Data completeness and quality in a community based and participatory epidemiologic study

Leah Schinasi

Intensive Animal Agriculture, Environmental Injustice, and Public Health

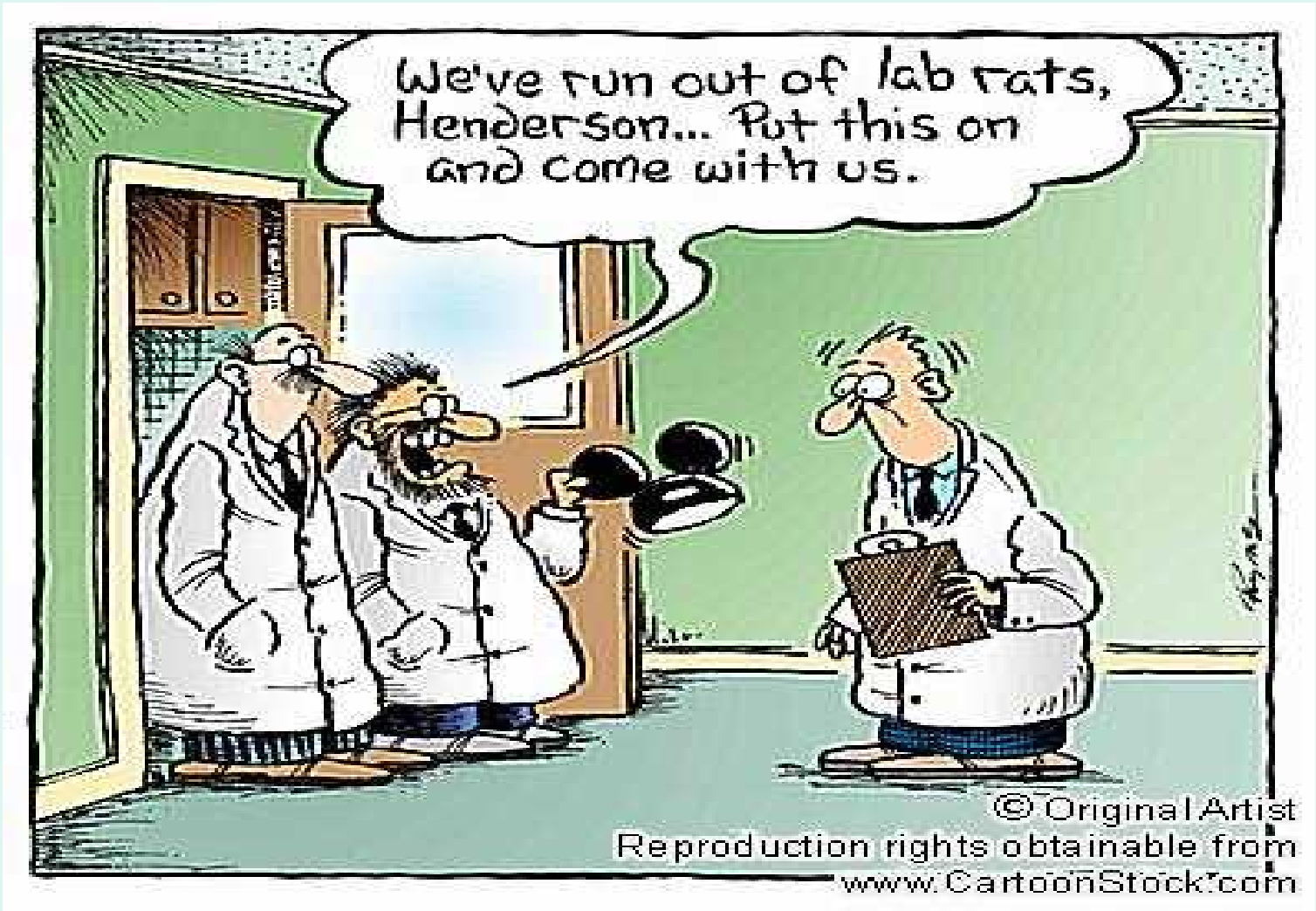
APHA

November 6, 2007

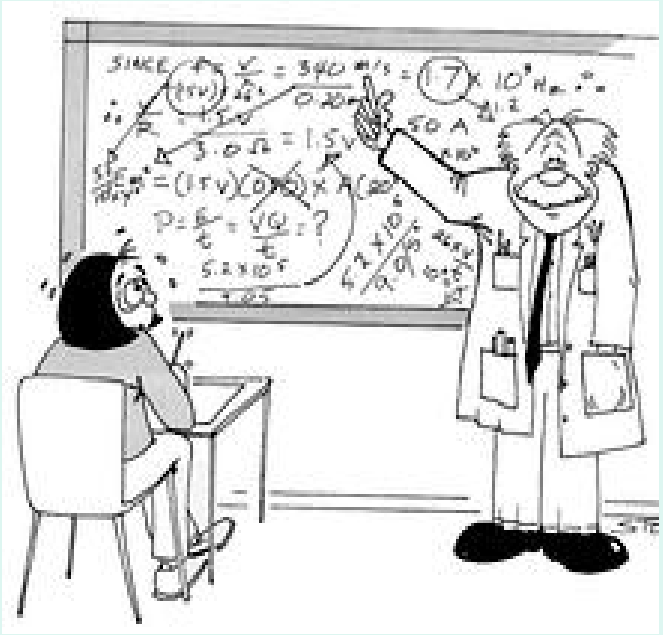
The plan

- Introduce the research question
- Describe the study
- Describe statistical methods
- Outline results and conclusions

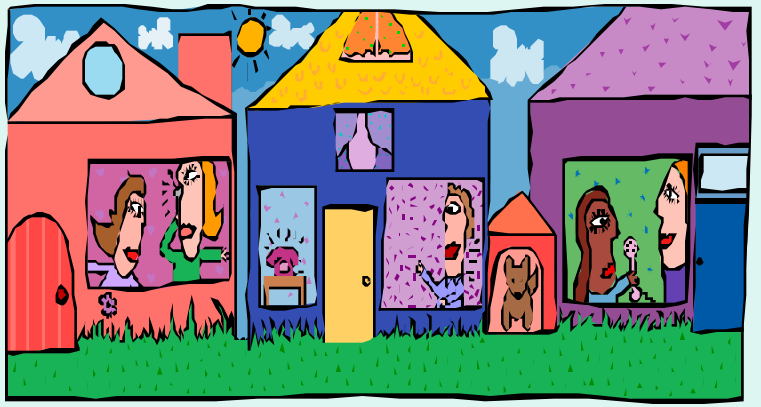
Epidemiology

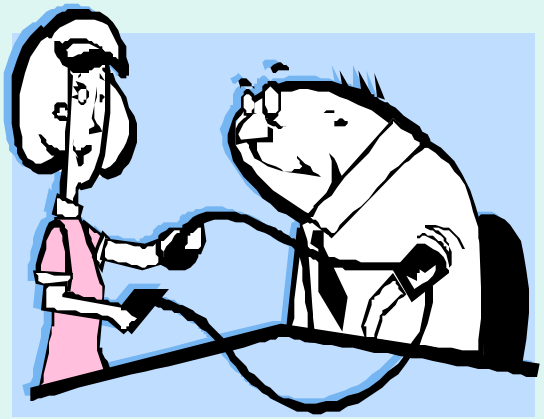


Scientific traditions vs. Community based participatory research (CBPR)



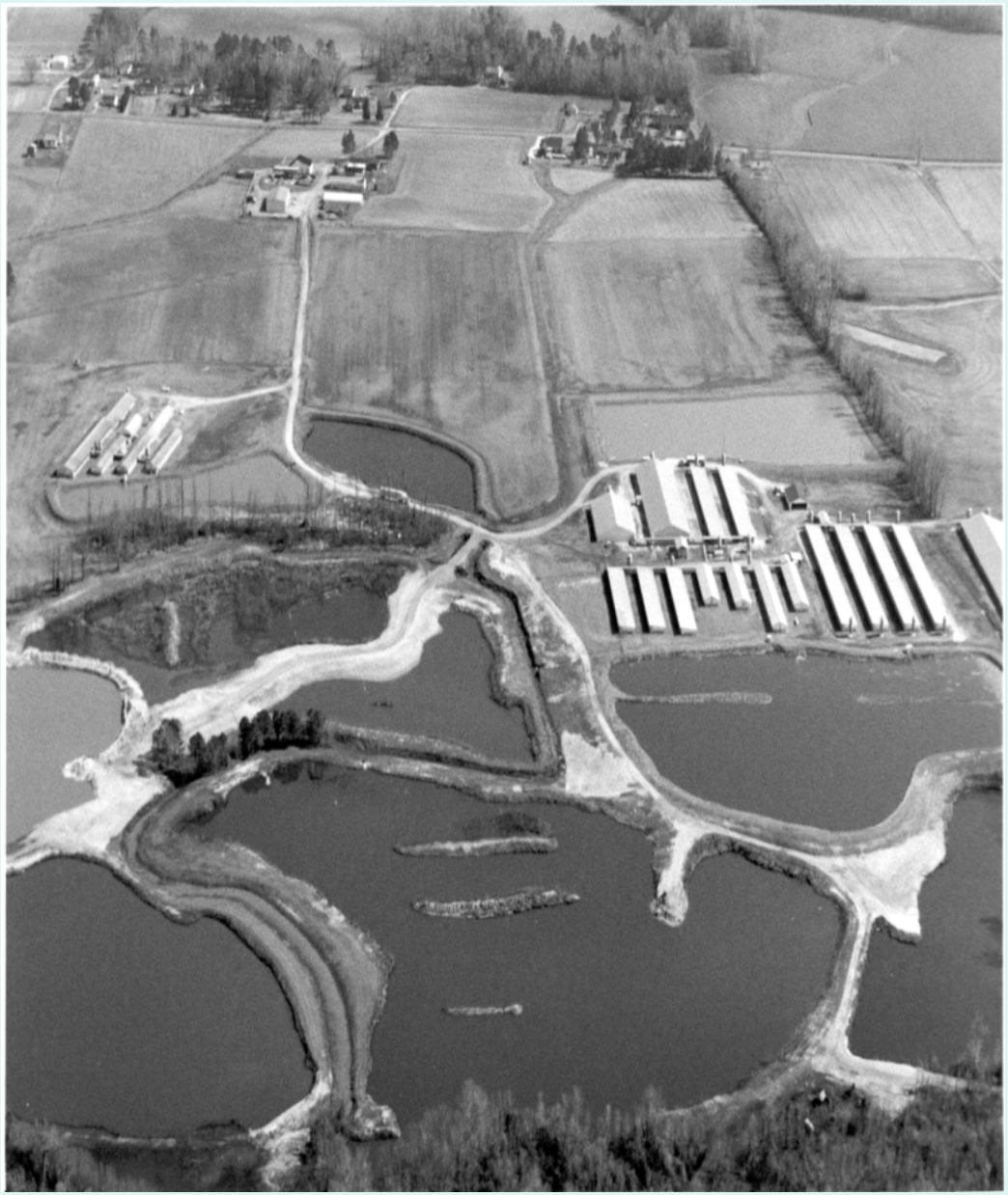
vs.

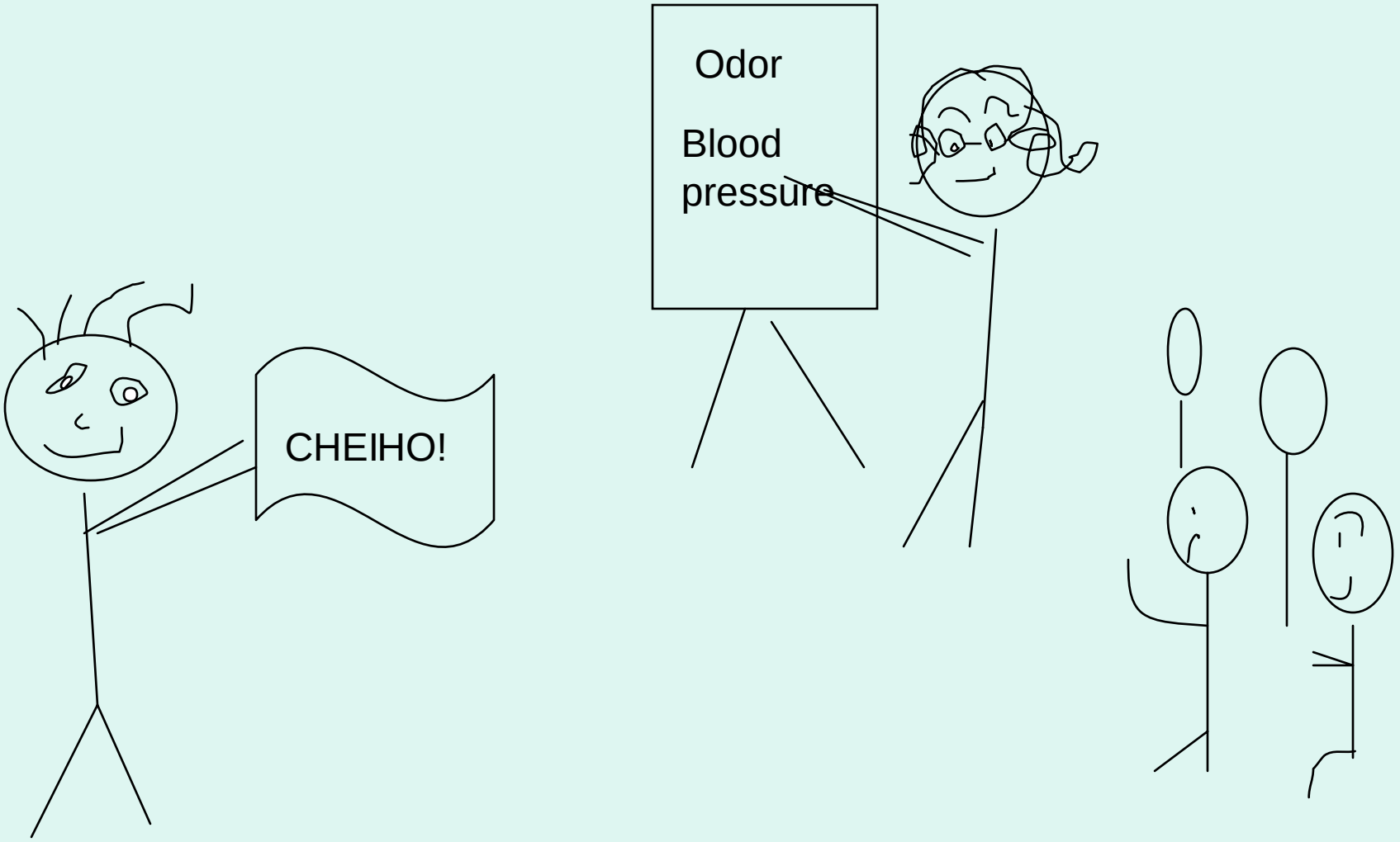


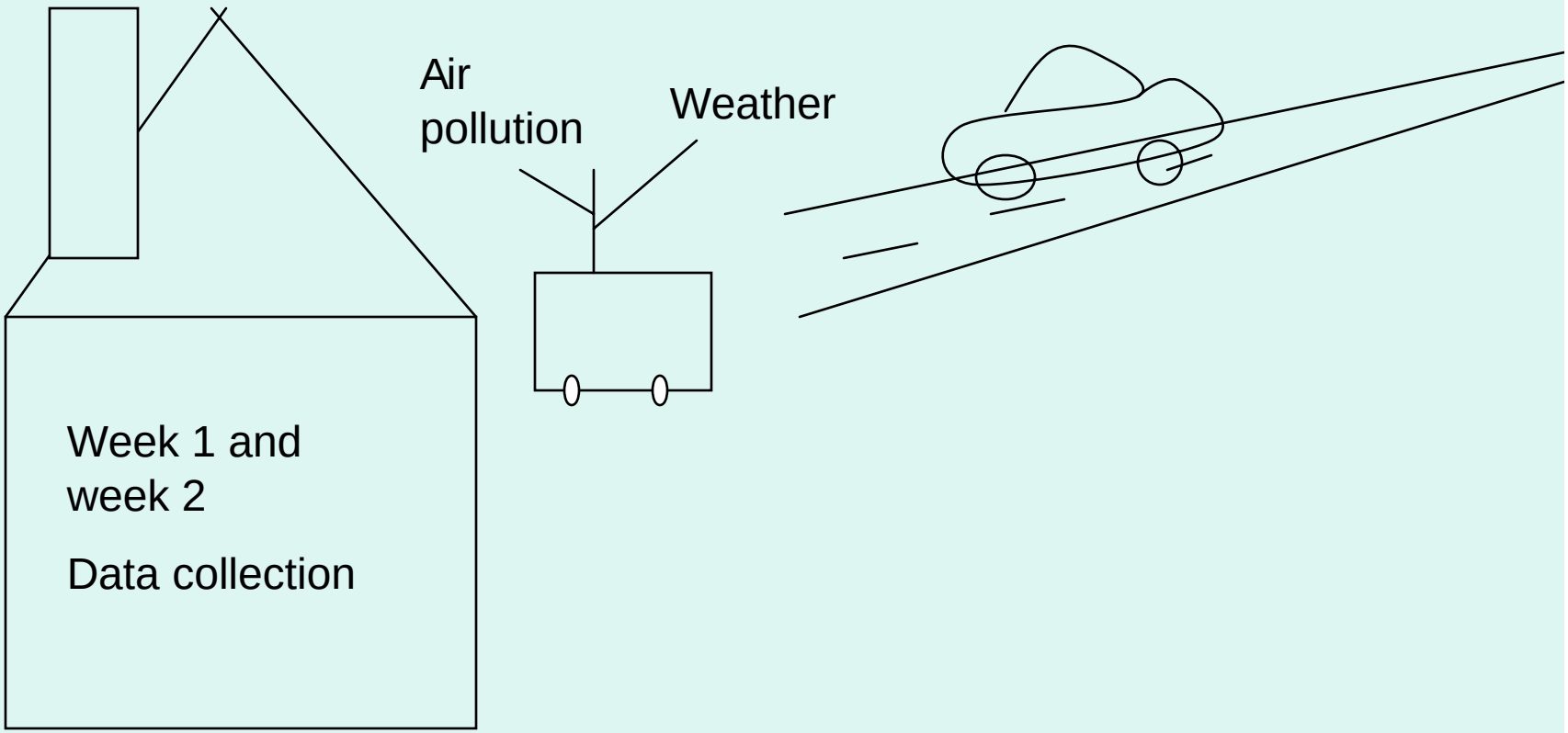


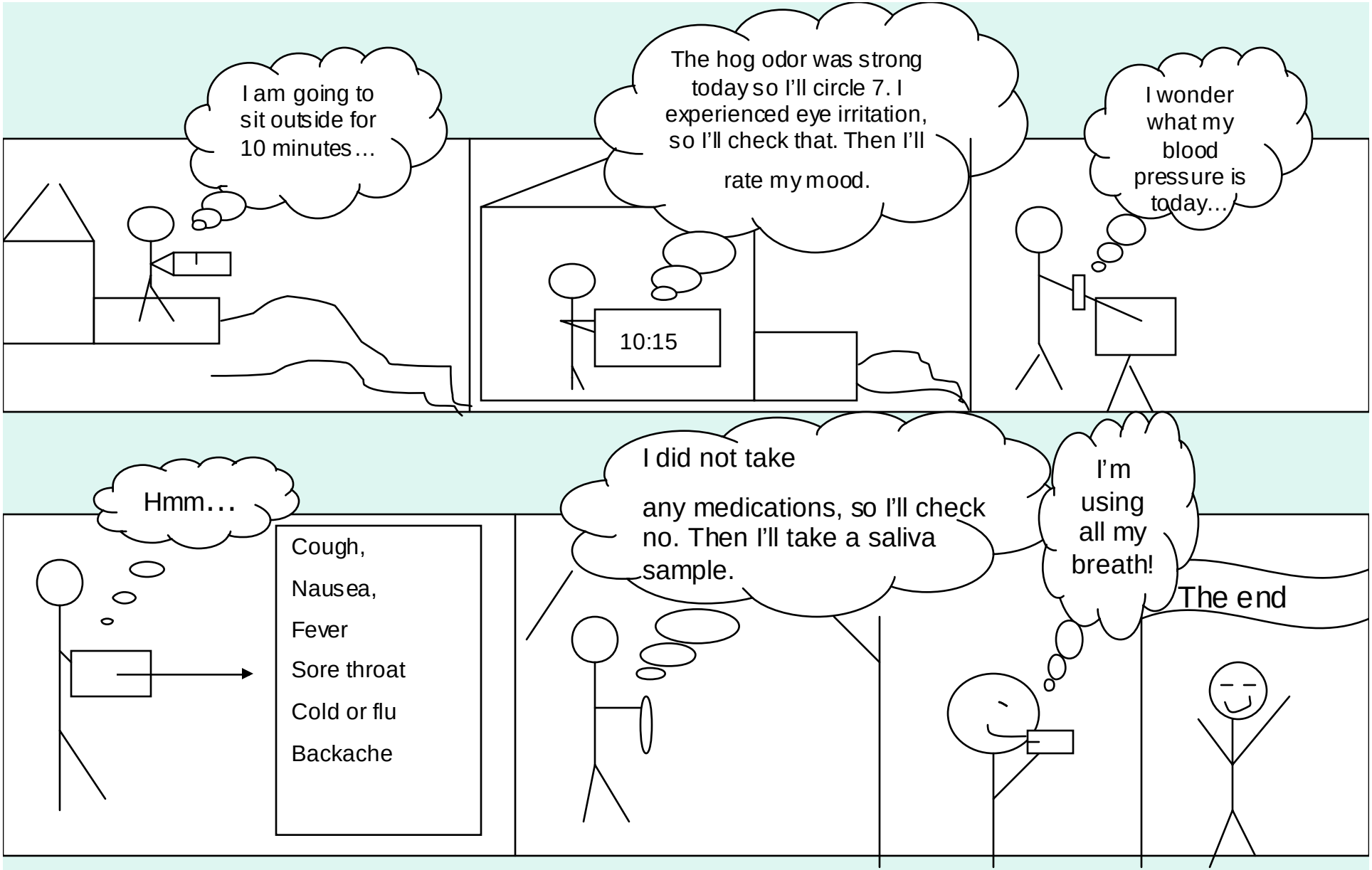
The Community Health Effects of Industrial Hog Operations Study

(CHEIHO)









A morning (or evening) in the life of a CHEIHO participant

Examination of data completeness and quality

Outcome variables

- Sequence error
  - If the participant recorded outcome information before they had exposed themselves to the outdoor air
  - Sequence error: 0=No sequence error 1=Sequence error
- Completeness of variables in journal records
  - 0=variable not missing, 1=variable missing
  - Summary variables based on activity or coding scheme
    - » Rating variables
    - » Response variables
    - » Machine use variables

Data analysis: Predictor variables

- Week in participation (Week 2 and beyond vs. week 1)
- Odor (Any odor vs. no odor)
  - Based on odor rating indicated in the journal record

Statistical analysis

- Estimated odds ratio estimates of association using hierarchical logistic regression models

- Considered a number of variables as potential effect measure modifiers, confounders

Results

CHEIHO

- Data collected between September 2003-September 2005
- 101 community members from 16 different neighborhoods in eastern, NC participated
- Each produced approximately 28 journal entries; 2 per day for 14 days
- 15 participants produced more than 28 records
- 2 participants left the study early
- In total, participants produced 2,949 journal records

Who participated

- 66 females, 35 males
- 86 African American, 15 Other
- Average age was 53, minimum age was 19, maximum age was 89

Percentage of journal records with incomplete variables or sequence errors

Error	Percentage of total records
Sequence error	2%
1 or more rating variable incomplete	12%
1 or more response variable incomplete	19%
1 or more machine use variable incomplete	26%
Mean and median percentage of records from which individual variables were missing	2%
Least frequently missing individual variable (nasal irritation variable)	1%
Most frequently missing individual variable (FEV-1 variable)	20%

Week-in-participation

Odds ratio estimate (95% Confidence Interval) Week 2 vs. week 1			
Sequence error	Response variables	Rating variables	Machine use variables
0.40 (0.13, 1.29)	1.17 (0.86, 1.59)	0.85 (0.62, 1.16)	0.41 (0.20, 0.84**)**

Odor

	Odds ratio (95% Confidence Interval) Any odor vs. no odor			
Sequence error	Response variables	Rating variables	Machine use variables
1.54 (0.74, 3.19)	1.11 (0.83, 1.48)	1.81 (1.22, 2.68)	1.18 (0.79, 1.78)

Conclusions

Summary of findings

- 98% retention rate
- Low percentages of records had incomplete variables and sequence errors
- Most incomplete variables were associated with use of machines, especially the Airwatch monitor

Summary of findings

- Lower relative odds that machine variables were incomplete in week 2 records compared to week 1
  - Might suggest the importance of mid-week check-ins, or of participant practice

Summary of findings

- Higher relative odds that circle variables were incomplete in records that participants produced during odor times, compared to non-odor times

- Potential implications regarding concerns about participant biases in data collection efforts

Conclusions

- Little work has been done to examine the quality of data that derives from CBPR studies
- Such work is useful
  - Responds to concerns about CBPR
  - Offers ideas for improving quality of data obtained from CBPR studies

Acknowledgements

- Steve Wing
- Rachel Avery
- Steve Hutton
- Amy Lowman
- NIEHS training grant
- Master’s paper support group
- Ghassan Hamra
- Community collaborators and participants