



Unlocking the Power of NLP in Pathology Casefinding

Complete Webinar Q & A

How are pathology reports sent to CRStar? Is it by fax, upload, or other?

Path reports are sent to CRStar electronically using HL7 messaging from the lab system.

Sometimes there are multiple radiology reports. How does the system know which ones to copy? Will the text box character limit be increased? It looks all copied.

The NLP Reportability Solution does not currently process radiology reports for casefinding; however, this is on the development pathway as a future feature.

Will this process still require manual review of the cases deemed reportable?

HLA's NLP engine's pipeline has multiple stages that maximizes reportability accuracies. The reportability engine makes an initial assessment of a document's reportability. Documents with a classification probability below a high confidence threshold are passed to a post-processing system that uses a number of different strategies to complete a second assessment of reportability. Documents are assigned one value from the set of {Reportable, Non-Reportable, Review}. The process means that the user can rely on the reportability results for Reportable and Non-Reportable while keeping in mind that there will be a marginal error rate that can be corrected over time and assuming that the system goes through an initial Quality Assurance (QA) on implementation.

The {Review} category is the system's way of ensuring that a report with a confidence rate below the threshold is specifically separated out for a manual review by an ODS (typically <10% of reports fall in this category but it varies from one facility to another).

Over time, missclassified and manual review reports are examined by the HLA ODS team to make manual determinations and then added to the training set of documents for the next cycle of model building using machine learning in rigorous process of continuous improvement in order to maintain high accuracy levels.

Is this software live and ready for purchase?

Yes, the CRStar software with NLP integration is currently available. The NLP service can always be utilized as a standalone service by central cancer registries (e.g. California Cancer Registry currently uses HLA-Global's NLP solution for reportability).

Is this doing any autocoding? If so, what is the specificity and sensitivity? Again if so, is there differences in accuracy depending on site?

The fields that are currently autocoded are listed on Slide 12 of the handouts. There is no difference in accuracy between sites. Additional auto coding fields are on the development pathway as a future feature.

Does installation cause any downtime?

There is minimal to no disruption as part of the installation.

Will it include outside path consults?

If the external path systems are electronically sending the reports to the hospital/facility, then it can be forwarded to CRStar and subsequently the NLP service; however, this would need to be assessed on a case by case basis depending on the current workflow of the facility.

What is the accuracy rate of the reportable cases? Do you have to review each path report or can you rely on reportable/non-reportable tags?

This is a statistical service so an error rate should be expected. However, HLA-Global's service has a market-leading accuracy rate that exceeds 95% in most installations for reportability. Specific accuracies on False Positives (FPs) and False Negatives (FNs) vary from site to site. The system also flags reports that it cannot determine with high confidence as "manual review" to reduce the risk of errors. Our approach includes a baseline test as part of the initial implementation to assess accuracy and any necessary adaptations to ensure accuracy is maximized. Regular audits are also expected to ensure continuous process improvement.

Will NLP be able to determine if the case is a new primary cancer before the end user deletes the case?

Yes, in the work queue there is an option to view patient details. See Slide 11 of the handouts. This will display the info on other primaries already in the database.

This seems to just copy your existing path that a registrar can easily find in their system anyway.

This is entirely different. This searches both positive and negative pathology records and will determine the reportability of a case for the user and put the case in the work queue. Reportability cases are then transferred automatically to the abstract as a suspense case without any manual entry.

We get a path-finding file that gets created each morning by our ERA Enterprise team. They run our pathology through a query and pulls reportable ICD-10 codes. When we run it through CRStar, it creates suspense cases. So what would the advantages of this service be since our method grabs "reportable" cases as well. Thanks!

Even as you address the reportability of the case finding effort, there is still a lag in the time it takes for patients to come through admission and be assigned an ICD10 code. Also, there could be future opportunities to reach deeper into the pathology report for further data mining with the NLP solution.

Do you have to be a CRStar client to utilize this service?

Yes. You do need to be a current CRStar client to utilize this solution. You are already under contract with us and no additional contract with HLA global is required.

What is the cost and how can I get a quote?

Pricing is based on the total number of pathology records, both positive and negative from each pathology lab in your health system. Our contact information is listed on the last slide of the handout. We would be happy to provide more information as well as a quote.

Does this include radiology reports?

Not at the moment, but it is in the pipeline for immediate development.

In addition to the data items you presented on the slide, can additional data fields be included such as info from the synoptic op reports?

Not at the moment. Synoptic reports are complicated and vary not only between health organizations but between clinical specialties within the one organization. This feature would require a specific project with the client as opposed to being part of the standard service.

What resources are needed from my facility, such as IT support?

From the client end, a pathology IT resource is needed. This person would have access to the pathology system and be able to report some high-level questions, such as providing CLIA numbers and be able to help select specific pathology records for the sample data sets needed in the development process. An Interface IT resource would also be necessary. This individual would be responsible for setting up and sending the HL7 messages to the ERS interface engine when needed. CRStar will provide a project manager. A project manager on the client side would be optional. And the time it takes to implement is highly dependent on the availability of the client's ODS and IT staff. We would set up a timeline during the planning phase of the project after identifying the necessary resources.

Additionally, there is usually an onboarding process for any new facility to the NLP service which includes a sampling request. Generally, we request a 1000 reportable report and a 1000 non-reportables report from your historical records to do a baseline test and model training if needed. We would then expect some form of communication with the cancer registry on any clarification of local rules and error analysis estimated at 1 week of an experienced ODS.

How much time savings would my registry see?

We estimate that registries spend 15% of their time on casefinding processes and manual methods. So, this would be a tremendous time saver. After the confidence level is reached, the solution allows users to accept the NLP findings, meaning users will no longer be required to review the pathology records and will not have to manually enter these records, saving a significant amount of time that can be spent on other import registry initiatives.

I currently use an automated pathology interface into CRStar. Would this replace that?

This would replace the automated pathology interface currently in use. The advantages of this solution vs your current interface is that the reportability is determined for you and the cases are put into a work queue for you where you can access the reportability before the case is put into suspense.

What is the accuracy rate of reportable vs non-reportables?

We expect our service to maintain a performance level above 90%. We generally say that it is 80%+ out of the box, and 90%+ with training. Additionally, reports that are determined to have a borderline classification are reviewed and added to the training set and then the machine learning model is re-calibrated to ensure continuous learning and improvement. The client can also request an accuracy review. HLA-G would advise on the size of the training materials needed to be added to the training set.

Support available at www.mycrstar.com