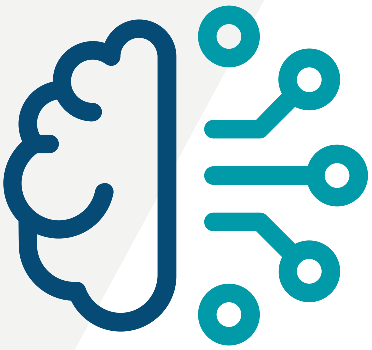
HEALTH PARTNERS

Increasing the efficiency of the medical record review process has traditionally focused on increasing the speed of traditional medical record reviews.

Learn how AI is disrupting this industry and how to harness this technology to increase the efficiency and ROI of the review process.

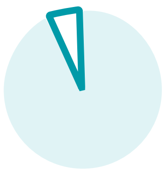


Reviewing medical records is a labor-intensive process

The medical record review process is labor-intensive and traditionally requires a reviewer that is a subject matter expert with respect to the specific type of review being performed. References to the application of artificial intelligence (AI) to detect clinical events and otherwise optimize this process can be seen in academia as far back as 2005. Still, the medical review industry has yet to see the same exponential improvements and disruptions caused by using AI and machine learning as other fields.

In 2020, 100 healthcare executives were polled about AI, and while 90% reported they had a strategy in place for AI technology, only 7% were fully implemented and operational.

90%
had an AI strategy in place



7%
were fully operational

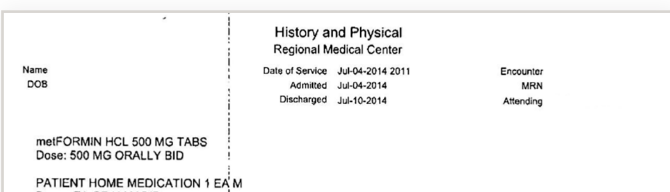
Medical record reviews present several unique challenges to AI application, and AI's capabilities have improved significantly since 2005. This whitepaper enumerates some of these challenges; it also details the algorithms and other approaches available to overcome difficulties and significantly improve the efficiency of the review process.

Cognitive Burden: What am I looking at?

The standard process for medical record review often involves scanning or otherwise referencing printed records. Electronic medical records and electronic health records have promised to address this issue, but adoption within the medical review field still has room for improvement. Paper records do not adhere to standard formats; this increases the cognitive burden of the review process, increases the risk of review errors, and requires significant effort to extract the desired information from a record printout.

Printed records lack standard formats

Every page in a document requires time to identify. Reviewers may look for a Medication Administration Record (MAR), but identifying documents isn't free. The mental load of determining if a page is a lab report, imaging diagnostics report, discharge summary, etc., requires mental effort and takes time. These two images are both pages from printouts of the history and physical portion of a medical record. However, at first glance, they do not appear to be similar.



History and Physical
Regional Medical Center

Name
DOB

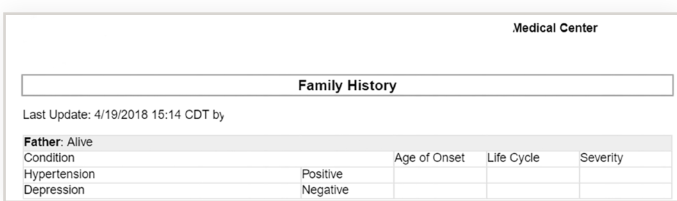
Date of Service Jul-04-2014 2011
Admitted Jul-04-2014
Discharged Jul-10-2014

Encounter
MRN
Attending

metFORMIN HCL 500 MG TABS
Dose: 500 MG ORALLY BID

PATIENT HOME MEDICATION 1 EA M
Dose: EA ORALLY BID

An example of a printout of a medical history portion of a medical record.



Medical Center

Family History

Last Update: 4/19/2018 15:14 CDT by

Father: Alive

Condition		Age of Onset	Life Cycle	Severity
Hypertension	Positive			
Depression	Negative			

An example of a printout of the family history portion of a medical record.

12

seconds for a reviewer to categorize a page

20%

of the time reviewers did not agree on page type

5

seconds to process a 250-page record with AI

How quickly can humans perform this task?

As part of developing an AI model for document classification, several reviewers and clinicians manually classified pages of medical records using tags like "discharge summary," "emergency room report," and many other possible categories. Tens of thousands of pages were viewed to determine the document's type (as opposed to interpreting information and measurements on the page). From our internal testing, on average, each page **took approximately twelve seconds to categorize.**

- When three reviewers were given an identical page, 20% of the time, reviewers disagreed about the page type.
- A group of three reviewers disagreed on roughly one out of every five pages.
- Three percent of the time, the three reviewers each picked a different page type; roughly three of every thirty-five pages reviewed were such that no two reviewers agreed with each other.

These results imply that, on average, page type identification for **a medical record of two-hundred and fifty pages will result in approximately thirty pages with identification errors**. It would take roughly fifty minutes for a single reviewer to process that record.

Can AI do it?

Yes. Modern AI techniques that fall under the natural language processing (NLP) category can use deep learning to extract a detailed numeric representation of a page. This numeric representation of a page (known in the field as an embedding) means little to humans. However, AI can use these page embeddings to learn a deep representation of the page content without being explicitly taught medical knowledge or concepts.

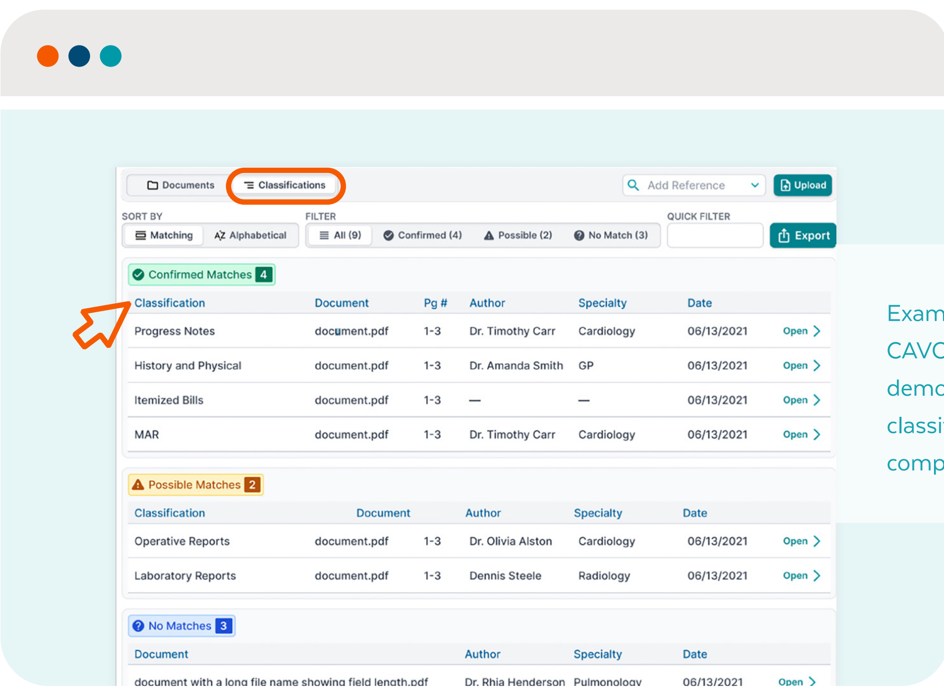
After training a model using approximately one-and-a-half million pages of medical records, Advent Health Partners developed our AI model for page classification. This model could process a two-hundred and fifty-page record in under five seconds, with an overall accuracy of around 85–90% (compared to the manual review group's 80% agreement rate). Our classification model is currently in production and being used by multiple clinical groups to increase the efficiency and accuracy of medical record reviews. Modern cloud architectures allow almost infinite scalability for models—Advent currently processes millions of pages each week and can scale up to handle just about any load.

Refer to the section covering "augmented AI" on page 14 for information about how portions of this task can be fully automated.

So what?

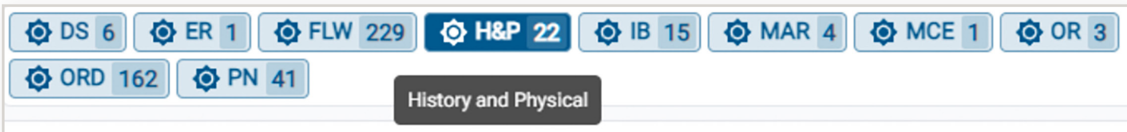
When reviewing medical records that can exceed a thousand pages, any amount of time saved per page becomes significantly more valuable. By automating the classification process, several efficiencies can be achieved:

Medical record completeness. When medical records are received, significant time pressures may require a rapid determination if the record is complete or if additional records must be requested. To manually determine that a record is missing a discharge summary, a reviewer has to identify every page and confirm that it is not a discharge summary. Using AI, reviewers can see the types of pages in a record at a glance.



Example view within CAVO platform demonstrating document classification for record completeness verification

Searching and navigating within records. Review tools can often apply Optical Character Recognition (OCR) to scanned pages which converts the images to searchable text. With prior knowledge of page types, searches can be performed across the entire record. Automated classification allows reviewers to restrict a search for vital signs to only laboratory reports, thus excluding potentially irrelevant documents like patient education packets.



Screenshot of the CAVO platform in classification view, used for navigating document types.

These use cases are not theoretical—Advent Health Partners' CAVO platform currently processes millions of pages of records each week. A twelve-thousand-page document takes less than 30 seconds to process.

Unstructured Structure: Tabular data lacks uniform representation

Several different types of information within a medical record are presented in a very similar, but not quite identical, format. One of the most common examples is an itemized bill (IB). Many review types require using itemized bills for policy validation purposes or otherwise interpreting items from an itemized bill.

For example, the two pictured IBs below both present an itemized list of charges. Minor differences like column orders and code formats increase the cognitive complexity of the review task and present an opportunity for errors to be introduced.

Date	Service	Rev Code	Procedure Code	Description	Qty	Amount
10/17/2016	049754	271		ELBOW IMMOBILIZER	5	330.00
10/17/2016	049768	271		LIMB HOLDER QUICK RELEASE 2	1	27.00
10/17/2016	049871	272		KIT VENTRICULOSTOMY 21512	1	886.00
10/17/2016	050052	272		CONNECTING TUBE SUCTION 6	2	46.00

A Raw IB example with a list of itemized charges. The columns read "Date, Service, Rev. Code, Procedure Code, Description, Qty, Amount."

Service Date	Rev. Code	Procedure Code	NDC	Description	Quantit y	Amount
07/09/2020	0636	J3490	63323013011	DOXYCYCLINE 100 MG SOLR 1 EACH VIAL	1	$264.99
07/09/2020	0636	J3490	63323073911	FAMOTIDINE (PF) 20 MG/2 ML SOLN 2 ML VIAL	1	$80.00
07/09/2020	0636	J3490	63323073911	FAMOTIDINE (PF) 20 MG/2 ML SOLN 2 ML VIAL	1	$80.00

A different style of a Raw IB with a list of itemized charges. The columns read "Service Date, Service, Procedure Code, NDC, Description, Quantity, Amount."

Can AI do it?

Yes. Efficiency and quality gains can be realized by using image processing and machine learning to extract the contents of an IB, normalize the representation into a standard format, and present the reviewer with a comprehensive view of all IB line items within an entire record at once. The images below illustrate a solution in which line items are collected in an interactive list that can then be used as a navigation tool for the record.

Date	Service	Rev Code	Procedure Code	Description	Qty	Amount
10/17/2016	049754	271		ELBOW IMMOBILIZER	5	330.00
10/17/2016	049768	271		LIMB HOLDER QUICK RELEASE 2	1	27.00
10/17/2016	049871	272		KIT VENTRICULOSTOMY 21512	1	886.00
10/17/2016	050052	272		CONNECTING TUBE SUCTION 6	2	46.00

Example of linking and highlighting within CAVO. The example is a photo of a physical copy of a Raw IB with a list of itemized charges with one line item highlighted. This links the IB item to the text.

# ↑	DOS	Revcode	Codes	NDC	Description	Qty	Unit Cost	Charge
0001	10/17/2016	0271			ELBOW IMMOBILIZER	5	$66.00	$330.00
0002	10/17/2016	0271			LIMB HOLDER QUICK RELEASE 2	1	$27.00	$27.00
0003	10/17/2016	0272			KIT VENTRICULOSTOMY 21512	1	$886.00	$886.00
0004	10/17/2016	0272			CONNECTING TUBE SUCTION 6	2	$23.00	$46.00

The photo shows the above example line item linked digitally within the CAVO platform.

So what?

Increasing IB review efficiency follows a similar pattern to that of document classification. The first application of AI allows for the page to be processed and its information presented to the reviewer in a more uniform and intuitive manner. Once the data is in a machine-readable format, AI can be applied to extract a deeper understanding of the content. The below image demonstrates a solution in which the reviewer is warned about potentially incorrect information.

# ↑	DOS	Revcode	Codes	NDC	Description	Qty	Unit Cost	Charge
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
0137	07/09/2020	0636	J3490		DOXYCYCLINE 100 MG SOLR 1 EACH VIAL	1	$264.99	$264.99
0138	07/09/2020	0636			AMOTIDINE (PF) 20 MG/2 ML SOLN 2 ML VIAL	1	$80.00	$80.00
0139	07/09/2020	0636			AMOTIDINE (PF) 20 MG/2 ML SOLN 2 ML VIAL	1	$80.00	$80.00
0140	07/09/2020	0636	J3490	67457043700	DOXYCYCLINE 100 MG SOLR 1 EACH VIAL	1	$166.39	$166.39

60% Unclassified drugs

A screenshot of the CAVO platform with a warning pop-up. The warning reads, "60% Unclassified Drugs."

By applying AI to the text in the itemized bill description and leveraging information about service and revenue codes, it is possible to perform an automated mapping of the line item text to an ontological representation such as HCPCS, LOINC, UMLS, SNOMED-CT, etc. From there, machine learning models can be built to identify billing anomalies and detect line items that are likely to need adjustment. These models can incorporate explicit reimbursement policies that may be defined in addition to being able to infer likely policies from historical behavior.

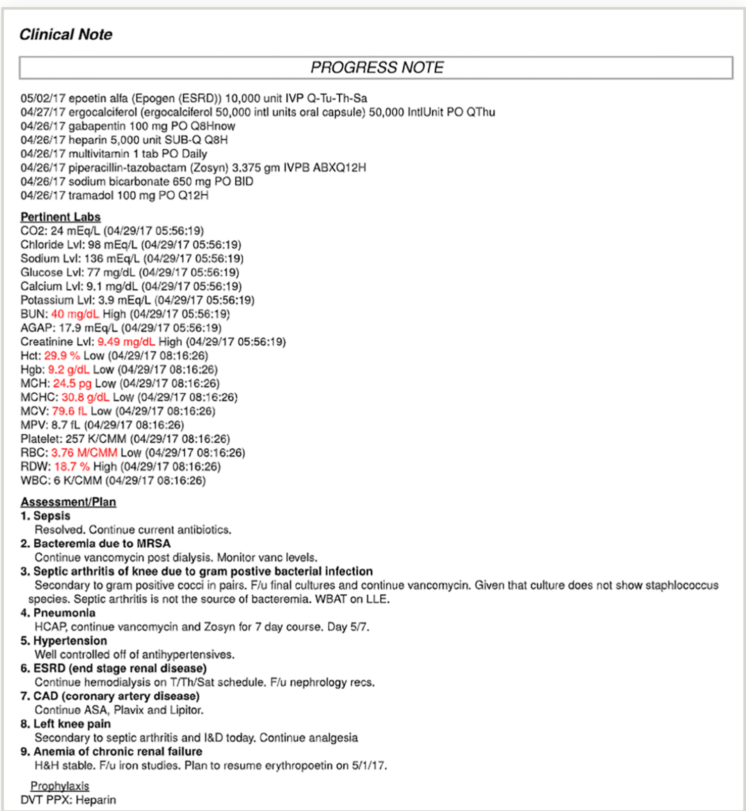
Refer to the section covering "augmented AI" on page 14 for information about how portions of this task can be fully automated.

These use cases are not theoretical—the CAVO platform currently processes millions of pages of records each week. A twelve-thousand-page document takes less than 30 seconds to process.

Time is spent searching records, not on making decisions

Existing workflows and review processes are built around the core assumption that a significant portion of review time is spent finding and validating information within the record. Reviewers leverage extensive collections of complex boolean queries and search terms to surface relevant information as fast as possible; traditional approaches to efficiency improvement tend to focus on making this search process more efficient.

It is now possible to automatically extract a significant amount of clinical information using NLP and an advanced AI. As an example, consider the progress note pictured here.



Clinical Note

PROGRESS NOTE

05/02/17 epoetin alfa (Epogen (ESRD)) 10,000 unit IVP Q-Tu-Th-Sa
04/27/17 ergocalciferol (ergocalciferol 50,000 intl units oral capsule) 50,000 IntlUnit PO QThu
04/26/17 gabapentin 100 mg PO Q8Hnow
04/26/17 heparin 5,000 unit SUB-Q Q8H
04/26/17 multivitamin 1 tab PO Daily
04/26/17 piperacillin-tazobactam (Zosyn) 3.375 gm IVPB ABXQ12H
04/26/17 sodium bicarbonate 650 mg PO BID
04/26/17 tramadol 100 mg PO Q12H

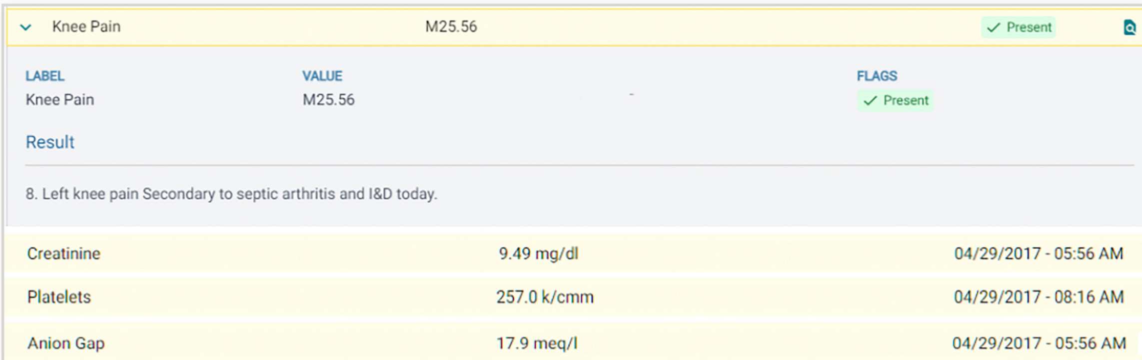
Pertinent Labs
CO2: 24 mEq/L (04/29/17 05:56:19)
Chloride Lvl: 98 mEq/L (04/29/17 05:56:19)
Sodium Lvl: 136 mEq/L (04/29/17 05:56:19)
Glucose Lvl: 77 mg/dL (04/29/17 05:56:19)
Calcium Lvl: 9.1 mg/dL (04/29/17 05:56:19)
Potassium Lvl: 3.9 mEq/L (04/29/17 05:56:19)
BUN: 40 mg/dL High (04/29/17 05:56:19)
AGAP: 17.9 mEq/L (04/29/17 05:56:19)
Creatinine Lvl: 9.49 mg/dL High (04/29/17 05:56:19)
Hct: 29.9 % Low (04/29/17 08:16:26)
Hgb: 9.2 g/dL Low (04/29/17 08:16:26)
MCH: 24.5 pg Low (04/29/17 08:16:26)
MCHC: 30.8 g/dL Low (04/29/17 08:16:26)
MCV: 79.6 fL Low (04/29/17 08:16:26)
MPV: 8.7 fL (04/29/17 08:16:26)
Platelet: 257 K/CMM (04/29/17 08:16:26)
RBC: 3.76 M/CMM Low (04/29/17 08:16:26)
RDW: 18.7 % High (04/29/17 08:16:26)
WBC: 6 K/CMM (04/29/17 08:16:26)

Assessment/Plan
1. Sepsis
Resolved. Continue current antibiotics.
2. Bacteremia due to MRSA
Continue vancomycin post dialysis. Monitor vanc levels.
3. Septic arthritis of knee due to gram postive bacterial infection
Secondary to gram positive cocci in pairs. F/u final cultures and continue vancomycin. Given that culture does not show staphlococcus species. Septic arthritis is not the source of bacteremia. WBAT on LLE.
4. Pneumonia
HCAP, continue vancomycin and Zosyn for 7 day course. Day 5/7.
5. Hypertension
Well controlled off of antihypertensives.
6. ESRD (end stage renal disease)
Continue hemodialysis on T/Th/Sat schedule. F/u nephrology recs.
7. CAD (coronary artery disease)
Continue ASA, Plavix and Lipitor.
8. Left knee pain
Secondary to septic arthritis and I&D today. Continue analgesia
9. Anemia of chronic renal failure
H&H stable. F/u iron studies. Plan to resume erythropoetin on 5/1/17.

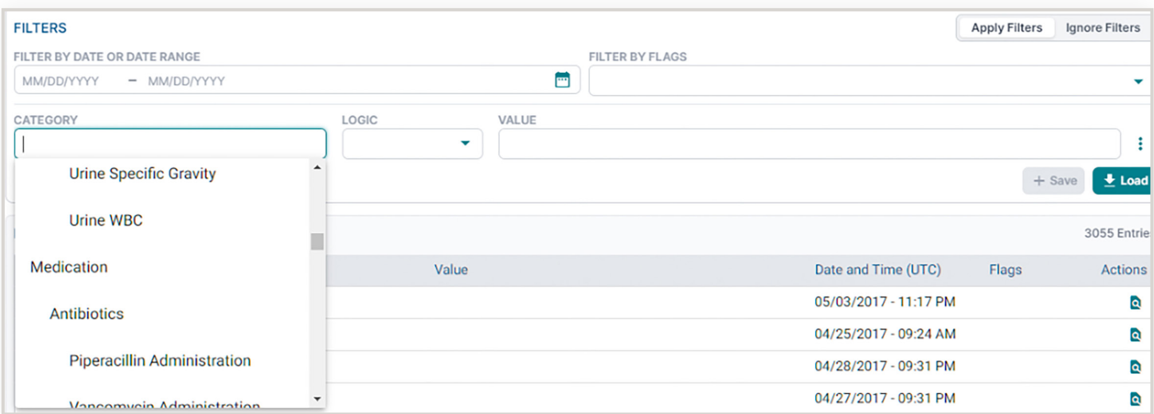
Prophylaxis
DVT PPX: Heparin

The photo is of a clinical progress note. It details raw notes taken by a clinician, lab results, and an assessment/plan.

This page presents a considerable amount of information, only some of which are relevant for a review. By processing and extracting information through our AI model, we can present the information in a uniform, searchable format, as shown in the images of sample extractions. Our model removes the need for a reviewer to spend time figuring out document formats, and finding data points yields significant benefits and efficiency gains.

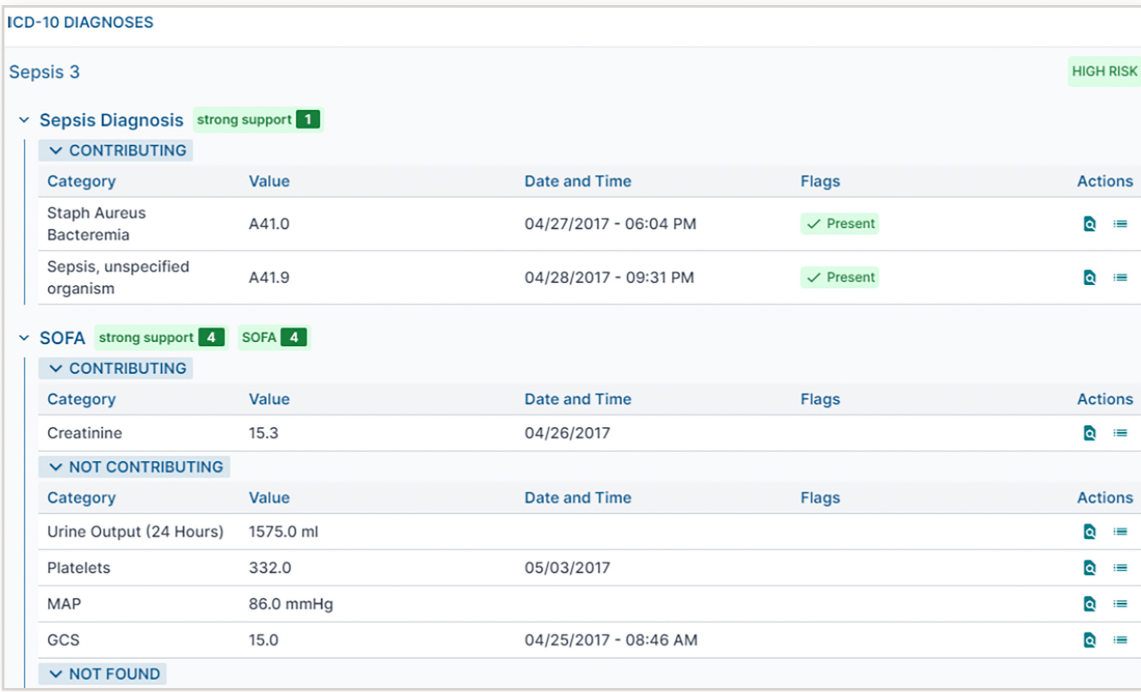


The screenshot shows extracted data from a medical record within the CAVO platform.



The screenshot shows a dropdown menu in the CAVO platform, showing the ability to filter and export medical records.

Extracting structured clinical information from raw paper records enables more advanced record review methods. Structured data is tagged with relevant ontological codes (e.g., SNOMED-CT, LOINC, UMLS, etc.) and can be automatically used in risk models and clinical guideline calculations. Focus shifts from reading records to applying clinical guidelines and policy; AI provides the reviewer with all the information needed to make a decision. This guided review process allows the reviewer to quickly verify a few relevant data points and move on to the creation of an appeal letter, review report, or other final artifacts.


ICD-10 DIAGNOSES

Sepsis 3 HIGH RISK

Sepsis Diagnosis strong support 1

CONTRIBUTING

Category	Value	Date and Time	Flags	Actions
Staph Aureus Bacteremia	A41.0	04/27/2017 - 06:04 PM	Present	
Sepsis, unspecified organism	A41.9	04/28/2017 - 09:31 PM	Present	

SOFA strong support 4 SOFA 4

CONTRIBUTING

Category	Value	Date and Time	Flags	Actions
Creatinine	15.3	04/26/2017		

NOT CONTRIBUTING

Category	Value	Date and Time	Flags	Actions
Urine Output (24 Hours)	1575.0 ml			
Platelets	332.0	05/03/2017		
MAP	86.0 mmHg			
GCS	15.0	04/25/2017 - 08:46 AM		

NOT FOUND

The screenshot is the CAVO platform showing how to use extracted data to calculate clinical guideline scores and automatically apply policy.

Automating the entire review process with AI

AI has disrupted a number of industries by automating processes that previously required human performance. Within the medical review industry, this proposal falls flat: review organizations don't want automated reviews. If the entire process isn't automated, what is the long-term value proposition of using AI within the field?

If AI isn't for automation, what are we even talking about?

The implicit assumption underlying many concerns is that AI is either on or off; it either can automate a process or it cannot. This is not an accurate depiction of how AI can help in practice. The field of Augmented AI focuses on how humans and AI systems can cooperatively perform tasks. Augmented AI allows human reviewers to continue making important decisions; **a reviewer with augmented AI is analogous to a reviewer with access to a nearly-infinite pool of intelligent assistants.** Instead of replacing the reviewer, these assistants interpret raw data and present the reviewer with the facts needed for decision-making.

The ability of AI to support and not supplant decision-making is boosted by another fundamental capability: confidence estimation.

Confidence values: You sure about that?

Popular depictions of successful AI often show an infallible system that knows the right answer and acts upon it. Unsuccessful AI is presented as a system that does not know the right answer but still acts as if it does—often with disastrous consequences. Humans don't make decisions this way, and thankfully, modern implementations of AI no longer function this way.

Consider a jar full of jelly beans. Often, people are asked to estimate the number of jelly beans in a jar, with the person that comes closest winning a prize. It's not easy to guess the exact number of jelly beans, but it is trivial to perform broad estimations of that number; one can be fairly certain that the jar contains more than one jelly bean and less than three billion jelly beans.

There are probably not three billion jelly beans in this jar.

Behind the scenes, AI does not operate in a world of black-and-white decisions; each decision is associated with a confidence value. Some industries have a greater tolerance for errors than others. A marketing AI that makes a mistake results in a consumer viewing an advertisement that is irrelevant to their interests. A medical review AI that makes a mistake may result in an extensive legal process with potentially devastating outcomes. Confidence values allow us to decide precisely how certain an AI must be before coming to a conclusion.

For example, confidence values are used in our itemized bill AI to decide whether the system has enough knowledge to indicate that a line item should be predicted as non-payable or needs further human review.

Future-proofing your review process is critical to long-term success.

Popular depictions of successful AI often show an infallible system that knows the right answer and acts upon it. Unsuccessful AI is presented as a system that does not know the right answer but still acts as if it does—often with disastrous consequences. Humans don't make decisions this way, and thankfully, modern implementations of AI no longer function this way.

About Advent Health Partners

At Advent Health Partners, our mission is to share and apply our clinical expertise, AI-driven technology, and best practices with plans, providers, and partners to increase review team productivity, accelerate appropriate billing and reimbursement and lower the total cost of healthcare. CAVO® from Advent streamlines intuitive access to medical records and other unstructured clinical data sources to support accurate, collaborative billing and reimbursement processes between plans and providers and remove administrative redundancy.