Biological Research: Missing Data in Biosocial Research

Tarani Chandola

NCRM, University of Manchester

Methodological considerations

- Biosocial research framework
- Biological data quality issues
- Missing data in biosocial research

Missing biological data

- Substantial proportions of missing biological data
- A number of processes that could result in missing biological data in biosocial surveys
  - No main interview
  - Main interview but no nurse visit
  - Nurse visit but no blood sample
  - Blood sample but no blood analytes

Blood sampling outcome in UKHLS- wave 2

blood sampling outcome- UKHLS wave 2

	Freq.	Percent
blood sample obtained	13021	64%
blood sample not obtained	1412	6.9%
refused blood sample or nurse	4271	21.0%
ineligible for blood sample	1570	7.7%
blood sample lost in transit	81	0.4%

Use survey data for correlates of missing biological data

Detailed survey information on participants

Standard practice for biological data collectors (usually nurses) to record reasons for missing biological samples.

These data can be useful in informing us about the reasons for missing biological samples.

Typical variables included in developing non-response weights

- Sample month/quarter
- Region or other geographical aggregates of postcode sectors
- Deprivation indices (IMD, Townsend, Carstairs) and other geographically-referenced indicators
- Interviewer observation variables regarding the dwelling and immediate surroundings
- Social, demographic and economic indicators from the household and individual questionnaires

English Longitudinal Study of Ageing (ELSA)- wave 6 hair analytes

- At wave 6, ELSA hair samples were collected for the first time.

- From the hair samples, hair analytes such as cortisol were processed. Hair cortisol is an integrated measure of Hypothalamic-Pituitary-Axis (HPA) axis activity, with higher levels indicating higher physiological stress responses.

- Around 2 cm of hair were collected, which is indicative of stress levels over the last 2-3 months.

ELSA Hair Sample Protocol

In order to measure Cortisol in a hair sample, the sample needs to be a minimum of 2cms in length and weigh a minimum of 10mg.

2cms long (10mg)

5cms long (10mg)

10cms long (10mg)

http://www.elsa-project.ac.uk/uploads/elsa/docs_w6/hair_sample_card.pdf
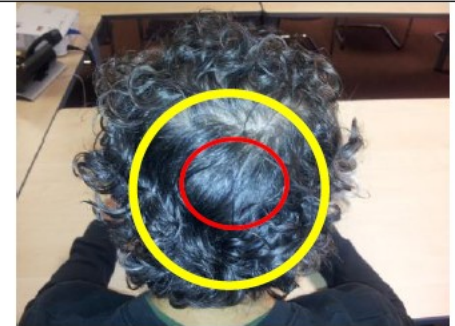
ELSA Hair Sample Protocol

The hair sample should be taken from an area on the back of the head, indicated by the yellow circles on the pictures

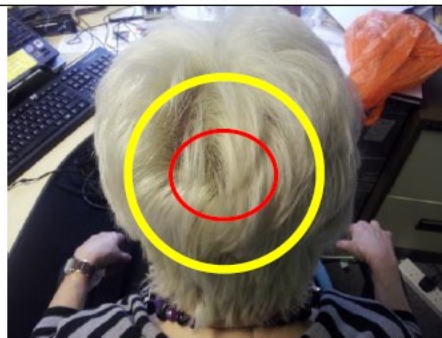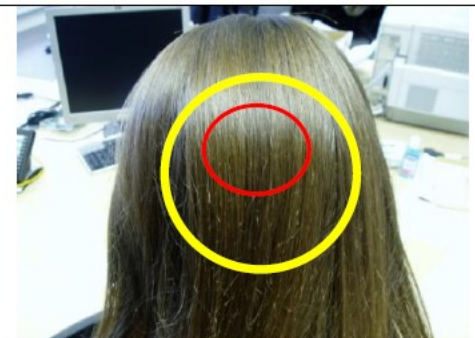|		
This respondent's hair is just 2cms long at the top of the sampling area.	This respondent would need a sample taken from the lower part of sampling area – if 2cms long	This male respondent's hair pattern is thinner at the top – so would need a sample taken from lower part of sampling area
This female respondent's hair pattern is also thinner at the top – so would need a sample taken from lower part of sampling area	This respondent has a dense hair growth pattern across the whole area (at about 5cms long) – resulting in a thinner sample of hair to	This respondent has long, but a fairly fine hair growth pattern across the whole area (at about 30 cms long) – resulting in a thinner

ELSA wave 6- missing hair cortisol data

- Out of 7,419 ELSA participants in the nurse data collection, there were only 2,558 participants with hair cortisol data.

- This is partly because some people were ineligible for the data collection (having less than 2 cm of hair).

- Others refused to give hair samples, mainly for reasons related to appearance.

- And funding constraints meant that only a subset of the hair samples could be processed to produce hair cortisol data.

# Possible characteristics of ELSA participants with missing hair data

- Baldness predominantly affects men and older adults

- Given the importance of appearance to some participants, it is likely that having a negative self-image is linked to missing hair cortisol data.

- ELSA survey asks detailed questions related to depressive symptoms (the CESD questionnaire)

- As stress and depression are interlinked, depressive symptoms may predict both missing hair cortisol data as well as higher levels of hair cortisol.

Predicted probability of missing hair cortisol data by age/gender and depressive symptoms
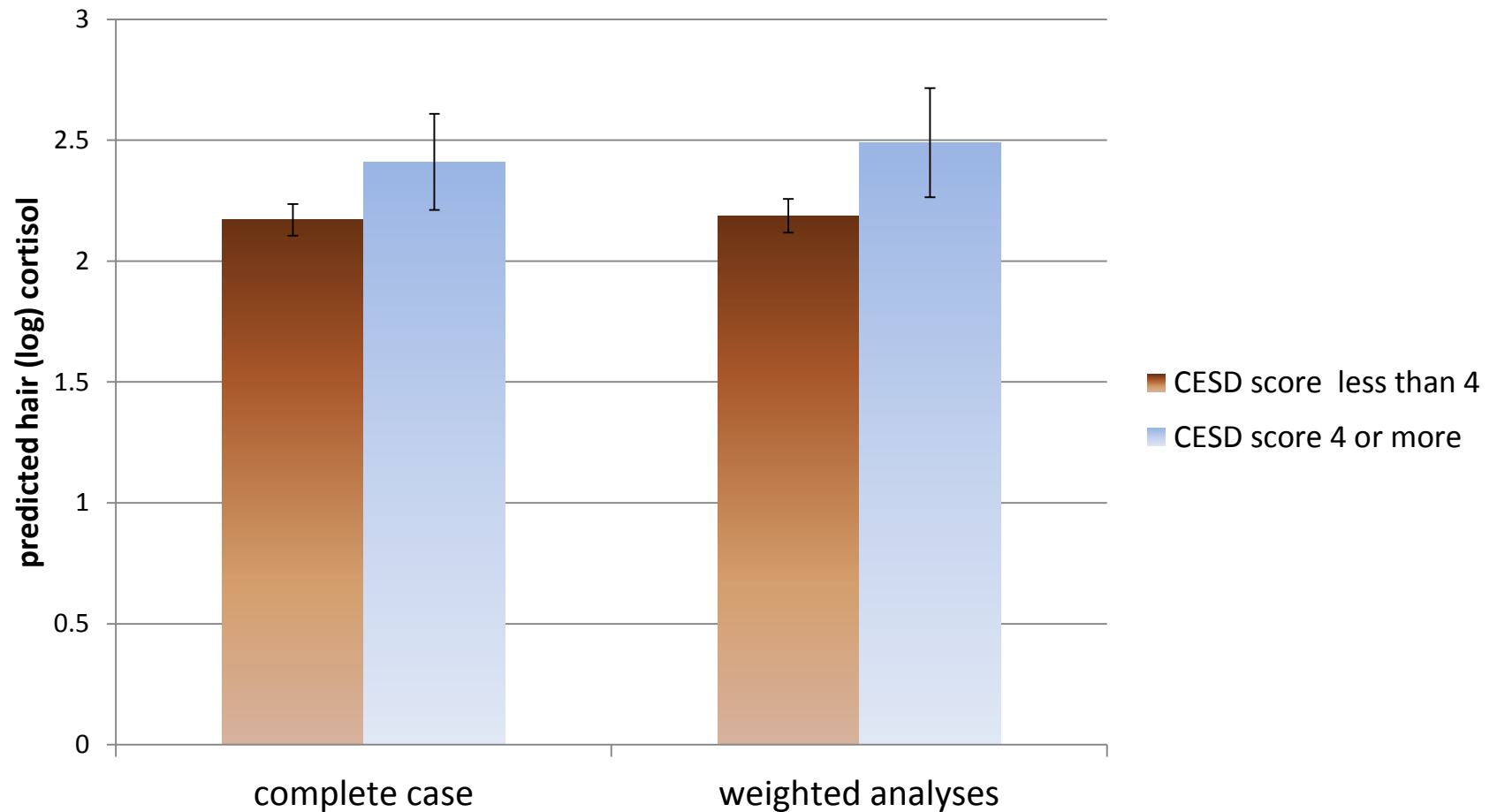
age/gender interaction
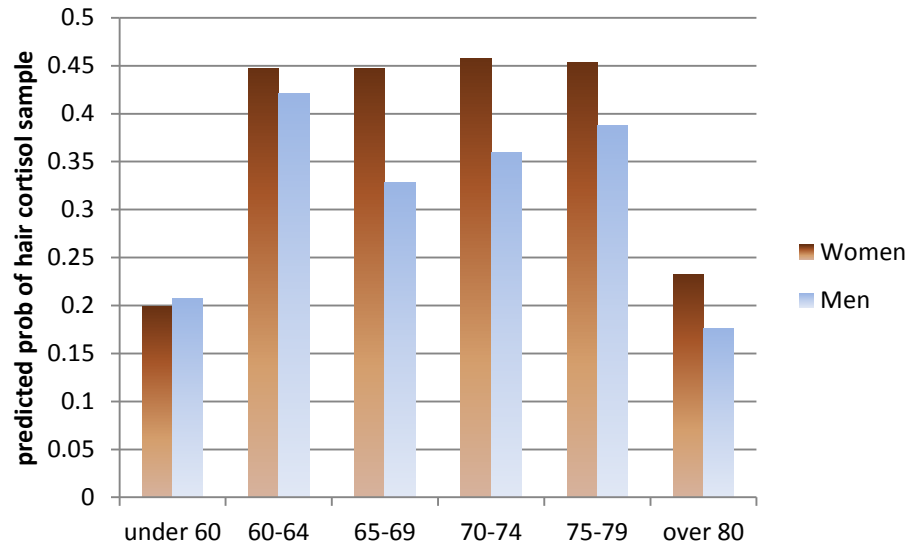
Depressive symptoms
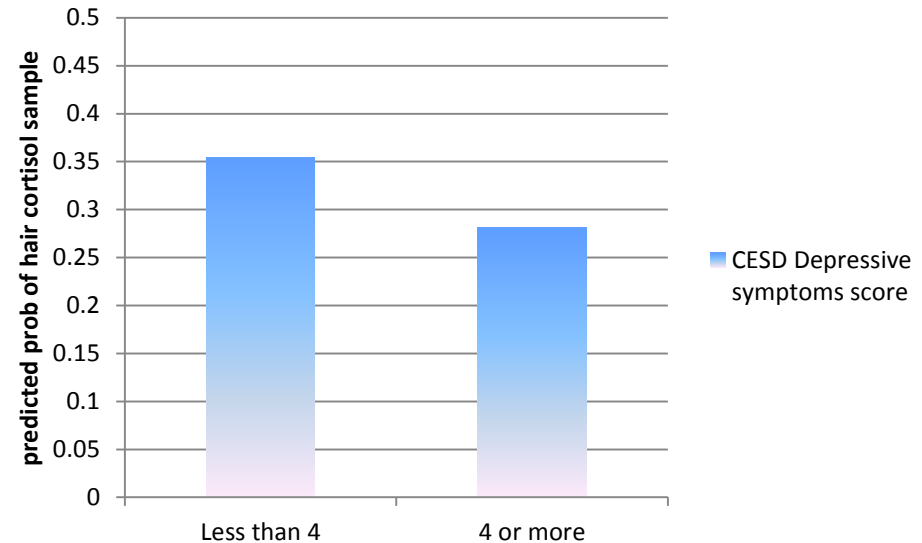
Differences in complete case and weighted regression estimates of (log) hair cortisol by depressive symptoms

Missing biosocial data: methodological considerations

- Rich survey and nurse observation data allows us to discover factors that are both correlated with the missingness mechanism as well as our outcome of interest.

- Inference based on complete case analyses may be biased if we don't take account of such factors.

- Need to investigate the reasons behind missing biological data and incorporate such information in their methods to deal with missing data.

Recap of methodological considerations

- Biosocial research framework

- Biological data quality issues

- Missing data in biosocial research

Methodological considerations when analysing biosocial data

- A biosocial theoretical framework is key
- Consider:
  - normal ranges of biological variables (if available)
  - identify outliers
  - relevant medication use
  - context of blood sampling like time of day, room temperature, recent operations, smoking, food & alcohol, etc
  - quality control processes in producing biological data
  - transformations (for skewed biological dependent variables)
- Identify relevant predictors of missing biological data to use in non-response methods from rich social/attitudinal data

Thank you for your attention