



Creating Master Patient Index Records: A Zero Based Data Quality Solution



The Missing Piece

When a North American state health department with over 15 years of patient data decided to move to an integrated disease/clinical management system their biggest challenge in creating master patient index records was bad data and a lack of appropriate software to assist them in the process. The existing system was a home grown legacy platform with data spread across multiple platforms and the data housed in the system had intrinsic issues including duplicate data, missing fields, incomplete information, and erroneous/incomplete addresses. The modern target system lacked the ability to handle the MPI process. The agency tried to use traditional approaches of scripts and custom programming to cleanse and match the data but the results were poor.

The Critical Piece: OpenDQ

Comprehensive. Easy to use. Cost effective.

OpenDQ™ solution is a complete data quality solution that provides data quality coverage for all enterprise data in a fully integrated solution. With its innovative design, OpenDQ™ bridges structured and unstructured data sources and integrates them into a single solution for comprehensive data quality processing. This integrated approach creates maximum efficiency for organizations with the elimination of complex interfaces, and custom programming previously required to address structured and unstructured data. OpenDQ™ identifies data quality problems and takes corrective actions to address issues now and in the future.

OpenDQ™ addresses key data quality concerns including:

- *Comprehensive Data Profiling*
- *Record linkage (Data matching) between data sources*
- *Elimination of duplicate records*
- *Data Standardization*
- *Address verification and correction*

With OpenDQ™, organizations have the power to develop, deploy and monitor comprehensive enterprise wide data quality programs, while preventing new data issues from occurring. The OpenDQ™ solution integrates all of the key components of data quality with ETL (extract, transform and load) software, creating unmatched efficiency in process control, data flow, and data processing.

The Complete Solution: Puzzle Solved

Infosolve Technologies experienced team of data quality specialists initiated the project with comprehensive data profiling of all the data elements in the Legacy patient management system to identify missing fields, incomplete data, incorrect data and bad formats. After the initial data cleansing removed any noise, the data was put through a comprehensive data cleansing process to create standard and consistent records. This process included:



- *Establishment of data dictionaries to standardize key data elements, such as names and abbreviations.*
- *Enhancement of data from both systems to implement address standardization and data appends through third party systems.*
- *Data matching and assignment of unique identifiers to patient records using Fuzzy Match techniques on multiple columns.*
- *Development of custom business rules to effectively merge the matched data.*
- *Comprehensive review and elimination of duplicate records to assure all records were unique and false positives were removed from the data sets.*

Once all the data completed the cleansing process, a global identifier was assigned to all the unique records. With a complete and consistent view of all of their patient data the agency could readily move the data into the modern Disease/Clinical management system. The highly complex process of creating the master patient index record was achieved with minimum effort at significant cost efficiencies.