Overview of CMF Guidebook
A Guide for Developing Quality CMFs

TRB International Workshop on CMFs
January 23, 2011
Objective

- Provide overview of guidebook
- Identify various methods for developing CMFs
- Set stage for afternoon
  - Issues with various methods
Overview of Guidebook

**Intent** – provide direction to agencies interested in developing CMFs

**Audience** – transportation safety practitioners, consultants, and researchers
Overview of Guidebook

**Prerequisites** – experience and/or education in the theory and practice of:

- Road safety engineering
- Basic analytical procedures
- Statistical concepts
Overview of Guidebook

Content

- Background on CMFs
  - Definitions, purpose, application, general issues

- Study Designs
  - Overview, sample size considerations, and strengths and weaknesses

- Resources
  - Identify appropriate analysis method
  - Improving completeness/consistency in reporting
Methods

- Before-After
  - Comparison Group
  - Empirical Bayes
  - Full Bayes

- Cross-Sectional

- Case-Control

- Cohort

- Alternative Approaches
Before-After – Comparison Group

- Account for changes in crashes unrelated to the treatment
  - Time trends
  - Traffic volume changes

<table>
<thead>
<tr>
<th>Time Period</th>
<th>Treatment Group</th>
<th>Comparison Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>( N_{\text{observed},TB} )</td>
<td>( N_{\text{observed},CB} )</td>
</tr>
<tr>
<td>After</td>
<td>( N_{\text{observed},TA} )</td>
<td>( N_{\text{observed},CA} )</td>
</tr>
</tbody>
</table>
Before-After – Empirical Bayes

- EB approach properly accounts for changes in crashes due to:
  - Regression-to-the-mean
  - Better accounts for traffic volume and time trends
Before-After – Full Bayes

- Not a type of evaluation study on its own
  - Modeling technique
  - Similar to generalized linear modeling procedure
  - Used to develop estimates of expected crashes
  - Can be used in before-after and development of cross-sectional models

- Before-After
  - Reference group is used similar to EB
  - Distribution is used instead of point estimate
    - Combined with observed crashes
    - Estimate expected crash frequency, variance, and variance of estimated CMF
Before-After – Full Bayes

- **Cross-sectional**
  - Applied similar to generalized linear regression to relate:
    - Geometric characteristics
    - Operational characteristics
    - Expected crash experience
  - More flexible modeling tool
    - Allows for complex model forms
    - Estimation of valid models with small sample size
    - Ability to consider spatial correlation
    - Opportunity to incorporate prior knowledge
Before-After – General Issues

- Observed change in crashes may be due to other factors
  - Traffic volume changes
  - Changes in reported crash experience
  - Regression-to-the-mean
  - Other improvements
Before-After – Comparison Group Issues

- Requires suitable comparison group
Before-After – Comparison Group Issues

- **Difficult to account for RTM**
  - Must match treatment and comparison sites based on crash occurrence

- **Difficult to confirm RTM is not an issue**
  - Crash frequency not considered in site selection
    - Blanket treatment
    - Truly random selection of sites for treatment
    - Safety evaluation is related to operational improvement
Before-After – Empirical Bayes Issues

- More complex than comparison group
- Identification of suitable reference group
  - Spillover
    - Treatment may affect the logical reference group
    - E.g., Red light camera programs
Before-After – Full Bayes Issues

- More complex than empirical Bayes method
  - Requires high level of statistical training
- Software developed for application of empirical Bayes method
  - Seems to be difficult for the full Bayes method
Cross-Sectional

- Compare with and without rather than before and after
- Useful for estimating CMFs where there are insufficient instances of actual treatments
  - Several instances of sites with and without, but few changed from before to after
  - E.g., Compare 4-ft shoulder to 6-ft shoulder
    - Few projects where the shoulder is widened from 4 to 6 feet
    - Many segments with 4-ft shoulders and many with 6-ft shoulders
Cross-Sectional

- **Question:**
  - What are the safety effects of signalization?

- **Scenario**
  - 100 two-way stop-controlled intersections
  - 100 signalized intersections
  - Rural, 4-legged with similar traffic volumes

\[
CMF = \frac{2.9}{3.4} = 0.85
\]

3.4 crashes/year  
2.9 crashes/year
Cross-Sectional Issues

- Comparison is between two distinct groups
  - Difference in crashes can be due to other factors (both known and unknown)
- Difficult to properly account for unknown, or known but unmeasured, factors
- Control for known factors through multiple variable regression
  - Science of assembling CMFs from multiple variable models is not fully developed
  - Inappropriate functional form, omitted variable bias, or correlation of variables
Case-Control

- Select sites based on outcome status and then determine prior treatment status
- Assess whether exposure to treatment is disproportionately distributed
  - Estimate odds ratio
  - Indicates likelihood of actual benefit

<table>
<thead>
<tr>
<th>Treatment</th>
<th># of Cases</th>
<th># of Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>With</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Without</td>
<td>C</td>
<td>D</td>
</tr>
</tbody>
</table>

\[ Odds \text{ Ratio (OR)} = \text{CMF} = \frac{A/B}{C/D} = \frac{AD}{BC} \]
Case-Control Issues

- Cannot be used to determine relative risk
  - Indicates likelihood of outcome given presence of specific feature

- Cannot demonstrate causality
  - No time sequence of events

- Does not recognize differences between locations with multiple crashes
Cohort

- Select sites based on treatment status and then determine outcome status over time
- Assess whether exposure (time until event) is disproportionate between cohorts
  - Estimate relative risk
  - Direct estimate of CMF

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Outcomes</th>
<th>Non-Outcomes</th>
<th>Total At-Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>With</td>
<td>A</td>
<td>B</td>
<td>A + B</td>
</tr>
<tr>
<td>Without</td>
<td>C</td>
<td>D</td>
<td>C + D</td>
</tr>
</tbody>
</table>

Relative Risk = CMF = \( \frac{A}{C \cdot (A + B)} \)
Cohort Issues

- Large samples are often required
  - Relatively expensive
- Site characteristics are subject to change
  - Particularly for lengthy study periods
- Does not recognize differences between locations with multiple crashes
Alternative Approaches

- **Meta-analysis**
  - Aggregate analysis of past research
  - Systematically combine knowledge on CMFs

- **Expert panels**
  - Critically evaluate findings of published and unpublished research
  - Derive CMFs through consensus

- **Surrogates**
  - Derive a CMF indirectly using data other than crash data
    - E.g., vehicle speeds, traffic conflicts, etc
## Surrogate Example

<table>
<thead>
<tr>
<th>Mean Pre-Treatment Speed (mph)</th>
<th>Speed Reduction (mph)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5</td>
</tr>
<tr>
<td>45</td>
<td>0.70</td>
</tr>
<tr>
<td>50</td>
<td>0.73</td>
</tr>
<tr>
<td>55</td>
<td>0.76</td>
</tr>
<tr>
<td>60</td>
<td>0.78</td>
</tr>
</tbody>
</table>

*NCHRP Report 617*
Alternative Approaches Issues

- **Meta-analysis**
  - Ensure all studies used are of sufficient quality
  - Sensitivity and publication bias

- **Expert panels**
  - Informal (based on expert consensus)
  - Difficult to derive estimates of uncertainty

- **Surrogates**
  - Need to establish relationship between surrogate and crashes
Which Method is Preferred?

- Before-after
  - Comparison group
  - Empirical Bayes
  - Full Bayes
- Cross-sectional
- Case-control
- Cohort
- Alternative methods
- It depends!
Scenario 1

- Jurisdiction implemented a 1.5 second all-red phase at 16 traffic signals in CBD
  - Blanket treatment
  - All 4-legged intersections
  - No other signalized intersections in vicinity
  - Several 2-way stop-controlled intersections along same two routes
  - Reasonable to believe that treatment does not impact safety at stop-controlled intersections

- Before-after with comparison group
Scenario 2

- Jurisdiction converted 2-way stop-controlled intersections to roundabouts
  - Suspected that safety benefits in this jurisdiction may be less than those found elsewhere
  - No new roundabout conversions until further study
  - Data will be used only from this jurisdiction
  - Limited before-after data exist for 10 conversions
  - All converted sites have similar characteristics
  - Large pool of 2-way stop-controlled intersections that have not been converted
  - Conversion is likely to change traffic volumes

- Empirical Bayes Before-After
Scenario 3

- Need to estimate CMF for flattening horizontal curves on rural, 2-lane roads
  - Agency’s crash data system has been updated
    Only latest 5 years of crash data are available
  - Few curves have undergone reconstruction and many were completed > 5 years ago
  - Data for 350 curves on rural, 2-lane roads
  - Data available for curve radii as well as other geometric and traffic volume data

- Cross-sectional
Scenario 4

- Develop CMF for increasing paved shoulder width on 2-lane rural roads
  - Several miles with narrow or no paved shoulders
  - Several more miles with 3 – 4 foot shoulders
  - Do not intend to implement treatment unless it is cost-effective to address run-off-road crashes
  - High frequency of ROR crashes, but spread-out over the network
    - Several segments with no crashes over 3 years
    - Several segments have only 1 or 2 crashes
  - Geometric and traffic volume data are available to control for factors other than the treatment

- Case-Control
Scenario 5

- Consider previous scenario, but now focus on mountainous regions instead of all 2-lane, rural roads
  - Crashes more prevalent in mountainous regions
  - Most segments experience at least 1 crash/year
  - Fewer miles for analysis
  - Do not intend to implement treatment unless it is cost-effective to address run-off-road crashes
  - Geometric and traffic volume data are available to control for factors other than the treatment

- Cohort
Resources

- Annotated outline
  - Improve completeness/consistency in reporting

- User must determine quality of CMF
  - Need complete and consistent information

- Highway Safety Manual and CMF Clearinghouse assess quality of CMFs
  - Need complete and consistent information
## Resources – Relative Quality of CMFs

<table>
<thead>
<tr>
<th>Relative Rating</th>
<th>Excellent</th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study Design</strong></td>
<td>Statistically rigorous study design with reference group or randomized experiment and control</td>
<td>Simple before / after study</td>
</tr>
<tr>
<td><strong>Sample Size</strong></td>
<td>Large sample, multiple years, diversity of sites</td>
<td>Limited homogeneous sample</td>
</tr>
<tr>
<td><strong>Standard Error</strong></td>
<td>Small compared to CRF</td>
<td>Large SE and confidence interval includes zero</td>
</tr>
<tr>
<td><strong>Potential Bias</strong></td>
<td>Controls for all sources of known potential bias</td>
<td>No consideration of potential bias</td>
</tr>
<tr>
<td><strong>Data Source</strong></td>
<td>Diversity in States representing different geographies</td>
<td>Limited to one jurisdiction in one State</td>
</tr>
</tbody>
</table>
Resources – Outline

- **Objective**
  - Identify treatment of interest
  - Discuss reason for conducting study
  - Identify target crash types and severities

- **Background**
  - Describe treatment of interest

- **Literature Review**
  - Summary of recent and salient literature
Resources – Outline

- Methodology
  - Discuss method used to develop CMF
    - Reason for selecting method
    - Strengths and weaknesses
  - Identify potential sources of bias
    - How these biases are addressed
    - Those that cannot be addressed
Resources – Outline

- **Data**
  - Data source(s)
  - Years of data
  - Number of sites (and/or miles if applicable)
  - Summary statistics
    - Average crashes per year
    - Annual, average, min, and max traffic volume
  - Applicability of CMFs
Resources – Outline

- **Results**
  - Present CMF(s)
  - Standard error of CMF
    - Calculate confidence interval
    - Judge quality and significance of results
Contact

Frank Gross, VHB
fgross@vhb.com
- Who covers general issues?
  - John, Frank, or Bhagwant?

- Applying Multiple CMFs

- CMFs Derived From High Crash Locations

- Considerations Related to Before-After and Cross-Sectional Designs

- Factors Affecting the Quality of CMFs