Overview of CMF Guidebook
A Guide for Developing Quality CMFs
Objective

- Provide overview of various methods for developing CMFs
  - General applicability
  - Strengths
  - Weaknesses
- Review flow chart to select appropriate method
Methods

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

- Cross-Sectional

- Case-Control

- Cohort

- Alternative Approaches
Before-After – Comparison Group

- General applicability
  - Treatment is similar among treatment sites
  - Before and after data are available for treated and untreated sites
  - Untreated sites used to account for changes other than the treatment
Before-After – Comparison Group

- **Strengths**
  - Simple
  - Accounts for time trends and traffic volume changes

- **Weaknesses**
  - Difficult to account for RTM
  - Difficult to confirm RTM is not an issue
Before-After – Empirical Bayes

- General applicability
  - Treatment is similar among treatment sites
  - Before and after data are available for treated and untreated reference group
Before-After – Empirical Bayes

- **Strengths**
  - Properly accounts for changes in crashes due to:
    - Regression-to-the-mean
    - Traffic volume and time trends

- **Weaknesses**
  - Relatively complex
  - May require separate comparison group if treatment impacts reference group
  - Cannot consider prior knowledge
  - Cannot account for spatial correlation
  - Cannot specify complex model forms
Before-After – Full Bayes

- General applicability
  - Useful for before-after or cross-sectional studies

- Before-After
  - Distribution is used instead of point estimate
    - Estimate expected crash frequency, variance, and variance of estimated CMF

- Cross-Sectional
  - Relate geometric and operational characteristics with expected crash experience
Before-After – Full Bayes

- **Strengths**
  - More flexible modeling tool
    - Complex model forms
    - Considers spatial correlation
    - Incorporate prior knowledge
    - Estimation of valid models with small sample size

- **Weaknesses**
  - Requires a high degree of statistical training
Cross-Sectional

- General applicability
  - Compare *with* and *without* rather than before-after
  - Useful when limited before-after data
  - Sites that are similar except for treatment of interest

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

3.4 crashes/year

2.9 crashes/year
Cross-Sectional

- **Strengths**
  - Possible to develop CMFunctions
  - Allows estimation of CMFs when conversions are rare

- **Weaknesses**
  - Difference in crashes can be due to other factors (both known and unknown)
  - Difficult to account for unknown, or known but unmeasured, factors
  - Inappropriate functional form, omitted variable bias, or correlation of variables
Case-Control

- General applicability
  - 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>

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

- **Strengths**
  - Useful for studying rare events
  - Can investigate multiple treatments per sample

- **Weaknesses**
  - Cannot demonstrate causality
  - Can only investigate one outcome per sample
  - Does not recognize differences between locations with multiple crashes
Cohort

- **General applicability**
  - 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>

\[
\text{Relative Risk} = CMF = \frac{A(A + B)}{C(C + D)}
\]
Cohort

- **Strengths**
  - Useful for studying rare treatments
    - Sample is selected based on treatment status
  - Can demonstrate causality

- **Weaknesses**
  - Large samples are often required (expensive)
  - Site characteristics are subject to change
  - 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
Which Method is Preferred?

- Before-after
  - Comparison group
  - Empirical Bayes
  - Full Bayes

- Cross-sectional

- Case-control

- Cohort

- Alternative methods

- It depends!
A GUIDE FOR DEVELOPING QUALITY CRASH MODIFICATION FACTORS

Are data available for the treatment in your jurisdiction? OR Can you install the treatment and collect data?

- Are there suitable locations to develop a comparison group or reference group?
  - Yes
  - No

- Are there sufficient existing or planned installations for a before-after study?
  - Yes
  - No

Select before-after method based on criteria in table below

<table>
<thead>
<tr>
<th>Study Criteria</th>
<th>CG</th>
<th>EB</th>
<th>FB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression-to-the-mean may be a factor</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Treatment is likely to impact traffic volumes</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Include spatial correlation (either among treated sites or among treated and comparison or reference sites)</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>A complex model form is required</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Include prior knowledge of model or CMF estimates in the analysis</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>

Are there previous evaluations for which published or unpublished material is available?

- Yes
- No

Is a formal statistical approach desired? If so, do the published research studies include sufficient information for a meta-analysis?

- Yes
- No

Select method based on criteria in table below

<table>
<thead>
<tr>
<th>Study Criteria</th>
<th>Cross-Sectional</th>
<th>Case-Control</th>
<th>Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crash type is rare</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Treatment is rare</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Accounts for locations with multiple crashes (rather than first occurrence)</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>CMFunction desired</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>
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
Contact

Frank Gross, VHB
fgross@vhb.com