Impact of varying data quality on statistical analyses of integrated datasets

Eugenio Andraca-Carrera
Mathematical Statistician
Division of Biometrics VII
Office of Biostatistics, CDER, FDA

ISCTM 11th Annual Scientific Meeting
Washington, DC
17 February 2015

Disclaimer

This presentation reflects the views of the author and should not be construed to represent FDA's views or policies

Motivation for integrating datasets

- Why? More data!
  - Obtain more precise answers
  - Quantify the risk of rare events
  - Investigate a potential signal
  - Assess effects in subgroups
  - Understand variation from different studies

- However… more data are not necessarily better data
  - Focus on retrospective analyses

Outline

- **Motivating example: PPI meta-analysis case study**
- Challenges of integrating randomized trials
- Additional challenges when the goal of the meta-analysis is outside the scope of the original data
- Conclusions

What do we mean by meta-analysis?

- Statistical technique to integrate findings from independent studies controlling for trial heterogeneity
- Estimate a common treatment effect
- In this talk we use the term meta-analysis for an integrated analysis of trials with subject-level data

Case study: the PPI meta-analysis

- Proton pump inhibitors (PPIs) are used to reduce the production of gastric acid
- Epidemiological studies suggested possible increase in risk of fractures of hip, wrist, and spine with use of PPIs
- FDA issued a Drug Safety Communication on May 25, 2010:

  planned meta-analysis of data from long-term, placebo-controlled trials of bisphosphonates (drugs intended for the treatment of **osteoporosis**) to evaluate risk of fractures associated with PPIs

Objectives of meta-analysis

1. To determine whether use of PPIs is associated with an increased risk of fractures among subjects enrolled in long-term, placebo-controlled trials of osteoporosis drugs. 3 types of fractures: vertebral, non-vertebral, hip.

2. To determine whether use of PPIs is associated with a decrease in Bone Mineral Density (BMD) among subjects enrolled in long-term, placebo-controlled trials of osteoporosis drugs. 3 BMD locations: femoral neck, lumbar spine, hip.

Why use trials for osteoporosis to study PPIs?

- **Advantages**
  - More fractures captured
  - Better assessment of fractures and BMD
  - Long-term trials (>1 year)
  - Common use of PPIs as concomitant medication (CM)
- **Disadvantages**
  - PPI exposure not randomized
  - Quality of PPI exposure data may be poor

This Meta-Analysis uses RCTs, but is closer to a MA of observational studies

Data Overview

- 15 trials for 7 different products
  - 45,032 randomized subjects (27,367 to active treatment)
  - Trials conducted between 1995 and 2004
  - Multicenter, international trials

Summary of trials

Product	Trial	N		Mean Follow-up (years)
		Placebo	Active Drug	
Reclast	2301	3876	3888	3
Forteo	GHAC	541	1085	1.7
	GHAJ	147	290	1
Boniva	mf4380	949	1912	3
	mf4411	982	1964	3
	mf4499	162	491	2
	mf4500	157	471	2
Preos	93001	1183	1241	1.5
Viviant	A1-300	347	1045	2
	A1-301	1914	3811	3
lasofoxifene	A2181002	2830	5646	3
	A2181003	233	691	2
	A2181004	245	737	2
denosumab	20030216	3933	3929	3
	20040132	166	166	2
Total:		**17665**	**27367**	

Outline

- PPI meta-analysis case study
- **Challenges of integrating randomized trials**
  - Covariates
  - Exposure
  - Outcome
- Additional challenges when the goal of the meta-analysis is outside the scope of the original data
- Conclusions

Challenges: Covariates

- Lack of consistency between studies
  - Different definitions
  - Different recording methods

- Missing data
  - Subject level
  - Trial level

Covariates in PPI Meta-Analysis:

- Demographics and baseline characteristics
  - Age, Sex, Race available in all datasets
  - Race was defined differently across trials
  - Smoking Status defined differently:

    Smoker/Non-Smoker, Smoker/Previous Smoker/Non-Smoker, Tobacco User/Non-User
  - BMI unavailable in one dataset (height was not captured)

Challenges: Exposure of Interest

- If exposure of interest is not part of a randomized treatment arm:
  - Dosage
  - Episodes
  - Duration

- Missing data

Exposure in PPI Meta-Analysis:

- Randomized treatment (osteoporosis drugs) exposure fully recorded

- PPI use was captured in concomitant meds datasets:
  - Missing dosage
  - Missing duration, start time, number of episodes
  - Captured in verbatim form:

    Example: Omeprazole = Acichek, Acifre, Acimed, Acipres, Acromon, etc…

- Possible misclassification as PPI user / non-user

Challenges: Outcome

- Lack of consistency between studies
  - Different definitions
  - Different recording methods
- Missing data

Outcome in PPI Meta-Analysis:

- Outcomes
  - Adjudication of fractures varied by location and trial:
    - Some used X-rays
    - Some adjudicated by a panel
    - Some only included fractures in AE datasets
  - Bone Mineral Density was available in all trials
  - Possible misclassification of fractures

The SAS System
The FREQ Procedure

FRTERM	Frequency
"R" ULNAR SYLOID FX	1
"R" DISTAL RADIUS COMMINUAL INTRAARTICUBE FX	1
# L SCAPHOID	1
# R FOOT	1
# R PERTROCHANTERIC NECK OF FEMUR	1
# RIBS R 9 & 10	1
'FIFTH METATARSAL FRACTURE	1
(BACK PAIN) THORAC	1
(INCIDENTAL) COMPR	1
(L) Collie fractur	1
(L) Collie's fract	1
(L) HIP FRACTURE	1
(L) Knee cap fracture	1
(L) Patellar fracture	1
(L) RIB FRACTURE	3
(L) ankle fracture - lateral malleolus	2
(L) collie's fracture	1
(L) foot 5th metatarsal spiral fracture (nondisplaced)	1
(L) foot fracture	1
(L) knee cap fracture	2
(L) knee cap refracture	2
(L) shoulder pain {(secondary to fracture)}	3
(R) 5TH TOE FRACTURED	4
(R) Colles fracture	1
(R) FOOT FRACTURE	1
(R) Fibula fractura	1
(R) Foot 5th Metatarsal fracture	1

~ 4200 different verbatim terms

Data Challenges - Summary

- Harmonizing covariates, exposure and outcomes
- Measuring exposure – particularly if not part of randomized treatment
- Capturing the right outcomes

- In our PPI meta-analysis significant resources were needed at each step - clinical and statistical input needed

Outline

- PPI meta-analysis case study
- Challenges of integrating randomized trials
- **Additional challenges when the goal of the meta-analysis is outside the scope of the original data**
- Conclusions

Moving Beyond the Data's Original Question

- Were the trials designed to answer the same research question?
  - Different endpoint of interest than in study?
  - Different treatment(s) of interest than in study?
- Randomization concerns
  - If exposure of interest is randomized in the trials, OK
  - If not, can't assume baseline covariates (including unobserved) are balanced
    - **Data is observational**
    - PPI study—treatment of interest is PPI, not osteoporosis medications
    - Potential confounding

Outline

- PPI meta-analysis case study
- Challenges of integrating randomized trials
- Additional challenges when the goal of the meta-analysis is outside the scope of the original data
- **Conclusions**

Lessons Learned:

- Prospective meta-analyses can anticipate many of these issues
- Plan research questions and Statistical Analysis Plan before collecting data

- For retrospective meta-analyses:
  - Consider heterogeneity of trials, endpoints, populations
  - Assumptions may be needed for harmonizing covariates and endpoints, defining exposure, etc…
  - Plan for missing data
  - Allocate enough resources for data cleaning and harmonization
  - Concomitant Medication data may provide limited information of exposure

Lessons Learned (continued):

- Test your assumptions - sensitivity analyses
- Understand randomization (or lack thereof)
- Standardization is important in clinical trials
- More data ≠ better data

- The interpretability and impact of an integrated analysis depend on:
  - Pre-specified plan for data collection, inclusion criteria, event definition and analysis
  - Quality of the data
  - Plausibility of your assumptions and generalizability of your results
  - Quality of your analysis
  - Clear and effective communication of your results

Thanks to:

- Mat Soukup
- Benjamin Neustifter
- Tamara Johnson
- Joyce Korvick
- Theresa Kehoe
- Marcea Whitaker
- Aloka Chakravarty
- Mark Levenson