Discussion: Reproducibility and Cross-study Replicability of Prognostic Signatures from High Throughput Genomic Data

Lee Dicker
Rutgers University

May 2, 2014
Rutgers Statistics Symposium
Discussion: Reproducibility and Cross-study Replicability of Prognostic Signatures from High Throughput Genomic Data

▶ “Leave-one-in” cross-study validation is a convincing principle for the evaluation of prognostic signatures.
Discussion: Reproducibility and Cross-study Replicability of Prognostic Signatures from High Throughput Genomic Data

- “Leave-one-in” cross-study validation is a convincing principle for the evaluation of prognostic signatures.
- What are the implications for the development of prognostic signatures?
Discussion: Reproducibility and Cross-study Replicability of Prognostic Signatures from High Throughput Genomic Data

- “Leave-one-in” cross-study validation is a convincing principle for the evaluation of prognostic signatures.
- What are the implications for the development of prognostic signatures?
  - Waldron et al. (2014)
    - Meta-analysis and validation of previously proposed prognostic signatures.
Discussion: Reproducibility and Cross-study Replicability of Prognostic Signatures from High Throughput Genomic Data

▶ “Leave-one-in” cross-study validation is a convincing principle for the evaluation of prognostic signatures.
▶ What are the implications for the development of prognostic signatures?
   ▶ Waldron et al. (2014)
     ▶ Meta-analysis and validation of previously proposed prognostic signatures.
     ▶ Aggregation of prognostic signatures for improved performance?
**Discussion:** Reproducibility and Cross-study Replicability of Prognostic Signatures from High Throughput Genomic Data

- “Leave-one-in” cross-study validation is a convincing principle for the evaluation of prognostic signatures.
- What are the implications for the development of prognostic signatures?
  - Waldron et al. (2014)
    - Meta-analysis and validation of previously proposed prognostic signatures.
    - Aggregation of prognostic signatures for improved performance?
  - Bernau et al. (2014)
    - Multistudy comparison of classification algorithms.
Discussion: Reproducibility and Cross-study Replicability of Prognostic Signatures from High Throughput Genomic Data

- “Leave-one-in” cross-study validation is a convincing principle for the evaluation of prognostic signatures.
- What are the implications for the development of prognostic signatures?
  - Waldron et al. (2014)
    - Meta-analysis and validation of previously proposed prognostic signatures.
    - Aggregation of prognostic signatures for improved performance?
  - Bernau et al. (2014)
    - Multistudy comparison of classification algorithms.
    - Stability of classification rules across studies?


**Discussion:** Reproducibility and Cross-study Replicability of Prognostic Signatures from High Throughput Genomic Data

- “Leave-one-in” cross-study validation is a convincing principle for the evaluation of prognostic signatures.
- What are the implications for the development of prognostic signatures?
  - Waldron et al. (2014)
    - Meta-analysis and validation of previously proposed prognostic signatures.
    - Aggregation of prognostic signatures for improved performance?
  - Bernau et al. (2014)
    - Multistudy comparison of classification algorithms.
    - Stability of classification rules across studies?
  - Trippa et al. (201X)
    - Clustering studies based on leave-one-in cross-study validation.
Discussion: Reproducibility and Cross-study Replicability of Prognostic Signatures from High Throughput Genomic Data

- “Leave-one-in” cross-study validation is a convincing principle for the evaluation of prognostic signatures.
- What are the implications for the development of prognostic signatures?
  - Waldron et al. (2014)
    - Meta-analysis and validation of previously proposed prognostic signatures.
    - Aggregation of prognostic signatures for improved performance?
  - Bernau et al. (2014)
    - Multistudy comparison of classification algorithms.
    - Stability of classification rules across studies?
  - Trippa et al. (201X)
    - Clustering studies based on leave-one-in cross-study validation.
    - Study-level covariates and standards for genomic studies?
Discussion: Reproducibility and Cross-study Replicability of Prognostic Signatures from High Throughput Genomic Data

- Cross-study validation is a more reliable statistical principle than cross-validation.
Discussion: Reproducibility and Cross-study Replicability of Prognostic Signatures from High Throughput Genomic Data

- Cross-study validation is a more reliable statistical principle than cross-validation.
- However, can clinically useful genomic signatures be derived from statistical principles alone or is scientific validation necessary?
Discussion: Reproducibility and Cross-study Replicability of Prognostic Signatures from High Throughput Genomic Data

- Cross-study validation is a more reliable statistical principle than cross-validation.
- However, can clinically useful genomic signatures be derived from statistical principles alone or is scientific validation necessary?
  - Should the need for scientific validation drive statistical methodology (e.g. hypothesis generation), as opposed to optimizing a statistical criterion?