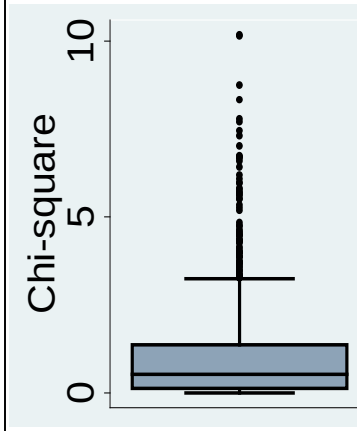
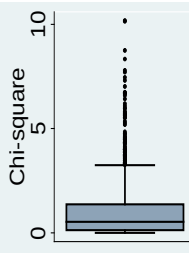
Repeated Measures
Lecture 1, Part 1

Charles E. McCulloch,
Division of Biostatistics,
Dept of Epidemiology and Biostatistics,
UCSF

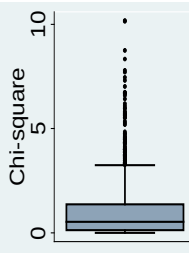


Outline



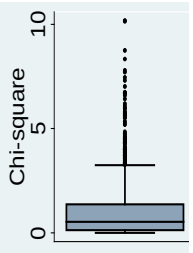
1. Motivating examples and introduction
2. Analysis of the fecal fat data
3. Accommodating correlated data
4. Summary

Example: Fecal fat



Lack of digestive enzymes in the intestine can cause bowel absorption problems. This will be indicated by excess fat in the feces. Pancreatic enzyme supplements can be given to ameliorate the problem. Does the supplement form make a difference? (Graham, Enzyme replacement therapy of exocrine pancreatic insufficiency in man. *NEJM*, **296**: 1314-17, 1977 – But note: sex information made up for illustration.)

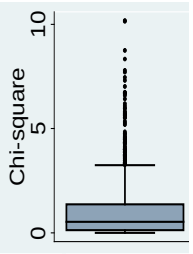
Example: Fecal fat



Considerable person to person variability

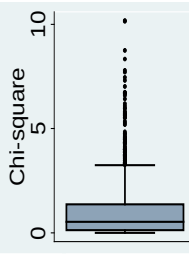
	Fecal Fat (g/day)				
	Pill type				
PatID / Sex	None	Tablet	Capsule	Coated Capsule	Avg
1 – M	44.5	7.3	3.4	12.4	16.900
2 – M	33.0	21.0	23.1	25.4	25.625
3 – M	19.1	5.0	11.8	22.0	14.475
4 – F	9.4	4.6	4.6	5.8	6.100
5 – F	71.3	23.3	25.6	68.2	47.100
6 – F	51.2	38.0	36.0	52.6	44.450
Avg	38.08	16.53	17.42	31.07	25.775

Example: Practice style and back pain (Korff, Barlow, Cherkin, and Deyo, 1994)



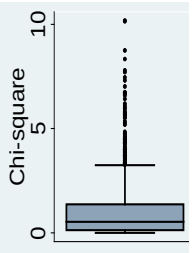
Forty-four primary care physicians in a large HMO were classified according to their practice style in treating back pain management (low, moderate or high frequency of prescription of pain medication and bed rest). An average of 24 patients per physician was followed for 2 years (1 month, 1 year and 2 year followups) after the indexed visit. Outcome measures included functional measures (pain intensity, activity limitation days, etc.), patient satisfaction (e.g., “After your visit with the doctor, you fully understood how to take care of your back problem”), and cost.

Example: Osteoarthritis Initiative (OAI): www.oai.ucsf.edu



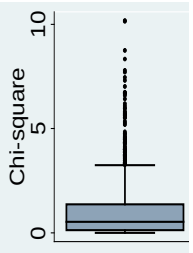
The OAI is a multi-center, longitudinal, prospective observational study of knee osteoarthritis (OA). 4,796 men and women ages 45-79 were enrolled between 2004 and 2006. Image (X-ray and MRI), demographic, and clinical data are being collected yearly. Some of the variables measured over time are: pain scores (one for each knee), presence of OA as judged from X-ray (one for each knee), functional limitation scores (one per person).

Example: Study of Osteoporotic Fractures (SOF): sof.ucsf.edu



The Study of Osteoporotic Fractures (SOF) is a longitudinal, prospective study of osteoporosis, breast cancer, stroke, and total and cause-specific mortality. In 1986, SOF enrolled 9,704 women and continues to track these women with clinical visits every 2 years. Data from the first seven visits are now available to the public. The data include measures of bone mineral density (BMD), sex and calcitropic hormones, tests of strength and function, cognitive exams, use of medication, health habits and much more.

Introduction: Hierarchical data

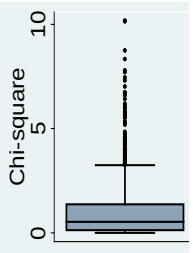


Hierarchical data (responses and/or predictors) are data collected from different levels within a study.

May be repeated measures data (e.g., fecal fat), clustered or multilevel data (e.g., back pain) or longitudinal data (over time, e.g., OAI or SOF).

A characteristic of hierarchical data is that predictors can be measured at any level in the hierarchy.

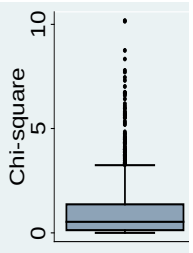
Some prototypical questions: Fecal fat example



Question 1: Does fecal fat depend on the repeated measures factor, pill type?

Question 2: Does fecal fat depend on the non-repeated measures factor, sex?

Some prototypical questions: Back pain example



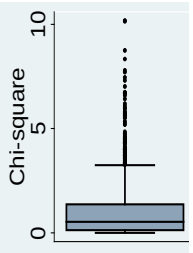
Question 1: Does log cost depend on the between physician factor, practice style?

Question 2: Does understanding of physician recommendation depend on practice style?

Question 3: Does log cost depend on the within physician, between patient factor, sex of the patient?

Question 4: Is there between physician variability in treatment of similar patients?

Some prototypical questions: SOF

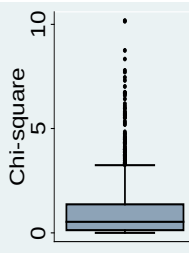


Question 1: Is change in BMD related to age at menopause? (time invariant predictor of change)

Question 2: Is change in BMD related to change in BMI? (time varying predictor of change)

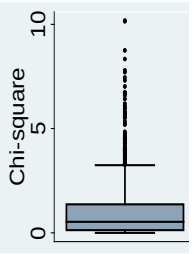
Question 3: Which participants are maintaining cognitive function into their 9^{th} and 10^{th} decades of life? (subject specific prediction)

Introduction



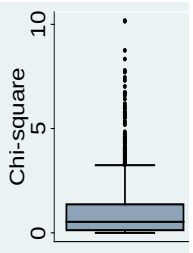
- Analysis technique depends on nature of the *outcome* variable and research question.
  - Binary: logistic regression (e.g., BMI>30)
    - Odds ratios, area under ROC curve
  - Numeric: linear regression (e.g., BMI, BMD)
  - Also – time to event (Cox model or pooled logistic regression), count outcomes (Poisson regression)
- Methods need to be modified for hierarchical data.

Accommodating hierarchical data



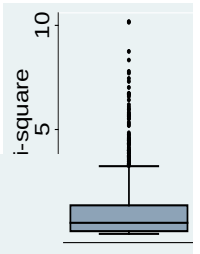
- Repeated measures/clustering in the *outcome* requires new methods of analysis. But not in the predictor.
- SOF: Is visit 8 cognitive status (outcome) related to previous (repeatedly measured across multiple visits) physical activity? Does not have repeated measures on the outcome.
- This situation can be accommodated by including multiple values of physical activity as predictors or by calculating summary measure(s) (e.g., average physical activity).

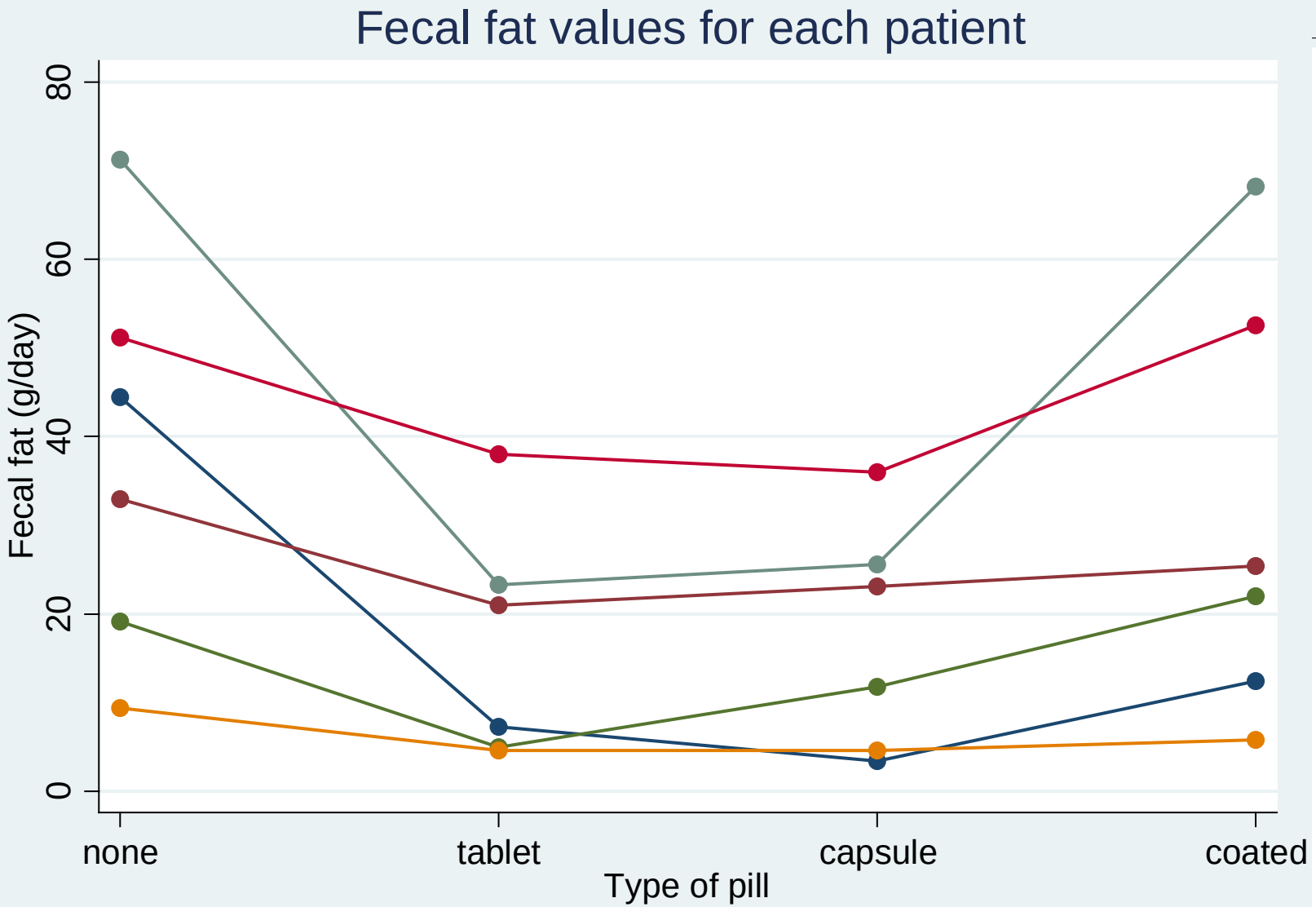
Fecal fat data analysis



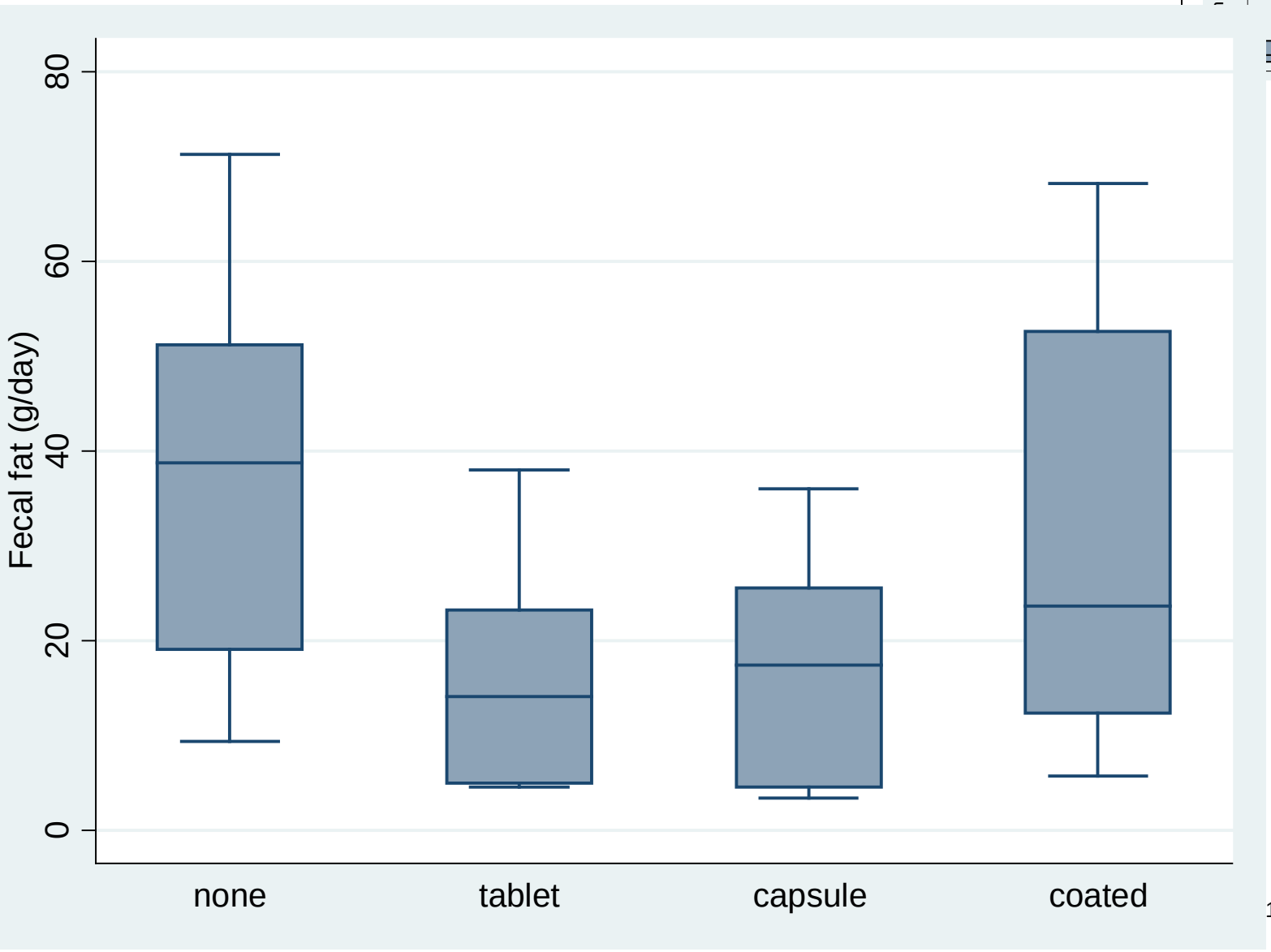

```
. tabstat fecfat, by(pilltype) stats(n mean sd min max)

Summary for variables: fecfat
     by categories of: pilltype (Type of pill)

pilltype |         N      mean        sd       min       max
---------+--------------------------------------------------
    none |         6  38.08333  22.47447       9.4      71.3
  tablet |         6  16.53333  13.32091       4.6        38
 capsule |         6  17.41667  12.93745       3.4        36
  coated |         6  31.06667   24.2641       5.8      68.2
---------+--------------------------------------------------
   Total |        24    25.775  20.00214       3.4      71.3
--------------------------------------------------------------
```

Fecal fat data analysis

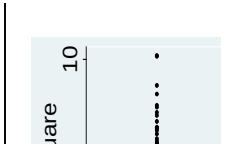




Fecal fat data analysis



Regression/ANOVA ignoring sex effects (a wrong analysis)



```
. regress fecfat i.pilltype

      Source |       SS           df       MS      Number of obs   =        24
-------------+----------------------------------   F(3, 20)        =      1.86
       Model |   2008.6017         3  669.533901   Prob > F        =    0.1687
    Residual |  7193.36328        20  359.668164   R-squared       =    0.2183
-------------+----------------------------------   Adj R-squared   =    0.1010
       Total |  9201.96498        23  400.085434   Root MSE        =    18.965

------------------------------------------------------------------------------
      fecfat |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    pilltype |
     tablet  |     -21.55    10.9494    -1.97   0.063    -44.39005     1.29005
    capsule  |  -20.66667    10.9494    -1.89   0.074    -43.50672    2.173384
     coated  |  -7.016668    10.9494    -0.64   0.529    -29.85672    15.82338
             |
       _cons |   38.08333   7.742396     4.92   0.000     21.93298    54.23369
------------------------------------------------------------------------------
```

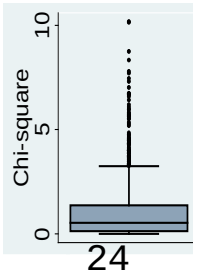
```
. testparm i.pilltype
 ( 1)  2.pilltype = 0
 ( 2)  3.pilltype = 0
 ( 3)  4.pilltype = 0
       F(  3,    20) =    1.86
            Prob > F =    0.1687
```

Difference between average for Tablet – None

(16.53 – 38.08 = -21.55)

Overall test, p=0.17

A hierarchical analysis

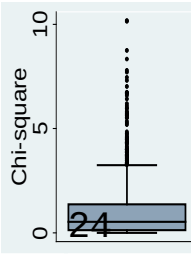



```
. xtgee fecfat i.pilltype, i(patid)

GEE population-averaged model                   Number of obs     =         24
Group variable:                      patid      Number of groups  =          6
Link:                              identity     Obs per group:
Family:                            Gaussian                   min =          4
Correlation:                   exchangeable                   avg =        4.0
                                                              max =          4
                                                Wald chi2(3)      =      22.53
Scale parameter:                   299.7235     Prob > chi2       =     0.0001

------------------------------------------------------------------------------
      fecfat |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    pilltype |
     tablet  |     -21.55   5.451781    -3.95   0.000    -32.23529   -10.86471
    capsule  |  -20.66667   5.451781    -3.79   0.000    -31.35196   -9.981373
     coated  |  -7.016668   5.451781    -1.29   0.198    -17.70196    3.668626
             |
       _cons |   38.08333   7.067808     5.39   0.000     24.23068    51.93598
------------------------------------------------------------------------------
. testparm i.pilltype
 ( 1)  2.pilltype = 0
 ( 2)  3.pilltype = 0
 ( 3)  4.pilltype = 0
           chi2(  3) =   22.53
         Prob > chi2 =    0.0001
```

A hierarchical analysis

```
. xtgee fecfat i.pilltype, i(patid)
GEE population-averaged model                   Number of obs     =         24
Group variable:                      patid      Number of groups  =          6
Link:                             identity      Obs per group:
Family:                           Gaussian                    min =          4
Correlation:                  exchangeable                    avg =        4.0
                                                              max =          4
                                                Wald chi2(3)      =      22.53
Scale parameter:                  299.7235      Prob > chi2       =     0.0001

------------------------------------------------------------------------------
      fecfat |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    pilltype |
     tablet  |     -21.55   5.451781    -3.95   0.000    -32.23529   -10.86471
    capsule  |  -20.66667   5.451781    -3.79   0.000    -31.35196   -9.981373
     coated  |  -7.016668   5.451781    -1.29   0.198    -17.70196    3.668626
             |
       _cons |   38.08333   7.067808     5.39   0.000     24.23068    51.93598
------------------------------------------------------------------------------
. xtcorr
Estimated within-patid correlation matrix R:
           c1      c2      c3      c4
r1    1.0000
r2    0.7025  1.0000
r3    0.7025  0.7025  1.0000
r4    0.7025  0.7025  0.7025  1.0000
```

A hierarchical analysis (variation)

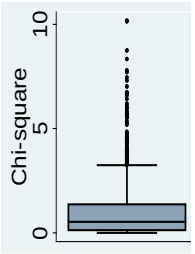



```
. xtgee fecfat i.pilltype, i(patid) robust

GEE population-averaged model                   Number of obs      =        24
Group variable:                      patid      Number of groups   =         6
Link:                             identity      Obs per group: min =         4
Family:                           Gaussian                     avg =       4.0
Correlation:                  exchangeable                     max =         4
                                                Wald chi2(3)       =     11.71
Scale parameter:                  299.7235      Prob > chi2        =    0.0084
                                  (Std. Err. adjusted for clustering on patid)
------------------------------------------------------------------------------
             |               Robust
     fecfat |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    pilltype |
     tablet  |     -21.55   6.931847    -3.11   0.002    -35.13617    -7.96383
    capsule  |  -20.66667   7.349407    -2.81   0.005    -35.07124   -6.262094
     coated  |  -7.016668   5.246295    -1.34   0.181    -17.29922    3.265881
             |
       _cons |   38.08333   9.175163     4.15   0.000     20.10034    56.06632
------------------------------------------------------------------------------
```

```
. testparm i.pilltype
 ( 1)  2.pilltype = 0
 ( 2)  3.pilltype = 0
 ( 3)  4.pilltype = 0
           chi2(  3) =   11.71
         Prob > chi2 =    0.0084
```

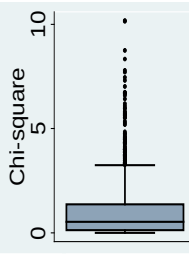
Coefficient unchanged, but SE is slightly different (6.93 versus 5.45 without robust)

Accommodating hierarchical data



- The usual statistical methods (multiple regression, basic ANOVA, logistic regression, and many others) assume observations are independent.
- *Important idea*: observations taken within the same subgroup in a hierarchy are often more similar to one another than to observations in different subgroups, other things being equal. [correlated]
- Failure to accommodate the hierarchical nature of the data can lead to incorrect SEs, p-values and confidence intervals, sometimes grossly incorrect.

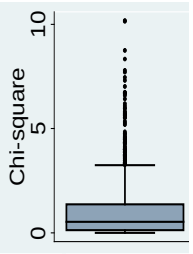
Regr/ANOVA with sex effects (incorrect analysis)




```
. regress fecfat i.pilltype i.sex

      Source |       SS           df       MS      Number of obs   =        24
-------------+----------------------------------   F(4, 19)        =      2.43
       Model |  3110.21668         4  777.554169   Prob > F        =    0.0837
    Residual |   6091.7483        19  320.618332   R-squared       =    0.3380
-------------+----------------------------------   Adj R-squared   =    0.1986
       Total |  9201.96498        23  400.085434   Root MSE        =    17.906

------------------------------------------------------------------------------
      fecfat |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    pilltype |
     tablet  |     -21.55   10.33793    -2.08   0.051    -43.18753    .0875334
    capsule  |  -20.66667   10.33793    -2.00   0.060     -42.3042    .9708671
     coated  |  -7.016668   10.33793    -0.68   0.505     -28.6542    14.62087
             |
       1.sex |      13.55    7.31002     1.85   0.079    -1.750047    28.85005
       _cons |   31.30833   8.172851     3.83   0.001     14.20236    48.41431
------------------------------------------------------------------------------
```

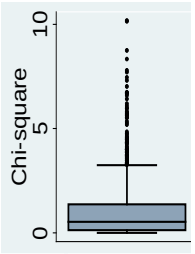
Sex effects are borderline statistically significant

A hierarchical analysis



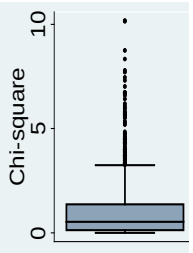

```
. xtgee fecfat i.pilltype i.sex, i(patid)

Iteration 1: tolerance = 9.971e-16

GEE population-averaged model                   Number of obs      =         24
Group variable:                      patid      Number of groups   =          6
Link:                              identity     Obs per group: min =          4
Family:                          Gaussian                   avg =        4.0
Correlation:                 exchangeable                   max =          4
                                                Wald chi2(4)       =      24.00
Scale parameter:                  253.8228      Prob > chi2        =     0.0001

------------------------------------------------------------------------------
      fecfat |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    pilltype |
           2 |     -21.55   5.451781    -3.95   0.000    -32.23529   -10.86471
           3 |  -20.66667   5.451781    -3.79   0.000    -31.35196   -9.981373
           4 |  -7.016668   5.451781    -1.29   0.198    -17.70196    3.668626
             |
       1.sex |       13.55   11.16389     1.21   0.225    -8.330816    35.43082
       _cons |   31.30833   8.570992     3.65   0.000      14.5095    48.10717
------------------------------------------------------------------------------
```

Fecal fat data analysis - summary



- Failing to accommodate the hierarchical nature of the data led to grossly incorrect statistical inferences.
- Estimates were unchanged, but SEs were affected. (In general estimates tend to be little affected, but do change slightly).
- The proper hierarchical analysis can lead to smaller or larger SEs compared to a naïve analysis.

Not just an academic concern …



J Assist Reprod Genet
DOI 10.1007/s10815-010-9518-0

EMBRYO BIOLOGY

A prospective randomized sibling-oocyte study of two media systems for culturing cleavage-stage embryos—impact on fertilization rate

Received: 30 July 2010 / Accepted: 19 November 2010

Not just an academic concern …

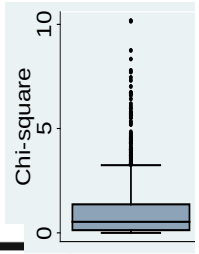
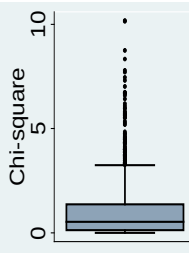


Table 2 Demographic data of patients included in the study is shown

Patients included (*n*)	110
Female age, mean±SD (years)	33.9±3.8
Infertility diagnosis, *n* (%)	
Tubal factor	8 (7.3)
Endometriosis	8 (7.3)
Unexplained	30 (27.3)
Anovulation	12 (10.9)
Male factor	37 (33.6)
Other	15 (13.6)
No. of oocytes retrieved per patient, mean±SD	10.9±4.8

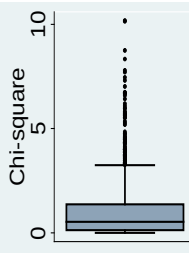
Not just an academic concern …



Statistical analysis

Comparison of normal fertilisation rates, polyploid rates, embryo quality and embryo utilization rates was performed using the Chi-square test. Positive hCG rates, clinical pregnancy rates, implantation rates and delivery rates were compared by using the Fisher's exact test (two-tailed). A value of $p<0.05$ was considered statistically significant.

Summary – Part 1



- Hierarchical data structures are common.
- They lead to correlated data.
- Ignoring the correlation can be a serious error.
- Not easy to predict whether a proper, correlated data analysis will yield larger or smaller standard errors compared to an incorrect analysis that assumes all the data are independent.