Systematic review and network meta-analysis of approved medicines for the treatment of idiopathic pulmonary fibrosis

Aristeidis Skandamis¹, Chara Kani¹, Sophia L. Markantonis¹ and Kyriakos Souliotis¹

¹Pharmacy Department, National and Kapodistrian University of Athens, Athens, Greece; ²Faculty of Social and Political Sciences, University of Peloponese, Corinth, Greece

ABSTRACT
Background: Clinical practice guidelines for the treatment of idiopathic pulmonary fibrosis (IPF) currently recommend pirfenidone and nintedanib. However, there is a lack of evidence from head-to-head comparisons.

Objectives: To perform a systematic review and network meta-analysis (NMA) to access the efficacy and tolerability of two new treatments for IPF, pirfenidone and nintedanib.

Methods: Randomized controlled trials (RCTs) selection (CENTRAL, MEDLINE, Embase), data extraction, risk of bias analysis, and GRADE assessment were carried out by two authors separately. Direct estimates were calculated using standard pairwise meta-analysis. A Bayesian mixed treatment comparison approach for NMA estimates, with 95% confidence intervals (CI), was used to compare the treatments, calculating odds ratios (OR) and number needed to treat (NNTB) or harm (NNTH).

Results: The NMA on 10 randomized controlled trials showed that each drug had a positive effect on percentage of forced vital capacity (FVC) decline ≥10% [pirfenidone OR = 0.54 [95% CI = 0.37–0.80], NNTB = 9 [95% CI = 7–22]; nintedanib OR = 0.59 [95% CI = 0.41–0.84], NNTB = 9 [95% CI = 6–23]), but no significant differences were noted when comparing pirfenidone and nintedanib with respect to acute exacerbations, mortality, and serious adverse events (FVC decline OR = 0.91 [95% CI = 0.45–2.03]) or dropouts (OR = 0.75 [95% CI = 0.33–1.27]). Nintedanib showed an effect on dropouts, OR = 1.61 (1.13–2.28) and NNTH = 14 (8–61).

Conclusions: Based on RCTs of 12-month duration in patients with IPF, a positive effect on FVC decline was noted for both treatments and on dropouts for nintedanib, but no significant differences were noted between treatments.

Introduction

Idiopathic pulmonary fibrosis is the most common type of idiopathic interstitial pneumonia, and its incidence and prevalence indicate that it is a disease which heavily impacts healthcare systems globally. It is a chronic progressive disease that primarily occurs in middle-aged and older adults and accounts for 20–30% of interstitial lung diseases¹.

In Europe, IPF prevalence has been reported to range from 1.25–23.4 cases per 100,000 population, and the annual incidence is between 0.22 and 7.4 per 100,000 population², but both increase with age, are higher among males, and appear to be on the increase in recent years. A recent epidemiological study in northern Italy estimated the mean annual incidence rate to be 2.3 using narrow case definitions, while the estimated annual prevalence rate was 12.6 per 100,000 person-years³.

Although its cause remains unknown, multiple genetic and environmental factors may play a role in the pathogenesis and progression of pulmonary fibrosis, including infectious agents (bacterial and viral pathogens), particulate inhalation (cigarette smoking, dusty environments), and several occupational factors, such as farming, livestock, hair-dressing, metal dust, raising birds, stone cutting/polishing, and vegetable dust/animal dust⁴,⁵.

Patients typically >45 years old present with non-specific symptoms of exertion dyspnea with or without dry cough, bilateral inspiratory crackles, and sometimes finger clubbing. Occasionally, patients will present acute symptoms, with days to weeks of respiratory worsening, often accompanied by fever and influenza-like symptoms⁶.

The 2013 clinical guidelines produced by the National Institute for Health and Clinical Excellence (NICE)⁶ recognize the fact that the initial assessment of individuals with IPF needs to be improved, to reduce the risk of delays in diagnosis and initiation of treatment, including monitoring and best supportive care (BSC). The assessment must include clinical
features, imaging (High-resolution CT, histopathology, and spirometry), and should be multidisciplinary at each stage of the diagnostic care pathway to include a consultant respiratory physician, consultant radiologist, and ILD specialist nurse. This multidisciplinary approach is considered the diagnostic gold standard, although a few patients remain unclassifiable. The clinical course of IPF may be variable and unpredictable and may be interspersed with acute exacerbations. The median survival time from diagnosis is 2–4 years. Three “pillars” of care are described for IPF: disease-centered management, symptom-centered management, and education and self-management. Best supportive, or palliative care is a proactive approach to symptomatic treatment and in IPF may include oxygen therapy, pulmonary rehabilitation, opiates, antireflux therapy, the withdrawal of steroids, and other immunosuppressants, early recognition of terminal decline, and liaison with palliative care specialists. Also, the NICE 2013 guideline suggests that ambulatory oxygen should be considered if a patient with IPF has breathlessness on exertion, and benzodiazepines and/or opioid therapy if a patient has breathlessness at rest.

Until recently, no intervention, other than lung transplantation, has demonstrated an enhanced survival in patients with IPF, but a few novel agents have demonstrated a decreased rate of disease progression as measured by forced vital capacity (%FVC). Two of these novel agents, Pirfenidone, an orally bioavailable synthetic molecule with antifibrotic and anti-inflammatory properties, and Nintedanib, an intracellular inhibitor of tyrosine kinases that received orphan status from the European Medicines Agency in 2013, have both marketing authorization for use in IPF in Europe. In the absence of head to head trials comparing these treatment regimens, systematic reviews, meta-analyses (MA), and network meta-analyses (NMA) are the suggested tools for their further exploration for the first line treatment of IPF. Recently, several NMAs have investigated the newer treatments for IPF. However, the aim of the present study was to perform a systematic review and network meta-analysis (NMA), using %FVC > 10% decline, acute exacerbations, dropouts, mortality and serious adverse events as endpoints, in order to access the efficacy and safety of pirfenidone and nintedanib to treat patients with IPF and, thus, generate a clinically useful ranking of these treatments.

**Methods**

Criteria for considering studies for inclusion

Parallel group Randomized Controlled Trials (RCTs) of pirfenidone or nintedanib for patients with IPF were included. Eligible studies were Phase II and III randomized controlled trials (RCTs), published in English. Quasi-randomized and cross-over trials were excluded. IPF was defined by the 2015 criteria. Studies that included patients with other confounding respiratory conditions and idiopathic interstitial idiopathic pneumonia other than IPF were excluded.

Search methods for the identification of clinical trials

A systematic review based on a protocol developed a priori up to September 2017 was performed. The following databases (MEDLINE, EMBASE, the Cochrane Library, and PubMed), conference proceedings, and previous systematic reviews were searched up to May 2017.

Types of participants

Individuals aged ≥ 18 years with suspected or diagnosed IPF.

Types of interventions

Pirfenidone or nintedanib compared with one of the other interventions or placebo. Study arms investigating the approved doses of pirfenidone and nintedanib (pirfenidone 1,800 or 2,403 mg/day and nintedanib 300 mg/day) were included.

Types of outcomes

Primary outcomes included the number of participants with an FVC decrease > 10%, acute exacerbations, dropouts, serious adverse events, and all-cause death. A > 10% decrease in FVC was considered clinically meaningful, even though a decrease of as little as 2–6% for %FVC has been estimated to be the minimal clinical requirement. This converts the continuous %FVC to a binary event. The relative ranking of the competing interventions was determined according to the above-mentioned primary outcomes. Other outcomes included the number of Serious Adverse Events.

Data collection and analysis

A brief RCT search strategy (BRSS), consisting of a search of CENTRAL, and then for variants of the word random across all fields (random$.af.) in MEDLINE and EMBASE, was devised and run. The search strategy was designed to search for IPF and RCTs, pirfenidone and nintedanib (see Supplementary material).

Data were extracted from the studies to facilitate the assessment of their suitability for comparison of trials within the NMA. Data were collected on key study characteristics, including study methods, populations, trial settings, treatments, and outcomes. One of the authors extracted key study characteristics using a standardized data extraction form, while another checked each extraction. Any discrepancies were resolved by discussion or with the input of a third author.

Quality assessment of included trials

Two authors independently evaluated the methodological quality of the included studies, as well as overall quality of evidence for each of the five outcomes using the Grading of Recommendations, Assessment, Development and Evaluations (GRADE), as summarized below.
Study quality
The quality of the studies was assessed using the Cochrane risk of bias tool, and two authors independently assessed the following risk of bias criteria for each included trial: random sequence generation, allocation concealment, presence of blinding (participants, personnel, and outcome assessors) in the studies, incomplete outcome data, and selective outcome reporting22.

The risk of bias was assessed as recommended: low risk, high risk, or unclear risk (i.e. either lack of information or uncertainty over the potential for bias). Disagreements were resolved by discussion between the authors22.

Overall quality of evidence
The GRADE approach specific to NMA (for each of the five outcomes) served to assess the quality of the evidence associated with specific comparisons, including direct, indirect, and final network meta-analysis estimates23,24.

Concordance assessment was not needed in this analysis as all estimates included only direct (interventions vs placebo) or only indirect evidence (for all other comparisons).

Data synthesis
Meta-analysis and network meta-analysis methods
Instances of heterogeneity in the study characteristics or outcome measures were highlighted, and a clinical perspective was obtained on whether the differences between studies warranted their exclusion from the analysis. In addition to this qualitative assessment of heterogeneity in the study characteristics and outcome measures, a quantitative assessment of heterogeneity was also undertaken to evaluate the variability in results from study to study. Heterogeneity in treatment effects was evaluated by estimating the variance between studies, and through Cochrane Q-test and I² when at least two studies were available for each pairwise comparison. I²-squared values of 25%, 50%, and 75% were defined as low, moderate, and high heterogeneity, respectively.

Random effects pairwise meta-analyses were conducted for every treatment comparison, with at least two studies using ADDIS27. The network meta-analysis model was implemented in the Bayesian framework and estimated using Markov Chain Monte Carlo (MCMC) methods using ADDIS27. This approach is recommended by the National Institute for Health and Clinical Excellence (NICE) Decision support unit.

Technical support documents on evidence synthesis28.

Synthesis of the results
The results of the assessments of heterogeneity and the principal NMAs are summarized in Table 1. Regarding FVC decline ≥10% both treatments were superior to placebo, with the results characterized by high, moderate heterogeneity, and statistical significance (pirf. vs placebo OR = 0.54 (95% CI = 0.37–0.80), I² = 53% and nint. vs placebo OR = 0.59 (95% CI = 0.41–0.84), I² = 48%). A notable difference was not found through the indirect comparison (pirf. vs nint. OR = 0.91 (95% CI = 0.45–2.03)). Acute exacerbations were another outcome where the therapies were effective (pirf. vs placebo OR = 0.77 (0.26–2.28) and nint. vs placebo OR = 0.61 (0.20–1.90)). Network results showed that pirfenidone was superior, but the difference was not statistically significant (pirf. vs nint. OR = 0.39 (0.00–15.53)). Dropouts were more prevalent in the treatments (pirf. vs placebo OR = 1.27 (0.96–1.68) and nint. vs placebo OR = 1.61 (1.13–2.28)). Pirfenidone exhibited less frequent discontinuations (pirf.
outcome assessment. Certain types of bias could not be adequately determined in the published studies. Those included the Ogura et al.39 and Huang et al.40 studies, where we could only detect the lack of bias in random sequence generation. Incomplete outcome reporting was noted for the Azuma et al.35 study. The other studies all lacked any presence of bias, while the CAPACITY trials14 had an undetermined outcome assessment bias.

When a significant Odds Ratio was calculated for the primary outcomes of decline in percent predicted FVC ≥ 10% and dropouts, we calculated an NNT (benefit or harm) for the different levels of risk as represented by control group event rates over a specified time period using the pooled Odds Ratio and its confidence interval (Visual Rx). The analysis is presented in Table 2. Regarding FVC decline, nine patients would have to be treated with either of the two therapies, for a beneficial effect to be present (pirf. vs placebo NNTB = 9 (7–23) and nint. vs placebo NNTB = 9 (6–22)). NNTH for discontinuations showed that 12 patients needed to be treated with pirfenidone for the harmful effect to show up, in comparison to 14 patients treated with nintedanib.

Regarding the other outcomes nintedanib had better values (smaller NNTBs and higher NNTHs), with the exception of acute exacerbations.

**Discussion**

The present network meta-analysis highlights potentially important differences between the two newest treatments for IPF, pirfenidone and nintedanib, compared to placebo with respect to the decline in percent predicted FVC ≥ 10%, acute exacerbations and mortality. No significant differences were noted when comparing these two treatments to each other apart from dropouts. The number needed to treat was estimated to be the same (NNTB = 9) for both treatments with regards to FVC decline.

The findings of the direct comparison showed that both drugs were superior to placebo. Other reviews have had similar results15–19.

The results of the Network meta-analysis showed less acute exacerbations with pirfenidone compared with nintedanib, but the difference was not statistically significant and was characterized by high heterogeneity. Rinciog et al.19 estimated acute exacerbations at 0.56 (95% CI = 0.35–0.89) for nintedanib and 1.10 (95% CI = 0.43–2.85) for pirfenidone, indicating predominance of pirfenidone. It is worth noting that the researchers did not include the study of Ogura et al.39 for nintedanib or the studies of Azuma et al.35, Huang et al.40, King37, or Taniguchi et al.36 for pirfenidone.

The NMA showed that placebo resulted in fewer dropouts. Pirfenidone caused fewer dropouts compared to nintedanib (OR = 0.75 (0.33–1.27)), the difference was statistically significant and with little heterogeneity. It is noteworthy that the profile of adverse events of the two treatments is similar, however treatment with nintedanib seems to cause more discontinuations. In terms of safety, these results indicate that pirfenidone may be a more stable option.
With regards to mortality the NMA did not show a difference between the two drugs [OR = 0.93 (0.38–1.94)], with the results being statistically significant and with little heterogeneity. Canestaro et al. estimated an OR (95% CI) of 0.87 (0.48–1.66) for pirfenidone vs nintedanib, which confirms our results. Loveman et al. estimated an OR (95% CI) = 1.39 (0.70–2.82) for nintedanib vs pirfenidone.

With regards to the decline in FVC, there was no difference between the two drugs, OR (95% CI) = 0.91 (0.45–2.03). Fleetwood et al. reported results similar to those of the present study, OR (95% CI) = 0.90 (0.50–1.66) pirfenidone vs nintedanib, while Loveman et al. estimated an OR (95%) of 0.67 (0.51–0.88).

For the serious adverse events (SAEs), there was no difference between the drugs (OR = 1.02, 95% CI = 0.62–1.62). Rochwerg et al. reported similar results (OR = 1.04, 95% = 0.51–2.24). Yoon et al. reported on the safety of nintedanib in patients with advanced IPF. They found that 62.7% of the patients with advanced IPF and 47.4% of those with non-advanced IPF suffered from SAEs. The advanced IPF group also had more dropouts due to adverse events (68% vs 40% in the non-advanced group).

Flaherty et al. examined the tolerability of pirfenidone-nintedanib combination therapy, assessing the occurrence of treatment-emergent adverse events (TEAEs). Of the 89 patients, 78% completed the 24 week single-arm, open label study, with the profile and rate of TEAEs proportionally similar to that of monotherapies. The researchers reported that nintedanib was linked to a higher rate of discontinuation due to TEAEs (10/89) compared to pirfenidone (0/89), but commented that this may be influenced by the fact that the included patients were already tolerable to pirfenidone. Additionally, they note that, due to the similar adverse event profile of the two therapies, there was uncertainty in identifying the treatment responsible for some of the TEAEs. Vancheri et al. also reported on the safety of combination

---

**Table 1. Synthesis of the results and network quality assessment.**

| Outcome Measure | Model type | Pirfenidone vs Placebo, estimate (95% CI) | Nintedanib vs Placebo, estimate (95% CI) | Pirfenidone vs Nintedanib, estimate (95% CI) | Mean residual deviance |
|----------------|------------|------------------------------------------|------------------------------------------|-------------------------------------------|------------------------|
| Decline in percent predicted FVC ≥ 10% | OR Random effect | 0.54 (0.37–0.80) | 0.59 (0.41–0.84) | 0.91 (0.45–2.03) | 12.17 |
| Acute exacerbations | OR Random effect | 0.77 (0.26–2.28) | 0.61 (0.20–1.90) | 0.39 (0.00–15.53) | 11.57 |
| Mortality | OR Random effect | 0.68 (0.46–1.03) | 0.70 (0.45–1.09) | 0.93 (0.38–1.94) | 14.76 |
| Dropouts | OR Random effect | 1.27 (0.96–1.68) | 1.61 (1.13–2.28) | 0.75 (0.33–1.27) | 24.11 |
| Serious adverse events | OR Random effect | 1.00 (0.78–1.29) | 0.99 (0.78–1.28) | 1.02 (0.62–1.62) | 11.51 |

Inconsistency; Imprecision; Poor quality of direct evidence.

---

**Table 2. Results of NNTB/NNTH estimation.**

| Decline in percent predicted FVC ≥ 10% | NNTB or NNTH | Pirfenidone vs Placebo | Nintedanib vs Placebo |
|-------------------------------------|--------------|------------------------|------------------------|
| NNTB | NNTH 9 (NNTB 7 to NNTB 22) | NNTH 12 (NNTH 7 to NNTH 58) | NNTH 9 (NNTB 6 to NNTB 23) |
| NNTH | NNTH 14 (NNTH 8 to NNTH 61) | NNTH 14 (NNTH 8 to NNTH 61) | NNTH 14 (NNTH 8 to NNTH 61) |
therapy, and found that gastrointestinal adverse events occurred in 69.8% of patients treated with nintedanib with add-on pirfenidone and 52.9% of patients treated with nintedanib.

Crestani et al.44 examined the long-term safety of nintedanib. In a placebo-controlled, open label, extension of the INPULSIS trials they found that diarrhea was the most common adverse event, with 15% of the patients discontinuing because of it. Progression of disease was the most common reason of discontinuation (26% of patients). In general, the rates of adverse events were maintained in the long-term study, and the researchers argue that the results prove the safety of nintedanib as a solution to the management of disease progression after initial treatment.

NNTH-NNTB indicated that nintedanib was better than pirfenidone in every outcome apart from acute exacerbations and FVC decline where there was no difference. NNT values have not been presented previously on these particular therapies for IPF, and they show that a different clinical interpretation of the outcomes may lead to the conclusion that nintedanib has an advantage over pirfenidone. Clinical professionals can study these statistics and assess how their own practice will affect their patients.

The quality assessment analysis showed that the results of the pairwise analysis were characterized by sufficient quality, with the exception of Nintedanib vs Placebo for acute exacerbations.

The results of the risk of bias analysis showed low risk for all clinical studies apart from the Taniguchi et al.36 study.

The strengths of this systematic review and NMA include the selection of those outcomes (clinically important decline in FVC, mortality, acute exacerbations, serious adverse events, and dropouts) that are important in policy decision-making. FVC decline \( \geq 10\% \) and acute exacerbations, in particular, have not been examined thoroughly in recent studies. Additionally, the calculation of NNT adds to its value as a study that can inform clinical practitioners. NNTH/NNTB is a measure of clinical effectiveness that has not been presented previously in other studies. Finally, the inclusion of all the most recent RCTs on the examined therapies for IPF makes the present study unique. This study uses all the eligible available information and focuses on the two best therapeutic options for IPF, thus providing a complete set of data regarding the effectiveness of nintedanib and pirfenidone.

The limitations of our review include the small number of studies included, resulting in low certainty in estimates for the comparisons. Additionally, there were no head-to-head trials between nintedanib and pirfenidone to compare the result of the indirect comparison.

**Conclusion**

In conclusion, the results of the NMA suggest that both pirfenidone and nintedanib are superior to placebo, but that they do not demonstrate a clear case of dominance of one over the other. The difference in the occurrence of dropouts indicates a slight clinical advantage of pirfenidone when it comes to safety, and this may be of interest when choosing an appropriate long-term solution for IPF, while the NNTH/NNT results favor nintedanib as a treatment option. These results show the presence of marginal differences between the two therapies, and a comparison of the costs for the two therapies may be able to compliment the process of making a cost-effective decision with respect to IPF therapy.

**Key points**

- Review of all the RCTs on the therapy of IPF with pirfenidone and nintedanib up to September 2017.
- Use of NNTH and NNTB to estimate the therapeutic impact and provide a statistical figure, more meaningful to a health professional.
- Indirect comparison of the two treatments indicated a decreased dropout rate associated with pirfenidone, while the number needed to treat was smaller in the case of nintedanib, suggesting a better overall clinical effectiveness.
- Proof that an economic comparison may be useful to the decision-making process regarding the reimbursement of therapy for IPF.

**Transparency**

**Declaration of funding**

This manuscript received no funding.

**Declaration of financial/other relationships**

The authors and JDA peer reviewers on this manuscript have no relevant financial or other relationships to disclose.

**Acknowledgements**

None reported.

**References**

[1] Ryu JH, Moua T, Daniels CE, et al. Idiopathic pulmonary fibrosis: evolving concepts. Mayo Clin Proc. 2014;89:1130–1142.
[2] Nalysnyk L, Cid-Ruzafa J, Routella P, et al. Incidence and prevalence of idiopathic pulmonary fibrosis: review of the literature. European Respiratory Review. 2012;21:355–361.
[3] Harari S, Madotto F, Caminati A, et al. Epidemiology of idiopathic pulmonary fibrosis in Northern Italy. Ciccozzi M, editor. PLoS One. 2016;11:e0147072.
[4] Baumgartner KB, Samet JM, Coultas DB, et al. Occupational and environmental risk factors for idiopathic pulmonary fibrosis: a multicenter case-control study. Collaborating Centers. Am J Epidemiol. 2000;152:307–315.
[5] Michaelson JE, Agauyo SM, Roman J. Idiopathic pulmonary fibrosis: a practical approach for diagnosis and management. Chest. 2000;118:788–794.
[6] National Clinical Guideline Centre (UK). Diagnosis and management of suspected idiopathic pulmonary fibrosis: idiopathic pulmonary fibrosis [Internet]. Natl Inst Heal Care Excell Clin Guidel. 2013.
[7] Raghu G, Collard HR, Egan JJ, et al. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based
