Longitudinal Data Analysis

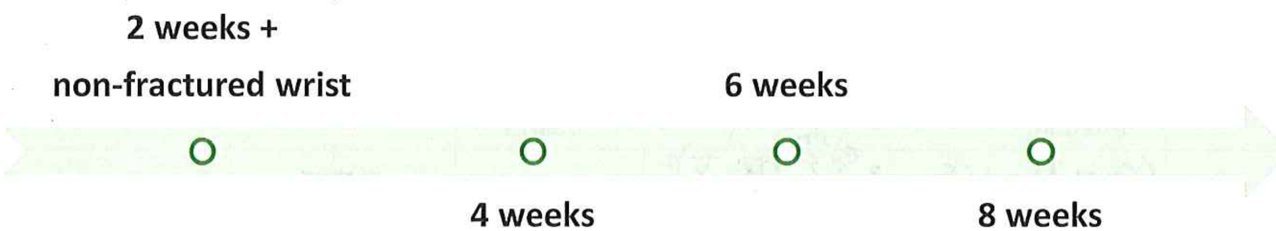
With focus on Repeated Measures ANOVA

McCaig Statistics Group

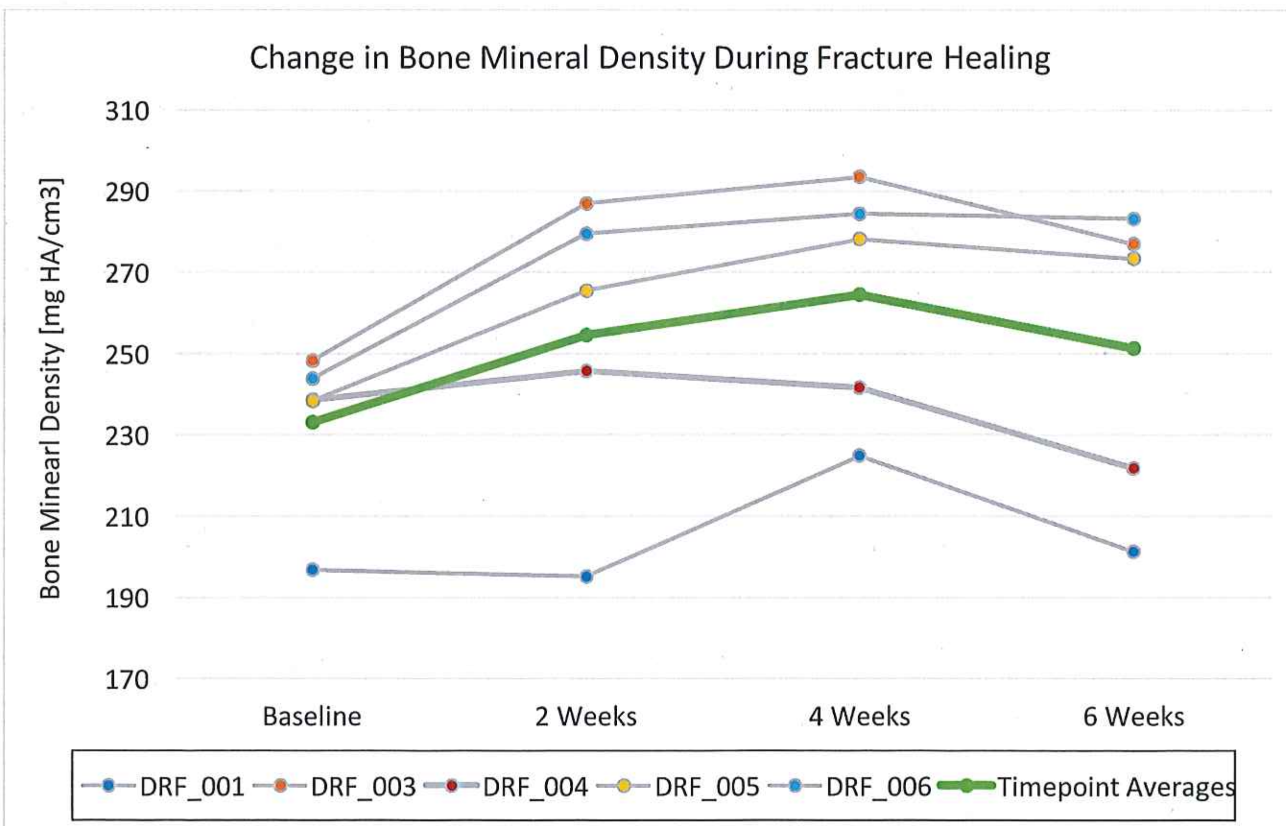
Danielle Whittier

Example: Suppose you've been roped into coming up with a way to analyze and quantify the healing process of typical wrist fractures using high-resolution CT. This is a new field of study, so your goal is to hypothesize and validate which measurable parameters show significant change during the healing process. Each participant's fractured wrist is scanned multiple times during healing, and the opposite non-fractured wrist is also scanned to represent the pre-fractured state of the bone. Very quickly you realize this is a longitudinal study and t-tests simply won't do the trick… at the very least you'll need to run a repeated measures ANOVA.

To date you've scanned the fracture region of **5 participants at the intervals shown below**. **You hypothesize that bone mineral density will change over time** as the bone transitions through callus formation, callus mineralization, then finally bone remodeling.



Below is the graph of your data, which looks promising, but *are these changes statistically significant*.



n = # subjects
$N = n \cdot G$ = # observations
G = # groups/conditions/time

Below is the data table summarizing the study

Table 1: Bone Mineral Density results for the first 5 participants of a fracture-healing study

StudyID	Bone Mineral Density [mg HA/cm^3] Baseline	2 Weeks	4 Weeks	6 Weeks	Subject Averages	
DRF_001	197	195	225	201	205	
DRF_003	248	287	294	277	276	
DRF_004	239	246	242	222	237	
DRF_005	238	266	278	273	264	
DRF_006	244	280	285	283	273	
Time-Point Averages	233	255	265	251	251	Grand Average

Table 2: A Somewhat Standard ANOVA Table

	Degrees of freedom	Sum of Squares	Mean Squares	F-ratio	p-value
Between Times	$(G-1)$ 3	$SS_{time(b)} = \sum_{k=1}^{G}\sum_{i=1}^{n_k}(\bar{Y}_k - \bar{Y})^2$ 2567	856	$\frac{MS_{time}}{MS_{time\ err.}} = 6.4$	0.008
Within Times	$(N-G)$ 16	$SS_w = \sum\sum(Y_{ik} - \bar{Y}_k)^2$ 16,176	1011		
Within Subjects	$(n-1)$ 4	$SS_{subjects} = \sum_{i=1}^{n}\sum(\bar{Y}_i - \bar{Y})^2$ 14571	~~14571~~ 3643		
Error (unexplained variance)	$(n-1)(G-1)$ 12	$SS_{error} = SS_w - SS_{subjects}$ 1604	~~1604~~ 134		

$F(3,12) = 6.4, p = 0.008$

→ Increased power by the reduction of MS error

Greenhouse Geisser correction – $F(1.68, 6.7) = 6.4, p = 0.03$

Assumptions: Normally distributed → do a test
Homogeneous variance → do a test Mauchly's Test for Sphericity

BOTH Failed

Table 3: Effect Size:

Effect Size Metric	Formula	Example Value
~~Partial~~ Eta Squared η^2	$\eta^2 = \frac{SS_{time}}{SS_{total}}$	$\frac{856}{18742} = 0.14$
Omega Squared ω^2	$\frac{SS_{time} - df_{time} \cdot MS_{error}}{SS_{total} + MS_{error}}$	$\frac{2567 - 3 \cdot 134}{18742 + 134} = 0.11$

large-medium effect.

Effect Size: