

Home monitoring improves endpoint efficiency in idiopathic pulmonary fibrosis

Supplementary Materials

Kerri A. Johansson MD

Eric Vittinghoff Ph.D

Julie Morisset MD

Joyce S. Lee MD

John R. Balmes MD

Harold R. Collard MD

Sample sizes required to detect treatment effects using a linear mixed model (LMM) for weekly home FVC measurements are sensitive to the within-subject (intraclass) correlation (ICC) of the repeated measurements (see Supplemental Table). For example, the sample size required to detect a 50% treatment effect over 24 weeks decreases from 2113 to 529 as the ICC increases from 0.8 to 0.95; these estimates assume that the ICC is the same for measurements taken 1, 2, 3, ..., 24 weeks apart. However, the sample size also strongly depends on whether the ICC is in fact constant. For example, if the correlation decreases from 0.95 for sequential measurements to 0.90 for measurements taken 24 weeks apart, the sample size increases from 529 to 3040. This happens because the LMM, with a correctly specified correlation structure, puts relatively little weight on measurements taken farther apart, which are most informative about changes. An incorrectly specified LMM assuming constant correlation can reduce the required sample size without biasing the treatment effect estimate. However, such a model will overstate precision and thus inflate the probability of false-positive results.

Supplemental Table S1: Sample sizes for linear mixed models with various correlation patterns.

Correlations of home FVC measurements		Decline in correlation from 1 to 24 weeks		
		0.00	0.025	0.05
Correlation between FVCs at sequential weeks	.80	2113	3415	4701
	.85	1585	2884	4163
	.90	1057	2350	3614
	.95	529	1806	3040

Abbreviations: FVC=forced vital capacity.

Methods

Estimating Sample Sizes

1 ANCOVA

To estimate the sample size for an analysis of baseline to week 24 changes on FVC, UCSD, or VAS, using an ANCOVA model adjusting for the baseline value of the outcome, we used the following formula

$$N = \frac{(z_{1-\alpha/2} + z_{1-\beta})^2 \sigma_y^2 (1 - \rho^2)}{f(1 - f)\Delta^2} \quad (1)$$

where

- α and β are the type-I and type-II error rates respectively; we set $\alpha = 0.05/3$ to account for three primary treatment comparisons, and $\eta = 0.2$ for 80% power

- z_p is the p th quantile of the standard normal distribution
- ρ is the correlation of the baseline and follow-up outcome measures
- $f=0.5$ is the proportion of the sample assigned to treatment
- Δ is the between-group difference in mean changes in the outcome.

2 Linear Mixed Model

To estimate sample size for a between-group comparison in trends using a linear mixed model (LMM) for repeated weekly or daily measures of FVC, UCSD, or VAS, we used the closely related formula

$$N = \frac{z_{1-\alpha/2} + z_{1-\beta})^2 \sigma_\gamma^2}{f(1-f)\gamma^2} \quad (2)$$

where α , β , and f are defined as in (1), γ is the regression coefficient for the treatment-by-time interaction in the LMM, and σ_γ^2 is its variance. The covariance matrix of the regression coefficients from the LMM can be approximated by

$$\sigma^2 \sum_{i=1}^N \mathbf{X}_i^T \mathbf{W} \mathbf{X}_i^{-1}, \quad (3)$$

where

- σ^2 is the variance of the outcome
- \mathbf{X}_i is the $\tau \times 4$ design matrix for observation i , with columns for the intercept, treatment, time, and the time-by-treatment interaction, and up to τ rows for weekly or daily measurements. We assume that the time variable is on the unit interval.
- \mathbf{W} is the inverse of the $\tau \times \tau$ correlation matrix of the repeated outcomes.

To accommodate drop-out we made the simplifying assumption of steady attrition week-by-week, with perfect adherence up to the time of drop-out. Then for patients lost after $k = 1, 2, \dots, \tau - 1$ weeks, the design matrix is taken to be the first k rows of \mathbf{X}_i , and W is obtained by inverting the first k rows and columns of the correlation matrix of the outcomes.

3 Implementation

Programs to carry out these computations were written in R (R Foundation for Statistical Computing, Vienna, Austria).

R-code for sample size calculations:

```
# sample size calculations based on ANCOVA
getn.ancova = function(
  sdy = .65,
  icc = .91,
  cch = c(0.05,.1,.15),
  txe = c(.2, .35, .5),
  f = 0.5,
  alpha = .05,
  power = .8,
  rho.adjust = 0.0,
  ltfu = 0.05
)
{
n = matrix(NA, length(cch), length(txe))
dimnames(n) = list(cch*1000, txe*100)
sdr = sdy*sqrt(1-icc^2)
for(i in 1:length(cch)) {
  for(j in 1:length(txe)) {
    n[i,j] <- ((qnorm(1-alpha/2)+qnorm(power))*(sdr/(txe[j]*cch[i])))^2/(f*(1-f))
  }
}
cat("\nSample size estimates based on ANCOVA, by control change over 24 weeks (rows) and
effect size in % (columns)\n")
print(round(n/(1-ltfu)))
}
# sampoe size calculations based on LMM for weekly measures
getn.dtrend = function(
  sdy = .65,
  icc = .91,
  cch = c(0.05,.1,.15),
  lambda = 0,
  corrtype = "compound symmetric",
  lt = 24,
  ltfu = 0.1, # net LTFU
  txe = c(.2, .35, .5), # net treatment effect as proportion of control change
  f = 0.5, # proportion in group 1
  alpha = 0.05,
  power = .8
)
{
time = seq(0,1,length=lt)
ploss <- 1-(1-ltfu)^(1/(lt-1))
pnv <- c(ploss*(1-ploss)^(0:(lt-2)), 1-ltfu)
```

```

X0 = cbind(rep(1, lt), rep(0, lt), time, rep(0, lt))
X1 = cbind(rep(1, lt), rep(1, lt), time, time)
N = matrix(NA, length(cch), length(txe))
dimnames(N) = list(cch*1000, txe*100)
if (corrtype == "compound symmetric") {
  S <- matrix(icc, lt, lt) + diag(rep(1-icc, lt))
}
else if (corrtype == "AR(1)") {
  S <- matrix(NA, lt, lt)
  for (k in 1:lt) {
    for (l in 1:lt) {
      S[k, l] <- icc^abs(k-l)
    }
  }
}
else if (corrtype == "exponential") {
  S <- matrix(1, lt, lt)
  for (l in 1:lt) {
    for (m in 1:lt) {
      if (l!=m) S[l, m] <- icc*exp(abs(time[l]-time[m])*lambda)
    }
  }
}
XWX = matrix(0, 4, 4)
for (k in 2:lt) {
  W = solve(S[1:k, 1:k])
  XWX = XWX + pnv[k]*((1-f)*t(X0[1:k, ]) %*% W %*% X0[1:k, ] +
    f*t(X1[1:k, ]) %*% W %*% X1[1:k, ])
}
var.beta <- solve(XWX)
for (i in 1:length(cch)) {
  for (j in 1:length(txe)) {
    N[i,j] = (qnorm(1-alpha/2)+qnorm(power))^2*var.beta[4, 4]*(sdy/(txe[j]*cch[i]))^2
  }
}
cat("\nSample size estimates based on LMM, by control change over 24 weeks (rows) and effect
size in % (columns)\n", sep="")
print(round(N))
}

# run calculations
cat("\nSample size for FVC\n")
getn.ancova()
getn.dtrend()

cat("\nSample size for VAS\n")

```

```

getn.ancova(cch=1,sdy=2,icc=.69)
getn.dtrend(cch=1,sdy=2,icc=.92,corrtype="exponential",lambda=log(69/92))

cat("\nSample size for UCSD\n")
getn.ancova(cch=6,sdy=28,icc=.85)
getn.dtrend(cch=6,sdy=28,icc=.98,corrtype="exponential",lambda=log(85/98))

cat("\nAppendix table: Sample sizes for ancova and LMM under varying correlation
structures\n")
icc <- seq(.8,.95,by=0.05)
dicc <- c(.025,.05)
for(i in 1:4) {
  getn.dtrend(cch=.05,sdy=.65,icc=icc[i],power=0.8,txe=.5)
  for(j in 1:2) {
    getn.dtrend(cch=.05,sdy=.65,icc=icc[i],power=0.8,txe=.5,
      corrtype="exponential",lambda=log((icc[i]-dicc[j])/icc[i]))
  }
}

```

Supplemental Figure S1: Dyspnea Visual Analog Score (VAS)

10-point Visual Shortness of Breath Scale (complete once a week)

Study ID: APPLE - ____

Week # ____

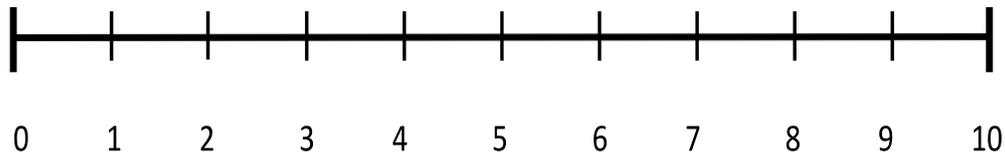
Today's Date (month/day/year): _____

During daily activities over the past week, how short of breath did you feel ? (Circle one)

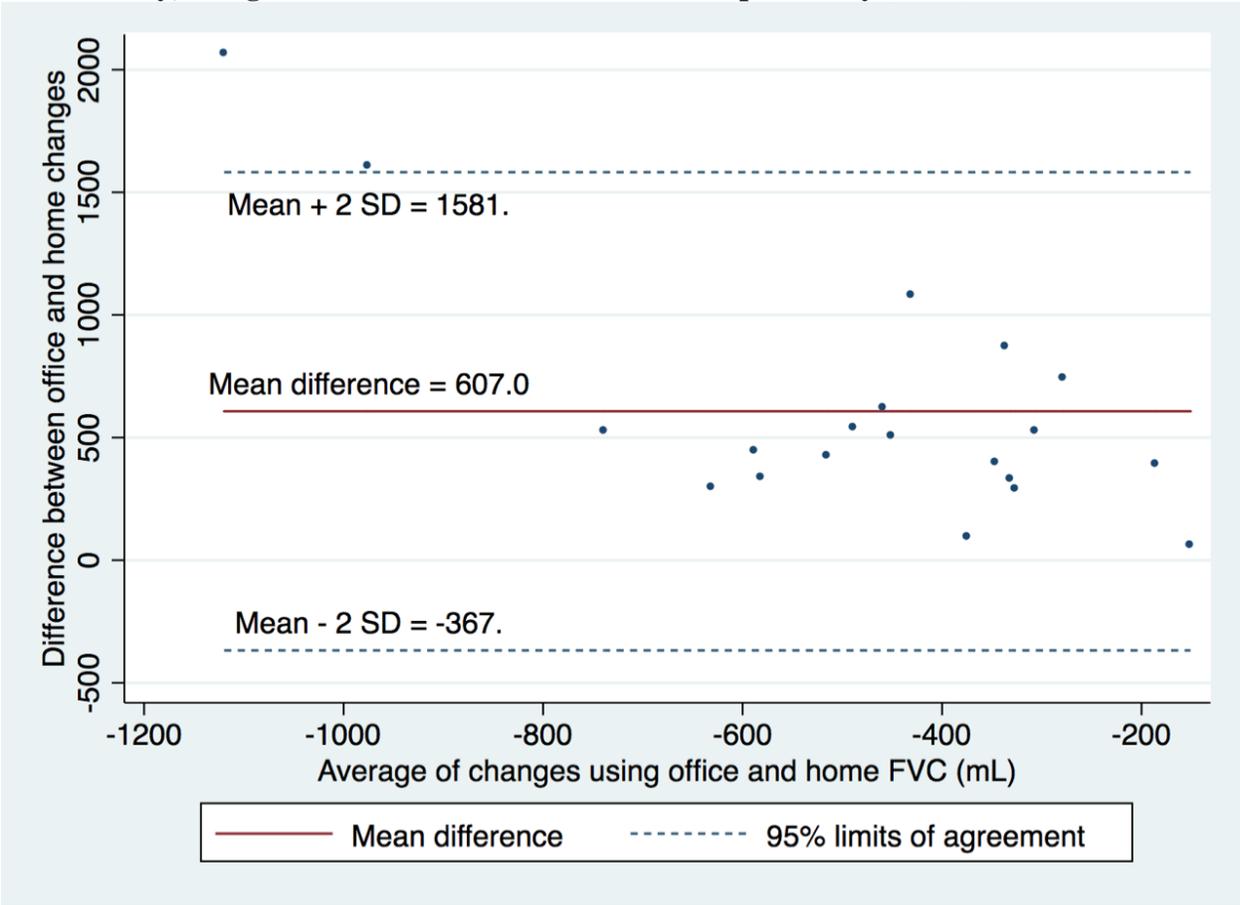
Not short of
breath at all

Some
shortness of
breath

Worst
breathlessness
imaginable

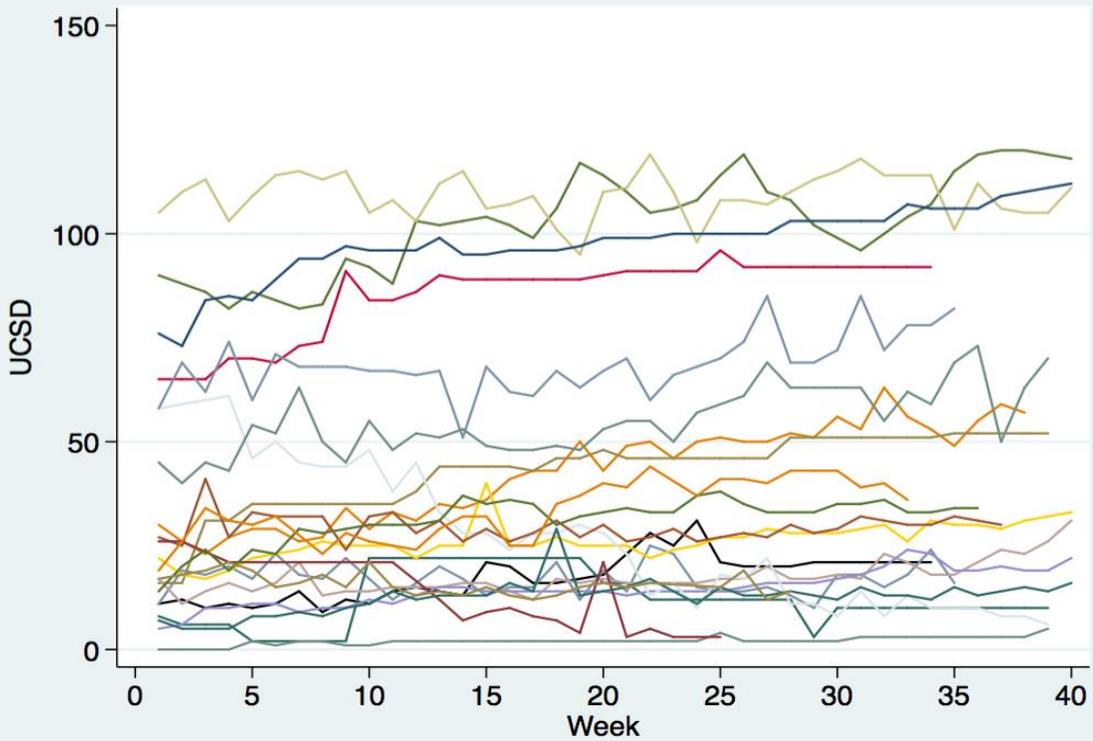


Supplemental Figure S2: Bland-Altman plot comparing change in FVC from baseline to end-of-study, using office-based and hand-held home spirometry (n=15).



Supplemental Figure S3: Individual patient (n=25) longitudinal changes in A) University of California, San Diego shortness of breath questionnaire (UCSD) and B) Visual analog scale score (VAS).

A



B

