

## Concurrent Session A1-1

### Performance Characteristics for a Master Patient Index (MPI) within the NM Health Information Collaborative (NMHIC)

Tracy Smith, BA, Lovelace Clinic Foundation; Scott Robinson, MA, MPH,  
Lovelace Clinic Foundation; Herbert T. Davis, PhD, Lovelace Clinic Foundation;  
Eva Lydick, PhD, Lovelace Clinic Foundation

**Aims:** To develop and test an MPI capable of linking healthcare records for a single individual from different healthcare systems.

**Background:** Healthcare in the US is delivered in a fragmented manner that promotes service duplication and errors related to missing information. Patients suffer waste and harm, while healthcare dollars are spent inefficiently. Considerable current initiatives are on-going to connect medical care practitioners with patient information which already exists in electronic health information silos. In October 2004, NMHIC was created as a 3-year project, with financial support from AHRQ and community partners. A major first step in developing this exchange was creation of an MPI or record locator service to match patient identifiers from multiple information sources to a single patient. Consultation with stakeholders indicated that a false positive (FP) error which returned the wrong patient's record was worse than a false negative (FN) error which failed to retrieve all of a single patient's records.

**Methods:** Numerous methods of linking records were considered. The final selection uses a probabilistic name-matching algorithm that is a Root Mean Square (RMS) of three commonly used scores, social security number (SSN), gender and date of birth (DOB). SSNs, DOB and gender were scored on an exact match while a RMS of 0.8 was required for name-matching.

A decision tree based on these elements provided the ultimate link/no-link decision.

**Results:** To test the algorithm, two health databases were used. The first database, with approximately 96,000 subscribers was used as the index database. A random sample of 40,000 patients in the second database was used as the incoming data. These two databases were selected as there is a reasonable overlap in their service populations. This created almost 4 billion comparisons, which were too many to manually review. We used a cut-off score of .76 on the RMS as true matches very rarely had an RMS score below that threshold. As a result 116,204 cases required manual review. A second independent reviewer adjudicated all of the 115 disagreements reducing the disagreements to 16 FN and no FP matches out of 4 billion comparisons. This resulted in essentially 100% sensitivity and specificity with a 95% confidence interval of (.9994,1) for the specificity.

**Conclusions:** The MPI developed for NMHIC had adequate performance characteristics for use in a regional healthcare information exchange.