

Residency Research Manual

2010

for the Osteopathic Medical
Education Consortium of
Oklahoma (OMECCO)

Assembled by:

Office of Educational
Development

Information in this manual was adapted from the following sources:

Beal, James & Sahmoun, Abe (2009-2010) Conducting Practice-Based Research: Third-year Medical Student Clinical Research Project, Clinical Epidemiology Course Med 8101, University School of Medicine & Health Sciences, North Dakota

<http://www.med.und.edu/familymedicine/research/documents/ResearchManual2009-10.pdf>

Residency Research Manual, University of Kansas School of Medicine, Department of Family Medicine (January 2009),

<http://www2.kumc.edu/fammed/resident/documents/KUFamilyMedicineResearchManual.pdf>

Palmer, Joy (2008) Research Manual for Residents: A Guide to the Path Less Traveled, University of New England College of Osteopathic Medicine, Maine

<http://www.une.edu/com/residency/upload/researchmanual.pdf>

Department of Family & Community Medicine, Faculty of Medicine, University of Toronto (September 2005), Manual for Preparation of the Residents' Academic Projects,

http://dfcm.utoronto.ca/postgrad/general_information/residents/pdf/resprojmanual.pdf

TABLE OF CONTENTS

TOPIC	PAGE
OMEKO Research Mission	4
Why should you care about research?	4
Resident requirements	5
Residency training competencies	5
Learning objectives	6
Example Timeline for a Typical Three Year Residency	6
Research support	7
Choosing a topic	8
Study design	8
Types of Research Questions	8
Types of Research Studies	9
Study feasibility – Checklist for Determining Feasibility of Research Project	10
Information on Sample Size	11
Refining the research question	12
Developing the research proposal	12
Approval of research project	13
Preliminary Considerations	14
Software for Data Collection and Analysis	14
Conducting a pilot test and revising the research proposal	14
Analyzing the findings	14
Drawing conclusions	16
Disseminating results	17
Poster Presentation or Oral Presentation Using PowerPoint [®]	17
Resources	18

APPENDICES

Appendix	PAGE
I. How to calculate 95% CI on a proportion	19
II. How to read an SPSS printout	20
III. Null hypothesis	25
IV. Research Project Decision Tree	27

OMECCO Research Mission

To support the clinical research and scholarly activities of OMECCO members.

OMECCO and its affiliated member institutions support clinical research and scholarly activities through:

- Assisting in the advancement of knowledge including contributions to medicine through scientific research
- Creating an association with the specialty colleges research-related educational objectives for interns and residents
- Providing support for facilitating resident participation in research requirements and/or projects initiated by an osteopathic specialty college
- Identifying and listing research resources and opportunities available to interns and residents within OMECCO
- Providing ready access to basic science and/or clinical research mentorship and identifying statistical support services available to interns and residents
- Providing opportunities for continuing study for the development of faculty, investigators, and physicians
- Establishing policies and guidelines governing scientific research activities

Why Should You Care About Research?

Research is essential in *improving* the delivery of primary care to patients. Only with research can we identify best practices, assess our effectiveness in improving health, and better understand patient needs. Research also provides practitioners with tools to critique and question published findings and reports that can impact practice. Honing research skills is an essential element to improving the practice of medicine globally and personally.

Participating in research and scholarly projects assists the resident to gain skills needed to meet several training competencies, especially *Competency 5: Practice-based Learning and Improvement*. Research also contributes to other competencies including *Competency 1: Knowledge* as well as *Competency 2: Patient Care*.



re·search: NOUN: 1. a detailed study of a subject, especially in order to discover (new) information or reach a (new) understanding.

Cambridge Dictionaries Online, © Cambridge University Press 2003.

Resident Requirements

The following lists the requirements for conducting research for your residency program.

1. **Each resident chooses one faculty member to act as a research mentor** (Contact the OMECO for a list of mentors- see page 18 for contact information) and assist with hypothesis generation, literature searching, project design, human subjects approval, and project implementation
2. **All residents must complete an online training series.** The purpose of this series is to receive training about various research topics intended to help them design and complete their research projects.
3. **All residents must complete an institutional online human subjects tutorial** which certifies them for work on research projects.
4. **All residents, either individually or in paired research teams, must complete an original research project** at least one semester prior to their planned graduation date.
5. **All residents, either individually or in paired research teams, must present** their research findings in an academic forum such as a conference, peer reviewed journal, or professional meeting.

Residency Training Competencies

Clinical care excellence requires the physician to weigh medical evidence to determine best management for a given health problem. Research plays a crucial role in this process.

Typically, information about our clinic's patient population and the care provided is gathered and interpreted routinely. Residents help identify specific clinical questions and help specify a study design, assist in data collection and data analysis. Findings are compared with literature retrieved using information technology such as PubMed. Published studies are critically appraised, and the strengths and limitations of the information gathered for the resident's own study are evaluated. Recommendations pertinent to patient care practice are made in the scientific papers written or posters presented.

Learning Objectives

After completing the research component of the residency program, the resident will be able to:

1. Appraise the strengths and weaknesses of scientific evidence from studies relevant to patient care.
2. Describe the study design features needed to generate valid data.
3. Conduct a practice-based or community-based research project
4. Draw appropriate conclusions about potential ways to improve patient outcomes or population health.
5. Complete a well-written project report
6. Present findings from the project to an academic forum (i.e. peer-reviewed journal article, oral presentation, or poster presentation).

Every resident is capable of completing a research project that contributes to the improvement of patient care, safety, community or public health, discovery of new information, and advancing the science of medicine.

Example Timeline for a Typical Three Year Residency – Contact your residency coordinator an individualized timeline

All residents should choose their research topic during PGY-1, and no later than the beginning of their PGY-2. Early starts are encouraged. Data collection and analysis should be done during the second year, with **most of the work done no later than the end of PGY-2**. Presentation of research findings must be shared in an academic forum during their PGY-2 or PGY-3 year.

A model timeline to guide your planning is presented below. While the exact nature of the research project and other residency requirements may dictate deviations from this pace, the timeline provides a pattern of accomplishment that should be emulated.

Year 1: Identifying and Refining the Research Project

- July - Obtain research requirements, guidelines, and list of mentors and possible research topics
- Oct. - Identify research mentor and define a need for new or continued research
- Oct. – Develop the research question
- Nov. – Meet with research mentor and begin literature search
- Dec. – Complete the literature search and written project proposal (including background, objective, and method)
- Jan. – Provide project proposal to mentor for comments/feedback
- Feb. - Present proposal to research mentor & faculty/resident committee
- Apr. - Complete Institutional Review Board (IRB) submission to obtain Human Subject Committee approval unless exempted
- May – Begin project; collect data

Year 2: Collecting Data

- Aug. - Meet with research mentor to review progress of data collection
- Dec. - Complete data collection
- Jan. – Meet with research mentor to present data and preliminary analysis to clinical and research mentor
- **Feb. – Presentation at annual Resident Research Day**

Year 3: Analyzing Results, Writing the Manuscript and Presenting at Resident Research Forum

- Apr –Complete and submit written manuscript and/or poster and PowerPoint® presentation to research mentor and Residency office
- Sep. - Apr. - Submit journal article for review and/or present poster or oral presentation at a professional conference.

Residents Interested in Fellowships should also consider the following:

- Begin developing research ideas by the end of the first year.
- Complete a research project suitable for publication by the second year.
- Conduct and/or participate in at least one additional scholarly project.
- Submit an abstract to a national meeting.
- Present one or more research projects

Research Support -

Technical support available from your mentors, the Office of Research and the Office of Educational Development may include:

- Consultation on study design features (e.g. control of bias and confounding)
- Assessment of statistical power and sample size needed
- Survey design and administration
- Data entry tips and database development
- Advice on data collection and management
- Scientific writing assistance
- Procedures for protecting human subjects (e.g. confidentiality)
- Other assistance with IRB applications
- Data analysis using statistical software
- Assistance making graphs of data findings (e.g. using Excel or PowerPoint)
- Feedback on drafts of the research report, copyediting
- Assistance submitting abstracts to scientific meetings
- Help with journal article submission.

Please understand that our primary role is consulting. It is expected that the resident will be ultimately responsible for performing these tasks. Contact OMECO for more details- see page 18 for contact information.

Choosing a Topic

To help select a topic, think about:

- Science: What gaps exist in the medical literature?
- Clinical mentor: What topics or projects are being conducted by a Faculty member with whom you would like to work?
- Personal: What interests you? Have you had a patient with a problem for which the literature is not adequate? Are you interested in applying for a fellowship in a specific field?

Study Design

Here are example study designs from previous resident projects:

1. **Cross-sectional studies**, e.g., a survey.
2. **Retrospective cohort studies or meta analysis** conducted in our longitudinal patient records.
3. Assessment of value of **a new test** compared to a standard diagnostic procedure.
4. Determining the **inter-observer variability for a test** when read by different physicians or technicians.

Types of Research Questions

The research question will suggest the type of study that should be done. There are 3 broad types of research questions to consider:

Question Type	Description	Example
Questions of Description	Research questions that describe what is going on or what exists	What is the incidence rate for H1N1 influenza among Native Americans in the state of Oklahoma for 2009?
Questions of Relationship	Research questions that look at the relationship between 2 or more variables	Is there a relationship between burnout and suicide ideation among ER residents?
Questions of Comparison	Research questions that often seek to establish cause and effect relationships	Does a cognitive behavioral therapy program reduce self-reported burnout symptoms among primary care physicians?

Types of Research Studies

DESIGN DESCRIPTION	STRENGTHS	WEAKNESSES
<p>Prospective Cohort</p> <ol style="list-style-type: none"> investigator defines and selects a sample from a population measures predictive variable(s) measures outcome variable(s) at follow-up some defined time later. 	<ol style="list-style-type: none"> powerful for defining incidence and investigating potential causes of a condition allows measurement of predictive variable(s) in the moment, no reliance on reconstructing patient after outcome variable(s) occurs 	<ol style="list-style-type: none"> often need to follow populations for long periods of time to achieve numbers sufficient for causal relationship power, therefore expensive and inefficient for general population questions – better for specific populations (i.e., outcome of colon ca. surgery)
<p>Retrospective Cohort</p> <ol style="list-style-type: none"> identifies a population that has been assembled from past events collects data on predictive variable(s) as available from past <ol style="list-style-type: none"> records collects data on outcome variable(s) as collected from past or present events 	<ol style="list-style-type: none"> can establish that predictor variable(s) precede outcomes less costly and time consuming than prospective cohort 	<ol style="list-style-type: none"> investigator has limited control over the sampling of population and what data was collected, thereby what data is available to serve as the predictive variable(s)
<p>Cross - Sectional</p> <ol style="list-style-type: none"> all measurements made at once with no follow up investigator selects sample from population measures predictor and outcome variables 	<ol style="list-style-type: none"> great at describing variables and their distribution patterns fast, inexpensive, no subjects lost to follow up may generate questions for future studies 	<ol style="list-style-type: none"> difficulty of establishing causal relationships limited in information they can produce on prognosis or natural history
<p>Case - Controlled</p> <ol style="list-style-type: none"> often retrospective sample group from a population of people with the disease (cases) sample group from a population at risk that is free of the disease (controls) measurement of predictor variable(s) 	<ol style="list-style-type: none"> provides descriptive information on cases provides odds ratio high yield of information with small subject size makes for efficient results and ability to generate hypotheses good for rare conditions 	<ol style="list-style-type: none"> increased susceptibility to bias: separate sampling of cases and controls, retrospective management of predictor values no direct way to estimate incidence or prevalence of disease or excess risk
<p>Meta Analysis</p> <ol style="list-style-type: none"> Identifies previous research examining the topic of interest Combines results of separate research studies to derive an overall treatment effect Seeks to uncover moderators that may be related to the treatment effect 	<ol style="list-style-type: none"> Can summarize from available studies the effects of interventions across many patients. Can reveal research designs as moderators of study results Can determine if the effect of the intervention is sufficiently large in practical and statistical terms Can allow more objective assessment of evidence and thereby reduce disagreement Can clarify heterogeneity between study results. Can suggest promising research questions for future study. 	<ol style="list-style-type: none"> Can be compromised by inclusion of non-peer-reviewed data (especially data derived from biased sources) Can pass along inflated estimates of size effects based on research design characteristics, published vs. unpublished composition, and small sample sizes Reader judgments about study quality are subjective in nature.

Study Feasibility

CHECKLIST FOR DETERMINING THE FEASIBILITY OF A RESEARCH PROJECT

The following questions will guide your thinking about the feasibility of a research project. These are not the only questions you should ask. These questions are a beginning. You should review these questions by yourself and again with your research mentor.

A. LITERATURE AND MENTORS

- Do I have some ideas from reading research articles or clinical case studies?
- Do I have some ideas from my clinical experience?
- Do I have a mentor who is conducting research in an area of interest to me?
- Has there been an idea generated by a journal club discussion/presentation?

B. WHAT KIND OF RESEARCH SHOULD I DO?

- Epidemiologic study?
- A clinical trial? (Use an experimental intervention and a control group?)
- A case control study?
- Retrospective cohort?
- Cross sectional?
- An arm of an existing research project?
- Replication of a study?
- Basic science bench research?
- Other? _____

C. TIME AND RESOURCES?

- Do I have easy access to data or patients?
- Who are the available experts or mentors with the time and experience to guide and support my research?
- What will I need and expect from that person or persons?
- How long do I have, or want, to take to conduct this research?
- Other: _____

D. DATA ACCESS

- Am I generating pilot data?
- Are pilot data available that can guide my research design development?
- Is there existing data that can be analyzed in a different way?
- Is there existing data in an electronic form that I can analyze such as a national survey or other national or local data base?
- Other _____

E. END POINTS AND PRODUCTS

- When will I publish or present the findings (Research Appreciation Day)?
- What is the benefit or value to the field of psychiatry, to our existing system of care, or to health care consumers?
- Is there a cost-benefit issue to address for this research project?
- What do I expect of myself during and after the project?
- Other _____

F. OTHER RESOURCES

- Do I have a biostatistician available to consult on data management and analysis?
- Am I trained in Human Subjects protection issues, IRB procedures, and/or have access to a clinical research coordinator to assist in data collection, and study reports/reviews?

- What expertise do I need to access?
- Do I need funding to conduct this research project?
- Other _____

G. MY OWN SKILLS

- What other knowledge do I need to complete this research project?
- Do I have or can I acquire the basic skill and knowledge to accomplish my goals for this research project?
- Am I committed to doing this research project?
- Other _____

Information on Sample Size

The research mentor will help you consider:

- Sample size: How many patients are needed for adequate statistical power?
How many patients are available in our practice?

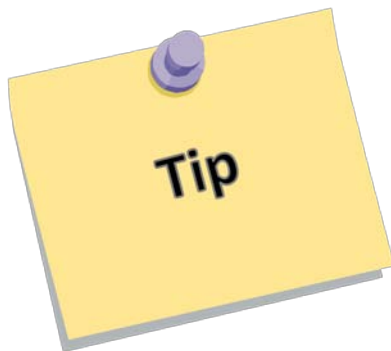
**SIZE OF SAMPLE NEEDED TO YIELD ERROR OF
5% OR LESS IN 95 OUT OF 100 SAMPLES**

Population Size	Obtained Sample Size	Percent of Population
50	45	90
100	80	80
200	132	66
300	168	56
400	196	49
500	218	43
600	234	39
800	240	32
1000	278	27
2000	323	16
3000	341	11
5000	357	7
6000	361	6
8000	367	6
10000	370	3
15000	375	2
20000	277	1.8
30000	379	1.2
50000	381	0.7
75000	382	0.4
100000	383	0.03

Source: Michigan State University, OMERAD, Medical Education Section

- Time: How much time will it take to collect the data?
- Costs: Are there special costs other than investigator time and routine copying?

The resident should discuss these types of issues with their research faculty mentors. A focused approach that answers a narrow question is better than a broad investigation of a larger topic with insufficient resources to complete the project.



A modest preliminary study can lay the groundwork for a future large study.

For large projects, separate the study into Phase I and Phase II. You do Phase I. Team up with a resident who will graduate a year or two after you for Phase II.

Collaborating with another institution to generate enough sample size is acceptable as long as you closely involve a clinical faculty member in our Department, your data is analyzed here, and all University and Departmental policies are followed.

Refining the Research Question

After choosing a general topic area, the resident should:

1. Conduct a literature search to evaluate the current research in the area
2. Talk with faculty about their interests and current projects
3. Decide what specific project in which he or she is interested

After the general question has been decided upon, it should be narrowed down into a specific answerable research question or hypothesis. For example:

General question: What factors influence whether patients stop smoking?

More specific research questions: Is the use of Chantix or Zyban more effective among low-income patients who are attempting to quit smoking? What indicators can be established among patients that smoke that would provide guidance for prescribing them the most effective drug?

Developing the Research Proposal

Resident should prepare a written research proposal in collaboration with their mentors. The proposal should provide a concise review of the relevant scientific literature pertaining to the research topic. The research plan section of the proposal should describe in adequate detail the question to be addressed, the study population(s) involved, the methods to be used and the analytic plan. It should include:

Title

Abstract (300 words)

Specific Aims (list of main deliverables including number of study subjects)

Background (introduction and literature review)

Methods:

Subject recruitment and selection

Data collection (attach data collection forms)

Statistical methods:

- Data management (spreadsheet software, confidentiality procedures)
- Analysis methods to be used (include control of confounders)
- Statistical tests to be applied
- Sample size and power estimates (for the major study outcomes)

References

Biosketches of investigators (refer to mentors for examples)



Use what you write for your research plan's introduction, methods and literature cited in your final paper or poster.

Approval of Research Project

After the research proposal has been developed, it should be approved by the resident's research faculty mentor. Depending on the topic or scope of the project, an additional meeting with other faculty and/or the residency director may be required.

NOTE: Institutional Review Board (IRB) approval is needed before the resident initiates the research activity.

Oklahoma State University Center for Health Sciences (OSU-CHS) requires justifying the participation of human subjects in research, and promotes the protection of the welfare, rights and privacy of those subjects. The resident (and co-investigators) need to complete the OSU-CHS on-line human subjects tutorial and other institutional requirements well in advance of the planned research project. In some cases, clearance both by the OSU-CHS and another IRB is needed, which is another reason for residents to start their projects early.

See Appendix IV for a decision tree

Guidelines for submitting a study to the Oklahoma State University Center for Health Sciences Institutional Review Board can be found at:

<http://www.healthsciences.okstate.edu/research/rsp/irb.cfm>

If you are submitting new research, you will need to complete the *Biomedical Research Investigators* modules BEFORE you may begin an approved study, scoring an overall 90% or better on the modules. To complete these online modules: [Go to www.citiprogram.org](http://www.citiprogram.org) to register for CITI online training. Click on "New Users Register Here". Click on "The Protection of Human Research Subjects". Under "Select your institution or organization" page select "Oklahoma State University Center for Health

Sciences" in the "Participating Institutions" drop down box. Create your own username and password and select the CITI Recommended Learner group; Biomedical Research Investigators. Then begin the modules.



IRB stands for "Institutional Review Board". The Scientific Review Committee (SRC) is designated as a subcommittee of the Institutional Review Board (IRB) for the Oklahoma State University Center for Health Sciences. The SRC is responsible for reviewing and making recommendations regarding a research project. The IRB is responsible for approving, modifying, rejecting and monitoring research involving human subjects.

Preliminary Considerations

Software Used for Data Collection and Analysis

If data are entered into Excel or Access, they can easily be exported into an SPSS[®] (Statistical Package for the Social Science, Chicago, IL) if the Excel[®] or Access[®] database was set up in the most efficient way. If not, exporting can be very time-consuming and prone to keying errors. Please consult with the research staff during design and planning to help minimize errors or need for time-consuming recoding or re-entry of data later on .

Conducting a Small Pilot Test and Revising the Research Proposal –

Once departmental and HSC approval have been obtained, the study may begin. It is useful to test the study plan on a small scale to identify process problems before the larger study is conducted. If a medical record audit is conducted, this pre-test may suggest additional variables to collect. If a survey is administered, unclear questions may need to be revised. If significant changes are needed to the research protocol, it is necessary to apply for a revision through the HSC.

Analyzing the Findings

Below is a highly simplified flowchart of some common statistical analysis procedures. An appendix shows how to read SPSS output, and provides the formula for calculating 95% confidence intervals (CI) for a proportion.

1. Descriptive Statistics (one variable at a time):

Categorical variables

What are they? Items coded as Yes/No or mild/moderate/severe

What do we calculate? **Frequencies, proportions, rates, 95% CI**

Continuous variables

What are they? Items like age, weight, systolic blood pressure

What do we calculate? **Means, standard deviations, medians**

2. Analytic Statistics for Two Variables

Categorical variables:

Degree of control of diabetes (excellent, fair, poor) among those with controlled and uncontrolled hypertension

Is preeclampsia more likely among teen moms?

If you have a *large sample*, generate a *P* value using the **Chi-Square test**

If you have *small numbers*, use **Fisher's Exact test**

Continuous variables: Is hypertension more common in smokers?

Is the variable *skewed*?

No (bell-shaped curve): Use **independent samples t-test**

Yes (a few non-smokers, a lot of smokers): Use a nonparametric test, e.g.

Mann-Whitney test

3. Analytic Statistics for More than Two Variables

Categorical variables

Same as for two variables.

Continuous variables

Is the variable **skewed**?

No: Use **ANOVA** (analysis of variance)

Yes: Use a nonparametric test, e.g. **Kruskal-Wallis test**

4. Controlling for Confounding Variables

Poor control of confounding can ruin a study. An incorrect conclusion is reached because the true cause of the health outcome was something other than what the data showed.

Suppose the data shows that patients who had a prior attempt to quit smoking did better with Chantix than patients who had never tried to quit (i.e., more successful quits occurred with the prior attempt to quit group, $P < .001$). But the patients who had never tried to quit were more likely to be depressed.

The patient's mental health is a confounder, and could be the true reason why the patients with prior attempts did better – it may have had little to do with Chantix.



Be sure you collect enough information about your study population. Then you will have the data to describe key demographic and medical characteristics of the group you investigated. You will also be able to control for key confounders. The audience expects this information during your presentation and in your paper.

There are three main ways to control for confounding.

Subject Eligibility Criteria:

Only let mentally healthy/non-depressed patients into the study.

Problem: You won't find out about depressed patients' needs or their response to Chantix or if depression is prevalent in your patient population, you may not have enough non-depressed patients to do the study.

Stratified Analysis:

Analyze results for depressed patients separately from non-depressed patients.

Problem: There is always more than one confounder (e.g., obesity). So you'd have to subdivide first by mental health/depression status, then by weight. Now you have four groups instead of only two. If small sample size, there will not be enough power to do many subgroup analyses.

Regression Analysis:

Calculate an odds ratio for the adjusted for the multiple risk factors (e.g. depressed and non-depressed, BMI \leq 30, age, gender, number of quit attempts, prior smoking cessation medications).

Problem: Need statistical software. Although a regression controls for several risk factors at a time, the number of confounders that can be put in a model is limited if sample size (number of patients in the study) is fairly small.

Drawing Conclusions

The conclusions of a research project should include:

1. What did the results show?
2. Were the results statistically significant?
3. What were the study's limitations?
4. Were the results clinically important?
5. What future research is needed?

Disseminating Results

Residents must prepare a written report summarizing their findings. This written report should be reviewed and approved by the faculty mentor and coauthors before it is submitted to the residency program director.

Resident papers are typically 7-10 single-spaced pages for the introduction, methods, results and conclusions.

Place the tables and figures at the **end of the manuscript** (after the references) instead of inserting them in the body of the results section. If your tables, figures or PowerPoint are giving you a hard time, call the research team.

Poster Presentation or Oral Presentation Using PowerPoint®

At a minimum, you should present your research in the form of a poster or oral presentation. For more information on creating a poster presentation or powerpoint presentation and to get tips/examples for presenting both a poster or oral presentation, visit the OMECO or CHS Faculty Development online classroom:

URL: <http://oc.okstate.edu>

USERNAME: omeco

PASSWORD: consortium

Click on the CHS OMECO link or the CHS Faculty Development link under communities and then click content to view information. If you experience problems with either of these sites, please contact Dr. Machele Davison at 918-561-5712 or machele.davison@okstate.edu.

NOTE: An electronic version of this manual and additional readings can be found under Content within the OMECO online classroom.

Resources

Studies designed and executed by our faculty members and residents have passed the test of peer review, resulting in numerous recent journal articles and abstract presentations at local, regional and national meetings. You can consult the following offices, faculty members, or staff to help as you begin your research endeavors:

OMEKO MAIN OFFICE
Jeffrey LeBoeuf, CAE 717 South Houston Avenue, Suite 301 Tulsa, OK 74127 918-586-4626
OMEKO Research Subcommittee:
Johnny Stephens, PharmD., Chair Christopher Thurman, D.O. David Hogan, D.O. Stanley Grogg, D.O. David Wallace, PhD Greg Martens, D.O. Richard Schooler, D.O.
OFFICE OF RESEARCH and IRB- OSU-CHS:
Bavette Miller, M.S. Director of Research Operations and Graduate Studies 918-586-4601 bavette.miller@okstate.edu
OFFICE OF EDUCATIONAL DEVELOPMENT- OSU-CHS:
Matt Vassar, Ph.D. OSU-CHS, Room 238 Tulsa, OK 74107 918-561-8492

Appendix I

How to calculate 95% CI on a proportion

Formula: $p \pm Z_{1-\alpha/2} \sqrt{p(1-p)/n}$

Proportion = p. For 37.8%, p = .378

Z has to do with the area under the bell shaped curve.

$Z_{1-\alpha/2} = 1.96$ when you want the typical 95% CI.

Example: In 37.8% of 185 cases, antibiotic guidelines were followed completely
“*” signifies multiplication in example below.

Lower 95% CI: Subtract the quantity from p

$$p - 1.96 * \text{square root of } [p * (1-p)/n]$$

$$0.378 - 1.96 * \text{square root of } [0.378 * 0.622/185]$$

$$0.378 - 1.96 * \text{square root of } 0.00127$$

$$0.378 - 1.96 * 0.0356$$

$$0.378 - 0.070$$

$$0.308 \text{ or } 30.8\%$$

Upper 95% CI: Add the quantity to p

$$P + 1.96 * \text{square root of } [p * (1-p)/n]$$

$$0.378 + 0.070$$

$$0.448 \text{ or } 44.8\%$$

Appendix II

How to Read an SPSS Printout

A. Frequencies: Percent and Valid Percent

Example I: Twin Deliveries at KU Hospital from 1998-2004. What was the race/ethnicity of the mothers of twins? SPSS gives you information to express the results in two ways, either of which is correct.

Output 1. Using the “Percent” column: You would say “the race/ethnicity of the moms was: White, 78.2%; Hispanic, 8.3%; Black, 7.3%; Other, 0.9%; **Unknown, 5.3%.**” These add up to 100%.

Output 1. Mother's race

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	White	506	78.2	82.5	82.5
	Hispanic	54	8.3	8.8	91.4
	Other	6	.9	1.0	92.3
	Black	47	7.3	7.7	100.0
	Total	613	94.7	100.0	
Missing	System	34	5.3		
Total		647	100.0		

Output 2. Using the “Valid Percent” column: You could also say “Among women whose race/ethnicity was reported, the distribution was: White, 82.5%; Hispanic, 8.8%; Black, 7.7%; Other, 1.0%.” These add up to 100%. You ignore the “Percent” column.

Output 2. Mother's race

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	White	506	78.2	82.5	82.5
	Hispanic	54	8.3	8.8	91.4
	Other	6	.9	1.0	92.3
	Black	47	7.3	7.7	100.0
	Total	613	94.7	100.0	
Missing	System	34	5.3		
Total		647	100.0		

Example II: Twin Deliveries at KU Hospital during 1998-2004. What instrument was used in these operative vaginal deliveries?

Output 3. This time, it is not correct to use the “Percent” column. That’s because there are missing values for moms who did not have an operative vaginal delivery because, for example, they had a cesarean. So use the “Valid Percent” column.

Output 3. Instrument used in AVD

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Vacuum	18	2.8	58.1	58.1
	Forceps	12	1.9	38.7	96.8
	Both	1	.2	3.2	100.0
	Total	31	4.8	100.0	
Missing	System	616	95.2		
Total		647	100.0		

B. Continuous Variables: Checking for Skewness.

A skewness test is an option that can be chosen when calculating means. If the skewness test says the sample is skewed, a non-parametric test should be used to calculate a P value, instead of the standard tests. The sample is skewed if “Skewness” is > 1.0 or is < negative 1.0.

Output 4. The BMI distribution is skewed at both KU Family Medicine clinic and JayDoc as shown below.

Output 4. BMI Pre-pregnancy

	Median	Mean	SD	Skewness	N
JayDoc	24.66	26.24	6.74	1.271	20039
Clinic	23.06	24.17	4.99	1.213	13061
Total	23.95	25.42	6.19	1.369	33100

C. Comparing Proportions: The Chi-Square or Fisher's Exact Test .

Output 5. The table on the next page shows how many cases, the row percents and the column percents for a group of women with preeclampsia who had Cesareans.

For MgSO₄, **0 means no and 1 means yes.** 72.1% of women with primary C/S received MgSO₄, compared to 64.0% of women with repeat C/S.

Output 5. Mode of delivery if C/S * MGSO4_GIVEN Crosstabulation

			MGSO4_GIVEN		Total
			0	1	
Mode of delivery if C/S	Primary C/S	Count	159	410	569
		% within Mode of delivery if C/S	27.9%	72.1%	100.0%
		% within MGSO4_GIVEN	66.0%	73.7%	71.4%
	Repeat C/S	Count	82	146	228
		% within Mode of delivery if C/S	36.0%	64.0%	100.0%
		% within MGSO4_GIVEN	34.0%	26.3%	28.6%
Total	Count	241	556	797	
	% within Mode of delivery if C/S	30.2%	69.8%	100.0%	
	% within MGSO4_GIVEN	100.0%	100.0%	100.0%	

Output 6. The table below shows the P values for comparing the proportions, 72.1% v s 64.0%. In SPSS-speak, "**Asymp Sig.**" means **P value**. The table gives us a lot of statistics. We want to use the **Pearson Chi-Square** unless any cells had expected counts less than 5. If so, we would use the Fisher's Exact test. For the table below, we would use the P value of **.026**.

Output 6. Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	4.965(b)	1	.026		
Continuity Correction(a)	4.592	1	.032		
Likelihood Ratio	4.874	1	.027		
Fisher's Exact Test				.027	.017
Linear-by-Linear Association	4.958	1	.026		
N of Valid Cases	797				

a Computed only for a 2x2 table

b **0 cells (.0%) have expected count less than 5.** The minimum expected count is 68.94.

D. Logistic Regression .

Output 7. This is a logistic regression to see how much of the risk of **shoulder dystocia** may be due to **diabetes** in our population. We are using the regression to control for other confounding influences, such as the mother's BMI.

The part of the logistic regression print out in the table below shows how many cases were included in the regression. If a case has missing data for any of the variables in the regression, it was excluded from the analysis. This regression only had 2.9% cases with missing data.

Output 7. Case Processing Summary

Unweighted Cases(a)		N	Percent
Selected Cases	Included in Analysis	4694	97.1
	Missing Cases	138	2.9
	Total	4832	100.0
Unselected Cases		0	.0
Total		4832	100.0

a If weight is in effect, see classification table for the total number of cases.

Output 8. The following table tells us how well our variables help to predict whether the patient had shoulder dystocia. The Cox & Snell R square value tells us that the regression model explained 7% of the variance in the data. Since that's not very much, it means we don't have information on some of the other important risk factors. Or maybe many of the true risk factors have not been discovered.

Output 8. Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1526.705(a)	.070	.212
2	1526.928(a)	.070	.212
3	1527.150(a)	.070	.212

a Estimation terminated at iteration number 9 because parameter estimates changed by less than .001.

Output 9. Results are shown on the next page. Unfortunately, they come in SPSS-speak. On the table below, the text originally in the SPSS print out is in the smaller, non-bold font and the translation that has been inserted is in the larger, bold font.

➤ Sig means P value.

- Exp(B) means adjusted odds ratio
- Upper/Lower are the 95% CI for the adjusted odds ratio.
- The brief variable name of each potential risk factor is listed instead of what it stands for, but there is another page that provides the translation.
- How the variables were coded determines what was the “reference value” to which the other levels of the variable were compared. For example, each level of parity was compared to ≥ 3 .

Findings: Diabetes was independently associated with a 22.6% increase in the risk of shoulder dystocia (Adj. OR 1.226), but the association was not statistically significant (P=.631).

Output 9. Variables in the Equation

		B	Sig. P Value	Exp(B) Adjusted OR	95.0% C.I. for EXP(B)	
					Lower	Upper
Step 2(a)	labor2 (vs. spontaneous)		.002			
	labor2(1) Augmented	.536	.004	1.709	1.190	2.454
	labor2(2) Induced	.603	.001	1.828	1.289	2.594
	delmode1(1)AVD (vs. CS)	4.259	.000	70.742	26.055	192.069
	paratot4 (compared to ≥ 3)		.000			
	paratot4(1) 0	-.157	.641	.854	.441	1.655
	paratot4(2) 1	.333	.341	1.395	.703	2.768
	paratot4(3) 2	.647	.083	1.909	.919	3.967
	bmi_pcat1a (vs. lean BMI < 18.5)		.052			
	bmi_pcat1a(1) ≥ 30	1.033	.014	2.810	1.228	6.428
	bmi_pcat1a(2) 25-29.9	.906	.028	2.475	1.102	5.558
	bmi_pcat1a(3) 18.5-24.9	.704	.079	2.022	.921	4.437
	fet_intol_any4_var(1)	.638	.000	1.893	1.415	2.531
	Diabetes(1)	.204	.631	1.226	.534	2.814
	Constant	-8.717	.000	.000		

a Variable(s) entered on step 1: year, labor2, delmode1, paratot4, bmi_pcat1a, fail_progress_avd_cs_reason, fet_intol_any4_var, Diabetes.

Note: the rows for “year” on the table above were deleted, and several columns were deleted to make the table simpler and smaller.

Appendix III

Null Hypothesis

The null hypothesis is a hypothesis which the researcher tries to disprove, reject or nullify. The 'null' often refers to the common view of something, while the alternative hypothesis is what the researcher really think is the cause of a phenomenon.

The null hypothesis, H_0 , is an essential part of any research design, and is always tested, even indirectly. The simplistic definition of the null is as the opposite of the alternative hypothesis, H_1 , although the principle is a little more complex than that.

An experiment conclusion always refers to the null, rejecting or accepting H_0 rather than H_1 . Despite this, many researchers neglect the null hypothesis, which is poor practice and can have adverse effects.

DEVELOPMENT OF THE NULL

Up until the 1500's most people thought that the world was flat (At the time: The Null hypothesis). Columbus challenged this idea with the alternative hypothesis: The world is round. Then most people thought that the earth was the center of the universe (Null hypothesis). Copernicus had an alternative hypothesis that the world actually circled around the sun, thus being the center of the universe. Eventually, people got convinced and accepted it as the null.

Later someone proposed an alternative hypothesis that the sun itself also circled around the something within the galaxy. This is how research works - the null hypothesis get's closer to the reality each time, even if it isn't correct, it is better than the last null hypothesis.

EXAMPLES OF THE NULL HYPOTHESIS

A researcher may postulate a hypothesis:

H1: Tomato plants exhibit a higher rate of growth when planted in compost rather than in soil.

And a null hypothesis

H_0 : Tomato plants do not exhibit a higher rate of growth when planted in compost rather than soil.

It is important to carefully select the wording of the null, and ensure that it is as specific as possible. For example, the researcher might postulate a null hypothesis:

H_0 : Tomato plants show no difference in growth rates when planted in compost rather than soil.

There is a major flaw with this null hypothesis. If the plants actually grow more slowly in compost than in soil, an impasse is reached. H_1 is not supported, but neither is H_0 , because there is a difference in growth rates.

If the null is rejected, with no alternative, the experiment may be invalid. This is the reason why science uses a battery of deductive and inductive processes to ensure that there are no flaws in the hypotheses. Many scientists neglect the null, assuming that it is merely the opposite of the alternative, but it is good practice to spend a little time creating a sound hypothesis. It is not possible to change any hypothesis retrospectively, including H_0 .

SIGNIFICANCE TESTS

If significance tests generate 95% or 99% likelihood that the results do not fit the null hypothesis, then it is rejected, in favor of the alternative. Otherwise, the null is accepted. These are the only correct assumptions, and it is incorrect to reject, or accept, H_1 .

Accepting the null hypothesis does not mean that it is true. It is still a hypothesis, and must conform to the principle of falsifiability, in the same way that rejecting the null does not prove the alternative.

PERCEIVED PROBLEMS WITH THE NULL

The major problem with the null hypothesis is that many researchers, and reviewers, see accepting the null as a failure of the experiment. This is very poor science, as accepting or rejecting any hypothesis is a positive result. Even if the null is not refuted, the world of science has learned something new. Strictly speaking, the term 'failure', should only apply to errors in the experimental design, or incorrect initial assumptions.

Appendix IV

Decision Tree

