

## POSTER ABSTRACTS

11<sup>th</sup> Annual HMO Research Network Conference

April 4-6, 2005 Santa Fe, NM

### Methods 52

#### **Probabilistic Corrections to Medical Record Abstract Data Using Inter/intra-rater Reliability Data from a Subset Sample**

Timothy L Lash, Rebecca A. Silliman, Soe Soe Thwin, Kerri Clough--Gorr

**Background:** The quality of data abstracted from medical records is often characterized by studies of intra- or inter-rater reliability in a subsample of the study population. Often these data are simply presented to demonstrate data quality. Less often, the observed error rates are used to calculate revised estimates of effect. The authors extend Monte Carlo methods to studies with internal reliability data to calculate a revised estimate of effect and a wider interval to account for both the reliability data and the uncertainty in the reliability dataset itself. The authors illustrate the method with an application to a study of breast cancer patients for which data collection was by medical record review.

**Methods:** Medical records were reviewed and abstracted for Stage I and II breast cancer patients diagnosed between 1990 and 1994 at six Cancer Research Network HMOs. For a subsample, inter-rater and intra-rater reliability data and an adjudicated gold standard were available. We calculate the conventional estimate of effect and its interval associating cancer stage with recurrence risk. This conventional estimate of effect is compared with the estimates and intervals corrected with the reliability data, and then with estimates and intervals that take account of the uncertainty in the reliability data.

**Results:** As expected, the risk of recurrence is associated with increasing cancer stage. Corrections to the estimate of effect using the reliability data show that the interval grows wider as the sources of uncertainty taken into account grows more complete.

**Conclusions:** While many studies use reliability data only as a metric for data quality, a more thorough approach can quantitatively depict the uncertainty in the estimates of effect.