NHS England

Evidence review: Infliximab for Progressive Pulmonary Sarcoidosis

NHS England

Evidence review: Infliximab for Progressive Pulmonary Sarcoidosis

First published:	September 2017
Updated:	Not applicable
Prepared by:	Solutions for Public Health (SPH) on behalf of NHS England Specialised Commissioning

Contents

1 Introduction

Introduction

- Sarcoidosis is characterised by non-caseating granulomas (non-necrotising nodules of inflammation and scarring) in the organs. The cause of sarcoidosis is not known (NICE 2017).
- The presentation of sarcoidosis varies considerably from mild, acute self-limiting disease to chronic disease involving several organs and causing severe symptoms and functional impairment (NICE 2017).
- Sarcoidosis usually has a benign course with high rates of spontaneous remission; 55% to 90% in patients presenting with stage I disease, 40% to 70% in patients presenting with stage II disease and 10 to 20% of patients presenting with stage III disease (NHS England 2017).It is possible that 40% of patients remit within six months (Gibson et al 1996), but the long term effects of treatment on the natural history of the disease are not known (NHS England 2017).
- Patients with chronic disease suffer from unremitting disease activity, risk of organ failure and symptoms which can severely reduce their quality of life (van Rijswijk et al 2013).
- The most commonly affected organ is the lungs, which are affected in more than 90% of people with sarcoidosis. The second most commonly affected organ is the skin, and other organs such as the eyes, brain, nervous system, liver and heart can also be affected (NICE 2017). A small number of patients experience end stage respiratory disease which may require lung transplant (NHS England 2017).
- Estimates for cases with neurological involvement vary from 5% to 15%. However this group commonly have a poor prognosis and present with severe acute events, e.g. optic neuritis and blindness, acute hydrocephalus and coma or progressive lower limb weakness (NHS England 2017).

Existing guidance from the National Institute of Health and Care Excellence (NICE)

- In recent evidence summaries, NICE clearly does not support the use of infliximab for pulmonary sarcoidosis except in cases where disease is refractory to standard treatments (NICE 2016). The conclusion for the use of infliximab for extrapulmonary sarcoidosis is less clear (NICE 2017).
- The NICE evidence summary on infliximab for pulmonary sarcoidosis (NICE 2016) concludes:

"The evidence supports British Thoracic Society guidance that immunosuppressants such as infliximab have only a limited role in pulmonary sarcoidosis because there are insufficient high-quality studies to confirm their place in therapy and they have significant adverse effects. The guidance advises that immunosuppressants should be used only when disease is refractory to standard treatments and when there are no pharmacological alternatives."

- The NICE evidence summary on infliximab for extrapulmonary sarcoidosis (NICE 2017) concludes:

"According to specialists involved in this evidence summary, infliximab may be an option for some patients with severe, refractory extrapulmonary sarcoidosis (particularly cutaneous or neurological sarcoidosis); for example, those affected by disabling or disfiguring disease, or whose life expectancy is likely to be reduced."

The indication and epidemiology

- General Practice data suggests an incidence of approximately three per 100,000 person years for sarcoidosis in the UK. Incidence is highest in people aged 20 to 50 years and appears to be higher in Afro-Caribbean people and marginally higher in women (NHS England 2017).
- In about 25% of patients sarcoidosis takes a chronic course (Jamilloux et al 2017).
- No figure for the proportion of patients who are refractory to treatments was identified. However in a survey of 19 of the world's leading sarcoidologists, 77% prescribed tumour necrosis factor alpha drugs (such as infliximab) to five or more sarcoidosis patients per year (Drent et al 2014). These experts also reported prescribing infliximab as an immunosuppressant in a mean of 7% (range 0% to 25%) of their total sarcoidosis population treated (Drent et al 2014).
- Disease-related mortality is about 5%, with the most common causes of death being lung, cardiac and neurological disease that is refractory to therapy (NICE 2017).

Standard treatment and pathway of care

- At present, treatment is not indicated for asymptomatic stage I to III sarcoidosis with mildly abnormal lung function and stable disease (NHS England 2017).
- For people with progressive disease or significant symptoms the first line treatment is oral corticosteroids, with a maintenance dose for six to 24 months (NHS England 2017).
- If corticosteroids are failing to control the disease or the side effects are intolerable to the patient, other immunosuppressive or anti-inflammatory second line treatments may be considered, such as Methotrexate or Azathioprine (NHS England 2017).
- Biological agents targeting the tumour necrosis factor (TNF) can be used as a third line treatment option in patients who are refractory to treatment or have developed adverse events (Jamilloux et al 2017).

The intervention (and licensed indication)

- Infliximab is a biological human monoclonal antibody, which inhibits tumour necrosis factor alpha (a cell signalling protein or cytokine involved in systematic inflammation) reducing disease activity (NICE 2017).
- There are four infliximab products: Remicade (the original brand name) and three biosimilar medicines, Flixabi, Inflectra and Remsima (NICE 2017). Infliximab is not licensed for treating any manifestation of sarcoidosis (NICE 2017).

Rationale for use

- Tumour necrosis factor (TNF) alpha elevations correlate with disease activity and progression in sarcoidosis (Russell et al 2013). Infliximab binds to TNF alpha inhibiting its contribution to granuloma formation (Russell et al 2013).

2 Summary of results

- This evidence review found one randomised controlled trial (RCT) comparing infliximab to placebo and seven uncontrolled studies with 15 or more patients. The uncontrolled studies consisted of one prospective study and six retrospective reviews. There were a large number of different outcomes reported, although none of these were reported across all of the studies.

- **The RCT** (Rossman et al 2006) found an improvement in pulmonary function (mean relative change in expected vital capacity 15% infliximab vs 8% placebo). However, there was no significant difference between the groups in the six-week randomised phase of the study. Dyspnea scores for functional impairment and magnitude of effort improved in both groups. No significance tests were reported. Radiologic improvement on chest x-ray was seen in 23% of the infliximab group. No patients in the placebo group showed improvement in chest x-ray in the six week randomised phase of the study but 33% showed improvement in the open label phase of the study when all patients received infliximab at weeks 6 and 14 with follow-up for 24 weeks. No significance tests between groups were reported. Quality of life scores on the SF-36 were similar in both the infliximab and placebo groups at baseline and after the six-week randomised phase.
- This RCT was closed early due to poor recruitment and was underpowered to detect any differences between the groups.
- Pulmonary function (four uncontrolled studies). Four studies reported improvements from baseline for pulmonary function outcomes in patients with pulmonary sarcoidosis. This improvement was statistically significant in two of the studies (Vorselaars et al 2015; van Rijswijk et al 2013) with mean changes from baseline of 6.6% predicted and 7.6% predicted for forced vital capacity (VC); 5.8% predicted and 7.9% predicted for forced expiratory volume in 1 second (FEV_1) and 4.1% predicted and 3.5% predicted for diffusing capacity of the lung for carbon monoxides, corrected for haemoglobin (DLCOc). In these studies approximately half of patients had an improvement of ≥10% in both forced VC and FEV_1. van Rijswijk et al (2013) reported that 37% patients with a pulmonary treatment indication had an improvement of ≥10% in DLCOc. In the third uncontrolled study (Russell et al 2013) the improvements in pulmonary function from baseline were not statistically significant and in the fourth uncontrolled study (Hostettler et al 2012) an improvement was seen but no statistical tests were reported. One study (Vorselaars et al 2015) reported the results of a 6-minute walking test with an improvement of 4.2 from a baseline of 61. No unit of measurement or significance test was reported.
- Physician assessed clinical response (five uncontrolled studies). The proportion of patients showing a response to infliximab varied from 58% (Russell et al 2013) to 94% (Chapelon-Abric et al 2015). Of these, the proportion showing a complete response varied from 18% (Jamilloux et al 2017) to 55% (Hostettler et al 2012) and the proportion showing a partial response varied from 26% (Russell et al 2013) to 56% (Chapelon-Abric et al 2015).
- Corticosteroid use (five uncontrolled studies). All five studies reported a reduction in corticosteroid use, with three studies reporting a statistically significant reduction in mean or median corticosteroid use from baseline. In two studies the mean reduction was 12 mg/day (Jamilloux et al 2017) and 8.8 mg/day (Vorselaars et al 2015). In Cohen Aubart et al (2017) the median dose was 50mg/day at baseline and 5mg/day at last visit. In a fourth uncontrolled study the median corticosteroid use reduced from 15 mg/day at baseline to 5.8 mg/day but this reduction was not statistically significant (Chapelon-Abric et al 2015). The fifth study (Russell et al 2013) reported that 73% patients discontinued corticosteroid use.
- Quality of life (two uncontrolled studies). The two uncontrolled studies assessing quality of life outcomes both reported statistically significant improvements from baseline. In Vorselaars et al (2015) the mean Patient Global Assessment score improved by 14.6 and the mean SF-36 score improved by 8.2. In van Rijswijk et al (2013) there was an improvement of 5.3 on a measure of fatigue severity (Checklist Individual Strength) and an improvement of 12.6 on the SF-36.

- Organ assessment (two uncontrolled studies). Both studies reporting organ assessment found a statistically significant improvement from baseline using a 6-point scale evaluating how severely individual organs were affected. One study (Jamilloux et al 2017) reported statistically significant improvements for the upper respiratory tract (by 0.5), central nervous system (by 1.2) and peripheral nervous system (by 0.9). The improvement for the lungs was not statistically significant. The second study (Chapelon-Abric et al 2015) assessed 24 index organs and reported a statistically significant improvement in median score of four points from baseline.
- Inflammatory response (two uncontrolled studies). Both Vorselaars et al (2015) and van Rijswijk et al (2013) reported a statistically significant improvement from baseline for all inflammatory response measures assessed. For F-fluorodeoxyglucose (FDG) by positron emission tomography (PET) maximum standardised uptake value (SUV_{max}) the mean improvement was 4.0 and 2.7 respectively for lung parenchyma, 3.0 and 2.3 for mediastinum and 5.8 (only assessed by Vorselaars et al) for index localisation. In Vorselaars et al the improvement in angiotensin-converting enzyme (ACE) was 28.2 and the improvement in soluble interleukin-2-receptor (sIL-2R) was 4,269. In van Rijswijk et al the improvement in ACE z score was 2.0 and the improvement in sIL-2R was 2,879.
- Patient-reported symptomatic response (one uncontrolled study). When all organs were included, some improvement was reported by 73% of patients (20% 'resolved' and 53% 'improved'). For lungs, an improvement was reported by 60% of patients (all 'improved'). For the central nervous system, an improvement was reported by 75% patients (25% 'resolved' and 50% 'improved') (Russell et al 2013).
- Composite overall response (one uncontrolled study). Composite overall response consisted of organ function, inflammation and quality of life (Vorselaars et al 2015). A response was reported in 96% of patients (40% 'excellent'; 39% 'good' and 17% 'partial').
- Modified Rankin score (one uncontrolled study). This score assessed degree of disability or dependence in daily activities (Cohen Aubart et al 2017). A statistically significant improvement of two points from baseline was reported.
- Infliximab trough levels (one uncontrolled study). There was no significant correlation between infliximab trough levels and response (Vorselaars et al 2015). The mean trough level was 18.0 µg/mL-1.
- **Safety.** In the RCT (Rossman et al 2006), four patients reported serious events (31%) in Group 1 (infliximab) and one (17%) in Group 2 (placebo). Adverse events were reported by 92% of Group 1 and 100% of Group 2. Infections were reported by 69% of Group 1 and 50% of Group 2. No significance tests comparing the two groups were reported.
- The reporting of adverse events varied in the uncontrolled studies. Four studies reported the proportion of patients who experienced adverse events which varied from 23% (Vorselaars et al 2015) to 58% (Jamilloux et al 2017). Five studies reported the proportion of patients who discontinued infliximab which ranged from 6% (van Rijswijk et al 2013) to 69% (Chapelon-Abric et al 2015).
- Four uncontrolled studies reported the percentage of patients who had infections which ranged from 19% (Russell et al 2013) to 44% (Chapelon-Abric et al 2015). Pneumonia and allergic reactions were the most commonly reported adverse events.
- **Cost –effectiveness.** No studies assessing the cost-effectiveness of infliximab for patients with refractory or progressive pulmonary and/or neurological sarcoidosis were identified.

- Infliximab was generally associated with improvements from baseline in uncontrolled studies on a range of outcome measures with these improvements often reaching statistical significance. The proportion of patients experiencing adverse events was generally fairly high.
- Overall, the evidence base is limited to one small, underpowered phase II RCT and uncontrolled, mostly retrospective studies, which are at risk of selection bias. The limitations of the evidence base limit the strength of the conclusions that can be drawn.

3 Methodology

- The methodology to undertake this review is specified by NHS England in their 'Guidance on conducting evidence reviews for Specialised Commissioning Products' (2016).
- A description of the relevant Population, Intervention, Comparison and Outcomes (PICO) to be included in this review was prepared by NHS England's Policy Working Group for the topic (see section 9 for PICO).
- The PICO was used to search for relevant publications in the following sources: PubMed, Embase and Cochrane Library (see section 10 for search strategy).
- The search dates for publications were between 1st January 2002 and 21st July 2017.
- The titles and abstracts of the results from the literature searches were assessed using the criteria from the PICO. Full text versions of papers which appeared potentially useful were obtained and reviewed to determine whether they were appropriate for inclusion. The higher quality papers which matched the PICO criteria were then selected for inclusion in this review.
- We excluded systematic reviews which included out of scope studies (e.g. Atkins et al 2017; Maneiro et al 2012), as well as studies where the population or intervention were out of scope, seeking confirmation from the CRG where necessary (e.g. Baughman et al 2006).
- One phase II randomised controlled trial met the criteria for inclusion. The other seven studies included are uncontrolled prospective and retrospective case series. As larger uncontrolled studies were available, studies with less than 15 patients are not included in this review.
- Evidence from all papers included was extracted and recorded in evidence summary tables, critically appraised and their quality assessed using the National Service Framework for Long Term Conditions (NSF-LTC) evidence assessment framework (see section 7).
- The body of evidence for individual outcomes identified in the papers was graded and recorded in grade of evidence tables (see section 8).

4 Results

A total of eight papers matching the PICO were included: one phase II randomised double-blind placebo controlled trial (Rossman et al 2006), one uncontrolled prospective study (Vorselaars et al 2015) and six uncontrolled retrospective studies (Jamilloux et al 2017; Cohen Aubart et al 2017; Chapelon-Abric et al 2015; van Rijswijk et al 2013; Russell et al 2013; Hostettler et al 2012).

The studies ranged in size from 16 to 132 participants with follow-up periods varying from 18 weeks to a median of 57 months. Full details of the study designs and outcomes are summarised in the evidence tables in section 7.

Clinical effectiveness

1. Is infliximab clinically effective in the treatment of refractory or progressive pulmonary and/or neurological sarcoidosis?

The outcomes reported in the studies included pulmonary function parameters, clinical response, corticosteroid use, organ assessment, inflammatory response, radiologic improvement on x-ray, dyspnea score, quality of life, patient-reported clinical response, composite overall response, modified Rankin score and infliximab trough levels. Further details of the outcomes reported are provided in the tables in sections 7 and 8.

Pulmonary function parameters

Pulmonary function parameters were reported in one RCT (Rossman et al 2006) and four uncontrolled studies (Vorselaars et al 2015; van Rijswijk et al 2013; Russell et al 2013; Hostettler et al 2012). Pulmonary function measures included forced vital capacity (VC), forced expiratory volume in 1 second (FEV_1) and diffusing capacity of the lung for carbon monoxides, corrected for haemoglobin (DLCOc)[1].

In the six-week randomised phase of the RCT there was no significant difference in mean relative change in expected VC for infliximab (15.2%) and placebo (8.4%) (p=0.65). The observed mean vital capacity was 2.5 litres at baseline and 2.7 litres at week 6 in Group 1 (infliximab) and 2.4 litres at both baseline and week 6 in Group 2 (placebo) (no significance test reported). In the phase 2 open label phase of the RCT, when all patients received infliximab, the percent expected VC ranged from 65.5% to 67.4% in Group 1 (infliximab) from a baseline of 59.6% and 70.7% to 72.5% in Group 2 (placebo and infliximab) from a baseline of 65.5%. The study authors combined the results for the first six weeks of taking infliximab for both groups (week 0 to 6 in Group 1 and week 6 to 12 in Group 2) and reported a significant improvement in mean VC ($p<0.02$)[2].

Four studies reported improvements from baseline for pulmonary function outcomes in patients with pulmonary sarcoidosis, although this improvement was not always statistically significant. Two of the four uncontrolled studies (Vorselaars et al 2015; van Rijswijk et al 2013) reported statistically significant improvements from baseline in percent predicted forced VC, FEV_1 and DLCOc ($p<0.05$) at 26 and 18 weeks follow-up respectively. In these studies the mean change was 6.6% predicted (baseline 73.6) and 7.6% predicted (baseline number not reported[3]) for forced VC; 5.8% predicted (baseline 55.8) and 7.9% predicted (baseline number not reported[3]) for FEV_1 and 4.1% predicted (baseline 56.6) and 3.5% predicted (baseline number not reported[3]) for DLCOc. Hostettler et al (2012) reported an improvement of 6% in forced VC percent predicted from baseline after a mean treatment duration of 29 months, but did not report a significance test. Hostettler et al also reported mean absolute forced VC which was 2.26 litres at baseline and 2.57 litres at follow-up (no significance test reported). However, the change from baseline was not statistically significant for forced VC (+4% from a baseline of 81%), FEV_1 (+2% from a baseline of 81%), DLCOc (0% from a baseline of 65%) or total lung capacity (+11% from a baseline of 75%) in the study by Russell et al (2013) which had a mean treatment duration of 46 months.

[1] As no minimal important difference for change in FVC has been defined (Vorselaars et al 2015) studies also report % predicted and percentage of patients with a specified increase e.g. 5%, 10% or 15%. An improvement in forced expiratory volume of ≥10% was considered to be a clinically relevant change for an individual patient (van Rijswijk et al 2013).
[2] Mean VC presented graphically by the authors but absolute values were not reported
[3] Mean values for the pulmonary sarcoidosis patients were presented graphically by the authors but absolute values were not reported.

The RCT and three of the uncontrolled studies also reported the proportion of patients who had achieved an improvement of ≥10% from baseline in patients with pulmonary sarcoidosis. In the RCT (Rossman et al 2006) three patients (23%) in the infliximab group had an improvement of ≥10% predicted VC compared to two patients (33%) in the placebo group. No significance test was reported. Two uncontrolled studies reported that 46% (Vorselaars et al 2015) and 52% (van Rijswijk et al 2013) of patients with a pulmonary treatment indication had an improvement of ≥10% in both forced VC percent predicted and FEV_1 percent predicted. Hostettler et al (2012) reported that 20% of patients with pulmonary involvement who received infliximab for >12 months had an improvement of ≥10% in forced VC percent predicted. In van Rijswijk et al (2013) 37% patients with a pulmonary treatment indication had an improvement of ≥10% in DLCOc. One study (Vorselaars et al 2015) reported an improvement of 4.2 on the 6-minute walking test from a baseline of 61. No unit of measurement or significance test was reported.

Clinical response

Clinical response, assessed by a physician, was reported in five uncontrolled studies (Jamilloux et al 2017; Cohen Aubart et al 2017; Chapelon-Abric et al 2015; Russell et al 2013; Hostettler et al 2012). The response definitions used for clinical response varied for each study so it is difficult to interpret or compare the results between studies. The proportion of patients who showed a response to infliximab varied from 58% (Russell et al 2013) to 94% (Chapelon-Abric et al 2015). Of these, the proportion with a complete response varied from 18% (Jamilloux et al 2017) to 55% (Hostettler et al 2012) and the proportion with a partial response varied from 26% (Russell et al 2013) to 56% (Chapelon-Abric et al 2015).

Corticosteroid use

The use of infliximab does appear to reduce the need for corticosteroid therapy with four studies reporting a reduction in mean or median corticosteroid use from baseline. In the two studies reporting mean reduction in corticosteroid use this was 12 mg/day from a baseline of 23 mg/day (median follow-up 21 months) (Jamilloux et al 2017) and 8.8 mg/day (baseline not reported, 26 weeks follow-up) (Vorselaars et al 2015) (both statistically significant, ($p>0.001$)). One study reported a statistically significant reduction in median corticosteroid use from 50 mg/day at baseline to 5 mg/day at last visit (median follow up 20 months) (Cohen Aubart et al 2017). In a fourth uncontrolled study the median corticosteroid use reduced from 15 mg/day at baseline to 8.5 mg/day at a median follow up of 47 months, but this difference was not statistically significant (Chapelon-Abric et al 2015). A fifth study (Russell et al 2013) reported that 73% patients discontinued corticosteroid use (average treatment duration 46 months).

Quality of life

Quality of life outcomes were reported in only one RCT and two uncontrolled studies (Vorselaars et al 2015; van Rijswijk et al 2013). In the six-week randomised phase of the RCT (Rossman et al 2006) quality of life scores (assessed by SF-36[4]) were similar in both groups at baseline (Group 1 26.7 vs. Group 2 26.4) and at the end of the 6-week randomised phase (Group 1 27.1 vs. Group 2 26.4). No significance tests were reported. Both uncontrolled studies reported a statistically significant improvement from baseline on quality of life measures ($p<0.05$). In Vorselaars et al (2015), the mean score of the Patient Global Assessment[5] improved by 14.6 from a baseline of 61.0 and the mean SF-36[4] score improved by 8.2 from a baseline of 40.6. In van Rijswijk et al (2013), there was an improvement of 5.3 from a baseline of 49.4 on the Checklist Individual Strength (a measure of fatigue severity[6]) and an improvement of 12.6 from a baseline of 30.9 on the SF-36[4].

4 The SF-36 is scored from 0-100 with higher scores indicating better functioning. An improvement of 10 points was considered clinically relevant (Vorselaars et al 2015).

5 The Patient Global Assessment (PGA) has a visual analogue scale ranging from 0 (best imaginable health status) to 100 (worst imaginable health status). An improvement pf 10 points was considered clinically relevant (Vorselaars et al 2015).

6 The Checklist Individual Strength has scores ranging from 8 (not fatigued) to 56 (severely fatigued)

Organ assessment

Organ assessment was reported in two uncontrolled studies (Jamilloux et al 2017; Chapelon-Abric et al 2015) using the Change in Extrapulmonary Physician Organ Severity Tool[7]. Jamilloux et al assessed individual organs and reported statistically significant improvements ($p<0.05$) between baseline and a median of 21 months follow-up for the upper respiratory tract (2.3 to 1.8), central nervous system (3.8 to 2.6) and peripheral nervous system (1.1 to 0.24). There was no significant improvement in the lungs (2.1 vs 1.9). Chapelon-Abric et al assessed 24 index organs and reported a statistically significant improvement in median score from 6 (range 3-12) at baseline to 2 (range 1-8) at a median follow-up of 57 months ($p<0.001$).

Inflammatory response

Inflammatory response was reported by two uncontrolled studies with both reporting statistically significant improvements. Vorselaars et al (2015) reported a statistically significant improvement from baseline for all of the inflammatory response measures assessed ($p<0.0003$). For F-fluorodeoxyglucose (FDG) by positron emission tomography (PET) maximum standardised uptake value (SUV_{max}) the mean improvement was 4.0 from a baseline of 6.6 for lung parenchyma, 3.0 from a baseline of 5.7 for mediastinum and 5.8 from a baseline of 9.0 for index localisation. For angiotensin-converting enzyme (ACE) the improvement after infliximab was 28.2 from a baseline of 89.7 and for soluble interleukin-2-receptor (sIL-2R) the improvement was 4,269 from a baseline of 8,824. van Rijswijk et al (2013) also reported a statistically significant response from baseline for all of the inflammatory response measures assessed ($p<0.0005$). For F-FDG PET SUV_{max} the mean improvement was 2.7 from a baseline of 4.3 for lung parenchyma and 2.3 from a baseline of 5.1 for mediastinum. For ACE z score the improvement after infliximab was 2.0 from a baseline of 2.6 and for sIL-2R the improvement was 2,879 from a baseline of 5,001.

Dyspnea score

Dyspnea score, assessed using the Baseline Dyspnea Index (BDI) and Transitional Dyspnea Index (TDI)[8], was reported in one RCT (Rossman et al 2006). The baselines scores assessed on the BDI were 2.17 for Group 1 and 2.08 for Group 2[9]. In the six-week randomised phase of the study both the infliximab and placebo groups showed an in improvement two of the three TDI domains[10] for mean functional impairment (0.38 vs. 0.17) and mean magnitude of effort (0.23 vs. 0.17). No change in magnitude of task was reported (figures not reported). No significance tests were reported.

Clinical response (patient reported)

Patient-reported symptomatic response was reported by one uncontrolled study (Russell et al 2013) using the response categories 'resolved', 'improved', 'unchanged' and 'progressed'. When all organs were included, some improvement was reported by 73% of patients (20% 'resolved' and 53% 'improved'). Of those that had not seen an improvement, 23% were 'unchanged' and 5% had 'progressed'. For lungs, an improvement was reported by 60% of patients, all of whom were 'improved'. For the central nervous system, an improvement was reported by 75% patients (25% 'resolved' and 50% 'improved').

7 The Extrapulmonary Physician Organ Severity Tool (ePOST) examines 17 extrapulmonary organs and assigns each a score from 0 (not affected) to 6 (very severely affected) to each organ. Intermediate scores are slight (1), mild (2), moderate (3), moderate to severe (4) and severe (5).

8 The Baseline Dyspnea Index (BDI) and Transitional Dyspnea Index (TDI) are used to assess breathlessness and the impact of intervention. The BDI and TDI have three domains: functional impairment, which determines the impact of breathless on the ability to carry out activities; magnitude of task, which determines the type of task that causes breathlessness; and magnitude of effort, which establishes the level of effort that results in breathlessness.

9 The BDI is scored from 0 (very severe impairment) to 4 (no impairment) for the three domains and summed to create a focal score (0-12).

10 The TDI is scored from -3 (major deterioration) to +3 (major improvement) for each domain and summed to create a focal score (-9 to +9).

Composite overall response
A composite overall response measure was devised by the authors in one study which included organ function, inflammation and quality of life[11] (Vorselaars et al 2015). A response was reported in 96% of patients, of which 40% showed an 'excellent' response, 39% a 'good' response and 17% a 'partial' response.

Modified Rankin score
Modified Rankin score, assessing degree of disability or dependence in daily activities[12], was reported in one uncontrolled study (Cohen Aubart et al 2017). A statistically significant improvement was reported from 3 at baseline to 1 at a median follow-up of 20 months ($p<0.0001$).

Infliximab trough levels
Infliximab trough levels were reported in one uncontrolled study (Vorselaars et al 2015). The mean trough level was 18.0μg/mL^{-1}. There was no significant correlation between trough level and response.

Safety

2. Is infliximab safe to use in the treatment of patients with refractory or progressive pulmonary and/or neurological sarcoidosis?

Adverse events were reported in one RCT and seven uncontrolled studies. In the RCT, serious adverse events were experienced by four patients (31%) in Group 1 (infliximab) and one patient (17%) in Group 2 (placebo). Adverse events were reported by 92% of Group 1 and 100% of Group 2. Infections were reported by 69% of Group 1 and 50% of Group 2. No significance tests comparing the two groups were reported.

The reporting of adverse events varied in the uncontrolled studies. The proportion of patients who experienced adverse events was reported in four studies and varied from 23% (Vorselaars et al 2015) to 58% (Jamilloux et al 2017). The proportion of patients who discontinued infliximab was reported in five studies and ranged from 6% (van Rijswijk et al 2013) to 69% (Chapelon-Abric et al 2015). The percentage of patients who had infections was reported in four studies and ranged from 19% (Russell et al 2013) to 44% (Chapelon-Abric et al 2015). Pneumonia and allergic reactions were the most commonly reported adverse events.

Cost effectiveness

3. Is infliximab a cost-effective treatment option for use in patients with refractory or progressive pulmonary and/or neurological sarcoidosis?
No studies assessing the cost-effectiveness of infliximab for patients with refractory or progressive pulmonary and /or neurological sarcoidosis were identified.

[11] Improvement in a category was scored only when one of the parameters improved significantly without deterioration of the others. A good or excellent response was a clinically relevant improvement in 2 or 3 categories, a partial response was a clinically relevant improvement in one category and no improvement in any category was a nonresponse.

[12] The modified Rankin score measures the degree of disability or dependence in the daily activities of people who have suffered a stroke or other causes of neurological disability. It is scored from 0 (no symptoms at all) to 6 (dead). A score of 3 is defined as moderate disability; requiring some help, but able to walk without assistance. A score of 1 is defined as no significant disability despite symptoms; able to carry out all usual duties and activities

5 Discussion

One phase II randomised double blind controlled trial compared infliximab to placebo. Improvements were reported in both groups of patients in this study, however no significant differences between the groups were reported. This trial closed early due to poor recruitment and was therefore underpowered to detect a difference between the groups. The results of this study should be treated with caution.

The seven uncontrolled studies varied in size from 16 to 132 participants with follow-up periods varying from 18 weeks to a median of 57 months. Five of the seven studies had average follow-up periods of more than 20 months.

The uncontrolled studies generally reported improvements from baseline with infliximab on a range of outcome measures. In the five studies reporting clinical response, the majority of patients were judged to have shown an improvement by a clinician and in one study 73% of patients reported an improvement in their symptoms. When measures of quality of life, inflammatory response and severity of organs affected were reported a statistically significant improvement was seen. Improvements in pulmonary function and reductions in corticosteroid use were also reported in multiple studies but these did not always reach statistical significance. The clinical meaningfulness of the improvements reported was not always clear.

The proportion of patients experiencing adverse events with infliximab was generally fairly high, as was the proportion of patients experiencing infections and discontinuing infliximab.

The populations in the uncontrolled studies were not always restricted to patients with only pulmonary and/or neurosarcoidosis. However they reported outcomes, such as pulmonary function parameters, separately for patients with a pulmonary treatment indication or reported outcomes for individual organs.

The only comparative study identified for the population of interest was published in 2006 and more recently published evidence for patients with refractory or progressive pulmonary and/or neurosarcoidosis is from uncontrolled studies only. It is not clear whether evidence from higher quality studies is likely to be published in the future.

6 Conclusion

The evidence identified for infliximab for refractory or progressive pulmonary sarcoidosis and/or neurosarcoidosis included one phase II randomised controlled study, one uncontrolled prospective study and six uncontrolled retrospective studies.

The RCT compared infliximab to placebo but failed to recruit enough participants leading to its early closure. The results of this RCT therefore do not provide sufficient evidence to draw conclusions about the comparative effectiveness of infliximab versus placebo.

Infliximab was generally associated with improvements from baseline in uncontrolled studies on a range of outcome measures with these improvements often reaching statistical significance.

Overall, the evidence base is limited to a small, underpowered phase II RCT and uncontrolled studies, most of which were retrospective studies which are at risk of selection bias. The limitations of the evidence base limit the strength of any conclusions that can be drawn.

7 Evidence Summary Table

For abbreviations see list after each table

Use of Infliximab Vs. Placebo for Sarcoidosis									
Study reference	**Study design**	**Population characteristics**	**Intervention**	**Outcome measure type**	**Outcome measures**	**Results**	**Quality of evidence score**	**Applicability**	**Critical appraisal summary**
Rossman et al 2006	P1 – Randomised, double-blind, placebo controlled phase II trial 5 US sites	Patients with: • Active symptomatic pulmonary sarcoidosis • Previous or current treatment with corticosteroids with a need for institution of another agent based on either suboptimal response or intolerance to corticosteroids • Stage II, III or IV[13] pulmonary parenchymal involvement on chest radiography • Functional abnormalities (defined by vital capacity ≤50% to ≤80% predicted using race-corrected standards) • Patients being treated with oral	N=19 Group 1: Infliximab 5mg/kg (n=13) Group 2: Placebo (n=6) Patients were stratified by corticosteroid use (current vs. prior) and randomised 2:1 infliximab: placebo The study included 2 phases. In the randomised phase 1, patients received infliximab or placebo at	Primary Clinical effectiveness	Pulmonary function parameters	**Phase 1 (randomised) –from baseline to week 6** **Mean ± SD relative % change in expected vital capacity** Group 1 (infliximab): 15.2% ± 9.9% Group 2 (placebo): 8.4% ± 3.3% No significant difference between the groups (p=0.65) **Mean ± SD observed vital capacity (litres)** Group 1 baseline: 2.5 ± 0.2 Group 1 week 6: 2.7 ± 0.2 Group 2 baseline: 2.4 ± 0.3 Group 2 week 6: 2.4 ± 0.3 No significance test reported **Mean ± SD vital capacity % expected** Group 1 baseline: 59.6 ± 3.7 Group 1 week 6: 64.7 ± 3.6 Group 2 baseline: 65.5 ± 3.0 Group 2 week 6: 67.7 ± 3.3 No significance test reported 2 patients (15%) in Group 1 had an improvement of ≥15% predicted vital capacity compared to none (0%) in Group 2. Significance test not reported	5	Direct	This phase II RCT compared infliximab to placebo. The investigators planned to recruit 42 patients with an 80% power to detect a mean relative change of 8.5% from baseline to week 6 in the infliximab group compared to placebo. Due to poor enrolment only 19 patients were recruited and the study was closed early. The study was therefore underpowered to detect any differences between the groups and its results should be treated with caution. 16 patients completed the study. 1 patient died during the study, 1 withdrew after a serious adverse event and 1 withdrew consent after week 8.

[13] Stage II is bilateral hilar lymphadenopathy (BHL); stage III is pulmonary infiltrates without BHL; stage IV is pulmonary fibrosis with honeycombing

		corticosteroids must have been treated for at least 3 months and have had a stable dose of prednisone up to 60mg/d (or equivalent corticosteroid) for at least 2 weeks prior to screening Patients were excluded if they had: • Arterial pO2 ≤55mg Hg at rest (if measured in the last 3 months) or oxygen saturation by pulse oximetry ≤88% at rest (if arterial blood gas has not been performed in the last 3 months) • Signs or symptoms of severe, progressive or uncontrolled renal, hepatic, hematologic, endocrine, cardiac or neurological disease • Previous or current treatment with infliximab • Immunosuppressive agents or other TNF-inhibiting agents within 4 weeks prior study treatment	weeks 0 and 2 and were assessed at week 6. In the open label phase 2, patients received infliximab at week 6 and 14 and were then assessed at weeks 22 and 38 Patients and site investigators were blinded to study group during the 6 week randomised phase			3 patients (23%) in Group 1 had improvement of ≥10% predicted vital capacity compared to 2 patients (33%) in Group 2. Significance test not reported **Phase 2 (open label) - from week 12 to week 38** **Mean ± SD observed vital capacity (litres)** Group 1 week 12: 2.7 ± 0.2 Group 1 week 14: 2.8 ± 0.2 Group 1 week 22: 2.7 ± 0.2 Group 1 week 38: 2.7 ± 0.2 Group 2 week 12: 2.6 ± 0.3 Group 2 week 14: 2.6 ± 0.3 Group 2 week 22: 2.6 ± 0.3 Group 2 week 38: 2.6 ± 0.2 No significance test reported **Mean ± SD vital capacity % expected** Group 1 week 12: 66.7 ± 3.9 Group 1 week 14: 67.4 ± 3.7 Group 1 week 22: 65.6 ± 4.4 Group 1 week 38: 65.5 ± 2.8 Group 2 week 12: 70.7 ± 2.6 Group 2 week 14: 72.2 ± 3.3 Group 2 week 22: 72.5 ± 4.2 Group 2 week 38: 72.5 ± 6.1 No significance test reported **Combined results for the first 6 weeks of taking infliximab (week 0 to 6 in Group 1 and week 6 to 12 in Group 2)** Significant improvement in change in mean vital capacity ($p<0.02$). Mean figures not reported.			
				Secondary Clinical effectiveness	Radiologic improvement on chest x-ray	Chest x-rays were assessed as 'markedly worse', 'slightly worse', 'unchanged', 'slightly improved' or 'markedly improved'. **Percentage showing 'radiologic improvement'** **From baseline to week 6** Group 1: 23.0% Group 2: 0%			

						At week 12 Group 1: 30.8% Group 2: 33.3% No significance tests reported.			
				Secondary Clinical effectiveness	Dyspnea scores	**At baseline (mean ± SD)** Group 1: 2.17 ± 0.40 Group 2: 2.08 ± 0.18 **Functional impairment at week 6 (mean ± SD)** Group 1: 0.38 ± 0.21 Group 2: 0.17 ± 0.17 **Magnitude of effort at week 6 (mean ± SD)** Group 1: 0.23 ± 0.26 Group 2: 0.17 ± 0.17 No change in magnitude of task (figures not reported) No significance tests reported.			
				Secondary Clinical effectiveness	Quality of life	**Mean ±SD SF-36 (scores range from 0-100 with higher scores indicating better functioning)** Group 1 baseline: 26.7 ± 0.5 Group 1 week 6: 27.1 ± 0.5 Group 2 baseline: 26.4 ± 0.8 Group 2 week 6: 26.4 ± 0.8 No significance test reported.			
				Safety	Adverse events	**Serious adverse events (SAE)** Group 1: 31% (n=4) Group 2: 17% (n=1) 1 patient in the infliximab group died during the study SAEs with infliximab: • Right leg cellulitis (n=1) • Acute renal failure (n=1) • Pulmonary emboli (n=1) • Cellulitis (n=1) • Decreased white blood cell count (n=1) • Elevated creatine phosphokinase (n=1) • Pneumonia (n=1) • Visual field defect (n=1)			

						SAEs with placebo: • Severe shortness of breath (n=1) **Adverse events:** Group 1: 92% Group 2: 100% **Infections:** Group 1: 69% Group 2: 50% **Discontinuation of treatment:** Group 1: 15% Group 2: 17% No significance tests reported.			

RCT – Randomised Controlled Trial; SAE – Serious Adverse Event; SD – Standard Deviation

Use of Infliximab for Sarcoidosis (No Comparator)									
Study reference	**Study design**	**Population characteristics**	**Intervention**	**Outcome measure type**	**Outcome measures**	**Results**	**Quality of evidence score**	**Applicability**	**Critical appraisal summary**
Jamilloux et al 2017	S2 – Retrospective multicentre review of patient records 25 medical departments in France 2014 to 2015	Patients with refractory sarcoidosis, a prescription of TNF antagonist and outcomes data available in patient records. Sarcoidosis was considered refractory when disease was not controlled despite prior use of ≥ 1 immunosuppressive drug.	N = 132 120 patients (91%) received infliximab as a 1st line anti-TNF. 5 (4%) patients received infliximab as a 2nd line anti-TNF. Other anti-TNFs used as 1st line treatment included adalimumab (n=8),	Primary Clinical effectiveness	Organ assessment	**Change in Extrapulmonary Physician Organ Severity Tool (ePOST). Scores range from 0 (not affected) to 6 (very seriously affected)** Lung (n=90): • Baseline 2.1 • Follow-up 1.9 No significant difference (p=0.46) Upper respiratory tract (n=25): • Baseline 2.3 • Follow-up 1.8 Significant improvement (p=0.04) CNS (n=63): • Baseline 3.8 • Follow-up 2.6	7	Direct	This uncontrolled retrospective review included patients from 25 centres and had a relatively large sample size. Patients had a range of clinical manifestations for sarcoidosis including pulmonary and/or CNS involvement. A small number of patients included in the sample (5%) did not receive infliximab as the anti-TNF treatment. The retrospective design of the study introduces the possibility of selection bias in the study population. This can result from the patients included in the analysis and the classification of patients from patient records. As the study does not include a comparator it is not possible to compare the outcomes for these patients with patients receiving alternative treatments.

		Pulmonary stage: 0 – 47 (36%) I - 38 (29%) II – 29 (22%) III – 7 (5%) IV – 11 (8%) 63 (47%) of patients had CNS clinical manifestations. 11 patients had possible neurosarcoidosis.	etanercept (n=3) and certolizumab pergola (n=1). 11 patients switched from infliximab to adalimumab as 2nd line treatment. Infliximab was given at 3-5mg/kg body weight at weeks 0, 2, 6 and then every 4-8 weeks. 113 patients (86%) received a concomitant corticosteroid. 97 patients (74%) received a concomitant immunosuppressant. Median treatment duration was 12 months. Median follow-up was 20.5 months (IQR 8 to 48).			Significant improvement (p=0.001) Peripheral nervous system (n=23) • Baseline 1.1 • Follow-up 0.24 Significant improvement (p=0.03) Specific skin lesions (n=44): • Baseline 3.4 • Follow-up 1.5 Significant improvement (p=0.004) Heart (n=28): • Baseline 2.5 • Follow-up 2.0 Significant improvement (p=0.02) Muscle (n=21) • Baseline 1.6 • Follow-up 1.3 No significant difference (p=0.19) Eye (n=25): • Baseline 3.9 • Follow-up 2.9 No significant difference (p=0.09)			
				Secondary Clinical effectiveness	Clinical response (physician evaluation)	**Responders:** • Complete response (disappearance of clinical signs (excluding sequelae) upon use of corticosteroids <10mg): 24 (18%) • Partial response (an improvement of clinical and para-clinical parameters upon >50% reduction of the initial corticosteroids dose): 61 (46%) **Non-responders:** • Stable (non-responders with no change): 33 (25%) • Progressive disease (non-responders with either new			

						organ involvement, worsening of an organ involvement or need for increased corticosteroid dosage): 14 (11%) In multivariate analysis comparing responders with non-responders, pulmonary involvement was associated with lower clinical response to anti-TNF (OR =0.38, 95%CI 0.14 to 0.92). No other parameters were significantly different between responders and non-responders.			
				Secondary Clinical effectiveness	Corticosteroid use	**Mean ± SD corticosteroid dose** Baseline: 23 ± 20 mg/day Follow-up: 11 ± 11 mg/day Significant improvement from baseline (p<0.001)			
				Safety	Adverse events	69 patients (52%) experienced an adverse event 31 patients (23%) experienced adverse events requiring treatment cessation 47 patients (36%) had an infection, with 25 (19%) requiring hospitalisation or treatment interruption. Infections included: • Pneumonia (n=28) • Bacterial sepsis (n=24) • Recurrent urinary tract infection (n=15) • Herpes zoster virus (n=5) • Legionellosis (n=1) • Invasive aspergillosis (n=1) • Pneumocystis pneumonia (n=1) • Cytomegalovirus primary infection (n=1) • Cryptococcosis (n=1) • Hepatitis B reactivation (n=1) • Non-tuberculosis mycobacterial infection (n=1)			

						9 patients (7%) had a severe allergic reaction. 4 patients (3%) paradoxical granulomatous reactions. 3 patients (2%) developed antibodies against anti-TNF. 2 patients (2%) had demyelinating lesions. 1 patient (1%) had serum sickness-like reaction. 1 patient (1%) experienced neoplasia during follow-up. 3 patients died, but these were not related to anti-TNF use.			
Cohen Aubart et al 2017	S2 – Retrospective multicentre review 3 hospitals in France 2010 to 2015	Patients with definite or probable refractory neurosarcoidosis who received ≥ 1 infusion of infliximab. All patients had received steroids with ≤1 immunosuppressive drug in 16 patients.	N=18 Infliximab was given at 3-7.5mg/kg body weight. All patients received 2 infusions at 2 week intervals, then 15 patients received an infusion every 4 weeks and 2 patients every 6 weeks for 6 months. After 6 months infusions were received every 4 to 8 weeks. 2 patients switched to adalimumab; 1 after 2 infliximab infusions and	Primary Clinical effectiveness	Clinical response (physician evaluation)	Clinical response assessed with the Extrapulmonary Physician Organ Severity Tool with scores ranging from 0 (not affected) to 6 (very seriously affected).A complete response was a score of 0. A partial remission was an improvement of ≥1 point. **At 6 months follow-up** • Complete neurological remission: 6 (33%) • Partial neurological remission: 10 (56%) • Stable: 2 (11%) **At final follow-up** • Complete neurological remission: 5 (31%) • Partial neurological remission: 10 (63%) • Stable: 1 (6%) 1 patient died during follow-up (unknown cause) and 1 patient was lost to follow-up.	7	Direct	This uncontrolled retrospective review included patients from 3 centres over a 5 year period but the sample of patients is small. The number of infusions and the dose of infliximab varied between patients. The retrospective design of the study introduces the possibility of selection bias in the study population. This can result from the patients included in the analysis and the classification of patients from patient records. As the study does not include a comparator it is not possible to compare the outcomes for these patients with patients receiving alternative treatments.
				Secondary Clinical effectiveness	Corticosteroid use	**Median dose corticosteroid mg/day** Baseline: 50 (range not reported) Follow-up: 5 (range 0 to 12.5)			

			1 after 6 infusions. All patients received concomitant corticosteroid and an immunosuppressive drug Median follow-up was 20 months (range 6 to 93).			Significant improvement from baseline (p<0.0001)			
				Secondary Clinical effectiveness	Modified Rankin score	Baseline: 3 Follow-up: 1 Significant improvement from baseline (p<0.0001)			
				Safety	Adverse events	8 patients (44%) experienced toxic side effects. 7 patients (39%) had an infection including: • Pulmonary infection requiring hospitalisation (n=5) • Cellulitis (n=1) • Cytomegalovirus primo-infection with fever and cytolysis (n=1) 1 patient experienced severe alopecia			
Vorselaars et al 2015	P1 – prospective open label study 1 centre, The Netherlands 2011 to 2013	Patients with severe sarcoidosis, unresponsive to 1st or 2nd line treatment or with severe side effects from these treatments. 34 (61%) had pulmonary sarcoidosis. 3 (5%) had CNS as the main indication. 93% had used ≥ 2 immunosuppressant drugs prior to infliximab. Disease severity was judged by the treating	N=56 Patients received 8 infusions of infliximab at 5mg/kg starting at week 0 and 2 and then every 4 weeks for 6 months. 19 patients had concomitant corticosteroids at the start of infliximab therapy. 8 patients discontinued infliximab before the 26 week follow-up.	Primary Clinical effectiveness	Pulmonary function parameters	**In patients with a pulmonary treatment indication and 26 week follow up (n=28)** **Forced vital capacity % predicted** Baseline:73.6 Change after infliximab: +6.6 Significant improvement from baseline (p=0.0007) 71% patients had improvement of ≥5% predicted 46% patients had improvement of ≥10% predicted. **Forced Expiratory volume in 1s % predicted** Baseline:55.8 Change after infliximab: +5.8 Significant improvement from baseline (p<0.0001) 64% patients had improvement of ≥5% predicted 46% patients had improvement of ≥10% predicted. **Diffusing capacity for carbon monoxide corrected for haemoglobin % predicted**	7	Direct	This uncontrolled prospective study included patients from a single centre over 2 years and had a relatively large sample size. Patients had a range of clinical manifestations for sarcoidosis including pulmonary and/or CNS involvement. The prospective design of the study reduces the possibility of selection bias in the study population. As the study does not include a comparator it is not possible to compare the outcomes for these patients with patients receiving alternative treatments.

		physician based on loss of function, impaired quality of life and disease activity on F-fluorodeoxyglucose (FDG) by positron emission tomography (PET).				Baseline:56.6 Change after infliximab: +4.1 Significant improvement from baseline (p=0.001) **6-minute walking distance % predicted** Baseline: 61 Change after infliximab: +4.2 Unit of measurement not reported. Significance test not reported **In patients with an extrapulmonary treatment indication (n not reported):** **Forced vital capacity % predicted** Baseline not reported Change after infliximab: +3.9 Significant improvement from baseline (p=0.027) 37% patients had improvement of ≥5% predicted **Forced Expiratory volume in 1s % predicted** Baseline not reported Change after infliximab: +3.5 Significant improvement from baseline (p=0.034) 37% patients had improvement of ≥5% predicted			
				Primary Clinical effectiveness	Inflammatory response	**Maximum standardised uptake value (SUV) on F-FDG PET lung parenchyma (mean ± SD)** Baseline: 6.6 ± 5.3 Change after infliximab: -4.0 Significant improvement from baseline (p<0.0001) **Maximum SUV on F-FDG PET mediastinum (mean ± SD)** Baseline: 5.7 ± 3.2 Change after infliximab: -3.0 Significant improvement from baseline (p<0.0001)			

						Maximum SUV on F-FDG PET index localisation (mean ± SD) Baseline: 9.0 ± 5.2 Change after infliximab: -5.8 Significant improvement from baseline (p<0.0001) **Angiotensin-converting enzyme (ACE) U.L^{-1}** Baseline: 89.7 ± 49.7 Change after infliximab: -28.2 Significant improvement from baseline (p=0.0003) **Soluble interleukin-2 receptor** (pg.mL^{-1}) Baseline: 8824 ± 8503 Change after infliximab: -4269 Significant improvement from baseline (p<0.0001) Linear regression analysis found significant correlations between change in pulmonary function and level of disease activity.			
				Primary Clinical effectiveness	Corticosteroid use	**Mean dose corticosteroid** Baseline not reported Change after infliximab: -8.8mg Significant improvement from baseline (p<0.001)			
				Primary Clinical effectiveness	Quality of life	**Mean Patient Global Assessment with scores on a visual analogue scale ranging from 0 (best imaginable health status) to 100 (worst imaginable health status)** Baseline: 61.0 Change after infliximab: -14.6 Significant improvement from baseline (p<0.0001) **Mean SF-36 (with scores ranging from 0-100 with higher scores indicating better functioning)** Baseline: 40.6 Change after infliximab: 8.2 Significant improvement from baseline (p=0.009).			

				Primary Clinical effectiveness	Composite overall response (organ function, inflammation and quality of life)	Functional category response: 69% Inflammation category response: 79% Quality of life category response: 67% • Excellent response (marked improvement in all 3 categories): 40% • Good response (marked improvement in 2 categories): 39% • Partial response (marked improvement in 1 category): 17% • No response: 4%.			
				Primary Clinical effectiveness	Infliximab trough levels	Mean trough level: 18.0µg/mL^{-1} No significant correlation between trough level and response.			
				Safety	Adverse events	34 of the 56 patients in this study (61%) experienced no side effects. Severe side effects (n=7, 12.5%): • Pneumonia requiring hospitalisation (n=3; 2 died of respiratory failure) • Peritonitis (n=1) • Severe gastrointestinal complaints (n=1) • Allergic reaction with antibody formation (n=2). Other side effects (n=13, 23%): • Mild infection upper or lower respiratory tract not requiring hospitalisation (n=5) • Headache (n=2) • Dizziness (n=1) • Oedema (n=3) • Joint pain (n=2). 1 patient stopped treatment for undisclosed reasons.			

Chapelon-Abric et al 2015	S2 – Retrospective review I centre in France 2005 to 2013	Patients with severe sarcoidosis (≥3 organ involvement), refractory to steroid and ≤ 1 immunosuppressive therapy. Organ involvement before infliximab CNS: 12 Cardiac: 4 Pulmonary: 3 Cutaneous: 1 Ocular: 1 Intra-abdominal lymph nodes: 1 Liver: 1 Joint: 1	N=16 Patients received infliximab at 3-5mg/kg at weeks 0,2 and 6, then every 8 weeks. In patients with suboptimal response the dose was increased to 7.5mg/kg every 4-6 weeks. Mean number of infliximab infusions 15 (range 2-42). 10 patients received concomitant corticosteroids or immunosuppressants. Median follow-up 57 months (range 2-91).	Primary Clinical effectiveness	Clinical response	Complete response (absence of functional signs and a normalisation of imaging) : 6 (38%) Partial response (persistence of physical signs and/or abnormal imaging): 9 (56%) No response (absence of clinical and/or laboratory improvement): 1 (6%)	6	Direct	This uncontrolled retrospective review included patients from 1 centre over an 8 year period but the sample of patients is small. Patients had a range of clinical manifestations for sarcoidosis including pulmonary and/or CNS involvement. The number of infusions and the dose of infliximab varied between patients. The retrospective design of the study introduces the possibility of selection bias in the study population. This can result from the patients included in the analysis and the classification of patients from patient records. As the study does not include a comparator it is not possible to compare the outcomes for these patients with patients receiving alternative treatments.
				Primary Clinical effectiveness	Organ assessment	Number of index organs: 24 Index organs assessed on a scale from 0 (not affected) to 6 (very severe) **Median (range) severity score** Baseline: 6 (3-12) Follow-up: 2 (1-8) Significant improvement (p<0.001)			
				Primary Clinical effectiveness	Corticosteroid use	**Median (range) dose (mg/day)** Baseline: 15 (0-30) Follow-up: 8.5 (0-30) No significant difference (p=0.11)			
				Safety	Adverse events	11 patients (69%) discontinued infliximab. This was due to: Acute infections (n=7; 44%) • Pulmonary infection (n=2) • Recurrent prostatitis (n=1) • Aspergillosis infection (n=1) • Atypic mycobacterium disease (n=1) • Bacterial sepsis (n=1) • Multiviral reactivation (n=1) Paradoxical granulotomas (n=1) Leucoencephalopathy (n=1) Non responsive neurosarcoidosis (n=1) Drop out (n=1)			

van Rijswijk et al 2013	S2 – Retrospective review 1 hospital in The Netherlands 2004-2010	Sarcoidosis patients with chronic disease activity who were refractory to corticoid and/or corticoid-sparing treatment or had severe side effects to this medication. Pulmonary indication: 23 Neurosarcoidosis: 9.	N=48 All patients received infliximab at 5mg/kg at weeks 0,2,6,10,14 and 18. Outcome data from 45 who completed 6 cycles of infliximab were reported. 41 (91%) had concurrent immunosuppressive medication. Efficacy assessed at week 18.	Primary Clinical effectiveness	Pulmonary function parameters	**All patients** **Mean ± SD vital capacity % predicted** Baseline: 85.7 ± 19.0 Change from baseline: +5.4 ± 7.6 Significant improvement (p<0.0001) **Mean ± SD forced expiratory volume in 1s % predicted** Baseline: 75.3 ± 22.9 Change from baseline: +5.3 ± 8.3 Significant improvement (p<0.001) **Mean ± SD diffusing capacity for carbon monoxide corrected for haemoglobin % predicted** Baseline: 66.7 ± 18.7 Change from baseline: +3.1 ± 7.3 Significant improvement (p=0.01) **Patients with pulmonary indication (n=23)** **Vital capacity % predicted** Baseline figures not reported Change from baseline: +7.6 Significant improvement (p<0.0001) 52% had a relative improvement of ≥ 10% **Forced expiratory volume in 1s % predicted** Baseline figures not reported Change from baseline: +7.9 Significant improvement (p<0.0001) 52% had a relative improvement of ≥ 10%	7	Direct	This uncontrolled retrospective review included patients from a single centre over 6 years and had a relatively large sample size. Patients had a range of clinical manifestations for sarcoidosis including pulmonary and/or CNS involvement. The authors considered an improvement in forced expiratory volume of ≥10% to be a clinically relevant change for an individual patient. The authors noted that the minimal clinically significant change for vital capacity in a sarcoidosis patient is not well established. The retrospective design of the study introduces the possibility of selection bias in the study population. This can result from the patients included in the analysis and the classification of patients from patient records. As the study does not include a comparator it is not possible to compare the outcomes for these patients with patients receiving alternative treatments.

						Diffusing capacity for carbon monoxide corrected for haemoglobin % predicted Baseline figures not reported Change from baseline: +3.5 Significant improvement (p=0.01) 37% had a relative improvement of ≥ 10%			
				Primary Clinical effectiveness	Inflammatory response	N= 40 **Maximum standardised uptake value (SUV) on F-FDG PET lung parenchyma (mean ± SD)** Baseline: 4.3 ± 3.6 Change from baseline: -2.7 ± 3.4 Significant improvement (p<0.00005) **Maximum SUV on F-FDG PET mediastinum (mean ± SD)** Baseline: 5.1 ± 3.9 Change from baseline: -2.3 ± 3.4 Significant improvement (p<0.0005) **Angiotensin-converting enzyme z-score U/ml (mean ± SD)** Baseline: 2.6 ± 3.9 Change from baseline: -2.01 ± 3.31 Significant improvement (p=0.0005) **Soluble interleukin-2 receptor pg/ml (mean ± SD)** Baseline: 5001 ± 3919 Change after infliximab: -2879 ± 3755 Significant improvement (p<0.00001)			
				Primary Clinical effectiveness	Quality of life	N=28 **Mean ± SD fatigue severity (Checklist Individual Strength, with scores ranging from 8 (not fatigued) to 56 (severely fatigued))** Baseline: 49.4 ± 9.2			

						Change after infliximab: -5.3 ± 8.5 Significant improvement (p=0.003) **Mean ± SD physical functioning (SF-36 with scores ranging from 0-100 with higher scores indicating better functioning)** Baseline: 30.9 ± 22.2 Change after infliximab: 12.6 ± 23.9 Significant improvement (p=0.01)			
				Safety	Adverse events	3 patients (6%) discontinued infliximab due to: Allergic reaction to infliximab (n=1) Progression of dyspnea (n=1) Severe methotrexate hepatitis (n=1) 1 patient was hospitalised due to pneumonia.			
Russell et al 2013	S2 – Retrospective review 1 US hospital 2000 to 2010	Adults with single or multi-organ sarcoidosis who received infliximab due to failure or intolerance of conventional therapy. Conventional therapy was defined as corticosteroid and/or disease modifying anti-rheumatic drug therapy such as methotrexate. 58% had lung involvement and 31% CNS involvement.	N=26 Average maximum infliximab dose 511mg (range 300 to 1000). Average dosing frequency 5.5 weeks (range 4-6 weeks). Details of individual dosage and number of infusions not reported. 62% had concurrent corticosteroids. 58% and 46% of	Primary Clinical effectiveness	Clinical response (physician evaluation)	• 'Resolved' = complete resolution of clinical disease activity • 'Improved' = organs with reduced sarcoid burden or reduced frequency in disease activity but still with evidence of disease • 'Unchanged' = disease activity clinically no different than prior to infliximab • 'Progressed' = clinical features of progressive disease despite infliximab **All organs (n=26)** • Resolved: 32% • Improved: 26% • Unchanged: 36% • Progressed: 6% **Lung (n=15)** • Resolved: 4 (27%) • Improved: 4 (27%) • Unchanged: 5 (33%)	6	Direct	This uncontrolled retrospective review included patients from 1 centre over a 10 year period but the sample of patients is small. Patients had a range of clinical manifestations for sarcoidosis including pulmonary and/or CNS involvement. Detail on the number of infusions and the dose of infliximab were not reported. The retrospective design of the study introduces the possibility of selection bias in the study population. This can result from the patients included in the analysis and the classification of patients from patient records. As the study does not include a comparator it is not possible to compare the outcomes for these patients with patients receiving alternative treatments.

			patients received concurrent methotrexate and hydroxychloroquine respectively. Average duration of infliximab therapy was 46.2 months.			• Progressed: 2 (13%) **CNS (n=8)** • Resolved: 3 (38%) • Improved: 2 (25%) • Unchanged: 3 (38%) • Progressed: 0 (0%) **Skin (n=11)** • Resolved: 4 (36%) • Improved: 7 (64%) • Unchanged: 0 (0%) • Progressed: 0 (0%) **Lymph node (n=11)** • Resolved: 2 (18%) • Improved: 1 (9%) • Unchanged: 7 (64%) • Progressed: 1 (9%) **Eye (n=3)** • Resolved: 1 (33%) • Improved: 0 (0%) • Unchanged: 2 (67%) • Progressed: 0 (0%) In other organs where n=1: • Liver : Unchanged • Kidney: Resolved • Sinus: Unchanged • Bone: Resolved • Muscle: Resolved			
				Primary Clinical effectiveness	Clinical response (patient-reported symptomatic response)	Same response categories as the physician evaluation. **All organs (n=26)** • Resolved: 20% • Improved: 53% • Unchanged: 23% • Progressed: 5% **Lung (n=15)** • Resolved: 0 (0%) • Improved: 9 (60%) • Unchanged: 4 (27%) • Progressed: 2 (13%) **CNS (n=8)** • Resolved: 2 (25%)			

						• Improved: 4 (50%) • Unchanged: 2 (25%) • Progressed: 0 (0%) **Skin (n=11)** • Resolved: 4 (36%) • Improved: 7 (64%) • Unchanged: 0 (0%) • Progressed: 0 (0%) **Eye (n=3)** • Resolved: 1 (33%) • Improved: 0 (0%) • Unchanged: 2 (67%) • Progressed: 0 (0%) In other organs where n=1: • Sinus: Unchanged • Bone: Resolved • Muscle: Improved			
				Primary Clinical effectiveness	Corticosteroid use	11 patients (73%) discontinued corticosteroids after initiation of infliximab.			
				Primary Clinical effectiveness	Pulmonary function parameters	In patients with pulmonary sarcoidosis (n=13) **Forced vital capacity % predicted** Baseline: 81% Follow-up: 85% No significant difference (p=0.06) **Forced expiratory volume in 1s % predicted** Baseline: 81% Follow-up: 83% No significant difference (p=0.72) **Total lung capacity % predicted** Baseline: 75% Follow-up: 86% No significant difference (p=0.07)			

						Diffusing capacity for carbon monoxide % predicted Baseline: 65% Follow-up: 65% No significant difference (p=1.0) In patients with pulmonary sarcoidosis and an abnormal pulmonary function test prior to infliximab (n=7) **Forced vital capacity % predicted** Baseline: 73% Follow-up: 78% No significant difference (p=0.17) **Forced expiratory volume in 1s % predicted** Baseline: 70% Follow-up: 73% No significant difference (p=0.53) **Total lung capacity % predicted** Baseline: 63% Follow-up: 80% Statistically significant improvement (p=0.04) **Diffusing capacity for carbon monoxide % predicted** Baseline: 54% Follow-up: 59% No significant difference (p=0.5)			
				Safety	Adverse events	15 patients (58%) had an adverse event Adverse events requiring discontinuation of infliximab (n=3; 12%): • Severe pneumonia (n=1) • Positive purified protein derivative (PPD) tuberculosis skin test (n=1) • Recurrent sinusitis attributed to infliximab (n=1) Other side effects:			

						• Minor infection (n=4; 19%) • Rash (n=4; 19%) • Pneumonia (n=3; 14%) • Sepsis (n=1; 5%) • Anaphylaxis (n=1; 5%) • Pustular psoriasis (n=1; 5%) • Leukopenia (n=1; 5%) • Psoriatic lesions (n=1; 5%) • Positive PPD (n=1; 5%) • Arthralgias (n=1; 5%) • HPV reactivation (n=1; 5%)			
Hostettler et al 2012	S2 – Retrospective review 1 hospital in Switzerland 2003 to 2010	Patients with chronic progressive, steroid resistant sarcoidosis 10 (36%) had predominantly pulmonary involvement. 18 (64%) had predominantly extrapulmonary sarcoidosis.	N=28 Outcomes reported for 16 patients who received infliximab ≥ 12 months. Patients typically received 3mg/kg infliximab in 4-8 weekly intervals. Patients remained on concurrent corticosteroids and/or other immunosuppressants The mean duration of treatment was 29 months (range 12 to 62)	Primary Clinical effectiveness	Pulmonary function parameters	**Patients with pulmonary involvement who received infliximab ≥ 12 months (n=5)** **Mean improvement in forced vital capacity % predicted (FVC%P):** 6% (range -6-23) Significance tests not reported. **Mean ± SD absolute forced vital capacity (litres)** • Baseline: 2.26 ± 1.25 • Follow-up: 2.57 ± 1.58 Significance tests not reported. • >10% improvement in FVC%P: 1 (20%) • 0-10% improvement in FVC1%: 3 (60%) • Decrease in FVC1%: 1 (20%)	6	Direct	This uncontrolled retrospective review included patients from 1 centre over a 7 year period but the sample of patients is small. Patients had a range of clinical manifestations for sarcoidosis including pulmonary and/or CNS involvement. No definitions were provided for the clinical response categories used. The total population is given as 28 patients, however outcomes are only reported for the 16 patients who received infliximab ≥ 12 months. In the 12 patients who received infliximab <12 months: • 5 stopped after treatment benefit • 2 stooped after treatment failure • 2 stopped to an adverse effect or suspected adverse effect • 2 dropped out • 1 received ongoing infliximab therapy The retrospective design of the study introduces the possibility of selection bias in the study population. This can result from the patients included in the analysis and the classification of patients from patient records. As the study does not include a comparator it is not possible to compare the outcomes for these patients with patients receiving alternative treatments.
				Primary Clinical effectiveness	Clinical response (physician evaluation)	**Patients with extrapulmonary involvement who received infliximab ≥ 12 months (n=11)** • Complete response: 6 (55%) • Partial response: 4 (36%) • No response: 1 (9%) **Patients with CNS involvement who received infliximab ≥ 12 months (n=6)** • Complete response: 3 (50%) • Partial response: 2 (33%) • No response: 1 (17%) No definitions for the response categories reported.			
				Safety	Adverse events	None of the 16 patients receiving infliximab for ≥ 12 months developed severe			

						infections or malignancies. 1 patient experienced a possible adverse effect (symptomatic bradyaahythmia). 2 patients receiving infliximab <12 months stopped due to adverse effect or suspected adverse effect.			

ACE - Angiotensin-Converting Enzyme; CI – Confidence Interval; CNS – central nervous system; DLCOc - Diffusing Capacity of the Lung for Carbon Monoxides, Corrected for Haemoglobin; ePOST – Extrapulmonary Physician Organ Severity Tool; FDG - F-fluorodeoxyglucose; FVC%P – Forced Vital Capacity % Predicted; IQR – Interquartile Range; OR – Odds Ratio; PET - Positron Emission Tomography; S – Second; SD – standard deviation; sIL-2R - Soluble Interleukin-2-Receptor; SUV_{max} - Maximum Standardised Uptake Value; TNF – tumour necrosis factor; US – United States

8 Grade of evidence table

For abbreviations see list after each table

Use of Infliximab Vs. Placebo for Sarcoidosis					
Outcome Measure	**Reference**	**Quality of Evidence Score**	**Applicability**	**Grade of Evidence**	**Interpretation of Evidence**
Pulmonary function parameters	Rossman et al 2006	5	Direct	C	Pulmonary function was assessed using observed and percent expected vital capacity (VC). As no minimal important difference for change in VC has been defined (van Rijswijk et al 2013) % predicted and percentage of patients with a specified increase e.g. 5%, 10% or 15%, are also reported. In the 6-week randomised phase of the RCT there was no significant difference in the mean ± SD relative % change in expected VC between Group 1 (infliximab) (15.2% ± 9.9%) and Group 2 (placebo) (8.4% ± 3.3%), (p=0.65). The observed mean vital capacity was 2.5 litres at baseline and 2.7 litres at week 6 in Group 1 (infliximab) and 2.4 litres at both baseline and week 6 in Group 2 (placebo) (no significance test reported). Three patients (23%) in Group 1 had an improvement of ≥10% predicted VC compared to 2 patients (33%) in Group 2. Two patients (15%) in Group 1 had an improvement of ≥15% predicted VC compared to no patients (0%) in Group 2. No significance tests were reported for these comparisons. In the phase 2 open label follow up from week 12 to week 38 the VC % expected ranged from 65.5 to 67.4 in Group 1 (baseline 59.6) and 70.7 to 72.5 in Group 2 (baseline 65.5). When the results for the first 6 weeks of taking infliximab were combined (week 0 to 6 in Group 1 and week 6 to 12 in Group 2) a significant improvement in mean VC was found (p<0.02) (mean figures not reported). An improvement in mean vital capacity % expected was seen in both groups. 23% of the patients in Group 1 (infliximab) and 33% of patients in Group 2 (placebo) achieved an improvement of ≥10% in the randomised phase of the study. No significant differences between infliximab and placebo were reported. This was a randomised, double-blind, placebo controlled phase II trial. However, the study was closed early due to poor enrolment and the study was therefore underpowered to detect any differences between the groups. The results should be treated with caution.
Radiologic improvement on chest x-ray	Rossman et al 2006	5	Direct	C	Chest x-rays were assessed as ‘markedly worse’, ‘slightly worse’, ‘unchanged’, ‘slightly improved’ or ‘markedly improved’. In the 6-week randomised phase of the RCT 23.0% of Group 1 (infliximab) had radiologic improvement compared to 0% of Group 2 (placebo). No significance test was reported. In the phase 2 open label the percentage of patients showing radiologic improvement appear similar in both groups (Group 1 30.8% vs. Group 2 33.3%). No significance test was reported. More patients in Group 1 showed radiologic improvement in the randomised phase of the study however it is not clear if this difference was statistically or clinically significant. It is not clearly stated if the reported improvement included both patients who had ‘slightly improved’ and ‘markedly improved’.

					This was a randomised, double-blind, placebo controlled phase II trial. However, the study was closed early due to poor enrolment and the study was therefore underpowered to detect any differences between the groups. The results should be treated with caution.
Dyspnea score	Rossman et al 2006	5	Direct	C	The Baseline Dyspnea Index (BDI) and Transitional Dyspnea Index (TDI) are used to assess breathlessness and the impact of intervention. The BDI and TDI have three domains: functional impairment, which determines the impact of breathless on the ability to carry out activities; magnitude of task, which determines the type of task that causes breathlessness; and magnitude of effort, which establishes the level of effort that results in breathlessness. The BDI is scored from 0 (very severe impairment) to 4 (no impairment) for each domain and summed to create a focal score (0-12). The TDI is scored from -3 (major deterioration) to +3 (major improvement) for each domain and summed to create a focal score (-9 to +9). An improvement of ≥1.0 has been suggested to be clinically important (Witek & Mahler 2003). The baselines scores assessed on the BDI were 2.17 for Group 1 and 2.08 for Group 2. Improvements on the TDI in the 6-week randomised phase were reported for both groups in mean functional impairment (Group 1 0.38 vs. Group 2 0.17) and mean magnitude of effort (Group 1 0.23 vs. Group 2 0.17). No significance tests were reported. No change in magnitude of task was reported (figures not reported). The meaning of the dyspnea scores are not clearly reported, however the results are presented by the authors as an improvement for Group 1. The scores of between 0 and 1 on the TDI would suggest a small, but not clinically important, improvement. This was a randomised, double-blind, placebo controlled phase II trial. However, the study was closed early due to poor enrolment and the study was therefore underpowered to detect any differences between the groups and the results should be treated with caution.
Quality of life	Rossman et al 2006	5	Direct	C	Quality of life was assessed using the SF-36 which is scored from 0-100 with higher scores indicating better functioning. An improvement of 10 points was considered clinically relevant in one study (Vorselaars et al 2015). SF-36 scores were similar in both groups at baseline (Group 1 26.7 vs. Group 2 26.4) and at the end of the 6-week randomised phase (Group 1 27.1 vs. Group 2 26.4). No significance tests were reported. No significant improvement in quality of life was reported for either group. This was a randomised, double-blind, placebo controlled phase II trial. However, the study was closed early due to poor enrolment and the study was therefore underpowered to detect any differences between the groups and the results should be treated with caution.
Safety	Rossman et al 2006	5	Direct	C	The percentage of patients reporting adverse events and serious adverse events were reported. Four patients reported serious events (31%) in Group 1 (infliximab) and 1 (17%) in Group 2 (placebo). Adverse events were reported by 92% of Group 1 and 100% of Group 2. Infections were reported by 69% of Group 1 and 50% of Group 2. No significance tests were reported.

					Most of the participants in the study reported adverse events, and serious adverse events were reported in approximately a third of the patients receiving infliximab. This was a randomised, double-blind, placebo controlled phase II trial. However, the study was closed early due to poor enrolment and the study was therefore underpowered to detect any differences between the groups and the results should be treated with caution.

RCT – Randomised Controlled Trial; SD – Standard Deviation; VC – Vital Capacity

Use of Infliximab for Sarcoidosis (No Comparator)					
Outcome Measure	**Reference**	**Quality of Evidence Score**	**Applicability**	**Grade of Evidence**	**Interpretation of Evidence**
Pulmonary function parameters	Vorselaars et al 2015	7	Direct	A	Pulmonary function was assessed using forced vital capacity (FVC), forced expiratory volume in 1 second (FEV_1) and diffusing capacity of the lung for carbon monoxides, corrected for haemoglobin (DLCOc). As no minimal important difference for change in FVC has been defined (Vorselaars et al 2015) studies also report % predicted and percentage of patients with a specified increase e.g. 5%, 10% or 15%. An improvement in forced expiratory volume of ≥10% was considered to be a clinically relevant change for an individual patient (van Rijswijk et al 2013). Function was also assessed by the 6-minute walking test which assesses how far someone can walk in 6 minutes. In patients with a pulmonary treatment indication there was a significant improvement from baseline for FVC % predicted (+6.6 from a baseline of 73.6), FEV_1 % predicted (+5.8 from a baseline of 55.8) and DLCOc (+4.1 from a baseline of 56.6) at 26 week follow-up. 71% of patients had an improvement of ≥5% FVC predicted and 46% of patients had an improvement of ≥10% FVC predicted. For FEV_1 % predicted, 64% had an improvement of ≥5% and 46% an improvement of ≥10%. The 6-minute walking distance % predicted improved by +4.2 from a baseline of 61. No unit of measurement or significance test was reported for this measure. In patients with an extrapulmonary treatment indication there was a significant improvement from baseline for FVC % predicted (+3.9) and FEV_1 % predicted (+3.5) (baseline figures not reported for these patients). 37% of patients had an improvement of ≥5% FVC predicted and FEV_1 % predicted. A statistically significant improvement was seen for all measures where significance testing was reported. An improvement of ≥10% was considered clinically relevant. This was achieved by 46% patients with a pulmonary treatment indication. This uncontrolled prospective study had a relatively large sample size (n=56). Patients had a range of clinical manifestations for sarcoidosis including pulmonary and/or CNS involvement. The lack of comparator in these studies limits the strength of the conclusions that can be drawn.
	van Rijswijk et al 2013	7	Direct		
	Hostettler et al 2012	6	Direct		
	Russell et al 2013	6	Direct		

Clinical response (physician evaluation)	Jamilloux et al 2017	7	Direct	A	Clinical response was assessed by physicians in five uncontrolled retrospective studies. The response definitions used varied for each study. In the largest of these studies (Jamilloux et al 2017) the following categories were used: • Complete response (disappearance of clinical signs (excluding sequelae) upon use of corticosteroids <10mg) • Partial response (an improvement of clinical and para-clinical parameters upon >50% reduction of the initial corticosteroids dose) • Stable (non-responders with no change) • Progressive disease (non-responders with either new organ involvement, worsening of an organ involvement or need for increased corticosteroid dosage). After a median follow-up of 20.5 months, a complete or partial response was reported in 18% and 46% of patients respectively. Non-responders included 25% of patients with no change and 11% of patients with progressive disease. In multivariate analysis comparing responders with non-responders, pulmonary involvement was associated with a lower clinical response (OR =0.38, 95%CI 0.14 to 0.92). A complete or partial response was seen in 64% patients, with most of these showing a partial response. The clinical meaningfulness of a partial response is not clear although the definition includes improvement with a reduction in corticosteroid use. This uncontrolled retrospective review included patients from 25 centres and had a relatively large sample size (n= 132). Patients had a range of clinical manifestations for sarcoidosis including pulmonary and/or CNS involvement and a small number of patients included in the sample (5%) did not receive infliximab as the anti-TNF treatment. The retrospective design and lack of comparator limit the strength of the conclusions that can be drawn.
	Hostettler et al 2012	6	Direct		
	Russell et al 2013	6	Direct		
	Cohen Aubart et al 2017	7	Direct		
	Chapelon-Abric et al 2015	6	Direct		
Corticosteroid use	Vorselaars et al 2015	7	Direct	A	The dose of corticosteroid at baseline and follow-up is compared to assess the corticosteroid sparing effect of treatment. In the prospective study (Vorselaars et al 2015, n= 56) with 26 weeks follow-up there was a statistically significant reduction in mean corticosteroid dose by 8.8mg for 19 patients who received concomitant corticosteroids (p=0.001). The baseline and follow-up dose per day was not reported. In the largest retrospective study (Jamilloux et al 2017, n=132) with a median of 20.5 months follow-up, there was a statistically significant reduction in the mean dose of corticosteroid from 23 ± 20 mg/day at baseline to 11 ± 11 mg/day in 113 patients who received concomitant corticosteroids. A reduction in corticosteroid use is a positive outcome. In these studies the mean reduction in corticosteroids was approximately 9-12 mg/ day. The clinical meaningfulness of a reduction of this magnitude in this population is not clear. The evidence comes from uncontrolled studies with relatively large sample sizes. Patients had a range of clinical manifestations for sarcoidosis including pulmonary and/or CNS involvement. The lack of comparator in these studies limits the strength of the conclusions that can be drawn.
	Jamilloux et al 2017	7	Direct		
	Russell et al 2013	6	Direct		
	Cohen Aubart et al 2017	7	Direct		
	Chapelon-Abric et al 2015	6	Direct		

Organ assessment	Jamilloux et al 2017	7	Direct	B	The Extrapulmonary Physician Organ Severity Tool (ePOST) examines 17 extrapulmonary organs and assigns each a score from 0 (not affected) to 6 (very severely affected) to each organ. Intermediate scores are slight (1), mild (2), moderate (3), moderate to severe (4) and severe (5). ePOST scores were provided for a range of organs. Pulmonary and neurological outcomes are of particular interest in this review. Statistically significant improvements between baseline and a median follow-up of 20.5 months were reported for the upper respiratory tract (2.3 to 1.8), central nervous system (CNS) (3.8 to 2.6) and peripheral nervous system (PNS) (1.1 to 0.24). There was no significant improvement for lungs between baseline (2.1) and follow-up (1.9). Statistically significant improvements were reported for the upper respiratory tract, CNS and PNS. The difference in scores between baseline and follow-up was greater for the nervous system scores. For the CNS this equates to an improvement from 'moderate to severe' to 'mild to moderate'. For the PNS this equates to an improvement from 'slight' to 'not affected to slight'. This uncontrolled retrospective review included patients from 25 centres and had a relatively large sample size (n= 132). Patients had a range of clinical manifestations for sarcoidosis including pulmonary and/or CNS involvement and a small number of patients included in the sample (5%) did not receive infliximab as the anti-TNF treatment. The retrospective design and lack of comparator limit the strength of the conclusions that can be drawn.
	Chapelon-Abric et al 2015	6	Direct		
Inflammatory response	Vorselaars et al 2015	7	Direct	A	Measures of inflammatory response included the biomarkers soluble interleukin-2-receptor (sIL-2R), angiotensin-converting enzyme (ACE) and F-fluorodeoxyglucose (FDG) by positron emission tomography (PET) maximum standardised uptake value (SUV_{max}). A statistically significant improvement from baseline to 26 week follow-up was reported for each of the inflammatory response measures in Vorselaars et al (2015). For F-FDG PET SUV_{max} the mean improvement was -4.0 from a baseline of 6.6 for lung parenchyma, -3.0 from a baseline of 5.7 for mediastinum and -5.8 from a baseline of 9.0 for index localisation. For ACE the improvement after infliximab was -28.2 from a baseline of 89.7 and for sIL-2R the improvement was -4269 from a baseline of 8,824. No information about the clinical relevance of the results reported was provided by the study authors. This uncontrolled prospective study had a relatively large sample size (n=56). Patients had a range of clinical manifestations for sarcoidosis including pulmonary and/or CNS involvement. The lack of comparator in these studies limits the strength of the conclusions that can be drawn.
	van Rijswijk et al 2013	7	Direct		
Quality of life	Vorselaars et al 2015	7	Direct	A	Quality of life was assessed using a Patient Global Assessment (PGA) with scores on a visual analogue scale ranging from 0 (best imaginable health status) to 100 (worst imaginable health status); and the SF-36 to assess physical functioning. The SF-36 is scored from 0-100 with higher scores indicating better functioning. An improvement of 10 points was considered clinically relevant (Vorselaars et al 2015). A statistically significant improvement was seen in both quality of life measures used from baseline to follow up at 26 weeks. For the PGA the mean score improved by -14.6 from a baseline of 61.0. For the SF-36 the score improved by 8.2 from a baseline of
	van Rijswijk et al 2013	7	Direct		

					40.6. An improvement of 10 points was considered clinically relevant. An improvement of >10 points was reported for the PGA but not for the SF-36. This uncontrolled prospective study had a relatively large sample size (n=56). Patients had a range of clinical manifestations for sarcoidosis including pulmonary and/or CNS involvement. The lack of comparator in these studies limits the strength of the conclusions that can be drawn.
Clinical response (patient reported)	Russell et al 2013	6	Direct	C	For patient-reported symptomatic response Russell et al (2013) used the following categories: • 'Resolved' = complete resolution of clinical disease activity • 'Improved' = organs with reduced sarcoid burden or reduced frequency in disease activity but still with evidence of disease • 'Unchanged' = disease activity clinically no different than prior to infliximab • 'Progressed' = clinical features of progressive disease despite infliximab Separate scores were provided for all organs and individual organs. Pulmonary and neurological outcomes are of particular interest in this review. When all organs were included, an improvement was reported by 73% of patients, consisting of 20% who were 'resolved' and 53% who were 'improved'. Of those that had not seen an improvement, 23% were 'unchanged' and 5% had 'progressed'. For lungs, an improvement was reported by 60% of patients, all of whom were 'improved'. For the central nervous system, an improvement was reported by 75% patients consisting of 25% 'resolved' and 50% 'improved'. The majority of patients reported an improvement in their symptoms with 20% reporting a complete resolution of disease activity. This uncontrolled retrospective review included patients from a single centre and included a small number of patients (n=26). The retrospective design and lack of comparator limit the strength of the conclusions that can be drawn.
Composite overall response	Vorselaars et al 2015	7	Direct	A	The composite overall response included organ function, inflammation and quality of life (Vorselaars et al 2015). This was an author-designed non-validated tool. Improvement in a category was scored only when one of the parameters improved significantly without deterioration of the others. A good or excellent response was a clinically relevant improvement in 2 or 3 categories, a partial response was a clinically relevant improvement in one category and no improvement in any category was a nonresponse. After 26 weeks follow-up a response was reported in 96% of patients. This included 40% showing an excellent response, 39% a good response and 17% a partial response. The definitions used for this outcome measure include the clinical relevance of the improvement observed. 40% of patients showed an 'excellent' response which equates to a clinically relevant improvement in all three categories. The composite tool used has not been validated and the study authors advised that it should be interpreted with care. This uncontrolled prospective study had a relatively large sample size (n=56) Patients had a range of clinical manifestations for sarcoidosis including pulmonary and/or CNS involvement. The lack of comparator in these studies limits the strength of the conclusions that can be drawn.

Modified Rankin score	Cohen Aubart et al 2017	7	Direct	A	The modified Rankin score measures the degree of disability or dependence in the daily activities of people who have suffered a stroke or other causes of neurological disability. It is scored from 0 (no symptoms at all) to 6 (dead). A score of 3 is defined as moderate disability; requiring some help, but able to walk without assistance. A score of 1 is defined as no significant disability despite symptoms; able to carry out all usual duties and activities[14]. A statistically significant improvement in modified Rankin score was reported from 3 at baseline to 1 at a median follow-up of 20 months. The modified Rankin scores reported translate to an improvement from moderate disability to no significant disability despite symptoms. It is likely that this improvement would be meaningful to the patient. This uncontrolled retrospective review included patients from three centres but included a small number of patients (n=18). The retrospective design and lack of comparator limit the strength of the conclusions that can be drawn.
Infliximab trough levels	Vorselaars et al 2015	7	Direct	A	The trough level is the lowest concentration reached by a drug before the next dose is administered. It can be used to evaluate appropriate dosage levels. The mean trough level was 18.0 µg/mL^{-1}. There was no significant correlation between trough level and response. This uncontrolled prospective study had a relatively large sample size. Patients had a range of clinical manifestations for sarcoidosis including pulmonary and/or CNS involvement. The lack of comparator in these studies limits the strength of the conclusions that can be drawn.
Safety	Vorselaars et al 2015	7	Direct	A	The percentage of patients reporting side effects and severe side effects were reported in the prospective study by Vorselaars et al (2015). Severe side effects were reported in 7 patients (13%). These included 3 patients with pneumonia requiring hospitalisation, 2 patients with allergic reaction with antibody formation,1 patient with peritonitis and1 patient with severe gastrointestinal complaints. Other side effects were reported for 13 patients (23%) including mild infection of the respiratory tract (n=5), oedema (n=3), headache (n=2), joint pain (n=2) and dizziness (n=1). 61% of the patients did not experience any side effects from infliximab. 13% of the study population experienced severe side effects. This uncontrolled prospective study had a relatively large sample size (n=56). Patients had a range of clinical manifestations for sarcoidosis including pulmonary and/or CNS involvement. The lack of comparator in these studies limits the strength of the conclusions that can be drawn.
	Jamilloux et al 2017	7	Direct		
	van Rijswijk et al 2013	7	Direct		
	Hostettler et al 2012	6	Direct		
	Russell et al 2013	6	Direct		
	Cohen Aubart et al 2017	7	Direct		
	Chapelon-Abric et al 2015	6	Direct		

ACE - Angiotensin-Converting Enzyme; CI - Confidence Interval; CNS – Central Nervous System; DLCOc - Diffusing Capacity of the Lung for Carbon Monoxides, Corrected for Haemoglobin; ePOST - Extrapulmonary Physician Organ Severity Tool; FDG - F-fluorodeoxyglucose; FEV_1 - Forced Expiratory Volume in 1 Second ; FVC - Forced Vital Capacity; OR – Odds Ratio; PET - Positron Emission Tomography; PGA - Patient Global Assessment; PNS - peripheral nervous system ; sIL-2R - Soluble Interleukin-2-Receptor; SUV_{max} - Maximum Standardised Uptake Value.

[14] https://www.mdcalc.com/modified-rankin-scale-neurologic-disability

9 Literature Search Terms

Search strategy 'infliximab', 'remicade', 'flixabi', 'inflectra', 'remsima', 'sarcoid' and their combinations	
P – Patients / Population Which patients or populations of patients are we interested in? How can they be best described? Are there subgroups that need to be considered?	Adults (18 years and above) with **refractory to oral corticosteroids or progressive pulmonary** and/or CNS sarcoidosis
I – Intervention Which intervention, treatment or approach should be used?	Use of infliximab (Remicade, Flixabi, Inflectra and Remsima) alone or as an adjuvant to current standard pharmaceutical treatments
C – Comparison What is/are the main alternative/s to compare with the intervention being considered?	Placebo, corticosteroid, methotrexate, azathioprine, lung transplant
O – Outcomes What is really important for the patient? Which outcomes should be considered? Examples include intermediate or short-term outcomes; mortality; morbidity and quality of life; treatment complications; adverse effects; rates of relapse; late morbidity and re-admission; return to work, physical and social functioning, resource use.	All outcome measures reported in studied should be included *Critical to decision-making:* 1. Adverse events 2. Disease progression As sarcoidosis affects a number of organs, outcome measures will be heterogenous but we are only interested **in pulmonary** and **neurological** endpoints therefore, particular need for data on: • changes on lung or CNS imaging; • for pulmonary function: mean % predicted change in FVC, FEV1 and/or Difussing capacity of Lungs for Carbon Monoxide (DLCO) or Transfer capacity of Lungs for Carbon Monoxide (TLCO), dyspnoea and 6 minute walking distance; • reduction in corticosteroid usage • improvement in secondary pulmonary hypertension • improvement in neurological function • improvements in cerebrospinal fluid abnormalities 3. Cost effectiveness 4. Quality of life 5. Survival
Assumptions / limits applied to search • Publication type Selection of studies should be restricted to include peer-reviewed articles published in journals. Conference abstracts, grey literature, anecdotal and unpublished evidence should be excluded from the assessment. • Language Only articles published in the English language should be included. • Time frame Only articles published in the last 15 years should be included. • Study design Only controlled studies, uncontrolled studies or case series of five or more patients should be included.	

10 Search Strategy

We searched PubMed, Embase and Cochrane Library limiting the search to papers published in England from **1St January 2002 to 21st July 2017**. We excluded conference abstracts, commentaries, letters, editorials and case reports.

Search date: 21st July 2017
Embase search:

#	**Narrow search**
1	lung sarcoidosis/
2	exp sarcoidosis/ and (lung function/ or lung/ or lung function test/)
3	((lung or pulmonary) and (sarcoid* or neurosarcoid*)).ti.
4	((lung or pulmonary) adj5 (sarcoid* or neurosarcoid*)).ti,ab.
5	((sarcoid* or neurosarcoid*) and (lung function* or pulmonary function*)).ti,ab.
6	1 or 2 or 3 or 4 or 5
7	infliximab/
8	(infliximab or flixabi or remicade or remisma or revellex or avakine).ti,ab.
9	7 or 8
10	6 and 9
11	limit 10 to (english language and yr="2002 -Current")
12	conference*.pt.
13	11 not 12

#	**Broader search**
1	exp *Sarcoidosis/
2	(sarcoid* or neurosarcoid*).ti,ab.
3	1 or 2
4	*infliximab/
5	(infliximab or flixabi or remicade or remisma or revellex or avakine).ti,ab.
6	4 or 5
7	3 and 6
8	limit 7 to (english language and "therapy (maximizes sensitivity)" and yr="2002 - Current")
9	conference*.pt.
10	8 not 9

11 Evidence Selection

- Total number of publications reviewed: 42
- Total number of publications considered potentially relevant: 27
- Total number of publications selected for inclusion in this briefing: 8

12 References

Atkins C. Wilson AM. (2017) Managing fatigue in sarcoidosis – a systematic review of the evidence. *Chron. Respir. Dis.* 14(2): 161-173.

Baughman RP. Drent M. Kavuru M. Judson MA. Costabel U. du Bois R. Albera C. Brutsche M. Davis G. Donohue JF. Muller-Quernheim J. Schlenker-Herceg R. Flavin S. Lo KH. Oemar B. Barnathan ES on behalf of the Sarcoidosis Investigators. (2006). Infliximab therapy in patients with chronic sarcoidosis and pulmonary involvement. *Am J Resp Crti Care Med* 174: 795-802.

Cohen Aubert F. Bouvry D. Galanaud D. Dehais C. Mathey G. Psimaras D. Haroche J. Pottier C. Hie M. Mathian A. Devilliers H. Nunes H. Valeyre D. Amoura Z. 2017. Long-term outcomes of refractory neurosarcoidosis treated with infliximab. *Journal of Neurology* 264: 891-897.

Chapelon-Abric C. Saadoun D. Biard L. Sene D. Resche-Rigon M. Hervier B. Costedoat-Chalumeau N. Drier A. Leger JM. Cacoub P. 2015. Long-term outcome of infliximab in severe chronic and refractory systemic sarcoidosis: a report of 16 cases. *Clinical and Experimental Rheumatology* 33: 509-515.

Drent M. Cremers JP. Jansen TL. Baughman RP. 2014. Practical eminence and experience-based recommendations for use of TNF-α inhibitors in sarcoidosis. *Sarcoidosis Vasculitis and Diffuse Lung Diseases* 31: 91-107.

Gibson GJ. Prescott RJ. Muers MF. Middleton WG. Mitchell DN. Connolly CK. Harrison BD. (1996) British Thoracic Society Sarcoidosis study: effects of long term corticosteroid treatment. *Thorax.* 51(3): 238-47.

Hostettler KE. Studler U. Tamm M. Brutsche MH. (2012). Long-term treatment with infliximab in patients with sarcoidosis. *Respiration* 83: 218-224.

Jamilloux Y. Cohen-Aubert F. Chapelon-Abric C. Maucort-Boulch D. Marquet A. Pérard L. Bouillet L. Deroux A. Abad S. Bielefeld P. Bouvry D. André M. Noel N. Bienvenu B. Proux A. Vukusic S. Bodaghi B. Sarrot-Reynauld F. Iwaz J. Amoura Z. Broussolle C. Cacoub P. Saadoun D. Valeyre D. Sève P. the Groupe Sarcoïdose Francophone. 2017. Efficacy and safety of tumor necrosis factor antagonists in refractory sarcoidosis: a multicenter study of 132 patients. *Seminars in Arthritis and Rheumatism* Mar 8 Epub ahead of print.

Maneiro JR. Salgado E. Gomez-Reino JJ. Carmona L. (2012) Efficacy and safety of TNF antagonists in sarcoidosis. Data from the Spanish Registary of Biologics BIOBADASER and a systematic review. *Seminars in Arthritis and Rheumatism* 42(1): 89-103.

National Institute for Health and Care Excellence. 2017. Refractory extrapulmonary sarcoidosis: infliximab. January 2017.

National Institute for Health and Care Excellence. 2016. Pulmonary sarcoidosis: infliximab. December 2016.

NHS England. Population, Intervention, Comparator and Outcomes (PICO): Infliximab for refractory, progressive, pulmonary sarcoidosis and neurosarcoidosis. URN 1673. July 2017.

Rossman MD. Newman LS. Baughman RP. Teirstein A. Weinberger SE. Miller Jr W. Sands BE. 2006. A double-blinded, randomized, placebo-controlled trial of infliximab in subjects with active pulmonary sarcoidosis. *Sarcoidosis Vasculitis and Diffuse Lung Diseases* 23:201-208.

Russell E. Luk F. Manocha S. Ho T. O'Connor C. Hussain H. 2013. Long term follow-up of infliximab efficacy in pulmonary and extra-pulmonary sarcoidosis refractory to conventional therapy. *Seminars in Arthritis and Rheumatism* 43: 119-124.

van Rijswijk HNAJ. Vorselaars ADM. Ruven HJT. Keijsers RGM. Zanen P. Korenromp IHE. Grutters JC. 2013. Changes in disease activity, lung function and quality of life in patients with refractory sarcoidosis after anti-TNF treatment. *Expert Opinion on Orphan Drugs* 1(6): 437-443.

Vorselaars ADM. Crommelin HA. Deneer VHM. Meek B. Claessen AME. Keijsers RGM. Van Moorsel CHM. Grutters JC. 2015. Effectiveness of infliximab in refractory FDG PET-positive sarcoidosis. *Eur Respir J* 46: 175-185.

Witek TJ. Mahler DA. (2003) Minimal important difference of the transition dyspnoea index in a multinational clinical trial. *European Respiratory Journal* 21: 267-272.