

Evaluating ASR Productivity

Productivity Calculations and Comparison Methods Help Evaluate Increasingly Popular Technology

by Mark Ivie

Automated speech recognition (ASR) has been around for two decades. However, after years of unfulfilled promise by early pioneers, recent technology developments have made ASR a viable technology for healthcare organizations. This article provides different productivity calculations and comparison methods for organizations to evaluate ASR's usefulness.

Productivity Factors

ASR creates a text file from a voice file using a speech recognition engine. It can provide quicker document turnaround time, highly accurate recognition, lower transcription costs, and faster access to medical information. ASR can increase transcription productivity by increasing output.

Since different suppliers use different ASR productivity calculation methods, it is impossible to make apples-to-apples comparisons without first understanding the factors influencing these calculations.

In order to accurately compare ASR productivity, you must first know the data collection starting point. Was the baseline data from the same transcription platform as that used after ASR was introduced?

As with any technology, user training and experience strongly influence a successful implementation. The baseline productivity measure should consider the background and experience level of the medical transcriptionists (MTs) and medical editors (MEs) who use ASR and their familiarity with the dictators, clients, work types, and platform for transcription and editing.

ASR is suitable for most dictators. The technology interprets what the physician says, not what the physician meant to say. As a result, dictation must be edited to ensure accuracy. The more corrections that MEs have to make, the lower the productivity. Correction effort statistics can be related to several technical aspects including the ASR engine's dictation audio quality filter,

which decides required audio quality for routing to an ME; the quantity of dictator audio files sent to the ASR for new dictator learning; and the ASR engine's process to "learn" the dictator's audio patterns.

In general productivity measures relate an input used to produce an output. An example of input is the time it takes to transcribe or edit a dictation. Examples of output include the number of lines produced or dictation minutes processed.

ASR suppliers may define inputs and outputs differently. It's risky to compare productivity values based on different input and output definitions. For example, does "time to edit" include inactive MT/ME time? If so, what's the definition of "inactive time"? Other issues to consider include the definition of "quantity of lines produced."

Productivity Calculations

Two common productivity representations are lines per hour and editing ratio. Lines per hour are the lines produced in one hour of transcription or editing. For example an MT who types 300 lines an hour would have a productivity level of 300 lines/hour. Increasing the quantity of lines produced increases the productivity values.

The editing ratio is the minutes spent editing or typing versus the corresponding minutes of dictation audio. An ME who spends 30 minutes editing 10 minutes of dictation would have a productivity level of 3 min editing time/min dictation audio time. Decreasing the time to edit dictation increases productivity.

To further interpret productivity values, consider:

- ▶ Sample size rationale. The sample size used in the productivity study and the sample's distribution of clients, work types, and dictators compared to typical distribution of the MT/MEs' workload.
- ▶ Data collection period. The time

period for data collection. Longer time periods may better represent the variety of dictators, work types, and clients normally encountered.

- ▶ Average, median, or aggregate used. Average productivity values may be skewed by unusual jobs (e.g., high line counts produced in a short period of time). Median represents the exact data midpoint and cannot be skewed by outliers. Aggregate productivity totals the inputs and outputs (e.g., total lines produced in a 40-hour week). This may fairly represent the total work normally encountered by the MT/ME.

By comparing typing and editing productivity data, the percentage productivity change from before versus after using ASR can be calculated. Potential ASR adopters must understand how a given ASR supplier calculated its percentage productivity value.

Three components are used in the percentage calculation: typing baseline productivity, editing productivity, and typing and editing comparison methods. To calculate the before-versus-after productivity related to typing productivity, subtract the editing edit ratio from the typing edit ratio and then divide by the typing edit ratio. To calculate before-versus-after productivity related to editing productivity, subtract the editing edit ration from the typing edit ratio, and then divide by the editing edit ratio.

When evaluating ASR productivity claims, it pays to ask questions about influencing factors and productivity measurement methods. For maximum results, make sure you are seeing the entire picture, including transcription platform, ASR engine, and best practices by physicians and editors. ♦

Mark Ivie is senior vice president and chief technology officer at MedQuist.

