A Controlled Trial of Sildenafil in Advanced Idiopathic Pulmonary Fibrosis

By Judy Mai and Mei Win

Introduction

The STEP-IPF (Sildenafil Trial of Exercise Performance in Idiopathic Pulmonary Fibrosis) was a multi-center, randomized, double-blind, placebo-controlled 12-week trial of oral sildenafil (20 mg orally three times daily) followed by a 12-week open-label period. The trial was designed to test if treatment with sildenafil may improve the exercise capacity and quality of life in patients with advanced idiopathic pulmonary fibrosis defined by a DLco (diffusing capacity of the lung for carbon monoxide) < 35% of the predicted value. The trial had two treatment periods in a permutated block design (each center was considered a block with the subjects randomly allocated in a 1:1 ratio into one of two groups).

The first period consisted of a 12-week double-blind comparison between sildenafil and an identical placebo control to determine the presence or absence of a 20% or greater improvement in the 6-minute walk distance (6MWD). Key secondary outcomes included changes in the 6MWD, degree of dyspnea and quality of life. The second period consisted of a 12-week open-label evaluation involving all patients receiving sildenafil to further determine the short-term efficacy (12-week) with the placebo group and longer-term safety profile (24-week) for the sildenafil group.

The study was sponsored by the Idiopathic Pulmonary Fibrosis Clinical Research Network (IPFnet) of the National Heart, Lung and Blood Institute. Pfizer donated sildenafil and its identical placebo but was not involved in the study design or data analysis.

Element	Criteria	Comments
Study Design Assessment	**Is the design appropriate to the research question? Is the research question useful?** • For **efficacy**, use of **experimental study design** (meaning study subjects and others were not allowed choice in determining interventions) • **Clinically significant area** for study (morbidity, mortality, symptom relief, functioning and health-related quality of life) and reasonable **definitions for clinical outcome such as response, treatment success or failure** • If **composite endpoints** used, reasonable combination used – and used for safety if used for efficacy	STEP-IPF was a randomized, double-blind placebo controlled trial utilizing a permutated block design and stratification in each treatment center. Primary outcome tested for a 20% or greater improvement in the 6-minute walk test (6MWT), in accordance with the 2005 American Thoracic Society guidelines. **THREAT:** 6MWT is an accepted measure of drug efficacy in studies of pulmonary arterial hypertension but less is known about its use as a measure in idiopathic pulmonary fibrosis. Therefore, this test may not be the most appropriate evaluation to measure for treatment success. **THREAT:** Although a double-blinded trial is considered the gold standard, it is challenging to blind patients who receive sildenafil due to possible side effects including headache, dyspepsia, flushing and diarrhea.
Internal Validity Assessment	**Can bias, confounding or chance explain the study results?** • Ensure prespecified and **appropriate** 1) research questions, 2) populations to analyze, 3) outcomes, 4) group assignment methods, 5) study conduct methods, 6) analysis methods, and 7) level for statistical significance	Comparing between the two arms, baseline characteristics were moderately different with the sildenafil arm having comparatively less healthy patients. These patients had a higher score in all 3 tests: Borg Dyspnea Index, Shortness of Breath Questionnaire and St. George's Respiratory Questionnaire. **THREAT:** By having less healthy patients in the sildenafil, as defined by the three test values, any improvement that occurred could be deemed more significant since the patients started with a lower baseline.
Selection Bias	• Groups are **appropriate** for study, of appropriate size, **concurrent** and **similar** in **prognostic variables** • Methods for generating the group assignment sequence are truly **random,** sequencing avoids potential for anyone **affecting assignment** to a study arm and **randomization remains intact** • **Concealment of allocation** strategies are employed to prevent anyone affecting assignment to a study arm	Eligible patients, with roughly similar baseline characteristics, were randomized in a 1:1 ratio to receive sildenafil or placebo with the use of a permutated block design with stratification according to the clinical center (each center is considered a block with the subjects randomly allocated into two equal size groups). Subjects were randomized to receive one of the two treatment regimens with equal probability, via telephone contact with a central interactive voice response system (IVRS), using a toll-free randomization number. Each subject was assigned an anonymous identification number to be used throughout the study.

Element	Criteria	Comments
Performance Bias	• **Double-blinding** methods employed (i.e., subject and all working with the subject or subject's data) and achieved • Reasonable **intervention** and reasonable **comparator** used (e.g., placebo) • **No bias or difference, except for what is under study, between groups during course of study** (e.g., intervention design and execution, co-interventions, concomitant medication use, adherence, inappropriate exposure or migration, cross-over threats, protocol deviations, measurement methods, study duration, etc.)	There is no available evidence of a dose–response relationship associated with the primary endpoint (exercise capacity) or with tolerability when using different doses of sildenafil. The reason for this is unclear but may be related to the complete inhibition of phosphodiesterase type 5 with the lowest dose. As a result, the study chose to use sildenafil 20 mg orally three times daily as opposed to higher doses of 25, 40 or 50 mg three times daily. Pfizer supplied sildenafil and its identical placebo. Medication adherence was evaluated via calculation of the pill count. Both arms reported a similar rate of compliance. All 6MWTs were performed by study personnel who were not directly involved in study coordination such that they were not able to identify which treatment arm a patient is in. **THREAT:** All patients received an initial dose of either sildenafil or placebo and subsequently were monitored for 60 minutes for adverse events. Some patients may recognize the adverse effects associated with the drug. As a result, one may recognize subsequently if he or she was on the placebo or active medication leading to an incomplete blinding for this study. **THREAT:** All subjects took part in a second 12-week open-label trial. The endpoint was compared from baseline to a 12-week treatment period whether the patient was taking sildenafil during the first study period or the second study period. The original group on sildenafil was blinded while data was being collected.The 12-week active sildenafil data obtained from the original placebo group was during a period when the study was open label. Although comparisons for the efficacy of sildenafil only included a 12 week period of active medication use, one group was blinded and another group was not when this data was obtained. Therefore, despite the result showing that there was a small statistically significant difference favoring the use of sildenafil in some secondary outcomes—degree of dyspnea and quality of life—it is difficult to determine if the result is solely due to the use of sildenafil. **Threat:** Any acute exacerbations will be treated at the discretion of the attending physician. The study medication will be held until the patient until the patient is able to continue. Data is not reported on how many patients the physicians had to additionally treat. Therefore, study results may favor the intervention.
Attrition Bias	• Zero or minimal **missing data points** or loss from randomization (e.g., approximately 5% with differential loss, or approximately 10% without differential loss) unless good ITT analysis (see ITT below)	**THREAT:** 16 of 89 subjects (18%) in the sildenafil arm and 21 of 91 subjects (23%) in the placebo arm did not continue in the study. A total loss rate of 20.55% between the two arms was due to the following factors: death, adverse events, lung transplants and loss to follow-up.
Assessment Bias	• Assessors are **blinded** • Low likelihood of findings due to **chance, false positive and false negative outcomes** (judgment call on statistical significance, including **confidence intervals**) • **Non-significant findings** are reported, but the **confidence intervals include clinically meaningful differences** • **Intention-to-Treat Analysis (ITT)** performed (all people are analyzed as randomized + reasonable method for imputing missing values which puts the intervention through a challenging trial or reasonable sensitivity analysis) • Use of **modeling** only with use of reasonable assumptions	In the intention-to-treat analysis (ITT), patients were deemed to have had no response if the rate of improvement was less than 20% at week 12 or if they died, withdrew from the study or were unable to complete the walk test for any reason at 12 weeks. Most results were not statistically significant (p-value>0.05) based on a 95% confidence interval. **THREAT:** There was no data regarding the blinding of the assessors who performed the statistical assessments for this study. **THREAT:** Utilization of a last-observation-carried-forward (LOCF) form of the ITT analysis may be biased and favor the intervention since there is no data regarding when the patient dropped out of the study. **THREAT:** Since the results are dichotomous (improvement >20% or no improvement), it doesn't factor in the consideration the magnitude of the improvement.
Usefulness Assessment	• **Clinically significant area + sufficient benefit size =** meaningful clinical benefit (consider efficacy vs effectiveness)	A > 20% improvement in 6MWD is expected to be a fairly rare event in an untreated population with advanced IPF. Over the 12 weeks of treatment, it is expected that fewer than 10% of placebo-treated subjects will have a clinically meaningful improvement in 6MWD. A response rate of 30% or more in the sildenafil-treated group is viewed as a clinically meaningful treatment effect. **THREAT:** Even with a "clinically meaningful treatment effect," 70% of patients taking sildenafil do not respond to the treatment. This raises the question of whether the benefit is sufficient to warrant treating a patient with a medication that will be ineffective during a majority of the time.

Element	Criteria	Comments
External Validity	**How likely are research results to be realized in the real world considering population and circumstances for care?** • Review n, inclusions, exclusions, baseline characteristics and intervention methods – this is a **judgment call**.	Idiopathic pulmonary fibrosis (IPF) affects more men than women and most commonly occurs between the fifth and seventh decades. It does not seem to favor one race or ethnicity over another. The median age at time of diagnosis is about 63 years old. Most recent estimates indicate that approximately 200,000 Americans have IPF. **THREAT:** The study consisted of mostly men (83%) and people who were caucasian (91%) with the average age around 69 years old, with a carbon monoxide diffusion capacity of less than 35%. The results might not be applicable for patients with milder physiological impairment. **THREAT:** Unlike the American Thoracic Society/European Respiratory Society consensus criteria, bronchoscopy is not required for diagnosis of IPF in this study. Therefore, the inclusion of patients in this study may vary from patients diagnosed outside of this study.
Patient Perspective	• Consider benefits, harms, risks, costs, uncertainties, alternatives, applicability to which patients, adherence issues, potential for abuse, dependency issues and patient satisfaction	**Advantages:** • There is currently no FDA approved treatment available for IPF so sildenafil can serve as a possible option to improve a patient's quality of life. **Disadvantages:** • There is currently no generic available. • Three times daily dosing may decrease adherence. • There is a possibility of hypotension and sexual side effects.
Provider Perspective	• Satisfaction, acceptability, likely appropriate application and actionability (e.g., FDA approval, affordability, external relevance, circumstances of care, able to apply, tools available)	**Advantages:** • There is currently no FDA approved treatment for IPF so sildenafil offers the provider an alternative treatment option. **Disadvantages:** • Sildenafil is not FDA approved for the treatment of idiopathic pulmonary fibrosis. • Sildenafil did not provide a significant improvement in the primary outcome test. Improvement in the 6-minute walk distance of at least 20% over baseline occurred in 9 of 80 subjects (10%) in the sildenafil groups and 8 of 91 (7%) in the placebo group (P = 0.39).

**Chart taken from the Delfini Group,LLC. Short Critical Appraisal Checklist: Updated 02/19/08*

Use of this tool implies agreement to the legal terms and conditions at www.delfini.org. © Delfini Group, LLC, 2006-2008. All Rights Reserved World Wide.

Author's Results and Conclusions

The use of sildenafil did not result in a statically significant difference (p=0.39) in the proportion of patients with an improvement of 20% or more in the 6MWD at week 12 (the primary outcome). At week 12, there was no significant difference between the two groups in mortality or rate of acute exacerbations of advanced idiopathic pulmonary fibrosis. There was also no difference in other adverse events between the two treatment arms. Although the study did not show that patients who were on sildenafil vs. placebo had an improved 6MWD, it reported that sildenafil was associated with a small but symptomatic improvement in arterial oxygenation, carbon monoxide diffusion capacity, degree of dyspnea and quality of life. These results may be of value for further study.

Reviewer's Analysis

There were several limitations to the STEP-IPF trial that threatens the validity of the study including:

- 6MWT is an accepted measure of drug efficacy in studies of pulmonary arterial hypertension but less is known about its use as a measure in IPF.
- Incomplete blinding of the sildenafil group compared to the control group due to adverse effects of the drug may compromise subjective results such as quality of life.
- A small effect size and a short study length (180 patients participated in the study for a period of 12 weeks) make it difficult to assess the efficacy and safety profile of sildenafil.

Data from this study may be used in clinical decision-making, however additional clinical data regarding sildenafil in treating IPF is necessary to confirm the efficacy of the drug. While such trials are being designed, sildenafil may be of value to certain patients since currently there is no FDA approved treatment for IPF. Depending on the patient's profile, sildenafil may be able to increase one's quality of life by providing symptomatic improvements.

Overall Grade: B-U=Possible to uncertain usefulness.

Reviewers: Craig Stern, Michael Stuart, Sheri Strite, Judy Mai, Mei Wu

References:

The Idiopathic Pulmonary Fibrosis Clinical Research Network. A Controlled Trial of Sildenafil in Advanced Idiopathic Pulmonary

Delfini Evidence Grading Scale

Grade A Evidence: Useful

The evidence appears strong and sufficient to use in making health care decisions - no significant threats to validity were ascertained

Grade B Evidence: Possibly Useful

The evidence appears potentially strong and is probably sufficient to use in making health care decisions - some threats to validity were identified

Grade B-U Evidence: Possible to uncertain usefulness

The evidence might be sufficient to use in making health care decisions; however, there remains sufficient uncertainty that the evidence cannot fully reach a Grade B and the uncertainty is not great enough to fully warrant a Grade U. Health care decision-makers should be fully informed of the evidence quality.

Grade U Evidence: Uncertain

There is sufficient uncertainty that caution is urged regarding its use in making health care decisions. Delfini does not use such information to inform clinical decisions regarding efficacy.

Fibrosis. N Engl J Med 2010. DOI: 10.1056/NEJMoa1002110.

Supplement to: The Idiopathic Pulmonary Fibrosis Clinical Research Network. A Controlled Trial of Sildenafil in Advanced Idiopathic Pulmonary Fibrosis. N Engl J Med 2010. DOI: 10.1056/NEJMoa1002110.

Additional Comments: By Michael E. Stuart, MD

There is an important "take home" message we can learn from the STEP-IPF trial. Readers should be extremely cautious about accepting the conclusions reported in studies of borderline quality. The quality of the STEP-IPF trial is borderline. By that I mean that there is enough uncertainty about the reliability of the results of this trial that caution should be exercised when using these results to inform decisions. In other words, there are threats to validity.

Threats to validity (the likelihood of true results) in the STEP-IPF trial include the likely possibility that the subjects in the two groups differed in important prognostic variables as seen in the table of baseline characteristic. When study groups are not sufficiently similar at the start of a trial the differences in study outcomes may be explained by the clinical or demographic differences between the study groups and not the intervention (selection bias). A second threat to validity is the possibility of the subjects knowing which agent they were receiving (performance bias). This is likely to have occurred because of the noticeable side effects of sildenafil such as flushing. When patients or clinicians know to which group a patient has been assigned there may be numerous differences in how those patients are treated. For example, clinicians may behave in a more supportive way to patients who are receiving the active agent because clinicians are always eager to use new, effective agents. Third, data was lost for a substantial number of patients (attrition bias). This may have an important effect on results because the patients remaining in the study may no longer sufficiently similar and the differences may be the explanation for the outcomes. Fourth, we are not given enough information about blinding of the assessors to know they were truly blinded (assessment bias). This is problematic because investigators, like clinicians are likely to be "rooting" for the intervention.

This trial illustrates problems in all four phases of a clinical trial—the formation of the study groups, the conduct of the trial so that everything is the same except the intervention, loss of data and assessment of results. Threats to validity often amplify the differences between study groups. Therefore, in this case, the true amount of benefit may have been substantially less than that reported. The bottom line is that readers need to be very cautious about accepting results from studies of borderline quality and may be well-advised to wait for confirmatory trials of high quality before using published results to inform clinical and organizational decisions.