in this issue >>>

Issue Four

A Quarterly Insight on the Services our Division Offers

DEB quarterly

EPI CORNER

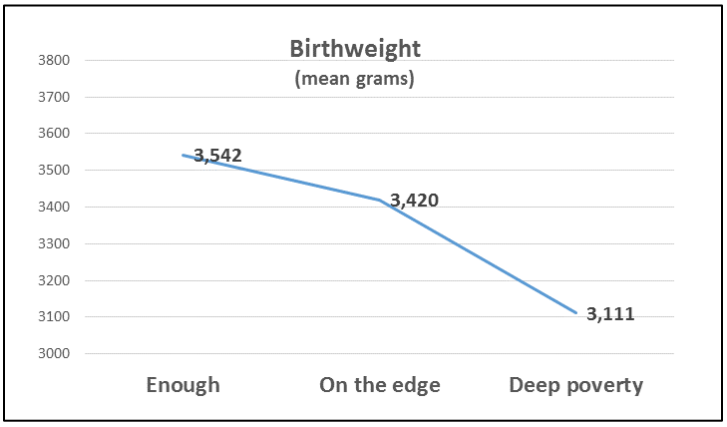
Capturing the Complexity of Socioeconomic Status

By Dr. Catherine Kothari, PhD Associate Professor

Socioeconomic status (SES) is a key component of the Social Determinants of Health framework; a framework that situates individual health within biological, environmental and socio-political structures and influences.[1] [2] Traditional SES measures are income, education and/or occupation.[3] Within health research, insurance status (often measured as private, public or no insurance) is a common, but unidimensional, method for operationalizing SES. A recent representative survey of 244 postpartum women (the Mom's Health Experiences Survey, Kothari PI) used multiple correspondence analysis (MCA) to create a contextualized SES measure. Six categorical variables were utilized, including: Insurance status, a 5-level household income measure and variables for transportation reliability, food sufficiency, housing stability and cohabitation/marital status. The resulting dimensions clustered into a three-level measure:

1. 44.7% (109) "Enough" – Adequate SES resources
2. 47.1% (115) "On the edge" – Low to medium income with safety net support
3. 7.8% (19) "Deep poverty" – Very low income and little support

These clusters were robust to subsequent psycho-social analyses, with stepwise statistical differences for: living in friendly neighborhoods, experiences of discrimination, self-reported stress and partner support (both practical and emotional). Finally, the graph below illustrates the association of these clusters with birth outcomes:



Creating a multi-dimensional SES metric produced a valid contextualized measure of a complex phenomenon. This measure not only has strong face validity, but provides a basis for developing targeted, risk-stratified interventions within medical and community settings.

[1]Link, B. G., & Phelan, J. (1995). Social conditions as fundamental causes of disease. *Journal of Health and Social Behavior,* 80.
[2]World Health Organization. (2014). Health status indicators. Retrieved from http://www.who.int/whosis/indicatordefinitions/en/
[3]National Center for Health Statistics. (2012). Multiple imputation of family income and personal earnings in the National Health Interview Survey: Methods and examples. Atlanta, GA: Centers for Disease Control and Prevention. Retrieved from http://www.cdc.gov/nchs/data/nhis/tecdoc11.pdf

Writing Accepted Abstracts

By: Dr. Laura Bauler, PhD

Writers Block
By: Dr. Laura Bauler, PhD

An abstract is a brief but complete summary of your work, providing enough information to convince the reader that the research/case is interesting, novel and relevant. You want to entice the readers to: 1) accept your work for inclusion into the conference, 2) read your manuscript or 3) visit your poster.

Identify your message: How does this work contribute knowledge to the field? How is it novel? What will others learn after hearing about your work?

Title: Should be interesting, informative, and accurate. Simple titles are best.

- **Descriptive:** Conveys essential info.
- **Interrogative:** Written as a question
- **Affirmative:** Summarizes the study results (not used for conferences).

Authors: All authors should contribute to the intellectual content of the abstract (not just be involved in the case or patient care).
Listed in order of contribution (First=Most work).

Three example titles for the same project:

Descriptive: Antigenic drift of Influenza A virus (H7N9) hemagglutinin.

Interrogative: How has the Influenza A virus H7N9 hemagglutinin changed over the past 5 years?

Affirmative: Increased mutational frequency in the Influenza A virus (H7N9) hemagglutinin protein over the past 5 years impacts vaccine effectiveness

Research Abstract:

- **Background/Introduction**: Big picture-problem you are studying.
- **Objective/Purpose/Rationale**: Specific question addressed, or topic discussed and why it is important for the reader to learn more about.
- **Materials and Methods:** Brief description of the methods used, include description of the patient population studied if applicable.
    - Ex. This was a prospective study utilizing the NCIS database, examining X and Y…
- **Results and Discussion:** Explain the findings of the study and put them into context for the reader. Present real data with numbers, frequencies, p-values, confidence intervals, odds ratios…
- **Conclusion:** Main finding (conclusion) of the study. Include future recommendations and/or important implications.

Case Report Abstract:

- **Introduction**: Big picture-problem of your case. Include a sentence about rationale for why this is an important/interesting/novel case.
- **Case Description:** Brief description of the case providing the essential details chronologically. It should include the history, physical examination, investigative studies (labs and tests), and the patients' progress and outcomes.
- **Discussion/Conclusions:** Short analysis of the importance of this case and what your reader should learn from this case.

Tips:

- Abstracts should be intelligible to a wide audience
- Be concise- long complex writing doesn't imply good science or medicine
- Omit needless words
- Use an active voice
- Avoid medical jargon and abbreviations
- Tailor the abstract to the application guidelines (For conferences: topic/goal of conference, structured vs unstructured, word limit, evaluation criteria, deadlines)

Need more help, join Laura Bauler, PhD our medical editor for a workshop on "Writing Award Winning Abstracts" on:
Monday January 14th at 12-1pm in Oakland Conference #3515
Wednesday January 24th at 12-1pm in Upjohn Classroom 112

Bulletproof Stats

Principal Component Analysis vs. Multiple Correspondence Analysis: Which is appropriate when?

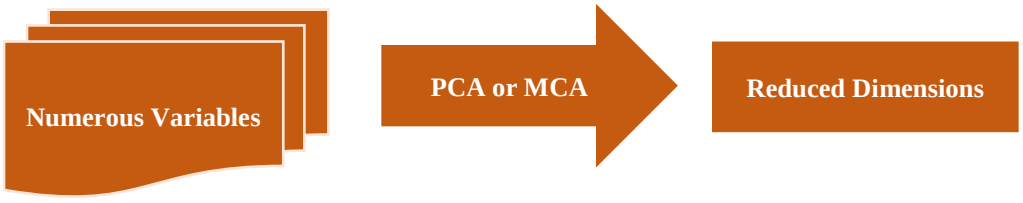
By: Alyssa Woodwyk, MS

Oftentimes when researchers have a large number of variables to consider for analysis, it is preferred to reduce the number of variables by transforming the data into a new space in which the information contained in the initial variables is preserved, prior to performing formal statistical comparisons. When we have quantitative variables, this can be done relatively simply using principal component analysis (PCA). However, what can one do when the data consist of numerous categorical variables? Given the frequent use of surveys in research, it is not surprising that one of my recent tasks was to find a way to reduce a large number of categorical variables.

PCA is appropriate only when we have all quantitative variables that need reducing. It uses eigenvalues, scalars corresponding to eigenvectors such that the eigenvalue linearly transforms the eigenvector in space, to determine the directions where input data have most variability. Using statistical software, PCA will generate a number of principal components, which are linearly uncorrelated variables, from the quantitative input variables. A typical rule of thumb used by statisticians is to take the first *x* principal components that account for at least 80% of variability in the data. Those *x* principal components can then be applied to each observation in the data. The idea here is that we will start with, for example, 10 quantitative variables as input to the PCA, and end up with a lesser number of principal components while still accounting for a sufficient amount of variability.

When we want to reduce categorical variables, multiple correspondence analysis (MCA) is more appropriate. MCA takes categorical variables as input, locates the categories in the Euclidean space, and assigns them distance values such that observations within the same category are closer than observations in different categories. The distance values are quantitative in nature, where values further from zero are more discriminative than values close to zero. After the distance values are assigned, the inertia (measure of variation each dimension explains) is considered in determining the optimal number of dimensions to utilize. The same rule of thumb applies where we will take the first *x* dimensions that account for at least 80% of the inertia. The result is a lesser number of quantitative dimensions compared to the number of input categories. MCA is useful, for example, in paring down a number of survey items analyzed. It provides numeric dimensions that can serve as input for quantitative analysis.

PCA and MCA are useful tools in reducing the number of variables for a data set. While PCA provides a well-known method for variable reduction, it is limited in use to only those quantitative in nature. When researchers find themselves in a situation where they wish to minimize a number of categorical variables, they can utilize MCA to generate quantitative dimensions for each subject. This enhances the utility of the data by expanding applicable statistical methods to those requiring quantitative data.



Q: *How do we contact you for project assistance?*

A: *That's easy, just reach out to us at epibio@med.wmich.edu or put in a request on our portal at Service Request Portal.*

REDCap Success Story – Early Introduction to Health Careers Pipeline Program

By: Anita Bell

DESCRIPTION

The Early Introduction to Health Careers pipeline program (EIH) exposes high school students to health and science related careers. Four major components work together to encourage students to become excited about learning science, as well as enhancing their ability to consider future careers in health and science. Core elements of the program include: science topic exploration, speakers from varied healthcare fields, ongoing mentorship with local medical students, and teaching basic research, with a student research presentation near the end of the program. The EIH program is delivered in half-day "Saturday Academies" once per month. Students practice critical thinking and problem-solving in an interactive learning format. The sessions include pre-and-post testing of the subject knowledge being presented.

CHALLENGE

The EIH team needed to capture student contact information, initial survey responses, and pre and post-activity quizzes. On-going data capture needs would be similar and last throughout the program, which is several months long. The team *could* have captured survey data manually, having the students fill out paper forms. This would have required that a team member distribute and collect paper forms, then enter data into a computer program such as Excel. This method of data collection could have been both time-consuming and error-prone, since there were many participants and multiple events planned over several months. Further, evaluation of the data could have been more difficult if hand-collected.

REDCap Solution?
See the next installment….

Data Bytes

final thoughts...

Secondary Data Research at WMed

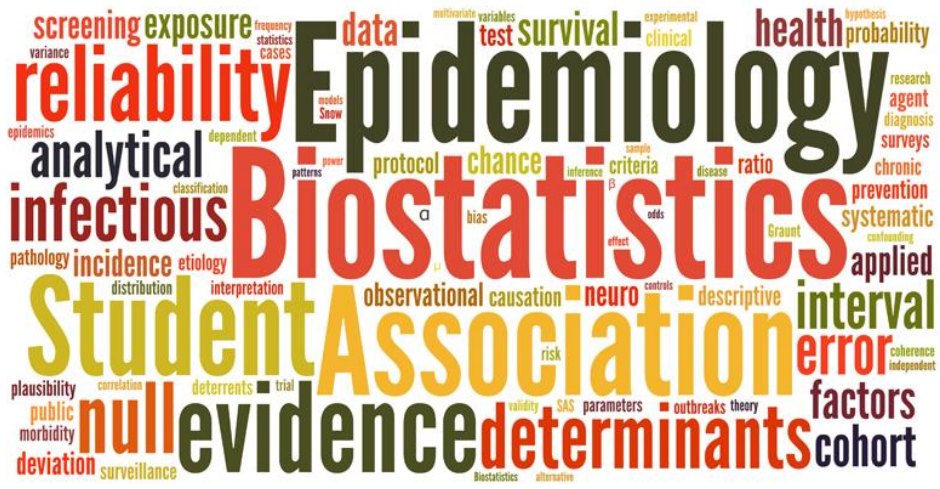
Good news! The Division of Epidemiology and Biostatistics has recently acquired new data sources from the HCUP family of databases (https://www.hcup-us.ahrq.gov/databases.jsp) thanks to a generous contribution of the School to assist in research. These complement the existing Nationwide Inpatient Sample (NIS), the largest claim database of inpatient care in the country. In Research Day 2018 presenters of session 1B showcased their projects based on NIS. The newly acquired databases correspond to the Kids' Inpatient Database (KID), the Nationwide Emergency Department Sample (NEDS), and the Nationwide Readmissions Database (NRD). As described by their names, these claim-based databases report on pediatric inpatient care, emergency visits, and readmission events in the US. These secondary sources of information are ideal for projects dealing with resource utilization and costs, epidemiology, practice patterns, care in rural America, rare conditions, and what your imagination can bring. To use these databases in your research, please contact us at epibio@med.wmich.edu or fill out the form at the EPIBIO Service Request Portal (link accessible at http://med.wmich.edu/node/88).

DEB

Epi & Bio
DIVISION

1000 Oakland Drive
Kalamazoo, MI 49008

Editor: Leah Bader
Contact us at 269-337-4609
Epi & Bio Website
epibio@med.wmich.edu



coming soon >>>

Important Upcoming Dates for 2019 Research Day:

Wednesday, January 23 - Deadline to submit exempt research studies to the IRB (irb@med.wmich.edu)

Wednesday, February 6 – Deadline for abstract submission at 11:30 p.m. (EST)

Friday, February 8 – Scoring begins

Friday, February 22– Complete judging of abstracts

Friday, March 1 – Authors notified of results

Friday, March 15 - Disclosures complete for all authors part of an oral presentation

Monday, April 1 – Poster submission deadline - MUST submit posters to researchday@med.wmich.edu (new process flow)

Monday, April 15 – Oral PowerPoint presentation deadline – MUST submit by 12:00pm (EST)

Tuesday, April 16 – 2019 Research Day – Poster Presentations (Schedule TBD)

Wednesday, April 17 – 2019 Research Day – Oral Presentations (Schedule TBD)

For Research Day related questions, please contact Leah Bader at:

researchday@med.wmich.edu