Automate Data Validation to Reduce Errors, Save Cost

Background and Challenge

A leading multi-state healthcare organization was converting and importing a large volume of data from a legacy lab system into Epic's Beaker application. Before importing this data into the Production environment, they needed a solution that could make sure the imported data matched the exported data and identify any potential errors.

The traditional manual method for an effort like this is time intensive, error-prone, and costly, requiring 40+ temporary chart reviewers to review the data for errors. Finding a solution they could trust not only for this first batch of lab data but also for future import and export matching required a team of technical and analytics import experts.

Results

50,000
records validated automatically

+99.7%
data validation matching

$150K–$250K
initial labor cost savings

520
working days of manual time saved

Solution

The client was seeking a cost-effective solution that would reduce effort, promote accuracy, save time, and be repeatable for future validation projects. If the client was unable to ensure a near-perfect matching of the imported data before moving it into the Production environment, the migration would be canceled. The data being imported is used to treat patients and track how they progress over time, and altered lab results could place a patient at risk.

The plan was to take a sampling of data from the legacy lab system, import that data into a copy of the Production system, and then use a raw pre-imported dataset to perform an automated validation. Results would be presented in a sortable format with a separate error log that could be used to make corrections. The automation would help remove human error, increase the rate at which this work could be completed, and provide an easily repeatable process.

Selecting the Tools

The client provided the raw legacy lab files as the source of truth and access to their Dev Clarity Beaker environment which housed the data needing to be compared. We built a standalone SQL database within the client's infrastructure and configured the appropriate database connections on a virtual desktop so this work could be performed remotely. The local IT team installed required tools on the virtual desktop, including Visual Studio, SSMS (used to deploy the SSIS package), Oracle SQL Developer, and Notepad++. These tools allowed our data architect and engineer to query and extract data from Clarity, import data into the standalone SQL database, code and run the automated validation job, and export results for review. We met with the client's lab conversion team after each round of validation to determine the validity of errors, help identify root causes, and receive feedback to make necessary changes for the next round of validation.

Testing and Honing the Process

The first round of validation was performed on approximately 10,000 cases with a matching percentage rate of 89%. The second round of validation was performed on 20,000 cases with a matching percentage rate of 99.74%. A third and final round of validation was performed on 20,000 cases with a matching percentage rate of 99.76%. A total of 50,000 cases (277,763 rows of data) were compared over a period of eight weeks. It was estimated that manual review would take about five minutes per record, including time to load the record in each system, review each data element, document discrepancies, and close the records. This amounts to 4,166 working hours (~520 days) to manually review all 50,000 records.

Providing Guidance for the Future

Along with delivering the automated solution, our team provided two training sessions and completed written instructions for running the automated process we developed. The training included how to make changes to the SSIS package and configuration files to allow the solution to be used for other Epic and non-Epic validation efforts. The flexibility of using this solution for any future validation effort will produce long-term cost savings. If the client needed support with this solution in the future, an hourly advisory provision was included.

> **❝ Definitely saved us a bunch of money (and) helped us identify a big gap in the data.**
> —Anonymous

Tegria

Outcomes

Go Live with Confidence

The automated solution, with a data validation matching percentage reaching 99.7%, gave the client complete confidence that the legacy data they migrated into the Epic Beaker Production environment was of the highest quality. To ensure that absolutely no errors fell through the cracks, the .3% with discrepancies were sent to an error log for correction before being imported.

Reduce Stress, Save Money Now and Later

The solution decreased the workload levied on an already strained project team by providing an automated system and repeatable process to be used for immediate and future validation efforts. The estimated initial labor cost savings were between $150,000 and $250,000, with projected exponential future savings as the solution is redeployed for other projects.

connect@tegria.com   tegria.com