Synthetic Arm Generation Utilizing Real-World Patient Data Demonstrates Treatment Effect in the Phase 2a AIR trial of Buloxibutid in IPF

Kiril R. Kirov[1], Muhunthan Thillai[1,2], Toby M. Maher[3,4], Bernt van den Blink[5], Bertil Lindmark[5], Simon L.F. Walsh[1,4]

[1]Qureight Ltd, Cambridge UK, [2]Royal Papworth Hospital, Cambridge UK, [3]University of Southern California, California USA, [4]National Heart and Lung Institute, Imperial College UK, [5]Vicore Pharma, Stockholm, Sweden

Introduction

Idiopathic pulmonary fibrosis (IPF) is a form of interstitial lung disease with a fatal outcome. **Buloxibutid** is a first-in-class angiotensin II type 2 (AT2) receptor antagonist designed to promote alveolar repair and maintenance, thereby reducing fibrosis formation. The AIR trial (ClinicalTrials.gov ID NCT04533022) was a multicentre, open-label, single-arm phase 2a trial investigating the safety, efficacy and pharmacokinetics of Buloxibutid (C21) in patients with IPF. This work aimed to perform synthetic arm generation (SAG) for efficacy testing in the AIR trial.

Methods

The AIR trial enrolled 52 patients at 21 sites in India, Russia, the Ukraine, and the UK (Figure 1.a). Participants were treated with 100mg BID (twice daily) of orally administered Buloxibutid for up to 36 weeks. The primary efficacy endpoint of the trial was the absolute change from baseline in Forced Vital Capacity (FVC).

Four patients who withdrew from the trial before week 5 of follow-up were not included in SAG and efficacy testing (Figure 1a). At week 36 of treatment, there were 28 patients with lung function measurements. Missing FVC values for 20 AIR patients were extrapolated to 36 weeks from the last follow-up measurement assuming the same rate of decline as that of the external control data (Figure 1.A).

SAG for 48 AIR patients was performed prior to efficacy testing (Figure 1.B). Random samples were drawn without replacement in a 1:1 treatment:control ratio from Qureight's pool of treatment-naive patients from around the world. The random samples were accepted as synthetic control arms (SCAs) if they passed statistical tests for similarity with the AIR cohort on 8 baseline characteristics (Figure 1.B).

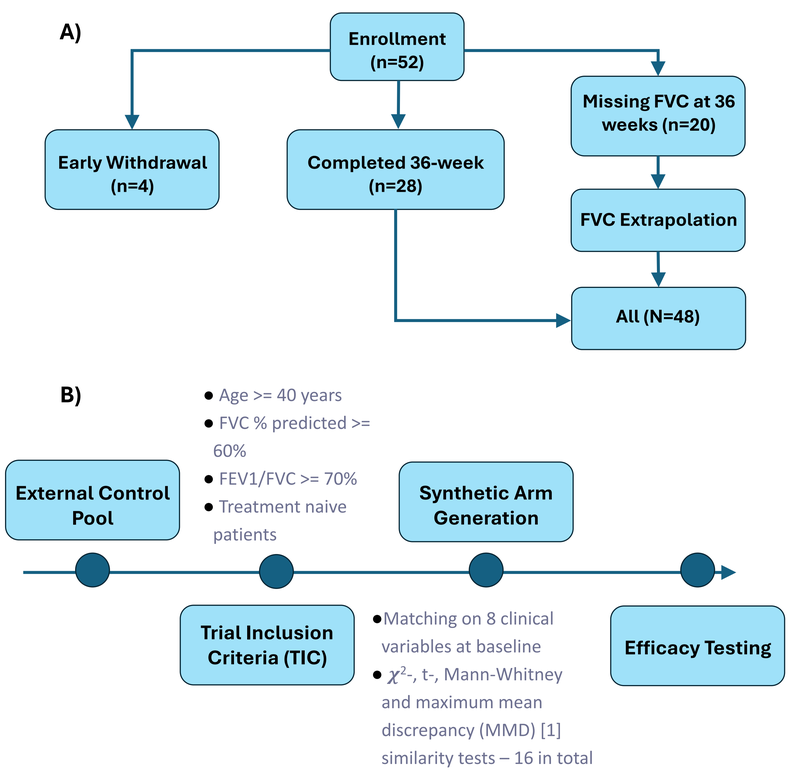

Figure 1. A) Overview of AIR trial data; **B)** Synthetic arm generation and efficacy testing workflow.

Baseline Characteristics of AIR and Control Pool Cohorts

Prior to SAG, the AIR cohort and the filtered control pool have statistically significant differences in the baseline features on which cohort matching was performed (Figure 2).

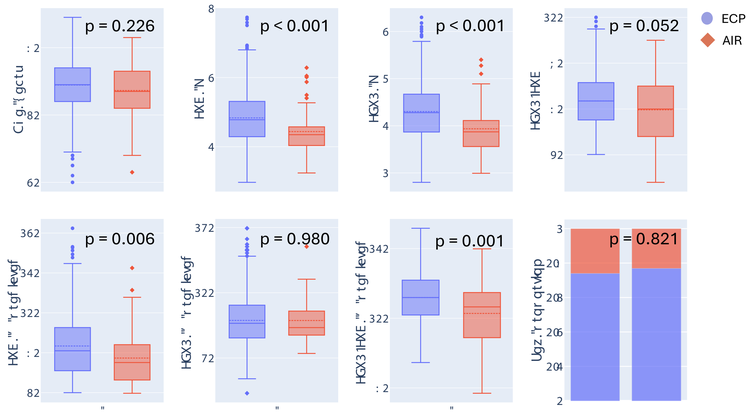

Figure 2. Comparison of the characteristics on which the AIR cohort and the external control pool (ECP) were matched at baseline. Mean values in the boxplots are represented with a dashed line, medians are represented with a solid line. The boxplot whiskers represent the box edges (25th or 75th percentile) +/- 1.5 * IQR (Interquartile Range). The p-values are from the t-tests for similarity between cohort means for the continuous variables and from a chi-squared contingency test for sex.

Synthetic Arm Generation

SCAs were generated by repeatedly drawing random samples from the filtered external control pool (Figure 3.A). The 408 SCAs had a mean multivariable MMD p-value for similarity at baseline of 0.144 (SD = 0.060) and an overall mean p-value for similarity tests on all baseline features of 0.239 (SD = 0.044). The distribution of p-values is presented in (Figure 3.B)

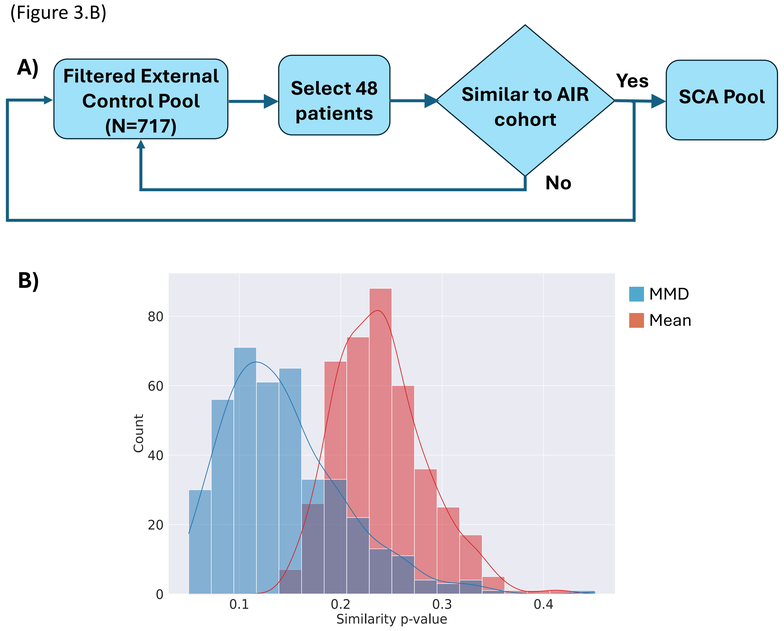

Figure 3. A). Flow diagram of synthetic control arm (SCA) generation; **B)** Distribution of the maximum mean discrepancy (MMD) and of the mean of all 16 similarity tests p-values of the generated 408 SCAs.

Lung Function Decline and Efficacy Testing

Missing FVC values for the AIR cohort were extrapolated to 36 weeks assuming the rate of change (-3.5 ml/week) observed for the external control pool prior to applying TIC. The mean FVC change from baseline was 216 ml (SD = 556 mL) for patients who completed week 36 and 23 ml (SD = 508 ml) when missing values were extrapolated (Figure 4).

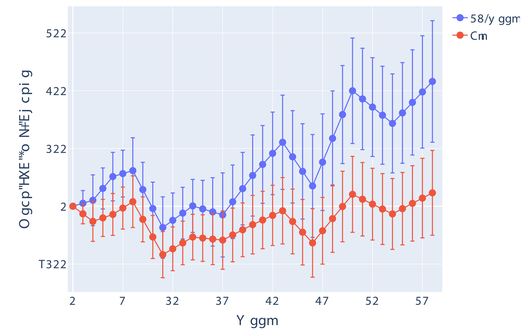

Figure 4. Mean FVC change of the AIR trial cohort as a function of treatment time. The data are presented for patients who have completed 36 weeks of therapy (36-week, N=28) and for all AIR patients (All, N=48) with extrapolation of missing values. The error bars represent the standard error of the mean.

The mean FVC change of the AIR cohort at Week 36 of treatment was compared with the distribution of mean FVC changes of the 408 SCAs (Figure 5). Only one of the SCAs (0.25%) had a greater mean FVC change – efficacy p-value of 0.0025.

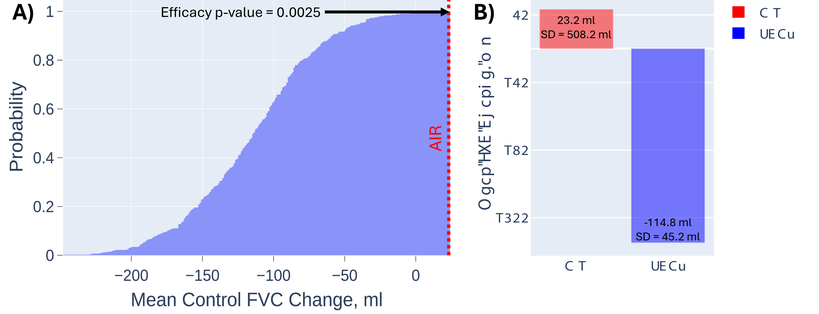

Figure 5. A) Empirical cumulative distribution function of mean control FVC change for the synthetic control arms (SCAs) at 36 weeks of follow-up. The dashed red line represents the mean FVC change of the AIR cohort. **B)** Bar plot of the mean FVC change of the SCAs (grand mean) and of the AIR cohort.

Conclusions

- ❑ We have shown a statistically significant difference between the mean FVC change at 36 weeks of treatment of the AIR cohort and of a large set of closely matched synthetic control arms (SCAs).
- ❑ This marks the first use of SCAs for efficacy testing in a trial of oral antifibrotic therapy and supports further investigation of Buloxibutid in IPF clinical trials.

References

1. Gretton A, Borgwardt KM, Rasch MJ, Schölkopf B, Smola A. A kernel two-sample test. Journal of Machine Learning Research 2012;13:723–773.

Conflicts of Interest

Kirov, Thillai, and Walsh are employees of Qureight Ltd. Lindmark and van den Blink are employees of Vicore Pharma AB. This work is funded by Vicore Pharma AB and Qureight Ltd.