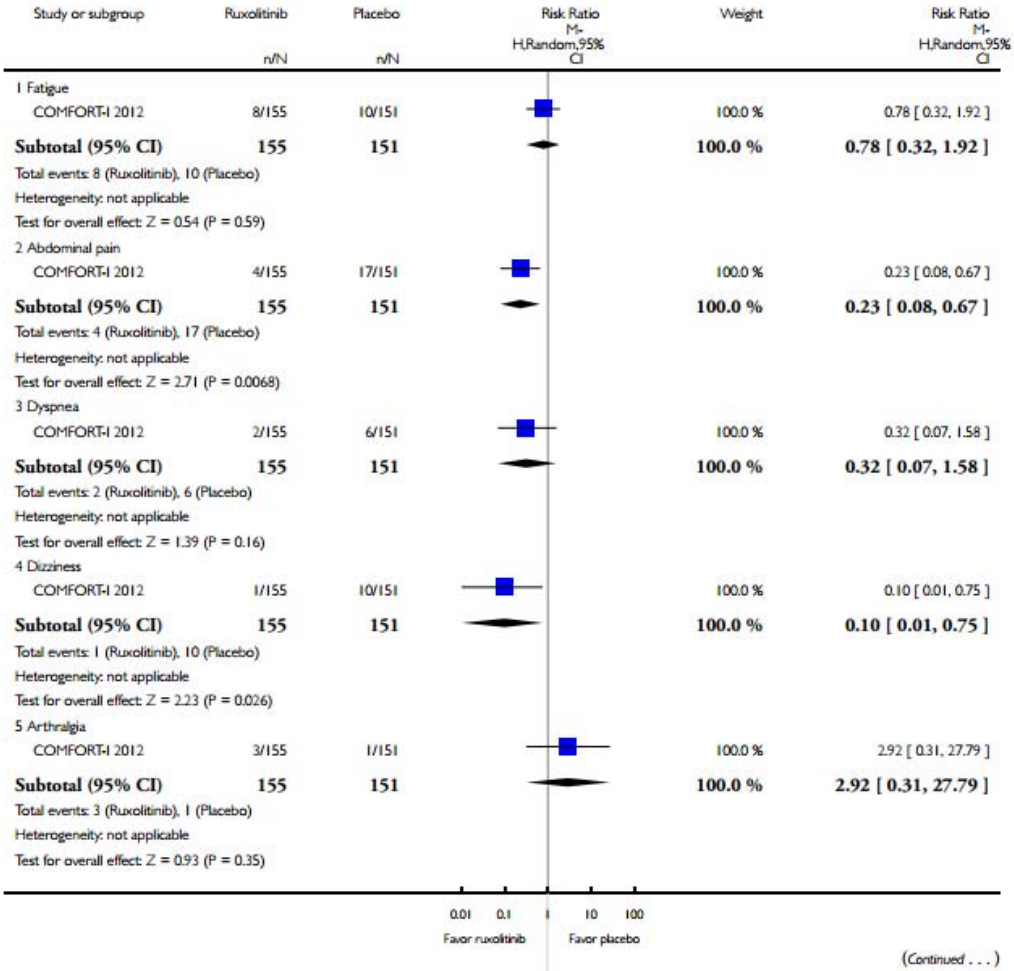


Analysis 1.3. Comparison 1 Ruxolitinib versus placebo, Outcome 3 Non-hematological adverse events (Adverse events observed in 10% or more of patients who received ruxolitinib. Harm (Grades 3 or 4). According to National Cancer Institute common terminology criteria for adverse events).

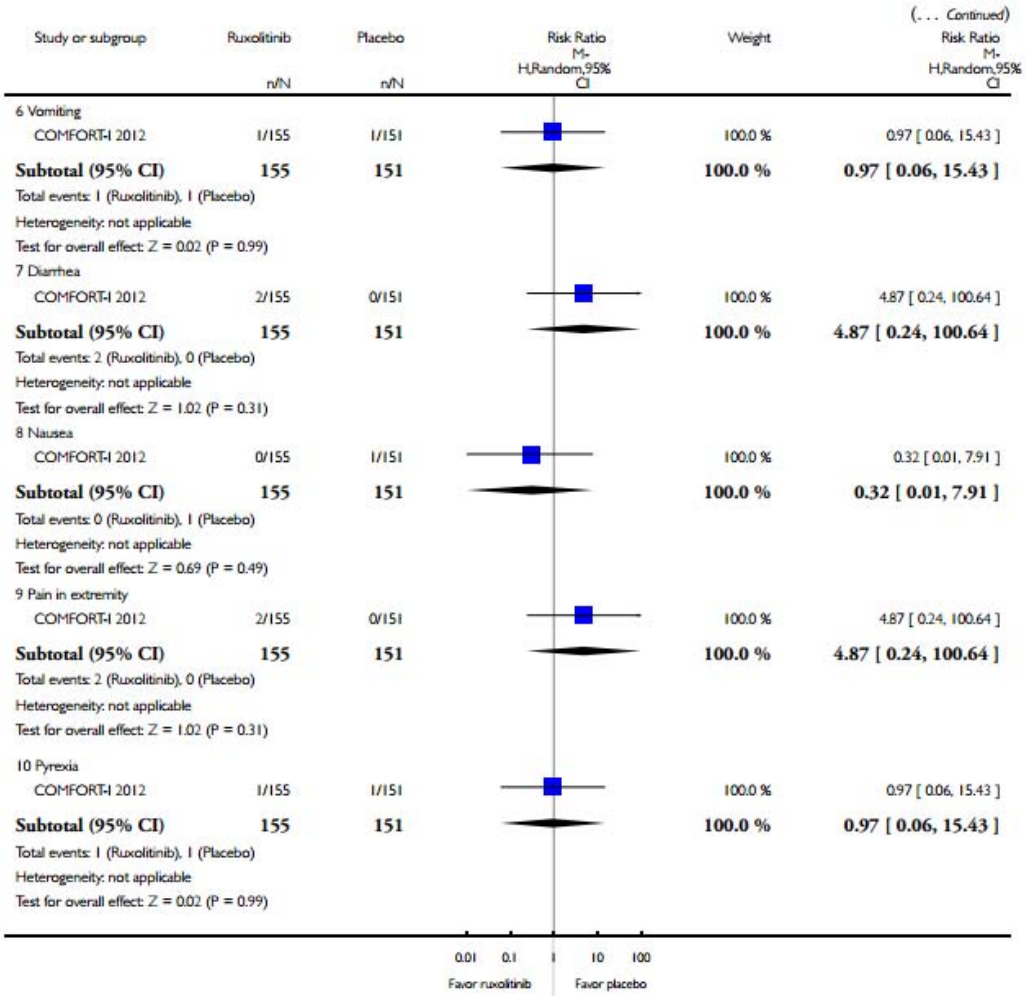
Review: Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis

Comparison: 1 Ruxolitinib versus placebo

Outcome: 3 Non-hematological adverse events (Adverse events observed in 10% or more of patients who received ruxolitinib. Harm (Grades 3 or 4). According to National Cancer Institute common terminology criteria for adverse events)



(Continued . . .)

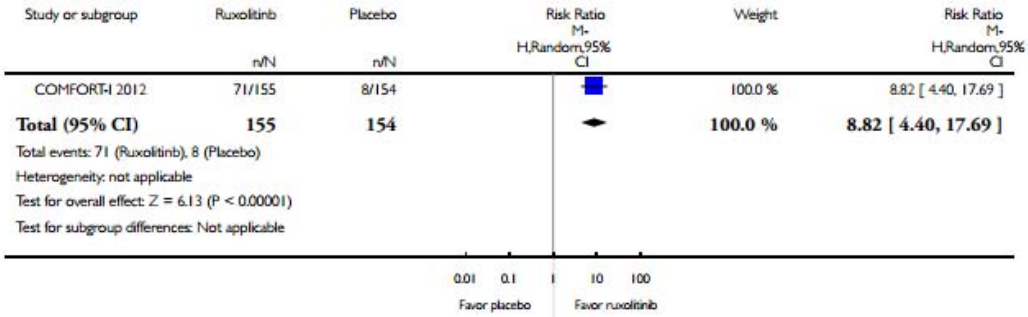


Analysis 1.4. Comparison 1 Ruxolitinib versus placebo, Outcome 4 Health-related quality of life: proportion of patients with a reduction of 50% or more in MFSAF scores at 24 weeks.

Review: Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis

Comparison: 1 Ruxolitinib versus placebo

Outcome: 4 Health-related quality of life: proportion of patients with a reduction of 50% or more in MFSAF scores at 24 weeks

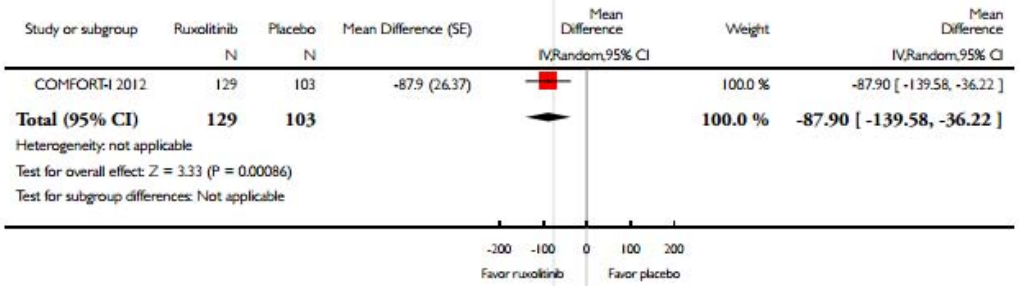


Analysis 1.5. Comparison 1 Ruxolitinib versus placebo, Outcome 5 Health-related quality of life: Mean difference in MFSAF at follow-up scores at 24 weeks.

Review: Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis

Comparison: 1 Ruxolitinib versus placebo

Outcome: 5 Health-related quality of life: Mean difference in MFSAF at follow-up scores at 24 weeks

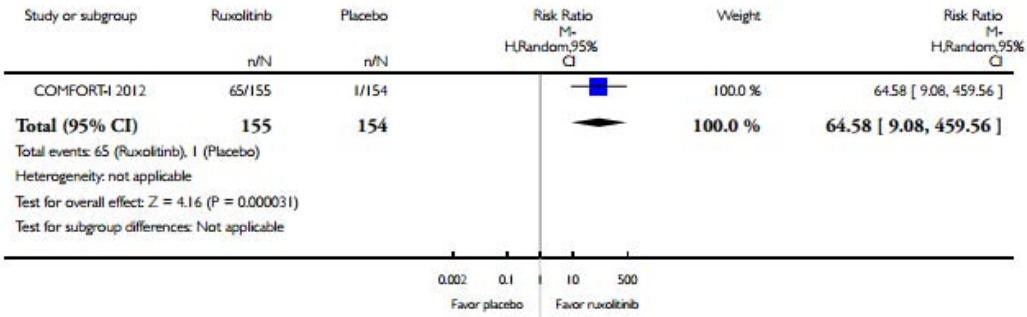


Analysis 1.6. Comparison 1 Ruxolitinib versus placebo, Outcome 6 Reduction in spleen size (≥ 35%) (at 48 weeks follow-up).

Review: Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis

Comparison: 1 Ruxolitinib versus placebo

Outcome: 6 Reduction in spleen size (≥ 35%) (at 48 weeks follow-up)

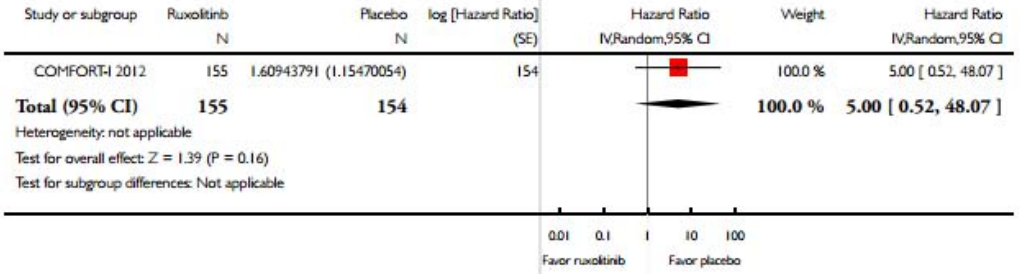


Analysis 1.7. Comparison 1 Ruxolitinib versus placebo, Outcome 7 Leukemia-free survival.

Review: Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis

Comparison: 1 Ruxolitinib versus placebo

Outcome: 7 Leukemia-free survival

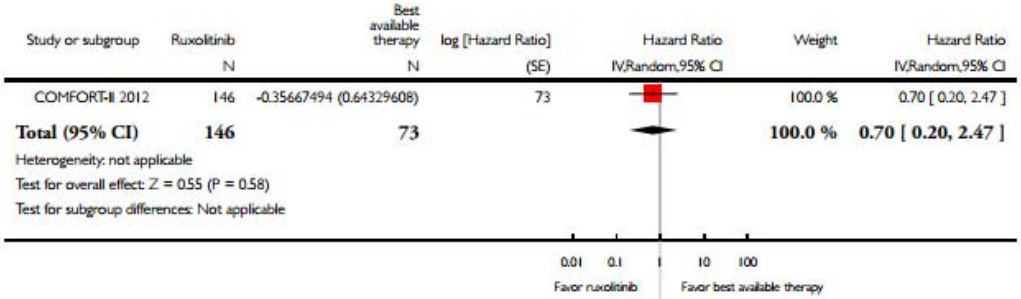


Analysis 2.1. Comparison 2 Ruxolitinib versus best available therapy, Outcome 1 Overall survival.

Review: Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis

Comparison: 2 Ruxolitinib versus best available therapy

Outcome: 1 Overall survival

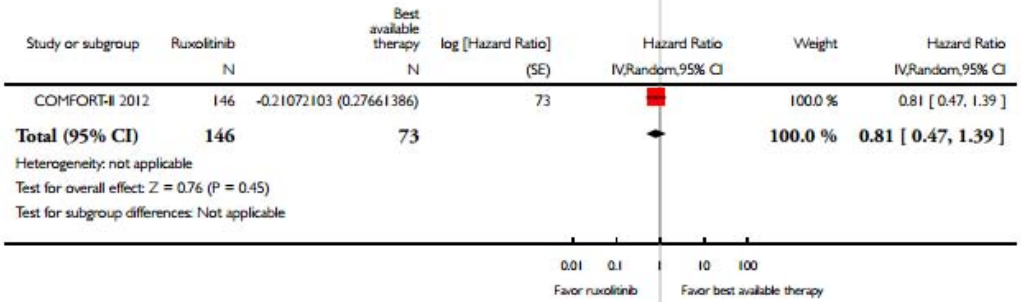


Analysis 2.2. Comparison 2 Ruxolitinib versus best available therapy, Outcome 2 Progression-free survival (at 48 weeks).

Review: Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis

Comparison: 2 Ruxolitinib versus best available therapy

Outcome: 2 Progression-free survival (at 48 weeks)

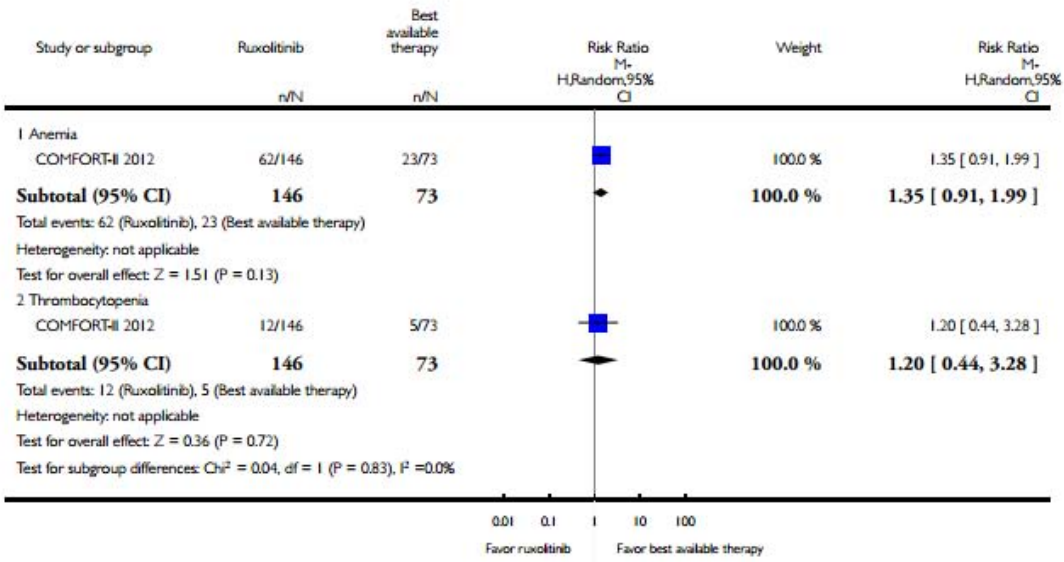


Analysis 2.3. Comparison 2 Ruxolitinib versus best available therapy, Outcome 3 Hematological adverse events.

Review: Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis

Comparison: 2 Ruxolitinib versus best available therapy

Outcome: 3 Hematological adverse events

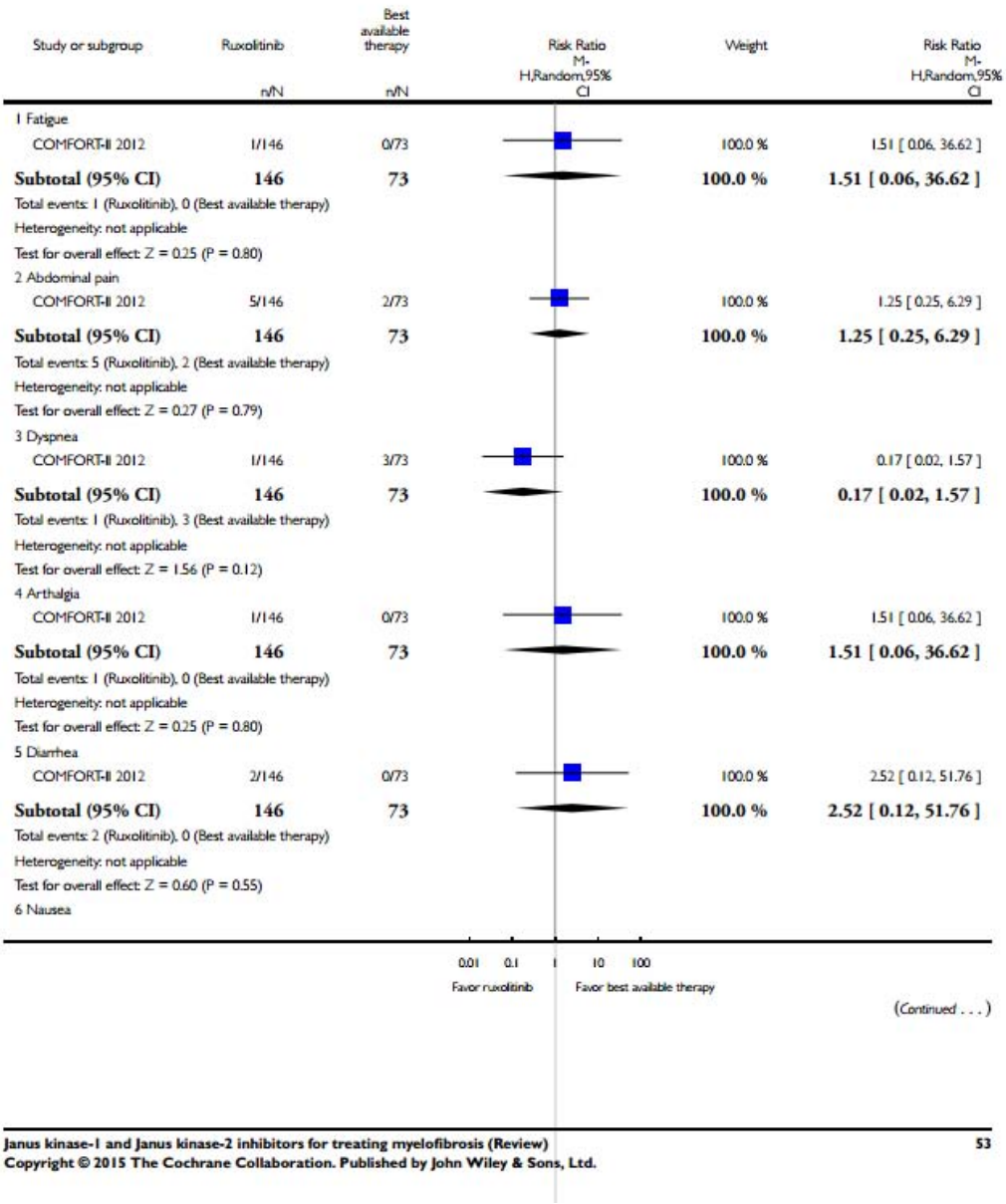


Analysis 2.4. Comparison 2 Ruxolitinib versus best available therapy, Outcome 4 Non-hematological adverse events.

Review: Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis

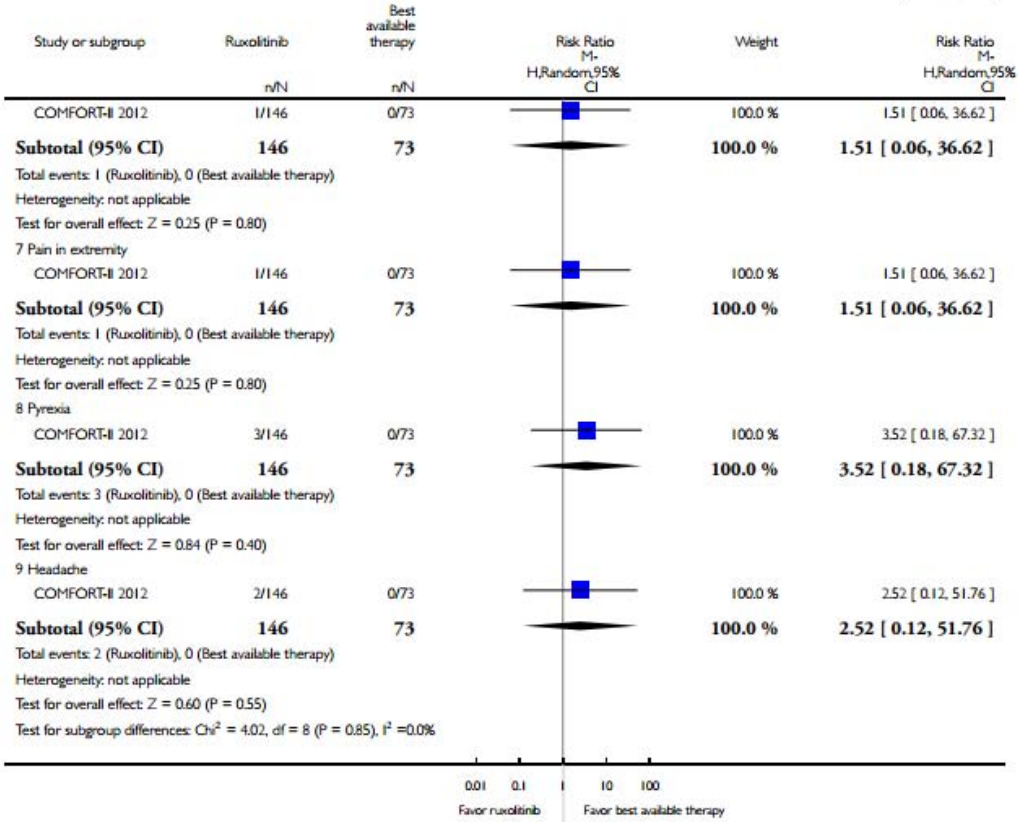
Comparison: 2 Ruxolitinib versus best available therapy

Outcome: 4 Non-hematological adverse events



(Continued . . .)

(. . . Continued)

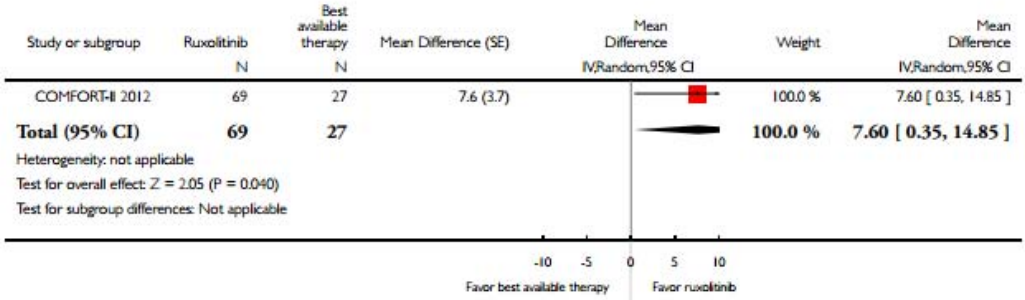


Analysis 2.5. Comparison 2 Ruxolitinib versus best available therapy, Outcome 5 Health-related quality of life.

Review: Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis

Comparison: 2 Ruxolitinib versus best available therapy

Outcome: 5 Health-related quality of life

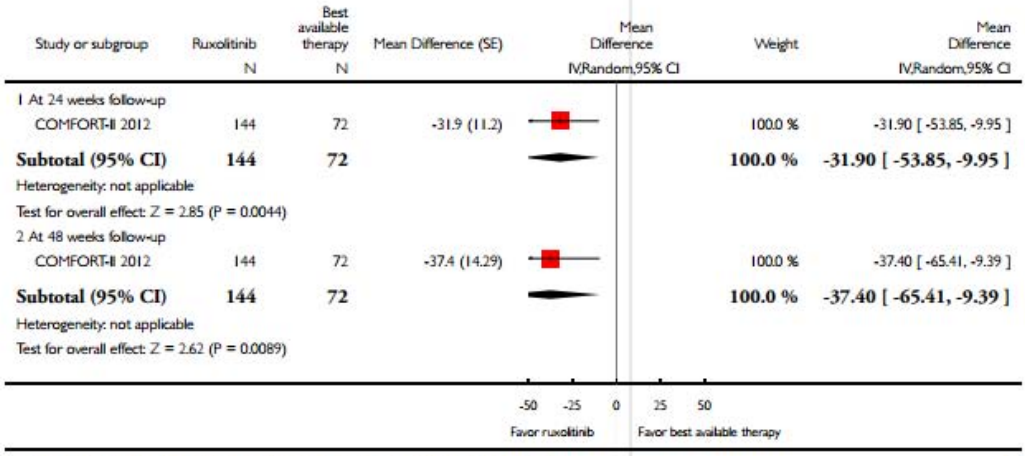


Analysis 2.6. Comparison 2 Ruxolitinib versus best available therapy, Outcome 6 Reduction in spleen size.

Review: Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis

Comparison: 2 Ruxolitinib versus best available therapy

Outcome: 6 Reduction in spleen size

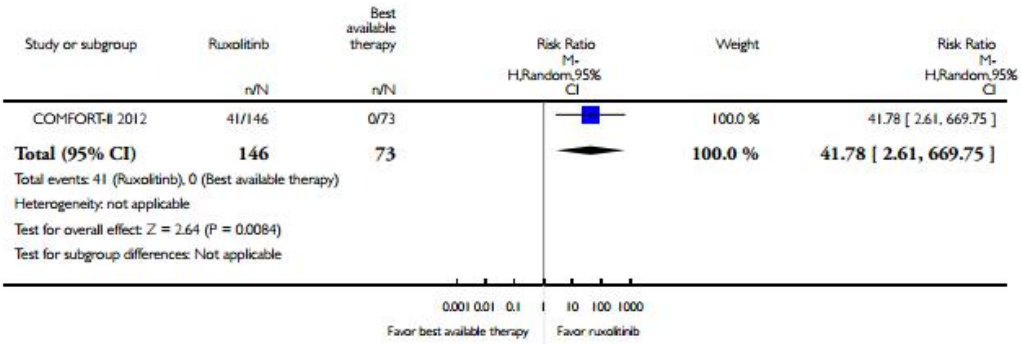


Analysis 2.7. Comparison 2 Ruxolitinib versus best available therapy, Outcome 7 Reduction in spleen volume (≥ 35%) (at 24 and 48 weeks follow-up).

Review: Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis

Comparison: 2 Ruxolitinib versus best available therapy

Outcome: 7 Reduction in spleen volume (≥ 35%) (at 24 and 48 weeks follow-up)

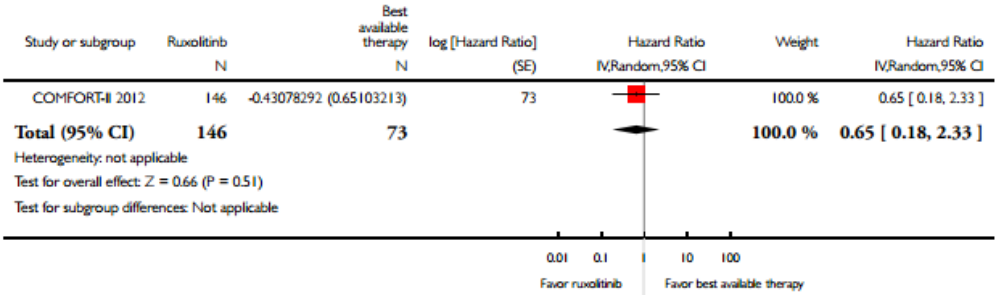


Analysis 2.8. Comparison 2 Ruxolitinib versus best available therapy, Outcome 8 Leukemia-free survival.

Review: Janus kinase-1 and Janus kinase-2 inhibitors for treating myelofibrosis

Comparison: 2 Ruxolitinib versus best available therapy

Outcome: 8 Leukemia-free survival



APPENDICES

Appendix 1. International Working Group for Myeloproliferative Neoplasms Research and Treatment (IWG-MRT) recommended criteria for post-polycythemia vera and post-essential thrombocythemia myelofibrosis

Criteria for post-polycythemia vera MF

Required criteria

- Documentation of a previous diagnosis of polycythemia vera as defined by the WHO criteria.
- Bone marrow fibrosis grade 2 to 3 (on 0 to 3 scale) or grade 3 to 4 (on 0 to 4 scale) (see footnote for details).

Additional criteria (two are required)

- Anemia or sustained loss of requirement for phlebotomy in the absence of cytoreductive therapy.
- A leukoerythroblastic peripheral blood picture.
- Increasing splenomegaly defined as either an increase in palpable splenomegaly of $\geq$ 5 cm (distance of the tip of the spleen from the left costal margin) or the appearance of a newly palpable splenomegaly.
- Development of $\geq$ 1 of three constitutional symptoms: > 10% weight loss in six months, night sweats, unexplained fever (> 37.5°C).

Criteria for post-essential thrombocythemia MF

Required criteria

- Documentation of a previous diagnosis of essential thrombocythemia as defined by the WHO criteria.
- Bone marrow fibrosis grade 2 to 3 (on 0 to 3 scale) or grade 3 to 4 (on 0 to 4 scale) (see footnote for details).

Additional criteria (two are required)

- Anemia and ≥ 2 g/dL decrease from baseline hemoglobin level.
- A leukoerythroblastic peripheral blood picture.
- Increasing splenomegaly defined as either an increase in palpable splenomegaly of ≥ 5 cm (distance of the tip of the spleen from the left costal margin) or the appearance of a newly palpable splenomegaly.
- Increased lactate dehydrogenase.
- Development of ≥ 1 of three constitutional symptoms: > 10% weight loss in six months, night sweats, unexplained fever (> 37.5°C).

Source: Tefferi 2011a.

Appendix 2. Medical glossary

Medical term	Definition	Source
Allogeneic stem cell transplantation	The transfer of stem cells from one individual (genetically non identical) to another within the same species The source and location of the stem cells determines their potency or pluripotency to differentiate into various cell types See stem cells. See Peripheral Blood Stem Cell Transplantation.	http://www.ncbi.nlm.nih.gov/mesh
Angiogenesis 1. Neovascularization, Physiologic 2. Neovascularization, Pathologic	1. The development of new blood vessels in restoration of blood circulation during the healing process. 2. A pathologic process resulting in proliferation of blood vessels in abnormal tissues or in abnormal positions.	http://www.ncbi.nlm.nih.gov/mesh
Ascites	Accumulation or retention of free fluid within the peritoneal cavity	http://www.ncbi.nlm.nih.gov/mesh
Atypia	Abnormal shape of any cell.	
Best available therapy	Selected therapy included a combination of available agents to treat the disease or its symptoms, or both	COMFORT-II 2012

(*Continued*)

Bone marrow	The soft tissue filling the cavities of bones. Bone marrow exists in two types: yellow and red. Yellow marrow is found in the large cavities of large bones and consists mostly of fat cells and a few primitive blood cells. Red marrow is a hematopoietic tissue and is the site of production of blood cells. Bone marrow is made up of a framework of connective tissue containing branching fibers with the frame being filled with marrow cells	http://www.ncbi.nlm.nih.gov/mesh
Cachexia	General ill health, malnutrition, and weight loss, usually associated with chronic disease	http://www.ncbi.nlm.nih.gov/mesh
Cytokine	Non-antibody proteins secreted by inflammatory leukocytes and some non-leukocytic cells, that act as intercellular mediators. They differ from classical hormones in that they are produced by a number of cell types rather than by specialized glands. They generally act locally in a paracrine or autocrine rather than endocrine manner	http://www.ncbi.nlm.nih.gov/mesh
Dacryocytes (Teardrop cell)	Distorted, drop-shaped cell.	Greer 2009
Eastern Cooperative Oncology Group (ECOG)	Scale for grading performance status.	Oken 1982
Eltrombopag	An oral, non peptide thrombopoietin receptor agonist.	http://www.ncbi.nlm.nih.gov/mesh
Erythropoietin	Glycoprotein hormone, secreted chiefly by the kidney in the adult and the liver in the fetus, that acts on erythroid stem cells of the bone marrow to stimulate proliferation and differentiation	http://www.ncbi.nlm.nih.gov/mesh
Essential thrombocythemia	Chronic myeloproliferative disorder characterized by a sustained proliferation of megakaryocytes, which leads to increased numbers of circulating platelets	Finazzi 2008
Extramedullary hematopoiesis	The formation and development of blood cells outside the bone marrow, as in the spleen; liver; or lymph nodes	http://www.ncbi.nlm.nih.gov/mesh

(*Continued*)

Janus kinases	A family of intracellular tyrosine kinases that participate in the signaling cascade of cytokines by associating with specific cytokine receptors. They act upon stat transcription factors in signaling pathway referred to as the JAK/STAT pathway. The name Janus kinase refers to the fact the proteins have two phosphate-transferring domains	http://www.ncbi.nlm.nih.gov/mesh
Janus kinase-1	A Janus kinase subtype that is involved in signaling from a broad variety of cytokine receptors	http://www.ncbi.nlm.nih.gov/mesh
Janus kinase-2	A Janus kinase subtype that is involved in signaling from growth hormone receptors; prolactin receptors; and a variety of cytokine receptors such as erythropoietin receptors and interleukin receptors. Dysregulation of Janus kinase 2 due to genetic translocations have been associated with a variety of myeloproliferative disorders	http://www.ncbi.nlm.nih.gov/mesh
Janus kinase-3	A Janus kinase subtype that is predominantly expressed in hematopoietic cell. It is involved in signaling from a broad variety of cytokine receptors including ones that utilize the interleukin receptors common gamma subunit	http://www.ncbi.nlm.nih.gov/mesh
Leukoerythroblastic condition	The leukoerythroblastic condition is characterized by the presence of nucleated red blood cells and immature myeloid elements in 96% of cases	Hoffman 2008
Hepatomegaly	Enlargement of the liver.	http://www.ncbi.nlm.nih.gov/mesh
Myelodysplastic syndromes	Clonal hematopoietic stem cell disorders characterized by dysplasia in one or more hematopoietic cell lineages. They predominantly affect patients over 60, are considered preleukemic conditions, and have high probability of transformation into acute myeloid leukemia	http://www.ncbi.nlm.nih.gov/mesh
MF	MF is a bone marrow disease characterized by excessive production of reticulin and collagen fibers	Ostojic 2012

(*Continued*)

Osteosclerosis	An abnormal hardening or increased density of bone tissue.	http://www.ncbi.nlm.nih.gov/mesh
Peripheral blood leukoerythroblastosis	Presence of nucleated red cells, immature granulocytes, and dacryocytes	Tefferi 2011a
Peripheral blood stem cell transplantation	Transplantation of stem cells collected from the peripheral blood. It is a less invasive alternative to direct marrow harvesting of hematopoietic stem cells. Enrichment of stem cells in peripheral blood can be achieved by inducing mobilization of stem cells from the bone marrow	http://www.ncbi.nlm.nih.gov/mesh
Pleural effusion	A pleural effusion is a buildup of fluid between the layers of tissue that line the lungs and chest cavity	http://www.ncbi.nlm.nih.gov/mesh
Poikilocytosis	It is an increased variation in cell shape.	Bain 2012
Polycythemia vera	A myeloproliferative disorder of unknown etiology, characterized by abnormal proliferation of all hematopoietic bone marrow elements and an absolute increase in red cell mass and total blood volume, associated frequently with splenomegaly, leukocytosis, and thrombocythemia. Hematopoiesis is also reactive in extramedullary sites (liver and spleen). It can lead to MF	http://www.ncbi.nlm.nih.gov/mesh
Portal hypertension	Abnormal increase of resistance to blood flow within the hepatic portal system, frequently seen in liver cirrhosis and conditions with obstruction of the portal vein	http://www.ncbi.nlm.nih.gov/mesh
Primary MF	It is a Philadelphia (Ph) chromosome-negative chronic myeloproliferative neoplasm (MPN) characterized by bone marrow fibrosis, extramedullary hemopoiesis with splenomegaly, and leukoerythroblastosis in blood	Cervantes 2011
Reticulin	A scleroprotein fibril consisting mostly of type III collagen. Reticulin fibrils are extremely thin, with a diameter of between 0.5 and 2 um. They are involved in maintaining the structural integrity in a variety of organs	http://www.ncbi.nlm.nih.gov/mesh

(*Continued*)

Romiplostim	Consists of a carrier Fc domain linked to multiple copies of Mpl-binding peptide; stimulates megakaryopoiesis in vitro by binding to Mpl	http://www.ncbi.nlm.nih.gov/mesh
Splenomegaly	Enlargement of the spleen.	http://www.ncbi.nlm.nih.gov/mesh
Splenic infarction	Insufficiency of arterial or venous blood supply to the spleen due to emboli, thrombi, vascular torsion, or pressure that produces a macroscopic area of necrosis	http://www.ncbi.nlm.nih.gov/mesh
Stem cells	Relatively undifferentiated cells that retain the ability to divide and proliferate throughout postnatal life to provide progenitor cells that can differentiate into specialized cells	http://www.ncbi.nlm.nih.gov/mesh
TYK2 kinase	A Janus kinase subtype that is involved in signaling from a broad variety of cytokine receptors. The TYK2 kinase is considered the founding member of the Janus kinase family and was initially discovered as a signaling partner for the Iinterferon alpha-beta receptor. The kinase has since been shown to signal from several interleukin receptors	http://www.ncbi.nlm.nih.gov/mesh
Variceal esophogeal bleeding	Bleeding esophogeal varices are enlarged veins in the walls of the lower part of the esophagus (the tube that connects your throat to your stomach) that bleed	http://www.nlm.nih.gov/medlineplus/ency/article/000268.htm

Appendix 3. CENTRAL search strategy

#1 MeSH descriptor: [Primary Myelofibrosis] explode all trees
#2 MeSH descriptor: [Myeloproliferative Disorders] explode all trees
#3 myelofibros*
#4 mielofibros*
#5 osteomyelofibros*
#6 (myeloid* near/1 metaplas*)
#7 (bone marrow near/1 fibros*)
#8 ((nonleukemic* or nonleukaemic*) near/2 myelos*)
#9 myeloscleros*
#10 myeloproliferativ*
#11 osteomyelofibros*
#12 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11
#13 MeSH descriptor: [Janus Kinase 1] explode all trees

#14 MeSH descriptor: [Janus Kinase 2] explode all trees
#15 (jak1* or jak-1*)
#16 (jak2* or jak-2*)
#17 (jakafi* or jakavi*)
#18 (jak* near/3 inhibit*)
#19 (janus* near/2 kinas*)
#20 (INCB-018424 or INCB018424)
#21 Ruxolitinib*
#22 (INC-424 or INC424)
#23 #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22
#24 #12 and #23
#25 #24 from 2012 to 2013, in Trials

Search update (13 November 2014)

MeSH descriptor: [Primary Myelofibrosis] explode all trees
#2 MeSH descriptor: [Myeloproliferative Disorders] explode all trees
#3 myelofibros*
#4 mielofibros*
#5 osteomyelofibros*
#6 (myeloid* near/1 metaplas*)
#7 (bone marrow near/1 fibros*)
#8 ((nonleukemic* or nonleukaemic*) near/2 myelos*)
#9 myeloscleros*
#10 myeloproliferativ*
#11 osteomyelofibros*
#12 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11
#13 MeSH descriptor: [Janus Kinase 1] explode all trees
#14 MeSH descriptor: [Janus Kinase 2] explode all trees
#15 (jak1* or jak-1*)
#16 (jak2* or jak-2*)
#17 (jakafi* or jakavi*)
#18 (jak* near/3 inhibit*)
#19 (janus* near/2 kinas*)
#20 (INCB-018424 or INCB018424)
#21 Ruxolitinib*
#22 (INC-424 or INC424)
#23 #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22
#24 #12 and #23
#25 #24 in Trials
#26 #24 Publication Year from 2013 to 2014, in Trials

Appendix 4. MEDLINE Ovid MEDLINE® In-Process & Other Non-Indexed Citations (13 November 2014), Ovid MEDLINE® (1946 to November 2014) search strategy

1	PRIMARY MYELOFIBROSIS/
2	exp MYELOPROLIFERATIVE DISORDERS/
3	myelofibros$.tw,kf,ot.
4	mielofibros$.tw,kf,ot.
5	osteomyelofibros$.tw,kf,ot.
6	(myeloid$ adj1 metaplas$).tw,kf,ot.
7	(bone marrow adj1 fibros$).tw,kf,ot.
8	((nonleukemic$ or nonleukaemic$) adj2 myelos$).tw,kf,ot.
9	myeloscleros$.tw,kf,ot.
10	myeloproliferativ$.tw,kf,ot.
11	osteomyelofibros$.tw,kf,ot.
12	or/1-11
13	JANUS KINASE 1/
14	JANUS KINASE 2/
15	(jak1$ or jak-1$).tw,kf,ot.
16	(jak2$ or jak-2$).tw,kf,ot.
17	(jakafi$ or jakavi$).tw,kf,ot.
18	(jak$ adj3 inhibit$).tw,kf,ot.
19	(janus$ adj2 kinas$).tw,kf,ot.
20	(INCB-018424 or INCB018424).tw,kf,ot.
21	ruxolitinib$.tw,kf,ot.
22	(INC-424 or INC424).tw,kf,ot.
23	or/13-22

(*Continued*)

24	12 and 23
25	randomized controlled trial.pt.
26	controlled clinical trial.pt.
27	randomi?ed.ab.
28	placebo.ab.
29	drug therapy.fs.
30	randomly.ab.
31	trial.ab.
32	groups.ab.
33	or/25-32
34	humans.sh.
35	33 and 34
36	24 and 35
37	limit 36 to ed=20121213-20131014

Appendix 5. EMBASE search strategy

#	Searches
1	'MYELOID METAPLASIA'/de
2	'MYELOPROLIFERATIVE DISORDER'/ de
3	myelofibros*:ab,ti
4	mielofibros*:ab,ti
5	osteomyelofibros*:ab,ti
6	(myeloid* NEXT/1 metaplas*):ab,ti
7	(bone marrow NEXT/1 fibros*):ab,ti

(*Continued*)

8	(nonleukemic* NEAR/2 myelos*):ab,ti
9	(nonleukaemic* NEAR/2 myelos*):ab,ti
10	myeloscleros*:ab,ti
11	myeloproliferativ*:ab,ti
12	osteomyelofibros*:ab,ti
13	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12
14	'JANUS KINASE 1'/de
15	'JANUS KINASE 2'/de
16	jak1*:ab,ti OR jak-1*:ab,ti
17	jak2*:ab,ti OR jak-2*:ab,ti
18	jakafi*:ab,ti OR jakavi*:ab,ti
19	(jak* NEAR/3 inhibit*):ab,ti
20	(janus* NEAR/2 kinas*):ab,ti
21	INCB-018424:ab,ti OR INCB018424:ab,ti
22	'RUXOLITINIB'/exp
23	INC-424:ab,ti OR INC424:ab,ti
24	#14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23
25	#13 and #24
26	'randomized controlled trial'/exp OR 'single blind procedure'/exp OR 'double blind procedure'/exp OR 'crossover procedure'/exp
27	random*:ab,ti OR placebo*:ab,ti OR allocat*:ab,ti OR crossover*:ab,ti OR 'cross over':ab,ti OR trial:ti OR (doubl* NEXT/1 blind*):ab,ti
28	#26 OR #27
29	'animal'/de OR 'animal experiment'/de OR 'nonhuman'/de
30	'human'/de

(*Continued*)

31	#29 OR #30
32	#29 NOT #31
33	#28 NOT #32
36	#25 AND #33

Appendix 6. LILACS search strategy

Search	Query
1	Ruxolitinib OR INCB-018424 OR INCB018424 OR jakafi OR jakavi OR jak1 OR jak-1 OR jak2 OR jak-2 OR jak OR janus OR INCB-018424 OR INCB018424 [Words] OR JANUS KINASE 1 OR JANUS KINASE 2 [Subject descriptor]

Appendix 7. *m*RCT search strategy

Search	Query
1	Ruxolitinib or INCB-018424 or INCB018424 or jakafi or jakavi or jak1 or jak-1 or jak2 or jak-2 or jak or janus or INCB-018424 or INCB018424

Appendix 8. Epistemonikos.org

Search	Query
1	Ruxolitinib

Appendix 9. Statistical terms related with potential biases in the review process

1. Excess significance

Excess significance is the phenomenon whereby the published literature has an excess of statistically significant results that are due to biases in reporting. Several mechanisms contribute to reporting bias, including study publication bias, where the results of statistically non-significant ('negative') studies are left unpublished; selective outcome reporting bias, where null results are omitted; and selective analysis bias, where data are analyzed with different methods that favor 'positive' results (Button 2013).

2. Winner's curse

The winner's curse refers to the phenomenon whereby the 'lucky' scientist who makes a discovery is cursed by finding an inflated estimate of that effect. The winner's curse occurs when thresholds, such as statistical significance, are used to determine the presence of an effect and is most severe when thresholds are stringent and studies are too small and thus have low power (Button 2013).

3. Vibration of effects

This refers to the situation in which a study obtains different estimates of the magnitude of the effect depending on the analytical options it implements. These
options could include the statistical model (Button 2013; Ioannidis 2008).

CONTRIBUTIONS OF AUTHORS

Arturo Martí-Carvajal: conceived and conducted the review, provided relevant contributions to its discussion and conclusions, rated the quality of evidence and provided the final discussion about the review findings. He provided comments and amended the review according the peer review comments.

Vidhu Anand: contributed to conducting the review and have provided relevant contributions to its discussion and conclusions.

Ivan Solà: contributed to the conducting of the review and have provided relevant contributions to its discussion and conclusions, rated the quality of evidence, prepared 'Summary of findings' tables and provided the final discussion about the findings of the review. He provided comments and amended the review according the peer review comments.

Arturo Martí-Carvajal is the guarantor of this Cochrane review.

DECLARATIONS OF INTEREST

In 2004 and 2007 Arturo Martí-Carvajal was employed by Eli Lilly to run a four-hour workshop on 'How to critically appraise clinical trials on osteoporosis and how to teach this'. This activity was not related to his work with Cochrane or any Cochrane Review.

Vidhu Anand: none known

Ivan Solà: none known

SOURCES OF SUPPORT

Internal sources

- Universidad Tecnológica Equinoccial, Ecuador.

Partially funded.

External sources

- Iberoamerican Cochrane Center, Barcelona, Spain.

Academic support

- Cochrane Haematological Malignancies Group, Germany.

Academic support

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

1. We reported spleen size reduction as both dichotomous and continuous approaches, as reported in COMFORT-I 2012; COMFORT-II 2012.

2. We included leukemia-free survival as an outcome after the protocol was published as it is a relevant clinical end point (Martí-Carvajal 2013).