National Personal Protective Technology Laboratory

Statistical Explanations for Development of the TIL Criteria

Doug Landsittel

Statistician/Senior Fellow

NIOSH NPPTL Public Meeting
The Embassy Suites Pittsburgh International Airport Hotel
June 26, 2007

Outline

- **Overall statistical objectives**
  - NIOSH test panel
- **Statistical justification for an optimal criteria**
- **Example calculations and proposed criteria**
  - Interpret results
- **Summary and conclusions**

Overall Statistical Objectives

- **Initial considerations:**
  - Test a representative panel
  - Specify an acceptable TIL (≤ 5%)
- **TIL ≠ APF**
- **View TIL criteria as a statistical test**
  - Define concepts for an optimal test
  - Adequate number of subjects
  - Minimum % of subjects with acceptable TIL

NIOSH Test Panel

- **Eligibility screening**
  - Facial dimensions fit within the PCA Panel
- **Bivariate panel based on face width and length**
  - 35 total subjects from 10 different cells
  - Cell frequencies representative of U.S. workforce
- **Random selection of available subjects from within each cell of the panel**
  - Goal: avoid systematic error in subject selection
  - Other facial dimensions may be significant

Statistical Justification: Overall Concepts

- **Assumption: for a given model, an unknown % of subjects achieve acceptable fit**
  - Effectiveness of the model is judged by the % of subjects with acceptable fit across the population
- **Overall Goal: formulate a criteria with the following characteristics**
  - A highly effective model almost always passes
  - An ineffective model almost always fails

Statistical Justification: Overall Concepts

- **Follow-up questions:**
  - What defines an effective versus ineffective model?
  - How many subjects do we test?
  - What defines 'almost always'?
- **Answers to all 3 questions are inter-related**
- **Use standard statistical calculations to assess results under different assumptions**
  - Calculate probabilities using the binomial distribution

CDC Workplace Safety and Health NPPTL NIOSH

Statistical Justification: Initial Assumptions

- **Consider a model that is > 80% effective**
  - Should almost always pass the test
- **Consider a model that is < 60% effective**
  - Should almost always fail the test
- **Between 60% and 80% effective**
  - Expect variability in results
- **Sample size and defining 'almost always'**
  - Larger sample size will give more certainty

Statistical Justification: Selected Results with 25 Subjects

- **Require 15/25 (60%) to achieve acceptable fit**
  - A model which is 85% effective will fail <0.1% of tests
  - A model which is 55% effective will fail 62% of tests
- **Require 19/25 (76%) to achieve acceptable fit**
  - A model which is 85% effective will fail 7% of tests
  - A model which is 55% effective will fail 97% of tests
- **19/25 provides a better criteria**
  - Models in the effective range fail more often
    - 90% effective will fail <1% of tests
  - Far more certainty in rejecting ineffective models

Statistical Justification: Selected Results with 35 Subjects

- **Require 21/35 (60%) to achieve acceptable fit**
  - A model which is 85% effective will fail <<0.1% of tests
  - A model which is 55% effective will fail 66% of tests
- **Require 26/35 (74%) to achieve acceptable fit**
  - A model which is 85% effective will fail 3% of tests
  - A model which is 55% effective will fail 98% of tests
- **26/35 provides a better criteria**
  - Models in the effective range fail more often, but still rare
    - 90% effective will fail <0.2% of tests
  - More certainty in rejecting ineffective models

Statistical Justification: Selected Results with 50 Subjects

- **Require 30/50 (60%) to achieve acceptable fit**
  - A model which is 85% effective will fail <<0.1% of tests
  - A model which is 55% effective will fail 71% of tests
- **Require 37/50 (74%) to achieve acceptable fit**
  - A model which is 85% effective will fail 1% of tests
  - A model which is 55% effective will fail >99% of tests
- **37/50 provides a better criteria**
  - Models in the effective range rarely fail
    - 90% effective will fail <0.1% of tests
  - More certainty in rejecting ineffective models

Summary of Results

- **Requiring about ¾ of subjects to achieve acceptable fit gives optimal results**
  - Lower criteria often pass ineffective models
  - Higher criteria often fail effective models
- **Larger sample sizes give more optimal results**
  - Increase from 25 to 35 gives a larger improvement
  - Need to balance practical and statistical issues
- **Proposed Criteria: 26/35 achieve a TIL ≤ 5%**

Reproducibility of a Single Test Result

- **26/35 represents a minimally passing result**
- **Designed to achieve optimal results**
  - Given an effective model → pass
  - Given an ineffective model → fail
  - Converse is not necessarily true
  - Either passing or failing may reflect a marginally effective model (say 70% effective)
- **Reproducibility requires a higher standard than 26/35 (76%)**

Summary and Conclusions

- **Select 35 subjects based on the NIOSH panels**
- **Specify ≤5% TIL as acceptable fit**
  - TIL ≠ APF
- **Specify 26/35 as the minimum fraction of subjects required to achieve acceptable fit**
- **Yields optimal statistical properties**
  - Models which provide acceptable fit for at least 80-85% of subjects will pass a high % of the tests
  - Models which provide acceptable fit for no more than 60% of subjects will fail a high % of the tests
- **Caution required in interpreting a given test result**

Statistical Explanations for Development of the TIL Criteria

Disclaimer: The findings and conclusions in this presentation have not been formally disseminated by the National Institute for Occupational Safety and Health and should not be construed to represent any agency determination or policy.

Thank you

Visit NPPTL at: http://www.cdc.gov/niosh/npptl/default.html