Statistical Reports: Motivation, Tips & Ideas

Jennifer Thompson, MPH

Department of Biostatistics, Vanderbilt University School of Medicine

February 19, 2010

An Analysis Report Should...

- Clearly present the analysis and results
  - Describe purpose and research question(s)
  - Explain methodology
  - Present results and their clinical meaning
- Serve as a manuscript base
- Be an archive of your work

Archiving Your Work

- My projects are finished the first time I send my report to the PI.

  above <- FALSE

- Methods & results should be explained enough that future users can reproduce or modify them, regardless of software package.
- The document needs to be **complete** and **clear**.
  - Strike a balance: too little vs. too much
  - Your audience:
    1. you
    2. PIs
    3. PIs' audience

Beginnings of a Manuscript

Typical progression:

1. Introduction
2. Methods
3. Results
4. Discussion

Following this structure saves effort in the long run.

- Knowing the information needed for the introduction ensures that you know the question(s) the PIs want to answer.
- Explaining methods and results thoroughly from the beginning reduces confusion and time writing the manuscript.

Beginnings of a Manuscript

Clear understanding of the clinical basis is absolutely necessary.

- To write the introduction, we must know the reason for the project and the questions we need to answer.
- To write the methods, we must have a clear understanding of why each method was used and what the resulting numbers mean in their clinical context.
- To write the results, we must know the main message our PIs are trying to get across.

Now: How?

- Know your audience.
- Keep the end result in mind.
- Make it easy to read.

Know Your Audience

- How much statistical background do they have?
  - If I say "proportional odds model," will they know what I mean? Will they be able to explain it to people who don't?
  - Are they familiar with terms we throw around - "beta coefficient," "longitudinal data," "case-control"?
- How much attention will they pay to the methodology?
  - Include a thorough explanation of what you did, but...
  - If it won't go in the manuscript and they don't care, put the emphasis on what is most important for them to know.

Where Is This Going?

- The amount of work and information you put into a med student's research day project will be different than one sent to NEJM.
- Should you use colors in graphics?
- What should tables and figures look like?

Don't Make Them Work For It

Make things easier on your PIs and on yourself. Keep your reports succinct and easy to read.

Clearly define sections; have a good flow to the report.

- Use organization tools: table of contents, lists of figures/tables, section headings, hyperlinks and cross references.
- Title tables, figures, and sections clearly. "Blood Pressure" is not as descriptive as "Systolic BP vs. Treatment (Adjusted Results)."
- Make it make sense, now and forever. Keep related sections together.
- But split them up when necessary. One long "Results" section, or "Primary Results," "Secondary Outcomes" and "Sensitivity Analyses"?

Outline Archiving Your Work Beginnings of a Manuscript How-Tos

Don't Make Them Work For It

and interaction terms). When an interaction with sepsis is present, two additional p-values (calculated using contrasts) are given for the association of biomarker and outcome separately for septic and non-septic patients. For odds and hazard ratios, the 25th and 75th percentiles of continuous variables were used as comparison points, rather than the traditional one-unit increment. We chose to focus on the outcomes of brain dysfunction (delirium/coma-free days), mortality, and kidney failure.

After adjusting for age, acute physiology score of the APACHE II, and treatment group, any relationship between PCT and DCFDs appears to be linear and unaffected by sepsis admission diagnosis. Therefore, nonlinear and interaction terms were removed from the proportional odds model. This model then showed a significant linear relationship between PCT and delirium/coma-free days during the MENDS study ($p = 0.04$, OR and 95% CI = 0.5 (0.3, 1.0); see Table 2). Therefore, we estimate that a patient with a baseline PCT value of 6.7 will have approximately half the odds of more DCFDs as a patient with a baseline PCT of 0.4 - in other words, higher PCT is associated with more delirium or coma. After adjusting for covariates, there seems to be no significant relationship between PCT and mortality ($p = 0.40$; see Table 5). There is no significant association between PCT and kidney failure-free days after adjusting for all covariates ($p = 0.23$; see Table 6). After adjusting for age, acute physiology score of the APACHE II, and treatment group, the relationship between CRP and DCFDs appears to be significant ($p = 0.03$), though linear and not modified by sepsis. With an odds ratio of 0.6 (0.4, 1.0), higher levels of CRP at baseline are associated with lower probability of more DCFDs - in other words, more delirium and/or coma (see Table 12). After adjusting for covariates, there is a marginally significant, linear relationship between CRP and mortality ($p = 0.12$; see Table 15). The hazard ratio and 95% CI for CRP is 1.9 (0.8, 4.2); in other words, we expect a patient with a CRP of 282 to have nearly twice the risk of death at any given point during their hospital stay as a patient with a CRP of 107.

Don't Make Them Work For It

where applicable, nonlinear and interaction terms). When an interaction with sepsis is present, two additional p-values (calculated using contrasts) are given for the association of biomarker and outcome separately for septic and non-septic patients.

For odds and hazard ratios, the 25th and 75th percentiles of continuous variables were used as comparison points, rather than the traditional one-unit increment.

We chose to focus on the outcomes of brain dysfunction (delirium/coma-free days), mortality, and kidney failure.

4.2 Results: PCT vs. DCFDs

After adjusting for age, acute physiology score of the APACHE II, and treatment group, any relationship between PCT and DCFDs appears to be linear and unaffected by sepsis admission diagnosis. Therefore, nonlinear and interaction terms were removed from the proportional odds model. This model then showed a significant linear relationship between PCT and delirium/coma-free days during the MENDS study ($p = 0.04$, OR and 95% CI = 0.5 (0.3, 1.0); see Table 2). Therefore, we estimate that a patient with a baseline PCT value of 6.7 will have approximately half the odds of more DCFDs as a patient with a baseline PCT of 0.4 - in other words, higher PCT is associated with more delirium or coma.

4.3 Results: PCT vs. Mortality

After adjusting for covariates, there seems to be no significant relationship between PCT and mortality ($p = 0.40$; see Table 5).

4.4 Results: PCT vs. Kidney Failure-Free Days

There is no significant association between PCT and kidney failure-free days after adjusting for all covariates ($p = 0.23$; see Table 6).

Outline Archiving Your Work Beginnings of a Manuscript How-Tos

Don't Make Them Work For It

We have spelling, grammar and punctuation rules because they make it easier to communicate ideas.

Pages

Name: 'Let's eat Grandma!' or, 'Let's eat, Grandma!' Punctuation saves lives.

Type: Other Business

Fans: 365,854 fans

- Pay attention to spelling, capitalization, and grammar. If you're not sure, ask.
- Edit.
- **Proofread.**

Don't Make Them Work For It

- "Statistical jargon": Either avoid it or explain it.
  - "Gaussian" = "normal"
  - "We used Markov regression, which means we adjusted for mental status on the previous day."
- Use examples to explain unfamiliar concepts.
  - Beta coefficient for treatment in a linear model
  - Using 25th and 75th percentiles to calculate odds ratios

"The hazard ratio and 95% CI for CRP is 1.9 (0.8, 4.2); in other words, we expect a patient with a CRP of 282 to have nearly twice the risk of death at any given point during their hospital stay as a patient with a CRP of 107."

Outline Archiving Your Work Beginnings of a Manuscript How-Tos

Don't Make Them Work For It

- Keep the eyes in mind: use clear fonts, left and right justification, and reasonable font sizes.
- Don't include information that isn't necessary or helpful. It only serves to distract. How many decimal places do you really need?
- Printing straight R/SAS/Stata output often a) is neither clear nor attractive and b) includes a lot of information your PI doesn't need.
  - Will your PI know what variable you mean by "wsbp.d1"? Use variable labels.
  - Do they need all the information that comes with a summary() of an lm() call?
  - Do they know how to find what they **do** need?
  - Bottom line: Making it hard for PIs to find what they need only makes it harder for you in the long run.

Don't Make Them Work For It

Table: Effects Response : mental.status.t

	Low	High	Δ	Effect	S.E.	Lower 0.95	Upper 0.95
age	46.533	68.168	21.636	0.19	0.26	-0.32	0.69
Odds Ratio	46.533	68.168	21.636	1.20		0.72	2.00
apache.aps	11.000	20.000	9.000	-0.15	0.26	-0.67	0.37
Odds Ratio	11.000	20.000	9.000	0.86		0.51	1.44
mental.status.y — Comatose:Normal	1.000	2.000		4.47	0.67	3.15	5.79
Odds Ratio	1.000	2.000		87.64		23.41	328.13
mental.status.y — Delirious:Normal	1.000	3.000		2.96	0.41	2.15	3.77
Odds Ratio	1.000	3.000		19.25		8.57	43.24
xigris.begin — Yes:No	1.000	2.000		-0.33	0.46	-1.22	0.57
Odds Ratio	1.000	2.000		0.72		0.29	1.76
group — Dexmedetomidine:Lorazepam	1.000	2.000		-1.14	0.40	-1.92	-0.37
Odds Ratio	1.000	2.000		0.32		0.15	0.69
sepsis.final — Septic:Non—septic	1.000	2.000		0.33	0.39	-0.44	1.09
Odds Ratio	1.000	2.000		1.39		0.65	2.98

Outline Archiving Your Work Beginnings of a Manuscript How-Tos

Don't Make Them Work For It

Table: Wald Statistics for `mental.status.t`

	χ^2	*d.f.*	*P*
mental.status.y	69.35	2	<0.0001
age	0.51	1	0.4750
apache.aps	0.32	1	0.5728
xigris.begin	0.52	1	0.4709
group	8.35	1	0.0039
sepsis.final	0.70	1	0.4016
TOTAL	82.64	7	<0.0001

Don't Make Them Work For It

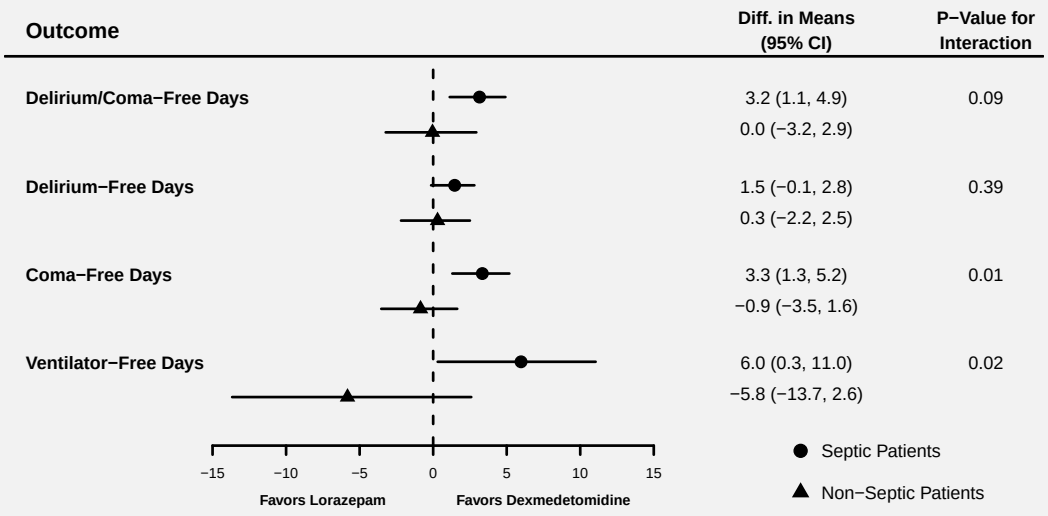
Table: Transition to Delirium vs. Treatment, Adjusted Results

	25th Pct	75th Pct	OR (95% CI)	Chi-Square	df	P-Value
Age at enrollment	46.5	68.2	1.2 (0.7, 2.0)	0.5	1	0.48
APACHE II acute physiology score	11.0	20.0	0.9 (0.5, 1.4)	0.3	1	0.57
Treatment group, DEX:LZ			0.3 (0.1, 0.7)	8.3	1	0.004
Previous day's mental status				69.3	2	<0.001
Comatose:Normal			87.6 (23.4, 328.1)			
Delirious:Normal			19.2 (8.6, 43.2)			
Sepsis at enrollment, Septic:Non-septic			1.4 (0.6, 3.0)	0.7	1	0.40
Xigris within 48 hours?, Yes:No			0.7 (0.3, 1.8)	0.5	1	0.47
Total				82.6	7	<0.001
Observations						262

White Space Is the Right Space

- Our eyes and brains need white space.
  - Will a figure help you get your point across better than more explanation?
  - Don't use extra words, numbers or graphical "stuff." Make what really counts be what stands out.
  - Use tables and figures to convey your message. Often, these are clearer and more concise than more words, and our eyes digest them more easily.
- Pretty is nice, but not the point.
  - Use graphics when they're helpful - and only then.
  - **Make every drop of ink count.** Do you really need grid lines, bar charts, multiple colors?
  - Make your figures attractive and eye-catching, but do it in ways which truly help your message come across.

White Space Is the Right Space



No, there is too much. Let me sum up.

- Understand the research.
- Know your audience, message and end result.
- Include all relevant information, including background, methodology and explanations of results.
- But edit! Take out unhelpful words, numbers, colors...
- Make it easy to find what you and your PIs need.