Controversies in Clinical Trials

Pirfenidone for Idiopathic Pulmonary Fibrosis (IPF)
Controversies to be highlighted by IPF Story

- Post-hoc analyses
- Primary end point selection
  - Changing prespecified endpoints
  - Surrogate endpoints
- Missing Data
- FDA approval/regulation
What is IPF

- Idiopathic pulmonary fibrosis (IPF) is a progressive and fatal disease characterized by scarring of the lungs that thickens the lining of the lungs, causing an irreversible loss of the ability to transport oxygen.
- IPF ultimately robs a patient of the ability to breathe.
- There is no known cause, no FDA approved treatments and no cure for IPF.
IPF Facts

- Median age at time of diagnosis is 63
- IPF affects about 128,000 people in the United States, with about 48,000 new cases diagnosed annually.
- 40,000 people die each year to IPF, the same as to breast cancer.
- 2/3 of IPF patients die within 5 years of diagnosis
- IPF is five times more common than cystic fibrosis and Lou Gehrig’s Disease (or ALS), yet the disease remains virtually unknown to general public and IPF receives a fraction of the research funding.
  - IPF: approx. $18 million per year
  - Cystic Fibrosis: $85 million per year
  - ALS: $48 million per year
IPF Symptoms/signs

- Dry, persistent cough lasting longer than 30 days
- Chronic dyspnea
- Inspiratory crackles on exam
- Restriction and diffusion impairment on PFT
- Hypoxemia
IPF Diagnosis

- In the absence of alternative causes, “classic” HRCT findings are sufficient for diagnosis.
  - Subpleural predominant reticulation favoring the lower lung fields,
  - Paucity of ground glass, and
  - Honeycombing
- If the HRCT is consistent, but not classic, then surgical biopsy showing usual interstitial pneumonia (UIP) is required.
  - Temporally heterogeneous pattern
  - Honeycombing
  - Fibroblastic foci
Treatment of IPF

- No FDA approved treatment
- Until recently, published guidelines suggest treatment with corticosteroids and cytotoxic agents.
  - Based upon opinion only – No supportive data
  - The majority of patients do not respond.
  - We do not recommend therapy unless patient is progressing rapidly.
- Lung transplantation
Clinical Trials

- A Placebo-Controlled Trial of Interferon Gamma-1b in Patients with Idiopathic Pulmonary Fibrosis
  - 330 IPF patients randomly assigned in a 1:1 ratio to receive interferon gamma-1b or placebo subcutaneously three times weekly
  - No difference in primary endpoint (progression free survival):
    - 10% decline in FVC, or
    - 5 mmHg increase in A-a gradient, or
    - Death
  - No difference in measures of lung function, gas exchange, or quality of life.
  - Trend towards benefit in survival (10% vs. 17%, p=0.08)
  - Post-hoc analysis suggested survival benefit in those with mild-moderate disease (4% vs. 12%, p=0.04)
Kaplan–Meier Estimates of Progression Free Survival among Patients with Idiopathic Pulmonary Fibrosis.

**Figure 1.** Kaplan–Meier Estimates of Progression-free Survival among Patients with Idiopathic Pulmonary Fibrosis.

Former InterMune CEO Convicted of Wire Fraud

Company Found to Have Disseminated Misleading Information About Results of Clinical Trial of Its Interferon Product

SAN FRANCISCO, CAL. 9/29/09—After 3 days of deliberations, a jury has convicted W. Scott Harkonen, M.D., the former CEO of InterMune, Inc., of wire fraud under 18 USC §1343 for the creation and dissemination of “false and misleading information” about the efficacy of drug Actimmune (interferon gamma-1b) as a treatment for idiopathic pulmonary fibrosis. The conviction arose from a press release issued by Harkonen on August 28, 2002, announcing the results of a clinical trial supposedly demonstrating that the biologic product enabled patients with IPF to live longer. (The headline read “InterMune Announces Phase III Data Demonstrating Survival Benefit of Actimmune in IPF: Reduces Mortality by 70% in Patients with Mild to Moderate Disease.”)

In fact, the trial did not demonstrate a benefit. Actimmune was not approved for the treatment of IPF, but that indication accounted for most of InterMune’s sales of the product, which can cost approximately $50,000 per year.

Further Troubles for Sequenom

—Investigations, Shareholder Lawsuit Follow Alleged Mishandling of Data on Fetal Test

SAN DIEGO, CAL. 10/6/09—The U.S. Attorney for the Southern District of California and the Nasdaq stock exchange have opened investigations into Sequenom over its mishandling of test data and results of its test for Down’s syndrome and other conditions in fetuses, called SEQureDx. The estimated annual sales of a successful test would be about $2 billion.

The company announced last week that it was firing five employees, including its President and CEO Harry Stylii and its Senior VP of Research Elizabeth Dragon. Two other employees, CFO Paul Hawran and VP for Commercial Development of Prenatal Diagnostics Steve Owings, resigned.

The company said it is cooperating with the probe and has met with the U.S. Attorney and the Federal Bureau of Investigation. It also has held discussions with the Securities and Exchange Commission.

UPDATE: On November 13, some Sequenom shareholders filed suit in U.S. District Court for the Southern District of California over the mishandling of data from SEQureDx. The complaint alleges breach of fiduciary duties by several people and insider trading by Owings. Other defendants are Stylii, Dragon, Hawran, the former VP of Regulatory Affairs, Qual...
Effect of interferon gamma-1b on survival in patients with idiopathic pulmonary fibrosis (INSPIRE): a multicentre, randomised, placebo-controlled trial

- 826 patients with mild to moderate IPF.
  - FVC 55-90%, DLCo 35-90%.
- Randomized 2:1
- Primary endpoint was survival
- Study stopped after 2nd interim analysis failed to show minimum benefit.
- No difference in measures of lung function, gas exchange, or quality of life.
Clinical Trials

- Other IPF trials:
  - Bosentan
    - Build-1: negative study, but post-hoc analysis showed benefit in progression free survival among patients with surgical lung biopsy.
    - Build-3: negative study.
  - Ambrisentan
    - stopped due to lack of efficacy at interim analysis
  - Sildenafil
    - negative study
  - NAC
    - Slowed progression of IPF when added to prednisone plus azathioprine
    - PANTHER –
      - multi-armed study including NAC alone and combined with other therapy ongoing.
      - Imuran + Prednisone arm stopped early due to increased mortality in treatment arm.
  - Warfarin
    - ACE – stopped early – lack of efficacy and increased adverse events.
Pirfenidone

- Experimental animal models of pulmonary fibrosis suggest anti-inflammatory, antioxidant, and antifibrotic effects

- A Japanese phase 2 study suggested benefit
  - Primary outcome was lowest O2 sat during 6MET
  - The study was stopped early given greater risk of IPF exacerbations in the placebo group.
Taniguchi Phase 3 study

- 275 patients randomized 2:1:2
- Primary endpoint: Change in FVC (compared HD to P)
  - Primary endpoint changed midway through trial
    - Initial primary endpoint was lowest O2 saturation during 6 min walk.
  - Missing data imputed by LOCF
- Secondary endpoints:
  - Progression free survival
    - Progression = 10% decline in FVC
  - Change in lowest SpO2 during 6MET
Change in FVC (Primary end-point)

**FIGURE 2.** Effects of pirfenidone on vital capacity (VC) at week 52. Data are presented as mean±SE. *: p<0.1, comparison of adjusted means based on ANCOVA (negative and positive of the changes represent deterioration and improvement from baseline, respectively). The last observation carried forward method was used for drop-outs in each group. Placebo group: n=103; high-dose group: n=104; low-dose group: n=54.

**TABLE 3** Comparison of changes in vital capacity

<table>
<thead>
<tr>
<th></th>
<th>Baseline L</th>
<th>Subjects n</th>
<th>52 weeks L</th>
<th>Subjects n</th>
<th>Comparison of adjusted means based on ANCOVA*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Crude mean ± se</td>
<td></td>
<td>Crude mean ± se</td>
<td></td>
<td>Difference from placebo mean ± se L</td>
</tr>
<tr>
<td>High-dose</td>
<td>2.42±0.61</td>
<td>107</td>
<td>2.36±0.53</td>
<td>67</td>
<td>104</td>
</tr>
<tr>
<td>Low-dose</td>
<td>2.44±0.66</td>
<td>53</td>
<td>2.04±0.71</td>
<td>38</td>
<td>54</td>
</tr>
<tr>
<td>Placebo</td>
<td>2.47±0.70</td>
<td>104</td>
<td>2.42±0.75</td>
<td>72</td>
<td>103</td>
</tr>
</tbody>
</table>

* negative and positive of the changes represent deterioration and improvement from baseline, respectively. Covariates: baseline vital capacity.

**FIGURE 3.** a) Transitional plots of “last observation carried forward imputed means”. b) “Crude means” of the changes in VC. Data are presented as mean±90% CI. ---: high-dose pirfenidone group; ---: low-dose pirfenidone group; --: placebo group.
Secondary and tertiary endpoints

- PFS better in the high dose pirfenidone compared with placebo
  - 11 patients died, 3 high dose, 4 low dose, and 4 placebo
- No difference in lowest SpO2
- No difference in acute exacerbations

**FIGURE 4.** Kaplan–Meier plot of progression-free survival time among idiopathic pulmonary fibrosis patient groups. ——: high dose; ———: low dose; ————: placebo. Symbols on the curve represent the censored points where patients discontinued the study treatment due to causes other than progression of the disease. Kaplan–Meier curves were compared with the log-rank test: $p = 0.0280$ between the high-dose group and placebo group; $p = 0.0655$ between the low-dose group and placebo group; $p = 0.9106$ between the high-dose group and low-dose group.
Taniguchi study

- What are the problems with this study?
- Based on this study, how convinced are you that Pirfenidone is effective for the treatment of IPF?
- Is there any additional data that you would request?
Change in Primary Endpoint

- Original primary end-point was change in lowest oxygen saturation during 6-min steady-state exercise test.
- During the course of this trial, views on appropriate primary end-points in IPF evolved.
- The decision to change end-points involved members of the DSMB who recommended change after a discussion of blinded interim comparative data (i.e. they had knowledge of whether there were significant differences between study groups with respect to the primary and secondary end-points).
- “the credibility and integrity of the trial is compromised. It is simply impossible for readers to assess the impact of this knowledge on the decision.”
Problems with LOCF

- LOCF may be appropriate, but it may not.
- For groups 2 and 3, LOCF artificially deflates or inflates outcome values for time-points after the last study visit.

Swigris and Fairclough et al., ERJ 10/2010
## Mixed model results

### TABLE 1
Mixed model analysis of changes in vital capacity from baseline: test of fixed effects

<table>
<thead>
<tr>
<th>Source</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>2.78</td>
<td>0.0623</td>
</tr>
<tr>
<td>Time (visit)</td>
<td>7.05</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Group x time</td>
<td>0.98</td>
<td>0.4939</td>
</tr>
</tbody>
</table>

*: both-sided p-value.

### TABLE 2
Mixed model analysis of changes in vital capacity from baseline: overall adjusted means of vital capacity changes in treatment groups and comparisons of the means

<table>
<thead>
<tr>
<th>Group</th>
<th>Estimate</th>
<th>se</th>
</tr>
</thead>
<tbody>
<tr>
<td>High dose</td>
<td>-0.04561</td>
<td>0.01653</td>
</tr>
<tr>
<td>Low dose</td>
<td>-0.03542</td>
<td>0.02277</td>
</tr>
<tr>
<td>Placebo</td>
<td>-0.09141</td>
<td>0.01641</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Difference</th>
<th>se</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>High versus Low dose</td>
<td>-0.02434</td>
<td>0.03752</td>
</tr>
<tr>
<td>High dose versus placebo</td>
<td>0.05060</td>
<td>0.03131</td>
</tr>
<tr>
<td>Low dose versus placebo</td>
<td>0.07494</td>
<td>0.03710</td>
</tr>
</tbody>
</table>

*: both-sided p-values.

### TABLE 3
Mixed model analysis of changes in vital capacity from baseline: adjusted means of vital capacity changes at the last visit in treatment groups and comparisons of the means

<table>
<thead>
<tr>
<th>Group</th>
<th>Estimate</th>
<th>se</th>
</tr>
</thead>
<tbody>
<tr>
<td>High dose</td>
<td>-0.09289</td>
<td>0.02248</td>
</tr>
<tr>
<td>Low dose</td>
<td>-0.06855</td>
<td>0.03003</td>
</tr>
<tr>
<td>Placebo</td>
<td>-0.14349</td>
<td>0.02179</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Difference</th>
<th>se</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>High versus Low dose</td>
<td>-0.02434</td>
<td>0.03752</td>
</tr>
<tr>
<td>High dose versus placebo</td>
<td>0.05060</td>
<td>0.03131</td>
</tr>
<tr>
<td>Low dose versus placebo</td>
<td>0.07494</td>
<td>0.03710</td>
</tr>
</tbody>
</table>

*: both-sided p-values.
Capacity studies

- PIPF 004: 435 pts
  - Patients randomized 2:2:1 to receive pirfenidone 2403 mg/d (174 patients), placebo (174 patients), or pirfenidone 1197 mg/d (87 patients)

- PIPF 006: 344 pts
  - Patients randomized 1:1 to receive pirfenidone 2403 mg/d (171 patients) or placebo (173 patients)

- Primary endpoint for each study was change in FVC at week 72 (compared with ranked ANCOVA)
Statistical Analyses

- Primary Endpoint: Change in FVC at 72 weeks
- Primary efficacy analysis: Ranked ANCOVA
- Missing data: Missing values as a result of death were assigned the worst rank in ANCOVA analyses, and worst possible outcome in mean change analyses (eg, FVC=0) and categorical analyses. Other missing data were imputed with the average value from three patients with the smallest sum of squared differences at each visit with data that were not missing.
Additional Analyses of the Primary Endpoint

A repeated measures analysis was prespecified to interrogate FVC across all study time points for purposes of inference and estimation. In the repeated measures model, in the case of patients who died, the first missing data value after time of death was replaced with a percent predicted FVC of 30. The FVC of 30 imputation, as opposed to zero as used in other analyses, was intended to preserve the normal distribution of the data while also assigning a “worst” outcome to death. Other missing values were not imputed. The mixed model included fixed effects for treatment, geographical region, assessment week (as a factor variable, not as a linear regression), treatment by week interaction, and Baseline percent predicted FVC as a covariate. Study was included in the model for the analysis of pooled data.

The protocol specified a sensitivity analysis for the primary outcome variable using lastobservation-carried-forward (LOCF) methodology for missing data. The LOCF analysis used a rank ANCOVA model for change from Baseline to Week 72 of percent predicted FVC. In this analysis, data missing due to death were ranked last by order of study day of death, but missing data for reasons other than death were imputed by the last observed value for that patient.
**Change in FVC**

![Graphs showing mean change from baseline in percentage predicted FVC in study 004 (A), study 006 (B), and the pooled population (C).](image)

**Figure 2:** Mean change from baseline in percentage predicted FVC in study 004 (A), study 006 (B), and the pooled population (C). FVC=forced vital capacity. *Pirfenidone 2403 mg/day versus placebo. †Rank ANCOVA (pirfenidone 2403 mg/day vs placebo). 95% CIs were only calculated for absolute differences for the week 72 timepoint in study 004 (0.7 to 9.1) and study 006 (-3.5 to 4.7).
## Secondary Endpoints

<table>
<thead>
<tr>
<th></th>
<th>Study 004</th>
<th></th>
<th>Study 006</th>
<th></th>
<th>Pooled data</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Absolute difference (95% CI)</td>
<td></td>
<td>Absolute difference (95% CI)</td>
<td></td>
<td>Absolute difference (95% CI)</td>
</tr>
<tr>
<td>Categorical change in FVC±10%</td>
<td>35 (20%)</td>
<td>14.4 (7.4 to 21.3)</td>
<td>39 (23%)</td>
<td>3.8 (2.7 to 10.2)</td>
<td>74 (21%)</td>
<td>106 (31%)</td>
</tr>
<tr>
<td>Progression-free survival time†</td>
<td>...</td>
<td>0.64 (0.44 to 0.95)</td>
<td>...</td>
<td>0.84 (0.58 to 1.22)</td>
<td>...</td>
<td>0.74 (0.57 to 0.96)</td>
</tr>
<tr>
<td>Mean change in 6MWT distance (m)</td>
<td>−60.4</td>
<td>−76.8</td>
<td>−45.1</td>
<td>−76.9</td>
<td>−52.8</td>
<td>−76.8</td>
</tr>
<tr>
<td>Mean change in DLco (% predicted)</td>
<td>−7.9</td>
<td>−9.9</td>
<td>−9.8</td>
<td>−9.2</td>
<td>−8.8</td>
<td>−9.6</td>
</tr>
<tr>
<td>Mean change in dyspnoea score¶</td>
<td>12.1</td>
<td>15.2</td>
<td>11.9</td>
<td>13.9</td>
<td>12.0</td>
<td>14.5</td>
</tr>
<tr>
<td>Mean change in worst SpO₂ during 6MWT (%)</td>
<td>−1.5</td>
<td>−2.3</td>
<td>−1.9</td>
<td>−1.3</td>
<td>−1.7</td>
<td>−1.8</td>
</tr>
<tr>
<td>Time to worsening in idiopathic pulmonary fibrosis</td>
<td>...</td>
<td>0.84 (0.50 to 1.42)†</td>
<td>...</td>
<td>0.73 (0.43 to 1.24)†</td>
<td>...</td>
<td>0.78 (0.54 to 1.14)†</td>
</tr>
<tr>
<td>Categorical change in HRCT-diagnosed fibrosis</td>
<td></td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

FVC = forced vital capacity. 6MWT = 6-minute walk test. DLco = haemoglobin-corrected carbon monoxide diffusing capacity. SpO₂ = peripheral oxygen saturation. HRCT = high-resolution CT. NA = not applicable.

* Rank ANCOVA (pipériférine 2403 mg/day vs placebo), unless otherwise indicated. † Cochran-Mantel-Haenszel row mean score test (pipériférine 2403 mg/day vs placebo) based on five categories (severe decline, ≥20%; moderate decline, <20% but ≥10%; mild decline, <10% but ≥0; mild improvement, >0 but <10%; and moderate improvement, ≥10%). ‡ Hazard ratio (95% CI) based on the Cox proportional hazard model with geographic region (USA vs non-USA) as a stratum. § Log-rank test (pipériférine 2403 mg/day vs placebo). ¶ Based on the University of California San Diego Shortness of Breath Questionnaire: total score ranges from 0 to 120, with larger scores indicating greater shortness of breath. || Cochran-Mantel-Haenszel row mean score test (pipériférine 2403 mg/day vs placebo) based on five categories (much better, better, same, worse, or much worse); assessed in study 006 only.

Table 2: Secondary efficacy endpoints at week 72
Progression Free Survival

Figure 3: Kaplan-Meier distribution of progression-free survival time in study 004 (A), study 006 (B), and the pooled population (C).

*Pirfenidone 2403 mg/day versus placebo.
## Mortality – Exploratory endpoint

<table>
<thead>
<tr>
<th></th>
<th>Pirfenidone 2403 mg/day (n=345)</th>
<th>Placebo (n=347)</th>
<th>Hazard ratio* (95% CI)</th>
<th>p value†</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overall</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All-cause mortality</td>
<td>27 (8%)</td>
<td>34 (10%)</td>
<td>0.77 (0.47–1.28)</td>
<td>0.315</td>
</tr>
<tr>
<td>Idiopathic-pulmonary-fibrosis-related mortality‡</td>
<td>18 (5%)</td>
<td>28 (8%)</td>
<td>0.62 (0.35–1.13)</td>
<td>0.117</td>
</tr>
<tr>
<td><strong>On-treatment§</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All-cause mortality</td>
<td>19 (6%)</td>
<td>29 (8%)</td>
<td>0.65 (0.36–1.16)</td>
<td>0.141</td>
</tr>
<tr>
<td>Idiopathic-pulmonary-fibrosis-related mortality‡</td>
<td>12 (3%)</td>
<td>25 (7%)</td>
<td>0.48 (0.24–0.95)</td>
<td>0.030</td>
</tr>
</tbody>
</table>

Data are number (%). *Based on the Cox-proportional hazard model. †Log-rank test (pirfenidone 2403 mg/day vs placebo). ‡Assessed by the investigator, who remained masked to treatment assignment. §Defined as the time from randomisation until 28 days after the last dose of study drug.

Table 3: All-cause and idiopathic-pulmonary-fibrosis-related mortality in the pooled population
What would you do?

- Moc FDA Vote
What has happened

- FDA advisory panel recommended approval
- FDA decision was to not approve Pirfenidone as of now
  - Requested another clinical trial to be the tie breaker.
  - Unsatisfied with FVC as surrogate endpoint
  - May require survival primary end-point