Manuscript Writing: “Publish or Perish” and “If it Wasn’t Published it Didn’t Happen”

Catherine A. McCarty, PhD, MPH
Senior Research Scientist
Center for Human Genetics
Learning Objectives

- Plan, organize and outline a written manuscript
- Identify opportunities for submission/publication
Why Publish?

• To share research results
• For promotion and tenure
• Project completion
Journal Consideration

- Intended audience
- Type of study
- Journal databases
- Impact factor
- Cost
- Turn-around time
- On-line availability
- Length
Components of a Manuscript Describing Original Quantitative Research

- Title page
- Abstract
- Introduction
- Methods
- Results
- Discussion
- Acknowledgements
- References
Title Page

• Title
  – Short and catchy
  – Long and descriptive

• Author list
  – Decide up-front
  – First and last most important in science generally

• Key words
  – MeSH words

• Corresponding author
Abstract

• Journals have different word limits
• Stand-alone
• Details in the next session
Introduction

• Set the stage
• Literature review
• Lead to justification for current study
Methods

• Data collection, management, analysis
• Detailed enough to allow replication
• Cite standard techniques without details
• IRB review and approval for human subjects research
• IACUC review and approval as necessary
Results

• Tables and figures labeled appropriately
  – Do not use red and green together in figures
• Consider the audience for presentation of results
Discussion

• Place the results in context of prior research
• Do not restate results
• Implications of results
• Next steps
• Strengths and weaknesses
Acknowledgements

- Funding support
- Support staff
- Colleagues who provided feedback
- Professional writers
- Journals sometimes require written permission for acknowledgement
References

• Journal style
• Some journals limit the number
  – Consider recent review papers
• Check spelling and citations!
• Don’t over-cite self
Submission Process

• Usually on-line
  – Requires a user name and password
• Cover letter
  – Originality
  – Appropriateness for journal
  – Suggested reviewers
• Conflicts of interest
Where Do You Start?

• Set aside undisturbed time to write
• Choose the journal prior to writing
• Outline the tables as you conduct the analyses
A Few Other Things to Consider

• Writing style – varies by journal and field
• Protection of intellectual property prior to publication (or presentation)
• Potential media releases
• Other types of manuscripts
  – Review articles
  – Qualitative research
Resources

• The Writing Center
  – writing.wisc.edu
• Classes and workshops
• Books on scientific writing
• Conference sessions
Example
Genome-Wide Association Study of Cataract in the eMERGE Consortium

CA McCarty, S Turner, L Rasmussen, C Waudby, R Berg, J Linneman, P Peissig, L Chen, J Starren, RA Wilke, M Ritchie
Aims

• Develop and validate electronic algorithms for cataract, smoking and statin use in PMRP
• Identify genetic predictors of age-related cataract through genome-wide association study (GWAS)
Personalized Medicine Research Project (PMRP) Summary

- Population-based biobank with 20,000
  - Average clinical history 29 years
  - Age 18-102, 57% female
- DNA, plasma and serum samples
- Dietary intake and physical activity data
- 50 polymorphisms available for all
- Consented to allow data/sample sharing
- Access to medical records
- Ability to recontact subjects
Cataract - Approach

Cataract - Based primarily on data from the electronic medical record:

- diagnoses dating back to 1960
- procedure codes dating back to 1985

Case determination:

- includes senile cataract diagnosis and surgery
- excludes congenital, traumatic and juvenile cataract diagnoses
- takes into consideration age & chronic steroid use
## Samples Genotyped

<table>
<thead>
<tr>
<th>HDL</th>
<th>Cataract surgery</th>
<th>Cataract diagnosis only</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;2 HDLs</td>
<td>596</td>
<td>302</td>
<td>538</td>
</tr>
<tr>
<td>Case</td>
<td>118</td>
<td>120</td>
<td>51</td>
</tr>
<tr>
<td>Control</td>
<td>182</td>
<td>337</td>
<td>303</td>
</tr>
<tr>
<td>Other</td>
<td>427</td>
<td>564</td>
<td>431</td>
</tr>
<tr>
<td>Total</td>
<td>1323</td>
<td>1323</td>
<td>1323</td>
</tr>
</tbody>
</table>

743 mixed, 882 nuclear, 122 cortical, 61% female
Genotyping

- 3968 DNA samples shipped to CIDR
- Illumina 660quad
- 3947 samples with GWAS data for analysis
  - 17 samples dropped due to poor performance (CIDR)
  - 4 samples dropped due to potential sample mix-up
- 21 blind duplicates (99.99% concordance)
- 85 HapMap control samples (99.8% concordance)
  - 44 CEU controls
  - 32 Yoruba controls
  - 5 Japanese
  - 4 Han Chinese
Analytic Approach

- Cataract case definitions
  - Case (including nuclear, cortical, and mixed)
  - Nuclear only cases
  - Cortical only cases
- Logistic regression analysis in PLINK
- Mixed model regression in genABEL
- Additional cases from Group Heath (n=105) and Vanderbilt (n=91)
- Little difference by type
- Little difference in logistic regression or mixed model regression
- Principal components analysis
Manhattan Plot

1 marker on chromosome 14 almost reaches genome-wide significance
Next Steps

- Add controls from the other sites
- Age at diagnosis as outcome variable
- Replication
- Gene/gene analyses
- Gene/environment analyses
Acknowledgements

• Center for Inherited Diseases Research, Johns Hopkins University
• Grant number U01HG004608-01
Manuscript Consideration

• Audience – clinicians, geneticists?
  – Background and context will vary

• Open access journal?

• Timing
  – No other GWAS data published previously
  – Data do not reach genome-wide significance

• Shared writing