Advancing Care, Boosting Physician Efficiency Through AI

Supported by PREMIER

ADVANCING CARE, BOOSTING PHYSICIAN EFFICIENCY THROUGH AI

As the tide of denials continues to rise, and as the prospect of aggressive regulatory oversight comes into view, the need for accurate and efficient clinical documentation integrity (CDI) practices couldn't be clearer.

CDI has a critical role to play across all domains of healthcare, including compliance, quality, and revenue. That role, however, can be substantially aided by the development of artificial intelligence (AI) technologies.

In 2022, WellSpan Health and Premier's Stanson Health collaborated on an AI-driven solution to help improve clinical documentation accuracy—in particular, Hierarchical Condition Category (HCC) capture—as well as increase coder efficiency and physician engagement.

Recently, this WellSpan-Premier collaboration was highlighted in an ACDIS-facilitated webinar. The following report is a summary of the information in that webinar and outlines the general features of the AI-driven solution, as well as the results derived from its usage.

WELLSPAN'S OBJECTIVES

WellSpan Health is an integrated health system based in central Pennsylvania. It has more than 23,000 team members, including 3,250 providers supporting a large accountable care organization.

Approximately 28% of WellSpan's revenue comes from value-based care arrangements, according to **Julie Platt, MD, MBA, MS, HCLM**, medical director of clinical transformation for WellSpan Medical Group.

"Our strategy prioritizes clinical accuracy and provider engagement. HCC risk capture is framed not as a coding exercise but as a reflection of the true patient complexity," Platt says. "This matters because our care management workflows are directly informed by the risk scores, enabling a more personalized proactive care."

When WellSpan reached out to Premier in 2022, it had a simple goal in mind: to utilize AI to "surface the right diagnosis at the right time with the requisite clarity and context" so that providers could "act with confidence," says Platt.

In practice, this meant developing an AI-driven solution grounded in clinical logic, delivered within the provider's workflow, and designed to reduce interdepartmental friction while improving the accuracy of HCC capture and documentation.

Because WellSpan's program operates in many aspects of the CDI space, this desired vision manifested differently in accordance with the demands of each corresponding workflow:

- *Pre-visit*: WellSpan envisioned "curated suspect and recapture diagnoses" routed to the CDI team for review. In particular, this included a targeted problem list cleanup (i.e., ensuring that only accurately validated conditions flow into Epic HCC alerts).
- *Post-visit*: WellSpan also envisioned "real time AI MEAT [Monitoring, Evaluation, Assessment, Treatmeat] detection" within the clinical note, which could then be forwarded to coders. From there, the diagnosis could be confirmed or a query could be sent if required.

"This model strikes a balance between accuracy, efficiency, and accountability while reducing friction across the care team. In summary, we wanted to leverage AI in all portions of the workflow," says Platt

According to Ryan Nellis, vice president and general manager for Premier's Stanson Health, approximately 30% of the data in a patient's chart is stored in electronic health records (EHR) as "discretely documented data," meaning structured items such as laboratory values, medications, dosages, and vital signs.

For an AI solution, structured data tends to be easier to process; however, the other type of data—unstructured data—tends to be much more difficult. And as Nellis notes, "the vast majority of information about patients is still stored in unstructured blurbs and blocks of text, dictations, transcriptions, free type, and sentences."

Consequentially, Premier's technology needed to employ natural language processing (NLP) to ensure that the unstructured portion of the patient's record was also processed; however, this was only one aspect of the problem.

In order for the NLP AI technology to be useful for CDI specialists, providers, and coders alike, it also needed to be context-sensitive: That is, Premier had to build a system that could understand both the structured and the unstructured data, as well as how these two aspects interacted with each other within the framework of the patient's chart.

"Our strategy prioritizes clinical accuracy and provider engagement. HCC risk capture is framed not as a coding exercise but as a reflection of the true patient complexity."

— Julie Platt, MD, MBA, MS, HCLM, medical director of clinical transformation at WellSpan Medical Group

For example, suppose a patient's chart contains a clinical note related to a "history of cancer." For the AI technology to be effective, that note needs to be further specified to include (or exclude): family history of cancer (e.g., grandmother, mother), a suspicion of cancer (e.g., malignant tumor removal), or a patient's individual experience with cancer (e.g., previous diagnosis).

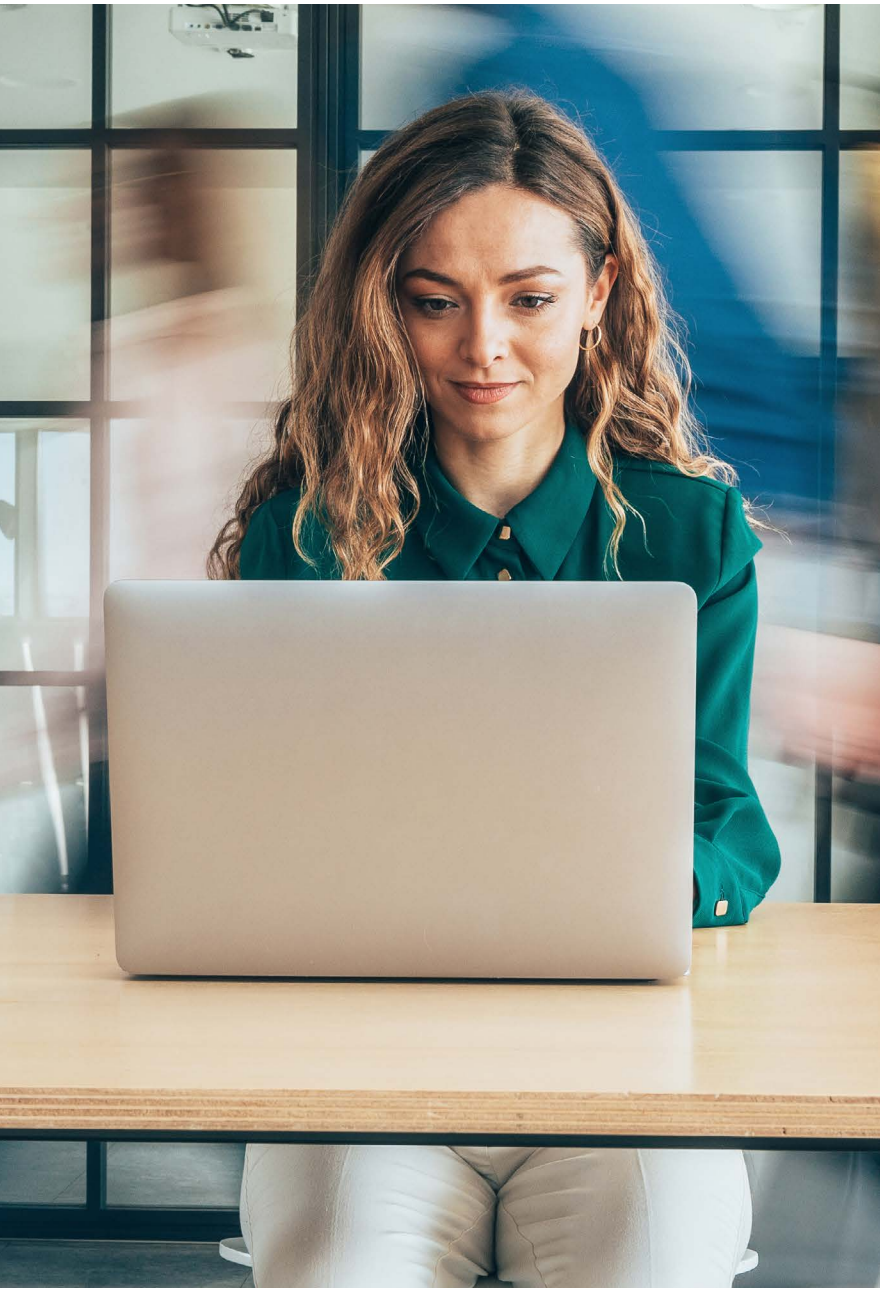
"We are using fine-tuned large language models built on vast data sets," Nellis explains. "But our goal is not only to have human users trust this, but to have higher compliance, higher accuracy, and, ultimately, see increasing and better follow-up rates."

To this end, Premier designed a pre-visit and post-visit capability called "CodingCare," which allowed the AI tool to scan through every chart, organizing them by physician, clinic, and insurance plan.

In particular, Premier developed a set of integrated capacities within the AI tool such as:

- *Recapture/revalidation opportunities*: Pre-visit, CDI specialists can utilize their EHR dashboard to see the list of patients, along with their upcoming visits and appointments. Based on a patient's data set, the AI tool examines "revalidation opportunities" where clinical differences between visits may necessitate a coding change.
- *Case flagging*: Post-visit, the AI tool can efficiently examine the clinical notes associated with a patient and flag cases where there is congruency between notes and documentation. For cases where there is a mismatch instead, the tool can flag them for further review.
- *Suspected condition logic*: As noted previously, for first-time coding reviews in the pre-visit space, the AI technology develops a work list of suspected conditions that is accessible to the CDI team and can be pushed to providers.
- *Clinical context*: Providers who receive a nudge to add relevant HCC codes for a visit can see all of the relevant information (e.g., prior visitation, specific notes) justifying a given recommendation.

PILOT RESULTS



According to Platt, utilizing Premier's tool has reaped dividends in terms of efficiency and response rates, while staying within the regulatory bounds of CDI compliance.

For instance, in WellSpan's post-visit space, AI MEAT-driven detection has increased by 15%, and the average number of charts reviewed per day has nearly doubled.

"Our average charts per day are significantly higher than the manual review, which we were at 43 charts per day. It just gets better and better as we go along," Platt says.

For WellSpan's new ambulatory CDI program, the "defect rate"—i.e., CDI messages sent to providers requesting a correction—has also declined rapidly.

In their pilot program, Platt specifically selected 10 providers who were "struggling with their coding," and since the implementation of Premier's program, these providers have started to "trend downward" this past April and May.

Additionally, since the inauguration of the pilot program in the fall of 2023, there has been a 74% increase in engagement among WellSpan's care managers.

Average risk adjustment has also seen noticeable improvements. According to Platt, "over the first two years of this pilot, our average risk adjustment data validation scores within WellSpan's population was 7.15% higher than patients seen in the same population but outside of WellSpan, demonstrating both the improved capture and clinical alignment."

However, Platt notes that the clearest gains for WellSpan are seen in post-visit coder efficiency.

"By leveraging AI, we've improved coder productivity by at least 15%. Fifteen percent of the charts are now skipped because they're compliant with at least 95% accuracy, and we have verified that a few times," Platt says.