

Stat645

More mondrian

Hadley Wickham

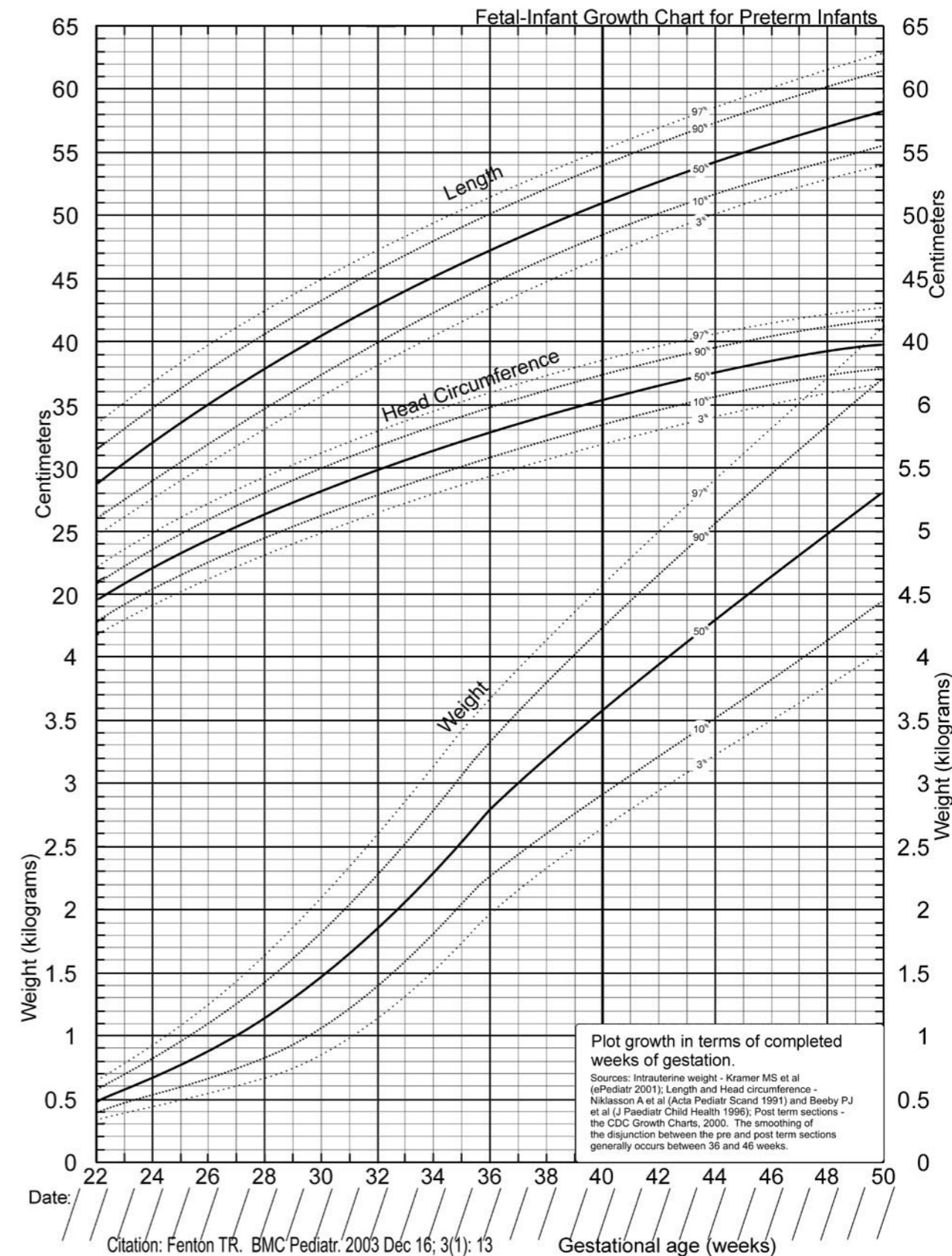
1. What affects birthweight?

2. New plots

3. Discussion

Background

- The influence of different substances such as alcohol, nicotine and drugs on the development of a fetus needed to be studied
- This study focuses on the influence of smoking on the fetal development measured by gestation time and birth weight
- Data collecting interview was conducted early in the pregnancy to avoid bias
- Subset of 1,236 male, single births, surviving at least 28 days



Description of Data

- *Child-related variables:*
- Id Number - identification number
- Date - date of birth
- Gestation
 - length of the gestation (in weeks, fractions represent days)
- Birthweight - weight (in grams)

Mother-related variables:

- Race
- Age (in years)
- Education - education (7 levels)
- Height (in inches)
- Weight (in pounds)
- Smoking
 - never, until pregnancy, etc.
- Smoking Amount
 - number of cigarettes smoked
- Quit
 - history of smoking (if quit, when)

Description of Data (cont.)

Father-related variables:

- Father Race
- Father Age (in years)
- Father Education - education (7 levels)
- Father Height (in inches)
- Father Weight (in pounds)

Family-related variables:

- Marital Status
- Income (in categories)

New plots

Missing value plot: select all variables and choose missing value plot

Boxplot y by x: select a continuous and a categorical variable

(Don't forget to reorder categorical variables)

Discussion

Topics

What are the advantages and disadvantages of interactive graphics?

Why do you think I taught ggplot2 in stat405 instead of Mondrian?

Next week

Two n-d displays

Discrete data: mosaic plots

Continuous data: parallel coordinate plots