NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282) [ID837]

The following documents are made available to the consultees and commentators:

1. **Final Appraisal Determination (FAD)** as issued to consultees and commentators in December 2016

2. **Appeal Decision**

3. **ERG addendum** (these results were presented at the last committee meeting for pirfenidone and informed the recommendations in the FAD)

Any information supplied to NICE which has been marked as confidential, has been redacted. All personal information has also been redacted.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Final appraisal determination

Pirfenidone for treating idiopathic pulmonary fibrosis

This guidance is a review of NICE's technology appraisal guidance for pirfenidone for treating idiopathic pulmonary fibrosis (TA282). The review looked at a different patient access scheme, and considered including people with a forced vital capacity above 80% predicted and removing the stopping rule. However, no changes to the recommendations in TA282 have been made.

1 Recommendations

1.1 Pirfenidone is recommended as an option for treating idiopathic pulmonary fibrosis in adults only if:

- the person has a forced vital capacity (FVC) between 50% and 80% predicted
- the company provides pirfenidone with the discount agreed in the patient access scheme and
- treatment is stopped if there is evidence of disease progression (an absolute decline of 10% or more in predicted FVC within any 12-month period).

1.2 This guidance is not intended to affect the position of patients whose treatment with pirfenidone was started within the NHS before this guidance was published. Treatment of those patients may continue without change to whatever funding arrangements were in place for them before this guidance was published until they and their NHS clinician consider it appropriate to stop.

2 The technology

Description of the technology	Pirfenidone (Esbriet, Roche) is an oral immunosuppressant with anti-inflammatory and antifibrotic effects.
Marketing authorisation	Pirfenidone has a marketing authorisation in the UK for treating mild to moderate idiopathic pulmonary fibrosis in adults.
Adverse reactions	The summary of product characteristics states that the very common adverse reactions (affecting 1 in 10 or more people) associated with using pirfenidone are nausea, rash, diarrhoea, fatigue, dyspepsia, anorexia, headache and photosensitivity reactions. For full details of adverse reactions and contraindications, see the summary of product characteristics.
Recommended dose and schedule	The recommended dosage of pirfenidone is three 267 mg capsules 3 times daily (that is, a total of 2,403 mg per day).
Price	The list price of pirfenidone is £501.92 for 63 capsules (excluding VAT; British national formulary online, accessed May 2016). This equates to a daily cost of £71.70. The company has agreed a patient access scheme with the Department of Health. This scheme provides a simple discount to the list price of pirfenidone, with the discount applied at the point of purchase or invoice. The level of discount is commercial in confidence. The Department of Health considered that the patient access scheme does not constitute an excessive administrative burden on the NHS.

3 Evidence

The appraisal committee (section 7) considered evidence submitted by Roche Products and a review of this submission by the evidence review group. See the committee papers for full details of the evidence.

4 Committee discussion

Review objectives

4.1 The appraisal committee reviewed existing and new data on the clinical and cost effectiveness of pirfenidone, having considered evidence on the nature of idiopathic pulmonary fibrosis and the value placed on the

benefits of pirfenidone by people with the condition, those who represent them, and clinical experts. It also took into account the effective use of NHS resources. The committee recognised that this appraisal reviewed NICE's previous technology appraisal guidance on pirfenidone, and that the company had proposed to expand the recommendation to include people with a forced vital capacity (FVC) above 80% predicted and remove the stopping rule, that is, to no longer stop pirfenidone after an absolute decline of 10% or more in predicted FVC within any 12-month period.

Current practice

4.2 The committee discussed the treatments for idiopathic pulmonary fibrosis in current NHS practice. The committee heard the clinical experts explain that they offer treatment with pirfenidone or nintedanib to people with an FVC between 50% and 80% predicted; this reflects NICE's previous technology appraisal guidance on pirfenidone and nintedanib for treating idiopathic pulmonary fibrosis. The clinical experts noted that they would offer best supportive care to people with an FVC above 80% predicted because NICE does not recommend pirfenidone or nintedanib in this population. The committee noted comments received during consultation that clinicians would like to offer active treatments to people with an FVC above 80% predicted. The committee concluded that the current treatment options are nintedanib and pirfenidone for people with an FVC between 50% and 80% predicted, and best supportive care for those with an FVC above 80% predicted.

4.3 The committee discussed the stopping rule and how this is implemented in clinical practice. It heard from clinical experts that they follow the stopping rule in NICE's previous technology appraisal guidance on pirfenidone and nintedanib, that is, an absolute decline of 10% or more in predicted FVC within any 12-month period, and that they often confirm the drop in percent predicted FVC with repeat testing. The committee heard

from the clinical experts during the committee meeting and from the comments received during consultation about the limitations of using percent predicted FVC to assess lung function in people with idiopathic pulmonary fibrosis. It heard that the disease course varies, both between patients and over time. It heard that there were periods of relative stability interspersed with acute exacerbations, and that acute exacerbations are associated with permanently reduced lung function and an increased risk of dying. The committee noted that the clinical experts could not suggest a better way of objectively defining treatment success or failure than using percent predicted FVC. It noted the limitations of FVC but understood that, in clinical practice, the wider patient characteristics would be taken into account in interpreting percent predicted FVC.

Population

4.4 The committee considered the population relevant to the appraisal. It recognised that the company had presented analyses with different subgroups of people with idiopathic pulmonary fibrosis. The committee agreed to consider subgroups relevant to the review objectives.

4.5 The committee discussed the appropriate population to determine whether the recommendation could be expanded to the population with a percent predicted FVC above 80%:

- The committee was aware that in NICE's previous technology appraisal for pirfenidone, the committee had heard that 'treatment decisions were made after taking multiple factors into account (such as symptoms and degree of deterioration) and would not be wholly based on lung function test results'. It had also heard that 'FVC 80% predicted or less was an arbitrary but acceptable threshold for initiating treatment for idiopathic pulmonary fibrosis and that this could be easily implemented in clinical practice'. In the current appraisal, the committee agreed that the treatment starting rule of FVC of 80% predicted was arbitrary, but recognised that this threshold was used in clinical practice.

- The committee was aware that NICE's guide to the methods of technology appraisal states that subgroups should be identified based on 'an expectation of differential clinical or cost effectiveness because of known biologically plausible mechanisms, social characteristics or other clearly justified factors'. The committee acknowledged that there were no known biologically plausible mechanisms for identifying subgroups, but noted that the company stated in its response to the appraisal consultation document that idiopathic pulmonary fibrosis 'is a complex disease that is not yet fully understood for subgroups'.
- The committee noted the evidence review group's (ERG) comments that Albera 2016 showed that over 12 months, more people with an FVC between 50% and 80% predicted died, or had a greater than 10% decline in predicted FVC than people with an FVC above 80% predicted. The ERG also advised that this difference in prognosis would affect the modelling of cost-effectiveness, because people with an FVC above 80% predicted are treated with pirfenidone for longer and generate relatively more costs than benefits. The committee also noted from the analyses presented that there was a consistent trend for higher ICERs when considering the FVC above 80% predicted group than the whole population (see sections 4.18 to 4.20). Given that one of the objectives of the review was to consider whether pirfenidone was cost effective in people with an FVC above 80% predicted, the committee agreed that the most accurate way to do this would be to consider the cost effectiveness in people with an FVC above 80% predicted, rather than for the whole population.
- The committee recognised that the company had presented some analyses with an upper limit of FVC of 90% predicted, because most of the data presented by the company was supported by patients with an FVC up to 90% predicted. The committee agreed that the evidence was only generalisable to people with an FVC of up to 90% predicted.

The committee concluded that the subgroup of people with an FVC between 80% and 90% predicted was the relevant population for decision-making.

4.6 The committee discussed the appropriate population to determine whether the stopping rule could be retained or removed for people with an FVC between 50% and 80% predicted, for which pirfenidone is currently recommended. The committee recognised it would need to consider whether to include the stopping rule if the recommendation was expanded to people with FVC predicted greater than 80%. The committee therefore considered the populations with an FVC between 50% and 80% predicted, and between 80% and 90% predicted when considering the stopping rule.

Comparators

4.7 The committee discussed whether best supportive care or nintedanib were relevant comparators:

- The committee recalled that people with an FVC between 80% and 90% predicted was the relevant population when considering whether to expand the population (see section 4.5), and that clinical experts stated they would currently offer best supportive care to people with an FVC above 80% predicted (see section 4.2).
- The committee recognised that in considering whether to retain or remove the stopping rule, the relevant population included those with FVC between 50% and 80% predicted, for whom nintedanib was also a comparator. The committee noted that the decision to include a stopping rule was based on a comparison with best supportive care, and agreed it was more appropriate to use this comparison to determine whether to retain or remove the stopping rule.

The committee agreed it was appropriate to compare pirfenidone with best supportive care and it did not consider the comparison with nintedanib further.

Clinical effectiveness

4.8 The committee considered the clinical evidence presented by the company. It noted that the evidence came from 4 randomised double-blind placebo-controlled phase III trials (CAPACITY 1, CAPACITY 2, ASCEND and SP3) and other observational data:

- The committee was aware that the results of SP3 and the CAPACITY trials were considered during NICE's previous technology appraisal of pirfenidone. It recognised that the new data presented by the company came from ASCEND, RECAP (an open label extension follow-up study of the CAPACITY trials) and observational data for best supportive care (the 'INOVA' registry).
- The committee noted that the populations differed across the trials in how they were defined by percent predicted FVC: the CAPACITY trials recruited patients with an FVC above 50% predicted with no upper limit, ASCEND recruited patients with an FVC between 50% and 90% predicted, and the investigators responsible for the SP3 trial did not specify the range, but reported an average baseline FVC of 77% predicted.
- The committee noted that the primary end point in both ASCEND and the CAPACITY trials was the change in percent predicted FVC from baseline, and that this was after 52 weeks in ASCEND and after 72 weeks in CAPACITY 1 and 2.
- The committee was aware that the company and the ERG included data from ASCEND in their network meta-analyses with data from CAPACITY 1, CAPACITY 2 and SP3. It concluded that the trials included in the ERG's meta-analysis were appropriate and relevant to clinical practice in England.

Evidence in the population with FVC between 80% and 90% predicted

4.9 The committee discussed the effectiveness of pirfenidone in people with an FVC between 80% and 90% predicted. It noted that the company presented:

- A pre-specified analysis using the CAPACITY trials in 3 pre-defined subgroups: predicted FVC at baseline of more than 80%; between 70% and 80%; and lower than 70%. This showed there was not a statistically significant estimate for better outcomes in the placebo group than in the pirfenidone group among people with a baseline FVC above 80% predicted.
- An analysis of covariance (ANCOVA) in 2 subgroups (people with an FVC above 80% predicted at baseline, and people with an FVC of 80% predicted or less). The outcome considered in ANCOVA was the change in predicted FVC at week 52, and this was presented for each clinical trial individually. The committee understood that the results of the treatment-by-subgroup interaction tests showed that there was no statistically significant difference in results between the 2 subgroups. The committee recognised, however, that a non-significant interaction test does not conclusively mean that there is no difference in treatment effect between subgroups because the test may not have been powered to detect a difference between the subgroups. It further noted that there was a smaller treatment effect in the FVC above 80% predicted subgroup than in the FVC 80% predicted or less subgroup in each trial. In the CAPACITY 1 trial, the treatment effect was no longer statistically significant when considering the FVC above 80% predicted group alone.

4.10 The committee considered the evidence presented for people with an FVC above 80% predicted. It agreed that it was only generalisable to people with an FVC between 80% and 90% predicted because most of the data presented by the company were supported by patients with an

FVC up to 90% predicted. The committee observed that none of the studies were designed to specifically determine the effectiveness of pirfenidone in people with FVC between 80% and 90% predicted, or to compare this group with those with an FVC between 50% and 80% predicted. The committee acknowledged the practical difficulties in designing studies to detect differences in outcomes between subgroups. However, it agreed that these results were relevant to the decision problem (see section 4.5). The committee concluded that:

- Pirfenidone may reduce disease progression, although the results from these analyses were not robust.
- There was no statistically significant reduction in mortality with pirfenidone compared with placebo.
- It was not clear whether pirfenidone was more, less or equally effective in the group with FVC above 80% predicted than in the group with FVC of 80% or less predicted. However, the committee agreed that, from the evidence presented, it was more likely to be less effective.
- It was unlikely that the results from the whole population were generalisable to those with FVC between 80% and 90% predicted, because this subgroup represented an earlier stage in the disease pathway and a different baseline mortality rate.

4.11 The committee also discussed the effectiveness of pirfenidone in the whole population. It was aware that in NICE's previous technology appraisal for pirfenidone, which considered the population with FVC between 50% and 80% predicted, the committee concluded that 'pirfenidone seemed to have a modest but measurable effect on slowing the decline in lung function'. The committee acknowledged that the company had provided new long-term and mature data in this appraisal, mainly relating to mortality. The committee considered the results of the ERG's meta-analyses. It agreed that pirfenidone reduced disease progression compared with placebo. It also agreed that there was evidence that it may reduce mortality. It concluded that it had not seen

anything to contradict the conclusion in NICE's previous technology appraisal. The committee concluded that pirfenidone remains effective in people with an FVC between 50% and 80% predicted.

Evidence on the stopping rule

4.12 The committee discussed whether to retain or remove the stopping rule, that is, stopping pirfenidone after disease progression, defined by an absolute decline of 10% or more in predicted FVC within any 12-month period. It recalled that clinicians follow the rule in clinical practice (see section 4.3), however, it noted the consultation comments from professional groups that clinicians would like to continue treating people after disease progression because the treatment may still be beneficial. The committee agreed that not all treatments are universally effective, and that stopping rules improve the cost effectiveness of a technology by stopping treatment when it is no longer considered clinically effective.

4.13 The committee considered the company's post-hoc subgroup analyses of patients who continued pirfenidone after a decline in predicted FVC of 10% or more within any 12-month period. These data showed that fewer people in the pirfenidone group (1 patient out of 24) experienced a further 10% decline in predicted FVC compared with those in the placebo group (15 patients out of 60; p=0.032). The committee was concerned with the results of this analysis because:

- the sample size of 84 patients was small, meaning that decisions based on this subgroup are uncertain
- the analysis broke the randomisation of the clinical trials
- to test the hypothesis that people benefit from continuing pirfenidone treatment after disease progression, it would be more informative to compare people who do not stop pirfenidone after disease progression with people who do stop it after disease progression, rather than to compare people continuing pirfenidone after disease progression with people who had been randomised to placebo at baseline.

The committee concluded that the company's evidence did not conclusively show that people continue to benefit from pirfenidone after disease progression. However, the committee recognised the comments from clinical experts that some patients may benefit after disease progression. It concluded it was appropriate to consider cost-effectiveness analysis with and without the stopping rule.

Cost effectiveness

4.14 The committee considered the company's partitioned survival model, which had 3 mutually exclusive health states: progression-free, progressed and dead. It heard from the ERG that the company's model did not capture the progressive nature of idiopathic pulmonary fibrosis. The committee noted that the company, in choosing the model structure, made several clinically implausible assumptions. For example:

- No relationship between time on treatment, time to disease progression (defined as a 10% decline in predicted FVC, a decline in 6–minute walking distance of 50 metres or more, or death) and mortality. The committee agreed that these were likely to be linked, so it was not appropriate to model them independently.
- Acute exacerbations were not connected to disease progression and mortality. Clinical experts advised that exacerbations had a substantial impact on quality of life and mortality. The committee agreed that the model may not fully represent the impact of idiopathic pulmonary fibrosis on patients.

The committee had serious concerns about the company's model and understood that the ERG could address only some of the issues in its exploratory analyses. The committee noted that, in its response to the appraisal consultation document, the company did not provide new clinical evidence but did provide revised analyses with:

- New parametric survival curves for mortality (see section 4.15 for discussion about the survival curves).
- New assumptions around how long the benefits of treatment last (see section 4.16 for discussion about the time the benefits of treatment last).
- A new subgroup with an FVC between 50% and 90% predicted. The committee considered the cost effectiveness of pirfenidone for the group with FVC between 80 to 90% predicted as more relevant (see sections 4.5 and 4.18 for the discussion about the subgroup).

In its response to consultation, the company suggested that it based its model on mortality because of the data available and, in its view, increasing complexity would not improve the fit to available data on costs and utilities. The committee acknowledged the limitations in the data and concluded that the model could be used for its decision-making.

4.15 The committee discussed how the company estimated the potential long-term mortality benefit with pirfenidone over a patient's lifetime by extrapolating from relatively short trials. It noted that the choice of parametric curve for mortality was a key driver of the cost-effectiveness results, and that the company estimated long-term mortality based on mortality data from ASCEND, CAPACITY 1 CAPACITY 2 and RECAP. The committee agreed that this was appropriate. It noted that the company had modelled mortality using the Weibull distribution, and that the ERG had used the Gompertz distribution.

- It heard from the ERG that, although the Weibull distribution fitted the observed data well, it predicted a lower probability of death for older people than in the general UK population; the ERG did not consider this to be clinically plausible.
- The ERG considered that the Gompertz distribution also fitted the data well but provided more clinically plausible long-term estimates for mortality beyond the observed data.

- In its response to the appraisal consultation document, the company adjusted the annual probability of death by age distribution in the UK population to show that the Weibull distribution no longer predicted a lower probability of death until the age of around 90 years. The company suggested that registry data (that is, the INOVA registry of people with idiopathic pulmonary fibrosis on best supportive care in the United States) supported the estimates of mortality using the Weibull rather than Gompertz distribution.
- The ERG accepted that the adjustment addressed one of their concerns, but remained concerned that INOVA had a high proportion of censoring after 10 years (that is, people for whom there were no additional data).
- The ERG acknowledged that there was uncertainty associated with the Gompertz distribution because it did not fit the data as well as the Weibull distribution, and advised the committee that both curves were plausible.
- However, the committee observed that, with the Weibull distribution, people with idiopathic pulmonary fibrosis had a risk of death that was unrelated to the length of time the person had had idiopathic pulmonary fibrosis. With the Gompertz distribution, the risk of death increased with disease duration. The committee agreed that the risk of death with idiopathic pulmonary fibrosis was likely to increase with length of time with the disease relative to the general population, and that the true risk of death of people with idiopathic pulmonary fibrosis might lie between the 2 curves, but closer to the Gompertz distribution.
- The committee also considered the company's new analysis that used a weighted parametric survival curve based on a statistical test of model fit (that is, the Akaike information criterion). The committee agreed with the ERG that it did not consider the approach credible because it included curves (such as a log-logistic distribution) with limited clinical plausibility, and agreed that it was not appropriate.

On balance, the committee acknowledged the company's different opinion on the choice of parametric survival curve and agreed to take the Weibull and Gompertz curves into account in its decision-making. However, the committee concluded that it was more plausible to use the Gompertz distribution to estimate mortality.

4.16 The committee discussed the company's modelling assumption that the mortality benefit of pirfenidone compared with best supportive care remains constant over a person's lifetime. It appreciated that the randomised trials were too short (that is, either 52 or 72 weeks) to provide evidence to support this assumption. In addition, the committee noted that, although follow-up data for pirfenidone were collected for over 8 years in RECAP, there was no best supportive care group in the study and so no long-term relative effectiveness could be estimated from the study. The committee was also aware that the model was very sensitive to the assumptions around duration of treatment benefit. It heard from 1 clinical expert that the treatment benefit of pirfenidone was likely to be constant over a person's lifetime. The committee did not agree that this was plausible, based on advice from the ERG that the trials showed a reduction in treatment effect over time for mortality (see section 4.10). In its response to the appraisal consultation document, the company used data from INOVA to support its argument that the treatment effect lasts for at least 8 years. The company provided a Kaplan–Meier plot and log-cumulative hazard plot for mortality based on data from the trials and INOVA. The committee noted that the log-cumulative hazard plots for pirfenidone and best supportive care were not parallel after 5 years. It determined that, although there was some maintained treatment effect, it was not constant. The committee concluded that the evidence did not justify assuming a constant mortality benefit for 8 years. It further concluded that it was reasonable to assume a constant benefit up to 5 years.

4.17 The committee discussed whether the company's model appropriately incorporated the treatment stopping rule. The ERG explained that incremental cost-effectiveness ratios (ICERs) including the stopping rule for pirfenidone would likely be biased in favour of pirfenidone when compared with best supportive care. This was because, in the model, the stopping rule reduced pirfenidone costs without affecting treatment outcomes. The committee concluded that analyses including a stopping rule for pirfenidone would underestimate the ICER because of the model structure.

Cost-effectiveness results and conclusions

4.18 The committee discussed whether to expand the current pirfenidone recommendation, and considered the cost effectiveness for the group with FVC between 80 and 90% predicted (see section 4.5). It noted that the analysis presented by the company included people with FVC above 80% predicted, and it had not been presented with cost-effectiveness results for the specific group of people with an FVC between 80% and 90% predicted. It noted that the ICER (including the stopping rule) that most closely matched the committee's preferred assumptions for pirfenidone compared with best supportive care ranged from £32,643 (Weibull) to £38,687 (Gompertz) per QALY gained, but the upper estimate was more plausible. The committee also recognised that this was likely to be an underestimate because:

- the method used to incorporate the stopping rule adjusted only costs, not benefits (see section 4.14) and
- uncertainty remained about how long treatment benefit would last; it could be less than 5 years (see section 4.15).

The committee recalled that it had not seen robust analyses showing that pirfenidone reduces mortality, or consistently reduced the decline in percent predicted FVC, in people with an FVC between 80% and 90% predicted (see sections 4.9 and 4.10). Therefore, the committee

concluded that pirfenidone could not be considered an effective use of NHS resources for people with an FVC between 80% and 90% predicted.

4.19 The committee considered the cost effectiveness of retaining or removing the requirement to stop treatment if a person's predicted FVC drops by 10%, in the population for whom pirfenidone is already available (that is, with an FVC between 50% and 80% predicted). It noted that the company submitted cost-effectiveness analyses with and without the stopping rule for this group. The ICERs without the stopping rule that most closely matched the committee's preferred assumptions for pirfenidone compared with best supportive care were between £35,905 (Weibull) and £40,110 (Gompertz) per QALY gained, but the upper estimate was more plausible. The committee also recognised that this was likely to be an underestimate because uncertainty remained about how long treatment benefit would last; it could be less than 5 years (see section 4.14). The committee concluded that pirfenidone was not cost effective without the stopping rule in this group.

4.20 The committee was aware that, in NICE's previous technology appraisal guidance, pirfenidone was regarded as cost effective for people with an FVC between 50% and 80% predicted if the company provided pirfenidone with the discount agreed in the patient access scheme. The committee agreed that it had not seen any evidence to alter its conclusion from NICE's previous technology appraisal on pirfenidone (see sections 4.16 and 4.17). It noted that the ICER (including the stopping rule) that most closely matched the committee's preferred assumptions for pirfenidone compared with best supportive care was between £24,933 (Weibull) and £27,780 (Gompertz) per QALY gained. The committee concluded that the recommendations in NICE's previous technology appraisal guidance on pirfenidone remained appropriate for people with idiopathic pulmonary fibrosis with an FVC between 50% and 80% predicted.

Innovation

4.21 The committee discussed whether pirfenidone was an innovative treatment. The committee noted the company's suggestion that pirfenidone is associated with health-related benefits that cannot be adequately captured in the QALY calculation. These benefits include: a reduction in breathlessness; improved patient choice based on a different adverse event profile; improved NHS capacity by reducing inpatient stays attributed to acute exacerbations; and the effect on people of working age. The committee acknowledged that, although these aspects were important to people with idiopathic pulmonary fibrosis, it did not consider that any sizeable health-related benefits had been excluded from the economic model and did not alter its conclusions.

Potential equality issues

4.22 The committee noted the potential equality issue raised by consultees that restricting treatment based on percent predicted FVC could discriminate against:

- minority ethnic groups, particularly people of south Asian family origin
- disabled people who have difficulty standing straight because FVC is expressed as a percentage of the predicted normal value for a person of the same height
- older people because the reference tables are derived from populations under the age of 70 years, whereas the average age of people with idiopathic pulmonary fibrosis is 72 years.

The committee discussed these issues with the clinical experts, noting that:

- The Global Lung Initiative has introduced equations to predict FVC values in minority ethnic groups and, when these equations were used, FVC values for these groups were comparable to the FVC values of

people in clinical trials. Thus, when using the newer equations, people would not be denied treatment because of their ethnicity.

- For people who cannot stand straight, their armspan (which approximates their height) can be used to calculate percent predicted FVC. Thus, when using this measure people would not be denied treatment because of their disability.
- According to clinical experts, it is difficult to compare the predicted FVC values of older people with the FVC values of people in clinical trials because older people show a wide range of predicted FVC.

The committee recognised the limitations of FVC but understood that, in clinical practice, wider patient characteristics would be taken into account when interpreting percent predicted FVC. It concluded that its recommendations did not discriminate against any groups of people protected by the Equality Act.

The Pharmaceutical Price Regulation Scheme

4.23 The committee was aware of NICE's position statement on the Pharmaceutical Price Regulation Scheme (PPRS) 2014 and, in particular, the PPRS payment mechanism. It accepted the conclusion 'that the 2014 PPRS payment mechanism should not, as a matter of course, be regarded as a relevant consideration in its assessment of the cost effectiveness of branded medicines'. The committee heard nothing to suggest that there is any basis for taking a different view about the relevance of the PPRS to this appraisal. It therefore concluded that the PPRS payment mechanism was not relevant in considering the cost effectiveness of the technology in this appraisal.

Summary of appraisal committee's key conclusions

TAXXX	**Appraisal title: Pirfenidone for treating idiopathic pulmonary fibrosis**	**Section**

Key conclusion	
Pirfenidone continues to be recommended as an option for treating idiopathic pulmonary fibrosis in adults only if: • the person has a forced vital capacity (FVC) between 50% and 80% predicted • the company provides pirfenidone with the discount agreed in the patient access scheme and • treatment is stopped if there is evidence of disease progression (an absolute decline of 10% or more in predicted FVC within any 12-month period). The committee concluded that pirfenidone was not cost effective without the stopping rule in this group. The incremental cost-effectiveness ratios (ICERs) without the stopping rule that most closely matched the committee's preferred assumptions for pirfenidone compared with best supportive care were between £35,905 (Weibull) and £40,110 (Gompertz) per quality-adjusted life year (QALY) gained in the population for whom pirfenidone is already available (that is, with an FVC between 50% and 80% predicted). The committee agreed that the recommendations in NICE's previous technology appraisal guidance on pirfenidone remained appropriate for people with idiopathic pulmonary fibrosis with an FVC between 50% and 80% predicted. It concluded that pirfenidone could not be considered an effective use of NHS resources for people with an FVC between 80% and 90% predicted. The ICER (including the stopping rule) that most closely matched the committee's preferred assumptions for pirfenidone compared with best supportive care ranged from £32,643 (Weibull) to £38,687 (Gompertz) per QALY gained.	1.1, 4.18 to 4.20

Current practice		
Clinical need of patients, including the availability of alternative treatments	The clinical experts noted that they would offer best supportive care to people with an FVC above 80% predicted. The committee concluded that the current treatment options are nintedanib and pirfenidone for people with an FVC between 50% and 80% predicted, and best supportive care for those with an FVC above 80% predicted.	4.2
The technology		
Proposed benefits of the technology How innovative is the technology in its potential to make a significant and substantial impact on health-related benefits?	The committee did not consider that any sizeable health-related benefits had been excluded from the economic model.	4.21
What is the position of the treatment in the pathway of care for the condition?	Pirfenidone has a marketing authorisation in the UK for treating mild to moderate idiopathic pulmonary fibrosis in adults.	2
Adverse reactions	The very common adverse reactions associated with using pirfenidone are nausea, rash, diarrhoea, fatigue, dyspepsia, anorexia, headache and photosensitivity reactions.	2

Evidence for clinical effectiveness		
Availability, nature and quality of evidence	The committee noted that the evidence came from 4 randomised double-blind placebo-controlled phase III trials (CAPACITY 1, CAPACITY 2, ASCEND and SP3) and other observational data.	4.8
Relevance to general clinical practice in the NHS	The committee concluded that the trials included in the meta-analysis were appropriate and relevant to clinical practice in England.	4.8
Uncertainties generated by the evidence	The committee concluded that pirfenidone may reduce disease progression in people with an FVC above 80% predicted, although the results from these analyses were not robust. It concluded that it was not clear whether pirfenidone was more, less or equally effective in the group with FVC above 80% predicted than in the group with FVC of 80% or less predicted.	4.10
Are there any clinically relevant subgroups for which there is evidence of differential effectiveness?	The committee concluded that the subgroup of people with an FVC between 80% and 90% predicted was the relevant population for decision-making. It agreed that the results were only generalisable to people with an FVC between 80% and 90% predicted.	4.5, 4.10
Estimate of the size of the clinical effectiveness	An analysis showed there was not a statistically significant estimate for better outcomes in the placebo group than in the	4.9, 4.10, 4.11

including strength of supporting evidence	pirfenidone group among people with a baseline FVC above 80% predicted. A different analysis showed that the treatment effect was no longer statistically significant when considering the FVC above 80% predicted group alone. It also concluded that pirfenidone may reduce disease progression and that there was no statistically significant reduction in mortality with pirfenidone compared with placebo in people with an FVC between 80% and 90% predicted. The committee concluded that pirfenidone remained effective in people with an FVC between 50% and 80% predicted.	
How has the new clinical evidence that has emerged since the previous appraisal (TA282) influenced the current recommendations?	The changes to NICE's technology appraisal guidance 282 proposed by the company in light of new clinical data are not recommended, specifically: • removing the recommendation to stop pirfenidone if the disease progresses • expanding the population to include people with an FVC above 80% predicted.	1.1, 4.1
Evidence for cost effectiveness		
Availability and nature of evidence	The company provided a partitioned survival model which had 3 mutually exclusive health states: progression-free, progressed and dead.	4.14
Uncertainties around and plausibility of	The committee acknowledged the limitations in the data and concluded that the model	4.14

assumptions and inputs in the economic model	could be used for its decision-making. The committee noted that there was no relationship between time on treatment, time to disease progression and mortality and that acute exacerbations were not connected to disease progression and mortality.	
Incorporation of health-related quality-of-life benefits and utility values Have any potential significant and substantial health-related benefits been identified that were not included in the economic model, and how have they been considered?	The committee did not consider that any sizeable health-related benefits had been excluded from the economic model.	4.21
Are there specific groups of people for whom the technology is particularly cost effective?	The committee concluded that the recommendations in NICE's previous technology appraisal guidance on pirfenidone (TA282) remained appropriate for people with idiopathic pulmonary fibrosis with an FVC between 50% and 80% predicted.	4.21
What are the key drivers of cost	• The estimate of long-term mortality benefit with pirfenidone over a patient's lifetime by extrapolating from relatively short trials	4.15 to 4.17

effectiveness?	• Whether or not pirfenidone is stopped after disease progression (the 'stopping rule').	
Most likely cost-effectiveness estimate (given as an ICER)	In people with an FVC between 80% and 90% predicted, the ICER (including the stopping rule) that most closely matched the committee's preferred assumptions for pirfenidone compared with best supportive care ranged from £32,643 (Weibull) to £38,687 (Gompertz) per QALY gained. In people with an FVC between 50% and 80% predicted, the ICERs without the stopping rule which most closely matched the committee's preferred assumptions for pirfenidone compared with best supportive care were between £35,905 (Weibull) and £40,110 (Gompertz) per QALY gained, but the upper estimate was more plausible. The ICER (including the stopping rule) that most closely matched the committee's preferred assumptions for pirfenidone compared with best supportive care was between £24,933 (Weibull) and £27,780 (Gompertz) per QALY gained.	4.18 to 4.20
How has the new cost-effectiveness evidence that has emerged since the previous appraisal (TA282) influenced the current	The committee concluded that the recommendations in NICE's previous technology appraisal guidance on pirfenidone (TA282) remained appropriate for people with idiopathic pulmonary fibrosis with an FVC between 50% and 80% predicted. The committee concluded that pirfenidone was not	4.19, 4.20

recommendations?	cost effective without the stopping rule in this group.	
Additional factors taken into account		
Patient access schemes (PPRS)	The committee concluded that the PPRS payment mechanism was not relevant in considering the cost effectiveness of the technology in this appraisal.	4.23
End-of-life considerations	Not applicable.	–
Equalities considerations and social value judgements	The committee noted the potential equality issue raised by consultees, that restricting treatment based on percent predicted FVC could discriminate against minority ethnic people, older people and disabled people. The committee discussed these issues with the clinical experts and concluded that its recommendations did not discriminate against any groups of people protected by the Equality Act.	4.22

5 Implementation

5.1 Section 7(6) of the National Institute for Health and Care Excellence (Constitution and Functions) and the Health and Social Care Information Centre (Functions) Regulations 2013 requires clinical commissioning groups, NHS England and, with respect to their public health functions, local authorities to comply with the recommendations in this appraisal within 3 months of its date of publication.

5.2 The Welsh Assembly Minister for Health and Social Services has issued directions to the NHS in Wales on implementing NICE technology appraisal guidance. When a NICE technology appraisal recommends the use of a drug or treatment, or other technology, the NHS in Wales must usually provide funding and resources for it within 3 months of the guidance being published.

5.3 When NICE recommends a treatment 'as an option', the NHS must make sure it is available within the period set out in the paragraphs above. This means that, if a patient has idiopathic pulmonary fibrosis and the doctor responsible for their care thinks that pirfenidone is the right treatment, it should be available for use, in line with NICE's recommendations.

5.4 The Department of Health and Roche Products have agreed that pirfenidone will be available to the NHS with a patient access scheme which makes it available with a discount. The size of the discount is commercial in confidence. It is the responsibility of the company to communicate details of the discount to the relevant NHS organisations. Any enquiries from NHS organisations about the patient access scheme should be directed to [NICE to add details at time of publication]

6 Review of guidance

The guidance on this technology will be considered for review 3 years after publication of the guidance. The guidance executive will decide whether the technology should be reviewed based on information gathered by NICE, and in consultation with consultees and commentators.

Amanda Adler
Chair, appraisal committee
September 2016

7 Appraisal committee members, guideline representatives and NICE project team

Appraisal committee members

The 4 technology appraisal committees are standing advisory committees of NICE. This topic was considered by committee B.

Committee members are asked to declare any interests in the technology to be appraised. If it is considered there is a conflict of interest, the member is excluded from participating further in that appraisal.

The minutes of each appraisal committee meeting, which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

NICE project team

Each technology appraisal is assigned to a team consisting of 1 or more health technology analysts (who act as technical leads for the appraisal), a technical adviser and a project manager.

Sophie Laurenson
Technical Lead

Rosie Lovett, Jasdeep Hayre
Technical Advisers

Jeremy Powell
Project Manager

ISBN: [to be added at publication]

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

APPEAL HEARING

Advice on pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282) [ID837]

Decision of the panel

Introduction

1. An Appeal Panel was convened on 2 December 2016 to consider an appeal against NICE's final appraisal determination on advice to the NHS on pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282).

2. The Appeal Panel consisted of:

• Dr Jonathan Fear	Chair
• Dr Mark Chakravarty	Industry representative
• Dr Biba Stanton	NHS representative
• Patrick Storrie	Lay representative
• Tim Irish	Non-executive director, NICE

3. Tim Irish declared that he had previously held shares in a company with an interest in lung imaging. Since April 2015 these investments are managed through a blind trust. This was a personal financial non-specific interest and as such does not preclude involvement in the appeals panel. None of the other appeal panel members had conflicts of interest to declare.

4. The Panel considered an appeal submitted by Roche Products Ltd.

5. Roche Products Ltd was represented by:

• Denzyl Cain	Head of Health Economics & Strategic Pricing
• Kevin Jameson	Group Health Economics Manager
• Dr James Mawby	Country Medical Lead – Rare Diseases
• Victoria Wakefield	Counsel, Brick Court Chambers
• Sarah Ellson	Partner, Fieldfisher LLP

6. In addition, the following individuals involved in the appraisal were present and available to answer questions from the Appeal Panel:

• Dr Amanda Adler	Chair, Technology Appraisal Committee B
• Meindert Boysen	Programme Director - CHTE
• Melinda Goodall	Associate Director - Committee B

• Professor John Cairns	Technology Appraisal Committee B representative
• Dr Nicky Welton	Technology Appraisal Committee B representative
• Sophie Cooper	Technical Analyst - CHTE

7. NICE's legal adviser Stephen Hocking (DACBeachcroft LLP) was also present.

8. Under NICE's appeal procedures members of the public are admitted to appeal hearings and several members of the public were present at this appeal.

9. There are two grounds under which an appeal can be lodged:

   Ground 1: In making the assessment that preceded the recommendation, NICE has:
   a) Failed to act fairly
   b) Exceeded its powers.

   Ground 2: The recommendation is unreasonable in the light of the evidence submitted to NICE.

10. The Vice Chair of NICE (Mr Andy McKeon) in preliminary correspondence had confirmed that:

    - Roche had potentially valid grounds of appeal as follows:
      - Ground 1a – NICE has failed to act fairly.
      - Ground 2 –the recommendation is unreasonable in the light of the evidence submitted to NICE.

11. Idiopathic pulmonary fibrosis is a chronic, progressive lung disease in which scarring (fibrosis) occurs. The cause is unknown, but it is thought to be related to an abnormal immune response. Symptoms may include breathlessness and cough. Over time, people can experience a decline in lung function, reduced quality of life, and death. The median survival for people with idiopathic pulmonary fibrosis in the UK from the time of diagnosis is approximately 3 years. People with mild-to-moderate disease live longer than people with severe disease.

12. The appraisal that is the subject of the current appeal provided advice to the NHS on the use of pirfenidone within its marketing authorisation for treating idiopathic pulmonary fibrosis.

13. Before the Appeal Panel inquired into the detailed complaints the following made a preliminary statement: Victoria Wakefield on behalf of Roche and Dr Amanda Adler on behalf of the Appraisal Committee.

14. Victoria Wakefield for Roche stated that Roche's arguments essentially all related to the same overarching error by the Committee: there was a failure to consider the totality of the data. More specifically:

    a. The Appraisal Committee had failed to answer the relevant question for the appraisal as set out in the scope by focusing on the subgroup of patients with Forced Vital Capacity (FVC) 80-90% predicted rather than all patients within the marketing authorisation.
    b. In defining this subgroup, the Appraisal Committee had arbitrarily and unjustifiably drawn a line through the whole population that did not have a scientific basis but was defined solely by NICE's own previous guidance.
    c. There was no statistically robust data to support the consideration of this subgroup.
    d. This approach to defining a subgroup was in breach of the Institute's policies, in particular, paragraph 5.10 of the Guide to the methods of technology appraisal 2013 (the Methods Guide).

15. Dr Amanda Adler for NICE, explained that:

    a. The scope of the appraisal was to examine the use of pirfenidone within its marketing authorisation for mild to moderate pulmonary fibrosis as well as considering subgroups by disease severity defined by FVC (such as above or below 80%). She stated that the Appraisal Committee had looked at the totality of data for patients within the marketing authorisation but had decided that it was more appropriate to consider subgroups separately because looking at the whole group might mask important differences in cost-effectiveness within the group. An average value for cost effectiveness was not the right approach in this case.
    b. The Appraisal Committee felt that examining the 80-90% subgroup was relevant because current practice is to treat patients with FVC 50-80% predicted. A previous committee had found it reasonable to use an 80% FVC threshold. The appellant itself had used an 80% threshold. The 80% value was arbitrary, but reflected current NHS practice. 90% was chosen because there was very little data for patients with FVC> 90% (and the new data included no one with FVC>90%).
    c. The Committee chose not to reverse the recommendation for the 50-80% group - although there was uncertainty surrounding it - because it felt that this would be unfair. The Committee therefore chose to concentrate on a subgroup to whom treatment could potentially be extended.
    d. The model had not assumed any difference in the clinical effectiveness of pirfenidone between the subgroups. However, it was judged that the greater cost of treating patients with milder disease could result in important differences in cost-effectiveness between these subgroups.
    e. The methods guide encourages the consideration of subgroups if there are potential differences in clinical or cost-effectiveness between them.

Appeal Ground 1a: In making the assessment that preceded the recommendation, NICE has failed to act fairly: *In failing to consider the totality of data in respect of 'adults with mild to moderate idiopathic fibrosis' (which is both the full licensed indication and the relevant population as identified in the final scope), and in particular determining that 'the subgroup with a FVC between 80% and 90% predicted was the relevant population for decision making (para 4.5 FAD) the Committee acted contrary to policy and procedures (in particular paragraphs 3.2.2, 5.1.4, 5.10 and 6.2.18 of the Methods guide) with inadequate reasons and unfairly.*

16. The appellant raised its concerns about this aspect of the appraisal under both ground 1(a) and ground 2.1. As there was no convenient way of disentangling the issues under each ground, the appeal hearing considered them together, and this letter will do likewise.

17. The Panel questioned the Committee as to what it considered the relevant population to be. Dr Amanda Adler replied that they took the population as defined in the scope, but noted that the scope said they were to look at subgroups if the data permitted. They then chose not to withdraw treatment from the existing treated population, (50-80% FVC) and published an Appraisal Consultation Document concluding that pirfenidone was not a good use of NHS resources in the mild patient population (>80%). The core question was whether pirfenidone was a good use of resources when compared to best supportive care in that population.

18. Dr Amanda Adler confirmed that the Committee had reconsidered the prior decision in the 50-80% FVC subgroup.

19. The Committee were asked about their view on the clinical effectiveness of pirfenidone in the different subgroups, with reference to their statement in the FAD (4.10) both that it was "not clear whether pirfenidone was more, less or equally effective" and that "the committee agreed that...it was more likely to be less effective". Sophie Cooper stated that the economic model used took the same hazard ratios for the 50-80% subgroup and for >80%: the same clinical effectiveness was assumed in all subgroups. Dr Amanda Adler stated she did not feel there was an inconsistency. Professor Cairns said there was uncertainty in all subgroups, but that for the >80% group it was more likely that treatment was less effective.

20. The Appraisal Committee was asked by the Panel whether it was the case that there was no substantial difference between the ICERs for the whole population (50-90% predicted FVC) versus the 50-80% predicted FVC subgroup. Dr Amanda Adler replied that the data for the FVC 80-90% group showed ICERs higher than the NICE threshold.

21. Professor John Cairns was asked by the Panel whether with ICERs for the whole population within, but at the top of, the NICE threshold (and with considerable uncertainty) the Appraisal Committee had felt it was within their remit to look for groups which might be less cost-effective. He replied that this was correct. He added that the Committee were aware that in the NICE appraisal of nintedanib, nintedanib had shown extended dominance over pirfenidone, although the

Committee appreciated that different manufacturers took different approaches to economic modelling.

22. Dr Amanda Adler added that it was not unusual to look for subgroups even if treatment in the whole population appeared cost effective. If the Committee had been starting anew the approach might have been different, but existing guidance recommended treatment for the 50-80% group. The Committee felt steered to look at mild disease. She felt that for an extension to guidance you would look at the extension group separately.

23. Professor John Cairns was asked to consider the modelling results that appeared to show that the cost effectiveness for the 50-80% group was essentially the same as for the 50-90% group. He explained that that was a standard result: the larger group masks the cost ineffectiveness of a smaller subgroup. The Committee believed that use would be cost-ineffective in the 80-90% group. Dr Amanda Adler added that the Committee knew that all the ICERs were likely to be underestimates because of its concern about the modelling of the stopping rule.

24. Meindert Boysen added that if the Committee knew that the ICERs for 50-80% and 50-90% were essentially the same, and that the ICER for the >80% group was above £30,000, then it was reasonable to recommend treating only patients with moderate disease. If those analyses did not exist would the Committee have called for them? - he felt probably not. If treatment was cost effective overall it would not be reasonable to "dig" for subgroups in which a treatment was not cost effective.

25. Sophie Cooper explained that the reason that cost effectiveness differed between the moderate and mild populations, despite an assumed equivalent clinical effect, was that the groups differed in how long they remained progression free, how long they were on treatment, and so on. The mild group received more benefit but incurred more cost and so received less favourable ICERs. Dr Amanda Adler added that there were far more data in the 50-80% group than the 80-90% group, and so it was more possible to be certain of the results for the 50-80% group.

26. The Appeal Panel appreciates that the Committee had an unusual task before it. It is also very wary of being seen to lay down general rules. That is not its role, and it cannot anticipate every twist and turn of a future appraisal. However, it has concluded that the Committee's approach was erroneous on the facts of this case, and it must give its reasons to guide the Committee on any reconsideration there may be.

27. In every appraisal, the starting point to define the question put to the Committee is the same: the scope. In this case the scope was in conventional form. The Committee were charged to " *appraise the clinical and cost effectiveness of pirfenidone within its marketing authorisation*". Only under "other considerations" did the scope record: " *If evidence allows, subgroup analysis by disease severity, defined by FVC (such as above and below or (sic) 80% FVC) and/or diffusing capacity for carbon monoxide, will be considered*".

28. Unless a scope specifies otherwise, the Appeal Panel considers that there is a soft presumption that the starting point for any Committee should be consideration of the whole patient group as one, with a view to making one recommendation for that group. Where different recommendations are to be made for different groups of patients, the reason for departing from one recommendation should be clear and adequate. The Panel does not suggest that this should be a particularly high hurdle to surmount.

29. In this case the Appeal Panel felt that the documentation and the evidence it heard in the hearing did not show that this Committee had adequately considered the possibility of a single recommendation before considering subgroups. The reasons are:

    a. Nowhere in the FAD is there a discussion of the whole patient population position. Although FADs are not to be read as if they are legal documents, and it is possible to supplement the reasoning in a FAD from other Committee papers, on such a central issue some discussion is needed in the FAD if the Panel is to be convinced that adequate consideration was given in Committee.
    b. The reference in the FAD committee meeting slides to a group of patients with FVC of >50% being the result of “combining subgroups" (rather than being the licensed population) suggested subgroups were being taken as the default position, and that it was a departure from that position that would need to be justified.
    c. The reference in the appeal hearing to considering an "extension to guidance" also suggested subgroups were being taken as the default.
    d. The reference in the appeal hearing to be "charged with looking at subgroups" also suggested that the Committee misunderstood its remit and may have considered that a recommendation based on subgroups was a requirement, rather than a question for its judgement.

30. As far as the reasonableness of considering subgroups is concerned, the Panel tended to agree with Meindert Boysen that in a case where it appeared that use of a product was acceptably cost effective in a whole population, it would not normally be reasonable to look for subgroups within that population where use was cost ineffective. However, it would go too far to make that a general rule. Hypothetically if a committee was aware that there existed an identifiable subgroup defined for a proper purpose and in a logical way and in which use was clearly not cost effective, then it might be difficult to say that taking account of that subgroup was unreasonable.

31. In this case the Panel was not yet persuaded that it was reasonable to divide this patient population into subgroups. This is because the modelled difference in cost effectiveness between the group 50-80% and 50-90% was small, and the Committee's conclusions on the 80-90% group were tentative at best. However, the Panel points out that it has already found above that the Committee's reasoning for and consideration of the use of subgroups is unfair, and as a result the Panel may not be fully sighted of the Committee's reasons. While the Panel finds the use of subgroups to be unreasonable on the evidence presented, it

does not rule out that in future a more fully reasoned approach to using these or other subgroups might be reasonable. That would be a matter to be considered at the time.

32. Turning to other considerations under this appeal ground, the Panel was concerned to learn that the Committee had taken account of the results of an economic model presented in another appraisal. The Panel does understand that committees conduct themselves administratively and not judicially. They are expected to bring their wide experience of relevant matters to an appraisal, and there is no need for them to spell out any of this background knowledge. However, they must conduct an appraisal on the evidence before them in that appraisal. If they rely on the economic modelling of the product being appraised that was presented in another appraisal to explain their doubts about the results of the modelling in the appraisal in front of them, then as a bare minimum they must say that this is in their minds before consultation. The Panel would have found that the failure to do so in this case rendered the appraisal unfair of itself.

33. Subject to its findings about the fairness and reasonability of using subgroups at all in this appraisal, the Appeal Panel considered that 80% predicted FVC was an acceptable threshold for defining subgroups that might have different effectiveness or cost-effectiveness. This threshold reflects clinical practice. The Panel was not persuaded that it was significant that that practice might be shaped by past NICE guidance; the Committee is entitled to take clinical practice as it finds it. Further the Panel accepted that the 80% threshold was "arbitrary", in the sense that there was no underlying clinical or scientific rationale for that instead of than some other value, rather, it seemed that if a line had to be drawn, then a line at 80% FVC was "as good as any". The Panel did not feel that such an approach was per se unfair or unreasonable; no doubt many thresholds for many biological markers are in clinical use on much the same basis.

34. The Appeal Panel did not consider it unfair to consider these subgroups on the grounds that there was insufficient data to analyse them with a high degree of certainty because an assumption of no difference in clinical effectiveness for the two subgroups was used in the model. Furthermore, the Panel considers that if a Committee has fairly and reasonably decided that consideration of subgroups is needed, then the fact that data is restricted for one or more subgroups would rarely if ever invalidate the decision to look at subgroups in the first place. Inevitably the data available to analyse a subgroup will be less than the data available to analyse the whole patient population, and this may have an impact on certainty.

35. The Appeal Panel did not agree with the Appellant on the effect of the NICE methods guide on the definition of subgroups. It believes that the intent of the methods guide advice on the definition of subgroups is to avoid data dredging, and the guide's advice on acceptable and unacceptable definitions for subgroups has to be seen in that light. The essential mischief that must be avoided is creating subgroups for the purpose of producing a particular outcome. A subgroup with a plausible biological basis will be one way to achieve this, but there will be others permitted by the guide. In this case, there is nothing suspect about a subgroup using an 80% cut off because the reason for that cut off can be

found in current clinical practice, nor was a 90% cut off suspect as that could be seen to derive from the trial data. The Panel concluded that the methods guide does allow NICE to use a treatment threshold to define clinical subgroups.

36. In conclusion, the Appeal Panel considered the Committee acted unfairly because they did not demonstrate consideration of the whole population as defined in the scope as one population before considering the use of subgroups, and their use of subgroups (on the reasoning currently presented) was unreasonable.

37. The Appeal Panel therefore upheld the appeal on this point under both grounds 1(a) and 2.1

Appeal Ground 1b: In making the assessment that preceded the recommendation, NICE has exceeded its powers.

38. There was no appeal under this ground.

Appeal Ground 2: The recommendation is unreasonable in the light of the evidence submitted to NICE.

Appeal point Ground 2.1: *Failing to consider the totality of data in respect of adults with mild to moderate idiopathic fibrosis (which is both the full licensed population and the relevant population in the Final Scope) and in particular determining that the 'sub group of people with an FVC between 80% and 90% predicted was the relevant population for decision making' (para 4.5 of the FAD), was perverse.*

39. Considered and upheld above.

Appeal point Ground 2.3: *The Committee's assessment of clinical effectiveness was perverse.*

40. Victoria Wakefield for Roche stated that:
    a. there was no robust evidence of a difference in the clinical effectiveness of pirfenidone for patients with FVC 50-80% versus >80% predicted.
    b. the FAD is internally inconsistent in stating both that "it was not clear whether pirfenidone was more, less or equally effective in the group with FVC above 80% predicted" but also that it was "more likely to be less effective" in this group.

41. Dr Amanda Adler for NICE explained that:
    a. the statement that pirfenidone was "more likely to be less effective" in this group was based on a pre-specified analysis from the CAPACITY trials suggesting better outcomes in the placebo group amongst patients with FVC >80% predicted but she acknowledged that this result was not statistically significant.
    b. the Appraisal Committee agreed that there was no robust evidence of a difference in clinical effectiveness between the subgroups.

c. the model had assumed equivalent clinical effectiveness for the subgroups, so any differences in cost-effectiveness were driven by the higher cost of treatment in patients with FVC >80% predicted.
d. numbers were small in the >80% population: a statement that there was no difference in clinical effect between the two groups could be a statement that there was in fact no difference, or that there were insufficient patients to establish a difference. The Committee lacked confidence in a claim that there was an equal effect. The FAD would have been better worded to refer to patient numbers being too small to demonstrate a difference.

42. The Appeal Panel noted above that a FAD must not be read as it if is a legal document. However, it is an important document that records the final recommendation and reasons of a Committee. In this case the FAD presents a confused picture as to what the Committee considered the efficacy of pirfenidone to be in the two subgroups, and its reasons for its conclusions. The Panel concluded that the apparent inconsistency between different statements in the FAD was unreasonable.

43. The Appeal Panel therefore upheld the appeal on this point.

Conclusion and effect of the Appeal Panel's decision

44. The Appeal Panel therefore upholds the appeal on the grounds that NICE has failed to act fairly and that the recommendation is unreasonable in the light of the evidence submitted to NICE.

45. The Appeal Panel suggests that the appraisal is remitted to the appraisal committee who must take all reasonable steps to demonstrate consideration of the effectiveness and cost-effectiveness of pirfenidone in the whole population as set out in the scope. Subgroups defined by predicted FVC could be considered if the treatment is not judged cost-effective in the whole population. The Appraisal Committee's assessment of the clinical effectiveness of pirfenidone in any subgroups should be clearly documented, including any uncertainty in the available evidence.

46. There is no possibility of further appeal against this decision of the Appeal Panel. However, this decision and NICE's decision to issue the final guidance may be challenged by applying to the High Court for permission to apply for a judicial review. Any such application must be made within three months of NICE publishing the final guidance.

2 years – without stopping rule

	PFN			BSC			Incremental			ICER vs. BSC
	Costs	LYs	QALYs	Costs	LYs	QALYs	Costs	LYs	QALYs	
ITT										
Weibull	£64,783	8.71	5.32	£36,711	8.22	5.00	£28,072	0.49	0.31	£90,273
Gompertz	£62,315	7.65	4.91	£34,258	7.18	4.59	£28,057	0.47	0.31	£89,253
Mild										
Weibull	£80,569	11.45	6.55	£43,615	11.02	6.29	£36,953	0.43	0.27	£138,840
Gompertz	£75,661	9.19	5.75	£38,934	8.81	5.49	£36,727	0.38	0.26	£141,482
Moderate										
Weibull	£59,877	7.71	4.84	£34,242	7.23	4.53	£25,635	0.48	0.31	£82,843
Gompertz	£57,834	6.91	4.52	£32,240	6.44	4.20	£25,594	0.47	0.32	£81,032
Truncated										
Weibull	£63,576	8.36	5.16	£35,992	7.89	4.86	£27,583	0.48	0.30	£90,778
Gompertz	£61,176	7.42	4.79	£33,579	6.95	4.48	£27,597	0.46	0.31	£88,621

5 years – without stopping rule

	PFN			BSC			Incremental			ICER vs. BSC
	Costs	LYs	QALYs	Costs	LYs	QALYs	Costs	LYs	QALYs	
ITT										
Weibull	£64,790	8.69	5.30	£34,168	7.30	4.50	£30,622	1.39	0.80	£38,351
Gompertz	£62,359	7.66	4.91	£32,297	6.52	4.21	£30,063	1.14	0.70	£42,960
Mild										
Weibull	£80,121	11.26	6.48	£40,910	9.91	5.73	£39,210	1.35	0.74	£52,794
Gompertz	£75,834	9.20	5.75	£37,373	8.20	5.14	£38,461	1.00	0.61	£62,772
Moderate										
Weibull	£59,946	7.74	4.86	£31,862	6.42	4.08	£28,084	1.33	0.78	£35,902
Gompertz	£57,784	6.90	4.51	£30,234	5.81	3.83	£27,550	1.10	0.69	£40,110
Truncated										
Weibull	£63,549	8.35	5.15	£33,393	7.00	4.37	£30,156	1.35	0.78	£38,529
Gompertz	£61,558	7.43	4.79	£31,898	6.31	4.11	£29,660	1.11	0.69	£43,062

2 years – with stopping rule

	PFN			BSC			Incremental			ICER vs. BSC
	Costs	LYs	QALYs	Costs	LYs	QALYs	Costs	LYs	QALYs	
ITT										
Weibull	£54,595	8.67	5.30	£36,654	8.18	4.99	£17,940	0.49	0.31	£57,568
Gompertz	£52,334	7.63	4.90	£34,245	7.16	4.59	£18,088	0.47	0.31	£57,548
Mild										
Weibull	£65,112	11.36	6.51	£43,522	10.92	6.24	£21,590	0.43	0.27	£80,217
Gompertz	£60,987	9.15	5.73	£38,892	8.77	5.47	£22,095	0.38	0.26	£86,250
Moderate										
Weibull	£51,125	7.71	4.84	£34,109	7.23	4.53	£17,016	0.48	0.31	£54,258
Gompertz	£49,099	6.89	4.50	£32,036	6.42	4.19	£17,063	0.47	0.32	£54,011
Truncated										
Weibull	£53,812	8.36	5.16	£36,147	7.89	4.86	£17,665	0.48	0.31	£57,773
Gompertz	£51,394	7.40	4.78	£33,569	6.94	4.47	£17,825	0.46	0.31	£57,504

5 years – with stopping rule

	PFN			BSC			Incremental			ICER vs. BSC
	Costs	LYs	QALYs	Costs	LYs	QALYs	Costs	LYs	QALYs	
ITT										
Weibull	£54,632	8.67	5.30	£34,141	7.29	4.50	£20,492	1.38	0.80	£25,706
Gompertz	£52,445	7.66	4.92	£32,246	6.52	4.22	£20,199	1.14	0.70	£28,870
Mild										
Weibull	£65,146	11.32	6.50	£40,964	9.97	5.76	£24,183	1.35	0.74	£32,643
Gompertz	£61,009	9.16	5.73	£37,275	8.16	5.11	£23,734	1.00	0.61	£38,687
Moderate										
Weibull	£51,269	7.72	4.84	£31,787	6.39	4.06	£19,483	1.32	0.78	£24,933
Gompertz	£49,102	6.86	4.49	£30,139	5.77	3.81	£18,963	1.09	0.68	£27,780
Truncated										
Weibull	£53,639	8.36	5.16	£33,395	7.01	4.38	£20,244	1.35	0.78	£25,914
Gompertz	£51,626	7.37	4.77	£31,808	6.27	4.09	£19,819	1.10	0.68	£29,036